Chandraratne, K.A.D.S.M., Devapriya, K.A.K. and Pandithawatta, T.P.W.S.I., 2019. Development of a computer model for cost estimation in educational buildings. In: Sandanayake, Y.G., Gunatilake, S. and Waidyasekara, A. (eds). *Proceedings of the 8th World Construction Symposium*, Colombo, Sri Lanka, 8-10 November 2019, pp. 266-275. DOI: doi.org/10.31705/WCS.2019.27. Available at: https://2019.ciobwcs.com/papers

DEVELOPMENT OF A COMPUTER MODEL FOR COST ESTIMATION IN EDUCATIONAL BUILDINGS

K.A.D.S.M. Chandraratne¹, K.A.K. Devapriya² and T.P.W.S.I. Pandithawatta³

ABSTRACT

Cost estimating in construction is critical at the early stage of the project in order to determine whether the project is feasible or not to the client. Cost estimators are facing numerous difficult moments due to incomplete project details and unavailability of cost estimation instruments in early stages of a project. This issue uplifts the necessity of novel and advanced cost model which would be simple, understandable and more reliable. Within the Sri Lankan context, the accuracy of the estimated amount is solely based on the experience and skills of the estimator due to unavailability of reliable cost estimation tool for educational institution buildings in Sri Lanka. Therefore, this paper aims to discuss about the development of a computer model for cost estimation in educational buildings with the current practices and related issues in preparing preliminary project estimates. Semi structured interviews were conducted between twelve experts from consultancy and contractor organisations and the necessary cost data were selected from twenty educational buildings. The Multiple Regression Analysis and Artificial Neural Network methods were utilised to analyse the collected data. Each method has a unique way of building relationships between predictors and responses. However, both the methods were succeeded only in estimating cost of limited number of sub-elements. Multiple Regression Analysis succeeded on five occasions and Artificial Neural Network method had presented efficacy in seven sub-elements only. Altogether eight elements were succeeded in estimating the cost of an educational institution buildings.

Keywords: Artificial Neural Networks; Cost Estimation Models; Educational Buildings; Elemental Cost Estimation; Multiple Regression Analysis.

1. INTRODUCTION

Construction projects need to be performed and delivered under specific constraints. These constraints can be listed as scope, time and cost. Out of that, the cost is one, which decides the initiation of the project. Moreover, in the competitive world, to maintain a profit margin and a market share, cost planning plays a significant role in the early stage of building design (Gunaydin and Dogan, 2004). Before the commencement of the construction, the investor needs to have an idea about the cost as early as possible (Ferry et al., 1999).

¹ Department of Building Economics, University of Moratuwa, Sri Lanka, sandunmc@ymail.com

² Department of Building Economics, University of Moratuwa, Sri Lanka, kakdevapriya@uom.lk

³ Department of Building Economics, University of Moratuwa, Sri Lanka, sonalitpw@gmail.com

According to Venkatachalam (1993), the field of Cost planning has been changed like all other areas, from its original path after the arrival of Information Technology (IT). Computer-based estimating models has many pros like effective data collection, accurate estimation, ability to deal with large volumes of data and to present results in an organized, neat, and consistent manner (Yaman and TAŞ, 2007). With its accuracy and consistency, construction cost estimation models are widely used in all over the world. However, those existing computer-aided and manual models are not matching for the Sri Lankan education buildings with its inherent complexities. Moreover, there's also a lack of computer-based estimation models in Sri Lankan Construction Industry. Therefore, it is a timely need to develop a cost estimation model for education buildings in Sri Lanka.

2. LITERATURE SYNTHESIS

2.1 CONSTRUCTION COST ESTIMATE

Since the industrial development directly affects the construction sector, the projects are getting more complicated to complete within quality standards, budgeted cost limits and expected time frame (Yaman and TAŞ, 2007). Therefore, the modern architectural and engineering firms submit cost estimates and participate in cost control activities for each phase to overcome this complexity (Division of Capital Assets Management, 2006). Latief *et al.*, (2013) showed that cost estimates evolve with the whole project life cycle and can be prepared using a variety of methods appropriate to the level of information and the time available to prepare the estimate. This estimation should provide the path for maximisation of the value of money.

The process of forecasting the cost of building or a physical structure is considered as the construction cost estimating process and it is similar in most cases. Moreover, both builders and clients worry about the financial impact of cost overruns and failing to complete a project. This process is critical because the structured estimating process should produce quality estimates, accuracy, confidence, reduced uncertainty, minimum costs and time savings (Adesola, 2012). Washington State Department of Transportation (2017) developed a flow chart to demonstrate cost estimation process, and it says the process should include following steps.

- 1. Determine Estimate Basis
- 2. Prepare Base Estimate
- 3. Review Estimate
- 4. Determine Risks and Set Total Project Cost
- 5. Determine Estimate Communication Approach
- 6. Obtain Management Endorsement (Approval)

2.2 BUILDING COST ESTIMATION MODELS

The cost estimation models came into act to improve the accuracy of estimation of construction costs in a construction project with minimum project information at the initial stage (Kim, An, and Kang, 2004). Cost modelling is a symbolic form of a system, and the content is defined with the factors affecting the cost (Holm, 2005). Further, Holm (2005) described that it is hard to understand the cause-effect relationship between the design parameters and the building cost, and to develop models to estimate the building cost. Models can be developed to provide various resource requirements within the cost and time (Lastra and Borja, 2014).

2.3 CONSTRUCTION COST DATABASE

For an estimation model development, cost database is essential. A cost database consisting of cost books (may be internal records at a particular company/agency /individual or public collection of data), the electronic equivalent of a cost book, or cost reference book, which used by cost estimators for more extended periods (Nardon, 1995). In a cost database estimators store data in a structured manner, which can be easily managed and retrieve in future (Samphaongoen, 2010).

2.4 SOFTWARE COST ESTIMATION

With the improvements in estimation techniques, formulas became more complex and spreadsheet errors became more frequent (Caulkins, Morrison, and Weidemann, 2008). As a result, cost estimating software applications were made with other benefits like reference databases for costs and other data, professional reports, speed, accuracy, and overall process standardisation. (Software Advice, 2014). After several years, more advanced features, such as saving data for reuse and trade-specific calculations, ability to work with multiple cost books, track project status, automatically compare estimates, clone, integrated sophisticated visual estimating and quantity take-off (QTO) tools have become available (Construction Software Review, 2017). Further, for the software development, programming languages give instructions to create the output from its input.

3. RESEARCH METHODOLOGY

At the outset a comprehensive literature survey was carried out with books, journals and articles to obtain the existing knowledge on various types of cost estimation models that used in the construction industry. Then semi-structured interviews have been carried out with twelve key individuals from consultancy and contractor organisations in Sri Lanka, in order to obtain expert opinions of the necessity of a cost estimation model, current practices in Sri Lanka. Further, the necessary information was gathered to determine the independent variables, dependent variables, relationships between parameters and methods to develop a sophisticated and accurate cost estimation model. After that, with the purposive sampling technique, twenty educational buildings, which constructed in last twenty years were selected. In order to collect data from these selected projects several documents (drawings and specifications, bills of quantities, preliminary project estimates, final bills and price catalogues) were reviewed. Then the MRA (Multiple Regression Analysis) and ANN (Artificial Neural Network) methods were utilised to analyse the collected data by using Minitab 18 and Neuroph Studio respectively. Finally, the cost estimation model was developed.

4. RESEARCH FINDINGS

4.1 CURRENT PRACTICES AND RELATED ISSUES IN PREPARING PRELIMINARY PROJECT ESTIMATE

All of the experts were practising similar type of methods in preparing Preliminary Project Estimates (PPE), which lead to similar kind of weaknesses, faults and laborious. Most of the respondents indicated that the method used to prepare the initial estimate mainly depends on the availability of the project data. Therefore, within the Sri Lankan context, there is no standard practice implemented in building cost estimation modelling.

However, during the process of predicting cost in the initial stage of the project, estimators do not have the actual details and specification for the structural elements, finishes etc. So, most of the time, estimators assume those details which can be different from the actual details. Most of the respondents identified this as a major issue which can affect to the precision of the cost estimate. Even though data is available, issues can arise when preparing PPE. Most of the time, rates derived from cost analysis are not adjusted according to the time and place. The main reasons for this are lack of knowledge of estimators and limited time frame. As per the respondents, most of the Architects make their basic plans based on their experience without even utilising analysed cost data. Further, most of the Quantity Surveyors use traditional and manual methods and do not use any framework or cost model to prepare PPEs. These traditional and manual methods require much time and are their precision is doubtful.

4.2 DETERMINATION OF INDEPENDENT AND DEPENDENT VARIABLES

To develop an accurate model, selection of the variables is a critical aspect. Each of the following variables ranked based on RII values of experts' opinions calculated based on 1 to 5 weighted scale. These based ranks of each variables are illustrated in Table 1.

Variable	RII (Relative index of inequality)	Based Rank
Gross Floor area	0.983	1
No of Floors	0.967	2
Structural Material	0.967	2
Building Height	0.917	4
Foundation Type	0.900	5
Type of Building	0.817	6
Project Location	0.817	6
Roof Type	0.817	6
Duration of the project	0.767	9
Construction Year	0.767	9
External Girth	0.650	11
Doors and Windows Area	0.617	12
Average Storey Height	0.617	13
Upper Floor Area	0.600	14
Wall Area	0.600	14
Ground Floor Area	0.517	16

Table 1: RII ranking of predictors

First 10 variables were selected as the predictors for the development of cost estimation model to avoid over fitting which can be caused by adding too many independent variables. Over fitting can cause more variance by decreasing the reliability of the model. Predictors can be categorised in to two types, continuous and categorical. Categorical variables contain a finite number of categories or distinct groups while Continuous variables are numeric variables that have an infinite number of values between any two values. Table 2 shows the category of the above selected variables.

Table 2: Types of variables

Continuous Predictors	Categorical Predictors
Duration	Type of Building
Gross Floor Area	Project Location
No of Floors	Construction Year
Building Height	Foundation Type
	Roof Type
	Structural Material

According to the experts, following items were identified as cost significant items in educational building.

• Pilling • Electrical Installations • Excavation and Earthwork • Floor/ Wall/ Ceiling Finishes • Concrete Works • Painting • Masonry Works • Drainage System • Water Proofing • Partitions / Linings / Suspended • Structural Metal Work Ceilings Metal Work Woodwork • Plumbing / Sanitary Installations • Roof Covering/ Roof Plumbing

Data of above-mentioned variables was collected through data retrieval for the development of the database. Table 3 shows the related input variables for each cost significant item.

Table 3: Parameters which are expected to be related

Sub Element	Input Variables					
	Continuous Predictors	Categorical Predictors				
Pilling	Gross Floor Area/ No of Floors/ Building Height	Type of Building/ Foundation Type				
Excavation and Earthwork	Gross Floor Area	Type of Building/Structure				
Concrete Works	Gross Floor Area/ No of Floors/ Building Height	Type of Building/ Province/ Construction Year/ Duration/ Foundation Type/ Roof Type/ Structure				
Masonry Works	Gross Floor Area/ No of Floors/ Building Height	-				
Water Proofing	Gross Floor Area	Foundation Type/ Roof Type				
Structural Metal Work	Gross Floor Area/ No of Floors/ Building Height	Type of Building/ Roof Type/ Structure				
Metal Work	Gross Floor Area/ No of Floors	-				
Roof Covering/ Roof Plumbing	Gross Floor Area/ Building Height	Type of Building/ Province/ Construction Year/ Duration/ Foundation Type/ Roof Type/ Structure				

Sub Element	Input Variables					
	Continuous Predictors	Categorical Predictors				
Electrical Installations	Gross Floor Area	Type of Building				
Floor/ Wall/ Ceiling Finishes	Gross Floor Area/ Building Height	Type of Building/ Roof Type/ Structure				
Painting	No of Floors/ Building Height	-				
Drainage System	Gross Floor Area/ No of Floors	Type of Building/ Structure				
Partitions/ Linings/ Suspended Ceilings	Gross Floor Area	Type of Building				
Woodwork	Gross Floor Area	Province				
Plumbing/ Sanitary Installations	Gross Floor Area/ No of Floors/ Building Height	Province/ Type of Building				

4.3 DATABASE DEVELOPMENT

Altogether, 20 educational buildings were considered for the development of cost database to perform multiple regression analysis and to train the neural network. This database consists of cost data of above selected cost significant items (sub building elements) and the characteristics of continuous and categorical variables. MS Excel, which is the most popular software in Sri Lanka to develop cost databases, was used to develop the cost database. MRA and ANN methods were selected because of its inherent strengths over weaknesses, such as accuracy of results, assessment tools, use of multivariable, input for new management trends.

4.4 MULTIPLE REGRESSION ANALYSIS

Table 4 illustrates the relationship between dependent and independent variables according to the multiple regression analysis. Separate equations for each successful categorical predictor are shown under every sub building element. Therefore, the most suitable equation should be selected for the predictions by considering characteristics of the building.

Fifteen cost significant items were analysed using MRA method to develop the estimation model. However, only five of them gave desired outcome with the equation. As a rule of thumb, typically R² values greater than 0.5 were considered as acceptable. Due to the overfitting issue, four of the elements were removed from the model development, and another three were eliminated due to insufficient database entries. Further, another three elements were discarded due to the low R-sq (pred).

Table 4: Summary of regression equations

Sub Building Element	R-sq(adj)	R-sq(pred)	Equation (Rs.)
Masonry Works	68.43%	60.83%	MASONARY WORKS = -29624179 + 1030 Gross Floor Area(m²) - 38826226 No of Floors + 10592149 Building Height (m)+ 5175993 Duration of the Project (months)

Sub Building Element	R-sq(adj)	R-sq(pred)	Equation (Rs.)
Water Proofing	85.35%	72.18%	WATER PROOFING = 4509627 + 1517 Gross Floor Area (m ²) [Pile/Roof Slab]
			WATER PROOFING = 1268479 + 1517 Gross Floor Area (m²)[Pile/Steel structure – ZnAl]
			WATER PROOFING = 1405491 + 1517 Gross Floor Area (m ²)[RCC Foundation/Roof Slab]
			WATER PROOFING = -1835657 + 1517 Gross Floor Area (m ²)[RCC Foundation/Steel structure - Zn Al]
Metal Works	79.66%	73.25%	METAL WORK = -2965626 + 10416 Gross Floor Area (m ²) - 2645482 No of Floors
Floor/ Wall/ Ceiling Finishes	76.59%	67.61%	FLOOR/ WALL/ CEILING FINISHES =19219407 + 13375 Gross Floor Area (m²) + 639111 Building Height (m) [Public/Roof Slab/Concrete]
			FLOOR/ WALL/ CEILING FINISHES =19891025 + 13375 Gross Floor Area (m²) + 639111 Building Height (m) [Public/Roof Slab/Steel]
			FLOOR/ WALL/ CEILING FINISHES = -24472810 + 13375 Gross Floor Area (m²) + 639111 Building Height (m) [Public/Steel structure - Zn Al/Concrete]
			FLOOR/ WALL/ CEILING FINISHES = -23801192 + 13375 Gross Floor Area (m²) + 639111 Building Height (m) [Public/Steel structure - Zn Al/Steel]
			FLOOR/ WALL/ CEILING FINISHES = 25387057 + 13375 Gross Floor Area (m²) + 639111 Building Height (m) [Technical College/Roof Slab/Concrete]
			FLOOR/ WALL/ CEILING FINISHES =26058675 + 13375 Gross Floor Area (m²) + 639111 Building Height (m) [Technical College/Roof Slab/Steel]
			FLOOR/ WALL/ CEILING FINISHES =-18305160 + 13375 Gross Floor Area (m ²)

Sub Building Element	R-sq(adj)	R-sq(pred)	Equation (Rs.)
			+ 639111 Building Height (m) [Technical
			College/Steel structure - Zn Al/Concrete]
			FLOOR/ WALL/ CEILING FINISHES =-17633542 + 13375 Gross Floor Area (m ²) + 639111 Building Height (m) [Technical College/Steel structure - Zn Al/Steel]
Painting	80.63%	74.09%	PAINTING = -5781796 + 2239 Gross Floor Area (m ²) - 8283350 No of Floors + 2909923 Building Height (m)

4.5 ARTIFICIAL NEURAL NETWORK

Artificial Neural Networks were created for each building sub-element of an educational institute building which was opted at literature synthesis and interviews. The Neural Networks were trained and tested by using the selected 20 numbers of educational institutional buildings. When analysing the data using ANN method, continuous predictors were ignored.

Element	Inputs	Outputs	Hidden Neurons	Learning rate	Momentum	Iteration	MSE
Concrete Works	4	1	2	0.1	0.6	1546	0.010186
Masonry Works	4	1	3	0.2	0.6	2380	0.009621
Metal Works	2	1	2	0.3	0.6	957	0.012009
Roof Covering/ Roof Plumbing	3	1	2	0.2	0.7	4995	0.010917
Floor/ Wall/ Ceiling Finishes	2	1	1	0.2	0.6	2030	0.014242
Painting	3	1	2	0.1	0.6	1267	0.012343
Plumbing/ Sanitary Installations	3	1	2	0.2	0.6	1296	0.010864

Table 5: Summary of ANN

Table 5 shows the summary of ANN and its learnings. During the analysis seven elements were identified with less than 5000 iterations and less than 0.01 Medium Standard Error (MSE). And for the other eight cases all the training attempts by Neuroph Studio failed to trained the neural network in 5000 iterations. When the number of hidden neurons is getting closer to the number of inputs, MSE figure become less. Finally, the equations and relationships were validated to determine the performance of the two techniques.

4.6 VALIDATION OF THE MRA AND ANN

To validate the results, trial and error method was used by estimating a sample project, which is not included in the cost database. In this case both the MRA and ANN models are having error percentage of 3.50 % and 2.78% respectively. However, the error range of MRA models (+-26.91 %) is pretty high compared to the ANN architectures (+-18.72%). In the MRA model, most of the positive errors and negative errors are settled

against each other. That is the reason for the low overall error. However, both the estimations have estimated with an overall estimation error less than 10%, which was established by the interviewees. Hence, both the methods have pros and cons when comes to the building cost estimations modelling. Table 6 shows validation of MRA and ANN.

Table 6: Validation of MRA and ANN

Elamant	Actual	Predicted	d Amount	Diffe	Error Percentage		
Element Amount MRA ANN		MRA ANN		MRA (%)	ANN (%)		
Concrete Works	57,442,450.00		55,248,365.00		(2,194,085.00)		-3.82
Masonry Works	15,317,000.00	13,585,937.10	16,458,124.25	(1,731,062.90)	1,141,124.25	-11.30	7.45
Water Proofing	6,873,200.00	7,148,371.00		275,171.00		4.00	
Metal Works	18,620,300.00	21,533,898.00	18,245,667.00	2,913,598.00	(374,633.00)	15.65	-2.01
Roof Covering/ Roof Plumbing	1,688,300.00		1,458,214.00		(230,086.00)		-13.63
Floor/ Wall/ Ceiling Finishes	35,954,085.00	38,169,665.90	42,684,246.00	2,215,580.90	6,730,161.00	6.16	18.72
Painting	3,248,145.20	2,374,166.70	2,854,621.00	(873,978.50)	(393,524.20)	-26.91	-12.12
Plumbing/ Sanitary Installations	8,038,615.00		7,254,698.00		(783,917.00)		-9.75
Total	147,182,095.20	82,812,038.70	144,203,935.25	2,799,308.50	3,895,040.05	3.50	2.78

5. CONCLUSIONS

The educational buildings are designed for different primary, secondary or higher education systems and often includes living areas for students. Majority of the universities, schools, technical colleges are constructed and controlled by the Sri Lankan government. Therefore, the accuracy and the transparency of financial transactions should be crystal clear. These projects are funded by tax money of public and foreign investments which need to be properly managed. With the limited amount of details available, it is arduous to make accurate predictions in early stages. Different parameters should identify from the initial drawings and available details. In Sri Lanka, the accuracy of the estimated amount is based on the experience and skills of the QS due to unavailability of reliable cost estimation tool for educational institution buildings in Sri Lanka. This is a must to improve the precision of the estimated amount. With the improvement of ICT and artificial intelligence lot of multi and nonlinear relationship estimation techniques comes to the stage. These methods and techniques will be the future of construction cost estimation. The computer model was developed using cost data of educational building in western province and southern province. Therefore, the model

cannot be used to estimate the construction cost of a buildings located in the different areas. Furthermore, this model can estimate only the cost of conventional buildings, i.e. universities and technical colleges. Both the MRA and ANN methods were succeeded only in estimating of cost of limited number of sub-elements. MRA succeeded on five occasions which delivered provided significant R-Squared predicting values. On the other hand, ANN method was presented efficacy in seven sub-elements in different combinations of hidden neurons, learning rates and momentums. Altogether eight elements were succeeded in estimating the cost of an educational institution building.

6. REFERENCES

- Adesola, A., 2012. The Estimating Process. Yangon: SPDC.
- Caulkins, J.P., Morrison, E.L. and Weidemann, T., 2009. Do Spreadsheet Errors Lead to Bad Decisions? Perspectives of Executives and Senior Managers. In *Evolutionary Concepts in End User Productivity and Performance: Applications for Organizational Progress* (pp. 44-62). IGI Global.
- Construction Software Review, 2017. Construction Estimating Software Buyer's Guide: Benefits. Available from: http://www.constructionsoftwarereview.com/learning_center/buyers_guides/constructionestimating-software-buyers-guide/construction-estimating-2 [Accessed 06 July 2017]
- Division of Capital Assets Management, Building for the Commenwealth, 2006. *Consultants Estimating Manual*. London: Commonwealth of Massachusetts.
- Ferry, D. J., Brandon, P. S., and Ferry, J. D., 1999. Cost Planning of Buildings. London: Blackwell Science Ltd.
- 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*, pp.595–602.
- Holm, L., 2005. Construction Cost Estimating: Process and Practices. Washington: Pearson/Prentice Hall.
- Kim, G.H., An, S.H., and Kang, K.I., 2004. Comparison of Construction Cost Estimating Models Based on Regression Analysis, Neural Networks, and Case-Based Reasoning. *Building and Environment*, 39, pp.1235 1242.
- Lastra, G. D., and Borja, 2014. A Methodology to Make Accurate Preliminary Estimates of Construction Material Quantities for Construction Projects.
- Latief, Y., Wibowo, A., and Isvara, W., 2013. Preliminary Cost Estimation Using Regression Analysis Incorporated With Adaptive Neuro Fuzzy Inference System. *International Journal of Technology*, 1, pp.63-72.
- Nardon, J. D., 1995. Bridge and Structure Estimating. New York City: McGraw-Hill Book Company.
- Samphaongoen, P., 2010. *A Visual Approach to Construction Cost Estimating*. Available from: http://epublications.marquette.edu/theses_open/28 [Accessed 02 July 2017]
- Software Advice, 2014. Construction Cost Estimating Software Comparison. Available from: http://www.softwareadvice.com/construction/cost-estimating-software-comparison/ [Accessed 01 May 2017]
- Venkatachalam, A.R., 1993, October. Software cost estimation using artificial neural networks. In *Proceedings of 1993 International Conference on Neural Networks (IJCNN-93-Nagoya, Japan)* (Vol. 1, pp. 987-990). IEEE.
- Washington State Department of Transportation, 2017. Cost Estimate Process. Available from: http://www.wsdot.wa.gov/Projects/ProjectMgmt/RiskAssessment/Process/ [Accessed 30 June 2017].
- Yaman, H., and Tas, E., 2007. A Building Cost Estimation Model Based on Functional Elements. *A*|*Z ITU Journal of the Faculty of Architecture*, 4, pp.73-87.