Structural Safety 67 (2017) 85–95

Contents lists available at ScienceDirect

Structural Safety

journal homepage: www.elsevier.com/locate/strusafe

# An efficient reliability method combining adaptive Support Vector Machine and Monte Carlo Simulation

Qiujing Pan, Daniel Dias *

*Laboratory 3SR, Grenoble Alpes University, CNRS UMR 5521, Grenoble, France*

ARTICLE INFO

*Article history:*
Received 20 September 2016
Received in revised form 25 April 2017
Accepted 30 April 2017

*Keywords:*
Reliability
Failure probability
Support Vector Machine
Adaptive
Monte Carlo Simulation

ABSTRACT

To enhance computational efficiency in reliability analysis, metamodeling has been widely adopted for reliability assessment. This work develops an efficient reliability method which takes advantage of the **A**daptive **S**upport **V**ector **M**achine (ASVM) and the **M**onte **C**arlo **S**imulation (MCS). A pool-based ASVM is employed for metamodel construction with the minimum number of training samples, for which a learning function is proposed to sequentially select informative training samples. Then MCS is employed to compute the failure probability based on the SVM classifier obtained. The proposed method is applied to four representative examples, which shows great effectiveness and efficiency of ASVM-MCS, leading to accurate estimation of failure probability with rather low computational cost. ASVM-MCS is a powerful and promising approach for reliability computation, especially for nonlinear and high-dimensional problems.



## 1. Introduction

The assessment of failure probability is very important for a structural system whose input parameters are generally subjected to different levels of uncertainty. Therefore, it is interesting to perform probabilistic analysis to account for the uncertainties of input parameters. Considering a structural system whose model response $Y$ is given by a computational model $T$:

$$Y = T(\boldsymbol{X}) \tag{1}$$

where $\boldsymbol{X} \in \mathbb{R}^d$ is a vector representing input parameters of this physical system and $d$ is the dimension of the problem. Given a set of values of input vector $\boldsymbol{X}$, Eq. (1) may give the corresponding model response. In reality, the uncertainty in the input parameters does exist, due to, for example, inherent soil variability or measurement errors. In order to account for these uncertainties, input parameters are often modelled as random variables following prescribed distributions, like the normal distribution or the lognormal distribution. The model response $Y$ is therefore also a random variable.

Under this circumstance, the failure probability of a structural system can be defined by the following multi-dimensional integral expression:

$$P_f = \int \cdots \int I[G(\boldsymbol{x})]f(\boldsymbol{x})d\boldsymbol{x} \tag{2}$$

where $\boldsymbol{x}$ is a realization of the $d$-dimensional random vector $\boldsymbol{X}$ sampled from the joint probability density function (PDF) $f(\boldsymbol{x})$; $G(\boldsymbol{x}) = 0$ is the limit state surface of the structural system, $G(\boldsymbol{x}) < 0$ defining the failure domain and $G(\boldsymbol{x}) > 0$ the safe domain. The indicator function $I[G(\boldsymbol{x})]$ is equal to 1 for $G(\boldsymbol{x}) \leqslant 0$, otherwise $I[G(\boldsymbol{x})] = 0$. The limit state surface $G$ is often determined by the computational model $T$ and a given threshold. For instance, if the computational model $T$ predicts the deformation of a specified position of a considered system, and the maximum allowable deformation at that position is denoted by $S'$, then the limit state function $G(x)$ can be written as:

$$G(\boldsymbol{x}) = T(\boldsymbol{x}) - S' \tag{3}$$

It is common to transform the random variables $\boldsymbol{x}$ into a standard space, which is composed of independent standard normal variables $\boldsymbol{u}$, by means of the Nataf transformation [40] or the Rosenblatt transformation [41]. Then the failure probability can be reformulated in the standard normal space as:

$$P_f = \int \cdots \int I[g(\boldsymbol{u})]\varphi(\boldsymbol{u})d\boldsymbol{u} \tag{4}$$

where $\varphi(\boldsymbol{u})$ is the standard normal PDF and $g(\boldsymbol{u}) = G(\boldsymbol{x}(\boldsymbol{u}))$ represents the transformed limit state function in the standard normal space.

* Corresponding author.
*E-mail address:* daniel.dias@3sr-grenoble.fr (D. Dias).

http://dx.doi.org/10.1016/j.strusafe.2017.04.006
0167-4730/

The exact calculation of the integral of Eq. (4) is generally not practically feasible, especially for high-dimensional problems. Alternatively, some approximated approaches have been developed to address this problem, such as the first-order and second-order reliability method (FORM/SORM), the Monte Carlo Simulation (MCS), the response surface method (RSM).

The most straightforward and robust one is the MCS, and it proceeds in three steps: (1) randomly sampling $n_{MC}$ sets of input parameters according to underlying PDFs, (2) repeatedly running $G$ for all samples and (3) post-calculating the failure probability or the statistical moments. Then, Eq. (4) can be approximated by:

$$\widehat{P}_f = \frac{1}{n_{MC}} \sum_{i=1}^{n_{MC}} I[G(\boldsymbol{u}_i)] \tag{5}$$

where $\boldsymbol{u}_i$ represents the $i$th MC samples. Eq. (5) is an unbiased estimate of failure probability and its coefficient of variation (COV) is evaluated by:

$$COV(\widehat{P}_f) = \sqrt{\frac{1-\widehat{P}_f}{n_{MC}\widehat{P}_f}} \tag{6}$$

The classical Monte Carlo Simulation is dimension-independent, and often regarded as a standard reference for the test of other probabilistic methods because of its versatility and robustness. However, it suffers from a quite low computational efficiency. This makes it hardly applicable for computationally expensive models, such as the finite element or finite difference numerical solvers, especially when the failure probability is rather small.

Recently, a metamodel-based Monte Carlo method has gained considerable attention in the community of reliability analysis. The core idea of this method is to firstly build an analytical metamodel (or the response surface) with limited calls to the original model based on the Design of Experiments (DoE), and then perform Monte Carlo Simulation on the obtained metamodel. The metamodel should be capable of catching the global behavior of the original model responses. There are several mathematical tools available to construct such a metamodel, e.g. polynomial chaos expansion (PCE), Kriging model and Artificial Neural Networks (ANN). The polynomial chaos expansion (PCE) is used to build a metamodel. A Collocation-based Stochastic Response Surface Method (CSRSM) on the basis of PCE was applied by researchers to perform probabilistic analysis [1,2]. However, the PCE suffers from the "curse of dimensionality". The size of the DoE rapidly increases with the number of the input variables and with the PCE order. In order to enhance the applicability of PCE, a sparse polynomial chaos expansion (SPCE) was proposed by Blatman and Sudret [3,4] which results in less terms than a full PCE. The Kriging model [5,6] and ANN [7,8] also attract more and more interest for probabilistic assessment.

Another powerful vehicle for the metamodel construction is the Support Vector Machine (SVM) developed by Vapnik [9,10] in the field of statistical learning theory. It has the following merits [11–15]: (1) In a SVM decision function, only support vectors which are a small fraction of the training set contribute to the model predictions (see Eq. (10)). This feature makes the model predictions more efficient compared with other surrogate models, the Kriging model for example. (2) The SVM makes advantage of the Structural Risk Minimization that minimizes an upper bound on the generalization error. This also gives SVM good performance on avoiding overfitting and on generalization. (3) The SVM deals with classifications of model responses, for example the safe state (labeled as +1) and failed state (labeled as -1) in structural stability, which makes it applicable for Monte Carlo Simulation for which only the "sign" of the model response instead of the exact value is of interest in the computation of failure probability. (4) The SVM is able to bypass the curse of dimensionality and handle highly nonlinear problems effectively, e.g. non-convex and disjoint limit state functions.

Many papers have been devoted to application of the Support Vector Machine (SVM) to reliability analysis recently. The pioneering work appears to be done by Rocco and Moreno [11] who firstly combined Monte Carlo Simulation and Support Vector Machine for reliability analysis. Hurtado, J. E. [16] developed a reliability algorithm based on Importance Sampling and SVM which is shown to be rather computationally efficient with respect to conventional Importance Sampling method. Li et al. [15], Zhao [17] and Zhao et al. [18] proposed a reliability method combining the SVM and the FORM and tested their methods on slope and tunnel reliability analysis. Tan et al. [19] reported four response surface methods which are respectively based on radial basis neural network (RBFN) and SVM for reliability analysis. No obvious difference was found between RBFN-based RSMs and SVM-based RSMs. Bourinet et al. [20] proposed an approach termed as "$^2$SMART" which combines Subset Simulation and SVM (SS-SVM). This approach is able to deal with reliability problems involving small failure probabilities and large numbers of random variables (up to a few hundreds). Ji et al. [21] applied the least-squares support vector (LS-SVM) in combination of Monte Carlo Simulation to assess slope system reliability. It is shown that the LS-SVM is effective to evaluate system probability of a complex slope involving several failure regions. The aforementioned research on application of SVM to reliability analysis mainly follows the procedures of firstly constructing the response surface using SVM based on the Design of Experiments (DoE) and then computing the failure probability or reliability index by using probabilistic approaches (e.g. FORM or MCS) thanks to the surrogate model obtained. Therefore, the effectiveness and efficiency of these methods highly depends on how to train the SVM model, with a required accuracy at the cost of the smallest size of DoE (training samples). The uniform design for the DoE is highlighted by Li et al. [15] and Ji et al. [21] for this purpose. Another more efficient technique is to use an adaptive SVM which aims to build a surrogate by sequentially selecting informative samples. Several adaptive SVM models can be found in Song et al. [14], Tong and Koller [22], Basudhar and Missoum [23].

This work aims to develop an efficient reliability method named ASVM-MCS which combines **A**daptive **S**upport **V**ector **M**achine in conjunction with **M**onte **C**arlo **S**imulation. The remainder of this paper is outlined as follows. The basic theory about how a SVM classifier works is presented in the Section 2. The following section introduces adaptive SVM and details the proposed ASVM-MCS algorithm. Four representative examples for the validation of the ASVM-MCS are given in Section 4. This paper ends up with a conclusion at Section 5.

## 2. A basic introduction to Support Vector Machines

The support vector machine (SVM) is an efficient classifier originated from pattern recognition in machine learning. For a two-category problem, the SVM training algorithm aims at building a hyperplane that separates all training data of one category from those of the other category. The SVM has been successfully applied to wide applications [22,24], ranging from pattern recognition, hand-written characters recognition, text classification to biological sciences. Recently, it has attracted more and more attention in reliability analysis. This section simply presents an overview of the SVM algorithm. For more details, the reader is referred to Vapnik [9,10] .