COMPUTER-AIDED PREDICTION OF PAVEMENT CONDITION INDEX (PCI) USING ANN

M. Jalal[1], I. Floris[2] and L. Quadrifoglio[3*]

[1]Zachry Department of Civil Engineering
Texas A&M University, USA
mjalal@tamu.edu

[2]Dipartimento di Ingegneria Civile, Ambientale e Architettura
Universita' di Cagliari, Italy
ignaziofloris89@gmail.com

[3]Zachry Department of Civil Engineering
Texas A&M University, USA
quadrifo@tamu.edu

ABSTRACT

The accurate prediction of pavement network status and performance is essential for development of pavement Maintenance and Rehabilitation (M&R) Plans and competent management of the transportation infrastructure system. Hence, pavement conditions need to be monitored and evaluated properly, so that adequate conditions, and, consequently, safety and comfort can be ensured during the entire road service life. With this respect, Pavement Condition Index (PCI), an indicator of great value in pavement engineering, which rates the surface condition and the structural integrity, is used in this research.

In this paper, an optimized artificial neural network (ANN) model has been used to predict the PCI based on the experimental study carried out on Texas A&M University campus. A multi-step approach has been carried out to find the optimized ANN model to predict PCI founded on the experimental data obtained. The results show that for PCI prediction, which can be considered as a complex civil engineering and management problem, the accuracy of optimal ANN based on this approach is well above the normal ANN models.

Keywords: Pavement Condition Index (PCI), engineering and management, performance prediction, optimal neural network

1 INTRODUCTION:

Departments of Transportation (DOTs) across the United States conduct pavement condition assessments each year to determine the health status of the roadway network and program improvements such as maintenance, resurfacing, rehabilitation and reconstruction. For various types of pavements, engineers or technicians inspect the pavement periodically and record the distresses occurred in different pavement sections in terms of type, severity and extent. Then in accordance with the procedure and equations established by ASTM D6433-11 [1], all distresses observed on sample, selected in the section considered, are combined and Pavement Condition

* Corresponding Author

Index (PCI) is calculated, for each sample, in order to evaluate road network conditions. PCI ranges from zero to one hundred; with zero indicating poor condition and one hundred being the excellent condition. Hand calculation of PCI is a tedious job and since many parameters influence PCI, its prediction would be a challenging issue.

In recent years, several computer-aided data mining techniques have found their way through science and engineering to solve a variety of design, prediction and optimization problems in various practical applications, for example Fathi [2], Garmsiri [3], Jodaei [4], and Jalal [6]. Artificial neural networks (ANNs), the most widely used pattern recognition and modeling systems, have been utilized to solve various problems in civil engineering materials and infrastructures such as Jalal [7], Ashrafi [8], and Jodaei [9]. Nevertheless, fine tuning and optimizing the ANN model and data refinement can significantly improve the robustness and accuracy of the model, which has not been taken into account in most of the engineering application of ANN. Consequently, this paper focuses on finding an optimal ANN model based on a multi-step approach to come up with an efficient procedure to be used as a computer-aided prediction for PCI based on an experimental assessment done at Texas A&M University campus.

2 EXPERIMENTAL PLAN

Pavement network of interest in this study is located in the Texas A&M University Campus. The area, about 22 km^2, was divided into zones, branches, sections and samples, as defined in ASTM D6433-11.

In order to calculate the PCI, all distresses were identified, evaluated and quantified for three consecutive years as 2014- 2016. Three pavement types were included in the study as Portland Cement Concrete (PCC), Hot Mix Asphalt (HMA), and Asphalt Concrete (AC), which were considered in the model as 1, 2, and 3 respectively. Numbers assigned to years associated with the survey were 4, 5, and 6 for 2014, 2015, and 2016 respectively. Since the sampling procedure may partly influence PCI evaluation, sample size, length and width of the samples were taken as 2 variables. Annual Average Daily Traffic (AADT), linked to traffic loads, was also considered as another variable. Consequently, the problem was defined with 5 inputs and 1 output which was PCI. A total number of 173 datasets were obtained in this experiment, which were used to build the model.

3 MODELING APROACH

A multi-step procedure was employed in this study as an optimization procedure to come up with optimal ANN model. The steps can be summarized as follows:

- Performance evaluation of training functions based on error.
- Performance evaluation of training functions based on speed (computationally expensive or not).
- Selection activation or transfer functions based on error.
- Selection of optimal ANN architecture based on error.
- Data refinement based on results from optimal ANN architecture.
- Final optimal ANN model based on refined data.

3.1 Data Processing, Performance Criteria, and ANN structure

To prepare the datasets for neural network model, all data need to be normalized as a part of pre-processing step. Equation (1) was used to normalize the PCI data which is as following:

$$\frac{X-X_{min}}{X_{max}-X_{min}} \qquad \textbf{(1)}$$

Data range and statistics are also presented in table 1.

Table 1. Statistics of the data base

	Length (m)	Width (m)	Pavement Type	AADT, (vehicle/day)	Year of Survey	Avg. PCI
max	1670	30	3	23310	6	100
min	16	5	1	50	4	5
Ave	327	9	-	3426	-	67
Std	250	4	-	3132	-	18

Since pavement type and year are qualitative variables, statistical values for theses variables are not meaningful and so are not reported in the table.

To evaluate errors, MATLAB considers the Mean Square Error (MSE) as the network performance criterion. Nevertheless three performance criteria were considered for the model as mean absolute percent error (MAPE), root mean square error (RMSE) and coefficient of determination (R^2) are calculated through equations (2), (3), and (4) respectively.

$$MAPE = \frac{100}{n}\sum_i \frac{|t_i - O_i|}{t_i} \quad \textbf{(2)}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_i (t_i - O_i)^2} \quad \textbf{(3)}$$

$$R^2 = 1 - \frac{\sum_i (t_i - O_i)^2}{\sum_i (O_i)^2} \quad \textbf{(4)}$$

The network architecture used in this study is called ANN 5-***n***-***m***-1, where the first digit is the number of input nodes, ***n*** is the number of nodes in the first hidden layer, ***m*** is the number of nodes in the second hidden layer, and fourth digit is the number of output nodes as shown in Fig. 1.

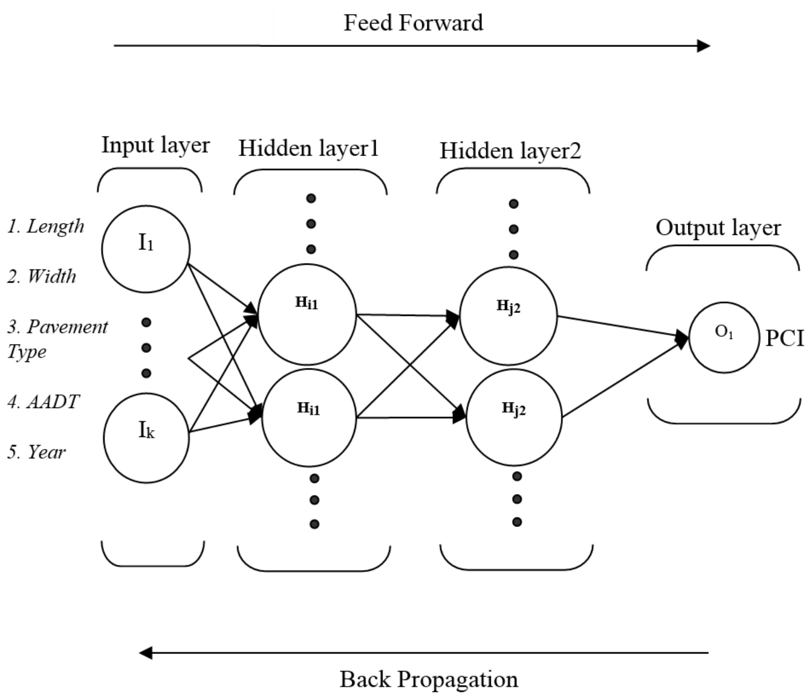


Fig.1 .schematic architecture of ANN model

Evaluation of the model performance based on error was implemented through a program coded in MATLAB. The program tries various number of hidden layers and number of neurons in the hidden layers for constant number of epochs for 10 times and selects the best NN architecture with the minimum MAPE. The flowchart of the whole process is shown in Fig. 2

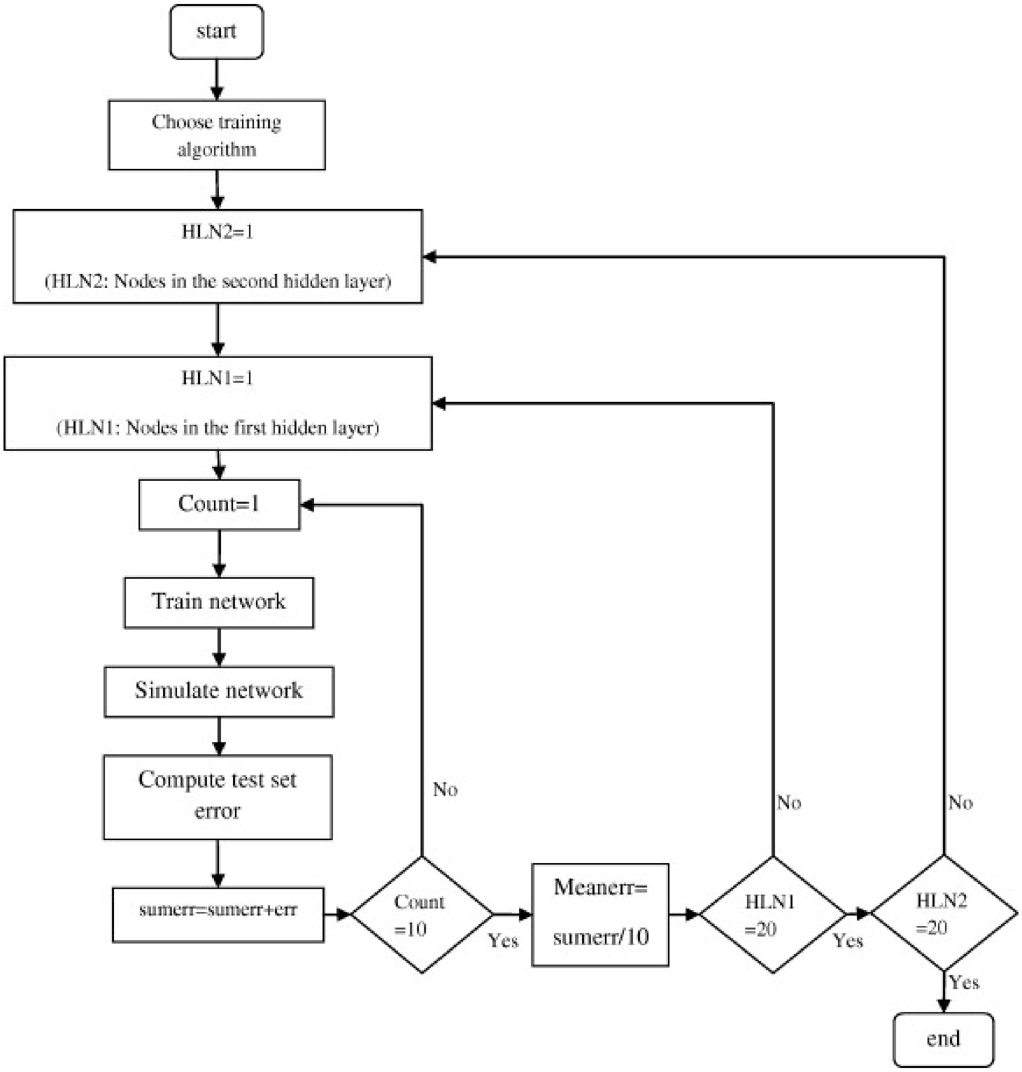
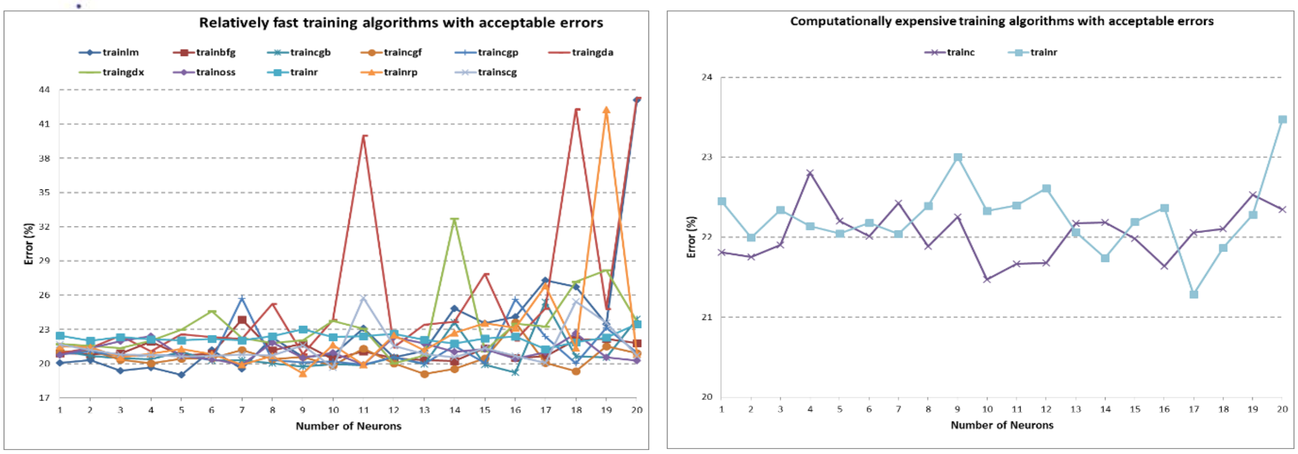


Fig.2 .Flowchart of ANN architecture selection process

3.2 Performance Evaluation Based on Error Level and Speed

For this step, an ANN structure as 5-n-1 including 5 inputs, 1 output and n neurons in the hidden layer was built. Number of neuron changes in an iterative loop from 1 up to 20 and for each structure 10 runs were implemented and MAPE was calculated. In this way, training algorithms with huge error were identified and eliminated in selection process and the functions with acceptable errors remained to be further evaluated.

Among remaining algorithms, computationally expensive algorithms were also identified to be eliminated from the model. The results of the training algorithms error for normal and computationally expensive ones are presented in fig. 3a, b.



(a) (b)

Fig. 3. Training errors for normal and computationally expensive algorithms

3.3 Selection of Optimal ANN Architecture and Transfer Function

A program was developed to evaluate the error of various ANN architecture and find the one with the highest accuracy. Either when the complexity of the data or their variation is high, an ANN architecture with two hidden layers is preferred. Various ANN architectures with two transfer functions as 'logsig' and 'tansig' were evaluated and their results are plotted in fig. 4.

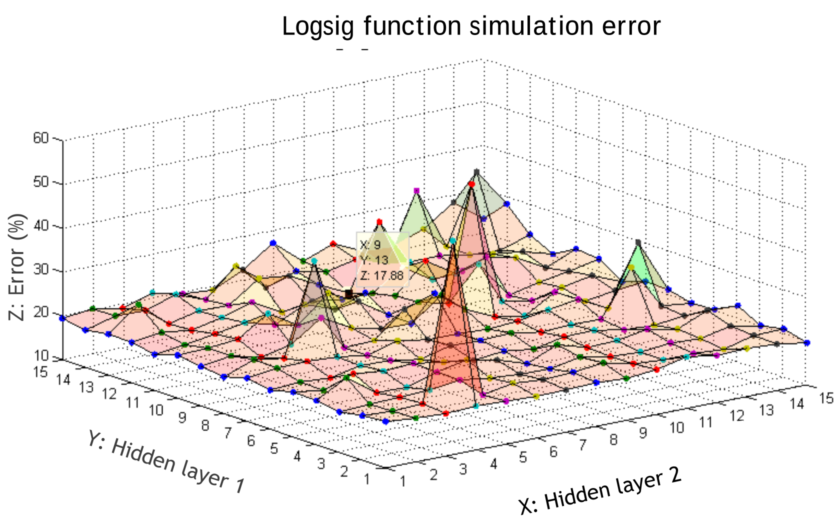


FIG. 4. Error plots of various ANN architecture

3.4 Data Refinement and Final Optimal ANN Model

For data refinement purpose, all data were modelled by optimal ANN obtained to eliminate some scattered data with high error ("outliers"). In this approach, data points with error higher than 50% were eliminated from the database, which constitute 16 datasets out of 173. So the refined database contains 157 datasets. Then all remaining data were modelled using optimal ANN architecture in this step as the final step. It is worth mentioning that modelling all refined data by optimal ANN resulted in average error (MAPE) of 9% which is approximately half of the error obtained by optimal ANN (17.88%) without data refinement. Consequently, the final ANN model with less than 10% error was selected and its results are presented in the following section.

4 RESULTS AND DISCUSSION

Modelling of the refined data was done by optimal ANN model in MATLAB in which training, validation and testing are automatically implemented and the error is evaluated based on MSE. However, for more verification, training, validation and testing of the network was also done manually and the model performance was assessed through MAPE and RMSE so that the accuracy and efficiency of the model can be ensured. Out of 157 datasets, about 70, 15, and 15% were used for training, validation and testing respectively. Performance and training state of the optimal ANN is depicted in fig. 5.

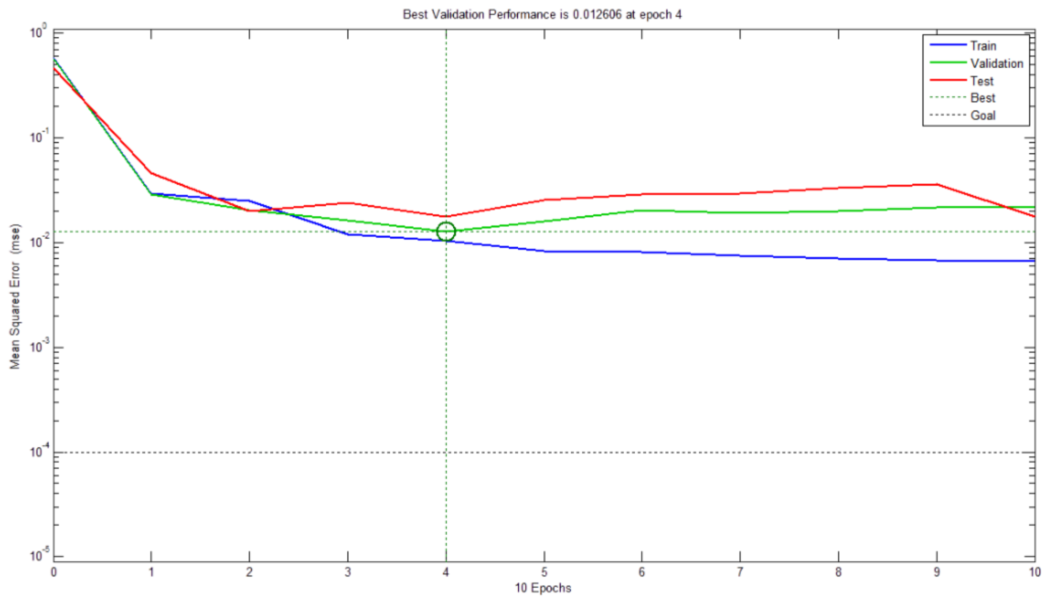


Fig. 5. Performance and training state of the optimal ANN

The errors of testing and validation datasets are displayed as a bar chart in fig. 6. As can be seen from the figure, test errors show greater values compared to those of validation indicating that model performance is quite good.

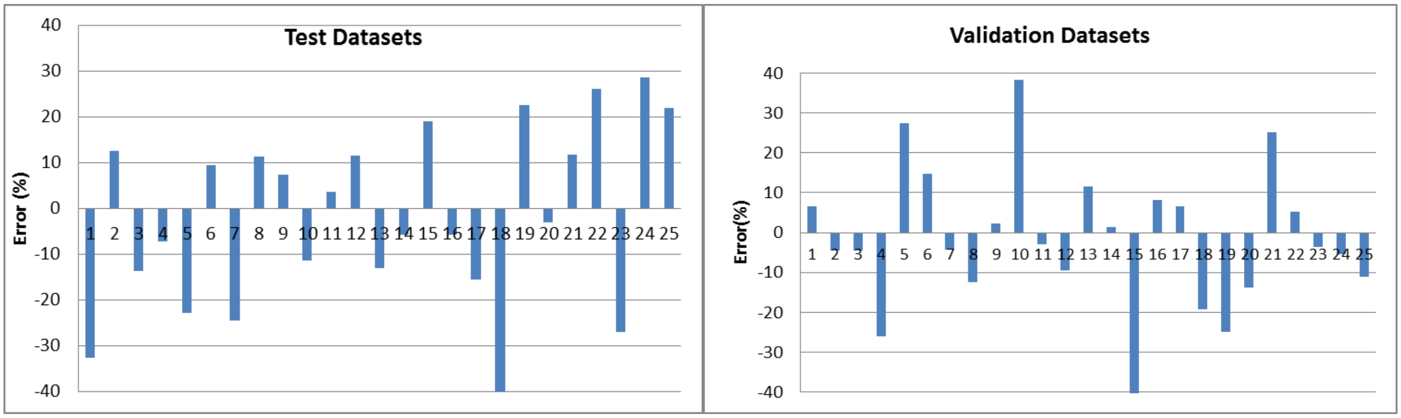


Fig. 6. Error of validation and test datasets

Error data points of training and all data for predicted and experimental PCI are also presented in fig. 7 for the purpose of comparison. For more accurate comparison of the optimal ANN performance, MAPE, RMSE, and R^2 values for training, testing, validation and all datasets are reported in table 2.

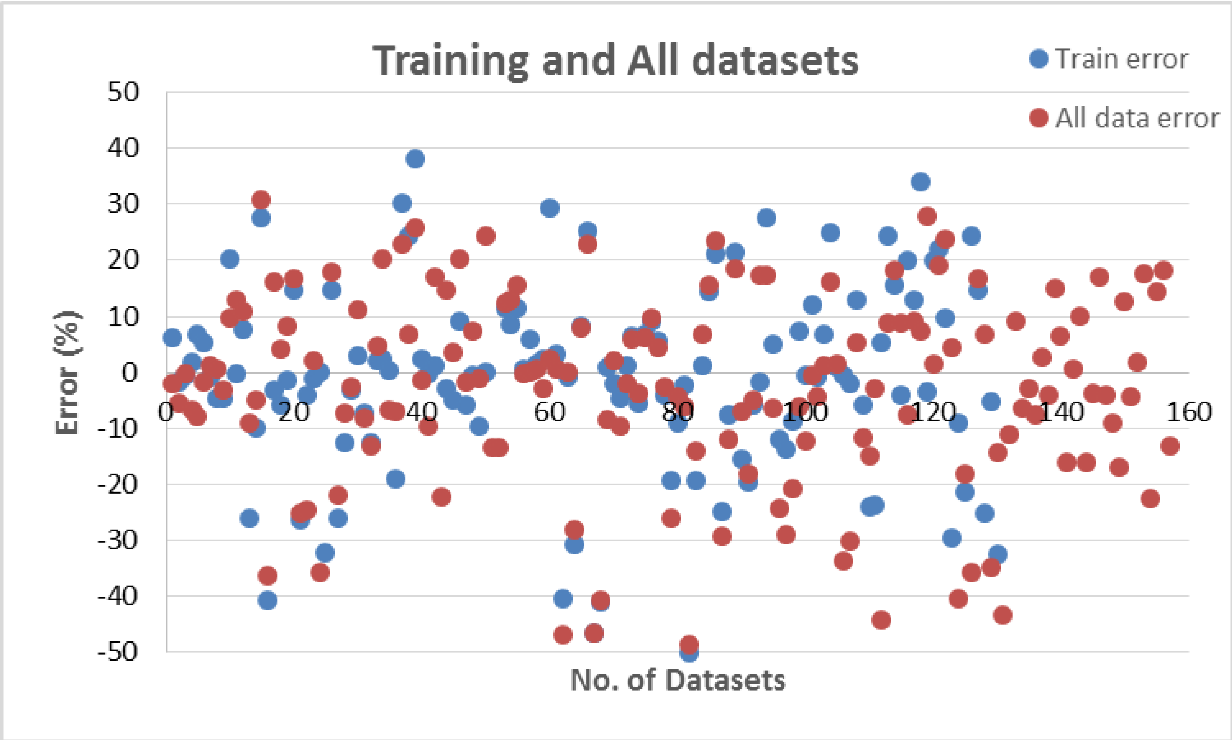


Fig. 7. Error for training set and all data

Table. 2. Error and correlation factor for various datasets

Datasets	train	validation	test	all
MAPE	13.10	13.20	16.78	13.39
RMSE	17.21	17.51	19.61	17.72
R2	0.978	0.973	0.965	0.974

As noted in table 2, the calculated coefficient of determination (R^2) for various datasets is quite high which supports the efficiency of the prediction model. The error values for various datasets can also be compared demonstrating the performance of the ANN prediction model.

5 CONCLUSION

In this paper, the results obtained experimentally for pavement condition index (PCI) prediction were modelled by using an optimal ANN. Qualitative variables such as year and type of pavement were also taken into account. Various ANN architectures were investigated through optimization process and the performance of the model was assessed through MAPE, RMSE and R^2. It was found that optimal ANN obtained by suggested multi-step approach can significantly enhance the performance and accuracy of the model and its error is well below half of the normal ANN model.

6 REFERENCES

[1] ASTM D6433 – 11. 2011. Standard Practice for Roads and Parking Lots Pavement Condition Index Surveys, West Conshohocken, PA.

[2] Fathi M, Jalal M, Rostami S. 2015. " Compressive strength prediction by ANN formulation approach for CFRP confined concrete cylinders", Earthquake and structures, 8(5): 1171-1190.

[3] Jalal M. 2015. "Soft computing techniques for compressive strength prediction of concrete cylinders strengthened by CFRP composites" Science and Engineering of Composite Materials, 22(1): 97-112.

[4] Garmsiri K, Jalal M. 2014. " Multiobjective optimization of composite cylindrical shells for strength and frequency using genetic algorithm and neural networks", Science and Engineering of Composite Materials, 21(4): 529-536.

[5] Jodaei A, Jalal M, Yas MH. 2013. " Three-dimensional free vibration analysis of functionally graded piezoelectric annular plates via SSDQM and comparative modeling by ANN", Mathematical and Computer Modelling, 57(5): 1408-1425.

[6] Jalal M, Ramezanianpour AA, Pouladkhan AR, Tedro P.2013. " Application of genetic programming (GP) and ANFIS for strength enhancement modeling of CFRP-retrofitted concrete cylinders", Neural Computing and Applications, 23(2): 455-470.

[7] Jalal M, Ramezanianpour AA. 2012. "Strength enhancement modeling of concrete cylinders confined with CFRP composites using artificial neural networks". Composites Part B: Engineering, 43 (8), 2990-3000.

[8] Jodaei A, Jalal M, Yas MH.2012. "Free vibration analysis of functionally graded annular plates by state-space based differential quadrature method and comparative modeling by ANN" Composites Part B: Engineering, 43 (2), 340-353.

[9] Ashrafi HR, Jalal M, Garmsiri K. 2010. "Prediction of load–displacement curve of concrete reinforced by composite fibers (steel and polymeric) using artificial neural network", Expert Systems with Applications, 37(12): 7663-7668