Earned schedule formulation using nonlinear cost estimates at completion

DOI NUMBER: 10.19255/JMPM01307

PROJECT TIME

KEYWORDS
Earned Value Management • Earned Schedule • Cost Estimate at Completion • Duration Estimate • Project Monitoring

ROGER D.H. WARBURTON
• Boston University, Metropolitan College - United States
• rwarb@bu.edu

ALBERTO DE MARCO
• Dept. of Management and Production Engineering
• Politecnico di Torino, Italy
• alberto.demarco@polito.it

FRANCESCO SCIUTO
• Dept. of Management and Production Engineering
• Politecnico di Torino, Italy
• francesco.sciuto89@gmail.com

EARNED SCHEDULE FORMULATION

using nonlinear cost estimates at completion

• ABSTRACT •

This work contributes to improving available methodologies for duration and cost estimates of ongoing projects with nonlinear cost profiles. It is demonstrated that accurate time estimates can be made when a generalized mathematical formulation of the Earned Schedule and the point estimate methodology are used. It also highlights the advantages of using these duration estimate methodologies to provide more accurate nonlinear schedule-based cost estimates at completion. This is shown via application and comparison of the proposed methodologies to datasets of eight real case projects from the construction industry. In particular, the defined methodologies tend to perform better, on average, than traditional index-based formulae, especially in the early stages of project development when the practical benefits are the greatest for project teams to take their corrective actions.

1. INTRODUCTION

Theoretical weaknesses in the existing Earned Value Management (EVM) method are observed when it comes to accurately estimating project duration at completion, especially during the early stages of a project when just a small amount of value has been earned. A current approach to duration estimation at completion is based on Earned Schedule (ES) and several research studies claim that ES often works well (Vanhoucke and Vandevoorde, 2006, 2007). However, the concept of ES is not without its problems, as several criticisms have been issued for this methodology (Book, 2006, Kim, 2000), and its theoretical foundations are still needed to assure its extended and wider use in project monitoring and control practice Batseliera and Vanhoucke (2015).

The underlying theory of ES is based on a geometrical construction: the intersection time of the horizontal projection from the earned value curve, at the current time point, to the planned value curve (Lipke, 2010). Evensmo and Karlsen (2006) pointed out that the existing duration formula is based on linear cumulative planned and earned cost curves. However, why should a linear theory work when real-world project cost curves are usually presented as nonlinear?

To this end, this research contributes to the debate by determining if improvements in project duration estimates at completion are possible by using a combination of the generalized mathematical formulation of the ES technique, as proposed by (Warburton and Cioffi, 2016), and nonlinear cost profiles. The goal is to determine if such a generalized ES method provides more accurate estimates for the final duration when nonlinear cost growth models are taken into account (Narbaev and De Marco, 2014, Warburton, 2014).

A complementary goal is to verify that the approach is practically useful by demonstrating its use on real-world projects. To do so, data sets for eight real case projects from the construction industry are used for test and comparison with the standard methodology. The work is an attempt to determine if one duration forecast methodology is superior to the others.

We also contribute to the exploration of the relationship that may exist between accuracies of the Time Estimate At Completion (TEAC) and of the Cost Estimate At Completion (CEAC) and to understand whether ES-based CEAC formulae provide for improved results when the approach of Narbaev and De Marco (2014) is used.

To this end, the paper is structured as follows. First, previous definitions and pertinent estimate at completion methods are summarized in the literature section with regard to both duration and cost predictions. Second, the research methodology is defined. Third, duration and cost estimates at completion are computed for the sample data sets and the research results are analyzed and discussed. Finally, implications are presented together with conclusions and future research directions to address unresolved issues.

--- 1.1. Review of Literature ---

1.1.1. Pertinent Literature in TEAC

The problems of applying EVM to duration prediction are well known (Marshall, 2006, Book, 2003, 2006) and Lipke (2003, 2010) was one of the first to address this issue by defining a geometrical construction procedure for ES for which Stratton (2007) later provided a formal definition.

Project Teams can compute TEAC and CEAC, at any stage of development of an ongoing project, through Earned Value Management (EVM), using schedule and cost performance indices (Project Management Institute, 2013). However, the accuracy and reliability of these index-based estimates can be questioned because, while most projects have nonlinear S-curve profiles of cumulative expenditures, the entire theory of EVM is based on linear planned value, earned value, and actual cost curves. (Vanhoucke and Vandevoorde, 2006, Batseliera and Vanhoucke, 2015).

Both deterministic and probabilistic approaches have been developed to overcome this problem and improve the final estimates. As far as time predictions are concerned, the main deterministic method is to replace the EVM method with the ES method to determine the TEAC. The ES-based method has been shown to outperform EVM-based methods of forecasting the final duration of the project (Colina and Vanhoucke, 2015). However, until recently, ES presented the same critical theoretical issues as other methods: it is based on linearity of the cumulative planned and earned value profiles. Most projects seem to follow a nonlinear S-shaped curve of cumulative costs, which leads to a theoretical concern (Cioffi, 2005, Warburton, 2014). The nonlinearity issue was resolved by Warburton and Cioffi (2016), who generalized the theory of ES and showed that it was valid for nonlinear cost profiles.

Therefore, to better fit the cumulative S-curve cost patterns and to more accurately predict the final duration and final cost, probabilistic approaches have been developed. These methods include usage of Kalman filter (Kim and Reinschmidt, 2010, Abdel Azeem, Hosny and Ibrahim, 2014), Bayesian approaches Gardoni, Reinschmidt and Kumar (2007), Kim and Reinschmidt (2009), and fuzzy logic (Naeini, Shadrokh and Salehipour, 2011, Mortaji, Bagherpour and Noori, 2013). Cioffi (2005) demonstrated a method of parameterizing the S-curve.

As part of the methodologies for TEAC predictions, the study by Warburton and Cioffi (2016) is a foundation of this research. They applied a generalized, theoretical definition of ES to several different nonlinear project cost profiles and showed how to derive expressions to estimate the final duration.

1.1.2. Pertinent literature in CEAC

Standard EVM techniques have been shown to yield reliable predictions for the final CEAC (Vanhoucke and Vandevoorde, 2006, Christensen, 1993.). In addition, similar research efforts to those above have been employed to improve the traditional, index-based CEAC methods (Narbaev and De Marco, 2013). Warburton (2011) proposed a deterministic adjustment to the CEAC formula that proved to converge quickly to the actual final cost. The method employs time dependent expressions for planned value, earned value, and actual costs and three parameters that represent the reject rate of activities, the cost overruns, and the time required to repair the rejected activities.

Vanhoucke and Vandevoorde (2007) reviewed the accuracy of CEAC forecasting methods and concluded that the Cost Performance Index (CPI) and the Schedule Performance Index (SPI) provide valuable information about trends in project performance. Lipke, Zwikael, K. and Anbari (2009) attempted to improve the accuracy of the CEAC method by merging the standard CPI with a statistical technique that provides upper and lower confidence bounds of the forecasts at different confidence levels. The incorporated EVM data into the Bayesian inference for cost estimates. Barraza et al. (2004) introduced Monte Carlo simulation techniques into the cost estimates: the simulation had to be run based on the, so-called, progress-based curves. The available actual cost data were inserted in the process in order to simulate the future development of the Stochastic S-curves and obtain the CEAC. In a similar approach, Naeini and Heravi (2011) applied the SS curves theory and Monte Carlo simulations to compute cost estimates.

Vanhoucke (2012) employed Monte-Carlo simulation and concluded that networks with more parallelism have more variability than networks with a more serial structure. Parallelism, which produces S-shaped curves, degrades forecasting accuracy, which suggests that a more powerful, nonlinear theory is required. Similar research efforts have also been made to improve traditional index-based CEAC methods (De Marco, Briccarello and Rafele, 2009, Narbaev and De Marco, 2017; De Marco, Rosso and Narbaev, 2016).

Furthermore, the fuzzy approach to duration estimates has been used for improving CEACs and provides an alternative option for evaluating the future conditions of the project in a reliable and robust way, bringing the fuzzy numbers principles into the analysis (Naeini et al., 2011).

As part of the stream of research aimed at refining CEAC methodologies, important groundwork for this research is the method proposed by Narbaev and De Marco (2014, 2013), who opened the field to ES-based nonlinear CEACs. They introduced the perspective that the schedule performance of the project is considered to be an important factor in predicting final cost overruns or underruns. A Completion Factor (CF), which is defined through the standard ES construction, is able to characterize the estimate of the project's final duration.

Narbaev and De Marco (2014) used several different distributions for the cost profile (e.g., Gompertz, Logistic, Weibull, and Bass growth S-curves), which can then be used to fit a combination of AC and PV data to obtain curves that represent the shape of the project's cumulative cost. The CF can be included in the selected growth model to predict the final cost of the project. The proposed forecasting method applies to all stages of project development and, in particular, to the early stages of the project when there is little actual progress and few actual cost data points are available.

Evensmo and Karlsen (2006) noted that the current ES approach is linear and proposed a cubic polynomial cost curve. Cioffi (2005) proposed an Scurve cost profile and demonstrated its use in predicting project costs and schedules. Chen, Chen and Lin (2016) attempted to improve the predictive power of planned value by using a logarithm linear transformation of the planned value data and linear regression modeling. Warburton (2014) used a trapezoidal labor profile, which often describes construction projects, to derive accurate TEACs early in the project.

2. RESEARCH METHODOLOGY

The objective of this research is twofold. First, determining if improvements in project duration estimates are possible by using the generalized mathematical formulation of the ES by Warburton and Cioffi (2016). Second, to verify that taking into account nonlinear cost growth models, for the TEAC estimate, can be useful in predicting more accurate CEACs, e.g., the models as proposed by Narbaev and De Marco (2014, 2013). In fact, when the project cumulative cost profile does not grow linearly, using a growth model to fit data leads to better results and minimizes errors in duration and cost predictions.

The generalized theoretical definition of ES begins with the definition of the delay, δ(t), which is defined as the horizontal intersection back from the current cumulative earned value, Ce(t) to the cumulative planned value, Cp(t), which is given mathematically by:

$$Cp[t - \delta(t)] = Ce(t) \qquad (1)$$

Positive values for δ(t) represent accelerations while negative values represent delays. We now define the Earned Schedule, ES(t), as the time from the start of the project to that of the above intersection:

$$ES(t) = t - \delta(t) \qquad (2)$$

We follow the standard definitions of EVM: as each activity is completed, it earns its planned value, even if there is a cost increase or a delay in completing the activity (Project Management Institute, 2013, 2011). Further, because these new definitions are completely general, they apply to any nonlinear cost profile.

We next explain how the generalized algebraic definition of ES given above can be obtained via linear regression to fit the cumulative planned value curve to a cumulative growth distribution function. We formulate the theoretical framework for the TEAC by developing different duration forecast formulae, using the generalized definition of ES above, for nonlinear cumulative cost profiles. The result is analytical expressions for estimates of both ES(t) and and the final duration.

To accomplish this, the planned value data are fit to an S-curve formula using the standard Least Squares (LS) approach, i.e., for all data points, ti, we minimize the sum of the squares of the errors between the selected equation at ti and the actual planned value data at ti. The nonlinear LS fit was accomplished using the The Oakdale Engineering Data Fit 9 software package. In particular, we present the method and formulae for a Gompertz Growth Model (GGM) cumulative distribution function for the fit to the cumulative planned value data. We then introduce three types of estimation procedures, namely, standard, point, and cumulative.

Then, each of the theoretical duration estimation formulas were validated by comparing them to eight real-world projects in the construction, infrastructure, and renovation industries, as shown in **Table 1**. The projects were selected so as to represent a variety of possible schedule and cost performances, and include projects that both experienced delays and were accomplished on time. **Table 2** summarizes the data sets and records, for each project, the planned and actual duration, and the planned and actual cost, both with percent deviations from the planned values.

Label	Type	Literature Source
A	Civil Construction	Warburton, 2014
B	Renovation	Vandevoorde and Vanhoucke, 2006
C	Renovation	Vandevoorde and Vanhoucke, 2006
D	Infrastructure	Khamidi et al., 2011
E	Infrastructure	Shokri-Ghasabeh and Akrami, 2009
F	Civil construction	Singletary, 2006
G	Renovation	Vandevoorde and Vanhoucke, 2006
H	Civil construction	Valle and Soares, 2006

TABLE 01. List of the sample case projects

Label	Duration			Cost			Time	Cost
	Plan	Act.	Dev. %	Plan	Act.	Dev. %	Unit	Unit
A	20	30	95%	1,600	2,000	25%	Weeks	K$
B	9	12	33%	2,875	3,247	13%	Months	K€
C	9	13	44%	361	320	-3%	Months	K€
D	10	12	20%	58,000	59,183	-2%	Quarters	KMR[a]
E	15	16	7%	57,717	61,561	7%	Months	MIR[b]
F	13	14	8%	12,592	12,585	0%	Months	K$
G	10	9	-10%	906	952	5%	Months	K€
H	10	10	0%	12,563	12,563	0%	Months	KBS[c]

TABLE 02. Summary project data

Finally, the accuracy of a duration-augmented cost estimate prediction method is evaluated to test the validity of the different TEAC prediction procedures. For this purpose, the TEACs were computed using the mathematical models of the CEAC formula proposed by Narbaev and De Marco (2014) and described in the next sections.

2.1. Theoretical Formulation of TEAC

The model uses the standard EVM approach: if there is no scope growth, the earned value at the end of the project is neither higher nor lower than the original total planned value, i.e., all the planned work is completed and, therefore, earned (Project Management Institute, 2013, 2011). Moreover, from an empirical point of view, if the structure of the project does not change, the topology of the network is also unlikely to change during the evolution of the project. This supports the idea that the same curve shape can be used for the earned value curve as was used for the planned value curve.

There is considerable literature support for this assumption, which we refer to as the "same shape" assumption. For example, multiple studies have shown that both the cumulative planned value and cumulative earned value data closely follow the Putnam-Norden-Rayleigh (PNR) curve. (Warburton, 1983, Basili and Beane, 1981, Lee, 2002, Gallagher and Lee, 1996). Davis, Christle and Abba (2009) demonstrated that the PNR curve explains cost variation as well now as it did in 1970.

That the planned and earned curves have the same shape is an implicit assumption that is actually built into standard EVM. For example, the standard linear TEAC and CEAC formulas both assume that the planned and earned curves are linear, but with different slope parameters. However, the assumption is rarely acknowledged. For nonlinear cost profiles, this assumption is again invoked, although we make it explicit by carefully specifying which parameters change and which remain constant.

Consider the case where the cumulative planned and earned values are represented by Gompertz curves:

$$G(t) = \alpha\, e{-}e(\beta{-}\gamma t) \qquad (3)$$

where α is the asymptote (G(t) for t → ∞); β is the y-intercept; and γ is the growth rate. For the planned value curve, the parameter, αp, represents the budget, or the final planned cost, i.e., the asymptote. Therefore, since the total value earned equals the total planned value (Project Management Institute, 2013), the asymptotic parameter for the earned value curve is equal to that of the planned value curve, αe = αp.

We note that from a practical perspective that if this constraint is not observed, very erratic estimates for the final duration occur, especially when few data are available in the early stages of the project. A changing value of αe, relative to αp, indicates modifications in the evolution of the earned value curve. The parameter, β, is the y-intercept and, since most projects start at zero staff, the value of β is usually small for both the planned and earned value curves.

The growth of the planned value data is represented by the parameter, γp. For a project that is delayed or accelerated, the earned value data will progress with a different growth rate, which is represented by the parameter, γe. Therefore, as the project proceeds, delays and accelerations relative to the planned curve mean that γe = γp, as the earned value data evolves differently from the planned value, reflecting the actual work in progress. Therefore, since α merely scales the curve in the vertical direction, for cumulative cost curves representing projects, the interesting aspects of the Gompertz curve will be dominated by the value of the growth rate, γ.

Cp(T1) We denote the actual project's planned cost data at time, ti, as Cp(ti) and the earned data as, Ce(ti). We define the planned end point of the project as T1 and assume that during execution, if the project is delayed, it ends at T1′, where T1′ > T1. If the project is accelerated, T1′ < T1. The total planned cost is then Cp(T1).

The planned data for the entire project is available before the project goes into execution. The first step is to fit a GGM curve to the planned value data, using the LS technique. Next we use the earned value data, but unlike the planned value data, the earned value data is only available up to the current time. Therefore, we fit the earned value data at different stages throughout the project: early, middle, and late. This allows us to analyze the accuracy of the final duration estimates over time.

--- 2.2. Theoretical Formulation for the CEAC ---

Narbaev and De Marco (2014) developed a method for calculating the CEAC that involved integrating the standard ES approach with growth models that use a nonlinear regression analysis for fitting the actual cost to the planned cost. This is referred to as the "AC-PV" fit, and is defined in the next section. Duration estimates are used to compute the Completion Factor (CF), which is defined as the ratio of the estimated duration at completion to the planned duration. Thus, CF is the inverse of the schedule performance index and CF > 1.0 indicates that a project is likely to be delivered late, while CF < 1.0 indicates an early finish. The CF is then used to predict the final cost, which, at any time, t, is determined by:

CEAC = Ca(t) + (GMM[CF] – GGM[t]) × BAC (4)

where Ca(t) is the actual cost at time, t, and BAC is the planned budget. CF is a specific value at time, t, that represents the nonlinear growth model function in the regression. In fact, the GGM is modified to consider possible influence of the work progress in the CEAC. The main assumption of this refinement is that a favorable schedule efficiency tends to improve the final cost, while a poor schedule progress may increase the final cost. This modification represents an integrated cost-schedule approach because the cost estimate reflects the schedule impact as a determinant factor in the cost behavior.

Three methods can be used to calculate the CF, as presented in **Table 3**.

	Method	T_1'	CF
Linear	Standard	$\frac{tT_1}{ES(t)}$	$\frac{t}{ES(t)}$
Linear	Linear $(ES_{GM^2}$	$\frac{tT_1}{ES_{GM}(t)}$	$\frac{t}{ES_{GM}(t)}$
Non Linear	Point Est.	$\frac{1}{\gamma^0}\left(\beta^0 - \ln\left(-\ln\left(1/X^0\right)\right)\right)$	$\frac{T_1^0}{T_1}$

TABLE 03. Completion Factors for different prediction models of project duration.

--- 2.3. The AC-PV Fit ---

In order to parameterize a particular GGM, Narbaev and De Marco (2013) fit the available actual cost data (i.e., up to the present time) to the planned value data through the end of the project. This is called as the "AC-PV Fit" and the aim is to provide a GGM[CF] point close to the actual CEAC. This technique proved to be effective on a range of real projects.

However, the AC-PV Fit fails to generate reliable estimates for the CEAC when the project is significantly delayed or experiences a significant cost overrun. In that case, the gap between the last Ca(t) data point available and the first Cp(t) data point is likely to be huge. An example is shown in Figure 1, where the cumulative normalized actual cost and planned value data, for Project A, are plotted. In such cases, the AC-PV Fit leads to unreliable cost estimates.

The PV curve's fitting parameters are known (αp, βp, and γp). When the project experiences a significant cost overrun, the parameter, αa, which represents the asymptote of the actual cost curve, must be different from αp. Therefore, we compute a new GGM asymptote αa. Then, we allow the software to compute new values for all three

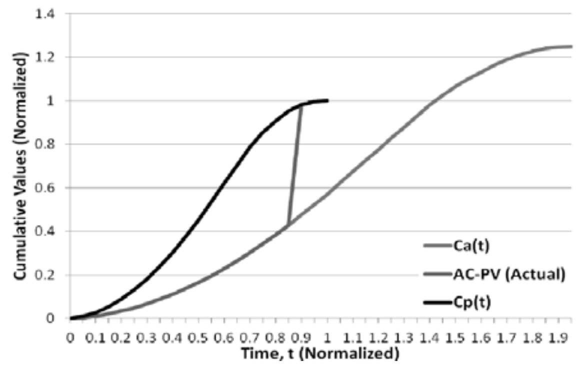

TABLE 03. Completion Factors for different prediction models of project duration.

parameters (αa, βa, and γa) using the actual cost data available to date.

When there is no scope creep, but a significant delay, the earned value curve is characterized by Ge(t), which is a GGM with parameters αe, βe, and γe. However, the same shape constraint above means that αe = αp and βe = βp, but γ_e represents a different growth of the earned value cost curve due to the delay. Since all the value planned is earned at the end of the project, αa = αe. We note that an indication of scope creep is, $\alpha_e = \alpha_p$. When there is a significant cost and/or scope growth, the actual cost curve is characterized by $G_a(t)$, which is a GGM with parameters αa, β_a, and γ_a.

Based on assumptions that in the latest stages of the project the growth tends to flatten and that Cp(CF) can be taken as a rough estimate of the cost at completion, Ca(CF), αa can be calculated using:

αa = (Cp(CF)-)/ Cp(T1) (5)

where Cp(T1) is the planned value at the originally scheduled completion time.

Cp(CF) is referred to as the value of the PV curve, Gp(t), dragged to the CF time point.

Next, both the AC-PV Fit and the new value of αa are applied to the set of case projects for comparison with the other methods. The analysis focuses on the possible connections between the accuracy of the TEAC, and the associated CEAC, at same stage of project development.

3. CALCULATING ES WITH THE GOMPERTZ GROWTH MODEL

As an example, we will use a Gompertz Growth Model (GGM) representation for the planned value, Gp(t), see Eq. (3). The earned value is Ce(t) and the delay, δ(t), at the point in time, t, can be derived, using Eq. (1) applied to the GGM planned value cost profile. ES is then computed from Eq. (2) and it is then possible to compute the TEAC. This is achieved in multiple ways, namely using the standard, point, and cumulative estimation methods, as presented in the following subsections.

--- 3.1. Standard Estimate ---

This is the standard earned schedule approach based on Warburton and Cioffi (2016), who demonstrated that the estimate of the TEAC is constant throughout the project and is given by:

T1′ = T_1(t) / ES(t) (6)

--- 3.2. Point Estimate ---

The second methodology is the, so-called, "point estimate." At any time, t, the cumulative planned value curve fit is defined as, Gp (αp, βp, γp, t). We next invoke the same shape constraint as described above, i.e., we assume that the earned value curve follows a GGM curve with the same values of αp and βp, but with a different value of γe. Using δ(t) from Eq. (1), we have,

Ge [αp, βp, γp, ti – δ(ti)] = Ce(ti) (7)

from which we calculate ES as,

$$ES(t_i) = \frac{1}{\gamma}\left(\beta - \ln\left[-\ln\left\{\frac{C_e(t_i)}{\alpha}\right\}\right]\right) \qquad (8)$$

At each point ti, we use the single, current, data point for the earned value, Ce(ti), to calculate a single value for γe:

$$\gamma_e = \frac{1}{t_i}\left(\beta_p - \ln\left[-\ln\left\{\frac{C_e(t_i)}{\alpha_p}\right\}\right]\right). \qquad (9)$$

Since γe only depends on the earned value data at one point, the resulting duration estimate will be referred to as a "point estimate." All the data are now available to estimate the final project duration:

$$T_1' = \frac{1}{\gamma_e}\left(\beta_p - \ln\left[-\ln\left\{\frac{1}{\alpha_p}\right\}\right]\right) \qquad (10)$$

--- 3.3. Cumulative Estimate ---

The next proposed model differs from the point estimate in the way the earned value curve parameters are determined: the cumulative estimate method uses all of the Ce(ti) data collected up to the current time in a curve fit. We again invoke the same shape constraints (αp = αe and βp = βe and the fit up to the current time generates a new value for γe for the earned value curve.

Hence, at any time, t, the cumulative earned value curve fit is defined as:

Ce(t) = G (αp, βp, γe, t) (11)

Once the parameters are determined, the TEAC is determined from Eq. (10).

4. APPLICATION TO THE CASE PROJECTS

We now analyze the accuracy of the above duration estimation methods for the project cost profiles described in **Table 2**. The key indicators used to compare the accuracy of the proposed methods are the Percentage Error (PE), and the Mean Absolute Percentage Error (MAPE), which are defined as per Eq. 12.

$$PE = \frac{T_1' - T_1}{T_1} \times 100 \qquad MAPE = \sum_{t=i}^{t=i+n} \frac{|PE_t|}{n} \qquad (12)$$

MAPE is used to compare the performances of the methodologies for the early, mid, and late stages. The stage of the project development is an important factor in evaluating the accuracy of methods. Particular emphasis is placed on results for early estimates, when from 10% to 35% of the planned budget has been earned. This is driven by the practical implications for project teams, who wish to estimate the reliability of predictions early enough in the project to enable timely and effective corrective actions. Mid estimates occur when the earned value is close to 50% of the planned budget. The MAPE for each stage includes more than one observation, made at consecutive time units and n represents the number of data points used.

Table 4 shows the MAPE values for early estimates, providing the average error and the standard deviation related to all the specific estimation methods. A similar investigation was conducted on errors for the middle estimates, see **Table 5**, and late estimates, see **Table 6**.

The declared main aim of this research is to prove that improvements in duration estimation can be realized if the overall shape of the cost profile is taken into account. The use of nonlinear profiles validates this claim for most of the projects. Although the sample of projects is not large enough to firmly establish the dominance of one method over the others, a number of useful conclusions can be drawn with regard to the project characteristics that can suggest which TEAC method is better or worse than the other methods.

In particular, as shown in **Table 2**, Project A ends up with a significant delay, almost doubling its planned duration. In this case, all methodologies give acceptable predictions in terms of PE, but the nonlinear estimates provide significant improvements in the TEAC. In particular, the cumulative duration estimate outperforms the others in the early and middle stages, while it is slightly worse in late stage. This could be due to the

Label	Std.	Lin.	Point	Cumlatv
A	4.32	3.12	3.53	2.40
B	12.66	9.70	3.75	7.59
C	23.44	36.78	30.72	27.62
D	14.26	25.34	13.43	23.37
E	1.55	3.16	5.04	23.66
F	54.49	54.91	5x.xx	15.63
G	7.32	20.05	9.98	50.30
H	13.04	8.84	7.72	31.19
I	10.90	4.95	5.48	3.68
Ave.	15.33	16.33	15.09	21.14
Std. Dev.	14.89	x6.33	14.88	14.60

TABLE 04. Early estimates of MAPE

Label	Std.	Lin.	Point	Cumlatv
A	2.91	4.53	2.88	0.42
B	9.99	10.52	9.57	9.28
C	26.52	20.09	31.14	26.38
D	9.24	13.94	8.65	12.35
E	1.98	4.34	4.29	5.47
F	19.85	20.36	27.43	1.24
G	4.92	4.93	4.18	6.82
H	7.98	7.99	5.13	7.22
I	6.41	6.04	9.33	4.92
Ave.	9.89	10.32	10.13	3.20
Std. Dev.	7.66	6.06	8.57	7.25

TABLE 05. Middle estimates of MAPE

Label	Std.	Lin.	Point	Cumlatv
A	A	B	C	D
B	E	F	G	H
C	26.52	20.09	31.14	26.38
D	9.24	13.94	8.65	12.35
E	1.98	4.34	4.29	5.47
F	19.85	20.36	27.43	1.24
G	4.92	4.93	4.18	6.82
H	7.98	7.99	5.13	7.22
I	6.41	6.04	9.33	4.92
Ave.	9.89	10.32	10.13	3.20
Std. Dev.	7.66	6.06	8.57	7.25

TABLE 06. Middle estimates of MAPE

erratic data towards the end of the project.

From the Tables of results, we observe that the estimates often improve as the project approaches completion. However, not all projects converge to the correct final duration and this could be due either to the weaknesses of the selected model or, simply, the unpredictable nature of the project data.

Project B has similar delayed behavior to Project A. However, Project B reports are monthly instead of weekly, making the forecast a more difficult task because there is less data to describe the growth curve. Moreover, the nonlinear cost shape is irregular at the beginning, which leads to erratic answers in the very early phase of the project development. Nevertheless, nonlinear estimates give relatively improved results and, in particular, the cumulative estimate methodology works better at all stages.

Project C was also behind schedule, by 44%, and the delay was caused by a dramatic drop in performance towards the end of the project execution, which made the AC curve deviate far from the fitted curve. In such a case, the standard ES methodology is dominant and it seems fundamental to have a precise fitting curve in order to obtain reliable answers from the set of nonlinear TEAC methodologies.

Project D was selected to study the behavior of duration estimates when the Ce(t) curve has a breaking point, by which we mean that the performance is suddenly reduced, in this case when the project is approximately 80% complete. Here, the point estimate appears to be a valid methodology that improves upon the other estimates.

Project E is an example of the cost curve following a quasi-linear shape.

Here the standard approach estimates are more reliable than alternative methods.

When the EV curve line grows differently in shape from what is planned, e.g., due to a particularly slow start of the actual progress, the linear and point estimates are not expected to work well, even though the Cp(t) fit is accurate. This behavior is exemplified by Project F, where the cumulative method is the only one that manages to avoid the overestimate in the final delay due to the initial lag. In fact, a very slow start can increase errors, if dragged linearly to completion.

Project G finished ahead of schedule by one month and both the standard and the nonlinear point estimates worked well. Project H was delivered on time and both the point estimate and the standard approach, using the GM ES, returned good answers.

5. IMPLICATIONS AND FUTURE RESEARCH

This work has both theoretical and practical implications. The theoretical perspective is that the models improve duration prediction by including lookahead capabilities: the models do not just consider past performance, as in look-back linear, index-based models, but also take into account S-curve cumulative cost profiles. The estimates also rely on the overall development of the project, taking into account the interaction between cost and schedule. Thus, the new methodologies emphasize the importance of combining the cost and schedule data.

Moreover, this work reaffirms the importance of the completion factor, CF. It extends its domain of applicability by introducing additional configurations and captures various scenarios of duration and cost performance at different stages of project development. The theory creates conditions for better estimates of schedule-based CEAC models.

In terms of its practical implication, this work is a contribution that should help project teams calculate accurate and reliable TEACs and CEACs. The cases cover various characteristics of nonlinear cost profiles and performance behaviors. This may also lead the project management software industry to add functionality to their packages, e.g., procedures to manage the fitting of the profiles and to automatically compute TEACs and CEACs if certain characteristics are specified by the user.

The results and open issues suggest several future research directions. First, improved theoretical work is required to explain the erratic behavior of the later estimates. Modified fitting procedures may be developed to avoid, or correct for, the behavior of these late estimates.

New indices could be proposed to represent various features of the project profile, such as, for example, the expected regularity of the growth curve and the expected productivity. This may help improve the reliability of the multivariate regression analysis to enhance the accuracy of the TEAC and CEAC. Clearly, this would require an even greater data set with consequent increased cost of the calibration activity. This may enhanced by statistical analyses that could help define the project behavior factors that play a significant role in the estimate at completion model.

Other growth models may be employed to compute the CEAC and may be tested and compared to the GGM model. Examples of other nonlinear growth models include the Putnam-Norden-Rayleigh model and Allen's trapezoidal model that decribes construction projects.

There is no one single prediction method that always returns the best estimate for the CEAC for any possible case. The accuracy of the estimate at completion depends on various factors, such as the specific project cost profile, the accuracy of the curve fitting procedure, the precision of the duration estimate when a schedule-based methodology is applied, and the relation between the actual project data and the assumed profile being used to represent it. However, we demonstrated that schedule-based CEACs are more accurate than traditional index-based methodologies. Therefore, this paper extends the results of Narbaev and De Marco (2013) by demonstrating that taking into account the nonlinear shape of the cost profile can improve the TEAC. Further, these nonlinear profiles lead to improvements in the CEAC when augmented by a schedule-based formula.

Still, there are some open points of discussion. Some of the proposed methodologies become erratic once the project is close to completion, which may well be due to manipulating small numbers. However, while the later estimates are not as reliable, neither are they as essential because project managers are more interested in early estimates of TEACs and CEACs when project teams can take appropriate actions.

The proposed refined nonlinear methodologies are not always more accurate than the linear model. In fact, this research demonstrates that the linear approach often gives acceptable estimates for a large range of projects, even if they are characterized by a nonlinear shaped cost curve. One explanation for this is that the standard ES definition results in the same formula for a variety of nonlinear cost curves. Therefore, more theoretical work is required to determine for which types of projects the linear formula is reliable.

6. CONCLUSION

This work demonstrates that improvements can be made to nonlinear TEAC methods when the generalized mathematical formulation of ES is used. It also highlights the advantages of using duration estimates to provide more accurate nonlinear schedule-based CEAC calculations. This was shown via application of the proposed methodologies to a sample of eight case projects.

The theoretically defined methodologies were proven effective on a range of projects by providing acceptable duration estimates. In particular, the point estimate duration methodology tends to perform better, on average, than index-based standard formulas, especially in the early stages of project development. This is when the practical benefits are greatest as it allows project teams to take timely management actions. ◆

• AUTHORS •

ROGER D. H. WARBURTON, PhD, PMP, is an Associate Professor in the Department of Administrative Sciences at Boston University's Metropolitan College. He earned a doctorate in astrophysics at the University of Pennsylvania, Philadelphia, PA, USA. He teaches graduate courses in Project Management, both online and in the classroom. Dr. Warburton lectures internationally about his research on Project Cost and Schedule Prediction, with specific emphasis on its practical applications and guidance for project professionals. He also writes about the theory of Project Management and its underlying models.

ALBERTO DE MARCO is an Associate Professor with the Dept. of Management and Production Engineering at Politecnico di Torino, Italy. He teaches Project Management, Project Finance and Operations Management at various institutions. He's been Visiting Professor at the Tongji University in Shanghai and the Massachusetts Institute of Technology. His research activities are in the areas of Project Management, Project Finance and Public-Private Partnership for infrastructures and services.

FRANCESCO SCIUTO, Francesco Sciuto was born on the 2nd of June 1989 in Novi Ligure (Italy). After completing the high-school, he attended the Management Engineering school at "Politecnico di Torino", getting his bachelor's degree in 2013; he then obtained his master's degree in Management Engineering at "Politecnico di Torino", in 2016. He developed his final thesis work (A generalized Earned Schedule definition: improvements in project Schedule and Cost Estimates using growth models) at Boston University, Boston (USA). Francesco is now based in Milano (Italy). He is a consultant at one of the major consultancy firms, working in the field of Program Management for multinational companies. Passionate for sports, in his free time Francesco enjoys playing everything which implies being involved in a team; he loves music and plays both the guitar and the piano.

• REFERENCES •

Abdel Azeem, S., Hosny, H.E., Ibrahim, A.H., 2014. Forecasting project schedule performance using probabilistic and deterministic models. HBRC Journal 10, 35–42.

Basili, V.R., Beane, J., 1981. Can the Parr curve help with manpower distribution and resource estimation problems? The Journal of Systems and Software 2, 59–69.

Batseliera, J., Vanhoucke, M., 2015. Evaluation of deterministic state-ofthe-art forecasting approaches for project duration based on earned value management. International Journal of Project Management In press.

Book, S., 2003. Issues associated with basing decisions on Schedule Variance in an Earned Value Management System. National Estimator Fall, 11–15.

Book, S., 2006. "Earned Schedule" and its possible unreliability as an indicator. The Measurable News Spring, 24–30.

C**hen, H.L., Chen, W.T., Lin, Y.L.,** 2016. Earned value project management: Improving the predictive power of planned value. International JournalofProjectManagement 34, 22–29.

Christensen, D.S., 1993. The estimate at completion problem: a review of three studies. Project Manage Journal 24, 37–42.

Cioffi, D.F., 2005. A tool for managing projects: An analytic parameterization of the S-curve. International Journal of Project Management 23, 215–222.

Colina, J., Vanhoucke, M., 2015. A comparison of the performance of various project control methods using earned value management systems. Expert Systems with Applications 6, 15.

Davis, D., Christle, G., Abba, W., 2009. Using the Rayleigh Model to Assess Future Acquisition Contract Performance and Overall Contract Risk (Volume I). Technical Report. CNA.

De Marco, A., Briccarello, D., Rafele, C., 2009. Cost and schedule monitoring of industrial building projects: Case study. J. Constr. Eng. Manage. 135, 853–862.

Narbaev, T., De Marco, A. (2017). Earned value and cost contingency management: A framework model for risk adjusted cost forecasting. Journal of Modern Project Management, 4(3), 12-19

Evensmo, J., Karlsen, J.T., 2006. Earned value based forecasts-some pitfalls. AACE International Transactions EV6, 1–5.

Gallagher, M.A., Lee, D.A., 1996. Final-cost estimates for research and development programs conditioned on realized costs. Military Operations Research 2, 2, 51–65.

Gardoni, P., Reinschmidt, K.F., Kumar, R., 2007. A probabilistic framework for bayesian adaptive forecasting of project progress. Computer-Aided Civil and Infrastructure Engineering 22, 182–196.

Kim, B.C., Reinschmidt, K., 2009. Probabilistic forecasting of project duration using bayesian inference and the beta distribution. Journal of Construction Engineering and Management 135, 178–186.

Kim, B.C., Reinschmidt, K., 2010. Probabilistic forecasting of project duration using kalman filter and the earned value method. Journal of Construction Engineering and Management 136, 834–843.

Kim, E.H., 2000. A Study of the Effective Implementation of Earned Value Management Methodology. Ph.D. thesis. The George Washington University. Washington, DC.

Lee, D., 2002. Norden-Raleigh analysis: A useful tool for EVM in development projects. The Measurable News March, 21–24.

Lipke, W., 2003. Schedule is different. The Measurable News Summer, 31–34.

Lipke, W., 2010. Earned Schedule. Lulu (R) Publishing, Lexington, KY.

Lipke, W., Zwikael, O., K., H., Anbari, F., 2009. Prediction of project outcome. The application of statistical methods to Earned Value Management and Earned Schedule performance indexes. International Journal of Project Management 27, 400–407.

Marshall, R.A., 2006. The contribution of Earned Value Management to project success on contracted efforts: A quantitative statistics approach within the population of experienced pratitioners. Ph.D. thesis. Lille Graduate School of Management.

Mortaji, S.T.H., Bagherpour, M., Noori, S., 2013. Fuzzy earned value management using fuzzy numbers. Journal of Intelligent & Fuzzy Systems: Applications in Engineering and Technology 24, 323–332.

Naeini, L.M., Shadrokh, S., Salehipour, A., 2011. A fuzzy approach for the earned value management. International Journal of Project Management 29, 764–772.

Naeini, M.E., Heravi, G., 2011. Probabilistic model development for project performance forecasting. World Academy of Science, Engineering and Technology 58, 396–401.

Narbaev, T., De Marco, A., 2013. Combination of growth model and earned schedule to forecast project cost at completion. Journal of Construction Engineering and Management, .

Narbaev, T., De Marco, A., 2014. An earned schedule-based regression model to improve cost estimate at completion. International Journal of Project Management 32, 1007–1018.

Narbaev, T., De Marco, A. (2017). Earned value and cost contingency management: A framework model for risk adjusted cost forecasting. Journal of Modern Project Management, 4(3), 12-19

Project Management Institute, 2011. Practice Standard for Earned Value Management. 2nd ed., Project Management Institute, Newtown Square, PA.

Project Management Institute, 2013. A Guide to the Project Management Body of Knowledge (PMBOK (R) Guide). 5th ed., Project Management Institute, Newtown Square, PA.

Stratton, R.W., 2007. Applying Earned Schedule analysis to EVM data for estimating Completion Date, in: The Association for Advancement of Cost Engineering Proceedings, pp. 41–45.

Vanhoucke, M., 2012. Measuring the efficiency of project control using fictitious and empirical project data. International Journal of Project Management 30, 2, 252–263.

Vanhoucke, M., Vandevoorde, S., 2006. A comparison of different project forecasting methods using Earned Value metrics. International Journal of Project Management 24, 289–302.

Vanhoucke, M., Vandevoorde, S., 2007. A simulation and evaluation of Earned Value Metrics to forecast the project duration. Journal of the Operational Research Society 58, 1361–1374.

Warburton, R.D.H., 1983. Managing and predicting the costs of real-time software. IEEE Transactions on Software Engineering SE-9, 562–569.

Warburton, R.D.H., 2011. A Time-Dependent Earned Value Model for software projects. International Journal of Project Management 29, 1082– 1090.

Warburton, R.D.H., 2014. A new schedule estimation technique for construction projects. Organization, Technology and Management in Construction, An International Journal 6(3), 1075 – 1082.

Warburton, R.D.H., Cioffi, D.F., 2016. Estimating a project's earned and final duration. International Journal of Project Management.