Structural and Multidisciplinary Optimization (2023) 66:127
https://doi.org/10.1007/s00158-023-03574-0

RESEARCH PAPER

A bootstrapping-based weighted average asymptotic sampling formulation for reliability estimation of highly safe structures

Gamze Bayrak[1] · Selver Tuğçe Nalkıran[2] · Erdem Acar[1]

Received: 30 December 2022 / Revised: 20 March 2023 / Accepted: 21 April 2023 / Published online: 17 May 2023

Abstract
Asymptotic sampling (AS) is an efficient simulation-based technique for estimating the small failure probabilities of structures. AS utilizes the asymptotic behavior of the reliability index with respect to the standard deviations of random variables. In this method, the standard deviations of random variables are progressively inflated using a scale parameter to obtain a set of scaled reliability indices. The collection of the standard deviation scale parameters and corresponding scaled reliability indices are called support points. Then, least squares regression is performed using these support points to establish a relationship between the scale parameter and scaled reliability indices. Finally, extrapolation is performed to estimate the actual reliability index. Various extrapolation models have been used in AS to improve accuracy. Moreover, a mean extrapolation formulation using the average value of different extrapolation models was proposed to further improve its accuracy. Although the mean extrapolation formulation protects against using the wrong extrapolation model, it did not guarantee a reliability estimation better than that of the best available extrapolation model. In this paper, we propose a weighted average AS formulation in which the weight factors are optimized to minimize the variance of the reliability index estimation through the bootstrapping method. In the weight factor determination, both convex and affine formulations are considered and the results are compared. The performance of the proposed method is evaluated using six benchmark example problems and a complicated engineering problem. It is found that the proposed weighted average formulation has higher accuracy than the mean extrapolation formulation. For weight factor optimization, the affine formulation yields more accurate results than the convex formulation in most cases.

Keywords Asymptotic behavior · Bootstrap · Extrapolation models · Reliability index · High reliability · Weight factor

1 Introduction

Structural reliability is predicted using a limit-state function (or performance function) to separate the safe and failure regions of an input space. The probability of failure estimation requires the calculation of the multi-dimensional integral of the joint probability density function of all random variables over the failure region. It is expressed as

$$P_f = \int \dots \int I\left[g(x) \le 0\right] f_X(x) dx, \quad (1)$$

where I is an indicator function that has a value of 1 when the condition is true and 0 when the condition is false, $f_{\mathbf{X}}(\mathbf{x})$ is the joint probability density function of the set of random variables $\mathbf{X}$, and $g(\mathbf{x}) \le 0$ defines the failure domain based on the limit-state function $g(\mathbf{x}) = 0$. For most real-life structural problems, the analytical integration of this multidimensional function is not possible; therefore, analytical and simulation-based approaches have been proposed to estimate the failure probability.

Analytical approaches require a small number of limit-state function calculations; therefore, they are typically computationally inexpensive compared to simulation-based approaches. The most popular analytical methods are first-order (Hasofer and Lind 1974; Rackwitz and Fiessler 1978) and second-order reliability methods (Breitung 1984; Tvedt

Responsible Editor: Chao Hu.

Erdem Acar
erdem.acar@gmail.com; acar@etu.edu.tr

[1] Department of Mechanical Engineering, TOBB University of Economics and Technology, Söğütözü, 06560 Ankara, Turkey

[2] Department of Research and Development, BVS Bülbüloğlu Crane Industry, 06935 Sincan, Ankara, Turkey

1990), based on first- and second-order expansions of the limit-state function at the most probable failure point (MPP), respectively. Although analytical approaches are computationally advantageous compared with other methods, they are not necessarily suitable for real-life problems that have complex and nonlinear limit-state functions (e.g., problems with multiple failure modes).

Simulation-based approaches can yield accurate results, provided that a sufficient number of simulations are applied. The most popular simulation-based approach is the Monte Carlo simulation (MCS) method (Rubinstein and Kroese 2016). However, MCS is computationally expensive for estimating small failure probabilities. Variance reduction techniques such as importance sampling (Melchers 1989) and adaptive importance sampling (Wu 1994) can be used to improve the accuracy of failure probability estimations. These methods rely on the concept of an MPP search, and most MPP search algorithms may fail or yield erroneous results when the limit-state function is highly nonlinear or discontinuous. In such cases, simulation-based methods that do not rely on an MPP search, such as stratified sampling (Iman and Conover 1980), subset simulation (Au and Beck 2001), or line sampling (Koutsourelakis et al. 2004), can be used.

Other alternatives include the utilization of metamodels, such as Kriging (Kaymaz 2005; Xiao et al. 2020; Zhou and Lu 2020), neural networks (Gondal and Lee 2012; Papadopoulos et al. 2012), support vector regression (Basudhar and Missoum 2010), radial basis functions (Zhou et al. 2019a, b), and polynomial chaos expansions (Diaz et al. 2018; Zhou et al. 2019a, b). Owing to the extremely high computational cost of implicit functions in certain problems, surrogate models are widely used as a replacement (Jiang et al. 2019; Chojaczyk et al. 2015). However, the performance of surrogate models is affected by high dimensionality. In other words, the computational effort required to construct a surrogate model increases dramatically with the number of random input variables.

For accurate estimation of high reliabilities (or small failure probabilities), various approaches have been developed. These approaches can be categorized into three: sampling-based approaches, surrogate-based approaches and statistics of extremes-based approaches (Lee et al. 2022). Advanced sampling approaches such as importance sampling (Tokdar and Kass 2010), subset simulation (Au and Beck 2001), asymptotic simulation (Bucher 2009) are few methods that have been proven to provide more accurate and more efficient estimations of small probabilities than the crude Monte Carlo method. Surrogate-based approaches include the use of traditional metamodels and machine learning models. Widely used surrogate models include Kriging, support vector regression and neural networks. The use of statistics of extremes for rare event probability estimation is often conducted using generalized Pareto distribution (Kim et al. 2006), extreme value distribution (Jenkinson 1955; Hosking et al. 1985) or other asymptotic distributions.

In the field of high-dimensional reliability analysis, deep learning (Erfani et al. 2016; Jampani et al. 2016) have gain lots of attention due to its capability of extracting critical features from high-dimensional space. Based on the information collected from data, deep learning utilizes a layered structure of algorithms to make predictions, and it has been applied for classification and regression problems in a variety of fields. By employing deep learning techniques, a high-dimensional data abstraction framework was first developed by training deep neural networks (Li and Wang 2020a, b). For reliability analysis of dynamic systems, a long short-term memory augmented deep learning framework was developed to handle time-dependent uncertainties (Li and Wang 2022a). The deep learning-based models were also utilized for reliability-based design optimization (RBDO) problems with limited data (Li and Wang 2022b). The high-dimensional reliability analysis, in which a surrogate model was built to approximate a performance function that is high dimensional, computationally expensive and implicit (Sadoughi et al. 2018). In the absence of sufficient statistical information about the input variables, Bayesian inference has been utilized to quantify the epistemic uncertainty due to lack of data and further formulate Bayesian reliability-based design optimization (Youn and Wang 2008; Srivastava and Deb 2013). A Bayesian-enhanced meta-model approach is used for managing the heterogeneous uncertainties which is due to model imperfection, lack of training data and input variations in RBDO (Li and Wang 2020a, b). The reader is referred to Ramu et al. (2022) for a more comprehensive literature review on machine learning techniques in structural and multidisciplinary optimization.

Amongst these aforementioned methods, this paper focuses on the asymptotic sampling (AS), which is an extrapolation-based method for estimating small failure probabilities of highly safe structures (Bucher 2009). This method extrapolates from low reliability indices to high-reliability indices based on the asymptotic behavior of the failure probability with respect to the standard deviation of the variables. Using a scale parameter, the standard deviations of random variables are progressively inflated to obtain various smaller-scaled reliability indices that can be predicted accurately using a small number of samples. Subsequently, least squares regression was used to establish a relationship between the standard deviation inflation parameter and scaled reliability index values. Finally, extrapolation was performed to estimate the actual reliability index. This method can reduce the computational cost for the estimation of a high reliability index because a low reliability index can be estimated at a lower computational cost. Inspired by the multiple tail median formulation of Ramu et al. (2010), where the median of multiple tail model predictions was used, Zhangchun et al. (2013, 2014) improved the accuracy of the AS method using the mean prediction of various extrapolation models. Acar (2016) increased the effectiveness of AS by reformulating the extrapolation formulation for highly safe structures with separable limit-state functions. Kaveh and Eslamlou (2019) used the flexibility of weighted simulation, which uses a uniform distribution for sampling, to decrease the required calls of the limit-state function with AS. Bayrak and Acar (2021a) critically evaluated the performance of the AS method for highly safe structures and provided guidelines for improving its performance. They showed that even though the mean extrapolation formulation of Zhangchun et al. (2013, 2014) protects against the use of the wrong extrapolation model, it does not guarantee a better reliability estimation than that of the best available extrapolation model. That is, assigning the same importance to all available extrapolation models may lead to a deterioration in accuracy. Therefore, we propose a weighted average AS formulation where the weight factors are optimized to minimize the variance of the reliability index estimated through the bootstrapping method. In the weight factor determination, both convex and affine formulations were considered, and the results were compared.

The remainder of this paper is organized as follows. The AS method is described briefly in Sect. 2. The mean extrapolation formulation used in the AS is presented in Sect. 3. The proposed weighted average formulation is presented in Sect. 4. Numerical examples used in this study are discussed in Sect. 5. The results obtained from these example problems are presented and discussed in Sect. 6. Finally, a summary of important conclusions is presented in Sect. 7.

2 Asymptotic sampling (AS)

Bucher (2009) developed an AS method that enables the accurate estimation of high reliability indices. In this method, the standard deviations of random variables are artificially inflated using a scale parameter to obtain smaller reliability indices known as "scaled" reliability indices. Subsequently, a functional relationship is established between the scale parameters and scaled reliability indices. Finally, the actual reliability index is predicted using the established functional relationship.

Bucher first considered a problem involving a linear limit-state function and suggested that this problem can be reduced to a single variable with a standard deviation of σ via an appropriate coordinate transformation. The reliability index can then be formulated as

$$\beta(f) = \frac{\beta_f}{f}, \tag{2}$$

where f is the scale factor, and β_f is the scaled reliability index computed for the scaled standard deviation of the random variable $\sigma_f = \sigma/f$. The actual reliability index was computed as $\beta_{act} = \beta(f = 1)$. For problems with multiple input random variables, the standard deviations of all random variables are scaled using the same scale factor f. To obtain a good estimate of β_{act}, the reliability index for a larger value of σ (a smaller value of scale factor f) can be computed using MCS and then simply extrapolated by multiplying the result by f.

Bucher also considered a second analytical problem with a hyper circular limit-state function in an n-dimensional Gaussian space, in which the failure domain is defined through $\mathrm{g(X)} = R^2 - \mathrm{X^T X} \le 0$. In this case, the reliability index is expressed in terms of the χ^2-distribution with n degrees of freedom as.

$$\beta = \Phi^{-1}\left[1 - \chi^2\left(f^2R^2, n\right)\right], \tag{3}$$

where $\Phi and \chi^2$ are the cumulative distribution functions of the standard normal distribution and chi-squared distribution, respectively. The relationship between the reliability index and standard deviation scale parameter f is shown in Fig. 1.

Based on the asymptotic behavior of the reliability index with respect to the standard deviation scale parameter, Bucher assumed the following functional relationship between the reliability index and standard deviation scale parameter f:

$$\beta = Af + \frac{B}{f} \tag{4}$$

As $f \to \infty$ (that is, $\sigma_f \to 0$), the reliability index $\beta \to \infty$, to ensure asymptotic behavior. Coefficients A and B are determined from a least squares regression analysis based on the

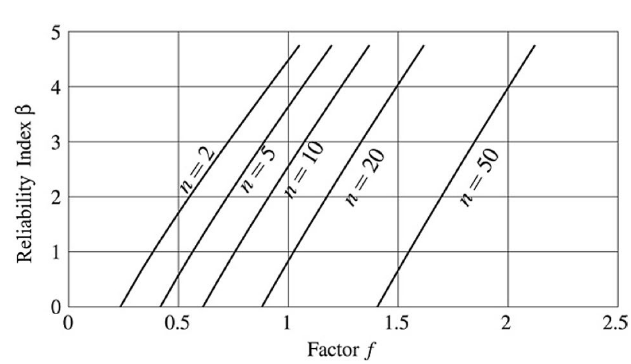


Fig. 1 Relationship between reliability index and standard deviation scale parameter f for hyper circular limit-state function (Bucher 2009)

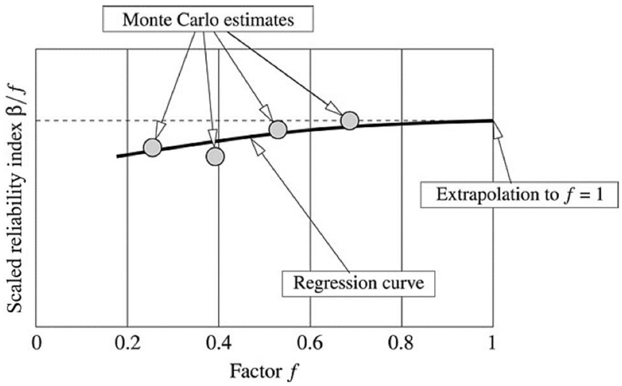


Fig. 2 Concept of asymptotic sampling (Bucher 2009)

estimates of β for different values of f smaller than 1. That is, a set of "support points" $[f_i, \beta(f_i)]$, shown in Fig. 2, is used in the regression.

To assign equal weights to all the support points for the regression analysis, Eq. (4) can be rewritten in terms of a scaled reliability index as

$$\frac{\beta}{f} = A + \frac{B}{f^2} \tag{5}$$

It is essential to use a sampling method that yields stable results. A typical choice is the Latin hypercube sampling (LHS) method (Iman and Conover 1982; Florian 1992). Alternatively, pseudo-random sequences with low-discrepancy sampling methods such as Sobol sequences (Bratley and Fox 1988), Halton sequences (Halton 1960), or good lattice-point sets (Fang et al. 1994) can be utilized. Bayrak and Acar (2021a) compared the use of LHS and Sobol sequences and found that Sobol sequences yielded better results. Therefore, Sobol sequences were used here.

Bucher (2009) initiated the AS algorithm using the scale parameter $f_0 = 1$. The required number of samples in the failure domain was set to $N_0 = 10$. In the first step, the actual number of samples N_F in the failure domain was inadequate (less than N_0). Therefore, f was decreased by a factor of 0.9, and the simulation was repeated until N_F was equal to or exceeded N_0. The support points and regression curves obtained from the extrapolation process are shown in Fig. 2. Bayrak and Acar (2021a) found that the initial scale parameter could be set between 0.3 and 0.4 for a reliability index range of 4–6. Here, the initial scale parameter was set to 0.4 for all example problems.

Bucher (2009) stated that five support points could be used. In his later studies, he used a different number of support points. In a follow-up study, Gasser and Bucher (2018) suggested that four or more support points yielded a more stable regression. However, this practice resulted in an increase in computational effort. Bayrak and Acar (2021a) found that using four support points provided the best compromise between accuracy and efficiency; if the reliability index was extremely high, five support points could be used to achieve an acceptable level of accuracy. Here, four support points were used for all reliability index values.

3 Mean extrapolation formulation

The AS method extrapolates a high reliability index from the obtained low reliability indices. This technique can decrease the computational cost for the evaluation of a high reliability index because a low reliability index can be estimated at a lower computational cost. However, Zhangchun et al. (2013) discovered that the use of a single extrapolation model was not robust. Inspired by the multiple tail median formulation of Ramu et al. (2010), Zhangchun et al. (2013) proposed the generation of multiple extrapolation models and used the mean value of the reliability predictions of these models. Specifically, they proposed using 10-extrapolation models expressed as

$$\beta_t(f) = A_t f + \frac{B_t}{f^{q_t}}\Big(t = 1, 2, 3, 4, 5; q_1 = 3, q_2 = 2, q_3 = 1, q_4 = 0.5, q_5 = \frac{1}{3}\Big) \tag{6}$$

$$\beta_t(f) = A_t f + \frac{B_t}{exp(f^{q_t})}\Big(t = 6, 7, 8, 9, 10; q_6 = 3, q_7 = 2, q_8 = 1, q_9 = 0.5, q_{10} = \frac{1}{3}\Big), \tag{7}$$

where $t = 1,\dots, 10$ represents the extrapolation model index, q_t $(t = 1,\dots, 10)$ is the exponent of the extrapolation model, and exp(.) is an exponential operation with a natural base e. The coefficients A_t and B_t were determined through least squares regression. Then, the actual reliability index was estimated using the average of these 10-extrapolation models and expressed as

$$\beta(1) = \frac{1}{10}\sum_{t=1}^{10}\beta_t(1) = \frac{1}{10}\Big(\sum_{t=1}^{5}(A_t + B_t) + \sum_{t=6}^{10}(A_t + B_t/e)\Big) \tag{8}$$

Here, the models in Eq. (6) are called "*nor* q_3," "*nor* q_2," "*nor* q_1," "*nor* $q_{0.5}$," and "*nor* $q_{1/3}$" in the order they appear in the equation. Similarly, the models in Eq. (7) are called "*exp* q_3," "*exp* q_2," "*exp* q_1," "*exp* $q_{0.5}$," and "*exp* $q_{1/3}$" in the order they appear in the equation.

In a follow-up study, Zhangchun et al. (2014) proposed a new mean extrapolation formulation that involved six extrapolation models to estimate the reliability index. In that study, only the models corresponding to q_2, q_3, and q_4 in Eq. (6) and to q_7, q_8, and q_9 in Eq. (7) were used. Zhangchun et al. (2014) did not compare the accuracy of the two versions.

Bayrak and Acar (2021a) compared the six- and ten-model mean extrapolation formulations and found the six-model mean extrapolation formulation to be more accurate. However, in this paper, a weighted average extrapolation model is used instead of the mean extrapolation formulation. Therefore, in our weighted average formulation, we use both six and ten different individual extrapolation models and compared the results.

Note also that parametric models are used in the extrapolation function. However, non-parametric surrogate or machine learning models can also be used. In an earlier work (Bayrak and Acar 2021b), we compared the performances of the existing parametric models to those of the Gaussian process, support vector regression, and Kriging models and found that the parametric models resulted in better accuracies. Therefore, parametric models are used in the extrapolation function in this study.

4 Proposed weighted average formulation

We showed in our earlier study (Bayrak and Acar 2021a) that even though the mean extrapolation formulation protects against using the wrong extrapolation model, it does not guarantee a better reliability estimation than that of the best available extrapolation models. That is, assigning the same importance to all available extrapolation models may lead to a deterioration in accuracy. Therefore, we propose a weighted average AS formulation inspired by an ensemble of metamodels (Acar and Rais-Rohani 2009). In the weighted average AS formulation, the prediction of the reliability index was obtained as

$$\beta = \sum_{i=1}^{n} w_i \beta_i, \tag{9}$$

where β is the predicted reliability index, n is the number of extrapolation models in the weighted average formulation ($n=6$ or 10), w_i is the weight factor for the ith reliability index, and β_i is reliability index prediction of ith model in the weighted average formulation.

4.1 Weight factor determination

The weight factors in Eq. (9) were optimized to minimize the variance in the reliability index estimation. The optimization problem for determining the weight factor can be formulated using an affine or convex formulation (Strömberg, 2021). In the case of the affine formulation, the weight factors are calculated from

$$\begin{aligned} &\textbf{Find } \boldsymbol{w} \\ &\textbf{Min } \mathrm{Var}(\beta), \\ &\textbf{S.t. } \boldsymbol{w}^T\mathbf{1} = 1 \end{aligned} \tag{10}$$

where $\mathbf{w}$ is the weight factor vector and $\mathbf{1}$ is the vector of ones. Additionally, as suggested by Breiman (1996) and discussed by Viana et al. (2009), a natural constraint to include in the optimization formulation is that the weight factors should be non-negative. For the convex formulation, the weight factors are calculated from

$$\begin{aligned} &\text{Find } w \\ &Min \;\; Var(\beta) \\ &S.t. \;\; \boldsymbol{w}^T 1 = 1, \; and \; w_i \geq 0 \; \text{for i} = 1, \ldots 6 \text{(or 10)}. \end{aligned} \tag{11}$$

4.2 Variance estimation using bootstrap method

The bootstrap method is an efficient numerical method for estimating the distribution of a statistical parameter from a sample set of results (Chernick 2011). The main idea of the bootstrapping method is to generate a number of bootstrap samples by resampling with replacement from the original samples and then approximating the distribution of the statistical parameter of interest (e.g., mean, standard deviation, probability of failure) from the bootstrap samples. Because the resampling procedure is based on selecting data randomly with replacement, the statistical properties of bootstrap samples are different from those of the original samples. Therefore, for any bootstrap sample, the value of the statistical parameter of interest is different, allowing the estimation of the statistical distribution of the statistical parameter of interest. The framework of the bootstrap method is shown in Fig. 3. In this study, the bootstrap method was used to estimate the variance in the reliability index prediction. First, we used $N=512$ Sobol samples in the AS method and stored the support points. Next, we created a $p=10$ set of bootstrap samples for each support point and computed the corresponding scaled reliability index values. Because we have four support points and 10 sets of bootstrap samples (10 scaled reliability index values), we performed least square regression $10^4=10{,}000$ times and obtained 10,000 different extrapolated reliability index values for each extrapolation model, leading to 10,000 reliability index estimations for mean and weighted average extrapolation formulations.

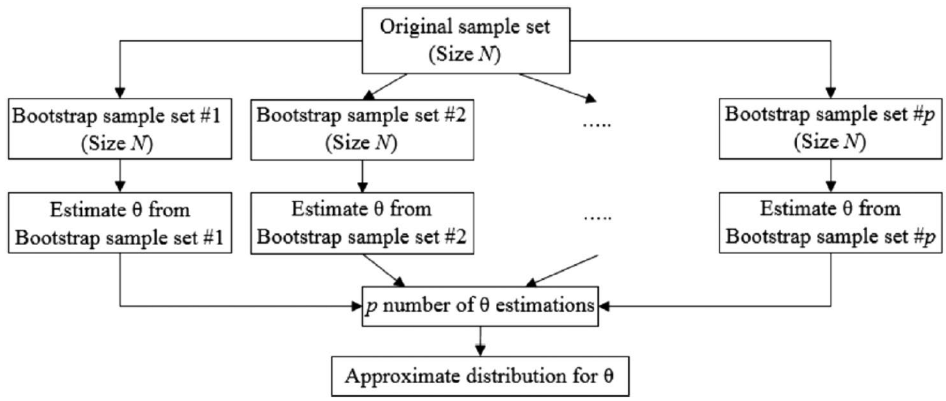


Fig. 3 Framework for bootstrap method (Picheny et al. 2010)

Table 1 Dimensionality of example problems

ID	Problem	Dimensionality (n_{var})
1	Connection rod	2
2	Cantilever beam	3
3	Central crack	4
4	Fortini's clutch	4
5	Roof truss	6
6	I-beam	8
7	Crane bridge	6

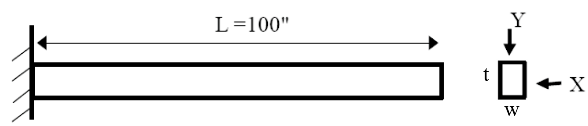


Fig. 5 Cantilever beam under vertical and lateral bending

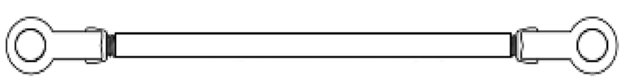


Fig. 4 Connecting-rod under axial loading

Table 2 Statistical properties of random variables in connecting-rod problem

Random variable	Distribution	Mean	Standard deviation
R	Normal	μ_R	6
C	Normal	100	8

5 Example problems

To test the performance of the proposed technique with optimized weight factors, six simple example problems and one complex engineering problem were selected as structural mechanics problems. The first example is a simple two-variable problem involving a linear limit-state function; therefore, an analytical solution can easily be obtained. Starting from this simple example, the dimensionality of the functions (i.e., the number of random variables) was varied (see Table 1). The complicated engineering design problem of a crane bridge was investigated under light (L) and heavy (H) loading conditions.

5.1 Connecting-rod problem

The connecting-rod problem under axial loading is illustrated in Fig. 4. The simple two-variable problem involved the linear limit-state function

$$g = C - R, \tag{12}$$

where R and C are the stress (response) and strength (capacity), respectively, and both are random variables. The statistical properties of the random variables are listed in Table 2.

The mean value of the stress μ_R can be changed to obtain various reliability index values. To solve this problem, the actual reliability index can be easily obtained using Eq. (13) because the limit-state function was linear and both random variables followed a normal distribution. In Eq. (13), μ and σ are the mean and standard deviation of the corresponding quantities, respectively.

$$\beta = \frac{\mu_C - \mu_R}{\sqrt{\sigma_C^2 + \sigma_R^2}} = \frac{100 - \mu_R}{\sqrt{8^2 + 6^2}} = 10 - \frac{\mu_R}{10} \tag{13}$$

Table 3 Statistical properties of random variables in cantilever-beam problem

Random variable	Distribution	Mean	Standard deviation
X [lb]	Normal	500	100
Y [lb]	Normal	1000	100
E [psi]	Normal	29×10^6	1.45×10^6

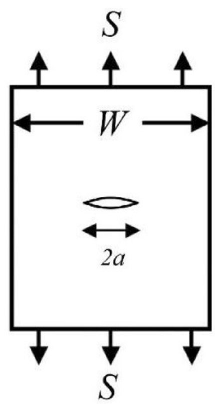


Fig. 6 Central-cracked plate with finite width

Table 4 Statistical properties of random variables in central-rack problem

Random variable	Distribution	Mean	Standard deviation
a [mm]	Normal	25	0.75
W [mm]	Normal	500	5
S [MPa]	Normal	100	10
K_{IC} [$MPa\sqrt{m}$]	Normal	$\overline{K}_{IC}$	0.1 $\overline{K}_{IC}$

5.2 Cantilever-beam problem

The cantilever-beam problem (Wu et al. 2001) is illustrated in Fig. 5. The limit-state occurs when the tip displacement exceeds the allowable D_o and is expressed as

$$g = D_0 - \frac{4L^3}{Ewt}\sqrt{\left(\frac{Y}{t^2}\right)^2 + \left(\frac{X}{w^2}\right)^2}, \tag{14}$$

where E is the modulus of elasticity, X and Y are mutually independent random loads, and the width $w = 2.7''$ and thickness $t = 3.4''$ are the design parameters. The definitions of random variables are presented in Table 3. The allowable displacement D_0 was varied to obtain various reliability levels as shown in Appendix A.

5.3 Central-crack problem

In this example (Bayrak and Acar 2018), a rectangular plate of finite width W with a central through-thickness crack of length $2a$ loaded in tension with a uniform stress S was considered (see Fig. 6). The limit-state function for this problem can be written as

$$g = K_{IC} - \sqrt{sec\left(\frac{\pi a}{W}\right)}S\sqrt{\pi a}, \tag{15}$$

where a is the half-crack length, W is the plate width, S is the applied stress, and K_{IC} is the fracture toughness, all of which were selected randomly. The probability distributions and the means and standard deviations of the random variables are given in Table 4. The mean value of the fracture toughness $\left(\overline{K}_{IC}\right)$ was varied to adjust the reliability level (see Appendix A).

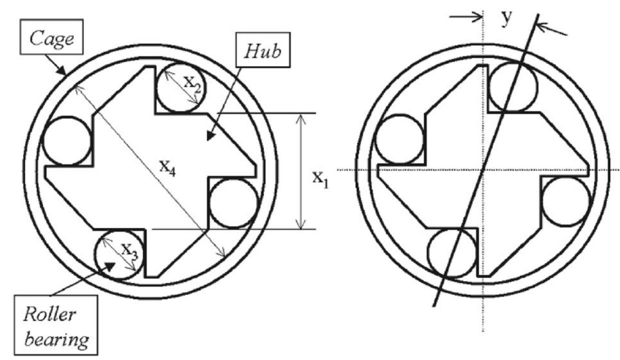


Fig. 7 Fortini's clutch (Lee and Kwak 2006)

Table 5 Statistical properties of random variables in Fortini's clutch problem

Random variable	Distribution	Mean	Standard deviation
X_1 [mm]*	Log-normal	55.29	0.0793
X_2 [mm]	Normal	22.86	0.0043
X_3 [mm]	Normal	22.86	0.0043
X_4 [mm]*	Extreme type I	101.6	0.0793

*For X_1, the scale parameter was $\lambda = 4.01$ and the shape parameter was $\zeta = 0.0014$. For X_4, the location parameter was $\mu = 101.6$ and the scale parameter was $\beta = 0.062$

5.4 Fortini's clutch problem

The Fortini's clutch used in many tolerance analysis studies (Creveling 1997) is illustrated in Fig. 7. The contact angle y is given in terms of the independent component variables X_1, X_2, X_3, and X_4 as

$$y = arccos\left(\frac{X_1 + 0.5(X_2 + X_3)}{X_4 - 0.5(X_2 + X_3)}\right) \tag{16}$$

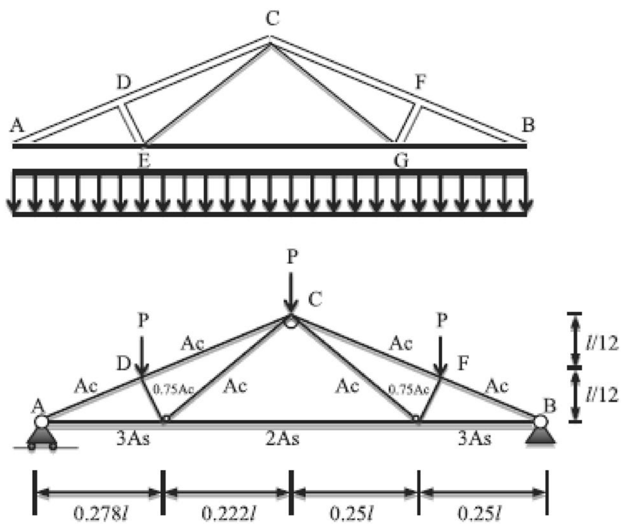


Fig. 8 Roof truss

Table 6 Statistical properties of random variables in roof-truss problem

Random variable	Distribution	Mean	Standard deviation
q [N]	Normal	20×10^3	1400
l [m]	Normal	12	0.12
As [m^2]	Normal	9.82×10^{-4}	5.892×10^{-5}
Ac [m^2]	Normal	0.04	4.8×10^{-3}
Es [Pa]	Normal	1×10^{11}	6×10^9
Ec [Pa]	Normal	2×10^{10}	1.2×10^9

The statistical properties of the random variables are listed in Table 5. The limit-state function for this problem is expressed as

$$g = y - y_{crit}, \tag{17}$$

where y_{crit} was adjusted to obtain various reliability levels as presented in Appendix A.

Table 7 Statistical properties of random variables in I-beam problem

Random variable	Distribution	Mean	Standard deviation
P	Normal	6070	200
L	Normal	120	6
a	Normal	72	6
S	Normal	$\underline{S}$	0.15 $\underline{S}$
d	Normal	2.3	1/24
b_f	Normal	2.3	1/24
t_w	Normal	0.16	1/48
t_f	Normal	0.26	1/48

5.5 Roof-truss problem

A roof truss subjected to uniform loads introduced by Song et al. (2009) is shown in Fig. 8. The top boom and compression members were made of concrete, and the bottom boom was made of steel. The limit-state function is expressed as

$$g = c - \left(\frac{ql^2}{2}\right)\left(\frac{3.81}{AcEc} + \frac{1.13}{AsEs}\right), \tag{18}$$

where c is the vertical deflection at the peak of the structure (node C in Fig. 8), q is the uniform load, l is the length, As and Ac are the sectional areas, and Es and Ec are the moduli of elasticity. The definitions of random variables are given in Table 6. The value of the vertical deflection c was changed to arrange the reliability level of the problem as shown in Appendix A.

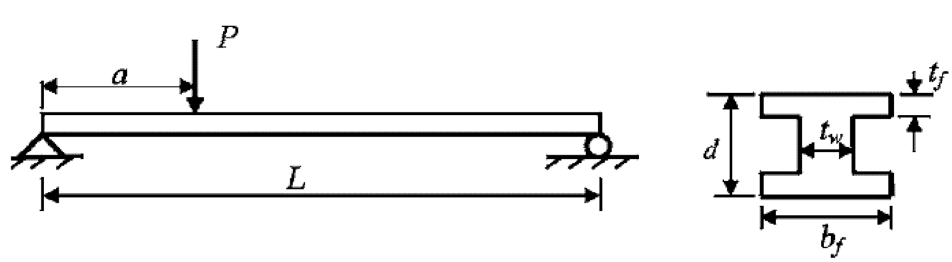
5.6 I-beam problem

A simply supported I-beam is illustrated in Fig. 9. The beam was subjected to a concentrated load as described by Huang and Du (2006). This problem involves a limit-state function defined as the difference between the strength (S) and maximum normal stress (σ_{max}) owing to bending and is expressed as

$$g = S - \sigma_{max}, \tag{19}$$

where

Fig. 9 Cross section and loading for simply supported I-beam



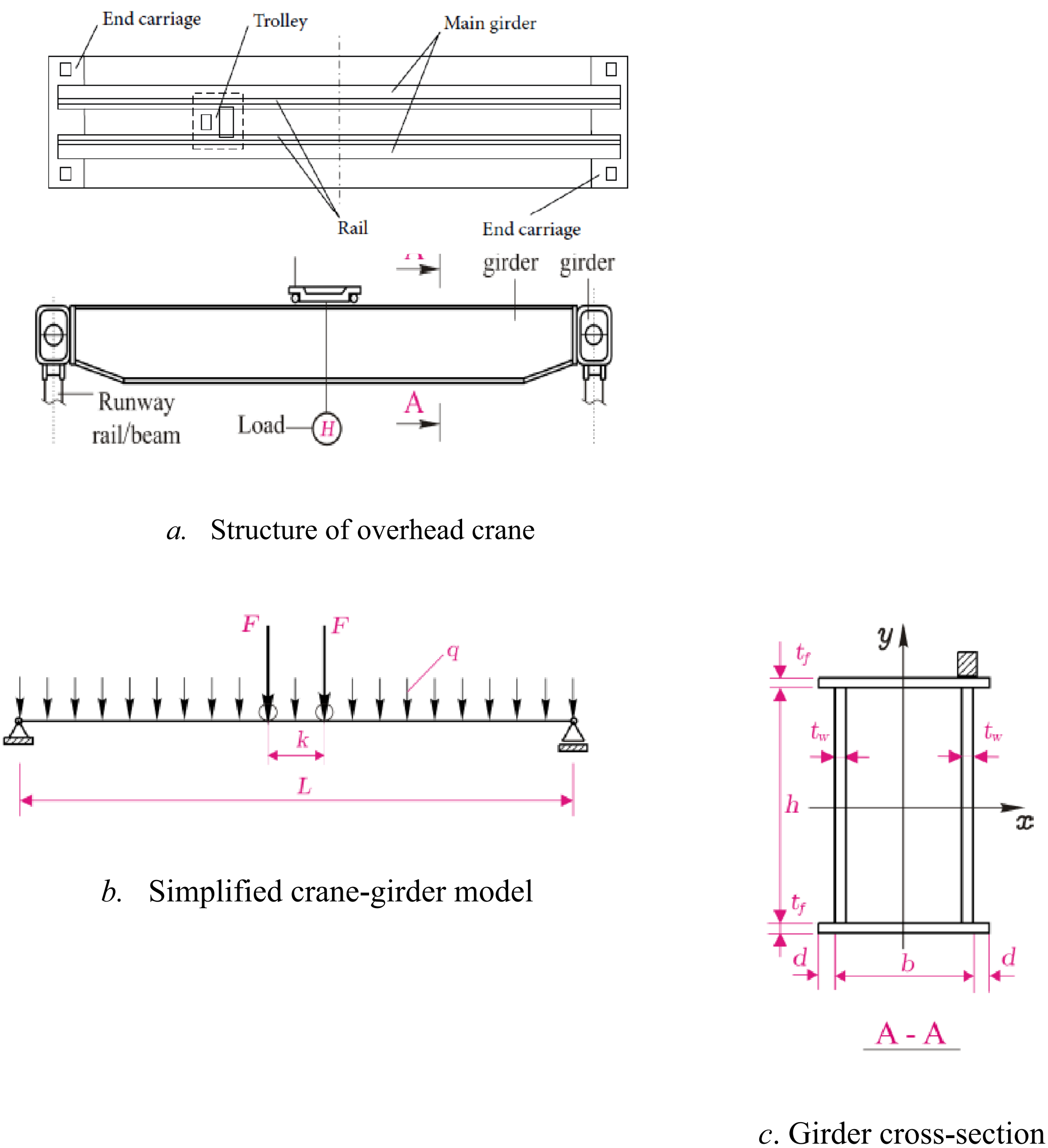


Fig. 10 Crane configuration and girder calculation schema

$$\sigma_{max} = \frac{Pa(L-a)d}{2LI}, I = \frac{b_f d^3 - (b_f - t_w)(d - 2t_f)^3}{12} \tag{20}$$

The statistical properties of the random variables used in this example are listed in Table 7. The mean value of strength was tailored to accommodate the reliability level of the problem as given in Appendix A.

5.7 Crane-bridge problem

The crane-bridge problem is a complicated engineering problem owing to the complexity of the loading and geometry. Farkas (1986) was the first to consider the overhead-crane problem. Van Hai et al. (2020) estimated the failure probabilities of overhead-crane girders with uncertain design parameters. To simplify the problem, we defined the crane design problem within a set of analytical stress constraints.

Here, the double-girder overhead crane consisting of two main bridge girders, two end girders, and a trolley hoist as shown in Fig. 10a is discussed. The trolley hook moves loads on rails mounted or welded on top of bridge beams. Figure 10b shows the loading conditions used to design the double-girder overhead crane. This design consisted of girders with a span of length L. For a single girder, the wheel load F is equal to one-quarter of the sum of the carriage mass G_t and the working load H. The linearly distributed load q represents the weight of the girder and the other distributed loads. The loads caused by wind and other external factors were neglected. The cross-section of the girder is shown in Fig. 10c. The section was assumed to be constant across the span, and the rail and reinforcement parts were neglected when calculating the section properties.

The static stress function on the bottom flange at the mid-span owing to biaxial bending was used as the limit-state function and is expressed as

$$\sigma_s = \frac{M_x}{W_x} + \frac{M_y}{W_y} \leq \sigma_{allow}, \tag{21}$$

where M_x and M_y are the bending moments, W_x and W_y are the section moduli, $\sigma_{allow} = \alpha_d(\alpha_s Y_s)$ is the allowable stress value, Y_s is the yield strength, α_d is the duty factor, and $\alpha_s = 0.59$.

The moments of inertia were calculated as

$$I_x = \frac{t_w h^3}{6} + 2\left[\frac{(b+2d)t_f^3}{12} + \left(\frac{h}{2} + \frac{t_f}{2}\right)^2 (b+2d)t_f\right] \tag{22}$$

and

$$I_y = \frac{t_f(b+2d)^3}{6} + 2\left[\frac{ht_w^3}{12} + \left(\frac{b}{2} - \frac{t_w}{2}\right)^2 ht_w\right] \tag{23}$$

The section moduli were calculated as

$$W_x = \frac{2I_x}{h+2t_f}, W_y = \frac{2I_y}{b+2d} \tag{24}$$

The bending moment owing to vertical loads was calculated as

$$M_x = \frac{L^2 q}{8} + \frac{\psi_d H + G_t}{8L}\left(L - \frac{k}{2}\right)^2, \tag{25}$$

where $(\psi_d H + G_t)/4$ is the wheel load, $q = (k_g A\rho + p_r + p_s)$ is the linear distributed weight with $k_g A\rho$ being the girder-distributed mass when considering stiffeners and diaphragms by a factor of $k_g = 1.05$, $A = 2[ht_w + (b+2d)t_f]$ is the given girder cross-sectional area, $K = 1.9$ m is the distance between the trolley axes, $g = 10$ m/s^2, and $\rho = 7850$ kg/m^3.

The bending moment owing to horizontal loads was calculated as

Table 8 Optimum geometric parameters x^{op} of girders (Farkas 1986)

Girder	h^{op} (mm)	t_w^{op} (mm)	b^{op} (mm)	t_f^{op} (mm)
L	950	5	375	14
H	1000	6	325	18

$$M_y = k_M\left[\frac{L^2 q}{8} + \frac{G_t}{8L}\left(L - \frac{k}{2}\right)^2\right], \tag{26}$$

where $k_M = 0.3 \times 0.5$, and a factor of 0.3 represents the effect of inertial forces, and a factor of 0.5 indicates that two of the four trolley wheels are driven.

Finally, the static stress limit-state function is given as

$$g = \sigma_{allow} - \sigma_s \tag{27}$$

Here, two types of girder overhead cranes with different geometrical properties and loading conditions were used. Because we investigated the effect of AS in systems with high reliability, girders with a high reliability index were selected from Farkas's (1986) study. The abbreviations L and H represent light and heavy loadings, respectively. The reliability index values of these loading cases were 5.82 and 4.58, respectively. The parameters of the girders under different loading conditions are listed in Table 8.

The deterministic data used for the beams are as follows: $L = 22.5$ m, $G_t = 42.25x10^3$ N, $p_r + p_s = 190$ kg/m, $E = 2.1x10^5$ N/mm^2, $d = 10$ mm. H, Y_s, h, t_w, b, and t_f are random variables with a normal distribution, mean μ_i, and variance coefficient COV$_i$. The mean of random variables $\mu_H = 200x10^3$ N, $\mu_{Ys} = 355$ N/mm^2, and $\mu_x = k^{geo} x^{op}$ ($k^{geo} = 1.05$) for the other geometric parameters. The coefficients of variation were COV$_F = 0.05$ for loadings (H and Y_s) and COV$_{geo} = 0.025$ for geometric parameters (h, t_w, b, and t_f).

6 Results

6.1 Numerical procedure

The numerical procedure in this study consisted of the following steps:

i. For each support point, $N = 512$ Sobol samples were generated, from which $p = 10$ sets of bootstrap samples were generated.
ii. $p = 10$ scaled reliability index values corresponding to each set of bootstrap samples were calculated. Because there were ten scaled reliability indices and four support points, there were 10^4 different [f, β] combinations.

iii. Extrapolation was performed for each [f, β] combination to obtain 10^4 extrapolated reliability index values corresponding to each individual extrapolation model.
iv. The variance of the reliability index estimation was obtained using the weighted average extrapolation formulation in terms of weight factors and the optimization of the weight factors to minimize this variance. For the optimization, we used both convex and affine formulations and compared the results.

To reduce the effect of random sampling, the above procedure was repeated 1000 times. The performances of different reliability index extrapolations (individual extrapolation models and mean and weighted average extrapolation formulations) were measured through the RMSE values obtained from these 1,000 runs. We investigated the performance of the AS for these example problems for reliability index values of 4, 4.5, 5, 5.5, and 6, which correspond to failure probabilities of 3.17×10^{-5}, 3.40×10^{-6}, 2.87×10^{-7}, 1.90×10^{-8}, and 9.87×10^{-10}, respectively. Note that the number of

Table 9 Bootstrapping-based weighted average AS results for cantilever-beam problems for a reliability index of 4.0

Bootstrapping results	Normal					Exponential					Mean Ext. Formulation	Weighted Average	
	$q=1/3$	$q=0.5$	$q=1$	$q=2$	$q=3$	$q=1/3$	$q=0.5$	$q=1$	$q=2$	$q=3$	Average	Convex	Affine
Std dev. of β	0.475	0.445	0.372	0.278	0.223	0.503	0.498	0.509	0.575	0.640	0.451	0.223	0.095
w_i (convex)	0	0	0	0	1	0	0	0	0	0	–	–	–
w_i (affine)	−1.501	−2.747	−3.375	−0.495	1.652	1.771	2.619	6.714	4.760	−8.397	–	–	–
AS (repeated 1000 times)													
	Normal					Exponential					Mean Ext. Formulation	Weighted Average	
	$q=1/3$	$q=0.5$	$q=1$	$q=2$	$q=3$	$q=1/3$	$q=0.5$	$q=1$	$q=2$	$q=3$	Average	Convex	Affine
Mean of β	4.139	4.115	4.057	3.980	3.933	4.161	4.157	4.166	4.218	4.268	4.119	3.933	3.780
Std dev. of β	0.362	0.342	0.295	0.237	0.207	0.380	0.377	0.384	0.428	0.471	0.345	0.207	0.148
RMSE	0.377	0.352	0.296	0.243	0.228	0.402	0.397	0.407	0.467	0.528	0.357	0.228	0.291
Bias	0.109	0.085	0.027	−0.050	−0.097	0.131	0.127	0.136	0.188	0.238	0.089	−0.097	−0.250

Table 10 Bootstrapping-based weighted average AS results for cantilever-beam problems for a reliability index of 4.0

Bootstrapping results	Normal			Exponential			Mean Ext. Formulation	Weighted Average	
	$q=0.5$	$q=1$	$q=2$	$q=0.5$	$q=1$	$q=2$	Average	Convex	Affine
Standard dev. of β	0.445	0.372	0.278	0.498	0.510	0.576	0.446	0.278	0.095
w_i (convex)	0	0	1	0	0	0	–	–	–
w_i (affine)	−16.766	−6.888	5.903	5.637	29.303	−16.190	–	–	–
AS (repeated 1000 times)									
	Normal			Exponential			Mean Ext. Formulation	Weighted Average	
	$q=0.5$	$q=1$	$q=2$	$q=0.5$	$q=1$	$q=2$	Average	Convex	Affine
Mean of β	4.115	4.057	3.982	4.157	4.166	4.218	4.115	3.980	3.780
Standard dev. of β	0.342	0.295	0.238	0.377	0.384	0.428	0.342	0.238	0.148
RMSE	0.352	0.296	0.243	0.397	0.407	0.467	0.352	0.243	0.291
Bias	0.085	0.027	−0.050	0.127	0.136	0.188	0.085	−0.050	−0.250

Table 11 Bootstrapping-based weighted average AS results for cantilever-beam problems for a reliability index of 5.0

Bootstrapping results	Normal					Exponential					Mean Ext. Formulation	Weighted Average	
	$q=1/3$	$q=0.5$	$q=1$	$q=2$	$q=3$	$q=1/3$	$q=0.5$	$q=1$	$q=2$	$q=3$	Average	Convex	Affine
Std dev. of β	0.626	0.5848	0.4845	0.354	0.278	0.664	0.657	0.672	0.762	0.852	0.591	0.278	0.110
Weight factors (convex)	0	0	0	0	1	0	0	0	0	0	–	–	–
Weight factors (affine)	−5.381	−7.828	−8.904	−1.428	10.531	−0.306	0.712	5.476	10.150	−2.020	–	–	–
AS (repeated 1000 times)	Normal					Exponential					Mean Ext. Formulation	Weighted Average	
	$q=1/3$	$q=0.5$	$q=1$	$q=2$	$q=3$	$q=1/3$	$q=0.5$	$q=1$	$q=2$	$q=3$	Average	Convex	Affine
Mean of β	5.082	5.061	5.011	4.946	4.906	5.102	5.098	5.106	5.151	5.192	5.066	4.906	4.834
Std dev. of β	0.471	0.443	0.374	0.289	0.243	0.499	0.494	0.506	0.570	0.628	0.449	0.243	0.204
RMSE	0.478	0.447	0.374	0.294	0.261	0.509	0.504	0.517	0.589	0.656	0.453	0.261	0.263
Bias	0.082	0.061	0.011	−0.054	−0.094	0.102	0.098	0.106	0.151	0.192	0.066	−0.094	−0.166

function evaluations required for the reliability prediction of a given problem is 2048. This increases the computational burden for real engineering black-box problems that require heavy computation.

6.2 Cantilever-beam results

The results obtained for the cantilever-beam problem are presented in Tables 9, 10, 11 for reliability index values of 4–6, respectively. The tables present the results for the individual extrapolation models, mean extrapolation formulation (proposed by Zhangchun et al. 2013, 2014), and proposed bootstrapping-based weighted average AS method with affine and convex formulations. The overall process was repeated 1000 times, and the RMSE values for the reliability index estimations were computed. Note that RMSE includes variance and bias error. In this study, the weight factors are selected such that the variance is minimized, because the bias error could not be measured through bootstrapping.

Bootstrapping-based weighted average AS results for cantilever-beam problems for a reliability index of 4.0 are presented in Tables 9 and 10 for ten-model and six-model extrapolations, respectively. Tables 9 and 10 show that the standard deviations of the reliability indices estimated through bootstrapping were larger than the standard deviations of the reliability indices obtained through AS. The bootstrapping successfully ordered the standard deviations of different extrapolation models. In Table 9, the "*nor q3*" model had the smallest standard deviation, the "*nor q2*" model had the second-smallest standard deviation, etc., according to the bootstrapping and repeated sampling. Similarly, in Table 10, the "*nor q2*" model had the smallest standard deviation, the "*nor q1*" model had the second-smallest standard deviation, etc., according to the bootstrapping and repeated sampling.

Table 9 shows that the standard deviation and the RMSE of the reliability index obtained using the weighted average formulation were smaller than those obtained using the mean extrapolation formulation. When the weight factors were optimized using a convex formulation, the weight factor of the extrapolation model with the smallest standard deviation had a value of 1, whereas the other extrapolation models had a value of 0. Table 9 shows that even though the affine weighted average formulation led to a smaller standard deviation of the reliability index than that of the convex weighted average formulation, the RMSE of the reliability index obtained using the convex weighted average formulation was smaller than that of the affine weighted average formulation.

Table 10 also shows that the standard deviation and the RMSE of the reliability index obtained using the weighted average formulation were smaller than those obtained using the mean extrapolation formulation. That is, both ten-model and six-model weighted average formulations provided more accurate results than the mean extrapolation formulation. Comparison of the RMSE results in Tables 9 and 10 reveals that the ten-model weighted average model leads to more accurate reliability index prediction than the six-model weighted average model.

The comparison of the accuracies of the ten-model and six-model weighted average extrapolation models over all numeric example problems are given in Appendix B. It is found that accuracy of the ten-model convex weighted average extrapolation formulation is always better than or that of the six-model convex weighted average extrapolation formulation. It is also found that the accuracies of the ten-model affine weighted average extrapolation and the six-model affine weighted average extrapolation are close.

Table 11 shows that the cantilever beam resulted in a reliability index of 5. The bootstrapping successfully ordered the standard deviations of different extrapolation models. The extrapolation model "*nor q3*" had the smallest standard deviation; therefore, its weight factor was 1, and the weight factors of the other models were 0 in the convex weighted average formulation. Table 10 also shows that the affine weighted average formulation led to a smaller standard deviation and the RMSE of the reliability index obtained using the convex weighted average formulation was smaller than that of the affine weighted average formulation.

Table 12 shows that the cantilever beam resulted in a reliability index of 6. The results presented in Table 12 were similar to those in Tables 9, 10, 11. Although the affine weighted average formulation led to a smaller standard deviation of the reliability index than that of the convex weighted average formulation, the RMSE of the reliability index obtained using the convex weighted average formulation was smaller than that of the affine weighted average formulation.

6.3 Results of other simple example problems

The results of the other example problems are given in detail in Appendix C. For all considered problems, the weighted average extrapolation formulation resulted in smaller standard deviations and smaller RMSE values in the reliability indices than the mean extrapolation formulation. A summary of the comparison of the performance of the affine and convex extrapolation formulations over all simple example problems is listed in Table 13. The affine formulation led to smaller RMSE values for the central crack and connecting rod and to smaller RMSE values. Out of 30 cases, the convex formulation provided the smallest RMSE values in 16 cases, whereas the affine formulation provided the smallest RMSE values in the remaining 14 cases. Although lower standard deviation values were obtained with the affine formulation in the 16 cases, the

Table 12 Bootstrapping-based weighted average AS results for cantilever-beam problems for a reliability index of 6.0

Bootstrapping results	Normal					Exponential					Mean Ext. Formulation	Weighted Average	
	$q=1/3$	$q=0.5$	$q=1$	$q=2$	$q=3$	$q=1/3$	$q=0.5$	$q=1$	$q=2$	$q=3$	Average	Convex	Affine
Std dev. of β	0.874	0.8109	0.6623	0.4838	0.388	0.941	0.935	0.969	1.096	1.183	0.831	0.388	0.167
Weight factors (convex)	0	0	0	0	1	0	0	0	0	0	–	–	–
Weight factors (affine)	−7.575	−8.223	−5.096	1.501	3.757	2.012	7.215	16.769	−8.758	−0.602	–	–	–
AS (repeated 1000 times)	Normal					Exponential					Mean Ext. Formulation	Weighted Average	
	$q=1/3$	$q=0.5$	$q=1$	$q=2$	$q=3$	$q=1/3$	$q=0.5$	$q=1$	$q=2$	$q=3$	Average	Convex	Affine
Mean of β	6.290	6.243	6.131	5.994	5.916	6.340	6.335	6.360	6.454	6.519	6.258	5.916	5.665
Std dev. of β	0.617	0.577	0.484	0.377	0.322	0.660	0.656	0.678	0.760	0.816	0.590	0.322	0.220
RMSE	0.662	0.608	0.490	0.381	0.349	0.721	0.715	0.745	0.860	0.941	0.625	0.349	0.443
Bias	0.240	0.193	0.081	−0.056	−0.134	0.290	0.285	0.310	0.404	0.469	0.208	−0.134	−0.385

Table 13 Results of weighted average AS regarding RMSE values

	Rel. ind. (β)	4	4.5	5	5.5	6
Problem	Cantilever beam	Convex	Convex	Convex	Convex	Convex
	Central crack	Affine	Affine	Affine	Affine	Convex
	Connecting rod	Affine	Affine	Affine	Affine	Affine
	Fortini's clutch	Convex	Convex	Convex	Convex	Convex
	I-beam	Affine	Affine	Affine	Affine	Affine
	Roof truss	Convex	Convex	Convex	Convex	Convex

Table 14 Bootstrapping-based weighted average AS results for crane-bridge problems for a reliability index of 4.58

Bootstrapping results	Normal					Exponential					Mean Ext. Formulation	Weighted Average	
	$q=1/3$	$q=0.5$	$q=1$	$q=2$	$q=3$	$q=1/3$	$q=0.5$	$q=1$	$q=2$	$q=3$	Average	Convex	Affine
Std dev. of β	0.517	0.4841	0.4038	0.2997	0.239	0.548	0.542	0.555	0.628	0.698	0.490	0.239	0.097
Weight factors (convex)	0	0	0	0	1	0	0	0	0	0	–	–	–
Weight factors (affine)	2.935	2.854	3.048	– 1.182	– 8.179	4.835	6.336	8.496	0.385	– 18.526	–	–	–

AS (repeated 1000 times)	Normal					Exponential					Mean Ext. Formulation	Weighted Average	
	$q=1/3$	$q=0.5$	$q=1$	$q=2$	$q=3$	$q=1/3$	$q=0.5$	$q=1$	$q=2$	$q=3$	Average	Convex	Affine
Mean of β	4.660	4.648	4.620	4.581	4.558	4.671	4.669	4.674	4.699	4.724	4.651	4.558	4.485
Std dev. of β	0.451	0.427	0.370	0.300	0.263	0.474	0.470	0.480	0.533	0.586	0.432	0.263	0.196
RMSE	0.458	0.433	0.372	0.300	0.264	0.483	0.478	0.488	0.546	0.603	0.437	0.264	0.217
Bias	0.080	0.068	0.040	0.001	– 0.022	0.091	0.089	0.094	0.119	0.144	0.071	– 0.022	– 0.095

reason for the lower RMSE value of the convex formulation was the higher bias error of the affine formulation.

6.4 Results of crane-bridge problem

The results obtained for the crane-bridge problem are presented in Tables 14 and 15 for reliability index values of 4.58 and 5.82, respectively. As in the earlier simple example problems, bootstrapping successfully ordered the standard deviations of different extrapolation models. In both cases, the RMSE of the reliability index obtained using the weighted average formulation was smaller than that obtained using the mean extrapolation formulation. The weight factor of the extrapolation model with the smallest standard deviation had a value of 1, whereas the other extrapolation models had values of 0. The standard deviation and RMSE for the dependability index were lower with the affine weighted average formulation than with the convex weighted average formulation.

7 Concluding remarks

A bootstrapping-based weighted average AS formulation in which the weight factors were optimized to minimize the variance of the reliability index estimation was proposed. For the weight factor determination, both convex and affine formulations were considered, and the results were compared. The performance of the proposed method was evaluated using six benchmark example problems and one complicated engineering problem. From the obtained results, the following conclusions can be drawn:

Table 15 Bootstrapping-based weighted average AS results for crane-bridge problems for a reliability index of 5.82

Bootstrapping results	Normal					Exponential					Mean Ext. Formulation	Weighted Average	
	$q=1/3$	$q=0.5$	$q=1$	$q=2$	$q=3$	$q=1/3$	$q=0.5$	$q=1$	$q=2$	$q=3$	Average	Convex	Affine
Std dev. of β	0.882	0.8923	0.8352	0.628	0.504	0.876	0.896	0.933	0.964	0.972	0.827	0.504	0.155
Weight factors (convex)	0	0	0	0	1	0	0	0	0	0	–	–	–
Weight factors (affine)	0.946	−0.190	0.399	−4.406	4.940	1.926	−2.389	0.123	−0.342	−0.006	–	–	–
AS (repeated 1000 times)	Normal					Exponential					Mean Ext. Formulation	Weighted Average	
	$q=1/3$	$q=0.5$	$q=1$	$q=2$	$q=3$	$q=1/3$	$q=0.5$	$q=1$	$q=2$	$q=3$	Average	Convex	Affine
Mean of β	5.824	5.810	5.777	5.736	5.712	5.839	5.837	5.845	5.873	5.893	5.815	5.712	5.628
Std dev. of β	0.661	0.621	0.525	0.410	0.350	0.703	0.698	0.718	0.804	0.872	0.632	0.350	0.232
RMSE	0.661	0.620	0.526	0.419	0.366	0.703	0.698	0.718	0.805	0.875	0.632	0.366	0.301
Bias	0.004	−0.010	−0.043	−0.084	−0.108	0.019	0.017	0.025	0.053	0.073	−0.005	−0.108	−0.192

- Bootstrapping successfully orders the standard deviations of different extrapolation models,
- the standard deviation and RMSE of the reliability index obtained using the weighted average formulation were smaller than those obtained using the mean extrapolation formulation,
- when the weight factors were optimized using a convex formulation, the weight factor of the extrapolation model with the smallest standard deviation had a value of 1, whereas the other extrapolation models had value of 0,
- of the 32 cases (including the crane-bridge problem) considered, the convex formulation provided the smallest RMSE values in 16 cases, whereas the affine formulation provided the smallest RMSE values in 16 cases,
- for the 16 cases the convex formulation provided the smallest RMSE values, it was observed that bias errors of the affine formulation were also larger.

In this paper, we aimed to enhance the accuracy of reliability estimation in highly safe structures. Investigation of the improvements in terms of computational efficiency is left for a future study.

Appendix

Reliability levels of numerical example problems

For the numerical example problems, five different reliability levels were considered by changing the proper term in the LSF (see Table 16). The reliability index values reported in Table 16 were predicted using crude MCSs with sample sizes of 10^7, 10^8, 10^9, 10^{10}, and 10^{11} for reliability indices 4, 4.5, 5, 5.5, and 6, respectively. The reliability indices of 4, 4.5, 5, 5.5, and 6 correspond to the failure probabilities of 3.17×10^{-5}, 3.40×10^{-6}, 2.87×10^{-7}, 1.90×10^{-8}, and 9.87×10^{-10}, respectively.

Comparison of the accuracies of ten-model and six-model weighted average extrapolation models over all example problems

Figure 11 provides a comparison of RMSE values of ten-model and six-model weighted average extrapolation models over all example problems. Figure 11a shows that the RMSE of the best individual model in the ten-model extrapolation formulation is always smaller than or equal to that of the six-model extrapolation formulation, as expected. Figure 11b shows that the RMSE values of the ten-model

Table 16 Reliability levels considered for example problems

ID	Problem	Term[a]	Value[b]	β[c]	Value[b]	β[c]	Value[b]	β[c]	Value[b]	β[c]	Value[b]	β[c]
1	Connection rod	μ_R	60	4.00	55	4.50	50	5.00	45	5.50	40	6.00
2	Cantilever beam	D_0	2.50	4.03	2.62	4.51	2.75	5.00	2.89	5.54	3.04	6.05
3	Central crack	$\overline{K}_{IC}$	52	4.01	57	4.52	63	5.01	70	5.52	79	6.04
4	Fortini's clutch	y_{crit}	4.05	4.02	3.55	4.53	3.02	5.01	2.31	5.50	1.20	6.04
5	Roof truss	c	0.0360	4.07	0.0378	4.53	0.0400	5.01	0.0425	5.50	0.0466	6.07
6	I-beam	$\overline{S}$	410×10^3	4.07	490×10^3	4.50	630×10^3	5.01	880×10^3	5.49	1700×10^3	6.06

[a]Term in limit-state function was varied to change reliability level

[b]Value of term

[c]Corresponding reliability index

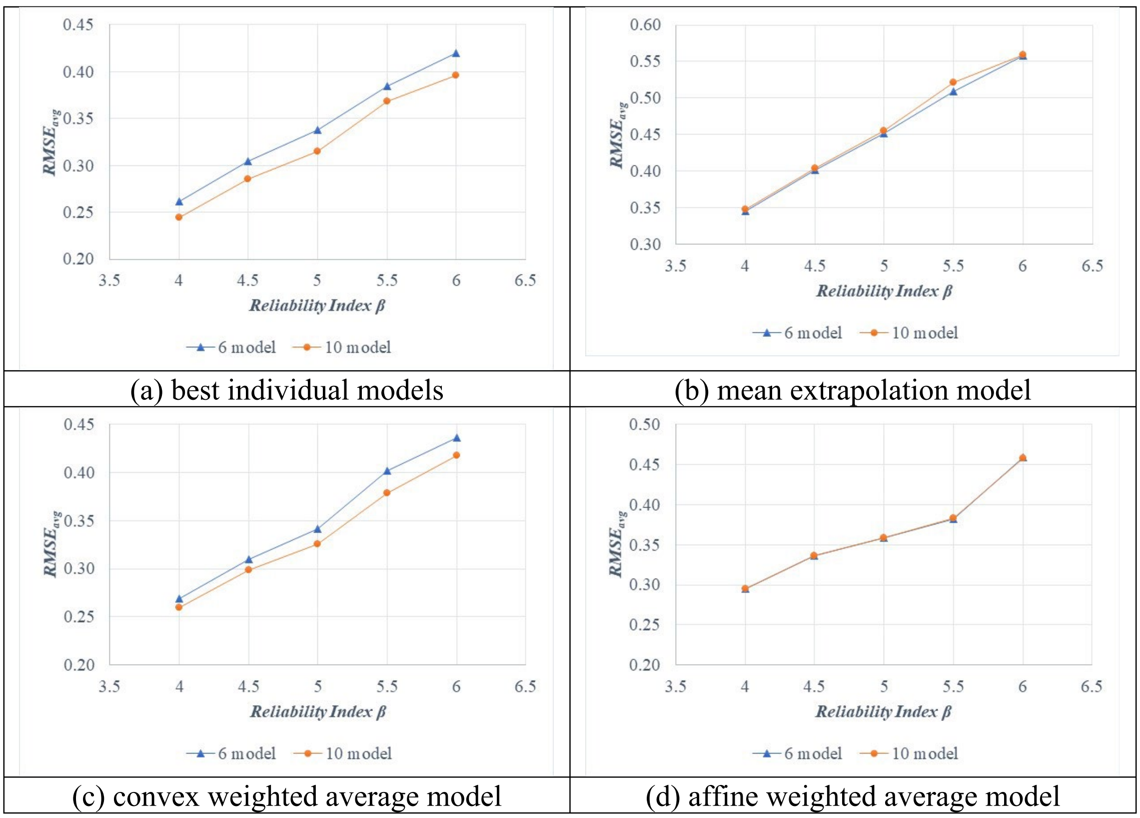


Fig. 11 Comparison of RMSE values of ten-model and six-model weighted average extrapolation models over all example problems

mean extrapolation and the six-model mean extrapolation are close, and the six-model mean extrapolation formulation is slightly more accurate. Figure 11c shows that the RMSE of the ten-model convex weighted average extrapolation formulation is always smaller than or that of the six-model convex weighted average extrapolation formulation. Finally, Fig. 11d shows that the RMSE values of the ten-model affine weighted average extrapolation and the six-model affine weighted average extrapolation are close.

Standard deviation and RMSE results for example problems

(See Tables 17, 18, 19, 20, 21, 22, 23).

Table 17 Results of cantilever-beam problem for all reliability levels

Rel. Index (β)		4.0	4.5	5.0	5.5	6.0
Average	Std of β in bootstrap	0.451	0.556	0.591	0.912	0.831
	Std of β	0.346	0.383	0.449	0.559	0.621
	RMSE	0.358	0.390	0.453	0.560	0.627
Convex	Std of β in bootstrap	0.223	0.269	0.278	0.449	0.388
	Std of β	0.207	0.227	0.243	0.302	0.322
	RMSE	0.228	0.246	0.261	0.333	0.349
Affine	Std of β in bootstrap	0.095	0.113	0.110	0.185	0.167
	Std of β	0.148	0.162	0.204	0.215	0.220
	RMSE	0.291	0.278	0.263	0.349	0.442
Best individual model ($nor\ q_3$)	RMSE	0.228	0.246	0.261	0.333	0.349

Table 18 Results of central-crack problem for all reliability levels

Rel. Index (β)		4.0	4.5	5.0	5.5	6.0
Average	Std of β in bootstrap	0.567	0.575	0.582	0.752	0.998
	Std of β	0.349	0.437	0.493	0.577	0.628
	RMSE	0.348	0.438	0.494	0.580	0.630
Convex	Std of β in bootstrap	0.287	0.266	0.260	0.334	0.426
	Std of β	0.208	0.231	0.267	0.284	0.305
	RMSE	0.208	0.231	0.268	0.285	0.305
Affine	Std of β in bootstrap	0.111	0.100	0.103	0.138	0.169
	Std of β	0.162	0.201	0.218	0.215	0.309
	RMSE	0.162	0.202	0.220	0.217	0.313
Best individual model ($nor\ q_3$)	RMSE	0.208	0.231	0.268	0.285	0.305

Table 19 Results of connecting-rod problem for all reliability levels

Rel. Index (β)		4.0	4.5	5.0	5.5	6.0
Average	Std of β in bootstrap	0.623	0.529	0.352	0.676	0.857
	Std of β	0.325	0.403	0.435	0.488	0.580
	RMSE	0.325	0.403	0.435	0.490	0.583
Convex	Std of β in bootstrap	0.276	0.247	0.165	0.316	0.364
	Std of β	0.182	0.204	0.220	0.271	0.275
	RMSE	0.182	0.204	0.220	0.272	0.275
Affine	Std of β in bootstrap	0.117	0.097	0.072	0.125	0.144
	Std of β	0.149	0.161	0.167	0.189	0.219
	RMSE	0.149	0.161	0.167	0.189	0.219
Best individual model ($nor\ q_3$)	RMSE	0.182	0.204	0.220	0.272	0.275

Table 20 Results of Fortini's clutch problem for all reliability levels

Rel. Index (β)		4.0	4.5	5.0	5.5	6.0
Average	Std of β in bootstrap	0.528	0.595	0.547	0.946	0.721
	Std of β	0.357	0.387	0.440	0.537	0.591
	RMSE	0.383	0.431	0.499	0.620	0.738
Convex	Std of β in bootstrap	0.295	0.299	0.282	0.482	0.327
	Std of β	0.197	0.235	0.244	0.280	0.294
	RMSE	**0.286**	**0.381**	**0.428**	**0.506**	**0.611**
Affine	Std of β in bootstrap	0.098	0.118	0.123	0.154	0.139
	Std of β	0.146	0.175	0.162	0.189	0.203
	RMSE	0.319	0.433	0.476	0.549	0.682
Best individual model (*nor* q_3)	RMSE	**0.286**	**0.381**	**0.428**	**0.506**	**0.611**

Table 21 Results of I-beam problem for all reliability levels

Rel. Index (β)		4.0	4.5	5.0	5.5	6.0
Average	Std of β in bootstrap	0.523	0.833	0.660	0.799	0.842
	Std of β	0.358	0.408	0.448	0.482	0.521
	RMSE	0.357	0.409	0.451	0.486	0.521
Convex	Std of β in bootstrap	0.267	0.420	0.306	0.372	0.399
	Std of β	0.211	0.245	0.248	0.275	0.289
	RMSE	0.219	0.255	0.255	0.283	0.293
Affine	Std of β in bootstrap	0.103	0.140	0.115	0.152	0.144
	Std of β	0.170	0.185	0.221	0.210	0.221
	RMSE	**0.208**	**0.215**	**0.248**	**0.232**	**0.231**
Best individual model (*nor* q_3)	RMSE	0.219	0.255	0.255	0.283	0.293

Table 22 Results of roof-truss problem for all reliability levels

Rel. Index (β)		4.0	4.5	5.0	5.5	6.0
Average	Std of β in bootstrap	0.409	0.521	0.711	0.568	0.900
	Std of β	0.361	0.419	0.482	0.523	0.624
	RMSE	0.362	0.422	0.489	0.554	0.678
Convex	Std of β in bootstrap	0.193	0.269	0.369	0.290	0.445
	Std of β	0.214	0.242	0.275	0.272	0.321
	RMSE	**0.436**	**0.474**	**0.519**	**0.594**	**0.671**
Affine	Std of β in bootstrap	0.080	0.118	0.155	0.130	0.174
	Std of β	0.165	0.169	0.183	0.251	0.217
	RMSE	0.640	0.731	0.779	0.761	0.859
Best individual model (*nor* q_3)	RMSE	**0.436**	**0.474**	**0.519**	**0.594**	**0.671**

Table 23 Results of crane-bridge problem for all reliability levels

Rel. Index (β)		4.0	4.5	5.0	5.5	6.0
Average	Std of β in bootstrap	–	0.490	–	–	0.827
	Std of β	–	0.432	–	–	0.632
	RMSE	–	0.437	–	–	0.632
Convex	Std of β in bootstrap	–	0.239	–	–	0.504
	Std of β	–	0.263	–	–	0.350
	RMSE	–	0.264	–	–	0.366
Affine	Std of β in bootstrap	–	0.097	–	–	0.155
	Std of β	–	0.196	–	–	0.232
	RMSE	–	**0.218**	–	–	**0.301**
Best individual model (*nor* q_3)	RMSE	–	0.264	–	–	0.366

Acknowledgements The authors acknowledge the funding provided by the BVS Bülbüloğlu Crane Industry, and their technical support on the crane-bridge problem.

Funding This study was supported by BVS Bulbuloglu Crane Industry.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Replication of results The results provided herein are replicable. Interested readers can contact the corresponding author to obtain the MATLAB codes used to generate the results.

References

Acar E (2011) Guided tail modelling for efficient and accurate reliability estimation of highly safe mechanical systems. Proc Inst Mech Eng C J Mech Eng Sci 225(5):1237–1251

Acar E (2016) A reliability index extrapolation method for separable limit states. Struct Multidisc Optim 53(5):1099–1111

Acar E, Rais-Rohani M (2009) Ensemble of metamodels with optimized weight factors. Struct Multidisc Optim 37(3):279–294

Au SK, Beck JL (2001) Estimation of small failure probabilities in high dimensions by subset simulation. Probab Eng Mech 16(4):263–277

Basudhar A, Missoum S (2010) An improved adaptive sampling scheme for the construction of explicit boundaries. Struct Multidisc Optim 42(4):517–529

Bayrak G, Acar E (2018) Reliability estimation using Markov chain Monte Carlo–based tail modeling. AIAA J 56(3):1211–1224

Bayrak G, Acar E (2021a) A critical evaluation of asymptotic sampling method for highly safe structures. Struct Multidisc Optim. https://doi.org/10.1007/s00158-021-03057-0

Bayrak, G. & Acar, E. (2021b). Asymptotic sampling regression with machine learning and surrogate modeling techniques. *International Conference on Engineering Technologies*, 2021b, Konya, Türkiye.

Bennett JA, Botkin ME (1986) The optimum shape. Plenum Press, New York, N.Y.

Bratley P, Fox BL (1988) Algorithm 659: implementing Sobol's quasirandom sequence generator. ACM Trans Math Softw (TOMS) 14(1):88–100

Breiman L (1996) Stacked regressions. Mach Learn 24(1):49–64. https://doi.org/10.1007/bf00117832

Breitung K (1984) Asymptotic approximations for multinormal integrals. J Eng Mech 110(3):357–366

Bucher C (2009) Asymptotic sampling for high-dimensional reliability analysis. Probab Eng Mech 24(4):504–510

Chaudhuri A, Haftka RT (2013) Separable Monte Carlo combined with importance sampling for variance reduction. Int J Reliab Saf 7(3):201–221

Chaudhuri A, Kramer B, Willcox KE (2020) Information reuse for importance sampling in reliability-based design optimization. Reliab Eng Syst Saf 201:106853

Chernick MR (2011) Bootstrap methods: a guide for practitioners and researchers. Wiley, Hoboken

Chojaczyk AA, Teixeira AP, Neves LC, Cardoso JB, Soares CG (2015) Review and application of artificial neural networks models in reliability analysis of steel structures. Struct Saf 52:78–89

Creveling CM (1997) Tolerance design: a handbook for developing optimal specifications. Prentice Hall, Hoboken

Dahlblom O, Peterson A, Petersson H (1986) CALFEM — a program for computer-aided learning of the finite element method. Eng Comput 3(2):155–160. https://doi.org/10.1108/eb023653

Diaz P, Doostan A, Hampton J (2018) Sparse polynomial chaos expansions via compressed sensing and D-optimal design. Comput Methods Appl Mech Eng 336:640–666

Erfani SM, Rajasegarar S, Karunasekera S, Leckie C (2016) High-dimensional and large-scale anomaly detection using a linear one-class SVM with deep learning. Pattern Recogn 58:121–134

Fang KT, Wang Y, Bentler PM (1994) Some applications of number-theoretic methods in statistics. Stat Sci 416–428. https://doi.org/10.1214/ss/1177010392

Farkas J (1986) Economy of higher-strength steels in overhead travelling cranes with double-box girders. J Constr Steel Res 6(4):285–301

Florian A (1992) An efficient sampling scheme: updated latin hypercube sampling. Probab Eng Mech 7(2):123–130

Gasser C, Bucher C (2018) An optimized strategy for using asymptotic sampling for reliability analysis. Struct Saf 71:33–40

Gondal ZUR, Lee J (2012) Reliability assessment using feed-forward neural network-based approximate meta-models. Proc Inst Mech Eng o 226(5):448–454

Halton JH (1960) On the efficiency of certain quasi-random sequences of points in evaluating multi-dimensional integrals. Numer Math 2(1):84–90

Hasofer A, Lind N (1974) Exact and invariant second-moment code format. J Eng Mech Div Proc ASCE 100(1):111–121

Hosking JRM, Wallis JR, Wood EF (1985) Estimation of the generalized extreme-value distribution by the method of probability-weighted moments. Technometrics 27(3):251–261

Huang B, Du X (2006) Uncertainty analysis by dimension reduction integration and saddlepoint approximations. J Mech Des 28:26–33

Iman RL, Conover WJ (1980) Small sample sensitivity analysis techniques for computer models with an application to risk assessment. Commun Stat-Theor Methods 9(17):1749–1842

Iman RL, Conover WJ (1982) A distribution-free approach to inducing rank correlation among input variables. Commun Stat-Simul Comput 11(3):311–334

Jampani, V., Kiefel, M., & Gehler, P. V. (2016). Learning sparse high dimensional filters: Image filtering, dense crfs and bilateral neural networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 4452–4461).

Jenkinson AF (1955) The frequency distribution of the annual maximum (or minimum) values of meteorological elements. Q J R Meteorol Soc 81(348):158–171

Jiang C, Qiu H, Yang Z, Chen L, Gao L, Li P (2019) A general failure-pursuing sampling framework for surrogate-based reliability analysis. Reliab Eng Syst Saf 183:47–59

Kaveh A, Dadras Eslamlou A (2019) An efficient method for reliability estimation using the combination of asymptotic sampling and weighted simulation. Scientia Iranica 26(4):2108–2122

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

Kim, N., Queipo, N., & Ramu, P. (2006). Tail modeling in reliability-based design optimization for highly safe structural systems. In *47th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference 14th AIAA/ASME/AHS Adaptive Structures Conference 7th* (p. 1825).

Koutsourelakis PS, Pradlwarter HJ, Schuëller GI (2004) Reliability of structures in high dimensions, part I: algorithms and applications. Probab Eng Mech 19(4):409–417

Lee SH, Kwak BM (2006) Response surface augmented moment method for efficient reliability analysis. Struct Saf 28(3):261–272

Lee I, Lee U, Ramu P, Yadav D, Bayrak G, Acar E (2022) Small failure probability: principles, progress and perspectives. Struct Multidisc Optim 65(11):326

Li M, Wang Z (2020a) Deep learning for high-dimensional reliability analysis. Mech Syst Signal Process 139:106399

Li M, Wang Z (2020b) Heterogeneous uncertainty quantification using Bayesian inference for simulation-based design optimization. Struct Saf 85:101954

Li M, Wang Z (2022a) LSTM-augmented deep networks for time-variant reliability assessment of dynamic systems. Reliab Eng Syst Saf 217:108014

Li M, Wang Z (2022b) Deep reliability learning with latent adaptation for design optimization under uncertainty. Comput Methods Appl Mech Eng 397:115130

Maute K (2009) Design element toolbox, center for aerospace structures. University of Colorado, Boulder, CO, USA

Melchers RE (1989) Importance sampling in structural systems. Struct Saf 6(1):3–10

Papadopoulos V, Giovanis DG, Lagaros ND, Papadrakakis M (2012) Accelerated subset simulation with neural networks for reliability analysis. Comput Methods Appl Mech Eng 223:70–80

Picheny V, Kim NH, Haftka RT (2010) Application of bootstrap method in conservative estimation of reliability with limited samples. Struct Multidisc Optim 41(2):205–217

Picheny V, Kim NH, Haftka RT, Quiepo NV (2008) “Conservative predictions using surrogate modeling,” 49th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials, April, Schaumburg, IL.

Qu X, Haftka RT (2004) Reliability-based design optimization using probability sufficiency factor. Struct Multidisc Optim 27(5):314–325

Rackwitz R, Flessler B (1978) Structural reliability under combined random load sequences. Comput Struct 9(5):489–494

Ramu P, Kim NH, Haftka RT (2010) Multiple tail median approach for high reliability estimation. Struct Saf 32(2):124–137

Ramu P, Thananjayan P, Acar E, Bayrak G, Park JW, Lee I (2022) A survey of machine learning techniques in structural and multidisciplinary optimization. Struct Multidisc Optim 65(9):266

Rubinstein RY, Kroese DP (2016) Simulation and the Monte Carlo method, vol 10. Wiley, New York

Sadoughi MK, Li M, Hu C, MacKenzie CA, Lee S, Eshghi AT (2018) A high-dimensional reliability analysis method for simulation-based design under uncertainty. J Mech Des. https://doi.org/10.1115/1.4039589

Smarslok BP, Haftka RT, Carraro L, Ginsbourger D (2010) Improving accuracy of failure probability estimates with separable Monte Carlo. Int J Reliab Safety 4:393–414

Song S, Lu Z, Qiao H (2009) Subset simulation for structural reliability sensitivity analysis. Reliab Eng Syst Saf 94(2):658–665

Srivastava R, Deb K (2013) An evolutionary based Bayesian design optimization approach under incomplete information. Eng Optim 45(2):141–165

Strömberg N (2021) Comparison of optimal linear, affine and convex combinations of metamodels. Eng Optim 53(4):702–718

Tokdar ST, Kass RE (2010) Importance sampling: a review. Wiley Interdiscip Rev 2(1):54–60

Tvedt L (1990) Distribution of quadratic forms in normal space—application to structural reliability. J Eng Mech 116(6):1183–1197

Van Hai T, Thu NH, Tuan HD, Van Hiu P (2020) Failure probability analysis of overhead crane bridge girders within uncertain design parameters. J Sci Technol Civil Eng (STCE)-HUCE 14(3):125–135

Viana FA, Haftka RT, Steffen V (2009) Multiple surrogates: how cross-validation errors can help us to obtain the best predictor. Struct Multidisc Optim 39(4):439–457

Wu YT (1994) Computational methods for efficient structural reliability and reliability sensitivity analysis. AIAA J 32(8):1717–1723

Wu YT, Shin Y, Sues R, Cesare M (2001). Safety-factor-based approach for probability-based design optimization. In *19th AIAA applied aerodynamics conference* (p. 1522).

Xiao NC, Yuan K, Zhou C (2020) Adaptive kriging-based efficient reliability method for structural systems with multiple failure modes and mixed variables. Comput Methods Appl Mech Eng 359:112649

Youn BD, Wang P (2008) Bayesian reliability-based design optimization using eigenvector dimension reduction (EDR) method. Struct Multidiscip Optim 36:107–123

Zhangchun T, Zhenzhou L, Wang P, Feng Z (2013) A mean extrapolation technique for high reliability analysis. Appl Math Comput 222:82–93

Zhangchun T, Zhenzhou L, Wang P (2014) Discussion on: applications of asymptotic sampling on high dimensional structural dynamic problems: MT Sichani, SRK Nielsen, and C. Bucher, Structural Safety, 33 (2011) 305–316. Struct Saf 46:8–10

Zhou Y, Lu Z (2020) An enhanced Kriging surrogate modeling technique for high-dimensional problems. Mech Syst Signal Process 140:106687

Zhou T, Peng Y, Li J (2019a) An efficient reliability method combining adaptive global metamodel and probability density evolution method. Mech Syst Signal Process 131:592–616

Zhou Y, Lu Z, Cheng K, Shi Y (2019b) An expanded sparse Bayesian learning method for polynomial chaos expansion. Mech Syst Signal Process 128:153–171