CORE
Metadata, citation and similar papers at core.ac.uk
Provided by International Institute for Science, Technology and Education (IISTE): E-Journals

Civil and Environmental Research
ISSN 2224-5790 (Paper) ISSN 2225-0514 (Online)
Vol.8, No.1, 2016
www.iiste.org

Developing Artificial Neural Network and Multiple Linear Regression Models to Predict the Ultimate Load Carrying Capacity of Reactive Powder Concrete Columns

Prof. Dr. Mohammed Mansour Kadhum Eng.Ahmed Abdulrazzaq Hussein Agha
Department of Civil Engineering, College of Engineering, Babylon University, Iraq.

Abstract
The study focuses on development a model to predict the ultimate load carrying capacity of Reactive Powder Concrete (RPC) columns. Two different statistical methods regression techniques (RT) and the artificial neural network (ANN) methods were used for determining the RPC columns ultimate load carrying capacity. The data is collected from three experimental studies the first used to develop the model and the other two used as a case study. Experimental results used as input data to develop prediction models. Two different techniques adopted to develop the models the first was Artificial Neural Network (ANN) and the second was multi linear regression techniques (RT). The models use to predict the ultimate load carrying capacity of RPC columns. To predict the ultimate load carrying capacity of RPC columns four input parameters were identified cross-section, micro steel fiber volume fraction content, compressive strength and main steel reinforcement area. Both models build with assistance of MATLAB software. The results exhibit that the cross section area has most significant effect on ultimate load carrying capacity. The performance of ANNs with different architecture was considered to adopt the pest ANN. An ANN with one layer consist of 7 neurons provide the best prediction. The results of this investigation indicate that ANNs have strong potential as statistical method for prediction the ultimate load carrying capacity of RPC columns.
Keywords: Reactive powder concrete, artificial neural network, multiple linear regressions, ultimate load carrying capacity, Statistical analysis.

1. Introduction
Reactive Powder Concrete also called Ultra-High Performance Concrete (UHPC) is a one of the most recent advances in concrete technology. It addresses the shortcomings of many concretes today (Therresa et al., 2008). RPC has attracted great research attention for its ultrahigh strength and high durability (Cheyrezy et al., 1995; Bonneau et al., 1996). RPC main features include a high percentage ingredient of Portland cement, silica fume, very low water-to-binder (cement + silica fume) ratio which ranges from 0.15 to 0.25, a high dosage of super plasticizer, fine sand with particle size ranges (150-600μm) and omitting the coarse aggregate. Among already built outstanding structures, **RPC** structures lie at the forefront in terms of innovation, aesthetics and structural efficiency. The unique properties for **RPC** make it extremely attractive for structural application. Many researchers have been carried out studies on RPC in the past years to assess the properties and its behavior. However, there is a lac in researches of RPC structural members in general and especially columns and that result in lac in design codes to estimate the behavior of RPC member. Knowing that columns occupy a vital place in the structural system weakness or failure of a column destabilizes the entire structure. From the above, we can recognize the necessity of estimating the ultimate load carrying capacity of RPC columns. In this research two statistical techniques will be investigated to estimate the RPC columns ultimate load carrying capacity.

2. Artificial Neural Networks (ANNs)
Artificial Neural Networks is a branch of artificial intelligence. Networks are modeled after the human brain consisting of brain cells and connections. As in the human brain, these networks are capable of learning from examples. Neural networks learn by adjusting their connection weights. Most networks are based on supervised learning algorithms in which pairs of input and desired output are shown to them during a training session.

3. Multiple Linear Regression Technique (RT).
Regression is a mathematical procedure for finding the best-fitting curve for a given set of points by minimizing the sum of the squares of the offsets ("the residuals") of the points from the curve. Multiple linear regression technique is a statistical technique that uses several explanatory variables (independent variables) to predict the outcome of a response variable (dependent variable). The goal of multiple linear regressions technique (RT) is to model the relationship between the explanatory and response variables. In the other terms it is model the relation between the independents and dependent variables.

4. Developing Neural Networks.

Developing the neural network required a number of steps summarized in the flow chart Fig.1. Building neural network consists basically of a number of simple processing units called neurons. Typically, the neurons are structured logically into groupings called layers. The network is hierarchical, consisting of three major layers: the input layer, hidden layer, and the output layer Fig. 2. Building ANN architecture also required number of selection to characterize the ANN structure. In general, there are two major neural network architectures the first is feed forward, and the second feed backward. Training of ANN could be supervised or unsupervised. Back propagation (BP) feed-forward multilayer perceptron is used extensively in engineering applications. A Back propagation network in which each neuron has one output and as many input as the neurons in the previous layer is the most common one. The network input is connected to every neuron in the first hidden layer while each network output is connected to each neuron in the last hidden layer Networks might consist of more than one hidden layer depending on the type of the studied problem. Theoretical one hidden layer is enough to approximately solve any function; however some problems might need more than one hidden layer, especially two hidden layers to be solved easily **(Partovi and Anandrajan, 2002; Carstenand Thorstein, 1993).** Each neuron in a given layer is linked to all the neurons in the following layer.

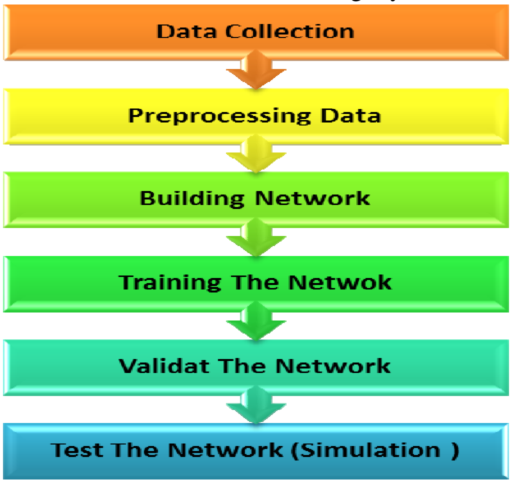


Figure (1): Basic flow for designing artificial neural network model.

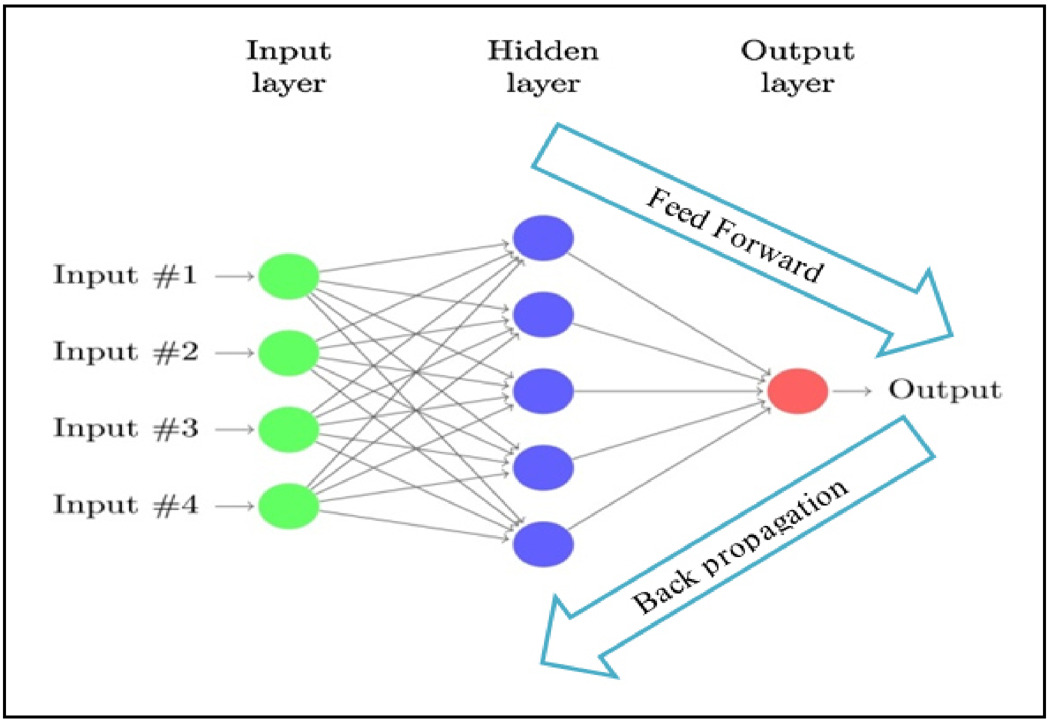


Figure (2): Basic Structure of Artificial Neural Network.

There are no clear rules to determine the number of neurons. It is most likely depend on the trial and error technique. However, there are many suggested criteria to specify the number of neurons depending on the number of inputs and the number of the outputs. Table 1 summarizes number of suggested function to determine the number of neurons in the first hidden layer. Network unit are connect through active function (transfer function) which is a mathematical function that a network unit uses to produce an output referring to its input value. The purpose of linear or nonlinear activation function is to ensure that the neuron's response is bounded-that is, the actual response of the neuron is conditioned or damped, as a result of large or small activating stimuli and thus controllable. Further, in order to achieve the advantages of multilayer nets compared with the limited capabilities of single layer networks, nonlinear functions are used, depending upon the paradigm and the algorithm used for training the network. There are some heuristic rules for the selection of the activation function. For Example, **(Klimasauskas, 1991)** suggest logistic activation function for classification problems which involve learning about average behavior and to use the hyperbolic tangent function for forecasting problem which involves in learning about deviations from the average. Three of the most commonly used activation functions are summarized in Table 2.

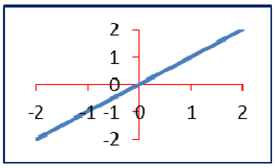
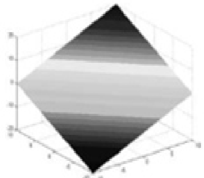
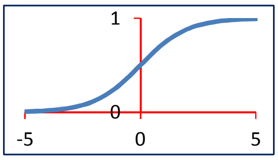
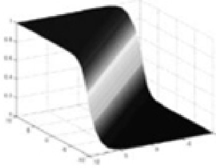
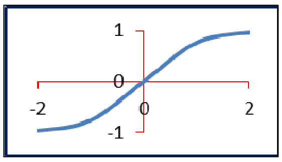
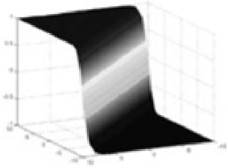
5. Training, Validation and Testing ANN.

The first step to train multi-layer ANN is to first divide the data into three subsets. The first subset is the training set, which is used for computing the gradient and updating the network weights and biases. The second subset is the validation set. The error on the validation set is monitored during the training process. The validation error normally decreases during the initial phase of training, as does the training set error. However, when the network begins to over-fit the data, the error on the validation set typically begins to rise. The network weights and biases are saved at the minimum of the validation set error. Training of the neural network is carried out using 21 data sets divided into 70% for training, 15% for testing and 15% for validation in addition to four data sets used as case study in simulation the ANN.

Table 1: Empirical Criteria to Determine the Number of Neurons in the Hidden Layer.

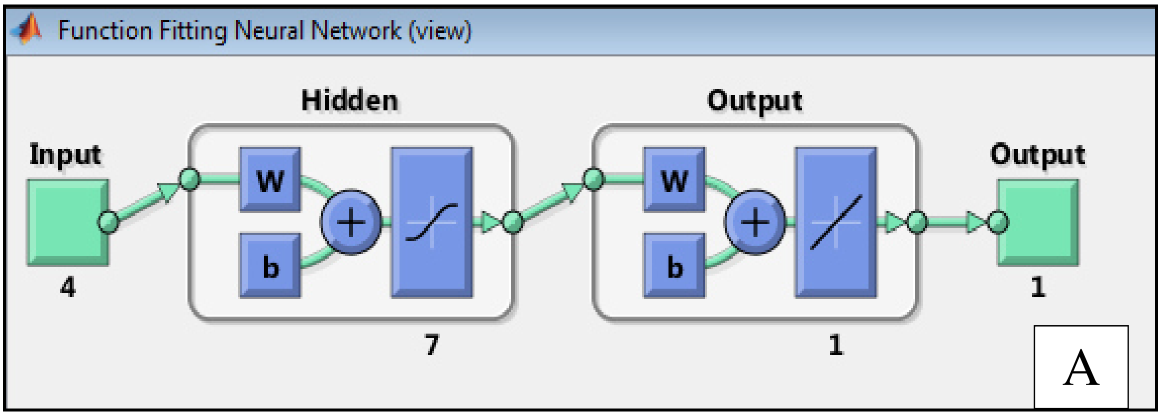
Number of Neurons in The First Hidden Layer	Number of Neurons for the Current Study	Researcher (Reference)
2*NI	8	GALLANT, S. I. (1993), Neural Network Learning and Expert Systems, MIT Press, Cam-bridge.
NI	4	LAI, S. and SERRA, M. (1997), „Concrete Strength Prediction by Means of Neural Net-work", Construction and Building Materials 11(2), 93–98.
NI+NO	5	NAGENDRA, S. (1998), Practical Aspects of Using Neural Networks: Necessary Prelimi-nary Specifications, Technical Paper, GE Research and Development Center.
0.75*NI	3	NEHDI, M., DJEBBAR, Y. and KHAN, A. (2001b), „Neural Network Model for Preformed Foam Cellular Concrete", ACI Materials Journal 98(5), 402–41
2*NI+1	9	NEVILLE, A. M. (1986), Properties of Concrete, Longman Scientific and Technical, Third Edition.
(NI+NO)/2	2-3	POPOVICS, S. (1990), „Analysis of Concrete Strength versus Water-Cement Ratio Relationship", ACI Materials Journal 87 (5), 517–529.

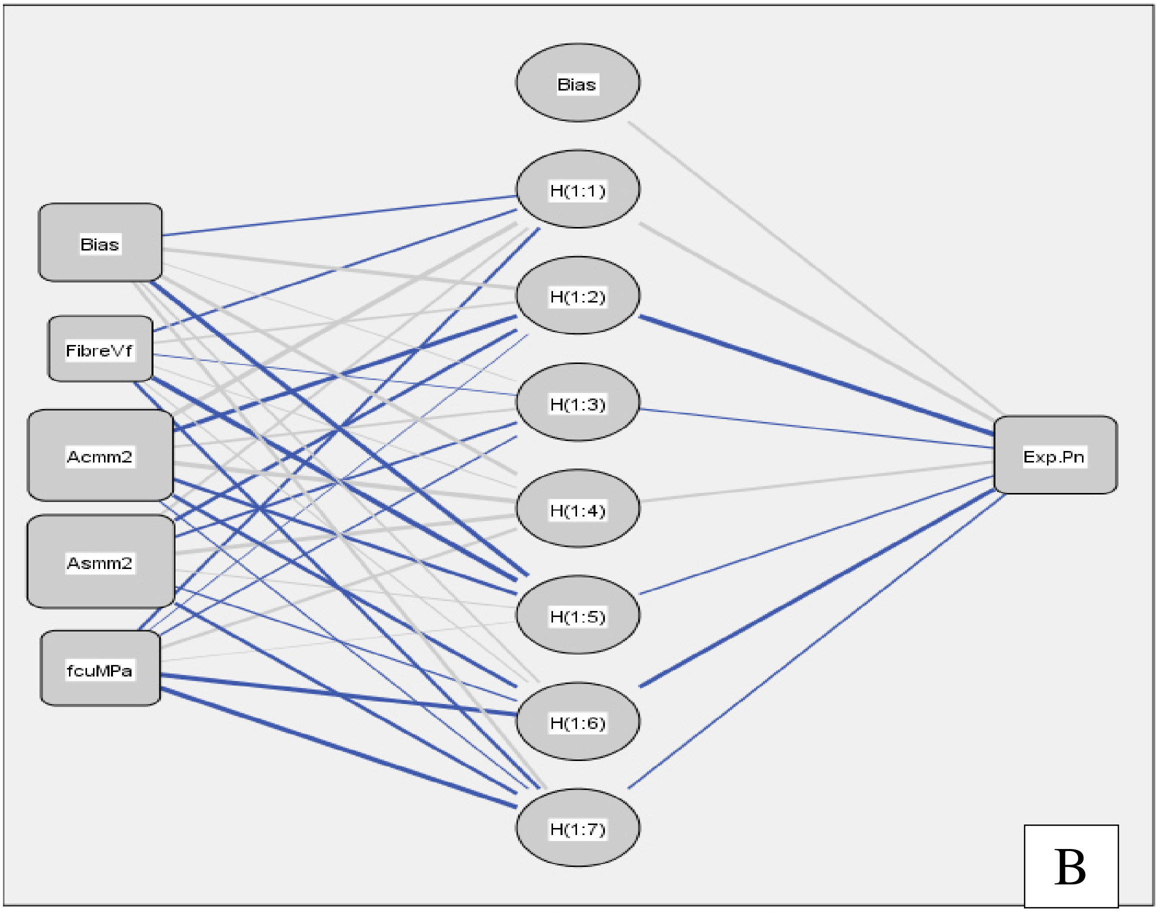
Table 2: Three of the Most Commonly Used Neuron Activation Functions.

Active Function Name	Formula	2D Graphical Representation	3D Graphical Representation	Description
Linear	$f(x) = x$, *for all x*			The activation of the neuron is passed on directly as the output
Logistic (or sigmoid)	$f(x) = \dfrac{1}{1+e^{-x}}$			A S-shaped curve, very popular because it is Monotonous and has a simple derivative, Range of logistic or sigmoid function is from 0 to 1
Hyperbolic Tangent	$f(x) = tanh(x)$ $f(x) = \dfrac{1+e^{-2x}}{1+e^{2x}}$			A sigmoid curve similar to the logistic function. Often performs better than the logistic function because of its symmetry. Ideal for multilayer Perceptrons, particularly the hidden layers. Output value is between -1 and +1

6. Prediction with ANN

ANN model developed in this research has four neurons in the input layer (independent variabels) and 1 neurons in the output layer (dependent variable) as shown in Fig.. 3. Depending on previous researches along with trial and error method the adopted network has one hidden layer with 7 neurons as it provided the best performance: minimum % error and maximum correlation values for training, validation and testing sets. The parameter used for prediction were, concrete cross-section area (AC), main steel reinforcement area (As), compressive strength (fcu) and micro steel fiber volume fraction content (Vf). These parameter used to predict the only independent variable which is the ultimate load carrying capacity. Table 3 present the ANN summary. The case processing includes twenty one cases for building, training and validation of the ANN and the final trained ANN tested with additional four cases from other researcher. Table 4 illustrates the used cases parameters. The values of the covariates rescaled by standardize method which is done by subtract the mean and divide by the standard deviation, (x−mean)/s for the input layer and output layer. Feed-forward model with hyperbolic tangent transfer function was used as the activation function for hidden and output layers. Because the back propagation network weights cannot be easily understood in the form of a numeric matrix, therefore they may be transformed into coding values in the form of a percentage by dividing the weights by the sum of all the input parameters. This gives the relative importance for each input parameter to the output parameter. Fig. 4 presents Neural Network Training, Testing and Validation Performance. .The relative importance for various input parameters is shown in Fig. 5 and the major dominant parameter is the concrete cross-section (46.8%). The relative importance of the other variables with respect to cross-section; As, fcu and Vf are: 86.6, 19.8 and 7.3%, respectively.





Hidden layer activation function: Hyperbolic tangent

Output layer activation function: Identity

Figure (3): A: Structure of the neural network for prediction of ultimate load carrying capacity.
B: Architecture schematic of the ANN for prediction of ultimate load carrying capacity.

Table 3: Adopted Artificial Neural Network Architecture Information.

Network Information			
Input Layer	Covariates	1	Steel Fiber Vf (Vol.%)
		2	Ac(mm2)
		3	As(mm2)
		4	fcu (MPa)
	Number of Units		4
	Rescaling Method for Covariates		Standardized
Hidden Layer	Number of Hidden Layers		1
	Number of Units in Hidden Layer 1		6
	Activation Function		Hyperbolic tangent
Output Layer	Dependent Variables	1	Exp.Pn
	Number of Units		1
	Rescaling Method for Scale Dependents		Standardized
	Activation Function		Identity
	Error Function		Sum of Squares error

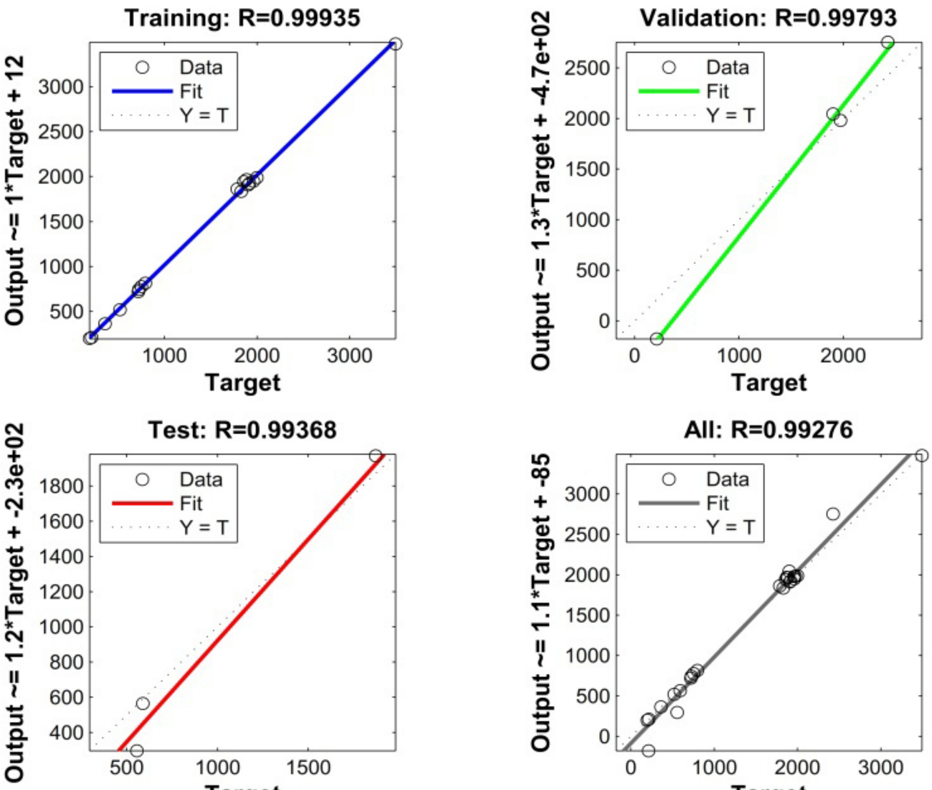


Figure (4): Neural Network Training, Testing and Validation Performance.

Civil and Environmental Research
ISSN 2224-5790 (Paper) ISSN 2225-0514 (Online)
Vol.8, No.1, 2016

Table 4: Values and Parameters of Database Used in This Study to Develop The Estimation Models.

Data Type	No.	Vf (Vol.%)	Ac(mm^2)	As(mm^2)	*fcu (MPa)*	*Exp.Pn (kN)*
Data Used to Develop the Models	1	1	22500	113	135.6	1859.4
	2	1	22500	201	135.6	1884.6
	3	1	22500	314	135.9	1901.3
	4	1.5	22500	113	141.5	1905.7
	5	1.5	22500	201	142.3	1918.5
	6	1.5	22500	314	142.5	1958.7
	7	2	22500	113	145	1962
	8	2	22500	201	144.4	1973.5
	9	2	22500	314	145.2	1998.5
	10	1	22500	0	134.9	1786.3
	11	1.5	22500	0	142.7	1829.5
	12	2	22500	0	144	1870.5
	13	1	10000	0	137.3	725
	14	1.5	10000	0	141.3	753
	15	2	10000	0	145.9	796.3
	16	1	5625	0	134.7	520.5
	17	1.5	5625	0	140.8	556.3
	18	2	5625	0	143.6	589.8
	19	1	2500	0	134.6	195.4
	20	1.5	2500	0	142.9	210.8
	21	2	2500	0	146.2	215.6
Case Study Data.	22	2	10000	113	132	720
	23	2	4900	50	132	360
	24	2	22500	1605	152	3493
	25	2	22500	905	152	2428
Data Ranges	Minimum=	1	2500	0	132	195.4
	Maximum=	2	22500	1605	152	3493
	Average=	1.58	15371	182.3	141.0	1376.5

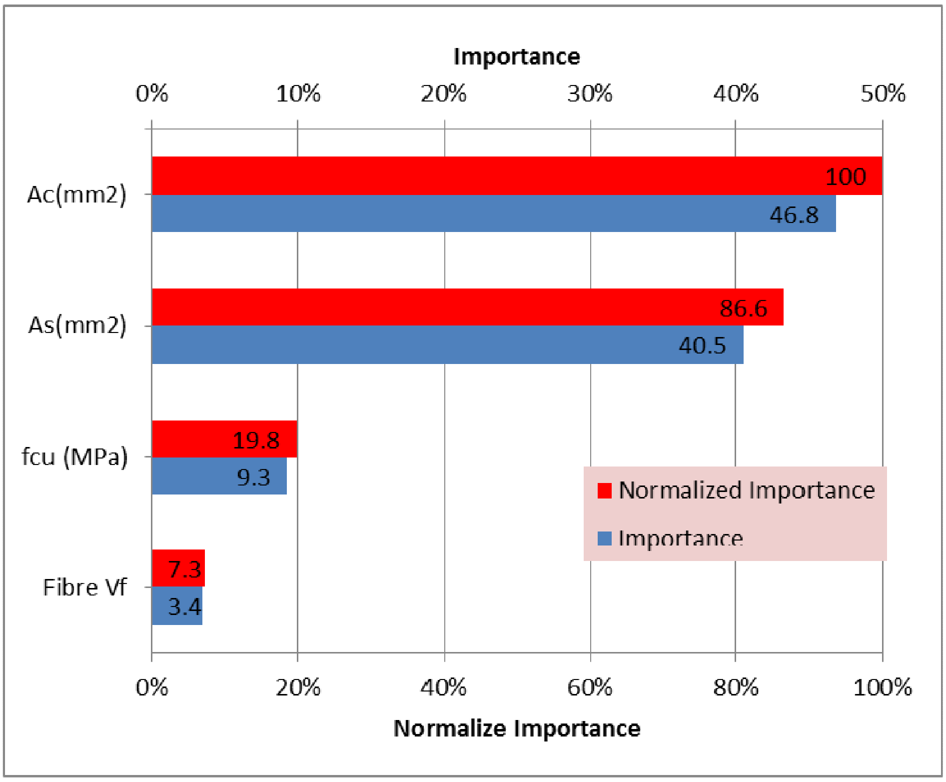


Figure (5): Relative importance of independent variables for the prediction of ultimate load carrying capacity.

7. Prediction with RT

Multiple linear regression model was used in the prediction of RPC columns ultimate load carrying capacity. Nonlinear regression model was not preferred here as there was no information about the data structure. Since the degree of the nonlinear regression model was unknown previously, this necessitated the use of the linear regression model. In the formulated model, the independent variables were: Ac, As, fcu and Vf; whereas the dependent variables was column ultimate load carrying capacity. A total of 21 data were used in developing the regression model and four data were used in testing the model equation obtained. The limit values of variables used in the multiple linear regression models were listed in Table 4. The prediction multiple linear regression model (variables and coefficients) used in this study is shown in Table 5.

Table 5: Multiple Linear Regression Model Parameters.

RT model **RT.Pn=a+b*Vf+c*Ac+d*As+e*fcu**					
Parameter	Intercept	Vf (Vol.)	Ac(mm2)	As(mm2)	fcu (MPa)
Value	-1206.032	-30.932	0.076	0.914	9.197

8. Models Performance

In general, there are two statistical routines for compression between prediction models. The first is to compare the two models to the same data. The first helps you compare two fitting functions to a single dataset in order to determine which function provides the best fit of the data. The second is to easily compare the same fitting function to different data sets to make a judgment about their statistical similarity.

In order to evaluate prediction models (ANN and RT) the observed experimental data was compared with predicted results for each model, in the other words the predicted results used as input. The same statistical measures applied for both models to standardize the comparison Table 6 illustrate the formulas that used to compute the statistical measures. The following statistical measures were used for comparison:

8.1 Coefficient of Determination (R^2)

R^2 is a statistical ratio that compares model forecasting accuracy with accuracy of the simplest model that just use mean of all target values as the forecast for all records. The closer this ratio to 1 the better the model is. Small positive values near zero indicate poor model. Since we are using multiple linear regressions it is preferable to adopt Adjusted R^2. Because R^2 increases with added predictor variables in the regression model, the adj.R^2 adjusts for the number of predictor variables in the model. This makes it more useful for comparing models with a different number of predictors.

8.2 Root Mean Square Error (RMSE)

The root mean square error is applicable to iterative algorithms and is a better measure for higher values. It offers a general representation of the errors involved in the prediction. The lower the value of RMSE, the better the fit is.

8.3 Mean Absolute Error (MAE)

The mean absolute error has the advantage that it does not distinguish between the over and underestimation and does not get too much influenced by higher values. It is generally engaged in addition to RMSE to get the average error without worrying about the positive or negative sign of the difference. Lower the value of MAE the better is the forecasting performance.

8.4 Absolute Relative Error. (ARE)

ARE is an error value that indicates the "quality" of the prediction model. This parameter is calculated by dividing the difference between actual and desired output values by the module of the desired output value. It offers a general representation of the prediction model relatively to the value of the observed variable.

8.5 Residuals (Re) and Sum of Residuals (SRe).

It is stand for the difference between the observed value of the dependent variable and the predicted value is called the residual (e). Each data point has one residual. It is offer general trends of the prediction model regarding been over or under estimating.

Table 6: Statistical Measures Calculation Formulas.

Statistical Measures	Symbol	Formula
R-Squared	R^2=	$1-\frac{SSE}{SST}$
Adjusted R-Squared	Adj.R^2=	$1-\frac{SSE(n-1)}{SST(n-p)}$
Root Mean Square Error	RMSE=	$\sqrt{\frac{SSE}{n}}$
Mean Absolute Error	MAE=	$\frac{1}{n}\sum_{i=1}^{n}\lvert Re_i\rvert$
Absolute Relative Error	ARE=	$\left\lvert\frac{Re_i}{y_i}\right\rvert$
Residuals	Re=	$y_i - f_i$

Where:
SSE: The sum of squared error.
SST: The sum of squared error with respect to the average.
n: The number of observations.
p: The number of regression coefficient including the intercept.
Re: Residual (error) equal to Observed value - Predicted value.
y_i *: observed dependent variable.*
f_i*: Predicted value.*

9. Comparison of Regression Techniques (RT) with Artificial Neural Network (ANN)

The ANN model performances are compared with Multiple Linear Regression Techniques. The models are evaluated by comparing the predicted results and measured strength values (statistical measures). Table 7 shows all the statistical measures. A comparison between ANN model and RT model statistical measures shows that ANN models provide better results than RT models. The Visual comparison illustrated in Figures 6 to 10. Fig. 6 illustrate the a compression between the ANN and RT models with respect to the observed result, it is quite obvious that ANN model had a better agreement and response to parameter variation rather than RT model.

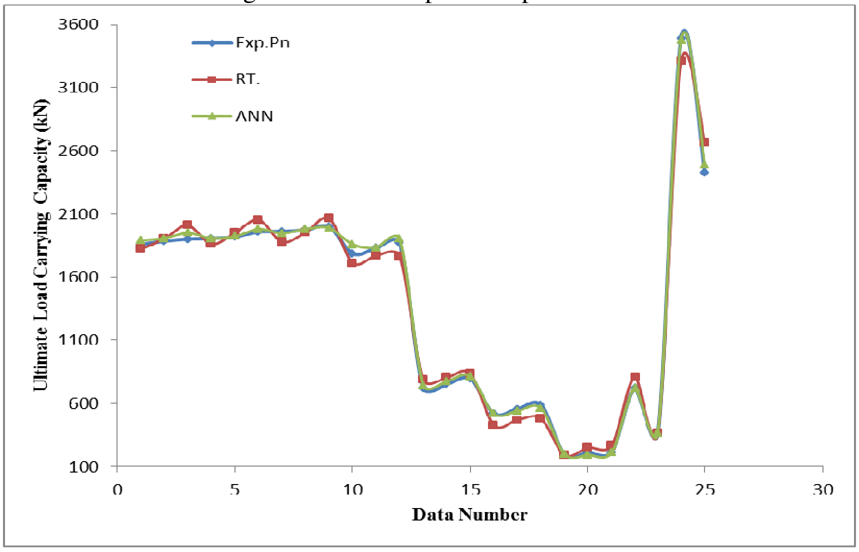


Figure (6): Comparison of predicated values for ANN and RT.

Figure 6 shows the models performance, it can be seen that ANN model results are closer to the observed results and the majority of ANN result are is located on the line of equality which means that the actual and the predicted values. This is quite true, because the model has adj.R^2 equal to 0.99 for the prediction of RPC columns ultimate load carrying capacity.

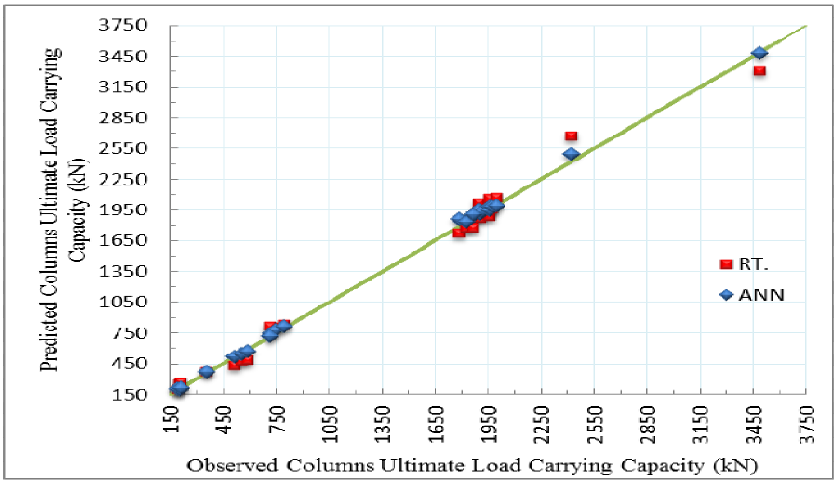


Figure (7): Performance comparison of observed values with predicted values obtained from ANN and RT.

Fig. 8 and Fig. 9 present the ANN and RT models prediction residuals (Re) and relative residuals (RRe) respectively. Both Figures shows that the errors are well distributed around the zero axis and the ANN model are generally closer to the zero axis. This conclusion are supported by the statistical measures evidence as shown in Table 7 ANN model has ARE and MEA equal to 0.48 and 19.66 respectively while RT model has ARE and MEA equal to 1.85 and 71.79 respectively. The Figures also present the extra four case study set of data and again the ANN models prediction results was obviously better than the RT model. Fig. 10 shows accuracy compression between ANN model and RT model. It is obvious that ANN model had uniform distribution of results declare the constant performance ANN model while RT model results show a wide range of fluctuation.

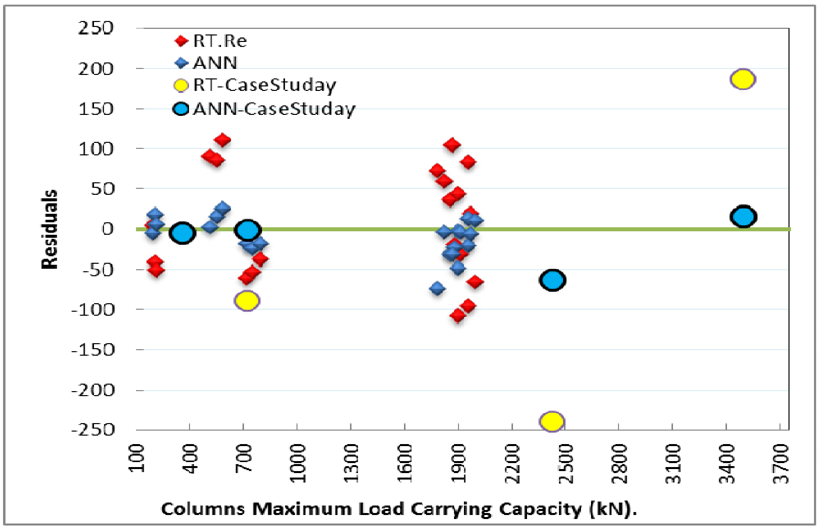


Figure (8): The distribution of the residual values with predicted values of RPC Columns Ultimate Load Carrying Capacity for ANN and RT models.

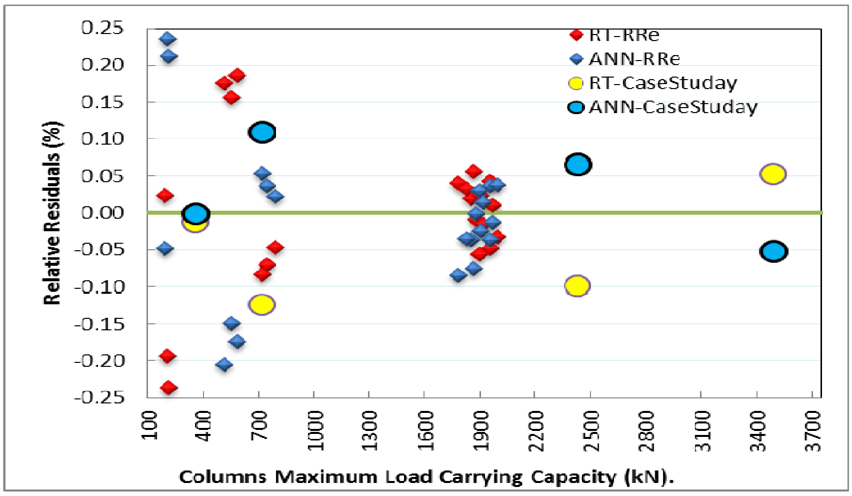


Figure (9): The distribution of the relative residual values with predicted values of RPC Columns Ultimate Load Carrying Capacity for ANN and RT models.

Table 7: The Prediction Performances and Results of Both Techniques for the Testing Set.

Data Type	**No**	***Ultimate Load Carrying Capacity***		
		Exp	**RT**	**ANN**
Data Used to Develop the Models	1	1859.4	1823.4	1890.7
	2	1884.6	1903.9	1908.2
	3	1901.3	2009.9	1950.4
	4	1905.7	1862.2	1908.5
	5	1918.5	1950.0	1922.9
	6	1958.7	2055.1	1980.0
	7	1962.0	1879.0	1948.5
	8	1973.5	1953.9	1980.1
	9	1998.5	2064.5	1988.4
	10	1786.3	1713.7	1860.8
	11	1829.5	1770.0	1833.6
	12	1870.5	1766.5	1902.5
	13	725.0	785.8	744.0
	14	753.0	807.1	777.5
	15	796.3	833.9	815.3
	16	520.5	429.4	518.0
	17	556.3	470.0	540.6
	18	589.8	480.3	564.3
	19	195.4	191.0	200.3
	20	210.8	251.8	192.6
	21	215.6	266.7	210.4
Case Study Data.	22	720.0	809.4	720.4
	23	360.0	364.2	364.2
	24	3493.0	3307.0	3476.0
	25	2428.0	2667.2	2490.0
Statistical Measures	**R^2=**		**0.988**	**0.9989**
	Adj.R^2=		**0.986**	**0.9987**
	RMSE=		**88.73**	**26.9**
	MAE=		**71.79**	**19.66**
	ARE=		**1.85**	**0.48**

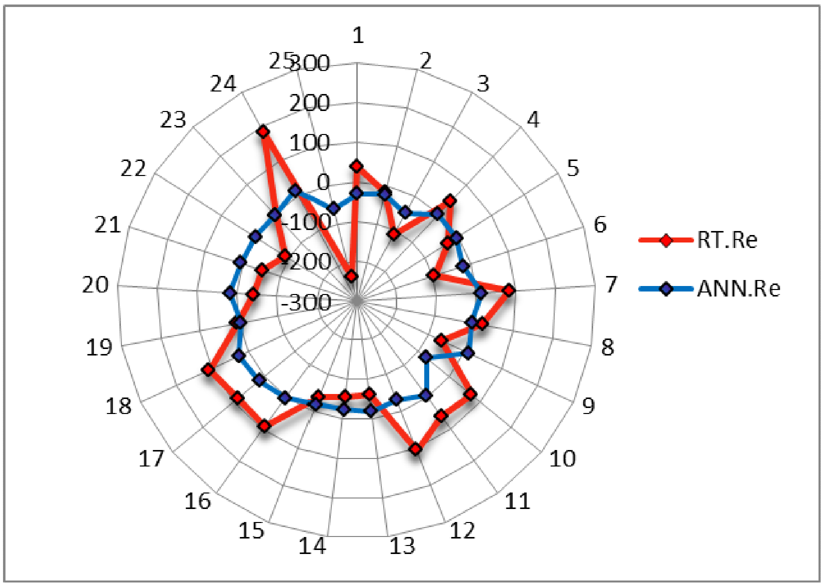


Figure (10): ANN and RT Residual Accuracy Distribution.

10. CONCLUSIONS

It has been shown that ANN models provide better performance in predication the ultimate load carrying capacity of the RPC columns rather than the multiple linear regression technique. ANN model results have shown that ANN base modeling can effectively be used in predicting the RPC columns ultimate load carrying capacity. The parametric study showed that the column concrete cross-section area is the most significant factor affecting the output of the model. Furthermore, the relative importance of values of other input parameters is insignificant with respect to the importance of column concrete cross-section. For selecting the best configuration of the networks, there are no specific criteria, and trial and error approach should be employed that takes into consideration the best network performance, average error and the best network performance for the testing data.

References

Bonneau, O., Poulin, C., Dugat, J., Richard, P., and Aitcin, P. C. (1996). "Reactive Powder Concretes: from Theory to Practice". Concrete International, 18 (4), 4749.

CARSTEN, P. and THORSTEIN R. (1993), Jetnet 3.0: A Versatile Artificial Neural Network Package, Sweden, http://www.thep.lu.se/activities/jetnet_30_manual/jetnet_30_manual.html, retrieved on 2007.

Cheyrezy, M., Maret, V., and Frouin, L. (1995). "Micro Structural Analysis of RPC (Reactive Powder Concrete)". Cement and Concrete Research, 25 (7), 1491-1500.

GALLANT, S. I. (1993), Neural Network Learning and Expert Systems, MIT Press, Cam-bridge. LAI, S. and SERRA, M. (1997), „Concrete Strength Prediction by Means of Neural Net-work", Construction and Building Materials 11(2), 93–98.

Jagannathasn Saravanan and Saranya Poovazhagan., Analitical Study On Compresisve Strength Of Reactive Powder Concrete. *International Journal of Recent Scientific Research Vol. 6, Issue, 5, pp.3872-3880, May, 2015.*

Klimasauskas, C.C., 1991. Applying neural networks. Part 3:138 Training a neural network, PC-AI, May/ June, 20–24.

NAGENDRA, S. (1998), Practical Aspects of Using Neural Networks: Necessary Prelimi-nary Specifications, Technical Paper, GE Research and Development Center.

NEHDI, M., DJEBBAR, Y. and KHAN, A. (2001b), „Neural Network Model for Preformed Foam Cellular Concrete", ACI Materials Journal 98(5), 402–41.

NEVILLE, A. M. (1986), Properties of Concrete, Longman Scientific and Technical, Third Edition. POPOVICS, S. (1990), „Analysis of Concrete Strength versus Water-Cement Ratio Relationship", ACI Materials Journal 87 (5), 517–529.

PARTOVI, F. Y. and ANANDRAJAN, M. (2002), „Classifying Inventory Using an Artificial Neural Network Approach", Computers and Industrial Engineering 41, 389–404.

Therresa, M., Erron, A., and Donald, M. (2008). "Ultra High-Performance Concrete for Michigan Bridges Material Performance Phase-I, Center for Structural Durability". Michigan Technological Transpiration Institute, Final Report-November, 152.