A Generalized Moving Least Square based Response Surface Method for Efficient Reliability Analysis of Structure

Sounak Kabasi*[1] Atin Roy[2] and Subrata Chakraborty[3]

[1]Department of Civil Engineering, Indian Institute of Technology, Bombay, Maharashtra, India, E-mail: sounak.kabasi001@gmail.com

[2]Department of Civil Engineering, Indian Institute of Engineering Science and Technology, Shibpur, West Bengal, India, **E-mail:** atin.3222@yahoo.com

[3]Department of Civil Engineering, Indian Institute of Engineering Science and Technology, Shibpur, West Bengal, India, **E-mail:** schak@civil.iiests.ac.in **ORCID 0000-0002-4792-3844 (Corresponding Author)**

Abstract

To improve the efficiency and accuracy of traditional least squares method based polynomial response surface method (RSM) for reliability analysis of structure, the application of various adaptive metamodeling approaches are notable. The moving least squares method (MLSM) based RSM is the simplest and found to be effective in this regard. But, its performance in reliability analysis of structure largely depends on proper choice of the parameter of weight function involved. In the present study, a generalized scheme to appropriately obtain the hyper-parameter of the MLSM based RSM to approximate implicit responses of structure for reliability analysis is proposed. The algorithm is hinged on the fact that for reliability analysis, one is interested on the sign of the approximated limit state function (LSF) rather than its magnitude. Thereby, it is sufficient to obtain the hyper-parameter for which the first derivative of probability of failure as obtained from the approximated LSF with respect to the hyper-parameter is zero. The effectiveness of the proposed algorithm is elucidated through three numerical examples. The improvement achieved by the proposed MLSM based RSM has been compared with the reliability results obtained by the MLSM based RSM considering the commonly recommended value of the hyper-parameter and also by the approach where the parameters are obtained by leave one out cross validation procedure.

Keywords: Reliability Analysis, Monte Carlo Simulation, Response Surface Method, Moving Least Squares Method, Hyper-parameter.

---

*Former Master Student, Department of Civil Engineering, Indian Institute of Engineering Science and Technology, Shibpur, West Bengal, India.

1. Introduction

The numerical simulation based structural reliability analysis (SRA) approach has emerged as an integrated platform for safety assessment of structures. The numerical simulation-based SRA involves comprehensive engineering analysis with rigorous treatment of uncertainty due to external loads and structural properties for computation of probability of failure (P_f). Such approaches can be classified into two groups namely the second moment based analytical methods i.e. the first order and the second order reliability analysis methods and the method based on Monte Carlo Simulation (MCS) technique. All such methods are well established and documented in numerous texts (Ditlevsen and Madsen 1996; Melchers 1990; Haldar and Mahadevan 2000). In fact, the first order reliability method is widely applied for estimating P_f values due to its simplicity and computational efficiency. However, it has known drawbacks with regard to the assumption about the shape of failure surface and accuracy of approximation. The most accurate and conceptually straightforward means of SRA is based on MCS technique. The approach is preferred as it does not require an assumption about the shape of the failure surface. But, such full simulation based approach requires a large number of repetitive evaluations of limit state function (LSF) for an acceptable confidence in reliability estimate. If the LSF is available in close form, the numbers of performance function calls do not play an important role. But, the performance behaviour of large complex structural system in explicit form is often unavailable. Though, reliability analysis of such structure can be carried out by MCS, each performance function evaluation typically involves analysis of a high fidelity numerical model e.g. a finite element model. This will require a high computational cost, especially for large complex structures. The enormous time requirement to extract necessary response quantity of interest for statistical analysis is studied by Kwon and Elnashai (2006). The numbers of simulations necessary might be of the order of several thousand for sufficient accurate estimate of reliability of a structure, depending on the function being evaluated and the magnitude of probability of failure (Mann et al. 1974). As an effective alternative, the polynomial Response Surface Method (RSM) based metamodeling approach is widely adopted to overcome the computational challenge of MCS based reliability analysis of large complex system involving implicit LSF. It simplifies the process by fitting a polynomial model i.e. 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 performing structural analysis for a smaller but sufficient number of times. By using polynomial approximations of the actual LSF, the computational involvement of analyses can be reduced drastically. The RSM adopted in most of the studies for SRA is primarily based on global approximation of scatter position data, obtained by the least squares method (LSM). But, the LSM is a major source of error in approximation (Kim et al. 2005). The predicted responses by the LSM based RSM may fail to capture the actual trend of the responses within a local domain as the efficiency of global approximation depends largely on the selection of basis functions. Such functions should represent as closely as possible to the actual variation of the response within a domain. But such a selection is not obvious as the response is implicit in nature (Gavin and Yau 2007). In the LSM based RSM, typically, the second order polynomials are employed. The higher order polynomials are usually avoided as this invites overfitting and demand huge numbers of data to overcome numerical singularity. The data requirement for linear polynomial is the minimum but it does not take into account the nonlinearity of the LSF. The second order polynomial can approximate the nonlinearity and adopted in RSM application as a balance choice with regard to accuracy and efficiency. But, the traditional polynomial RSM needs a considerable numbers of training data to ensure necessary accuracy and becomes computationally intensive for real engineering problems involving many variables. To improve the efficiency and accuracy of LSM based RSM, the application of various adaptive metamodeling approaches e.g. quadratic polynomial based moving least square method (MLSM) (Breitkopf et al. 2005; Kim et al. 2005; Most and Bucher 2005; Kang et al. 2010; Taflanidis and Cheung 2012; Goswami et al. 2016; Ghosh and Chakraborty 2017), artificial neural network (Chojaczyk 2015; Elhewy 2006), Kriging method (Kaymaz 2005; Echard et al. 2011), polynomial Chaos Expansion (Notin et al. 2010; Chakraborty and Chowdhury 2016; Schöbi et al. 2017), support vector machine (Hurtado and Alvarez 2003; Guo and Bai 2009; Roy et al. 2019) etc. are notable. The MLSM based RSM is the simplest amongst these and found to be effective in this regard (Kim et al. 2005; Kang et al. 2010; Goswami et al. 2016). The present study focuses on the MLSM based adaptive RSM for SRA application.

The MLSM based metamodeling approach primarily follows the basic idea of the LSM based RSM but overcomes the disadvantages of global approximation nature. This is achieved by making a local approximation of the polynomial coefficients around each point in the interpolation domain by using a varying weight function with respect to the position of approximation. The weight associated with a particular sampling point decays as the point of approximation moves away from the sampling point. Thereby, introduction of dependence of response approximation on the weights helps to reduce the approximation error at each point by performing a weighted local averaging of the information obtained from the support points closer to it. This leads to a smaller dependence of fit on the type of basis functions used (Youn and Choi 2004; Bucher and Most 2008). Thus, such adaptive capability of providing higher weights to response values at nearby data points makes the approach advantageous over the LSM based RSM. In this regard, it is important to note that the accuracy of MLSM based RSM largely depends on suitable choice of weight function and its parameters. The weight function should prioritize support points closer to the interpolation point and should vanish after a certain influence radius. To achieve this, various weight decay functions e.g. Gaussian function, cubic, fourth, fifth and seventh order polynomials are investigated with regard to their effectiveness. The Gaussian function is noted to be the most wide choice (Toropov *et al*. 2005) and many of the literatures suggested the use of Gaussian function (Haussler-Combe and Korn 1998; Karutz 2000; Haussler-Combe 2001; Most and Bucher 2005). The parameter involved in the Gaussian function representing its extent to control the variation of weight value is considered as 0.4 in all of these studies. A more generic exponential form of the Gaussian type was proposed by Taflanidis and Cheung (2012) for SRA problems. For a well-fitted metamodel, the best choice of this parameter can be attained by minimizing the actual measure of fit which can only be obtained by performing direct MCS study. Thereby, to avoid further computational burden, the parameters should be obtained by using the response values already obtained at the design of experiment (DOE) points only. As the actual measure of fit of the model is unknown, a generalized root mean squared error (GRMSE) may be taken as a substitute which can be obtained by the cross validation method. The GRMSE value can be minimized to obtain the best possible choice of the hyper-parameter. The applications of such algorithm to obtain the parameters in various soft computing-based metamodeling approaches are well known (Jiang et al. 2015; Su et al. 2016). It is also explored for MLSM based metamodel (Roy et al. 2018; Ghosh et al. 2018). The approach has the capability to estimate optimum model parameter. However, the use of GRMSE as a precision measure is found to be suitable for global response approximation and useful for optimization study (Goel et al. 2007). But, it may not be appropriate for SRA applications as proper estimation of probability of failure requires approximation precision at local domain i.e. around the region where the value of the LSF is close to zero and the contribution of probability of failure is the maximum.

In the present study, a generalized scheme to appropriately obtain the hyper-parameter of the weight function involved in the MLSM based RSM for SRA application is explored. The algorithm is hinged on the fact that for SRA application, one is interested about the sign of an approximated LSF rather than its magnitude. Thus, it is sufficient to obtain the parameter for which the first derivative of the probability of failure (obtained based on the approximated LSF) with respect to the hyper-parameter is zero (Kabasi and Chakraborty 2019). Following this, a search algorithm is proposed to obtain the hyper-parameter. The effectiveness of the proposed algorithm to construct the MLSM based response surface for SRA is elucidated by considering three numerical examples. The improved capability of the proposed approach to estimate reliability is demonstrated by comparing with the reliability results obtained by the MLSM based RSM based on commonly recommended value of the hyper-parameter and also based on the hyper-parameter obtained by the leave one out cross validation procedure considering the direct MCS based reliability results as the benchmark.

2 The Moving Least Squares Method Based RSM

The MLSM based RSM is a weighted LSM that uses different weights for different positions of approximations. Thus, the coefficients of metamodel change with the change of approximation point of interest. This procedure is interpreted as local approximation. The polynomial function is a function that can be expressed as $f\left(x,\beta\right)$, where $\beta=\left[\beta_0,\beta_1,\ldots,\beta_{m-1}\right]^T$ is the vector of the m coefficients to be tuned and $x=\left[x_1,x_1,\ldots,x_n\right]^T$, where x_i is the vector of the n input parameters. The data set consists of $\left(s_i,y(s_i)\right)$ pairs, $i=1,2,...,p,$ where s_i is the i^{th} DOE point associated with response variable $y(s_i)$. The response surface model in terms of the observations can be written in matrix notation as,

$$Y=\hat{Y}+e=\mathbf{F\beta}+e \tag{1}$$

Where, Y is a vector of actual responses, **F** is the design matrix consists of $\left(s_i,y(s_i)\right)$, $\boldsymbol{\beta}$ is a vector of unknown parameters and e is a vector of error terms. The method of least squares is typically used to estimate the coefficients so that it best fits a data set. For the best fit of the polynomial to scatter data, the LSM evaluates the unknown coefficients by minimizing the sum of the squared error (S_e). However, in the MLSM approach, the sum of the squared error S_e is defined as the sum of the weighted errors as following,

$$S_e=\sum_{i=1}^{N}w_ie_i^2=\left(y_i-\mathbf{F\beta}\right)^TW\left(x\right)\left(y_i-\mathbf{F\beta}\right) \tag{2}$$

The coefficient **β**(**x**) can be obtained by the matrix operation as below,

$$\boldsymbol{\beta}=\left(\mathbf{F}^T\mathbf{W}(x)\mathbf{F}\right)^{-1}\mathbf{F}^T\mathbf{W}(x)y \tag{3}$$

In the above equation, **W(x)** is the diagonal matrix of the weight function and it depends on the location of the associated approximation point of interest (**x**); *N* is the total number of sample points. The weight matrix **W(x)** can be constructed by using the weighting function in the diagonal terms. It can be noted from Eq. (3) that the coefficient **β(x)** is a function of the location or position **x**. Thus, the procedure to calculate '**β(x)**' is a local approximation, and "moving" processes performs a global approximation over the entire design domain. Note that in the MLSM based approach, there exist separate approximation functions for each calculation point of interest. Whereas, the LSM yields a fixed approximation function for a set of training data set. As already mentioned, the effectiveness of the MLSM based RSM largely depends on the weighing function chosen.

2.1 The Gaussian Weight Function

The Gaussian type function has been used extensively as a suitable choice for weighing function necessary in the MLSM approach and (Toropov *et al*. 2005; Haussler-Combe and Korn 1998; Karutz 2000; Haussler-Combe 2001; Most and Bucher 2005). The weight matrix, **W(x)** can be obtained by utilizing the Gaussian weighting function of the following form,

$$w(x-x_I)=w(d)=\frac{e^{-\left(d/cD_I\right)^2}-e^{-1/c^2}}{1-e^{-1/c^2}} \tag{4}$$

In the above equation, d is the distance of the point where the approximate response is required (**x**) to the point of DOE, x_I represents the ordinate of the Ith DOE point and D_I is the approximate radius of sphere of influence. An appropriate support size D_I should be selected at any point **x** so that a minimum number of neighbouring support points are included to avoid ill-conditioning in the system of equations. This implies that D_I should include at least sufficient number of neighbouring experimental points so as to avoid singularity (Taflanidis and Cheung 2012). The more details about the calculation of support size can be found elsewhere (Kim et al. 2005; Kang et al. 2010). The parameter, c in the above is taken as 0.4 by all these studies. A more generic exponential form of the above as following proposed by Taflanidis and Cheung (2012) for SRA problem,

$$w(x-x_I)=w(d)=\frac{e^{-\left(d/cD_I\right)^{2k}}-e^{-(1/c)^{2k}}}{1-e^{-(1/c)^2}} \tag{5}$$

The above equation exactly mimics the former i.e. Eq. (4) if the parameter k is equal to unity. Taflanidis and Cheung (2012) also recommended the value of the parameter, c as 0.4

as suggested in the earlier studies. The selection of such value of the parameter appears to be a balance choice. If a higher value (>1) of the parameter is chosen, the MLSM approaches towards the conventional LSM based RSM i.e., the local approximation is no longer working. On the other hand, when a very low value of c (<0.1) is taken then very few DOE points close to the point of interest get largely weighted compared to the other DOE points. Hence, the value of 0.4 seems to be a balance between a local approximation and accuracy by avoiding near-singular problem. However, it cannot be considered to be the best choice.

2.2 The Cross Validation Approach

The cross validation approach to obtain the hyper-parameters to construct metamodels is frequently applied. Usually, a search algorithm by solving an optimization sub-problem where the mean square error value obtained by the leave one out cross validation method is minimized to obtain the best possible optimum choice of the parameter. In the leave one out cross validation approach, only one data point is taken out from the original DOE points to build a metamodel based on all the remaining DOE points and evaluate the error on the single-point held out. The GRMSE is obtained by repeating this procedure for each of the training points available for cross validation i.e. it is obtained as:

$$GRMSE = \sqrt{\frac{1}{p}\sum_{i=1}^{p}\left(f_i - \hat{f}_i^{i-1}\right)^2} \tag{6}$$

Where, $\hat{f}_i^{i-1}$ represents the prediction at the i-th sample point using the metamodel constructed using all the sample points except the i-th sample point; p is the total number of sample points and f_i is the actual response obtained from the numerical model at the i-th sample point. The optimization sub-problem is solved where the GRMSE value as obtained from Eq. (6) is minimized to obtain the best possible choice of the weight parameter.

3. The Proposed approach of hyper-parameter tuning

As already discussed, the hyper-parameter obtained by minimizing the GRMSE as a measure of forecasting error obtained by the cross validation method may not be appropriate for SRA. In this regard, it is important to realize that for successful computation of reliability by RSM, the sign of the approximated LSF is important, not its magnitude (Kabasi and Chakraborty 2019). Relying on this, a search algorithm is investigated to obtain the hyper-parameter. Specifically, a generalized scheme to obtain the optimum value of the hyper-parameter of the MLSM based RSM to approximate the implicit responses of a structure for reliability analysis is proposed. To elucidate the procedure, a variable F_i is defined as following,

$$F_i = \frac{g(\mathbf{x}_i)}{|g(\mathbf{x}_i)|} \times \frac{\tilde{g}(\mathbf{x}_i)}{|\tilde{g}(\mathbf{x}_i)|} \tag{7}$$

Where, $g(\mathbf{x}_i)$ denote the values of the actual LSF (not known beforehand) and $\tilde{g}(\mathbf{x}_i)$ is its approximated value obtained from the metamodel at a simulation point, $\mathbf{x}_i$. The absolute values of $g(\mathbf{x}_i)$ and $\tilde{g}(\mathbf{x}_i)$ are denoted by $|g(\mathbf{x}_i)|$ and $|\tilde{g}(\mathbf{x}_i)|$, respectively. It can be noted that at a particular point $\mathbf{x}_i$, the value of F_i can be either +1 i.e. $g(\mathbf{x}_i)$ and $\tilde{g}(\mathbf{x}_i)$ are of the same sign or -1 i.e. $g(\mathbf{x}_i)$ and $\tilde{g}(\mathbf{x}_i)$ are of the opposite sign. It is important to note here that to obtain the probability of failure, one is interested only on the sign of $\tilde{g}(\mathbf{x}_i)$ rather than its magnitude. If $g(\mathbf{x}_i)$ and $\tilde{g}(\mathbf{x}_i)$ are of the same sign at a particular simulation point, $\mathbf{x}_i$, then it can be concluded that $\tilde{g}(\mathbf{x}_i)$ has successfully predicted the response to decide the failure or safety of the system at the said point. Thus, one gets F_i as +1 in case of a successful prediction and F_i as -1 in case of an unsuccessful prediction. Now, a new variable *S* is further defined as following,

$$S = \frac{\sum_{i=1}^{n} F_i}{n} \tag{8}$$

Where, *n* is the total number of simulations performed for reliability estimate. The variable *S* gives a measure of accuracy of predicting the sign of the LSF based on the metamodel at all the MCS points. Thus, if $g(\mathbf{x}_i)$ and $\tilde{g}(\mathbf{x}_i)$ are of the same sign at a MCS trial point $\mathbf{x}_i$, the accuracy of prediction by the metamodel is acceptable. It can be readily realized from Eq. (8) that if most of the F_i values take the value of +1; the best prediction of sign of the LSF approximated by the metamodel, $\tilde{g}(\mathbf{x}_i)$ is achieved. Thus, to achieve the best possible prediction of sign of the LSF approximated by the metamodel, $\tilde{g}(\mathbf{x}_i)$, the value of *S* should be maximum. It can be noted from Eqs. (7) and (8) that the accuracy of approximation by the response surface at a point depends on the hyper-parameter involved in the response surface i.e. the function $\tilde{g}(\mathbf{x}_i)$ depends on *c*. Thus, to satisfy the condition that *S* be maximum, the parameter *c* of the metamodel should be such that $\frac{\partial S}{\partial c} = 0$, which gives,

$$\frac{\partial\left(\sum_{i=1}^{n}\frac{g(\mathbf{x}_i)}{|g(\mathbf{x}_i)|}\times\frac{\tilde{g}(\mathbf{x}_i)}{|\tilde{g}(\mathbf{x}_i)|}\right)}{\partial c}=0 \quad \text{i.e. } \sum_{i=1}^{n}K_i\times\frac{\partial\left(\frac{\tilde{g}(\mathbf{x}_i)}{|\tilde{g}(\mathbf{x}_i)|}\right)}{\partial c}=0 \tag{9}$$

In the above, $K_i=\frac{g(\mathbf{x}_i)}{|g(\mathbf{x}_i)|}$ is a constant and independent of c and the value of K_i is not known beforehand. However, $\frac{\partial S}{\partial c}=0$ indicates that the rate of change of sign prediction accuracy with respect to c is zero i.e. $\frac{\partial S}{\partial c}=0$ is valid for such values of c for which the rate of change of probability of failure with respect to the parameter is zero. As evaluating P_f can be readily performed based on the metamodel; the parameter, c can be obtained by satisfying

$$\frac{\partial P_f}{\partial c}=0 \tag{10}$$

Note that the above cannot be applied analytically to obtain the best value of c. But, the value of P_f can be readily obtained by MCS technique utilizing the metamodel once a particular value of the parameter, c is known. Thus, a search scheme can be employed to obtain the value of c satisfying Eq. (10). For this, the search domain i.e. the bounds of c are first selected. For Gaussian weight function, the value of the parameter is approximately around the value of 0.4. Thus, a choice of the parameter, c in the range of 0.1 to 0.8 will be reasonable. Now, for a suitable step size of h (if it is evident from the rough trend that P_f varies sharply with c, then a smaller step size is necessary), the following is evaluated,

$$T_c=\frac{\sum_{i=1}^{n}\left|P_{f_c}^{i}-P_{f_{c+h}}^{i}\right|}{P_{f_c}^{i}} \tag{11}$$

Where, $P_{f_c}^{i}$ and $P_{f_{c+h}}^{i}$ represent the probabilities of failures obtained by the metamodel with parameter values of c and $c+h$, respectively for i-th MCS seed. To satisfy Eq. (10), the value of T_c should be ideally zero. However, the numerical approximation in Eq. (11) will not exactly satisfy the condition stated in Eq. (10) and need to settle for a very small value of T_c for such approximation. In this regard, it may be noted here that multiple values of c may satisfy the above conditions. In such cases, a leave one out cross validation scheme can be employed in order to choose the hyper-parameter appropriately. The GMRSE is evaluated by leave one out cross validation approach for each such values of the parameter, c. The value of *c* for which the minimum value of GRMSE is obtained is used to generate the final metamodel for reliability analysis of structure.

4. Numerical Study

The effectiveness of the proposed algorithm to appropriately obtain the hyper-parameter of the Gaussian weight function to construct the MLSM based metamodel for SRA application is elucidated numerically by considering three numerical examples. The improved capability to estimate reliability by the proposed approach with respect to the two existing approaches is studied. For this, the reliability results obtained by the MLSM based RSM by adopting commonly recommended value of the hyper-parameter (i.e. $c = 0.4$) and also by obtaining the hyper-parameter by the leave one out cross validation procedure are also obtained and compared with the most accurate direct MCS based reliability results.

4.1 Example 1: A sphere subjected to internal pressure

This first example concerns with the approximation of Von-Mises stress of a sphere under internal pressure as shown in Fig.1. The material is homogenous without spatial variation. The equivalent Von-Misses stress equation can be expressed as,

$$\sigma_{eq}(r = r_0) = \sqrt{(\sigma_r^2 + \sigma_\theta^2 - 2\sigma_r\sigma_\theta)} = (|\sigma_r - \sigma_\theta|) = \frac{3p_0 r_1^3}{2(r_1^3 - r_0^3)} \tag{12}$$

Where, r_0 and r_1 are the internal and external radiuses, respectively, p_0 is the internal pressure, σ_r and σ_θ are the radial and circumferential stresses, respectively.

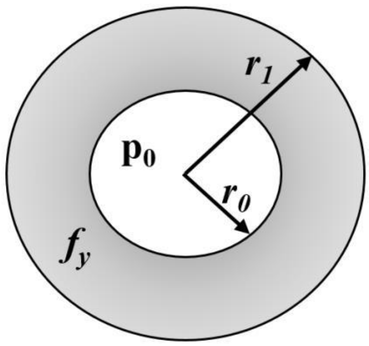

Fig. 1 The sphere under internal pressure

The variables, p_0, r_0 and r_1 are assumed to follow lognormal distribution having mean values of 120 MPa, 50mm and 100 mm, respectively. Unless mentioned otherwise, the coefficient of variation (COV) of all the random parameters are assumed to be 7.5%. The LSF considered is with regard to the elasto-plastic failure i.e. the Von-Mises stress reaches the yield stress f_y at a point of the sphere which can be expressed as,

$$g = f_y - \sigma_{eq} = f_y - \frac{3p_0 r_1^3}{2(r_1^3 - r_0^3)} \tag{13}$$

To study the effectiveness of the proposed algorithm, the reliability results are obtained by the proposed approach are compared with that of obtained by the leave one out cross validation scheme and also by the mostly adopted value of the hyper-parameter i.e. c =0.4. The results are presented in terms of reliability index, β defined as, $\beta = \phi^{-1}(1-P_f)$ where, $\phi^{-1}(1-P_f)$ is the standard normal variate at the probability level (1-P_f) (Haldar and Mahadevan 2000). The reliability results with varying yield stress obtained by the proposed approach and the two existing approaches considering 20 and 25 training data obtained by the Uniform Design (UD) scheme (Fang et al. 2000) are compared in Fig. 2. The mean values of r_0 and r_1 are taken as 50mm and 100mm, respectively. The improved capability of the proposed approach with respect to the existing MLSM based RSM approaches when compared with the most accurate direct MCS based results can be readily noted from the plots. As expected, the accuracy of the estimated reliabilities improved (considering the direct MCS as the benchmark) as the number of training data points increases. The corresponding hyper-parameters obtained by the proposed approach for varying mean value of f_y considering 20 and 25 UD training data set are shown in Table 1. It may be observed that there is notable variation in the optimum value of the hyper-parameters and as such there is no specific trend.

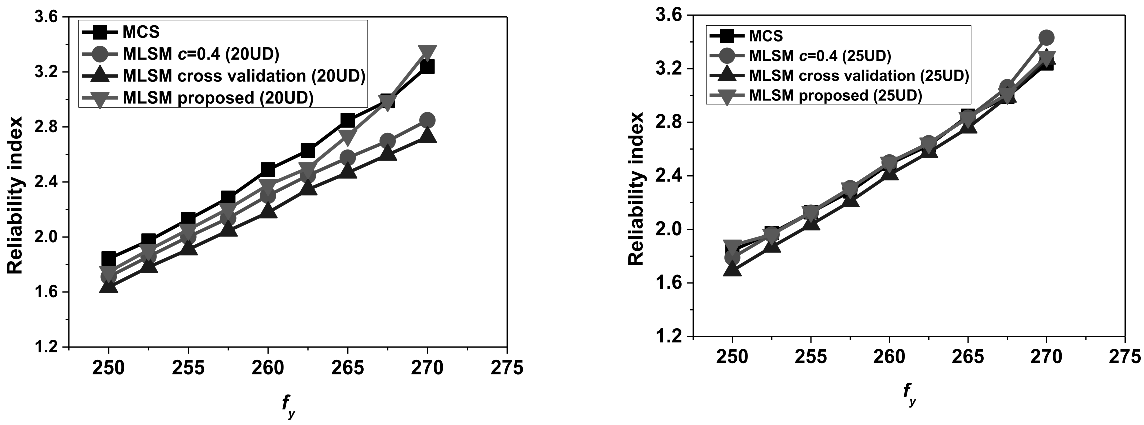

Fig. 2 The comparison of reliability index of the sphere for varying mean value of f_y considering 20 and 25 UD training data set.

Table 1 The hyper-parameters obtained by the proposed approach for varying mean value of f_y considering 20 and 25 UD training data set

UD	f_y				
	250	255	260	265	270
20	0.21	0.29	0.31	0.22	0.22
25	0.23	0.19	0.23	0.30	0.35

To study the effectiveness of the proposed approach to estimate reliability with respect to the two existing approaches, the comparison of percent error for varying mean value of f_y considering different order of the polynomial basis function with 25 UD training data are shown in Table 2. The percent errors are calculated by considering the direct MCS based results as the benchmark. As expected, the percent errors reduce in all the three approaches when the higher order polynomials are used. However, third order polynomial is not a wise choice as it may results in overfitting problem. It is of worth noting here that the results clearly demonstrate improved performance of the proposed approach with respect to the existing approaches for both higher and lower order polynomial basis.

Table 2 The comparison of percent error for varying mean value of f_y considering different order of polynomial basis function with 25 UD training data

f_y		**Percent error**		
		MLSM *c*=**0.4**	**MLSM cross validation**	**MLSM proposed**
250	1st order	18.6	18.6	18.4
	2nd order	3.87	9.26	3.42
	3rd order	3.10	4.47	2.23
257.5	1st order	12.8	12.8	10.03
	2nd order	7.24	3.56	2.25
	3rd order	2.71	3.19	1.89
267.5	1st order	9.6	9.32	8.97
	2nd order	8.47	4.43	1.99
	3rd order	1.82	2.26	1.78
Mean Error	1st order	13.67	13.57	12.47
	2nd order	6.53	5.75	2.55
	3rd order	2.54	3.31	1.97

In metamodeling approach, the effect of choice of training data based on a chosen DOE scheme is an important issue (Mohammed and Felician 2020). To study the sensitivity of the effectiveness of the proposed MLSM scheme of reliability estimate for different training data set, the scatter in the results with different sets of training data obtained by the UD scheme are studied further. Table 3 compares the percent error of estimated reliability of the sphere for varying mean value of f_y for ten different sets of 25 UD training data. It can be noted that the percentage of errors in estimated reliabilities vary for different sets of training data by the proposed approach as well as by the two existing approaches. However, it has been consistently observed that the percentage of errors in estimated reliabilities by the proposed approach are less compare to that of obtained by the existing approaches for all the training data set. This clearly indicated the robustness of the proposed algorithm in improve estimate of reliability.

Table 3 The comparison of percentage of error of estimated reliability of the sphere for varying mean value of f_y for ten different sets of 25 UD training data

f_y	Set	1	2	3	4	5	6	7	8	9	10
250	*c*=0.4	2.899	2.999	3.326	3.869	3.360	3.435	3.197	3.271	3.406	3.532
	Cross validation	8.156	8.156	8.174	9.260	8.599	8.603	8.274	8.382	8.432	8.713
	Proposed approach	2.078	1.535	2.877	3.420	2.532	2.746	2.369	2.315	2.817	2.880
260	*c*=0.4	0.453	0.462	1.503	2.748	1.479	1.688	1.105	1.232	1.645	1.917
	Cross validation	3.229	3.229	3.229	4.435	3.711	3.711	3.350	3.470	3.519	3.832
	Proposed approach	0.453	0.051	0.378	0.780	0.455	0.481	0.335	0.295	0.433	0.513
270	*c*=0.4	5.951	5.971	5.626	6.555	6.166	6.099	5.887	5.984	5.900	6.160
	Cross validation	1.038	1.656	3.575	3.883	2.615	3.061	2.522	2.677	3.349	3.345
	Proposed approach	1.595	0.977	2.267	2.576	1.869	2.065	1.776	1.684	2.160	2.163
Mean Error	*c*=0.4	3.594									
	Cross validation	4.940									
	Proposed approach	1.629									

Further, the comparison of the estimated reliability index with varying mean value of p_0 is shown in Fig.3. The value of f_y has been considered to be 260 kN/mm^2 to develop this plot. The improved performance of the proposed approach can be readily noted from the plot.

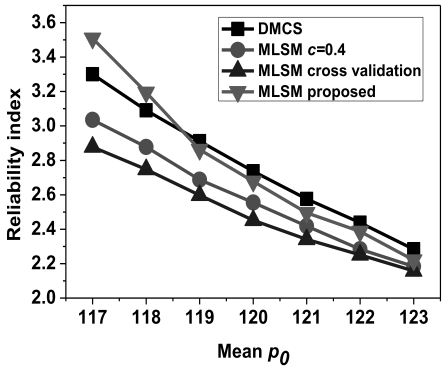

Fig. 3 The comparison of reliability indices of the sphere for varying mean value of p_0 considering 20 UD training data set

4.2 Example 2: A simply supported I-beam

The second example considered is a simply supported I-beam subjected to a concentrated load P at a distance A from the left end as shown in Fig.4. For reliability analysis, the LSF is considered in terms of the maximum stress under bending moment as following (Keshtegar, 2017),

$$g = SS - \frac{12PA}{2L}\left[\frac{(L-A)d}{b_f d^3 - (b_f - t_w)(d - 2t_f)^3}\right] \tag{14}$$

Where, SS is the moment resisting capacity of the beam and the other parameters are as depicted in the Fig. 4. The details of the various uncorrelated random parameters are furnished in Table 4.

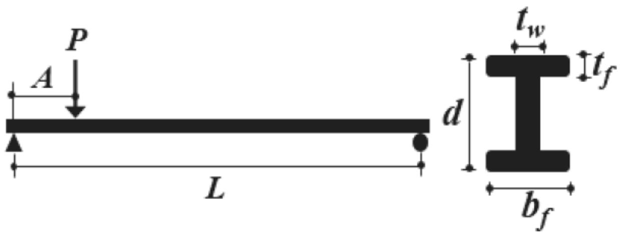

Fig. 4 The simply supported I-Beam

Table 4 The details of the random parameters

Variable	Distribution	Mean	SD
P (N)	Normal	6500	300
L(mm)	Normal	120	5
A(mm)	Log-normal	72	5
SS(MPa)	Gumbel	170000	4760
d(mm)	Normal	2.3	0.04
b_f(mm)	Gumbel	2.97	0.075
t_w(mm)	Normal	0.16	0.02
t_f(mm)	Normal	0.26	0.02

The reliability results obtained by the proposed approach and the other two commonly used approaches considering the UD scheme with 55 and 59 training data are shown in Fig. 5 for varying moment resisting capacity (*SS*) of the beam. The corresponding hyper-parameters obtained by the proposed approach for varying mean value of *SS* considering 55 and 59 UD training data set are given in Table 5. The improved performance of the proposed approach is noted for this case also.

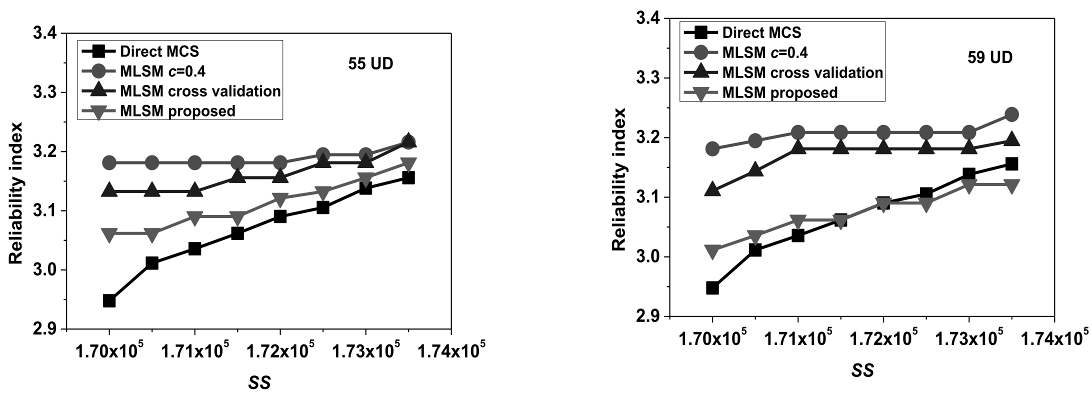

Fig. 5 The comparison of reliability indices of the simply supported I-Beam for varying moment resisting capacity considering 55 and 59 UD training data set

Table 5 The hyper-parameters obtained by the proposed approach for varying mean value of *SS* considering 55 and 59 UD training data set

UD	*SS*							
	170000	170500	171000	171500	172000	172500	173000	173500
55	0.32	0.30	0.35	0.43	0.33	0.36	0.33	0.35
59	0.29	0.32	0.37	0.36	0.37	0.32	0.38	0.43

The comparison of percent error for varying mean value of *SS* considering different order of the polynomial basis function with 55 UD training data are shown in Table 6. The improved performance of the proposed approach with respect to the existing approaches is noted for lower as well as for higher order polynomial for this example also.

Table 6 The comparison of percent error for varying mean value of *SS* considering different order polynomial basis functions with 55 UD training data

SS		**Percent error**		
		MLSM, *c*=0.4	**MLSM cross validation**	**MLSM proposed**
170500	1st order	16.87	17.10	15.95
	2nd order	4.29	4.93	2.57
	3rd order	4.15	4.60	1.81
172000	1st order	13.01	13.20	10.89
	2nd order	5.26	5.26	3.64
	3rd order	3.34	4.19	3.35
173000	1st order	11.27	10.37	10.42
	2nd order	5.56	5.27	4.25
	3rd order	4.89	2.30	2.42
Mean Error	1st order	13.72	13.56	12.42
	2nd order	5.04	5.15	3.49
	3rd order	4.13	3.70	2.53

To study the sensitivity of the effectiveness of the proposed MLSM scheme, the percent errors in the estimated reliabilities of the simply supported I-Beam with varying moment resisting capacity for ten different sets of 55 UD training data are compared in Table 7. Like the previous example, the percentage of errors in estimated reliabilities vary for different sets of training data by all the three approaches. However, the percentages of errors in estimated reliabilities by the proposed approach are noted to be less compared to that of obtained by the existing approaches for all the training data set.

Table 7 The comparison of percent error of estimated reliability of the simply supported I-Beam with varying moment resisting capacity for ten different sets of 55 training data.

SS	**Set**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
170500	*c*=0.4	5.637	5.139	4.285	5.504	5.019	5.127	5.128	4.822	5.261	5.078
	Cross validation	4.019	4.162	4.930	3.797	4.327	4.192	4.218	4.436	4.036	4.206
	Proposed approach	1.672	2.004	2.572	3.001	2.208	2.384	2.507	2.587	2.679	2.417
172000	*c*=0.4	2.944	2.944	5.264	3.915	3.737	3.699	4.028	4.395	3.838	3.699
	Cross validation	2.125	2.449	5.264	3.419	3.326	3.271	3.617	4.148	3.400	3.303
	Proposed approach	1.008	1.128	3.643	1.656	1.863	1.729	2.058	2.520	1.746	1.729
173500	*c*=0.4	1.903	1.898	4.990	3.482	2.985	2.992	3.460	3.919	3.266	2.992
	Cross validation	1.903	1.901	4.990	2.851	2.923	2.804	3.208	3.731	2.882	2.804
	Proposed approach	0.802	0.802	3.793	2.069	1.826	1.780	2.206	2.677	1.964	1.780

Mean Error	c=0.4	4.045
	Cross validation	3.555
	Proposed approach	2.094

Further, Fig. 6 compares the reliability index of the simply supported I-Beam for varying mean value of P considering 50 UD training data set. The improved performance of the proposed approach compared to the existing approaches is quite apparent from the plot.

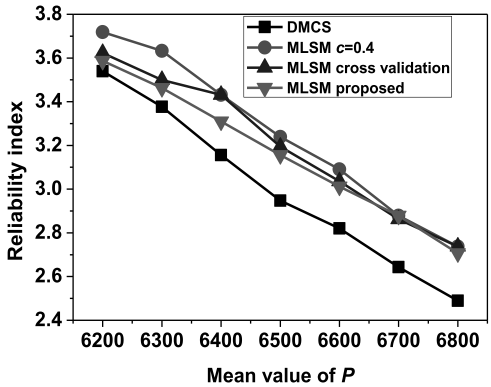

Fig.6 The comparison of reliability index of the simply supported I-Beam for varying mean value of P considering 50 UD training data set

4.3 Example 3: A space dome truss

A space dome truss as shown in Fig. 7 is further taken up as the final example to demonstrate the effectiveness of the proposed algorithm (Keshtegar, 2017). The LSF for this problem is implicit in nature which is defined with respect to the maximum vertical displacement of the node under load P_1 as,

$$g = \Delta_{allow} - \left|\Delta_{P1}^{z}\right| \tag{15}$$

Where, Δ_{P1}^{z} is the maximum vertical displacement of the node and Δ_{allow} is the allowable displacement. The six independent random variables considered are: the Young's modulus of the material of all the bars, the cross section areas of the top radial bars, A_1 (bars 1 to 6), the peripheral bars, A_2 (bars 7 to 12), the bottom inclined bars, A_3 (bars 13 to 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 statistical properties of the random variables are depicted in the Table 8. The response of the finite element model of the space truss is obtained by using ANSYS mechanical APDL module.

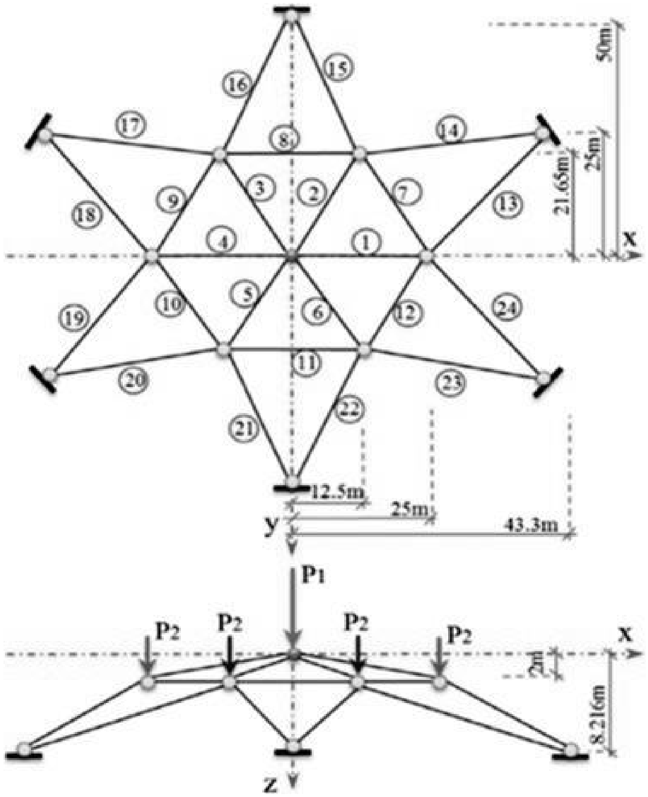

Fig. 7 The space dome truss truss (redrawn from Keshtegar, 2017)

Table 8 The characteristics of the uncorrelated random variables of the space dome truss

Variable	Distribution	Mean	Covariance	Comment
P_1	Gumbel Max	20(1+U/100) kN	0.15	Point load at centre node
P_2	Gumbel Max	10(1+U/100) kN	0.12	Point load at 6 nodes of mid hexagon
A_1	Normal	0.013 m^2	0.1	Top radial member's cross section
A_2	Normal	0.01 m^2	0.1	Peripheral member's cross section
A_3	Normal	0.016 m^2	0.1	Bottom inclined member's cross section
E	Normal	205 GPa	0.05	Young's Modulus

The reliability results obtained by the proposed approach and the two existing approaches considering 30 DOE points following the UD scheme are compared in Fig. 8 for varying allowable displacement. The mean values of P_1 and P_2 are taken as 20 kN and 10 kN, respectively. The reliability indices obtained by all the three approaches are compared in Fig. 9 for varying mean values of P_1 and P_2 (defined by factor, U) considering Δ_{allow} as 0.0112. It can be seen from the plots that the proposed approach yields consistently better results than the two existing approaches for all the cases for this example also.

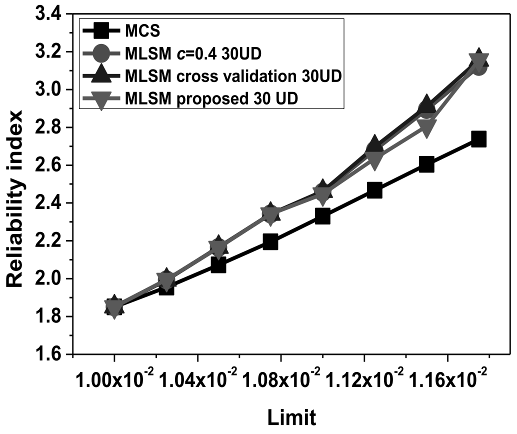

Fig. 8 The comparison of reliability index with varying allowable displacement with 30 UD training data

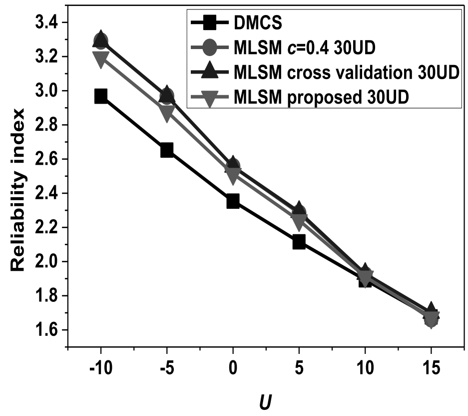

Fig. 9 Comparison of reliability index with varying mean of loads P_1 and P_2 with 30 UD training data

The comparison of percent errors for varying allowable displacement limit considering different order of the polynomial basis function with 30 UD training data set are shown in Table 9. Like the previous two examples, the results clearly show the improved performance of the proposed approach compared to the existing approaches.

Table 9 The comparison of the percent error in estimating reliability index for varying allowable displacement considering different polynomial basis functions.

Displacement Limit (m)		Percent error		
		MLSM *c*=0.4	MLSM cross validation	MLSM proposed
0.01025	1st order	12.291	9.537	8.092
	2nd order	2.714	2.402	1.891
	3rd order	2.669	2.387	1.881
0.01100	1st order	11.766	10.250	5.212
	2nd order	4.389	5.020	3.909
	3rd order	4.319	4.927	3.862
0.01175	1st order	15.227	18.110	13.689
	2nd order	10.029	10.029	6.404
	3rd order	9.964	10.001	6.358
Mean Error	1st order	13.095	12.632	8.998
	2nd order	5.711	5.817	4.068
	3rd order	5.651	5.772	4.034

The results of the sensitivity of the effectiveness of the proposed MLSM scheme in reliability estimate of the space dome with varying allowable displacement limit for ten different sets of 30 training data are shown in Table 10. The robustness of the proposed algorithm to better estimate reliability compare to the existing approaches is clearly revealed for this case as well.

Table 10 The comparison of the percent errors in reliability estimates of the space dome with varying allowable displacement for ten different sets of 30 training data.

Displacement Limit (m)	**Set**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
0.01025	*c*=0.4	1.98	1.94	4.162	2.714	2.708	2.637	2.928	3.291	2.711	2.637
	Cross validation	1.98	1.95	4.382	2.402	2.743	2.587	2.869	3.308	2.55	2.587
	Proposed approach	2.074	2.063	3.937	1.891	2.614	2.391	2.559	2.95	2.204	2.391
0.011	*c*=0.4	5.457	5.436	3.188	4.389	4.669	4.683	4.349	4.002	4.511	4.683
	Cross validation	5.678	5.478	2.802	5.020	4.749	4.906	4.552	4.043	4.903	4.906
	Proposed approach	5.022	5.012	2.244	3.909	4.077	4.132	3.743	3.299	3.982	4.132
0.01175	*c*=0.4	14.044	14.014	8.429	10.029	11.958	11.716	10.753	10.032	10.865	11.716
	Cross validation	15.305	15.315	9.203	10.029	12.947	12.502	11.364	10.671	11.293	12.502
	Proposed approach	15.305	15.295	8.838	6.404	12.475	11.341	9.805	9.401	9.035	11.341
Mean Error	*c*=0.4	6.221									
	Cross validation	6.518									
	Proposed approach	5.796									

With regard to the computational involvement, it is to be noted that the number of training samples by the usual MLSM with c=0.4, the MLSM with cross validation and the proposed MLSM approach are same. Thus, the major concern of computational involvement in metamodeling approach i.e. to execute a high fidelity numerical code (usually a finite element model) to prepare the training samples to construct the metamodel remains same for all the three approaches. However, the MLSM with c=0.4 needs no further computing involvement whereas the cross validation and the proposed approach needs to determine the value of the parameter, c. This needs to solve an optimization sub problem. But, the computational requirement for this will be negligible with respect to the first part as the later part of computing will be entirely on the metamodel (just explicit function evaluation) without any further execution of the high fidelity numerical model.

4. Summary and Conclusions

The tuning of the hyper-parameter of the Gaussian weight function appropriately in a generic way to construct the MLSM based metamodel for reliability analysis application is investigated. The effectiveness of the proposed algorithm to construct the MLSM based metamodel for SRA applications is studied by comparing with the reliability results obtained by the MLSM based RSM by using commonly recommended value of the hyper-parameter (c=0.4) and also by using the hyper-parameter obtained by the leave one out cross validation approach considering the direct MCS based results as the benchmark. The numerical results of all the three examples clearly indicate the effectiveness of the proposed approach in reliability estimation. The improved accuracy of the estimated reliability by the proposed approach compared to the reliability results obtained by the two existing MLSM based RSM approaches is clearly noted when compared with the most accurate direct MCS based results. The sensitivity study with regard to the effectiveness of the proposed scheme for ten different training data set showed that the reliability prediction accuracy is enhanced by the proposed compared to that of obtained by the other two approaches. It is observed from these results (Table 2 and 3 for example 1, Table 6 and 7 for example 2 and Table 9 and 10 for example 3) that the proposed approach is consistently better and yields substantially less mean error with respect to that of obtained by the existing two MLSM based RSM approaches. This clearly shows the robustness of the proposed algorithm for improved estimate of reliability. The approach is generic in nature. Though, it is illustrated for the Gaussian weight function; but can be applied for any other weight function involving a hyper parameter. However, the level of accuracy is problem dependent and also depends on the numbers of training samples.

5 Replication of results

All the necessary information for generating the data sets for the problems are presented in the manuscript. The first author of this manuscript is pursuing his doctoral study on the subject. The codes can be readily available from the first author on request.

Compliance with ethical standards

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

References

Breitkopf P, Naceur H, Rassineux A, Villon P (2005) Moving least squares response surface approximation: formulation and metal forming applications. Compu and Struct 83:1411–28

Bucher C, Most T (2008) A comparison of approximate response functions in structural reliability analysis. Probab Engg Mech 23:154-63

Chakraborty S, Chowdhury R (2016) Assessment of polynomial correlated function expansion for high-fidelity structural reliability analysis. Struct Safety 59 9-19

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

Ditlevsen O, Madsen HO (1996) Structural reliability methods. John Wiley and Sons Ltd., Chichester

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

Elhewy AH, Mesbahi E, Pu Y (2006) Reliability analysis of structures using neural network method. Probab Eng Mech; 21:44-53

Fang KT, Lin DK, Winke P, and Zhang Y (2000) Uniform Design: Theory and Application. Technometrics 42(3):237-248

Gavin HP, Yau SC (2007) High-order limit state functions in the response surface method for structural reliability analysis. Struct Safety 30:162-179

Ghosh S, Chakraborty S (2017) Simulation based efficient seismic fragility analysis of existing structures. Earthq and Struct 12 (5):569-581

Ghosh S, Roy A, Chakraborty S (2018) Support vector regression based metamodeling for seismic reliability analysis of structures. Appl Math Modelling 64, 584-602

Goel T, Haftka RT, Shyy W, Queipo NV (2007) Ensemble of surrogates. Structural and Multidisciplinary Optimization 33(3), 199-216

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

Guo Z, Bai G (2009) Application of Least Squares Support Vector Machine for Regression to Reliability Analysis. Chinese J of Aeronautics 22(2):160-166

Haldar A, Mahadevan S (2000) Reliability assessment using stochastic finite element analysis. John Wiley and Sons, NY

Haussler-Combe U. Elementfreie G. Verfahren (2001) Grundlagen und Einsatzmoglichkeitenzur Berechnung von Stahlbetontragwerken. Habilitation-Thesis, University of Karlsruhe, Germany

Haussler-Combe U, Korn C (1998) An adaptive approach with the element-free-Galerkin method. Compu Methods in Appl Mech and Engg 162;203-222

Hurtado JE , Alvarez DA (2003) A classification approach for reliability analysis with stochastic finite element modelling. J of Struct Engg ASC, 129: 1141– 49

Jiang Y, Luo J, Liao G , Zhao Y , Zhang J (2015) An efficient method for generation of uniform support vector and its application in structural failure function fitting. Struct Safety 54:1-9.

Kabasi S, Chakraborty S (2019) An Efficient Moving Least Squares Based Response Surface Method for Reliability Analysis of Structures, 29th European Safety and Reliability Conference, Hannover, Germany.

Kang SC, Koh HM, Choo JF (2010) An efficient response surface method using moving least squares approximation for structural reliability analysis. Probab Engg Mech 25: 365-371

Karutz H (2000) Adaptive Kopplung der Elementfreien Galerkin-Methodemit der Methode der FinitenElementebeiRissfortschrittsproblemen. Ph.D. Thesis Ruhr-Universitat Bochum, Germany 2000

Kaymaz I (2005) Application of Kriging method to structural reliability problems. Struct Safety 27:133-151

Keshtegar B (2017) A hybrid conjugate finite-step length method for robust and efficient reliability analysis. Appl Math Modelling, 45,226-237

Kim C, Wang S, Choi KK (2005) Efficient response surface modeling by using moving least-squares method and sensitivity. AIAA J 43(1):2404-2411

Kim SH, Na SW (1997) Response surface method using vector projected sampling points. Struct Safety 19: 3-19

Kwon OS, Elnashai AS (2006) The effect of material and ground motion uncertainty on the seismic vulnerability curves of RC structure, Eng Struct 28(2):289-303

Mann NR, Schafer RE, Singpurwalla ND (1974) Methods for Statistical Analysis of Reliability and Life Data, John Wiley & Sons Inc, New York

Melchers RE (1999) Structural reliability analysis and prediction. John Wiley and Sons, Chichester.

Mohammed RK and Felician C(2020) Performance evaluation of metamodelling methods for engineering problems: towards a practitioner guide, Struct & Multidis Optim 61:159–186

Most T, Bucher C (2005) A Moving Least Squares weighting function for the Element-free Galerkin Method which almost fulfils essential boundary condition. Struct Engg and Mech 21(3):315-332.

Notin A, Gayton N, Dulong JL, Lemaire M, Villon P, Jaffal H (2010) RPCM: a strategy to perform reliability analysis using polynomial chaos and resampling. Euro J of Comp Mech19(8):795–830

Roy A, Manna R, Chakraborty S (2019) Support vector regression based metamodeling for structural reliability analysis. Probab Engg Mech 55:78-89

Schöbi R, Sudret B, Marelli S (2017) Rare event estimation using Polynomial Chaos-Kriging. ASCE-ASME J. Risk Uncertainty Eng. Syst., Part A: Civ. Eng. 3(2): D4016002

Su H, Chen Z, Wen Z (2016) Performance improvement method of support vector machine-based model monitoring dam safety. Structl Control and Health Monitoring 23: 252–266

Taflanidis AA, Cheung SH (2012) Stochastic sampling using moving least squares response surface approximations. Probab Engg Mech 2: 216-224.

Toropov VV, Scharamm U, Sahai A, Jones RD, Zeguer T (2005) Design optimization and Stochastic Analysis based on the Moving Least Squares Method. 6th World Cong. of Struct. Multidisc. Opti., Rio de Janeiro, Brazil

Youn BD, Choi KK (2004) A new response surface methodology for reliability-based design optimization. Compu and Struct 82: 241–5