RESEARCH ARTICLE OPEN ACCESS

Polynomial Function and Fuzzy Inference for Evaluating the Project Performance under Uncertainty

A.S. Abdel Azeem, A.H. Ibrahim, and H.E.Hosny
Construction Engineering and Utilities Dept., Faculty of Engineering, Zagazig University, Egypt.

ABSTRACT
The objectives of this paper are two folds. The first one is to improve the time forecasting produced from the well known Earned Value Management (EVM), using the polynomial function. The time prediction observed from the polynomial model, which is compared against that observed from the most common method for time forecasting (critical path method), is a more accurate (mean absolute percentage of error is less than 2%) than that observed from the conventional deterministic forecasting methods (CDFMs). The second is to evaluate and forecast the overall project performance under uncertainty using the fuzzy inference. As the uncertainty is inherent in real life projects, the polynomial function and fuzzy inference model (PFFI) can assist the project managers, to estimate the future status of the project in a more robust and reliable way. Two examples are used to illustrate how the new method can be implemented in reality.
Keywords-Cost Control; Earned Value; Fuzzy; Forecasting; Polynomial; Probabilistic.

I. Introduction

Earned Value Management (EVM) is a project management technique that can be used to measure the project progress in an accurate manner. In a more precious definition by PMI (2005) [1], EVM combines measurements of technical performance (i.e. accomplishment of planned work), schedule performance (i.e. behind/ahead of schedule), and cost performance (i.e. under/over budget) within a single integrated methodology. In the EVM, the schedule and cost performance of a project are analyzed in terms of four performance indicators: (1) cost variance (CV) $=EV_{AD}-AC_{AD}$; (2) cost performance index (CPI) $=EV_{AD}/AC_{AD}$; (3) schedule variance (SV_{AD}) $=EV-PV$; and (4) schedule performance index (SPI_{AD}) $=EV_{AD}/PV_{AD}$. Where PV_{AD} is the planned value cost, EV_{AD} is the earned value cost, AC_{AD} is the actual cost for work performed, and AD is the actual data date. The estimated cost at completion ($ECAC_{AD}$) at actual data date AD is then equal to the cost already spent (AC) plus the adjusted cost for the remaining work $(BAC-EV_{AD})/CPI_{AD}$, i.e. $ECAC_{AD} = AC_{AD} + (BAC-EV_{AD})/CPI_{AD} = BAC/CPI_{AD}$, where BAC=budget at completion.

In EVM, the schedule performance index (SPI_{EV}) can be obtained as EV_{AD}/PV_{AD} and the schedule variance (SV_{AD}) can be obtained in monetary units as $EV_{AD}-PV_{AD}$ [2].

Using EVM as a tool for cost and schedule forecasting has been criticizing many researchers [2], [3], [4] & [5]. EVM forecasting methods are typically unreliable especially in the early stages of a project [3] & [4]. EVM formulas for cost or schedule forecasting are deterministic and do not provide any information about the range of possible outcomes and the probability of meeting the project objectives [5]. Moreover, its mathematical formulas forecast the final outcome of a project based on the assumption that the current status measurements are accurate and without any errors, which is unrealistic because there are inherent errors in both measuring time to report and measuring performance at the reporting times [6]. The schedule variance of the EVM does not also measure time but expressed it in a monetary unit [2].

Over the past years, many alternative ways to improve the schedule and cost forecasting performance of the earned value method have been studied. A set of S-curves generated from a network-based simulation as a forecasting tool and extended the concept to a probabilistic forecasting method by adjusting the parameters of probability distributions of future activities with performance indices of finished activities [7] & [8]. A probabilistic model on the basis of the beta distribution developed as a curve fitting technique and the Bayesian inference to forecast the estimated duration at completion and the probability of success [9]. Useful they do, the progress S-curves generated from the beta S-curve function are smooth, regular, and are not satisfactory to simulate the uneven nature inherent in the progress S-curves. Also, their model was developed mainly to forecast the duration at completion neglecting the uncertainty inherent in the actual measurements of the project progress. A fuzzy-based earned value model presented to forecast both the time and the cost estimates at completion under uncertainty [6]. But, their approach was developed on the basis of earned schedule method that produces larger errors for the time forecasting especially at the early stages of the project [5].

This study introduces a new integrated methodology on the basis of both the polynomial function, as an S-curve tool to improve the time forecasting generated from the earned value method, and the fuzzy inference to evaluate the overall performance of the project under the inherent uncertainty in the actual measurements. It is expected that the fuzzy inference is very useful in evaluating the overall performance of a project where uncertainty arises in real-life projects.

Measuring the earned value (EV) is one of the first stages in implementing the EV management. There are some different techniques to measure the EV; the fixed formula, weighted milestone, and percent complete techniques. Selecting the best approach during the project planning stage is based on the task duration and the number of measurement periods through its duration [1]. The fixed formula technique, where a fixed percentage of work performance is approved at the start of the work and the remaining percentage is approved at the completion of the work, is most effectively used on shorter duration tasks (less than two periods). While, the weighted milestone and percent complete techniques are more suitable for longer duration tasks (greater than two periods). The percent complete approach is the simplest and easiest technique for the project managers to measure of the percentage of the work performed (WP) for each activity, which can be calculated through dividing the quantity of work performed by the total quantity of the work. The percent complete technique is a more useful technique to estimate the earned value for each activity. The percent complete is the simplest and the most implemented technique for measuring the EV; however it has the disadvantage of using subjective judgments to describe the percent of the completed work [6].

The ultimate goals of this paper are to improve the schedule performance forecasting of the earned value management using the polynomial function and to evaluate the overall progress of the project under uncertainty using the fuzzy model. This paper presents a new method on the basis of polynomial function and fuzzy inference with the advantage, over the traditional EVM and its extensions, for developing and analyzing the time and the cost at completion when uncertainty arises. The level of uncertainty in measuring of the actual progress may influence decisions of project managers to control their projects. The ultimate goal of project performance forecasting is to provide decision makers with objective and refined forecasts in a timely manner. However, actual performance data, which are probably the most objective and reliable source of predictive performance information, are even limited and depends on people's judgments.

The next sections of this paper are organized as follows. Section 2 presents an introduction for the conventional deterministic methods. Section 3 explains briefly the polynomial function and its application to improve the schedule performance forecasting for the EV method. Section 3explains briefly the fuzzy theory and its application to the earned value management. Developing the new fuzzy based EV technique and its interpretations are covered in Section 4. For clarification purposes, two numerical examples illustrate how the new model can be implemented in reality are presented in details in section 5, the paper ends with the conclusion.

II. Conventional Deterministic Forecasting Methods

2.1 Earned Value Method

Earned value method (EVM) was originally developed for cost management and it has not widely been used for forecasting project duration [2]. In EVM, the schedule performance index (SPI_{EV}) can be obtained as EV_{AD}/PV_{AD} and the schedule variance (SV_{AD}) can be obtained in monetary units as EV_{AD}-PV_{AD}. Where PV_{AD} is the planned value cost, EV_{AD} is the earned value cost, and AD is the actual data date.

At the end of a project, the EV_{AD} = PV_{AD} = BAC (budget at completion), and hence, the SV_{AD} and SPI_{EV} always equals 0 and 1 respectively. If SV_{AD} = 0 and SPI_{EV} =1, the earned value is exactly as planned, regardless of the real project status (behind, on schedule or ahead) [10]. It means that SV_{AD} and SPI_{EV} cannot give appropriate information at the late stages of the project. The estimated duration at completion generated from the EVM can be calculated, on the basis of the two indicators: PV_{AD} and EV_{AD} , can be defined as [5]:

$$EDAC_{EV} = PD/SPI_{EV} \ldots\ldots\ldots\ldots\ldots\ldots.(1)$$

Where PD is the original estimate of the project duration.

Different alternatives of ways have been studied to improve the schedule performance forecasting of the earned value method; the earned schedule, planned value, and the earned duration methods [11], [12], [13] & [14]. The next section reviews these different techniques to improve the time forecasting of the EVM.

2.1.1 The Earned Schedule Method

The earned schedule method (ESM) was introduced as an extended EV metric to overcome the deficiencies of the EVM schedule indicator SPI [11]. In ESM, the term ES is the planned time to perform the EV and is resulted by projecting each value of EV on the baseline curve, as shown in Fig.

1. Then the linear interpolation between two successive planned values is generated. The term ES_{AD} is defined as [2]:

$$ES_{AD} = k + \frac{EV_{AD} - PV_k}{PV_{k+1} - PV_k} \dots\dots\dots.(2)$$

Where k is the longest time interval in which PV_k is less than EV_{AD}, PV_k is the planned value at time k, and PV_{k+1} is the planned value at the next time interval, i.e. time k+1. The general framework of the earned value and the earned schedule methods is shown in Fig. 1.

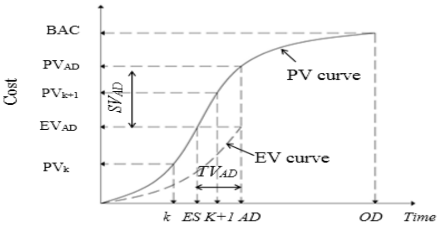


Figure 1: The PV and EV Metrics [10]

Thus, the schedule performance index (SPI_{ES}) can be obtained as ES_{AD}/AD and the time variation (TV_{AD}) can be obtained in time units as ES_{AD}-AD. The estimated duration at completion generated from the ESM can be calculated as defined as [11]:

$$EDAC_{ES} = PD/SPI_{ES} \dots\dots\dots\dots\dots\dots.(3)$$

2.1.2 The Planned Value Method

The planned value method (PVM) was introduced and assumes that the time variation can be translated into time units through dividing the schedule variance by the planned value rate (PVRate), i.e. PVRate = BAC/PD [12]:

$$EDAC_{PV} = PD/BAC^*(BAC - SV_{AD}) \dots.(4)$$

Where BAC is the budget cost at completion and AD is the actual data date.

2.1.3 The Earned Duration Method

The earned duration method (EDM) can be calculated as the product of the actual duration AD and the schedule performance index SPI, i.e. ED = AD * SPI_{EV}. The estimated duration at completion generated from the method is [14]:

$$EDAC_{ED} = PD + AD\ (1 - SPI_{EV}) \dots\dots.(5)$$

Where AD is the actual data date.

It should be concluded that the last two techniques (PVM and EDM) use the two efficiency indicators of EVM, SPI and SV, for making the time forecasting. But, at the end of the project, the SPI_{EV} tends to be 1 and SV_{EV} tends to be 0, thus, these two techniques are not reliable at the end of the project. Therefore, we modify the ES technique to develop the polynomial-based time estimate at completion.

III. The General Framework of the Polynomial Model

The polynomial function has a long history of application in engineering and project management [15], [16] & [17]. The polynomial forecasting model is mainly developed on the basis of the nonlinear regression techniques, fits an S-curve function to the cumulative progress S-curve of a project, and is used to provide better forecasting for the expected completion dates and their confidence bounds. The underlying strategy of the polynomial forecasting model is to simulate the progress curve of a project and to use that curve to forecast future progress of the project.

The general framework of the polynomial forecasting method is shown in Figs. 2a and 2b. A library of polynomial S-curves were generated and then evaluated using the least square method to select the best polynomial S-curve, which is then used to estimate the planned time for achieving a certain work performed $E(T|W=WP_j)$ through updating its coefficients vector (C) in the light of new actual performance data. In this case, the time variation (TV_j) can be measured in time dimension as the difference between the planned time $E(T|W=WP_j)$ and actual data date AD_j. Consequently, the time prediction generated from the polynomial function can be calculated basically from adding the time variation TV_j to the planned project duration (PD).

The dash line, as shown in Figs. 2a and 2b, represents the extension of the earned value curve of on-going project after the actual data date AD. This extension will follow the best polynomial S-curve, which fits the uneven nature of the progress S-curve, to ensure that the remaining activities will be executed according to the detailed plans of the project.

The cumulative normal function is developed on the basis of historical data and subjective judgments and is used to assess the probability of meeting the project duration. The polynomial forecasting model consists of four steps: (1): Creating a library of polynomial S-curves, (2): Selecting the best polynomial S-curve, (3): Updating the coefficients vector of the best polynomial S-curve, (4): Forecasting the project duration.

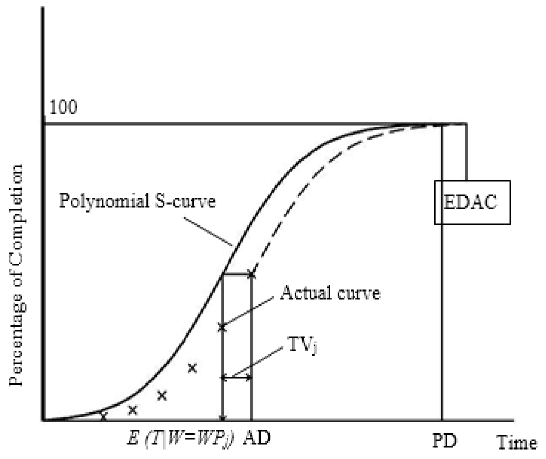


Figure 2a: Case of Behind Schedule

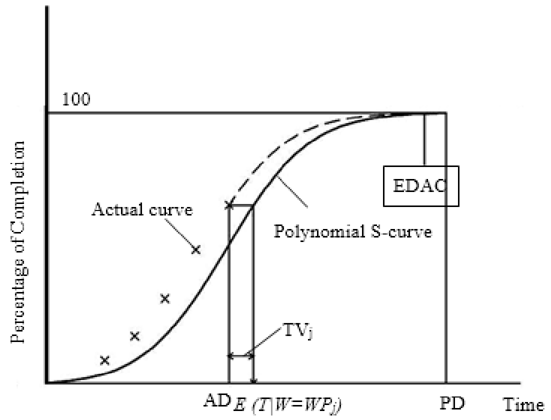


Figure 2b: Case of Ahead Schedule

3.1 Creating a Library of Polynomial S-curves

The library of polynomial S-curves can be constructed on the basis of various degrees of polynomial functions. The degree of a polynomial function (D) plays an important role in simulating the irregular nature inherited in the progress S-curve of a project. For each degree of a polynomial function, the goodness of fit for the progress S-curve of the project can be quantified in a systematic way using the least square method. In this study, ten polynomial S-curves were evaluated to select a model that can simulate the uneven and irregular nature inherited in the progress S-curve. It should be noted that the degree of a polynomial function D should be less than the number of planned measurement points of the progress S-curve (N) [18], i.e. D < N. The polynomial S-curve of a project can be represented over a range of planned times values ti, is defined as [18]:

$$E(ws_i \backslash T = t_i) = \sum_{d}^{D} C_d * t_i^{D-d} \dots \dots \dots .(6)$$

i = 1,2,3 …….N
d= 0,1,2…......D

Where ws_i is the planned percentages of completion at a measurement point i and D is the degree of a polynomial function. For example, if the degree of a polynomial function D is 3, the polynomial function can therefore be represented as:

$$\boldsymbol{E(WS_i \backslash T = t_i) = C_0 * t_i^3 + C_1 * t_i^2 + C_2 * t_i + C_3 \dots .(7)}$$

The coefficients vector C, with a length (D+1)for a polynomial degree D, can fit the progress S-curve of the project by minimizing the sum of the squares of the deviations (least-square method). The coefficients vector C of a polynomial function is measured basically from two vectors, planned times and their corresponding schedule complete percentages, and is ordered in descending powers (from the highest to the lowest degree), is defined as:

$$C = \begin{bmatrix} C_0 \\ C_1 \\ C_2 \\ . \\ . \\ . \\ C_D \end{bmatrix} \dots \dots \dots ..(8)$$

Where C_o is the coefficient of the highest power of planned time and C_D is the constant term. Each polynomial S-curve that is represented by its coefficients C is evaluated to obtain the best polynomial S-curve, which will be updated in the light of new actual data.

3.2 Selecting the Best Polynomial S-curve

The ultimate goal is to select a polynomial S-curve that can acceptably simulate the uneven nature inherited in the progress S-curve of a project. A simple way of evaluating and comparing the fitness of different polynomial S-curves to the progress S-curve of a project is the least square method. The deviations are defined as the differences between the progress S-curve of a project (ws_i) and the generated polynomial S-curve $E(ws_i|T= t_i)$. The least square method is used to minimize these deviations for each generated polynomial S-curve, is defined as:

$$\boldsymbol{\sum_{i=1}^{N} E\{[ws_i - E(ws_i \backslash T = t_i]^2\} = min \dots \dots .(9)}$$

Two indicators are used to measure the accuracy or strength of the generated polynomial S-curves in terms of closeness of fit and to provide a basis for the model performance evaluation: coefficient of determination (R^2) and root-mean-square error (RMSE), and are defined as:

$$R^2 = \frac{\sum_{i=1}^{N}\{[E(ws_i \backslash T = t_i) - \overline{ws}]^2\}}{\sum_{i=1}^{N}[(ws_i - \overline{ws})^2]} \dots (10)$$

$$\boldsymbol{RMSE} = \sqrt{\frac{\sum_{i=1}^{N}\{[\boldsymbol{ws_i} - \boldsymbol{E}(\boldsymbol{ws_i} \backslash \boldsymbol{T} = \boldsymbol{t_i})]^2\}}{\boldsymbol{N}}} .. (\mathbf{11})$$

Where ws bar is the mean of schedule complete percentages. Both Equs. (10) And (11) are consistent accuracy indicators for selecting the best polynomial S-curve. Table 1 indicates the best-fit coefficients vector C of the best polynomial S-curve model, R^2, and RMSE obtained for the projects.

Based on the R^2 and RMSE results that are indicated in Table 1, the tenth-degree of a polynomial function can efficiently visualize the irregular nature of progress S-curve of all projects and can produce a better accuracy of schedule performance forecasting for on-going projects. It should be noted that the constant term C_D (indicated in Table 1) will play an important role during updating the coefficients vector of the best polynomial S-curve to estimate the planned time to achieve the work performed.

3.3 Updating the Coefficients Vector of the Best Polynomial S-curve

The ultimate goal is to solicit the planned information; which are the original project duration, budget cost at completion, and the planned time to achieve the work performed; from the polynomial S-curve and to update the baseline plan in the light of new actual performance data. The coefficients vector of the best polynomial S-curve, which has the highest R^2 and lowest RMSE, will be updated in the light of new actual performance data and used later to forecast the estimated duration at completion of the project. Once a project gets started, the actual percentage of work performed WP_j is reported periodically and its actual data AD_jcan be represented as a series of discrete values

$$W: (WP_j, AD_j), j = 1,2 \ldots M \ldots \ldots \ldots . (12)$$

Where M is the number of actual data dates points and WP_j is the percentage of work performed that is measured at the actual data date AD_j. The ultimate goal is to find the planned time for achieving a specified work performed $E(T|W=WP_j)$, which can be obtained by projecting each percentage of work performed (WP_j) on the best polynomial S-curve, as shown in Figs. 2

$$\sum_{d}^{D} C_d * t_i^{D-d} = WP_j \ldots \ldots \ldots .. (13)$$

Then, the coefficients vector C of the best polynomial S-curve can be updated in the light of new actual performance data, can be shown as

$$\sum_{d}^{D} C_d * t_i^{D-d} - WP_j = e_i \ldots \ldots \ldots .. (14)$$

i=1,2,3……10,0000
j= 1,2,3………… M

Where e_i is the error term corresponding to the planned time t_i. The left hand side of Equ. (15) Refers to the updated coefficients vector C_{adj} and can be obtained by subtracting the vector of work performed WP_j from the vector of the best-fit coefficients C, can be expressed as:

$$C_{adj.} = \begin{bmatrix} C_0 \\ C_1 \\ C_2 \\ . \\ . \\ . \\ C_D - WP_i \end{bmatrix} = \begin{bmatrix} C_0 \\ C_1 \\ C_2 \\ . \\ . \\ . \\ C_D \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ 0 \\ . \\ . \\ . \\ WP_i \end{bmatrix} \ldots (15)$$

It should be noted that the updated coefficients vector C_{adj} of the best polynomial S-curve changes periodically in the light of new reporting data. The updated coefficients vector C_{adj} is used to estimate the planned time for each work performed $E(T|W=WP_j)$ through minimizing the error term e_i in Equ. (15).

Table 1: The Coefficients of the Tenth-Degree of A Polynomial Function

Coefficients	Project A	Project B
Co	-4.5E-13	1.15E-15
C1	1.22E-10	-5.1E-13
C2	-1.4E-08	9.45E-11
C3	8.89E-07	-9.5E-09
C4	-3.5E-05	5.8E-07
C5	0.000886	-2.3E-05
C6	-0.01496	0.000577
C7	0.160974	-0.00874
C8	-0.81804	0.068205
C9	1.946997	0.261061
CD	0.001717	0.016841
R^2	0.998232	0.999163
RMSE	0.238764	0.445291

To estimate the planned time for each work performed $E(T|W=WP_j)$ and minimize the error term e_i, a column vector *(t)* of planned times should be assumed. This column vector *t* starts from a value PD/10,000 with a small value of increment PD/10,000 and ends at the PD value. It should be noted that the smaller values of the increment step is preferred to minimize the error term e_i. Then, the column vector of the error term *(e)*can then be obtained from substituting by the corresponding vector of planned times in Equ. (15). consequently, both the planned times and errors vectors are combined together in one matrix (referred as te) with size (10,000*2), then this matrix takes an ascending order according to the absolute error column e. In this case, the conditional expectation of T at given $W=WP_j$ and its corresponding error e_j can be expressed as:

$$E\left(T\middle|W = WP_j\right) = te(1,1) \ldots\ldots(16)$$
$$e_j = te(1,2) \ldots\ldots\ldots\ldots\ldots(17)$$

It should be noted that the error term e_i in Equ. (14) Was substituted by $e_j$in Equ. (17) Because the lowest error e_j is always corresponding to the term $E(T|W=WP_j)$. A graphical display of the error profile of the term e_j for the two projects A and B showed that the error term e_j appeared randomly scattered around the line of zero error (max value of error is 0.0006), indicating that a higher accuracy could be obtained during the calculations of the planned time for each work performed $E(T|W=WP_j)$.

3.4 Measuring the Time Variation in Time Units

The polynomial forecasting method is developed mainly on the basis of time variation in the time dimension between the PV and EV curves. The time variation TV_j can be measured, in time unit rather in monetary one, as the difference between the planned times $E(T|W=WP_j)$ and actual data dates AD_j, as shown below

$$TV_j = AD_j - E\left(T\middle|W = WP_j\right) \ldots\ldots(18)$$

IV. Numerical Examples

The PFFI model formulated in the previous sections has been programmed in a graphical user interface (GUI) [18]. The GUI of Matlab is to make the programming of PFFI easier for the interested users and project managers. The required inputs to run the PFFI are the PV_{AD}, EV_{AD}, AC_{AD}, and AD. The possible outputs of the probabilistic PFFI are the $EDAC_j$ and the $ECAC_j$ of a project at each data date, on the other hand, the possible outputs of the deterministic methods CDFMs are the expected project duration and estimated cost at completion. Moreover, the PFFI provides prediction bounds for EDAC, four point estimate. The CDFMs, on the other hand, don't provide prediction bounds of EDAC forecasting. Two artificial projects (A and B) are presented in Table 2 to overcome limited and often incomplete real project data and to validate the results observed from all the models.

Table 2: The Data of the Two Artificial Projects

Project	BAC (EGP.)	PD (weeks)	Number of activities	Number of actual observations
A	2,040,000	56	20	14
B	2,750,000	100	13	20

The planned value cost PV_{AD} at each data date can be directly computed from the baseline Gantt chart. Hence, the percentage of schedule complete (ws) at each data date can be determined through dividing the planned value cost PV_{AD} by the budget at completion cost (BAC), i.e. ws = PV_{AD}/ BAC. Once, a project gets started, and then each percentage of work performed WP for each activity can be measured from the construction site. Then, the earned value cost for each activity can be calculated on the basis of the percent complete technique as the product of the percentage of work performed WP and the budget cost for each activity, i.e. EV_{AD} = WP * budget cost for each activity. The cumulative earned value cost EV_{AD} is therefore calculated by summation of the earned value cost for each activity. The estimated duration at completion of CPM can be directly calculated through updating process for the durations and dates of the remaining activities, based on the actual reporting data. However, the time prediction produced from the polynomial function and the other CDFMs will be assessed and compared from the start to the end of the project to that produced from the CPM.

V. Comparative Study

The time prediction produced from the best polynomial model and the other CDFMs will be assessed and compared to the forecasted duration by CPM that makes the time forecast at the activity level through updating process. Then, the accuracies of the all models are assessed and compared through the different project periods of work performed WP_j. The absolute percentage of error (APE_j) and mean absolute percentage of error (MAPE) are then determined as:

$$APE_j = \frac{100}{EDAC_{cpm}} * \left|EDAC_j - EDAC_{cpm}\right| \ldots\ldots(19)$$

$$MAPE = \frac{100}{M} * \sum_{j=1}^{M} APE_j \ldots\ldots\ldots(20)$$

Where $EDAC_j$ is the estimated duration at completion produced from a specified model and $EDAC_{cpm}$ is the estimated duration at completion produced from the CPM.

Table 3 represents a comparative study for the models and provides the values of the absolute percentage of errors at different periods of actual completion for the two projects. The columns (APE_{EV}; $APE_{ES;}$ $APE_{PV;}$ APE_{ED}) indicated the APE results produced from CDFMs for the two projects, while column (APE_{PF}) indicated the APE profile produced from the PFM. The absolute percentage of error APE_j of the all models can be measured at different evaluation periods. These evaluation periods are estimated to 0-10% BAC, 10-20% BAC, 20-30% BAC, 30-40% BAC, 40-50% BAC, 50-60% BAC, 60-70%BAC, 70-80% BAC, 80-90% BAC, and 90-100% BAC. If the evaluation period has more than one value of APE_j, the average value can then be calculated. For example, the evaluation period 30-40% BAC for the Project A has two values of APE_j produced from the PFM, 0.19% and 0.56%, and then the average value can be measured as:

(0.19% + 0.56%)/2=0.38%, as indicated in Table 3. If the APE_j value produced from a specified model is greater than 5% (allowable error), the cell will therefore appear in yellow color to demonstrate more erratic responses of time forecasting. In other words, the forecast of a specified model can be accepted if its APE_j is less than 5%.

Table 3: APE of the All Models for the Two Projects [19]

Project	Error	0-10%	10-20%	20-30%	30-40%	40-50%	50-60%	60-70%	70-80%	80-90%	90-100%
A	APE_{PF}	1.10%	3.37%	0.11%	0.02%	------	2.66%	0.40%	1.41%	0.33%	0.80%
	APE_{EV}	126%	184.44%	80.65%	44.52%	------	22.30%	6.22%	0.56%	3.88%	10.32%
	APE_{ES}	67.3%	53.73%	29.63%	20.43%	------	16.15%	9.76%	6.19%	8.04%	3.25%
	APE_{PV}	2.25%	13.01%	11.57%	11.57%	------	9.80%	1.39%	2.83%	5.23%	10.48%
	APE_{ED}	3.47%	11.25%	6.45%	4.84%	------	4.39%	1.95%	5.05%	6.69%	10.59%
B	APE_{PF}	0.28%	0.71%	1.35%	3.10%	2.86%	0.45%	0.23%	1.02%	0.28%	0.71%
	APE_{EV}	29.4%	27.79%	40.27%	30.29%	26.77%	27.85%	15.33%	9.86%	2.19%	9.77%
	APE_{ES}	28.1%	24.95%	18.74%	12.81%	7.86%	7.54%	4.78%	6.98%	5.06%	1.60%
	APE_{PV}	1.76%	3.95%	3.16%	4.48%	6.20%	10.36%	5.78%	4.44%	1.41%	10.23%
	APE_{ED}	0.28%	1.10%	14.84%	11.62%	11.44%	15.01%	5.77%	2.60%	2.80%	11.94%

Based on such comparative study presented in Table 3, the two projects reveal that the polynomial forecasting method is the only method which showed satisfying and reliable results during the whole project duration. The limited sample and short lifecycles cannot make universal conclusion for the conventional deterministic methods. But the results of these two projects confirms the previous ones [19]. Consequently, the results produced from the earned value, the planned value, and the earned duration methods are unreliable at the end of the project. Instead, the earned value and earned schedule methods seem to provide unreliable forecasts at the first-half of actual completion. But, the advantage of the earned schedule over the earned value, the planned value, and the earned duration methods is that it provides valid and reliable results at the end of the project [19].Another advantage for the earned schedule is apparently indicated in Table 3 and it produces less erratic responses for the time forecasting than those produced from the earned value method during the early stages of the projects. On the other hand, the advantage of the planned value and the earned duration methods over the earned value and the earned schedule is that they provides valid and reliable results at the early stages of the projects. All CDFMS, however, provides no information about the range of possible outcomes and the probability of meeting the project objectives.

VI. Measuring the Actual Percent of Completion Using Fuzzy Theory

The actual reporting data regarding the actual measurements come from people's judgments; hence they carry some degree of uncertainty. Considering this uncertainty into interpretations and calculations, not only helps in measuring better the performance and the progress of a project, but also in extending the applicability of the EV techniques under the real-life and uncertain conditions. In this paper, the fuzzy theory is applied as an efficient tool to deal with the uncertainly in real case projects to quantify the vagueness of actual data in reality [20]. The fuzzy variables are obtained as normal distribution fuzzy numbers due to ease in representation and calculation. In the proposed model in this study, the actual percent of completion of projects are considered fuzzy parameters described as fuzzy normal distribution numbers.

Assume that the measurement of the completion percent for a project includes uncertainty. Therefore, the project expert should transform this uncertainty into a fuzzy number by assigning a membership function to take in consideration the uncertainty inherent in the measurement of the actual completion, to express the fuzzy number (like the one showed in the Fig. 3). In this figure the horizontal axis refers to the actual percent). The general framework for the proposed model is clarified as shown in Fig. 3.

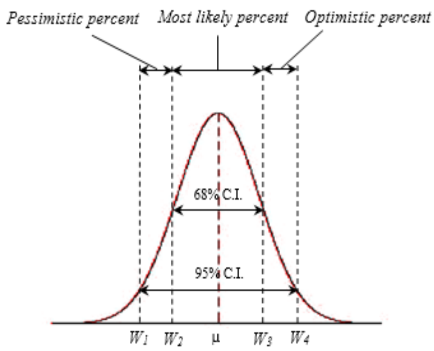


Figure 3: The Membership Function of a Fuzzy Number

In fact, the membership of fuzzy number helps to estimate the activity progress easier by providing better vision for the uncertain nature of the project. It is reasonable to model and treat the uncertainty using the confidence intervals (C.I) with the fuzzy theory. For example, if a project progress cannot be stated in certainty, using the confidence intervals 68 and 95%, it may be calculated as

$$WP_k = WP_j - 1.96 * e_j \dots\dots\dots\dots (21)$$

j= 1,2,3………M
k= 1,2,3,and 4

Where WP_j is the actual percent of completion measured at actual data date AD_j. The value of measurement error(e_j)is determined by the project manager to adjust the sensitivity of predictions and can be calculated statistically using the different people's judgments at the same the data date. Typically, the project experts perform this transformation in accordance with their knowledge and their experience about the project and according to the project attributes. Then we should modify the EV mathematics to consider fuzzy numbers, as defined next

$$EV_k = WP_k * BAC \dots\dots\dots (22)$$

Then, the generic equations to estimate the cost at the completion of a project are defined next

$$ECAC_k = AC_{AD} + (BAC - EV_k)/CPF \dots (23)$$

Where the BAC is the planned duration to complete the project, EV_1, EV_2, EV_3, EV_4 is a membership of a fuzzy number EV, the AC is the actual cost at each data date, and CPF is the cost performance factor which depends on the project status. If the CPF equals 1 or (CPI), then the remaining activities will be done according to the baseline plan or (according to the past performance). The generic equations to estimate the time at the completion of a project, or the project duration are defined next

$$EDAC_k = AD + (PD - E(T|W = WP_k)/SPF \dots (24)$$

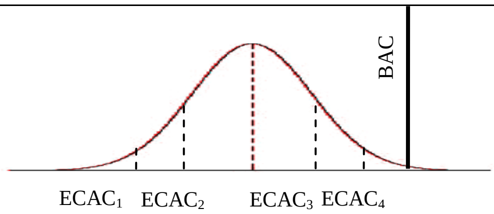
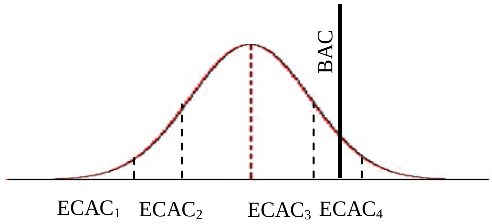
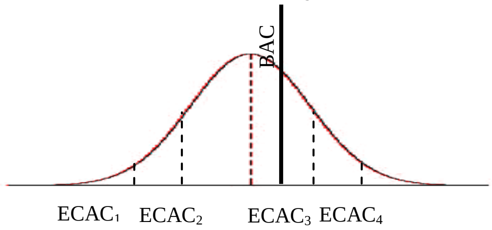
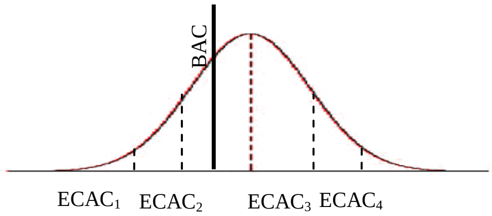
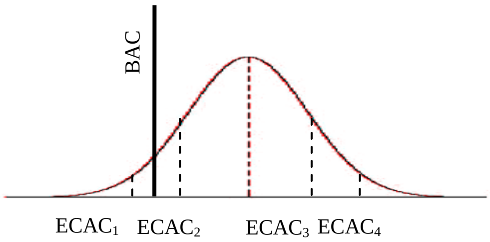
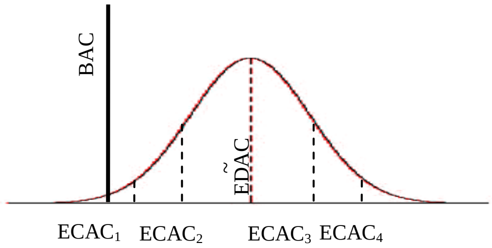
Where the PD is the planned duration to complete the project and CPF is the cost performance factor which depends on the project status.

The term, $E(T|W=WP_k)$, is the planned time for each work performed and is computed corresponding at the earned value EV_1. If the SPF equals 1 or (SPI), then the remaining activities will be done according to the baseline plan or (according to the past performance).

VII. Interpretation of Fuzzy Estimates

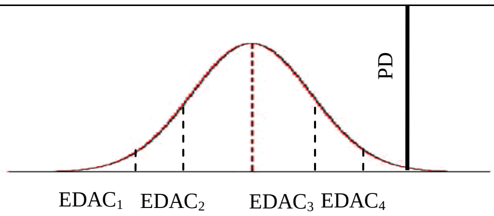
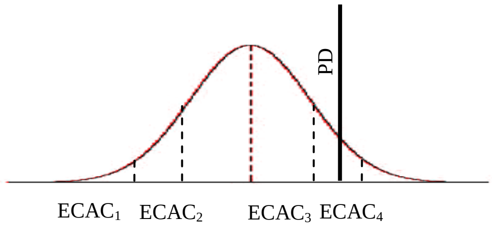
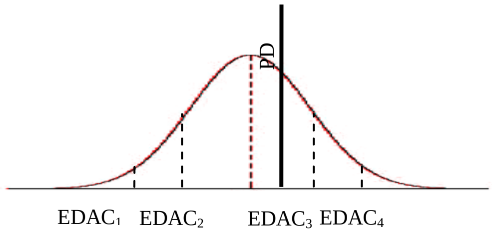
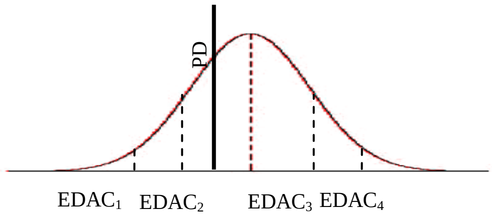
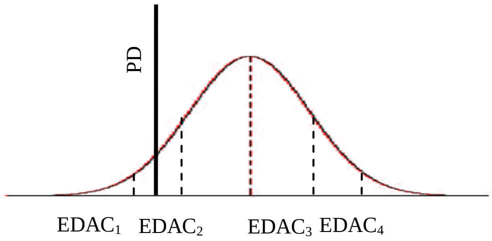
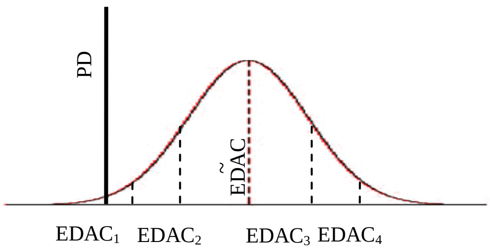
After developing the polynomial and fuzzy-based EV model, we should interpret both the expected cost and time at completion (ECAC and EDAC) to have an inference regarding the project progress and its status at a specified probability level. Similar to the traditional EV, the comparison is made against the value BAC and PD, the same story goes here, i.e. the new fuzzy numbers ECAC and EDAC must be compared against values BAC and PD. But as the new forecasting values of ECAC and EDAC are fuzzy numbers, we should compare them using the proposed method for evaluating the project progress probabilistically at each data date. The proposed method leads to probabilistic not deterministic conclusions for the expected project duration and cost as identified in Tables 4 and 5 respectively, which might be of interest for the project managers to manage their project under uncertainty.

Table 4: Interpretation of Fuzzy Estimates and Evaluation The Cost Status

No. of Scenario	State of ECAC	Graphical Representation	Decision Making
1	$ECAC_4 \leq BAC$		Under-run the budget with probability level greater than 95%
2	$ECAC_3 \leq BAC < ECAC_4$		Under-run the budget with probability level in between 68-95%
3	$ECAC_m{}^a \leq BAC < ECAC_3$		On the budget with probability level in between 50-68%
4	$ECAC_2 \leq BAC < ECAC_m{}^a$		On the budget with probability level in between 32-50%
5	$ECAC_1 \leq BAC < ECAC_2$		Behind the budget with probability level in between 95-68%
6	$ECAC_1 \leq BAC$		Behind the budget with probability level greater than 95%

[a]*$ECAC_m$ is the mean of the estimated cost at completion*

Table 5: Interpretation of Fuzzy Estimates and Evaluation The time Status

No. of Scenario	State of ECAC	Graphical Representation	Decision Making
1	$EDAC_4 \leq PD$		Ahead schedule with probability level greater than 95%
2	$EDAC_3 \leq PD < EDAC_4$		Ahead schedule with probability level in between 68-95%
3	$EDAC_m{}^a \leq PD < EDAC_3$		On the schedule with probability level in between 50-68%
4	$EDAC_2 \leq PD < EDAC_m{}^a$		On the schedule with probability level in between 32-50%
5	$EDAC_1 \leq PD < EDAC_2$		Behind schedule with probability level in between 68-95%
6	$EDAC_1 \leq PD$		Behind the schedule with probability level greater than 95%

[a]*$EDAC_m$ is the mean of the estimated duration at completion*

VIII. Evaluating the Overall Performance of the Two Projects

The ECAC and EDAC profiles can be collectively used to judge and forecast the cost and time performance through identifying the project status and time forecast for the projects at each actual data date. The next sections review the status of each project and the cost and time forecasts for the two projects.

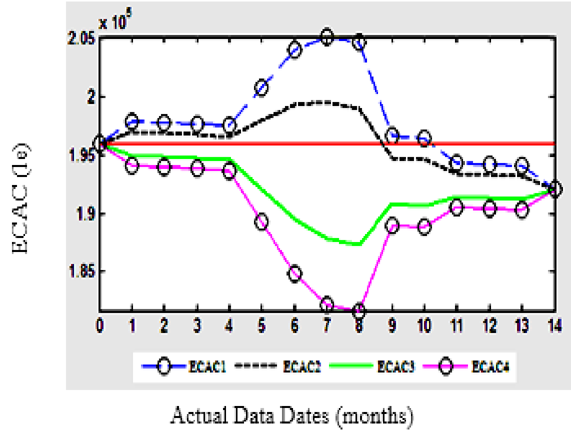


Figure 4a: Cost Profile for the Project A

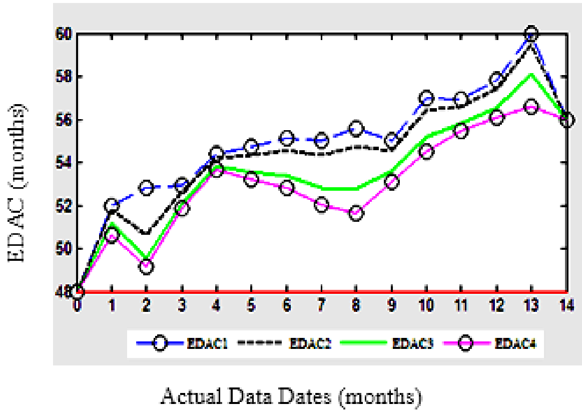


Figure 4b: Time Profile for the Project A

Project A

The status of the Project A regarding the cost profile, as shown in Fig. 4a, has been on the budget until the end of actual data date no. 9th , thereafter, the project status over-run the budget with probability level within 68-95% in-between the actual data dates no. 10th and 11th. The project status at the end of the Project A is under-run the budget with probability level greater than 95%. The status of the Project A regarding the time profile, as shown in Fig. 4b, has been behind the schedule with probability level greater or equal than 95% along the whole project duration.

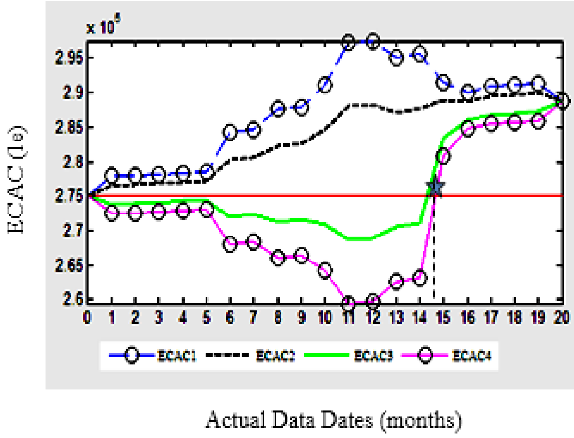


Figure 5a: Cost Profile for the Project B

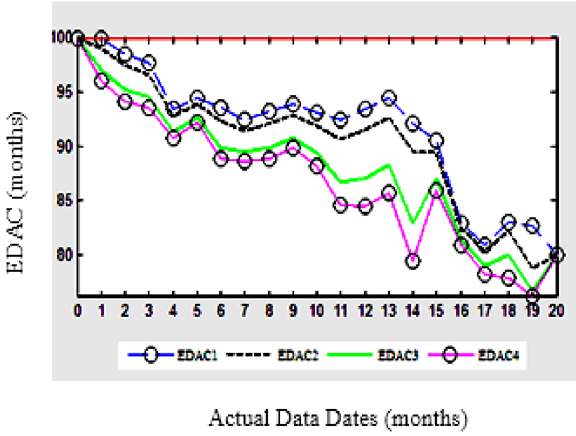


Figure 5b: Time Profile for the Project B

Project B

The status of the Project B regarding the cost profile, as shown in Fig. 5a, has been on the budget with probability level within 32-68% by the end of reporting period no. 14th. Thereafter the project status tended to be over-run the budget at probability level greater than 95%. An early warning of over-run the budget for the Project B should have been issued at a 5% risk level by the reporting period no. 15th.

The status of the Project B regarding the time profile, as shown in Fig. 5b, has been ahead schedule with probability level greater or equal than 95% along the whole project duration.

The cost and time profiles can be collectively used to judge the cost and schedule performance through identifying the current status and forecast for the projects at each actual data date. These charts, used separately or in combination, can serve as cost and time-saving project management dashboard, which is ideal for real-time monitoring of the overall cost and schedule risks from various perspectives, for example: (1) What is the probability level of finishing the project on the scheduled time and original budget? (2) Is the overall cost and time performance improving or deteriorating? Most of all, project managers using the PFFI will be able to identify when projects are in control and when attention is needed long before the project has deteriorated in a quantitative way and, literally, as shown in Figs. 4 and 5.

IX. Conclusions and Recommendations

The objective of this paper is enhance the capability of project managers for making informed decisions through introducing a new method, which is developed on the basis of the polynomial function as a curve fitting technique and fuzzy inference as a decision making tool for the uncertain conditions. The main role for the polynomial function is to improve schedule performance forecasting of the earned value method through fitting the uneven nature inherent in the progress S-curve of a project and updating the coefficient vector of the best polynomial S-curve in the light of new actual performance data. The main role for the fuzzy inference, on the other hand, is to evaluate the overall progress of a project when uncertainty arises. The fuzzy estimates of both the completion cost and the completion time can assist project managers to estimate the future status of the project in a more robust and reliable way.

The time prediction generated from the polynomial function and conventional forecasting method were compared, and assessed, against that generated from the CPM that is considered the most reliable forecasting method. Then, the absolute percentage of error APE of all the models can be presented at the different periods of actual completion. Such a comparative study reveals that the polynomial forecasting method is the only model that shows reliable results during the different periods of actual completion and predicts the expected completion dates with smaller errors than those generated from the conventional methods. Moreover, the PFFI generates the time prediction at the summary of the project-level and does not rely on the activity-level performance data and analysis as the CPM does.

The PFFI was applied to two projects, which are used to demonstrate how effectively the proposed model can be used to assist project managers to estimate the future status of the project in a more robust and a more reliable way, improve upon deterministic forecasting methods by adaptively developing the forecasts of project duration, by developing quantitative intervals around these forecasts, and by providing project managers with early warnings of possible project overruns. That is, probabilistic outputs from the PFFI are visualized with profile. Used separately or in combination, these charts can serve as a time and cost saving project management dashboard, which is ideal for real-time and cost monitoring of the overall risk from various perspectives, for example: (1) What is the probability level of finishing the project on a baseline time and cost? (2) Is the overall schedule and cost performance improving or deteriorating? Most of all, project managers using the PFFI will be able to identify when projects are in control and when attention is needed long before the project has deteriorated in a quantitative way and, literally, as shown in Figs. 4 to 8. The bounds for detecting such an early warning point can be determined by the project management according to a predetermined acceptable risk level for the project.

The proposed method PFFI has been programmed in a graphical user interface (GUI) for [18] and it can be applied to all kinds of projects. The software of the PFFI program is available on request by the authors. The limitation of the PFFI is that it is applicable when the uncertainty arises in the actual measurements, not the uncertainty arises in the project planning. Due to space limitations, this research cannot be reported here but will be covered in a subsequent publication. However, the overall uncertainties inherent in the actual measurements and project planning can be integrated within a consistent methodology to assist project managers to estimate the future status of the project in a more robust and a more reliable way**.**

Declaration:

This paper is based on a PHD Thesis [19], prepared by the first author, and supervised by the second two authors.

References

[1] PMI. (2005). "*Practice Standard for Earned Value Management*."Proj. Manage. Ins., Newtown Square, Pa.

[2] Vandevoorde, S. and Vanhoucke, M. (2006). "*A Comparison of Different Project Duration Forecasting Methods Using Earned Value Metrics*." Int. J. Proj. Manage. 24(4), 289–302.

[3] Christensen, D. S., and Heise, S. R. (1993). "*Cost Performance Index Stability*." National Contract Manage. J., 25(1), 7–15.

[4] Zwikael, O.; Globerson, S.; and Raz, T. (2000). "Evaluation of Models for Forecasting the Final Cost of a Project" Proj. Manage. J., 31(1), 53–57.

[5] Kim, B. C.; and Reinschmidt, K. F. (2010). "*Probabilistic Forecasting of Project Duration Using Kalman Filter and the Earned Value Method*." J. Constr. Eng. Manage., 130(1), 25–32.

[6] Naeni, L. M.; Shadrokh, S.; and Salehipour, A. (2011). *"A Fuzzy Approach for the Earned Value Management."* International J. of Proj. Manage. 29, 764–772.

[7] Barraza, G. A.; Back, W. E.; and Mata, F. (2000). *"Probabilistic Monitoring of Project Performance Using SS-curves*." J. Constr. Eng. Manage., 126(2), 142–148.

[8] Barraza, G. A.; Back, W. E.; and Mata, F. (2004). *"Probabilistic Forecasting of Project Performance Using Stochastic S Curves."* J. Constr. Eng. Manage., 130(1), 25–32.

[9] Kim, B. C.; and Reinschmidt, K. F. (2009). *"Probabilistic Forecasting of Project Duration Using Bayesian Inference and the Beta Distribution."* J. Constr. Eng. Manage, 135(3), 178-186.

[10] Vanhoucke, M. (2012*). "Earned Value Management."*Proj. Manage. With Dynamic Scheduling, Part 3, 215-238.

[11] Lipke, W. H. (2003). *"Schedule is Different."*The Measurable News, summer, 31–34.

[12] Anbari, F. (2003) *"Earned Value Method and Extensions*." Proj. Manage. J., 34(4), 12–23.

[13] Jacob, D. (2003). "*Forecasting Project Schedule Completion with Earned Value Metrics*." The Measurable News, March, 7–9.

[14] Jacob, D. S. and Kane, M. (2004). *"Forecasting Schedule Completion Using Earned Value Metrics Revisited*." The Measurable News, summer, 11–7.

[15] Skitmore, M. (1992). "*Parameter Prediction for Cash Flow Forecasting Models*." J. Constr. Manage. Econom, 10(5), 397–413.

[16] Navon, R. (1996). *"Cash Flow Forecasting and Updating for Building Projects*."Proj. Manage. J., 27(2), 14–23.

[17] Chao, L. C.; and Chien, C. F. (2009). "*Estimating Project S-Curves Using Polynomial Function and Neural Networks*." J. Constr. Eng. Manage., 135(3), 169–177.

[18] Matlab (R2009a). "*User Guides of Matlab Software*", MathWorks, License Number 161051.

[19] Abdelazeem, S. A. (2014) "*Forecasting of Project Cost and Schedule Performance Using Probabilistic and Deterministic Models*" Ph.D. Dissertation, Faculty of Engineering, Zagazig University, (Incomplete).

[20] Zadeh, L., (1965). "*Fuzzy sets. Information and Control*."Proj. Manage. Body of Knowledge (PMBOK), Proj. Manage. Ins. PMI, 2000, 338–353.