Research Article

Three-Dimensional Site Characterization Model of Bangalore Using Support Vector Machine

Pijush Samui

Centre for Disaster Mitigation and Management, VIT University, Vellore 632014, India

Correspondence should be addressed to Pijush Samui, pijush.phd@gmail.com

The main objective of site characterization is the prediction of *in situ* soil properties at any half-space point at a site based on limited tests. In this study, the Support Vector Machine (SVM) has been used to develop a three dimensional site characterization model for Bangalore, India based on large amount of Standard Penetration Test. SVM is a novel type of learning machine based on statistical learning theory, uses regression technique by introducing ε-insensitive loss function. The database consists of 766 boreholes, with more than 2700 field SPT values (N) spread over 220 sq km area of Bangalore. The model is applied for corrected N (N_c) values. The three input variables (x, y, and z, where x, y, and z are the coordinates of the Bangalore) were used for the SVM model. The output of SVM was the N_c data. The results presented in this paper clearly highlight that the SVM is a robust tool for site characterization. In this study, a sensitivity analysis of SVM parameters (σ, C, and ε) has been also presented.

1. Introduction

In general, geotechnical engineers characterize a site based on a limited number of tests. They interpret a site in terms of working soil profiles, which are generally prepared based on soil properties. So, they have to predict in-situ soil properties at any half-space point at a site, based on a limited number of tests. The prediction of soil property is a difficult task due to uncertainty. Spatial variability, measurement "noise," measurement and model bias, and statistical error due to limited measurements are the sources of uncertainty [1]. Prediction of soil properties using geostatistics has been reported by many researchers [2–6]. However, several reasons appear to hinder the use of geostatistics in geotechnical engineering [7]. In probabilistic site characterization, random field theory has been used by many researchers in geotechnical engineering [8–19]. One of the most important assumptions of random field theory is that the soil property is to be statistically homogeneous with the chosen layer. In addition, these models have also assumed that soil property consists of a constant mean or a global mean trend with a stationary stochastic portion. To model the constant mean or global mean trend, regression analysis with polynomial functions has been used by the above researchers. Autocovariance function, autocorrelation function, autoregressive processes, power spectra functions, variance function, and scale of fluctuation are available for modelling the stationary stochastic portion. Statistically homogeneous soil layers have been determined by using "Modified Bartlett Statistics" [20]. However, random field methods and geostatistics have been applied in site characterization modelling with limited success [21]. Recently, artificial neural network has been used for site characterization [21]. A major disadvantage of ANN models is that there is no information about the relative importance of the various parameters [22]. In ANN, as the knowledge acquired during training is stored in an implicit manner, it is very difficult to come up with reasonable interpretation of the overall structure of the network [23]. This leads to the term "black box" which many researchers use while referring to ANN's behavior. In addition, ANN has some inherent drawbacks such as slow convergence speed, less generalizing performance, arriving at local minimum, and overfitting problems.

The support vector machine (SVM) based on statistical learning theory has been developed by Vapnik (1995) [24]. It provides a new, efficient novel approach to improve the

generalization performance and can attain a global minimum. In general, SVMs have been used for pattern recognition problems. Recently, it has been used to solve nonlinear regression estimation and time series prediction by introducing ε-insensitive loss function [25–27]. The SVM implements the structural risk minimization principle (SRMP), which has been shown to be superior to the more traditional empirical risk minimization principle (ERMP) employed by many of the other modelling techniques [28, 29]. SRMP minimizes an upper bound of the generalization error, whereas ERMP minimizes the training error. In this way, SVM produces the better generalization than traditional techniques.

The standard penetration test (SPT) is a well-established and unsophisticated method of soil test, which was developed in the United States around 1925. It has been established as the most popular field testing method to characterize the subsurface soil profiles, despite its limitations. Field SPT (N) value is used to determine the bearing capacity, settlement, and liquefaction potential, and it is also correlated to many soil properties such as shear wave velocity, angle of internal friction, and cone tip resistance. The objective of this paper is to use SVM for three-dimensional (3D) site characterization model for Bangalore, India based on a large amount N_c values in this area. Further, sensitivity analysis of SVM parameters (σ, C, and ε) has been carried out and results are presented to highlight their influence on the results.

2. Site Description

The city of Bangalore covers an area of over 220 square kilometres, and ground reduced levels (GRLs) also vary a lot in the city. It varies from 810 m in the north-eastern part to 940 m in the south-western part of Bangalore. Ground reduced levels do not vary much in the other parts of the city. There were more than 450 lakes once upon a time, and more than 340 lakes have dried up due to erosion and encroachments for construction of layouts and buildings. The population of greater Bangalore region is over 6 million, and it is the fifth biggest city in India. It is growing very fast and is situated on latitude of $12°8'$ north and longitude of $77°37'$ east.

From geology, the most part of Bangalore falls in gneiss complexes, which are formed due to several tectonic-thermal events with large influx of sialic material and are believed to have occurred between 3400 and 3000 million years ago giving rise to an extensive group of gray gneisses designated as the "older gneiss complex." These gneisses act as the basement for a widespread belt of schist's. The younger group of gneissic rocks mostly of granodiorite and granite composition is found in the eastern part of the state, representing remobilized parts of an older crust with abundant additions of newer granite material, for which the name "younger gneiss complex" has been given [31]. The soil is mostly a residual soil from granite gneiss due to weathering action. In the old tank beds, silty sand/clay is also found as overburden.
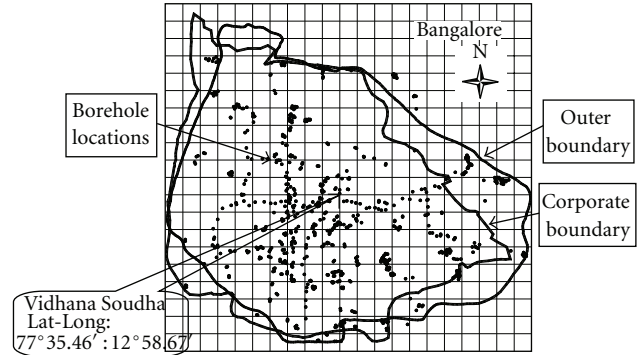

FIGURE 1: Borehole location in Bangalore map (scale: 1 : 20000).

3. Geographic Information System (GIS) Model and Geotechnical Data

The Bangalore map forms the base layer for the development of GIS model (see Figure 1). The map entities have been developed in view of two aspects, firstly for locating the borelogs to the utmost accuracy on a scale of 1 : 20000 and secondly for identification of borelogs by end user. The digitized map has several layers of information. Some of the important layers considered are the boundaries (outer and Administrative), highways, major roads, minor roads, streets, rail roads, water bodies, drains, ground contours, and borehole locations. A large amount of geotechnical data consisting of 766 boreholes has been collated along with index and engineering properties of subsoil layers at different locations in Bangalore (location of boreholes is shown in Figure 1). Geotechnical data were evaluated for geotechnical investigations of several major projects in Bangalore. In total, 766 borelogs information has been entered into the database using a GIS with ARCINFO package. The latitudes and longitudes were confirmed using global positioning system (GPS) stations at selected locations. In total, 2722 "N" values are available in 766 boreholes in the three-dimensional GIS model. Distribution of collected boreholes in Bangalore is shown in Figure 2, indicating a very good distribution of the boreholes in each quadrant of Bangalore from the city center. Figure 1 depicts a grid of 1 km × 1 km within the corporate boundary of Bangalore along with outer boundary circumscribing the ring road also with location of boreholes. It gives a clear view of the spatial distribution of boreholes in Bangalore region. An average of about four boreholes data is available within the grid of 1 km × 1 km.

Geotechnical data was collated from archives of Torsteel Research Foundation in India and Indian Institute of Science for geotechnical investigation carried out for several major projects in Bangalore. The data collected are of very high quality for important projects in Bangalore during the years 1995–2003. The data in the model are on average to a depth of 30 m below the ground level. The borelogs contain information about depth, density of the soil, total stress, effective stress, fines content, and N values and depth of ground water table. For the purpose of general identification of soil layers, the Bangalore map area is divided into
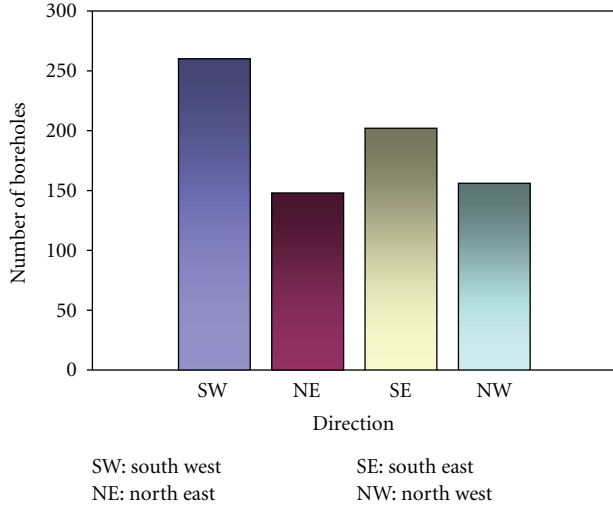
FIGURE 2: Distribution of boreholes in quadrants for Bangalore.

SW: south west SE: south east
NE: north east NW: north west

four parts (four quadrants) in north-south and east-west directions as shown in Figure 2. The typical soil profile in the north-western part of the Bangalore has three layers of soil deposition. The first layer contains brownish silty sand with clay or red soil in some location up to 3 m, after which up to 6 m, medium dense to very dense silty sand is present. The third layer has weathered rock varying from 6 m to 17 m depth and followed by hard rock. The south-western part contains red soil or reddish silty sand with gravel up to 1.7 m depth, yellowish clayey sand from 1.7 m to 3.5 m, yellowish silty sand with clay from 3.5 m to 8.5 m, and hard rock below 8.5 m. The soil in the south-eastern part can be classified into 4 layers. The first layer up to 1.5 m contains brownish clayey sand, brownish clayey sand with gravel from 1.5 m to 4 m, yellowish silty sand with gravel up to 5.5 m, different stages of weathered rock from 5.5 m to 17.5 m, and hard rock beneath. North-eastern side has 4 layer depositions, filled up soil to 1.5 m, reddish silty clay from 1.5 m to 4.5 m, sandy clay from up to 7.5 m, weathered rock form 7.5 m to 18.5 m, and hard rock below. The corrections for field N values (shown in Tables 1 and 2) are applied for overburden pressures (C_N), hammer energy (C_E), borehole diameter (C_B), presence or absence of liner (C_S), rod length (C_R), and correction for fines content (C_{fines}) as per standard procedures existing in literature [32–37].

4. Support Vector Machine Model

SVM has originated from the concept of statistical learning theory pioneered by Boser et al. (1992) [38]. In this section, a brief introduction is presented on the construction process of SVM for regression problems. There are three distinct characteristics of SVM when they are used to estimate the regression function. First of all, SVM estimates the regression using a set of linear functions that are defined in a high-dimensional space. Secondly, SVM carries out the regression estimation by risk minimization where the risk is measured using Vapnik's ε-insensitive loss function. Thirdly, SVM uses

TABLE 1: Different types of SPT corrections used to obtain N_c values.

Different types of correction	Correction factor
Correction for hammer	
Donut hammer	0.5–1.0
Safety hammer	0.7–1.2
Automatic-trip donut hammer	0.8–1.3
Sampler correction	
Without liner	1.00
With liner: dense sand, clay	0.80
Loose sand	0.90
Rod length correction	
Length > 10 m	1.0
6–10	0.95
4–6	0.85
0–4	0.75
Borehole diameter correction	
Hole diameter 60–120 mm	1.00
150 mm	1.05
200 mm	1.15
Correction for overburden pressure, σ'_{v0} (C_N)	$2.2/(1.2 + \sigma'_{v0}/P_a)$, where $P_a = 100$ kPa
Correction for fines content (C_{fines})	$1 + 0.004\mathrm{FC} + 0.055(\mathrm{FC}/N_{60})$, where FC = percent fines content (percent dry weight finer than 0.074 mm) and N_{60} = SPT value for 60% energy ratio

a risk function consisting of the empirical error and a regularization term which is derived from the SRMP. This study uses the SVM as a regression technique by introducing an ε-insensitive loss function. The ε-insensitive loss function ($L_\varepsilon(y)$) can be described in the following way:

$$L_\varepsilon(y) = \begin{cases} 0, & \text{for } |f(x) - y| < \varepsilon, \\ |f(x) - y| - \varepsilon, & \text{otherwise.} \end{cases} \tag{1}$$

This defines an ε tube (Figure 3) so that if the predicted value is within the tube, the loss is zero, while if the predicted point is outside the tube, the loss is the magnitude of the difference between the predicted value and the radius, ε, of the tube. Assume that the training dataset consists of l training sample $\{(x_1, y_1), \ldots (x_l, y_l)\}$ where x is the input and y is the output. For site characterization model for Bangalore, $x = [x, y, z]$ and $y = [N_c]$.

The main aim in 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

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

where w = an adjustable weight vector, b = the scalar threshold, R^n = n-dimensional vector space, and r = one-dimensional vector space.

Flatness in the case of (2) means that one seeks a small w. One way of obtaining this is by minimizing the Euclidean

TABLE 2: Typical calculation showing N_c values obtained from N values.

| Depth (m) | SPT value (N_{field}) | Unit weight (kN/m³) | Total stress (T.S. (kN/m²)) | Effective stress (kN/m²) | C_N | Correction factor | | | Sample method | F.C. | C_{fines} | SPT value ($N_{corrected}$) |
						Hammer effect	Boreholedia	Rod length				
1.50	11	19.10	28.65	28.65	1.48	1	1.05	0.75	1	54.5	1.431	18
4.00	12	19.10	76.40	76.40	1.12	1	1.05	0.85	1	59	1.482	18
6.00	30	19.10	114.60	114.60	0.94	1	1.05	0.95	1	53.6	1.309	37
7.50	74	19.10	143.25	143.25	0.84	1	1.05	0.95	1	29.8	1.143	71
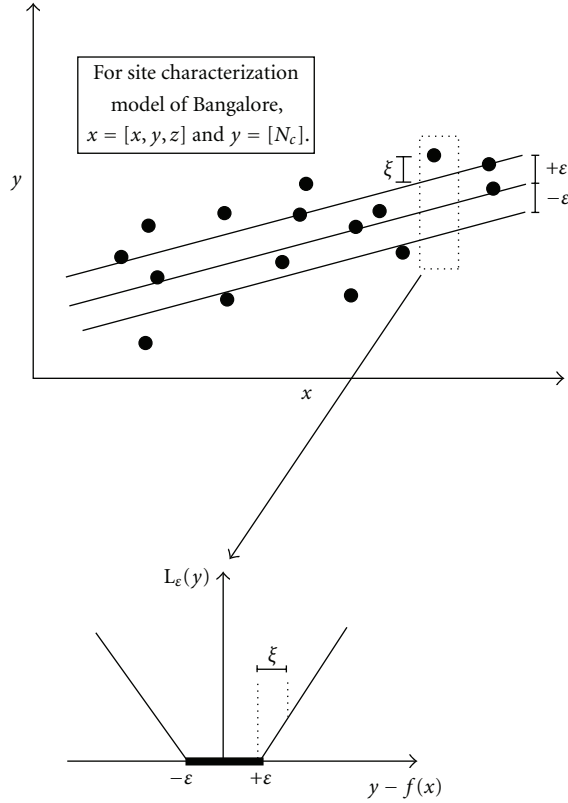
FIGURE 3: Prespecified accuracy ε and slack variable ξ in support vector regression [30].

FIGURE 4: Concept of nonlinear regression.

norm $\|w\|^2$. This is equivalent to the following convex optimization problem:

$$\text{minimize: } \frac{1}{2}\|w\|^2$$
$$\text{subjected to: } y_i - (\langle w \cdot x_i \rangle + b) \leq \varepsilon, \quad i = 1, 2, \ldots, l \qquad (3)$$
$$(\langle w \cdot x_i \rangle + b) - y_i \leq \varepsilon, \quad i = 1, 2, \ldots, l.$$

The above convex optimization problem is feasible. Sometimes, however, this may not be the case, or we also may want to allow for some errors, analogously to the "soft margin" loss function [39] which was used in SVM by Cortes and Vapnik (1995) [40]. As shown in Figure 1, the parameters ξ_i, ξ_i^* are slack variables that determine the degree to which samples with error more than ε are 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)

$$\text{minimize: } \frac{1}{2}\|w\|^2 + C \sum_{i=1}^{l} (\xi_i + \xi_i^*)$$
$$\text{subjected to: } y_i - (\langle w \cdot x_i \rangle + b) \leq \varepsilon + \xi_i, \quad i = 1, 2, \ldots, l$$
$$(\langle w \cdot x_i \rangle + b) - y_i \leq \varepsilon + \xi_i^*, \quad i = 1, 2, \ldots, l$$
$$\xi_i \geq 0, \quad \xi_i^* \geq 0, \quad i = 1, 2, \ldots, l. \qquad (4)$$

The constant $0 < C < \infty$ determines the trade-off between the flatness of f and the amount up to which deviations larger than ε are tolerated [41]. This optimization problem (4) is solved by Lagrangian multipliers [42], and its solution is given by

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

where $b = -(1/2)w \cdot [x_r + x_s]$, α_i, α_i^* are the Lagrangian multipliers, and nsv is the number of support vectors. 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 has to be mapped into a high-dimensional feature space through some nonlinear mapping [38] (see Figure 4). 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
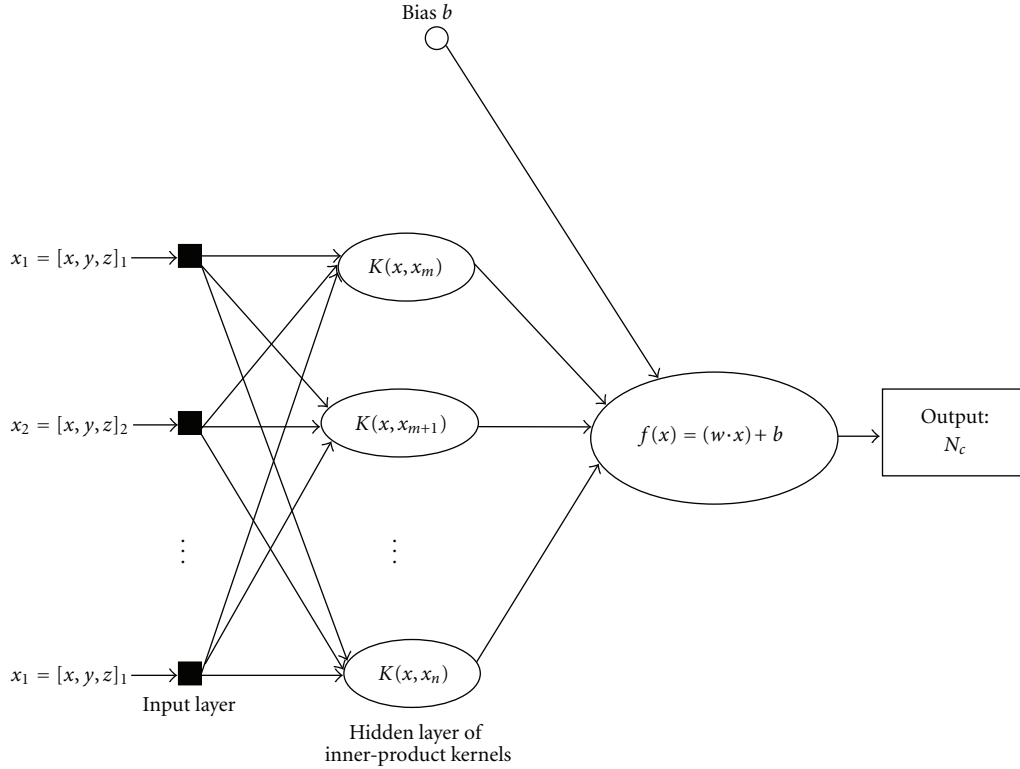
FIGURE 5: Architecture of support vector machine for N_c prediction.

input data is mapped onto the feature space by a map Φ (see Figure 4). 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 $[K(x_i, x_j) = \Phi(x_i) \cdot \Phi(x_j)]$ has been introduced to reduce the computational demand [40, 43]. So, (5) becomes written as

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

In this study, radial basis function has been used as a kernel function.

5. SVM Implementation for Site Characterization Model

Figure 5 shows the architecture of SVM for N_c prediction in 3D subsurface of Bangalore. In SVM, each of the input variables (x, y, and z) is first normalized with respect to their respective maximum value. The output variable N_c was also normalized with respect to the maximum N_c value. For implementing the SVM, the data has been divided into two subsets:

(1) a training dataset: this is required to train the model. In this study, 90% of total boreholes (number of total boreholes = 766, number of N_c values = 2429, and 90% of total boreholes = 689.4≈690) are considered for training dataset.

(2) a testing dataset: this is required to examine the model performance. In this study, the remaining 10% of the total boreholes is considered as testing dataset, which consists of 76 boreholes of 293 N_c data.

The training and testing datasets have been chosen using sorting method to maintain statistical consistency. The application of SVM for this study requires the proper selection of design parameters (C and ε). The identification of optimal values of C and ε is largely a trial and error process. However, there are guidelines that can be used for selecting these parameters. A large C assigns higher penalties to errors so that the regression is trained to minimize error with lower generalization, while a small C assigns fewer penalties to errors; this allows the minimization of margin with errors, thus higher generalization ability. If C goes to be infinitely large, SVM would not allow the occurrence of any error and result in a complex model, whereas when C goes to zero, the result would tolerate a large amount of errors, and the model would be less complex. With regards to the selection of ε, if ε is too large, too few support vectors are selected which leads to a decrease of the final prediction performance. If ε is too small, many support vectors are selected which leads to the risk of overfitting. The optimum values of C and ε obtained in this study are presented in Section 6. The program of SVM is constructed using MATLAB.
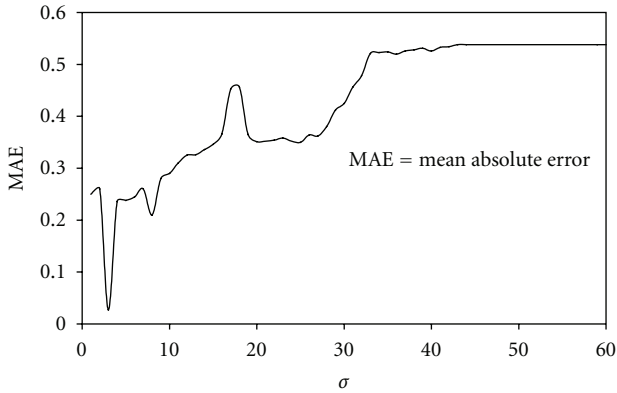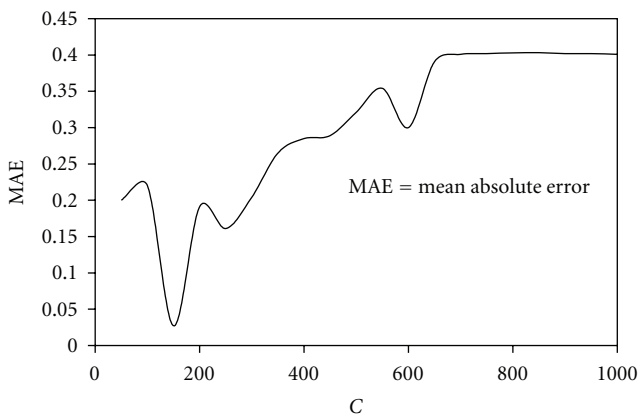
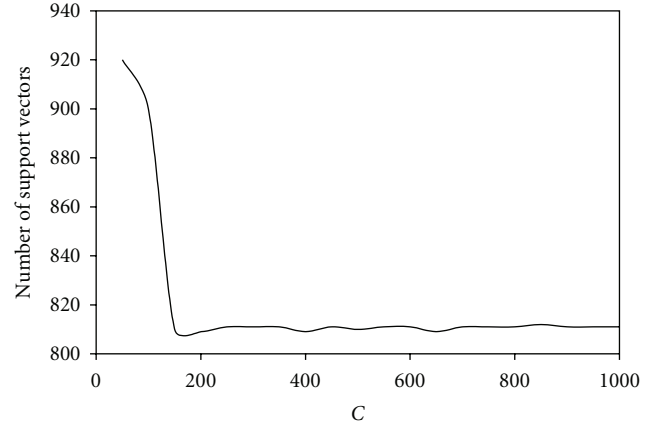Figure 6: Variation of σ with MAE.

Figure 8: Variation of number of support vectors with C values.

Figure 7: Variation of MAE with C values.
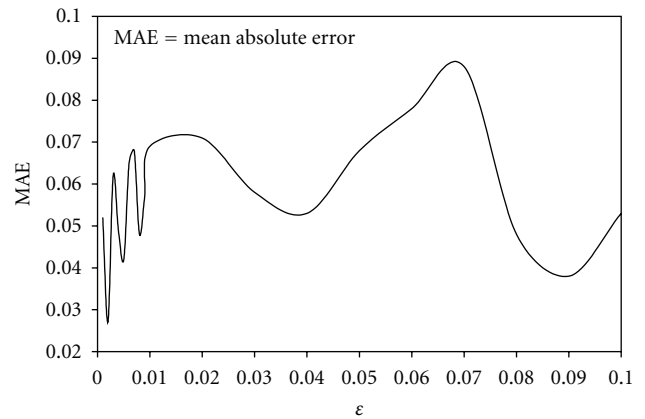
Figure 9: Variation of MAE with ε.

6. Result and Discussion

In this analysis as a first step, the free parameters of Gaussian kernel function σ, C, and ε have been chosen arbitrarily. So it is necessary to investigate the impact of these free parameters on the generalization error and number of support vectors. Firstly, the influence of σ on the prediction performance is studied. It is known to us that the level of predicting accuracy is greatly influenced by the value of σ. Using too small σ (i.e., $\sigma \rightarrow 0$) or too large σ (i.e., $\sigma \rightarrow \infty$) will be not well suited for good model. Figure 6 represents the impacts of σ on the testing results. The mean absolute error (MAE) (MAE $= (1/n) \sum_{i=1}^{n} |a_i - p_i|$, where a_i is the actual data, p_i is the predicted data, and n is the number of data) achieves minimum value of 0.0271 at $\sigma = 3$ for N_c values. It can be seen from Figure 6 that the MAE values change sharply when $\sigma <$ 40 and tend to flatten after $\sigma \geq 40$. In this study, a σ value 3 has been used for N_c. Figure 7 shows the variation between the MAE and the C values. The MAE has a minimum value of 0.0271 at $C = 150$ for the N_c values. Figure 8 shows the variation of a number of support vectors with the C values. It can be seen from Figure 8 that the number of support vector values changes sharply when $C < 150$ and tends to flatten after ≥ 150. In order to make the learning process robust, C has been assigned a value of 150. Figure 9 depicts

the variation MAE value with ε values. The MAE has a minimum value at $\varepsilon = 0.002$. Figure 10 shows the relation between the number of support vectors and the ε values. It is also found that the number of support vectors is decreasing with increasing ε. In general, ε should be set at small value, specified as $\varepsilon = 0.002$ in this analysis. To produce the best possible result, the σ value should be 3. The SVM was found to generalize well by setting the capacity factor C as 150 and ε value as 0.002. Figure 11 represents the performance of SVM model for training dataset (coefficient of correlation, $R = 0.994$), and the results are almost identical to the original data. In order to evaluate the capabilities of the SVM model, the model is validated with new N_c data that are not part of the training dataset. Figure 12 shows the performance of the SVM model for testing dataset ($R = 0.986$). From Figure 12, it is clear that the SVM model has predicted the actual values of N_c very well, and it can be used for 3D site characterization model of Bangalore. Figures 13 and 14 show the N_c values with depth corresponding to borehole nos. BH 176-2 and BH 276-2, respectively. From Figures 13 and 14, it is clear that the predicted values match very well with the actual values of N_c. Figures 15 and 16 shows three-dimensional and two-dimensional surface of N_c using SVM model, respectively.
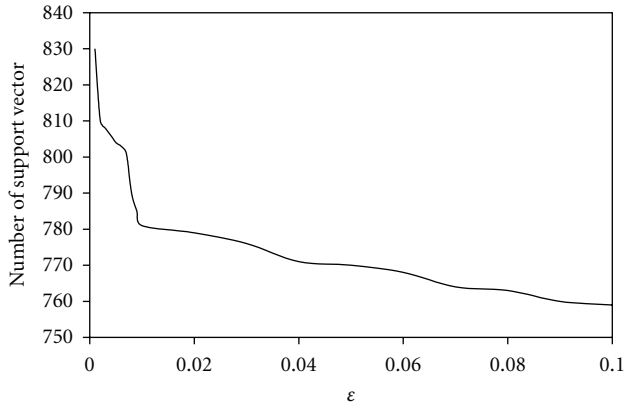
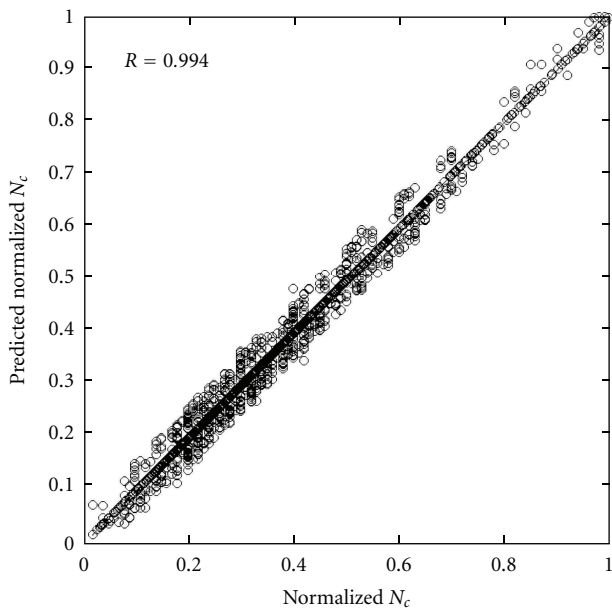Figure 10: Variation of number of support vectors with ε.

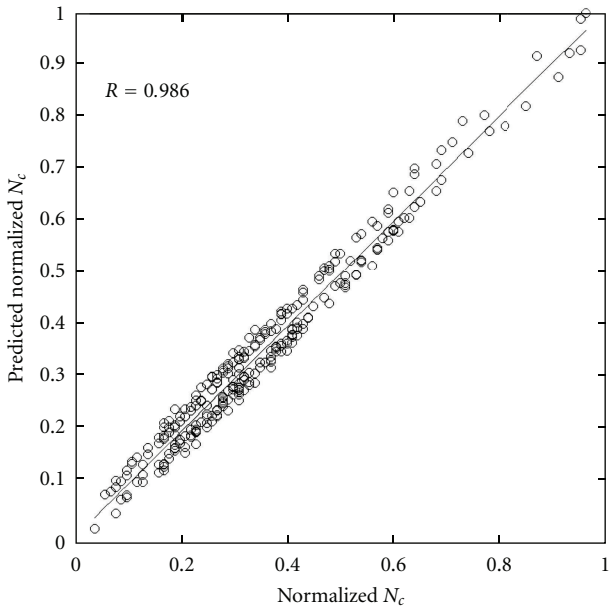Figure 11: Performance of SVM model for training dataset.

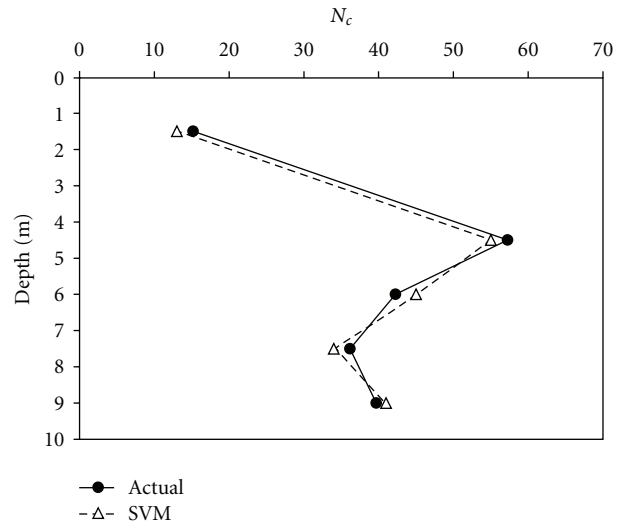Figure 12: Performance of SVM model for testing dataset.

Figure 13: Comparison of N_c with predicted values at BH 176-2 by SVM.
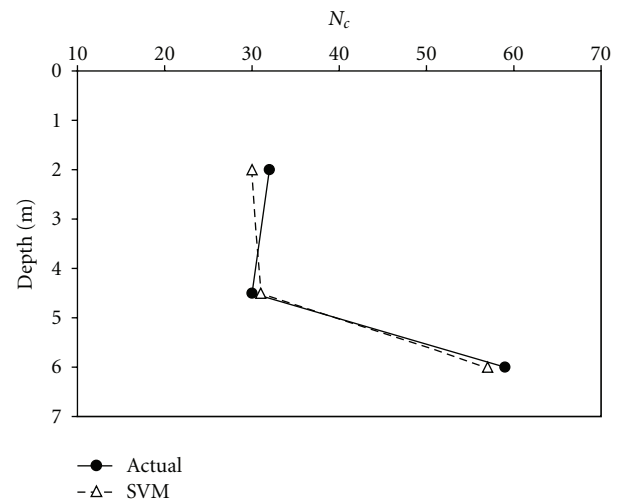

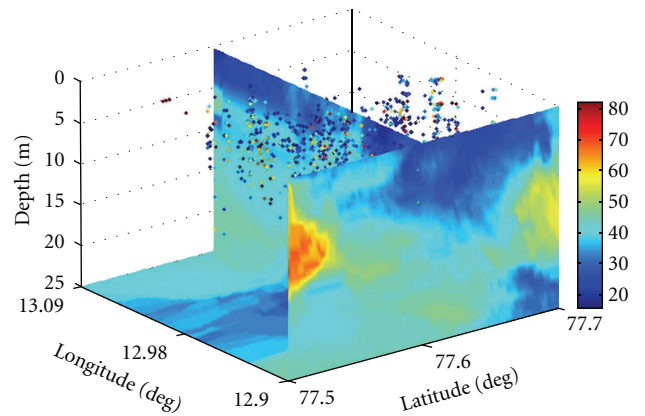Figure 14: Comparison of N_c with predicted values at BH 276-2 by SVM.

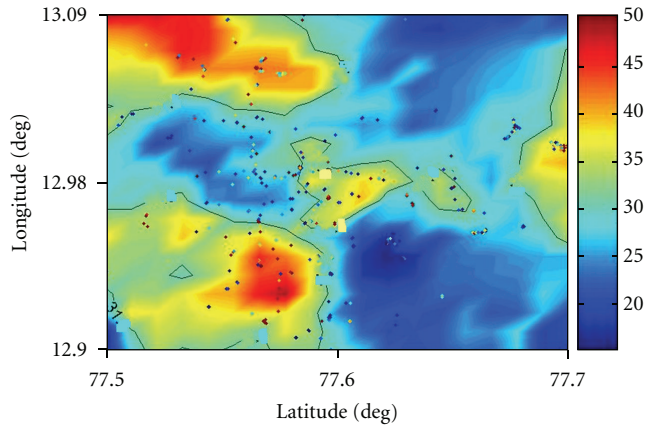Figure 15: Three-dimensional surface of N_c using SVM model.

Figure 16: Two-dimensional surface of N_c on the plane of $z = 1.5$ m using SVR model.

7. Conclusions

The three-dimensional site characterization model has been developed for Bangalore using SVM technique. SVM technique has shown to be a promising tool for site characterization. SVM training consists of solving a-uniquely solvable-quadratic optimization problem and always finds a global minimum. In this study, C and ε factors are considered in SVM method by using a kernel function. A detailed parametric analysis of these parameters on the predictive performance has been carried out. The SVM was found to generalize well by setting the capacity factor C as 150 and ε value as 0.002. The result obtained shows that the SVM model is accurate in predicting N_c values. In general, SVM is shown to provide a general site characterization model of Bangalore. This has a potential for seismic hazard analysis, site response, and liquefaction studies for the development of microzonation maps for an area. The predicted N_c values from the developed model can also be used to estimate the subsurface information, allowable bearing pressure of soils, and elastic modulus of soils.

Acknowledgment

References

[1] G. B. Baecher, "Geotechnical error analysis," *Transportation Research Record*, vol. 1105, pp. 23–31, 1986.

[2] J. P. Delhomme, "Spatial variability and uncertainty in groundwater flow parameters: a geostatistical approach," *Water Resources Research*, vol. 15, no. 2, pp. 269–280, 1979.

[3] P. H. S. W. Kulatilake, "Probabilistic Potentiometric surface mapping," *Journal of Geotechnical & Geoenvironmental Engineering*, vol. 115, no. 11, pp. 1569–1587, 1989.

[4] M. Soulie, P. Montes, and V. Silvestri, "Modelling spatial variability of soil parameters," *Canadian Geotechnical Journal*, vol. 27, no. 5, pp. 617–530, 1990.

[5] P. Chiasson, J. Lafleur, M. Soulie, and K. T. Law, "Characterizing spatial variability of a clay by geostatistics," *Canadian Geotechnical Journal*, vol. 32, no. 1, pp. 1–10, 1995.

[6] D. J. DeGroot, "Analyzing spatial variability of in situ soil properties," in *Proceedings of the Conference on Uncertainty in the Geologic Environment (UNCERTAINTY '96)*, vol. 85, pp. 210–238, 1996.

[7] R. B. Kulkarni, "Bayesian kriging in geotechnical problems," in *Geostatistics for Natural Resources Characterization, Part 2*, NATO ASI Series, pp. 775–786, Reidel, Dordrecht, The Netherlands, 1983.

[8] A. M. Yaglom, *Theory of Stationary Random Functions*, Prentice-Hall, Englewood Cliffs, NJ, USA, 1962.

[9] P. Lumb, "Spatial variability of soil properties," in *Proceedings of the 2nd International Conference on Applications of Statistics and Probability in Civil Engineering*, pp. 397–421, Aachen, Germany, 1975.

[10] E. H. Vanmarcke, "Probabilistic modeling of soil profiles," *Journal of Geotechnical and Geoenvironmental Engineering*, vol. 103, no. 11, pp. 1227–1246, 1977.

[11] W. H. Tang, "Probabilistic evaluation of penetration resistances," *Journal of the Geotechnical Engineering Division*, vol. 105, no. 14902, pp. 1173–1191, 1979.

[12] T. H. Wu and K. Wong, "Probabilistic soil exploration: case history," *Journal of the Geotechnical Engineering Division*, vol. 107, no. 16764, pp. 1693–1711, 1981.

[13] A. Asaoka and D. Athanasiou-Grivas, "Spatial variability of the undrained strength of clays," *Journal of the Geotechnical Engineering Division*, vol. 108, no. 5, pp. 743–756, 1982.

[14] E. H. Vanmarcke, *Random Fields: Analysis and Synthesis*, The MIT Press, Cambridge, Mass, USA, 1983.

[15] G. B. Baecher, "On estimating auto-covariance of soil properties," in *Proceedings of the 4th ASCE Joint Specialty Conference on Probabilistic Mechanics and Structural Reliability*, vol. 110, pp. 214–218.

[16] H. S. W. Kulatilake and K. M. Miller, "A scheme for estimating the spatial variation of soil properties in three dimensions," in *Proceedings of the 5th International Conference on Application of Statistics and Probabilities in Soil and Structural Engineering*, pp. 667–677, Vancouver, British Columbia, Canada, 1987.

[17] G. A. Fenton, "Random field characterization NGES data," in *Proceedings of the Workshop on Probabilistic Site Characterization at NGES*, Seattle, Wash, USA, 1998.

[18] K. K. Phoon and F. H. Kulhawy, "Characterization of geotechnical variability," *Canadian Geotechnical Journal*, vol. 36, no. 4, pp. 612–624, 1999.

[19] M. Uzielli, G. Vannucchi, and K. K. Phoon, "Random filed chracterisation of strees-normalised cone penetration testing parameters," *Geotechnique*, vol. 55, no. 1, pp. 3–20, 2005.

[20] K. K. Phoon, S. T. Quek, and P. An, "Identification of statistically homogeneous soil layers using modified bartlett statistics," *Journal of Geotechnical and Geoenvironmental Engineering*, vol. 129, no. 7, pp. 649–659, 2003.

[21] C. H. Juang, T. Jiang, and R. A. Christopher, "Three-dimensional site characterisation: neural network approach," *Geotechnique*, vol. 51, no. 9, pp. 799–809, 2001.

[22] D. Park and L. R. Rilett, "Forecasting freeway link ravel times with a multi-layer feed forward neural network," *Computer Aided Civil and Infrastructure Engineering*, vol. 14, pp. 358–367, 1999.

[23] V. Kecman, *Leaming and Soft Computing: Support Vector Machines, Neural Networks, and Fuzzy Logic Models*, The MIT Press, Cambridge, Mass, USA, 2001.

[24] V. Vapnik, *The Nature of Statistical Learning Theory*, Springer, New York, NY, USA, 1995.

[25] S. Mukherjee, E. Osuna, and F. Girosi, "Nonlinear prediction of chaotic time series using support vector machine," in *Proceedings of the 7th IEEE Signal Processing Society Workshop*, pp. 511–519, IEEE, New York, NY, USA, 1997.

[26] K. R. Muller, A. Smola, G. Ratsch, B. Scholkopf, J. Kohlmorgen, and V. Vapnik, "Predicting time series with support vector machines," in *Proceedings of the International Conference on Artificial Neural Networks (ICANN '97)*, p. 999, Springer, Berlin, Germany, 1997.

[27] V. Vapnik, S. Golowich, and A. Smola, "Support method for function approximation regression estimation and signal processing," in *Advances in Neural Information Processing Systems*, M. Mozer and T. Petsch, Eds., vol. 9, MIT Press, Cambridge, Mass, USA, 1997.

[28] E. Osuna, R. Freund, and F. Girosi, "An improved training algorithm for support vector machines," in *Proceedings of the 7th IEEE Workshop on Neural Networks for Signal Processing*, pp. 276–285, IEEE, New York, NY, USA, 1997.

[29] S. Gunn, "Support vector machines for classification and regression," Tech. Rep., University of Southampton, Southampton, UK, 1998, Image Speech and Intelligent Systems Research Group.

[30] B. Scholkopf, *Support Vector Learning*, R. Oldenbourg, Munich, Germany, 1997.

[31] B. P. Radhakrishna and R. Vaidyanadhan, *Geology of Karnataka*, Geological Society of India, Bangalore, India, 1997.

[32] H. B. Seed, K. Tokimatsu, L. F. Harder, and R. M. Chung, "Influence of SPT Procedures in soil liquefaction resistance evaluation," *Journal of Geotechnical Engineering*, vol. 111, no. 12, pp. 1425–1445, 1985.

[33] A. W. Skempton, "Standard penetration test procedures," *Geotechnique*, vol. 36, no. 3, pp. 425–557, 1986.

[34] C. O. Riggs, "American standard penetration test practice," in *Proceedings of the 14th PSC ASCE Insitu Tests in Geotechnical Engineering*, vol. 124, pp. 949–967.

[35] P. K. Robertson and C. E. Wride, "Cyclic liquefaction and its evaluation based on the SPT and CPT," in *Proceedings of the NCEER Workshop on Evaluation of Liquefaction Resistance of Soils*, 1998, Directed by T. L. Youd and I. M. Idriss.

[36] J. H. Schmertmann, "Statics of SPT," *Journal Geotechnical Engineering Division*, vol. 105, no. 14573, pp. 655–670, 1979.

[37] L. Finn and C. Ventura, "Challenging issues in local microzonation," in *Proceedings of the 5th International Conference on Seismic Zonation*, pp. 1554–1561, Nice, France, 1995.

[38] B. E. Boser, I. M. Guyon, and V. N. Vapnik, "A training algorithm for optimal margin classifiers," in *Proceedings of the 5th Annual Workshop on Computational Learning Theory*, D. Haussler, Ed., pp. 144–152, ACM Press, Pittsburgh, Pa, USA, 1992.

[39] K. P. Bennett and O. L. Mangasarian, "Robust linear programming discrimination of two linearly inseparable sets," *Optimization Methods and Software*, vol. 1, no. 1, pp. 23–34, 1992.

[40] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995.

[41] A. J. Smola and B. Schölkopf, "A tutorial on support vector regression," *Statistics and Computing*, vol. 14, no. 3, pp. 199–222, 2004.

[42] V. Vapnik, *Statistical Learning Theory*, John Wiley & Sons, New York, NY, USA, 1998.

[43] N. Cristianini and J. Shawe-Taylor, *An Introduction to Support Vector Machine*, Cambridge University Press, London, UK, 2000.