International Journal of Engineering & Technology
CENTRE OF PROFESSIONAL RESEARCH PUBLICATIONS
ISSN 2049-3444

International Journal of Engineering and Technology Volume 3 No. 1, January, 2013

Expert System-Based Exploratory Approach to Cost Modeling of Reinforced Concrete Office Building

Amusan Lekan Murtala, Joshua Opeyemi, Owolabi James D, Tunji-Olayeni Patricia F, Anosike Michael N
Building Technology Department College of Science and Technology,Covenant University, Cannanland, PMB 1023 Ota. Nigeria

ABSTRACT

Expert system is a conventional method that is in use in cost modeling, considering its advantage over traditional regression method. It is based on this fact, that this study aimed at deploying neural network in cost modeling of reinforced concrete office building. One hundred (100) samples were selected at random and divided into two; one part was used to develop network algorithm while the second part was used for model validation. Neural network was used to generate the model algorithm; the model is divided into 3 modules: the data optimization module, criteria selection with initializing and terminating modules. Regression analysis was carried out and model validated with Jackknife re-sampling technique. The colinearity analysis indicates high level of tolerance and -0.07403 lowest variation prediction quotients to 0.66639 highest variation quotients. Also the Regression coefficient (R-square) value for determining the model fitness is 0.034 with standard error of 0.048 this attest to the fitness of the model generated. The model is flexible in accommodating new data and variables, thus, it allows for regular updating.

Keywords: *Expert system, Model, Colinearity, Validation, Office building.*

***Corresponding Author**

1. INTRODUCTION

Construction industry is one of the sectors considered as vital in every nation's economy. It is noted to have been responsible for employment provision for active group of nation's economy. The contributory ratio of construction sector has however risen in recent time, with more jobs provision on account of high demand for house unit. This trend is believed to be instigated on account of global economic meltdown, which has resulted in more demand for space usable for commercial activities since attention is drawn to service sector more than ever before in the history of Nigerian economy. This development has resulted in space conversion into office buildings in order to meet ever increasing demand for commercial outlets. However, in order to match delivery speed with demand, factors such as efficiency of building, cost delivery system, time-cost quality target among others need to be taken into consideration [1]. Meanwhile cost has been considered as most critical in measuring the efficiency of building project delivery, therefore efforts geared toward creating an improved efficiency through an effective cost modeling will be worthwhile [2]. In modeling, the framework of modeling is as important as the model itself. Series of modeling framework had been adopted in the past which are regression based; paradigm thus needs to be shifted in the direction of conventional method that compliments the regression method shortcomings such as case base reasoning and expert system. Expert system (Artificial Neural Network) are patterned after the natural biological neurons which has ability to map input to output and deduct a meaningful inference, it has capability of studying data trend even if the series is inconsistent, once the pattern is mastered the network can generalize the trend to predict a consistent series having mastered a previous trend. It is against this background that the study carried out an exploratory approach to cost modeling of office buildings in Nigeria using an expert - based system (ANN).

2. COST MODELLING: HISTORICAL PERSPECTIVE

Building cost model may be defined as the symbolic representation of a building system expressing the content of the system in terms of cost influencing parameters. Cube method according to [3] was the first known cost model which was invented 200 years ago, floor area method was developed in 1920, while storey enclosure method was developed in 1954. According to [3]; Storey enclosure method was found to be more accurate in cost estimating than cube and floor area methods. Statistical cost modeling technique was evolved in the mid 1970's; this includes approximate quantity and optimized models. However, during this era, research efforts were delivered in the direction of validating the applicability of developed model given the seemingly applicable nature of models generated. The developed models are called regression-based models, the models are found to be limited in application as a result of non-flexibility and margin of error between input and output. Paradigm later shifted in the direction of application of expert system as advocated by [4], given the expert system background of good attributes such as capacity to accommodate large data input, consistent output, output and input mapping, consistent output, low variation error between input and output. It is to this end that this research work has generated an expert-system based cost model for office building in Nigeria.

2.1 Cost Models in Use

There are different schools of thought in cost modeling. A school of thought classified model as product-based while other classified it as process-based. Product-based model according to [5] is defined as the one that models finished product. Process-based model on the other hand synthesizes a model through the modeling process of such model. [6] and [7] presented approaches in modeling as elemental, regression, heuristics and expert system. Modelers had been using regression model since early 18th century, and this system relies on historical cost and has as its shortcomings reliance on historical cost of projects, inability to capture intervening variables that impact project such as price change, inflation change among others [5]. Similarly, [1] submitted that, area method is deficient in its cost measurement; the cost is usually influenced by factors other than floor area. Heuristic on the other hand, which has roots in Monte Carlo simulation, is also deficient in dependence on comprehensive study of systems antecedent. However, expert–based system has been found to generate less error between input and expected output, it tends to have variation error within the range of 2% to 4% while parametric model(regression model) often have variation error greater than 7% [8], [9] , and [5]. In the light of this, a robust expert system-based model that incorporates economic and environmental parameters capable of generating an accurate project cost was developed in this study and the study is limited to cost modeling of reinforced concrete office building in Nigeria.

2.2 Review of Related Works on Non-Traditional Models [Neural Networks]

There has been a number of researches carried out on the modeling of building cost variables with the aid of Artificial neural networks, some of the selected articles covers highway cost modeling, actual construction cost modeling, cost and risk estimating among others. [10] worked on risk identification using neural network, the study predicted the percentage change in the estimated cost from final cost as the index of risk measurement. Similarly, [11] carried out analysis of different methods of estimating model in use at early stage of construction works, such as regression analysis and neural network, the study concluded that neural network performed best in term of prediction accuracy. Also, [12] developed a neural network based cost estimating model and used combination of regression and neural network model to generate a regression-based model. In the same vein, [13] deployed neural network in stock market return forecasting, the study submitted that neural network can be used when an accurate results and higher trading results are desirable. It is on this premise that this study used neural network in model formulation.

3. RESEARCH METHODOLOGY

The objective of this study is to carry out an exploratory study of cost modeling of reinforced office building project in Nigeria.

3.1 Data Source

One hundred (100) samples were picked at random from projects completed within the past four (4) years at four selected locations: Ogun State, Lagos State and Federal Capital Territory (FCT) in Nigeria, these areas are regarded as economic nerve center and region of high construction activities. Initial and final cost of the sampled projects were extracted and adjusted with price index to 2008 price and prevailing inflation index to be able to capture economic variable that influences building cost. Multi Layer Perceptron Neural network with Back Propagation system and Levenberg Marqua was used as configuration frame work, from Table 1.1 Thirty-six (36) percent of the samples was used in model testing, while fourteen(14) percent was used in model training for configuration.

3.2 Model Configuration Development and Validation

The method used in model generation with Artificial neural network involves three (3) stages: the design, modeling (training) and cross validation stage.

The Design Stage: The first stage involves the design of suitable neural network topology. Neural network architecture and multi-layer perceptron with back propagation from Neuro Solution Software (MATLAB) were used to design a suitable algorithm.

Data Description: Cost significance work package was used in this context; it involves combining the bill of quantities with similar description, construction methodology together into a package, this towed the line of submission of [9] which finds base in Pareto principle. However, in this context, the work package that belongs to 40% (items with high cost) and 60% (items with low cost) were combined. This is to ensure a holistic estimation or prediction whenever the model is being used.

The Modeling Stage: The adjusted initial and final construction cost were fed into the Multilayered Perceptron System with internal guiding principles and one layer. The principles are: data characteristics, nature of problem, data complexity, and sample data. A number of hidden layers were selected after several iterations to obtain an optimum output. An optimized output was obtained after a stable and consistent output emerged. This is often determined by trials sine there is no rule to determine it. Further configuration parameters were set as presented in Tables 1.1, 1.2 and 1.3 the parameters include the means through which the data input, output error would be displayed, display format for performance matrix and validation window. These were set before the network building button was activated.

The Model Training Stage: The model was trained after configuration; the training was stopped when the mean square error was very low. The Back propagation technique was used in this context, since it tends to reduce error between model input and output. Back propagation method develops output from input while minimizing mapping error, that is, mean square error (MSE). This is given by the following relation.

MSE = [(square root of[((summation).sub(i=1)sup.n)[(xi-E(i)]sup.2])]/n ……….. 1

Where MSE = Mean square Error, n = number of projects to be evaluated at the training phase

[x.subi] = the model output related to the sample, E = target output. Mean square error is the measure of fitness of an output, the lower the figure the fitted the output. It is as well an index of training session success. The error was noted for each of the training epoch carried out, and was stopped when the value remain constant for a given iterations of epoch. This is to prevent technical dogmatism and output over fitting when the network is presented with unseen set of data.

4. RESULTS AND DISCUSSION

4.1 Neural Network Algorithm Synthesized Output

The output of developed model is presented in Table 1.4. One hundred samples (100) of Reinforced Concrete Office Structures were used and categorized according to the period of execution that spans 2006 to 2009 as presented in Table 1.4. Highest contract sum was obtained among the projects executed in 2009 while the lowest was obtained among the 2008 projects with highest occurrence of variation is noticeable in 2006 projects and lowest among 2007 projects. Economic meltdown could be adduced as responsible for trend. Radial diagram in Fig. 1.1 was used for visualization of synthesized output for the sampled office building projects.

Distribution pattern of the As-built cost, Bill of quantities (BOQ) value and neural network predicted cost on a stretched-line radar diagram is illustrated in Fig 1.1. As-built cost value overlapped the initial value of the projects (BOQ value), this occurred from project one (1) to twenty-nine (29), where a noticeable variation occurred. Significance difference was noticed between As-built cost and neural network predicted project cost. The projects were discovered to have been completed during the economic meltdown period, this tend to tow the line of occurrence as observed in the case of 2/3 bedroom projects presented in Table 1.1. Reason suggested as responsible for this is data variation margin generalization by the neural network system used in data training for fitness so as to obtain an optimum and stabilized value.

4.2 The Testing Analysis Phase of the Developed Model

Fourteen (14) percent of the samples after network topology configuration were used in model training, the resultant model was analysed for relationship among variables. Stepwise regression analysis is carried out to investigate the relationship between a number of independent variables(initial contract sum, as-built sum and neural network output). The orrelation coefficient is presented in Table 1.5. Correlation matrix in Tables 1.5 and 1.6 indicates value of Spearman and Kendalls tau Test. The analysis indicates perfect and positive correlation between independent variables neural output and initial contract sum in spearman analysis while positive correlation exist between As-built sums and Initial contract sum. Neural output is a little higher as a result of econometric factors added unto it. Generally, linear relationship exists between the two independent variables determined by the extent of the colinearity. Summary of collinearity statistics is presented in Table 1.8, tolerance limit is large for the model variables; neural network output has value of 1.08 while contract sum has 1.00 tolerance values. In this model the two variables are regarded as very important.

4.3 Re-Sampling

Re-sampling test was conducted on the model in order to ascertain the stability and the influence of outliers on the models' stability. The results are presented in Tables 1.7 and 1.8; two models are presented here, model of as-built sum and neural network model. Neural model has standard error of 0.197 while as-built sum's model has 0.312. Generally, the two models showed stability with high level of tolerance.

4.4 Cross Validation Test on the Model

Twenty three (23) samples of one hundred (100) projects executed in 2009 were used in the model cross validation to ascertain the accuracy level, according to the analysis of report presented in Table 1.11, -0.07403 lowest variation quotients to 0.66639 highest variation quotients are obtained. Also the Regression coefficient (R-square) value for determining the model fitness is 0.034 with standard error of 0.048 this indicate the fitness of the model as good.

In modeling, variation error and prediction error determination are important. The results of analysis presented give an indication as regard validity expectation of the model. Regression analysis through the Jackknife technique also produced results revalidating stability verdict earlier obtained at network configuration stage. This method is deployed to ascertain how the model will perform when being influenced by new set of variables. Also, at all the stages, neural output has shown stable and consistent output when compared with as-built cost of projects.

5. CONCLUSION

Modeling reinforced concrete office building cost using expert system approach is presented in this study. The model is flexible in accommodating new data and variables, thus, it allows for regular updating. Neural network was used to generate the model algorithm and divided into modules, the data optimization module, criteria selection with initializing and terminating modules. The model parameters include bill of quantity value of a project, as-built sum and neural network generated output.

The neural output represents a predicted cost range for the office projects with regards to prevailing economic situation like inflation and building price index, this was factored into the as-built cost of the project and predicted upward for the period of six (6) months. Thus the specified range of prediction expressed for the model is six (6) month subject to constant economic variables; however, if economic variables change before the six month prediction window period, the cost should be adjusted with the current economic variables. Cross validation analysis indicates -0.07403 lowest variation prediction quotients to 0.66639 highest variation quotients. Also the Regression coefficient (R-square) value for determining the model fitness is 0.034 with standard error of 0.048 this attest to the fitness of the model generated.

REFERENCES

[1] Li Heng, Shen Q.P and Love Peter (2005). Cost Modeling of Office Building in Hong Kong: An exploratory study." *Facilities.* United Kingdom: Taylor and Francis. 23(9/120), PP 438-452.

[2] Ogunsemi, D.R and Jagboro, G.O (2006) "Time-Cost Model for Building Projects in Nigeria". *Journal of Construction Management and Economics.* United Kingdom: Taylor and Franchis Group. 24, PP 253-258.

[3] Skitmore, R.M, and Ng, S.T. (2003) Forecast Models for Actual Construction Time and Cost", *Building and Environment.* 38 (8), Pp 1075.

[4] Brandon, A. (1987) *Cost modeling for Estimating.* London: The Royal Institution of Chartered Surveyors.

[5] Moore C.F; Lees, T. and Fortunes, C (1999) The Relative Performance of Traditional and New Cost Models in strategic Advice for Clients. *RICS Paper Series.* London: Prentice Hall Publishers. PP 45-57.

[6] Mawdesley, M.J; Williams, A and O'Reilly M (1999). *Planning and Controlling Construction Project.* Harlow, England: Addison Wesley Longman.

[7] Ashworth, A (1994) *Cost studies of Buildings*, Third Edition. New Yorks: Longman Scientific and Technical.

[8] David, P and Seer, M (2004). Implementation Issues in Neural Networks Development. *Neurovest Journal.* United States of America: Idea Group Incorporated. 10 (7). Nov/Dec, PP 8-12.

[9] Dissanayaka, S.M and Kumaraswamy, M.M (2007) Comparing Contributory Time and Cost Performance in Building Projects. *Building and Environment.*31 (6), PP 569-578.

[10] McKim, R (2005a) Neural Network Application to Cost Engineering. Cost Engineering. *AACB International.* Vol. 35 (7), Pp 31-35.

[11] Gwang-Hee K, Sang-Hoon A and Kyung-ink, (2004). Comparison of Construction Cost Estimating Models Based on Regression Analysis, Neural Networks, and Case-Based Reasoning. *Journal of Building and Environment.* 39, PP1235 – 1242.

[12] Setyawati, B. R, Sahirman S and Creese R. C, (2007). Neural Networks for Cost Estimation. *ASCE International Transactions;* 13(9) PP 22

[13] Thawornwong, S and Enke, D (2004). Forecasting Stock Returns with Artificial Neural Networks. *Intelligent Engineering Systems through Artificial Neural Networks.* 12, PP 695-705.

Table 1.1 Selection Criteria Matrix

Data Read From Existing file	Office Building
Percentage of Training Data for Cross Validation	14
Percentage of Data for Model Testing	36
Cross Validation Exemplar	16
Test Exemplar	43
Multilayer Perceptron Input	4
Multi Layer Perceptron Processing Elements	26
Multi Layer Perceptron Exemplars	62
Hidden Layer	1

Source: 2012 Survey

Table 1.2 Supervised Learning Control Attributes [Hidden Layers]

Input layer	Output Layer
Processing Elements: 22	Processing element: 1
Transfer Tanhaxon	Transfer Tanhaxon
Learning Rule Levenberg Marqua	Learning Rule Levenberg Marqua
Momentum Step Size:1.00	Momentum Step Size:1.00
Momentum Step Size 0.70	Momentum Step Size 0.70

Source: 2012 Survey Neuro Tool

Table 1.3 Active Cross Validation Performance for Office Building

Parameters	Active Cross Validation Performance	Cross Validation Performance
Mean Square Error	0.032	0.00003
Normal Mean Square Error	0.098	346521.81
Regression Value 'r'	0.950	0.023

Source: 2012 Survey

Table 1.4 Summary of Project Adjusted Bill of Quantity and As-built Value of Office Projects

Period	Highest Initial Contract Sum (NMillion)	Highest As-built Sum (NMillion)	Lowest As-built Sum (NMilion)	Lowest Initial Contract Sum (NMillion)	Highest Variation (NMillion)	Lowest Variation (NMillion)
2009	296571798	478787280	155238227	141138227	155433571	141000000
2008	294693872	296700622	215321000	213241563	81452309	81379622
2007	276896223	282873000	114450000	111320500	165575723	3129500
2006	297323000	309873000	114450000	111320500	186002500	99875500

Source: 2010 Survey

Table 1.5 Coefficients Matrix of Reinforced Concrete Office Buildings

			Initailcontsum	Asbuiltsum	Neuraloutput
Kendall's tau_b	Initailcontsum	Correlation Coefficient	1.000		
		Sig. (2-tailed)	.		
		N	18		
	Asbuiltsum	Correlation Coefficient	.827**	1.000	
		Sig. (2-tailed)	.000	.	
		N	18	18	
	Neuraloutput	Correlation Coefficient	-.020	.140	1.000
		Sig. (2-tailed)	.909	.424	.
		N	18	18	18
Spearman's rho	Initailcontsum	Correlation Coefficient	1.000		
		Sig. (2-tailed)	.		
		N	18		
	Asbuiltsum	Correlation Coefficient	.907**	1.000	
		Sig. (2-tailed)	.000	.	

		N	18	18	
	Neuraloutput	Correlation Coefficient	-.027	.145	1.000
		Sig. (2-tailed)	.914	.565	.
		N	18	18	18

Source: Data Analysis 2012 Note: Correlation is significant at the 0.01(2-tailed)

Table 1.6 Summary of Analysis of 100 Samples of Office Building

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	.960[a]	.987	.979	22.42611	.0024	0.000	2	15	.033

Source: Data Analysis 2012

Table 1.7 Regression Coefficients of the Developed Model

Model		Unstandardized Coefficients		Standardized Coefficients	t	Significance	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	4.1398	4.1587		9.953	.000		
	As built sum	-.808	.312	.965	-2.587	.021	1.00	1.00
	Neural network cost.	.574	.197	1.089	2.919	.011	.804	1.08

Table 1.8 Model Statistics

Model	Number of Predictors	Model Fit statistics	Ljung-Box Q(18)		
		Stationary R-squared	Statistics	DF	Sig.
Asbuiltsum-Model_1	1	.008	.000	0	.000
Neural Network-Model_2	1	.034	.000	0	.000

Source: Data Analysis 2012

Table 1.9 Collinearity Diagnostics[a]

Model	Dimension	Eigen value	Condition Index	Variance Proportions		
				(Constant)	As built Sum	Neural network Sum
1	1	2.923	1.000	.01	.00	.00
	2	.064	6.759	.58	.01	.023
	3	.013	14.995	.42	.019	.077

Source: Data Analysis 2012 Notes: Dependent Variable: Neural Networks

Table 1.10 Model Fit

Fit Statistic	Mean	Square Error	Minimum	Maximum
Stationary R-squared	.021	.019	.008	.034
R-squared	.034	.048	.000	.068
Root Mean Square Error	8.1367	2.4257	6.4227	9.8517
Mean Average Percentage Error	30.184	3.878	27.442	32.926
Maximum Average Percentage Error	92.134	.936	91.472	92.796

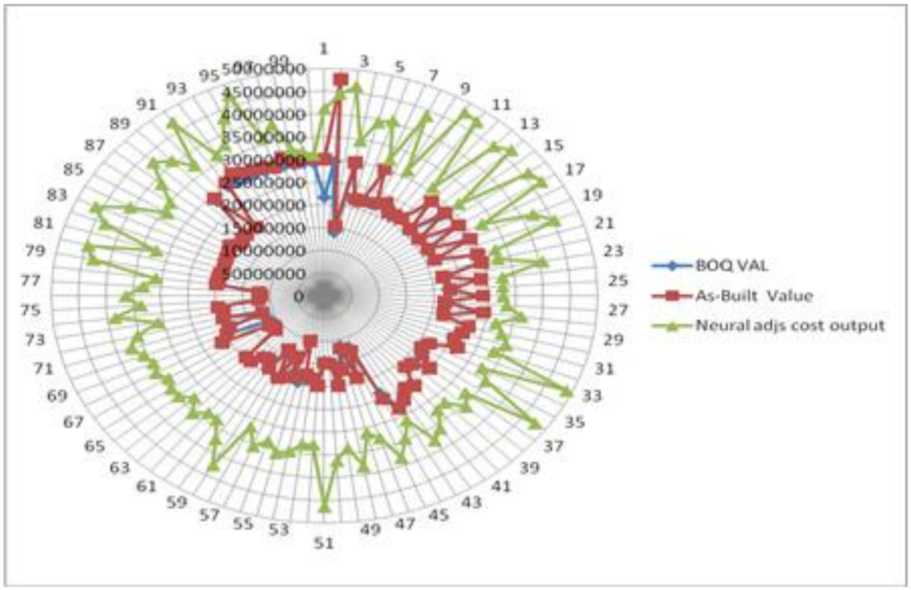
Source: Data Analysis 2012

Table 1.11 Summary of 100 sampled Reinforced Concrete Office Buildings

		1	2	3	4	5
	Project	A	B	C	D	E
Cost Centers		Boq Value [NMillion]	As-Built Value [NMillion]	Neural Adjs Cost Output [NMillion]	Variation [NMillion]	Variation Quotient
Project 1-20	1	217093854	300814387	412,797,416	111983029	0.271278416
Residential	2	296571798	478737280	445,738,080	-32999200	-0.07403271
Building	3	141138227	155238227	465,329,444	310091217	0.666390707
2009	4	290928823	298956814	348,432,150	49475336	0.141994176
	5	216996254	220856000	394,547,922	173691922	0.440230229
	6	219887135	219887136	405,878,924	185991788	0.458244508
	7	220768961	299672863	323,622,889	23950026	0.074005971
	8	220768961	225138124	438,200,127	213062003	0.48622077
	9	231136821	233268148	315,232,642	81964494	0.260012712
	10	215783222	218112136	478,307,495	260195359	0.543991808
	11	218444863	219000125	474,091,263	255091138	0.53806336
	12	219564813	221136000	310,324,221	89188221	0.287403351
	13	285763822	286144368	452,405,229	166260861	0.367504287
	14	210703023	215231000	469,007,811	253776811	0.541092931
	15	276813043	286144268	318,401,000	32256732	0.101308513
	16	211973388	213142000	460,833,922	247691922	0.537486305
`	17	288764472	290166500	470,407,364	180240864	0.383159104
	18	213671123	215850000	328,522,228	112672228	0.342966833
	19	291773632	294650000	421,535,709	126885709	0.301008209
	20	214685684	216720000	453,063,634	236343634	0.521656598
	21	293886923	294986520	328,522,229	33535709	0.102080487
	22	294693872	296700622	327,022,716	30322094	0.092721675
	23	219784963	220825120	406,183,226	185358106	0.456341114

Source: Data Analysis 2012

LIST OF FIGURES



Source: 2012 Survey

Fig. 1.1: Radar Diagrammatic Visualization of Input and Neural Output for Office Accommodation

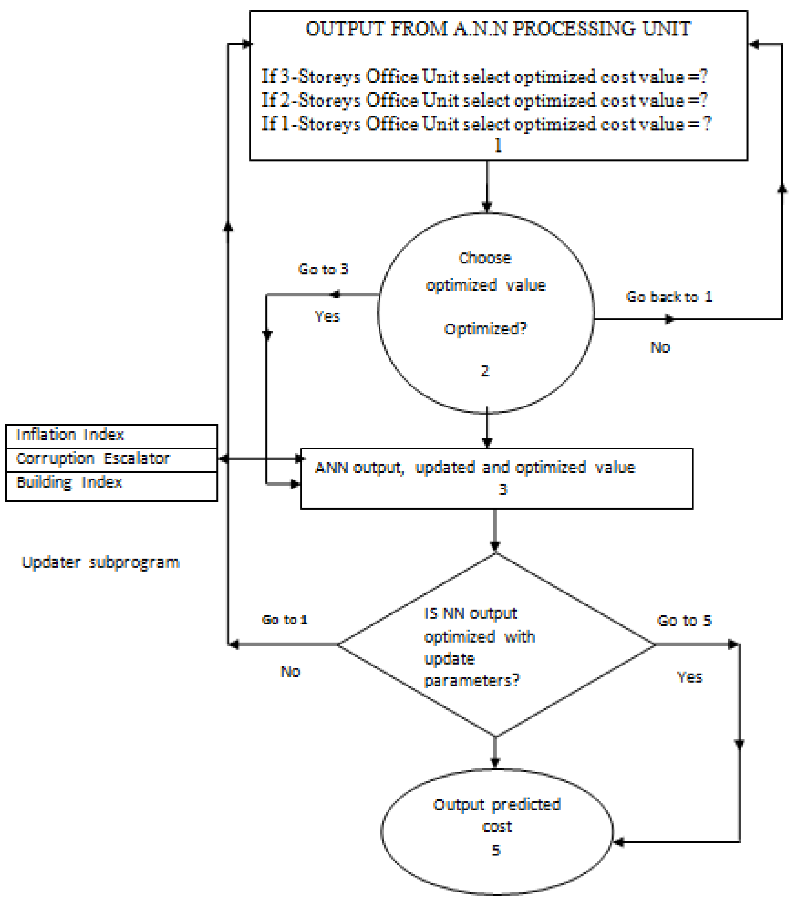


Fig. 1.2: Cost Variable Prediction Algorithm (Flow Chart)