International Journal of Architectural, Civil and Construction Sciences
ISSN: 2415-1734
Vol:9, No:1, 2015

Construction Unit Rate Factor Modelling Using Neural Networks

Balimu Mwiya, Mundia Muya, Chabota Kaliba, Peter Mukalula

Abstract**—**Factors affecting construction unit cost vary depending on a country's political, economic, social and technological inclinations. Factors affecting construction costs have been studied from various perspectives. Analysis of cost factors requires an appreciation of a country's practices. Identified cost factors provide an indication of a country's construction economic strata. The purpose of this paper is to identify the essential factors that affect unit cost estimation and their breakdown using artificial neural networks. Twenty five (25) identified cost factors in road construction were subjected to a questionnaire survey and employing SPSS factor analysis the factors were reduced to eight. The 8 factors were analysed using neural network (NN) to determine the proportionate breakdown of the cost factors in a given construction unit rate. NN predicted that political environment accounted 44% of the unit rate followed by contractor capacity at 22% and financial delays, project feasibility and overhead & profit each at 11%. Project location, material availability and corruption perception index had minimal impact on the unit cost from the training data provided. Quantified cost factors can be incorporated in unit cost estimation models (UCEM) to produce more accurate estimates. This can create improvements in the cost estimation of infrastructure projects and establish a benchmark standard to assist the process of alignment of work practises and training of new staff, permitting the on-going development of best practises in cost estimation to become more effective.

Keywords**—**Construction cost factors, neural networks, roadworks, Zambian Construction Industry.

I. Introduction

CONSTRUCTION is a vital activity in the Zambian economy. The national roads system in Zambia is experiencing a period of exceptional activity. The roads' share of GDP increased from 1.5% to 4.1% in just two years [1]. There is concern as to whether the existing road sector in Zambia is operating efficiently and whether it can handle the rising changes effectively. In Zambia, the construction industry is perceived to exhibit high margins of profit. Consequently, key stakeholders have in the recent past called for an informed position on prevailing market rates in the Zambian construction industry. As a result attention has been focused on pricing of construction unit rates.

Unit cost estimating, aided by sound engineering judgement, is the most definitive estimate technique and uses information down to the lowest level of detail available. This is the most common approach to cost estimation used in Zambia. Reference [2] describes the unit cost estimation approach as, where a unit cost is assigned to each task as represented by the bill of quantities and the total cost is the summation of the products of the quantities multiplied by the corresponding unit costs. Estimating in Zambia is carried out by a wide range of personnel who subscribe to protocols that are broadly understood, but are not consistently well documented. Approaches to estimating usually vary between the contractor's, service providers and Client's organisation and is not reflected in a documented or accepted industry standard resulting in continued inconsistencies. Industry regulators and public institutions have indicated that there was a notable trend in varying costs of construction from project to project and from one public institution to another, that it had become increasingly difficult to ascertain the true cost of projects and thereby unable to guarantee value for [3]. Construction regulatory bodies require reliability in estimation of project costs to understand prevailing market rates in the Zambian construction industry.

A review of literature showed that there was no single approach to developing construction unit rates (CUR). Generally, road construction works can be considered as a combination of two types of items. Firstly, those that can be estimated through some form of calculation, for example, the direct labour, material and equipment inputs. In this category the cost relationships between labour, material and plant are known. And secondly, indirect items such as those that cannot be calculated directly from labour, material and equipment costs. The difference between actual direct costs and prevailing market rates is referred to as the 'economic strata' or 'cost structure' which reflects the peculiarities of the local setting usually qualitative in nature such as prevailing project conditions, contractor capacity and other risk factors. Though construction cost factors may be identical internationally, it is the impact on the unit cost that varies according to a country's political, economic, social and technological inclinations.

The study focuses on deriving an objective scientific prediction of the economic strata using neural networks. Neural networks (NN) are a form of artificial intelligence capable of capturing the relations between independent and dependent variables. One of the most important and exciting characteristics of NN is their ability to learn and self-organize [4]. Reference [5] indicated that neural networks have advantages when dealing with data that does not adhere to the generally chosen low order polynomial forms, or data for

B. Mwiya is with the Department of Civil and Environmental Engineering, University of Zambia, Lusaka, Zambia (phone: +260 965 837400; e-mail: balimu.mwiya@ unza.zm).

M. Muya is Dean with the School of Engineering, University of Zambia, Lusaka, Zambia (e-mail: mmuya@ unza.zm).

C. Kaliba is with the Department of Civil and Environmental Engineering University of Zambia, Lusaka, Zambia (e-mail: chabota.kaliba@gmail.com).

P. Mukalula is with the Copperbelt University, Kitwe, Zambia (e-mail: musomuko@yahoo.com).

which there is little a priori knowledge of the appropriate cost estimating relationship (CER). A CER is functional relationship between changes in cost and the factor or factors upon which the cost depends resulting in a mathematically-fitted function [6].

To achieve the aim of the research, 25 identified construction cost factors were subjected to a questionnaire survey. The factors were reduced to eight (8) using SPSS factor analysis. The 8 factors were then analysed using NN to calculate the influence on unit costs. The resulting ratio of the factors namely political environment, contractor capacity, financial delays, project feasibility and overhead & profit was 4:2:1:1:1. Project location, material availability and corruption perception index had minimal impact on the unit cost. Quantified cost factors can be incorporated in unit cost estimation models to produce more accurate estimates. This can create improvements in the cost estimation of infrastructure projects and establish a benchmark standard. This standard will assist the process of alignment of work practises and training of new staff, allowing the on-going development of best practises in cost estimation to become more effective.

II. Literature Review

Accuracy of the estimate depends on the accuracy of available information. Foundation of unit cost-based estimating, sometimes referred to as first principle estimating, is the calculation of project-specific costs based on a detailed study of the resources, such as labour hours, material costs, equipment costs, subcontractor costs, or other unit-cost-type items required to accomplish each activity of work contained in the project work breakdown structure (WBS) or bill of quantities (BOQ). Indirect costs, overhead costs, contingency, and escalation are then added as necessary. Drawings, specifications, and project scope are used to identify activities that make up the BOQ. A BOQ serves three purposes: first and foremost it must be prepared with the objective of providing the estimator with as accurate a picture of the project as possible so as to provide a proper basis for pricing; second, it should enable the employer to compare tenders on an equal basis; and third it will be used to evaluate the work executed for payment purposes [7].

A study carried out by [8] investigated various cost estimation methods used in the Zambia construction industry. It revealed that the most common method was the use of rates based on past contracts with an allowance for inflation followed by building up of unit rates from first principles and finally use of computer software. Reference [9] observed that use of previous tender rates was common because of lack of experienced cost estimators in the bidding firm and the reduced duration of coming up with an estimate. This trend in Zambia indicated that estimate accuracy was the least requirement by the bidders perhaps due to the lack of cost estimate accuracy classification as established by professional quantity surveying bodies in other countries.

The construction economic strata consist of various factors that influence prevailing market rates used. Reference [10] stated that in cost estimation choosing cost drivers was the most important step since the model's accuracy was based upon selecting the relevant and appropriate cost drivers. Literature identifies various factors that affect construction cost with varying impacts in different parts of the world.

In Palestine, the top three factors were location of project (hot areas), segmentation of Gaza strip and closure of Gaza strip [11]. Closure and blockade of borders indicate security concerns from contractors' perspective. In the United States of America, the top three factors were project scope, land acquisition and utility relocation [12]. These factors reflect the developed nature of the country. Reference [13] revealed that fluctuation in prices of materials, cash flow & financial difficulties faced by contractors and shortage of site workers were the top three factors in Malaysia. But it was shortage of materials that Nigerians had to contend with followed by financing methods and payments for completed works and poor contract management [14]. These factors reveal an insight into the construction economic strata of the various countries. It can be deduced from literature that periodic review of cost factors was essential because of country's constant political, economic social and technological transformation.

Though there is no limit on the number of factors or variables to be used in NN. Reference [15] stated that the number of attributes assumed to have an effect on cost should be small because the architectural complexity increases with the number of attributes, requiring more training samples to reach a given accuracy, yet training samples were usually scarce in cost estimation. Literature on similar studies indicated factors of ten (10) or less. Reference [16] used 9 factors to determine preliminary estimate of time and cost in urban road construction using neural networks. Reference [17] also used 9 factors when investigating parametric cost estimation of road projects using artificial neural networks. 5 factors were employed in modelling construction labour production rates using artificial neural network [18].Reference [19] used 9 factors in developing a preliminary cost estimate of highway construction using neural networks.

TABLE I
LIST OF COST FACTORS AFTER PARETO ANALYSIS

1.	Location	8.	Exchange Rate	17.	Topography
2.	Hauling distance	9.	Contractor type	18.	Contractor selection method
3.	Delayed payment	10.	Duration	19.	Material Shortages
4.	Project planning	11.	Project scope	20.	Contractor size
5.	Material Source	12.	Detour construction	21.	Overhead & Profit
6.	Equipment availability	13.	Contract financing	22.	Corruption Perception Index
7.	Project need	14.	Labour	23.	Fuel
		15.	Contractor cash-flow	24.	Project supervision & management
		16.	Political risk	25.	Client type

Determination of the cost factors for the study used a two-step process. Firstly 45 cost factors were identified from literature. These were subjected to expert opinion through structured interviews by 10 industry experts. From the interviews, factors were ranked using frequency statistics. The

Pareto Analysis was then used to determine the vital few. Pareto Analysis is a statistical technique in decision making used for the selection of a limited number of tasks that produce significant overall effect. From the Pareto Principle (also known as the 80/20 rule) the influencing factors were reduced to 22. Three (3) factors were added to the list following expert opinion recommendation. The 25 factors subjected to a questionnaire survey are shown in Table I.

From Table I, factor 1 was the most influential and factor 22 the least influential with 23-25 as the added factors.

III. Research Methodology

To obtain a conclusive result on factors influencing road construction unit rates, the 25 factors were subjected to a questionnaire survey. Respondents were asked questions regarding the impact of cost factors on unit rates based on the Likert scale of 1 to 5 where 1 meant no impact and 5 meant extremely high impact.

A. Population and Sample Size

The targeted research population consisted of civil engineers with experience in cost estimation of roadworks. Equation (1) was used to determine the sample size of unlimited population [20], [21].

$$ss = \frac{z^2 \times p \times (1-p)}{c^2} \quad (1)$$

where: z = Z value (e.g. 1.96 for 95% confidence level); p = percentage picking a choice, expressed as decimal (0.5 used for sample size needed); c = confidence interval, expressed as decimal (0.5 = ±5); ss = sample size

$$ss = \frac{1.96^2 \times 0.5 \times (1-0.5)}{0.5^2} = 384 \quad (2)$$

The correction for finite population is:

$$NewSS = \frac{ss}{1+\frac{ss-1}{pop}} \quad (3)$$

where: pop = population

$$NewSS = \frac{384}{1+\frac{384-1}{108}} = 84 \quad (4)$$

TABLE II
TOTAL VARIANCE EXPLAINED

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	6.518	26.073	26.073	6.518	26.073	26.073	3.738	14.953	14.953
2	3.652	14.608	40.681	3.652	14.608	40.681	3.328	13.311	28.265
3	2.665	10.659	51.340	2.665	10.659	51.340	3.105	12.418	40.683
4	2.139	8.556	59.896	2.139	8.556	59.896	2.711	10.846	51.529
5	1.876	7.506	67.402	1.876	7.506	67.402	2.168	8.674	60.203
6	1.574	6.295	73.697	1.574	6.295	73.697	2.158	8.631	68.833
7	1.239	4.957	78.654	1.239	4.957	78.654	1.928	7.711	76.544
8	1.171	4.684	83.338	1.171	4.684	83.338	1.698	6.794	83.338
9	.976	3.905	87.242						
10	.866	3.463	90.706						
11	.568	2.271	92.977						
12	.485	1.939	94.916						
13	.345	1.380	96.296						
14	.293	1.172	97.468						
15	.276	1.103	98.571						
16	.180	.720	99.291						
17	.130	.519	99.810						
18	.047	.190	100.000						
19	3.664E-16	1.466E-15	100.000						
20	1.570E-16	6.279E-16	100.000						
21	1.010E-16	4.041E-16	100.000						
22	5.574E-17	2.230E-16	100.000						
23	-7.602E-17	-3.041E-16	100.000						
24	-2.913E-16	-1.165E-15	100.000						
25	-4.500E-16	-1.800E-15	100.000						

The research population targeted all the 72 large scale contractors in National Council for Construction (NCC) grades 1 – 3 in the R (road) category. At the time of the survey there were 26 registered civil engineering consulting firms with the Association of Consulting Engineers of Zambia (ACEZ) and 10 clientele organisations that dealt with estimation of roadworks. The total population size was 108 and the calculated sample size using (4) was 84. The questionnaires were distributed proportionately to the three groups as follows:

- Contractors – 84 x 72/108 = 56 (actual distribution = 30);
- Consultants – 84 x 26/108 = 20 (actual distribution = 20); and
- Clients – 84 x 10/108 = 8 (actual distribution = 8).

A total of 58 questionnaires were distributed because some contractors could not be located and the contact details with the registration body did not work. The response rate was 69%. Statistical analysis was conducted using Statistical Package for Social Sciences (SPSS). Reference [22] shows that factor analysis can yield good quality results for sample sizes less than 50. After establishing the influencing factors, prediction of the breakdown of the factors was done using NeuroShell2, an artificial neural network software.

IV. Findings And Discussion

A. Factor Analysis

In this study, SPSS factor analysis was performed to reduce the factors further by analysing the correlation between the variables and that the grouped the factors were statistically significant. The principal components extraction was used. Each eigenvalue represents the amount of variance that has been captured by one component. Table II shows the eigenvalues and proportions of variance for the components. From Table II, the total of the rotation sums of squared loadings was 1 or more in eight components.

To classify the components, an orthogonal factor rotation analysis was conducted, and the rotated component matrix was analysed, as given in Table III. The rotation method used was Varimax with Kaiser Normalization supressing variable values less than 0.400.

TABLE III
ROTATED COMPONENT MATRIX

	Component							
Variable	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Contractor_Size	.837							
Contractor_Type	.819							
Financial_Status_or_Cashflow_of_Contractor	.815							
Client_Type	.755							
Hauling_Distance		.815						
Location_of_the_Project		.814						
Fuel		.756						
Exchange_Rate		.634	.503					
Procurement	.536	.631						
Payments			.804					
Construction_Workers			.768					
Topography			.766					
Detour_Construction			.662					
Project_Scope				.817				
Contract_Financing				.751				
Project_Planning				.701				
Project_Need	.499			.527				
Project_Management	.445				.722			
Overhead_and_Profit					-.710			
Plant_and_Equipment					.680			
Material_Sources						.881		
Material_Shortages						.836		
Corruption_Perception_Index							.846	
Project_Duration							-.709	
Political_Interference								.844

The eight components were reclassified and named according to the loading of the variables in the rotated solution ensuring that the factor name is brief and communicates the nature of the underlying factors. The contractor variables loaded well on the first component reflecting how client perceives the contractor. The principal first factor thus labelled 'contractor capacity' accounts for 26.073% of the total variance and contains seven variables. The second component appears to be reflecting location of the project versus economy. It is labelled 'project location' and has five variables representing 14.608% of the total variance. The third component labelled 'financial delays' accounts for 10.659% of the total variance, with five variables and shows the effect of delayed payments on the exchange rate, construction workers and the type of work to be carried out. The fourth component is focused on the projects pre-planning activities. It is labelled 'project feasibility' with four variables and represents 8.556% of the total variance. The fifth component is more interesting, with a negative loading on overheads and profit. It has three variables reflecting overheads and profit with regards to plant, equipment and the projects management's capacity. It is labelled 'overhead and profit' and represents 7.506% of the total variance. Component six labelled 'material availability' is straight forward and reflects material factors with two variables representing 6.295% of the total variance. The seventh component is also interesting, with a negative loading reflecting concern for corruption versus duration of the project. It has been labelled 'corruption profile' with two variables and accounts for 4.957% of the total variance. The last component eight labelled 'political risk' reflects concern for political intrusion. It has one variable and accounts for 4.684% of the total variance. The renamed eight factors are shown in Table IV.

TABLE IV
NAMING OF COMPONENTS AND ASSOCIATED INPUT NODES

Component	Cost factor	Factor Input for NN
1	Contractor capacity	Capacity of contractor 1 = Grade 1, 2 = Grade 2, 3 = Grade 3, 4 = Grade 4, 5 = Grade 5, 6 = Grade 6, 7 = ungraded
2	Project Location	Distance in km from Lusaka or urban location 1 = very near (<100km), 2 = near (101-300km), 3 = average (301-500km), 4 = far (501-700km), 5 = very far (701-900km), 6 = extremely far >901km
3	Financial delays	Number of days payment is delayed 1= 0 days 2= up to 30 days, 3= 31-60 days, 4= 61-90 days, 5= 91-180days, 6= above 180 days
4	Project feasibility (pre-construction)	1= very good, 2= good, 3= satisfactory, 4= poor, 5= unacceptable
5	Overheads and profit	1 = 0%, 2 = 5%, 3 = 10%, 4 = 15%, 5 = 20%, 6 = 25%, 7 = 30%, 8 = above 35%
6	Materials availability	1 = available locally, 2 = available imported, 3 = shortages, 4= severe shortages
7	Corruption profile	CPI compiled by Transparency International 1=very clean (100-80), 2= clean (79-60), 3= moderate (59-40), 4= corrupt (39-20), 5= highly corrupt (19-0)
8	Political risk	Political interference 1=none, 2= low, 3= moderate, 4= high, 5= very high

B. Neural Network Analysis

NeuroShell2 was selected because of its classic neural network paradigms, its popularity amongst researcher and its user friendly graphical user interface (GUI) as shown in Fig. 1. NeuroShell2 has five network architectures that include different learning paradigms namely: Backpropagation (BP); Kohonen; Probabilistic Neural Network (PNN); General Regression Neural Network (GRNN); and Group Method of Data Handling or Polynomial Nets (GMDH Network). All the networks are supervised type of network, trained with both inputs and outputs except the Kohonen network.

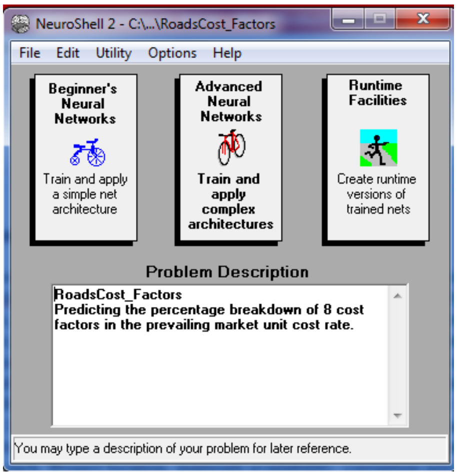

Fig. 1 Screen shot of NeuroShell 2 GUI

The Kohonen architecture was selected because it is unsupervised and has the ability to learn without being shown correct outputs in sample patterns. The 8 factors were analysed. The training epochs were increased steadily from 1000, 5000, 10,000 and 50,000. The results after 10,000 epochs remained the same. The Kohonen Self Organizing Map network was able to separate data patterns as shown in

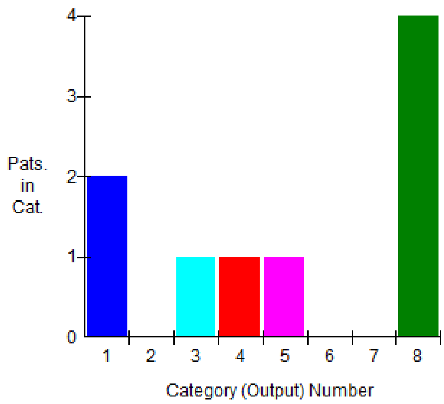

Fig. 2 Graphical representation of NN output
1 = Contractor capacity, 2 = Project Location, 3 = Financial delays, 4 = Project feasibility, 5 = Overheads and profit, 6= Materials availability, 7 = Corruption profile, 8 = Political risk

REFERENCES

[1] G. Raballand and A. Whitworth "The Crisis in the Zambian Road Sector", Working Paper No. 5, Zambia Institute for Policy Analysis and Research (ZIPAR), 2012, Lusaka.
[2] C. Hendrickson, Project Management for Construction, Fundamental Concepts for Owners, Engineers, Architects and Builders, 2008, (available online: http://pmbook.ce.cmu.edu/ (accessed 05/11/20139))
[3] Zambia Public Procurement Authority, Consulting services to undertake an assessment of the prevailing market rates in the construction industry, Request for proposals, ZPPA, Zambia, 2014.
[4] J. De la Garza and K. Rouhana, Neural Network versus parameter-based application. Cost Engineering, 37(2) 1995, pp.14-18.
[5] A. K. Mason and A. E. Smith, Cost Estimation Predictive Modeling: Regression versus Neural Network. Engineering Economist 42, (2) 1997, pp. 137-161.

[6] AACE International, Cost Engineering Terminology, Recommended Practice No. 10S-90, TCM Framework: General Reference, Association for the Advancement of Cost Engineering (AACE),2013
[7] Langdon, D., Spon's Civil Engineering and Highway Works Price Book 2009, Spon Press, 2009
[8] J. Mashilipa, An investigation on the effects and their implications of using British outputs in Estimating in the Zambian Construction Industry, BSc (Building) Unpublished thesis, Copperbelt University, Zambia, 2004.
[9] *Ibid*
[10] E. J Blocher, E. David, G. C. Stout and K. H. Chen, Cost management; A strategic emphasis 4th ed., McGraw Hill, 2008.
[11] A. Enshassi, S. Mohamed, and I. Madi, Contractors' Perspectives towards Factors Affecting Cost Estimation in Palestine. Jordan Journal of Civil Engineering, 1 (2) 2007, pp. 186-193.
[12] A. Trombka and S. Downey, A Study of County Road Project Cost and Schedule Estimates, Office of Legislative Oversight, Montgomery County, Maryland Report Number 2008-04, 2008, (available online https://www.montgomerycountymd.gov/olo/resources/files/2008-4.pdf (accessed 05/07/2014))
[13] A.H. Memon, I.A. Rahman, M.R. Abdullah, and A.A.A. Azis, Factors Affecting Construction Cost Performance in Project Management Projects: Case of Mara Large Projects. Proceedings of Post Graduate seminar on Engineering, Technology and Social Science. 29-30 November 2010, Universiti Tun Hussein Onn Malaysia, 2010
[14] A.U. Elinwa and S.A. Buba, Construction cost factors in Nigeria, Journal of Construction Engineering and Management. 119 (4), 1993, pp. 698-713.
[15] J. Bode, Neural Networks for Cost Estimation, Cost Engineering Journal 40 (1) 1998, pp.25-30
[16] I. Peško, M. Trivunić, G. Cirović and V. Mučenski, A preliminary estimate of time and cost in urban road construction using neural networks Technical Gazette 20 (3), 2013, pp. 563-570 (accessed 16/01/2014)
[17] A. J. Hasan, Parametric Cost Estimation of Road Projects Using Artificial Neural Networks, M.Sc. Unpublished thesis, The Islamic University, Gaza, 2013
[18] S. Muqeem, A. Idrus, F. M. Khamidi, J.B. Ahmad, and S.B. Zakaria, Construction labor production rates modeling using artificial neural network, Journal of Information Technology in Construction (ITcon), Vol. 16, 2011, pp. 713-726, (available online http://www.itcon.org/2011/42 (accessed on 09/04/2014))
[19] A. Hashem, P.A. Alex, and M. Tantash, Preliminary Cost Estimation of Highway Construction Using Neural Networks, Cost Engineering 41, (3) 1999, pp. 19-24
[20] D. R. Cooper and P.S. Schindler, Business research methods 11th ed., McGraw Hill, 2011, pp.390-397
[21] Creative Research Systems, Sample Size Formula, (available online: http://www.surveysystem.com/sample-size-formula.htm (accessed on 09/04/2014))
[22] J. C. F. de Winter, D. Dodou, and P. A. Wieringa, Exploratory Factor Analysis with Small Sample Sizes, Multivariate Behavioral Research, 44, 2009, pp147–181