Geotech Geol Eng (2008) 26:503–517
DOI 10.1007/s10706-008-9185-4

ORIGINAL PAPER

Spatial Variability of Rock Depth in Bangalore Using Geostatistical, Neural Network and Support Vector Machine Models

T. G. Sitharam · Pijush Samui · P. Anbazhagan

Received: 7 October 2006 / Accepted: 8 March 2008 / Published online: 11 April 2008

Abstract Geospatial technology is increasing in demand for many applications in geosciences. Spatial variability of the bed/hard rock is vital for many applications in geotechnical and earthquake engineering problems such as design of deep foundations, site amplification, ground response studies, liquefaction, microzonation etc. In this paper, reduced level of rock at Bangalore, India is arrived from the 652 boreholes data in the area covering 220 km^2. In the context of prediction of reduced level of rock in the subsurface of Bangalore and to study the spatial variability of the rock depth, Geostatistical model based on Ordinary Kriging technique, Artificial Neural Network (ANN) and Support Vector Machine (SVM) models have been developed. In Ordinary Kriging, the knowledge of the semi-variogram of the reduced level of rock from 652 points in Bangalore is used to predict the reduced level of rock at any point in the subsurface of the Bangalore, where field measurements are not available. A new type of cross-validation analysis developed proves the robustness of the Ordinary Kriging model. ANN model based on multi layer perceptrons (MLPs) that are trained with Levenberg–Marquardt backpropagation algorithm has been adopted to train the model with 90% of the data available. The SVM is a novel type of learning machine based on statistical learning theory, uses regression technique by introducing loss function has been used to predict the reduced level of rock from a large set of data. In this study, a comparative study of three numerical models to predict reduced level of rock has been presented and discussed.

Keywords Rock depth · Geostatistical · Ordinary Kriging · Artificial Neural Network · Support Vector Machine

T. G. Sitharam · P. Samui (✉) · P. Anbazhagan
Department of Civil Engineering, Indian Institute of Science, Bangalore 560 012, India
e-mail: pijush@civil.iisc.ernet.in

T. G. Sitharam
e-mail: sitharam@civil.iisc.ernet.in

P. Anbazhagan
e-mail: anbazhagan@civil.iisc.ernet.in

Notations

a	Range of the variogram
b	The scalar threshold
C	Capacity factor (for learning machine)
C (0)	σ Sill of the variogram
C_0	Nugget of the variogram
l	The number of training sets
R^n	*n*-Dimensional real vector space
w	The boundary
w_i	Weight assigned to each scater point
x	The input vector
y	A binary value representing the two classes
δ_k	Actual error
ε	Error insensitive zone
ε_k	Normalized error

ρ(w,b)	Margin
γ (h)	Semi-variogram
σ	The width of radial basis function
Γ	Gamma function

1 Introduction

Spatial variability of the bed/hard rock with reference to ground surface is vital for many applications in geosciences. Rock depth in a site is very useful parameter to the geotechnical earthquake engineers to find their basic requirement of hard strata and ground motion at rock level. In most of the geotechnical investigations, knowledge of the hard strata or rock is essential to decide the type of foundations and design a suitable foundation for a structure. In the ground response analysis, Peak Ground Acceleration (PGA) and response spectrum for the particular site is evaluated at the rock depth levels and further on at the ground level considering local site effects. This is an essential step to evaluate site amplification and liquefaction hazards of a site and further to estimate induced forces on the structures. In ground response analysis, the response of the soil deposit is determined from the motion at the bed rock level. In all these problems, it is essential to evaluate the depth of the hard rock from the ground level. With an objective of predicting the spatial variability of the reduced level of the bed/hard rock in Bangalore, an attempt has been made to develop models based on Ordinary Kriging technique, Artificial Neural Network (ANN) and Support Vector Machine (SVM). It is also aimed at comparing the performance of these developed models for the available data in Bangalore.

The kriging method was developed during the 1960s and 1970s and has been acknowledged as a good spatial interpolator (Matheron 1963; Isaaks and Srivastava 1989; Davis 2002). The most important features of this method are (1) the unbiased estimate of results, (2) the minimum estimation error, and (3) uncertainty evaluation of interpolation data points. This technique is widely used in the field of earth sciences, including mining, geochemistry, remote sensing, etc. The main goal of kriging is to predict the unknown properties from the knowledge of semi-variogram. Semi-variogram is the analytical tool used to evaluate and quantify the degree of spatial autocorrelation. The semi-variogram is an appreciation of the dispersion of the parameters, which equates to the variance and also gives an autocorrelation distance that represents the radius of influence of a measurement made at a given point. Further, it provides the type of variability that indicates how values fluctuate in space. A new method for cross-validation analysis of developed models has been also proposed and validated. The cross-validation of the model has been done based on the examination of residuals.

ANN model is one of the data processing models made up of highly interconnected nodes (neurons) that map a complex input pattern with a complex output pattern (Kohonen 1988; Khanna 1989; Aleksandar and Morton 1990; Hertz et al. 1991; Dowla and Rogers 1995; Hagan et al. 1996). One of the promising characteristics of the ANN is its ability to learn and generalize from experience and example and to adapt for changing situations. ANN was originated from the work of McCulloch and Pitts (1943), who demonstrated the ability of interconnected neurons to calculate some logical functions. Hebb (1949) pointed out the importance of the synaptic connections in the learning process. Later, Rosenblatt (1958) presented the first operational model of a neural network: the 'Perceptron'. The perceptron, built as an analogy to the visual system, was able to learn some logical functions by modifying the synaptic connections. In this paper, ANN with multi-layer perceptrons (MLPs) that are trained with Levenberg–Marquardt backpropagation algorithm (Hagan and Menhaj 1994) has been developed for predicting reduced level of rock.

The Support Vector Machine (SVM) based on statistical learning theory has been developed by Vapnik (1995). Originally SVM was developed for pattern recognition problem. Recently it has been used to solve non-linear regression estimation and time series prediction by introducing ε-insensitive loss function (Mukherjee et al. 1997; Muller et al. 1997; Vapnik 1995; Vapnik et al. 1997). 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 (Osuna et al. 1997; Gunn 1998). SVM is trained through optimisation of a convex, quadratic cost function, which ensures the uniqueness of the SVM solution. The SVM model depends explicitly on the most "informative" data (the support vectors). In this study, SVM has been used to study the spatial variability of reduced level of rock by introducing ε-insensitive loss function.

2 Subsurface of Bangalore and GIS Model Development

Bangalore covers an area of over 220 km^2 and Ground Reduced Levels (GRL) also vary a lot in the city. It varies from 810 m in north-east part to 940 m in south-western part of Bangalore. GRL does not vary much in the central and northwestern parts of the city. There were over 150 lakes, though most of them are dried up due to erosion and encroachments leaving only 64 at present in an area of 220 km^2. The population of greater Bangalore region is over 6 million and it is the fifth biggest city in India. It is situated on latitude of 12°8′ N and longitude of 77°37′ E. From geology, subsurface of Bangalore region covers in Gneiss complexes, which is formed due to several tectonic-thermal events with large influx of sialic material, are believed to have occurred between 3000 and 3400 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 granodiomitic and granitia 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 (Radhakrishna and Vaidyanadhan 1997). 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.

A Geographic Information System (GIS) model (see Fig. 1) of Bangalore with several layers on a scale of 1:20000 has been developed with a purpose of carrying out microzonation of Bangalore. The Bangalore map forms the base layer for GIS. The map entities have been developed for locating the boreholes to the utmost accuracy and at each location borelogs have been attached along with geotechnical data of each layer upto the hard rock. 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. The locations of boreholes are shown in Fig. 1 along with ground reduced level with an interval of 10 m (see Fig. 2). Distribution of collected boreholes in Bangalore is shown in Fig. 3, indicating a very good distribution of the boreholes in each quadrant of Bangalore from the city center. Figure 1 also depicts grids of 1 km × 1 km along with the corporate boundary of Bangalore and outer boundary circumscribing the ring road. Figure 1 gives a clear view of the spatial distribution of boreholes in Bangalore region. An average of about three boreholes data is available within the grid of 1 km × 1 km.

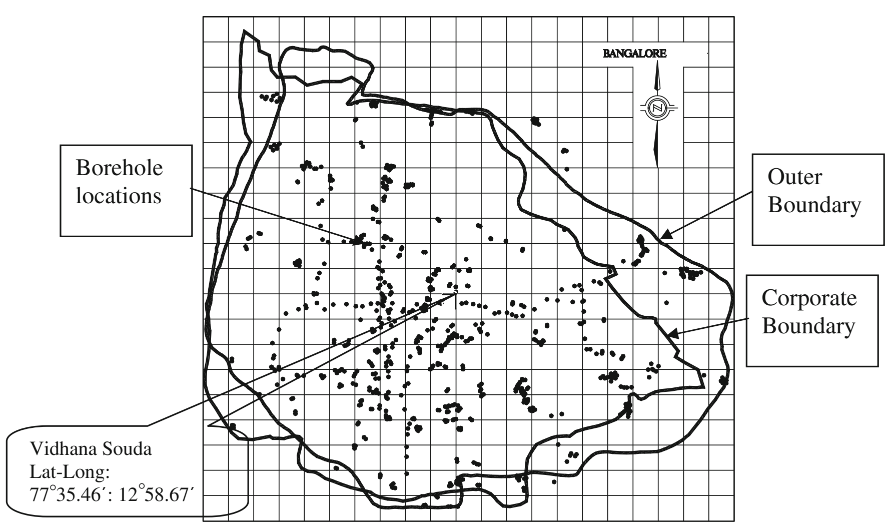

Fig. 1 Borehole location in Bangalore Map (scale: 1:20000)

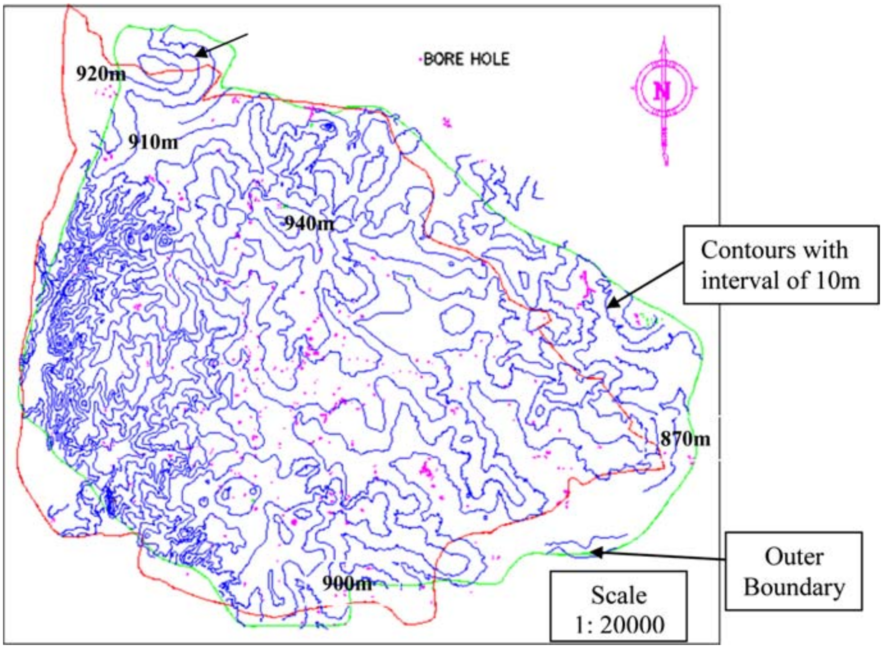

Fig. 2 GIS model of borehole locations with respect to contours

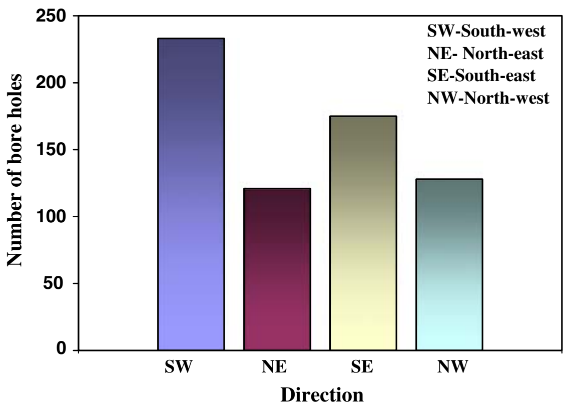

Fig. 3 Distribution of boreholes in quadrants for Bangalore

Geotechnical data for 652 boreholes was collated from archives of only two organizations; Torsteel Research Foundation in India and Indian Institute of Science. This data was generated for geotechnical investigations carried out for several major projects in Bangalore including Bangalore metro project. The data collected is of very high quality and collected during the years 1995–2003. The data in the model is on average to a depth of 30 m below the ground level. Each borelog contains information about depth, density of the soil, total stress, effective stress, fines content and N values, depth of ground water table and rock depth. In GIS model, the boreholes are represented as three dimensional object spanning below the map layer. These three dimensional boreholes are generated with several layers with a bore location in each layer overlapping one below the other and each layer representing 0.5 m interval of the subsurface. Each layer of this model is attached with borelog data at that depth. The data consists of visual soil classification, borehole location, ground water level, date and time during which test has been carried out, other physical and engineering properties of soil and rock depth. As such when this model is viewed in three dimensional subsurface information on any borehole at any depth can obtained by clicking at that level. The hard rock has been identified by visual observation of the cores taken at these locations. Rock depth from ground level is the difference between the ground reduced level at borehole location and reduced level of the hard rock at the same borehole location. The reduced level of the hard rock at borehole location is the difference between the ground reduced level at borehole location and depth of overburden thickness up to hard rock for the same borehole. The depth of overburden is estimated from the available borelogs.

3 Ordinary Kriging Model

In this paper, Ordinary Kriging has been adopted for predicting reduced level of rock in the subsurface of

Bangalore. For this method, there is a need to introduce some terminologies such as covariance function and semi-variogram. The covariance function between two points is defined as:

$$\mathrm{C(h)} = \mathrm{E[(d(x) - m)(d(x') - m)]} \tag{1}$$

where m is the mean of reduced level of rock, d(x) and d(x′) are reduced level of rock values of at points x and x′ respectively and C (h) is the covariance function with a lag h, with h being the distance between two points x and x′:

$$\mathrm{h} = \|\mathrm{x} - \mathrm{x}'\| = \sqrt{(\mathrm{x} - \mathrm{x}')^2 + (\mathrm{y} - \mathrm{y}')^2} \tag{2}$$

and E is the expectation.

The experimental semi-variogram (Matheron 1972; Guillaume 1977) is defined as:

$$\gamma(\mathrm{h}) = 0.5 * \mathrm{E}\left[(\mathrm{d(x)} - \mathrm{d(x')})^2\right] \tag{3}$$

Figure 4 represents the different components of the semi-variogram. The relation between the covariance function and the semi-variogram is as follows:

$$\gamma(\mathrm{h}) = \mathrm{C(0)} - \mathrm{C(h)} \tag{4}$$

where, γ(h) is the semi-variogram and C(0) is sill.

For semi-variogram, the model used in this analysis is spherical model. Once the model of semi-variogram is constructed, the weights are computed for kriging. The details of Ordinary Kriging are given by many researchers (Journel and Huijbregts 1978; Rendu 1978; Clark 1979; Burgess and Webster 1980a, b; Rubeis et al. 2005) and thus it is not discussed here.

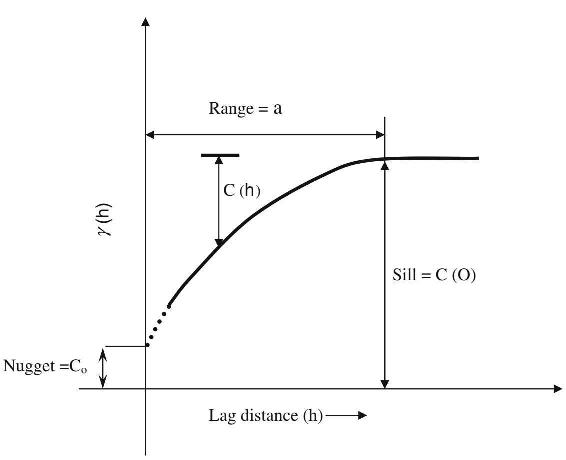

Fig. 4 A typical semi-variogram

A new type of cross-validation analysis for kriging has been presented in this study. In practice, cross-validation is based on statistical tests involving the residuals. The detailed description of residuals in the case of kriging is given by Kitanidis (1991). It has been assumed that the n measurements are available at a time, in a given sequence. The kriging estimate of z at the second point x_2 from the first measurement x_1 is calculated. So, one can write $\hat{Z}_2 = z(x_1)$ and $\sigma_2^2 = 2\gamma(x_1 - x_2)$. where, $\hat{Z}_2$ is the kriged value at the point x_2. The actual error $(\delta_2) = z(x_2) - \hat{Z}_2$ is normalized by the standard error (σ_2) and this normalized value of the error is given by:

$$\varepsilon_2 = \frac{\delta_2}{\sigma_2} \tag{5}$$

For the k-th measurement location, the actual error (δ_k) and normalized error (ε_k) can be written as, respectively:

$$\delta_k = z(x_k) - \hat{Z}_k, \quad \text{for} \quad k = 2, \ldots, n \tag{6}$$

$$\varepsilon_k = \frac{\delta_k}{\sigma_k}, \quad \text{for} \quad k = 2, \ldots, n \tag{7}$$

A cross-validation Q1 and Q2 are used to check the statistical distribution of the residuals between the observed data and kriged values at the original observation location by using the same kriging parameters and semi-variogram model parameters. To perform Q1 and Q2 cross validation, a normalized residual array (ε_k) is constructed as suggested by Kitanidis (1997). Q1 is the mean of the residual (ε_k) and it is written as:

$$\mathrm{Q1} = \frac{1}{n-1} \sum_{k=2}^{hboxn} \varepsilon_k \tag{8}$$

Under the null hypothesis, Q1 is normally distributed with mean 0 and variance $\frac{1}{n-1}$. The probability density function (pdf) of Q1 is:

$$f(\mathrm{Q1}) = \frac{1}{\sqrt{\frac{2\pi}{(n-1)}}} \exp\left(-\frac{\mathrm{Q1}^2}{\frac{2}{(n-1)}}\right) \tag{9}$$

where, n is the number of data. If the experimental value of Q1 turns out to be acceptable close to zero then this test gives no reason to question the validity of the model. The Q2 is the variance of ε_k and it is written as:

$$Q2 = \frac{1}{n-1}\sum_{k=2}^{n} \varepsilon_k^2 \quad (10)$$

(Q2)*(n − 1) approximately follows the chi-square distribution with parameter (n − 1). Where, n is the number of data points. The mean and variance of Q2 are 1 and $\frac{2}{n-1}$ respectively. The pdf of Q2 is given by the following equation:

$$f(Q2) = \frac{(n-1)^{(n-1)/2} Q2^{(n-3)/2} \exp\left(-\frac{(n-1)Q2}{2}\right)}{2^{(n-1)/2}\Gamma\left(\frac{n-1}{2}\right)} \quad (11)$$

where, Γ is the gamma function. For robust model, the experimental value of Q2 should be close to one. In this work, Ordinary Kriging model and cross-validation have been programmed using MATLAB software (MathWork, Inc. 1999).

4 ANN Model

In this study, ANN has been used to approximate the function of reduced level of rock $(d) = d(X, Y)$ where, d(X,Y) corresponds to a reduced level of rock corresponding to X and Y, the coordinates of a point. ANN uses multi-layer perceptrons (MLPs) that are trained with Levenberg–Marquardt backpropagation algorithm. A typical structure of MLPs consists of a number of processing elements, which are usually arranged in layers: an input layer, an output layer and one or more hidden layers. The connections between the neurons in the different layers are such that the output from one neuron is one of the inputs to all the neurons in the next layer and the inputs are the outputs from all the neurons in the previous layer. In backpropagation 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 have to be stopped. In a MLP model, once the system is trained, network can calculate outputs as a functional mapper using last updated network parameters. This is also the reason why MLPs are called as "universal functional approximaters"(Haykin 1999). Levenberg–Marquardt Backpropagation algorithm is a variation of Newton's method and is well suited to ANN training. The theory and implementation of Levenberg–Marquardt Backpropagation has given by More (1977).

For predicting reduced level of rock in a given space, the two input variables(X, Y) are used for the neural network models in this study. Hence, the input layer has two neurons. The only output is the reduced level of rock or rock depth and therefore the output layer has only one neuron. In ANN analysis, normalization of the data is very important. There are many ways of normalizing data, but the method used in this analysis is normalizing the data against their maximum values (Sincero 2003). In ANN modeling, the data has been divided into two subsets; a training dataset, to construct the model, and a testing dataset to estimate the model performance. So, the reduced level of rock data has been divided into training and testing datasets using sorting method, to maintain statistical consistency. In this study, only 10% of the total data selected randomly are considered as testing dataset, which consists of 65 reduced levels of rock values. The remaining 90% reduced level of rock values are considered as training dataset. The statistical consistency of training and testing datasets improve the performance of the backpropagation model and later helps in evaluating them better (Shahin et al. 2000). In backpropagation model, the optimum backpropagation networks that can be obtained in the present study are a four-layer feed forward network. Different types of architecture have been tested to get optimum architecture [For example, one hidden layer with 3,4,5,6,7 and 8 neurons and two hidden layers (1st hidden layer with 3,4,5,6,7,8,9 and 10 neurons and 2nd hidden layer with 3,4,5,6,7,8,9,10,11,12,13,14 and 15 neurons) with different neurons in hidden layers]. The optimum architecture of the backpropagation model with two hidden layers is shown in Fig. 5. In this study, the transfer function used in first and second hidden layers is tansig and logsig respectively. 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 few

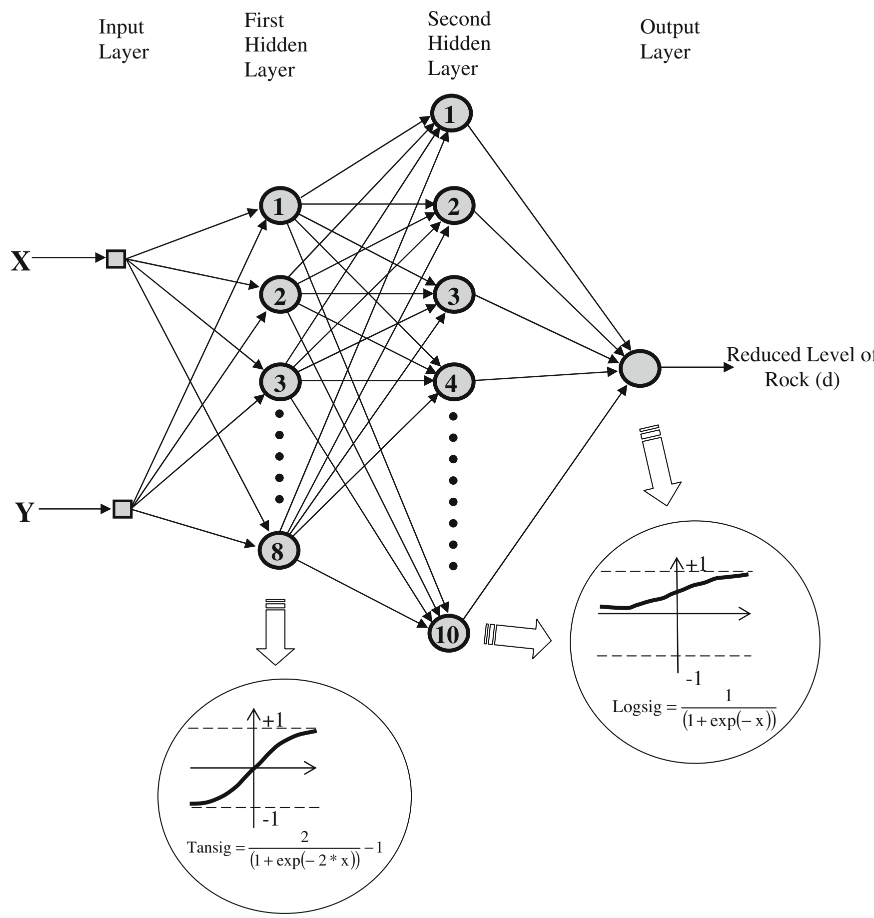

Fig. 5 ANN architecture and transfer functions

hidden neurons could result in huge training errors and errors during testing, due to underfitting 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 overfitting and high variance. In this study, first hidden layer with 8 neurons and the second hidden layer with 10 neurons have been used.

5 Support Vector Machine (SVM) Model

Support Vector Machine (SVM) is a method for pattern classification and regression technique, which has been used successfully to number of applications (Drucker et al. 1999; Furey et al. 2000; Dibike et al. 2001; Guyon et al. 2002; Foody and Mathur 2004). The function can be a classification function or a general regression function. In SVM, the main goal is to separate the two classes by a function which is done by placing a boundary between the two different classes and orient it in a way the margin is maximized. For example in Fig. 6, there are many possible linear classifiers that can separate the data

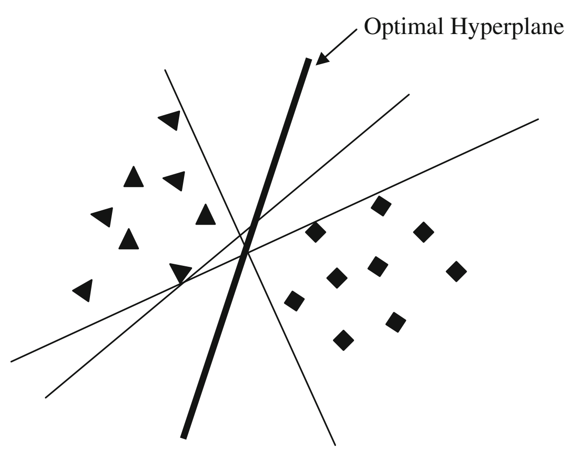

Fig. 6 Optimal separating hyperplane

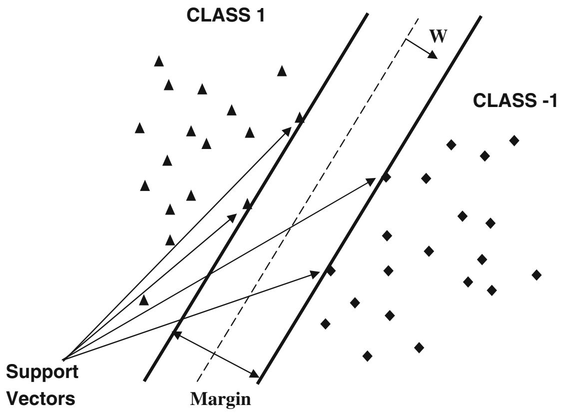

Fig. 7 Support vectors with maximum margin

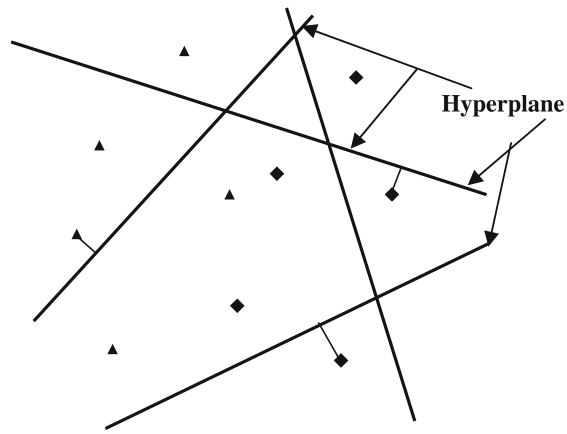

Fig. 8 Canonical Hyperplane

but there is only one that can be maximized for the margin (maximizes the distance between the nearest data point of each class). This linear classifier is called the optimal separating hyperplane. Maximum margin has good generalization capability. The nearest data points are used to define the margin and are known as support vectors (see Fig. 7).

Consider the problem of separating the set of training vectors (D) belonging to two separate classes as:

$$D = \left\{ \left(x^1, y^1\right), \ldots, \left(x^n, y^n\right) \right\}, \quad x \in R^n, \; y \in \{-1, \; 1\} \tag{12}$$

where, y = a binary value representing the two classes (class +1 and class −1), x = the input vector, R^n = n-dimensional vector space.

The hyper plane (a boundary line) considered can be expressed as:

$$\langle w \cdot x \rangle + b = 0, \quad w \in R^n, b \in R \tag{13}$$

where, w = the boundary, x = the input vector, b = the scalar threshold and R = one-dimensional vector space. Equation 13 has some redundancy. To remove redundancy, it is appropriate to consider a canonical hyperplane (Vapnik 1995) satisfying the following condition,

$$\langle w \cdot x \rangle + b = 1, \tag{14}$$

$$\langle w \cdot x \rangle + b = -1. \tag{15}$$

Figure 8 illustrates this idea, where the distance from the nearest point to each hyper plane is shown. A separating hyper plane in canonical form must satisfy the following condition.

$$y_i \lfloor (w \cdot x_i) + b \rfloor \geqslant 1, \quad i = 1, \ldots, n. \tag{16}$$

where, n = the number of training sets.

The distance $d(w, b; x)$ of a point x from the hyperplane (w, b) is,

$$d(w, b; x) = \frac{|(w \cdot x_i) + b|}{\|w\|} \tag{17}$$

The margin (ρ(w,b)) i.e., the sum of the absolute distance between the hyperplane and closest training data points in each class +1 and −1, is given by:

$$\rho(w, b) = \min_{\text{class}-1} d\left(w, b; x^i\right) + \min_{\text{class}+1} d\left(w, b; x^j\right) = \frac{2}{\|w\|} \tag{18}$$

The hyper plane has to be separated the set of vectors without error and the distance between the closest vectors in the plan is maximum. The maximization of this margin leads to the following constrained optimization problem under the inequality constraints of Eq. 16:

$$\min \left\{ \frac{1}{2} \|w\|^2 \right\} \tag{19}$$

In case where linear supporting hyper plane is inappropriate, SVM applies 'kernel trick' to support hyperplane (Boser et al. 1992). The kernel trick is a method for easily converting a linear classification learning algorithm into a non-linear one, by mapping the original observations into a higher-dimensional non-linear space so that linear classification in the new space is equivalent to non-linear classification in the original space. The transformation may be non-linear. Some common kernels have been used such as polymomial(homogeneous), polynomial(nonhomogeneous), radial basis function, gaussian function, sigmoid function, etc, for non-linear cases. SVM can also be applied to regression problems by the introduction of an alternative loss function that is modified to include a distance measure (Smola 1996). The possible loss functions are quadratic, Laplace, ε-insensitive and Huber. In Support Vector Regression (SVR), the basic idea to map the original data, into a feature space, with high dimensionality by a non-linear mapping unknown function, and further to carry on linear regression in this space.

In the present study, SVR has been used for prediction of reduced level of rock values in the subsurface of Bangalore. ε-insensitive loss function has been used in this analysis. In SVR modelling, the data has been divided into two sub-sets; a training dataset, to construct the model, and a testing dataset to estimate the model performance. So, for our study the 90% reduced level of rock data (No of data = 652, 90% of total data = 587) are considered for training dataset. The remaining 10% of the total data is considered as testing dataset, which consists of 65 reduced level of rock values. The data is normalized against their maximum values. The coordinates(X, Y) of each data were prepared as input of the model, while reduced level of rock value was the output from this model. When applying SVM, in addition to the specific kernel parameters, the optimum values of the capacity factor C and the size of the error-insensitive zone ε should be determined during the modeling experiment. In this study, radial basis function, polynomial and spline function are used as the kernel function of the SVM.

6 Code Availability

In this work, kriging model and cross-validation scheme developed have been programmed using MATLAB software (MathWork, Inc. 1999 and homepage: www.mathworks.com). For ANN, the training and testing of backpropagation model is carried out using neural network tool box in MATLAB (Demuth and Beale 1999). The SVM toolbox (Gunn 2003) in MATLAB has been used. Codes developed and carry out the analysis are available based on request. These can be made available from this website: http://www.isis.ecs.soton.ac.uk/resources/svminfo/.

7 Results and Discussion

In case of Ordinary Kriging, the semi-variogram of reduced level of rock depth obtained from field borelogs shown in Fig. 9. The spherical model has been plotted in Fig. 9 and gives a reasonable fit to the values obtained. The range, sill and nugget of the semi-variogram are 0.95, 1.216 and 0.057 respectively. One of the most important finding of this study is that the semi-variogram is free from white noise or a pure nugget effect. The pure nugget effect corresponds to the total absence of auto-correlation. The semi-variogram stops increasing beyond a certain distance. This semi-variogram is called "transition" models, and corresponds to a random function which is not only intrinsic but also second-order stationary. In this study, semi-variogram has no nugget effect. If the semi-variogram is continuous (no nugget effect), a unique solution cannot be obtained because the determinant of the matrix of coefficients of the kriging system vanishes. There is only one possible solution to this problem, adding a nugget term to the semi-variogram. So, a nugget term has been introduced in the analysis. As a result, spherical model shows nugget effect. In case of cross-validation of kriging model, the acceptable region is defined in the Figs. 10 and 11 (between the two vertical lines). For a good model, the Q1 as well as Q2 must fall in this acceptable region as shown in Figs. 10 and 11. In case of ordinary kriging, the value of Q1 and Q2 is 0.002, 1.069 respectively. The Q1 and Q2 values are well

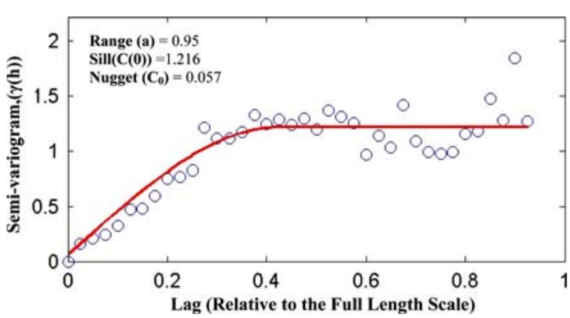

Fig. 9 Semi-variogram model for Ordinary Kriging

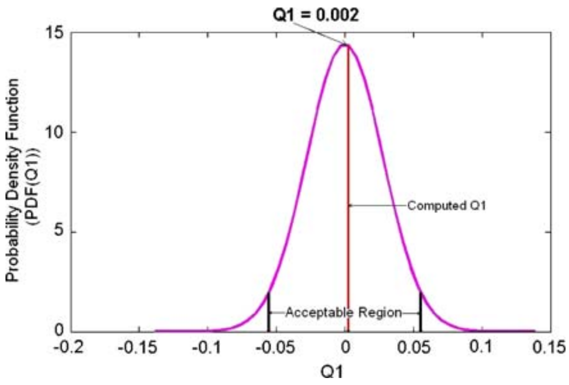

Fig. 10 Distribution of Q1 for Ordinary Kriging

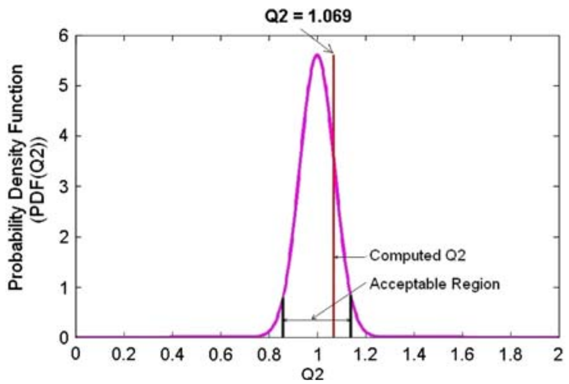

Fig. 11 Distribution of Q2 for Ordinary Kriging

within the acceptable region (shown in Figs. 10, 11). For Ordinary Kriging model, the value of Q1 and Q2 are close to 0 and 1 respectively. The cross-validation indicates that the developed Ordinary Kriging model is robust model for the estimation of the reduced level of rock in the subsurface of Bangalore.

For backpropagation ANN model, the converged results have been achieved at 650 epochs (an epoch is one complete presentation of the entire set of training patterns during the training process). The value of mean square error (MSE) for the backpropagation model has been computed and monitored during training. Figure 12 shows how the MSE for backpropagation model reduces as training proceeds. Figure 13 shows the performance of the backpropagation model for training dataset (Regression Value(R) = 0.941). 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 capabilities of the

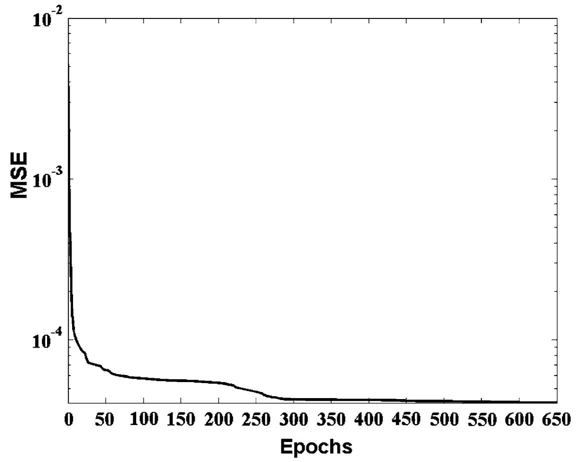

Fig. 12 MSE versus Epochs for Backpropagation model

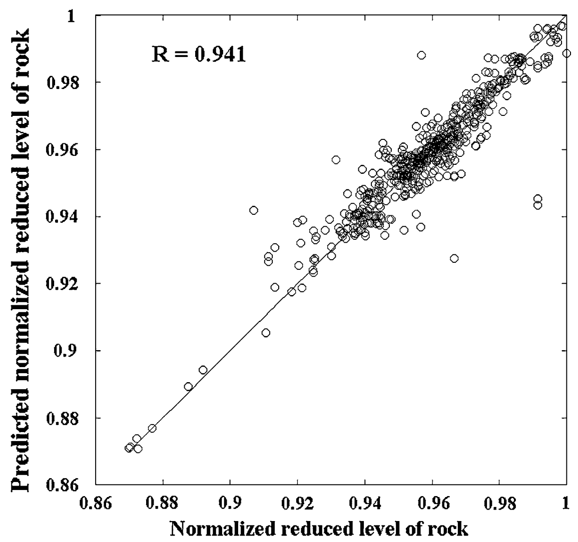

Fig. 13 Performance of Backpropagation model for training dataset

backpropagation model, the model is validated with new reduced level of rock values that are not part of the training dataset. Figure 14 shows the performance of the backpropagation model for testing dataset (R = 0.938). The result indicates that backpropagation model predicts reasonably well reduced level of rock values in the subsurface of the Bangalore. However, a major perceived disadvantage of ANN models is that, unlike other statistical models, they provide no information about the relative importance of the various parameters. In ANN, as the knowledge acquired during training is stored in an implicit manner, it is very difficult to come up with reasonable

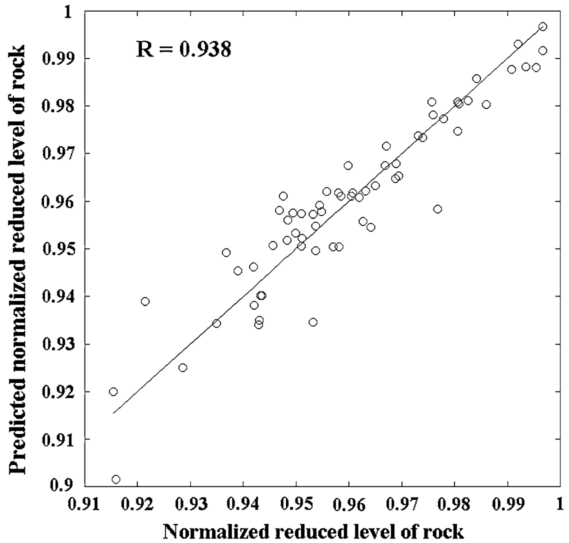

Fig. 14 Performance of Backpropagation model for testing dataset

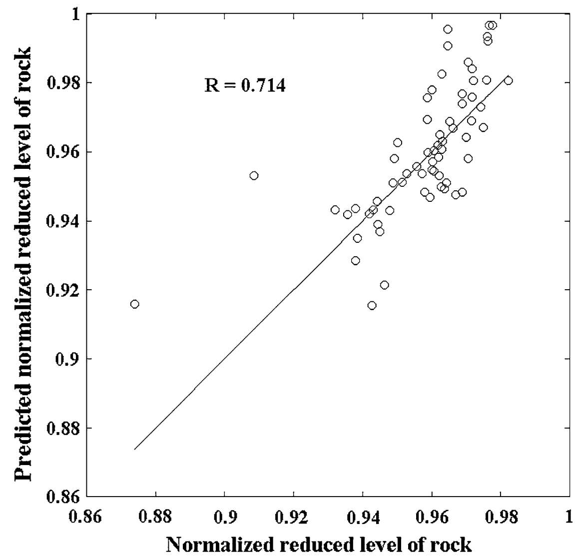

Fig. 16 Performance of SVM model for testing dataset using radial basis function kernel

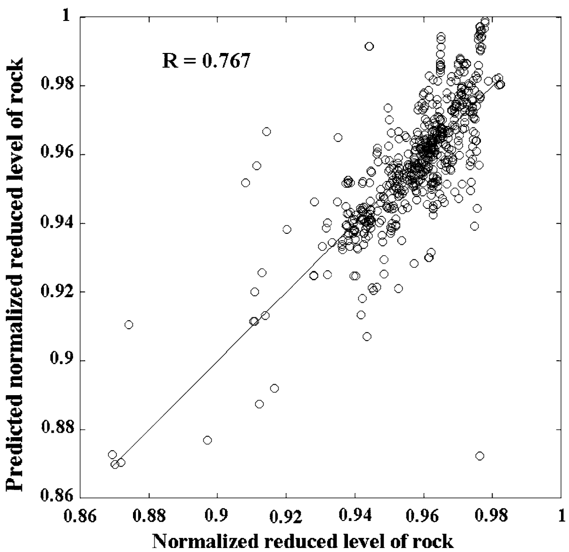

Fig. 15 Performance of SVM model for training dataset using radial basis function kernel

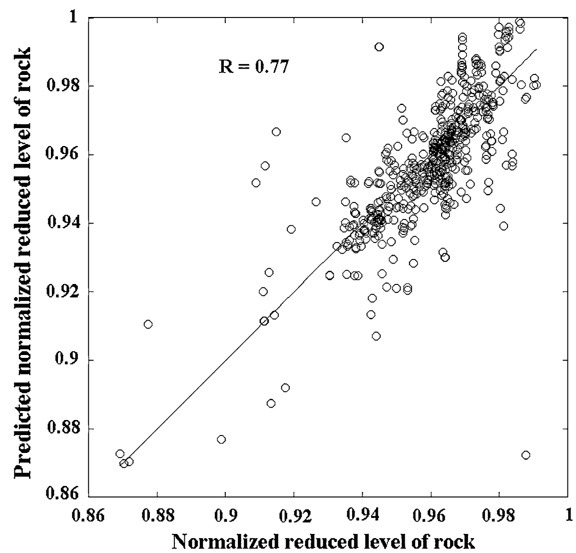

Fig. 17 Performance of SVM model for training dataset using polynomial kernel

interpretation of the overall structure of the network. This lead to the term "black box" which many researchers use while referring to ANN's behavior.

In SVM method, the SVR was found to generalize well by setting the capacity factor C as 100 and ε value as 0.001. At the moment, identification of the optimal values for these parameters is largely a trial-and-error process, which does, however, become much easier with practice. Figures 15, 16 and 17 show the performance of the SVM model for training dataset for radial basis function, polynomial and spline kernel respectively. In order to evaluate the capabilities of the SVR model, the model is validated with new reduced level of rock values that are not part of the training dataset. Figures 18, 19 and 20 show the performance of the SVR model for testing

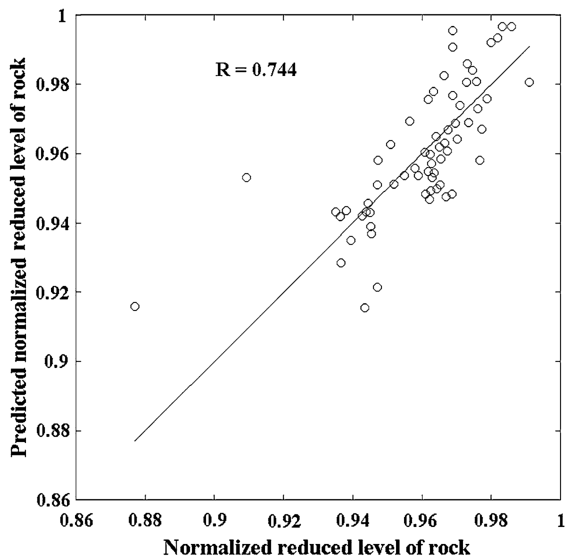

Fig. 18 Performance of SVM model for testing dataset using polynomial kernel

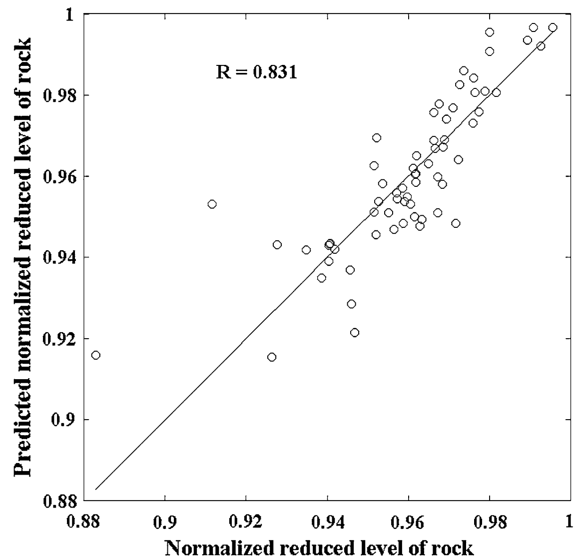

Fig. 20 Performance of SVM model for testing dataset using spline kernel

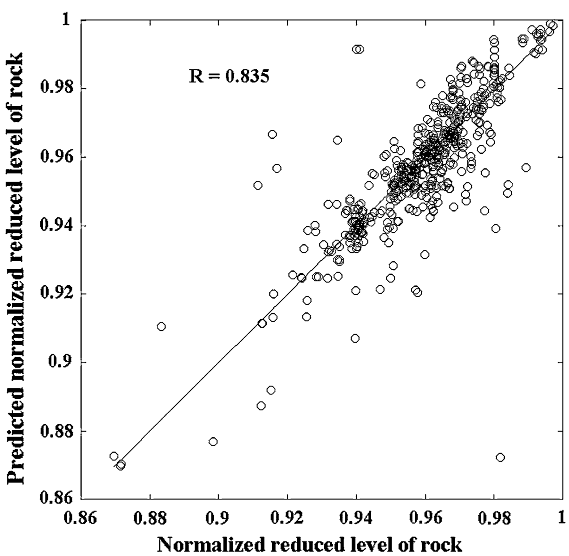

Fig. 19 Performance of SVM model for training dataset using spline kernel

dataset for radial basis function, polynomial and spline kernel respectively. The width (σ) of radial basis kernel is 1 and the degree of polynomial kernel is 4. Table 1 shows the R of SVM for each kernel type with the corresponding numbers of support vectors. From the Table 1, it is clear that spline kernel gives better result than other kernels. The other observation is that an increase the performance of SVM on the training and testing dataset usually corresponds to a decrease in the number of corresponding support vectors.

Table 1 General performance of SVM with different kernel functions

Kernel function	Training performance (R)	Testing performance (R)	No. of support vectors
Radial basis function	0.767	0.714	527
Polynomial	0.77	0.744	518
Spline	0.835	0.831	503

In order to compare between the Ordinary Kriging, ANN and SVM models, five points have been chosen randomly from known reduced level of rock values of 652 points in the subsurface model of Bangalore. The predicted values of these points are shown in Table 2. It can be seen from the table that the ANN model has given best prediction. SVM model gives better prediction than Ordinary Kriging model. Figures 21, 22 and 23 are the surface of reduced level of rock values in the subsurface of Bangalore by Ordinary Kriging, ANN and SVM respectively. Figures clearly indicate that the results obtained from different methods are comparable.

Table 2 Comparison between ANN, SVM and Ordinary Kriging model

Bore hole no.	Longitude (°)	Latitude (°)	Actual reduced level of rock (m)	Predicted reduced level of rock (m) by ANN	Predicted reduced level of rock (m) by SVM (spline kernel)	Predicted reduced level of rock (m) by Ordinary Kriging
71–2	77.5765	12.9448	907.64	908.85	909.13	900.03
53–6	77.6237	12.9447	901	899.74	898.62	910.45
725–39B	77.6641	12.9924	905.86	904.85	901.50	898.65
51–9	77.5874	12.9331	927	926.80	925.87	932.48
87–4	77.5368	13.0293	889	889.13	890.16	882.54

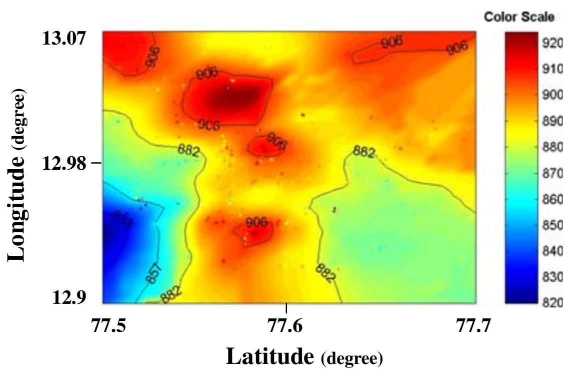

Fig. 21 Surface of reduced level of rock depth using Ordinary Kriging

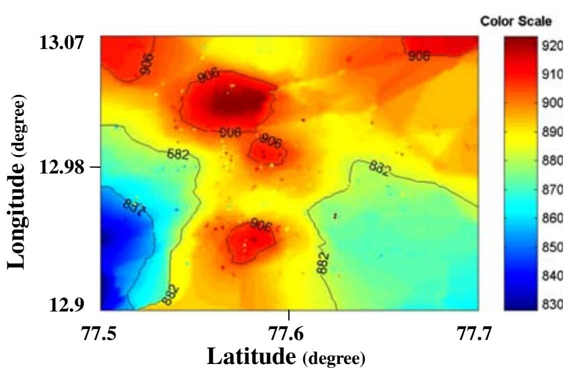

Fig. 22 Surface of reduced level of rock depth using ANN

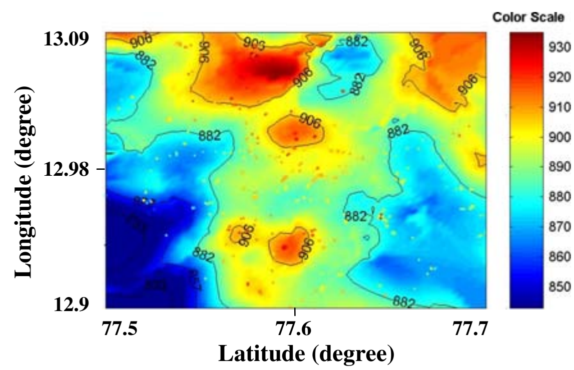

Fig. 23 Surface of reduced level of rock depth using SVM

8 Conclusion

This study has demonstrated the usefulness of Ordinary Kriging, ANN and SVM as tools to predict reduced level of rock or rock depth values of Bangalore considering a large data distributed over 220 km^2 area. Geostatistics has permitted the development of a semi-variogram model for predicting reduced level of rock in Bangalore. The power of geostatistics has become even more apparent through the estimated reduced level of rock in a way that is consistent with what is obtained from actual reduced level of rock values. By the use of the semi-variogram, it is possible to make estimation of the reduced level of rock at any point of the site where reduced level of rock depth was not known. A new type of cross-validation analysis ($Q1 = 0.002$ and $Q2 = 1.069$) which proves the robustness of the developed Ordinary Kriging model has been also presented in this study. For ANN model, the procedures to determine data division, data normalizing technique, network architecture selection, transfer function and no of epochs are outlined. 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 different kernel functions. The SVM was found to generalize well by setting the capacity factor C as 100 and ε value as 0.001. For SVM, spline kernel gave best result. Comparison between the ANN, Ordinary Kriging and SVM models developed with the available data indicates that ANN model is superior to Ordinary Kriging and SVM models developed with the available data for predicting reduced level of rock values in the subsurface of Bangalore.

Acknowledgements Authors thank Seismology division, Department of Science and Technology, Government of India for funding the project titled "Geotechnical site characterization of greater Bangalore region". Ref no. DST/23(315)/SU/2001 dated October 2003.

References

Aleksandar I, Morton H (1990) An introduction to neural computing. Chapman and Hall, London

Boser BE, Guyon IM, Vapnik VN (1992) A training algorithm for optimal margin classifiers. In: Haussler D (ed) 5th Annual ACM workshop on COLT. ACM Press, Pittsburgh, pp 144–152

Burgress TM, Webster R (1980a) Optimal interpolation and isarithmic mapping of soil properties I. The semivariogram and punctual kriging. J Soil Sci 31:315–331

Burgress TM, Webster R (1980b) Optimal interpolation and isarithmic mapping of soil properties II. Block kriging. J Soil Sci 31:333–341

Clark I (1979) Practical geostatistics. Applied Science Publishers, Ltd., London, 129 pp

Davis JC (2002) Statistics and data analysis in geology, 3rd edn. Wiley, New York

Demuth HB, Beale M (1999) Neural Network Toolbox, users guide. The Mathworks, Inc., Natick

Dibike YB, Velickov S, Solomatine D, Abbot MB (2001) Model induction with support vector machine: introduction and application. J Comput Civil Eng 15(3):208–216

Dowla FU, Rogers LL (1995) Solving problems in environmental engineering and geoscience with artificial neural networks. MIT, Cambridge

Drucker H, Donghui W, Vapnik VN (1999) Support vector machine form spam categorization. IEEE Trans Neural Netw 10(5):1048–1054

Foody GM, Mathur A (2004) A relative evaluation of multiclass image classification by support vector machines. IEEE Trans Geosci Remote Sens 42(6):1335–1343

Furey TS, Cristianini N, Duffy N, Bednarski DW Bednarski, Schummer M Haussler D (2000) Support vector machine classification and validation using microarray expression data. Bioinformatics 16(10):906–914

Guillaume A (1977) Introduction a la géologie quantitative: Masson, Paris

Gunn S (1998) Support vector machines for classification and regression. Image Speech and Intelligent Systems Technical Report, University of Southampton, UK

Gunn R (2003) Support vector machines for classification and regression. http://www.ecs.soton.ac.uk/~srg/publications/pdf/SVM.pdf

Guyon I, Weston J, Steohen B, Vapnik V (2002) Gene selection for cancer classification using support vector machines. Mach Learn 46(1–3):389–422

Hagan MT, Menhaj MB (1994) Training feedforward networks with the Marquardt algorithm. IEEE Trans Neural Netw 5(6):989–993

Hagan MT, Demuth HB, Beale M (1996) Neural network design. PWS, Boston

Haykin S (1999) Neural networks: a comprehensive foundation. Prentice-Hall Inc., New Jersey

Hebb DO (1949) The organization of behavior. Wiley, New York

Hertz J, Krogh A, Palmer R (1991) Introduction to the theory of neural computation. Addison-Wesly, Reading

Isaaks EH, Srivastava RM (1989) An introduction to applied geostatics. Oxford University Press, New York

Journel AG, Huijbregts CJ (1978) Mining geostatistics. Academic Press, New York

Khanna T (1989) Foundations of neural networks. Addison-Wesly, Reading

Kitanidis PK (1991) Orthonormal residuals in geostatistics: model criticism and parameter estimation. Math Geol 23(5):741–758

Kitanidis PK (1997) Introduction to geostatistics: applications in hydrogeology. Cambridge University Press, pp 86–95

Kohonen T (1988) An introduction to neural computing. Neural Netw 1(1):3–16

Matheron G (1963) Principles of geostatistics. Econ Geol 58:1246–1266

Matheron G (1972) Théorie des variables régionalisées in Traité d'Informatique Géologique. Masson, Paris, pp 306–378

MathWork, Inc. (1999) Matlab user's manual, version 5.3. The MathWorks, Inc., Natick

McCulloch WS, Pitts W (1943) A logical calculus in the ideas immanent in nervous activity. Bull Math Biophys 5:115–133

More JJ (1977) The Levenberg-Marquardt algorithm: implementation and theory. In: Watson GA (ed) Numerical analysis. Springer, Heidelberg, pp 105–116

Mukherjee S, Osuna E, Girosi F (1997) Nonlinear prediction of chaotic time series using support vector machine. In: Proceedings of the IEEE workshop on neural networks for signal processing 7. Institute of Electrical and Electronics Engineers, New York, pp 511–519

Muller KR, Smola A, Ratsch G, Scholkopf B, Kohlmorgen J, Vapnik V (1997) Predicting time series with support vector machines. In: Proceedings of the international conference on artificial neural networks. Springer-Verlag, Berlin, 999

Osuna E, Freund R, Girosi F (1997) An improved training algorithm for support vector machines. In: Proceedings of the IEEE workshop on neural networks for signal processing 7. Institute of Electrical and Electronics Engineers, New York, pp 276–285

Radhakrishna BP, Vaidyanadhan R (1997) Geology of Karnataka. Geological Society of India, Bangalore

Rendu JM (1978) An introduction to geostatistical methods of mineral evaluation. S African Inst of Min and Metal, Kimberly, 84 pp

Rosenblatt F (1958) The perceptron: a probabilistic model for information storage and organization in the brain. Psychol Rev 68:386–408

Rubeis VD, Tosi P, Gasparini C, Solipaca A (2005) Application of Kriging technique to seismic intensity data. Bull Seismol Soc Am 95(2):540–548

Shahin MA, Jaksa MB, Maier HR (2000) Predicting the settlement of shallow foundations on cohesion less soils using back-propagation neural networks. Department of Civil and Envi Eng, University of Adelaide, Australia, R167

Sincero AP (2003) Predicting mixing power using artificial neural network. EWRI World Water and Environmental

Smola A (1996) Regression estimation with support vector learning machines. Technische Universitat Munchen, Munchen

Vapnik V (1995) The nature of statistical learning theory. Springer, New York

Vapnik V, Golowich S, Smola A (1997) Support method for function approximation regression estimation and signal processing. In: Mozer M, Petsch T (eds) Advance in neural information processing system 9. MIT Press, Cambridge