Application of Soft Computing Techniques to Expansive Soil Characterization

Pijush Samui[1], Sarat Kumar Das[2], and T.G. Sitharam[3]

[1] Postdoctoral fellow, Department of civil engineering,
Tampere University of Technology, Tampere, Finland.
pijush.phd@gmail.com
[2] Assistant Professor, Department of Civil Engineering,
National Institute of Technology Rourkela, India-769 008.
saratdas@rediffmail.com
[3] Professor, Department of civil engineering,
Indian Institute of Science, Banglore, India-560012
sitharam@civil.iisc.ernet.in

Abstract. Very often it is difficult to develop mechanistic models for pavement geotechnical engineering problems due to its complex nature and uncertainty in material parameters. The difficulty in mechanistic analysis has forced the engineers to follows certain empirical correlations. The artificial neural network (ANN) is being as an alternate statistical method, mapping in higher-order spaces, such models can go beyond the existing univariate relationships. The applications of ANNs in pavement geotechnical engineering problems is mostly limited to constitutive modeling, with few applications on prediction of soil layer properties using Falling Weight Deflectometer (FWD), prediction of swelling potential and compute the remaining life of flexible pavements. However, ANN is considered as a 'Black box' system being unable to explain interrelation between inputs and output. The ANNs also have inherent drawbacks such as slow convergence speed, less generalizing performance, arriving at local minimum and over-fitting problems. Recently support vector machine (SVM) is being used due to its, better generalization as prediction error and model complexity are simultaneously minimized. SVM is based on statistical learning theory unlike ANNs (biological learning theory). The application of SVM in pavement geotechnical engineering is very much limited and to best of the knowledge such methods have not been applied to pavement geotechnical engineering. However, engineering application of numerical methods is a science as well as an art. This juxtaposition is based on the fact that even though the developed algorithms are based on scientific logic and belong to the special branch of applied mathematics, their successful application to new problems is problem oriented and is an art. As no method can be the panacea to solve all problems to the last details, their application to new areas needs critical evaluation. With above in view, an attempt has been made to develop the art of applying the above artificial intelligence techniques (ANN and SVM) to different pavement engineering problems such as prediction of compaction characteristics, permeability, swelling potential, coefficient of subgrade reaction etc. The parameters associated with the model developments are discussed in terms of guide line for its future

1 Introduction

The pavement geotechnical engineering is a complex problem involving three phase system. The difficulty in mechanistic analysis and uncertainty in soil

K. Gopalakrishnan et al. (Eds.): Intel. & Soft Comp. in Infra. Sys. Eng., SCI 259, pp. 305–323.
springerlink.com © Springer-Verlag Berlin Heidelberg 2009

parameters has forced the engineers to follows certain empirical correlations. Potential of artificial neural network (ANN) has been realized as an alternate tool to handle such cases and have been successfully applied in various complex problems. The ANN is being as an alternate statistical method, for solving certain types of problems too complex, too poorly understood, or too resource-intensive to tackle using more-traditional computational methods. The ANN is capable of mapping in higher-order spaces, and such models can go beyond the existing univariate relationships. The applications of ANNs in pavement geotechnical engineering problems is mostly limited to constitutive modeling (Ghaboussi 1992), with a few applications on prediction of soil layer properties using Falling Weight Deflectometer (FWD) (Meier and Rix 1994), prediction of swelling potential (Najjar et al. 1996) and compute the remaining life of flexible pavements(Abdallah et al. 2000). However, ANN is considered as a 'Black box' system being unable to explain interrelation between inputs and output. The ANNs also have inherent drawbacks such as slow convergence speed, less generalizing performance, arriving at local minimum and over-fitting problems.

The biggest challenge in successful application of ANN is when to stop training. If training is insufficient then the network will not be fully trained, where as if training is excessive then it will memorize the training patter or learn noise. When the numbers of data points are scanty the training set is driven to a very small value, but when new data is presented to the network the error is too large which is known as overfitting. The network needs to be equally efficient for new data during testing or validation, which is called as generalization. There are different methods for generalization like early stopping or cross validation (Basheer 2001; Shahin et al. 2002, Das and Basudhar 2006). In case of early stopping criteria the error on the validation/testing set is monitored during the training process and the training is stopped when the error on the testing set begin to rise. In cross validation an independent test set is used to asses the performance of the model at various stages of learning. However, this method is not suitable if data points are scanty.

Recently machine learning algorithms like support vector machine (SVM) and relevance vector machine (RVM) are being used due to its, better generalization as prediction error and model complexity is simultaneously minimized. The SVM and RVM are based on statistical learning theory unlike ANNs (biological learning theory). The application of SVM and RVM in geotechnical engineering is very much limited and to best of the knowledge such methods have not been applied to pavement geotechnical engineering. Engineering application of numerical methods is a science as well as an art. This juxtaposition is based on the fact that even though the developed algorithms are based on scientific logic and belong to the special branch of applied mathematics, their successful application to new problems is problem oriented and is an art. As no method can be the panacea to solve all problems to the last details, their application to new areas needs critical evaluation. There are no fixed rules for developing an ANN model, even though a general framework can be followed based on previous successful applications in such problems. With above in view some of problems related to pavement geotechnical engineering are discussed as follows with introduction to the methodology used.

2 Methodology

2.1 Basic Principles of Artificial Neural Network

A typical structure of ANN consists of a number of processing elements or neurons that are usually arranged in layers; an input layer, an output layer and one or more hidden layers (Figure 1). The input from each processing element in the previous layer is multiplied by an adjustable connection weight (w_{ji}). At each neuron, the weighted input signals are summed and a threshold value (b_j) is added. The combined input (I_j) is then passed through a nonlinear transfer function {f()} to produce the output of processing element. Hence the output (y_k) from the output node can be written as Equation (1).

$$y_k = F(v_k) = F\left[\sum_{j=1}^{N_h} w_{kj}\, f\left[\sum_{l=1}^{N_l} w_{jl} x_l + b_{j0}\right] + b_{k0}\right] \tag{1}$$

The 'learning' or 'training' process in ANN in general, is a nonlinear optimization of an error function. The aim of the training is to minimize the error function to get the optimized weight vectors. This is equivalent to the parameter estimation phase in conventional statistical models. The most commonly used error function is the mean squared error (MSE) function. The error associated with weights and sigmoid function is a highly non-linear optimization with many local minima. Local and global optimization methods are carried out for finding out the weight vectors. As the characteristic of traditional nonlinear programming based optimization method are initial point dependent, the results obtained using back propagation algorithm are sensitive to initial conditions (weight vector) (Shahin et al. 2002). The use of global optimization algorithms like genetic algorithm and simulated annealing though being widely used in other field of engineering (Morshed and Kaluarachchi 1998), in geotechnical engineering use of GA for training ANN is limited (Goh 2002; Goh et al. 2005). In recent past another heuristic global optimization called differential evolution (DE), introduced by Storn and Price (1995) is being used successfully in aerodynamics shape optimization and mechanical design.

The steepest descent algorithm and Levenberg-Marquardt (LM) algorithm which are gradient search algorithms are mostly used in ANNs applied to geotechnical engineering problems (Das 2005). The magnitudes of the weights and biases (parameters) are responsible for the poor generalization of the ANN rather than the number of network parameters.

In the present study, the ANN models are trained with differential evolution and Bayesian regularization method and are defined as DENN and BRNN respectively. The results are compared with that obtained from commonly used Levenberg-Marquardt trained neural networks (LMNN) to discuss the prediction efficiency of the networks. The above neural network models have been developed using MATLAB tool boxes (Math Works 2001). A brief description about the BRNN and DENN is presented here for completeness.

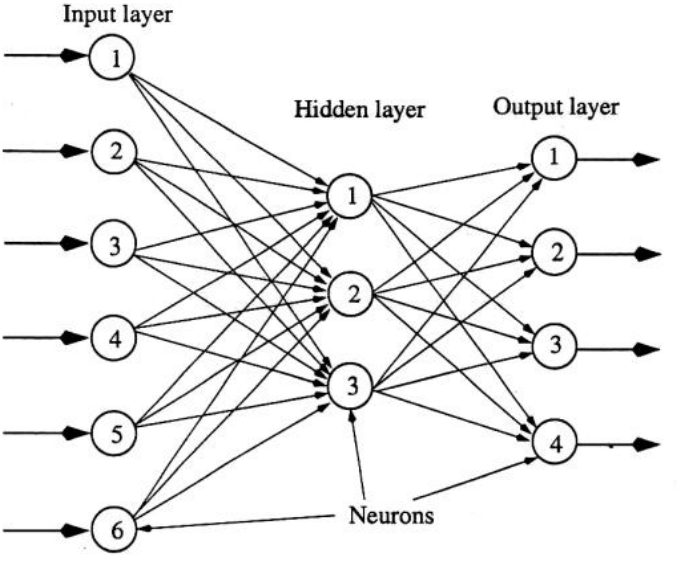


Fig. 1. Typical architecture of a Neural Network

2.1.1 Bayesian Regularization Neural Network (BRNN)

The most commonly used error function is the mean squared error (MSE) function. In LMNN, overfitting is due to unbounded values of weights (parameters) during minimization of the error function, mean square error (MSE). The other method called as regularization, in which the performance function is changed by adding a term that consist of mean square error of weights and biases as given below.

$$\text{MSEREG} = \gamma\,\text{MSE} + (1-\gamma)\;\text{MSW} \tag{2}$$

Where MSE is the mean square error of the network, γ is the performance ratio and

$$\text{MSW} = \frac{1}{n}\sum_{j=1}^{n} w_j^2 \tag{3}$$

This performance function will cause the network to have smaller weights and biases there by forcing networks less likely to be overfit. The optimal regularization parameter λ is determined through Bayesian framework (Demuth and Beale 2000) as the low value of λ will not adequately fit the training data and high value of it may result in over fit. The number of network parameters (weights and biases) are being effectively used by the network can be found out by the above algorithm. The above combination works best when the inputs and targets area scaled in the range [-1, 1] (Demuth and Beale 2000). The above neural network models have been developed using MATLAB tool boxes (Math Works Inc. 2001).

2.1.2 Differential Evolution Neural Network (DENN)

The training of the feed-forward BPNN using DE optimization is known as differential evolution neural network (DENN) (Ilonen *et al*. 2003). The DE optimization is a population based heuristic global optimization method. Unlike other evolutionary optimization, in DE the vectors in current populations are randomly sampled and combined to create vectors for next generation. The real valued cross over factor and mutation factor governs the convergence of the search process. The detail of DENN is available in Ilomen et al. (2003). However, the DENN has not been applied in geotechnical engineering. In the present study, DENN has been implemented using the MATLAB (Math Works Inc. 2001) modeling environment.

In the present study single hidden layer is used and number of hidden layer neuron was obtained by trial and error. In the present study, the generalization was given priority and hence, the model with minimum error for the testing data was considered. The best ANN model was obtained with three neurons in the hidden layers.

2.2 Support Vector Machine

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 $\{(x_1, y_1),....,(x_l, y_l)\}, x \in R^n$, $y \in r$. Where x is the input, y is the output, R^N is the N-dimensional vector space and r is the one-dimensional vector space.

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

$$L_\varepsilon(y) = 0 \text{ for } |f(x) - y| < \varepsilon \text{ otherwise } L_\varepsilon(y) = |f(x) - y| - \varepsilon \tag{4}$$

This defines an ε tube (Figure 2) 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 $f(x)$ 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

$$f(x) = (w.x) + b \quad w \in R^n, \; b \in r \tag{5}$$

Where, w = an adjustable weight vector and b = the scalar threshold. Flatness in the case of (5) 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

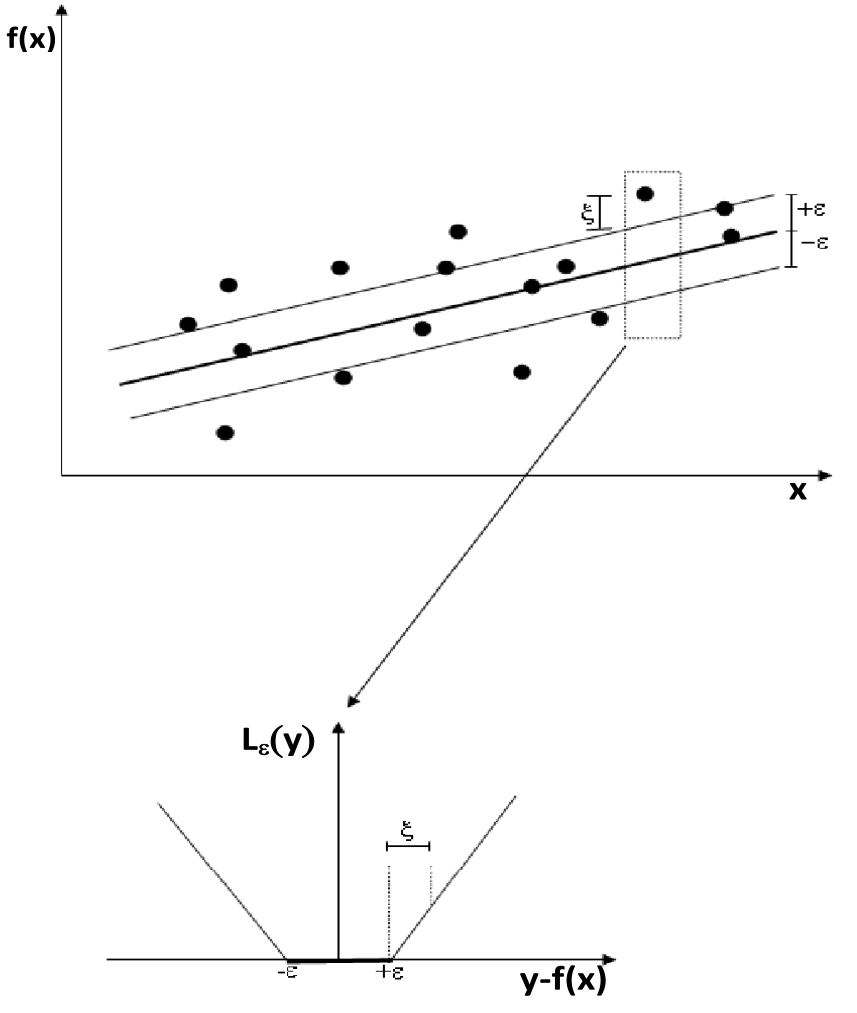


Fig. 2. Prespecified Accuracy ε and Slack Variable ξ in support vector regression [Scholkopf (1997)].

$$\text{Minimize: } \frac{1}{2}\|w\|^2$$

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

$$\left(\langle w.x_i \rangle + b\right) - y_i \le \varepsilon \text{ , } i = 1, 2,..., l \tag{6}$$

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 2, 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 (6).

$$\text{Minimize: } \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{l}\left(\xi_i + \xi_i^*\right)$$

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

$$\left(\langle w.x_i \rangle + b\right) - y_i \le \varepsilon + \xi_i^* \text{ , } i = 1, 2,..., l$$

$$\xi_i \ge 0 \text{ and } \xi_i^* \ge 0 \text{ , } i = 1, 2,..., l \tag{7}$$

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 (7) is solved by Lagrangian Multipliers (Vapnik 1998), and its solution is given by

$$f(x) = \sum_{i=1}^{nsv}\left(\alpha_i - \alpha_i^*\right)\left(x_i.x\right) + b \tag{8}$$

Where $b = -\left(\frac{1}{2}\right)w.\left[x_r + x_s\right]$, 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

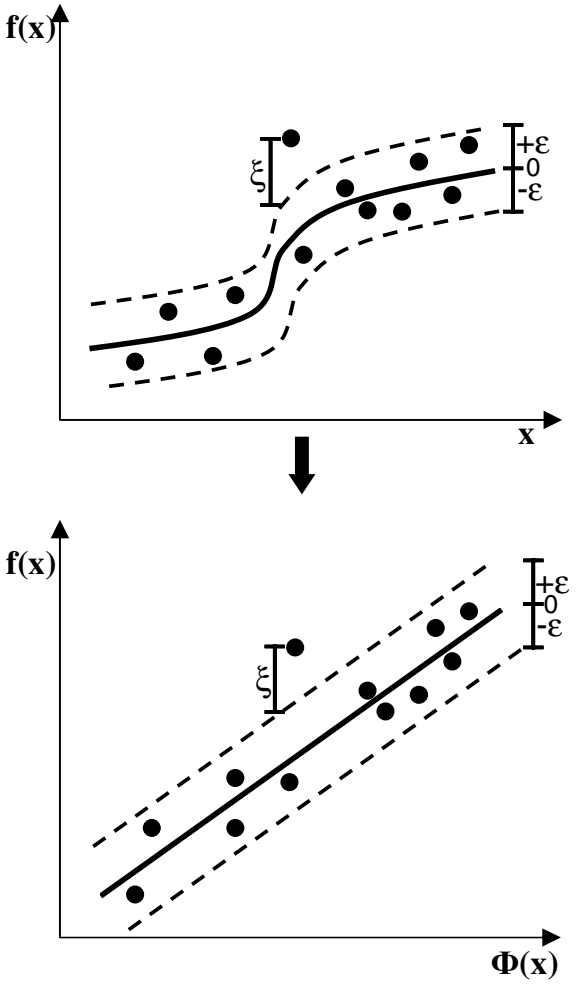


Fig. 3. Concept of nonlinear regression.

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) (see Figure 3). 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(see Figure 3). 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

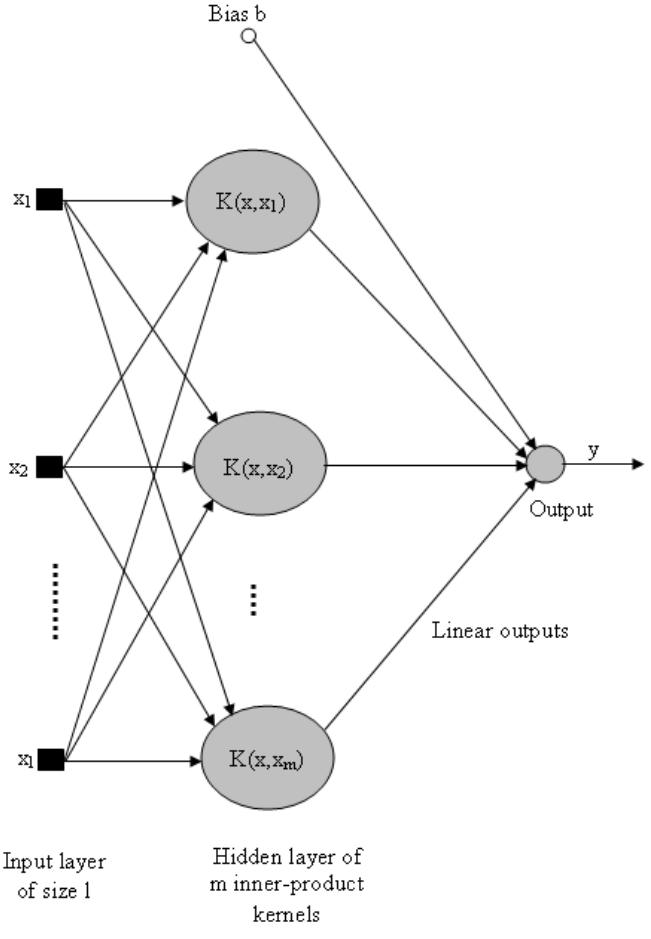


Fig. 4. Architecture of Support Vector Machine (Haykin, 1999).

function [$K(x_i, x_j) = \Phi(x_i).\Phi(x_j)$] has been introduced to reduce the computational demand (Cristianini and Shawe-Taylor 2000, Cortes and Vapnik 1995). So, equitation (5) becomes written as

$$f(x) = \sum_{i=1}^{nsv} (\alpha_i - \alpha_i^*) K(x_i . x_j) + b \tag{9}$$

Some common kernels have been used such as polynomial (homogeneous), polynomial (non homogeneous), radial basis function, Gaussian function, sigmoid etc for non-linear cases. Figure 4 shows a typical architecture of SVM.

The successful application of SVM models depends upon suitable parameters like type of kernel function and the parameters C and ε is obtained by trial and error. 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 general zation ability. If C goes to 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(Thissen et al. 2004). To train the SVM model, three types of kernel function have been used: They are

1. Polynomial
2. Radial basis function
3. Spline

3 Prediction of Swelling Pressure of Expansive Soil

Expansive soil and bedrock underlie more than one third of world's land surface. Each year, damage to buildings, roads, pipelines, and other structures by expansive soils is much higher than damage that are caused by floods, hurricanes, tornadoes, and earthquakes combined Jones and Holtz (1973). The estimated annual cost of damage due to expansive soils is $1000 million in the USA, £150 million in the UK, and many billions of pounds worldwide Gourley et al. (1993). However, as the hazards due to expansive soils develop gradually and seldom present a threat to life, these have received limited attention, despite their severe effects on the economy. Much of the damage related to expansive soils is not due to a lack of appropriate engineering solutions but to the non recognition of expansive soils and expected magnitude of expansion early in land use and project planning. The damage to foundation on expansive soil can be avoided / minimized by proper identification, classification, quantification of swell pressure and provision of an appropriate design procedure. Swelling potential of clayey soil is a measure of the ability and degree to which such a soil might swell if its environments were changed in a definite way. Hence, the expansive soil is classified based on its potential for swelling. However, there is not a definite expression of swell potential for classification of expansive soils (Nelson and Miller 1992). Holtz (1959) referred to swell potential as the volume change of air-dried undisturbed sample, whereas, Seed et al. (1962) defined it as change in volume of a remoulded sample. Though factors like clay content, Atterberg's limits and mineral types are found to affect the swelling potential, the available literature presents contradicting results. McCormack and Wilding (1975) observed that for soil dominated by illite, clay content to be as reliable in predicting swelling potential, where as Yule and Ritchie (1980) and Gray and Allbrook (2002) reported, there being no relationship

between clay percentage and soil swelling. The cation exchange capacity (CEC), saturation moisture and plastic index (PI) are also important indices for estimation of swelling potential Gill and Reaves (1957). Parker et al. (1977) concluded swell index and PI as superior to other indices.

The swelling pressure depends upon various soil parameters such as mineralogy, clay content, Atterberg's limits, dry density, moisture content, initial degree of saturation, etc along with structural and environmental factors. The parameters are interrelated in a complex manner, and it is difficult to model and analyze effectively taking all the above aspects into consideration. However, it can be measured easily with relevant data pertaining to soil, structure and environment. So various statistical/empirical methods have been attempted to predict the swelling pressure based on index properties of soil (Das 2002).

4 Results and Discussion

The data from various sources available in literature (Aciroyd et al. 1988; Savana et al. 1978; Abdujauwad 1994; Abdujauwad et al. 1994) are taken with input parameters, natural moisture content (w_n), dry density (γ_d), LL, PI, clay fraction (CF) and swelling pressure (SP) as output. The total number of data points considered is 230 out of which 167 are taken for training and 63 are taken for testing. The data is normalized between 0 to 1. The maximum, minimum, average and standard deviation for the data used are shown in Table 1 and it can be seen that it covers a wide range of values. The successful application of a method depends upon the identification of suitable input parameters. Table 2 shows the cross correlation between the inputs and output, it can be seen that CF, LL, PI are found to be important input parameters.

The results of different ANN models using the above parameters are shown in Table 3. The correlation coefficient (R) and root means square error (RMSE) are mostly for performance criteria evaluation of ANN models. However, R is a

Table 1. Parameters of the data considered for the present study

	w_n (%)	γ_d (kN/m^3)	LL	PI	Clay Fraction	Swelling pressure (kN/m^2)
Maximum	63.90	15.70	193.00	165.00	97.00	805.00
Minimum	2.70	1.04	26.00	12.00	19.00	3.00
Average	18.31	9.33	90.74	59.06	41.19	122.61
Std Dev.	9.58	5.89	47.77	43.42	14.04	140.90

Table 2. General performance of different neural network models.

ANN models	Training data		Testing data		Overfitting ratio
	R	E	R	E	
DENN	0.95	0.91	0.87	0.75	1.37
BRNN	0.98	0.96	0.90	0.79	2.12
LMNN	0.95	0.90	0.88	0.74	1.43
SVM	0.98	0.96	0.94	0.88	1.40

biased parameter and sometimes, higher values of R may not necessarily indicate better performance of the model because of the tendency of the model to be biased towards higher or lower values (Das and Basudhar 2006), the coefficient of efficiency (E) is also considered. The E is defined as

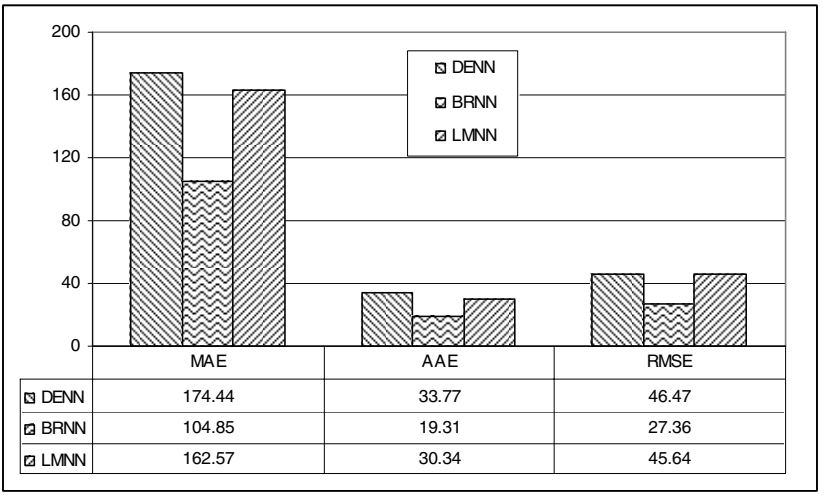
$$E = \frac{E_1 - E_2}{E_1} \tag{10}$$

Where

$$E_1 = \sum_{t=1}^{N} (SP_m - \overline{SP_m})^2$$
$$E_2 = \sum_{t=1}^{N} \left(SP_p - SP_m\right)^2 \tag{11}$$

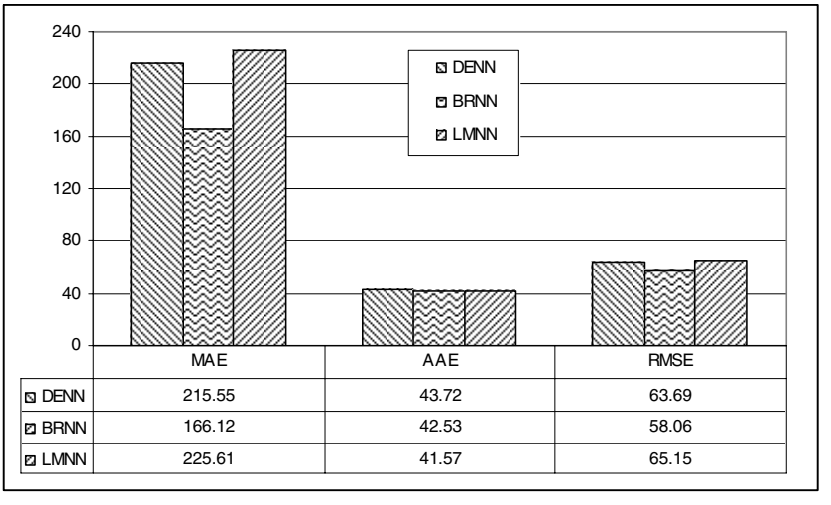
and SP_m, $\overline{SP_m}$ and SP_p are the measure, average and predicted swelling pressure respectively. The E value compares the modeled and measured values of the variable and evaluates how far the network is able to explain total variance in the data set. The overfitting ratio is defined as the ratio of RMSE for testing and training data and it defines the generalization. It can be seen that comparing the values of R and E values for training and testing data, BRNN is found to better than DENN and LMNN. However, DENN is having good generalization with small overfitting ratio, followed by LMNN and BRNN.

The RMSE value only defined the efficiency of a model as overall; however MAE can reveal the presence of regional areas of poor prediction. Figure 5 and 6 show the value of MAE, AAE and RMSE for different ANN models for training and testing data respectively.



	MAE	AAE	RMSE
DENN	174.44	33.77	46.47
BRNN	104.85	19.31	27.36
LMNN	162.57	30.34	45.64

Fig. 5. Comparison of prediction capabilities of ANN models for training data.



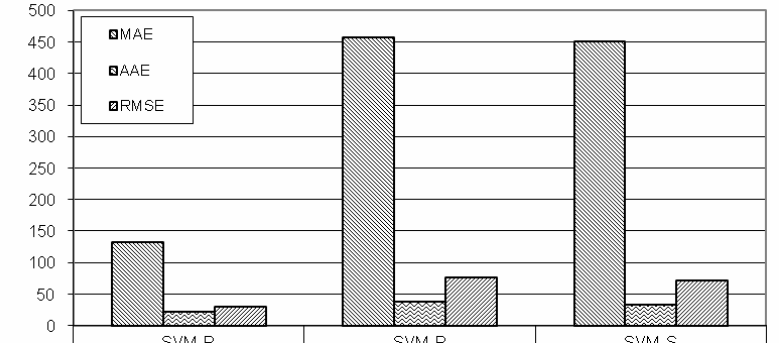
	MAE	AAE	RMSE
DENN	215.55	43.72	63.69
BRNN	166.12	42.53	58.06
LMNN	225.61	41.57	65.15

Fig. 6. Comparison of prediction capabilities of ANN models for testing data.

It can be seen that for training data BRNN is having lowest values of MAE, AAE and RMSE. However, for testing data AAE is comparable for all the methods, but based on MAE and RMSE values BRNN performs better than DENN and LMNN. Hence, based on different statistical performance criteria for the present study it can be concluded that BRNN is better followed by DENN and LMNN.

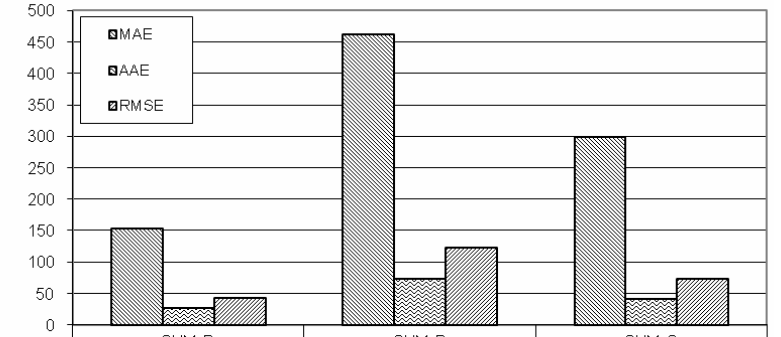
The ANN is considered as a 'Black box' system due to insufficient explanations to the weight vectors, but methods like Garson's algorithm and connection weight approach have been used utilizing the weight vector to identify the important input vectors (Das and Basudhar 2006). Such a study also made here to compare the above two methods in identifying the important parameters. Table 3 shows the ranking of important input parameters as calculated from Garson's algorithm and connection weight approach with the weights obtained from DENN, BRNN and LMNN. It can be seen from that the ranking of important input parameters as obtained by Garson's algorithm and Connection weight approach are different for BRNN and LMNN, where as for DENN the ranking of 1st and 2nd parameters are same by both the methods.

Table 5 presents the results of SVM models developed and based on R and E values SVM model with radial basis kernel function (SVM-R) found to be more efficient compared to models developed with other kernel functions (SVM-P and SVM-S). From Table 5, it is clear that SVM model employs 65 to 75 % (radial basis function=74.85%, Polynomial kernel=65.26% and spline kernel = 66.46%) of the training patterns as support vectors. So, SVM is remarkable in producing an excellent generalization level while maintaining the sparsest structure. Sparseness means that a significant number of the weights are zero (or effectively zero), which has the consequence of producing compact, computationally efficient models, which in addition are simple and therefore produce smooth functions. In SVM, support vectors represent prototypical examples. The prototypical examples exhibit the essential features of the information content of the data, and thus are able to transform the input data into the specified targets. Figure 7 and 8 show the value of MAE, AAE and RMSE for different SVM models for training and testing



	SVM-R	SVM-P	SVM-S
MAE	131.84	456.62	450.85
AAE	21.60	37.41	33.35
RMSE	29.64	76.58	70.51

Fig. 7. Comparison of prediction capabilities of SVM models for training data.



	SVM-R	SVM-P	SVM-S
MAE	152.47	461.06	298.75
AAE	26.89	73.43	40.85
RMSE	41.78	123.16	72.74

Fig. 8. Comparison of prediction capabilities of SVM models and for testing data

data respectively. It can be seen that for training data SVM-R is having lowest values of MAE, AAE and RMSE. In comparison to ANN models SVM-R model is found to better than all the ANN models. The use of the SRM principle in defining cost function provided more generalization capacity with the SVM compared to the ANN, which uses the empirical risk minimization principle. SVM uses only three parameters (radial basis function: σ, C and ε; polynomial kernel: degree of polynomial, C and ε; spline kernel: C and ε). In ANN, there are a larger number of controlling parameters, including the number of hidden layers, number of hidden nodes, learning rate, momentum term, and number of training epochs, transfer functions, and weight initialization methods. Obtaining an optimal combination of these parameters is a difficult task as well. Another major advantage of the SVM is its optimization algorithm, which includes solving a linearly constrained quadratic programming function leading to a unique, optimal, and global solution compared to the ANN. In SVM, the number of support vectors has determined by algorithm rather than by trial-and-error which has been used by ANN for determining the number of hidden nodes.

In this study, a sensitivity analysis has been carried out to extract the cause and effect relationship between the inputs and outputs of the SVM model. The basic idea is that each input of the model is offset slightly and the corresponding change in the output is reported. The procedure has been taken from the work of Liong et al. (2000). According to Liong et al. (2000), the sensitivity (S) of each input parameter has been calculated by the following formula

$$S = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{\%\text{Change in output}}{\%\text{Change in input}}\right)\times 100 \tag{11}$$

Table 4. Relative importance of different input parameters

Input Parameters	Garson's algorithm			Connection weight approach			Ranking of inputs as per relative importance
	Ranking of inputs as per relative importance			Ranking of inputs as per relative importance			
	DENN	BRNN	LMNN	DENN	BRNN	LMNN	SVM-R
w_n	5	5	4	4	4	4	3
γ_d	3	3	3	5	3	3	2
LL	2	1	1	2	2	2	4
PI	1	2	2	1	1	1	1
CF	4	4	5	3	5	5	5

Table 5. General performance of SVM for different kernels

Kernel	C	ε	Training performance		Testing performance		Number of support vector
			Correlation coefficent (R)	Coefficient of determination (E)	Correlation coefficent (R)	Coefficient of determination (E)	
Radial basis function, width(σ) = 2.6	20	0.009	0.979	0.958	0.941	0.887	125
Polynomial, degree = 2	10	0.01	0.865	0.726	0.652	0.018	109
Spline	4	0.01	0.890	0.768	0.859	0.657	112

Where N is the number of data points. The analysis has been carried out on the trained model for radial basis function by varying each of input parameter, one at a time, at a constant rate of 20%. The result of the above analysis is also presented in Table 4. It is observed that similar to ANN analysis using connection weight approach PI is found to be more important parameters followed by γ_d and w_n.

5 Conclusions

The different ANN techniques and SVM model examined here have shown the ability to build accurate models with high predictive capabilities for prediction of swelling pressure of soil from the inputs; natural moisture content ($w_{n)}$, dry density (d), liquid limit (LL), plasticity index (PI) and clay fraction (CF). Based on different statistical performance criteria, the Bayesian regularization neural network (BRNN) model found to be more efficient compared to DENN and LMNN. However, the DENN model found to better in terms of generalization. The performance of the developed SVM model is better than the developed ANN models.

The ranking of important input parameters found to be consistent as per connection weight approach for the ANN model considered here. However, while using Garson's algorithm the ranking found to be different for different ANN models. Developed ANN and SVM models have the advantage that once the model is trained, it can be used as an accurate and quick tool for predicting swelling pressure without a need to perform any manual work such as using tables or charts. Comparison between the ANN and SVM model indicates that SVM model is superior to ANN model for predicting swelling pressure.

References

Abdallah, I., Ferregut, C., Nazarian, S., Melchor Lucero, O.: Prediction of remaining life of flexible pavements with artificial neural network models. In: Tayabji, S.D., Lukanen, E.O. (eds.) Nondestructive testing of Pavements and Backcalulcation of Moduli, ASTM, STP, 1375, vol. 3, pp. 484–498 (2000)

Abdujauwad, S.N.: Swelling behaviour of calcareous clays from the Eastern Provinces of Saudi Arabia. Quarterly Journal of Engg. Geology 27, 333–351 (1994)

Abdujauwad, S.N., Al-Sulaimani, B, Sl-Buraim: Response of structure to heave of expansive soil. Geotechniques (1994)

Aciroyd, L.W., Husain, R.: Residual and Lacustrine black cotton soil of North- East Nigeria. Geotechnique, 113–118 (1986)

Basheer, I.A.: Empirical modeling of the compaction curve of cohesive soil. Canadian Geotechnical Journal 38(1), 29–45 (2001)

Boser, B.E., Guyon, I.M., Vapnik, V.N.: A training algorithm for optimal margin classifier. In: The proceedings of the Fifth Annual ACM Workshop on Computational Learning theory, Pittusburgh, PA, USA, July 27–29 (1992)

Cortes, C., Vpanik, V.N.: Supportvector networks. Machine Learning 20(3), 273–297 (1995)

Cristianini, N., Shawe-Taylor, J.: An introduction to Support vector machine. Cambridge University Press, London (2000)

Dakhsnamurthy, Raman: A simple method of identifying expansive soil. Soils and Foundation 13(1), 97–104 (1973)

Das, B.M.: Principles of Geotechnical Engineering. Brookes- Cole, New York (2002)

Das, S.K.: Application of genetic algorithm and artificial neural network to some geotechnical engineering problem, Ph.D Thesis submitted to Indian Institute of Technology, Kanpur, India (2005)

Das, S.K., Basudhar, P.K.: Undrained lateral load capacity of piles in clay using artificial neural network. Computer and Geotechnics 33(8), 454–459 (2006)

Demuth, H., Beale, M.: Neural Network Toolbox. The Math Works Inc., USA (2000)

Ghaboussi, J.: Potential application of neuro-biological computational models in geotechnical engineering. In: Pande, G.N., Pietruszezak, S. (eds.) Numerical models in geomechanices, Balkemma, Rotterdam, The Netherlands, pp. 543–555 (1992)

Gill, W.R., Reaves, C.A.: Relationships of Atterberg limits and cation-exchange capacity to some physical properties of soil. Soil Sci. Soc. Am. Proc. 21, 491–494 (1957)

Goh, A.T.C.: Probabilistic neural network for evaluating seismic liquefaction potential. Canadian Geotechnical Journal 39, 219–232 (2002)

Goh, A.T.C., Kulhawy, F.H., Chua, C.G.: Bayesian neural network analysis of undrained side resistance of drilled shafts. J. of Geotech. and Geoenv. Engineering, ASCE 131(1), 84–93 (2005)

Gourley, C.S., Newill, D., Shreiner, H.D.: Expansive soils: TRL's research strategy. In: Proc. 1st Int. Symp. on Engineering Characteristics of Arid Soils (1993)

Gray, C.W., Allbrook, R.: Relationships between shrinkage indices and soil properties in some New Zealand soils. Geoderma 108(3-4), 287–299 (2002)

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

Holtz, W.G.: Expansive clays—properties and problems. Q. Colo. Sch. Mines 54(4), 89–117 (1959)

Ilonen, J., Kamarainen, J.K., Lampinen, J.: Differential Evolution training algorithm for feed-forward neural network. Neural Processing Letters 17, 93–105 (2003)

Jones, D.E., Holtz, W.G.: Expansive soils – The hidden disaster. Civil Engineering ASCE 43(8) (1973)

Liong, S.Y., Lim, W.H., Paudyal, G.N.: River stage forecasting in Bangladesh: neural network approach. Journal of Computing in Civil Engineering 14(1), 1–8 (2000)

MathWork, Inc.: Matlab user's manual. Version 6.5. Natick, The MathWorks, Inc., MA (2001)

McCormack, D.E., Wilding, L.P.: Soil properties influencing swelling in Canfield and Geeburg soils. Soil Sci. Soc. Am. Proc. 39, 496–502 (1975)

Meier, R.W., Rix, G.J.: Backcalculation of flexible pavement moduli using artificial neural networks. In: Transportation research record, 1448, TRB, National Research Council, Washington, DC, pp. 75–82 (1994)

Morshed, J., Kaluarachchi, J.J.: Parameter estimation using artificial neural network and genetic algorithm for free-product migration and recovery. Water resource research AGU 34(5), 1101–1113 (1998)

Najjar, M.Y., Basheer, I.A., McReynold, R.: Neural mod ing of Kansas soil swelling. Transportation Research Record (1526), 14–19 (1996)

Nelsonm, J.D., Miller, D.J.: Expansive Soils: Problem and Practice in Foundation and Pavement Engineering. Wiley, New York (1992)

Okasha, T.M., Abdujauwad, S.N.: Expansive soil in Al-Madinha. Saudi Arabia. Quarterly Journal of Engg. Geology 27, 333–351 (1994)

Park, D., Rilett, L.R.: Forecasting freeway link ravel times with a multi-layer feed forward neural network. Computer Aided Civil and infastructure Engineering 14, 358–367 (1999)

Parker, J.C., Amos, D.F., Kaster, D.L.: An evaluation of several methods of estimating soil volume change. Soil Soc. Am. J. 41, 1059–1064 (1977)

Haykin, S.: Neural Networks. A Comprehensive Foundation. Prentice Hall, Englewood Cliffs (1999)

Samui, P.: Slope stability analysis: a support vector machine approach. Environ. Geol. 56, 255–267 (2008)

Savana, G.S., Rajan, C.R., Srinivasan, B.S., Subrahmanyam, N., Sampathkumar, T.S.: Swelling Characteristics of Black cotton soils of Karnatak. In: Proc. CBIP, 47th Research Session, vol. 3, pp. 109–111 (1978)

Scholkopf, B.: Support vector learning. R. Oldenbourg, Munich (1997)

Seed, H.B., Woodward, J.R., Lundgren, R.J.: Prediction of swelling potential for compacted clays. J. Soil Mech. Found. Div., Am. Soc. Civ. Eng. 88 (SM3), 53–87 (1963)

Shahin, M.A., Maier, H.R., Jaksa, M.B.: Predicting settlement of shallow foundations using neural network. Journal of Geotechnical and Geoenvironmental Engineering ASCE 128(9), 785–793 (2002)

Smola, A.J., Scholkopf, B.A.: A tutorial on support vector regression. Statistics and Computing 14, 199–222 (2004)

Storn, R., Price, K.: Differential Evolution- A simple and efficient adaptive scheme for global optmization over continuous spaces, Technical Report TR-95-012, International Computer Science Institute, Berkeley, CA, USA (1995)

Thissen, U., Pepers, M., Ustuna, B., Melssena, W.J., Buydensa, L.M.C.: Comparing support vector machines to PLS for spectral regression applications. Chemom. Intell. Lab. Syst. 73(3), 169–179 (2004)

Vapnik, V.: Statistical learning theory. Wiley, New York (1998)

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

Yule, D.F., Ritchie, J.T.: Soil shrinkage relationships of Texas vertisols: 1 small cores. Soil Sci. Soc. Am. J. 44, 1285–1291 (1980)