- Article type: paper
- Date text written or revised: November 27, 2020
- Number of words in our main text and tables: 4413
- Number of tables: 01
- Number of figures: 11

Title

Reliability Analysis of Structure by Iterative Sequential Sampling based Response Surface

Author 1

- Sounak Kabasi, PhD candidate
- Department of Civil Engineering, Indian Institute of Technology Bombay, Mumbai, India
- ORCID number: 0000-0001-6466-7371

Author 2

- Atin Roy, PhD candidate
- Department of Civil Engineering, Indian Institute of Engineering Science and Technology, Shibpur, Howrah, India
- ORCID number: 0000-0002-7353-3903

Author 3

- Subrata Chakraborty, Professor
- Department of Civil Engineering, Indian Institute of Engineering Science and Technology, Shibpur, Howrah, India
- ORCID number: 0000-0002-4792-3844

Corresponding author: Atin Roy, Department of Civil Engineering, Indian Institute of Engineering Science and Technology, Shibpur, Howrah, India. atin.3222@yahoo.com

Abstract

In the metamodeling based Monte Carlo simulation (MCS) framework for reliability analysis of structures, the training samples to construct response surface should be as close to the failure plane as possible to ensure sufficient accuracy in reliability estimate. For this, an algorithm based on maximin distance criterion combined with a leave-one-out cross-validation based error norm is proposed to construct a moving least squares based response surface for improved reliability estimate. The algorithm hinges on the fact that the MCS points whose predicted responses are less than the maximum absolute error obtained by the leave-one-out cross-validation approach are likely to have the maximum effect on the accuracy of the reliability estimate. Relying on this, two points are added at each iteration by ensuring that the new data points are sufficiently close to the actual limit state as well as adequately away from the existing data points. The improved reliability estimation capability of the proposed algorithm considering the direct MCS based results as the benchmark is elucidated numerically by considering two example problems.

Keywords chosen from ICE Publishing list

Risk & probability analysis; computational mechanics; mathematical modelling

List of notations

$\hat{y}(\mathbf{X})$	is the approximated response
$\mathbf{X}$	is the input vector
n	is the number of input variables
x_i	is the i-th variable
β_i, β_{ii}	are the unknown polynomial coefficients
$\mathbf{f}(\mathbf{X})$	is the vector of basis functions
$\boldsymbol{\beta}$	is the vector of unknown coefficients
$\mathbf{S}$	is the matrix of training samples
$\mathbf{y}$	is the output vector for actual responses
y^l	is the actual response at l-th training point
$\mathbf{S}^l$	is the input vector of l-th training point
$\mathbf{F}$	is the design matrix
p	is the number of training points
$\hat{\boldsymbol{\beta}}$	is the estimate for coefficients vector
SSE	is the weighted sum squared errors

w_l is the weight function corresponds to l th training point

$\mathbf{W}$ is the diagonal matrix of the weight function

d is the Euclidean distance between the points

R is the radius of the hypersphere of influence

c is a free parameter for the weight function

k is another free parameter for the weight function

$\hat{y}_{CV}^{l}$ is the response at l th point approximated by the training set of a cross-validation approach

e_{CV}^{l} is the cross-validation error magnitude at l th point

$\mathbf{e}_{CV}$ is the vector of cross-validation error magnitudes

$e_{CV}^{\max}$ is the maximum value of elements of vector, $\mathbf{e}_{CV}$

Ω is the set of candidate points for adaptive sampling

N is the number of candidate points in the set, Ω

d_{ij} is the Euclidean distance between *i*-th candidate point, $\mathbf{X}^i$ and *j*-th data point, $\mathbf{S}^j$.

D is the $N \times p$ matrix of Euclidean distances

$\left(d^{\min}\right)_i$ is the minimum value of the *i*-th row of the matrix D

$P_{f,i}$ is the probability of failure estimated at i th iteration

σ_{eq} is the equivalent Von-Misses stress

p_0 is the internal pressure

r_0 , r_1 are the internal and external radius

f_y is the yield stress

g is the limit state function

E is the Young's modulus of the material of the space-dome truss

A_1 is the section area of the top radial bars of the space-dome truss

A_2 is the section area of the peripheral bars of the space-dome truss

A_3 is the section area of the bottom inclined bars of the space-dome truss

P_1 is the load at the centre node of the space-dome truss

P_2 is the load at each of the six nodes of the middle hexagon of the space-dome truss

Δ_{P1}^{z} is the maximum vertical displacement of the node under load P_1

Δ_{allow} is the allowable maximum displacement of the top node of the space-dome truss

1. Introduction

The uncertainty is introduced in a structure due to uncertainty in the parameters required to characterize the structure e.g. geometry, material properties, boundary conditions, etc., and loads acting on it. Structural reliability analysis (SRA) is a theoretical framework that accounts for the effect of such parameter uncertainties. The primary task of SRA is to estimate the probability of failure which requires the computation of a multidimensional integral over the unsafe domain involving the joint probability distribution function (PDF) of the related random input variables. The joint PDF of the input variables is rarely available. Moreover, the exact computation of such integral is often computationally demanding. Several methods have been developed to estimate the probability of failure which can be classified into two groups: approximate analytical techniques and statistical simulation-based methods (Ditlevsen and Madsen, 1996; Haldar and Mahadevan, 2000; Kwon and Elnashai, 2006; Faravelli, 1989). Analytical approximate techniques are based on the second moment method which suffers criticism with regard to the accuracy of the estimated reliability index. Moreover, such approaches require evaluating not only the limit state function (LSF) but also its gradients during iterations which are computationally challenging in the case of implicit LSF of a large complex structure. The most accurate and conceptually straightforward means of SRA is based on the Monte Carlo simulation (MCS) technique (Shinozuka, 1983; Melchers, 1999; Dymiotis et al., 1999; Au and Beck, 2001). The approach is preferred as it does not require an assumption about the shape of the failure surface. However, such a full simulation approach requires a large number of repetitive evaluations of LSF for an acceptable confidence in the reliability estimates. If the performance function is explicitly available in closed-form, the numbers of performance function calls do not play an important role. However, the performance behaviour of real large complex structures is usually defined by LSF in implicit form as an explicit form of LSF is often unavailable. Although reliability analysis of structures involving implicit LSF can be performed by direct MCS technique, each performance function evaluation typically requires analysis of a finite element model involving high computational cost, especially for large complex structures (Ditlevsen and Madsen, 1996; Haldar and Mahadevan, 2000). The computational involvement for the analysis of structures to extract necessary responses for statistical analysis is studied by Kwon and Elnashai (2006). The number of simulations required may be in the order of several thousand for an acceptable estimate of reliability, depending on the function being evaluated and the magnitude of the probability of failure (Faravelli, 1989). As an effective solution to such problems, the polynomial response surface method (RSM) based metamodeling approach is widely used to overcome the computational challenges of MCS based reliability analysis of large complex structures involving implicit LSF (Faravelli,1989; Bucher and Bourgund, 1990; Liu and Moses, 1994; Rajashekhar and Ellingwood, 1993). The RSM simplifies the simulation process by fitting a response surface model to approximately replace the implicit LSF. The response surface is an approximated polynomial function of random variables. The coefficients of each term in the polynomial could be obtained by the least squares method (LSM) using the actual structural responses at a lesser but sufficient number of times. Thereby, the computational involvement of analysis can be reduced drastically. The applications of such LSM based RSM in reliability analysis of structures are based on global approximation of scatter position data. However, the LSM is one of the major sources of error in prediction by the RSM and needs a considerable number of training points to ensure necessary accuracy. Thus, it becomes computationally intensive for practical engineering problems involving too many variables.

To improve the efficiency and accuracy of RSM, the application of various adaptive metamodeling approaches e.g. polynomial based moving least square method (MLSM) (Kim et al., 2005; Kang et al., 2010), artificial neural network (ANN) (Elhewy et al., 2006; Lagaros et al., 2009), Kriging (Kaymaz, 2005), polynomial chaos expansion (Blatman and Sudret, 2011), support vector machines (SVM) (Li et al., 2006; Ghosh et al., 2018; Roy et al., 2019), etc. are notable. The MLSM based RSM is noted to be the simplest amongst these. In this regard, it is important to note that for accurate estimation of reliability, the DOE points should be as close to the LSF as possible. Based on this fact, Bucher and Bourgound (1990) proposed a sampling method to construct an improved LSM based response surface for SRA. The approach is further modified by Rajashekhar and Ellingwood (1993) for the iterative improvement of RSM. Such iterative improvements are not only limited to LSM based RSM (Farag and Haldar, 2016; Gaxiola-Camacho et al., 2017) but also studied in the framework of adaptive metamodeling approaches e.g. by MLSM (Goswami et al., 2016) and SVM (Richard et al., 2012) based metamodels. However, these approaches require the complete replacement of DOE sets after each iteration step. The number of iterations required may be large in many cases and the replacement of DOE sets may demand much computational time especially when the LSF is implicit and highly non-linear. Thereby, such approaches reduce the efficiency of the RSM significantly. The iterative improvement by adding a new data point to the existing training dataset after each iteration step to improve reliability estimate is appealing in this regard (Echard et al., 2011). The approach selects the adaptive sampling points based on the uncertainty in the prediction and the magnitude of predicted LSF by the Kriging method. But, such an adaptive sampling procedure can be employed in metamodeling only when the prediction variance is available. However, most of the metamodels not being Gaussian process based regression are unable to provide the prediction variance. Sequential adaptive sampling for such metamodels to improve MCS based SRA seems to be attractive in this regard. Roussouly et al. (2013) attempted a sequential adaptive sampling for polynomial RSM by searching a hypercube with the most probable failure point as its centre and the hypercube is considered as reduced space. A similar sequential adaptive sampling is proposed by Guimarães et al. (2018) with little modifications. However, reduced space with a regular hypercube shape includes a huge amount of unimportant regions. Xiao et al. (2018a) developed three learning functions for selecting the most suitable sample point at each step of iteration for all types of metamodels. However, the relative weights for these learning functions are heuristic and need experiences. Xiao et al. (2018b) proposed another learning function to select sequential training samples combining the cross-validation method, weighted Euclidean-distance, and the weights of sample qualities in the input parameter space. The cross-validation method is employed to estimate the average probabilistic classification error function on the candidate sample point. But, the probabilistic classification function requires both the prediction and its variance at the candidate sample point. Thus, the application of this learning technique is also limited to Kriging or other Gaussian process based regression methods. Recently, Roy and Chakraborty (2020) developed a sequential sampling for SVR based on the maximin distance criterion.

In the present study, an algorithm based on maximin distance criterion combined with a leave-one-out cross-validation based error norm is proposed to construct a moving least squares based response surface for improved estimate of reliability. Specifically, the proposed algorithm attempts to select training samples to construct a response surface as close to the failure plane as possible to ensure sufficient accuracy in the reliability estimate. It hinges on the fact that the MCS points whose predicted responses are less than the maximum absolute error obtained by the leave-one-out cross-validation approach are likely to have the maximum effect on the accuracy of the reliability estimate. In this regard, it can be realized that the misclassification of a point can only occur if the error in approximating an LSF at any sample point is more than the magnitude (irrespective of the sign) of the approximated LSF. Hence, it can be intuitively anticipated that the points corresponding to a magnitude of approximated LSF less than the value of the noted absolute error obtained by the leave-one-out cross-validation approach are most likely to get misclassified. Hence, the accuracies of MLSM based metamodel in approximating the LSF at these points are of paramount interest for an improved estimate of the probability of failure. Thus, a specific error norm based on the leave-one-out cross-validation is utilized to decide a reduced input space that is concise and contains only the important regions. Relying on this, two points are added at each iteration from the reduced space by ensuring that the new data points are sufficiently close to the actual limit state as well as adequately away from the existing data points to avoid clustering effect. The improved reliability estimation capability of the proposed algorithm considering the direct MCS based results as the benchmark is elucidated numerically by considering two example problems.

2. Response surface methods

The present study deals with the sequential updating of MLSM based RSM for improved estimates of the reliability of structures. Thus, to explain the proposed algorithm, the fundamentals of the usual LSM and MLSM based response surfaces are first discussed in this section for an effective presentation of the proposed algorithm in the subsequent section.

2.1 Least Squares Method Based RSM

The RSM is used to uncover an unknown analytical relationship (an empirical model) between several inputs and outputs. The reduced quadratic polynomial model (without cross-terms) is mostly employed to replace the unknown LSF for reliability analysis due to its trade-off between simplicity and accuracy (Wong et al., 2005). The reduced quadratic polynomial RSM can be expressed as,

$$\hat{y}(\mathbf{X}) = \beta_0 + \sum_{i=1}^{n} \beta_i x_i + \sum_{i=1}^{n} \beta_{ii} x_i^{\,2} = \mathbf{f}(\mathbf{X})^{\mathrm{T}} \boldsymbol{\beta}$$

1.

Where, $\hat{y}(\mathbf{X})$ is the approximated response for input vector $\mathbf{X}$ consisting of n number of variables; x_i denotes the i th variable; β_0 , β_i and β_{ii} are unknown polynomial coefficients; $\mathbf{f}(\mathbf{X}) = \{1, \cdots, x_i, \cdots, x_i^{\,2}, \cdots\}^{\mathrm{T}}$ is the vector of basis functions; and $\boldsymbol{\beta} = \{\beta_0, \cdots, \beta_i, \cdots, \beta_{ii}, \cdots\}^{\mathrm{T}}$ is the vector of $2n+1$ number of unknown coefficients. The unknown coefficients are obtained by the usual LSM. The training data is obtained by constructing a DOE and evaluating corresponding actual responses of structure. The LSM minimizes the sum squared error (SSE) at all the sample data points considered for training to estimate the polynomial coefficients. The error norm, SSE can be expressed as,

$$SSE = \sum_{l=1}^{p} \left(y^l - \beta_0 - \sum_{i=1}^{n} \beta_i x_i^l - \sum_{i=1}^{n} \beta_{ii} \left(x_{ii}^l \right)^2 \right)^2 = (\mathbf{y} - \mathbf{F}\boldsymbol{\beta})^T (\mathbf{y} - \mathbf{F}\boldsymbol{\beta})$$

2.

Where, $\mathbf{y} = \{y^1 \cdots y^l \cdots y^p\}^T$ is the output vector of actual responses at p number of training data points, $\mathbf{S} = \{\mathbf{S}^1 \cdots \mathbf{S}^l \cdots \mathbf{S}^p\}$; $\mathbf{F}$ is the design matrix composed of p number of rows containing $\mathbf{f}(\mathbf{X})^T$ for each training points. The LSM based estimate of coefficients vector, $\hat{\boldsymbol{\beta}}$ is obtained as follows,

$$\hat{\boldsymbol{\beta}} = \left(\mathbf{F}^T \mathbf{F}\right)^{-1} \mathbf{F}^T$$

3.

Once the vector $\boldsymbol{\beta}$ is obtained, the LSF at any point from the input space can be approximated by Equation 1.

2.2 Moving Least Squares Method Based RSM

The MLSM based RSM is a weighted LSM having various weights with respect to the position of approximation. Thus, the coefficients of response surface function change with the change of approximation point of interest. This procedure is interpreted as a local approximation. To fit the polynomial function to scattered data, the LSM evaluates the unknown coefficients of the function by minimizing the SSE. However, in the MLSM approach, the SSE is defined as the sum of the weighted errors as following,

$$SSE(\mathbf{X}) = \sum_{l=1}^{p} w_l \left(y^l - \beta_0 - \sum_{i=1}^{n} \beta_i x_i^l - \sum_{i=1}^{n} \beta_{ii} \left(x_{ii}^l \right)^2 \right)^2 = \left(\mathbf{y} - \mathbf{F}\boldsymbol{\beta}\right)^T \mathbf{W}\left(\mathbf{X}\right)\left(\mathbf{y} - \mathbf{F}\boldsymbol{\beta}\right)$$

4.

Where, $SSE(\mathbf{X})$ is the weighted sum square errors at training points which depend on the location of the point of interest, $\mathbf{X}$; w_l is the weight corresponds to l-th training point; $\mathbf{W}\left(\mathbf{X}\right)$ is the diagonal matrix of the weight function for point of interest $\mathbf{X}$. The weight matrix $\mathbf{W}\left(\mathbf{X}\right)$ is constructed by using the weighting function in the diagonal terms. It may be obtained by utilizing weight functions such as constant, linear, quadratic, higher-order polynomials, exponential functions, etc. (Kim et al., 2005; Kang et al., 2010). In the present study, the weight matrix, $\mathbf{W}\left(\mathbf{X}\right)$ is obtained by utilizing a Gaussian weighting function of the following form,

$$w(d) = \frac{e^{-\left(\frac{d}{cR}\right)^{2k}} - e^{-\left(\frac{1}{c}\right)^{2k}}}{1 - e^{-\left(\frac{1}{c}\right)^{2k}}}$$

5.

Where, $d = \left\| \mathbf{X} - \mathbf{S}^l \right\|$ is the Euclidean distance between the point of interest, $\mathbf{X}$ and l-th training point $\mathbf{S}^l$; R is the approximate radius of hypersphere of influence for $\mathbf{X}$. In Equation 5, the parameter, k is taken as unity to ensure the Gaussian nature of the weighting function. The value of the parameter, c is taken as 0.4 (Taflanidis and Cheung, 2012). It can be noted from Equation 5 that the weight associated with a particular sampling point $\mathbf{S}^l$ decays as the point $\mathbf{X}$ moves away from $\mathbf{S}^l$. The weighting function has its maximum value of 1.0 at a normalized distance, d/R = 0, and a minimum value of 0.0 outside the influence hypersphere, i.e. $w\left(d/R > 1\right) = 0.0$. The function value decreases smoothly from 1.0 to 0.0. The value of R is so chosen in order to secure a sufficient number of neighbouring experimental points to avoid singularity. To calculate the weights using Equation 5, the radius of hypersphere of influence, R is taken as the distance between $\mathbf{X}$ and the farthest training point from it. Now, the coefficient vector $\boldsymbol{\beta}\left(\mathbf{X}\right)$ can be obtained as below,

$$\hat{\boldsymbol{\beta}}(\mathbf{X}) = \left(\mathbf{F}^T \mathbf{W}(\mathbf{X})\, \mathbf{F}\right)^{-1} \mathbf{F}^T \mathbf{W}(\mathbf{X})\, \mathbf{y}$$

6.

It can be noted from the above equation that the coefficient $\boldsymbol{\beta}\left(\mathbf{X}\right)$ is a function of the location or position of $\mathbf{X}$. Thus, the procedure to calculate $\boldsymbol{\beta}\left(\mathbf{X}\right)$ being a local approximation and moving processes performs a global approximation over the entire design domain.

3. The proposed adaptive MSLM algorithm for SRA

The present study is intended to explore an algorithm based on maximin distance criterion combined with a leave-one-out cross-validation based error norm to construct a MLSM based response surface model for improved reliability estimate. The related formulation is presented in this section.

The algorithm starts with an initial DOE. As the failure plane is not known a priori, training samples for the initial DOE should be uniformly selected from the entire input space. Thus, a single-shot DOE is obtained by Latin hypercube sampling (LHS) considering the upper and lower boundaries of all the related random variables. The algorithm is hinged on the fact that the training points close to the failure plane should be added to the DOE for the improved accuracy of approximation of the LSF close to the failure region. This, in turn, is expected to improve the accuracy of estimated reliability by the RSM based MCS approach. Once, an approximate LSF is constructed by the MLSM based on the initial DOE, a reduced domain containing candidates for adaptive samples can be identified. To provide a measure of fit of a model to a data set, the cross-validation approach used in regression analysis is frequently applied (Kohavi, 1995; Xiao et al., 2018b; Roy et al., 2019; Roy and Chakraborty, 2020). A fitted model having been constructed, each data, in turn, is held out and the model is reconstructed using the remaining data. The process of evaluating error at each training point by excluding that point from the entire training data set to build a model based on the remaining data is referred to as the leave-one-out cross-validation approach. Let, the entire training data set is represented by $\{\mathbf{S}, \mathbf{y}\}$ with p number of training points, $\mathbf{S} = \{\mathbf{S}^1 \cdots \mathbf{S}^l \cdots \mathbf{S}^p\}$ and corresponding actual responses, $\mathbf{y} = \{y^1 \cdots y^l \cdots y^p\}^T$. In the leave-one-out cross-validation approach, l-th data pair is held out $\{\mathbf{S}^l, y^l\}$ from the training data set $\{\mathbf{S}, \mathbf{y}\}$ and a metamodel is constructed with remaining data to predict the response at $\mathbf{S}^l$ (say, $\hat{y}_{CV}^l$ is the approximated response). The cross-validation error magnitude (i.e., absolute value), e_{CV}^l at $\mathbf{S}^l$ can be expressed as, $e_{CV}^l = \left| y^l - \hat{y}_{CV}^l \right|$. Thus, a vector of cross-validation error magnitudes, $\mathbf{e}_{CV} = \{e_{CV}^1 \cdots e_{CV}^l \cdots e_{CV}^p\}^T$ at all training points is obtained. The maximum absolute error at data points obtained by the leave-one-out cross-validation approach is the maximum value of elements of vector, $\mathbf{e}_{CV}$ and can be expressed as, $e_{CV}^{\max} = \max\{\mathbf{e}_{CV}\}$. In this regard, it is noted that the cross-validation error magnitude is usually higher than the magnitude of actual error in prediction by the metamodel. Intuitively, it is unlikely that the actual prediction error magnitude at any MCS point will be greater than $e_{CV}^{\max}$. Thereby, the sign of LSF can only be misrecognized if the error magnitude is higher than the magnitude of approximated LSF. Hence, it can be assumed that the MCS points for which the predicted absolute values of the approximated LSF are less than $e_{CV}^{\max}$ are the most probable to be on the other side of the actual limit state as compared to that predicted by the metamodel. Such points are expected to have the maximum effect on the accuracy of reliability estimates by a metamodel. Thus, the accuracy of the predictions at these MCS points should be improved. Therefore, these MCS points are considered as candidates for adaptive sampling in the proposed algorithm. It is desirable to include a minimum number of adaptive training points into the DOE for efficiency of the metamodeling approach. To effectively fill the sub-domain with adaptive samples avoiding clustering, the minimum inter-distance of the training samples can be maximized similar to the widely used maximin distance criterion (Johnson et al., 1990). If *p* number of data points are already in the DOE, then to add (*p*+1)-th data point, first a set of candidate points, Ω is selected. A $N \times p$ matrix can be obtained as,

$$D = \begin{pmatrix} d_{11} & \cdots & d_{1p} \\ \vdots & \ddots & \vdots \\ d_{N1} & \cdots & d_{Np} \end{pmatrix}, \qquad d_{ij} = \left\| \mathbf{X}^i - \mathbf{S}^j \right\|$$

7.

Where, *N* is the number of candidate points in the set, Ω and d_{ij} is the Euclidean distance between the *i*-th candidate point, $\mathbf{X}^i$ and the *j*-th data point, $\mathbf{S}^j$. The minimum value of the *i*-th row of the matrix D i.e. $\left(d^{\min}\right)_i = \min\{d_{i1} \cdots d_{ip}\}$ is the Euclidean distance between *i*-th candidate point and the nearest data point. The candidate point having the maximum $d^{\min}$ value is the next best choice for the new data point and it is included in the DOE. At each iteration step, two new data points are added. The MCS points which are situated in the approximated unsafe domain having the magnitude of the approximated LSF less than the value of $e_{CV}^{\max}$ are considered as the candidate sample to add as the first point. Based on the mentioned maximin distance criterion, one point is added from this domain to the existing DOE. To add the next point, the MCS points having the magnitude of the approximated LSF less than the value of $e_{CV}^{\max}$ are selected from the approximated safe domain as the candidate sample. Thus, the latest two adaptive samples (one having the predicted LSF greater than zero and another one less than zero) are included in the DOE. The actual responses corresponding to the two new points are evaluated to update the training dataset. Subsequently, the approximation of LSF by the MLSM is updated with the current MLSM based response surface. Now, a new reduced domain of candidate samples is obtained based on the current approximation of LSF and modified error norm obtained by leave-one-out cross-validation approach. Subsequently, two new points can be included in the DOE. This enrichment of DOE is continued until a suitable stopping criterion is satisfied. In the present study, the stopping criterion is decided based on the convergence in the estimated probability of failure as follows,

$$\left| P_{f,i-1} - P_{f,i} \right| \Big/ P_{f,i} \leq 0.05$$

8.

Where, $P_{f,i-1}$ and $P_{f,i}$ are the probability of failure estimated at $i-1$ and i-th iteration, respectively.

The algorithm is summarized as below:

Step1: An initial single-shot DOE is constructed by LHS over the entire physical domain of the input variables.

Step2: The metamodel is trained by all the data points of the DOE.

Step3: Using the leave-one-out cross-validation approach, note the absolute maximum error, $e_{CV}^{\max}$.

Step4: The LSF at all the MCS points are predicted by the metamodel to estimate the probability of failure.

Step5: The reduced space is built by the MCS points having the absolute value of the predicted response less than $e_{CV}^{\max}$.

Step6: Two points from the reduced space are selected based on the maximin distance criterion, each one on either side of the approximate limit state.

Step7: The actual responses corresponding to the new points are evaluated and these two data points are added to update the DOE set.

Step8: Go to the **Step2** till convergence of probability of failure is observed.

Step9: The final reliability index is calculated corresponding to the converged probability of failure.

A flow chart explaining the implementation of the proposed algorithm is described in Figure 1.

4. Numerical Study

The effectiveness of the proposed adaptive MLSM approach based on sequential updating of the training data set is elucidated numerically by considering two examples which are presented in this section.

4.1 Example 1: Sphere subjected to internal pressure

This first example is to approximate the Von-Mises stress of a hollow sphere under internal pressure as shown in Figure 2. The material is homogenous without spatial variety. The LSF is defined based on the equivalent Von-Misses stress (σ_{eq}) and can be described as,

$$g = f_y - \sigma_{eq}; \quad \sigma_{eq} = \frac{3p_0}{2} \times \frac{r_1^3}{r_1^3 - r_0^3}$$

9.

In the above equation, f_y is the yield stress of the material of the hollow sphere; p_0 is the internal pressure, r_0 and r_1 are the internal and external radii, respectively. p_0, r_0 and r_1 are assumed to follow lognormal distribution having the mean values of 1200 MPa, 50mm, 100 mm, respectively. The coefficients of variations (COV) of all the random parameters are assumed to be 7.5%. The physical limits of the random variables are taken as mean ± 3×standard deviation (SD).

To study the effectiveness of the proposed adaptive MLSM scheme, the response of the sphere under the internal pressure is approximated by the proposed MLSM based response surface. To perform the MCS study, 10^5 numbers of random samples are generated according to the assumed PDF for each variable. The reliability results are obtained for varying f_y. To study the effect of size of initial DOE, the reliability results are obtained by the proposed MLSM based response surface starting with three different sizes of initial DOE composed of 11, 14, and 17 samples and results are shown in Figures 3, 4, and 5, respectively. The results obtained by the MLSM based metamodel using the initial DOE and those obtained from the proposed MLSM based metamodel with converged DOE after necessary iterations are denoted as MLSM-adaptive_initial and MLSM-adaptive_final, respectively. For a meaningful comparison, the reliability results are also obtained by the MLSM trained by an equivalent single-shot DOE which is composed of the same number of training points as has been used to obtain the final converged results by the proposed adaptive MLSM (denoted as MLSM-Equivalent-DOE). The results of the most accurate direct MCS (denoted as DMCS) are considered as the benchmark. It can be readily noted that the reliability results obtained by the proposed approach are much better than those obtained by the equivalent single-shot DOE with the same numbers of training points as required by the proposed approach. It is important to note that the improved performance of the proposed MLSM based approach is observed for all the DOE configurations which show the robustness of the proposed algorithm. In Figure 6, the absolute percentage errors in estimating reliability indices are shown for each step of the proposed MLSM approach (denoted as MLSM-adaptive_*i* for *i*-th step) starting with 11 initial training samples and the corresponding MLSM-Equivalent-DOE. It is observed that the errors reduce drastically after the first iteration by the proposed MLSM based metamodeling approach. The improved capability of reliability estimate by the proposed approach with only 4 iterations can be readily noted from the plot. The final estimate of reliability by the proposed approach is very close to that obtained by the direct MCS technique with a deviation of around 1%-2%. Though the comparisons of the absolute percentage of errors for the other two initial DOE are not shown for brevity, the observations are found to be similar.

4.2 Example 2: A space dome truss

A space dome truss problem is taken as the second example involving implicit LSF (Keshtegar, 2017). The Young's modulus (E) of the material of all the bars, the section area of the top radial bar numbers 1-6 (A_1), the peripheral bar numbers 7-12 (A_2), the bottom inclined bar numbers 13-24 (A_3), the point load P_1 at the centre node and the point load P_2 at each of the six nodes of the middle hexagon (as shown in Figure 6) are considered as the six independent random variables and statistical properties of those are furnished in Table 1. The implicit LSF is defined as,

$$g = \Delta_{\text{allow}} - \left|\Delta_{\text{P1}}^{z}\right| \tag{10}$$

where, Δ_{P1}^{z} is the maximum vertical displacement of the node under load P_1 and Δ_{allow} is the allowable maximum displacement. ANSYS mechanical APDL module is employed to obtain the maximum displacement Δ_{P1}^{z} which is necessary for evaluating the LSF.

Like the previous example, three initial DOE consists of 20, 25, and 30 training data points are constructed within the physical domain (mean ± 0.3×mean) of the random variables. For MCS, 10^5 numbers of random simulation samples are generated. The reliability results for different Δ_{allow} values estimated by the proposed MLSM based metamodels starting with 20, 25, and 30 initial samples are compared in Figures 8, 9, and 10, respectively. The absolute percentage error in obtaining the reliability results for varying Δ_{allow} when compared with the direct MCS by using the proposed adaptive approach with 30 initial data after each iteration and the corresponding equivalent single-shot MLSM based reliability results are shown in Figure 11. Similar observations are noted for this example also. It is observed that the errors are drastically reduced by the proposed MLSM approach even after the first enrichment of the DOE for this example also.

5. Summary and conclusions

A new adaptive MLSM based response surface methodology based on maximin space-filling design criterion combined with a leave-one-out cross-validation based error norm is proposed for improved reliability estimate of structure. The algorithm relies on two points being added after each iteration. The maximin distance criterion combined with a cross-validation based specific error norm ensures that the new data points added sequentially are sufficiently close to the actual limit state and at the same time sufficiently away from the existing data points. The algorithm proposes to add one new data from the safe and another from the unsafe domain to reduce the bias in response approximation near the limit state. The results of the numerical study of both examples reveal improved reliability estimation capability of the proposed algorithm with regard to the conventional approach considering the direct MCS based results as the benchmark solution. The improved capability of estimating reliability by the algorithm with different sets of training data clearly reveals the robustness of the proposed approach. It is worth mentioning here that for satisfying both the criteria, the algorithm does not use any heuristic weighting scheme rather uses a step-by-step approach. Thus, the algorithm can be applied generically to any reliability analysis problem. Though the present algorithm is tested for MLSM based adaptive metamodel, it can be readily applied for other adaptive metamodeling approaches e.g. Kriging, ANN, SVM, etc. However, this needs further study.

References

Au SK and Beck JL (2001) Estimation of small failure probabilities in high dimensions by subset simulation. Probabilistic Engineering Mechanics **16(4):** 263–277.

Blatman G and Sudret B (2011) Adaptive sparse polynomial chaos expansion based on least angle regression. Journal of Computational Physics **230(6)**: 2345–2367.

Bucher CG and Bourgund U (1990) A fast and efficient response surface approach for structural reliability problems. Structural Safety **7(1)**, 57–66.

Ditlevsen O, and Madsen HO, (1996) *Structural Reliability Methods*. John Wiley and Sons Ltd., Chichester, UK.

Dymiotis C, Kappos AJ and Chryssanthopoulos MK (1999) Seismic reliability assessment of RC frames with uncertain drift and member capacity. Journal of Structural Engineering ASCE **125(9)**.

Echard B, Gayton N and Lemaire M (2011) AK-MCS: An active learning reliability method combining Kriging and Monte Carlo Simulation. Structural Safety **33(2)**: 145–154.

Elhewy AH, Mesbahi E and Pu Y (2006) Reliability analysis of structures using neural network method. Probabilistic Engineering Mechanics **21(1)**: 44–53.

Farag R and Haldar A (2016) A novel reliability evaluation method for large engineering systems. Ain Shams Engineering Journal **7(2)**: 613–625.

Faravelli L (1989) Response-Surface Approach for Reliability Analysis. Journal of Engineering Mechanics **ASCE 115(12)**: 2763–2781.

Gaxiola-Camacho JR, Azizsoltani H, Villegas-Mercado FJ and Haldar A (2017) A novel reliability technique for implementation of Performance-Based Seismic Design of structures. Engineering Structures **142**: 137–147.

Ghosh S, Roy A and Chakraborty S (2018) Support vector regression based metamodeling for seismic reliability analysis of structures. Applied Mathematical Modelling **64**: 584–602.

Goswami S, Ghosh S and Chakraborty S (2016) Reliability analysis of structures by iterative improved response surface method. Structural Safety **60**: 56–66.

Guimarães, H, Matos JC and Henriques AA (2018) An innovative adaptive sparse response surface method for structural reliability analysis. Structural Safety **73**: 12–28. doi:10.1016/j.strusafe.2018.02.001

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

Johnson ME, Moore LM and Ylvisaker D (1990) Minimax and maximin distance designs. Journal of Statistical Planning and Inference **26(2)**: 131–148.

Kang SC, Koh HM and Choo, JF (2010) An efficient response surface method using moving least squares approximation for structural reliability analysis. Probabilistic Engineering Mechanics **25(4)**: 365–371.

Kaymaz, I (2005) Application of kriging method to structural reliability problems," Structural Safety **27(2)**: 133-151.

Keshtegar B (2017) A hybrid conjugate finite-step length method for robust and efficient reliability analysis. Applied Mathematical Modelling **45**: 226-237.

Kim C, Wang S, and Choi KK (2005) Efficient Response Surface Modeling by Using Moving Least-Squares Method and Sensitivity. AIAA Journal **43(11)**: 2404–2411.

Kohavi R (1995) A study of cross-validation and bootstrap for accuracy estimation and model selection, In: *Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence*- Vol. 2 (Mellish CS (ed.)), Montréal, Québec, Canada.: Morgan Kaufmann Publishers Inc. San Francisco, CA, USA , pp. 1137–1145.

Kwon OS and Elnashai A (2006) The effect of material and ground motion uncertainty on the seismic vulnerability curves of RC structure. Engineering structures **28(2):** 289-303.

Lagaros ND, Tsompanakis Y, Psarropoulos PN and Georgopoulos EC (2009) Computationally efficient seismic fragility analysis of geostructures. Computers & Structures **87(19–20)**: 1195–1203.

Li H, Lü Z and Yue Z (2006) Support vector machine for structural reliability analysis. Applied Mathematics and Mechanics **27(10)**: 1295–1303.

Liu YW and Moses F (1994) A sequential response surface method and its application in the reliability analysis of aircraft structural systems. Structural Safety **16**:39-46.

Melchers RE (1999) *Structural reliability analysis and prediction*. John Wiley and Sons Limited, Chichester, West Sussex, England, second edition.

Rajashekhar MR and Ellingwood BR (1993) A new look at the response surface approach for reliability analysis. Structural Safety **12(3)**: 205–220.

Richard B, Cremona C and Adelaide L (2012) A response surface method based on support vector machines trained with an adaptive experimental design. Structural Safety **39**: 14–21.

Roussouly N, Petitjean F and Salaun M (2013) A new adaptive response surface method for reliability analysis. Probabilistic Engineering Mechanics **32**: 103–115.

Roy A and Chakraborty S (2020) Support vector regression based metamodel by sequential adaptive sampling for reliability analysis of structures. Reliability Engineering & System Safety **200**: 106948.

Roy A, Manna R and Chakraborty S (2019) Support vector regression based metamodeling for structural reliability analysis. Probabilistic Engineering Mechanics **55**: 78–89.

Shinozuka M (1983) Basic analysis of structural safety. Journal of Structural Engineering ASCE **109(3)**: 721-740.

Taflanidis AA and Cheung SH (2012) Stochastic sampling using moving least squares response surface approximations. Probabilistic Engineering Mechanics **28**: 216–224.

Wong SM, Hobbs RE and Onof C (2005) An adaptive response surface method for reliability analysis of structures with multiple loading sequences. Structural Safety **27**, 287–308.

Xiao NC, Zuo MJ and Zhou C (2018a) A new adaptive sequential sampling method to construct surrogate models for efficient reliability analysis. Reliability Engineering and System Safety **169**: 330–338.

Xiao NC, Zuo MJ and Guo W (2018b) Efficient reliability analysis based on adaptive sequential sampling design and cross-validation. Applied Mathematical Modelling **58**: 404–420.

Table 1: Details of the random variables of the space dome truss

Random Variables	Probability distribution		
	Type	Mean	COV
A_1	Normal	0.013 m^2	0.1
A_2	Normal	0.01 m^2	0.1
A_3	Normal	0.016 m^2	0.1
E	Normal	205 GPa	0.05
P_1	Gumbel Max.	20 kN	0.15
P_2	Gumbel Max.	10 kN	0.12

Figure captions

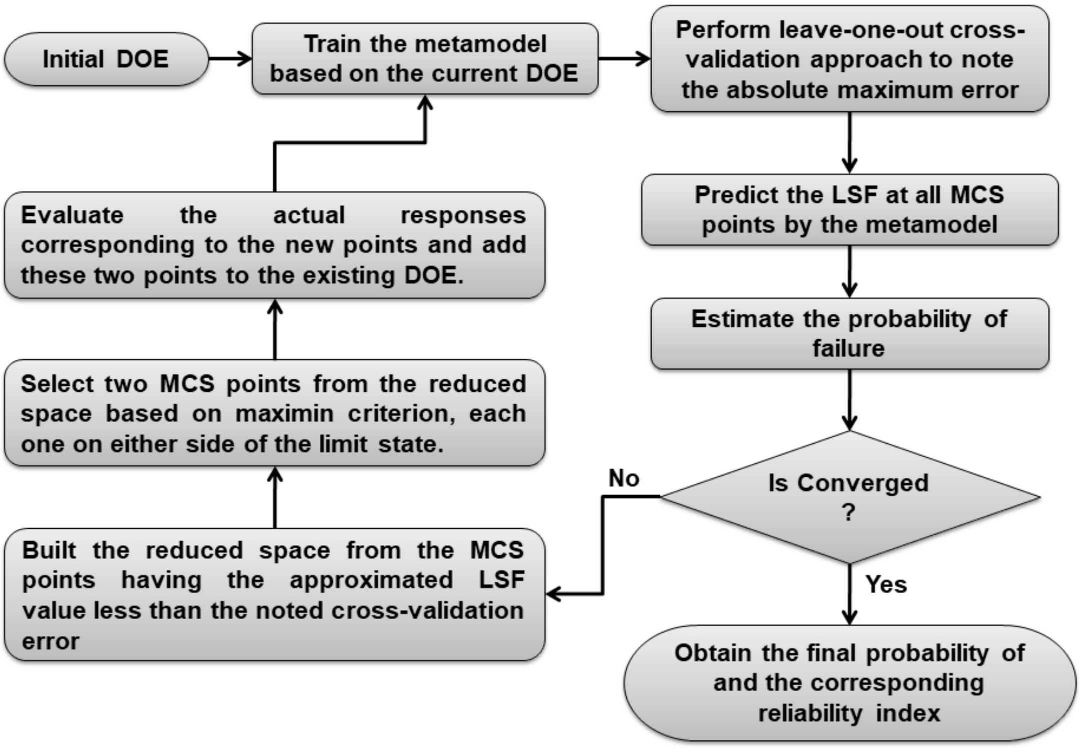


Figure 1. A flow chart of the proposed algorithm

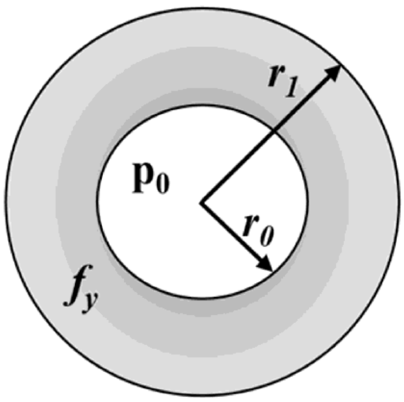


Figure 2. Schematic diagram of the hollow sphere under internal pressure

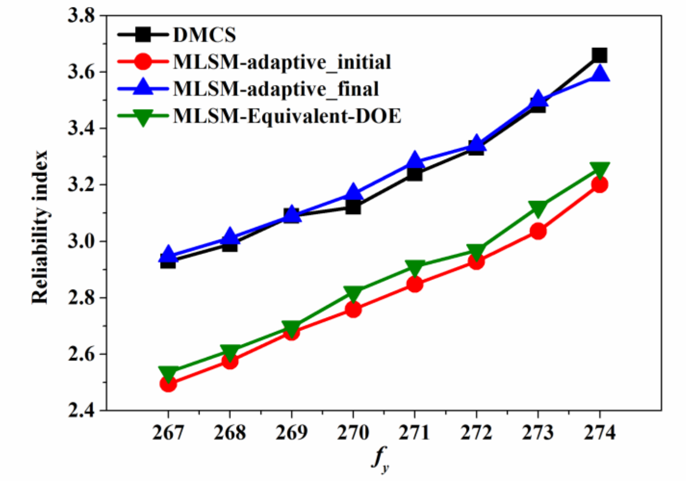


Figure 3. The comparison of estimated reliability indices for varying f_y considering 11 initial training data.

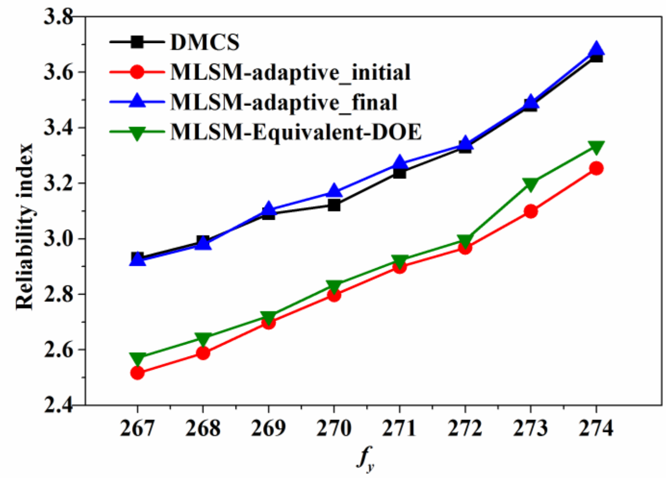


Figure 4. The comparison of estimated reliability indices for varying f_y considering 14 initial training data.

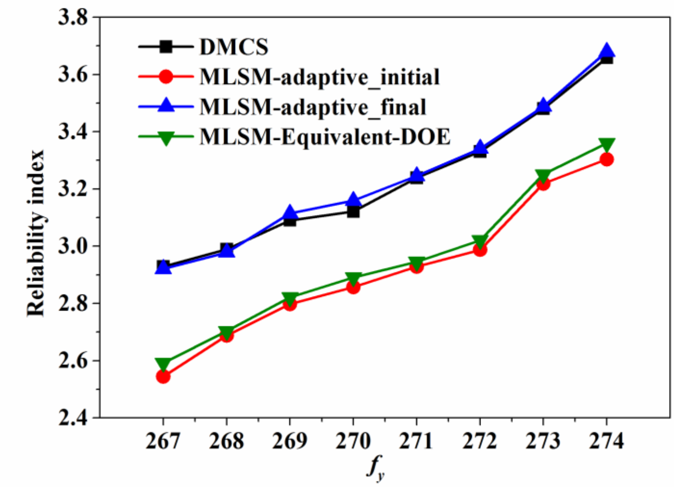


Figure 5. The comparison of estimated reliability indices for varying f_y considering 17 initial training data.

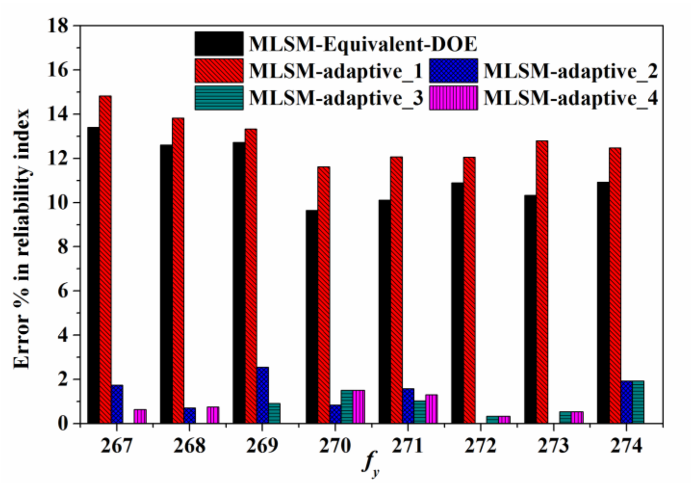


Figure 6. The comparison of absolute percentage error in estimating reliability indices for each iteration step of the proposed MLSM approach considering11 initial training data for varying f_y

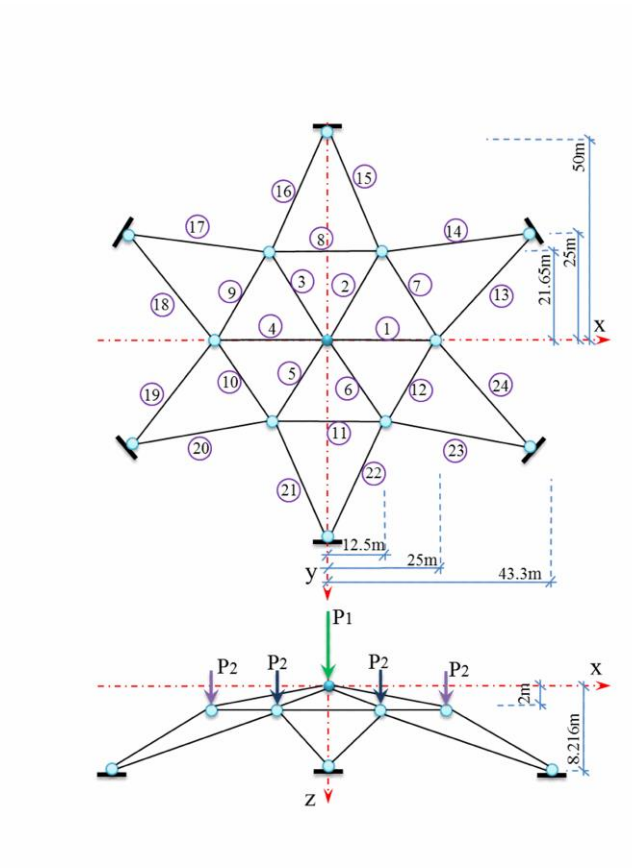


Figure 7. Schematic diagram of the space dome truss (Redrawn from Keshtegar, 2017)

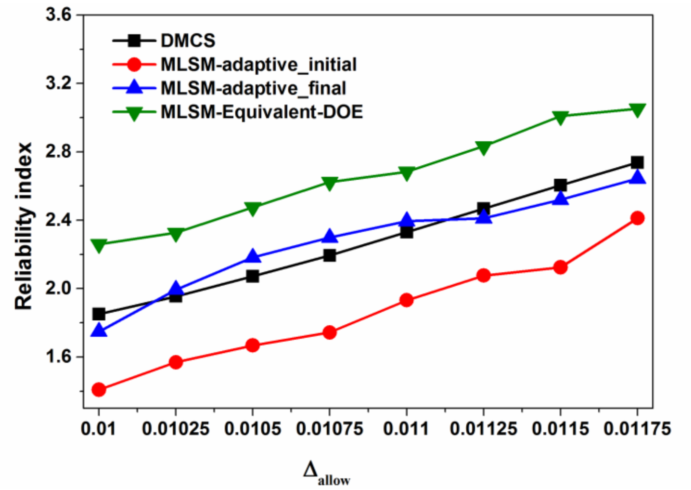


Figure 8. The comparison of reliability indices for different Δ_{allow} considering 20 initial training data.

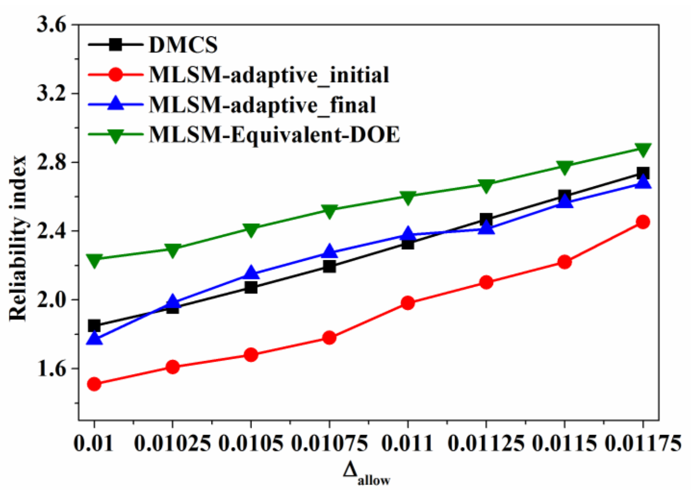


Figure 9. The comparison of reliability indices for different Δ_{allow} considering 25 initial training data.

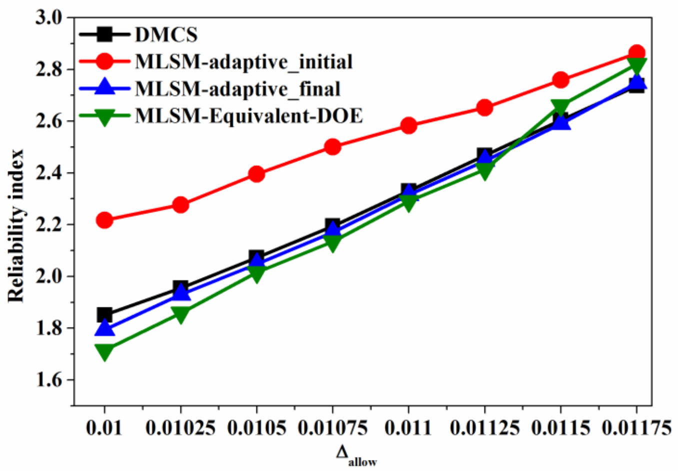
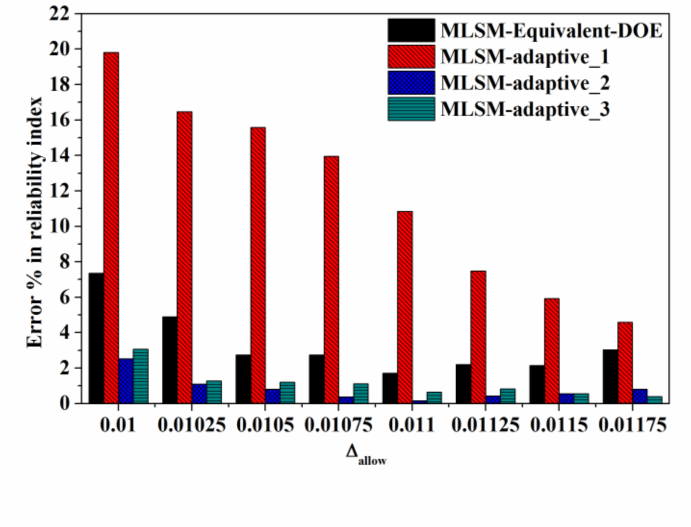


Figure 10. The comparison of reliability indices for different Δ_{allow} considering 30 initial training data.



495

496 Figure 11. The comparison of absolute percentage error in obtaining reliability indices for

497 different Δ_{allow} considering 30 initial training data.