Multiple response surfaces method with advanced classification of samples for structural failure function fitting

Youbao Jiang[a], Jun Luo[a], Michael Beer[b], Edoardo Patelli[b], Matteo Broggi[b], Yihua He[a], Jianren Zhang[a]

(a School of Civil Engineering and Architecture, Changsha University of Science and Technology, Changsha 410114, China;

b Institute for Risk & Uncertainty, University of Liverpool, Liverpool L69 3GQ, UK)

Abstract: The current response surface methods based on classifier usually fail to classify all samples correctly, thus neglect the effects of the misclassified samples on the fitting function. To overcome this issue, an improved multiple response surfaces method is proposed. It is mainly based on the techniques of sector division and correct classification of samples. The main steps are: (1) compute a normalized inner product coefficient between the closest sample to the origins and any other one, and sort samples by the coefficient values; (2) select a reasonable number of sorted samples (i.e. range of normalized inner product coefficient) for each sector to assure that the samples in the sector can be classified correctly; (3) divide the overall space into multiple sectors based on such ranges and execute an approximation sector by sector based on support vector machines. A main merit of this method is that it can approximate implicit failure functions well as the number of samples is large enough due to the features of the correct classification of all samples. In addition, it can be applied to both single failure functions and multiple failure functions (explicit ones and enveloped ones). Numerical examples show that the proposed method can achieve a good fitting of implicit failure functions, and the reliability results are accurate, too.

Keywords: structural reliability; multiple response surfaces; support vector; correct classification; failure function; sector division

1 Introduction

For a mechanical structural system with uncertainty, the estimation of its reliability provides valuable information. For a simple structure, the failure function would be explicit and the reliability analysis can be performed effectively by the first order reliability method (FORM), the second order reliability method (SORM) or Monte Carlo simulation (MCS). However, for a large and complex structure, such failure function is usually implicit, complex (e.g. piecewise and nonlinear) and in a high-dimensional space. In this case, the conventional FORM, SORM and MCS would be less efficient or accurate.

To overcome these difficulties, Der Kiureghian et al [1,2] proposed a search algorithm and strategies for finding the multiple design points, and Katafygiotis [3,4] et al developed a spherical subset simulation method for solving high-dimensional reliability problems. Numerical examples indicate that these

measures need no surrogate model and can increase the efficiency and accuracy of reliability analysis.

Another strategy for dealing with the difficulties is to obtain an explicit approximation (i.e. surrogate model) of the implicit failure function of the structure before performing a reliability estimation. As a useful tool for modelling and analyzing, the response surface method (RSM) has attracted significant attention due to its computational efficiency and convenience in combination with common software.

Faravelli and Bigi [5,6] discussed a stochastic finite element method based on response surface approximation to analyze the reliability of structural and mechanical systems whose geometrical and material properties have spatial random variability. Bucher and Bourgund [7] studied a new adaptive interpolation scheme of updating polynomial to increase the efficiency and accuracy of the response surface method in reliability calculations. Quite a number of measures have been proposed to improve the efficiency and accuracy of the conventional

1

response surface method. These improvements mainly concern approaches that use more complex function models, such as complete quadratic polynomials [8], higher-order polynomials [9], adaptive models with selected terms [10], and artificial neural networks (ANN) [11-14]; and approaches that use efficient techniques to allow the approximation to be closer to the limit state function at the design point, such as the weighted regression method [15,16] and the experimental points moving schemes [17,18]. For a low dimensional case (e.g. less than 4 input variables), it is stated that such improvements are quite capable of approximating failure functions of structural systems [19,20]. However, the fitting accuracy would be largely affected by the number of sample points, and the generalization error would also increase largely as the number of input variables increases, because these improvements are mainly based on the principle of empirical risk minimization (i.e. fitting residual minimization) [21].

Based on the statistical learning theory, an optimal way to minimize the generalization error of a learning machine is following the principle of structural risk minimization for a high-dimensional case. Support vector machines (SVM) are one of the best options to follow this principle because they only use the support vectors rather than any other samples to fit a function. With this unique property of SVM, an accurate fitting function and reliability results can often be achieved [22-24].

As mentioned earlier, the real failure function of a large structural system would usually be of a complex structure and high-dimensional. It is clear that a single response surface, whether it is based on SVM, or polynomials, or ANN, cannot approximate the real failure function well in this case. Thus, a reasonable way is adopting multiple response surfaces to obtain an accurate approximation [25].

Recently, the multiple response surfaces method has attracted attention in slope reliability analysis [26,27]. In these applications, each possible failure mode (i.e. slip surface) of the slope can be identified by the Bishop method before using a quadratic polynomial model to perform an approximation, and thus multiple failure functions can be obtained one (failure mode) by one. Following this way, an integral reliability can be estimated easily if the assumption of a series system is adopted. Unlike the reliability problems in a slope system, each failure mode of a large structural system is difficult to be identified and is generally unknown before using RSM to perform an approximation. Thus, Neves et al [28] recommended to regard the system failure function as an integral enveloped one with complex and high-dimensional characteristics and use RSM to approximate the enveloped function directly.

A simple measure to achieve a good approximation of a complex function is dividing the overall space into many hypercubes based on the divided ranges of each variable, and obtaining an approximation to the complex function in each hypercube (see [29]). Note that such measure cannot be applied well to a high-dimensional case because the needed number of hypercubes (i.e. response surfaces) would increase exponentially as the number of variables increases, resulting in a time-consuming computation.

To reduce the computational cost, Mahadevan and Shi [30] proposed an approach to approximate the real failure function with multiple hyperplanes, and a way to calculate the failure probability through the union or intersection of the failure domains corresponding to each segment. Liu and Lv [31] proposed a similar approach for response surfaces combination in reliability analysis. One of the main advantages of these approaches is that they can be useful for both component and system reliability problems. However, it is difficult to determine whether a failure domain defined by the corresponding response surface contributes through a union operation or an intersection operation to the overall failure domain for reliability estimation, because such operation may vary largely in different domains when the real failure function is of a complex nature and high-dimensional. Thus, the multiple response surfaces method still needs to be improved further in efficiency and accuracy.

Herein, we propose a method for correct

classification of samples to fulfill this demand, which is mainly based on the techniques of sector divisions of the overall high-dimensional space and SVM. The proposed method as well as an iterative algorithm is used to achieve a converged solution in function fitting. The computational efficiency and accuracy are also studied for the proposed method.

2 Classifying models

2.1 Short review of SVM

This section is devoted to a short description of the SVM method of calssification. More details can be found in [32,33].

Given is a set of N training samples $(\boldsymbol{x}_i, h_i)$ ($i=1,2, \cdots ,N$) with binary outputs $h \in \{+1,-1\}$ corresponding to the two classes. Assume that both classes can be separated through a hyperplane, as shown in Fig.1. Then, the optimal hyperplane is expressed as $G(\boldsymbol{x}) = \boldsymbol{w} \cdot \boldsymbol{x} + b$, which obeys $|G(\boldsymbol{x})| \geq 1$ for all $\boldsymbol{x}_i$. Thus, the margin with respect to both classes is $2/\|\boldsymbol{w}\|$, and the optimum linear classifier is solved and given by

$$G(\boldsymbol{x}) = \mathrm{sgn}[\sum_{i=1}^{n} \alpha_i^* h_i (\boldsymbol{x}_i \cdot \boldsymbol{x}) + b^*] \qquad (1)$$

where $\mathrm{sgn}(\cdot)$ means the sign function; $(\boldsymbol{x}_i \cdot \boldsymbol{x})$ means the inner product operation; α_i^* and b^* are two relevant parameters to define the optimum linear classifier. For most samples, $\alpha_i^* = 0$. By comparison, for support vectors, $\alpha_i^* \neq 0$.
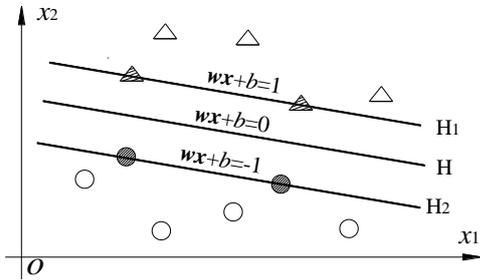

Fig.1 Optimal hyperplane for linearly separable case

2.2 Quadratic function model

Let $\boldsymbol{x}$ denote the normalized vector of random variables, $\boldsymbol{x} = [x_1, x_2, \cdots, x_n]$, where n is the number of variables. For a nonlinear function fitting, a model of quadratic polynomials without cross terms is often

used. In this paper, such model is also selected and a corresponding transformation from $\boldsymbol{x}$ space to $\boldsymbol{z}$ space is given by

$$\boldsymbol{z} = [x_1 \cdots x_n \quad x_1^2 \cdots x_n^2] \qquad (2)$$

Thus, in the $\boldsymbol{z}$ space, the optimum linear classifier is expressed by

$$G(\boldsymbol{z}) = \mathrm{sgn}[\sum_{i=1}^{n} \alpha_i^* h_i (\boldsymbol{z}_i \cdot \boldsymbol{z}) + b^*] \qquad (3)$$

It is known that such a classifier is actually a quadratic support vector machine (QSVM) in $\boldsymbol{x}$ space.

3 Sample classification strategy

3.1 Efficient technique for sector divisions of a high-dimension space

Let all variables be standard normal (achieved through a normalized transformation). In the standard normal space, the closer a point on limit state surface is to the origin, the more it would contribute to the failure probability. Therefore, higher attention should be paid to points, which have a smaller distance from the origin. Let $\boldsymbol{x}_0$ be the closest point to the origin. A normalized inner product coefficient $\rho_0(\boldsymbol{x})$ between $\boldsymbol{x}_0$ and any other point $\boldsymbol{x}$ is given by

$$\rho_0(\boldsymbol{x}) = (\boldsymbol{x}_0 \cdot \boldsymbol{x}) / \|\boldsymbol{x}_0\| / \|\boldsymbol{x}\| \qquad (4)$$

For a high-dimensional space, as mentioned earlier, it is less efficient to divide the overall space by hypercubes based on the divisions of each variable. However, it would be more efficient if the overall space is divided by such a normalized inner product coefficient, as shown in Fig.2. For example, if s values of the coefficient are selected, namely $\rho_1, \rho_2, \cdots, \rho_s$, the overall space can be divided into s sectors.
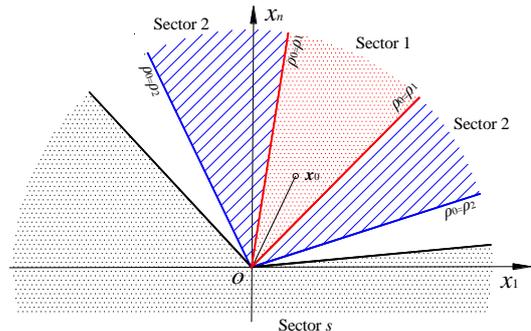

Fig.2 Sector divisions in high-dimensional space

3.2 Correct classification of samples

Suppose that x_0 is the closest point from the origin among all N sample points. Then, the normalized inner product coefficient between x_0 and x_i is given by

$$\rho_0(i) = (x_0 \cdot x_i)/\|x_0\|/\|x_i\| \qquad (5)$$

It is known that the sample points which have a similar coefficient can be clustered into a subclass of the overall samples. Suppose that the overall samples are classified into s subclasses and the numbers of sample points in the subclasses are N_1, N_2, $\cdots$, N_s, respectively, so that

$$N = N_1 + N_2 + \cdots + N_s \qquad (6)$$

If the number of sample points is selected carefully for every subclass, then the corresponding sample points are separable for the selected SVM model. For example, in an n-dimensional space, n sample pairs (see Fig.3 and Eq.(9), n safe points and n failure points, $2n$ points totally) are linearly separable because a linear hyperplane can be determined perfectly through n sample points. Similarly, to make them separable, the number of sample points can be set as $4n$ (i.e. $2n$ sample pairs) for each subclass when an SVM model of quadratic polynomials without cross terms is used.

In order to classify all sample points correctly, the following steps can be adopted:

(1)Calculate $\rho_0(i)$ for each sample with Eq.(5) and sort them in descending order based on their normalized inner product coefficients, which is given by

$$1 = \rho_0(1) \geq \rho_0(2) \geq \cdots \geq \rho_0(N) \qquad (7)$$

(2)Select initial numbers of samples for the (s-1) subclasses (e.g. N_1, N_2, $\cdots$, N_{s-1} all equal $4n$).
(3)Obtain the number N_s for the last subclass with Eq.(6) and execute a QSVM analysis for each subclass.
(4)If all the samples in each subclass are classified correctly (i.e. G(x_i)=h_i for any sample x_i), go to the step (5). Otherwise, reselect N_1, N_2, $\cdots$, N_{s-1} by a minor decrement operation, and go to the step (3).

For example, assume that some samples in the lth subclass are misclassified, where l=1, 2, $\cdots$, (s-1), and

select 2 as the decrement. Then, N_l=N_l-2 and repeat this operation until all samples in the subclass are classified correctly. Our experience shows that in most cases this operation needs to be done only once.
(5)Divide these coefficient series into s ranges, $[\rho_1, 1]$, $\cdots$, $[\rho_{l+1}, \rho_l]$, $\cdots$, $[-1, \rho_{s-1}]$, where ρ_l can be obtained with N_1, N_2, $\cdots$, N_{s-1}, and given by

$$\rho_l = [\rho_0(\sum_{i=1}^{l} N_i) + \rho_0(\sum_{i=1}^{l} N_i + 1)]/2 \qquad (8)$$

Thus, as shown in Fig.3, s sectors and the corresponding approximations are obtained.
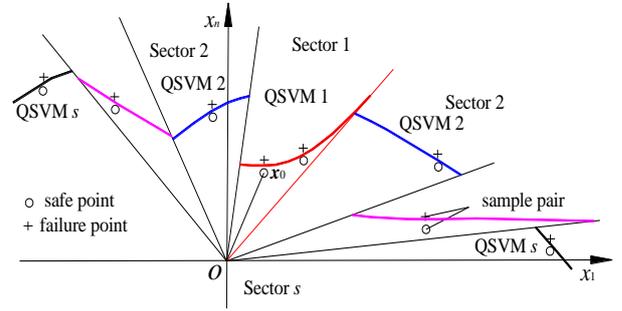

Fig.3 Diagram for correct classification of samples

4 Generation of sample pairs

4.1 Data transformation

Let y denote the pre-normalization vector of random variables. Then, the sample pair SV$_1$ and SV$_2$ in the safe domain and in the failure domain, respectively, (more details in [24]) for function fitting can be obtained by

$$SV_1 = \left\{ y_1, y_2, \cdots, y_{n_1}, F_{m-1}[1, \frac{y_{n_1+2}}{y_{n_1+1}}, \cdots, \frac{y_{n_1+n_2}}{y_{n_1+1}}] \right\} (9a)$$

$$SV_2 = \left\{ y_1, y_2, \cdots, y_{n_1}, F_m[1, \frac{y_{n_1+2}}{y_{n_1+1}}, \cdots, \frac{y_{n_1+n_2}}{y_{n_1+1}}] \right\} (9b)$$

where y_1, y_2, $\cdots$, y_{n1} are n_1 resistance variables; y_{n1+1}, y_{n1+2}, $\cdots$, y_{n1+n2} are n_2 load variables, $n_1+n_2=n$; F_{m-1} and F_m are safe load and failure load, respectively, corresponding to a serviceability limit state or ultimate limit state; m is the number of load steps for solving the limit load. It is known that more load steps are usually needed to obtain the value of F_{lim} than to obtain the load range [F_{m-1}, F_m]. Thus, the

computational cost can be reduced moderately since only F_{m-1} and F_m are involved for the generation of sample pairs.

The random variable y (e.g. load or material strength) in a practical case is often not standard normal. However, most of the relevant operations (e.g. sector division, samples classification) in this paper are performed in this space. Thus, a data preparation is usually needed before a multi-response surface analysis can be performed.

Let x_j be the corresponding normalized variable to y_j. Then, a normalized expression is given by

$$x_j = \Phi^{-1}[F_{cd}(y_j)] \quad j=1,2,\cdots,n \qquad (10)$$

Moreover, the corresponding inverse transformation is also required and expressed by

$$y_j = F_{cd}^{-1}[\Phi(x_j)] \quad j=1,2,\cdots,n \qquad (11)$$

where $\Phi(\cdot)$ and $F_{cd}(\cdot)$ are the cumulative distribution functions of the standard normal variable and y_j, respectively; $\Phi^{-1}(\cdot)$ and $F_{cd}^{-1}(\cdot)$ are their corresponding inverse functions, respectively.

4.2 Initial sample pairs

In order to produce initial sample pairs, methods for design of experiments can be employed, such as Latin hypercube sampling (LHS) [34,35] or a uniform design method [24,36]. With these methods it is possible to distribute the training data as uniformly as possible over the entire design space. Herein, the uniform design method (i.e. uniform table) is selected to achieve this goal. For a uniform table denoted by $U_n(q^s)$, n means the number of samples; q means the number of levels of each variable; s means the maximum number of variables.

Generally, such method is available for a normalized space. However, the load range [F_{m-1}, F_m] in Eq.(9) for solving sample pairs is obtained in the original space by a structural analysis or a performance function call. Therefore, to obtain a comparable range of each variable, a data transformation should be performed as follows:

(1)Use the uniform design method (e.g. select a proper uniform table in [37]) to establish a reasonable experiment design for the initial samples based on

statistical properties of variables.

For example, assume that the interested range of the jth variable x_j is [$-k$, k] for function fittings, where k is a range parameter (e.g. k=3.0 for normal cases), and u_{ij} is the data value in uniform table and often in range [1, q]. Then, for the ith experiment design, the relation between x_{ij} and u_{ij} is given by

$$x_{ij} = k[2(u_{ij}-1)/(q-1)-1] \qquad (12)$$

(2)Perform an inverse transformation with Eq.(11) to obtain the corresponding samples in the original y space, and execute a structural analysis or call a performance function to establish the load range [F_{m-1}, F_m].

(3)Calculate the sample pairs in the original space based on Eq.(9), and then perform a transformation based on Eq.(10) to obtain the initial sample pairs for response surface approximation.

4.3 Additional sample pairs

It is known that only the real failure function can classify all points in failure and safe domains correctly. The more sample points an approximation can classify correctly, the more accurate it would be. However, due to consideration of computational cost, the obtained number of sample points is often limited. To acquire a good quality for the training samples, particular attention needs to be paid to the points around the design point

Usually, such set can be obtained by continuously adding extra samples around the approximate design point in an iterative calculation process. Herein, a practical method for generation of additional sample pairs (ASP) in divided sectors is proposed to achieve this goal. Its main steps are as follows:

(1)Solve the corresponding design points (e.g. D_1, D_2) for the currently acquired response surfaces (e.g. QSVM1, QSVM2).

(2)Use the design points to perform an inverse transformation based on Eq.(11), and then obtain the corresponding sample points in the original y space.

(3)Perform a structural analysis or call a performance function to calculate the load range [F_{m-1}, F_m], and obtain an ASP based on Eqs.(9-10), as shown in Fig.4.

Note that the obtained design points may be beyond the corresponding boundaries of sectors, like

the design point D_2 in Fig.4, and thus the ASP should be generated carefully. The corresponding reliability index β_i to the design point D_i can be calculated as:
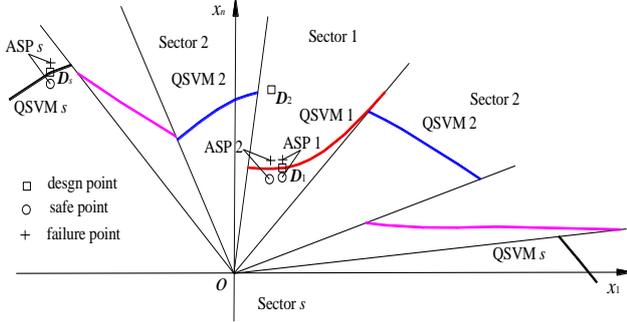
$$\beta_i = \|OD_i\|, \ i=1, 2, \cdots, s.$$

Fig.4 Generation of ASP in sectors

5 Multiple response surfaces method for reliability analysis

If the failure function is implicit, of a complex structure and high-dimensional, then the design point approximated with a response surface is generally not very accurate even if all sample points can be classified correctly. Thus, the proposed approximation still needs to be improved. Herein, a practical iterative algorithm is proposed for convergence in reliability analysis. Its main steps are as follows:

(1)Use the proposed approach in Section 4.2 to collect N_0 initial samples, and $i=0$.

(2)Divide the overall space into $s^{(i)}$ sectors with the currently obtained samples based on the method given in Section 3.1.

(3)Perform a multi-response surface analysis with the technique of correct classification of samples as proposed in Section 3.2.

(4)Use the method in Section 4.3 to obtain $s^{(i)}$ ASPs, and check whether they can all be classified correctly by the currently used response surfaces.

(5)If yes, stop the calculation and go to the step (6). Otherwise, add the obtained $s^{(i)}$ ASPs to update the current set of samples, $i=i+1$, and go to the step (2).

(6)Use the converged response surfaces to perform a reliability analysis.

Herein, the conventional MCS is used because it is accurate and simple (e.g. needs no skills for construction of importance sampling density function). The converged response surfaces are presented in multiple sectors of the standard normal space, thus the MCS should be operated in converged sectors, too.

The main points are: firstly, generate a random sample point in the standard normal space, and calculate the coefficient with Eq.(4) and the converged x_0; secondly, check which sector the sample point is in, and call the corresponding QSVM to calculate its performance function value; finally, repeat this procedure with a sufficient number of points and count the number of failure to obtain the failure probability P_f. Then, the reliability index β is given by

$$\beta = -\Phi^{-1}(P_f) \tag{13}$$

Note that the reliability analysis method above involves the methods of sector division, correct classification of sample pairs, generation of sample pairs and an iterative algorithm. As a main advantage of this method, the approximation is accurate as the number of samples becomes large enough, because all the samples can be classified correctly. This feature is demonstrated in the following examples.

6 Example analysis

Three examples are introduced to show the usefulness of the proposed method.

6.1 Examples with multiple failure modes

6.1.1 Series system with 8 failure modes

Consider a series system with 8 failure modes:

$$g_1 = 2x_1 - 2x_2 + 8 \qquad g_2 = 2.6x_1 - 2x_2 + 9.3$$
$$g_3 = 1.4x_1 - 2x_2 + 7.2 \quad g_4 = 4x_1 - 2x_2 + 14$$
$$g_5 = 0.7x_1 - 2x_2 + 6.8 \quad g_6 = -0.5x_1 - 2x_2 + 8$$
$$g_7 = -2x_1 - 2x_2 + 11 \qquad g_8 = -1.5x_1 - 2x_2 + 10$$

where x_1, x_2 are both standard normal variables.

For this case, there are only 2 standard normal variables. Thus, following the instructions in [37], the first and the third column data of $U_6(6^4)$ is selected (see all data of the table in [37]). Then, using Eq.(12), the corresponding values of x_1 and x_2 can be computed with $k=3.0$, as shown in Table 1.

Moreover, the transformation and inverse transformation in Eqs.(11-12) are not required for this

example since both variables are standard normal. Assume x_1 and x_2 can be regarded as a resistance and a load variable, respectively. Then, for a given x_1, two values of x_2, one in a safe state and another one in a failure state, can be obtained by computing the assumed load range $[F_{m-1}, F_m]$. Herein, $F_{m-1}=0.95F_{lim}$, and $F_m=1.05F_{lim}$, where F_{lim} (i.e. limit value of x_2) can be obtained with a given x_1 by calling the 8 series failure functions. Then, the initial sample pairs are established, as shown in Table 2.

Table 1 Uniform design of x_1 and x_2

No.	u_1	u_2	x_1	x_2
1	1	3	-3.0	-0.6
2	2	6	-1.8	3.0
3	3	2	-0.6	-1.8
4	4	5	0.6	1.8
5	5	1	1.8	-3.0
6	6	4	3.0	0.6

Table 2 Initial sample pairs for Example 1

No.	Safe samples (SV$_1$)		Failure samples (SV$_2$)	
	x_1	x_2	x_1	x_2
1	-3.0	0.71	-3.0	0.79
2	-1.8	2.09	-1.8	2.31
3	-0.6	3.02	-0.6	3.34
4	0.6	3.43	0.6	3.79
5	1.8	3.37	1.8	3.73
6	3.0	2.38	3.0	2.63

The computing process for convergence is presented in Table 3, and the comparisons between the converged fitting function, and the real piecewise function are shown in Fig.5. It can be seen that the obtained fitting function (i.e. response surfaces QSVM 1 and QSVM 2) approximates the real piecewise function very well in the domains of interest, where the strongest contributions to the failure probability are made.

Table 3 Computing process for Example 1

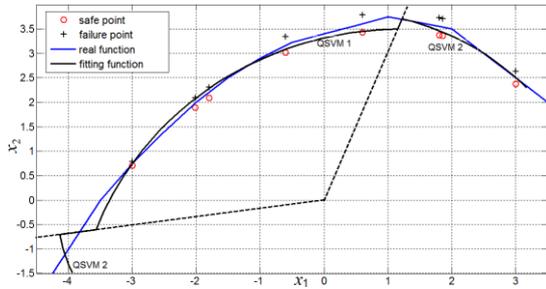| Iterative No. 0	$[\rho_{l+1}, \rho_l]$	Fitting function	Design point	ASP		Classifying check
				Safe	Failure	
QSVM 1	[0.516, 1.0]	$104.51+13.14x_1-19.63x_2-5.37x_1^2-3.58x_2^2=0$	(-2.01, 2.08)	(-2.01,1.89)	(-2.01,2.09)	Yes
QSVM 2	[-1.0, 0.516]	$24.95-0.31x_1-0.34x_2-1.50x_1^2-1.53x_2^2=0$	(1.86,3.43)	(1.86,3.36)	(1.86,3.71)	Yes

Fig.5 Comparisons of fitting function and real one

Using the converged fitting function, the failure probability is computed by MCS, as shown in Table 4. It is seen that the results with different methods are very close to each other. This shows that the proposed method is accurate for a piecewise function fitting.

Table 4 Results with different methods for the Example 1

Method	Analysis Settings	P_f
MCS	10^5 samples	0.0300
Proposed method	8 sample pairs for fitting	0.0306
Ditlevsen [38]	-	[0.0303, .0307]

6.1.2 Integral capacity of a truss string structure

Consider a truss string structure (see [24], herein more variables involved) with a span of 128m, as shown in Fig.6.
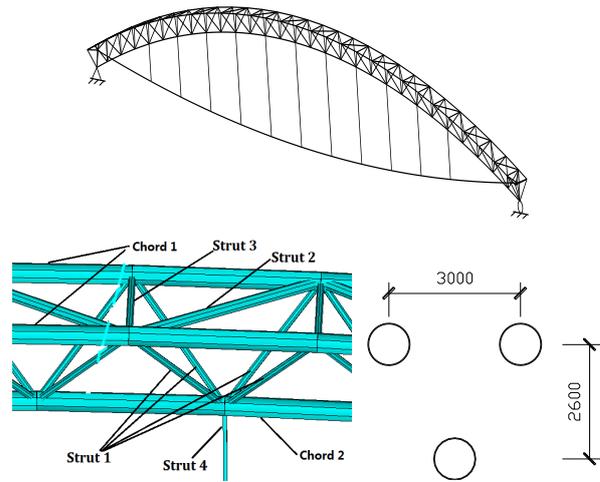

Fig.6 A truss string structure and its truss section

It is fully supported by a lateral brace and subjected to a half-span distributed load F_q and a

full-span distributed load F_g. The truss arch rise-to-span ratio and the cable sag-to-span ratio are 0.08 and 0.03, respectively. Assume that the steel stress-strain relationship is perfectly elastic-plastic, the elastic modulus of each member is 2.0×10^5MPa, and the sectional dimensions of all members are shown in Table 5. Further, the applied pretension is 400MPa for cable, and the space of Strut 4 is 9.2m. Suppose that the sectional thickness t, cross areas of cable and Strut 4, steel yielding strength f_y, half-span load F_q and full-span load F_g are random variables, and their statistics are shown in Table 6.

Table 5 Sectional parameters of a truss string structure

Member	D /mm	t /mm	Member	D /mm	t /mm	Member	A /m^2
Chord 1	480	t_1	Strut 2	273	t_3	Strut 4	A_1
Chord 2	480	t_2	Strut 1,3	168	t_4	Cable	A_2

Note: D and t means sectional outer diameter and sectional thickness, A means cross area.

Table 6 Statistics of variables for Example 2

Variable	Distribution	Mean	COV
t_1	Normal	18mm	0.075
t_2	Normal	16mm	0.075
t_3	Normal	7mm	0.075
t_4	Normal	6mm	0.075
A_1	Normal	0.00796m^2	0.05
A_2	Normal	0.016895m^2	0.05
f_y	Normal	376MPa	0.07
F_q	Type-I largest	50kN/m	0.3
F_g	Normal	50kN/m	0.1

Note: COV means coefficient of variance.

If the limit state function is defined as: the half-span load F_q should be equal to its ultimate capacity $F_{\lim}(t_i, A_1, A_2, f_y)$ (i=1,$\cdots$,4) under all possible load ratio vectors $\boldsymbol{r}$=[1, F_g/F_q], then it is given by

$$Z = \left\{ \left[F_{\lim}(t_i, A_1, A_2, f_y) - F_q \right] \middle| \boldsymbol{r} = [1, F_g / F_q] \right\} = 0 \quad (14)$$

It is known that Eq.(14) would involve multiple failure modes due to the random properties of the load ratio vector and resistant variables.

For this case, there are 9 variables and they are all not standard normal, thus the transformation and inverse transformation in Eqs.(10-11) are needed. If each variable adopts 20 levels, then a uniform table is selected, as shown in Table 7.

Using k=3.0 and Eqs.(11-12), the corresponding values of both resistance and load variables in the original space are obtained for the initial uniform samples. Then, based on the load values (to obtain the load ratio) and resistance values of each initial uniform sample, the load range [F_{m-1}, F_m] can be solved through a structural analysis with a commercial software (e.g. ANSYS). Then, using Eq.(9), the sample pairs (with the same resistance values as uniform samples) are obtained. Both the initial uniform samples and the sample pairs in the original space are shown in Table 8.

Table 7 Uniform table for Example 2

No.	u_1	u_2	u_3	u_4	u_5	u_6	u_7	u_8	u_9
1	2	7	16	3	5	17	16	10	8
2	16	10	5	9	2	4	13	3	3
3	7	11	19	17	1	14	14	16	16
4	19	3	3	4	15	15	9	5	12
5	13	20	8	5	3	12	8	12	20
6	18	12	9	2	10	8	19	20	15
7	3	2	7	18	9	6	10	7	19
8	11	9	6	20	11	16	1	11	1
9	17	6	17	14	8	1	6	17	10
10	12	1	18	11	12	11	20	2	14
11	9	14	20	1	14	9	5	8	4
12	10	15	1	13	17	2	17	9	17
13	6	4	11	6	19	5	15	15	2
14	8	5	2	12	4	10	4	19	6
15	5	17	13	7	6	3	2	4	13
16	4	18	4	8	13	20	12	18	11
17	1	13	10	15	20	13	7	1	9
18	14	16	12	16	7	19	18	6	5
19	20	19	15	19	16	7	11	13	7
20	15	8	14	10	18	18	3	14	18

The normalized sample pairs for function fitting are obtained with Eq.(11), as shown in Table 9. Using the data, the overall space can be divided into 2 sectors. Then, a QSVM analysis is performed and the response surface function is also obtained in each sector. The corresponding design points are solved accordingly to generate 2 ASPs. It is found that the generated ASPs cannot be classified correctly by the previously obtained fitting function. Thus, such 2 ASPs should be added to the normalized initial sample

pairs for updating and an iterative procedure is needed to achieve an accurate approximation. After 7 iterative steps, the 2 newly generated ASPs can be classified correctly by the currently obtained response surfaces. Therefore, the solutions have converged with 36 sample pairs (i.e. 72 samples) and the finally obtained fitting function can classify all the sample pairs correctly, as shown in Fig.7.
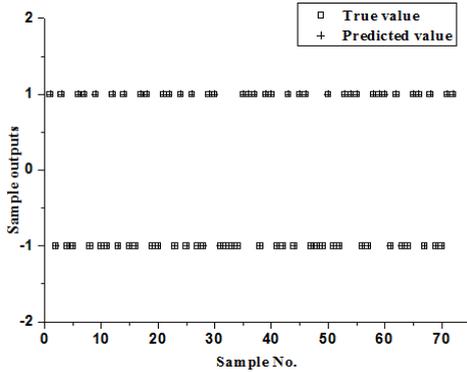
The values of the reliability index are also computed in each iterative step, as shown in Fig.8. It is seen that the reliability index corresponding to QSVM 1 (i.e. β_1), which contributes mostly to the overall reliability index β, converged faster (only needs 2 iterative steps) than that corresponding to QSVM 2 (i.e. β_2).
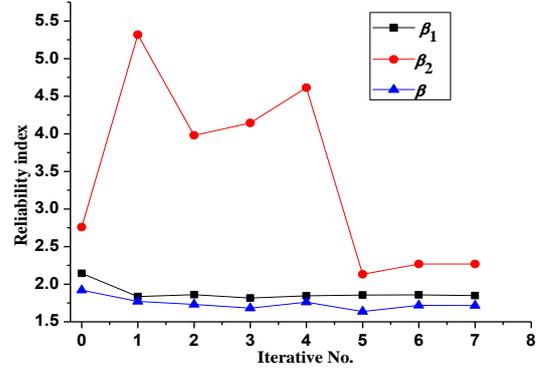

Fig.7 Classifications of the fitting function for Example 2

Fig.8 Values of reliability indexes in iterative steps

Table 8 Initial uniform samples and sample pairs in original space for Example 2

No.	Uniform samples									Loads values for sample pairs				
	Resistance variables							Load variables		Safe load		Failure load		
	t_1/mm	t_2/mm	t_3/mm	t_4/mm	A_1/m^2	A_2/m^2	f_y/MPa	F_q/(kN/m)	F_g/(kN/m)	F_q/(kN/m)	F_g/(kN/m)	F_q/(kN/m)	F_g/(kN/m)
1	14.38	14.67	7.91	4.93	0.00727	0.01863	421.714	45.47	46.05	69.35	70.25	70.75	71.67
2	20.34	15.81	6.09	5.79	0.00689	0.01516	396.779	25.10	38.16	68.40	103.87	69.78	105.97
3	16.51	16.19	8.41	6.92	0.00677	0.01783	405.091	80.32	58.68	82.41	60.17	84.07	61.38
4	21.62	13.16	5.76	5.08	0.00853	0.01810	363.533	29.68	52.37	60.16	106.17	61.37	108.32
5	19.07	19.60	6.59	5.22	0.00702	0.01730	355.221	54.49	65.00	69.08	82.40	70.48	84.07
6	21.20	16.57	6.75	4.79	0.00790	0.01623	446.648	120.55	57.11	109.85	50.56	112.06	51.58
7	14.80	12.78	6.42	7.07	0.00777	0.01569	371.844	35.11	63.42	45.52	82.19	46.44	83.85
8	18.21	15.43	6.25	7.35	0.00802	0.01836	297.040	49.73	35.00	64.51	45.39	65.81	46.31
9	20.77	14.29	8.08	6.50	0.00765	0.01436	338.598	88.87	49.21	72.65	40.22	74.12	41.03
10	18.64	12.40	8.24	6.07	0.00815	0.01703	454.960	23.06	55.53	56.13	133.95	57.27	136.66
11	17.36	17.33	8.58	4.65	0.00840	0.01649	330.286	38.22	39.74	62.62	65.10	63.89	66.41
12	17.79	17.71	5.43	6.36	0.00878	0.01463	430.025	41.65	60.26	67.99	98.34	69.36	100.33
13	16.08	13.54	7.08	5.36	0.00903	0.01543	413.402	72.68	36.58	82.97	41.79	84.65	42.63
14	16.93	13.92	5.59	6.21	0.00714	0.01676	321.975	108.95	42.89	68.00	27.05	69.37	27.60
15	15.66	18.46	7.41	5.50	0.00739	0.01489	305.352	27.30	53.95	37.92	74.94	38.68	76.45
16	15.23	18.84	5.92	5.64	0.00827	0.01943	388.467	98.40	50.79	86.18	44.52	87.92	45.42
17	13.95	16.95	6.92	6.64	0.00915	0.01756	346.909	21.16	47.63	35.92	82.87	36.65	84.55
18	19.49	18.08	7.25	6.78	0.00752	0.01916	438.337	32.27	41.32	83.42	106.68	85.10	108.84
19	22.05	19.22	7.75	7.21	0.00865	0.01596	380.156	59.85	44.47	98.84	73.44	100.84	74.93
20	19.92	15.05	7.58	5.93	0.00890	0.01890	313.663	65.89	61.84	65.86	61.86	67.19	63.11

Table 9 Normalized initial sample pairs for Example 2

No.	Resistance variables							Safe load		Failure load	
	t'_1	t'_2	t'_3	t'_4	A'_1	A'_2	f'_y	F'_q	F'_g	F'_q	F'_g
1	-2.68	-1.11	1.74	-2.37	-1.74	2.05	1.74	1.27	4.05	1.34	4.33
2	1.74	-0.16	-1.74	-0.47	-2.68	-2.05	0.79	1.23	10.77	1.29	11.19
3	-1.11	0.16	2.68	2.05	-3.00	1.11	1.11	1.82	2.03	1.88	2.28
4	2.68	-2.37	-2.37	-2.05	1.42	1.42	-0.47	0.81	11.23	0.87	11.66
5	0.79	3.00	-0.79	-1.74	-2.37	0.47	-0.79	1.26	6.48	1.32	6.81
6	2.37	0.47	-0.47	-2.68	-0.16	-0.79	2.68	2.71	0.11	2.77	0.32
7	-2.37	-2.68	-1.11	2.37	-0.47	-1.42	-0.16	-0.15	6.44	-0.08	6.77
8	0.16	-0.47	-1.42	3.00	0.16	1.74	-3.00	1.04	-0.92	1.10	-0.74
9	2.05	-1.42	2.05	1.11	-0.79	-3.00	-1.42	1.42	-1.96	1.48	-1.79
10	0.47	-3.00	2.37	0.16	0.47	0.16	3.00	0.57	16.79	0.64	17.33
11	-0.47	1.11	3.00	-3.00	1.11	-0.47	-1.74	0.94	3.02	1.00	3.28
12	-0.16	1.42	-3.00	0.79	2.05	-2.68	2.05	1.21	9.67	1.27	10.07
13	-1.42	-2.05	0.16	-1.42	2.68	-1.74	1.42	1.84	-1.64	1.90	-1.47
14	-0.79	-1.74	-2.68	0.47	-2.05	-0.16	-2.05	1.21	-4.59	1.27	-4.48
15	-1.74	2.05	0.79	-1.11	-1.42	-2.37	-2.68	-0.82	4.99	-0.74	5.29
16	-2.05	2.37	-2.05	-0.79	0.79	3.00	0.47	1.96	-1.10	2.02	-0.92
17	-3.00	0.79	-0.16	1.42	3.00	0.79	-1.11	-1.02	6.57	-0.94	6.91
18	1.11	1.74	0.47	1.74	-1.11	2.68	2.37	1.86	11.34	1.92	11.77
19	3.00	2.68	1.42	2.68	1.74	-1.11	0.16	2.38	4.69	2.44	4.99
20	1.42	-0.79	1.11	-0.16	2.37	2.37	-2.37	1.10	2.37	1.17	2.62

Note: superscript ' means a normalized one.

As a result, using the obtained fitting function for MCS, the reliability index is calculated as 1.716. For comparison, line sampling method and subset simulation (as advanced sampling methods [39,40]) are applied using OpenCossan software [41,42], whereby no failure function fitting was applied. Their settings used for reliability analysis are given in Table 10. The obtained reliability index is 1.795 and 1.811, respectively. The comparisons are shown in Table 11. It indicates that the proposed method is as accurate as line sampling method and subset simulation, and needs a smaller number of finite element analysis (FEA) calls in the example.

Table 10 Settings of line sampling and subset simulation

Method	Settings Description
Line Sampling	No. of samples to compute important direction: 9, No. of lines: 5, Total number of samples: 39
Subset Simulation	Intermediate failure probability: 0.1, Proposal distribution: uniform [-0.5, 0.5], No. of initial samples: 100, Total number of samples: 190

Table 11 Comparisons of different methods for Example 2

Method	β	No. of FEA	Classifying ratio
Proposed method	1.716	36	100%
Line sampling method	1.795	39	100%
Subset Simulation	1.811	190	100%
Monte Carlo method	1.762	3000	100%

6.2 Example 3 (single failure mode)

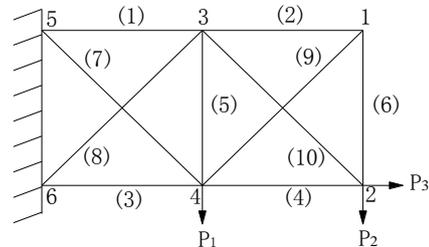A ten bar truss structure as illustrated in Fig.9 is analyzed here.

Fig.9 A ten bar truss structure

It involves 15 basic variables which describe the various properties of structural components, including the horizontal and vertical bar length L, bar areas A_i

($i=1,2,\cdots,10$), modulus of elasticity E and loads P_1, P_2 and P_3. They are assumed to be independent and their statistical parameters are shown in Table 12.

Table 12 Parameters of random variables for Example 3

Variable	Distribution	Mean	COV
L	Normal	1m	0.05
A_i	Normal	0.001m^2	0.15
E	Normal	100GPa	0.05
P_1	Normal	80kN	0.05
P_2	Normal	10kN	0.05
P_3	Normal	10kN	0.05

The limit state is defined as: the maximum vertical displacement equal to 0.004m, and it can be expressed as

$$Z = 0.004 - v(L, A_i, E, P_1, P_2, P_3) = 0 \quad (15)$$

where v is the vertical displacement of node 2. It is known that Eq.(15) is a complex single equation due to high-dimensional variables and nonlinear relations between v and the random variables.

For this example, a uniform table with 40 samples and 40 levels of each variable is selected for the generation of initial samples. Following the way in Example 2, the fitting function finally converged after 8 iterative steps, as shown in Fig.10, though it varies largely during the first 5 iterative steps. The correct classifying ratio is 100% for all samples and the overall failure probability P_f is calculated as 0.004398. This indicates that an accurate approximation can usually be obtained as long as the number of samples is large enough.
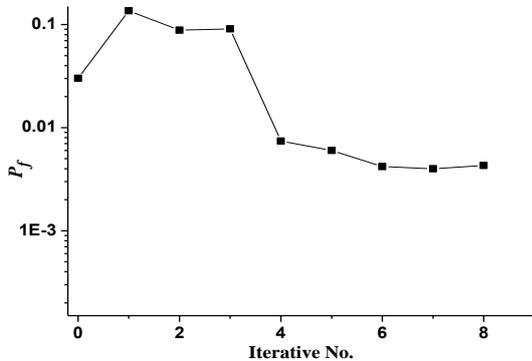

Fig.10 Values of total failure probability in iterative steps

Similarly, the line sampling method and subset simulation are also applied using OpenCossan software for this example. However, their settings are different from those in Example 2. Herein, the settings for line sampling method are: No. of samples to compute important direction: 15, No. of lines: 10, total number of samples: 75. It is noteworthy that a similar reliability analysis of this structure is also performed in [43]. Thus, the comparisons of different methods are shown in Table 13.

It is seen that the subset simulation can achieve a better accuracy with a larger number of FEA for this example, and the response surface methods (e.g. the proposed method and the fitting method [43]) need a smaller number of FEA to achieve a similar accuracy.

Table 13 Results of different methods for Example 3

Method	P_f	Error Ratio	No. of FEA
Proposed method	4.398×10^{-3}	-5.28%	58
Line sampling method	3.770×10^{-3}	-18.80%	75
Subset Simulation	4.700×10^{-3}	1.22%	280
Fitting method [43]	4.405×10^{-3}	-5.13%	55
Monte Carlo method	4.643×10^{-3}	0	10^6

6.3 Summary

The fitting accuracy of the conventional response surface methods is usually largely affected by the number of samples used for constructing the surfaces. For example, sometimes it would lead to a problem called "over fitting" for the conventional approaches when the number is very large. However, the proposed method can effectively avoid such problem by increasing the number of sectors adaptively to achieve a correct classification of all samples. Thus, as long as the number of samples is large enough, the fitting function can be obtained accurately with the proposed method. In this sense, the obtained multiple response surfaces can be regarded as a good approximation with a given set of samples, because the contribution of each sample to fitting accuracy is considered reasonably.

The numerical examples presented have shown the applicability and efficiency of the proposed approach. In fact, it can be applied to analyze structures with multiple enveloped failure modes (e.g. Example 2), or systems with multiple explicit failure functions (e.g. Example 1), or structures with a single implicit failure function under a given failure mode (e.g. Example 3). Compared to the conventional approaches, it needs a smaller number of FEA to

achieve a similar accuracy for reliability analysis in most cases.

Therefore, the proposed method is suitable for approximating the implicit complex failure function of large structures.

7 Conclusions

A multiple response surfaces method with correct classification of samples is proposed to obtain a good approximation to implicit structural failure functions. The method is mainly based on the techniques of sector division and correct classification of samples. The following conclusions may be drawn.

(1)Based on sectors established by the normalized inner product coefficient involving the closest sample to the origins, other than hypercubes established by ranges of each variable, the overall space can be efficiently divided into multiple parts, especially for a high-dimensional case.

(2)With a selection of a reasonable number of samples (i.e. range of normalized inner product coefficient) for each sector, all the given samples can be classified correctly using a quadratic model in the support vector machine analysis.

(3)Combining the sector division of the overall space and correct classification of samples with other relevant techniques (e.g. uniform design, data transformation), an iterative algorithm is proposed for function fitting. It can achieve an accurate approximation as long as the number of samples is large enough.

(4)Example analysis indicates that the proposed fitting method can be applied well to cases of both multiple failure modes and single failure modes, and needs a smaller number of FEA to achieve an accurate reliability result. Thus, it is suitable for large structures, especially.

Acknowledgement

References

[1] Der Kiureghian A, Dakessian T. Multiple design points in first and second-order reliability. Structural Safety, 1998, 20(1): 37-49.

[2] Haukaas T, Der Kiureghian A. Strategies for finding the design point in non-linear finite element reliability analysis. Structural Safety, 2006, 21(2): 133-147.

[3] Katafygiotis LS, Cheung SH. Application of spherical subset simulation method and auxiliary domain method on a benchmark reliability study. Structural Safety, 2007, 29(3):194-207.

[4] Katafygiotis LS, Zuev KM. Geometric insight into the challenges of solving high-dimensional reliability problems. Probabilistic Engineering Mechanics, 2008, 23(2-3): 208-218.

[5] Faravelli L. Response-surface approach for reliability analysis. Journal of Engineering Mechanics, 1989, 115(12):2763-2781.

[6] Faravelli L, Bigi D. Stochastic finite elements for crash problems. Structural Safety, 1990, 8(1):113-130

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

[8] Gayton N, Bourinet J M, Lemaire M. CQ2RS: a new statistical approach to the response surface method for reliability analysis. Structural Safety, 2003, 25(1): 99-121.

[9] Gavin HP, Yau SC. High-order limit state functions in the response surface method for structural reliability analysis. Structural Safety, 2008, 30(2):162-179.

[10] Roussouly N, Petitjean F, Salaun M. A new adaptive response surface method for reliability analysis. Probabilistic Engineering Mechanics 2013, 32(2): 103-115.

[11] Cheng J, Li Q S, Xiao R. A new artificial neural network-based response surface method for structural reliability analysis. Probabilistic Engineering Mechanics, 2008, 23(1):51-63.

[12] Deng J, Gu D, Li X, et al. Structural reliability

analysis for implicit performance functions using artificial neural network. Structural Safety, 2005, 27(1):25-48.

[13] Elhewy A H, Mesbahi E, Pu Y. Reliability analysis of structures using neural network method. Probabilistic Engineering Mechanics, 2006, 21(1):44–53.

[14] Schueremans L, Van Gemert D. Benefit of splines and neural networks in simulation based structural reliability analysis. Structural Safety, 2005, 27(3): 246-261.

[15] Kaymaz I, McMahon CA. A response surface method based on weighted regression for structural reliability analysis. Probabilistic Engineering Mechanics. 2005, 20(1):11–17.

[16] Kang SC, Koh HM, Choo JF. An efficient response surface method using moving least squares approximation for structural reliability analysis. Probabilistic Engineering Mechanics 2010, 25(2): 365-371.

[17] Das PK, Zheng Y. Cumulative formation of response surface and its use in reliability analysis. Probabilistic Engineering Mechanics, 2000, 15 (4): 309–315.

[18] Zhao W, Qiu Z. An efficient response surface method and its application to structural reliability and reliability-based optimization. Finite Elements in Analysis and Design, 2013, 67:34-42.

[19] Gomes HM, Awruch AM. Comparison of response surface and neural network with other methods for structural reliability. Structural Safety, 2004, 26(1): 49-67.

[20] Bucher C G, Most T. A comparison of approximate response functions in structural reliability analysis. Probabilistic Engineering Mechanics 2008, 23(2): 154-163.

[21] Hurtado JE. An examination of methods for approximating implicit limit state functions from the viewpoint of statistical learning theory. Structural Safety, 2004, 26(3):271-29.

[22] Hurtado JE. Filtered importance sampling with support vector margin: A powerful method for structural reliability analysis. Structural Safety, 2007, 29(1): 2-15.

[23] Guo Z, Bai G. Application of least squares support vector machine for regression to reliability analysis.

[24] Jiang Y, Luo J, Liao G, et al. An efficient method for generation of uniform support vector and its application in structural failure function fitting. Structural Safety, 2015, 54:1-9.

[25] Donald E. Brown, Jeffrey B. Schamburg. A generalized multiple response surface methodology for complex computer simulation applications. Proceeding of the 2004 Winter Simulation Conference[C], 2004:958-966.

[26] Li DQ, Jiang SH, Cao ZJ, et al. A multiple response-surface method for slope reliability analysis considering spatial variability of soil properties. Engineering Geology, 2015, 187:60–72.

[27] Li L, Chu X. Multiple response surfaces for slope reliability analysis. International Journal for Numerical and Analytical Methods in Geomechanics, 2015, 39(2):175–192.

[28] Neves RA, Mohamed-Chateauneuf A, Venturini WS. Component and system reliability analysis of nonlinear reinforced concrete grids with multiple failure modes. Structural Safety, 2008, 30(3):183-199.

[29] Yang C, Zhang M, Bai X. Multiple response surface method in analysis of structural reliability. Journal of Northern Jiaotong University, 2001, 25(1):1-4. (in Chinese)

[30] Mahadevan S, Shi P. Multiple linearization method for nonlinear reliability analysis. Journal of Engineering Mechanics, ASCE, 2001, 127(11): 1165-1173.

[31] Liu C, Lv Z. Response surface combination improved by higher order term for structural reliability analysis. Acta Aeronautica et Astronautica Sinica, 2006, 27(4):594-599. (in Chinese)

[32] Cristianini N, Shawe-Taylor J. An introduction to support vector machines and other kernel-based learning methods. Cambridge: Cambridge University Press, 2000.

[33] Geisser S. Predictive Inference: An introduction. New York: Chapman and Hall, 1993.

[34] Olsson A M J, Sandberg G E. Latin hypercube sampling for stochastic finite element analysis. Journal of Engineering Mechanics, 2002, 128(1): 121-124.

Chinese Journal of Aeronautics, 2009, 22(2):160-166.

[35] Roshanian J, Ebrahimi M. Latin hypercube sampling applied to reliability-based multidisciplinary design optimization of a launch vehicle. Aerospace Science and Technology, 2013, 28(1): 297-304.

[36] Cheng J, Li Q S, Xiao R. A new artificial neural network-based response surface method for structural reliability analysis. Probabilistic Engineering Mechanics, 2008, 23(1):51–63.

[37] Fang K T, Wang Y. Number-theoretic methods in statistics. London: Chapman & Hall, 1994.

[38] Ditlevsen O. Narrow reliability bounds for structural system. Journal of Structural Mechanics, 1979, 7(4): 453-472.

[39] Schueller GI, Pradlwarter H J, Koutsourelakis P. A critical appraisal of reliability estimation procedures for high dimensions. Probabilistic Engineering Mechanics, 2004, 19(4):463-474.

[40] Au SK, Beck J. Estimation of small failure probabilities in high dimensions by subset simulation. Probabilistic Engineering Mechanics, 2001, 16(4): 263-277.

[41] Patelli E, Broggi M, De Angelis M, *et al*. OpenCossan: an efficient open tool for dealing with epistemic and aleatory uncertainties. Vulnerability, Uncertainty, and Risk, American Society of Civil Engineers, 2014, 2564-2573.

[42] Patelli E, Panayirci H M, Broggi M, *et al*. General purpose software for efficient uncertainty management of large finite element models. Finite Elements in Analysis and Design, 2012, 51:31-48.

[43] Li H, Lv Z. A support vector machine response surface method for structural reliability analysis. Chinese Journal of Computational Mechanics, 2009, 26(2):199-203. (in Chinese)