INTERNATIONAL JOURNAL FOR NUMERICAL AND ANALYTICAL METHODS IN GEOMECHANICS
Int. J. Numer. Anal. Meth. Geomech. 2012; **36**:100–110
Published online 6 December 2010 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/nag.997

Application of statistical learning algorithms to ultimate bearing capacity of shallow foundation on cohesionless soil

Pijush Samui*,†,‡

Centre for Disaster Mitigation and Management, VIT University, Vellore 632 014, India

SUMMARY

This study employs two statistical learning algorithms (Support Vector Machine (SVM) and Relevance Vector Machine (RVM)) for the determination of ultimate bearing capacity (q_u) of shallow foundation on cohesionless soil. SVM is firmly based on the theory of statistical learning, uses regression technique by introducing ε-insensitive loss function. RVM is based on a Bayesian formulation of a linear model with an appropriate prior that results in a sparse representation. It also gives variance of predicted data. The inputs of models are width of footing (B), depth of footing (D), footing geometry (L/B), unit weight of sand (γ) and angle of shearing resistance (ϕ). Equations have been developed for the determination of q_u of shallow foundation on cohesionless soil based on the SVM and RVM models. Sensitivity analysis has also been carried out to determine the effect of each input parameter. This study shows that the developed SVM and RVM are robust models for the prediction of q_u of shallow foundation on cohesionless soil. Copyright © 2010 John Wiley & Sons, Ltd.

Received 16 June 2010; Accepted 17 October 2010

KEY WORDS: support vector machine; relevance vector machine; artificial intelligence; ultimate bearing capacity; shallow foundation; cohesionless soil; sensitivity analysis

1. INTRODUCTION

Ultimate bearing capacity (q_u) is an important parameter for designing shallow foundation on cohesionless soil. Geotechnical engineers use different methods for the determination of q_u of shallow foundation on cohesionless soil [1–3]. Experimental techniques have also been used to study q_u of shallow foundation on cohesionless soil [3–16]. The available experimental techniques have some limitations [17].

Recently, different Artificial Intelligence (AI) techniques (Artificial Neural Network (ANN); Fuzzy Inference System (FIS) and Adaptive Neuro Fuzzy Inference System (ANFIS)) have been successfully adopted for the prediction of q_u of shallow foundation on cohesionless soil [17]. ANN has been successfully used to solve different problems in geotechnical engineering [18–24]. But, the developed ANN model has some limitations, such as black box approach, arriving at local minima, less generalization capability, slow convergence speed, overfitting problem and absence of probabilistic output [25, 26]. Furthermore, there is no proper method to determine the number of hidden layers in the ANN model [27]. The developed FIS model determines the fuzzy rules with difficulty [27]. The developed AI models did not give any equation for the prediction of q_u

*Correspondence to: Pijush Samui, Centre for Disaster Mitigation and Management, VIT University, Vellore 632 014, India.
†E-mail: pijush.phd@gmail.com
‡Associate Professor.

of shallow foundation on cohesionless soil. It did not perform any sensitivity analysis to show the effect of each input parameter on q_u.

This study examines the capability of two statistical learning algorithms (Support Vector Machine (SVM) and Relevance Vector Machine (RVM)) for the determination of q_u of shallow foundation on cohesionless soil. The study uses the database collected by Padmini *et al.* [17]. The First statistical learning algorithm adopts SVM, a novel type of learning machine based on the statistical learning theory [28]. The details of SVM and its application to geotechnical engineering problems can be found in the literature [28–32]. The second statistical learning algorithm uses RVM. RVM introduced by Tipping [33] produces sparse solutions using an improper hierarchical prior and optimizing over hyperparameters. RVM is referred to as Bayesian kernel methods that choose sparse basis sets using an 'Automatic Relevance Determination' [34] style prior that pushes nonessential weights to zero. The paper has the following aims:

- To examine the capability of SVM and RVM models for the prediction of q_u of shallow foundation on cohesionless soil.
- To determine the variance of predicted output based on the developed RVM model.
- To develop equations for the determination of q_u of shallow foundation on cohesionless soil based on the developed SVM and RVM models.
- To make a comparative study between SVM, RVM and the other available methods.
- To do sensitivity analysis for the determination of the effect of each input parameter on q_u.

2. DETAILS OF SVM

SVM has originated from the concept of statistical learning theory pioneered by Boser *et al.* [35]. This study uses SVM as a regression technique by introducing an ε-insensitive loss function. In this section, a brief introduction about how to construct SVM for regression problem is presented. More details can be found in many publications [28, 35–37]. SVM uses a risk function consisting of the empirical error and a regularization term, which is derived from the structural risk minimization (SRM) principle. Considering a set of training data $(x_1, y_1), \dots, (x_l, y_l), x \in R^n, y \in r$, where x is the input, y is the output, R^N is the N-dimensional vector space and r is the one-dimensional vector space. The five input variables used for the SVM model in this study are width of footing (B), depth of footing (D), footing geometry (L/B), unit weight of sand (γ) and angle of shearing resistance (ϕ). The output of the SVM model is q_u. So, in this study, $x = [B, D, (L/B), \gamma, \phi]$ and $y = q_u$.

The ε-insensitive loss function can be described in the following way:

$$L_\varepsilon(y) = 0 \quad \text{for } |f(x) - y| < \varepsilon \ \text{ otherwise } L_\varepsilon(y) = |f(x) - y| - \varepsilon. \tag{1}$$

This defines an ε tube so that if the predicted value is within the tube the loss is zero, whereas if the predicted point is outside the tube, the loss is equal to the absolute value of the deviation minus ε. The main aim of SVM is to find a function $f(x)$ that gives a deviation of ε from the actual output and at the same time is as flat as possible. Let us assume a linear function

$$f(x) = (w \cdot x) + b \quad w \in R^n, \quad b \in r, \tag{2}$$

where w is an adjustable weight vector and b is the scalar threshold.

Flatness in the case of (3) means that one seeks a small w. One way of obtaining this is by minimizing the Euclidean norm $\|w\|^2$. This is equivalent to the following convex optimization problem:

$$\begin{aligned} &\text{Minimize} && \tfrac{1}{2}\|w\|^2 \\ &\text{Subjected to} && y_i - (\langle w \cdot x_i \rangle + b) \leqslant \varepsilon, \quad i = 1, 2, \dots, l \\ &&& (\langle w \cdot x_i \rangle + b) - y_i \leqslant \varepsilon, \quad i = 1, 2, \dots, l \end{aligned} \tag{3}$$

DOI: 10.1002/nag

The above convex optimization problem is feasible. Sometimes, however, this may not be the case, or one may also want to allow for some errors, analogous to the 'soft margin' loss function [38], which was used in SVM by Cortes and Vapnik [36]. The parameters ξ_i, ξ_i^* are slack variables that determine the degree to which samples with error more than ε be penalized. In other words, any error smaller than ε does not require ξ_i, ξ_i^* and hence does not enter the objective function because these data points have a value of zero for the loss function. The slack variables (ξ_i, ξ_i^*) have been introduced to avoid infeasible constraints of the optimization problem (3).

$$\begin{aligned} &\text{Minimize} && \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{l}(\xi_i+\xi_i^*) \\ &\text{Subjected to} && y_i - (\langle w.x_i\rangle + b) \leqslant \varepsilon + \xi_i, \quad i=1,2,\ldots,l \\ & && (\langle w.x_i\rangle + b) - y_i \leqslant \varepsilon + \xi_i^*, \quad i=1,2,\ldots,l \\ & && \xi_i \geqslant 0 \quad \text{and} \quad \xi_i^* \geqslant 0, \quad i=1,2,\ldots,l \end{aligned} \tag{4}$$

The constant $0<C<\infty$ determines the trade-off between the flatness of f and the amount to which deviations larger than ε are tolerated [29]. This optimization problem (4) is solved by the Lagrangian Multipliers [28], and its solution is given by

$$f(x) = \sum_{i=1}^{l}(\alpha_i - \alpha_i^*)(x_i \cdot x) + b, \tag{5}$$

where $b = -(1/2)w[x_r + x_s]$, α_i, α_i^* are the Lagrangian Multipliers and l is the number of dataset. An important aspect is that some Lagrange multipliers (α_i, α_i^*) will be zero, implying that these training objects are considered to be irrelevant for the final solution (sparseness). The training objects with nonzero Lagrange multipliers are called support vectors.

When linear regression is not appropriate, then input data have to be mapped into a high dimensional feature space through some nonlinear mapping [35]. The two steps that are involved are first to make a fixed nonlinear mapping of the data onto the feature space and then carry out a linear regression in the high dimensional space. The input data are mapped onto the feature space by a map Φ. The dot product given by $\Phi(x_i)\cdot\Phi(x_j)$ is computed as a linear combination of the training points. The concept of kernel function $[K(x_i, x_j) = \Phi(x_i).\Phi(x_j)]$ has been introduced to reduce the computational demand [36, 39]. So, Equation (2) becomes

$$f(x) = \sum_{i=1}^{l}(\alpha_i - \alpha_i^*)K(x_i \cdot x_j) + b. \tag{6}$$

Some common kernels, such as polynomial (homogeneous), polynomial (nonhomogeneous), radial basis function, Gaussian function and sigmoid, have been used for nonlinear cases. Figure 1 shows a typical architecture of SVM for q_u prediction.

This study uses the above SVM for the prediction of q_u of shallow foundation on cohesionless soil. The database has been collected from the work of Padmini *et al.* [17]. The database contains the information about B, D, L/B, γ, ϕ and q_u. The data have been further divided into two subsets; a training dataset, to construct the model, and a testing dataset, to estimate the model performance. So, for our study a set of 78 data is considered as the training dataset and remaining set of 19 data is considered as the testing dataset. The data are normalized between 0 and 1. Radial basis function has been used as the kernel function. The design value of C, ε and width (σ) of radial basis function has been determined by a trial-and-error approach.

In this study, a sensitivity analysis has been carried out to extract the cause and effect relationship between the inputs and outputs of the SVM model. The basic idea is that each input of the model is offset slightly and the corresponding change in the output is reported. The procedure has been taken from the work of Liong *et al.* [40]. According to Liong *et al.* [40], the sensitivity (S) of

DOI: 10.1002/nag

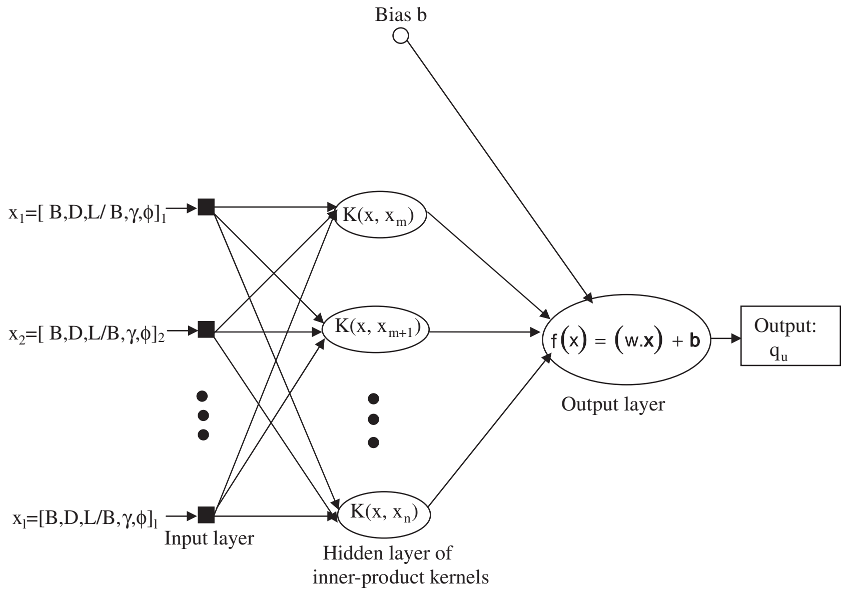

Figure 1. The SVM architecture for the q_u prediction.

each input parameter has been calculated by the following formula:

$$S(\%) = \frac{1}{78}\sum_{j=1}^{78}\left(\frac{\%\text{ change in ouput}}{\%\text{ change in input}}\right)_j \times 100. \tag{7}$$

The analysis has been carried out on the trained model by varying each of the input parameter, one at a time, at a constant rate of 30%. In the present study, training, testing and sensitivity analysis of SVM have been carried out using MATLAB.

3. DETAILS OF RVM

The RVM, introduced by Tipping [33], is a sparse linear model.

Let $D=\{(x_i,t_i), i=1,\ldots,N\}$ be a dataset of observed values, where x_i is the input, t_i is the output, $x_i \in R^d$ and $t_i \in R$. In this study, the input parameters are B, D, L/B, γ and ϕ. So, $x=[B, D, L/B, \gamma, \phi]$. The output of the RVM model is q_u of shallow foundation on cohesionless soil. So $t=[q_u]$. One can express the output as the sum of an approximation vector $y=(y(x_1),\ldots,y(x_N))^{\mathrm{T}}$, and zero mean random error (noise) vector $\varepsilon=(\varepsilon_1,\ldots,\varepsilon_N)^{\mathrm{T}}$, where $\varepsilon_n \sim \mathbf{N}(0,\sigma^2)$ and $\mathbf{N}(0,\sigma^2)$ is the normal distribution with mean 0 and variance σ^2. So, the output can be written as

$$t_n = y(x_n,\omega)+\varepsilon_n, \tag{8}$$

where ω is the parameter vector. Let we assume

$$p(t_n|x) \sim \mathbf{N}(y(x_n),\sigma^2), \tag{9}$$

where $\mathbf{N}(y(x_n),\sigma^2)$ is the normal distribution with mean $y(x_n)$ and variance $\sigma^2 \cdot y(x)$ can be expressed as a linearly weighted sum of M nonlinear fixed basis function,

$$\begin{aligned}&\{\Phi_j(x)|j=1,\ldots,M\}:\\ &y(x;\omega)=\sum_{i=1}^{M}\omega_i\Phi_i(x)=\mathbf{\Phi}\omega.\end{aligned} \tag{10}$$

DOI: 10.1002/nag

The likelihood of the complete dataset can be written as

$$p(t|w,\sigma^2)=(2\pi\sigma^2)^{-N/2}\exp\left\{-\frac{1}{2\sigma^2}\|t-\Phi w\|^2\right\}, \tag{11}$$

where $t=(t_1,\ldots,t_N)^{\mathrm{T}}$, $\omega=(\omega_0,\ldots,\omega_N)$ and

$$\Phi^{\mathrm{T}}=\begin{bmatrix}1 & K(x_1,x_1) & K(x_1,x_2) & \cdots & K(x_1,x_n)\\ 1 & K(x_1,x_2) & K(x_2,x_2) & \cdots & K(x_2,x_n)\\ \vdots & \vdots & \vdots & \vdots & \vdots\\ 1 & K(x_n,x_1) & K(x_n,x_2) & \cdots & K(x_n,x_n)\end{bmatrix},$$

where $K(x_i,x_n)$ is a kernel function.

To prevent overfitting, automatic relevance detection (ARD) prior is set over the weights

$$p(w|\alpha)=\prod_{i=0}^{N}N(\omega_i|0,\alpha_i^{-1}), \tag{12}$$

where α is a hyperparameter vector that controls how far each weight is allowed to deviate from zero [41]. Consequently, using the Bayes rule, the posterior over all unknowns could be computed given the defined noninformative prior distribution

$$p(w,\alpha,\sigma^2/t)=\frac{p(y/w,\alpha,\sigma^2)p(w,\alpha,\sigma)}{\int p(t/w,\alpha,\sigma^2)p(w,\alpha,\sigma^2)\,\mathrm{d}w\mathrm{d}\alpha\mathrm{d}\sigma^2}. \tag{13}$$

Full analytical solution of this integral (6) is obdurate. Thus, decomposition of the posterior according to $p(w,\alpha,\sigma^2/t)=p(w/t,\alpha,\sigma^2)p(\alpha,\sigma^2/t)$ is used to facilitate the solution [42]. The posterior distribution over the weights is thus given by

$$p(w/t,\alpha,\sigma^2)=\frac{p(t/w,\sigma^2)p(w/\alpha)}{p(t/\alpha,\sigma^2)}. \tag{14}$$

The resulting posterior distribution over the weights is the multivariate Gaussian distribution

$$p(w/t,\alpha,\sigma^2)=\mathbf{N}(\mu,\textstyle\sum), \tag{15}$$

where the mean and the covariance are, respectively, given by

$$\textstyle\sum=(\sigma^{-2}\Phi^{\mathrm{T}}\Phi+A)^{-1}, \tag{16}$$

$$\mu=\sigma^{-2}\textstyle\sum\Phi^{\mathrm{T}}t, \tag{17}$$

with diagonal $A=\mathrm{diag}(\alpha_0,\ldots,\alpha_N)$.

For uniform hyperpriors over α and σ^2, one needs only to maximize the term $p(t/\alpha,\sigma^2)$:

$$\begin{aligned}p(t/\alpha,\sigma^2)&=\int p(t/w,\sigma^2)p(w/\alpha)\mathrm{d}w\\ &=\left((2\pi)^{-N/2}/\sqrt{|\sigma^2+\Phi A^{-1}\Phi^{\mathrm{T}}|}\right)\times\exp\left\{-\frac{1}{2}y^{\mathrm{T}}(\sigma^2+\Phi A^{-1}\Phi^{\mathrm{T}})^{-1}y\right\}.\end{aligned} \tag{18}$$

Maximization of this quantity is known as the type II maximum likelihood method [43, 44] or the 'evidence for hyperparameter' [45]. Hyperparameter estimation is carried out in iterative formulae, e.g. gradient descent on the objective function [42]. The outcome of this optimization is that many elements of α go to infinity such that w will have only few non-zero weights that will be considered as relevant vectors.

This paper also uses the above methodology for the prediction of q_u of shallow foundation on cohesionless soil. In RVM, the same training, testing, normalization technique, sensitivity analysis and kernel function have used in the SVM model. The design value of σ has been determined by the trial-and-error approach. The RVM model has been developed by MATLAB.

4. RESULTS AND DISCUSSION

Coefficient of correlation (R), coefficient of efficiency (E), root mean square error (RMSE) and mean bias error (MBE) have been used to assess the performance of the SVM and RVM models. The design value of C, ε and σ is 10, 0.02 and 0.5, respectively. The number of support vectors is 27. Figures 2 and 3 show the performance of SVM for training and testing datasets, respectively. The value of R is close to 1 for both training and testing datasets. So, the developed SVM model has the ability to predict q_u of shallow foundation on cohesionless soil. Table I summarizes the values of different performance indexes for the SVM model. The following equation (by putting $K(x_i, x) = \exp\{-((x_i - x)(x_i - x)^{\mathrm{T}})/2\sigma^2\}$, $l = 97$, $\sigma = 0.5$ and $b = 0$ in Equation (6)) has been developed from the developed SVM model:

$$q_u = \sum_{i=1}^{97} (\alpha_i - \alpha_i^*) \exp\left\{-\frac{(x_i - x)(x_i - x)^{\mathrm{T}}}{0.5}\right\}. \tag{19}$$

The value of $(\alpha_i - \alpha_i^*)$ has been given in Figure 4.

For RVM, the design value of σ is 0.1. The number of relevance vector is 39. Figure 5 illustrates the performance of RVM for the training dataset. From Figure 5, it is clear that the developed RVM has successfully captured the input and output relations. The performance of testing dataset

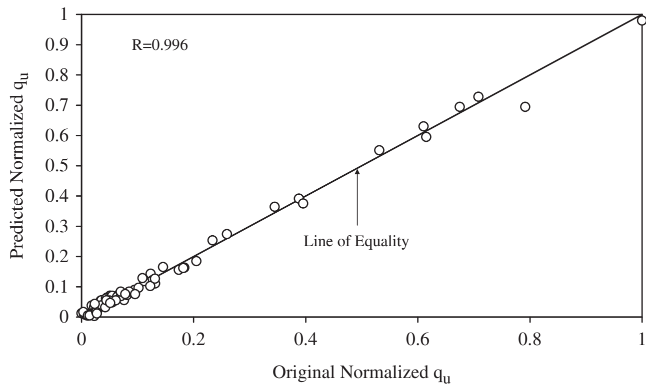

Figure 2. Performance of the SVM model for the training dataset.

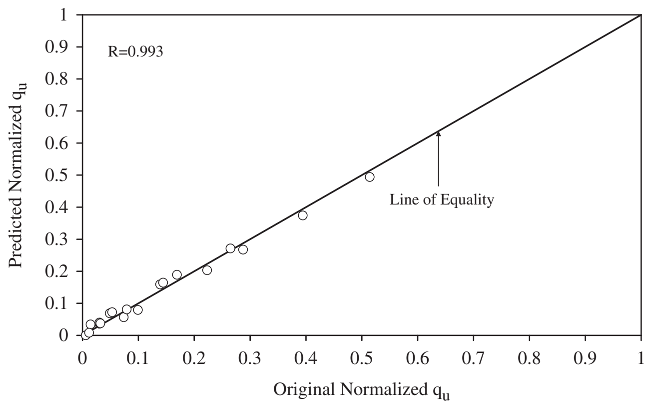

Figure 3. Performance of the SVM model for the testing dataset.

Table I. Comparison between different models.

Performance index	SVM		RVM		ANN		ANFIS		FIS	
	Training	Testing	Training	Testing	Training	Testing	Training	Testing	Training	Testing
R	0.996	0.993	0.998	0.996	0.995	0.992	0.9986	0.996	0.990	0.9899
E	0.992	0.986	0.997	0.994	0.989	0.983	0.997	0.992	0.980	0.972
RMSE (kPa)	46.5873	50.0396	28.1928	29.9289	52.9	77.2	26.4	52.3	71.1	98
MBE (kPa)	−0.8796	2.6027	1.9326	5.7296	−1.78	−12.04	0	11.50	0	13.93

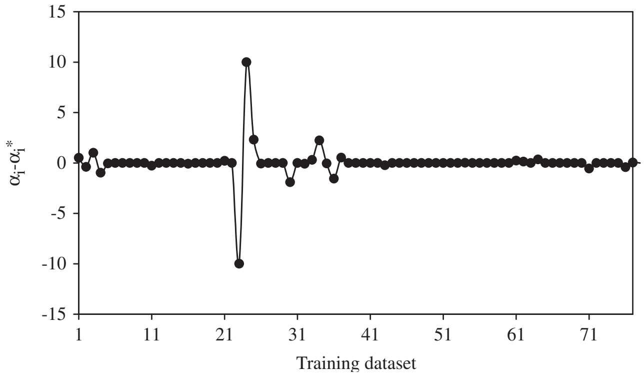

Figure 4. Values of $(\alpha_i - \alpha_i^*)$ for the SVM model.

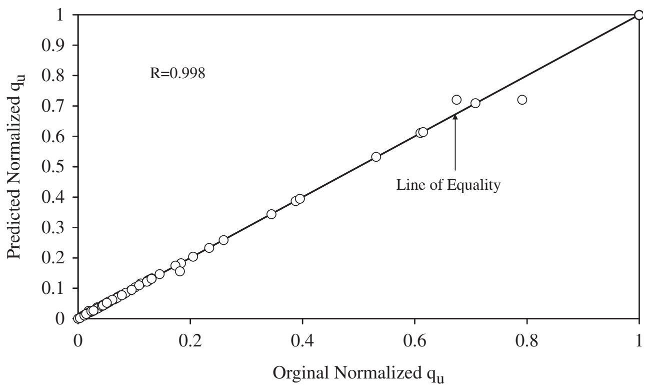

Figure 5. Performance of the RVM model for the training dataset.

is shown in Figure 6. Figure 6 shows that the developed RVM model has the ability to predict q_u of shallow foundation on cohesionless soil. Table I shows the values of performance index for the RVM model. The developed RVM model gives the following equation for prediction of q_u of shallow foundation:

$$q_u = \sum_{i=1}^{97} w_i \exp\left\{-\frac{(x_i - x)(x_i - x)^{\mathrm{T}}}{0.02}\right\} \tag{20}$$

Figure 7 shows the value of w for the RVM model.

Table I shows that the performance of the developed RVM model is slightly better than other AI models. The developed RVM model also gives the variance of the predicted data. Figures 8 and 9 represent the variance of training and testing datasets, respectively. The obtained prediction variance allows one to assign a confidence interval about the model prediction.

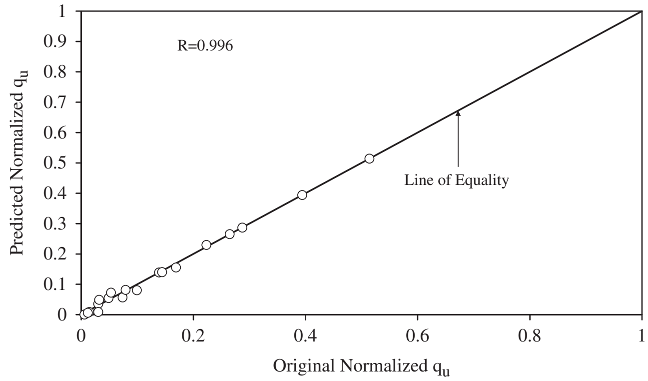

Figure 6. Performance of the RVM model for the testing dataset.

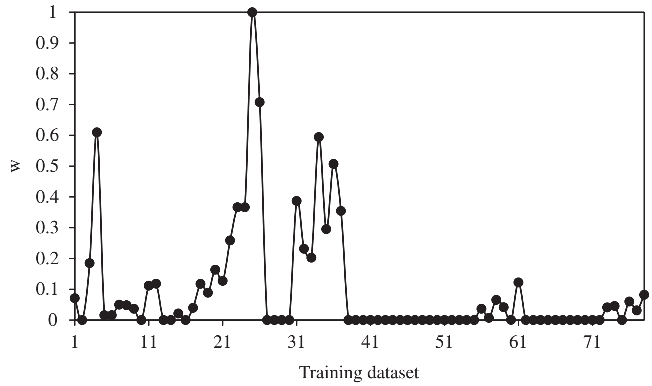

Figure 7. Values of w for the RVM model.

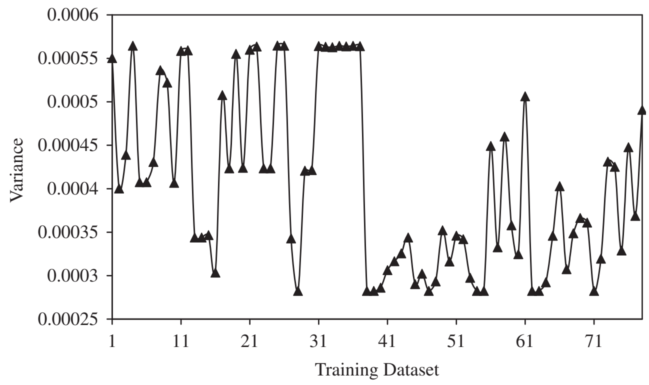

Figure 8. Variance of the training dataset for the RVM model.

Sensitivity results of the SVM and RVM models have been shown in Figure 10. Figure 10 shows that D has the greatest effect on the bearing capacity of all the physical properties of the shallow foundation. Of all the properties of soil, ϕ has greatest influence on the bearing capacity. Therefore, the sensitivity results match well with the result of Meyerhof [4] and Padmini *et al.* [17]. The sensitivity results of SVM and RVM are almost similar.

A comparative study has been carried out between the developed models (SVM and RVM) and traditional methods (Meyerhof [4], Vesic [3] and Brinch Hansen [46]) for the prediction of q_u of

DOI: 10.1002/nag

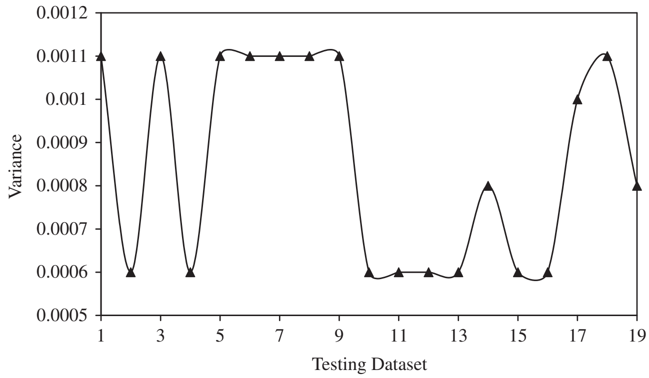

Figure 9. Variance of the testing dataset for the RVM model.

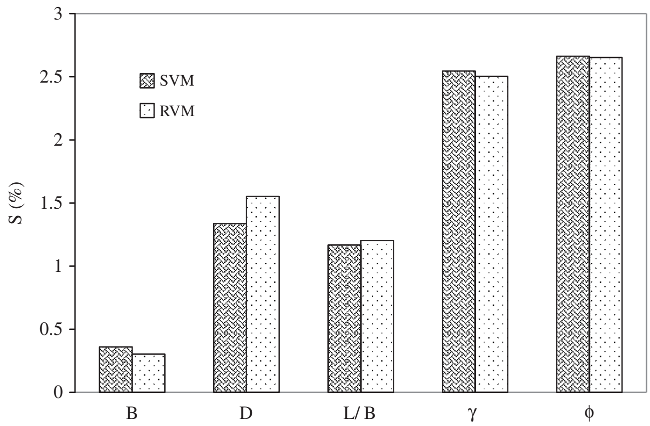

Figure 10. Sensitivity analysis of input parameters.

Table II. Performance indexes of traditional methods.

	Traditional methods		
Performance index	Meyerhof [4]	Vesic [3]	Brinch Hansen [46]
R	0.9412	0.9496	0.9457
E	0.874	0.815	0.727
RMSE (kPa)	207.3	251.3	305.3
MBE (kPa)	−45.27	−105.30	−161.29

shallow foundation on cohesionless soil. The comparison was done for the testing dataset. Table II shows the values of performance indexes for the traditional methods. It is clear from Tables I and II that the developed SVM and RVM models outperform the traditional methods. The RVM model uses 50% of training data as relevance vectors. About 34.61% of training data has been used as support vector by SVM. So, both the models (SVM and RVM) give sparse solution. The developed SVM produces more sparse solution than RVM. Sparseness means that a significant number of the weights are zero (or effectively zero), which has the consequence of producing compact, computationally efficient models, which in addition are simple and therefore produce smooth functions. SVM uses three tuning parameters (C, ε and σ). But, RVM uses only one tuning parameter (σ).

5. CONCLUSION

Ultimate bearing capacity of shallow foundation on cohesionless soil has been determined by SVM and RVM. Both the developed models give promising result and show good generalization capability. A comparative study has been carried out between the developed models and the available methods. The performance of the RVM model is slightly better than the available methods. The developed RVM has also been used for the measurement of uncertainty. Geotechnical engineers can use the developed equations for the prediction of ultimate bearing capacity of shallow foundation on cohesionless soil. The results of sensitivity analysis match well with the available results. It can be concluded that SVM and RVM can be further employed in several kinds of geotechnical engineering problems with inherent uncertainties and imperfections.

REFERENCES

1. Terzaghi K. *Theoretical Soil Mechanics*. Wiley: New York, 1943.
2. Meyerhof GG. Some recent research on the bearing capacity of foundations. *Canadian Geotechnical Journal* 1963; **1**(1):16–26.
3. Vesic AS. Analysis of ultimate loads of shallow foundations. *JSMFD*, *ASCE* 1973; **99**(1):45–73.
4. Meyerhof GG. The bearing capacity of sand. *Ph.D. Thesis*, University of London, London, 1950.
5. Meyerhof GG. Factors influencing bearing capacity of foundations. *Géotechnique* 1955; **5**:227–242.
6. Muhs E. Uber die Zulässige Belastung nichtibindiger Boden. Mitteilungen der Deutschen Forschungsgesellshaft Fu/r Bodenmechanik (Degebo), Berlin, Heft 16. 1963; 102–121.
7. Muhs E. Discussion. *Proceedings of the Sixth International Conference on Soil Mechanics and Foundation Engineering*, vol. 3. University of Toronto Press: Toronto, 1965; 419–440.
8. De Beer EE. Bearing capacity and settlement of shallow foundations on sand. *Proceedings*, *Symposium on Bearing Capacity and Settlements of Foundations*. Duke University: Durham, NC, 1965; 15–33.
9. Ovesen NK. Centrifugal testing applied to bearing capacity problems of footings on sand. *Géotechnique* 1975; **25**(1):394–401.
10. Yamaguchi H, Kimura T, Fujii N. On the scale effect of footings in dense sand. *Proceedings of the Ninth International Conference on Soil Mechanics and Foundation Engineering*, Tokyo, Japan, vol. 1, 1977; 795–798.
11. Kimura T, Kusakabe O, Saitoh K. Geotechnical model tests of bearing capacity problems in a centrifuge. *Géotechnique* 1985; **35**(1):33–45.
12. Kutter BL, Abghari A, Cheney JA. Strength parameters for bearing capacity of sand. *Journal of Geotechnical Engineering* 1988; **114**(4):491–497.
13. Pu JL, Ko HY. Experimental determination of bearing capacity in sand by centrifuge testing. *Proceedings of the International Conference on Centrifuge '88*. Balkema: Rotterdam, The Netherlands, 1988; 293–299.
14. Larsen P, Ovesen NK. Bearing capacity of square footings on sand. *Proceedings of the 12th International Conference on Soil Mechanics and Foundation Engineering*, vol. 3. A. A. Balkema: Rotterdam, 1989; 2077–2080.
15. Kusakabe O, Yamaguchi H, Morikage A. Experiment and analysis on the scale effect of Ng for circular and rectangular footings. *Proceedings of the International Conference on Centrifuge '91*. Balkemma: Rotterdam, The Netherlands, 1991; 179–186.
16. Aiban SA, Znidarcic D. Centrifugal modeling of bearing capacity of shallow foundations on sand. *Journal of Geotechnical Engineering* 1995; **121**(7):704–712.
17. Padmini D, Ilamparuthi K, Sudheer KP. Ultimate bearing capacity prediction of shallow foundations on cohesionless soils using neurofuzzy models. *Computers and Geotechnics* 2008; **35**:33–46.
18. Ghaboussi J. Potential application of neuro-biological computational models in geotechnical engineering. In *Numerical Models in Geomechanics*, Pande GN, Pietruszezak S (eds). Balkemma: Rotterdam, The Netherlands, 1992; 543–555.
19. Goh ATC. Seismic liquefaction potential assessed by neural networks. *Journal of Geotechnical Engineering* 1994; **120**(6):1467–1480.
20. Goh ATC. Neural-netwok modeling of CPT seismic liquefaction data. *Journal of Geotechnical Engineering* 1996; **122**(1):70–73.
21. Mayoraz F, Vulliet L. Neural networks for slope movement prediction. *International Journal of Geomechanics* 2002; **2**:153–173.
22. Das SK, Basudhar PK. Undrained lateral load capacity of piles in clay using artificial neural network. *Computer and Geotechnics* 2006; **33**(5):454–459.
23. Samui P, Kumar B. Artificial neural network prediction of stability numbers for two-layered slopes with associated flow rule. *Electronics Journal of Geotechnical Engineering* 2006; **11**. Available from: http://www.ejge.com/2006/Ppr0626/Abs0626.htm.
24. Shahin MA, Jaksa MB, Maier HR. Artificial neural network applications in geotechnical engineering. *Australian Geomechanics* 2001; **36**(1):49–62.

DOI: 10.1002/nag

25. Park D, Rilett LR. Forecasting freeway link travel times with a multi-layer feed forward neural network. *Computer Aided Civil and Infrastructure Engineering* 1999; **14**:358–367.
26. Kecman V. Learning and soft computing: support vector machines. *Neural Networks and Fuzzy Logic Models*. MIT Press: Cambridge, MA, London, England, 2001.
27. Yan K, Shi C. Prediction of elastic modulus of normal and high strength concrete by support vector machine. *Construction and Building Materials* 2010; **8**:1479–1485.
28. Vapnik VN. *Statistical Learning Theory*. Wiley: New York, 1998.
29. Smola AJ, Scholkopf B. A tutorial on support vector regression. *Statistics and Computing* 2004; **14**:199–222.
30. Goh ATC, Goh SH. Support vector machines: their use in geotechnical engineering as illustrated using seismic liquefaction data. *Computers and Geotechnics* 2007; **34**(2):410–421.
31. Samui P, Kurup P, Sitharam TG. OCR prediction using support vector machine based on piezocone data. *Journal of Geotechnical and Geoenvironmental Engineering* 2008; **134**(3):894–898.
32. Samui P. Support vector machine applied to settlement of shallow foundations on cohesionless soils. *Computers and Goetechnics* 2008; **35**(3):419–427.
33. Tipping ME. The relevance vector machine. *Advances in Neural Information Processing Systems* 2000; **12**: 625–658.
34. Bishop C. *Neural Networks for Pattern Recognition*. Oxford University Press: Oxford, 1995.
35. Boser BE, Guyon IM, Vapnik VN. A training algorithm for optimal margin classifiers. In *Fifth Annual ACM Workshop on COLT*, Haussler D (ed.). ACM Press: Pittsburgh, PA, 1992; 144–152.
36. Cortes C, Vapnik VN. Support vector networks. *Machine Learning* 1995; **20**:273–297.
37. Gualtieri JA, Chettri SR, Cromp RF, Johnson LF. Support vector machine classifiers as applied to AVIRIS data. *Summaries of the Eighth JPL Airbrone Earth Science Workshop*, Jet Propulsion Laboratory, Pasadena, CA, 1999.
38. Bennett KP, Mangasarian OL. Robust linear programming discrimination of two linearly inseparable sets. *Optimization Methods and Software* 1992; **1**:23–34.
39. Cristianini N, Shawe-Taylor J. *An Introduction to Support Vector Machine*. Cambridge University Press: London, 2000.
40. Liong SY, Lim WH, Paudyal GN. River stage forecasting in bangladesh: neural network approach. *Journal of Computing in Civil Engineering* 2000; **14**(1):1–8.
41. Scholkopf B, Smola AJ. *Learning with Kernels*: *Support Vector Machines*, *Regularization*, *Optimization*, *and Beyond*. MIT Press: Cambridge, MA, 2002.
42. Tipping ME. Sparse Bayesian learning and the relevance vector machine. *Journal of Machine Learning Research* 2001; **1**:211–244.
43. Berger JO. *Statistical Decision Theory and Bayesian Analysis* (2nd edn). Springer: New York, 1985.
44. Wahba G. A comparison of GCV and GML for choosing the smoothing parameters in the generalized spline-smoothing problem. *Annals of Statistics* 1985; **4**:1378–1402.
45. MacKay DJ. Bayesian methods for adaptive models. *Ph.D. Thesis*, Department of Computation & Neural Systems, California Institute of Technology, Pasadena, CA, 1992.
46. Brinch Hansen J. A revised and extended formula for bearing capacity. *Danish Geotechnical Institute Bulletin No. 28*, 1970.

