Preliminary Cost Estimating Models for Road Construction Activities

Ibrahim MAHAMID and Amund BRULAND, Norway

Keywords: cost estimating, cost engineering, road management, regression models, construction management.

SUMMARY

This study describes the development of linear regression models to predict the cost of road construction activities based on 100 set of data collected in the West Bank – Palestinian National Authority. The projects were implemented over the years 2005-2008. The used procedure resulted in the formulation of regression models which provides a means of determining the cost of road construction activities as a function of project's characteristics. The prediction models were developed for three major road construction activities which are earthworks, base coarse works, and asphalt works. Three groups of models for each activity were developed based on used dependent variable; they are total cost of construction activity, cost per meter length, and cost per meter square. The proposed independent variables are road length, pavement width, base coarse width, terrain condition, soil drill ability, and soil suitability

Preliminary Cost Estimating Models for Road Construction Activities

Ibrahim MAHAMID, Amund BRULAND, Norway

1. INTRODUCTION

The term preliminary estimate is used to describe the process of predicting a project's cost before design of the project is completed (Sanders et al, 1992). The technique is used to estimating one characteristic of a system, usually its cost, from other physical and/or performance characteristics of the system (Rose, 1982). Construction clients require early and accurate cost advice prior to site acquisition and commitment to build in order to enable them to take a right decision regarding the feasibility of proposed project. The accuracy of an estimate is highly dependent on the available information about the proposed project at the time of estimate. Preliminary cost estimates are required at the early phase in which minimal information is available about the project, which makes the estimate less accurate. However, preliminary estimates are important to construction clients to examine them before deciding to continue with a project.

The objective of this research is to develop simple preliminary cost estimation for contracted road construction activities that can be handled using calculator or simple computer program. To address that objective, mathematical models which describe the cost of each activity as a function of project's characteristics were developed.

In this study, the estimation models were developed based on collected data for implemented road construction projects in West Bank. As these cost estimates are required at early stage of the project, consideration was given to the fact that the input data for the required mathematical model could be easily extracted from sketches or scope definition of the projects.

2. BACKGROUND

Preliminary cost estimation is considered as the most significant starting process to influence the fate of a new road construction project (Jamshid, 2005). Estimating is the heart of the cost engineer's work and consequently it has received appropriate attention over the years (Jelen and Black, 1983).

Major difficulties which arise while conducting cost estimation during the conceptual phase are lack of preliminary information, lack of database of road works costs, and lack of up to date cost estimation methods. Additional difficulties arise due to larger uncertainties as result of engineering solutions, socio-economical, and environmental issues. Parametric cost estimation or estimation based on historic database during the conceptual estimate phase is widely used in developed countries. However, developing countries face difficulties related to the creation of a road work costs database, which may be used for cost estimation in either the conceptual stage or the feasibility study of a project cycle (Jamshid, 2005).

Some cost prediction models have been developed, Wilmot and Cheng, (2003) developed application of the model in forecasting to highway construction costs in Louisiana, it shows that the model closely replicates past construction costs for the period 1984–1997. When applied to forecasting future highway construction costs, the model predicts that highway construction costs in Louisiana will double between 1998 and 2015. Applying cost-cutting policies and assuming input costs are 20% less than anticipated, the model estimates highway construction costs will increase by 75% between 1998 and 2015. Jamshid (2005) discussed how to handle data missingness, analyzed the impact of a different set of variables on the project cost, and evaluated the proposed cost estimation model. It proposes a cost estimation technique for developing countries and criteria for using an artificial neural network.

Oberlender and Trost (2001) concentrated more on improving prediction accuracy by deriving four determinants influencing the accuracy of an early cost estimates. Following this, they established a multiple regression model which derived significant relationship between estimate accuracy and influencing factors such as basic process design, team experience and cost information, time allowed to prepare the estimate, site requirements, and bidding and labor climate.

Han et al (2008) investigated the actual budgeting process in the highway construction project, under the research collaboration of the Korean Ministry of Construction and Transportation. Then they developed the two-tiered cost estimation models of highway construction projects, considering the target goals for forecasting, allowable accuracy, and available information level at each phase of a project budgeting and initiation.

3. RESEARCH METHODOLOGY

The methodology of the study followed the steps shown in Figure 1. After establishing the objectives of the study, the needed data to achieve these objectives were collected from road construction projects implemented in West Bank over the years 2005 - 2008. The data analysis is carried out using an array of statistical methods that include descriptive statistics, correlation analysis, F-test, and statistical modeling.

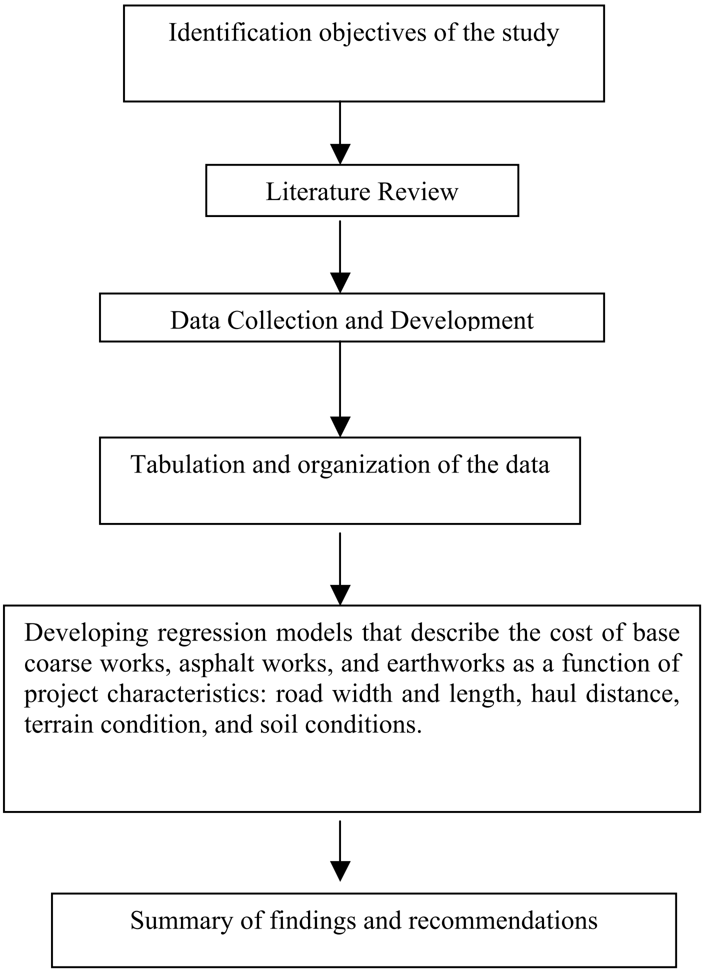

Figure 1 Research methodology

3.1 Data Collection

Estimating technique requires an extensive historical data base; the data were collected from contracts awarded in West Bank. The data collected comprised of 100 projects awarded over the years 2005-2008. The data were tabulated to ensure that all costs were considered, and that none is double-counted, and that all are clearly defined. All the data were deflated using the construction cost index for 2008 obtained from the Palestinian Central Bureau of Statistics.

3.2 Models Development

Once the variables to be included in the preliminary estimate equation have been identified, a series of mathematical models were developed using the multiple and linear regression analysis techniques. The analysis of data and the creation of models was done using excel. The explanation of developed models is as the following:

Cost model of asphalt works

The developed model of asphalt works is of the following form:

$$Y = \beta_o + \beta_1 X_1 + \beta_2 X_2 + \ldots\ldots.. + \beta_n X_n$$

Where;
Y is the dependent variable, three dependent variables were considered, the first one is the total cost of asphalt works, the second is the cost/m, and the last one is the cost/m^2.

$X_1 \ldots X_n$ are the dependent variables, the used dependent variables are road length (km), pavement width (m), pavement thickness after compaction (cm), haul distance (km) which is the distance from asphalt source to the center of project, and the pavement area (m^2).

β_i's are regression parameters to be estimated.

Cost model of base coarse works

The developed model of asphalt works is of the following form:

$$Y = \beta_o + \beta_1 X_1 + \beta_2 X_2 + \ldots\ldots + \beta_n X_n$$

Where;

Y is the dependent variable, three dependent variables were considered, the first one is the total cost of base coarse works, the second is the cost/m, and the last one is the cost/m^2.

$X_1 \ldots X_n$ are the dependent variables, the used dependent variables are road length (km), base coarse width (m), base coarse thickness after compaction (cm), haul distance (km) which is the distance from base coarse source to the center of project, and the base coarse area (m^2).

β_i's are regression parameters to be estimated.

Cost model of earthworks

The developed model of asphalt works is of the following form:

$$Y = \beta_o + \beta_1 X_1 + \ldots\ldots\ldots.. + \beta_n X_n + \gamma 1D1 + \gamma_2 D_2 + \gamma_3 D_3$$

Where;

Y is the dependent variable, two dependent variables were considered, the first one is the total cost of earthworks, and the second is the cost/m.

$X_1 \ldots X_n$ are the dependent variables, the used dependent variables are road length (m), road width (m), and the road area (m^2).

B_i's and γ_i's: regression parameter to be estimated.

D_i's: qualitative variables (dummy variables) which are as following:

D_1 is a regressor for terrain condition where;

Terrain condition	**D_1**
semi even	0
hilly	1

The semi even terrain is the terrain with good visibility and hilly is the terrain with bad visibility.

D_2 is a regressor for soil drillability where;

Soil drillability	**D_2**
good	0
poor	1

The drillability of sol is considered to be good when it could be excavated easily with high production rate.

D_3 is a regressor for soil suitability to be used for fill where;

Soil suitability	**D_3**
good	0
poor	1

The suitability of soil is good when it can be used as fill material in situ.

4. RESULTS AND DISCUSSION

Once a set of probable predictors was identified for each activity, mathematical models were developed using multiple linear regression technique. The results are discussed below.

4.1 Cost estimating models of asphalt works

Three groups of regression models that predict the cost of asphalt works were developed as shown in the following sections.

4.1.1 Total cost as dependent variable

In this section, the total cost of asphalt works in US dollars used as the dependent variable, four models were developed using different independent variables in each one, the results are shown in table 1.

Table 1 Cost estimating models using total cost as dependent variable

Model #	Independent Variables	Coefficients	r^2	p-value	*F*
1	road length(km)	61198.03	0.89	1.91E-49	809.22
2	road length(km)	55214.21	0.91	4.74E-39	467.27
	pavement width(m)	2775.93		0.000239	
3	Intercept	7382.43	0.94	0.010812	1592.44
	pavement area	11.10		2.07E-62	
4	pavement area	11.62	0.97	3.91E-79	3511.64

4.1.2 Cost per meter length as dependent variable

In this section, the cost of asphalt works in US dollars per meter length used as the dependent variable, two models were developed using different independent variables in each one, the results are shown in table 2.

Table 2 Cost estimating models using cost/m as dependent variable

Model #	Independent Variables	Coefficients	r^2	p-value	*F*
1	pavement width(m)	13.11	0.96	3.32E-70	2284.83
2	Intercept	-45.62	0.80	0.001548	127.53
	pavement width(m)	12.78		2.22E-20	
	pavement thickness (cm)	5.47		0.044441	
	pavement haul (km)	1.10		0.006997	

4.1.3 Cost per meter square as dependent variable

In this section, the cost of asphalt works in US dollars per meter square used as the dependent variable, three models were developed using different independent variables in each one, the results are shown in table 2.

Table 3 Cost estimating models using cost/m^2 as dependent variable

Model #	Independent Variables	Coefficients	r^2	p-value	*F*
1	Intercept	8.93	0.06	2.34E-07	5.89
	pavement thickness (cm)	0.67		0.017057	
2	pavement thickness (cm)	2.17	0.95	1.99E-65	1810.05
3	Intercept	6.19	0.11	0.001995	5.89
	pavement thickness (cm)	0.80		0.004217	
	pavement haul (km)	0.14		0.019779	
4	pavement thickness (cm)	1.59	0.96	8.38E-23	1141.94
	pavement haul (km)	0.25		2.03E-06	

4.2 Cost estimating models of base coarse works

The following sections show three groups of regression models that developed to predict the cost of base coarse works.

4.2.1 Total cost as dependent variable

In this section, the total cost of base coarse works in US dollars used as the dependent variable, four models were developed using different independent variables in each one, the results are shown in table 4.

Table 4 Cost estimating models using total cost as dependent variable

Model #	Independent Variables	Coefficients	r^2	P-value	*F*
1	road length(km)	31086.94	0.75	1.2E-31	298.45
2	Intercept	-44448.26	0.57	0.01759	63.09
	road length(km)	33071.67		7.2E-19	
	base coarse width(m)	5346.99		0.00832	
3	area (m^2)	4.17	0.78	5.4E-34	344.24
4	base coarse volume (m^3)	19.81	0.85	7.6E-43	569.48

4.2.2 Cost per meter length as dependent variable

In this section, the cost of base coarse works in US dollars per meter length used as the dependent variable, one model was developed using base coarse width as independent variable, the results are shown in table 5.

Table 5 Cost estimating models using cost/m as dependent variable

Model #	Independent Variables	Coefficients	r^2	P-value	*F*
1	base coarse width(m)	5.26	0.84	3.7E-41	519.01

4.2.3 Cost per meter square as dependent variable

In this section, the cost of base coarse works in US dollars per meter square used as the dependent variable, one model was developed using base coarse thickness as independent variable, the results are shown in table 6.

Table 6 Cost estimating models using cost/m^2 as dependent variable

Model #	Independent Variables	Coefficients	r^2	P-value	*F*
1	base coarse thickness (cm)	0.23	0.87	3.9E-45	644.38

4.3 Cost estimating models of earthworks

Only two groups of regression models to predict the cost of earthworks were developed. Group one that use total cost as dependent variables and group two that use cost per meter length as dependent variables. In the third group, the group use the cost per meter square as dependent variable, the regression analysis resulted with r^2 less than 0.30 which give indication of weak linear relationship between dependent and independent variables, also the p-value always more than 0.05 which is another indication of not fitted and significant models.

4.3.1 Total cost as dependent variable

Table 7 shows eight developed models that use the total cost of earthworks in US dollars as dependent variables. Different collections of independent variables were used in each model.

Table 7 Cost estimating models using total cost as dependent variable

Model #	Independent Variables	Coefficients	r^2	p-value	*F*
1	Intercept	-13574.95	0.72	1.25E-05	245.97
	road length (m)	15.35		1.81E-28	
2	road length (m)	15.48	0.79	1.39E-23	189.32
	road width (m)	-1784.32		0.000317	
3	area (m^2)	1.34	0.81	6.62E-38	432.43
4	Intercept	-12118.28	0.81	3.94E-06	99.18
	area (m^2)	1.27		1.29E-27	
	terrain condition	11429.90		0.001758	
	soil drill ability	6666.81		0.075784	
	soil suitability	7256.52		0.032191	
5	Intercept	-9729.30	0.79	0.000128	179.23
	area (m^2)	1.36		1.35E-30	
	terrain condition	15941.58		3.03E-06	
6	area (m^2)	1.20	0.83	4.41E-29	242.62
	terrain condition	9968.51		0.001515	
7	area (m^2)	1.20	0.83	1.86E-27	236.64
	soil drill ability	9553.12		0.00458	
8	area (m^2)	1.23	0.82	1.86E-27	226.00
	soil suitability	7136.61		0.00458	

4.3.2 Cost per meter length as dependent variable

Table 8 shows five developed models that use the t cost of earthworks in US dollars per meter length as dependent variables. Different collections of independent variables were used in each model.

Table 8 Cost estimating models using cost/m^2 as dependent variables

Model #	Independent Variables	Coefficients	r^2	p-value	*F*
1	road width (m)	0.452431709	0.86	2.9E-07	151.5272
	terrain condition	4.175524566		5.91E-06	
	soil drill ability	3.174638837		0.000648	
	soil suitability	3.452925565		6.05E-05	
2	road width (m)	1.119914153	0.71	4.6E-28	237.6938
3	road width (m)	0.723727828	0.81	1.51E-14	207.8216
	terrain condition	6.405405883		8.33E-11	
4	road width (m)	0.74769453	0.80	6.17E-15	196.883
	soil drill ability	6.22191769		7.32E-10	
5	road width (m)	0.786114472	0.77	1.18E-13	163.4631
	soil suitability	5.255770461		1.13E-06	

5. CONCLUSION

It is possible to develop a linear regression models which can be used for preliminary cost estimating. Such type of models is very useful, especially in its simplicity and ability to be handled by simple computer program or calculator. It has a good benefit in estimating project cost at early stages of the project.

It must be remembered that an estimated project cost is not an exact number. Rather it is an opinion of probable cost, not an exact calculation. The accuracy and reliability of an estimate is totally dependent upon how will the project scope is defined and the time and effort expended in preparation the estimate.
The type of estimate to be made and its accuracy depends upon many factors including the purpose of the estimates, how much is known about the project, and how much time and effort is spent in preparing the estimate.

The coefficients of determination r^2 for the developed models ranging from 0.57 to 0.96 which means that the relationship between the independent and dependent variables of the developed models is strong and the predicted values from a forecast models fit with the real-life data, also the f-test for all models show that the relation between dependent and independent variables was not by chance and p-value for each independent variables is less than 0.05 which means that the use of dependent variable in the model is significant.

REFERENCES

Han, S. H., Kim, D. Y., & Kim, H., 2008. Two-staged Early Cost Estimation for Highway Construction Projects. Report (2008), Civil & Environmental Engineering, Yonsei University, Seoul, Korea.

Jelen, F.C. & Black, J.H., 1983. Cost and Optimization engineering, 2nd ed. , McGraw-Hill, New York.

Oberlender, G. D. & Trost, M., 2001. Predicting Accuracy Early Cost Estimates Based on Estimate Quality. Journal of Construction Engineering and Management, Vol. 127, No. 3, pp. 173–182.

Palestinian Centre Bureau of Statistics (PCBS). Personal communication, September, 2009. West Bank, Palestine.

Palestinian Economic Council for Development and Reconstruction (PECDAR). Personal communication, November, 2009. West Bank, Palestine.

Rose, Aaron, 1982. An Organized Approach to Parametric Estimating. Transactions of the seventh International Cost Engineering Congress, London, England, October.

Sanders, S.R., Maxwell, R.R., & Glagola, C.R., 1992. Preliminary estimating models for infrastructure projects. Cost Engineering 34 (8).

Sodikov, Jamshid, 2005. Cost Estimation of Highway Projects in Developing Country: Artificial Neural Network Approach. Graduate Project, Department of Civil and Environmental Eng., Saitama University.

Wilmot, C. G. & Cheng, G., 2003. Estimating Future Highway Construction Costs. Journal Construction Engineering and Management. Volume 129, Issue 3, pp. 272-279.

CONTACTS

Ibrahim MAHAMID[1]
Norwegian University of Science and Technology (NTNU),
7034 Trondheim,
Norway.
(+47) 73 59 46 40
(+47) 73 59 70 21
imahamid@yahoo.com