SITE CHARACTERIZATION OF (IGCAR) KALPAKKAM USING SOFT COMPUTING TECHNICES

Pijush Samui[1], Edison. T[2], Harikumar. C[3] and Sivathanu Pillai. C[4]

[1]Director In Charge, Centre for Disaster Mitigation and Management, VIT University, Vellore 632014, India. Email: pijush.phd@gmail.com
[2]Junior Research Fellow, Centre for Disaster Mitigation and Management, VIT University, Vellore 632014, India. Email: edison.gis@gmail.com
[3]Scientist, Civil Engineering Group, IGCAR, India
Email: chk@igcar.gov.in
[4]Associate Director, Civil Engineering Group,IGCAR, India
Email: csp@igcar.gov.in

Abstract

This study employs Artificial Neural Network (ANN) and Support Vector Machine (SVM) for developing site characterization model of Indira Gandhi Centre for Atomic Research (IGCAR) (Kalpakkam) based on Standard Penetration Test (SPT) value (N). The input variables of ANN and SVM are latitude, longitude and depth of the boreholes. N is the output of the ANN and SVM models. The developed SVM gives an equation for prediction of N value at any point in IGCAR. The performance of ANN and SVM is comparable. The results show that the developed ANN and SVM is reliable model for prediction of N value at any point in IGCAR.

Keywords: *Artificial Neural Network, Support Vector Machine, Site Characterization, SPT.*

1   Introduction

Site characterization is an imperative task in every geotechnical engineering project. Geotechnical engineers use laboratory and in situ tests for site characterization. There are different type of in situ tests such as cone penetration test (CPT), standard penetration test (SPT), dilatometer test (DMT), pressuremeter test (PMT). Geophysical tests such as spectral analysis of surface waves (SASW) and multi channel analysis of surface waves (MASW) are also becoming popular

as they are inexpensive and non-destructive tests. It is never possible to know the detailed geotechnical properties at every location beneath an actual site because, in order to do so, one would need to sample and/or test the entire subsurface profile. So, one has to predict geotechnical properties at any point of a site based on a limited number of tests. Prediction of geotechnical properties of a site is a difficult task due to uncertainty. Based on finite set of in situ data, in probabilistic site characterization, random field theory has been used by many researchers in geotechnical engineering (Yaglom, 1962; Lumb, 1975; Vanmarcke, 1977; Tang, 1979; Wu and Wong, 1981; Asaoka and Grivas, 1982; Vanmarcke, 1983; Baecher, 1984; Kulatilake and Miller, 1987; Kulatilake, 1989; Fenton, 1998; Phoon and Kulhawy, 1999; Uzielli et al., 2005). The science of prediction in the presence of correlation between samples is not at all well developed in a random field method. The interpretation of trends from random field method in the data as true trends in the mean or simply as large scale fluctuations is a question. This question can only be answered by engineering judgment. Therefore, the statistical parameters which have been used to model a random field are generally uncertain.

This study adopts Artificial Neural Network (ANN) and Support Vector Machine (SVM) for developing site characterization models of Indira Gandhi Centre for Atomic Research (IGCAR)(Kalpakkam) based on SPT value (N). ANN has been used to solve the different problems in engineering (Mutlu et al., 2008; Sivasankari and Thanushkodi, 2009; Leal et al., 2009; Naseri and Elliott, 2010; Osman et al.,2010; Mardi et al., 2011; Kuok et al., 2011; Nurmaini et al., 2009; Buratti et al., 2012). Researchers have solved the different problems by using SVM (Lingling and Kuihe, 2008; Zhao and Ding, 2009; Du et al., 2010; Ludwig et al, 2011; Segata et al, 2012). This article has the following aims:

•   To examine the capability of SVM and ANN for developing site characterization models of IGCAR based on N values

•   To develop an equation for predicting N values at any point in IGCAR based on the SVM

•   To make a comparative study between the ANN and SVM models

2   DATASET

This article uses 95 N values for developing the ANN and SVM models. The latitude and longitude of boreholes have been determined by using Global

Positioning System (GPS). Figure 1 shows the borehole locations. Table 1 shows the different statistical parameters of the dataset.
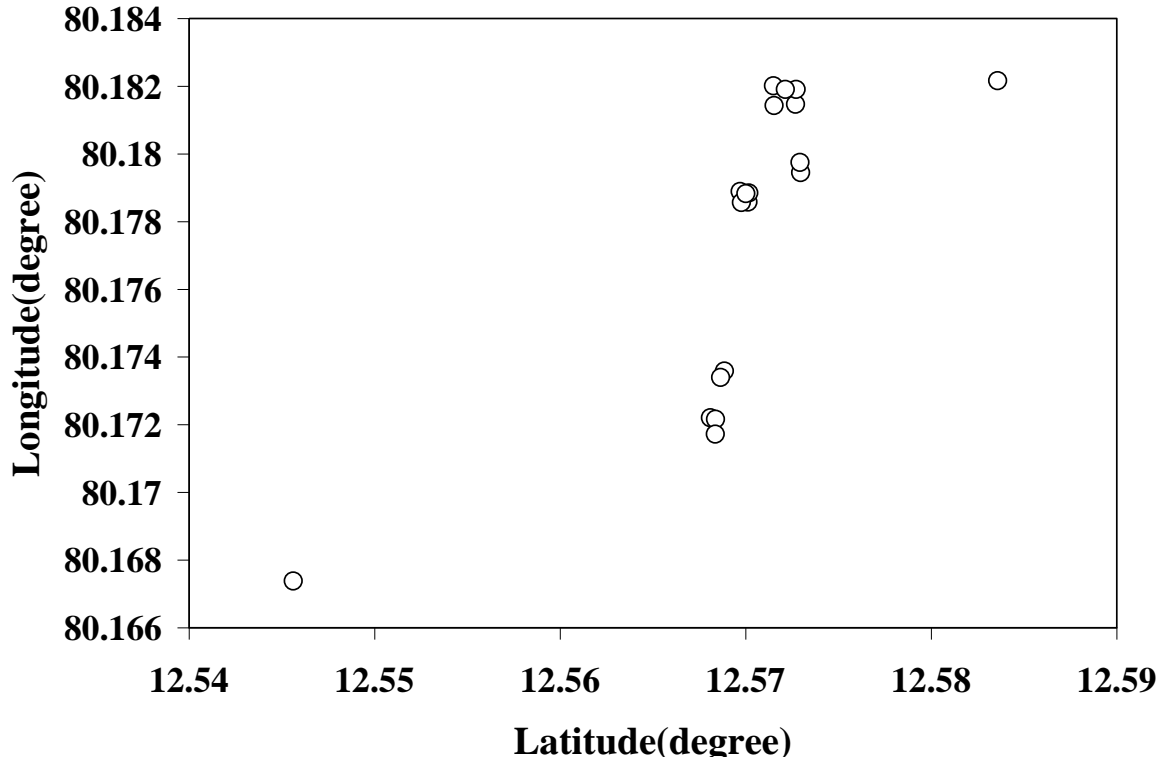

Figure 1. Locations of the boreholes.

Table 1. Statistical parameters of the dataset.

Mean	Standard Deviation	Skewness	Kurtosis
69.75	36.36	-0.58	1.62

3   DETAILS OF ANN

ANN has been used with multi-layer perceptrons (MLPs) that are trained with Levenberg-Marquardt Backpropagation(BP) algorithm. MLPs are perhaps the best-known type of feed forward networks (Hagan and Menhaj, 1994). It has generally three layers: an input layer, an output layer and an intermediate or hidden layer. MLPs are perhaps the best-known type of feed forward networks.

MLP has generally three layers: an input layer, an output layer and an intermediate or hidden layer.

In back propagation training process, the network error is back propagated into each neuron in the hidden layer, and then continued into the neuron in the input layer. The modification of connection weights and biases depend on the distribution of error at each neuron. The global network is reduced by continuous modifications of connection weights and biases. An error goal is set before the network training, and if the network during the training becomes less than the error goal, the training has to be stopped. The theory and implementation of Levenberg-Marquardt Backpropagation has given by More (1977). In this study, N is considered in the site characterization. The problem here is to learn characteristic of the site using the measured N data. In 3D analysis, the function is to be approximated with which N value at any half space point in IGCAR can be determined.

To develop the ANN mode, the datasets have been divided into the following two groups:

Training Dataset: In this study, 66 out of 95 are considered for training dataset. This is used for developing the ANN.

Testing Dataset: This is required to assess the model's performance. In this study, the remaining 29 data are considered as testing dataset.

The data are normalized between 0 and 1. In this study, the training and testing of BP model is carried out using neural network tool box in MATLAB (Demuth and Beale, 1999).

4   DETAILS OF SVM

Support Vector Machine (SVM) has originated from the concept of statistical learning theory pioneered by Boser et al. (1992). This study uses the SVM as a regression technique by introducing a ε-insensitive loss function. In this section, a brief introduction on how to construct SVM for regression problem is presented. More details can be found in many publications (Boser et al. 1992; Cortes and Vapnik 1995; Gualtieri et al. 1999; Vapnik 1998). There are three distinct characteristics when SVM is 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 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 $\{(x_1, y_1), \ldots, (x_1, y_1)\}, x \in R^n, y \in r$

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 four input variables used for the SVM model in this study are latitude, longitude, and depth. The output of the SVM model is N. So, in this study, $x = [Latitude, longitude, depth]$ and $y = N$.

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

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

This defines an ε tube so that if the predicted value is within the tube the loss is zero, while if the predicted point is outside the tube, the loss is equal to the absolute value of the deviation minus ε. The main aim in SVM is to find a function 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

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

Where, w = an adjustable weight vector and b = 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

$$\text{Minimize: } \frac{1}{2}\|w\|^2$$
$$\text{Subjected to: } y_i - \left(\langle w.x_i \rangle + b\right) \le \varepsilon \ , i = 1, 2,..., l$$
$$\left(\langle w.x_i \rangle + b\right) - y_i \le \varepsilon, i = 1, 2,...,l \tag{3}$$

The above convex optimization problem is feasible. Sometimes, however, this may not be the case, or I also may want to allow for some errors. Analogously to the "soft margin" loss function (Bennett and Mangasarian 1992) which was used in SVM by Cortes and Vapnik (1995).As shown in the Figure 1 for location of boreholes, 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^*) has been introduced to avoid infeasible constraints of the optimization problem (3).

$$\text{Minimize: } \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{l}\left(\xi_i + \xi_i^*\right)$$
$$\text{Subjected to: } y_i - \left(\langle w.x_i \rangle + b\right) \le \varepsilon + \xi_i \ , i = 1, 2,...,l$$
$$\left(\langle w.x_i \rangle + b\right) - y_i \le \varepsilon + \xi_i^*, i = 1, 2,...,l$$

$$\xi_i \geq 0 \text{ and } \xi_i^* \geq 0 , \text{ i = 1, 2,...,l} \tag{4}$$

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 (Smola and Scholkopf 2004). This optimization problem (4) is solved by Lagrangian Multipliers (Vapnik, 1998), and its solution is given by

$$f(x) = \sum_{i=1}^{nsv} \left(\alpha_i - \alpha_i^* \right)(x_i.x) + b \tag{5}$$

Where $b = -\left(\dfrac{1}{2}\right) w.[x_r + x_s]$, 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 (Boser et al. 1992). 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 is mapped onto the feature space by a map Φ. The dot product given by $\Phi(x_i).\Phi(x_j)$ is computed as a linear combination of the training points. The concept of kernel function [$K\left(x_i,x_j \right) = \Phi(x_i)\Phi\left(x_j \right)$] has been introduced to reduce the computational demand (Cristianini and Shwae-Taylor 2000; Cortes and Vapnik 1995). So, equitation (5) becomes written as

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

Some common kernels have been used such as polynomial (homogeneous), polynomial (nonhomogeneous), radial basis function, Gaussian function, sigmoid etc for non-linear cases.

SVM uses the same training dataset, testing dataset and normalization technique as used by ANN model. Radial basis function has been used as kernel function.

5   RESULTS AND DISCUSSION

For predicting N_c values, the three input variables (latitude, longitude and depth) are used for the neural network model in this study. Hence, the input layer has three neurons. The only output is the N_c and therefore the output layer has only one neuron. In BP model, the optimum BP network that is obtained in the present study is a four-layer feed forward network for N. Figure 2 shows the final architecture of the BP model with one hidden layers.
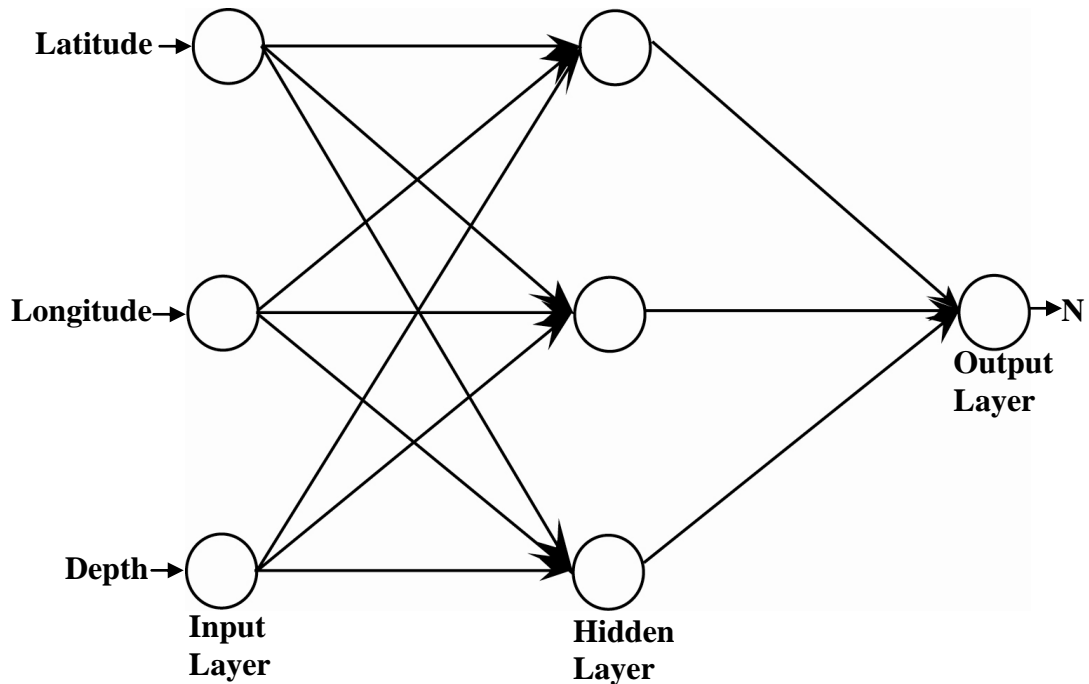
Figure 2. ANN architecture for predicting N.

In this study, the transfer function used in first and second hidden layer is tansig and logsig respectively for N. The logsig transfer function has been used in the output layer. The number of neurons in the hidden layer is determined by training several networks with different numbers of hidden neurons and comparing the predicted results with the desired output. Using too few hidden neurons could result in huge training errors and errors during testing, due to under fitting and high statistical bias. On the other hand, using too many hidden neurons might give low training errors but could still have high testing errors due to over fitting and high variance. In this study, hidden layer with 3 neurons have been used. For BP model, the converged results have been achieved at 500 epochs (an epoch is one complete presentation of the entire set of training patterns during the training process).
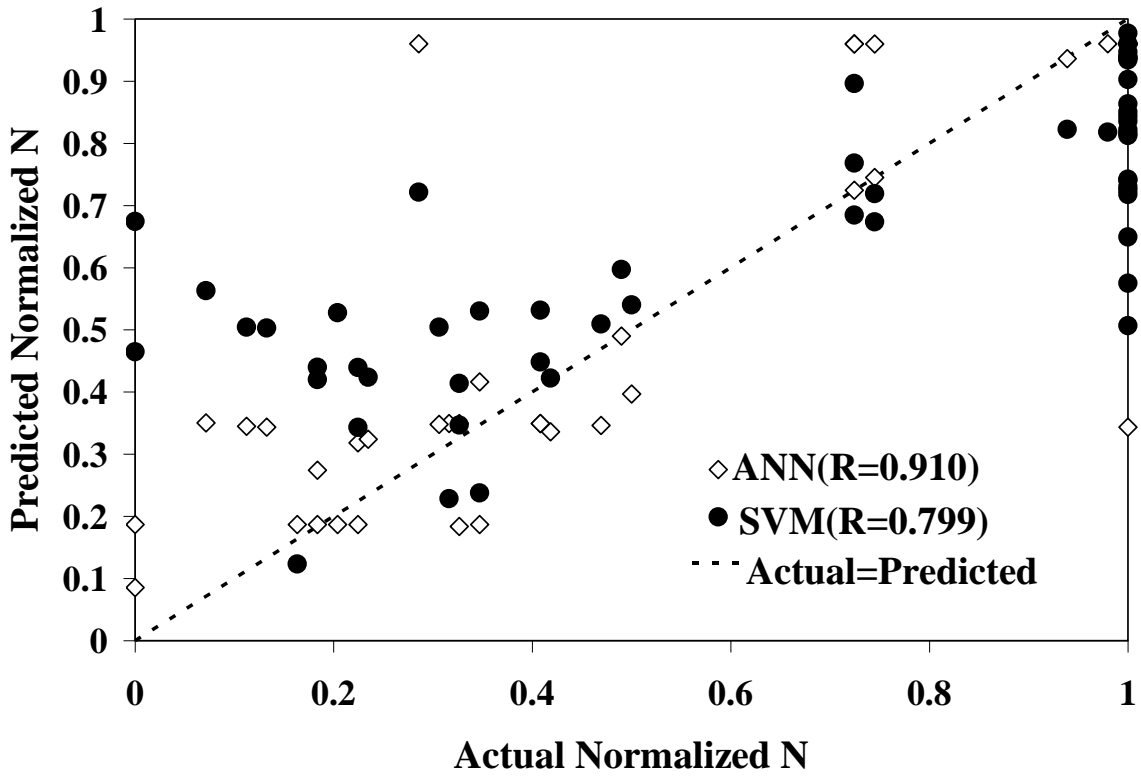
Figure 3. Performance of training dataset.

Figure 3 shows the performance of the back propagation model for training dataset (coefficient of correlation (R) = 0.910). For good model, the value of R should be close to one. According to the results of network training, the network has successfully captured the relationship between the input parameters and output. In order to evaluate the capability of the BP model, the model is validated with testing N data that are not part of the training dataset. Figure 5 shows the performance of the BP model for testing dataset(R=0.716). The result indicates that BP model predicts reasonably well N values in the 3D subsurface of IGCAR.
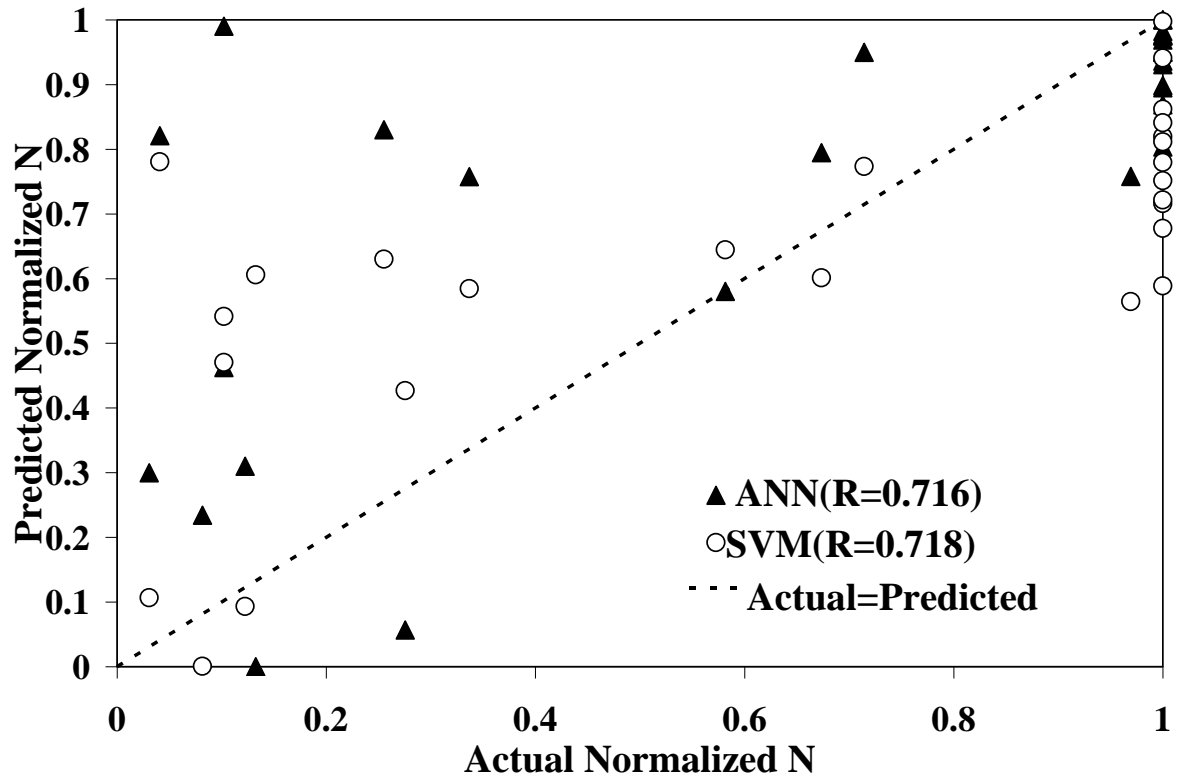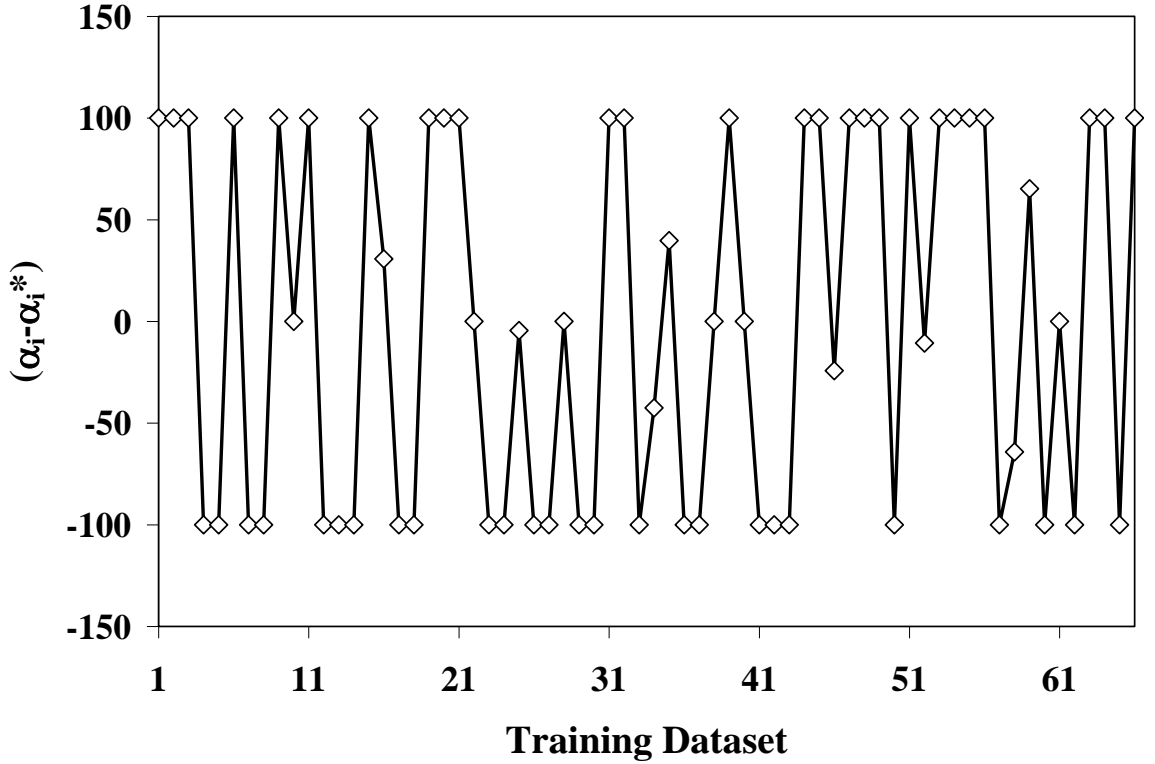
Figure 4. Performance of testing dataset.

For SVM mode, the design value of C, ε and σ has been determined by trail and error approach. The design value of C, ε and σ is 100, 0.04 and 2 respectively. Number of support vector is 60. The performance of training and testing dataset have been determined by using the design values of C, ε and σ. Figure 4 and 5 show the performance of training and testing dataset respectively. The developed SVM gives the following equation for prediction of N value at any point in IGCAR.

$$N = \sum_{i=1}^{66}\left(\alpha_i - \alpha_i^*\right)\exp\left\{-\frac{(x_i - x)(x_i - x)^T}{8}\right\}$$

The value of $\left(\alpha_i - \alpha_i^*\right)$ has been given by figure 5.

Figure 5. Values of $\left(\alpha_i - \alpha_i^*\right)$.

The performance of ANN and SVM is almost same. Table 2 shows Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) of the ANN and SVM models.

Table 2. RMSE and MAE values of the ANN and SVM models.

Error Parameter	Training(ANN)	Training(SVM)	Testing(ANN)	Testing(SVM)
RMSE	0.147	0.256	0.290	0.274
MAE	0.085	0.105	0.286	0.261

4   Conclusion

This article successfully applied ANN and SVM for developing site characterization model of IGCAR. 95 datasets have been used to develop the ANN and SVM models. The developed ANN and SVM give almost same performance. User can use the developed equation for predicting N value at any point in IGCAR. The developed SVM produces sparse solution. In summary, it can be concluded that the developed ANN and SVM is robust models for predicting N value at any point in IGCAR.

References

[1] A. Asaoka, D.A. Grivas, "Spatial variability of the undrained strength of clays", *Journal of Geotechnical. Engineering*, ASCE, Vol. 108, No. 5. (1982), pp.743-745.

[2] G.B. Baecher, "On estimating auto-covariance of soil properties", *Specialty Conference on Probabilistic Mechanics and Structural Reliability*, ASCE, Vol. 110, (1984), pp.214-218.

[3] K.P. Bennett, O.L. Mangasarian, "Robust linear programming discrimination of two linearly inseparable sets", *Optimization Methods and Software*, (1992), 1, 23–34.

[4] B.E. Boser, I.M. Guyon, V.N. Vapnik, "A training algorithm for optimal margin classifier", *In the proceedings of the Fifth Annual ACM Workshop on Computational Learning theory,* Pittusburgh, PA, USA. (1992), 27-29.

[5] C. Buratti, L. Barelli, E. Moretti, "Application of artificial neural network to predict thermal transmittance of wooden windows", *Applied Energy,* No.98, (2012), pp.425-432.

[6] C. Cortes, V.N. Vpanik, "Support-vector networks", *Machine Learning*, Vol. 20, No. 3, (1995), pp.273-297.

[7] N. Cristianini, J. Shawe-Taylor, *An introduction to Support vector machine,* University press, London, Cambridge, (2000),

[8] H.B. Demuth, M. Beale, *Neural Network Toolbox, users guide,* The Mathworks, Inc., Natick, MA, (1999).

[9] X. Du, L. Wang, J. Song, Y. Zhang, "Application of neural network and support vector machines to power system short-term load forecasting", *Proceedings - International Conference on Computational Aspects of Social Networks,* CASoN'10 , No. 5636947, (2010), pp.729-732.

[10] G.A. Fenton, "Random field characterization NGES data", *Paper presented at the workshop on Probabilistic Site Characterization at NGES*, Seattle, Washington. (1998),

[11] J.A. Gualtieri, S.R. Chettri, R.F. Cromp, L.F. Johnson, "Support vector machine classifiers as applied to AVIRIS data", *In the Summaries of the Eighth JPL Airbrone Earth Science Workshop*, (1999).

[12] M.T. Hagan, M.B. Menhaj, "Training Feedforward Networks with the Marquardt Algorithm", *IEEE Trans on Neural Networks*, 5, (1994).

[13] P.H.S.W. Kulatilake, "Probabilistic Potentiometric surface mapping", *Journal of Geotechnical Engineering*, ASCE, Vol. 115, No.11, (1989), pp.1569-1587.

[14] P.H.S.W. Kulatilake, K.M. Miller, "A scheme for estimating the spatial variation of soil properties in three dimensions," *Proc. Of the Fifth Int. Conf. On Application of Statistics and Probabilities in Soil and Struct Engnieering, Vancouver,* British Columbia, Canada, (1987), pp.669-677.

[15] A.G. Leal, J.A. Jardini, S.U. Ahn, "Distribution transformer technical losses estimation with aggregated load curve analytical methodology and artificial neural network approach implementation", *Frontiers in Artificial Intelligence and Applications*, 186 (1), (2009), pp.270-284.

[16] Z. Lingling, Y. Kuihe, "A new engine fault diagnosis model based on support vector machine, 2008 International Conference on Wireless Communications", *Networking and Mobile Computing*, WiCOM, No.4679179, (2008)

[17] O. Ludwig, C. Premebida, U. Nunes, R. Araújo, "Evaluation of Boosting-SVM and SRM-SVM cascade classifiers in laser and vision-based pedestrian detection", *IEEE Conference on Intelligent Transportation Systems, Proceedings*, ITSC, No.6082909, (2011), pp.1574-1579.

[18] P. Lumb, "Spatial variability of soil properties", *Proc. Second Int. Conf. On Application of Statistics and Probability in Soil and struct. Engrg, Aachen,* Germany, (1975), pp.397-421.

[19] M. Mardi, H. Nurozi, S. Edalatkhah, "A water saturation prediction using artificial neural networks and an investigation on cementation factors and saturation exponent variations in an Iranian oil well", *Petroleum Science and Technology*, 30(4), (2011), pp.425-434.

[20] J.J. More, *The Levenberg-Marquardt algorithm: Implementation and theory,* Heidelberg, Springer, Numerical Analysis, Watson, G.A, ed., 105-116, (1977).

[21] E. Mutlu, I. Chaubey, H. Hexmoor, S.G. Bajwa, "Comparison of artificial neural network models for hydrologic predictions at multiple gauging stations in an agricultural watershed", *Hydrological Processes*, 22(26), (2008), pp.5097-5106.

[22] M.B. Naseri, G. Elliott, "Targeting online customers a comparative analysis of artificial neural networks and logistic regression", *Journal of Decision Systems*, 19(3), (2010), pp.291-312.

[23] K.K. Phoon, F.H. Kulhawy, "Characterization of geotechnical variability", *Canadian Geotechnial Journal*, Vol. 36, No. 4, (1999), pp.612-624.

[24] N. Segata, E. Pasolli, F. Melgani, E. Blanzieri, "Local SVM approaches for fast and accurate classification of remote-sensing images", *International Journal of Remote Sensing*, 33(19), (2012), pp.6186-6201.

[25] A.J. Smola, B. Scholkopf, "A tutorial on support vector regression", *Statistics and Computing*, Vol.14, (2004), pp.199-222.

[26] W.H. Tang, "Probabilistic evaluation of penetration resistance", *Journal of Geotechnical Engineering,* ASCE, Vol.105, No.GT10, (1979), pp.1173-1191.

[27] M. Uzielli, G. Vannucchi, K.K. Phoon, "Random filed characterization of strees-normalised cone penetration testing parameters", *Géotechnique,* Vol.5, No.1, (2005), pp.3-20.

[28] E.H. Vanmarcke, "Probabilistic Modeling of soil profiles", *Journal of Geotechnical Engineering,* ASCE, Vol.102, No.11, (1977), pp.1247-1265.

[29] E.H. Vanmarcke, *Random fields: Analysis and synthesis*, The MIT Press, Cambridge, Mass, (1983).

[30] V.N. Vapnik, *The nature of statistical learning theory,* Springer, New York, (1995).

[31] V.N. Vapnik, *Statistical learning theory,* Wiley, New York. (1998).

[32] T.H. Wu, K. Wong, "Probabilistic soil exploration: a case history", *Journal of Geotechnical Engineering,* ASCE, Vol. 107, No. GT12, (1981), pp.1693-1711.

[33] A.M. Yaglom, *Theory of stationary random functions,* Prentice-Hall, Inc., Englewood Cliffs, N.J, (1962).

[34] P.-Y. Zhao, Y.-S. Ding, "Using a fuzzy support vector machine classifier to predict interactions of membrane protein", *3rd International Conference on Bioinformatics and Biomedical Engineering,* iCBBE, No. 5163735, (2009).

[35] Kuok King Kuok, Sobri Harun, Chiu Po Chan, "Hourly Runoff Forecast at Different Leadtime for a Small Watershed using Artificial Neural Networks", *Int. J. Advance. Soft Comput. Appl.*, Vol.3, No.1, (2011), pp.68-86.

[36] Siti Nurmaini, Siti Zaiton Mohd Hashim, Dayang Norhayati Abang Jawawi, "Modular Weightless Neural Network Architecture for Intelligent Navigation", *Int. J. Advance. Soft Comput. Appl.,* Vol.1, No.1, (2009), pp.1-18.

[37] N. Sivasankari, Dr.K. Thanushkodi, "Automated Epileptic Seizure Detection in EEG Signals Using Fast ICA and Neural Network", *Int. J. Advance. Soft Comput. Appl.,* Vol.1, No.2,(2009), pp.91-104.

[38] Mohd Haniff Osman, Choong-Yeun Liong, Ishak Hashim, "Hybrid Learning Algorithm in Neural Network System for Enzyme Classification", *Int. J. Advance. Soft Comput. Appl.,* Vol.2, No.2, (2010), pp.209-220.