 ISBN: 978-93-86171-25-2

PREDICTION OF CONSOLIDATION CHARACTERISTICS OF CLAYS USING MACHINE LEARNING TECHNIQUES

Harsh Deep Prasad

M. Tech Student, Department of Civil Engineering, NIT Kurukshetra, Haryana, (India)

ABSTRACT

Compression index (C_c) and coefficient of consolidation (C_v) values are one of the most important engineering properties of soil. These characteristics are frequently used in design of foundations for the estimation of settlement in clay layers. These characteristics can be determined in laboratory by carrying out oedometer test, however, the test itself is very time consuming and involves lots of calculations. In the present study, various correlations between the index properties and consolidation parameters have been reviewed. Also, the various soil index parameters that influence the consolidation characteristics of soil have been studied using various machine learning techniques. Techniques such as Multiple linear regression (MLR), Artificial neural network (ANN), Support vector machines (SVM), Forrest and M5 Trees were studied. As per the result of the observation these techniques have been proven to be more useful and efficient in the determination of consolidation parameters than other conventional methods.

Keywords*: *Consolidation characteristics, Index properties, Machine learning Techniques, Settlement.

I. INTRODUCTION

One dimensional consolidation test is generally conducted to forecast the settlements expected due to primary consolidation of highly compressible soil layers. The amount and rate of settlement of a structure founded on clays are expressed in terms of Compression Index (C_c) and Coefficient of Consolidation (C_v) respectively.

1.1 Compression Index (C_c)

Settlement occurs when a structure is founded on a compressible soil layers. The amount of settlement is related to the compression index (C_c). Settlement can inflict damage to super structure as well as domestic utilities such as gas pipe lines, electric cables, sanitary fittings and water pipelines. Hence it is very important to calculate the consolidation settlement of the normally consolidated and over consolidated saturated fine grained soils. Compression Index plays and important role in secondary consolidation settlement under expected stresses.

It represents the slope of a straight-line portion of the e vs $\bar{\sigma}$ curve and can be calculated using the following relationship:

 ISBN: 978-93-86171-25-2

$$Cc = \frac{\Delta e}{\log\left(\frac{\sigma_2}{\sigma_1}\right)} \quad (1)$$

where C_c = Compression Index, e = change in void ratio under varying effective stresses ( ).

1.2 Coefficient of Consolidation (C_v)

Coefficient of consolidation is used to predict the rate of settlement in the range of primary consolidation. Information about the rate at which compression of the soil layer takes place is important for design considerations. This parameter assumes more importance in pre-loading techniques for ground improvement. It is used to predict the time required to achieve a desired degree of consolidation for a given clay layer. The time required for settlement to occur during the life span of the structure is an important consideration. It gives an idea about the settlement a structure will undergo after it is constructed, and whether such a settlement will impair its functioning or not. For assessing time rate of consolidation of statically compacted specimens, coefficient of consolidation can be calculated using Taylor's square root of time fitting method or Casagrande's logarithm of time fitting method.

The test itself is very time consuming and requires skilled supervision and in depth knowledge. Moreover, the calculation of compressibility parameters is based on graphical methods that directly depends on personal experience. The accuracy of these parameters is also influenced by the quality of samples used in the tests. Due to these limitations, many researchers tried to develop alternative solutions. In the present study, several correlations and techniques developed by various investigators have been reviewed.

II. MACHINE LEARNING TECHNIQUES

2.1 Regression Analysis

In statistical modeling, regression analysis is a statistical process for estimating the relationships among variables. It includes many techniques for modeling and analyzing several variables, when the focus is on the relationship between a dependent variable and one or more independent variables (or 'predictors'). Regression analysis is widely used for prediction and forecasting, where its use has substantial overlap with the field of machine learning. Regression analysis is also used to understand which among the independent variables are related to the dependent variable, and to explore the forms of these relationships. Several relationships between the compressibility parameters and the basic soil properties have been developed over the years. Many different correlations based on multiple linear regression have been proposed for the determination of compression index (C_c) and coefficient of consolidation (C_v) of soil by the researchers.

2.1 Artificial Neural Network (Ann)

An artificial neural network (ANN), usually called neural network (NN), is a mathematical model or computational model that is inspired by the structure and/or functional aspects of biological neural networks. A neural network consists of an interconnected group of artificial neurons, and it processes information using a connectionist approach to computation. They are the calculating or operating systems which consists of a large number of interconnected simple processors. Process elements of the ANN are non-linear circuits which are also

 ISBN: 978-93-86171-25-2

called nodes. Every node can have numerous input connections. However, there should be only one connection as their output. ANNs are divided into subsets which include neurons and are named as layers. Input layer is a layer which has the information come from the external world to ANN. In this layer, process elements transfer information to the hidden layers as receiving from the external world. Incoming information from the input layer are processed in the hidden layer and forwarded to the output layer. The number of hidden layers can be changed according to the neural structure. The increase in the number of neurons in the hidden layer boosts the complexity and calculation time. Nevertheless, this structure also enables the use of ANN in solving more complex problems. Output layer is the layer that produces output that corresponds to the data from the input layer of the network by processing information from hidden layers.

2.2 Support Vector Machine (Svm)

In machine learning, support vector machines (SVMs, also support vector networks) are supervised learning models with associated learning algorithms that analyze data used for classification and regression analysis. SVMs have originated from the concept of statistical learning theory pioneered by Boser et al. In machine learning, support vector machines (SVMs, also support vector networks) are supervised learning models with associated learning algorithms that analyze data used for classification and regression analysis. Training, testing and sensitivity analysis of SVM has been carried out using the SVM toolbox in WEKA, MATLAB Gunn. An SVM model is a representation of the examples as points in space, mapped so that the examples of the separate categories are divided by a clear gap that is as wide as possible. More formally, a support vector machine constructs a hyperplane or set of hyperplanes in a high- or infinite-dimensional space, which can be used for classification, regression, or other tasks. In addition to performing linear classification, SVMs can efficiently perform non-linear classifications.

2.3 Random Forest

Tree based learning algorithms are considered to be one of the best and mostly used supervised learning methods. Tree based methods empower predictive models with high accuracy, stability and ease of interpretation. Unlike linear models, they map non-linear relationships quite well. Random forests or random decision forests are an ensemble learning method for classification, regression and other tasks, that operate by constructing a multitude of decision trees at training time and outputting the class that is the mode of the classes (classification) or mean prediction (regression) of the individual trees. It works for both categorical and continuous input and output variables. In this technique, we split the population or sample into two or more homogeneous sets (or sub-populations) based on most significant splitter / differentiator in input variables. The first algorithm for random decision forests was created by Tin Kam Ho. The use of this algorithm is unexcelled in accuracy and runs efficiently on large data bases. It is an alternative and fairly new pattern recognition tool that can also be considered.

2.4 M5 Tree

Regression Trees is a term introduced by Leo Breiman to refer to Decision Tree algorithms that can used for classification or regression predictive modeling problems. Tree-based regression models are known for their simplicity and efficiency when dealing with domains with large number of variables and cases. Regression trees

are obtained using a fast divide and conquer greedy algorithm that recursively partitions the given training data into smaller subsets. The use of this algorithm is the cause of the efficiency of these methods. However, it can also lead to poor decisions in lower levels of the tree due to the unreliability of estimates based on small samples of cases. Methods to deal with this problem turn out to be nearly as important as growing the initial tree. Despite their advantages regression trees are also known for their instability. A small change in the training set may lead to a different choice when building a node, which in turn may represent a dramatic change in the tree, particularly if the change occurs in top level nodes. Spite of this drawback, regression trees do not assume any particular form for the function being approximated thus being a very flexible regression method. Moreover, the obtained models are usually considered easily comprehensible.

III. REVIEW OF LITERATURE

3.1 Regression Analysis

Burland and Burbridge (1985) carried out analysis of over 200 records of settlement of foundations, tanks and embankments on sands and gravels. A remarkably simple picture has emerged relating the settlement to the bearing pressure, the breadth of loaded area and the average SPT blow count or cone resistance over the depth of influence. The influence of a number of factors such as shape and depth of foundation, depth of water table, grain size and time have been investigated. The study briefly describes the application of the results in prediction of settlement with particular emphasis on the limits of accuracy [1].

Kim et al (2005) carried out the prediction of relative crest settlement of concrete-faced rockfill dams (CFRD) using regression analysis from 30 databases of field data from seven countries (of which 21 were used for training and 9 for testing). It demonstrated that the model was capable of predicting accurately the relative crest settlement of CFRDs and is potentially applicable for general usage with knowledge of the three basic properties of a dam (void ratio, e; height, H; and vertical deformation modulus, E_V) [2].

Yildirim and Gunaydin (2011) estimated the California Bearing Ratio (CBR) of soils from different parts of Turkey using regression analysis. He concluded that, the correlation equations obtained as a result of regression analysis are in satisfactory agreement with the test results and recommended that the proposed correlations will be useful for preliminary design of project where there is a financial limitation and limited time. It is evident in literature that the prediction of compression index with regression analysis has proved to be successful and widely accepted [3].

Abasi et al (2012) used regression analysis to predict the compression behaviour of normally consolidated fine grained soil and concluded that, the proposed empirical models predict the compression index accurately in comparison with the existing equations [4].

3.2 Artificial Neural Network (Ann)

Sivakugan et al. (1998) explored the possibility of using neural networks to predict the settlement of shallow foundations on granular soils. A neural network was trained with five inputs representing the net applied

pressure, average blow count from the standard penetration test, width of foundation, shape of foundation and depth of foundation. The output was the settlement of the foundation [5].

Shi et al. (1998) presented a study of neural networks for predicting settlements of tunnels. A general neural network model was trained and tested using data from the 6.5 km Brasilia Tunnel, Brazil. The study identified many factors to be used as the model inputs and three settlement parameters as the model outputs. The study surfaces a good coefficient of determination (R^2 value) for the prediction of settlement of tunnels [6].

Shahin et al. (2000) carried out similar work for predicting the settlement of shallow foundations on cohesionless soils using the ANN technique. In this work, 272 data records were used for modelling. The input variables considered to have the most significant impact on settlement prediction were the footing width, the footing length, the applied pressure of the footing and the soil compressibility [7].

Solanki (2011) studied compressibility characteristics of locally available highly plastic clay from Surat city and Suda, treated with different percentages of rice husk using the machine learning technique namely ANN and found that the coefficient of consolidation has a better result with the plasticity index (P.I). It is inferred from his work that coefficient of correlation depends on a varying set of parameters [8].

Kurnaz et al (2016) predicted the compressibility parameters namely compression index and coefficient of consolidation of soils using artificial neural network. For this purpose, input parameters were selected as natural water content, initial void ratio, liquid limit and plasticity index. The data set consisted of 247 laboratory oedometer and index test results of fine grained soils obtained from different locations of Turkey. It has been shown that the ANN model is successful in prediction of compression index and values obtained by this technique is very close to the values obtained from the laboratory [9].

3.3 Support Vector Machine (Svm)

Wang et al (2011) predicted the subgrade settlement based on SVM and BP neural network technique. By comparing with the traditional forecasting algorithms and BP neural networks, the results showed that SVM can obtain high prediction precision and good generalization capability in few training samples comparing to other algorithms and provided a more secure and reliable solution for ballast less track settlement [10].

Pijush Samui (2012) used Support Vector Machine for the determination of Settlement of Shallow Foundation on Cohesionless Soils. It provided much sparser regressors without compromising performance, and kernel bases give a small but worthwhile improvement in performance. The study showed that compared to the available methods for determining the settlement, SVM was better at determining the settlement of shallow foundation on cohesionless soils [11].

Yu and Shangguan (2016) predicted the settlement of road foundation using support vector machine and BP neural networks. This case study based on road foundation engineering project showed that the forecasted results were in good agreement with the measured data. The results indicated that that it was effective and feasible to use this method for estimation of foundation settlement and its influence factor could be expressed well. Therefore, settlement prediction based on SVM model reflected actual settlement process more correctly than the other model [12].

 ISBN: 978-93-86171-25-2

3.4 Random Forest (Rf)

Ocak and Seker (2013) carried out the calculation of surface settlements caused by EPBM tunneling using machine learning techniques including RF technique. In this study, 18 different parameters have been collected by municipal authorities from field studies pertaining to EPBM operation factors, tunnel geometric properties, and ground properties. The data source has a pre-process phase for the selection of the most effective parameters for surface settlement prediction. This paper focuses on surface settlement prediction using three different methods: artificial neural network (ANN), support vector machines (SVM), and Random Forest (RF). The success of the study has decreased the standard error in the work which was relatively higher in the earlier contemporary research [13].

Zhou et al (2016) check the feasibility of random forest approach to prediction of ground settlement induced by the construction of shield driven tunnel. To achieve this goal, tunnel geometry, geological properties, and construction parameters were investigated as input variables to utilize in the RF modelling, resulting in the maximum surface settlement value (S_{max}) and trough width (i) as the ground surface settlement index. A fivefold cross-validation procedure was then applied to identify the optimal parameter values during modelling, and an external testing set was employed to validate the prediction performance of the model. Two performance measures, R^2 and RMS error, were employed. The relative importance of different parameters in the prediction of ground settlements was also investigated. Findings demonstrate that the RF method provides promising results and offers an alternative means in predicting ground settlements induced by tunnelling [14].

3.5 M5 Tree

Heshmati et al (2013) predicted the compression ratio for municipal solid waste using the M5 tree technique. A reliable database was retrieved from the literature was used to develop a practical model that relates Cc to waste composition and properties, including dry density, dry weight water content, and percentage of biodegradable organic waste using the decision tree method. The performance of the developed model was examined in terms of different statistical criteria, including correlation coefficient, root mean squared error, mean absolute error and mean bias error. The obtained results demonstrated that the suggested model was able to evaluate the compression ratio of municipal solid waste effectively [15].

 ISBN: 978-93-86171-25-2

Table 1. Various Correlations of C_c and C_v with Index Properties of Soil

Author(s) Name	Correlation		Soil Type
	Compression Index (C_c)	Coefficient of consolidation (C_v) in m^2/sec	
Terzaghi and Peck (1956)	$C_c = 0.009\ (w_L - 10\%)$	-	CL-CM
Nishida (1956)	$C_c = 0.54\ (e_o - 0.35)$	-	Clay
Azzouz (1976)	$C_c = 0.4\ (e_o - 0.25)$	-	Clay
Rendon-Herrero (1980)	$C_c = 0.3\ (e_o - 0.27)$	-	Clay
Koppula Wroth Model (1981)	$C_c = 1.325\ PI$	-	Clay
Carrier (1985)	-	$C_v = 9.09^{*}/(1.192 + ACT^{-1})^{0.999}$ ($IL + 1)^{3.46}$)	Clay
Serajjudin (1987)	$C_c = 0.0102\ (W_n - 9.15)$	-	Alluvial Silt & Clay
Skempton (1994)	$C_c = 0.007\ (w_L - 10\%)$	-	Clay
Raju et al (1995)	-	$C_v = 1 + e_L\ (1.23 - 0.276\ \tau v\ )\ *\ \frac{}{}$	Clay
Sridharan and Nagaraj (2000)	$C_c = 0.014\ (PI + 3.6)$	-	Clay
Sridharan and Nagaraj (2000)	-	$C_v = \frac{}{}$	Clay
K N Prasad and Achintya (2013)	$C_c = 0.0124\ I_p$	-	Silty Clay
Binu Sharma and Padma K Bora (2015)	$C_c = G_s$ *	-	Clay
Guruprasad Jhadav (2016)	-	$C_v = 0.1105/(LL + SL) - 0.0009$	Clay
Guruprasad Jhadav (2016)	-	$C_v = (5.4 * PL)/I_s * 3.54 + 0.0002$	Clay

IV. CONCLUSION

This paper provides a review of correlation of consolidation parameters of soil using its index properties by various machine learning techniques. Different research papers presented the use of different index properties for the determination of consolidation parameters. Liquid limit and void ratio were the major used parameters among the rest of the index properties for the development of correlations. Several machine learning techniques namely Linear Regression, ANN, SVM, RF and M5 Tree have been studied for the determination of the correlations. Different techniques have given better results for compression index and coefficient of consolidation. The result calculated using the several techniques and the observed compression index and coefficient of consolidation values were found to be close. This study also concludes that these machine learning techniques offer distinct advantages over conventional hand calculation and laboratory tests. Further researchers are recommended to observe the influence of other soil properties on the compression index and coefficient of consolidation values.

 ISBN: 978-93-86171-25-2

REFERENCES

[1.] Burland, J.B., Burbidge, M.C., Wilson, E. J., and Terzaghi, K.V., Settlement of foundations on sand and gravel, Proceedings of the Institution of Civil Engineers-Part 1, 1985, 18(6), 1985, 1325-1401.

[2.] Kim, Y.S., Kim, M.W., Ha, I.S. and Olson, S.M., Prediction of relative crest settlement of concrete-faced rockfill dams using linear regression, Journal of Geotechnical and Geo Environmental Engineering, 135(8) 2005.

[3.] Yildirim, A.S and Gunaydin I.K, Prediction of California bearing ratio (CBR) of fine grained soils by regression analysis, Journal of Traffic and Transportation Engineering, 2(2), 2011, Pages 107–115.

[4.] Abasi, S.M., Nejad, S.M.V., Khavari, E., Regression analysis of soil compressibility, Scientific research and essays 7.31, Soils and Foundation (2), 2012, pp 19-29.

[5.] Sivakugan, N. and Das, B.M., Settlement Predictions Using Neural Networks, Australian Civil Engineering Transactions, Vol. 40, 1998: 49-52.

[6.] Shi, P., Qiu, J., Chen J. and Feng, Z., Prediction of Settlement in Tunnelling Using Artificial Neural Networks, Proceedings of the 2nd International Conference on Soil Mechanics and Foundation Engineering, 2(1), 1998, pp 561-596.

[7.] Shahin, M.A., Jaksa, M.B. and Maier, H.R., Predicting the Settlement of Shallow Foundation on Cohesionless Soil Using Back Propagation Neural Network, Soils and Foundations, 55(4), 2002, Pages 737–748.

[8.] Solanki, A., Study on Consolidation and Correlation with Index Properties of Different Soils in Surat City, International Journal of Engineering Research and Development, 11(5), 2011, pp 57-63.

[9.] Kurnaz, T.F., Dagdeviren U., Yildiz M. and Ozkan O., Prediction of compressibility parameters of the soils using artificial neural network, Springer Plus, 2016; 18:6135–6144.

[10.] Wang L., Man C., Wang W. and Zhao J., Subgrade settlement prediction based on Support Vector Machine, Strategic Technology (IFOST), 2011 6th International Forum, Vol.1, pp. 113-116.

[11.] Samui P and Sitharam T.G., Least-square support vector machine applied to settlement of shallow foundations on cohesionless soils, International Journal for Numerical and Analytical Methods in Geomechanics, 32(17), 2012, Pages 2033–2043.

[12.] Yu H. and Shangguan Y., Settlement Prediction of Road Soft Foundation Using a Support Vector Machine (SVM) Based on Measured Data, International Symposium on Materials Application and Engineering (SMAE 2016), Volume 67, 2016.

[13.] Ocak A. and Seher L.K., Selection of Tunnelling Method and Parameters Effecting Ground Settlements, Electronic Journal of Geotechnical Engineering, 2013, ISSN: 2248-9622.

[14.] Zhou J., Shi X., Du K. and Qui X., Feasibility of Random-Forest Approach for Prediction of Ground Settlements Induced by the Construction of a Shield-Driven Tunnel, International Journal of Geomechanics 40(11), 2016, 4653–4659.

[15.] Heshmati K., Mokhtari M. and Shariatmadari N., Compression ratio of municipal solid waste simulation using artificial neural network and adaptive neurofuzzy system, Earth Sciences Research Journal, 2013.