PREDICTION OF HYDRAULIC CONDUCTIVITY OF CLAY LINERS USING ARTIFICIAL NEURAL NETWORK

S. K. Das[1] and P. K. Basudhar[2]

ABSTRACT: This paper pertains to prediction of hydraulic conductivity of soil used as clay liners using artificial neural networks based on soil classification test results like Atterberg's limit, grain size and compaction characteristics. Feed forward back propagation neural network has been used and trained with different combination of input parameters of laboratory test results available in literature. Statistical performances criteria like root mean square error, correlation coefficient, coefficient of determination and overfitting ratio are used to compare different neural network models, the available statistical model and the results obtained using group method of data handling (GMDH) neural network. The neural network models are found to be more efficient and reliable compared to statistical model. Identification of important soil parameters affecting the hydraulic conductivity of soils is discussed. A model equation is presented with weights of the trained neural network as model parameter.

Keywords: Hydraulic conductivity of clay liners, artificial neural network, group method of data handling

INTRODUCTION

Compacted clay in isolation or in combination with geo-membrane liners are used as hydraulic barrier in waste containment facilities. It is required to have desired hydraulic conductivity of the clay liners for proper functioning of the containment system and as per standard it should be less than 10^{-7}cm/sec. The hydraulic conductivity (coefficient of permeability) of soil varies with many factors such as soil density, molding water content, degree of saturation, void ratio, composition, soil structure, permanent properties etc. interdependent complexly with each other (Lambe and Whitman 1976). The hydraulic conductivity of soil can be predicted using hydraulic radius theories (Lambe and Whitman 1976), empirical relationship (Carrier and Beckman 1984), capillary models and statistical models (Wang and Huang 1984; Benson et al. 1994). Kozeny-Carman equation (Lambe & Whitman 1976), which is developed assuming soil pores as pipes and taking into account tortousity of the flow, pipes and specific surfaces is also frequently used. However, the main difficulty in using original Kozeny-Carman equation lies in determining the specific surface of soil. Though specific surface can be estimated or measured, in geotechnical engineering practice it is generally not used, as there is no ASTM standard to find the same for soil (Carrier 2003). Chapuis and Aubertin (2003) have used the Kozeny-Carman

equation for fine-grained soil by estimating specific surface of soil from liquid limit. Carrier and Beckman (1984) presented a semi-empirical correlation for permeability of soil based on liquid limit (LL), plastic limit (PL) and void ratio of soil based on some field and laboratory data. Wang and Huang (1984) identified that bentonite content, liquid limit, plastic limit, fineness modulus, effective and mean diameter controls the permeability of bentonite mixed soil. Benson et al. (1994) presented a stepwise regression statistical model based on field measurement of hydraulic conductivity and other soil properties for natural soil used for clay liners at different sites. They suggested that factors like plasticity index (PI), gravel content (Gr), clay content (C), initial degree of saturation (S_i) and weight of compactor governs the hydraulic conductivity of compacted clays.

There are numerous studies for development of equations based on statistical models correlating the index properties of soil with its permeability. These types of equations/functions are known as Pedotransfer functions and mostly used by agricultural and environmental scientists (Wosten et al. 2001). The Pedotransfer functions can be defined as predictive functions of certain soil properties from other easily-, routinely-, or cheaply-measured properties (McBratney et al., 2002).

[1] Research Scholar, Civil Engineering Department, Indian Institute of Technology Kanpur, INDIA
[2] Professor; Civil Engineering Department, Indian Institute of Technology Kanpur, INDIA
Note: Discussion on this paper is open until December 31, 2007

Application of artificial neural network (ANN) in civil engineering problems is getting popular due to its reliable predictive capability in complex problems (Maier and Dandy 2000; Shahin et al. 2001; Zaheer and Bai 2003; Das 2005). ANNs are not very different from statistical methods, but it has the ability to develop a model/correlation based on the input and output data only (Maier and Dandy 2000).

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. As such, there is a need to explore their successful application to new problems from the experience gained, to frame some guidelines for future applications.

The statistical performances used for comparing different ANN models and other empirical models developed with the statistical/empirical analysis are based on correlation coefficient (R) or coefficient of determination/efficiency (R^2) between the predicted and observed values. But, 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 2005). The R^2 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. However, this may not identify the region where the model is deficient and cannot explain in overall, whether the model under predicts or over predicts compared to other models.

With the above in view, ANN models have been presented for prediction of hydraulic conductivity of soil based on laboratory data results. The data available in literature (Wang and Huang 1984) has been taken and the ANN has been implemented using Matlab (MathWork Inc., 2001) and its neural network toolbox (Demuth and Beale 2000). The results of the ANN models have been compared with the available statistical method and the results obtained using GMDH neural network. After obtaining the results from the ANN modeling, the trained weights has been used to carryout sensitivity analysis in finding out the important input parameters. Model equations have been developed based on the weights of the ANN model and GMDH neural network.

METHODOLOGY

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 out put layer and one or more hidden layers (Fig 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 Eq. (1).

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

The learning process is nothing but the nonlinear optimization of the error function to find out the above weights and biases.
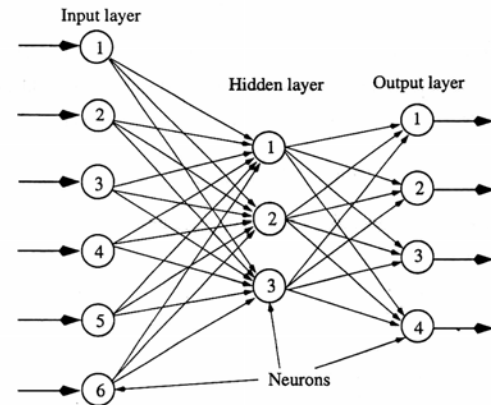

Fig. 1 Typical architecture of a Neural Network

Successful application of ANN depends upon factors like number of hidden nodes, data division, data normalization, transfer function, learning algorithm etc. In the present study three layers (one hidden layer) back propagation neural network is used. The number of hidden layer neurons is determined through a trial- and-error process and the smallest number of neurons that yield satisfactory results (based on performances criteria) is used. In the present investigation as only limited data could be collected, they are divided randomly into training and testing set, without any validation data set. All the variables (inputs and output) are normalized in the range [-1, 1] before training. The network is trained (learning) with Levenberg-Marquardt (LM) algorithm as it is efficient in comparison to gradient descent back propagation algorithm (Demuth and Beale 2000; Goh et al. 2005).

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. So it will not generalize for new set of data. Methods like early stopping or cross validation can be used to avoid overfitting of ANN models (Basheer 2001; Shahin et al. 2002). In the present study as the data points are limited; data for cross validation could not be considered, so early stopping criteria has been used.

Group Method of Data Handling (GMDH)

Group Method of Data Handling (GMDH) algorithms represent sorting-out methods that can be used for analysis of complex objects having no definite theory (Madala and Ivakhnenko 1994). Most GMDH algorithms use the polynomial basis functions and the connectivity configuration is not limited to adjacent layers like ANN. Such generalization of network's topology provides optimal networks in terms of hidden layers and/or number of neurons so that a polynomial expression for dependent variable of the process can be achieved consequently. When constructing a GMDH network, all combinations of the inputs are generated and sent into the first layer of the network. The outputs from this layer are then classified and selected for input into the next layer with all combinations of the selected outputs. If there are m number of inputs then there are m(m-1)/2 number of neurons in 1st layer and number of neurons grows exponentially with number of layers. This process is continued as long as each subsequent layer(n+1) produces a better result than layer(n). When layer (n+1) is found to not be as good as layer(n), the process is halted. The details of the method have been described in Madala and Ivakhnenko (1994). Comprehensive testing of GMDH implemented by Dolenko et al. (1996) proved that it is a powerful tool for mathematical modelling that can be used to solve a wide variety of different real-life problems. The most pronounced feature of GMDH is that it can choose the really significant input variables among dozens of these, thus actually reducing the dimension of the solved problem. However, they notice that GMDH is hardly suitable for very large problems with a great number of inputs which are nearly equally significant and the order of polynomial grows exponentially with number of layers.

Data Base and Preprocessing

The data used in this study to calibrate and validate the neural network models were obtained from the literature and includes laboratory measurement of permeability and the corresponding soil parameters (Wang and Huang, 1984). The data base consists of total 57 cases describing the gravel (Gr), sand (S), silt (Si) and clay (C), specific gravity (G), liquid limit (LL), PL, grain size D_{50} and D_{10}, percent by weight of particles smaller than 0.001 mm ($F_{0.001}$), coefficient of uniformity (C_u), fineness modulus (FM), void ratio at 95% compaction (e_{95}), optimum moisture content (OMC), maximum dry density (MDD) and corresponding permeability (k_l). But 42 data points are considered for the present study as other data points does not have all the above parameters.

Table 1 Results of different ANN models for prediction of hydraulic conductivity of soil comparing the R^2 values

Parameters	Maximum	Minimum	Average
Gravel (%)	20	0	8.67
Clay (%)	84	10	38.69
Silt (%)	83	3	30.33
Sand (%)	81	0	22.85
G	2.87	2.69	2.75
LL (%)	495	24	243.21
PL (%)	47	10	32.64
Log(D50)	-0.32	-3.46	-2.09
Log(D10)	-2.66	-4.42	-3.72
FM	3.55	0.37	1.96
Log Cu	3.79	0.26	2
F0.001 (%)	71	5	30.45
OMC (%)	32.5	8	21.9
MDD (kN/m3)	19.33	12.65	15.72
e_{95}	1.25	0.45	0.83
Log k_l	-5.347	-10.602	-8.205

Table 1 shows the maximum, minimum and average values of the above parameters used in the present study. The data points are divided randomly with 32 data as training set and 10 data as testing set. The results of the different ANN models developed are presented as follows.
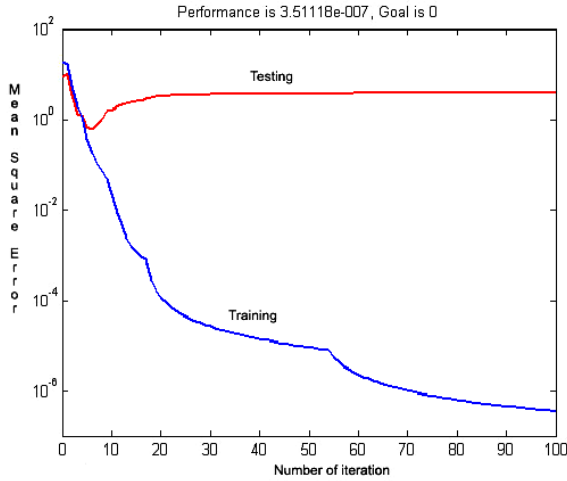
Fig. 2 Performances of proposed model during training and testing

Fig. 4 The observed hydraulic conductivity vs. predicted hydraulic conductivity for ten iterations

RESULTS AND DISCUSSION

Wang and Huang (1984) presented a stepwise regression model (Eq. 2) for prediction of hydraulic conductivity of the soil to be used as clay liners, based on the above data.

$$Log_{10} k_l = -3.601 - 0.0059 \times LL \qquad (2)$$
$$-0.011 \times PL \ Log_{10} D_{50} - 0.28 \times Log_{10}^2 D_{10}$$

where the k_1 is the hydraulic conductivity of clay liners and is correlation is expressed in $Log_{10} k_1$ as it has less skewness value compared to that data taken as k_1.

Figure 2 shows a typical stage of ANN modeling showing the reduction in error due to training and testing data. It can be seen that as the number of epoch (iteration) increases there is decrease in error during training but for testing set data initially there is decrease
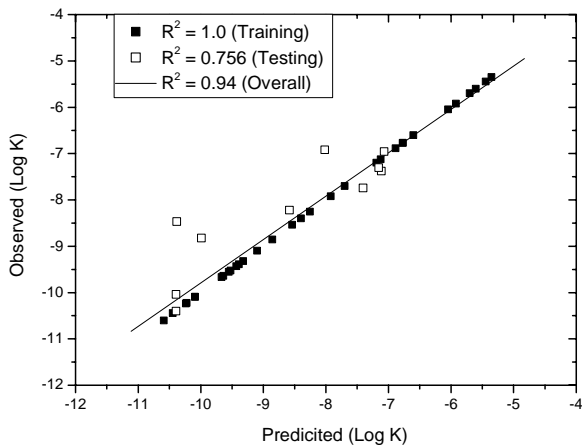

Fig. 3: The observed hydraulic conductivity vs. predicted hydraulic conductivity for hundred iterations
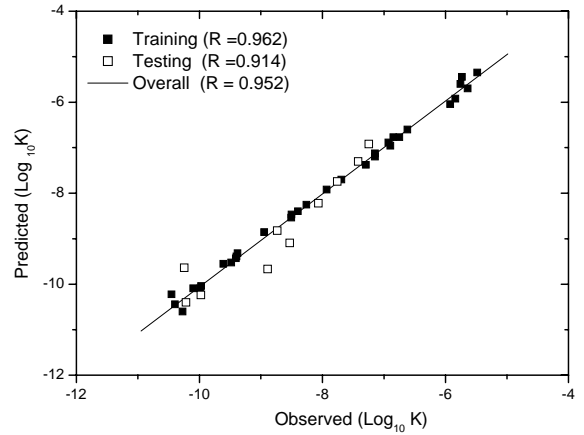
in error up to certain iteration then after the error goes on increasing or remains constant. The correlation between predicted and observed value of soil permeability for training and testing data with 100 iterations is shown in Figure 3. It can be seen that the correlation coefficient (R) value for training data is 1.0 where as for testing data it is 0.756. This shows poor generalization of the model for data other than the training set.

Figure 4 shows the agreement between predicted and observed permeability value when the network training is stopped after ten iterations. Though there was decrease in R (0.962) value for training set, the results of testing phase suggest that ANN model was capable of generalization and gave very good prediction (R = 0.914) for testing set. This is known as early stopping criteria i.e the training is stopped when testing phase error increases even though error during training phase may goes on decreasing.

Different ANN models are developed with different sets of input variables to find out the 'best' model (Basheer 2001; Shahin et al. 2002). For the present study it is observed that optimum results are found with two hidden layer neurons. Some of the successful ANN models are summarized in Table 2. The different models as described in Table 2, are compared in terms of correlation coefficient (R), coefficient of determination (R^2), root mean square error (RMSE) value and overfitting ratio. The overfitting ratio is defined as the ratio of RMSE value during testing and training. The overfitting value close to one shows good generalization of the model. From Table 2, it can be seen that the performances of ANN model depends upon the input parameters. Based on R value (0.98) during training Model -4 is found to have good predictions; however it shows very poor prediction for testing data signified by high overfitting ratio (2.04). The other models are found

to have good generalization with overfitting ratio varying from 1.02 to 1.30. Though Model -5, 8 and 9 are found to have very close values of R, RMSE and overfitting ratio, Model -9 is the 'best' model having high values of R for training (0.977) and testing (0.961), minimum RMSE values (0.362) for testing data and minimum overfitting ratio (1.02). So the Model-9 with eight (8) inputs (C, Gr,LL,PL,D_{50}, D_{10}, C_u, and FM), two (2) hidden layer neurons and one(1) output ($Log_{10}k_l$) variable described as a 8-2-1 ANN architecture, can be considered as the 'best' model.

The R values as per statistical model i.e. Equation 2 (Wang and Huang 1984) for the training and testing data are found to be 0.929 and 0.903 respectively (Table 2). From Table 2 it also can be seen that as per R values, all the ANN models considered for training data and Models 1,2,5,8 and 9 for testing data are more efficient than the statistical model. From Table 2 it can be seen that, Models 5, 8 and 9 are more efficient compared to other ANN models, having high values of R^2 for both training and testing data. Comparing with statistical model the R^2 values (0.862 and 0.897) are less than that of Models 5, 8, and 9. As per R^2 values also, Model-9 found to be the most efficient model. It is noteworthy to mention that in statistical model, both training and testing data were used for model development where as ANNs are not trained with testing data. The weights and biases of the final network are presented in Table 3. The weights and biases can be utilized for selection of important input parameters and framing an equation based on ANN model.

The GMDH was developed with the same inputs as per Model 9. It was observed that the optimum value is obtained after 5th layer with RMSE values as 0.221 and 0.400 during training and testing respectively (Table 2). It can be seen that the RMSE values are comparable to that of ANN values, but the ovefitiing ratio is very high compared to ANN and hence poor generalization.

SELECTION OF IMPORTANT INPUT VARIABLES

The ANN is data driven approach unlike statistical approach and the important inputs are selected based on the performances of the ANN models or by sensitivity analysis using Garson's algorithm (Goh 1994). The different methods for selection of input variables, used in the present study are discussed below:

Correlation Criteria (Cross correlations)

The Pearson correlation coefficient is defined as one of the variable ranking criteria in selecting proper inputs

for the ANN (Wilby et al. 2003; Guyon and Elisseeff 2003). The linear relationship between x_i and y is defined in terms of Pearson coefficient as:

$$R(i) = \frac{\text{cov}(X_i, Y)}{\sqrt{\text{var}(X_i)\text{var}(Y)}} \qquad (3)$$

cov designates the covariance and *var* the variance. The estimate of the R(i) is given as

$$R(i) = \frac{\sum_{k=1}^{m}(x_{k,i} - \overline{x}_i)(y_k - \overline{y})}{\sqrt{\sum_{k=1}^{m}(x_{k,i} - \overline{x}_i)^2 \sum_{k=1}^{m}(y_k - \overline{y})^2}} \qquad (4)$$

Correlation criteria have been used extensively in water resources engineering to select the suitable input variables (Wilby et al. 2003). High value of R_i of an input with the output indicates good correlation between the corresponding input and the output. However, correlation criteria can detect only linear dependencies between input variables and the output.

Garson's Algorithm

Garson (1991) proposed a method of partitioning the neural network weights to determine the relative importance of each input variable in the network which has been modified and used by Goh (1994), Shahin et al.(2002) etc. The input-hidden and hidden-output weights are partitioned and the absolute values of the weights are taken to select the important input variables. The details of the algorithm with an example have been described in Goh (1994).

Table 4 shows the cross correlation of inputs with the $Log_{10}k_l$ value. From the table it can be seen that $Log_{10}k_l$ is highly correlated to LL value (cross correlation value 0.88) followed by clay content (C), D_{10}, PL and Cu. It can also be observed that FM (cross correlation value 0.05), Gr and D_{50} are poorly correlated to $Log_{10}k_l$ values. The sensitivity analysis for the model, as per Garson's method, is presented in Table 5. The LL is found to be the most important input parameter followed by Cu, D_{10}, Gr, PL, FM, C and D_{50}.

Wang and Huang (1984) statistical model also considers LL, PL, D_{50} and D_{10} as important input parameters. Similarly as per the GMDH it was observed that the important Gr and LL are more important followed by D_{10}, FM, C, PL, Cu and D_{50}.

From the above results it can be seen that using the above sensitivity analysis method LL is found to be the most important parameters in predicting $Log_{10}k_l$ value. The relative ranking of other parameters are found to

Table 2 Results of different ANN models for prediction of hydraulic conductivity of soil

ANN Models	Model Inputs	Correlation coefficient (R)		Coefficient of determination (R^2)		RMSE value		Over fitting ratio
		Training	Testing	Training	Testing	Training	Testing	
Model 1	LL,PI,D_{50},D_{10},OMC, MDD	0.962	0.914	0.924	0.736	0.460	0.598	1.30
Model 2	LL,PI,D_{50},D_{10},OMC,MDD, e_{95}	0.962	0.914	0.923	0.736	0.460	0.598	1.30
Model 3	LL,PI,D_{50},D_{10}, e_{95}	0.945	0.892	0.892	0.743	0.548	0.589	1.075
Model 4	C, LL,PI,D_{50},D_{10}, e_{95}	0.980	0.808	0.96	0.64	0.342	0.698	2.04
Model 5	C, Gr, LL,PL, D_{50}, D_{10}, Cu, FM,$F_{0.001}$	0.977	0.955	0.954	0.891	0.358	0.384	1.072
Model 6	C, LL,PL, D_{50}, D_{10},Cu, FM	0.949	0.867	0.899	0.738	0.531	0.595	1.120
Model 7	C, LL,PI, D_{50},D_{10}, Cu, FM	0.949	0.865	0.899	0.738	0.529	0.596	1.127
Model 8	G, LL,PI, D_{50},D_{10}, Cu, FM	0.976	0.947	0.953	0.892	0.362	0.382	1.055
Model 9	C, Gr, LL,PL, D_{50},D_{10},Cu, FM	0.977	0.961	0.955	0.903	0.355	0.362	1.02
GMDH	C, Gr, LL,PL, D_{50},D_{10},Cu, FM					0.221	0.40	1.81
Statistical model	LL,PL, $Log_{10}D_{50}$, $Log_{10}D_{10}$,	0.929	0.903	0.862	0.789	0.619	0.534	

Table 3 Weights and biases for hydraulic conductivity of soil

Neuron	Weights									Biases	
	Input 1	Input 2	Input 3	Input 4	Input 5	Input 6	Input 7	Input 8	Output	b_{hk}	b_0
Hidden neuron 1 ($k=1$)	0.1081	0.04	-0.590	-0.078	-0.026	0.438	-0.437	-0.134	1.017	-0.012	-0.207
Hidden neuron 2 ($k=2$)	-0.091	0.470	-0.813	-0.199	-0.116	-0.148	-0.448	0.066	1.012	0.298	

Table 4 Cross correlation between inputs and the measured hydraulic conductivity

	C	Gr	LL	PL	D_{50}	D_{10}	C_u	FM	$Log_{10}k_l$
C	1.00	-0.16	0.93	0.77	-0.49	-0.65	-0.79	-0.17	-0.75
Gr		1.00	-0.01	0.04	0.11	0.12	0.46	0.25	0.19
LL			1.00	0.74	-0.28	-0.69	-0.55	0.06	-0.88
PL				1.00	-0.46	-0.74	-0.61	-0.11	-0.62
D_{50}					1.00	0.36	0.74	0.47	0.05
D_{10}						1.00	0.47	-0.15	0.71
Cu							1.00	0.45	0.36
FM								1.00	-0.21
$Log_{10}k_l$									1.00

Table 5 Relative Importance of different inputs as per Garson's algorithm

Inputs	Relative importance	Ranking of inputs as per relative importance
C	4.86%	7
Gr	11.07%	4
LL	33.25%	1
PL	6.34%	5
D_{50}	3.15%	8
D_{10}	14.97%	3
Cu	21.32%	2
FM	5.03%	6

vary with the methods used. So there is need for further study in this respect to interpret the connection weights for finding the important input variables.

ANN MODEL EQUATION FOR THE HYDRAULIC CONDUCTIVITY BASED ON TRAINED NEURAL NETWORK

After the ANN is trained, a model equation can be established with the weights as the model parameters. The mathematical equation relating the input variables and the output can be written as,

$$Log_{10}k_{ln} = f_{sig}\left\{b_0 + \sum_{k=1}^{h}\left[w_k \times f_{sig}\left(b_{hk} + \sum_{i=1}^{m}w_{ik}X_i\right)\right]\right\}$$

(5)

where, $Log_{10}k_{ln}$ is the normalized (in the range -1 to 1 in this case) $Log_{10}k_l$ value,

b_0 = bias at the output layer;

w_k = connection weight between k^{th} neuron of hidden layer and the single output neuron; b_{hk} = bias at the kth neuron of hidden layer;

h = number of neurons in the hidden layer;

w_{ik} = connection weight between i^{th} input variable and k^{th} neuron of hidden layer;

X_i = normalized input variable i in the range [-1, 1] and

f_{sig} = sigmoid transfer function.

Using the values of the weights and biases tabulated in Table 3 the following expression can be written to finally arrive at a correlation of hydraulic conductivity of soils with the input parameters.

$A_1 = -0.012 + 0.1081 \, C + 0.04 \, Gr - 0.5907LL - 0.0783PL - 0.0257D_{50} + 0.4384D_{10} - 0.4367Cu - 0.1341 \, FM$

(6)

$A2 = 0.2984 + 0.0913 \, C + 0.4699 \, Gr - 0.8136LL - 0.1988PL - 0.1157D_{50} - 0.1475D_{10} - 0.4484Cu + 0.0664 \, FM$

(7)

$$B_1 = 1.0167 \times \frac{e^{A_1} - e^{-A_1}}{e^{A_1} + e^{-A_1}}$$

(8)

$$B_2 = 1.012 \times \frac{e^{A_2} - e^{-A_2}}{e^{A_2} + e^{-A_2}}$$

(9)

$$C_1 = 0.2067 + B_1 + B_2$$

(10)

$$Log_{10}k_{ln} = \frac{e^{C_1} - e^{-C_1}}{e^{C_1} + e^{-C_1}}$$

(11)

The $Log_{10} k_{ln}$ value as obtained from Eq. 11 is in the range [-1, 1] and this needs to be denormalizeld as,

$$Log_{10} k_l = 0.5 (Log_{10} k_{ln} +1) (Log_{10} k_{l\,max} - Log_{10} k_{l\,min}) + Log_{10} k_{l\,min} \qquad (12)$$

where, $Log_{10} k_{l\,max}$ and $Log_{10} k_{l\,min}$ are the maximum and minimum values of $Log_{10} k_l$ respectively in the data set.

CONCLUSIONS

The following conclusions can be drawn from the above studies:

(1)   The generalization of the model could be improved by early stopping criterion. The overfitting ratio found to vary from 1.02 to 1.3 for different models considered here except Model -4.

(2)   The developed ANN model is found to be more efficient compared to available statistical models based on R, R^2 and RMSE both for training and testing data. The ANN model was also found to be better than GMDH as per overfitting ratio.

(3)   Sensitivity analysis using cross correlation method, Garson's algorithm approach and GMDH reveal that liquid limit (LL) is the most important parameter for hydraulic conductivity of clay liners. The ranking of other input parameters are found to be different for different sensitivity methods.

(4)   A model equation is presented for prediction of hydraulic conductivity of soil, with the weights and biases obtained from the ANN model.

REFERENCES

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

Benson, C.H., Zhai, H., and Wang, X. (1994). Estimation of hydraulic conductivity of compacted clay liners. J. Geotech. Engrg. ASCE, 20 (2): 366–387.

Carrier ,W. D. (2003). Goodbye, Hazen; Hello, Kozeny-Carman. J. Geotech. and Geoenv Engrg. ASCE, 129(11) : 1054–1056.

Carrier, W.D., and Beckman, J.F. (1984). Correlation between test and the properties of remolded clays. Geotechnique, 34(2): 211-228.

Chapuis, R.A., and Aubertin, M.(2003). On the use of the Kozeny-Carman equation to predict the hydraulic conductivity of soils. Can. Geotech. Journal, 40(3): 616-618.

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

Das, S.K., and Basudhar , P.K. (2005). Prediction of coefficient of lateral earth pressure using artificial neural networks. Electronic Journal for geotechnical engineering, 10(1).

Demuth, H., and Beale, M. (2000). Neural Network Toolbox. The MathWorks Inc.USA.

Dolenko S.A., Orlov Yu.V., & Persiantsev I.G. (1996), Practical implementation and use of Group Method of Data Handling (GMDH): Prospects and Problems, *Proceedings of ACEDC'96*, PEDC, University of Plymouth, UK

Garson, G. D. (1991). Interpreting neural-network connection weights. Artificial Intelligence Expert, 6(7): 47–51.

Goh, A. T. C. (1994). Seismic liquefaction potential assessed by neural network. J. Geotech. Engrg. ASCE, 120(9) : 1467-1480.

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

Guyon, I., and Elisseeff, A. (2003). An Introduction to variable and feature selection. Journal of Machine learning Research. 3 : 1157-1182.

Lambe, T.W., and Whitman, R.V. (1976). Soil Mechanics SI version. John Wiley & Sons, Inc., New York.

Madala, H. R., Ivakhnenko, A. G. (1994). Inductive Learning Algorithms for Complex Systems Modelling, CRC Press Inc., Boca Raton.

Maier, H.R., and Dandy, G.C. (2000). Neural networks for prediction and forecasting of water resource variables: a review of modeling issues and applications. Environemental Modeling software. 15: 101-123.

MathWork Inc. (2001). Matlab User's Manual. Version 6.5, Natick, MA: MathWork Inc.

McBratney, A.B., Minasny, B., Cattle, S.R., Vervoort, R.W. (2002). From pedotransfer function to soil inference system. Geoderma, 109 : 41-73.

Shahin, M. A., Jaksa, M. B., and Maier, H. R. (2001). Artificial neural network applications in geotechnical Engineering. Australian Geomechanics, 36(1): 49-62.

Shahin, M. A., Maier, H. R., and Jaksa, M. B. (2002). Predicting settlement of shallow foundations using neural network. J. Geotech. and Geoenv Engrg. ASCE, 128(9) : 785-793.

Wang, X., and Huang, C. (1984). Soil Compaction and permeability prediction models. J. Geotech. Engrg. ASCE, 110 (6): 1063–1083.

Wilby, R. L., Abrahart, R. J., and Dawson, C. W (2003). Detection of conceptual model rainfall-rainoff process inside an artificial neural network. Hydrological Sciences, 48(2): 163-181.

Wosten, J.H.M., Pachepsky, Ya, A., and Rawls, W.J. (2001). Pedotransfer function: bridging the gap between available basic soil data and missing soil hydraulic characteristics. Journal of Hydrology. 251: 123-150.

Zaheer, I. and Bai, C. G. (2003). Application of artificial neural network for water quality management. International Journal of Lowland Technology, 5(2): 10-15.