UTILIZING SITE CHARACTERISTICS IN NEURAL NETWORK MODELLING OF PERCENTAGE COST-TIME OVERRUN OF BUILDING PROJECTS

A. O. Ujene[1] and A. A. Umoh[1]
[1]Department of Building, University of Uyo, Uyo, Akwa-Ibom State, Nigeria

Abstract- This study evaluated the site characteristics influencing the time and cost delivery of building projects, determined the range of percentage cost and time overrun and developed a neural network model for predicting the percentage cost and time overrun using the site characteristics of building projects. The study evaluated twelve site characteristics and two performance indicators obtained from records of construction costs, contract documents, and valuation reports of 126 purposively sampled building projects spread across several cities in Nigeria. Analyses were with descriptive and artificial neural network. It was concluded that with fairly favourable site characteristics, cost overrun range reached 77.95% with a mean variation of 44.36%, while time overrun range reached 51.23% with a mean variation of 26.77%. It was found that the accuracy performance levels of 91.93% and 91.43% for the cost and time overrun predictions respectively were very high for the optimum models. Building projects have eight significant site characteristics which can be used to reliably predict the percentage overrun, among which the ground water level, level of available infrastructure and labour proximity around the site are the most important predictors of cost and time overrun. The study recommended that project owners, consultants, contractors and other stakeholders should always use the eight identified site characteristics in predicting percentage cost and time overrun, with more priority on the first three characteristics. The study also recommended the neural network prediction approach due to its prediction accuracy.

Keywords: Building Projects, Construction Industry, Cost Overrun, Site Characteristics, Time Overrun.

1.0 INTRODUCTION

Buildings are major products of the construction industry which serve as life support systems, provide shelter, encourage productivity, and embody our culture [1, 2, 3]. Housing development has been a concern of individuals, families, groups and government since the dawn of urban civilization due to its importance [4], because buildings are incredibly expensive to build and maintain, hence their economics have become a vital and complex issue throughout their lives. Some studies have reported the poor performance of the industry in cost and time of projects due to many problems [5, 1, 2]. Consequently, individual efforts to build or own a house are affected by a number of constraints which include; land

acquisition problems, low income of individuals, high cost of building materials, statutory regulation, and non-use of local materials among others [4].

Among these problems, land acquisition and its related characteristics pose the greatest difficulty to urban housing production in many developing countries [6]. The location of a building, unforeseen ground condition of the site and soil undoubtedly contribute a lot to the cost of buildings [7, 8]. The knowledge of site condition is important to reduce construction costs, minimize risks of natural hazards arising from the impact of proposed development on natural resources such as soils, vegetation and water systems [9]. Similarly, it has been observed that location will affect construction cost as the characteristics of the area will not only determine resource availability, but also the monetary value of the resources [10]. A previous study by Al-Juwairah [11] noted that geo-technical formations such as limestone, cavities, high water table and others are some site related problems that increase building costs.

Some studies have reported that the high construction costs of buildings add to the frustration of most low and medium income families [12, 13]. Ujene [3] noted that the high cost of construction and inability of construction managers to accurately forecast the cost and time behaviour of building projects have resulted in many abandoned projects, poor quality jobs, collapsed buildings, disputes among stakeholders and housing shortages among others [4]. Literature appears scarce on utilization of site characteristics to predict the time and cost performance of building in Nigeria, hence this study utilized locational features to predict the cost and time overrun of building projects in Nigeria using an artificial neural network approach.

1.1 OBJECTIVES

For the purpose of achieving the aim of this study, the first objective evaluated site characteristics influencing time and cost delivery of building projects, the second assessed the range of percentage cost and time overrun of building projects and the third developed a neural network model for the prediction of percentage cost and time overrun of building projects using the site characteristics of the projects.

The result of this study generally will broaden stakeholders' knowledge of the nature and extent to which site characteristics influence cost and time overruns. This will help to ensure planning and execution strategies that can provide the necessary construction management support for enhanced project delivery.

1.2 LITERATURE REVIEW

1.2.1 OVERVIEW OF SITE CHARACTERISTICS

The locational characteristics of sites are very vital to the planning and construction of a facility [14, 15]. Site topography is critical in evaluating storm water quality and discharge, as well as quantities of soil cut and fill [16, 17]. Other important characteristics are ground water level, bearing capacity of soil, size of the site, access to the site, land use compatibility, level of available infrastructure, proximity of material/labour and confined nature of site [3], while Amusan [8] and Memon *et al.* [7] noted that the location of a building, unforeseen ground condition of the site and soil certainly contribute a lot to the cost of building.

Oduwaye [18] identified cost of land, Okali, Okpara and Olawoye [10] identified location of projects, while Al-Juwairah [11] identified geo-technical formations such as limestone, cavities, high water table and the swell of clay fills as site related problems that influence building costs. These constituted the sources of the site characteristics used in this study.

1.2.2 PERFORMANCE AND OVERRUN IN CONSTRUCTION

Project performance is about meeting or exceeding stakeholders' needs and expectations from a project. It invariably involves placing consideration on major project elements of time, cost, quality, health and safety, profitability, productivity and others [19]. It has been pointed out that in today's highly competitive and uncertain business environment, the client, who is the major stakeholder, wants speedier delivery of their project with early start of construction work, certainty of performance in terms of cost, quality and time, value for money for their investment, minimal exposure to risk, and early confirmation of design and price or cost [20].

Studies have noted that client needs are generally in terms of time, cost and quality. Thus, they serve as bases of measuring the success of the projects [21, 22]. Cost/time overrun simply refers to the difference between the value originally envisaged for a project and the value reflected in the final certificate, serving as measures of assessing the cost and time performance of projects [23] as adopted in this study.

1.2.3 ARTIFICAL NEURAL NETWORK

Artificial neural network (ANN) is a computational model that is inspired by the structure or functional aspects of biological neural networks. They are characterized in principle by a network topology, a connection pattern, neural activation properties, training strategy and ability to process data [24]. Jamous [25] further noted that ANN are distributed, adaptive, generally nonlinear learning machines built from many different processing elements (PEs) which communicate by sending signals to each other over a large number of weighted connections. It is a preferred tool for many predictive data mining applications because of their power, flexibility, and ease of use. Every parallel-distributed processing model such as the neural network has eight components, namely the processing elements (PE) or neurons, the activation function f, the output function, the connectivity pattern (wi,R), the propagation rule, the activation rule, the learning rule and the system's operating environment [25].

The most common neural network model is Multilayer Perceptron (MLP), known as a supervised network because it requires a desired output to learn so as to create a model that correctly maps the input to the output using historical data. In a single hidden layer MLP, inputs are fed and get multiplied by interconnection weights as they are passed from the input layer to the hidden layer. Within the hidden layer, they get summed then processed by a nonlinear function (usually the hyperbolic tangent). Finally the data is multiplied by interconnection weights then processed one last time within the output layer to produce the neural network output [24].

2.0 METHODOLOGY

The research method is qualitative and quantitative in nature, and is appropriately in line with the positivist belief. It conforms to a construction management field that appears to be firmly rooted within the positivist tradition that facilitates the application of either pure or applied

research in the built environment [26]. Data were obtained from records of construction costs, contract documents and valuation reports of 216 purposively sampled building projects completed between 2010 and 2013 in Nigeria. Cost and time overrun as well as twelve site characteristics identified from literature and focused group discussions were extracted from the documents. The variables are described in Table 1.

The independent variables comprised five factors and seven covariates. The factors which were ordinal variables were coded using the thermometer coding. This is implemented by specifying a set of binary input nodes, each of which is either on (1) or off (0). The numbers of nodes are usually one less than the numbers of categories [27]. The covariates which were continuous variables were normalized to take values between -1 and 1 as applicable in artificial neural networks. In this study several algorithms resulting from various architectures and units of one hidden layer were obtained by dropping less important variables. The 216 cases were partitioned into 70% for training, 20% for testing and 10% for holdout. The outputs were saved and compared for optimum prediction performance. The performance assessment was based on the coefficient of determination (R^2), mean absolute error (MAE), the mean absolute percentage error (MAPE) and relative squared error (RSE) [24, 28].

Table 1: Descriptions of variables and their measurement

Variable (Independent)	Variable Measurement	classification	Unit Entry
Soil Strength	Soil bearing capacity measured in Newton per square millimetre	Covariates	KN/m^2
Site Topography	percentage of height to horizontal distance in metres, categorised into: Flat slope, medium slope and steep slope	Factor	Binary coding
Material Proximity	Availability and nearness of site to supply source, categorised into: nil, low, moderate, high and very high availability	Factor	Binary coding
Labour Availability	Availability and nearness of site to supply source, categorised into: nil, low, moderate, high and very high availability	Factor	Binary coding
Ratio of Plot Undeveloped	Ratio of area of plot undeveloped to total area of site	Covariate	Ratio
Level of Infrastructure	Nominal ratio of the available infrastructure around site to the expected stock of infrastructure	Covariate	Ratio
Ground Water Level	Depth of water level from ground surface	Covariate	metre
Accessibility	Ease of transport and supply of resources, categorised into: nil, low, moderate, high and very high accessibility	Factor	Binary coding
Size of Site	Total area of site	Covariate	Square metre
Area of Site Developed	Total portion of site occupied by the developed building	Covariate	Square metre
Land Use Compatibility	Integration with adjoining properties/environment, categorised into: nil, low, moderate, high and very high compatibility	Factor	Binary coding
Cost of Land	Monetary value of acquiring site	Covariate	Naira
(Dependent)			
Cost Overrun	Difference between initial and final total construction cost	Continuous	Naira
Time Overrun	Difference between initial and final construction duration	Continuous	Weeks

The coefficient of determination measured the variance that is interpreted by the model. The model has healthy predictive ability when it is near to 1 and poor prediction when it is near to 0. Mean absolute error (MAE) measures the average deviation of the predicted values from the corresponding observed values; the lower MAE the better is the long term model prediction. Mean absolute percent error (MAPE) corrects the 'cancelling out' results and also keeps into basis the different scales at which this measure can be calculated and thus can be used to analyze different predictions. Relative squared error (RSE) is the aggregate squared error produced relative to what the error would have been if the prediction had been the average absolute value. Lower RSE is the better model prediction. According to Wilmot and Mei [29], accuracy performance (AP) is defined as (100−MAPE) %, hence a smaller MAPE and higher R^2 implies better accuracy performance. The R^2 was obtained from the prediction chart, while the expressions of other measures are shown in Equations 1, 2 and 3.

$$MAE = \frac{1}{N} \sum_{i=1}^{N} \left| \hat{y}_i - y_i \right| \qquad \text{…………….....1}$$

$$MAPE = \frac{1}{N} \sum_{i=1}^{N} \frac{|\hat{y}_i - y_i|}{y_i} \times 100 \qquad \text{…………….2}$$

$$RSE = \frac{\sum_{i=1}^{N} [\hat{y}_i - y_i]^2}{\sum_{i=1}^{N} [\text{Mean}(y_i) - y_i]^2} \qquad \text{…………….3}$$

Where y_i and $\hat{y}_i$ are observed and predicted values [28].

3.0 RESULTS AND ANALYSIS

3.1 EVALUATION OF SITE CHARACTERISTICS

The results of the descriptive analysis of the five characteristics categorized as factors and measured qualitatively, and the seven characteristics categorized as covariates measured quantitatively are presented in Tables 2 and 3.

Table 2 shows that over 90% of the project sites had flat and moderate slope with very few sites having steep slopes. This implies that building developments are scarcely carried out on sites with steep slopes in line with the observation by Omole and Owoeye [17]. The results indicate that the buildings were sited at places with moderate and high material/labour availability and proximity with over 80% of the projects with these characteristics. The majority of the sites were found to fall within medium and high accessibility with about 83% of the sites falling within the stated range. Similarly about 82% of the sites evaluated were within the range of medium and high compatibility.

The results generally support the observation by Al-Juwairah [11] that developers most often strive to embrace favourable site factors. The implication, however, is that when such choice is beyond the control of the developer, the performance of the project will be affected.

Table 2: Descriptive analysis of site characteristics (factors)

Characteristics	Sub-characteristics	Frequency	Percentage

Site Topography	Flat slope	128	59.3
	Moderate slope	79	36.6
	Steep slope	9	4.2
	Total	216	100.0
Material Availability and Proximity	Low availability	20	9.3
	Medium availability	105	48.6
	High availability	72	33.3
	Very high availability	19	8.8
	Total	216	100.0
Labour Availability and Proximity	Low availability	20	9.3
	Medium availability	110	50.9
	High availability	71	32.9
	Very high availability	15	6.9
	Total	216	100.0
Site Accessibility	Low accessibility	24	11.1
	Medium accessibility	96	44.4
	High accessibility	84	38.9
	Very high accessibility	12	5.6
	Total	216	100.0
Site Compatibility	Low compatibility	23	10.6
	Medium compatibility	96	44.4
	High compatibility	81	37.5
	Very high compatibility	16	7.4
	Total	216	100.0

Table 3: Descriptive analysis of site characteristics (covariates) and percentage overrun

Characteristics	N	Minimum	Maximum	Sum	Mean	Std. Deviation
Soil Strength	216	130.00	300.00	43524.00	201.500	41.256
Ratio of Plot Undeveloped	216	0.45	0.95	151.92	0.699	0.129
Level of Infrastructure	216	0.43	1.00	155.65	0.721	0.199
Ground Water Level	216	1.00	10.00	763.20	3.533	1.928
Site Area	216	452.00	2778.76	271525.69	1257.063	516.210
Land Value	216	1399734.00	9725660.00	8.46E8	3.9152E6	1.698E6
Area of Plot Developed	216	58.56	1298.14	86302.05	399.5465	284.19337
Percentage Cost Overrun	216	12.61	90.56	9582.00	44.361	17.376
Percentage Duration Overrun	216	10.08	61.31	5782.37	26.770	10.305

The results in Table 3 show that the soil strength of the sites investigated ranged between $130 KN/m^2$ and $300 KN/m^2$ with a mean value of $201.5 KN/m^2$. This implies that a majority of the sites had adequate bearing capacity above the value of $140 KN/m^2$ adequate for low rise buildings [30]. The majority of the sites investigated had adequate working space with average of about 69.9% of undeveloped area in most of the sites which is greater than the 40% specified by most development authorities in Nigeria as noted by Aluko [31]. The level of available stock of infrastructure also showed that most of the sites were located in areas with about 72% of expected stock of infrastructures. This indicates that most sites were located around urban areas in line with the observation by Oisasoje and Aidelunuoghene [32], that facilities location in Nigeria is very lopsided with the urban centres having more concentration of infrastructure than the rural areas. The evaluation of the average depth of site water level across the seasons shows a mean value of 3.53m from surface level. The

results indicate that a majority of the sites have good aquifers presence mostly found in the Tertiary and Quaternary sediments of the southern coastal areas (Benin and Niger Delta Basins), where the water level varies between 0 and 9 m [33].

The results show an average plot size of 1257 square meters per project, with average value of ₦3,115.00 per square meter. The high cost of land may be attributable to the demand as well as the fact that customary owners both within and outside city boundaries sell their land without reference to the public authorities, who are in charge of land administration and control in Nigeria [34]. The result in Table 3 also indicates that a majority of the sites investigated had about 31.5% of the total area developed which is less than the maximum of 60% specified by most development authorities in Nigeria [31]. The study generally showed that with fairly favourable site characteristics, cost overrun range reached 77.95% with a mean value of 44.36% variation, while time overrun range reached 51.23% with a mean value of 26.77% variation. It therefore implies that with harsh site characteristics percentage overrun may rise by more than a hundred percent [35].

3.2 ARTIFICIAL NEURAL NETWORK MODEL

The model analysis in this study utilized the ANN algorithm built into SPSS17 software. The algorithm implements automatically the multilayer perceptron neural network with gradient descent learning, hyperbolic tangent hidden layer activation function and identity output layer activation function, the process of which has been discussed in the methodology. The output result of the software displays the following useful information presented only for the best performed model in appendix.

(a) Network information –Appendix I displays information about the neural network, including the 2 dependent variables, 5 input factors and 3 input covariates, which were rescaled by normalization. The network had one hidden and 5 units in the hidden layer with identity activation function. The result is similar to the form by Chukwu and Nwachukwu [24], hence can be reliable for adoption.

(b) Diagram - Appendix III displays the network diagram as a non-editable chart. It shows that as the number of covariates and factor levels increases, the diagram becomes more complex and difficult to interpret, an indication of some complex relationship between the variables.

(c) Synaptic weights - Displays the coefficient estimates that show the relationship between the units in a given layer to the units in the following layer. The synaptic weights are based on the training sample, the number of which can become rather large. These weights are generally not used for interpreting network results hence not shown.

(d) Network performance - Displays results used to determine whether the model is "good". The charts in this group which are based on the combined training and testing samples have fairly linear goodness, with high coefficient of determination as shown in Appendix IV.

(e) Model summary – This displays a summary of the neural network results by partition and overall, including the error, the relative error or percentage of incorrect predictions, the stopping rule used to stop training, and the training time. The error is the sum-of-squares error when the hyperbolic tangent activation function is applied to the output layer. It is the cross-entropy error when the softmax activation function is applied to the output layer.

Relative errors or percentages of incorrect predictions are displayed depending on the dependent variable measurement levels. The model summary results were used for the basis of the prediction performance discussed in the next section.

3.3 COMPARISON OF PREDICTION PERFORMANCE OF SEVERAL MODEL ALGORITHMS

The result for the assessment of prediction performance of cost overrun algorithms based on the criteria described in the methodology is presented in Table 4.

Table 4: Results of prediction performance of cost overrun algorithms

S/no of algorithm	No of inputs	No of units in jidden layer	R^2	MAE	MAPE	RSE	AP
1	12	3	0.942	3.494	9.286	0.006	90.714
2	11	3	0.825	5.520	15.581	0.006	84.419
3	10	4	0.839	5.578	15.151	0.005	84.849
4	9	4	0.928	3.471	8.921	0.009	91.079
5	8	5	0.943	2.947	8.073	0.008	91.927
6	7	5	0.904	4.249	10.64	0.010	89.360
7	8	6	0.733	7.522	19.171	0.004	80.829

Coefficient of determination (R^2), mean absolute error (MAE), mean absolute percentage error (MAPE), Relative squared error (RSE), accuracy prediction (AP)

The results in Table 4 show that the fifth algorithm obtained with eight variable inputs and five hidden layers had an overall accuracy performance of 91.93%. This prediction accuracy level for cost overrun prediction is observed to be higher than the other algorithms.
Similarly, the result for the assessment of prediction performance of duration algorithms based on the same criteria described in the methodology is presented in Table 5.

Table 5: Results of prediction performance of time overrun algorithms

S/no of algorithm	No of inputs	No of units in hidden layer	R^2	MAE	MAPE	RSE	AP
1	12	3	0.921	2.321	9.895	0.003	90.105
2	11	3	0.785	3.663	15.966	0.011	84.034
3	10	4	0.857	2.954	12.505	0.010	87.495
4	9	4	0.903	2.421	10.141	0.020	89.859
5	8	5	0.936	1.965	8.555	0.010	91.445
6	7	5	0.866	2.979	12.344	0.015	87.656
7	8	6	0.699	4.581	17.760	0.008	82.240

Coefficient of determination (R^2), mean absolute error (MAE), mean absolute percentage error (MAPE), relative squared error (RSE), accuracy prediction (AP)

The result in Table 5 shows that the fifth algorithm obtained with eight variable inputs and five hidden layers had an overall accuracy performance of 91.43% for the duration overrun prediction, which is higher than the other algorithms. These accuracy performance levels for the cost and time overrun predictions are better than the 80% accuracy obtained by ElSawy, Hosny and Razek [36] from a neural network model for construction projects site overhead cost estimating in Egypt and comparable to the 93% accuracy level achieved by Gunaydin and Dogan [37] from a neural network model built to estimate the cost in early phases of building design process, as well as the 93.7% accuracy performance obtained by Abujamous,

Rustom and Abukmail [38] from the investigation of influencing cost factors in road projects in Gaza Strip using artificial neural network. A potential application of this research is to cut down surplus budget amount to complete a construction project.

The model summary of the fifth algorithm with the optimum accuracy performance (Table 6) shows that the average overall relative errors of the sample partitions are training (0.057), testing (0.082), and holdout (0.148). This gives a range of 0.091 which is less than the 0.190 range obtained by Kitikidou and Iliadis [49]. This implies that the average overall relative errors are roughly equal across the samples which give some confidence that the model is not over-trained and that the error in future cases scored by the network will be close to the error reported in this result. The result also shows that the estimation algorithm stopped because the error did not decrease after a step in the algorithm. These are generally indications that the model has been obtained under expected conditions required to give reliable results applicable under the circumstances of the study.

Table 6: Model summary for the optimum algorithm (No 5)

Characteristics	Sub-characteristics	Values
Training	Sum of Squares Error	0.422
	Average Overall Relative Error	0.057
	Relative Error for Scale Dependents of Percentage Cost Overrun	0.052
	Relative Error for Scale Dependents of Percentage duration Overrun	0.064
	Stopping Rule Used	1 consecutive step(s) with no decrease in error[a]
	Training Time	0:00:00.172
Testing	Sum of Squares Error	0.110
	Average Overall Relative Error	0.082
	Relative Error for Scale Dependents of Percentage Cost Overrun	0.091
	Relative Error for Scale Dependents of Percentage Duration Overrun	0.068
Holdout	Average Overall Relative Error	0.148
	Relative Error for Scale Dependents of Percentage Cost Overrun	0.099
	Relative Error for Scale Dependents of Percentage duration Overrun	0.196

a. Error computations are based on the testing sample.

In order to appreciate the contributions of the independent variables to prediction of the percentage cost-time overrun, the result of the variable importance is presented in Table 7.

Table 7: Result of the independent variable importance

Site Characteristics	Importance	Normalized Importance	Rank
Ground Water Level	0.169	100.0%	1
Level of Infrastructure	0.147	87.3%	2
Labour Proximity	0.138	81.8%	3
Accessibility	0.127	75.0%	4

Site Compatibility	0.121	71.9%	5
Site Topography	0.105	62.4%	6
Ratio of Plot Undeveloped	0.105	62.3%	7
Material Proximity	0.088	52.1%	8

Table 7 and Appendix II show that the ground water level, level of available infrastructure and labour proximity around the site are the most important predictor of cost and time overrun, while the ratio of plot developed and the material proximity to site were least important predictors. Although the result shows that all eight variables had more than 50% importance, signifying some level of indispensability of all the eight variables to the prediction of cost and time overrun, yet developers may lay more emphasis on ground water level which affects the soil bearing capacity and selection of foundation, level of available infrastructure provided by the government and labour availability/proximity which reduce the burden of development cost and facilitate prompt delivery.

4.0 CONCLUSIONS

This study has evaluated site characteristics influencing the time and cost delivery of building projects, determined the range of percentage cost and time overrun and developed a neural network model for predicting the percentage cost and time overrun using the site characteristics, with a view to broadening and deepening stakeholder understanding of the behaviour pattern of some site related factors for enhanced prediction of project performance. The study evaluated twelve characteristics and two performance indicators. It was concluded that building developments are scarcely carried out on sites with steep slopes as site development is largely dependent on the absence of obstacles such as steep topography. Buildings were mostly sited at places with moderate to high material/labour availability and proximity with medium to high accessibility and compatibility.

The study also concluded that majority of the sites had adequate bearing capacity, working space and stock of infrastructure around sites. A majority of the sites have good aquifers presence mostly found in the tertiary and quaternary sediments of the southern coastal areas with high likelihood of flooding at the peak of the raining season. The sites which were of high costs had their total developed area within regulation. The percentage cost overrun range was found to be greater than the percentage duration overrun, an indication that cost variation responds more to site characteristic than duration variation. The study generally showed that with fairly favourable site characteristics, cost overrun range reached 77.95% with a mean value of 44.36% variation, while time overrun range reached 51.23% with a mean value of 26.77% variation. It therefore implies that with harsh site characteristics percentage overrun may reach or exceed a hundred percent. It was found that the accuracy performance levels of 91.93% and 91.43% for the cost and time overrun predictions respectively were very high which can help to cut down surplus budget cost and time meant to take care of overrun in building projects. The study concluded that the model with eight site characteristics had the best prediction accuracy, with ground water level, level of available infrastructure and labour proximity around the site being the most important predictors of cost and time overrun, while the ratio of plot developed and the material proximity to site were least important predictors.

The study recommends that project owners, consultants, contractors and other stakeholders should make use of the identified eight site characteristicsof ground water level, level of

infrastructure, labour proximity, accessibility, site compatibility, site topography, ratio of plot undeveloped and material proximity to predict percentage cost and time overrun, with more priority on the first three characteristics. The study also recommends the neural network prediction approach due to its prediction accuracy.

The field of neural networks according to Chukwu and Nwachukwu [24] is very diverse and opportunities for improvement exist in many facets, including data pre-processing and representation, architecture selection, and application. This study therefore suggests that further research can be carried out to improve the performance of Neural Networks for percentage cost-time overrun prediction, perhaps through better training methods, better architecture selection, or better additional inputs of site or other characteristics.

REFERENCES

[1] Ijigah, E.A., Ogunbode, E.B., and Ibrahim, M.O. 2012. ''Analysis and Prediction of Cost and Time Overrun of Millennium Development Goals (MDGS) Construction Projects in Nigeria,'' *Developing Country Studies*, 2(10): 140-147.

[2] Ujene, A.O. ''Integrating Environmental Priority Concerns in Building Planning and Production in Niger Delta, Nigeria,'' *Journal of Architecture, Planning & Construction Management*, 4(2): 37-56, 2014.

[3] Ujene, A.O. 2012. ''Dynamics of Direct Costs of Building Elements in South-South, Nigeria. An Unpublished PhD Thesis, Department of Building, University of Uyo, Nigeria,''.

[4] Aliyu, A.A., Kasim, R. and Martin, D. 2011. ''Factors Affecting Housing Development in Makama Jahun Area of Bauchi Metropolis, Nigeria,'' *International Journal of Trade, Economics and Finance,* 2(4): 263-268.

[5] Achuenu, E. and Ujene, A.O. 2006. ''Evaluation of Material and Labour Cost of Building Elements in Nigeria, *Nigerian Journal of Construction Technology and Management*, 7(1): 99-110.

[6] Ugonabo, C.U. and Emoh, F.I. 2013. ''The Major Challenges to Housing Development and Delivery in Anambra State of Nigeria,'' *Civil and Environmental Research*, 3(4): 1-19.

[7] . Memon, A. H., Rahman, I. A., Abdullah, M. R and Abdu Azis, A. A. 2010. ''Factors Affecting Construction Cost in Mara Large Construction Project: Perspective of Project Management Consultant,'' *International Journal of Sustainable Construction Engineering & Technology, 1*(2):41-54.

[8] Amusan, L.M. 2011. ''Study of Factors Affecting Construction Cost Performance in Nigerian Construction Sites,'' Unpublished Article, assessed from: http://eprints. Covenant university. edu.ng/121/1/COST_PERFORMANCE2.doc, date assessed 28/05/2011.

[9] Wakita, O. A and Linde, R. M. 2003. ''*The Professional Practice of Architectural Working Drawing*,'' (3rd ed.), New York: John Willie & Sons, Inc., pp. 240-260.

[10] Okali, D., Okpara, E and Olawoye, J. 2001. ''Rural-Urban Interactions and Livelihood Strategies a Case of Aba and its Region, Southeastern Nigeria,'' A Working Paper Series on Human Settlements Programme, International Institute for Environment and Development (IIED).

[11] Al- Juwairah, Y.A. 1997. **''**Factors Affecting Construction Cost in Saudi Arabia,'' *Unpublished Master Thesis*, King Fahd University of Petroleum and Minerals. Saudi Arabia. Retrieved from; http://eprints.kfupm.edu.sa/10044/, accessed: 20/ 10/ 2010**.**

[12] Daramola, S. A., Alagbe, O. A., Aduwo, B. and Ogbiye, S.A. 2005. ''Public-Private Partnership and Housing Delivery in Nigeria,'' *Proceedings of African Union of Architects Congress,* Abuja, Nigeria. 23 – 28, May.

[13] Alagbe, A. O. 2010. ''Knowledge as a Determinant of Acceptability of Compressed Stabilized Laterite Bricks for Affordable Housing in Nigeria,'' Paper presented at XXXVII IAHS World Congress on Housing, October 26 – 29, Santander, Spain.

[14] UOFPP, 2003. ''Site Selection Process,'' A Report Prepared by the University Office for Facilities Planning and Programs, University of Illinois, Urbana–Champaign.

[15] Li, H., Yu, L and Cheng, E. W. L. 2005. ''A GIS- Based Site Selection System for Real Estate Projects,'' *Construction Innovation*, 5(4): 231 – 241.

[16] Monoliadis, O. G., Pantouvakis, J. P. and Tsolas, I. E. 2008. ''Compromise Programming Model in Site Selection for Construction Temporary Facilities, *International Journal of Operation Research*, 7 (3): 381-400.

[17] Omole, F. K., and Owoeye, J.O. 2012. ''Impact of Rocky Topography on Settlement Pattern and Housing Development in Idanre, Nigeria,'' *Mediterranean Journal of Social Sciences, 3* (3): 517-527.

[18] Oduwaye, L. 2009. ''Spatial Variations of Values of Residential Land Use in Lagos Metropolis,'' *An International Multi-Disciplinary Journal, Ethiopia,* 3(2): 381-403.

[19] Ogunsemi, D. R., and Aje, I. O. 2006. ''A Model for Contractors' Selection in Nigeria,'' *Journal of Financial Management of Property and Construction*, 11(1): 33-44.

[20] Takim, R. and Akintoye, A. 2002. ''Performance Indicators for Successful Construction Project Performance,'' In: Greenwood, D (Ed.), 18th Annual ARCOM Conference, 2-4 September, University of Northumbria. *Association of Researchers in Construction Management*, 2: 545-555.

[21] Fagbenle, O. I., and Oluwunmi, A. O. 2010. ''Building Failure and Collapse in Nigeria: The Influence of the Informal Sector,'' *Journal of Sustainable Development,* 3(4): 268-278.

[22] Adindu, C. C. and. Ibironke, O. T. 2012. ''Effect of Construction Risks and Uncertainties on Time and Cost Strategies of Building Contracts in Nigeria,'' *Journal of Environmental Management and Safety*, 3(2): 105 -117.

[23] Love, P. E. D., Wang, X., Sing, C. and Tiong, R. L. K. 2013. ''Determining the Probability of Project Cost Overruns'' *Journal of Construction Engineering and Management*, 139 (3): 321-330.

[24] Chukwu, S. C. and Nwachukwu, A. N. 2012. ''Analysis of Some Meteorological Parameters Using Artificial Neural Network Method for Makurdi, Nigeria,'' *African Journal of Environmental Science and Technology*, 6(3):182-188.

[25] Jamous, H. A. 2013. ''Parametric Cost Estimation of Road Projects Using Artificial Neural Networks,'' An M.Sc Thesis submitted to The Islamic University – Gaza Deanery of Graduate Studies Faculty of Engineering Civil Engineering Department Construction Management.

[26]   Dainty, A. R. J. 2008. ''Methodological Pluralism in Construction Management Research,'' In: Knight, A. and Ruddock, L. eds. *Advanced Research Methods in the Built Environment*. Oxford: Wiley-Blackwell. Pp.1-11.

[27]   Fletcher, L. 2002. ''Statistical Modeling by Neural Networks,'' A PhD thesis in Statistics, University of South Africa, South Africa.

[28]   Doreswamy, H. and Vastrad, C. M 2013. ''Performance Analysis of Neural Network Models for Oxazolines and Oxazoles Derivatives Descriptor Dataset,'' *International Journal of Information Sciences and Techniques (IJIST)*, 3(6): 1-14.

[29]   Wilmot, C.G. and Mei, B. 2005. ''Neural Network Modeling of Highway Construction Costs,'' *Journal of Construction Engineering and Management*, 131(7):765–771.

[30]   Alhassan, M., Adejumo, T. W. and Boiko, I. L. 2012. ''Classification of Subsoil Bases in Nigeria,'' *The Electronic Journal of Geotechnical Engineering*, 17: 1407- 1413.

[31]   Aluko, O. 2011. ''Development Control in Lagos State: An Assessment of Public Compliance to Space Standards for Urban Development,'' *International Multidisciplinary Journal, Ethiopia*, 5 (5): 169-184.

[32]   Oisasoje, O. M. and Aidelunuoghene, O. S. 2012. ''The Role of Public Infrastructure in Poverty Reduction in the Rural Areas of Edo State, Nigeria,'' *Research on Humanities and Social Sciences,* 2(7): 109-120.

[33]   Edet, A., Nganje, T. N., Ukpong, A. J. and Ekwere, A. S. 2011. ''Groundwater Chemistry and Quality of Nigeria: A status review'' *African Journal of Environmental Science and Technology*, 5(13): 1152-1169.

[34]   Ikejiofor, U.C. 2009. ''Planning within a Context of Informality: Issues and Trends in Land Delivery, Enugu, Nigeria. Case study prepared for Revisiting Urban Planning: Global Report on Human Settlements,'' Available: http://www.unhabitat.org/grhs/2009, accessed on 15/5/2015.

[35]   Oladapo, A.A. 2007. ''A quantitative assessment of the cost and time impact of variation orders on construction projects,'' *Journal of Engineering, Design and Technology*, 5(1), 35 – 48.

[36]   ElSawy, I., Hosny, H. and. Razek, M. A. 2011. ''A Neural Network Model for Construction Projects Site Overhead Cost Estimating in Egypt,'' *International Journal of Computer Science Issues*, 8 (3): 273-283.

[37]   Gunaydın, H. M. and Dogan, S. Z. 2004. ''A Neural Network Approach for Early Cost Estimation of Structural Systems of Buildings,'' *International Journal of Project Management*, 22: 595–602.

[38]   Abujamous, H. K., Rustom, R. N. and Abukmail, M. Y. 2014. ''Influencing Cost Factors in Road Projects in Gaza Strip Using Artificial Neural Network,'' *Journal of Engineering Research and Technology,* 1(4): 24-33.

[39]   Kitikidou, K. and Iliadis, L. 2012. ''Developing Neural Networks to Investigate Relationships Between Air Quality and Quality of Life Indicators,'' In: *Air Pollution - Monitoring, Modelling and Health*, Khare, M. (Ed.), Rijeka, Croatia: InTech. Available from: http://www.intechopen.com/books/air-pollutionmonitoring-modelling-and-health/   developing   -neural-networks-to-investigate-relationships-between-air-qualityand-quality-of-life-indi.

APPENDIX I: Network Information

Input Layer	Factors	1	Site Topography	
		2	Material Proximity	
		3	Labour Proximity	
		4	Accessibility	
		5	Site Compatibility	
	Covariates	1	Ratio of Plot Undeveloped	
		2	Level of Infrastructure	
		3	Ground Water Level	
		Number of Units[a]		22
		Rescaling Method for Covariates	Normalized	
Hidden Layer(s)		Number of Hidden Layers		1
		Number of Units in Hidden Layer 1[a]		5
		Activation Function	Hyperbolic tangent	
Output Layer	Dependent Variables	1	Percentage Cost overrun	
		2	Percentage duration Overrun	
	Number of Units		2	
	Rescaling Method for Scale Dependents		Normalized	
	Activation Function		Identity	
	Error Function		Sum of Squares	

a. Excluding the bias unit

APPENDIX II: Independent Variable Importance Chart

13
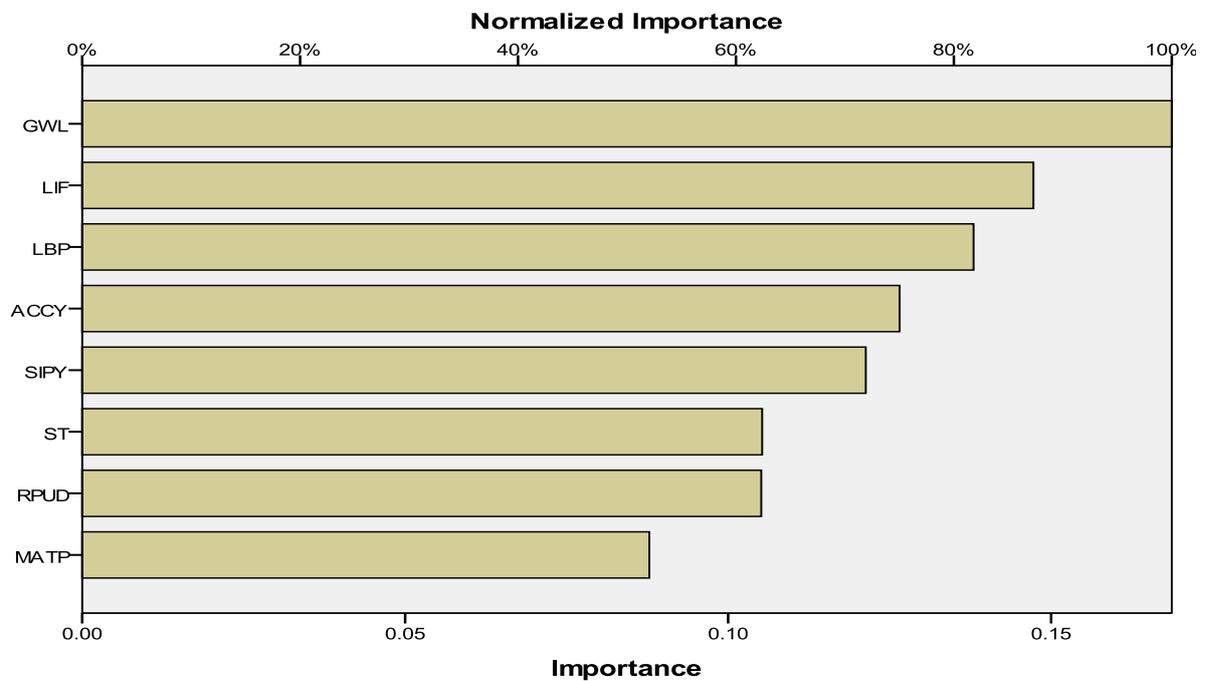
APPENDIX III: Network Diagram
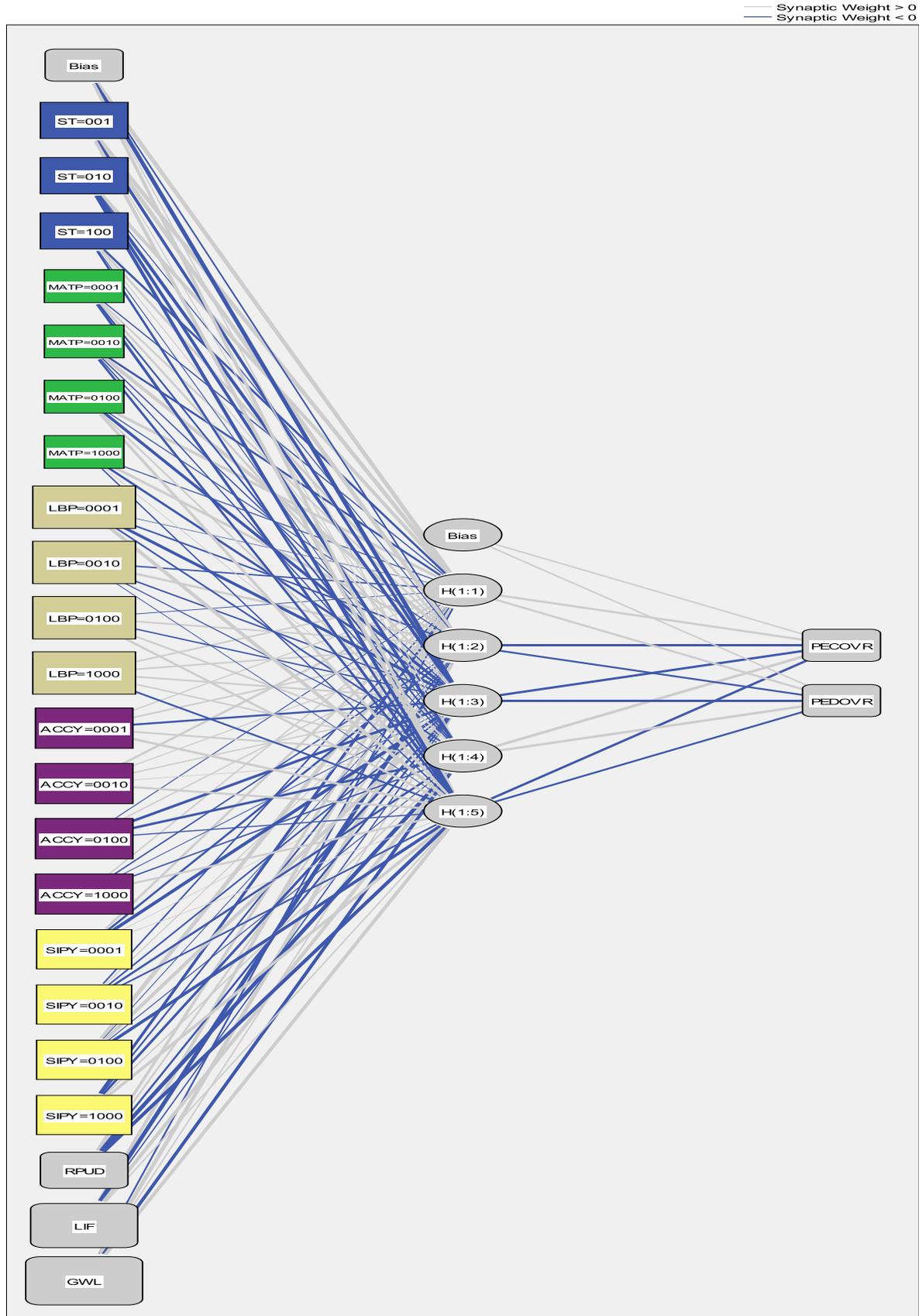
Hidden layer activation function: Hyperbolic tangent

Output layer activation function: Identity

APPENDIX IV: Predicted by Observed Charts