Support Vector Machine in Structural Reliability Analysis: A Review

Atin Roy and Subrata Chakraborty*

Indian Institute of Engineering Science and Technology, Shibpur, India.

atin.3222@yahoo.com, *Corresponding author schak@civil.iiests.ac.in

Abstract

Support vector machine (SVM) is a powerful machine learning technique relying on the structural risk minimization principle. The applications of SVM in structural reliability analysis (SRA) are enormous in the recent past. There are review articles on machine learning-based methods that partly discussed the development of SVM for SRA applications along with other machine learning methods. However, there is no dedicated review on SVM for SRA applications. Thus, a review article on the implementation of various SVM approaches for SRA applications will be useful. The present article provides a synthesis and roadmap to the growing and diverse literature, specifically the classification and regression-based support vector algorithms in SRA applications. In doing so, different advanced variants of SVM in SRA applications and hyperparameter tuning algorithms are also briefly discussed. Following the detailed review studies, future opportunities and challenges in the area of applications are also summarized. The review in general reveals that the SVM in SRA is getting thrust as it has an excellent capability of handling high-dimensional problems utilizing relatively lesser training data. The review article is expected to enhance the state-of-the-art developments of support vector algorithms for SRA applications.

Keywords: Review, Reliability of Structures, Support vector machine, Support vector regression, Hyperparameter, Design of experiments.

1. Introduction

The support vector algorithm is a nonlinear generalization of the generalized portrait algorithm developed in the 1960s [1]. It is firmly grounded in the framework of statistical learning theory. Its present form was developed at AT&T Bell Laboratories by Vapnik and co-workers in the early 1990s [2]. Statistical learning theory characterizes the properties of learning machines which enable them to generalize well to unseen data [3,4]. The present article focuses on reviewing on applications of support vector algorithms in structural reliability analysis (SRA).

Most of the real structures involved complex geometry and nonlinear material behaviours that required a computationally demanding finite element (FE) analysis or other numerical techniques for response evaluation. Different metamodeling approaches e.g., response surface method (RSM) [5,6], radial basis functions networks (RBFN) [7], polynomial chaos expansion (PCE) [8,9], multivariate adaptive regression splines (MARS) [10], Kriging method [11,12], artificial neural networks (ANN) [13,14], etc., were developed to address the computational challenge of large complex SRA problems. However, such metamodels were developed following the empirical risk minimization principle. The precisions of such approaches to approximate responses for SRA largely depend on the number of training data and usually suffer from the overfitting and curse of dimensionality. On the contrary, the support vector machine (SVM) based on the structural risk minimization principle and small sample learning [4,15] could estimate implicit function with better accuracy and generalization capability. The SVM initially developed for solving classification problems is further extended to solve regression problems. The SVM for regression known as support vector regression (SVR) has revealed superior performance due to its inherent capability to circumvent the overfitting problem in regression and improved response approximation ability [3,4]. Clarke et al. [16] investigated the performance of SVR in comparison to four commonly used metamodeling techniques namely, RSM, Kriging, RBFN, and MARS for approximating responses of complex engineering systems. The application of SVR and its improvement for structural response approximation are vast and multidisciplinary. For example, recently, the SVR and RSM are coupled based on two calibrating strategies to predict the load capacity of shear walls [17].

The early applications of SVM for SRA [18–21] treated reliability analysis as a classification problem. The developments in the SVM-based classification approach for reliability analysis in the recent past are also noted [20,22–26]. Besides classification, the applications of SVR-based metamodeling in SRA are quite prominent [27–31]. It is expected that exploiting the real-valued output of regression is more informative than just a sign from the binary classification and, for this reason, SVR has been preferred over SVM-based classification [31]. However, the performances of SVM-based metamodels (both classification and regression) are largely governed by the proper selection of hyperparameters involved. Several algorithms based on different optimization techniques were developed to search the SVM hyperparameters e.g., gradient descent algorithm [32], coordinate descent method [33], grid search [34], five-fold cross-validation method [24], real-value genetic algorithm [35], particle swarm optimization [36], dynamic particle filter [37], cross-entropy method [31] etc.

The relevance vector machine (RVM) is introduced to avoid the setting of hyperparameters by forming sparse Bayesian inference-based learning [38]. However, RVM involves the selection of a kernel function similar to SVM and parameter tuning is unavoidable if the selected kernel has a free parameter(s). RVM has also been successfully applied for SRA [39–41]. Besides RVM, several modifications of SVM were also attempted for improved SRA, e.g., the applications of least squares support vector machine (LS-SVM) for regression [42,43], particle filter-SVR [37], extended SVR [44], Bayesian SVR [45] and support vector density-based importance sampling method [46].

The applications of support vector algorithms in SRA are developing rapidly in the recent past. There are also some general review articles which briefly covered the applications of SVM-based metamodeling for approximating structural responses. For example, Dey et al. [47] reviewed metamodeling approaches for high-fidelity stochastic analysis of composite laminates. In this regard, two excellent review articles [48,49] on machine learning-based methods for reliability analysis, which partly discussed the development of SVM for SRA applications along with other different machine learning-based metamodels like ANN, Kriging etc., are notable [48,49]. However, the specific details of reliability estimation methods, sampling strategies for training and selection of hyperparameters of SVM models are not well covered in the existing reviews [48,49]. Nevertheless, there is no mention of the number of input parameters involved or the total number of training samples required in different studies. It is noteworthy that there is no review article exclusively on SVM for SRA applications. Thus, a dedicated review of the various SVM algorithms employed for SRA applications will enhance the state-of-the-art developments of support vector algorithms for SRA applications.

The present article attempts to provide a synthesis and roadmap to the growing and diverse literature on support vector algorithms. Specifically, various SVM and SVR-based reliability analysis methods for SRA are critically assessed with regard to computation cost, dimensionality, order of failure probability, applications, advantages and disadvantages. This is expected to be useful to understand the nature of engineering problems various SVM approaches can tackle. Furthermore, different variants of SVM in reliability analysis and the factors that significantly affect the performance of SVM models, e.g., sampling techniques, hyperparameter tuning, selection of kernel and loss function are also discussed. The review study first searches articles with the keywords 'Structural reliability analysis', 'Support vector machine', and 'Support vector regression'. Then, articles that develop or implement support vector algorithms for SRA are critically reviewed. The development of the SVM algorithm itself is not the subject of the present review; rather its application and related issues of implementing those algorithms for SRA are focused on. However, a brief theoretical background of SVM is presented in section 2 for an easy transition from the introduction to the subsequent sections. The applications of other advanced variants of SVM in SRA are presented in section 5 followed by various searching methods of prerequisite hyperparameters of SVM in section 6. The summary of observations and conclusions is made with the future direction of research in section 7.

2. Support Vector Machine

The foundation of SVM was developed by Vapnik [2] and is gaining acceptance due to its various attractive features and promising performance. The formulation embodies the structural risk minimization principle and is found to be superior to the traditional empirical risk minimization principle employed by conventional machine learning methods. The SVM was initially developed to solve classification problems. Subsequently, it has been extended to the domain of regression problems as well (presented in the next section).

2.1 SVM for classification

The SVM primarily describes classification with support vector methods [4]. In the classification problem, the goal is to separate two classes by a function that is induced from the available examples and the classifier works well on unseen examples, i.e., it generalises well. The concept is elucidated by a simple example in Figure 1 (a). Note that many possible linear classifiers can separate the data, but there is only one that maximises the margin. This linear classifier is known as the optimal separating hyperplane.

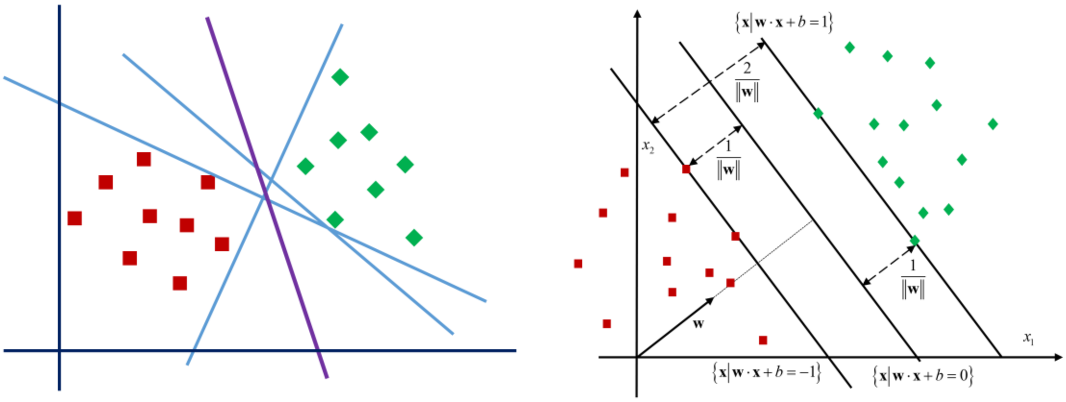

Figure 1. (a) Representation of separating hyperplanes for two-class data and (b) the optimal hyperplane to separate two-class data.

Consider the problem of separating a set of training vectors belonging to two separate classes, $\left\{(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), \cdots, (\mathbf{x}_p, y_p)\right\}$, $\mathbf{x} \in \mathbf{R}^n$, $y \in \{+1, -1\}$ with a hyperplane, $\mathbf{w}.\mathbf{x} + b = 0$. The set of vectors is said to be optimally separated by the hyperplane if it is separated without error and the distance of the closest vector to the hyperplane is maximal. The optimal hyperplane is given by maximizing the margin 2/||w|| (see Figure 1 (b)). Maximizing the margin is equivalent to minimizing ||w||/2, leading to the following quadratic programming (QP) problem,

$$\text{Min.} \quad \frac{1}{2}\|\mathbf{w}\|^2 \quad \text{s.t.,} \; y_i\left[\mathbf{w}\cdot\mathbf{x}_i + b\right] \geq 1 \tag{1}$$

Apart from the classification in linear space, the SVM approach can be readily applied for nonlinear classification as well by utilizing kernel tricks to implicitly map the inputs into high-dimensional feature spaces. Figure 2 explains one such mapping where a two-dimensional input space (x_1, x_2) is mapped into a three-dimensional feature space (z_1, z_2, z_3) by using the mapping, $z_1 = x_1^2, z_2 = \sqrt{2}x_1x_2, z_3 = x_2^2$. It can be observed from Figure 2 that the two-class data are not linearly separable in the original input space but so is possible in a three-dimensional feature space. This aspect will be further discussed later.

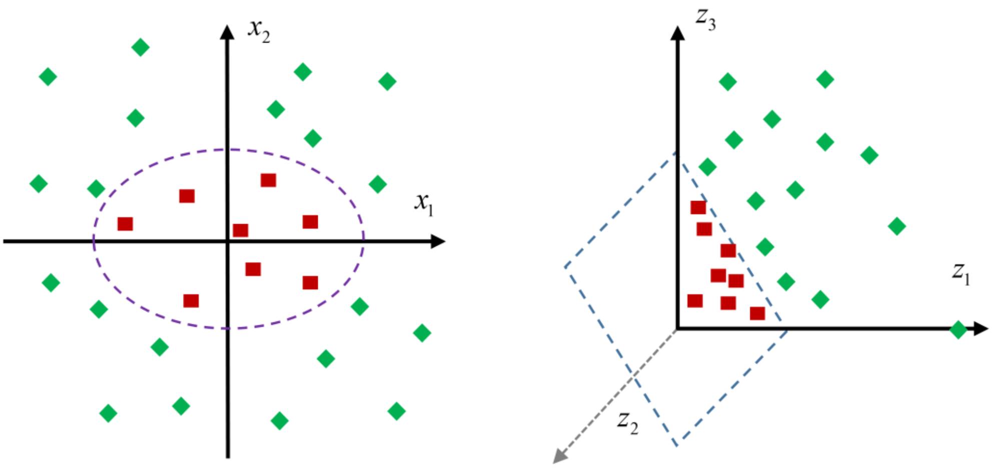

Figure 2. Mapping of input space into high-dimensional feature space.

2.2 Support vector regression

The SVM classification essentially searches the maximal margin to separate two classes by an optimal hyperplane. For this, a QP problem with inequality constraint is solved. However, the SVM method will be time-consuming and huge space demanding for reliability analysis involving a large size of training data as the size of the matrix of the QP problem is directly proportional to the number of training samples and random variables. The classification problem of SVM can be also applied to solve regression problems known as SVR [27].

For a given set of training data, $\{(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), \cdots, (\mathbf{x}_p, y_p)\}$, $\mathbf{x} \in \mathbf{R}^n$, $y \in \mathbf{R}$, if there is a set of functions that map a point in the space $\mathbf{R}^n$ onto the space $\mathbf{R}$ i.e.,

$$F = \{ f(\mathbf{x}, \mathbf{w}), \mathbf{w} \in \Lambda \mid f : \mathbf{R}^n \rightarrow \mathbf{R} \} \tag{2}$$

where Λ is a set of parameters and $\mathbf{w}$ is an unknown parameter vector that needs to be determined. Then, the regression is a function $f \in F$ that corresponds to the lowest expected risk as follows,

$$R(f) = \int_D e(y - f(\mathbf{x}, \mathbf{w}))\, d\, P(\mathbf{x}, y) \tag{3}$$

In which, $e\left(y-f\left(\mathbf{x},\mathbf{w}\right)\right)$ is an error function and defined in SVR as [3],

$$e\left(y-f\left(\mathbf{x},\mathbf{w}\right)\right)=\max\left\{0,\left|y-f\left(\mathbf{x},\mathbf{w}\right)\right|-\varepsilon\right\},\quad \varepsilon>0 \tag{4}$$

The above is known as the ε-insensitive loss function. It neglects the error if the difference between the predicted value $f\left(\mathbf{x},\mathbf{w}\right)$ and the observed value y is less than ε. Different types of error functions like Gaussian or quadratic, Laplacian or least modulus, and Huber's robust loss functions are applied for SVR applications. Note that the advantage of a sparse decomposition will be lost unless $\varepsilon \neq 0$ i.e. for using loss functions other than the ε-insensitive one [3]. Following this, the ε-insensitive loss function is quite commonly used.

2.2.1 SVR for linear regression

The linear regression in SVR is expressed as,

$$f\left(\mathbf{x}\right)=\mathbf{w}\cdot\mathbf{x}+\mathbf{b} \tag{5}$$

Here, $\mathbf{w}\cdot\mathbf{x}$ represents the dot product and $\mathbf{w}$ decides the orientation of a separating plane from the origin and *b* is bias. The SVR algorithm illustrated here considers an ε-insensitive loss function. To approximate all data points within $y_i \pm \varepsilon$, the problem can be expressed as a convex optimization problem [3]. As follows,

$$\text{Min.}\quad \frac{1}{2}\left\|\mathbf{w}\right\|^2 \quad \text{s.t.,}\quad \begin{cases} y_i-\mathbf{w}.\mathbf{x}_i-b\le\varepsilon \\ \mathbf{w}.\mathbf{x}_i+b-y_i\le\varepsilon \end{cases} \tag{6}$$

It is to be noted that ensuring the above perfectly may not be true for all samples. Thus some error allowance is anticipated [4] for generalization. To accomplish this, by introducing two slack variables, the optimization problem is modified as follows,

$$\text{Min.}\ \frac{1}{2}\left\|\mathbf{w}\right\|^2+C\sum_{i=1}^{p}\left(\xi_i+\xi_i^*\right)\ \text{s.t.,}\begin{cases} y_i-\mathbf{w}.\mathbf{x}_i-b\le\varepsilon+\xi_i \\ \mathbf{w}.\mathbf{x}_i+b-y_i\le\varepsilon+\xi_i^* \end{cases}\quad \xi_i,\xi_i^*\ge 0 \tag{7}$$

In the above, C represents the regularization constant, ξ_i, ξ_i^* are the two slack variables for the ith sample and p is the total number of samples. The corresponding primal Lagrange function can be defined as,

$$L_{\text{prim}}\left(\mathbf{w},b,\xi,\xi^*,\alpha,\alpha^*,\eta,\eta^*\right) = \frac{1}{2}\|\mathbf{w}\|^2 + C\sum_{i=1}^{p}\left(\xi_i+\xi_i^*\right) - \sum_{i=1}^{p}\alpha_i\left(\varepsilon+\xi_i-y_i+\mathbf{w}.\mathbf{x}_i+b\right)$$
$$-\sum_{i=1}^{p}\alpha_i^*\left(\varepsilon+\xi_i^*-\mathbf{w}.\mathbf{x}_i-b+y_i\right) - \sum_{i=1}^{p}\left(\eta_i\xi_i+\eta_i^*\xi_i^*\right) \quad (8)$$

where, α,α^*,η and η^* are the Lagrange multipliers. Hence, the dual variables in Eq. (8) need to satisfy positivity constraints i.e., $\alpha,\alpha^*,\eta,\eta^* \geq 0$. According to the Karush–Kuhn–Tucker (KKT) stationarity condition [50], the partial derivatives of L_{prim} with respect to the primal variables $\left(\mathbf{w},b,\xi,\xi^*\right)$ vanish at the primal optimal point i.e.

$$\frac{\partial L_{\text{prim}}}{\partial \mathbf{w}} = \mathbf{w} - \sum_{i=1}^{p}\left(\alpha_i-\alpha_i^*\right)\mathbf{x}_i = 0, \frac{\partial L_{\text{prim}}}{\partial b} = \sum_{i=1}^{p}\left(\alpha_i-\alpha_i^*\right) = 0 \text{ and } \frac{\partial L_{\text{prim}}}{\partial \xi_i^{(*)}} = C - \alpha_i^{(*)} - \eta_i^{(*)} = 0 \quad (9)$$

where the set of vectors $\left(\xi_i^{(*)},\alpha_i^{(*)},\eta_i^{(*)}\right)$ represents both the sets $\left(\xi_i,\alpha_i,\eta_i\right)$ and $\left(\xi_i^*,\alpha_i^*,\eta_i^*\right)$. Moreover, the KKT complementary slackness gives,

$$\alpha_i\left(\varepsilon+\xi_i-y_i+\mathbf{w}.\mathbf{x}_i+b\right)=0,\ \alpha_i^*\left(\varepsilon+\xi_i^*-\mathbf{w}.\mathbf{x}_i-b+y_i\right)=0, \text{ and } \left(C-\alpha_i^{(*)}\right)\xi_i^{(*)}=0. \quad (10)$$

Now using Eq. (9) and (10) along with the primal feasibility (constraints of Eq.(7)) and the dual feasibility ($\alpha,\alpha^*,\eta,\eta^* \geq 0$), the Lagrange dual problem can be obtained as,

$$\text{Max. } L_{\text{dual}}\left(\alpha,\alpha^*\right) = -\frac{1}{2}\sum_{i=1}^{p}\sum_{j=1}^{p}\left(\alpha_i-\alpha_i^*\right)\left(\alpha_j-\alpha_j^*\right)\mathbf{x}_i.\mathbf{x}_j - \varepsilon\sum_{i=1}^{p}\left(\alpha_i+\alpha_i^*\right) + \sum_{i=1}^{p}y_i\left(\alpha_i-\alpha_i^*\right)$$
$$\text{s.t.,}\quad \sum_{i=1}^{p}\left(\alpha_i-\alpha_i^*\right)=0, \quad \alpha_i,\alpha_i^* \in [0, C] \quad (11)$$

The weight vector can be also derived from Eq. (9) as,

$$\mathbf{w} = \sum_{i=1}^{p} \left(\alpha_i - \alpha_i^* \right) \mathbf{x}_i \, , \tag{12}$$

and, substituting the above in Eq. (5), the function $f(\mathbf{x})$ can be obtained as,

$$f(\mathbf{x}) = \sum_{i=1}^{p} \left(\alpha_i - \alpha_i^* \right) \mathbf{x}_i \cdot \mathbf{x} + b. \tag{13}$$

2.2.2 SVR for nonlinear regression

For solving real problems, a linear SVR may not be always appropriate. The linear SVR can be readily extended to nonlinear regression. The key idea is to map the input vector $\mathbf{x}$, into a high dimensional feature space, $\mathbf{z}$ through a function $\varphi(\mathbf{x})$ as explained in Figure 3 and then to solve a linear regression in $\mathbf{z}$ [27]. Now, the dot product, $\mathbf{x}_i \cdot \mathbf{x}_j$ in Eq. (11) required to be replaced by, $\varphi(\mathbf{x}_i) \cdot \varphi(\mathbf{x}_j)$; the computation of which in high dimensional space is quite expensive. By defining the kernel function $K(\mathbf{x}_i, \mathbf{x}_j)$ to implicitly map $\mathbf{x}$ into $\mathbf{z}$ [4], an expensive computational requirement can be avoided.

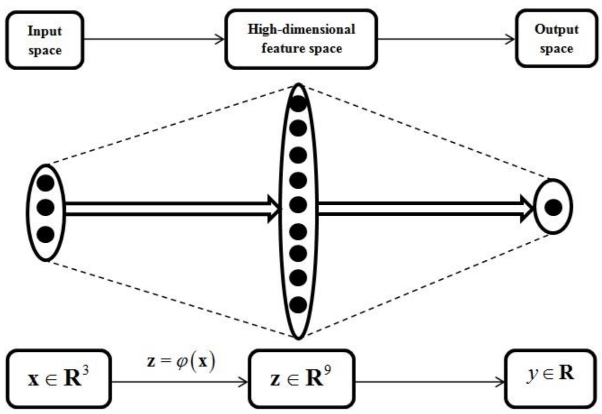

Figure 3. Mapping of input space into a high-dimensional feature space (redrawn from [51]).

Once, the kernel function $K(\mathbf{x}_i, \mathbf{x}_j)$ is introduced in the SVR, the selection of a proper mapping function and computation of the dot product $\varphi(\mathbf{x}_i) \cdot \varphi(\mathbf{x}_j)$ is no more required. The various Kernel functions used in SVR applications include homogeneous and inhomogeneous polynomials, multi-layer perceptron or sigmoid, spline, Laplacian (or exponential) and Gaussian radial basis functions (RBF). The Fourier series, ANOVA-spline, B-splines, additive kernels and tensor product kernels can be also used in many applications [3]. The polynomial and RBF type functions are the most commonly used kernel function for response approximation. Though, the polynomial kernel functions have no free parameters; they may not capture highly nonlinear responses properly. Whereas, the free parameter of RBF kernels allows the SVR to efficiently approximate wide variations of nonlinear responses [23]. By proper selection of the parameter, the RBF kernel can be employed for samples of any distribution. There are two types of RBF Kernel i.e., the Laplacian or exponential and Gaussian. Due to the capability of smoothening the derived function, the Gaussian RBF (GRBF) is a favoured choice. Smola and Schölkopf [3] highlighted the justification of assuming a general smoothness if small information can be obtained from the given data. Thus, the use of GRBF kernel is widely adopted in various literatures on SVM applications.

Now, by replacing the dot product with the kernel function, the optimization problem in Eq. (11) can be represented as,

$$\begin{aligned}&\text{Max.}\quad -\frac{1}{2}\sum_{i=1}^{p}\sum_{j=1}^{p}\left(\alpha_i - \alpha_i^*\right)\left(\alpha_j - \alpha_j^*\right)K\left(\mathbf{x}_i,\mathbf{x}_j\right)-\varepsilon\sum_{i=1}^{p}\left(\alpha_i + \alpha_i^*\right)+\sum_{i=1}^{p}y_i\left(\alpha_i - \alpha_i^*\right)\\ &\text{s.t.,}\ \sum_{i=1}^{p}\left(\alpha_i - \alpha_i^*\right)=0\quad \alpha_i,\alpha_i^*\in\left[0,\mathrm{C}\right]\end{aligned} \tag{14}$$

Finally, the SVR metamodel for approximation of nonlinear responses can be obtained as,

$$f\left(\mathbf{x}\right)=\sum_{i=1}^{p}\left(\alpha_i - \alpha_i^*\right)K\left(\mathbf{x}_i,\mathbf{x}\right)+b\ . \tag{15}$$

It is worth noting here that if either of α_i or α_i^* is positive, the corresponding $\mathbf{x}_i$ will contribute to the computation of the regression function and these contributing input vectors are the so-called support vectors. The above-mentioned SVR algorithm can readily be implemented in the MATLAB platform with the Gunn toolbox [52] available at http://www.isis.ecs.soton.ac.uk/resources/svminfo/.

It is important to note that the loss function parameter, ε and the regularization parameter, C are involved in the SVR algorithm. The latter controls the complexity and degree to which a deviation larger than the former is tolerated. If C is assigned too large i.e. infinity, the parameter cannot introduce any additional capacity control [52]. Consequently, the most possible complex SVR model is formed allowing a very small value of tolerance.

3. SVM Classification in Reliability Analysis

In reliability analysis, samples can be divided into two classes based on whether the value of the considered limit state function (LSF) exceeds its threshold or not. Accordingly, any classification tool can be trained by a limited number of data samples to predict the class of unseen data. Thus, SVM being a classification tool is successfully employed for the same.

3.1 SVM-based reliability analysis methods

In SRA framework, the probability of failure (P_f) can be estimated by a multi-dimensional integral as follows,

$$P_f = \iiint_{\cdots} \int_{g(\mathbf{X})<0} f_{\mathbf{X}}(\mathbf{X})\, d\mathbf{X} \tag{16}$$

where $f_{\mathbf{X}}(\mathbf{X})$ represents the joint probability density function (PDF) of $\mathbf{X}$ and $g(\mathbf{X})$ is the LSF considered as unsafe if $g(\mathbf{X}) < 0$ and safe otherwise. However, it is a formidable task to solve the above. Therefore, various analytical methods (e.g., the first-order reliability method (FORM), the second-order reliability method (SORM)) and simulation methods based on the Monte Carlo simulation (MCS) technique are employed. SVM was utilized to replace the implicit LSF evaluations for reliability analysis by various analytical and simulation methods.

It is first implemented by Rocco and Moreno [18] for reliability evaluation. They performed reliability analysis by treating it as a problem of two-class classification. Based on the information from a few data samples, an SVM model is trained for classifying the samples of an MCS population (say, containing N_{MC} samples) into safe and unsafe categories. Then, P_f is obtained as $P_f = N_f / N_{MC}$, where N_f is the number of unsafe samples. Thus, the approach provides fast execution of the MCS technique for reliability evaluation of complex systems involving implicit LSF. Another classification approach for reliability analysis with stochastic FE modelling was developed by Hurtado and Alvarez [53]. A kernel method for classification usually applied in the fields of pattern recognition or image analysis is employed for this purpose. They developed a greedy sequential procedure to minimize the number of evaluations of the involved LSF. The algorithm is based on the key concept of support vectors. The approach ensures that the evaluation of LSF is only required at the points closest to the failure boundary. Li et al. [20] developed an SVM-based MCS method for SRA. The training of SVM provides the output which indicates only the location of samples in the failure or safe regions and the probability of failure was estimated accordingly by the MCS method. Jiang et al. [24] employed SVM by generating uniform support vectors for MCS-based reliability analysis. This method can also be applied to reliability analyses of large structures involving multiple failure modes. The relevant techniques of data normalization and optimization of kernel parameters of the SVM model are presented with the proposed method for efficient SRA. Pan and Dias [25] proposed a pool-based adaptive SVM approach of metamodel construction to estimate failure probability. The MCS was employed to compute the failure probability based on the obtained SVM classifier. Hurtado and Alvarez [54] developed a method of generating an optimal population for learning SVM to estimate failure probability in MCS framework. Alibrandi et al. [23] proposed a novel sampling strategy for the training of SVM for MCS-based SRA.

Instead of a two-class problem, Zhang et al. [55] developed a three-class classification problem utilizing SVM for SRA involving both random and interval variables. The separating hyperplanes of the three-class classification problem were described by projection outlines on the limit-state surface. Very limited regions are covered by projection outlines. Thereby, they concentrated only on improving the approximation of projection outlines for reducing computational demand. Thereby, a new reliability evaluation technique, where the projection outlines are adaptively and locally approximated, was developed based on the refined SVM model in the MCS framework.

Besides the brute force MCS method, various advanced MCS techniques were also integrated with SVM for reliability estimation, particularly for small failure probability cases. Hurtado [21] proposed a method of SRA by combining importance sampling with an SVM classifier. The training samples from the regions of little probabilistic interest are avoided by utilizing the importance sampling technique. Equally, the importance sampling also exploits the support vector margin to exclude samples that are positioned outside of it. They numerically illustrated that the proposed importance sampling approach filtered with a support vector margin drastically reduces the number of samples required by the conventional importance sampling method. Ling and Lu [56] proposed a novel two-stage SVM based on an importance sampling technique to efficiently deal with SRA problems involving multiple failure regions and small failure probabilities. Bourinet et al. [22] developed a new approach based on SVM classification for estimating small failure probability by employing the subset simulation procedure proposed by Au and Beck [57]. An SVM classifier is built for the boundary of each intermediate failure event of the subset simulation technique. The new method is referred to as ^{2}SMART.

Apart from the simulation-based reliability analysis method, SVM was also integrated with the Taylor series expansion-based analytical reliability analysis methods. Li et al. [20] developed an SVM-based FORM where the partial derivatives of the final expression of SVM (before sign determination) with regard to each variable were calculated. Alibrandi et al. [58] developed the secant hyperplane method by enhancing the FORM to treat SRA as a classification problem. They emphasized that a metamodel for efficient SRA should be directly based on the failure surface in the region of higher probability density instead of the LSF. In this method, a suitable secant hyperplane to the failure plane provides an approximation which is significantly improved with respect to a tangent hyperplane. Alibrandi et al. [59] noted that the limit-state approximated by the existing nonlinear SVM-based methods does not have good computational efficiency in high dimensions. Thus, they proposed a novel linear SVM with implications of high-dimensional geometry for SRA. The starting model is built based on the selection of a set of sample points along the direction of the design point. Therefore, the limit state can be approximated with a hyperplane secant to the limit state based on the design point. This provides an alternative linear response surface based on SVM and the distance of this surface from the origin of the standard normal space is considered equivalent to the reliability index. It is noted that the performances are not affected by the number of random variables.

3.2 DOE schemes for SVM-based reliability analysis

Besides methods adopted for reliability analysis, the performance of SVM in SRA also depends on the selection of training samples. This section focuses on various sampling schemes explored for successful applications of SVM in SRA. Li et al. [20] generated training samples for SVM from a uniform distribution with mean ± K × SD, where, K is a positive number and SD represents the standard deviation of each random variable. The data scaling is applied to improve the stability of the SVM training process and generalization ability. The implemented data scaling for the basic variables was expected to alleviate the effect of different physical properties and dimensions on the training of SVM. Hurtado [21] proposed a modification of the Markov chain Monte Carlo (MCMC) simulation method to obtain the most probable failure point (MPFP) that accepts only those states of the Markov chain having greater joint PDF value than that of the previous state and the LSF was evaluated accordingly. These samples on the failure side along with a few safe samples were used to fit the initial SVM model. A new training sample lying inside the margins of the current SVM is selected iteratively until the failure estimate stabilizes. Basudhar and Missoum [60] proposed an algorithm that selects a new training sample having the highest probability of being misclassified by the SVM decision function and constrained to a minimum distance from the existing training samples which is determined by a function of the hypervolume of the design space, the problem dimensionality, and the number of training samples. Basudhar and Missoum [61] proposed an improved adaptive sampling scheme for the construction of explicit boundaries. Basically, they presented substantial modifications to their previous adaptive scheme [60]. Basudhar and Missoum [61] further improved the choice of a new sample such that it removes the locking of SVM, a phenomenon that was not taken care of in the previous version of the algorithm. The locking of SVM in the previous scheme means the selection of new samples only on the SVM boundary and the modification of the SVM boundary due to such a sample may be negligible if the margin is thin. The approach can be applied to define decision boundaries for reliability analysis and optimization of complex systems. Pan and Dias [25] proposed an adaptive sampling based on a learning strategy which gives more weight to the samples in the vicinity of the failure surface, far away from the existing training samples and located in the margin. A pool-based adaptive SVM approach was employed for metamodel construction with a minimum number of training samples, for which a learning function is proposed to select informative training samples sequentially. Zhang et al. [55] studied an adaptive local approximation method where the initial SVM model was sequentially updated by adding new training samples located around the projection outlines. The ^{2}SMART method [22] first builds an initial DOE by selecting cluster centres of 200,000 standard Gaussian samples. Then the active learning method adds multiple new points at each iteration. The new points are the cluster centres of samples of the work population (obtained by the usual modified Metropolis algorithm) which, either, lie within the margins of SVM, or, change their classes in successive two iterations or are very close to the classifier for two consecutive iterations. Ling and Lu [56] employed the K-means clustering in the first stage of their two-stage method to select multiple new training points (the cluster centres of candidate samples) at each iteration for updating the SVM model with a fast convergence of the algorithm. In the second stage, a new learning function is proposed to select training points sequentially (one per iteration) for further refinements of the SVM model. The proposed learning function give importance on samples laying inside the SVM margin, located very close to failure boundary, having high chances of misclassification and sufficiently away from existing training points.

Hurtado and Alvarez [54] proposed an optimization method for learning statistical classifiers like SVM for SRA. The approach produces an optimal population by solving an unconstrained optimization problem based on the LSF to maximize the entropy. They found that Sobol quasi-random numbers among several proposals assured a high entropy from the initial step. The optimal learning population for SVM-based SRA was achieved by a sequential minimization program using particle swarm optimization. Jiang et al. [24] proposed an efficient method for generating uniform support vectors which are composed of safe and failure samples near the failure surface. Using the uniform design scheme, initial samples are generated first. Then, each initial sample was transformed into a uniform sample pair based on the failure load and safe load close to the limit load. The method can increase the proportion (compared to the whole DOE) and uniformity of support vectors in the input space. Consequently, it reduces the number of required training samples in response approximation. Alibrandi et al. [59] selected the starting linear SVM model based on a set of training points along the direction of the design point and, then, a new training point inside the margin of SVM was added iteratively until the convergence of estimated failure probability. Alibrandi et al. [23] further proposed a novel sampling strategy based on sampling directions, instead of sampling points, which was specifically designed for SVM-based SRA. Suitable sampling cone(s) were introduced to determine the sampling directions.

Though SVM is quite common in machine learning-based applications, it is too much sensitive to outliers in training samples due to the unboundedness of the convex loss [62]. Robust SVMs by replacing the convex loss with a non-convex bounded loss [62] and with generalized quantile loss [63] were developed to deal with noise and outliers. Data imbalance, a common issue for the classification problem, is addressed in three ways [64] i.e. by assigning a distinct cost, modifying traditional algorithms and by data pre-processing that includes undersampling and oversampling methods [64]. Besides two-class and three-class SVM, there is one-class SVM [65]. The applications of one-class SVM are notable for learning in presence of class imbalance [66]. However, one-class SVM is not widely applied to SRA and only a single literature [67] is found on this topic.

The notable contributions to SVM classification-based reliability analysis are summarized in Table 1. Besides the employed reliability estimation methods and DOE schemes, the information on the numerical example problems studied are also mentioned in the table. Among the input dimensions of demonstrated numerical examples, the highest number is reported. The order (power of 10) of the lowest probability estimated is mentioned and the maximum number of training samples or the number of actual function evaluations (N_E) required for SVM classification-based reliability analysis among all the illustrated cases of each study is also cited. The maximum input dimension of the problems and the estimated minimum probability of failure provided in table 1 will be useful to choose which kind of system or engineering is suitable for an approach.

Table 1. SVM-based classification approaches for reliability analysis

Reliability method	DOE	Ref.	Illustrated Examples	Input dim. (max.)	Order of P_f (min.)	N_E (max.)	Advantages and disadvantages
FORM	Samples generated from uniform distribution with mean ± K×SD	[20]	Quadratic limit state, fourth order limit state, three-span continuous beam.	3	10^{-4}	100	Easy to implement, not enough accurate in many cases.
Secant Hyperplane Method (enhanced version of FORM)	Sample points along the direction of the design point + adaptive sampling inside the margin of SVM	[59]	Analytical limit state, stochastic dynamic analysis: oscillator with nonlinear damping	49	10^{-5}	500	Suitable for reliability analysis in high dimensions, complex computational procedure.
	Sample points along the direction of the design point + adaptive sampling inside the margin of SVM	[58]	Dynamic analysis of a frame, stochastic dynamic analysis of an oscillator with nonlinear damping	66	10^{-2}	1580	Suitable in very high-dimensional domains, provides bounds on the P_f, complex computational procedure.
Brute-force MCS	Samples generated from	[20]	Quadratic limit state, fourth	3	10^{-4}	100	Easy to implement, not enough accurate

uniform distribution with mean ± *K*×SD		order limit state, three-span continuous beam				in many cases.
Sobol sequence + adaptive sampling by the entropy maximization using a particle swarm optimization	[54]	A 2D function with multiple design points, A highly concave function, A 2D series system with multiple failure points, A 3D series system, A 5D parallel system, A function with failure probability independent of dimensionality	30	10^{-4}	4000	Generalized method for any statistical learning classifier, applicable for multiple LSFs and multiple design points, number of samples needed for training slowly increases with dimensionality, requiring 500-1000 evaluations to obtain optimum training samples.
Uniform design + generation of sample pairs	[24]	Multiple failure modes (integral capacity of a truss string structure), single failure mode (displacement of a frame structure, and displacement of a truss)	15	10^{-3}	40	Applied to both single and multiple enveloped failure modes, generation method of sample pair increases accuracy, needs a small number of FE analysis to achieve an accurate SRA, number of samples and number of levels of each variable for generation of uniform support vector is experience-based.
Sampling strategy based on sampling directions by introducing sampling cones	[23]	Elastic analysis, limit state where SORM is not accurate enough, limit state with multiple design points, buckling analysis of systems	5	10^{-4}	87	Require reduced number of samples, effectiveness to high-dimensional problems is not verified.
LH + adaptive sampling by learning strategy by giving	[25]	A series system with multiple design points, dynamic response of a	250	10^{-3}	2363	Easily implementable, accurate with moderate number of training data,

	importance to the points near the limit-state and away from existing training samples		nonlinear oscillator, high-dimensional example, tunnel face stability				converges before the stopping criterion, not suitable for very small P_f.
Importance Sampling	Initial samples nearby the MPFP were obtained by MCMC + adaptive sampling inside the margin of SVM	[21]	A 2D case, a 4D case and a 7D series (plastic frame)	7	10^{-4}	37	Very less training data required, number of solver calls is independent of P_f, applicable for small P_f, experience needed to select stating point of Markov chain.
	Adaptive multi-point enrichment by the K-mean clustering in the first stage + one point selected by a learning function augmented sequentially in the second stage	[56]	2D example (single, two and four failure branches), A nonlinear undamped single degree of freedom (DOF) system, A latch lock mechanism of hatch	6	10^{-7}	103	Can estimate rare failure event, reduce the number of function evaluations, suitable size of initial DOE depends on problem complexity (prior knowledge about order of P_f and nonlinearity of LSF are needed).
Subset Simulation	Initial samples from a uniform distribution + adaptive multi-point sampling by selecting informative cluster centres of samples of the work population	[22]	2D LSF with multiple design points, Two DOF. primary/secondary damped oscillator, High dimensional example	250	10^{-7}	10707	Can estimate rare failure event, can handle problems involving multiple design points suitable for moderately high dimensional spaces (up to a few hundred random variables) but not for very large ones.

4. SVR in Reliability Analysis

The SVR is extensively employed for SRA with various reliability estimation methods. Various schemes for selecting training data to construct an SVR-based metamodel were also proposed. This section provides an overview of the contributions to SVR-based metamodel in SRA applications.

4.1 SVR-based reliability analysis methods

Different SVR-based metamodeling approaches were developed to obtain an approximate LSF which can be used to obtain partial derivatives necessary for Taylor series expansion-based reliability analysis methods. Richard et al. [29] developed an adaptive SVR-based metamodeling approach for reliability analysis based on Taylor series expansion. The gradient and Hessian matrix required for FORM and SORM were extracted from the SVR-based metamodel. Li et al. [68] proposed an SVR-based metamodel for reliability analysis of tunnel structures by FORM.

In simulation-based reliability analysis methods, an indication function is only required to know whether a simulation sample is unsafe or not. Once the LSF is approximated by an SVR model, the value of the indication function can be obtained based on the approximate value of the LSF at any sample point. Then, P_f is obtained as,

$$P_f = \frac{1}{N_{MC}} \sum_{i=1}^{N_{MC}} I\left[g\left(\mathbf{X}_i\right)\right] \tag{17}$$

where, the indication function, $I[g(\mathbf{X}_i)]$, is equal to 1, if, LSF at $\mathbf{X}_i$ is unsafe and 0, otherwise. Most of the studies on SVR-based reliability analysis employed the MCS method. SVR-based MCS method is applied in SRA [28,30,51,69], structural system reliability analysis of a cable-stayed bridge [70], and seismic reliability analysis of structures [71].

SVR is also integrated with various advanced MCS techniques for reliability analysis problems involving low failure probabilities. Bourinet [31] proposed an SVR-based reliability analysis method hinged on a subset simulation technique to assess low failure probabilities. The method constructed a sequence of SVR models to reach the failure domain gradually. A highly curved failure surface at a single MPFP, a smooth high-dimensional LSF and a parallel system were taken to verify the efficiency of the method. Roy and Chakraborty

[72] developed a three-stage adaptive SVR-based approach which can be implemented with both MCS and importance sampling techniques for SRA. The importance sampling was centred at MPFP which was obtained by the FORM. The necessary gradients were evaluated using the SVR model.

4.2 DOE schemes for SVR-based reliability analysis

Various DOE schemes have been developed for efficient reliability estimation by SVR-based metamodeling approach. Li et al. [73] adopted the median Latin hypercube (LH) sampling to obtain the training samples for the multi-input multi-output SVR model which was used for SRA of problems involving multiple LSFs. Further, Li et al. [68] proposed a hybrid approach combining uniform design and SVR-based metamodel for reliability analysis of tunnel structures. The approach integrates the merits of both the uniform design and the SVR model for response approximation. Noting the lowest discrepancy of uniform design from the theoretical uniform distribution, Roy et al. [51] also employed uniform design to construct DOE for SVR-based MCS method of SRA.

Besides single-shot DOE schemes, various adaptive DOE schemes for the training of the SVR model have also been proposed by different researchers. The primary aim of such adaptive sampling schemes is to improve the accuracy of an SVR model near the limit-state instead of the whole input domain. Richard et al. [29] developed an adaptive DOE scheme for the training of SVR-based metamodel and applied it to SVR-based FORM and SORM for reliability analysis. The key feature of the adaptive DOE scheme is that it can rotate according to the direction of the gradient of the SVR approximation to capture the nonlinearity of LSF with reduced computational cost. However, the structure of experimental designs (being based on the classical DOE) depends on the input dimension. Thereby, the computational cost increases with the number of random input variables which was reported in their study. Dai et al. [28] proposed a new local approximation method based on SVR and adaptive Markov chain MCS for SRA. The adaptive Markov chain simulation was combined with the importance sampling method to generate samples near the most likely failure region. The SVR model was trained by the generated samples. The SVR approximation was iteratively updated with additional training samples generated by adaptive Markov chain simulation. The iteration process continues until convergence is achieved on the estimated probability of failure. The stopping criterion was judged by a relative deviation of the two successive values of the probability of failure. Bourinet [31] proposed an adaptive approach based on an SVR-based metamodel for assessing low failure probabilities. The method hinges on a subset simulation technique and iterative construction of a sequence of adaptive SVR models. The adaptive training samples were progressively added for moderate accuracy in the safe domain and filled the failure domain as much as possible to improve the accuracy of SVR models in the failure domain. In this regard, Bourinet [31] claimed that distance-based and other space-filling criteria frequently used in literature could not be efficient in higher dimension problems. Therefore, additional training points were generated from the existing SVR models with the modified Metropolis-Hastings MCMC algorithm [57]. To reduce the training cost, some training points were excluded at each iteration based on the evaluated LSF value.

Roy and Chakraborty [69] proposed an adaptive sequential sampling approach of SVR-based metamodel construction for SRA in the MCS framework. In this approach, an initial DOE was built over the entire input space by a space-filling design. Adaptive samples were selected from a subset of MCS points consisting of half of the failure and safe points closer to the limit-state. SVR-based metamodel was found superior in the comparative study with the moving least-square method, Kriging and RVM approaches. However, the algorithm suffers from the scarcity of candidate samples for adaptive sampling in case of low failure probability. Roy and Chakraborty [72] further developed an adaptive SVR-based reliability analysis method modifying the previous approach. The reduced space for adaptive sampling was constructed by simulation samples having a magnitude of the approximated LSF less than the generalized root mean squared error of the cross-validation method employed for selecting perquisite hyperparameters. The issue of scarcity of candidate samples was circumvented by generating samples near the limit state based on the importance sampling technique.

Like SVM, SVR also faces difficulty to deal with outliers or noisy data [74]. The negative effects of outliers on SVR can be reduced by two distinct strategies [75]. In the first strategy, a pre-processing step to remove outliers is applied before employing SVR. Incorporating a robust loss function in the SVR formulation is another strategy. The second direction is noted to be used frequently e.g., a novel LS-SVR in combination with robust 3σ and adaptive weight [76], hybrid robust SVM for regression [77] were attempted to deal with outliers. The robust SVM with generalized quantile loss [63], by converting the constraints of the standard SVM to fuzzy inequalities [75] can also deal with noise and outliers in SVR applications.

4.3 Selection of kernel and loss functions for SVR model

The selection of kernel functions and loss functions to construct an SVR model is an important issue for the success of reliability analysis. Noting the frequent use of RBF in the context of SVR due to its capability of approximating the resulting surface for a wide range of strongly nonlinear implicit functions, Richard et al. [29] used RBF as the kernel function. It is reported that the polynomial kernel function requires more iteration to reach the same level of accuracy as obtained using RBF. Dai et al. [28] selected a fourth-order polynomial as the kernel function for SVR approximation. However, it requires 500 to 1000 training samples. Dai et al. [30] further presented a new multiwavelet linear programming SVR method for reliability analysis. An innovative and efficient SVR model for response approximation was developed by addressing the issue of model sparsity of classical SVR from two different perspectives. Firstly, linear programming SVR was employed to take the advantage of model sparsity, the flexibility of using more general kernel functions, and the computational efficiency of linear programming SVR over classical SVR. Secondly, by constructing the autocorrelation function of multiwavelets, a novel multiwavelet kernel was employed in the context of linear programming SVR for structural response approximation to yield a more compact and sparse representation by leveraging the flexibility of linear programming SVR in choosing the kernels. In most of the studies [31,51,69,72], the ε-insensitive loss function and GRBF kernel were chosen.

The notable contributions to SVR-based reliability analysis are summarized in Table 2 with similar information as provided in Table 1 for SVM classification.

Table 2. Notable contributions to reliability analysis based on SVR metamodel

Reliability method	DOE	Ref.	Illustrated Examples	Input dim. (max.)	Order of P_f (min.)	N_E (max.)	Advantages and disadvantages
FORM	Star-shaped DOE + adaptive sampling by rotating according to the direction of the gradient of the SVR approximation	[29]	A 4D linear explicit LSF, a 7D explicit series, a 21D nonlinear implicit LSF	7	10^{-4}	220	Avoid numerical instabilities related to the numerical estimations of the curvatures directly from the LSF, DOE size depends on the dimension
	Uniform design	[68]	Tunnel stability, implicit LSF for real-life tunnel via a numerical method	7	10^{-2}	28	Require a small number of sampling points, the heuristic selection process of model parameters, validated only for

							P_f higher than 10^{-2}
Brute-force MCS	Adaptive sampling by adaptive Markov chain simulation	[28]	A series system with multiple design points, 10-bar planar truss, Two-bay six-story frame, two DOF Primary/secondary damped oscillator	10	10^{-5}	1000	Good local approximation capability near the failure plane, takes 500-1000 training samples even for less complex problems
	Uniform design	[51]	10D test problem, Space-bar truss	10	10^{-2}	30	Easy to implement, Not enough accurate especially for low P_f cases
	LH/Uniform design + adaptive sampling based on a maximin criterion and the number of failure samples	[69]	Ten-bar truss, space-bar truss, 29D test problem, heat conduction (random fields), six-storey building	75	10^{-4}	124	Enough accurate with limited data, Suffers from the scarcity of samples in case of low P_f
	LH sampling	[30]	Robustness against noises in LSF, 10-bar planar truss, a four-storey building excited by a single period sinusoidal pulse of ground motion	10	10^{-2}	250	Robust to noisy LSF, C and ε were chosen constant, poor accuracy even with a higher number of training data
Importance Sampling	Uniform design + adaptive sampling based on a maximin criterion and a cross-validation error norm	[72]	Ten-bar truss, 29D test problem, space-bar truss	75	10^{-4}	124	Enough accurate with limited data, can estimate low P_f, limited to single LSF-based SRA problems
Subset Simulation	Adaptive DOE by modified Metropolis-Hastings MCMC algorithm [57]	[31]	Two DOF primary/secondary damped oscillator, smooth equally curved high dimensional example, 5D parallel system	250	10^{-4}	1264	Able to estimate low P_f, require a few hundred evaluations of LSF for high accuracy, not suitable for high dimensional problems involving random fields or random processes.

5. Variants of SVM in Reliability Analysis

Apart from SVM and SVR, other variants e.g., RVM, LS-SVM etc. are also applied widely in the field of reliability analysis. These are reviewed separately in this section. Studies on RVM and LS-SVM are presented first in the following two sub-sections. Then, the studies related to the application of other variants are discussed.

5.1 RVM

RVM is the particular specialization of the general Bayesian framework to obtain sparse solutions for regression and classification tasks utilizing models with linear parameters introduced by Tipping [38]. In RVM, the same data-dependent kernel basis as the functional form of SVM is used. Thus, it provides probabilistic predictions and greater sparsity than that of SVM. Unlike the SVM, RVM does not involve the free regularization parameter C and the loss function parameter. The kernel basis used for RVM may or may not satisfy Mercer's condition which is the essential criterion for the selection of kernel function in SVM. RVM has been successfully applied to solve different reliability analysis problems. RVM is employed to predict the implicit LSF for FORM-based slope reliability analysis [78,79]. An adaptive reliability method combining RVM and importance sampling was developed by Changcong et al. [40]. An active learning algorithm-based adaptive RVM within a probabilistic Bayesian learning framework to perform reliability analysis was developed by Li et al. [80]. Adaptive RVM was also combined with Markov-chain-based importance sampling for reliability analysis [81]. Ghosh and Chakraborty [41] proposed an RVM-based Bayesian framework for seismic fragility analysis of structures where demand prediction models were efficiently constructed utilizing limited numbers of training data.

5.2 LS-SVM

Guo and Bai [42] noted that SVR is time-consuming and huge space demanding for reliability analysis involving a large number of simulation samples. They introduced the LS-SVM for regression to overcome those shortcomings. Seismic reliability assessment of reinforce-concrete structures including soil-structure interaction was investigated using wavelet weighted LS-SVM which was designed by combining the weighted LS-SVM and a wavelet kernel function [43]. An LS-SVM-based response surface was combined with FORM for reliability analyses of tunnels [82]. An effective sampling strategy for adaptive reliability analysis based on LS-SVM was developed to improve the efficiency of reliability analysis in practical rock engineering problems [26]. LS-SVM was also applied to slope reliability analysis [83]. Reliability analysis of the settlement of a pile group in clay was also performed using LS-SVM [84].

5.3 Other miscellaneous variants of SVM

Apart from the various SVM approaches discussed in the previous sub-sections, there are some more variants also. For example, Song et al. [85] proposed an adaptive virtual SVM-based method for reliability analyses of high-dimensional problems. In this method, virtual samples were obtained by the universal Kriging method to improve the accuracy of SVM classification for highly nonlinear problems. Cheng and Lu [45] developed Bayesian SVR models which can provide a point-wise probabilistic prediction for active learning algorithms-based SRA. A support vector density-based importance sampling method was developed for reliability assessment [46]. Kriging regression and SVM classification were combined for damage tolerance reliability analysis [86]. The extended support vector regression (X-SVR) was employed for dynamic reliability analysis [44].

6. SVM Hyperparameter Tuning

The construction of an SVM model involves different prerequisite hyperparameters. Generally, a suitable kernel function is selected first for an SVM model. Then, the regularization parameter (C), the tube size (ε) for the ε-insensitive loss function and the free parameter(s) of the selected kernel function are tuned to construct an efficient SVM model. The success of response approximation by SVM largely depends on the proper selection of these parameters. This section briefly reviews the related developments.

Chapelle [32] proposed a methodology to automatically tune multiple parameters (i.e. the regularization parameter, C and the radius of the Gaussian kernel, σ) by gradient descent algorithm to construct an SVM model. This is based on the possibility of computing the gradient of various bounds on the generalization error with respect to these parameters. Ito and Nakano [33] proposed a method to optimize the hyperparameters of an SVR model. The method is based on the minimization of leave-one-out cross-validation error (mean squared) by using a coordinate descent method. Clarke et al. [16] manually optimized the free kernel parameter (i.e. the radius of Gaussian kernel, σ) for a given training data to build an efficient SVR model. Chen [35] proposes a new method termed genetic algorithms-SVR, which optimizes all the SVR parameters (C, ε and bandwidth of the GRBF kernel function, σ^2) simultaneously. The real-valued genetic algorithms were employed to determine the optimal parameters of the SVR model to minimize the generalized mean absolute percentage error of the five-fold cross-validation method on the training data. Hsu et al. [34] recommend a logarithmic grid search on two parameters i.e. C and the GRBF kernel parameter γ for selecting their optimum values to obtain an SVM classification model. The parameter γ is directly related to the bandwidth of the GRBF kernel function, σ^2 as, $\gamma = 1/(2\sigma^2)$. The pair of (C, γ) with the best cross-validation accuracy, i.e., the percentage of data that are correctly classified were selected. The ^{2}SMART method [22] employed a three-fold cross-validation technique for selection of an appropriate kernel function parameter, σ at the initial step and this selected value of σ was used for all further steps of the algorithm.

Demyanov et al. [87] proposed two new approaches using the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC) to estimate the best values of SVM model parameters. This study employed the SVM method for classification problems and selected the GRBF as the kernel Function. The first approach is margin-based which operates using distances of points from the hyperplane. The second one is density-based and it analyses the disposition of support vectors. Therefore, four different algorithms: Margin-AIC, Margin- BIC, Density-AIC and Density-BIC were presented in the study. Among these, the Density-AIC is observed to outperform the others. Lins et al. [36] employed particle swarm optimization to choose the most suitable values of SVM model parameters aiming at minimizing prediction error. The developed SVM model was applied to deal with time series data-based reliability prediction problems. Zhao et al. [37] proposed a novel parameter selection method that combines the SVR and particle filter. The initial values of the parameters were set first, and then a particle filter was used to update these values as new reliability data are available. The method can adapt the hyperparameters according to the new training data. The dynamic particle filter-SVR method was applied for the reliability prediction of time-series data. Zhao et al. [88] proposed a method of SVR parameters selection by combining an analytic selection method and a genetic algorithm. Prior selection by the analytic selection method enables the use of available prior knowledge for guiding the optimization process by genetic algorithm. This avoids divergence and local optima and accelerates convergence. The constructed SVR model was applied for system reliability prediction problems based on available time series data. Jiang et al. [24] chose a quadratic polynomial as the kernel function for SVM classification. The possible ranges of C and polynomial kernel parameter γ are set as [2^{-20}, 2^{30}] for obtaining optimum values of C and γ by using the five-fold cross-validation method.

Bourinet [31] used the GRBF kernel function for the construction of an SVR model and to find the optimal values of SVR hyperparameters, a stochastic search algorithm known as the cross-entropy method introduced by Rubinstein [89] was applied. The hyperparameter space was explored in a logarithmic (base-10) scale within carefully preselected ranges. Noting the huge computation cost to perform leave-one-out cross-validation even for not so large dataset, true leave-one-out cross-validation was avoided by obtaining bounds on approximations of the leave-one-out cross-validation error. For this purpose, the span bound approximations derived for ε-insensitive SVR by Chang and Lin [90] was applied. The optimal values of SVR model parameters (C, ε, γ) were obtained by minimizing an estimate of the leave-one-out error.

Roy et al. [51] first searched the optimum GRBF kernel parameter σ at each node of a logarithm (base-10) grid of C-ε by minimizing the leave-one-out cross-validation error. Then, the final choice of the three parameters was obtained by selecting the grid point corresponding to the lowest error norm. Noting the computational demand of the leave-one-out cross-validation method, they used the same hyperparameter searching algorithm by replacing the leave-one-out cross-validation method with a two-fold [69] and a holdout [72] cross-validation method in the latter steps of iteration for adaptive SVR-based reliability analysis. The notable contributions to hyperparameter tuning for constructing SVM models are summarized in Table 3.

Table 3. Various methods of hyperparameter tuning for SVM

Kernel function	Free parameters searched	Method of optimization	Ref.
Polynomial	C, γ	grid search in the range [2^{-20}, 2^{30}] by five-fold cross-validation method	[24]
Gaussian RBF	σ	manual optimization for $\varepsilon = 10^{-4}$	[16]
		three-fold cross-validation technique	[22]
	C, γ	grid search on C and γ using cross-validation	[34]
	C, σ	minimizing estimates of the generalization error of SVMs using a gradient descent algorithm	[32]
	C, σ^2, ε	using real-valued genetic algorithms	[35]
	C, γ, ε	combination of an analytic selection method of prior selection followed by a genetic algorithm for intelligent optimization	[88]
		logarithmic (base-10) grid search by minimizing an estimation of the leave-one-out error with the cross-entropy method	[31]
	C, σ, ε	minimizing the leave-one-out cross-validation (mean squared) error with a coordinate descent method	[33]
		particle swarm optimization	[36]
		dynamic particle filter	[37]
		minimizing the leave-one-out cross-validation (root mean squared) error for searching optimum σ in different nodes of the logarithm (base-10) grid of C-ε	[51]

Note: γ represents two separate parameters for Gaussian RBF and polynomial kernel functions.

7. Summary and Conclusions

SVM has emerged as a powerful metamodel for its foundation based on the structural risk minimization principle. The detailed reviews of the literature presented clearly reveal that the SVM-based SRA is getting wide attention for its capability to excellently deal the high dimensional problems with fewer samples. Based on the detailed review of the literature presented here, critical observations are summarized in this section.

The SVM for SRA was initiated by Rocco and Moreno [18]. Subsequently, the SVM classification approach was followed by many researchers. Nevertheless, SVM classification does not provide the approximate value of an LSF. Rather, it predicts the sign of the LSF and the distance from the approximate failure plane. On the contrary, SVR can approximate the value of the LSF. The use of SVM in SRA is not only limited to classification and regression approaches but also, includes several other variants like RVM, LS-SVM, Bayesian SVR, X-SVR, virtual SVM etc.

The SVR model generally involves free parameters i.e. a loss function parameter, ε and a regularization parameter, C. In addition, adopting kernel function to deal with nonlinear regression problems involved additional parameter(s). The accomplishment of an SVM-based model significantly hinges on the appropriate selection of such parameters. Generally, the cross-validation approach is applied with an appropriate optimization procedure to obtain the SVR model parameters without further function evaluations.

In both the classification and regression-based SVM models, most of the studies preferred the GRBF as the kernel function due to its capability of approximating the resulting surface for a wide range of strongly nonlinear implicit functions. Uses of a polynomial kernel are also noted in SVM applications as it is simple and free from kernel parameter tuning. However, it was reported that if a polynomial kernel function is used, more iteration is required to reach the same level of accuracy as obtained in the case of GRBF.

Apart from the hyperparameters, the DOE scheme has a significant impact on the performance of the SVM model. Several adaptive sampling schemes have been developed and applied successfully for efficient SRA. There are DOE schemes which are specifically designed for different SVM-based metamodels. Active learning-based algorithms are also combined with the advanced SVM variants which can provide predictive variance.

The significant developments in the application of SVM toward reliability analysis are clearly observed in the existing literature. Yet, there are scopes of further developments in this field. The existing hyperparameter searching algorithms preselect a kernel function and a loss function. A new search algorithm may be explored that will include kernel and loss function selections. The ensemble of metamodels is a new trend in reliability analysis. Along with PCE, Kriging and other metamodels, SVM has already been used as a component metamodel for an ensemble. New ensembles, where SVM with various hyperparameter settings or kernel or loss function or its advanced variants will only be used, may be explored for SRA. Relevant adaptive sampling schemes for such ensembles are also an important area of research.

References

[1] Vapnik V, Lerner A. Pattern recognition using generalized portrait method. Autom Remote Control 1963.
[2] Vapnik VN. The nature of statistical learning theory. Springer; 1995.
[3] Smola AJ, Schölkopf B. A tutorial on support vector regression. Stat Comput 2004;14:199–222. doi:10.1023/B:STCO.0000035301.49549.88.
[4] Vapnik VN. Statistical learning theory. Wiley; 1998.
[5] Gaxiola-Camacho JR, Azizsoltani H, Villegas-Mercado FJ, Haldar A. A novel reliability technique for implementation of Performance-Based Seismic Design of structures. Eng Struct 2017;142:137–47. doi:10.1016/J.ENGSTRUCT.2017.03.076.
[6] Goswami S, Ghosh S, Chakraborty S. Reliability analysis of structures by iterative improved response surface method. Struct Saf 2016;60:56–66. doi:10.1016/J.STRUSAFE.2016.02.002.
[7] Deng J. Structural reliability analysis for implicit performance function using radial basis function network. Int J Solids Struct 2006;43:3255–91. doi:10.1016/J.IJSOLSTR.2005.05.055.
[8] Zhou Y, Lu Z, Yun W. Active sparse polynomial chaos expansion for system reliability analysis. Reliab Eng Syst Saf 2020;202:107025. doi:10.1016/j.ress.2020.107025.
[9] Cheng K, Lu Z. Active learning polynomial chaos expansion for reliability analysis by maximizing expected indicator function prediction error. Int J Numer Methods Eng 2020;121:3159–77. doi:10.1002/nme.6351.
[10] Metya S, Mukhopadhyay T, Adhikari S, Bhattacharya G. System reliability analysis of soil slopes with general slip surfaces using multivariate adaptive regression splines. Comput Geotech 2017;87:212–28. doi:10.1016/j.compgeo.2017.02.017.
[11] Kaymaz I. Application of kriging method to structural reliability problems. Struct Saf 2005;27:133–51.
[12] Thapa A, Roy A, Chakraborty S. Reliability analysis of underground tunnel by a novel adaptive Kriging based metamodeling approach. Probabilistic Eng Mech 2022;70:103351. doi:10.1016/J.PROBENGMECH.2022.103351.
[13] Papadrakakis M, Papadopoulos V, Lagaros ND. Structural reliability analyis of ela’stic-plastic structures using neural networks and Monte Carlo simulation. Comput Methods Appl Mech Engrg 1996;136:145–63.
[14] Lagaros ND, Fragiadakis M. Fragility assessment of steel frames using neural

networks. Earthq Spectra 2007;23:735–52. doi:10.1193/1.2798241.
[15] Vapnik VN. The nature of statistical learning theory. Springer; 2000.
[16] Clarke SM, Griebsch JH, Simpson TW. Analysis of support vector regression for approximation of complex engineering analyses. J Mech Des Trans ASME 2005;127:1077–87. doi:10.1115/1.1897403.
[17] Keshtegar B, Nehdi ML, Trung NT, Kolahchi R. Predicting load capacity of shear walls using SVR–RSM model. Appl Soft Comput 2021;112:107739. doi:10.1016/J.ASOC.2021.107739.
[18] Rocco CM, Moreno JA. Fast Monte Carlo reliability evaluation using support vector machine. Reliab Eng Syst Saf 2002;76:237–43. doi:10.1016/S0951-8320(02)00015-7.
[19] Hurtado JE. An examination of methods for approximating implicit limit state functions from the viewpoint of statistical learning theory. Struct Saf 2004;26:271–93.
[20] Li HS, Lü Z, Yue ZF, Lu ZZ, Yue ZF. Support vector machine for structural reliability analysis. Appl Math Mech (English Ed 2006;27:1295–303. doi:10.1007/s10483-006-1001-z.
[21] Hurtado JE. Filtered importance sampling with support vector margin: A powerful method for structural reliability analysis. Struct Saf 2007;29:2–15.
[22] Bourinet JM, Deheeger F, Lemaire M. Assessing small failure probabilities by combined subset simulation and Support Vector Machines. Struct Saf 2011;33:343–53. doi:10.1016/j.strusafe.2011.06.001.
[23] Alibrandi U, Alani AM, Ricciardi G. A new sampling strategy for SVM-based response surface for structural reliability analysis. Probabilistic Eng Mech 2015;41:1–12. doi:10.1016/J.PROBENGMECH.2015.04.001.
[24] Jiang Y, Luo J, Liao G, Zhao Y, Zhang J. An efficient method for generation of uniform support vector and its application in structural failure function fitting. Struct Saf 2015;54:1–9. doi:10.1016/j.strusafe.2014.12.004.
[25] Pan Q, Dias D. An efficient reliability method combining adaptive Support Vector Machine and Monte Carlo Simulation. Struct Saf 2017;67:85–95. doi:10.1016/j.strusafe.2017.04.006.
[26] Zhao H, Li S, Ru Z. Adaptive reliability analysis based on a support vector machine and its application to rock engineering. Appl Math Model 2017;44:508–22. doi:10.1016/J.APM.2017.02.020.
[27] Moura M das C, Zio E, Lins ID, Droguett E. Failure and reliability prediction by support vector machines regression of time series data. Reliab Eng Syst Saf 2011;96:1527–34. doi:10.1016/J.RESS.2011.06.006.
[28] Dai H, Zhang H, Wang W, Xue G. Structural Reliability Assessment by Local Approximation of Limit State Functions Using Adaptive Markov Chain Simulation and Support Vector Regression. Comput Civ Infrastruct Eng 2012;27:676–86. doi:10.1111/j.1467-8667.2012.00767.x.
[29] Richard B, Cremona C, Adelaide L. A response surface method based on support vector machines trained with an adaptive experimental design. Struct Saf 2012;39:14–21.
[30] Dai H, Zhang B, Wang W. A multiwavelet support vector regression method for efficient reliability assessment. Reliab Eng Syst Saf 2015;136:132–9. doi:10.1016/j.ress.2014.12.002.
[31] Bourinet JM. Rare-event probability estimation with adaptive support vector regression surrogates. Reliab Eng Syst Saf 2016;150:210–21.

doi:10.1016/j.ress.2016.01.023.
[32] Chapelle O. Choosing Multiple Parameters for Support Vector Machines. Mach Learn 2002;46:131–59.
[33] Ito K, Nakano R. Optimizing Support Vector Regression Hyperparameters Based on Cross-Validation. Proc Int Jt Conf Neural Networks 2003;3:2077–82. doi:10.1109/ijcnn.2003.1223728.
[34] Hsu C-W, Chang C-C, Lin C-J. A Practical Guide to Support Vector Classification. BJU Int 2008;101:1396–400.
[35] Chen KY. Forecasting systems reliability based on support vector regression with genetic algorithms. Reliab Eng Syst Saf 2007;92:423–32. doi:10.1016/j.ress.2005.12.014.
[36] Lins ID, Moura MDC, Zio E, Droguett EL. A particle swarm-optimized support vector machine for reliability prediction. Qual Reliab Eng Int 2012;28:141–58. doi:10.1002/qre.1221.
[37] Zhao W, Tao T, Ding Z, Zio E. A dynamic particle filter-support vector regression method for reliability prediction. Reliab Eng Syst Saf 2013;119:109–16. doi:10.1016/j.ress.2013.05.021.
[38] Tipping ME. Sparse Bayesian Learning and the Relevance Vector Machine. J Mach Learn Res 2001;1:211–44. doi:10.1162/15324430152748236.
[39] Zhou C, Lu Z, Yuan X. Use of relevance vector machine in structural reliability analysis. J Aircr 2013;50:1726–33. doi:10.2514/1.C031950.
[40] Changcong Z, Zhenzhou L, Feng Z, Zhufeng Y. An adaptive reliability method combining relevance vector machine and importance sampling. Struct Multidiscip Optim 2015;52:945–57. doi:10.1007/s00158-015-1287-z.
[41] Ghosh S, Chakraborty S. Seismic fragility analysis of structures based on Bayesian linear regression demand models. Probabilistic Eng Mech 2020;61:103081. doi:10.1016/J.PROBENGMECH.2020.103081.
[42] Guo Z, Bai G. Application of Least Squares Support Vector Machine for Regression to Reliability Analysis. Chinese J Aeronaut 2009;22:160–6. doi:10.1016/S1000-9361(08)60082-5.
[43] Khatibinia M, Javad Fadaee M, Salajegheh J, Salajegheh E. Seismic reliability assessment of RC structures including soil–structure interaction using wavelet weighted least squares support vector machine. Reliab Eng Syst Saf 2013;110:22–33. doi:10.1016/J.RESS.2012.09.006.
[44] Feng J, Liu L, Wu D, Li G, Beer M, Gao W. Dynamic reliability analysis using the extended support vector regression (X-SVR). Mech Syst Signal Process 2019;126:368–91. doi:10.1016/j.ymssp.2019.02.027.
[45] Cheng K, Lu Z. Adaptive Bayesian support vector regression model for structural reliability analysis. Reliab Eng Syst Saf 2021;206:107286. doi:10.1016/j.ress.2020.107286.
[46] Dai H, Zhang H, Wang W. A support vector density-based importance sampling for reliability assessment. Reliab Eng Syst Saf 2012;106:86–93. doi:10.1016/J.RESS.2012.04.011.
[47] Dey S, Mukhopadhyay T, Adhikari S. Metamodel based high-fidelity stochastic analysis of composite laminates: A concise review with critical comparative assessment. Compos Struct 2017;171:227–50. doi:10.1016/j.compstruct.2017.01.061.
[48] Xu Z, Saleh JH. Machine learning for reliability engineering and safety applications:

Review of current status and future opportunities. Reliab Eng Syst Saf 2021;211:107530. doi:10.1016/j.ress.2021.107530.

[49] Afshari SS, Enayatollahi F, Xu X, Liang X. Machine learning-based methods in structural reliability analysis: A review. Reliab Eng Syst Saf 2022;219:108223. doi:10.1016/J.RESS.2021.108223.

[50] Boyd S, Vandenberghe L. Convex optimization. Cambridge University Press; 2004.

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

[52] Gunn S. Support Vector Machines for Classification and Regression. 1998.

[53] Hurtado JE, Alvarez DA. Classification Approach for Reliability Analysis with Stochastic Finite-Element Modeling. J Struct Eng 2003;129:1141–9. doi:10.1061/(asce)0733-9445(2003)129:8(1141).

[54] Hurtado JE, Alvarez DA. An optimization method for learning statistical classifiers in structural reliability. Probabilistic Eng Mech 2010;25:26–34. doi:10.1016/j.probengmech.2009.05.006.

[55] Zhang J, Xiao M, Gao L, Chu S. Probability and interval hybrid reliability analysis based on adaptive local approximation of projection outlines using support vector machine. Comput Civ Infrastruct Eng 2019;34:991–1009. doi:10.1111/mice.12480.

[56] Ling C, Lu Z. Support vector machine-based importance sampling for rare event estimation. Struct Multidiscip Optim 2021;63:1609–31. doi:10.1007/s00158-020-02809-8.

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

[58] Alibrandi U, Ma C, Koh CG. Secant Hyperplane Method for Structural Reliability Analysis. J Eng Mech 2016;142:04015098. doi:10.1061/(asce)em.1943-7889.0001024.

[59] Alibrandi U, Alani A, Koh CG. Implications of High-Dimensional Geometry for Structural Reliability Analysis and a Novel Linear Response Surface Method Based on SVM. Int J Comput Methods 2015;12:1540016. doi:10.1142/S0219876215400162.

[60] Basudhar A, Missoum S. Adaptive explicit decision functions for probabilistic design and optimization using support vector machines. Comput Struct 2008;86:1904–17. doi:10.1016/J.COMPSTRUC.2008.02.008.

[61] Basudhar A, Missoum S. An improved adaptive sampling scheme for the construction of explicit boundaries. Struct Multidiscip Optim 2010;42:517–29. doi:10.1007/s00158-010-0511-0.

[62] Kanamori T, Fujiwara S, Takeda A. Breakdown point of robust support vector machines. Entropy 2017;19:83. doi:10.3390/e19020083.

[63] Yang L, Dong H. Robust support vector machine with generalized quantile loss for classification and regression. Appl Soft Comput J 2019;81:105483. doi:10.1016/j.asoc.2019.105483.

[64] Elreedy D, Atiya AF. A Comprehensive Analysis of Synthetic Minority Oversampling Technique (SMOTE) for handling class imbalance. Inf Sci (Ny) 2019;505:32–64. doi:10.1016/j.ins.2019.07.070.

[65] Schölkopf B, Platt JC, Shawe-Taylor J, Smola AJ, Williamson RC. Estimating the support of a high-dimensional distribution. Neural Comput 2001;13:1443–71. doi:10.1162/089976601750264965.

[66] Klikowski J, Woźniak M. Employing one-class svm classifier ensemble for imbalanced data stream classification. Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 12140 LNCS, Springer Science and Business Media Deutschland GmbH; 2020, p. 117–27. doi:10.1007/978-3-030-50423-6_9.

[67] Zhan L, Liu J, Zhang M, Zhou C, Zhang L, Shi T. One-class support vector machine based schemes for structural reliability assessment under imbalanced sample conditions. IEEE Access 2020;8:184350–9. doi:10.1109/ACCESS.2020.3027815.

[68] Li X, Li X, Su Y. A hybrid approach combining uniform design and support vector machine to probabilistic tunnel stability assessment. Struct Saf 2016;61:22–42. doi:10.1016/J.STRUSAFE.2016.03.001.

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

[70] Liu Y, Lu N, Yin X, Noori M. An adaptive support vector regression method for structural system reliability assessment and its application to a cable-stayed bridge. Proc Inst Mech Eng Part O J Risk Reliab 2016;230:204–19. doi:10.1177/1748006X15623869.

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

[72] Roy A, Chakraborty S. Reliability analysis of structures by a three-stage sequential sampling based adaptive support vector regression model. Reliab Eng Syst Saf 2022;219:108260. doi:10.1016/j.ress.2021.108260.

[73] Li HS, Zhao AL, Tee KF. Structural reliability analysis of multiple limit state functions using multi-input multi-output support vector machine. Adv Mech Eng 2016;8:1–11. doi:10.1177/1687814016671447.

[74] Sabzekar M, Naghibzadeh M. Fuzzy c-means improvement using relaxed constraints support vector machines. Appl Soft Comput 2013;13:881–90. doi:10.1016/J.ASOC.2012.09.018.

[75] Sabzekar M, Hasheminejad SMH. Robust regression using support vector regressions. Chaos, Solitons & Fractals 2021;144:110738. doi:10.1016/J.CHAOS.2021.110738.

[76] Cui W, Yan X. Adaptive weighted least square support vector machine regression integrated with outlier detection and its application in QSAR. Chemom Intell Lab Syst 2009;98:130–5. doi:10.1016/j.chemolab.2009.05.008.

[77] Chuang CC, Lee ZJ. Hybrid robust support vector machines for regression with outliers. Appl Soft Comput J 2011;11:64–72. doi:10.1016/j.asoc.2009.10.017.

[78] Samui P, Lansivaara T, Kim D. Utilization relevance vector machine for slope reliability analysis. Appl Soft Comput 2011;11:4036–40. doi:10.1016/J.ASOC.2011.03.009.

[79] Li S, Zhao H, Ru Z. Relevance vector machine-based response surface for slope reliability analysis. Int J Numer Anal Methods Geomech 2017;41:1332–46. doi:10.1002/NAG.2683.

[80] Li TZ, Pan Q, Dias D. Active learning relevant vector machine for reliability analysis. Appl Math Model 2021;89:381–99. doi:10.1016/j.apm.2020.07.034.

[81] Wang Y, Xie B, E S. Adaptive relevance vector machine combined with Markov-chain-based importance sampling for reliability analysis. Reliab Eng Syst Saf

2022;220:108287. doi:10.1016/J.RESS.2021.108287.

[82] Zhao H, Ru Z, Chang X, Yin S, Li S. Reliability analysis of tunnel using least square support vector machine. Tunn Undergr Sp Technol 2014;41:14–23.

[83] Samui P, Lansivaara T, Bhatt MR. Least Square Support Vector Machine Applied to Slope Reliability Analysis. Geotech Geol Eng 2013;31:1329–34. doi:10.1007/s10706-013-9654-2.

[84] Kumar M, Samui P. Reliability Analysis of Settlement of Pile Group in Clay Using LSSVM, GMDH, GPR. Geotech Geol Eng 2020;38:6717–30. doi:10.1007/s10706-020-01464-6.

[85] Song H, Choi KK, Lee I, Zhao L, Lamb D. Adaptive virtual support vector machine for reliability analysis of high-dimensional problems. Struct Multidiscip Optim 2013;47:479–91. doi:10.1007/s00158-012-0857-6.

[86] Chocat R, Beaucaire P, Debeugny L, Lefebvre JP, Sainvitu C, Breitkopf P, et al. Damage tolerance reliability analysis combining Kriging regression and support vector machine classification. Eng Fract Mech 2019;216:106514. doi:10.1016/j.engfracmech.2019.106514.

[87] Demyanov S, Bailey J, Ramamohanarao K, Leckie C. AIC and BIC based approaches for SVM parameter value estimation with RBF kernels. J Mach Learn Res 2012;25:97–112.

[88] Zhao W, Tao T, Zio E. System reliability prediction by support vector regression with analytic selection and genetic algorithm parameters selection. Appl Soft Comput J 2015;30:792–802. doi:10.1016/j.asoc.2015.02.026.

[89] Rubinstein RY. Optimization of computer simulation models with rare events. Eur J Oper Res 1997;99:89–112. doi:10.1016/S0377-2217(96)00385-2.

[90] Chang MW, Lin CJ. Leave-one-out bounds for support vector regression model selection. Neural Comput 2005;17:1188–222. doi:10.1162/0899766053491869.