Journal of Engineering and Applied Sciences 13 (7): 1676-1682, 2018
ISSN: 1816-949X

Utilization Multifactor Linear Regression Technique for Prediction the Earned Value in Bridges Projects

[1]Faiq M.S. AL-Zwainy, [2]Jumaa Awad AL-Somaydaii and [3]Hadi S.M. Aljumaily
[1]College of Engineering, Al-Nahrain University, Baghdad, Iraq
[2]College of Engineering, Al-Anbar University, Ramadi Iraq
[3]Faculty of Engineering, University of Mustansiriya, Baghdad, Iraq

Abstract: Earned value analysis is an extremely significant part in the operations of bridges projects management principally with respect to EAC, this study aims at developing EAC Model for bridges projects using Multifactor Linear Regression Technique (MLRT). One MLR Model was built up based upon 43 set of database gathering between 2013 and 2016 from bridges and roads directorate in Iraq construction sector. Only five affect factors were applied with regard to prediction by MLRT as independent factors F1:AC, Actual Cost, F2:EV, Earned Value, F3:PV, Planning Value, F4:BAC, Budget at Completion and F5:D, Duration of project. Only single model was constructed for forecasting with EAC for bridges projects. It was found that MLR have the ability to predict the EAC with excellent rank of accuracy for the correlation coefficient (R) was 93.91% with Average Accuracy (AA%) equal to 93.86%.

Key words: MLRT, earned value, bridges projects, EAC, R, R^2, AA%

INTRODUCTION

Earned Value Management (EVM) is an approach of management for evaluation of the construction project performance. EVM has the ability to merge measurements of the project management triangle, time, cost and quality. EVM applications for complex and large construction projects have some indicators and prediction of schedule performance (ahead of schedule or behind schedule) and cost performance (under budget or over budget). Earned Value analysis philosophy is not a specific philosophy or device set but rather a group of guidelines that lead a construction project's management control system (Sarhan *et al*., 2015).

Main purpose of this research was to manage a careful mensuration for earned value in Iraq construction project through developing multifactor linear regression model for forecasting earned value for bridges projects in Iraqi construction sector. Structure of this study consisting of:

- Justifications of research
- Hypothesis of research
- Methodology of research
- Theoretical part review
- Parameters (factors) influenced on the earned value
- Data collection
- Developed regression models
- Validation of model
- Conclusion, recommendations

Justifications of research: Causes that stand beyond the adopted of this research as following:

- Lack of application of earned value approach in Iraqi bridges projects
- Scarcity of local research related to the topic of earned value for bridges projects
- Application of multifactor linear regression technique as an effective technique in Iraqi construction sector is necessary to ensure successful management for bridges projects with the need of such technique in project management

Hypothesis of research: The research hypothesis is formulated based on only one question: is Multifactor Linear Regression Technique (MLRT) can be applied in construction sector in order to calculate the earned value of bridges projects? Where, MLRT has powerful modeling mechanization with dynamic learning technique and active assessment capacity to predicting the earned value beneath any specified situation.

MATERIALS AND METHODS

Research aims were accomplished using the following stages.

Literature review: It is to narrative the earned value and application multifactor linear regression technique in construction project. The concept of the earned value and

Corresponding Author: Faiq M.S. AL-Zwainy, College of Engineering, Al-Nahrain University, Baghdad, Iraq

regression analysis, benefits, structures and applications in project management field which include the review of literatures involving references, thesis, papers, books and web-site relating to the subject of research, especially which are related to construction industry.

Data collection: Data collection method used in this study based on a lot of historical data for bridges projects collected which were done between 2013 and 2016 from bridges and roads directorate in Iraq construction sector. Historical data assists in providing a relation between the main factors affecting the earned value parameters of the bridges projects to make estimates for new bridges projects. Statistical analysis using computer software, such as Statistical Package for the Social Sciences (SPSS) Version 24. Various statistical analysis methods including nonparametric tests, adjectival statistics, regression and correlation were also hired for modeling over this study.

Developed model and validation: Based on the conclusion of the data analysis, it can be preparing multifactor linear regression model to forecasting earned value for bridges projects and discussion the results from developed model. Finally, can be validity MLR Model.

Conclusions and recommendations: It can be obtained from outcome of the multifactor linear regression model and also, validation model.

Application earned value approach in project management: Earned value management is art and it depending on three key points.

Planned Value (PV): It also known as the Budgeted Cost of Work Schedule (BCWS). PV presents how far along project work is supposed to be at any given point in the project schedule in. PV is a numeric of the budgeted work that is schedule to be performed and it is the established baseline (also known as the Performance Measurement Baseline or PMB) against which the actual progress of the project is measured. Once established, this baseline may only change to reflect cost and schedule changes necessitated by changes in the scope of research.

Earned Value (EV): Known as a Budgeted Cost of Work Performance (BCWP). Earned value is zooming of picture of work progress at a given point in time. EV reflected the amount of work that has actually been accomplished to date, expressed as the planned value for that work.

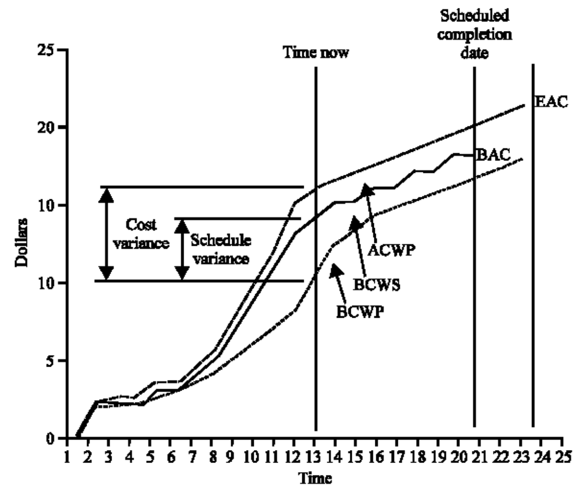

Fig. 1: Earned value curves (Czernigowska, 2008)

Actual Cost (AC): It is known as a Actual Cost of Work Performance (ACWP) is an indication of the level of resources that have been expended to achieve the actual work performance to date (or in given time period). Figure 1 displays the concept of the earned value management. Earned value analysis requires six point as following.

Cost Variance (CV): Calculate of variance between planned cost and actual cost of works done until the date of recording progress in dibs units. If negative, it indicates that the project is over budget:

$$CV = AC\text{-}EV \quad (1)$$

To catch the scale of deviation, it is oftentimes expressed as a part of the budgeted of works performed:

$$CV\% = CV/EV\times 100\% \quad (2)$$

Schedule Variance (SV): Mensuration of variances between the actual progress and the planned progress. Though, it is interpreted as time deviation, it is expressed in money units. In other words, it is the difference between the planned cost of works that have been done and planned cost of works that should have been done by the reporting date. If negative, it indicates a delay:

$$SV = EV\text{-}PV \quad (3)$$

To address any distortion caused by the relative value of activities, it is expressed as a fraction of BCWS:

$$SV\% = SV/PV\times 100\% \quad (4)$$

Cost Performance Index (CPI): Equiponderate the actual value and planned value of works done, if CPI<1, it marks the project has consumed more money than planned, if CPI>1, there had thrift, it can be expressed as:

$$CPI = EV/AC \quad (5)$$

Schedule Performance Index (SPI): It is a difference between the planned cost of works done with planned cost of works planned, if SPI<1, it is mark a delay:

$$SPI = EV/PV \quad (6)$$

Budget at Completion (BAC): It equals BCWS at the planned finish also, it is total planned cost of the whole project.

Estimate at Completion (EAC): It is determined at the date of reporting progress to serve as an estimate of the effect of deviations cumulated from the construction project's begin on the total construction project cost, so, it informs how much the project is going to be in the finish, when the Cost Performance Index (CPI) stays the same (Christensen, 1998):

$$EAC = BAC/CPI \quad (7)$$

During the last few years or so, the use of Earned Value Management (EVM) has increased in many construction engineering problems and has demonstrated some degree of access. Fleming and Koppelman (2005) explained an overall form of earned value management and application this system on the large construction projects. Jigeesh and Bhat (2006) studied the construction project control system and explain earned value management concept for dynamic ambience based on computer simulation. Noori *et al*. (2008) developed a fuzzy control technique to monitoring earned value performance indexes.

Also, a novel method depended on a Multi Period-Multi Product (MPMP) production control problem was clarified and successfully accomplished. By developing an efficient forecasting approach, Lipke *et al*. (2009) developed project manager's efficiency to make better managerial decisions.

In order to evaluation of the schedule performance, must be used a well-established project management technique. Bagherpour *et al*. (2010) modeled the uncertainty associated with activity time in earned value analysis. The approach incorporated a control approach, which is viable within project control. Costs were assumed to be directly associated with fuzzy activity duration estimated through a bottom up hierarchy process. Chou *et al*. (2010) proposed a design, visual architecture and construction based on web to evaluate a knowledgebase management system and project performance concerning combination of earned value management. Pajares and Lopez-Paredes (2011) presented combination of construction project risk management with earned value management to control and monitor the construction projects. Model was built by Warburton (2011) in order to demonstrating calculate final cost for construction projects with faster coverage to the suitable result and less deviation than Estimate At Completion (EAC).

In Iraq, Sarhan *et al*. (2014) proposed EVA as a better index for future performance utilizing follow-up approach and it is possible to predicting cost or schedule overruns at quite an early phase in an infrastructure project. One highway could be utilized as actual implementation for explain the advantages of earned value management in extracting more facts from data about the situation, direction and future project schedule performance and related cost.

Application regression technique in project management: Two technical can be investigating the relationships between factors (variables). First technique called the multivariable (multifactor) linear regression. It attempts to map the relationships between the influential variables and the depended variable with the clear mathematical functions. Mapping functions are initially presumed and later evaluated. They could be non-linear functions (multivariable non-linear regression) or linear functions (multivariable linear regression). Second technique that has been vastly utilized in modern studies for determine the relationships are artificial neural network (Al-Zwainy *et al*., 2012).

In this study, multifactor linear regression can be utilized to evaluation the statistical relationship between an explanatory factors and Estimate at Completion (EAC). Multifactor linear regression model requires some assumptions. It is of the form:

$$Y_i = \beta_0+\beta_1F_1+\beta_2F_2+, ..., +\beta_pF_{ip} \quad (8)$$

where, Y_i the response that corresponds to the levels of the explanatory factors F_1, F_2, ..., F_p at the ith observation.

Table 1: Descriptive statistics

Variables	N	Range	Minimum	Maximum	Mean	SD
Duration	34	318.00	60.00	378.00	209.1765	100.20781
BAC	34	339475450	959343750	1298819200	1111971805.88	97686283.542
PV	34	449750000	1000000000	1449750000	1177402000.00	138077877.510
EV	34	1199750000	250000000	1449750000	718822426.47	347291883.814
AC	34	992451562	239835938	1232287500	675882096.35	314687732.292
EAC	34	735169692	719507813	1454677504	1066120463.53	190683295.497

$\beta_0, \beta_1, \ldots, \beta_p$ are the coefficients in the linear relationship. For a single factor (p = 1), β_0 is the intercept and β_1 is the slope of the straight line defined. Below display several of the past studies that have used regression models in project management.

Thomas (2009) executes statistical analysis procedures to conduct cause-effect analysis on past cumulative productivity measurements, so as to evaluate the significance of the learning curve effects on construction activities. Al-Zwainy *et al*. (2013) developed MLR Model to predicting the construction productivity for marble works of floors. One model was built based on one hundred set of data collected in republic of Iraq for different kinds of construction project. Ten influencing factors are used in MLR Model. Developed MLR Model has ability to forecasting the construction productivity with excellent degree of accuracy of the correlation coefficient (R) was 90.6% with AA% equal to 96.3%. Rasool and Al-Zwainy (2016) founded that Multiple Linear Regression (MLR) have the ability to forecasting construction productivity for brickworks with a very good degree of accuracy of the correlation coefficient equal to 87.28% and average accuracy equal to 92.5%. In the second part of this research was development another mathematical model for the same variables using Logistic Regression Technique (LRT). Results of the study showed that the use of a binary logistic regression gave a logical response results are consistent with the studied case with a moderate degree of average accuracy percentage of 95.8%. Al-Zwainy and Hadhal (2016) carried out a study to building a mathematical model for cost estimating for communication towers projects in Iraq. Multifactor linear regression technique is utilized for cost estimating for communication towers projects. Only seven factors used for cost predicting using MLR Model, involve high and types of tower, security situation, foundation, experience, main cable and site area. MLR Model has flexibility to predicting the cost with an excellent degree of accuracy AA% = 90.1%, R = 98.6% and MAPE = 9.891%.

Factors affecting Estimate at Completion (EAC): Estimate at Complete (EAC) was known as the subject factor and each individual project is used as the basic unit of the oversight. Investigation and assessment of the factors affecting Estimate at Completion (EAC) had become a crucial case facing owner, contractors and construction project managers for a long time, so as to intension control for bridges projects. After fixation the subject factors which are to be forecasting by regression technique, it was necessary to investigation the independent factors to demonstrate each change in earned value. There are many factors as independent factors such as:

- F1:AC, Actual Cost
- F2:EV, Earned Value
- F3:PV, Planning Value
- F4:BAC, Budget at Completion
- F5:D, Duration of project

Researcher has been able to get 34 bridges in Republic of Iraq. Table 1 show the results of the field study where that data and information collected from the different bridges projects sites were usable and it guaranteed the verification and validity of the data used.

Development of multifactor linear regression model: Multifactor Linear Regression Technique (MLRT) is adopted in the study because MLRT is the most widely utilized kind and employing the construction project features in a mathematical model to forecasting EAC. Software (SPSS) is a Statistical Package for the Social Sciences Vertion. About 24 is utilized to build the accurate mathematical model, Table 2-4 shows results of the statistical analysis. The researchers used Enter method as a type of regression analysis. All requested (factors) variables entered as shown in Table 2.

The correlation among input factors is tested symbol (R), titled a linear correlation coefficient, calculating a direction and strength the linear relationship between dependent and independent factors, also, symbol (R^2) called coefficient of determination which exemplify amount of data that is the close to line of best fit as shown in Table 3, correlation coefficient (R) equal to (90.9%) and coefficient of determination (R^2) equal to (82.6%). These results indicate a positive and excellent relationship and a high correlation between dependent factor (EAC) and independent factors (Duration, PV, EV, AC and BAC).

Table 2: Variables entered

Model	Variables entered	Variables removed	Type method
1st	AC, PV, BAC, duration, EV[a]	None	Method of enter

Table 3: Model summary

Model	R	R^2	Adjusted R^2	SE of the estimate
1	0.909	0.826	0.826	9479911.457

Table 4: Unstandardized coefficients of variables

Model 1	Unstandardized coefficients B	t-values	Sig.
Constant	-1297145300.342	-4.6	0.000
Duration	-224146.191	-7.0	0.000
BAC	1.893	64	0.000
PV	-0.732	-20.9	0.000
EV	0.021	0.53	0.600
AC	-0.025	-0.57	0.568

The values in Table 4 show that all of the model coefficients was nonzero, therefore, this model seem auxiliary for forecasting the EAC. Also, MLR Model content all power independent factors that have been investigated. MLR Model can be summarized as following:

$$\text{EAC} = -129714530-22416.191\text{Duration}+1.893\text{BAC}-0.732\text{PV}+0.021\text{EV}-0.025\text{AC} \quad (9)$$

where, EAC: Estimate at Completion as output (dependent) factor.

Validation MLR Model: Four main tools can be used to validating of the MLR Model as following.

Using a table of critical values: The 95% critical values of the sample correlation coefficient as shown in Table 5 may be utilized to assign a valid idea of whether the calculated value of R is significant or not significant. Emulate (R) to the suitable critical value in Table 5. If (R) is not between the negative and positive critical values and then correlation coefficient is significant. If (R) is significant can be used to forecasting.

Value (R) Computed equal to 0.909 using number pf bridges projects (n) equal to 34 data. Degree of freedom (df = n-2 = 34-2 = 32). Critical values correlated with degrees of freedom equal to 32 ranging between (±0.34). If value (R) less than negative critical value or (R) greater than positive critical value, then correlation coefficient (R) is significant. Since, correlation coefficient (R) equal to 0.909 and 0.909>0.34, R is significant and the MLR Model can be applied for estimation the EAC for bridges projects.

"t" test: There is another test of significance of coefficient of correlation in which the value of "t" is computed by the following Eq. 10:

Table 5: Critical values

Degrees of freedom n-2	Critical value (+ and -)
1.00	0.9970
2.00	0.9500
3.00	0.8780
4.00	0.8110
5.00	0.7540
6.00	0.7070
30.0	0.3490
31.0	0.3400
32.0	0.3310
40.0	0.3040
50.0	0.2730
60.0	0.2500
70.0	0.2320
80.0	0.2170

$$t = \frac{R^*\sqrt{N-2}}{\sqrt{1-R^2}} \quad (10)$$

$$t = \frac{0.909^*\sqrt{43-2}}{\sqrt{1-0.826}} = 13.9808 \quad (11)$$

If the computed value of "t" is greater than the table value (t-distribution), the correlation is taken as significant. For 95|% confidence and df = n-1 = 33, the t-table = 2.040 and t computed = 13.9808. Therefore, t computed t-table. This means that the correlation coefficient is highly significant.

Numerical example: Equation was tested against the information used in the MLR Model (Model-EAC) as a numerical example is provided to better demonstrate the application for the mathematical formula. Following data can be given:

- Duration = 150 days
- AC = 974,114.400 ID
- EV = 869,745.000 ID
- PV = 1,159,660.000 ID
- BAC = 1,298,819.200 ID

Estimated value utilizing mathematical Eq. 2 equal to (1470628452 ID), compares well with actual value (measured value) (EAC = 1454677504 ID). The difference is very little equal to (15950947.95 ID), this shows the strength of the predicted MLR developed in this study, Table 6 shows, five new bridges projects as a historical data for each factors of MLR Model. These new historical data which were not included in the MLR Model calibration processes were used as in independent verification examination. While, the actual EAC of bridges projects and the predicted (estimation) values which have presented in Table 7. Figure 2 show the comparison between the predicated value and actual value of the EAC

Table 6: Five new bridges projects historical data

	No. bridges projects				
Factors	1	2	3	4	5
F1 (AC)	974114400	255977500	337500000	1171828125	696742050
F2 (EV)	869745000	269450000	337500000	1562437500	749185000
F3 (PV)	1159660000	1077800000	1350000000	2083250000	1498370000
F4 (BAC)	1298819200	1023910000	1350000000	1562437500	1393484100
F5 (D)	150	210	180	300	270

Table 7: Actual EAC and predicted EAC

No. bridges projects	Actual EAC	Predicated EAC
1	1454677504	1470628452
2	972714500	1014149112
3	1350000000	1432250556
4	1171828125	1299831285
5	1295940213	1403605993

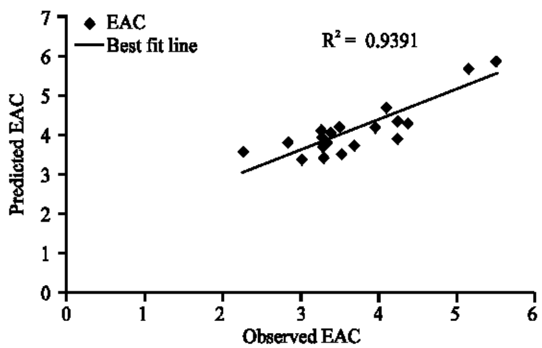

Fig. 2: Comparison of predicted and observed EAC

for bridges projects where, coefficient of correlation (R) is 96.9% and determination coefficient (R^2) equal to 93.91%, therefore, it can be concluded that the developed MLR Model show excellent agreement with actual data.

Average accuracy degree: Statistical measures can be used to evaluation the performance of the MLR Model as following.

Mean Absolute Percentage Error (MAPE):

$$MAPE = \frac{\left(\sum_{i=1}^{n}\frac{|A-E|}{A}*100\%\right)}{n} \quad (12)$$

Average Accuracy (AA%):

$$AA\% = 100\%-MAPE \quad (13)$$

Finally, Mean Absolute Percentage Error (MAPE) created by MLR Model (EAC) equal to 6.14 %. Therefore, Average Accuracy (AA%) equal to 93.86%. Thus, EAC Model can be presented an excellent with the actual data.

CONCLUSION

This study aims to developing MLR Model for EAC of bridges projects. One MLR Model was built up based upon 43 set of database composed from Iraqi bridges and roads directorate. Five factors (Duration, PV, EV, AC and BAC) was examined in MLR Model. MLR Model has the efficiency to estimate the EAC for bridges projects with very high grade of reliability with 93.86% and the determination coefficients R^2 = 0.9391 and correlation coefficient R = 0.9690 for the built model. Those results show the relationship between objective and subjective variables of built model was excellent.

REFERENCES

Al-Zwainy, F.M.S. and N.T. Hadhal, 2016. Building a Mathematical Model for Predicting the Cost of the Communication Towers Projects Using Multifactor Linear Regression Technique. Int. J. Constr. Eng. Manage., 5: 25-29.

Al-Zwainy, F.M.S., H.A. Rasheed and H.F. Ibraheem, 2012. Development of the construction productivity estimation model using artificial neural network for finishing works for floors with marble. ARPN J. Eng. Appl. Sci., 7: 714-722.

Al-Zwainy, F.M.S., M.H. Abdulmajeed and H.S.M. Aljumaily, 2013. Using multivariable linear regression technique for modeling productivity construction in Iraq. Open J. Civil Eng., 3: 127-135.

Bagherpour, M., A. Zareei, S. Noori and M. Heydari, 2010. Designing a control mechanism using earned value analysis: An application through production environemnt. Int. J. Adv. Manuf. Technol., 49: 419-428.

Chou, J.S., H.M. Chen, C.C. Hou and CW. Lin, 2010. Visualized EVM system for assessing project performance. Automation Construc., 19: 596-607.

Christensen, D.S., 1998. The costs and benefits of the earned value management process. J. Parametrics, 18: 1-16.

Czernigowska, A., 2008. Earned value method as a tool for project control. Constr. Archit., 3: 15-32.

Fleming, Q.W. and J.M. Koppelman, 2005. Earned Value Project Management. 3rd Edn., Project Management Institute, Pennsylvania.

Jigeesh, N. and M.S. Bhat, 2006. Modeling and simulating the dynamics in project monitoring and earned value analysis. J. Adv. Manage. Res., 3: 26-43.

Lipke, W., O. Zwikael, K. Henderson and F. Anbari, 2009. Prediction of project outcome: The application of statistical methods to earned value management and earned schedule performance indexes. Intl. J. Project Manage., 27: 400-407.

Noori, S., M. Bagherpour, A. Zareei, 2008. Applying fuzzy control chart in earned value analysis: A New Application. World Applied Sci. J., 3: 684-690.

Pajares, J. and A. Lopez-Paredes, 2011. An extension of the EVM analysis for project monitoring: The cost control index and the schedule control index. Int. J. Project Manage., 29: 615-621.

Rasool, S.H. and F.M.S. Al-Zwainy, 2016. Estimating Productivity of Brickwork item using Logistic and Multiple Regression Approaches. Scholars J. Eng. Technol., 4: 234-243.

Sarhan, A.F.M., A.M. Ibrahim and D.S. Mohsen, 2015. Earned Value Management in Construction Project. OmniScriptum GmbH & Co. KG, Germany, Europe, ISBN:9783659750786, Pages: 140.

Sarhan, F.M., I.A. Mohammed and D. Sameer, 2014. Evaluation The Performance of The Infrastructure Project Using Earned Value Management. Int. J. Civil. Eng., 5: 145-155.

Thomas, H.R., 2009. Construction learning curves. Pract. Period. Struct. Des. Constr., 14: 14-20.

Warburton, R.D., 2011. A time-dependent earned value model for software projects. Intl. J. Project Manage., 29: 1082-1090.