ITcon

www.itcon.org - Journal of Information Technology in Construction - ISSN 1874-4753

COMPARATIVE STUDY IN THE USE OF NEURAL NETWORKS FOR ORDER OF MAGNITUDE COST ESTIMATING IN CONSTRUCTION

SUBMITTED: January 2014
REVISED: September 2014
PUBLISHED: October 2014 at http://www.itcon.org/2014/27
EDITOR: Rezgui Y.

Hany El-Sawah, PhD, PMP, Professor
Faculty of Engineering, Helwan University, Cairo Egypt;
Visiting Scientist, Department of Building Civil & Environmental Engineering, Concordia University, Montreal, Canada;
email: helsawah@hotmail.com

Osama Moselhi, PhD, P.Eng., Professor
Department of Building Civil & Environmental Engineering, Concordia University, Montreal, Canada;
email: Moselhi@encs.concordia.ca

SUMMARY: *This paper presents a study on the use of artificial neural networks (ANNs) in preliminary cost estimating. The choice and the design of the ANN model significantly affect the results obtained from the model and, hence, the accuracy of the estimated cost. The study considered Back Propagation Neural Network (BPNN), Probabilistic Neural Network (PNN) and Generalized Regression Network (GRNN) as well as regression analysis. Models were developed for order of magnitude cost estimating of low-rise structural steel buildings and short-span timber bridges. The study was conducted on actual data for 35 low-rise structural steel buildings and their respective cost was estimated using the developed regression and ANN models. These models were also applied to estimate the cost of a timber bridge extracted from the literature. The results showed that the mean absolute percentage error (MAPE) for the neural network models ranges from 16.83% to 19.35% whereas was equal to 23.72% for the regression model. Moreover, the linear regression model was more sensitive to the change of the number of the training data and that the PNN network was the most stable network among all the other estimating models as the maximum difference in MAPE percentage was only 2.46%. Whereas, the maximum difference in MAPE was 19.47%, 17.91%, and 61.45% for BPNN, GRNN and regression models respectively.*

KEYWORDS: *Cost estimating, artificial neural networks, structural steel buildings.*

REFERENCE: *Hany El-Sawah, Osama Moselhi (2014). Comparative study in the use of neural networks for order of magnitude cost estimating in construction, Journal of Information Technology in Construction (ITcon), Vol. 19, pg. 462-473, http://www.itcon.org/2014/27*

1. INTRODUCTION

At the conceptual stage of design, it is important to arrive at reliable early cost estimates (Carr, 1998; Dysert, 2003; Moselhi and Siqueira, 1998; Petroutsatou et al., 2012). The challenge facing estimators is to produce an estimate that is an accurate reflection of reality. The estimator then uses the best information available to evaluate the costs of performing the required work (Carr, 1989; Hendrickson and Au 2003). According to the Association for the Advancement of Cost Engineering International (AACE International, 2005), at each phase in the project life cycle different levels of cost estimating accuracy can be achieved based on the information available. In particular, the order of magnitude cost estimate is used at the conception phase of the project. The estimated cost may vary from –30% to +50% of the actual cost. Practically, the order of magnitude cost estimate is calculated using RS Means construction cost data (RS Means 2012). This estimated cost is based essentially on the project area while in some cases, the estimated cost is adjusted for perimeter and height of the building.

Traditional parametric cost estimating techniques have been widely used due to their simple formulation. They lead to the development of cost models based on regression analysis to historical data of similar projects to the one at hand. A major disadvantage of these techniques is that the mathematical form has to be defined before any analysis can be performed (Creese and Li, 1995). Another disadvantage is their unsuitability to account for the interaction among the large number of variables present in a construction project (Hegazy and Ayed, 1998). These limitations may contribute to the low accuracy of the traditional models and their limited use in construction (De la Garza and Rouhana, 1995). On the other hand, ANNs offer an alternative approach to cost estimation modeling. They overcome the shortcomings of the traditional cost estimating techniques as they are nonparametric estimators and they learn by detecting the hidden relationship between the input parameters and the output cost of the training data set.

Several researches demonstrated the potential use of ANN in construction (Moselhi et al, 1992; Hegazy and Ayed, 1998) and their superior performance over traditional regression analysis (De la Garza and Rouhana, 1995; Creese and Li, 1995; Emsley et al. 2002; Setyawati et al. 2002). There are 3 commonly used ANNs in different engineering research fields namely; BPNN, PNN, and GRNN. The choice and the design of the ANN model significantly affect the results obtained from the model and, hence, the accuracy of the estimated cost. The BPNN is widely used for cost estimating of construction projects. Despite its capabilities, BPNN suffers from several problems that make the development of a neural network model a difficult task that is neither simple nor straightforward (Moselhi et al., 1991; Hegazy and Moselhi, 1994; Petroutsatou et al., 2012; Setyawati et al., 2002). The main problem is that there are no fixed rules to determine the appropriate architecture or its parameter values. The development of high-quality BPNN is difficult. The process of developing BPNN models often involves experimentation and multiple simultaneous development tracks. It often requires iterative refinements of network parameters, network redesign, and problem reformulation. Several design factors forecasts including: selection of input variables, significantly architecture of the network, and quantity of the training data, significantly impact the accuracy of the neural network.

There are reported advantages of using GRNN compared to the commonly used BPNN (Petroutsatou et al., 2012; Specht, 1991). First, it is known for its ability to train quickly with sparse data sets. Second, the output converges as the number of sample increases. Finally, the estimate is always bounded by the minimum and the maximum of the observations. In addition, PNN has a fast learning scheme and can be retained or updated. It has a unique feature that under certain easily met conditions, the decision boundary implemented by PNN asymptotically approaches the optimal decision surface (Demuth and Beale, 1998).

The main objective of this paper is to evaluate the performance of the three types of ANN cited above for preliminary cost estimating of construction projects namely; BPNN, PNN and GRNN. This leads to a better design of ANN models for cost estimating at the pre-design stage when there is insufficient definition of scope and characteristics for detailed estimating. The results of the different NN models are compared with the regression analysis and the actual costs of a real data.

The work in this research has been carried out in two stages:

- A pilot study was made using data extracted from the literature, where cost variables were identified and data is available for (12) timber bridge projects.
- A full scale study was done using a real data for 35 low-rise structural steel buildings, in which more sophisticated models and analysis were developed.

2. NEURAL NETWORK APPLICATION FOR COST ESTIMATING PURPOSES

NNs are widely used for cost estimating of different types of construction projects. BPNN was widely used for cost estimating of building projects (Emsley et al., 2002; Kim et al., 2005; Moselhi et Siqueira, 1998; Setyawati et al. 2002) and highway projects (Hegazy and Ayed, 1998; Pewdum et al., 2009). BPNN was also used for cost estimating of other construction projects such as timber bridges (Creese and Li, 1995), drainage projects (Alex et al., 2010). Moreover, Petroutsatou et al. (2012) developed two cost estimating models for road tunnels using two types of neural networks, namely BPNN and GRNN. On the other hand, radial basis neural network was used by Williams (2002 and 2005) to estimate the cost of highway projects. The findings of these articles confirmed that the effectiveness of NNs for cost estimating of construction. They also demonstrated that the mean average percentage error calculated using NN models ranged from 4.63% to 16.6%; depending on selection of input variables, NN topology, and quantity of the input data. Moreover, Petroutsatou et al. (2012) showed that GRNN provided higher level of accuracy than BPNN.

Several articles compared the estimates obtained from NNs with those obtained from various linear regression models (Emsley et al. 2002; Setyawati et al. 2002; Creese and Li, 1995). The results demonstrated that NN models outperform linear regression models given the same training data and the same variables as the major benefits of the NNs are their learning and generalization capabilities as well as their ability to model the nonlinearity of the data to predict the estimated cost.

3. PILOT STUDY

The data used in this part was obtained from previous research conducted by (Creese and Li, 1995) 0for timber bridges. Table 1 shows the original timber bridge data. The data includes three input variables namely: the volume of the webs, the volumes of the bridge decks, and the weight of the steel used and one output which is the actual cost. All data were normalized using the same scale applied by Creese and Li (1995) (i.e. 0.1 to 0.9) to enable comparison of results. Three NN models were developed using different input variables to determine the practicality of using NNs for cost estimation i.e. PBNN, GRNN and PNN. The same input variables were used in the regression analysis as well as the NN models. The BPNN topology was kept as that of Creese and Li (1995) 0to enable comparison of results. Table 2 summarizes the design characteristics of the different neural networks. To compare the accuracy of estimates obtained from the different models, the coefficient of determination R^2 and the Mean Absolute Percentage Error (MAPE) were calculated and presented in Fig. 1 and Fig. 2 and summarized in Tables 3 and 4. Fig. 1 shows that R^2 values for the neural network models were always greater than that in the linear regression model, i.e. the performance of the neural network models outperformed the performance of the linear regression model based on least squared error analysis. It also shows that model III was better than models I and II, i.e. the estimating error decreased as more input variables were introduced to the training.

TABLE 1: *Cost and Primary Parameters for Timber Bridges*

#	Name of bridge	Y1 Actual Cost ($)	X1 Web Volume (ft^3)	X2 Deck Volume (ft^3)	X3 Steel Weight (Ib)
1	Trace Fork Timber	74,982	662.86	542.34	527.98
2	Trout Run	87,602	791.15	566.72	651.08
3	Six Mile Creek	45,400	265.58	254.54	352.67
4	Left Hand Run	92,850	781.41	737.70	676.12
5	King Lear	75,000	336.88	753.38	434.06
6	Island Run	60,894	348.05	830.25	394.41
7	George Branch	61,354	455.18	567.50	535.27
8	Camp Arrowhead	79,512	1,164.17	892.97	834.72
9	Dunloup #1	201,600	1,661.65	2,825.00	1,316.25
10	Dunloup #2	194,599	1,665.04	2,484.38	1,168.81
11	Nebo	55,113	383.90	403.30	367.00
12	Light-burn	174,000	2,320.00	1,444.00	1,331.00

TABLE 2: *Design Characteristics for the different Neural Networks*

Model	Input Variables	BPNN	GRNN	PNN
I	X3	1 input, 2 neurons in the hidden layer, and one output	Spread = 0.1	No training parameters
II	X1, X2	2 input, 2 neurons in the hidden layer, and one output	Spread = 0.1	No training parameters
III	X1, X2, X3	3 input, 3 neurons in the hidden layer, and one output	Spread = 0.1	No training parameters

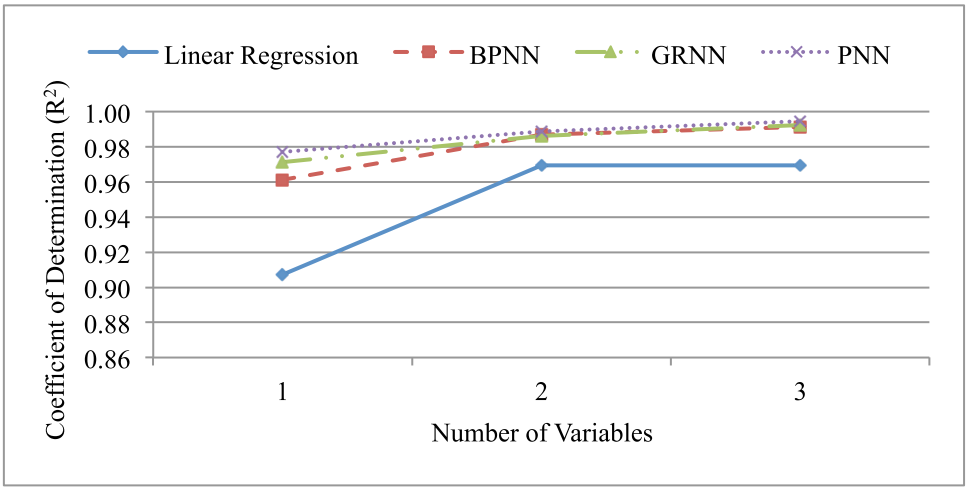

FIG. 1: *Coefficient of Determination R^2 for Different Models*

TABLE 3: *Coefficient of Determination R^2 for Different Models*

Model	Input Variables	R^2 values			
		Linear Regression	BPNN	GRNN	PNN
I	Variable : X3	0.9072	0.9613	0.9712	0.9772
II	Variables: X1, X2	0.9697	0.9870	0.9864	0.9888
III	Variables: X1, X2, X3	0.9697	0.9914	0.9926	0.9946

Fig. 2 presents the values of Mean Absolute Percentage Error (MAPE) of the estimated costs for the different models compared with the actual costs. It can be seen that the PNN showed the best performance among all models. Table 4 shows that the MAPE for the PNN ranges from 1.91% to 6.86% for Models III and I respectively. The estimated costs for the best regression model, e.g. model III, neural network models and the actual cost are illustrated in Table 5.

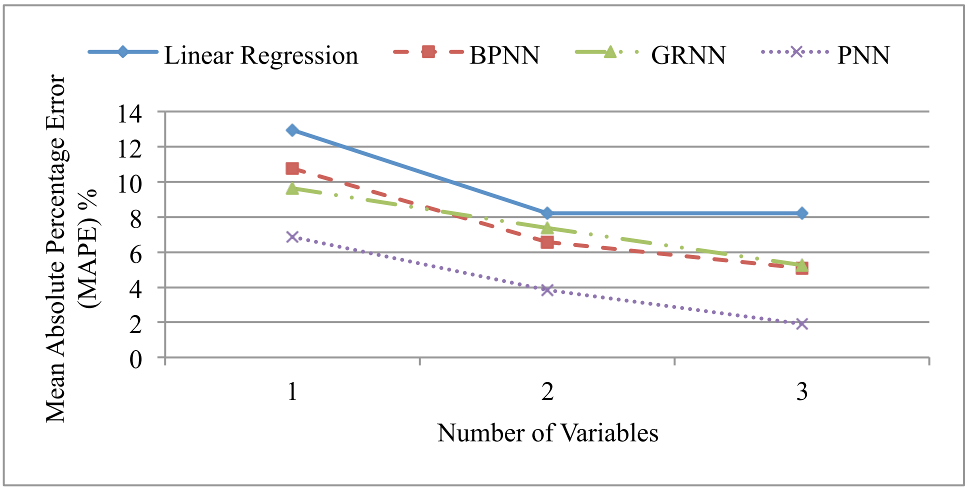

FIG. 2: *Mean Absolute Percentage Error (MAPE) for Different Models*

TABLE 4: *Mean Absolute Percentage Error (MAPE) for Different Models*

Model	Input Variables	Mean Absolute Percentage Error (MAPE) %			
		Linear Regression	BPNN	GRNN	PNN
I	Variable : X3	12.95	10.78	9.65	6.86
II	Variables: X1, X2	8.21	6.56	7.37	3.85
III	Variables: X1, X2, X3	8.23	5.09	5.25	1.91

TABLE 5: *Cost estimates from the different models ($)*

#	Name of bridge	Actual Cost	Linear Regression	BPNN	PNN	GRNN
1	Trace Fork Timber	74,982	71,136	76,835	75,078	73,719
2	Trout Run	87,602	77,010	84,060	87,574	85,158
3	Six Mile Creek	45,400	44,355	45,400	45,400	52,421
4	Left Hand Run	92,850	83,780	85,817	92,846	87,013
5	King Lear	75,000	67,789	65,706	75,078	65,282
6	Island Run	60,894	71,349	65,511	60,825	65,185
7	George Branch	61,354	64,539	70,978	75,078	68,207
8	Camp Arrowhead	79,512	104,504	84,450	79,569	80,389
9	Dunloup #1	201,600	203,603	201,991	201,600	200,595
10	Dunloup #2	194,599	189,418	194,181	194,571	195,603
11	Nebo	55,113	54,907	52,624	55,163	57,035
12	Light-burn	174,000	170,516	173,679	174,070	174,000

4. CASE STUDY: COST ESTIMATION OF STRUCTURAL STEEL BUILDINGS

This case study also evaluates the performance of the three ANN models against the regression analysis and the actual costs using a real data 35 low-rise structural steel buildings, in which more sophisticated models and analysis were developed. The data used in developing the neural network models were thirty nine real data for low-rise structural steel buildings fabricated and built between 1993 and 1997. A large manufacturer of low-rise structural steel buildings in Canada provided all the documents related to the building projects. Four types of project documents were collected containing cost data and describing the characteristics of each building for data extraction: the contract including project specifications and change orders, blueprint, detailed estimate, and the final cost report. The final cost report includes the actual costs of a project with and without the markup. Markups may vary considerably due to market conditions, the contractors' need for work, the number of bidders etc. (El-Sawah, 1994; Hegazy and Moselhi, 1994; Hegazy and Moselhi, 1995). It was agreed upon that only the direct cost of the buildings (i.e. material, labor and subcontractors' costs) would be used in the development of the neural network models in order to ensure the consistency in cost estimating.

The collected documents were presented in different formats, reflecting the company's estimating and accounting standards over time. Special care was dedicated to the extraction of data to ensure that the variation in type reporting, over the years, would not impact the costs actually incurred. A data entry sheet was designed for standardized data collection, organization, indexing, recording and analysis. Data were analyzed for consistency and parameter values checked for reasonableness in order to ensure same definition, in terms of content. In doing so, individual calculations of the total structural cost per square foot of building area were performed. Projects with unique characteristics were identified and excluded. Consequently, non-representative projects (4 projects) were accordingly rejected. Accordingly, the data of only 35 projects were used in the development of the proposed neural network model. Building area, perimeter, joist span and height were identified as the main parameters directly correlated to the fabrication cost of structural steel buildings. Table 6 shows the statistics of the input and output data used in the model development.

TABLE 6: *Data Characteristics*

	Area (ft^2)	Perimeter (ft.)	Height (ft.)	Joist Span (ft.)	Building Cost ($)
Minimum	2,356.2	194.4	10.0	25.7	27,838
Maximum	46,480.0	944.0	45.4	100.0	501,328
Mean	14,328.6	464.8	20.3	49.9	142,949
Standard deviation	11,126.9	179.9	5.7	20.9	121,869

In order to increase the accuracy of the developed models and to take account for the effects of inflation on the historical data, the data were indexed to July 1997, using (RS Means, 2012). Moreover, in order to improve the performance of the NNs and obtain better results, the original input data were normalized to a scaled data – transformed to the interval (1, 10) networks as shown in Fig. 3 This normalization was reported to improve the accuracy of the estimates generated (Moselhi and Sequeira, 1998). Upon completing the training, the scaled output is transformed back to the original data format. Fig. 3 shows the normalized cost of the buildings versus the normalized building area.

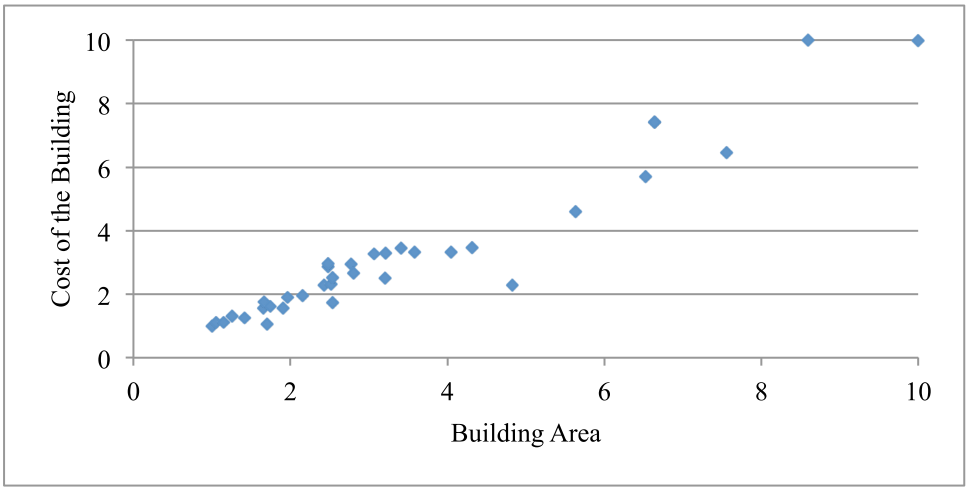

FIG. 3: *Normalized Building Cost versus Normalized Building Area*

5. DESIGN OF NEURAL NETWORK MODELS

A two-layer architecture with four nodes in the input layer and one node in the output layer was used for the BPNN design. The selected parameters used for the design of the BPNN were proved to be the most effective parameters by (Demuth H. and Beale M., 1998, Yuce et al., 2014). For example the Levenberg-Marquardt (LM) training algorithm used in the present study was reported to have the fastest convergence and best performance among other training algorithms. The number of nodes in the hidden layer was calculated according to suggested rules introduced by (Hegazy et al., 1994, Moselhi and Sequeira, 1998). Fourteen different architectures of BPNN architectures, with the number of nodes in the hidden layer varying from 3 to 16, were developed for the identification of the network structure with the best performance. Input and output parameters were normalized for confidentiality and effective training of the model being developed (Hegazy and Moselhi, 1994, Moselhi and Sequeira, 1998, Setyawati et al., 2002). The input parameters were kept constant in all networks.

It was decided to use 60% (18), 20% (6), and 20% (6) of the training and testing data to generate training, validation, and test sets respectively. The data were split in this way to maximize the cases available to train and validate the neural network while still providing sufficient cases to provide an independent test of the network's performance. The data was divided randomly by MATLAB. The performance of the networks was evaluated based on the Mean Squared Error (MSE) observed in the validation set. The coefficient of determination (R^2) was then calculated for the fourteen networks. The training of the networks was interrupted when the error falls below the user-specified level or when the user-defined number of training iterations was reached. The training of the network was based on 10,000 iterations. Fig. 4 and Fig. 5 show the MSE and R^2 values for the ten networks being considered.

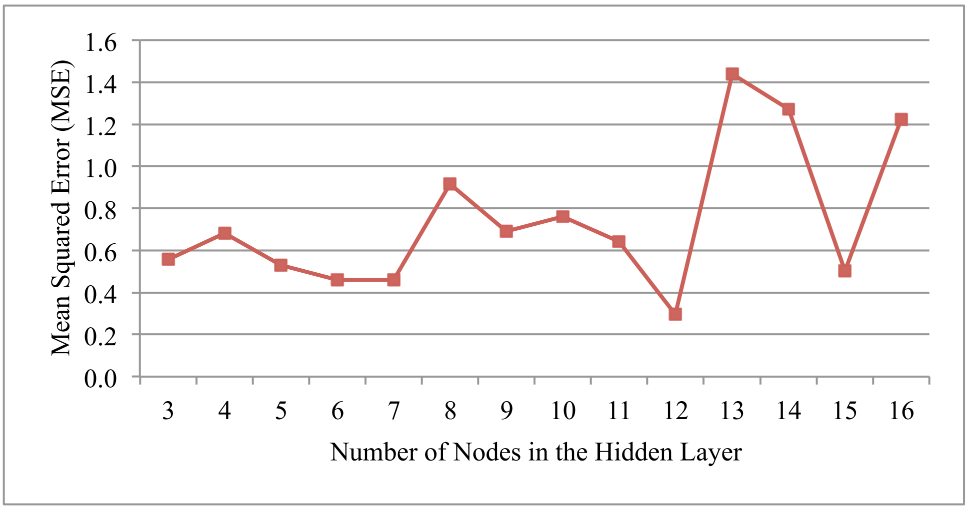

FIG. 4: *Mean Squared Error (MSE) for different number of nodes in the hidden layer*

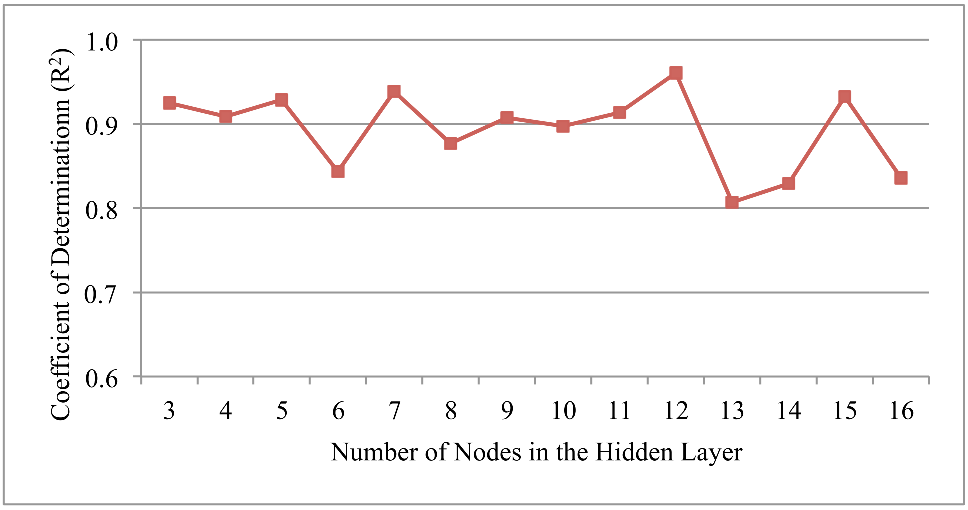

FIG. 5: *R^2 Values for different number of nodes in the hidden layer*

It can be seen that the network with 12 nodes in the hidden layer presented the least MSE and the highest R^2 value among all networks indicating the best network performance. Accordingly, it was adopted for the developed cost estimating model. The network's training time was 37 seconds on an IBM compatible computer with 2MB RAM and the best validation check was 0.11395. The interruption of the network's training occurred after 1,005 learning epochs. The least MSE associated with the training set was 0.00497.

Unlike the BPNN, the design of the PNN is straightforward and does not depend on training parameters as in BPNN (Demuth and Beale, 1998; Sinha and Pandey, 2002). Similar to PNN is, the design of the GRNN; straightforward and does not depend on training parameters, but a smoothing factor is applied after the network has been trained. GRNN measures how far a given sample pattern deviates from patterns in the training set. When a new pattern is presented to the network; that input pattern is compared to all the patterns in the training set to determine its distance from these patterns. The output of the network is a proportional amount of all the outputs in the training set (Demuth and Beale, 1998; Specht, 1991).

The development of the NN models was performed using MATLAB Neural Network Toolbox (Demuth and Beale, 1998). For comparison purposes, a regression analysis was also performed using the data set used in the trained examples. The performance of the neural network models compared with the actual cost and the results of the regression analysis is shown in Fig. 6 and summarized in Table 7.

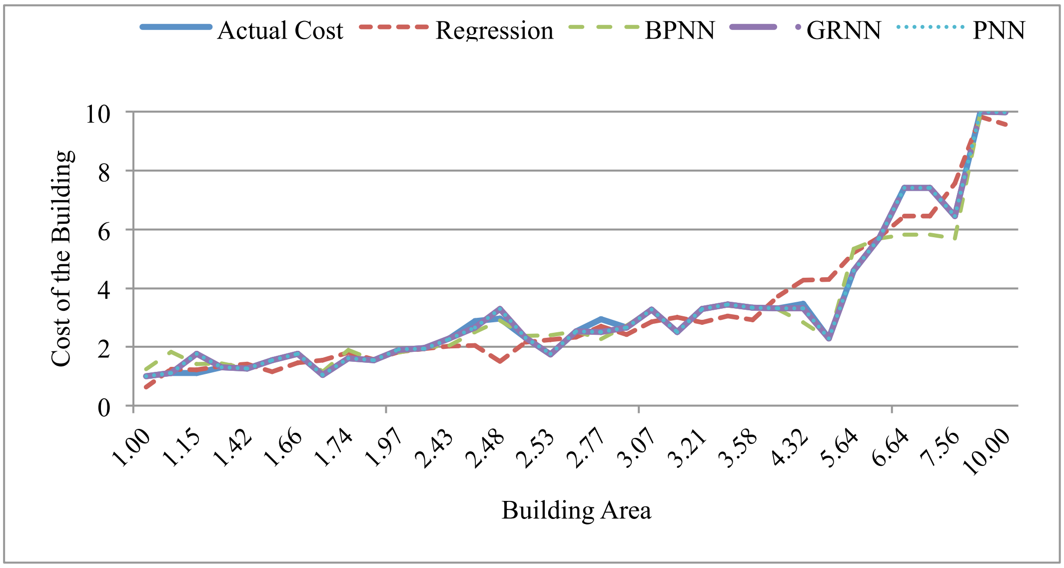

FIG. 6: *Cost Comparisons as a Function of Building Area*

Table 7 shows the R^2 values for the training set compared with the linear regression model. It can be seen that the performance of the neural network models is better than the linear regression model as one of the major benefit of NNs is their ability to model the nonlinearity of the data (Emsley et al. 2002). It can also be shown that the PNN and the GRNN successfully described the variability of the trained data.

TABLE 7: *Performance of Different Trained Models*

	Regression	BPNN	GRNN	PNN
R^2 Values	0.94622	0.9605	1.0000	1.0000

6. VALIDATION OF THE NEURAL NETWORK MODELS

For validation of the developed neural network models, the available sample data were divided into two groups. The first group, a training and testing data set, was used for the development of the neural network models. The second group, a validation data set, was then used for model to examine the predictive capability of the developed models. It was decided that 85% of the data (30 projects) would be a training and testing data set and 15% of the data (5 projects) is selected randomly for model validation. The data were split in this way to maximize the cases available to train and validate the neural network while still providing sufficient cases to provide an independent test of the network's performance.

For comparison purposes, the trained networks were used to estimate the cost of the projects in the verification set. The performance of the neural network models compared with the actual cost and the results of the regression analysis is shown in Fig. 7 and summarized in Table 8. Fig. 7 and Table 8 show that the overall neural network models performance was more accurate than the linear regression model for the data contained in the validation set. The mean absolute percentage error (MAPE) was equal to 16.83%, 19.35%, and 19.29% for the BPNN, GRNN, and PNN respectively, whereas, it was equal to 23.72% for the regression model.

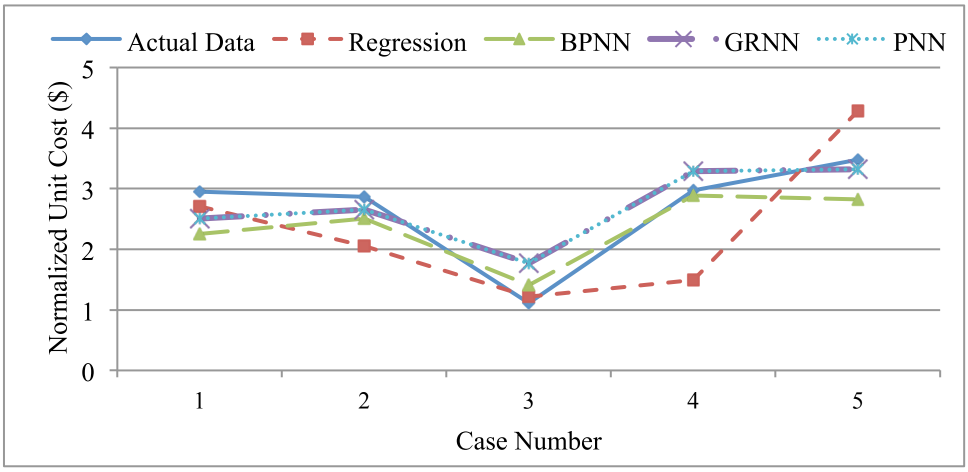

FIG. 7: *Cost Comparisons for Validated Data Set*

TABLE 8: *Normalized Cost Estimated by Different Models*

Project Area	Actual Cost	BPNN		GRNN		PNN		Regression	
		Cost ($)	%	Cost ($)	%	Cost ($)	%	Cost ($)	%
2.7701	2.9536	2.2577	-23.56	2.5110	-14.99	2.5118	-14.96	2.7134	-8.13
2.4775	2.8717	2.5086	-12.64	2.6580	-7.44	2.6576	-7.46	2.0513	-28.57
1.1517	1.1113	1.4081	26.71	1.7690	59.18	1.7656	58.88	1.2135	9.20
2.4775	2.9690	2.8929	-2.56	3.2880	10.74	3.2881	10.75	1.4982	-49.54
4.3153	3.4765	2.8268	-18.69	3.3230	-4.42	3.3227	-4.42	4.2817	23.16

7. EFFECT OF THE NUMBER OF INPUT TRAINING CASES

It has been known that the number of input training cases plays an important role in the estimating accuracy (Bode, 1998; Bode, 2000; Wang et al., 2000). This statement was investigated by using the same data and eliminating some input data each time and using the same network structures. The number of input data was changed from 10 to 30 cases and results was compared with the targeted actual outputs for the validated data. The performance of the different models is presented in Fig. 8 and summarized in Table 9. Fig. 8 shows that the linear regression model is very sensitive to the change of change of the number of the training data as the MAPE decreased from 68.36% for 10 training samples to 23.72% for 30 training samples. However, the results obtained from neural network models showed that the neural network models less sensitive to the change of change of the number of the training data than the linear regression model as the MAPE decreased from 26.28%, 19.59% and 37.20% for 10 training samples to 16.83%, 19.35% and 19.29% for 30 training samples for BPNN, PNN and GRNN network respectively.

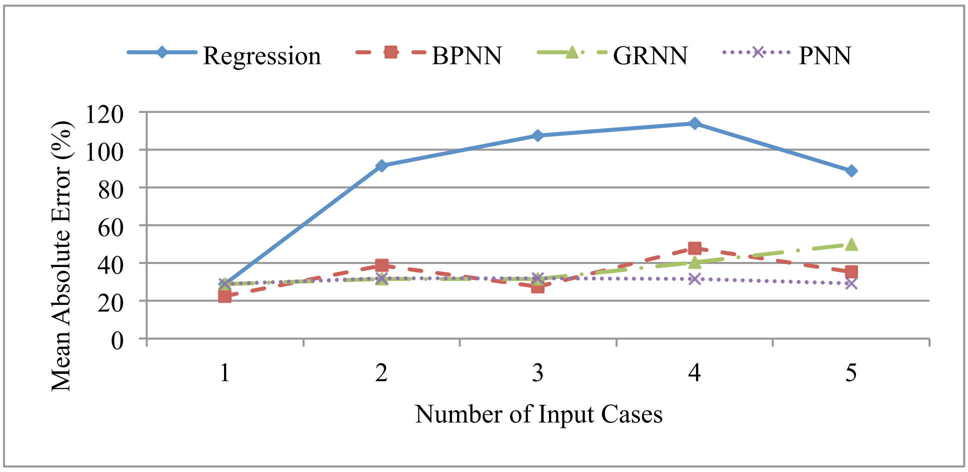

FIG. 8: *Mean Absolute Percentage Error (MAPE) for the Validated Data*

Table 9 shows that the PNN network was the most stable network among all the other estimating models as the maximum difference in MAPE percentage was only 2.46%. However, the maximum difference in MAPE was 19.47%, 17.91%, and61.45% for BPNN, GRNN and regression models respectively. Therefore, it can be concluded that the PNN network is the most suitable model for cost estimating for cost estimating when there is a small number of input data.

TABLE 9: *Range of MAPE for Different Models*

	Mean Absolute Percentage Error (MAPE)			
	Regression	BPNN	GRNN	PNN
Minimum MAPE %	23.72	16.83	19.29	19.35
Maximum MAPE %	85.17	36.30	37.20	21.81
Maximum difference in MAPE%	61.45	19.47	17.91	2.46

8. CONCLUSIONS

An initial pilot study was made, where potentially cost significant variables were identified and data is available for a relatively small number (12) of timber bridge projects. Three different NN models were designed and trained with different number of input variables i.e. BPNN, GRNN and PNN. The results are then compared with those obtained from a linear regression model using the same input variables. The results showed that the estimating error for all models decreased as more input variables were used i.e. 3 variables. It also showed that the PNN model showed the best performance among all models. The MAPE for the PNN ranges from 1.91% to 6.86% and that MAPE for the best models were 8.23%, 5.09%, 5.25%, and 1.91% for regression, BPNN, GRNN, and PNN respectively.

A full scale study was made using a real data for 35 low-rise structural steel buildings fabricated and built between 1993 and 1997 in which more sophisticated analysis were developed. Three different neural network models have been developed for estimating the direct cost at the pre-design stage when there is insufficient definition of scope and characteristics for detailed estimating using BPNN, PNN, and GRNN. The models were developed and trained using MATLAB Neural Network Toolbox. Fifteen percent of the projects data were randomly extracted to validate the performance of the trained networks. A linear regression analysis was also performed using the same trained examples. The results obtained from the neural network models showed that the mean absolute percentage error (MAPE) was equal to 16.83%, 19.35%, and 19.29% for the BPNN, GRNN, and PNN respectively, whereas, it was equal to 23.72% for the regression model.

The effect of the number of training input cases was also studied on the performance of the different models using the same data and eliminating some input data each time. The number of input data was changed from 10 to 30 cases and results was compared with the targeted actual outputs for the validated data. The linear regression model was more sensitive to the change of the number of the training data than the neural network models. The results showed that the PNN network was the most stable network among all the other estimating models as the maximum difference in MAPE percentage was only 2.46%. However, the maximum difference in MAPE was 19.47%, 17.91%, and61.45% for BPNN, GRNN and regression models respectively. Therefore, it is recommended to use the PNN network for cost estimating when there is a small number of input data.

9. ACKNOWLEDGEMENTS

The authors thankfully acknowledge Canam Manac for providing the data utilized in this study. Special thanks and acknowledgements also to Ines Siqueira for all her efforts in collecting project documents, designing the data sheet, recording project data and conducting interviews with different project professionals.

10. REFERENCES

AACE International. (2005). AACE International Recommended Practice No. 18R-97: Cost estimate classification system – As applied in engineering, procurement, and construction for the process industries. *AACE Inc., Morgantown, WV 26505-1876 USA.*

Alex D. P., Al Hussein M., Bouferguene A. and Fernando S. (2010). Artificial neural network model for cost estimation: City of Edmonton's water and sewer installation services. *Journal of Construction Engineering and Management*, Vol. 136, No. 7, July 1, pp. 745 – 756.

Bode J. (1998). Neural networks for cost estimation. *Cost Engineering*, Vol. 40, No. 1, pp. 25–30.

Bode J. (2000). Neural networks for cost estimation: Simulations and pilot application. *International Journal of Production Research*, Vol.38, No.6, pp. 1231–1254

Carr R. I. (1989). Cost Estimating Principles. *Journal of Construction Engineering and Management*, ASCE, Vol. 115, No. 4, pp. 545-551.

Creese R. C. and Li L. (1995). Cost estimation of timber bridges using neural networks. *Cost Engineering*, Vol. 37, No. 5, pp. 17–22.

De la Garza, J., and Rouhana. K. (1995). Neural network versus parameter based application. *Cost Engineering*, Vol. 37, No. 2. Pp. 14–18.

Demuth H. and Beale M. (1998). Neural network toolbox for use with MATLAB. *The MathWorks, Inc.*

Douglas M. and Runger C. (2006). Applied statistics and probability for engineers 4th edition. *John Wiley and Sons, Inc.*

Dysert L. R. (2003). Sharpen Your Cost Estimating Skills. *Cost Engineering*, Vol. 45, No. 6, pp. 22– 30.

El-Sawah H. (1994). Management Applications in the Construction Industry in Egypt (With Particular Reference to Bidding Strategy). *Ph.D. thesis, civil engineering department, Helwan University, Cairo, Egypt.*

Emsley M. W., Lowe D. J., Roy D. A., Harding A. and Hickson, A. (2002). Data modeling and the application of a neural network approach to the prediction of total construction cost. *Construction Management and Economics*, Vol. 20, No. 6, pp. 465–472.

Hegazy T. and Ayed, A. (1998). Neural network model for parametric cost estimation of highway projects. *Journal of Construction Engineering and Management*, Vol.124, No. 3, pp. 210–218.

Hegazy T. and Moselhi O. (1994). Analogy-Based Solution to Markup Estimation Problem. *Journal of Computing in Civil Engineering*, Vol. 8, No. 1, pp. 72 – 87.

Hegazy T. and Moselhi, O. (1995). Elements of Cost Estimation: A survey in Canada and the United States. *Cost Engineering*, Vol. 37, No. 5, pp. 27–33.

Hegazy T., Fazio P. and Moselhi O. (1994). Developing practical neural network applications using back-propagation. *Microcomputing in Civil Engineering*, Vol. 9, No. 2, pp. 145–159.

Hendrickson C. and Au, T. (2003). Project management for construction. *Edition 2.1, Prentice Hall, Inc., N.J.* http://www.ce.cmu.edu/~cth/

Kim G. H., Seo D. S. and Kang K. I. (2005). Hybrid models of neural networks and genetic algorithms for predicting preliminary cost estimates. *Journal of Computing in Civil Engineering*, Vol. 19, No. 2, pp. 208–211.

Moselhi O. and Siqueira, I. (1998). Neural networks for cost estimating of structural steel buildings. *Proceedings of the AACE international transactions*, IT/IM.06, American association of cost engineers (AACE) International, Morgantown, WV.

Moselhi O., Hegazy, T. and Fazio P. (1991). Neural networks as tools in construction. *Journal of Construction Engineering and Management*, Vol. 117, No. 4, pp. 606-625.

Petroutsatou K., Georgopoulos E., Lambropoulos S. and Pantouvakis J. P. (2012). Early cost estimating of road tunnel construction using neural networks. *Journal of construction engineering and management*, Vol. 138, No. 6, pp. 679-687.

Pewdum W., Rujirayanyong T. and Sooksatra V. (2009). Forecasting final budget and duration of highway construction projects. *Engineering, Construction and Architectural Management*, Vol. 16, No. 6, pp. 544–557.

RS Means (2012). RS Means Building Construction Cost Data. RS Means Company Inc. Kingston, MA, USA.

Setyawati B. R., Sahirman, S. and Creese, R. C. (2002). Neural networks for cost estimation. *Proceeding, 2002 AACE International Transactions*, EST.13, American Association of Cost Engineers (AACE) International, Morgantown, WV.

Sinha S. K. and Pandey M. D. (2002). Probabilistic neural network for reliability assessment of oil and gas Pipelines. *Computer-Aided Civil and Infrastructure Engineering*, Vol. 17, No. 5, pp. 320–329.

Specht D. F. (1991). A general regression neural network. *IEEE transaction Neural Network*, Vol. 2, No. 6, pp. 568 – 578.

Wang Q., Stockton, D. and Baguley, P. (2000). Process cost modeling using neural networks. *International Journal of Production Research*, Vol. 38, No. 16, pp. 3811–3821.

Williams T. P. (2002). Predicting completed project cost using bidding data. *Construction Management and Economics*, Vol. 20, No. 3, pp. 225 – 235.

Williams T. P. (2005). Bidding ratios to predict highway project costs. *Engineering, Construction, and Architecture Management*, Vol. 12, No. 1, pp. 38 – 51.

Yuce B., Li H., Rezgui Y., Petri I., Jayan B., Yang C. (2014). Utilizing artificial neural network to predict energy consumption and thermal comfort level: an indoor swimming pool case study. *Energy and Buildings*, Vol. 80, pp. 45–56.