Roy, A. and Chakraborty, S. (2022) Reliability analysis of structures by a three-stage sequential sampling based adaptive support vector regression model. *Reliability Engineering and System Safety*, 219, 108260. (doi: 10.1016/j.ress.2021.108260)

The material cannot be used for any other purpose without further permission of the publisher and is for private use only.

There may be differences between this version and the published version. You are advised to consult the publisher's version if you wish to cite from it.

https://eprints.gla.ac.uk/295698/

Deposited on 20 April 2023

Enlighten – Research publications by members of the University of Glasgow
http://eprints.gla.ac.uk

Reliability analysis of structures by a three-stage sequential sampling based adaptive support vector regression model

Atin Roy [1] and Subrata Chakraborty[2]
[1, 2]Department of Civil Engineering, Indian Institute of Engineering Science and Technology, Shibpur, India. E-mail: [1]*atinroy.rs2017@civil.iiests.ac.in* [2]*schak@civil.iiests.ac.in*

Abstract

A three-stage adaptive support vector regression (SVR) based metamodel by sequential sampling of training data close to the limit state function (LSF) is proposed that alleviates the difficulty of scarcity of samples in the reduced space for reliability evaluation of structures involving implicit LSF. Specifically, an importance sampling technique-based algorithm is proposed to ensure sufficient number of simulation points near the approximated failure plane. An initial design of experiments is first constructed by space-filling design over the entire domain. The optimum choices of the hyper-parameters of SVR model are determined by minimizing generalized root mean square error (GRMSE) and the corresponding minimum GRMSE is noted. A subset of Monte Carlo simulation samples having the magnitude of approximated LSF less than the noted GRMSE values are selected next. Subsequently, the data points are added sequentially from the subset based on maximin criterion. Finally, the SVR model is iteratively updated to improve reliability estimation by adding more data from the apparent safe and unsafe updated domain until convergence by an improved convergence criterion to avoid false convergence. The effectiveness of the proposed approach including estimating very small probability of failure is elucidated through three numerical examples.

Keywords: Support vector regression; three-stage sequential sampling; Monte Carlo simulation, reliability analysis, small probability; importance sampling.

1. Introduction

The reliability analysis of structure involves evaluation of a multiple integral which is a daunting task and various analytical i.e., the first order reliability method (FORM), the second order reliability method [1–3] and the simulation-based approximations are made to estimate probability of failure, P_f . The most accurate and conceptually straightforward Monte Carlo simulation (MCS) technique is preferred choice in this regard. However, the brute-force MCS technique requires a large number of

MCS samples for estimating small failure probability. Hence, various advanced MCS techniques requiring fewer samples e.g., importance sampling [4–7], directional sampling [8,9], subset simulation [10,11], line sampling [12,13] and local domain MCS [14] have been attempted. The number of simulation necessary in such approach is definitely less than that required by the brute-force MCS. But, still it remains an important issue [15] as structural reliability analysis (SRA) involving implicit LSF needs to be evaluated by computationally expensive finite element (FE) method [16]. Various metamodeling techniques have emerged as a viable alternative in this regard. The polynomial response surface method (PRSM) introduced by Faravelli [17] is widely used as a metamodeling technique due to its simplicity and computational efficiency [18–20]. The PRSM adopted in most of the studies relies on the least squares method based global approximation of scatter position data [21] and may not be enough accurate in a local domain. The efficiency of local approximation largely depends on the selection of basis functions which should closely represent the actual variation of the response within a local domain. The basis functions are typically fixed second order polynomials. But such a selection is not obvious as the response is implicit in nature [22]. Therefore various advanced metamodeling techniques such as moving least square method based PRSM [21,23], Kriging [24], artificial neural network [25,26], radial basis function [27,28], polynomial chaos expansion [29] etc. are gaining momentum due to their adaptive nature of approximation. But, such empirical risk minimization principle-based techniques usually suffer from overfitting and curse of dimensionality. On the other hand, the support vector machine (SVM) based on structural risk minimization principle and small sample learning that could estimate implicit function with better accuracy and generalization capability is worth noting [30,31]. The SVM based classification approach has been successfully applied for SRA problems involving implicit LSF [15,32,33]. Besides classification, SVM can also be utilized for regression which is referred as support vector regression (SVR). The application of SVR based metamodeling in SRA is quite notable [34–41]. In this regard, the advances in suppressing the curse of dimensionality with other metamodeling techniques e.g. sliced inverse regression-based sparse polynomial chaos expansions [42], surrogate modelling immersed probability density evolution method for reliability analysis in high dimensions are notable [43].

In this regard, it is important to note that the accuracy of reliability analysis predominantly depends on the accuracy of prediction of sign of LSF at the MCS sample points. To this end, including more data points near the limit state to construct design of experiments (DOE) for metamodel training is very important. A metamodel can be updated by reconstructing its DOE by shifting the centre and spread of the same based on the information of failure plane gathered from the previous metamodel [18–20,44,45]. The concept is also applied for SVR based SRA [37]. Instead of one-shot DOE, adaptive DOE by enriching an initial DOE with additional samples sequentially for metamodel training is an effective means to address this issue. It seems to be more effective than the former approach with regard to number of actual response evaluations. Liu et al. [46] presented a detailed survey on sequential sampling based adaptive DOE for global metamodeling. Echard et al. [47] proposed adaptive Kriging combined with MCS (AK-MCS), basically an active learning based sequential sampling specially effective with Kriging metamodel for improved SRA. Echard et al. [48] further proposed an active learning based reliability analysis method for small probability estimation and termed as adaptive Kriging combining with importance sampling (AK-IS). These methods are getting wide attention [49,50]. But such approach is not applicable for metamodels which are unable to provide prediction variance. The learning function for sequential sampling was developed as well for such metamodels [51]. Other sequential sampling based adaptive DOE are also implemented successfully with several metamodeling techniques for solving SRA problems [52–54]. Successful application of sequential sampling is also noted to construct adaptive SVR model. For example, Dai et al. [36] employed polynomial kernel function which essentially invites the curse of dimensionality problem even by using SVR model. As a result, five hundred to thousand numbers of total training data are required. Bourinet [39] proposed a subset simulation based adaptive SVR procedure where the accuracy of SVR model prediction is important near the actual limit state and also in estimating probabilities of each subset. Consequently, the total number of training points becomes very high. It has been generally noted that the requirement of total training sample in such adaptive SVR model is still seems to be quite high [36]; particularly for successful estimation of small failure probability [39]. Recently, Roy and Chakraborty [55] proposed a two stage sequential updating approach of SVR based metamodel for SRA with desired accuracy utilizing comparatively smaller number of samples. However, it is not so effective in estimating very small probability of failures.

Keeping the above in view, a three-stage adaptive SVR model based on sequential sampling of training data is proposed in the present study for SRA with limited number of training samples. It basically improves the approach proposed by Roy and Chakraborty [55] by duly addressing the difficulty of estimating very low failure probability case. The approach of Roy and Chakraborty [55] suffers from scarcity of samples in the reduced space particularly for low probability of failures problems. This limitation is circumvented in the present study but retaining the advantages of the previous approach. For this, a modification to tackle scarcity of samples is proposed by relying on importance sampling technique. For importance sampling method, the most probable failure point (MPFP) is evaluated by the FORM algorithm and a joint normal distribution with the evaluated MPFP as the mean vector and the original standard deviation (SD) values of the random variables as the SD vector is considered to obtain the importance sampling density function. Furthermore, the stopping condition in the sequentially updating procedure is enhanced by introducing an improved convergence criterion to avoid false convergence. The effectiveness of the proposed approach including estimating very small probability of failure is elucidated through three numerical examples.

2. Support vector regression

The present study hinges on SVR based metamodeling approach. The SVR is based on the principle of structural risk minimization of statistical learning theory [30,31]. The basic concept of SVR is briefly explained in this section and more details can be found elsewhere [56]. For a given set of training data, $\left\{(\mathbf{x}_1, y_1),(\mathbf{x}_2, y_2),\cdots,(\mathbf{x}_p, y_p)\right\}$, $\mathbf{x}\subset\mathbf{R}^n$ and $y\subset\mathbf{R}$ (**x** is the input vector; y is the corresponding output; p is the number of data pairs; **R** represents the set of real numbers; n is the dimension of input vector), the problem of regression is to find the flattest function f that map a point in the space $\mathbf{R}^n$ onto the space $\mathbf{R}$ with the lowest expected risk. The SVR can be applied for both linear and nonlinear regressions. For, a linear mapping, the regression function is expressed as,

$$f(\mathbf{x})=\langle\mathbf{w},\mathbf{x}\rangle+b,\qquad \mathbf{w}\in\mathbf{R}^n,\qquad b\in\mathbf{R} \tag{1}$$

Where, $\langle \mathbf{w}, \mathbf{x} \rangle$ is the dot product of $\mathbf{w}$ and $\mathbf{x}$; $\mathbf{w}$ and b are the two parameters to define a canonical hyperplane representing weight vector and bias, respectively. Mathematically, flatness of $f(\mathbf{x})$ can be ensured by minimizing the norm $\|\mathbf{w}\|^2 = \langle \mathbf{w}, \mathbf{w} \rangle$ and can be expressed as a convex optimization problem:

$$\min \frac{1}{2}\|\mathbf{w}\|^2 \qquad \text{s.t.} \quad \begin{cases} y_i - \langle \mathbf{w}, \mathbf{x}_i \rangle - b \le \varepsilon \\ \langle \mathbf{w}, \mathbf{x}_i \rangle + b - y_i \le \varepsilon \end{cases} \tag{2}$$

The optimization problem in Eq. (2) is feasible if the function $f(\mathbf{x})$ can approximate all the training points within $y_i \pm \varepsilon$, where, ε is a non-negative precision tolerance. However, the above may not be true for all the training points and for a more generalized approach some error allowance is desired. Thus, two slack variables, ξ_i, ξ_i^* are introduced to obtain a modified optimization problem [31] as,

$$\min \frac{1}{2}\|\mathbf{w}\|^2 + C\sum_{i=1}^{p}\left(\xi_i + \xi_i^*\right) \qquad \text{s.t.} \quad \begin{cases} y_i - \langle \mathbf{w}, \mathbf{x}_i \rangle - b \le \varepsilon + \xi_i, \\ \langle \mathbf{w}, \mathbf{x}_i \rangle + b - y_i \le \varepsilon + \xi_i^*, \end{cases} \qquad \xi_i, \xi_i^* \ge 0 \tag{3}$$

Where, C is the regularization constant which regulates the trade-off between the flatness of $f(\mathbf{x})$ and the amount by which the fitting error magnitudes of $f(\mathbf{x})$ exceed ε. For brevity, the detailed solution procedure [56] of Eq. (3) is not included here; instead the result is directly provided. The weight vector and subsequently, the regression function $f(\mathbf{x})$ can be obtained as,

$$\mathbf{w} = \sum_{i=1}^{p}\left(\alpha_i - \alpha_i^*\right)\mathbf{x}_i, \qquad f(\mathbf{x}) = \sum_{i=1}^{p}\left(\alpha_i - \alpha_i^*\right)\langle \mathbf{x}, \mathbf{x}_i \rangle + b, \qquad \alpha_i, \alpha_i^* \in [0, C] \tag{4}$$

Where, α_i and α_i^* are the Lagrange dual variables [56]. The use of SVR for linear regression as discussed above can be readily extended to nonlinear regression cases. The SVR approximation function for nonlinear responses is obtained by replacing the dot product $\langle \mathbf{x_i}, \mathbf{x_j} \rangle$ in Eq. (4) by a kernel function $K(\mathbf{x_i}, \mathbf{x_j})$ as,

$$f(\mathbf{x}) = \sum_{i=1}^{p}\left(\alpha_i - \alpha_i^*\right)K(\mathbf{x_i}, \mathbf{x}) + b. \tag{5}$$

The kernel function must satisfy the Mercer's condition [57]. In the present study, the following Gaussian radial basis function (GRBF) kernel is adopted for approximation of LSF,

$$K\left(\mathbf{x_i}, \mathbf{x_j}\right) = \exp\left(-\frac{\| \mathbf{x_i} - \mathbf{x_j} \|^2}{2\sigma^2}\right) \tag{6}$$

Where, σ is a parameter of the GRBF kernel function. In most of the cases, very little information about the given data is available and a general assumption for smoothness of the derived function is reasonable which can be well achieved by GRBF [56]. The SVR is implemented here in MATLAB using Gunn toolbox readily available at http://www.isis.ecs.soton.ac.uk/resources/svminfo/.

The fitting of SVR model by optimizing ε-insensitive loss function as presented here, besides ε, it has a regularization parameter, C, the GRBF kernel parameter, σ. The values for C, ε and σ are necessary to build an SVR model. The optimum values of these parameters can be obtained by cross-validation techniques. The search algorithm applied for selection of hyper-parameters for SVR model is based on Roy et al. [41]. Basically, it solves an optimization sub-problem where the generalised root mean square error (GRMSE) value obtained by the cross-validation method is minimized. The optimum value of σ is obtained for the pairs of parameters C and ε from a logarithmic grid. Then, the values of the three parameters corresponds to the minimum cross-validation error norm are selected. The details of the algorithm can be seen in Roy et al. [41].

3. Proposed three-stage adaptive SVR model-based reliability analysis

The proposed adaptive SVR based metamodel for SRA is basically a three-stage procedure composed of initial, intermediate and final stages which are elaborated in the present section. The important issue to address the difficulty of scarcity of samples in the reduced space in case of reliability analysis of problems involving very small probability of failure by applying importance sampling technique is elaborated separately.

3.1 Initial stage

The proposed three stages adaptive SVR based metamodeling approach starts with an initial DOE which is constructed by space-filling design over the entire physical space of the random variables. To construct an initial SVR model, the optimum choice of the hyper parameters are obtained by minimizing the GRMSE by leave-one-out cross-validation approach [41] and the corresponding minimum GRMSE value (say, $e_{GRMSE}^{\min}$) is noted. Once an initial SVR model is obtained, the approximate value of an LSF at each MCS point can be obtained to decide whether it is safe or unsafe. Fig. 1 (a) shows the initial DOE and the approximate safe and unsafe MCS samples by the initial SVR model for a 2-D problem. If y is the actual value of an LSF, $\hat{g}(\mathbf{X})$ is its approximated value at point $\mathbf{X}$ and e is the error of approximation at the point, then one can write the following,

$$\begin{aligned} &|e| = |y - \hat{g}(\mathbf{X})| \;= |y| + |\hat{g}(\mathbf{X})| \quad \textit{if } sign(y) \neq sign(\hat{g}(\mathbf{X})) \\ &i.e., \;\; |e| \geq |\hat{g}(\mathbf{X})| \end{aligned} \tag{7}$$

In the above, $sign(*)$ and $|*|$ represent the sign and magnitude of a variable. From Eq. (7), it can be realized that if the absolute error magnitude in approximating an LSF at any sample point is more than the magnitude of the approximated value of the LSF then only misclassification of that point may occur. But, the actual magnitude of error ($|e|$) at any sample point is unknown as the value of the implicit LSF ($|y|$) at the said point is not known. It has been noted that the cross-validation error norms i.e. GRMSE for SVR based metamodel are comparable with the corresponding prediction error norm i.e. root mean square error [41,58]. Hence, it can be anticipated intuitively that the points corresponding to a magnitude of the approximated LSF less than the value of the noted GRMSE value are most likely be get misclassified. Hence, the accuracies of SVR model in approximating the LSF at these points are of paramount interest for SRA application.

3.2 Intermediate stage

In the second stage, a set of MCS points, Ω is identified based on the magnitudes of the LSF approximated by the initial SVR model at MCS points and the previously noted GRMSE value, $e_{GRMSE}^{\min}$. If $\mathbf{X}$ represents any MCS point and $|\hat{g}(\mathbf{X})|$ is the magnitude of the approximated LSF at that point then the set Ω can be identified as,

$$\Omega = \left\{ \mathbf{X} \,|\, |\hat{g}(\mathbf{X})| < e_{GRMSE}^{\min} \right\} \tag{8}$$

It can be readily realized that improved failure estimation by a metamodel depends on its better approximation capability near the failure boundary. Thus, the SVR model should be constructed so that the approximation region near the limit state gets more importance. For this, the GRMSE value obtained by a hold-out validation approach [59] is minimized to select the SVR hyper-parameters. In a hold-out validation, usually two-third of the data are considered as the training set and the remaining as the test set [59]. Following this, if an initial DOE contains p_0 numbers of data, then $p_0/2$ numbers of new training sample from **Ω** set are included in the DOE. These data points are selected from the set **Ω** sequentially based on maximin criterion to effectively fill the reduced space represented by the set **Ω** [60] to ensure that a new training sample is positioned at a maximum distance from its nearest existing training sample. For this, the minimum distance of a point $\mathbf{X}$ from the existing training samples is obtained as,

$$d(\mathbf{X}) = \left\| \mathbf{X} - \mathbf{S}_{nearest} \right\| \tag{9}$$

Where, $\mathbf{S}_{nearest}$ is the nearest existing training sample to the point $\mathbf{X}$. The maximin criterion selects the point having the maximum value of $d(\mathbf{X})$ from the set **Ω** and this process of adding new training samples is repeated sequentially to effectively fill the reduced space represented by the set **Ω**.

It is important to note here that an SVR model trained with DOE with most of the training samples selected from unsafe (or safe) domain may be biased in LSF approximation. Therefore, more or less equal numbers of adaptive points from the safe and the unsafe domains should be included in the DOE to obtain an unbiased metamodel. To ensure this, the set **Ω** is partitioned into two sets $\mathbf{\Omega}_{\text{safe}}$ and $\mathbf{\Omega}_{\text{unsafe}}$ based on the sign of the approximated LSF, i.e.,

$$\Omega_{safe} = \left\{ \mathbf{X} \,|\, \left| \hat{g}(\mathbf{X}) \right| < e_{GRMSE}^{\min},\ \hat{g}(\mathbf{X}) > 0 \right\}, \qquad \Omega_{unsafe} = \left\{ \mathbf{X} \,|\, \left| \hat{g}(\mathbf{X}) \right| < e_{GRMSE}^{\min},\ \hat{g}(\mathbf{X}) < 0 \right\} \tag{10}$$

Now, based on the maximin criterion, two training samples are selected from the set $\mathbf{\Omega}_{\text{unsafe}}$ and another two from the set $\mathbf{\Omega}_{\text{safe}}$. Thus, four new training samples are added to the existing DOE. This process of adding four new training samples is repeated sequentially until $p_0/2$ numbers (or, the nearest multiple of four) of training samples are included into the DOE. Fig. 1 (b) shows the MCS points of the reduced space, **Ω** along with the initial DOE and new adaptive training samples selected from the set **Ω**. The two sets $\mathbf{\Omega}_{\text{safe}}$ and $\mathbf{\Omega}_{\text{unsafe}}$ are also distinguished by separate colours in Fig. 1 (b). In the hold-out validation approach, all the training samples of the initial DOE are considered as training set and the newly added training samples are kept as the validation set. Thus, the hyper-parameters of the SVR model so obtained are expected to improve the LSF approximation at MCS points of the set

Ω than those obtained by the leave-one-out cross-validation where the regions far away from the failure plane unwisely get equal importance. Once $p_0 / 2$ numbers of new training samples are added, the second SVR model is constructed and the sets, Ω_{safe} and Ω_{unsafe} are updated accordingly. Fig. 1 (c) depicts the updated approximation of safe and unsafe domain based on the second SVR model.

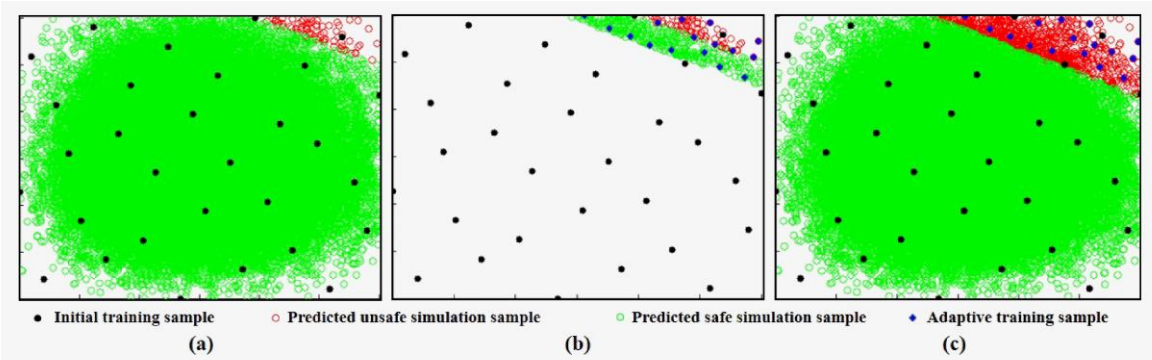

Fig. 1. The proposed adaptive sampling procedure showing (a) the initial DOE and the approximate safe and unsafe MCS samples, (b) the MCS points of the reduced space and the new adaptive training samples and (c) the updated approximation of safe and unsafe domain based on the second SVR model.

It may be noted that selecting by maximin criterion, the new training sample is ensured to be placed away from the existing training samples. Thus, adding alternatively one training sample from each of the sets Ω_{safe} and Ω_{unsafe} may place the training samples from one set near the approximated limit state and the other one from the remaining set far away from the limit state to ensure maximin criteria. On the contrary, adding more than two data alternatively from each of the sets can increase the chance to position the new data near the approximated limit state. However, adding more new data from each set will effectively fill the space of one set chosen first. But this process may not effectively cover the other space as no point of later set get the chance to positioned near the approximated limit state as already training samples are placed in the former set to maintain maximin criterion. Hence, adding two training samples alternatively from each of the set is adopted as an optimal solution.

3.3 Final stage

At the final stage, the SVR model is updated iteratively for further improvement of failure estimation by adding four new training samples sequentially from the updated set Ω (as obtained in the previous stage) into the DOE. In this regard, it may be noted here that a good number ($\approx p_0 / 2$) of adaptive training samples are already added to the DOE at the intermediate stage. Now, at the final stage, the DOE is iteratively enriched where accuracy of prediction needs to be improved further for obtining the correct sign of the LSF. The hyper-parameters to construct the updated SVR model at any iteration are obtained by hold-out validation approach and the corresponding minimum GRMSE is noted to update the existing set Ω. The updated set Ω at the end of each iteration is used to select four new training samples for next iteration and the process is continued until convergence. It is important to note here that in the hold-out validation approach at this stage, if the updated DOE contains a total of p numbers of training samples, then the most recent $p/3$ (or, nearest integer) numbers of training samples are hold-out as validation set and the remaining are taken as the training set. However, all the p numbers of training samples are utilized to train the SVR model at each iteration. Thereby, the SVR hyper-parameters are obtained to emphasize better approximation of the LSF at the MCS points surrounded by the recent $p/3$ numbers of training samples. As the newest $p/3$ number of training samples of the updated DOE are selected from the set Ω which is reconstructed at each iteration with the MCS points most vulnerable for misclassification. Thus, the proposed searching scheme of hyper-parameters is expected to further reduce the chances of misrecognition of the actual sign of the LSF at these MCS points.

The convergence criterion proposed in the present study is based on the variation of P_f values over the previous two iterations instead of the usual practice of considering value of the last iteration. Mathematically, this convergence criterion can be expressed as,

$$\max\left(\frac{\left|P_f^{i-2} - P_f^{i}\right|}{P_f^{i}}, \frac{\left|P_f^{i-1} - P_f^{i}\right|}{P_f^{i}}\right) \le 0.05 \tag{11}$$

Where, P_f^{i-2}, P_f^{i-1} and P_f^{i} are the values of the probability of failures at (*i*-2), (*i*-1) and *i*-th iterations, respectively. The above intuitively assumes that if three consecutive iterations give very close estimate of P_f values then it can be presumed to be very close to the actual one. Once, the values of P_f for three consecutive iterations bounded within ±5% of the most updated value, the newest value is considered as the converged P_f value.

3.4 Proposed modification to deal with small failure probabilities

The success of the proposed three stage sequential updating algorithm depends on the availability of sufficient number of MCS samples in the sets Ω_{safe} and Ω_{unsafe} which need to be selected for updating the SVR model to improve response approximation for reliability estimate. However, in case of very small probability of failure, the set Ω_{unsafe} may suffer from scarcity of samples and a very few points may be available having lesser magnitude of approximated LSF than the noted GRMSE value. To circumvent this difficulty, advanced MCS technique e.g., importance sampling and subset simulations capable of generating more simulation samples in the failure regions can be employed. However, the accuracy of P_f estimated by subset simulation technique depends on the accuracies of failure estimations of all the subsets involved in the entire process which demands prediction accuracy not only near the final limit state but also in all such subset boundaries. Thus, the proposed three-stage adaptive SVR approach needs to be applied for each subset and this will largely increase the requirement of total number of training samples. But importance sampling is free from such difficulty. Thus, a modification to tackle scarcity of samples for very small probability problem is proposed in the present study relying on importance sampling technique.

To obtain a quasi-optimal density function for importance sampling method, the MPFP is evaluated first by Rackwitz-Fiessler FORM algorithm [1]. In doing so, the value of the LSF and its gradient at any iteration are evaluated from one single regression model constructed to approximate the LSF at the corresponding iteration step. The two parameters equivalent normal transformation [61] is adopted to deal with the non-normal random variables in the FORM algorithm. Now, a joint normal distribution with the evaluated MPFP as mean vector (say, $\boldsymbol{\mu}^{MPFP}$) and the original SD values of the random variables as the SD vector is considered to obtain the importance sampling density function [62,63]. Then the joint PDFs for the importance sampling density function, f_{IS} at any point $\mathbf{X}$ is obtained as,

$$f_{IS}(\mathbf{X}) = \prod_{k=1}^{n} \phi\left(\frac{x_k - \mu_k^{MPFP}}{\sigma_k}\right) \tag{12}$$

Where, $\phi(*)$ represents the standard normal PDF; n is dimension of the input space; σ_k is the original SD value of the kth random variable; x_k and μ_k^{MPFP} represent the values of the kth random variable and the corresponding ordinate of the MPFP point, respectively. In this regard, it may be noted that the reliability is not computed by the FORM rather used to obtain a suitable centre for importance sampling. Thus, the failure probability accuracy will not be affected much as the failure probability is finally estimated by importance sampling [48].

At the initial stage, the MPFP and subsequently, the importance sampling density function are obtained based on the initial SVR model. Certain number of simulation samples (say, N_{IS}) are generated from the quasi-optimal density function and the LSF is approximated at all these samples based on the initial SVR model. The sets Ω, Ω_{safe} and Ω_{unsafe} are built by the eligible simulation samples selected based on the criteria detailed in section 3.2. Once the sets Ω_{safe} and Ω_{unsafe} are obtained, the DOE can be enriched with new sequential data points. Subsequently, the SVR model is updated with new training data set. Based on the updated SVR model, a new MPFP is evaluated and subsequently, the mean values of the random variables for importance sampling density function are updated. The previous N_{IS} number of simulation samples are replaced by the new N_{IS} number of simulation samples generated from the updated importance sampling density function. Based on the approximated values of the LSF at new N_{IS} number of samples obtained based on the updated SVR model, the sets Ω, Ω_{safe} and Ω_{unsafe} are rebuilt. Thus, the proposed three-stage adaptive SVR algorithm continued at the intermediate and the final stages accordingly.

3.5 Outline of the proposed SVR approach

The implementation of the proposed three-stage adaptive SVR algorithm for reliability evaluation is explained through a flow chart in Fig. 2. To start the algorithm, the maximum allowable number of brute-force MCS samples, (say, N_{MC}) is decided. Based on an acceptable COV of P_f value, a minimum value of P_f (say, P_f^{min}) can be obtained up to which the value of P_f can be estimated following brute-force MCS based estimation procedure. The procedure to obtain the reduce space for sequential sampling is altered by changing the reliability evaluation method automatically between the brute-force MCS and the importance sampling method based on the value of COV of failure probability. If the estimated P_f is less than P_f^{min} at any stage then the simulation samples are replaced by N_{IS} number of samples generated from the importance sampling density function (as discussed in section 3.4) and the P_f value is estimated by importance sapling method. In any later iteration step, if the estimated P_f is more than P_f^{min} then the value of P_f is estimated again by the brute-force MCS method from the next iteration. If the estimated P_f is more than P_f^{min}, even in case of scarcity of samples, the brute-force MCS remains as the reliability estimation method. But, to find the eligible candidates for the set Ω, N_{IS} number of samples are generated as per the quasi-optimal density function and the values of the LSF at these sample points are evaluated by the SVR model. It is to mention here that the present SRA approach is restricted to aleatory uncertainty only and consideration of epistemic uncertainty [64,65] is beyond the scope of the present study and needs separate consideration.

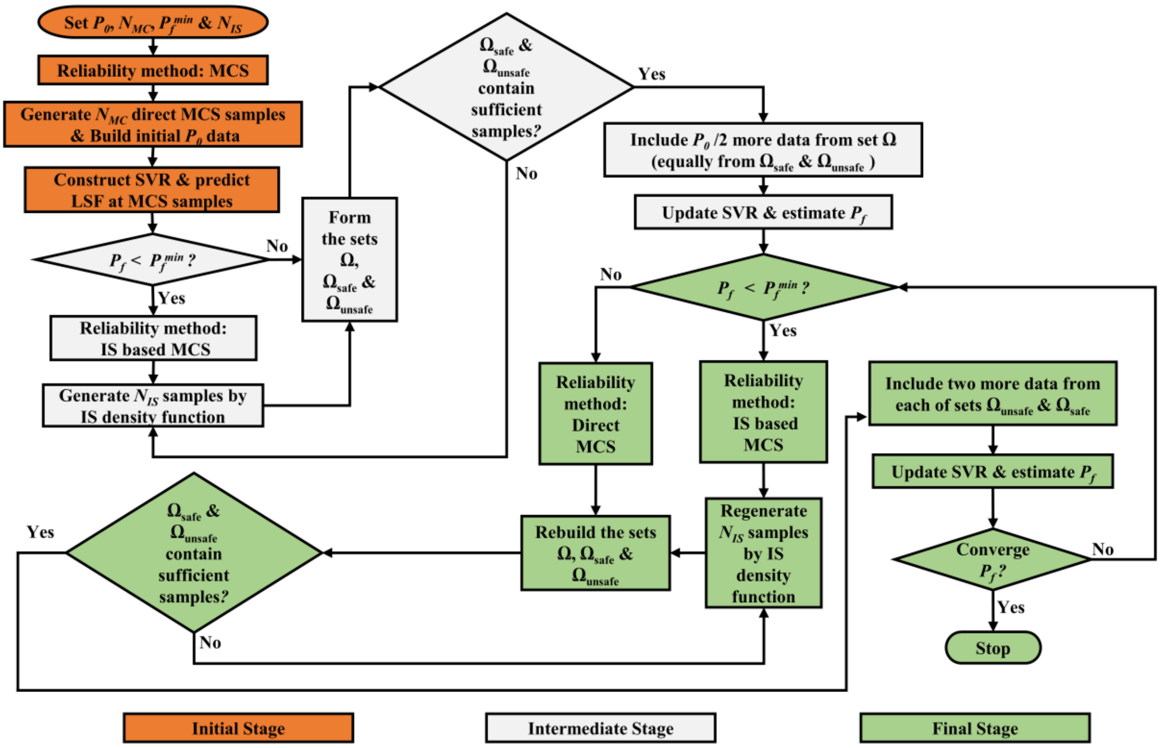

Fig. 2. The flowchart of the proposed three-stage adaptive SVR algorithm for SRA.

4. Numerical Study

The proposed adaptive SVR approach based on sequential updating of training data set in three stages for reliability analysis by duly addressing the small failure probability issue is illustrated by considering three examples. The first example problem is a ten-bar truss problem for which explicit

LSF is available. Thus, the reliability computation by brute-force MCS technique using the actual LSF can be easily performed. This will enable to study the performance of the proposed algorithm by comparing with the most accurate MCS based reliability results. The second example is a twenty-nine-dimensional standard test problem involving a purely nonlinear mathematical LSF used to demonstrate the effectiveness of the proposed approach for higher dimensional problem. The last example is a more realistic one i.e., a space dome truss problem requiring evaluation of an implicit LSF involving FE analysis of the structure.

The P_f evaluation in the present study involves three approaches: (i) the brute-force MCS (ii) the importance sampling technique and (iii) the proposed SVR based approach. The brute-force MCS technique proceeds in three steps: (a) random sampling of N_{MC} sets of input parameters according to underlying PDFs, (b) evaluating the values of the LSF at all the sample points and (c) post-calculating the failure probability. The estimated P_f value obtained by brute-force MCS technique and its variability measured by its COV, δ_{MC} can be obtained as [66],

$$P_f = \frac{1}{N_{MC}} \sum_{i=1}^{N_{MC}} I\left[g(\mathbf{X}_i)\right], \qquad \delta_{MC} = \sqrt{\frac{1-P_f}{N_{MC}\, P_f}} \tag{13}$$

Where, $\mathbf{X}_i$ represents the ith sample. The indicator function, $I[g(\mathbf{X})]$ is equal to 1 for $g(\mathbf{X})<0$ and 0, otherwise. In importance sampling method, the P_f value and its COV are obtained as [66],

$$P_f = \frac{1}{N_{IS}} \sum_{i=1}^{N_{IS}} I\left[g(\mathbf{X}_i)\right] \frac{f_X(\mathbf{X}_i)}{f_{IS}(\mathbf{X}_i)}, \qquad \delta_{IS} = \frac{1}{P_f} \sqrt{\frac{1}{N_{IS}} \left(\frac{1}{N_{IS}} \sum_{i=1}^{N_{IS}} I\left[g(\mathbf{X}_i)\right] \left(\frac{f_X(\mathbf{X}_i)}{f_{IS}(\mathbf{X}_i)} \right)^2 - P_f^{\,2} \right)} \tag{14}$$

Where, f_X and f_{IS} are the joint PDFs for the actual input space and the generated importance samples, respectively. N_{IS} is the total numbers of importance samples generated. For SVR based reliability analysis, one needs LSF value from the SVR model. Once the SVR model is finalized, linking it to the MCS or importance sampling method-based reliability analysis is straightforward. For this, a set of values of all the random variables are generated and the trained SVR model is invoked to provide responses and reliability is estimated accordingly. The result of the MCS based failure estimate using the actual LSF is considered as the reference result.

For comparative study, the reliability results are also obtained by the most widely used active learning reliability methods. In doing so, the AK-MCS method [47] is first attempted and if it is found that the method demands very large population of samples then the AK-IS method [48] is employed. The U-function proposed by Echard et al. [47] is used as the learning function for all the active learning methods for numerical study. In addition, the results are also obtained by the sequential updating based SVR approach proposed by Roy and Chakraborty [55] for comparison and it is referred as ‘Sequential SVR’ in the rest of the article. In cases of very small probabilities, only AK-IS method is compared with the proposed there-stage SVR method as the other two methods (AK-MCS and ‘Sequential SVR’) are not capable with computationally feasible number of simulation samples.

4.1 Example 1: A ten-bar truss

The ten-bar truss as shown in Fig. 3 is taken as the first example. The cross-sectional areas of the horizontal members (A_1), vertical members (A_2), diagonal members (A_3), Young’s modulus (E), length (L) and load (P) are considered as independent random variables. The statistical characteristics and physical boundaries of the considered six random variables are detailed in Table 1. This example is adopted from Choi et al. [67] where reliability analysis was demonstrated considering displacement at nodes as the criteria. The LSF for reliability is considered with respect to the tip displacement at node 2 as following [67],

$$\begin{aligned} g_{disp} = d_{allow} - \frac{PL}{A_1 A_3 E D_T}[4\sqrt{2}A_1^3(24A_2^2 + A_3^2) + A_3^3(7A_1^2 + 26A_2^2) \\ +4A_1A_2A_3\{(20A_1^2 + 76A_1A_2 + 10A_3^2) + \sqrt{2}A_3(25A_1 + 29A_2)\}] \end{aligned} \tag{15}$$

Where, $D_T = 4A_2^2(8A_1^2 + A_3^2) + 4\sqrt{2}A_1A_2A_3(3A_1 + 4A_2) + A_1A_3^2(A_1 + 6A_2)$ and d_{allow} is the allowable limit of the tip displacement at node 2.

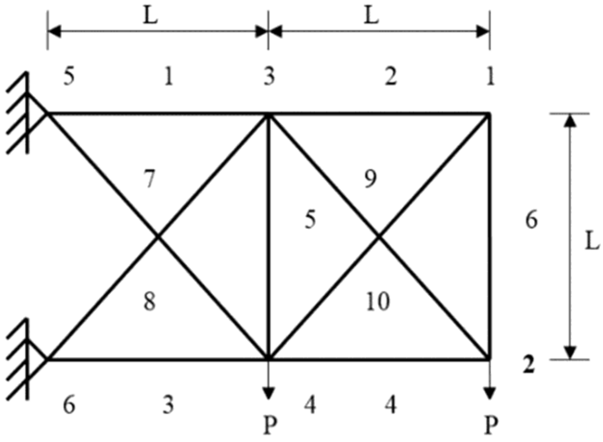

Fig. 3. The planar ten-bar truss (redrawn from [67])

Table 1.The details of the random variables of the ten-bar truss

Random variables (unit)	Probability distribution			Physical boundary	
	Type	Mean	COV	Lower limit	Upper limit
A_1 (m^2)	Normal	7.5×10^{-3}	0.1	5.25×10^{-3}	9.75×10^{-3}
A_2 (m^2)	Normal	1.5×10^{-3}	0.1	1.05×10^{-3}	1.95×10^{-3}
A_3(m^2)	Normal	5.0×10^{-3}	0.1	3.5×10^{-3}	6.5×10^{-3}
E (N/m^2)	Normal	7.0×10^{10}	0.05	5.95×10^{10}	8.05×10^{10}
L (m)	Lognormal	9.0	0.05	7.65	10.35
P (N)	Gumbel Max.	3.5×10^{5}	0.1	2.45×10^{5}	4.55×10^{5}

To study the effectiveness of the proposed three-stage adaptive SVR approach for reliability estimation, an initial DOE consists of 30 training data points are constructed over the entire physical domain of the random variables according to the uniform design (UD) table, $U_{30}(30^6)$ (readily available at http://www.math.hkbu.edu.hk/UniformDesign). The UD is a space-filling design and has the distinctive feature of accommodating the largest possible number of levels for each variable and the discrepancy for UD is the smallest amongst all the space-filling designs [68]. Thus, in order to construct an efficient metamodel, the UD scheme [69] is adopted in the present study. With these 30 initial training data, the intermediate stage adds 16 new data (the nearest multiple of four to half of 30). Then in the final stage, four data are added iteratively until convergence. The prerequisite parameters to implement the algorithm are set as: $N_{MC} = 10^5$; $P_f^{min} = 10^{-3}$ (compliant with COV of P_f < 0.1) and $N_{IS} = 10^4$.

The result of brute-force MCS based failure estimate with 10^5 MCS samples using the actual LSF is considered as the reference result if the COV of P_f is less than 0.1; otherwise, the reference P_f value is estimated by importance sampling method. The estimated P_f values by the proposed approach after each update of the SVR model for different d_{allow} are shown in Fig. 4. The reference results are shown in the same plot for comparative study. The probability of failures estimated by the 'Sequential SVR' method are also shown in the plot, wherever, the method works with reasonable number of MCS samples. Regarding the AK-MCS method, Echard et al. [47] suggested that the algorithm can start from an initial DOE composed of a dozen of random samples from the MCS set. The results obtained by U-function based AK-MCS method starting with 12 MCS samples as training data is denoted by 'AK-MCS+U'. However, the proposed SVR method and the 'Sequential SVR' start from an identical initial DOE with 30 UD samples. But, the 'AK-MCS+U' method starts with an initial DOE with lesser samples. It is to be noted here that the performance of any adaptive method is expected to be influenced by the initial design. Thus, for meaningful comparative study, the performances of the two adaptive methods are compared with an identical initial DOE. Thereby, another AK-MCS method starting from the same initial DOE i.e., with 30 UD samples (denoted by '30UD+AK-MCS+U') is also considered for comparison.

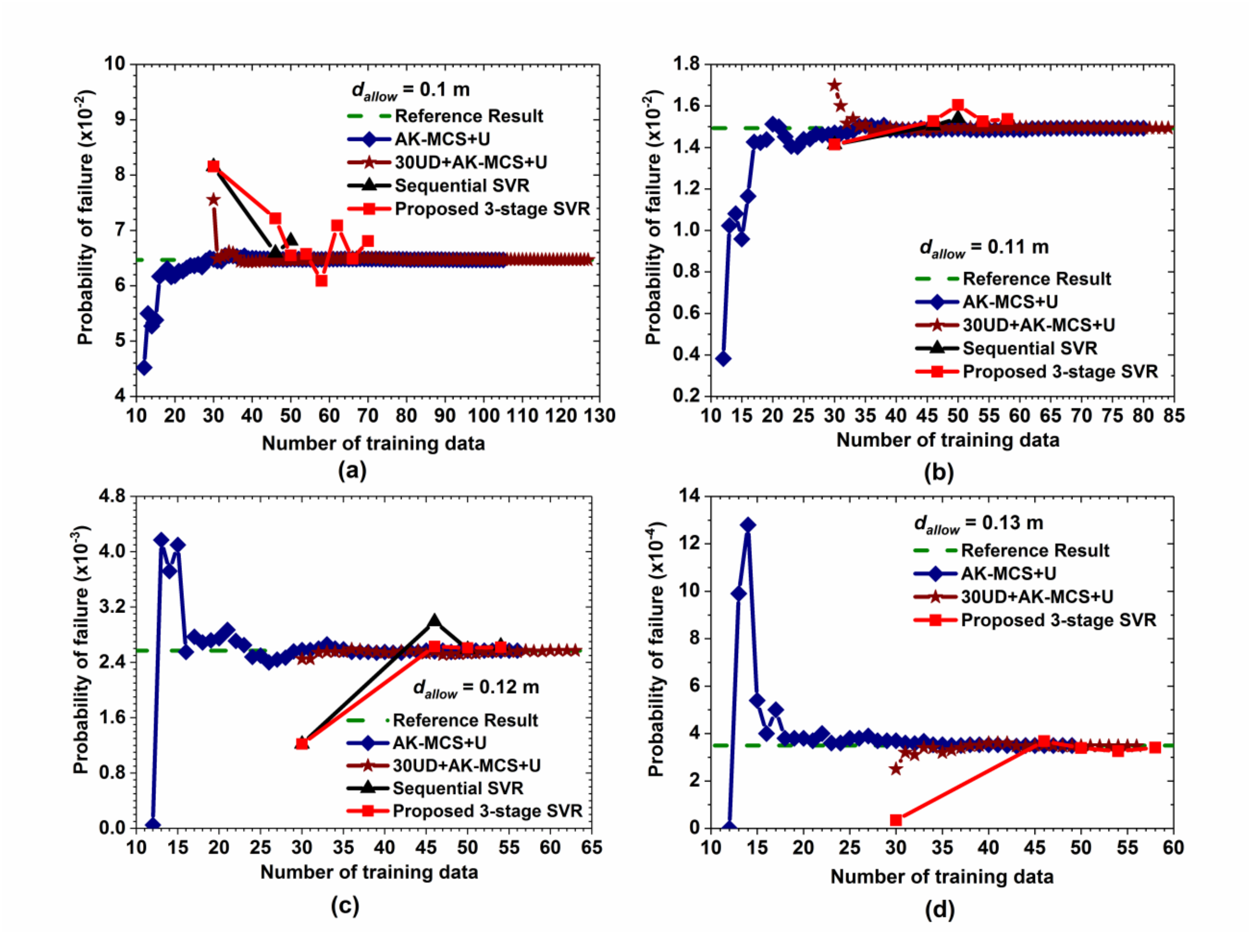

Fig. 4. The comparison of estimated P_f values using various approaches for different d_{allow} values

It can be observed from the Fig. 4 that the P_f values estimated by the proposed SVR model are quite close to the reference results. The computational efficiency of the proposed method is studied with respect to the numbers of actual function evaluation to obtain the actual value of the LSF. It can be noted that more numbers of actual function call represent more computational time requirement irrespective of problem size and computing platform. The proposed three-stage SVR approach takes only 54 to 70 numbers of training data for different d_{allow} values. The results of the 'Sequential SVR' method have similar accuracy with little lesser number of data (i.e., 47 to 49). But, the 'Sequential SVR' method is unable to estimate reliability in case of d_{allow} = 0.13 m due to scarcity of candidate points in the reduced space. The 'AK-MCS+U' and '30UD+AK-MCS+U' estimate P_f with good accuracy. But, the AK-MCS methods require more numbers of iterations as well as more numbers of training data than the proposed three-stage adaptive SVR approach. The '30UD+AK-MCS+U' method takes a smaller number of iterations e.g., 97 and 26 than the 'AK-MCS+U' method e.g., 99 and 37 for d_{allow}= 0.1 and 0.13 m, respectively. However, the 'AK-MCS+U' method requires a smaller number of total training samples e.g., 12+99=111 and 12+37=49 than the 30UD+AK-MCS+U' method e.g., 30+97=127 and 30+26=56 in case of d_{allow}= 0.1 and 0.13 m, respectively. It shows that the initial DOE data points (30 UD samples) reduce the number of iterations but there is no benefit with respect to total number of training samples. Hence, the initial DOE based on 30 UD samples does not facilitate the AK-MCS method for this problem.

The value of d_{allow} is further increased (i.e., 0.14 and 0.15 m) to study the performance of the proposed three-stage adaptive SVR approach for small failure probability case. Instead of AK-MCS, AK-IS is employed for reliability estimation. The corresponding results are shown in Fig. 5. It can be observed that the AK-IS method takes lesser number of training data (53 and 68) than the proposed three stage adaptive SVR method (62 and 90) for this particular problem. However, it can be noted that the AK-IS method is unable to start with small initial data (unlike AK-MCS) as it requires to find the MPFP first by employing FORM which involves several iterations with gradient evaluations [48]. In detail, initial number of samples predominantly depends on number of iterations in FORM and the input dimension. For example, FORM with seven iterations with six input variables required 50 (i.e., 7×(1+6)+1 = 50) numbers of data for evaluation of the actual implicit function. The AK-IS method requires 50 (i.e., seven iterations for FORM) and 64 (i.e., nine iterations for FORM) number of initial data for d_{allow} = 0.14 and 0.15 m, respectively. Though, AK-IS method requires marginally lesser samples in this problem than the proposed method; the initial data requirement of AK-IS method might be very high for high dimensional problems. This is demonstrated in the next example.

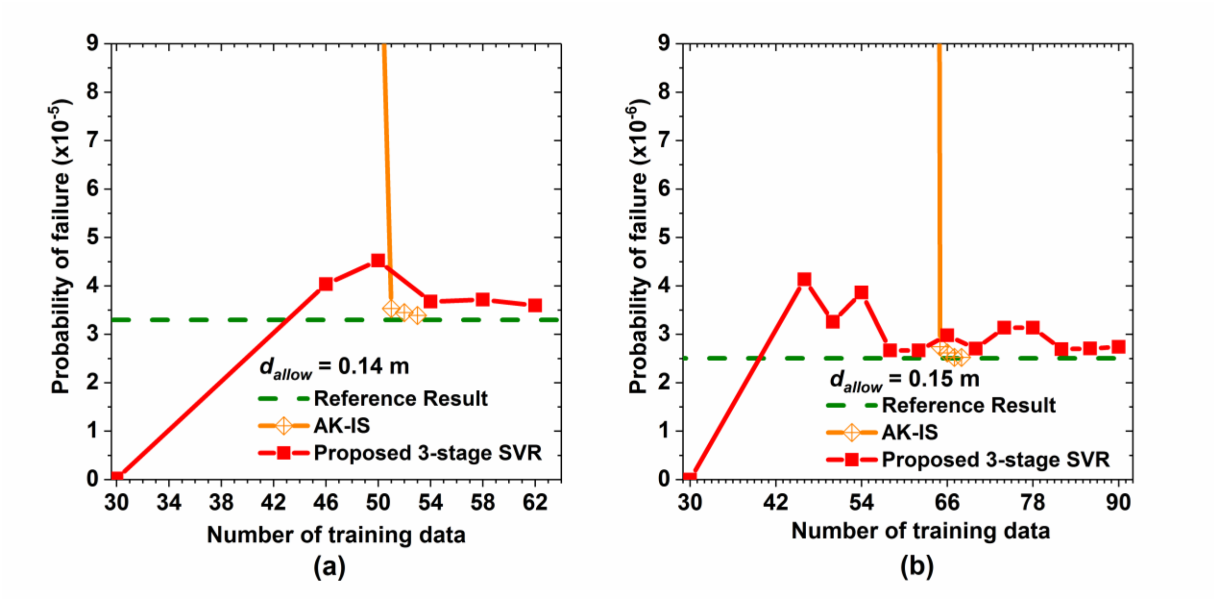

Fig. 5. The comparison of estimated P_f values for d_{allow}= 0.14 and 0.15 m.

4.2 Example 2: A twenty-nine-dimensional test problem

A twenty-nine dimensional (29-D) standard test problem is taken to demonstrate the effectiveness of the proposed approach for comparatively higher dimensional problem. The LSF is expressed as,

$$g\left(\mathbf{X}\right)=Y_{allow}-\sum_{i=1}^{29}x_i^2-\left(\sum_{i=1}^{29}\left(\frac{1}{2}\right)i\,x_i\right)^2-\left(\sum_{i=1}^{29}\left(\frac{1}{2}\right)i\,x_i\right)^4,i=1,2\ldots,29. \quad (16)$$

Where, Y_{allow} is the allowable value of the function. Each random variable is considered to be lognormal with mean and COV of 10.0 and 0.1, respectively and truncated between 7.0 and 13.0. The initial DOE is constructed using UD table, $U_{30}(30^{29})$ over the entire input space. The necessary parameters are taken as: $N_{MC}=10^6$; $P_f^{min}=10^{-4}$ and $N_{IS}=5\times10^4$. The estimated P_f values obtained by the proposed approach, the two AK-MCS methods and the 'Sequential SVR' method along with the reference results are shown in Fig. 6 for different Y_{allow}. It has been noted that for the case of $Y_{allow}=3.0\times10^{13}$, the proposed algorithm changes from brute-force MCS to importance sampling method for reliability estimate at the initial stage; but it returns to brute-force MCS again at intermediate stage and continues to employ brute-force MCS up to the final stage. This switching of reliability method is based on the estimated P_f values at any iteration stage. The observation on the performance of 'Sequential SVR' method is similar to the previous problem. In case of $Y_{allow}=2.8\times10^{13}$, 2.9×10^{13} and 3.0×10^{13}, learning of the 'AK-MCS+U' method with 10^6 brute-force MCS samples is stopped at the initial step as no MCS points are detected as unsafe even with such large numbers of MCS sample size. On the other hand, in the case of '30UD+AK-MCS+U' method, the stopping criterion for learning is not satisfied even after adding 120 new data (i.e., use of total 150 training data) for Y_{allow} = 2.7×10^{13}, 2.8×10^{13}, 2.9×10^{13} and 3.0×10^{13} values.

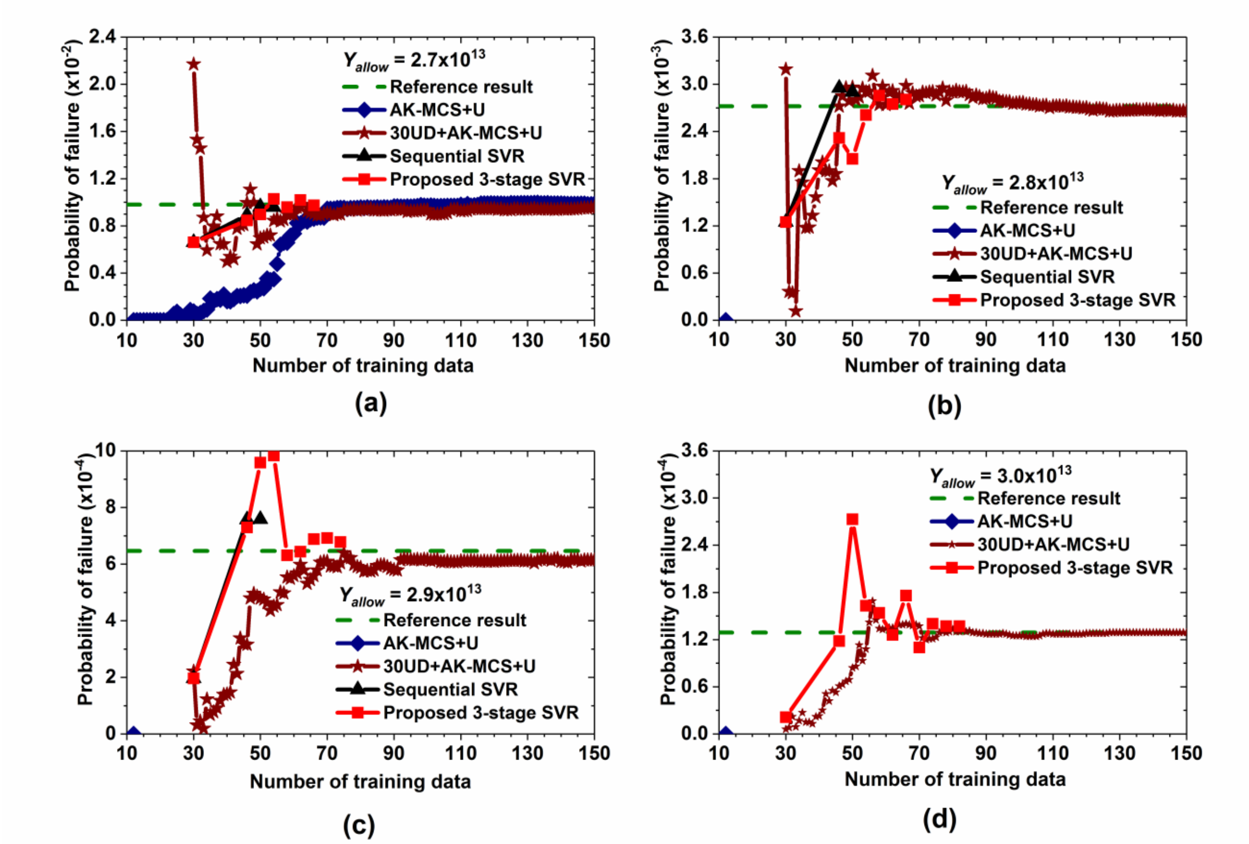

Fig. 6. The comparison of estimated P_f values for different Y_{allow} values.

Further, to study the performance of the proposed three-stage adaptive SVR approach for low probabilities, the P_f values are also estimated for comparatively higher values of Y_{allow} e.g., Y_{allow} $=3.5\times10^{13}$ and 3.6×10^{13}. The corresponding results are shown in Fig. 7. It can be observed from the figure that the proposed three-stage adaptive SVR method estimates the failure probabilities successfully with maximum of 118 training data. However, the AK-IS method consumed 121 data before starting of active learning iterations. These 121 data are basically used by four iterations in FORM to obtain the MPFP (i.e., 4×(29+1)+1=121). Moreover, the convergence of active learning is not reached even after 120 iterations i.e., after utilization of total 240 training data in cases of Y_{allow}= 3.5×10^{13} and 3.6×10^{13}. It can also be observed that for several times, the estimated P_f values by the proposed SVR approach do not change significantly for two successive iterations in case of Y_{allow} = 3.5×10^{13}. This reveals the efficiency and accuracy of utilizing the convergence criterion based on the results of consecutive three iterations over that of two iterations.

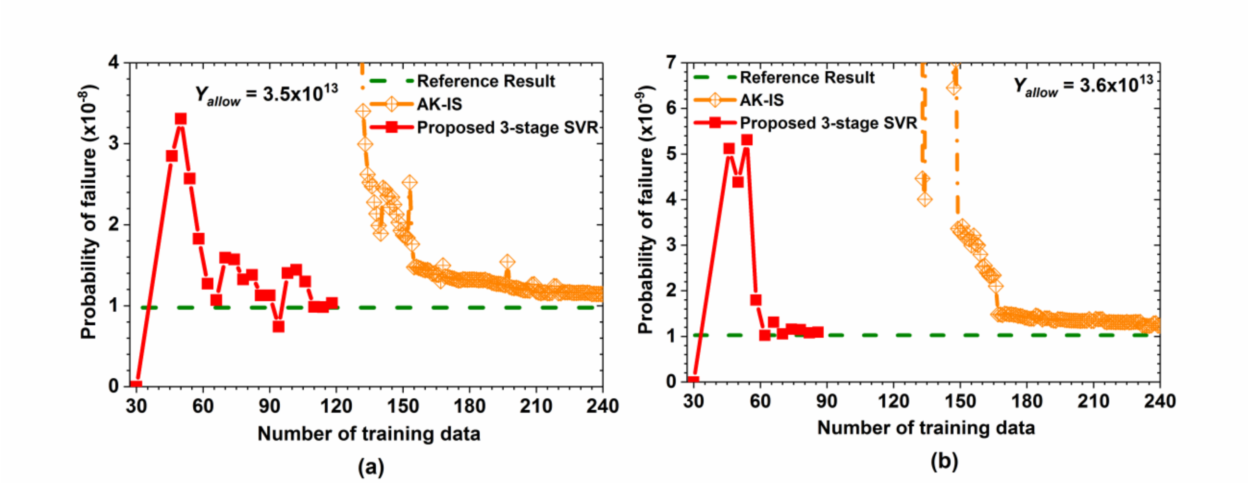

Fig. 7. The comparison of estimated P_f values by various approaches for Y_{allow}= 3.5×10^{13} and 3.6×10^{13}.

4.3 Example 3: A space-dome truss

The third example problem is reliability analysis of a space-dome truss involving implicit LSF. The schematic diagram of the space truss is shown in Fig. 8 [70]. The six independent random variables considered are: the material Young's modulus (E) of all the bars, the cross section areas of the top radial bars (A_1 for bar numbers 1-6), the peripheral bars (A_2 for bar numbers 7-12), the bottom inclined bars (A_3 for bar numbers 13-24), the point load P_1 at the centre node and the point load P_2 at the six nodes of the middle hexagon. The reliability analysis is performed with respect to the maximum vertical displacement of the node under load P_1, the implicit LSF for which can be expressed as,

$$g = \Delta_{\text{allow}} - \left|\Delta^{z}_{\text{P1}}\right| \tag{17}$$

Where, Δ^{z}_{P1} is the vertical displacement of the node under load P_1 and Δ_{allow} is the allowable maximum displacement of the same. The displacement Δ^{z}_{P1} which is necessary to evaluate the LSF is required to be computed based on FE analysis of the structure. The statistical properties and physical boundaries of the random variables are shown in Table 2.

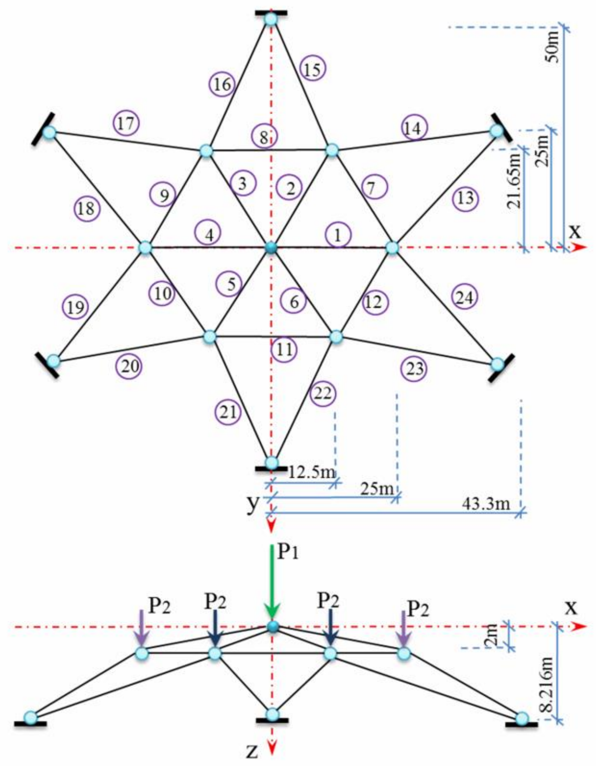

Fig. 8. Schematic diagram of the space truss (redrawn from [70])

Table 2. The details of the random variables of the space truss

Random variables (unit)	Probability distribution			Physical boundary	
	Type	Mean	COV	Lower limit	Upper limit
A_1 (m^2)	Normal	0.013	0.1	0.0091	0.0169
A_2 (m^2)	Normal	0.01	0.1	0.007	0.013
A_3(m^2)	Normal	0.016	0.1	0.0112	0.0208
E (GPa)	Normal	205	0.05	143.5	266.5
P_1(kN)	Gumbel Max.	20	0.15	14	26
P_2(kN)	Gumbel Max.	10	0.12	7	13

An initial DOE consisting of 30 data points is constructed using UD table, $U_{30}(30^6)$ over the entire physical space to estimate reliability by the proposed approach. The prerequisite parameters considered are: N_{MC} = 5×10^4; P_f^{min} = 2×10^{-3} and N_{IS} = 10^4. For comparative study, failure probabilities are also estimated by the previously mentioned two AK-MCS methods and 'Sequential SVR'. The estimated values of P_f by the different methods at each step of iterations for different Δ_{allow} are shown in Fig. 9. The reference results are obtained by evaluating the actual LSF for 5×10^4 brute-force MCS samples if the COV of estimated P_f is less than 0.1; otherwise, importance sampling method is employed. For comparison, the reference results are shown in the same plot. The P_f values estimated by the final SVR models are observed to be very close to the reference results for this problem also. It can be noted from Fig. 9 that the proposed method takes only 54 to 62 numbers of training data for Δ_{allow} = 0.011, 0.0115, 0.012 and 0.0125 m. Though, the 'Sequential SVR' method requires little lesser number of training data (45 to 49); like the previous examples, it is unable to produce any result in case of Δ_{allow} = 0.0125 m due to scarcity of candidate samples in the reduced space. As earlier, both 'AK-MCS+U' and '30UD+AK-MCS+U' methods take more number of training data for most of the cases. The value of Δ_{allow} is further increased to 0.013 and 0.0135 m for studying the performance of the proposed method for small probability of failure case. The corresponding results are compared with the reference values in Fig. 10. Only 54 numbers of total training samples are required by the proposed method for both the cases of Δ_{allow} = 0.013 and 0.0135 m.

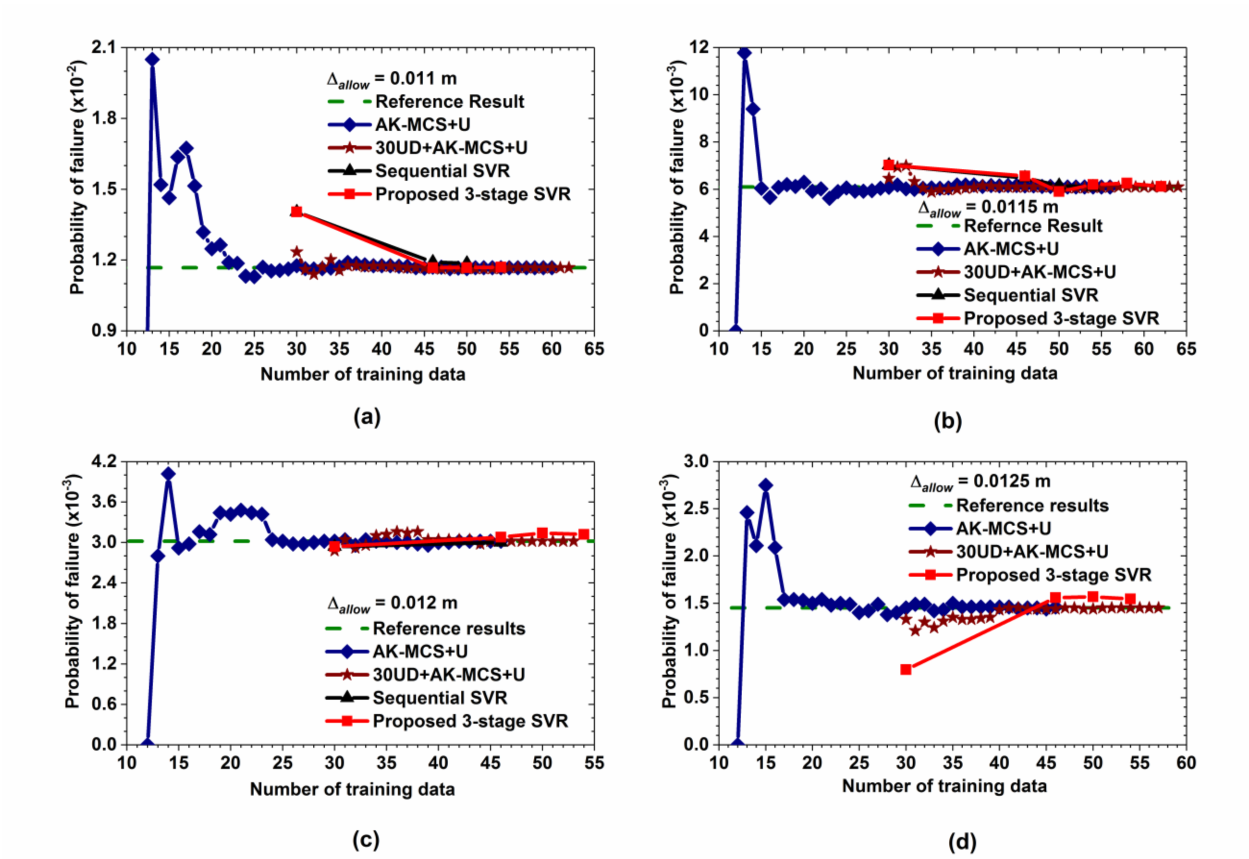

Fig. 9. The comparison of estimated P_f values for different Δ_{allow} values.

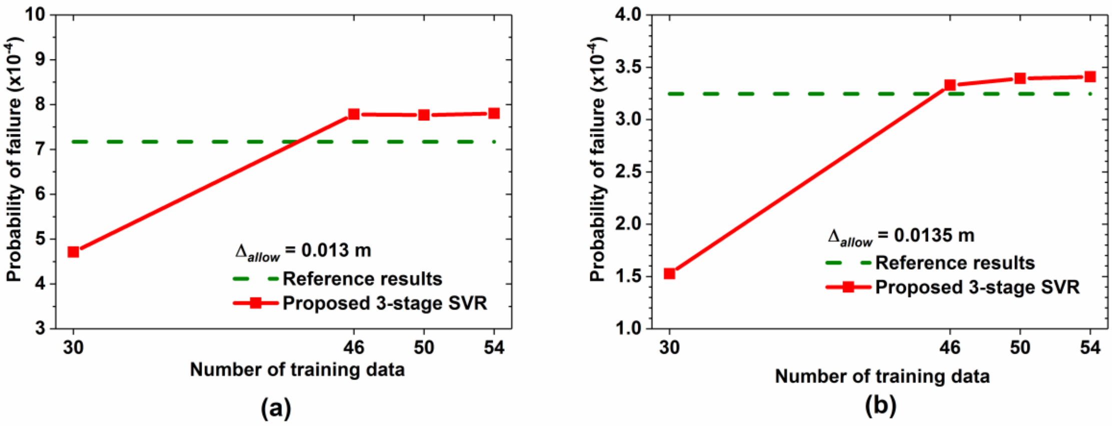

Fig. 10. The comparison of estimated P_f values obtained by the proposed approach and the reference value for Δ_{allow} = 0.013 and 0.0135 m.

In general, it is observed that the final estimates of P_f values are quite accurate with respect to the reference results even for very small P_f values. The computational efficiency of the proposed approach studied with respect to the numbers of actual function evaluation shows that the proposed three-stage SVR approach and the 'Sequential SVR' method have similar accuracy with less numbers of data by the 'Sequential SVR'. But, the 'Sequential SVR' is unable to estimate small failure probability cases. The 'AK-MCS+U' and '30UD+AK-MCS+U' estimate P_f with good accuracy. But, the AK-MCS method involves more number of iterations and training data than the proposed three-stage adaptive SVR approach. The performance of AK-IS method is found to be largely depended on the input dimension of the problem. In this regard, it is to be noted that the initial Kriging model of AK-IS method is trained with all the points required to obtain MPFP by FORM. Thereby, the number of initial training data and subsequently the total number of training data is expected to be very large for high dimension problem. It may be noted that in case of small probabilities, brute-force MCS is avoided to obtain the reference results due to huge computational involvement for problems involving FE analysis. In such cases, the importance sampling technique is employed. This may introduce error in the actual estimate of P_f values resulting in some differences of the results of the proposed SVR method with the reference results. Better efficiency and accuracy of utilizing the proposed convergence criterion based on three iterations is observed over the conventional one. i.e., based on two iterations only. It is quite obvious that the proposed convergence criterion demands at least one additional iteration than the conventional stopping criterion. Thus, the number of total training data required by the ‘Sequential SVR’ method (based on the conventional stopping criterion) is marginally less than that of required by the proposed adaptive SVR approach. But, as already discussed, the ‘Sequential SVR’ method suffers from the issue of scarcity of candidate points in the reduced space for small failure cases.

5. Summary and conclusion

A three-stage adaptive SVR based metamodel is explored where the training data are sequentially sampled in three-stage for improved estimate of reliability of structures. In particular, the algorithm employs importance sampling method to address the non-availability of sufficient number of simulation points near the approximated failure plane to deal with small failure probability cases. The proposed approach automatically chooses the appropriate simulation method between the brute force MCS technique and the importance sampling method based on a threshold COV of the estimated probability of failure. The advantage of the proposed approach is that it reduces the sample size requirement to estimate reliability with reasonable accuracy for very small probability of failure case also. The improved performance of the proposed approach in reliability estimation is demonstrated through three numerical examples. The results of all the three examples clearly indicate the effectiveness of the proposed algorithm in reliability estimation including estimating very small probability of failure for wide ranges of allowable limits. The effectiveness of the proposed improved stopping criterion based on the results of three consecutive iterations to avoid false convergence in estimating reliability is clearly noted in all the numerical examples. Further, failure probabilities are estimated by sequential sampling based adaptive SVR (‘Sequential SVR’) and active learning based adaptive Kriging (AK-MCS and AK-IS) methods for comparative study. The proposed approach is found to be superior over the reliability results obtained by the active learning methods in most of the cases, especially the observation is much prominent for higher dimension problem. In general, the proposed approach is found to be effective when judges with respect to the computational efficiency along with the accuracy to estimate very small probability of failure. The proposed approach being generic in nature can be readily extended to reliability analysis of nonlinear structural system. The present study is applied for SRA involving single LSF. However, the approach can be extended for SRA involving multiple LSFs which needs further study.

References

[1] Rackwitz R, Flessler B. Structural reliability under combined random load sequences. Comput Struct 1978;9:489–94. https://doi.org/10.1016/0045-7949(78)90046-9.

[2] Keshtegar B, Meng Z. A hybrid relaxed first-order reliability method for efficient structural reliability analysis. Struct Saf 2017;66:84–93. https://doi.org/10.1016/j.strusafe.2017.02.005.

[3] Kiureghian A Der, Stefano M De. Efficient Algorithm for Second-Order Reliability Analysis. J Eng Mech 1991;117:2904–23. https://doi.org/10.1061/(asce)0733-9399(1991)117:12(2904).

[4] Shinozuka M. Basic Analysis of Structural Safety. J Struct Eng 1983;109:721–40. https://doi.org/10.1061/(ASCE)0733-9445(1983)109:3(721).

[5] Engelund S, Rackwitz R. A benchmark study on importance sampling techniques in structural reliability. Struct Saf 1993;12:255–76. https://doi.org/10.1016/0167-4730(93)90056-7.

[6] Ibrahim Y. Observations on applications of importance sampling in structural reliability analysis. Struct Saf 1991;9:269–81. https://doi.org/10.1016/0167-4730(91)90049-F.

[7] Melchers RE. Radial Importance Sampling for Structural Reliability. J Eng Mech 1990;116:189–203. https://doi.org/10.1061/(ASCE)0733-9399(1990)116:1(189).

[8] Ditlevsen O, Melchers RE, Gluver H. General multi-dimensional probability integration by directional simulation. Comput Struct 1990;36:355–68. https://doi.org/10.1016/0045-7949(90)90134-N.

[9] Nie J, Ellingwood BR. Directional methods for structural reliability analysis. Struct Saf 2000;22:233–49. https://doi.org/10.1016/S0167-4730(00)00014-X.

[10] Au S-K, Beck JL. Estimation of small failure probabilities in high dimensions by subset simulation. Probabilistic Eng Mech 2001;16:263–77. https://doi.org/10.1016/S0266-8920(01)00019-4.

[11] Au SK, Ching J, Beck JL. Application of subset simulation methods to reliability benchmark problems. Struct Saf 2007;29:183–93. https://doi.org/10.1016/J.STRUSAFE.2006.07.008.

[12] Pradlwarter HJ, Schuëller GI, Koutsourelakis PS, Charmpis DC. Application of line sampling simulation method to reliability benchmark problems. Struct Saf 2007;29:208–21. https://doi.org/10.1016/J.STRUSAFE.2006.07.009.

[13] Lu Z, Song S, Yue Z, Wang J. Reliability sensitivity method by line sampling. Struct Saf 2008;30:517–32. https://doi.org/10.1016/J.STRUSAFE.2007.10.001.

[14] Pradlwarter HJ, Schuëller GI. Local Domain Monte Carlo Simulation. Struct Saf 2010;32:275–80. https://doi.org/10.1016/J.STRUSAFE.2010.03.009.

[15] Hurtado JE. Filtered importance sampling with support vector margin: A powerful method for structural reliability analysis. Struct Saf 2007;29:2–15.

[16] Lemaire M, Mohamed A. Finite element and reliability: a happy marriage? In: Nowak A., editor. Proc. 9th~IFIP WG 7.5 Work. Conf. Reliab. Optim. Struct. Syst. (Keynote Lect., 2000, p. 3–14.

[17] Faravelli L. Response-Surface Approach for Reliability Analysis. J Eng Mech 1989;115:2763–81. https://doi.org/10.1061/(ASCE)0733-9399(1989)115:12(2763).

[18] Bucher CG, Bourgund U. A fast and efficient response surface approach for structural reliability problems. Struct Saf 1990;7:57–66. https://doi.org/10.1016/0167-4730(90)90012-E.

[19] Rajashekhar MR, Ellingwood BR. A new look at the response surface approach for reliability analysis. Struct Saf 1993;12:205–20. https://doi.org/10.1016/0167-4730(93)90003-J.

[20] Gaxiola-Camacho JR, Azizsoltani H, Villegas-Mercado FJ, Haldar A. A novel reliability technique for implementation of Performance-Based Seismic Design of structures. Eng Struct 2017;142:137–47. https://doi.org/10.1016/J.ENGSTRUCT.2017.03.076.

[21] Kim C, Wang S, Choi KK. Efficient Response Surface Modeling by Using Moving Least-Squares Method and Sensitivity. AIAA J 2005;43:2404–11. https://doi.org/10.2514/1.12366.

[22] Gavin HP, Yau SC. High-order limit state functions in the response surface method for structural reliability analysis. Struct Saf 2008;30:162–79.

https://doi.org/10.1016/j.strusafe.2006.10.003.

[23] Kabasi S, Roy A, Chakraborty S. A generalized moving least square–based response surface method for efficient reliability analysis of structure. Struct Multidiscip Optim 2020:1–13. https://doi.org/10.1007/s00158-020-02743-9.

[24] Kaymaz I. Application of kriging method to structural reliability problems. Struct Saf 2005;27:133–51.

[25] Hosni Elhewy A, Mesbahi E, Pu Y. Reliability analysis of structures using neural network method. Probabilistic Eng Mech 2006;21:44–53. https://doi.org/10.1016/J.PROBENGMECH.2005.07.002.

[26] Lagaros ND, Tsompanakis Y, Psarropoulos PN, Georgopoulos EC. Computationally efficient seismic fragility analysis of geostructures. Comput Struct 2009;87:1195–203. https://doi.org/10.1016/J.COMPSTRUC.2008.12.001.

[27] Deng J. Structural reliability analysis for implicit performance function using radial basis function network. Int J Solids Struct 2006;43:3255–91. https://doi.org/10.1016/J.IJSOLSTR.2005.05.055.

[28] Keshtegar B, Kisi O. RM5Tree: Radial basis M5 model tree for accurate structural reliability analysis. Reliab Eng Syst Saf 2018;180:49–61. https://doi.org/10.1016/j.ress.2018.06.027.

[29] Sudret B, Der Kiureghian A. Comparison of finite element reliability methods. Probabilistic Eng Mech 2002;17:337–48. https://doi.org/10.1016/S0266-8920(02)00031-0.

[30] Vapnik VN. Statistical learning theory. Wiley; 1998.

[31] Vapnik VN. The nature of statistical learning theory. Springer; 2000.

[32] Hurtado JE. An examination of methods for approximating implicit limit state functions from the viewpoint of statistical learning theory. Struct Saf 2004;26:271–93.

[33] Rocco CM, Moreno JA. Fast Monte Carlo reliability evaluation using support vector machine. Reliab Eng Syst Saf 2002;76:237–43. https://doi.org/10.1016/S0951-8320(02)00015-7.

[34] Li HS, Lü Z, Yue ZF, Lu ZZ, Yue ZF. Support vector machine for structural reliability analysis. Appl Math Mech (English Ed 2006;27:1295–303. https://doi.org/10.1007/s10483-006-1001-z.

[35] Moura M das C, Zio E, Lins ID, Droguett E. Failure and reliability prediction by support vector machines regression of time series data. Reliab Eng Syst Saf 2011;96:1527–34. https://doi.org/10.1016/J.RESS.2011.06.006.

[36] Dai H, Zhang H, Wang W, Xue G. Structural Reliability Assessment by Local Approximation of Limit State Functions Using Adaptive Markov Chain Simulation and Support Vector Regression. Comput Civ Infrastruct Eng 2012;27:676–86. https://doi.org/10.1111/j.1467-8667.2012.00767.x.

[37] Richard B, Cremona C, Adelaide L. A response surface method based on support vector machines trained with an adaptive experimental design. Struct Saf 2012;39:14–21.

[38] Dai H, Zhang B, Wang W. A multiwavelet support vector regression method for efficient reliability assessment. Reliab Eng Syst Saf 2015;136:132–9. https://doi.org/10.1016/j.ress.2014.12.002.

[39] Bourinet JM. Rare-event probability estimation with adaptive support vector regression surrogates. Reliab Eng Syst Saf 2016;150:210–21. https://doi.org/10.1016/j.ress.2016.01.023.

[40] Ghosh S, Roy A, Chakraborty S. Support vector regression based metamodeling for seismic reliability analysis of structures. Appl Math Model 2018;64:584–602. https://doi.org/10.1016/j.apm.2018.07.054.

[41] Roy A, Manna R, Chakraborty S. Support vector regression based metamodeling for structural reliability analysis. Probabilistic Eng Mech 2019;55:78–89. https://doi.org/10.1016/J.PROBENGMECH.2018.11.001.

[42] Pan Q, Dias D. Sliced inverse regression-based sparse polynomial chaos expansions for reliability analysis in high dimensions. Reliab Eng Syst Saf 2017;167:484–93. https://doi.org/10.1016/j.ress.2017.06.026.

[43] Peng Y, Zhou T, Li J. Surrogate modeling immersed probability density evolution method for structural reliability analysis in high dimensions. Mech Syst Signal Process 2021;152:107366. https://doi.org/10.1016/j.ymssp.2020.107366.

[44] Goswami S, Ghosh S, Chakraborty S. Reliability analysis of structures by iterative improved

response surface method. Struct Saf 2016;60:56–66. https://doi.org/10.1016/J.STRUSAFE.2016.02.002.

[45] Farag R, Haldar A. A novel reliability evaluation method for large engineering systems. Ain Shams Eng J 2016;7:613–25. https://doi.org/10.1016/j.asej.2016.01.007.

[46] Liu H, Ong Y-S, Cai J. A survey of adaptive sampling for global metamodeling in support of simulation-based complex engineering design. Struct Multidiscip Optim 2018;57:393–416. https://doi.org/10.1007/s00158-017-1739-8.

[47] Echard B, Gayton N, Lemaire M. AK-MCS: An active learning reliability method combining Kriging and Monte Carlo Simulation. Struct Saf 2011;33:145–54.

[48] Echard B, Gayton N, Lemaire M, Relun N. A combined Importance Sampling and Kriging reliability method for small failure probabilities with time-demanding numerical models. Reliab Eng Syst Saf 2013;111:232–40. https://doi.org/10.1016/j.ress.2012.10.008.

[49] Zhang X, Wang L, Sørensen JD. REIF: A novel active-learning function toward adaptive Kriging surrogate models for structural reliability analysis. Reliab Eng Syst Saf 2019;185:440–54. https://doi.org/10.1016/j.ress.2019.01.014.

[50] Zhang X, Wang L, Sørensen JD. AKOIS: An adaptive Kriging oriented importance sampling method for structural system reliability analysis. Struct Saf 2020;82:101876. https://doi.org/10.1016/j.strusafe.2019.101876.

[51] Pan Q, Dias D. An efficient reliability method combining adaptive Support Vector Machine and Monte Carlo Simulation. Struct Saf 2017;67:85–95. https://doi.org/10.1016/j.strusafe.2017.04.006.

[52] Roussouly N, Petitjean F, Salaun M. A new adaptive response surface method for reliability analysis. Probabilistic Eng Mech 2013;32:103–15.

[53] Chakraborty S, Chowdhury R. Sequential experimental design based generalised ANOVA. J Comput Phys 2016;317:15–32. https://doi.org/10.1016/J.JCP.2016.04.042.

[54] Razaaly N, Congedo PM. Novel algorithm using Active Metamodel Learning and Importance Sampling: Application to multiple failure regions of low probability. J Comput Phys 2018;368:92–114. https://doi.org/10.1016/J.JCP.2018.04.047.

[55] Roy A, Chakraborty S. Support vector regression based metamodel by sequential adaptive sampling for reliability analysis of structures. Reliab Eng Syst Saf 2020;200:106948. https://doi.org/10.1016/j.ress.2020.106948.

[56] Smola AJ, Schölkopf B. A tutorial on support vector regression. Stat Comput 2004;14:199–222. https://doi.org/10.1023/B:STCO.0000035301.49549.88.

[57] Schölkopf B, Burges CJC, Smola AJ. Advances in kernel methods : support vector learning. MIT Press; 1999.

[58] Acar E, Rais-Rohani M. Ensemble of metamodels with optimized weight factors. Struct Multidiscip Optim 2009;37:279–94. https://doi.org/10.1007/s00158-008-0230-y.

[59] Ron Kohavi. A study of cross-validation and bootstrap for accuracy estimation and model selection. In: Mellish CS (Christopher S., editor. Proc. 14th Int. Jt. Conf. Artif. Intell. - Vol. 2, Montréal, Québec, Canada.: Morgan Kaufmann Publishers Inc. San Francisco, CA, USA ; 1995, p. 1137–43.

[60] Johnson ME, Moore LM, Ylvisaker D. Minimax and maximin distance designs. J Stat Plan Inference 1990;26:131–48. https://doi.org/10.1016/0378-3758(90)90122-B.

[61] Rackwitz R, Fiessler B. Note on discrete safety checking when using non-normal stochastic models for basic variables. Cambridge, MA, USA: 1976.

[62] Harbitz A. Efficient and accurate probability of failure calculation by the use of importance sampling technique. Int. Conf. Appl. Stat. Probab. Soil Struct. Eng., vol. 4, 1983, p. 825–86.

[63] Melchers RE. Importance sampling in structural systems. Struct Saf 1989;6:3–10. https://doi.org/10.1016/0167-4730(89)90003-9.

[64] Kiureghian A Der, Ditlevsen O. Aleatory or epistemic? Does it matter? Struct Saf 2009;31:105–12. https://doi.org/10.1016/j.strusafe.2008.06.020.

[65] Castaldo P, Gino D, Mancini G. Safety formats for non-linear finite element analysis of reinforced concrete structures: discussion, comparison and proposals. Eng Struct 2019;193:136–53. https://doi.org/10.1016/j.engstruct.2019.05.029.

[66] Au S, Wang Y. Engineering Risk Assessment with Subset Simulation. vol. 9781118398043.

Wiley; 2014. https://doi.org/10.1002/9781118398050.

[67] Choi S-K, Grandhi R V., Canfield RA. Reliability-based structural design. Springer; 2007.

[68] JMP 10 Design of Experiments Guide | Guide books n.d. https://dl.acm.org/doi/10.5555/2331344 (accessed March 22, 2021).

[69] Fang K-T, Lin DKJ, Winker P, Zhang Y. Uniform Design: Theory and Application. Technometrics 2000;42:237–48. https://doi.org/10.1080/00401706.2000.10486045.

[70] Keshtegar B. A hybrid conjugate finite-step length method for robust and efficient reliability analysis. Appl Math Model 2017;45:226–37. https://doi.org/10.1016/j.apm.2016.12.027.