FEDERAL UNIVERSITY OF TECHNOLOGY, AKURE
JOURNAL OF ENGINEERING AND ENGINEERING TECHNOLOGY

ASSESSMENT OF SELECTED MACHINE LEARNING SYSTEMS IN PREDICTING THE RESILIENT MODULUS OF COHESIVE SOILS

Adebajo, O. M., Oluyemi-Ayibiowu, B. D.* and Falola, K. E.

Department of Civil and Environmental Engineering, Federal University of Technology Akure, PMB, 704, Ondo State, Nigeria

**Correspondence author: bayibiowu@yahoo.com*

Adebajo, O. M., Oluyemi-Ayibiowu, B. D., Falola, K. E. (2025): Assessment of Selected Machine Learning Systems in Predicting the Resilient Modulus of Cohesive Soils. *FUTA Journal of Engineering and Engineering Technology /19(2), 74-79.*

Received Date: 28.03.25 **Accepted Date: 02.10.25**

Abstract
Characterizing subgrade soils in terms of resilient modulus (M_R) is crucial for pavement design. However, the process can be expensive and time-consuming, leading to the need for more efficient alternatives. This research investigates the effectiveness of using machine learning systems to predict the resilient modulus of cohesive subgrade soils. A dataset of 400 resilient modulus measurements obtained from in-situ cone penetration tests waere collected, along with an additional 200 prediction data points from past research. The combined dataset were then divided into training and testing subsets and used to develop two machine learning predicting systems: Support Vector Machine (SVM) and Back Propagation Neural Network (BPNN). The performance of these systems was evaluated using the sum of residuals (R) and the Mean Square Error (MSE) performance index. Among the SVM models tested, the one trained with a medium Gaussian kernel demonstrated the best predictive capability. For the BPNN, the most effective configuration included four input variables, fifteen hidden neurons, and thirty-three epochs. The BPNN system outperformed the SVM system, yielding higher predictability with training and testing R values of 0.98 and 0.92, and MSE training and test values of 19.84 and 90.16, respectively. In contrast, the SVM system produced training and testing R values of 0.92 and 0.72, with MSE training and test values of 47.75 and 223.73. This research demonstrates that machine learning systems can accurately predict the resilient modulus of soils, providing acceptable error levels for pavement design and construction.

Keywords: *Resilient modulus, support vector machine, machine learning, neural network, predictions, cone penetration test*

Introduction

When subjected to moving loads, the underlying soil layers exhibit two types of deformation: resilient (recoverable) and plastic (irrecoverable or permanent). These deformations contribute to distress in the pavement structure above. This distress typically manifests as cracking and rutting, which adversely affect both the functional condition and structural integrity of flexible pavements.

In designing pavement layers, two key criteria are considered, which are fatigue cracking at the bottom of the surface layer and permanent deformation at the surface of the subgrade soils. Fatigue cracking is closely related to the resilient behavior of the pavement materials in response to traffic loading. Therefore, it is essential to thoroughly understand the resilient response of foundation soils under moving loads, as this knowledge can inform more effective foundation design (Sun *et al.*, 2017).

The American Association of State Highway and Transportation Officials (AASHTO) introduced the modulus of resilience (M_R) of subgrade soils to characterize the stress-strain behavior (i.e., resilient response) of subgrade soils under normal cyclic traffic loading. This parameter is crucial for estimating transient soil deformation under repeated traffic loads (Yang *et al.*, 2017).

The practical assessment of resilient modulus is often costly, time-intensive, and not readily accessible to all agencies. Furthermore, due to constrained budgets and inadequate planning, especially in developing countries, many cases suffer from a lack of sufficient soil investigation data. This situation is frequently exacerbated by engineers relying on their experience to estimate the M_R value rather than conducting the necessary tests (Oluyemi-Ayibiowu and Omomomi, 2021). Therefore, there is an urgent need for the development of predictive models that are both quicker and more economical, utilizing readily determinable parameters that can be applied across various subgrade soil categories. Such models could provide a solid basis for evaluating the validity of M_R values for subgrade soils. This study seeks to investigate the use of Back Propagation Neural Networks (BPNN) and Support Vector Machines (SVMs) as machine learning techniques for predicting the resilient modulus of subgrade soils.

Materials and Method

Data Collection

A total of four hundred (400) soil samples were collected from forty (40) randomly selected locations in the southern regions of Akure State, with each location contributing ten samples. The predominant soil types in these areas include lateritic soil, pure clay, silty clay, and sandy clay. Laterite soil exhibits characteristics of both cohesive and non-cohesive soils, largely depending on the particle size. Silt and clay-based lateritic soils demonstrate cohesive behavior, while sand-sized lateritic soils exhibit non-cohesive behavior (Oyelami and Van Rooy, 2018). For the purposes of this study, only the clay-silt-based lateritic soils were utilized.

The samples were collected from a depth of 1.5 m using an undisturbed sampling technique. Geotechnical property tests were performed on these samples, and in-situ cone penetration tests were conducted at the sites. Furthermore, to strengthen the data's robustness, two hundred (200) additional sample data points were gathered from a literature review of previous research on similar tests conducted on cohesive soils globally.

Experimental Procedure

Laboratory tests were conducted to determine the moisture content (w), dry density (δ_d), and resilient modulus (M_R). The resilient modulus test followed the AASHTO T 307 standard (2003).

For the cone penetration test (CPT), a cylindrical cone penetrometer with a circular area of 10 cm² and a tip angle of 60° was utilized. The depths of the groundwater table (GWT), which varied between 0.4 m and 4.5 m at the test sites, were recorded after the CPT test. To predict the resilient modulus (M_R) under in-situ stress, CPT tests were performed at locations adjacent to the boreholes from which soil specimens were extracted. The horizontal distances between the boreholes and the CPT test sites for each data point were kept to less than 2 m.

(i) Artificial Neural Network Methodology

For the development of the Artificial Neural Network (ANN) model, a total of 600 datasets were collected and divided into three groups. The first group contained 420 datasets (70 % of the total) which were used for training the ANN model. The second group consisted of 90 datasets (15%) that were used for testing, while the remaining 90 datasets (15 %) were utilized to validate the accuracy of the developed ANN model. The Levenberg-Marquardt (LMNN) algorithm was employed to train the data within the ANN model.

The input variables for the ANN model included w (%), δ_d (kN/m³), fs (MPa), and q_c (MPa), while the output variable was the resilient modulus (M_R in MPa). The ANN architecture used was a Back Propagation Neural Network (BPNN), which comprises one input layer, one or more hidden layers, and one output layer. To determine the most appropriate BPNN architecture, various configurations of neurons in the hidden layer and different BPNN networks utilizing distinct transfer functions were tested to predict the resilient modulus (M_R). Additionally, several training parameters, including the number of iterations (epochs), learning rate, error goal, and number of hidden layers were varied until satisfactory convergence of the BPNN training was achieved. The ANN toolbox of MATLAB software was used to conduct the necessary computations for ANN development.

(ii) Support Vector Machine (SVM) Methodology

The Support Vector Regression (SVR) method was utilized for this study. SVR is a type of support vector machine that aims to find a function that best predicts continuous output values based on given input values. The input and output parameters used were the same as those in the Artificial Neural Network (ANN) model, and the data divisions were consistent with those employed for ANN.

SVR employs kernels, which are functions that determine the similarity between input vectors; these can be either linear or non-linear. For this study, the radial basis function (RBF) kernel, linear kernel, and polynomial kernel were used to predict the resilient modulus. The kernel that resulted in the lowest error was selected for the final analysis. The SVM Toolbox in MATLAB was used to perform the necessary computations for developing the SVR model.

(iii) Performance Evaluation

In order to evaluate the performance of the proposed ANN and SVM models, the sum of residuals (R) and mean square error (MSE) error indices was used as the criteria between the measured and predicted values. The formula is given in Equations (1) and (2):

$$R = 1 - \sum_{i=1}^{n}(h_i - t_i) \Big/ \sum_{i=1}^{n}(h_i - \overline{h_i}) \quad (1)$$

Where h_i and t_i are, respectively, the measured output and the predicted output value for the ith output, $\overline{h_i}$ is the average of the actual outputs and n is the number of sample.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \widehat{Y}_i)^2 \quad (2)$$

Where, n is the number of data points, Y_i represents the observed values and $\widehat{Y}_i$ represent the predicted values.

Results and Discussion

Optimal SVM Predictive Performance Choice

Five different types of Support Vector Machine (SVM) models, distinguished by their kernel functions, were utilized to predict the resilient modulus of cohesive soil. Among these, the best-performing models were selected. The types of SVMs employed included quadratic SVM, cubic SVM, fine Gaussian SVM, medium Gaussian SVM, and coarse Gaussian SVM. To evaluate the predictive performance of each machine learning model, error indices were calculated based on the discrepancies between predicted and actual values. Table 5 presents the predictive performance of the SVM models used in the study.

Where: RMSE is Root mean square error, MSE is Mean square error, MAE is Mean absolute error.

For RMSE (Root Mean Square Error), values closer to one (1) indicate better predictive capability of the model. Similarly, for the R error indices, a value closer to one (1) also signifies improved predictive ability. The negative R score observed for the cubic SVM model suggests that it is under-forecasting, meaning the model's predicted values are significantly lower than the actual observed values. In contrast, for both MSE (Mean Square Error) and MAE (Mean Absolute Error), lower values indicate better predictive performance. Based on the information in Table 1, the optimal Support Vector Machine (SVM) model for estimating resilient modulus is the medium Gaussian SVM.

The SVM Prediction Regression Line

Figures 1 and 2 display the predicted and actual response plots for the training and test data. The regression line represents the model selected for estimating the resilient modulus of cohesive soils. Most data points are close to this line, indicating reasonable accuracy in estimation, as shown by the error performance indices in Table 1. However, the largest discrepancies between the data points and the regression line occur at lower resilient modulus values: 10 MPa to 40 MPa for the training set and 10 MPa to 33 MPa for the test set. This suggests that the model is less effective for lower values but performs better at higher ones. The data gathered may have skewed results, as most points reflect higher modulus values suitable for subgrade material in road construction, leaving minimal representation

Table 1: SVMs predictive performance

Error Indices	Quadratic SVM	Cubic SVM	Fine Gaussian SVM	Medium Gaussian SVM	Coarse Gaussian SVM
RMSE	40.11	15.0001	11.345	6.9105	8.5595
R	-1.85	-21322284.53	0.77	0.92	0.87
MSE	1608.8	1.20E+10	128.7	47.754	73.265
MAE	7.1939	4864.3	7.8291	4.715	5.9165

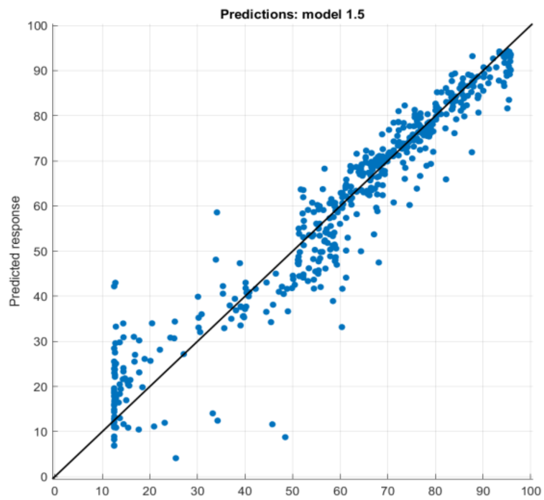

Figure 1: SVM Predictive Regression Line Plot for the training data

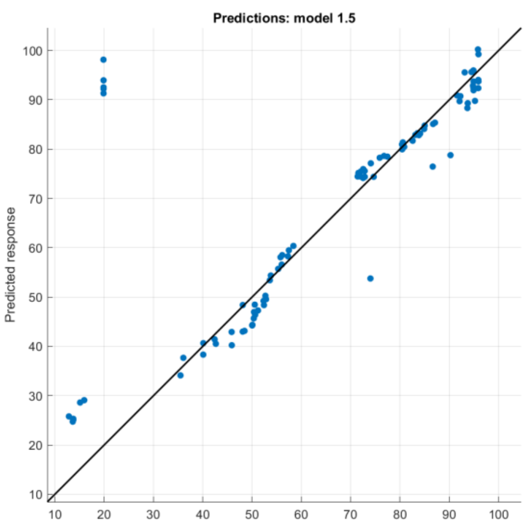

Figure 2: SVM Predictive Regression Line Plot for the test data

Table 2: SVM and BPNN predictive performance comparison

Indices	SVM	BPNN
	Training Phase	
R	0.92	0.98
Mean Square Error (MSE)	47.754	19.84
	Testing Phase	
R	0.72	0.92
Mean Square Error (MSE)	223.73	90.158

for unsuitable lower values. Geotechnical characterization found soils suitable for subgrade purposes in seventeen of the twenty study areas, indicating that the Akure South region has a higher prevalence of suitable cohesive soils.

Artificial Neural Network Prediction Performance Result

The data were divided into training data which was 420 (70%), validation data 15 (90%) and test data which was 90(15%), and was inputted into the ANN module of the machine learning application on Mathlab 2021 software. The result observed were as follows**:**

Artificial Neural Network Architecture

Figure 3 illustrates the Back Propagation Neural Network (BPNN) used for predicting the resilient modulus of cohesive soil. The performance of the neural network is influenced by the number of epochs and hidden neurons. Various combinations of these parameters were tested, and their performance was monitored until optimal results were achieved, which were then recorded. The model utilized four input variables, employed fifteen hidden neurons, and was trained over thirty-three epochs.

ANN Regression Performance Plot

Figure 4 displays the residuals for the regression line fitted to the training, validation, testing, and overall phases of the dataset. It indicates the accuracy of the regression through the R-squared value. Values above the regression line represent correct predictions, while those below indicate incorrect ones. The distance from the line measures prediction accuracy.

In the training phase, all values fell within the prediction line, leading to an accuracy of 98.30 %. In the validation phase, most values aligned closely with the regression line, except resilient modulus values between 10 MPa and 40 MPa, resulting in a slightly lower accuracy of 97.81 %. The testing phase showed an accuracy of 92.12 %, mainly due to wider gaps in predicting lower resilient modulus values.

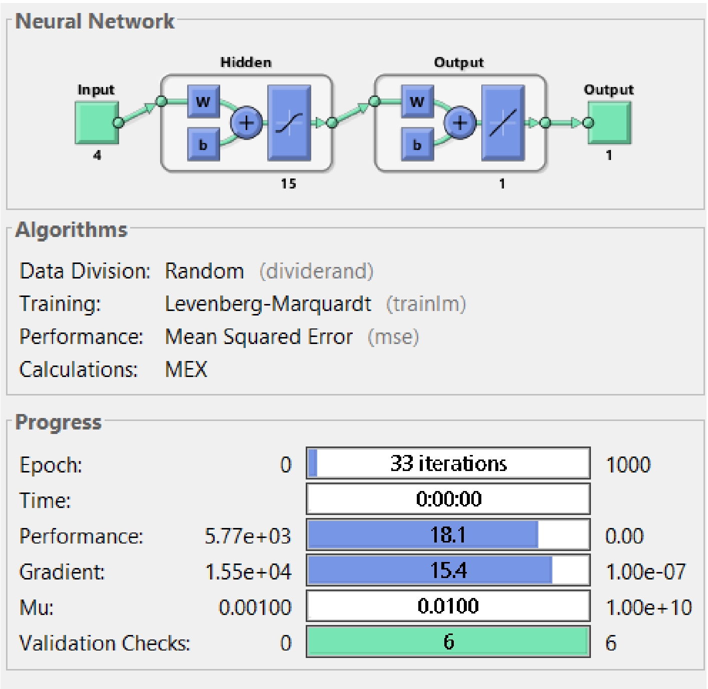

Figure 3: SVM Predictive Regression Line Plot for the test data

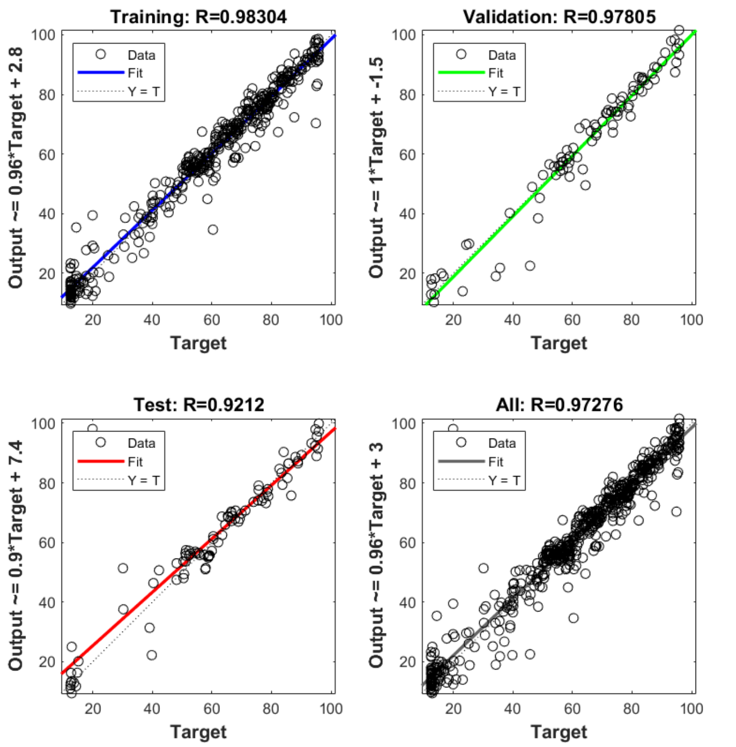

Figure 4: Resilient Modulus Prediction BPNN Regression Line Plot

When using all available data, the predictive accuracy was 97.28%, demonstrating the BPNN neural network's strong capabilities using the Levenberg-Marquardt algorithm. Similar to the SVM system, the ANN struggled with lower resilient modulus predictions, impacting accuracy in the lower data range. This may be due to the dataset's demographics, which contain more suitable soils for subgrade purposes. The study area, Akure South, therefore appears to have more cohesive soils suitable for road construction.

Comparative Analysis of the SVM and the BPNN resilient modulus prediction performance

Table 2 shows the performance indices of the coefficient of determination (R value) and mean square error (MSE) for the Support Vector Machine (SVM) and Back Propagation Neural Network (BPNN) on training and test datasets. The results indicate that the BPNN outperformed the SVM, achieving R values of 0.98 for training and 0.92 for testing, compared to the SVM's 0.92 and 0.72. Furthermore, the BPNN had lower MSE values of 19.84 for training and 90.158 for testing, while the SVM recorded MSE values of 0.72 and 223.73. These findings highlight the BPNN as a superior predictor for the resilient modulus of cohesive soils.

This conclusion contrasts with previous studies, such as that of Nasrin et al. (2021), where Support Vector Machines (SVM) demonstrated superior predictive capacity compared to Artificial Neural Networks (ANN). The differences observed in this study may stem from the specific type of ANN used, namely the Back Propagation Neural Network (BPNN). Additionally, the training algorithm employed—Levenberg-Marquardt (LMNN)—along with the ability to adjust the number of hidden neurons, epochs, and performance monitoring of the gradient function associated with the BPNN, may have provided the ANN with a distinct advantage in estimating capabilities over the Medium Gaussian SVM developed in the machine learning system.

Conclusions

The following conclusions were made from the study:

- Among the five Support Vector Machine (SVM) systems evaluated, namely: quadratic SVM, cubic SVM, fine Gaussian SVM, medium Gaussian SVM, and coarse Gaussian SVM, the medium Gaussian SVM demonstrated the best predictive performance for resilient modulus data.
- The regression line plot indicated that the SVM performed better at predicting high resilient modulus values (50 MPa and above) compared to low resilient modulus values for both training and test datasets.
- The optimal prediction model for resilient modulus using a Back Propagation Neural

Network (BPNN) consisted of four input features, fifteen hidden neurons, and one output. The best predictive model was achieved after thirty-three epochs.

- The BPNN machine learning system exhibited higher predictability, with R values of 0.98 for training and 0.92 for testing, and Mean Squared Error (MSE) values of 19.84 for training and 90.16 for testing. In contrast, the SVM system had R values of 0.92 for training and 0.72 for testing, along with MSE values of 47.75 for training and 223.73 for testing.

References

Nasrin, H., Ali, R. G. and Ali, B. (2021). Prediction of the resilient modulus of non-cohesive subgrade soils and unbound subbase materials using a hybrid support vector machine method and colliding bodies optimization algorithm. *Construction and Building Materials*, 1-14.

Oluyemi-Ayibiowu, B. D. and Omomomi, J. (2021): Predicting the California Bearing Ratio of Chemically Stabilized Expansive Soil Using Soft Computing Techniques (Case Study of the Artificial Neural Network Model). *International Journal of Scientific Research and Innovative Technology*, 8(8), 104-122.

Oyelami, C. A. and Van Rooy, J. L. (2018). Mineralogical characterisation of tropical residual soils from south-western Nigeria and its impact on earth building bricks. *Environmenntal and Earth Science, 77*(5), 178-185

Sun, L., Gu, C. and Wang, P. (2017). Effects of cyclic confining pressure on the deformation characteristics of natural soft clay. *Earthquake Engineering, 78*, 99-109.

Yang, S. R., Huang, W. H. and Tai, Y. T. (2017). Variation of Resilient Modulus with Soil Suction for Compacted Subgrade Soils. *Journal of the Transportation Research Board*, 8, 1-8.