RESEARCHERID
THOMSON REUTERS

[Neena Pookot *et al.*, 6(6): June, 2017]
IC™ Value: 3.00

ISSN: 2277-9655
Impact Factor: 4.116
CODEN: IJESS7

IJESRT
INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY

A COMPARISON OF DIFFERENT COST FORECASTING METHODS USING EARNED VALUE METRICS

Neena Pookot
M.Tech (Construction Engineering & Management) Student,
Cochin College of Engineering and Technology, India.

DOI: 10.5281/zenodo.809176

ABSTRACT
Scope, Time and Cost are the three main attributes of a Project and they should be continuously monitored on their performance. Efficient decision making mandates the accuracy of forecasted estimations of a Project's final value termed as Cost Estimate At Completion (CEAC) in Earned Value Management. The greatest benefit gained by use of EVM is its ability to predict project outcome and potentially prevent project failure. To develop project management proficiency, organizations need to look at the critical elements of EVM as pragmatic stepping stones to prioritizing which project management processes are most important for successful delivery. In this paper the most commonly used predictive tools based on the performance indices, which are compared with a nonlinear regression based CEAC. Gompertz growth model is adopted, the input data is modified with cost variance and schedule variance. The output is modified with earned schedule which helps in predicting CEAC more accurately. Five data sets are used in the comparative study of CEAC methods. The model based on nonlinear regression is found to be the most accurate and precise method in the early stages of the Project as compared to other Index based methods.

KEYWORDS:EVM, ES, Nonlinear Regression Analysis, Performance Management

I. INTRODUCTION
The success of a project is obtained through the proper management from the beginning to the end of the project, Project management is an art and a science; it demands the proper management of people and the knowledge about the various techniques, methods, processes etc. involved in a project. Each and every project undergo through the triple constraints of time, money and quality; the balancing of which results in the satisfactory project output. While a project is being executed the monitoring and controlling is very essential. The performance reporting is a part of the project monitoring and control process. The earned value method is a widely accepted performance measurement technique. This methodology is useful for reporting the status of the project as well as to predict the future performance based upon the past performance.
Earned Value Management system is emerged out of these ambiguities and showed the way for better performance management system. The Performance Measurement Baseline is highly important to earned value management because it provides the baseline plan to measure the project's performance. It is the sum of the project's planned cost over time and establishes the scope, schedule and budget for a project. A baseline should accurately represents the only authorized work on the contract. It includes a realistic network schedule baseline, and a realistic time phased spread of budget to the baseline schedule.
The Project Management Institute defines earned value management (EVM) as "a management methodology for integrating scope, schedule, and resources, and for objectively measuring project performance and progress. Performance is measured by determining the budgeted cost of the work performed (i.e. earned value) and comparing it to the actual cost of the work performed (i.e. actual cost). Progress is measured by comparing the earned value to the planned value." (PMI 2004, p. 359). Earned Value Management has got any modifications as time proceeds. This paper also modifies the EVM using nonlinear regression method which incorporates Cost Variance and Earned Schedule.

ISSN: 2277-9655
Impact Factor: 4.116
CODEN: IJESS7

This project incorporates the earned value management and earned schedule data in nonlinear regression analysis and develops a model for better performance analysis. Five infrastructure projects were selected which are comparable in project cost and duration.

II. METHODOLOGY

This chapter describes the procedure followed to evaluate the growth models as predictive tools for Estimates at Completion (EAC). The objective of this evaluation is to determine the best method in predicting EAC commonly used Index methods, including the CPI, the SCI, Composite Index and the modified Regression model. Gompertz growth model is taken and modified for this evaluation. This chapter provides the methodology required for solving these models. This portion includes the introduction of software utilized in this study.

Data collection

For the comparative study of different methods for calculating EACs five completed projects were selected.

About the software

Microsoft Excel was used as the basic tool for all sorts of calculations. For the analysis of non-linear regression Minitab 17 was used. Minitab is a statistics package which provides a simple, effective way to input statistical data, manipulate that data, identify trends and patterns, and then extrapolate answers to the problem at hand.

CEAC methods adopted

Methods for forecasting of Cost Estimate At Completion adopted are,

CEAC1 = AC + (BAC-EV) (1)

CEAC2 = BAC/CPI (2)

CEAC3 = BAC/(CPI*SPI) (3)

Using Regression
CEAC4 = AC(x) + [growth model(1)-growth model(x)] (4)

Modified regression method
CEAC5 = AC(x) + [growth model(1/SPI(t)-growth model(x)] (5)

Modification over regression model for CEAC calculation

This section develops the new methodology in following steps.

Step 1 : Developing S-Curve

The first step is to develop the S-curve incorporating the actual cost and time with the planned values. The new curve should reflect the past and the future. Actual Time Vs Actual Cost curve will be merged with remaining Planned Value Curve. Remaining planned value is identified from Earned Value and the remaining time is identified with the help of Earned Schedule. Each next y-coordinate is the cumulative value of the planned value over the actual cost and each next x-coordinate is the cumulative value planned value month over the modified time.

Step 2 : Finding GGM parameters

The next step is to develop the non-linear regression model that fit the cumulative S-curve prepared. Gompertz Growth Model (GGM) has three parameters α, β and γ.

$$GGM(x) = \alpha e^{[-e^{(\beta-\gamma x)}]} \quad (6)$$

1) α-asymptote represents the never-attained value of the final cost as time tends to infinity,
2) β is the y-intercept (which represents the initial size of a project cumulative cost),
3) the γ-scale parameter that governs the rate of the cost growth.

In equation (6)the predictor x represents normalized time points with its maximum value equal to 1.00 (100% time-complete); i.e., the PD of a project. The corresponding value of the response variable is normalized points of the combined-to-date AC (from a project beginning to AT) and PV (from AT and onto 100%) with its maximum value (BAC) of 1.00. This allows building the combined data set of AC

ISSN: 2277-9655
Impact Factor: 4.116
CODEN: IJESS7

and PV with respect to time points used in the model-fitting and determining the three parameters for the GGM nonlinear equation. For this purpose the Minitab software package is used. Unlike linear regression, building a nonlinear regression model requires specifying the starting values for a model parameters and the algorithm for approximation of these values. However, no standard procedure exists to define good starting values. Unless one knows the starting values of the model parameters based on prior information, to define a good set of these values in nonlinear model fitting is difficult because it is because of the existence of a nonlinear relationship between a predictor and a response variable and the phenomena the model represents. To estimate values of these parameters, both linear and nonlinear regression use the least squares (LS) method of approximation.

The most common assumption in curve-fitting is that data points are randomly scattered around an ideal curve with the scatter in accordance with a Gaussian distribution. Taking into account this consideration, the LS approach minimizes the sum of the squared errors (SSE, the difference between the estimated values and actual input values of the parameters) of the vertical distances of the points from a curve. The research reported in this paper applies a Gauss-Newton algorithm for the LS approximation, which obtains convergence iteratively close-to-linear regression that are not heavily dependent on the starting values. The iteration process continues until the algorithm converges to determine the parameter values within the specified tolerance on the minimum SSE. Minitab17 iterates the values till it can't give a smaller value and it can take up-to 200 iterations. In this paper regression analysis, the levels of the confidence interval (CI) are set as 95% which is an accepted standard. The CI gives a range of estimate values between two limits in which the actual values are more likely to fall. After inputting all the required data in Minitab data sheet, it plots the GGM curve and gives the output data which includes the values of α, β, γ.

Step 3 : Calculating the Project CEAC

Using the parameters α, β, γ in equation 3.6 the response variables are fitted into GGM. To compute the CEAC equation 3.7 is used, which assumes the values of the growth model when a project is to-date and 100% complete. This formula is similar to the classical index based formula in Eq. (1) because in both equations new estimate is added to actual cost. However, in Eq. (7) the remaining portion of the CEAC is computed based on the nonlinear growth model results, whereas Eq. (1) corrects the remaining portion of the BAC by CPI. Trahan (2009) presents the generic form of CEAC calculation using growth model in Eq. (7) and developed the nonlinear growth model by regressing the response values of the AC for the entire project life cycle against the corresponding time increments. Unlike the approach of Trahan (2009), the research reported in this paper combines values of current AC and PV, as presented previously. The GGM curve is made and the values at completion and the actual time x are found out by using Eq. (6) and CEAC(x) is calculated as Eq. (1).

$$CEAC(x) = AC(x)+[GGM(1)-GGM(x)]*BAC \qquad (7)$$

The CEAC equation used above calculates the final value of GGM for the planned duration (PD) but the SPI(t) of the project may not be 1.00. If the SPI(t) is less than 1.00 the planned duration (PD) has to be modified. Lipke (2003) introduced Earned Schedule which can predict the Time Estimate At Completion (TEAC).

$$TEAC = PD/SPI(t) \qquad (8)$$

Here PD=1.00 therefore we modify Eq. (7) as follows by incorporating the extra cost that can be incurred because of poor Schedule Performance Index.

$$CEAC(x) = AC(x)+\{GGM[1/SPI(t)]-GGM(x)\}*BAC \qquad (9)$$

Data comparison

To determine which method is best to forecast CEAC in early, middle and later stages of the project. Respective forecasts were compared with and errors found by MAPE and MPE.

This section provides the framework for assessing the quality of the proposed methodology and other index based formulae. This study measures CEAC accuracy by a percentage error (PE) and the mean absolute percentage error (MAPE) for early, middle, and late stages. PE is the difference between CEAC and Cost at Completion (CAC) expressed as a percentage of CAC with a negative value suggesting underestimation and a positive value suggesting overestimation. MAPE is referred to as the average of the absolute values of differences between CEAC and CAC over the number of projects tested Eq. (10) and Eq. (11)) are used to compute these measures.

$$PE\% = \frac{CEAC - CAC}{CAC} 100\% \tag{10}$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n} \frac{|CEACi - CACi|}{CACi} = \frac{1}{n}\sum_{i=1}^{n} |PEi|\% \tag{11}$$

Where, CAC—Cost at Completion; n—number of projects.

The precision of the results are evaluated with standard deviation. Standard deviation is a quantity expressing by how much the members of a group differ from the mean value for the group. It is expressed as equation 3.12, it calculates the deviation of the percentage errors from the Mean percentage errors (MPE). The smallest value gives the more reliable method over different projects.

$$SD\% = \sqrt{\frac{\sum_{i=1}^{n} (PEi - MPE)^2}{n}} \% \tag{12}$$

III. DATA ANALYSIS

The conventional CEAC methods and proposed CEAC methodology is demonstrated through five infrastructure projects. Five past projects were selected for the analysis of this project because they have all the necessary data readily available at ones disposal and one would know the endpoint of the project which helps significantly in understanding the method in which CEAC methods works on projects. These projects all have medium-sized budgets with an average BAC close to 3 Crore INR and planned duration varying from 9 to 20 months.

Table 1. List of projects and their average performance indicators

Project	Planned value		CPI	SPI
	Time(Months)	Cost(Lakhs)		
Project A	18	231.28	0.844	0.957
Project B	20	480.00	0.888	0.960
Project C	9	287.50	0.912	0.857
Project D	9	360.73	1.040	0.913
Project E	10	90.60	0.970	1.060

Table.1 provides the list of these projects along with their associated information. For projects having CPI<1.00 are cost overrun and those projects having CPI>1.00 are cost underrun. Similarly for projects having SPI<1.00 are late finishing projects whereas SPI>1.00 are early finishing projects. Here out of five selected projects four of them experienced cost overrun. Only Project 4 was cost underrun with an average CPI=1.04 and Project 5 was early finishing with an average SPI=1.06. Other four projects reported a schedule delay.

For the demonstration purpose Project A is selected as a numerical example, which is a cost overrun schedule delayed project. This paper calculates CEAC at early stage, middle stage and later stage. The CEAC calculation at middle stage is showing here and then show the results for all cases. Table 2. shows periodic and cumulative data for Planned Value, Earned Value and Actual Cost.

Index based forecasting methods use performance indicators to forecast future which only depends on the past performance. Regression analysis is useful early in a project life when little or unreliable EV performance and progress data are available.

The advantage of regression modeling in this context is that it extrapolates past available data with future planned data, whereas the conventional EV approach solely relies on past performance and progress. This extrapolation is achieved through the development of the growth model in which the values of its parameters show a relationship between past, current, and future project performance and progress. Using the growth models CEAC for projects with schedule delays are computed for its estimated completion time.

Table 2. List of projects and their average performance indicators

TIME	PLANNED VALUE(Lakhs)		EARNED VALUE(Lakhs)		ACTUAL COST(Lakhs)	
Months	Periodic	Cumulative	Periodic	Cumulative	Periodic	Cumulative
1	6.06	6.06	6.04	6.04	8.23	8.23
2	6.93	12.99	6.69	12.73	9.89	18.12
3	9.06	22.05	8.62	21.73	12.03	30.15
4	12.86	34.91	12.09	33.44	13.89	44.04
5	14.87	49.78	14.39	47.83	16.45	60.49
6	15.65	65.43	14.87	62.7	15.72	76.21
7	16.39	81.82	15.41	78.11	16.85	93.06
8	16.24	98.05	15.04	93.15	15.92	108.98
9	17.42	115.48	-	-	-	-
10	17.66	133.14	-	-	-	-
11	17.89	151.03	-	-	-	-
12	17.00	168.03	-	-	-	-
13	16.44	184.47	-	-	-	-
14	14.25	198.72	-	-	-	-
15	12.77	211.49	-	-	-	-
16	9.00	220.49	-	-	-	-
17	6.74	227.23	-	-	-	-
18	4.05	231.28	-	-	-	-

Developing S-Curve

According to our methodology the first step is to create the S-curve for that the planned value till month 9 has to be modified with Earned Value and Actual Cost. The Actual Cost gives the total cost incurred till date and the Earned Value shows the portion of the PV which is completed. The remaining PV data has to be merged with past available data. The EV value which is earned till month 9 is 109.35 but the cost incurred for the same is 126.85 which is more than EV value. For merging past data into PV data, the EV value has to be modified to AC and the periodic values will be added up to get the cumulative Planned Value. The periodic values are available for each month only so the actual cost incurred for earning Earned Value corresponding to 9 month and the time taken for the same are calculated as in Table.3. Linear interpolation Eq .(1) is used to find the actual cost and time taken for the same.

$$y = y_1 + (x - x_1)\frac{(y_2 - y_1)}{(y - y_1)} \tag{13}$$

So the time taken to earn 9 months schedule is 9.369 months and the cost incurred for it is 133.157 with a cost variance of -17.677. In order to prepare new planned value curve the periodic values of planned value should be added to the actual cost-actual time curve. For this the actual cost incurred for earning a particular month schedule is calculated using linear equations as shown in Table 3. The periodic planned values in hand can be easily added to these values to get the cumulative S-curve.

Table 3. Comparison table for motoring mode

ES	Actual Time Taken	SV(t)	Actual Cost(AC)	CV=EV-AC
1	1.003	-0.003	8.260	-2.200
2	2.030	-0.030	18.483	-5.493
3	3.058	-0.058	30.954	-8.904

ISSN: 2277-9655
Impact Factor: 4.116
CODEN: IJESS7

4	4.102	-0.102	45.720	-10.810
5	5.131	-0.131	62.551	-12.771
6	6.177	-0.177	79.195	-13.765
7	7.247	-0.247	96.987	-15.167
8	8.32	-0.302	114.385	-16.335
9	9.369	-0.369	133.157	-17.677

Finding GGM Parameters

The next step of the methodology is about developing the regression-based nonlinear growth model that will be used to fit its s-shaped curve to the cumulative cost curve of Project A. For the nonlinear regression analysis the predictor and response are made by normalizing PD and BAC to unity.

In this paper Gompertz growth model is used for curve fitting. Minitab finds values for Gompertz equation Eq. (6) based on the options the user set for the non nonlinear regression analysis.

$$GGM(x) = \alpha e^{[-e^{(\beta-\gamma x)}]} \qquad (6)$$

Table 4. shows both the initial input data and the results obtained from Minitab17. After 9 months of execution, the project is 47% complete and therefore this is the period in which the CEAC is computed. The periodic planned values are added to the actual cost as well the cumulative is tabulated. Predictor is made by normalising PD to unity and Response by normalising BAC to unity. The fig.1. shows the Minitab window in which predictor and response is inputted and the growth model function is loading.

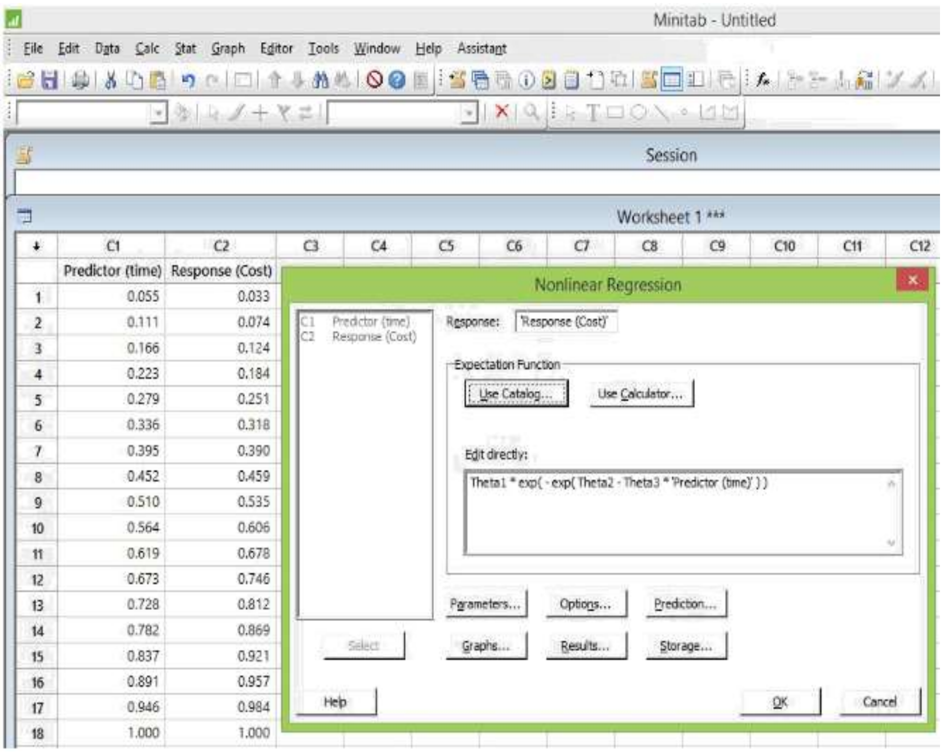

Figure.1. Inputting predictor, response and loading growth model equation.

For nonlinear regression unlike linear regression we have to define the starting values for the three parameters. Taking into account the normalization to unity of both the predictor and response variables 1.0 is used as a starting value for all parameters. Fig. 2. shows the options in Minitab the confidence level is set at 95% with the Gauss-Newton algorithm to converge on the minimum SSE. The maximum number of iterations is 200 with a default-set convergence tolerance of 1 × 10−5.

ISSN: 2277-9655
Impact Factor: 4.116
CODEN: IJESS7

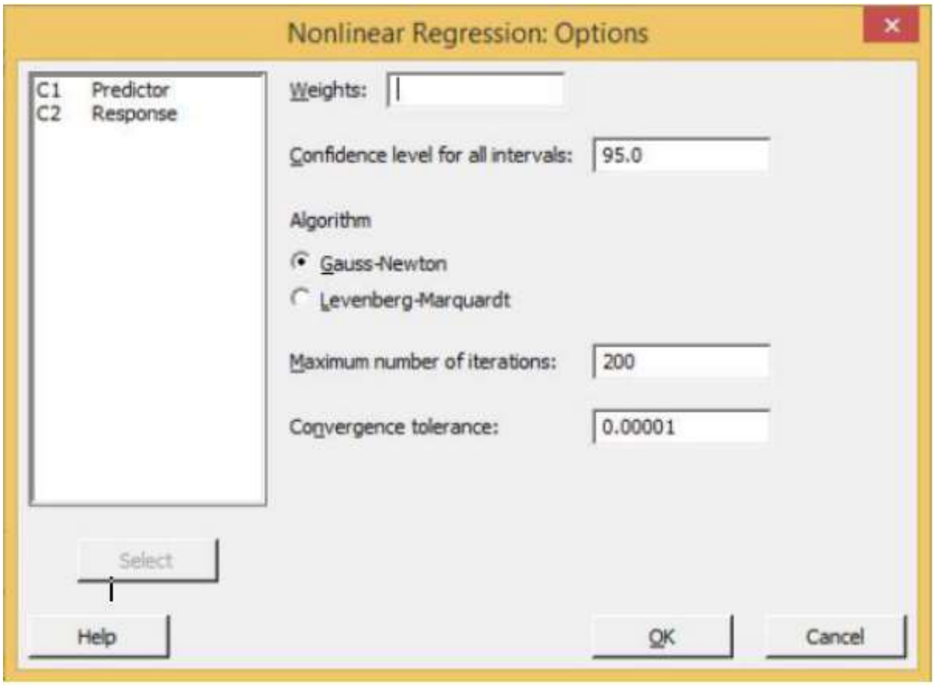

Figure.2. Options for non linear regression in Minitab

Table 4. Curve fitting by non-linear regression model

ES	PV Periodic(Lakhs)	Actual Time Taken(months)	Actual Cost(Lakhs)	Predictor	Response	Fitted Curve
1	6.06	1.003	8.260	0.055	0.033	0.048
2	6.93	2.030	18.483	0.111	0.074	0.080
3	9.06	3.058	30.954	0.166	0.124	0.122
4	12.86	4.102	45.720	0.223	0.184	0.1766
5	14.87	5.131	62.551	0.279	0.251	0.239
6	15.65	6.177	79.195	0.336	0.318	0.310
7	16.39	7.247	96.987	0.395	0.390	0.388
8	16.24	8.32	114.385	0.452	0.459	0.466
9	17.42	9.369	133.157	0.510	0.535	0.545
10	17.66	10.369	150.817	0.564	0.606	0.617
11	17.89	11.369	168.707	0.619	0.678	0.685
12	17.00	12.369	185.707	0.673	0.746	0.748
13	16.44	13.369	202.147	0.728	0.812	0.806
14	14.25	14.369	216.397	0.782	0.869	0.859
15	12.77	15.369	229.167	0.837	0.921	0.906
16	9.00	16.369	238.167	0.891	0.957	0.948
17	6.74	17.369	244.907	0.946	0.984	0.985
18	4.05	18.369	248.957	1.000	1.000	1.018

Fig. 3 presents the Gompertz-fitted s-curve of project A. The curve fits the AC-PV data of the project very well; i.e., all response values are in the CI (upper and lower dashed curves). The writers are confident with a probability of 95% that the CI estimates the real AC-PV data. The values for the parameters are obtained from the equation developed by Minitab as shown in Eq. (14)

$$GGM(x) = 1.21768e^{[-e^{(1.3416-3.05972x)}]} \qquad (14)$$

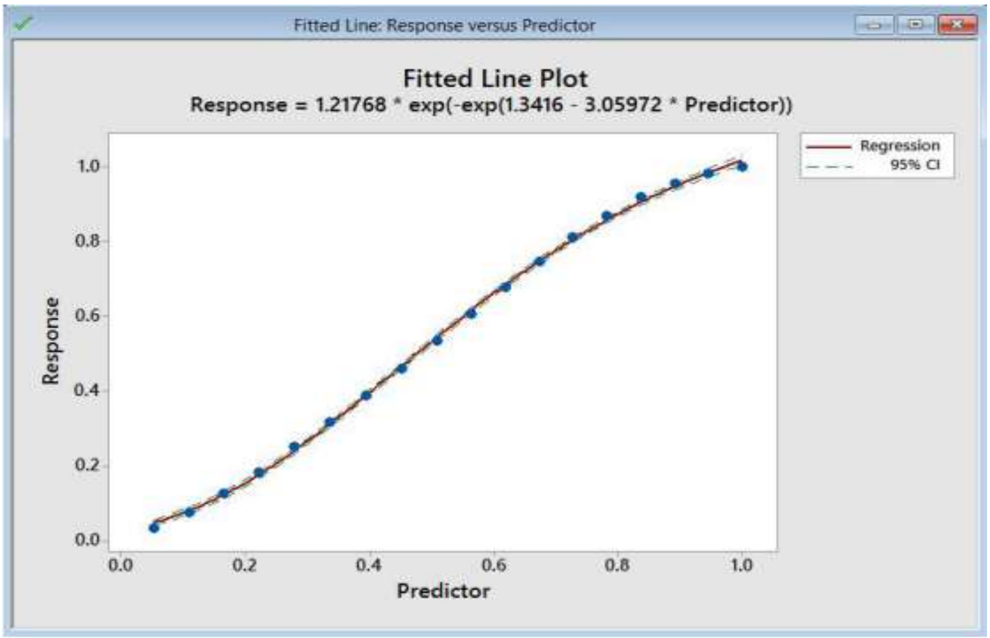

Figure. 3. Fitted Line Plot with 95% CI (Minitab 17)

Calculating The Project CEAC

The next step is to find out the Cost Estimate at Completion using regression method. For the comparative study five methods are chosen, i.e., three classical index based formula, regression based and the modified regression based.

The PMBOK Guide provides different cost estimates, based on different assumptions. In this section, these estimates are reviewed, simplified and enhanced. They are given a sequential subscript to differentiate among them.

When current analysis shows that the assumptions underlying the Original estimate are flawed, or no longer applicable due to changed conditions affecting the activity, work package, or project, a new ETC needs to be developed; CEAC is the sum of the cumulative AC plus the ETC.

As a formula, CEAC = AC + ETC. This method needs a re-estimate so this method is not chosen for the study. When current analysis shows that past performance is not a good predictor of future performance, that problems or opportunities which affected performance in the past will not occur in the future, and that future performance will parallel the original plan, the EAC2 is the sum of the cumulative AC plus the original budget for the remaining work(BAC-EV) is shown in Eq. (1).

CEAC1 = AC + (BAC-EV) (15)

We know that, CV = EV- AC

i.e., CEAC1 = BAC - CV (1)

Where BAC is the Budget At Completion and CV is the Cost Variance

When current analysis shows that past performance is a good predictor of future performance, that performance to date will continue into the future, and that efficiencies or inefficiencies observed to date will prevail to completion, the EAC2 as given in eq. (2) is the sum of the cumulative AC plus the original budget for the remaining work (BAC-EV), modified by a performance factor, which is usually the cumulative CPI. (Anbari 2003).

CEAC2 = BAC/CPI (2)

ISSN: 2277-9655
Impact Factor: 4.116
CODEN: IJESS7

Table 5. Estimate at Completion using IB methods

Months	CEAC1= BAC-CV	CEAC2=BAC/CPI	CEAC3=BAC(CPI*SPI)
1.003	233.470	315.138	316.182
2.030	236.670	329.206	335.930
3.058	240.080	326.609	337.317
4.102	241.880	304.592	317.982
5.131	243.940	292.497	304.422
6.177	244.790	281.114	293.354
7.247	246.230	275.546	288.634
8.302	247.110	270.584	284.818
9.369	248.780	268.293	283.333

Another performance factor which is generally used is the schedule cost index which is the product of Cost performance Indicator and Schedule performance Indicator (CPI*SPI).

CEAC3 = BAC/(CPI*SPI) (3)

CEAC calculation based on regression analysis Eq. (7) takes the sum of the actual cost and the product of the differences between the growth model at final time and growth model at actual time.

CEAC4(x) = AC(x)+[GGM(1)-GGM(x)]*BAC (7)

The Eq. (7) takes the final duration in calculating CEAC but the final duration can be calculated using earned schedule method and for that time the growth model value has to be found out. This value has to be incorporated in finding CEAC as in equation. (9).

CEAC5(x) = AC(x)+{GGM[1/SPI(t)]-GGM(x)}*BAC (9)

The index based methods are taken as EAC1, EAC2 and EAC3 and the regression based models are denoted as EAC4 and EAC5. Where EAC4 is the regression based CEAC calculation as in Eq. (7) and EAC5 is the modified form of the regression based CEAC as in Eq. (8).

At the middle stage of the construction project the S-curve is modified and fitted with Gompertz regression which gives the parameters as in Eq. (14). Using this equation CEAC at 47% completion is calculated as follows.

CEAC4(x) = AC(x)+[GGM(1)-GGM(x)]*BAC (7)

= 133.15[GGM(1.00) - GGM(0.51)]* 248.957

= 250.758

Here the final time used is difference between planned duration and schedule variance (PD-SV). But the final duration can be different and the widely adopted method for finding final duration is Earned Schedule method by Lipke. Incorporating the estimated duration the CEAC is found out as,

CEAC5(x) = AC(x)+{GGM[1/SPI(t)]-GGM(x)}*BAC (9)

= 133.15 + [GGM(1.00/0.862) - GGM(0.51)* 248.957

= 258.742

ISSN: 2277-9655
Impact Factor: 4.116
CODEN: IJESS7

Table 6. Comparing CEAC Methods

PROJECTS	% Complete	CEAC1	CEAC2	CEAC3	CEAC4	CEAC5
PROJECT 1 231,280	25%	241.880	304.592	317.982	249.114	253.057
	50%	248.780	268.293	283.333	250.759	258.742
	75%	255.820	261.678	278.510	264.659	273.353
PROJECT 2 480	25%	496.590	560.299	585.723	519.031	532.273
	50%	508.910	541.002	576.367	518.486	533.340
	75%	523.490	534.524	564.178	547.539	562.559
PROJECT 3 287.5	25%	293.600	311.361	360.077	273.622	319.355
	50%	298.400	309.067	376.071	296.495	318.894
	75%	310.700	320.455	388.219	304.297	325.278
PROJECT 4 360.738	25%	349.896	316.859	325.910	354.431	366.279
	50%	353.351	347.331	381.210	353.003	360.722
	75%	353.881	352.280	389.816	315.307	325.501
PROJECT5 906	25%	911.000	932.805	866.570	882.398	871.272
	50%	932.000	945.457	869.441	944.142	934.471
	75%	941.000	943.795	897.674	955.885	947.709

Percentage error (PE) shows the effectiveness of each method in finding CEAC. The accuracy of the estimates of the equations is based on a comparison of percentage error (PE), which is termed the difference between the actual and estimated values of final cost expressed as a percentage; and on the mean absolute percentage error (MAPE) of the number of valid projects tested. These two measures were computed in accordance with Eq. (10) and Eq. (11).

$$PE\% = \frac{CEAC - CAC}{CAC} 100\% \qquad (10)$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n} \frac{|CEACi - CACi|}{CACi} = \frac{1}{n}\sum_{i=1}^{n} |PEi|\% \qquad (11)$$

Table 7. Percentage Error of CEAC from from Cost at Completion

PROJECTS	EVALUATION	EAC1	EAC2	EAC3	EAC4	EAC5
PROJECT 1	Early	-6.858	17.291	22.447	-4.073	-2.554
	middle	-4.201	3.313	9.104	-3.439	-0.365
	late	-1.490	0.766	7.247	1.914	5.261
PROJECT 2	Early	-6.189	5.847	10.649	-1.949	0.344
	middle	-3.861	2.201	8.882	-2.052	0.754
	late	-1.107	0.977	6.579	3.436	6.274
PROJECT 3	Early	-9.578	-4.108	10.895	-15.731	-1.646
	middle	-8.100	-4.814	15.821	-8.686	-1.788
	late	-4.312	-1.308	19.562	-6.284	0.178
PROJECT 4	Early	0.148	-9.308	-6.717	1.446	4.837
	middle	1.137	-0.586	9.111	1.037	3.247
	late	1.289	0.830	11.574	-9.752	-6.834
PROJECT 5	Early	-4.307	-2.016	-8.974	-7.311	-8.480
	middle	-2.101	-0.687	-8.672	-0.825	-1.841
	late	-1.155	-0.862	-5.707	0.408	-0.451

The Table 7. shows the percentage error of the estimated cost from the original cost at completion. The positive sign shows the estimated cost is less than the original cost at completion and the negative sign shows estimated cost is more than that of the original cost. But we will be more interested to know

ISSN: 2277-9655
Impact Factor: 4.116
CODEN: IJESS7

which estimate is closer to original cost at the end. So we take the absolute value of the percentage error which is given in Table 8.

MAPE –Mean Absolute Percentage Error

Mean Absolute Percentage Error of the CEAC methods is found out from the percentage error of the CEAC methods in five projects. The Table 8. shows the MAPE values and the figure.4 depicts the MAPE at each stages of evaluation.

Table 8. MAPE of the methods chosen over five projects

EVALUATING STAGE	EAC1	EAC2	EAC3	EAC4	EAC5
EARLY	5.416	7.714	11.937	6.102	3.572
MIDDLE	3.880	2.320	10.318	3.208	1.599
LATER	1.871	0.949	10.134	4.359	3.800

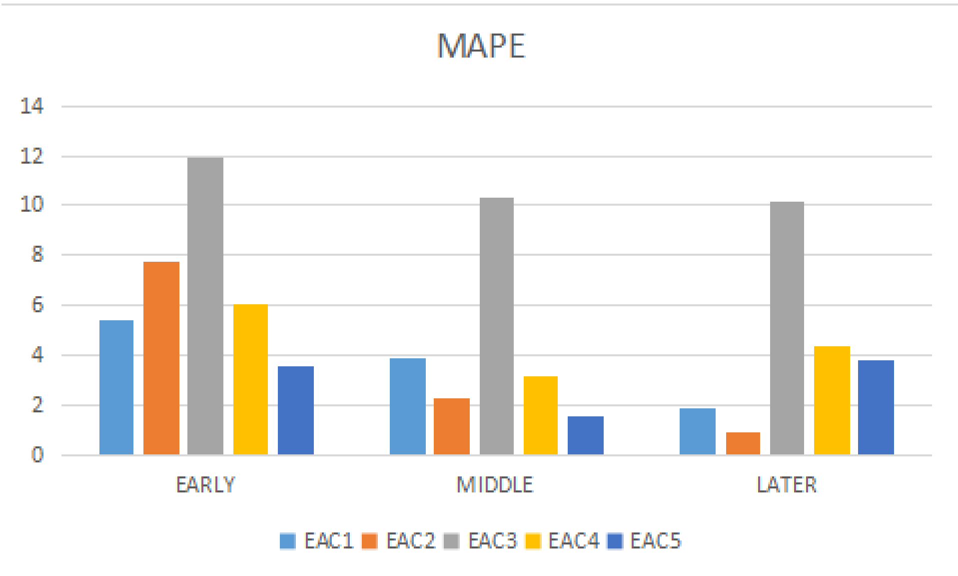

Figure .4. MAPE of CEAC calculation for five Projects.

Standard Deviation (SD)

The narrowness of the forecast error is measured by the standard deviation, which is an indicator of the statistical dispersion of the values of prediction errors from the average forecast within the population. Standard deviation is computed by Eq. (12). A smaller value of SD indicates that the cost estimates calculated by a particular model are closer to its MPE and hence produce more precise CEACs.

$$SD\% = \sqrt{\frac{\sum_{i=1}^{n}(PEi - MPE)^2}{n}}\% \qquad (12)$$

Table 9. Standard Deviation of percentage errors from MPE

EVALUATION STAGE	EAC1	EAC2	EAC3	EAC4	EAC5
EARLY	1.47	8.74	5.50	6.06	1.21
MIDDLE	1.92	3.60	3.22	2.86	1.04
LATER	1.43	1.03	5.97	4.27	2.67

ISSN: 2277-9655
Impact Factor: 4.116
CODEN: IJESS7

The standard deviation shown in table 4.10 is the standard deviation of the projects with delay and cost underrun. The deviation of the percentage errors from the mean percentage error is calculated by standard deviation. The smallest value gives the best method of cost forecasting.

IV. RESULTS AND DISCUSSION

Time and cost overruns are the key challenges for the infrastructure industry. About 40% of the ongoing infrastructure projects are running behind schedule. As per project managers, most of the delays are due to extraneous reasons that can be avoided by adhering to appropriate risk management, time management and change management processes. Proper performance analysis enables project managers to execute and review their projects in a more structured manner. Figure.5. shows the status of the 564 infrastructure projects in India. Almost half of them are delayed projects and one-third of it has not fixed the date of commissioning. Among them only 1% of the projects are ahead of schedule.

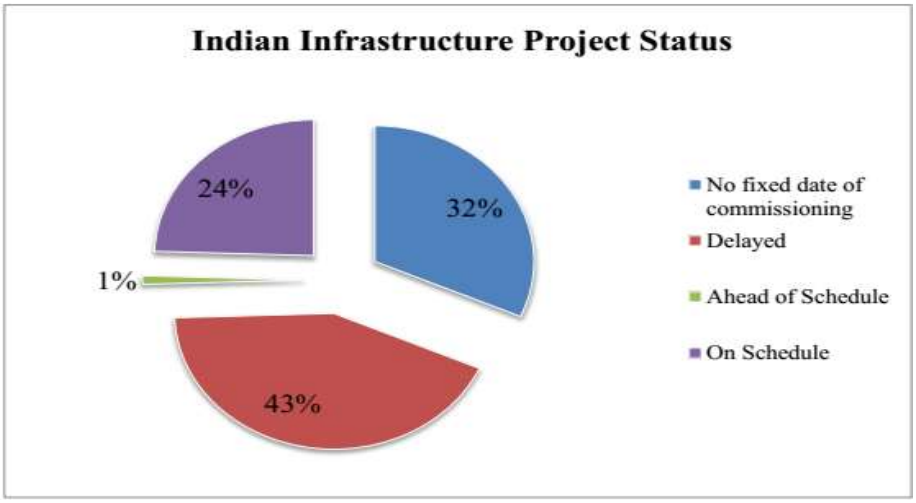

Figure . 5. Project status of 564 infrastructure projects as of May 2012 (PMI study)

So the study narrows down to the projects with delay. The comparative study of the CEAC methods shows that there is no particular method to find CEAC which stands as the best method. Even though the method introduced by modifying nonlinear regression shows to be a good predictor for projects which has CPI and SPI less than unity. The percentage error for projects with cost underrun and running ahead of schedule shows more error in regression based method. But for the projects having delay and cost overrun the regression based method shows the best results. The percentage error for the projects is shown in Table 10. and the light shaded cells showing least error. It is clear from the table that the EAC5 is the best method. The accuracy of the method is tested by MAPE and the precision of the method is checked by Standard deviation.

ISSN: 2277-9655
Impact Factor: 4.116
CODEN: IJESS7

Table 10. Percentage Variance at Completion from Cost at Completion

PROJECT	Evaluation Stage	EAC1	EAC2	EAC3	EAC4	EAC5
PROJECT 1 231.28	Early	6.858	17.291	22.447	4.073	2.554
	Middle	4.201	3.313	9.104	3.439	0.365
	Late	1.490	0.766	7.247	1.914	5.261
PROJECT 2 480	Early	6.189	5.847	10.649	1.949	0.344
	Middle	3.861	2.201	8.882	2.052	0.754
	Late	1.107	0.977	6.579	3.436	6.274
PROJECT 3 287.5	Early	9.578	4.108	10.895	15.731	1.646
	Middle	8.100	4.814	15.821	8.686	1.788
	Late	4.312	1.308	19.562	6.284	0.178

The cost and time forecasting is best needed when it is in the initial growing stages and the new method has a stronghold in the early and middle stages where the past data are only a few. The MAPE of EAC5 in early stage is 1.5149 for the projects with cost overrun and schedule delay and for middle stage is 0.9689 which shows it is the best method to apply during this period. Table 12. shows the MAPE values for projects with both cost overrun and schedule delays. Among five methods compared other than EAC5 only EAC2 shows a better value that even has an MAPE equals 3.442 at middle stage and the early stage its MAPE is 9.08 which is not an acceptable value. But during later stage EAC2 shows more acceptable values with an MAPE of 1.016 whereas EAC5 has 3.904 as MAPE.

Table 11. MAPE for cost overrun and schedule delay projects

EVALUTION STAGE	EAC1	EAC2	EAC3	EAC4	EAC5
EARLY	7.542	6.343	14.664	7.251	1.285
MIDDLE	5.387	1.233	11.269	4.726	0.466
LATE	2.303	0.145	11.130	0.311	3.904

The figure .6. shows the difference in MAPE at different stages of evaluation for the projects with CPI and SPI less than unity. The EAC5 method has huge difference from other index based and regression method.
The test results shows that the index based method is not a strong method for forecasting the final cost and time of the project as it relies only on past performance. It finds the performance index for a particular time and it assumes that the future will hold the same performance for that particular project. But that assumption usually fails because as the project grows, the performance index changes. So it is necessary to consider future performance also. The advantage of nonlinear regression is that it takes the past performance data and merge it with future planned data and create an equation for the time-cost pattern. Using the equation it is easy to find any value on the curve and so, the cost at the forecasted final duration of the project is calculated and the average error on this method is very less. Hence this is the most adoptable method for the early and middle stage of the project.

ISSN: 2277-9655
Impact Factor: 4.116
CODEN: IJESS7

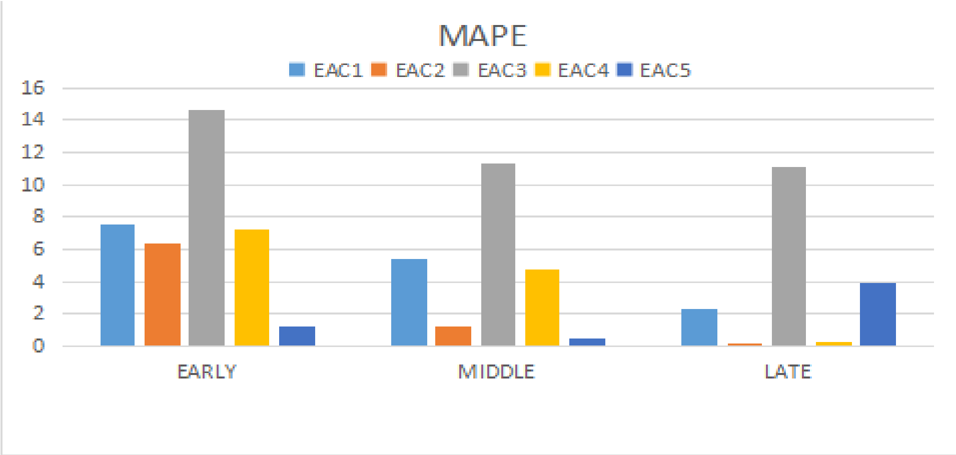

Figure .6. MAPE for three projects

The standard deviation of the projects with delay shown in Table 9. shows the precision of the method EAC5, ie., the modified nonlinear regression method. Overall the results shows that the EAC5 method is dominant over other index based formulae. The integration of cost variance and schedule progress into the model equation leads to improving the CEAC accuracy.

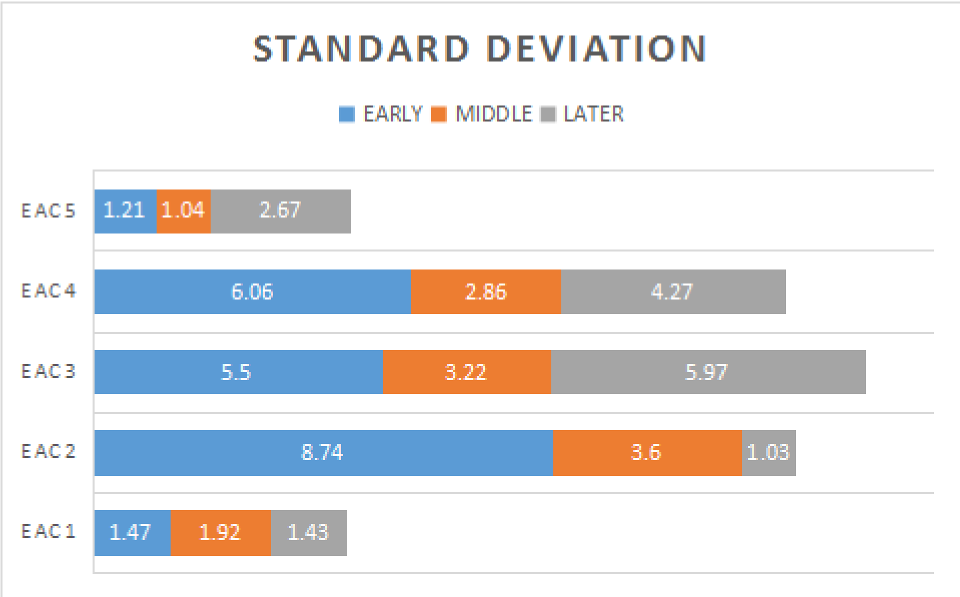

Figure.7. Standard deviation of percentage errors.

As shown in figure.7. the total standard deviation at three stages for EAC5 is very less as compared to other methods. The standard deviation at early and middle stage for EAC5 is very less and for late stage the CPI method gives more reliable values than regression method.
The cumulative value of CPI, which gets stabilized as the project progresses. It ensures for more stable values of CEAC by the end of a project. For a project in its early life, when few EVM data are at hand, this technique is unreliable as it makes extrapolations from few time points for the rest of the project: this is risky and provides inaccurate estimates. But when the project attains some maturity the stabilized CPI with more past performance to support its forecasting it gives more reliable data for the late stage cost forecast. In most projects regardless of their nature, budget, and duration, estimates by a traditional approach stabilize by the second half of the project life or at late stage. Previous studies showed that the PI values (CPI and SPI(t)) converged to their respective final values as the project gets closer to completion.

ISSN: 2277-9655
Impact Factor: 4.116
CODEN: IJESS7

One major finding of this study is that considering schedule progress as a factor of future cost improves both accuracy and precision of the developed model. EAC4 method doesn't include the earned schedule factor and the cost forecasted by this method is more erroneous than EAC5 which incorporates the earned schedule in CEAC calculation. EVM is a system that integrates project cost, schedule, and scope. In this regard, schedule is known as a factor of project cost performance. Delay in work progress has its influence on cost behaviour. If the project is ahead of schedule it follows the estimated planned value so there won't be any extra cost due to delay rather it sticks to the planned data. The majority of projects experience impact of schedule progress on their final cost. Therefore, EAC5 methodology makes explicit use of ES concepts in calculation of CEAC. This practical contribution of the ES method into the forecasting formula reflects schedule impact and, hence, provides more reliable CEAC.

V. CONCLUSION

Everyone is concerned about future especially project managers are very much worried about the future of the project. So the CEAC forecasting was always a subject for many researches and comparative studies. Accurate CEAC forecasting helps the project managers to take required decisions on time which helps the project to complete on time and within budget. The methodology used in this paper holds well in the early and middle stages of a project which has cost overrun and are running late. In this paper the flaws in EVM is studied and a better performance analysis is suggested which uses the nonlinear regression analysis. The regression modified method uses the earned schedule suggested by Lipke which helps in forecasting CEAC for delay in projects. The input for the nonlinear regression analysis is modified by adding cost variance to the planned cost and adding earned schedule to the baseline. The output of the nonlinear regression analysis is then used to find the CEAC of the project and the commonly used equation is modified by earned schedule which gives better result.

The accuracy of the different cost forecasting methods are compared and EAC5 method gives the most accurate results. Also the EAC5 method is more precise compared to other methods with very less standard deviation.

This paper suggests that the use of the regression modified method for the first half and CPI based index method in the second half of project gives more reliable information in cost forecasting. IB method produce inaccurate and unreliable data in the first half of the project duration but in second half the information available about the project is more and hence provide the best reliable CEAC forecast. The index based CEAC forecasting which started a few decades before was initially focused on large and complex projects but the method used in this paper is suitable for projects with any budget and duration.

APPENDIX - I

The building status of the projects at each months with PV, EV, and AC are shown in the following table.

PROJECT 1						
Time (Months)	Planned Value (Lakhs)		Earned Value (Lakhs)		Actual Cost (Lakhs)	
	periodic	cumulative	periodic	cumulative	periodic	cumulative
1	6.06	6.06	6.04	6.04	8.23	8.23
2	6.93	12.99	6.69	12.73	9.89	18.12
3	9.06	22.05	8.62	21.35	12.03	30.15
4	12.86	34.91	12.09	33.44	13.89	44.04
5	14.87	49.78	14.39	47.83	16.45	60.49
6	15.65	65.43	14.87	62.7	15.72	76.21
7	16.39	81.82	15.41	78.11	16.85	93.06
8	16.24	98.05	15.04	93.15	15.92	108.98
9	17.42	115.48	16.2	109.35	17.87	126.85
10	17.66	133.14	16.61	125.96	17.09	143.94
11	17.89	151.03	16.44	142.4	18.63	162.57

ISSN: 2277-9655
Impact Factor: 4.116
CODEN: IJESS7

12	17	168.03	15.98	158.38	17.25	179.82
13	16.44	184.47	15.13	173.51	16.94	196.76
14	14.25	198.72	13.2	186.71	14.49	211.25
15	12.77	211.49	12.03	198.74	13.55	224.8
16	9	220.49	8.73	207.47	9.28	234.08
17	6.74	227.23	6.6	214.07	7.34	241.42
18	4.05	231.28	3.96	218.03	4.22	245.64
19	0	231.28	3.28	221.31	3.52	249.16
20	0	231.28	2.38	223.69	2.46	251.62
21	0	231.28	3.6	227.29	3.83	255.62
22	0	231.28	3.99	231.28	4.24	259.69

PROJECT 2						
Time (Months)	Planned Value (Lakhs)		Earned Value (Lakhs)		Actual Cost (Lakhs)	
	periodic	cumulative	periodic	cumulative	periodic	cumulative
1	18.45	18.45	18.39	18.39	20.67	20.67
2	19.73	38.18	19.57	37.96	22.16	42.83
3	20.83	59.01	19.62	57.58	24.98	67.81
4	21.79	80.8	19.94	77.52	23.61	91.42
5	22.87	103.67	21.65	99.17	24.34	115.76
6	24.71	128.38	22.98	122.15	25.98	141.74
7	25.51	153.89	23.78	145.93	26.87	168.61
8	27.97	181.86	25.98	171.91	26.04	194.65
9	29.34	211.2	27.52	199.43	28.87	223.52
10	31.15	242.35	28.05	227.48	32.87	256.39
11	33.27	275.62	31.56	259.04	35.98	292.37
12	34.82	310.44	33.57	292.61	36.29	328.66
13	33.05	343.49	32.81	325.42	34.61	363.27
14	31.16	374.65	29.71	355.13	32.21	395.48
15	29.45	404.1	27.73	382.86	30.87	426.35
16	24.32	428.42	23.01	405.87	25.07	451.42
17	20.45	448.87	19.97	425.84	21.56	472.98
18	16.5	465.37	15.59	441.43	18.23	491.21
19	9.38	474.75	12.67	454.1	11.34	502.55
20	5.25	480	7.89	461.99	6.98	509.53
21	0	480	5.01	467	6.25	515.78
22	0	480	4.74	471.74	5.01	520.79
23	0	480	4.51	476.25	4.76	525.55
24	0	480	3.75	480	3.8	529.35

PROJECT 3						
Time (Months)	Planned Value (Lakhs)		Earned Value (Lakhs)		Actual Cost (Lakhs)	
	periodic	cumulative	periodic	cumulative	periodic	cumulative
1	37.5	37.5	32.5	32.5	34.4	34.4
2	15	52.5	10.2	42.7	10.8	45.2

ISSN: 2277-9655
Impact Factor: 4.116
CODEN: IJESS7

3	32.5	85	30.8	73.5	34.4	79.6
4	50.5	135.5	29	102.5	26	105.6
5	41.3	176.8	42.8	145.3	50.6	156.2
6	35.7	212.5	32.1	177.4	36	192.2
7	32.7	245.2	25	202.4	33.4	225.6
8	17.3	262.5	16.6	219	19.5	245.1
9	25	287.5	16.6	235.6	22.5	267.6
10	0	287.5	20.9	256.5	24.9	292.5
11	0	287.5	17	273.5	21.3	313.8
12	0	287.5	14	287.5	10.9	324.7

PROJECT 4						
Time (Months)	Planned Value (Lakhs)		Earned Value (Lakhs)		Actual Cost (Lakhs)	
	periodic	cumulative	periodic	cumulative	periodic	cumulative
1	28.975	28.975	25.645	25.645	25.567	25.567
2	52.706	81.681	42.429	68.074	40.726	66.293
3	10	91.681	21.061	89.135	12.000	78.293
4	46.905	138.586	36.109	135.244	45.78	124.073
5	79.555	218.141	73.51	198.754	67.294	191.367
6	84.337	302.478	70.009	268.763	68.478	259.845
7	21.154	323.632	23.7-6	292.469	25.767	285.612
8	22.244	345.876	14.2256	306.725	5.231	290.843
9	14.862	360.738	6.139	312.864	12.646	303.489
10	0	360.738	14.83	327.694	12.942	316.431
11	0	360.738	10.978	338.672	4.259	320.69
12	0	360.738	11.189	349.861	16.066	336.756
13	0	360.738	10.877	360.738	12.623	349.379

PROJECT 5						
Time (Months)	Planned Value (Lakhs)		Earned Value (Lakhs)		Actual Cost (Lakhs)	
	periodic	cumulative	periodic	cumulative	periodic	cumulative
1	34	34	36	36	35	35
2	53	87	57	93	60	95
3	70	157	76	169	79	174
4	216	373	233	402	238	412
5	176	549	195	597	211	623
6	124	673	138	735	131	754
7	125	798	104	839	120	874
8	44	842	48	887	58	932
9	34	876	19	906	20	952
10	30	906	0	0	0	952

VI. REFERENCES

[1] Abba W.F, "Earned value management- reconciling government and commercial practices" *Program Management*, pp 58 – 63, January (1997).

[2] Anbari F.T , "Earned value project management method and extensions".*Project management Management Journal*, 34(4). pp 12 – 23, (2003).

[3] Bates, Douglas M, and Watts D G, "Nonlinear regression analysis and its applications" John Wiley & Sons, New York, (1999).

[4] Bhosekar. S.K and Vyas. G, "Cost controlling using earned value analysis in construction industries". *IJEIT*., 1(4), pp324 – 332, (2012).

[5] Chandrasekaran, R. and Venkatesh Kumar, R, "Application of logistic regression to predict over target baseline".*International Journal of Computer Applications* (0975 – 8887), Volume 44– No10, April (2012).
[6] Christensen, D.S. "Using performance indices to evaluate the estimate at completion" *Journal of cost analysis and management*, pp 17 – 24, (1994)
[7] Christensen, David S. and Scott Heise, "cost performance index stability." *National Contract Management Journal 25* (Spring), pp. 7 – 15, (1993).
[8] Christensen, David S, Richard C. Antolini, and John McKinney, "A review of EAC research." *Cost Estimating and Analysis - Balancing Technology and Declining Budgets*. New York: Springer- Verlag (July), pp. 207 – 224, (1994).
[9] Cummings E.G, Sehneidar K.A, "Cost/schedule control systems criteria a reference guide information". Thesis, school of systems and logistics of the Air Force Institute of Technology, Air University.(1992)
[10] Henderson, K, "Earned schedule in action", *The Measurable News*., pp 23 – 30, (2005).
[11] Henderson, K., "Further developments in earned schedule."*The Measurable News*., 15 – 22, (2004).
[12] Lipke.W. and Henderson, K., "Earned schedule- an emerging enhancement to evm" *Projects & Profits*., pp 67 – 75, (2007).
[13] Lipke. W, "Schedule is different". *The Measurable News. pp* 31–34 ,Summer.(2003)
[14] Lipke. W., (2005). "A re-examination of project outcome prediction-using the earned value management methods" *The Measurable News*, pp 16-24, Summer (2005).
[15] Lipke. W, "Project duration forecasting. A comparison of earned value management methods to earned schedule." *The Measurable News*., Issue 2,pp 24 – 31, (2009).
[16] Lipke. W, "Schedule adherence: a useful measure for project management". *The Measurable News*., Issue 3, 1, pp 9 – 15.
[17] Lipke. W, "Earned schedule application to small projects*". PM World Today.*, XIII (I), (2012).
[18] Lipke. W, (2012). "Speculations on project duration forecasting". *The Measurable News*, Issue 3, 1, pp 4 – 7, (2012).
[19] Lipke, W., (2014). "Examining project duration forecasting reliability". *PM World Journal*., Volume III, Issue III , March(2014).
[20] Naderpour. A.and Mohd, M,"Improving construction management of an educational centre by applying earned value technique", *Procedia Engineering*, 14, pp 1945 – 1952,(2012).
[21] Narbaev.T, De Marco.A, "Combination of growth model and earned schedule to forecast project cost at completion". *Journal of Construction Engineering Management*. 140 (1). May (2013).
[22] PMI, 2013. "A guide to the project management body of knowledge, fifth edition". *Project Management Institute, Inc.*, Newtown Square, PA.
[23] PMI, 2014. "PMI - study on perspectives of Indian organizations on project management skills." *Project Management Institute*, Ernst and Young LLP.
[24] Tracy S. P, "Estimate at completion: A regression approach to earned value." M.S. thesis, Air Force Institute of Technology, Wright-Patterson Air Force Base, OH, (2005).
[25] Trahan E. N, "An evaluation of growth models as predictive tools for estimates at completion (Estimate At Completion)." M.S. thesis, Air Force Institute of Technology, Wright-Patterson Air Force Base, OH, (2009).
[26] Vandevoorde .S and Vanhoucke M, "A comparison of different project duration forecasting methods using earned value metrics" *IJPM*, 24(4),pp 289 – 302, (2006).

CITE AN ARTICLE

Pookot , N. (2017). A COMPARISON OF DIFFERENT COST FORECASTING METHODS USING EARNED VALUE METRICS. *INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY, 6*(6), 302-319. doi:10.5281/zenodo.809176