Stability and accuracy of deterministic project duration forecasting methods in Earned Value Management

Alexis Barrientos-Orellana[1] ; Pablo Ballesteros-Pérez[2]; Daniel Mora-Melià[3]*;

Maria Carmen González-Cruz[4] ; Mario Vanhoucke[5,6]

[1] PhD student. +56 968 442 022, alexis.barrientos@utalca.cl

Doctorado en Sistemas de Ingeniería. Dpto. de Ingeniería y Gestión de la Construcción.

Facultad de Ingeniería, Universidad de Talca. Camino Los Niches km.1, 3340000 Curicó, Chile

https://orcid.org/0000-0003-2243-9738

[2] Associate Professor, +34 622 478 882, pablo.ballesteros.perez@gmail.com

Centro de Investigación en Dirección de Proyectos, Innovación y Sostenibilidad (PRINS)

Dpto. de Proyectos de Ingeniería, Universitat Politècnica de València

Camino de Vera s/n, 46022 Valencia, Spain.

https://orcid.org/0000-0002-4629-9664

[*3] (*corresponding autor*) Associate Professor. +56 951 265 044, damora@utalca.cl

Dpto. de Ingeniería y Gestión de la Construcción. Facultad de Ingeniería

Universidad de Talca. Camino Los Niches km.1, 3340000 Curicó, Chile

https://orcid.org/0000-0002-6191-7299

(continues)

[4] Associate Professor. +34 963 879 866, mcgonzal@dpi.upv.es

Centro de Investigación en Dirección de Proyectos, Innovación y Sostenibilidad (PRINS)

Dpto. de Proyectos de Ingeniería, Universitat Politècnica de València

Camino de Vera s/n, 46022 Valencia, Spain.

https://orcid.org/0000-0002-6987-5732

[5] Professor, +32 926 435 69, mario.vanhoucke@ugent.be

Ghent University, Tweekerkenstraat 2, 9000 Gent (Belgium)

Vlerick Business School, Reep 1, 9000 Gent (Belgium)

[6] UCL School of Management, University College London, 1 Canada Square, London E14 5AA (UK)

https://orcid.org/0000-0001-6702-3563

Stability and accuracy of deterministic project duration forecasting methods in Earned Value Management

Structured abstract

Purpose: Earned Value Management (EVM) is a project monitoring and control technique that enables the forecasting of a project's duration. Many EVM metrics and project duration forecasting methods have been proposed. However, very few studies have compared their accuracy and stability.

Design/methodology/approach: This paper presents an exhaustive stability and accuracy analysis of 27 deterministic EVM project duration forecasting methods. Stability is measured via Pearson's, Spearman's, and Kendall's correlation coefficients; while accuracy is measured by Mean Squared and Mean Absolute Percentage Errors. These parameters are determined at ten percentile intervals to track a given project's progress across 4,100 artificial project networks with varied topologies.

Findings: Findings support that stability and accuracy are inversely correlated for most forecasting methods, and also suggest that both significantly worsen as project networks become increasingly parallel. However, the $AT+PD\text{-}ES_{min}$ forecasting method stands out as being the most accurate and reliable.

Originality: Unlike previous research comparing EVM forecasting methods, this one includes all deterministic methods (classical and recent alike) and measures their performance in accordance with several parameters. Activity durations and costs are also modelled akin to those of construction projects.

Practical implications: Implications of this study will allow construction project managers to resort to the simplest, most accurate and most stable EVM metrics when forecasting project duration. They will also be able to anticipate how the project topology (i.e. the network of

activity predecessors) and the stage of project progress can condition their accuracy and stability.

Keywords: Earned Value Management; construction projects; project duration; deterministic forecasting; metrics stability; time estimates accuracy.

Article classification: Research article.

Introduction

Earned Value Management (EVM) is a deterministic project monitoring and control technique widely adopted in many industries, including construction (Batselier and Vanhoucke, 2015a). EVM can measure whether a project is behind (or ahead) of schedule and whether it is costing more (or less) than planned. These applications are extremely useful for project managers taking corrective actions to bring the project back on track (Vanhoucke, 2010). Another common application of EVM metrics is generating forecasts of when the project will finish and how much it will cost by the time it is completed.

EVM has been included in the Project Management Body of Knowledge (PMBOK) Guide since its first edition in 1987. However, EVM has also been adopted by many other US government agencies [e.g. the National Aeronautics and Space Administration (NASA) and the United States Department of Energy]. More recently, EVM has also been standardized in other regions such as Australia (e.g. AS 4817-2003 and AS 4817-2006) and Europe (ISO 21508:2018). In parallel with its adoption by governments, practitioners, certification and professional bodies, EVM has also received extensive research attention (Vanhoucke, 2011, 2013).

Due to its relative simplicity, EVM also allows conveying in simple terms to top management how the project is performing. To do this, it relies on a series of *metrics* whose values can be updated whenever the *actual* activity durations, costs and percentages of completion are measured. These measurements should take place at approximately regular time intervals during project execution and are commonly known as *tracking periods*.

The present study will focus on project *duration* methods. These are commonly known in EVM as Estimated time At Completion or just EAC(t). EAC(t) methods have received much less research attention than their cost counterpart (Batselier and Vanhoucke, 2015b). Yet, many EVM metrics and EAC(t) expressions have been proposed over the last 20 years. Earlier EAC(*t*) expressions mostly resorted either directly to classical EVM metrics [Planned Value (PV), Actual Cost (AC) and Earned Value (EV)] or their derived indicators [the Schedule Performance Index (SPI) or the Schedule Cost Index (SCI)]. In 2003, though, a fourth metric named Earned Schedule (ES) was proposed by Lipke (2003). Several EAC(t) expressions have used the ES metric since, also in combination with its derived performance indicators [e.g. the SPI(t) or the SCI(t)]. Finally, and much more recently, other EAC(t) expressions have been proposed based on exponential smoothing and further reformulations of the ES metric (e.g. ES(e), ES_{min}, ES_{max}) (Ballesteros-Pérez et al., 2019).

Among recent variants of EAC(t) methods, some have not been intrinsically deterministic [e.g. (Acebes *et al.*, 2015; Elshaer, 2013; Nadaf *et al.*, 2019)]. However, this study will focus on *deterministic* methods. Deterministic methods are the most common nowadays in the EVM framework and they are also simpler and easier to use (Vanhoucke, 2013). These make deterministic EVM quite appealing to most construction practitioners.

Apart from EVM project duration forecasting methods, other more advanced techniques have also been proposed (e.g. fuzzy logic, neural network analysis, Bayesian inference, Monte Carlo simulation, statistical learning and artificial intelligence methods, Kalman filter algorithms, and endless variants of PERT) (Bai *et al.*, 2020; Ballesteros-Pérez, Cerezo-Narváez, *et al.*, 2020). However, most of these techniques are also more computer demanding and/or data-intensive than the EVM deterministic expressions analyzed here.

Hence, this paper will perform a *stability* and *accuracy* performance comparison of all deterministic EVM project duration forecasting methods found in the literature. We understand by use of the term *accuracy* that this refers to the capability of an EAC(t) method to anticipate the final project duration before the project is actually completed. Similarly, *stability* is defined here as the ability of an EAC(t) method to experience low volatility (as the project progresses and new information comes in), while showing a high correspondence between the current (forecast) and eventual (actual) project duration values. Both indicators —accuracy and stability— will be numerically defined later within the *Research methods*. A representative artificial projects' dataset will be used to measure the performance of both indicators for all EAC(t) methods.

We will analyze both stability and accuracy because, among the few studies that have compared EAC(t) expressions so far, most have neglected the stability criterion and/or have resorted to an *ad hoc* definition that cannot be easily generalized. Similarly, most comparative studies have resorted to a single parameter for measuring accuracy (generally the Mean Absolute Percentage Error, MAPE), which does not necessarily gauge the occasional existence of significant deviations (unlike the Mean Squared Error, MSE).

Literature review

Before continuing, the 27 EAC(t) methods to be compared are presented in Table 1. Due to the large number of metrics and variables involved, further details on all methods' variables are provided as *Supplemental material* (see Table S1).

< Insert Table I here >

Table 1 classifies the project duration forecasting expressions by authors [e.g. (Anbari, 2004; Ballesteros-Pérez et al., 2019; Batselier and Vanhoucke, 2017a; Jacob, 2003; Khamooshi and Golafshani, 2014; Lipke, 2003, 2011)], and type of formula. The formulae are all numbered on the first column for easier reference in the upcoming sections.

Broadly speaking, most of these methods (but not all) resort to one of these three standard EAC(t) formula configurations: *PD/PF*, MAX(*PD*, *AT*)/*PF* or *AT*+(*PD*-ES_x)/*PF*. Where: *PD* is the project Planned Duration (before it starts); *AT* is the Actual (current) Time; MAX(·) is the maximum of two variables; ES_x is one of the variants of the Earned Schedule metric (e.g. *ES*, *ED*, *ES(e)*, ES_{min}, etc.); and *PF* is a Performance Factor that can be 1 or another indicator such as *SPI*, *SPI(t)*, *SPI(t)(e)*, $SPI(t)_{ESmin}$, $SPI(t)_{ESmax}$, *SCI*, *SCI(t)*, *SCI(t)(e)*, $SCI(t)_{ESmin}$, $SCI(t)_{ESmax}$, even an exponential smoothing factor (e.g. $T_{t,SPI(t)}$, $T_{t,AT}/T_{t,ES}$ or $T_{t,EDI}$).

Previous research comparing EAC(t) methods' accuracy

The first comparison of EAC(t) project duration methods was performed by Vanhoucke (2010). He compared methods 1-9 from Table 1. In his comparison, method 7 was the most accurate but considering Mean Absolute Performance Errors (MAPE) only.

Five years later, Batselier and Vanhoucke (2015b) compared the same methods 1-9, but this time with 23 real projects. Again, method 7 turned out to be the most accurate. In the present study we will not use real projects, though. This is because *stability* calculations require multiple realizations of the same project (which can only be obtained by simulation). Then, Batselier and Vanhoucke (2015c) compared methods 7, 15, 16 and some of Eslhaer's (2013) non-deterministic EAC(t) expressions. Again, method 7 was the most accurate.

In 2017, Khamooshi and Abdi (2017) compared methods 10-13, 22 and 24 with another 19-real-project dataset. In their analyses, methods 22 and 24 were the top performers. Both methods 22 and 24 were proposed by the authors and resorted to exponential smoothing. In the same year, Batselier and Vanhoucke (2017b) proposed method 20, which was also coincidentally based on exponential smoothing. However, they only compared it against method 9, observing just a marginal accuracy improvement.

Also in 2017, de Andrade and Vanhoucke (2017) compared methods 7 and 15 on a 14-real-project subset extracted from Batselier and Vanhoucke's (2015b) dataset. Method 7 was again the top performer (regarding accuracy as measured by MAPEs only).

More recently, Ballesteros-Pérez et al. (2019) compared methods 1-22 and 25. In his study, method 22 was the best but very closely followed by methods 7, 14, 16 and 25. This study, however, omitted methods 23, 24, 25 and 27, and neglected the stability analysis. Also, it only calculated the MAPEs (not the MSEs), nor did it use activity durations and costs similar to those of real construction projects.

As can be seen, no previous study has compared all EVM project duration methods, nor used activity durations and cost representations of real construction projects. These aspects justify the current comprehensive comparative analysis. Additionally, our analysis will include measurements of the methods' Mean Squared Errors (MSEs) besides MAPEs (the only error metric previously reported). This is relevant as MSEs allow for the detection of *sporadic* but significant deviations from the project duration estimates. Namely, MSEs, on measuring *quadratic* deviations instead of *absolute*, grow quicker when significant deviations appear and are more difficult to dilute in error averages. MAPEs, on the other hand, tend to *blur* sporadic deviations as long as the method is accurate most of the time.

Previous research comparing EAC(t) methods' stability

Regarding the stability of EAC(t) methods within the temporal dimension, significantly fewer studies can be found in the literature. Stability of the cost dimension, on the other hand, was defined early on as the time point in the project life-cycle at which the Cost Performance Index (CPI) is already deemed to be accurate and constant (De Koning and Vanhoucke, 2016). Several formative studies referred to a rule of thumb which suggested that, from approximately the 20% stage of project completion onwards, cost estimates do not differ much from final project costs (Sato *et al.*, 2017). However, later studies have consistently denied the existence of such a putative point (Henderson and Zwikael, 2008; Khafri *et al.*, 2018; Kim *et al.*, 2019; Petter *et al.*, 2015).

Among those studies that addressed cost stability, a few also included a description of the variability that some EVM *time* indicators experience during project execution. The SPI and the SPI(t) were the commonly preferred indicators [e.g. (Henderson and Zwikael, 2008; De Koning and Vanhoucke, 2016; Ladeeda and Jeevan, 2020; De Marco *et al.*, 2016; Petter *et al.*, 2015)]. In those studies, researchers determined whether these indicators remained within an arbitrary interval (e.g. ±0.1 or ±0.2) from a given point until the end of the project execution (Khafri *et al.*, 2018; Petter *et al.*, 2015). Most of them also resorted to real projects, rather than artificial ones which, as stated earlier, remains the only way of assessing multiple potential outcomes of the same project (Kim, 2016; Ladeeda and Jeevan, 2020). However, some of these studies did focus on specific project factors potentially influencing stability: namely project duration and budget, project network topology, the S-curve of the project's baseline, and so forth. Some of their results have been at least partially inconclusive (Henderson and Zwikael, 2008; De Koning and Vanhoucke, 2016), but they did find that project duration and budget do not seem to influence project stability in real projects. They also suggested that the steeper the project expenditure S-curve line was in the middle, the more likely the expected stability of the CPI, SPI, and SPI(t) later in the project.

We can only find three studies focusing on the time stability of EAC(t) methods. The first was Wauters and Vanhoucke (2015) who compared 12 EAC(t) methods' stability and accuracy. The methods' accuracy was measured with MAPEs, whereas stability was measured with an ad-hoc *Mean Lags* estimator. This *Mean Lags* estimator was similar to the MAPE in the sense that it measured the differences in the EAC(t) estimates from one tracking period to the next in absolute percentage values. The authors claimed that their estimator had the advantage of being independent of each method's accuracy. That is indeed true, but that is not the only way of detaching stability from accuracy. In this regard, an EAC(t) method that consistently claimed that project duration equals the Planned Duration (PD) would be perfectly stable, but utterly useless for project managers. Additionally, it seems more advisable comparing the possible EAC(t) estimates at one tracking period, not with the next tracking period's, but with the actual project durations, that is, with the final set of possible Real Duration (RD) values. We are of course assuming a probabilistic approach here as, by definition, it is the only way of assessing the stability of multiple possible project realizations. This is, as mentioned earlier, the very reason why this study cannot resort to a real projects dataset to perform the EAC(t) methods stability analysis.

The other two studies which compared some EAC(t) methods' stability (but this time not their accuracy) were Batselier and Vanhoucke's (2017a) and de Andrade et al.'s (2019). They proposed a Regularity Index (RI) which measured deviations in the project value accrued during execution. Namely, this index measured the Planned Value (PV) curve's degree of closeness to a perfectly linear curve. However, this RI indicator conceives the evolution of the PV as linear, when indeed it is usually closer to an S-shaped curve in most construction projects (Hürol *et al.*, 2020). This is because, in most construction projects, more activities are being performed (and therefore generating associated costs) in the middle stages of the project as compared to the initial and final stages (Narbaev and De Marco, 2017). Also, this RI indicator conflates accuracy with stability.

Hence, due to the limitations highlighted in the three previous studies mostly regarding their ad hoc stability indicators, our study will resort to parametric order correlation coefficients (e.g. Spearman's and Kendall's) for measuring EAC(t) methods' stability. These order correlation coefficients are also more statistically robust, and their interpretation easier.

To serve as a summary of the works reviewed in this section, Table II is included.

This table summarizes the scope and limitations of previous studies, comparing some EVM project duration forecasting methods.

< Insert Table II here >

Research methods

Artificial project networks dataset

In our study, we use an artificial projects dataset which boasts 4,100 activity-on-node networks with varied topologies (different configurations of activities' predecessors). Each project has 30 activities plus two dummy activities (zero cost and duration) indicating the start and end of each project. This dataset was generated with RanGen2, a robust random network generator validated in previous research (Demeulemeester *et al.*, 2003; Vanhoucke *et al.*, 2008). This same dataset has been used by other studies on EVM (e.g. (Ballesteros-Pérez, Sanz-Ablanedo, *et al.*, 2019; Batselier and Vanhoucke, 2015c; Colin and Vanhoucke, 2014; 2010, 2011)). The complete project dataset is currently curated by the Ghent University Operations Research & Scheduling Research Group. All project files can be downloaded at: https://www.projectmanagement.ugent.be/research/data (MT set). For each project, a file containing all activities' predecessors information can be found. Files also contain information on resource allocation that will not be used in this study.

More precisely, these 4,100 projects were generated by setting staggered values of the Serial-Parallel (SP) indicator. The SP indicator measures how close a project is to a serial network (SP=1) or a parallel network (SP=0). The SP is calculated as:

$$SP = \frac{q-1}{n-1} \tag{1}$$

Where q is the number of activities in the path with the highest number of activities (which is not necessarily the longest in duration), and n is the total number of (non-dummy) activities in the network (30 in our case for all projects).

This way our project dataset contains networks with SP={7%, 17%, 28%, 38%, 48%, 59%, 69%, 79%, 90%}. More rounded values (e.g. SP=10%, 20%,30%…) were not possible due to the fixed amount of activities within each network (30). Also, we did not consider perfectly serial (SP=100%) or parallel (SP=0) projects, as those configurations are not representative of real construction projects [see some average SP values of building, civil engineering, industrial and services projects in Ballesteros-Pérez et al. (2020), Table 9].

Finally, while the 4,100 projects were generated by increasing values of the SP indicator, other topological indicators were also calculated. They can be found in the Excel file of the *Supplemental material* for each network. These other indicators are the Coefficient of Network Complexity (CNC), the Order Strength (OS), the Activity Distribution (*AD*), the Length of Arcs (*LA*), and the Topological Float (*TF*). Further details on the meaning and calculations of these topological indicators can be found in Vanhoucke (2010). However, we will not refer to them anymore as our analyses did not find any significant stability nor accuracy correlation with any of them.

Activity durations and costs

Despite previous studies having assumed different distributions for modeling activity durations and costs, we use correlated Lognormal distributions whose variability resembles that of real construction projects. Several studies support, in fact, that the distribution of activity durations and costs closely resembles Lognormal distributions in real construction projects (Colin and Vanhoucke, 2016; Trietsch *et al.*, 2012; Vanhoucke, 2015). Furthermore, a recent analysis of construction activities also measured that the activities' duration-cost correlation can vary between 0 (no correlation whatsoever) and 100% (perfect correlation) (Ballesteros-Pérez, Sanz-Ablanedo, *et al.*, 2020).

Hence, we generate Lognormally-distributed activity duration and cost values with some partial correlation between both values. Mathematical details on how these durations and values are generated are included in the *Appendix*.

Additionally, when calculating the project schedules, all activities are scheduled to start as soon as possible (ASAP scheduling). Activity preemption is not allowed.

Accuracy and stability measurement

For measuring the *accuracy* of the 27 EAC(t) project duration forecasting methods we calculate the Mean Squared Errors (MSEs) and Mean Absolute Percentage Errors (MAPEs) at 10% of project progress tracking intervals (respect the Real project Duration RD). Namely, for each EAC(t) method and tracking interval (AT), we calculate:

$$MSE_{AT} = \frac{1}{M}\sum_{m=1}^{M}\left\{\frac{1}{K}\sum_{k=1}^{K}\frac{\left(RD_{mk} - EAC(t)_{AT}\right)^2}{RD_{mk}}\right\} \quad (2)$$

$$MAPE_{AT} = \frac{1}{M}\sum_{m=1}^{M}\left\{\frac{1}{K}\sum_{k=1}^{K}\frac{\left|RD_{mk} - EAC(t)_{AT}\right|}{RD_{mk}}\right\} \quad (3)$$

Where:

M is the total number of projects analyzed (M=4100) and m=1, 2…4100.

K is the total number of simulation runs per project (K=100) with k=1, 2…100.

RD_{mk} is project m's Real Duration in simulation run k.

$EAC(t)_{AT}$ is the project duration estimate at tracking period AT=10%, 20%…100% of RD_{mk}. For the sake of simplicity, though, we will not use an extra subscript to refer to each of the 27 EAC(t) forecasting methods.

As justified earlier, the MAPE allows us to directly anticipate the order of magnitude of our errors when we use EAC(t) methods in real contexts. The MSE, on the other hand, allows detecting occasional, but significant errors of the EAC(t) methods.

For measuring the methods' *stability*, we resort to three correlation coefficients:

Pearson's linear correlation (*R*), Spearman's rank correlation (ρ), and Kendall's rank correlation (τ). The use of Spearman and Kendall's coefficients was justified in the *Literature review* section. Pearson's correlation calculations are only included for comparative purposes.

Namely, Pearson's *R* correlation of the EAC(t) estimates at every tracking period AT with the final Real project Duration (RD) is calculated as:

$$R_{AT} = \frac{1}{M}\sum_{m=1}^{M}\left\{\frac{\sum_{k=1}^{K}\left(EAC(t)_{AT} - \overline{EAC(t)_{AT}}\right)\left(RD_{mk} - \overline{RD_{mk}}\right)}{\sqrt{\sum_{k=1}^{K}\left(EAC(t)_{AT} - \overline{EAC(t)_{AT}}\right)^2}\sqrt{\sum_{k=1}^{K}\left(RD_{mk} - \overline{RD_{mk}}\right)^2}}\right\} \tag{4}$$

Where $\overline{EAC(t)_{AT}}$ represents the average of the EAC(t) estimates at tracking period AT in the K=100 simulation runs; and $\overline{RD_{mk}}$ represents the average of project *m*'s RD values in the K=100 simulation runs.

Pearson's linear correlation coefficient *R* has the advantage of having a straightforward interpretation, i.e., the higher *R*, the more (linearly) proportional RD values are expected to be respect to the EAC(t) estimates. However, it also has some important limitations. First, it mixes accuracy and stability. Hence, it is useful as a combined indicator but misleading for measuring stability only. Second, the relationship between EAC(t) and RD is expected to be nonlinear, at least for early tracking periods. Third, Pearson's correlation assumptions (that both variables are normally distributed, linearity and homoscedasticity) do not seem realistic in our analysis. Hence, we propose using rank correlation coefficients.

Spearman's rank correlation coefficient ρ is calculated at every tracking period AT as:

$$\rho_{AT} = \frac{1}{M}\sum_{m=1}^{M}\left\{1-\left(6\sum_{k=1}^{K} d_{mk}^{\ 2} \Big/ K\left(K^2-1\right)\right)\right\} \tag{5}$$

Where d_{mk} is the difference between the ranking (order) values of $EAC(t)_{AT}$ and RD_{mk} for project *m*, for the K=100 simulation runs and at a particular tracking period *AT*, that is,

$d_{mk} \equiv rank\left(EAC(t)_{AT}^{k}\right) - rank\left(RD_{mk}\right)$ for k=1,2…K.

Spearman's correlation has the advantage of being parametric, thus, capable of detecting nonlinear relationships among the EAC(t) estimates and the RD values. Additionally, Spearman's correlation does not carry any assumptions about the distribution of the data and works very well with ordinal data (as in our analysis). The calculation of ρ is not computationally demanding either. However, Spearman's ρ is a little more sensitive to errors and less robust than Kendall's τ.

That is why we also suggest the calculation of Kendall's τ correlation coefficient for measuring the stability of the EAC(t) estimates at any tracking period AT:

$$\tau_{AT} = \frac{1}{M}\sum_{m=1}^{M}\left\{\frac{4P_m}{K\left(K-1\right)} - 1\right\} \tag{6}$$

Where P_m represents the number of concordant pairs between the EAC(t) estimates and the RD values for each project m, that is:

$$P_m = \sum_{k=1}^{K-1}\sum_{\ell=k+1}^{K}\mathbf{1}\left\{\left(EAC(t)_{AT}^{\ell} - EAC(t)_{AT}^{k}\right)\left(RD_{mk}^{\ell} - RD_{mk}^{k}\right) > 0\right\} \tag{7}$$

Where $\mathbf{1}(\cdot)$ is a binary operator that equals 1 when the condition is fulfilled and 0 when it is not.

Kendall's rank correlation coefficient is the more robust of all three correlation coefficients for measuring the methods' stability. However, its values tend to be lower than Spearman's and it also is much more computationally expensive. Particularly, the calculations involved in Kendall's τ are proportional to K^2, that is, 10.000 assuming K=100. (number of simulation runs per project). However, since we are evaluating 4,100 projects, 27 EAC(t) methods and 10 tracking periods, the number of calculations quickly skyrockets ($10^4 \cdot 4100 \cdot 27 \cdot 10 = \cdot 1.107 \cdot 10^{10}$). This is the reason why only 100 iterations were performed per project. Otherwise, for K=1000 simulation runs per project, for example, the computational effort would have required months of computing time for an average computer.

Still, as the total number of simulations is circa half a million (4,100 projects·100 simulations/project=$0.41\cdot10^6$) this will provide us with enough accuracy up to the second decimal place in our results. The latter as the error of Monte Carlo estimates are proportional to $1\Big/\sqrt{n^{o}\,simul.\,runs} = 1\Big/\sqrt{M\cdot K} = 1\Big/\sqrt{0.41\cdot10^{6}} = 0.00156 < 1\%$ (Koehler *et al.*, 2009).

Finally, for the sake of clarity, Figure 1 includes a graphical description of both Spearman and Kendall's correlation coefficients interpretation.

< Insert Figure 1 here >

For the examples shown in Figure 1, Pearson's correlation coefficient would have turned out low values. This is because this coefficient measures relative numerical deviations between the EAC(t) estimates and the project's possible Real Duration RD outcomes.

However, both Spearman's ρ and Kendall's τ coefficient values depend on relative order differences, not on pairwise numerical differences. Hence, it does not matter whether each pair of (EAC(t), RD) values is numerically close, nor whether their deviations follow a nonlinear relationship. What matters is the degree of coincidence between the relative orders of both sets of values. If there is a good match, then EAC(t) estimates will remain in an approximately constant percentile until the project finishes. This means each EAC(t) estimate will point towards the correct RD value from that tracking period onwards. This calculation becomes much more accurate of course, as we count on more data points. That is why we used K=100 (iterations) points per project.

Results

Once the artificial projects dataset, the simulation framework and the analysis variables have all been described, Figure 2 and 3 present the 27 EAC(t) project duration forecasting methods' stability and accuracy results for the 4,100 artificial projects. As justified earlier, the methods' stability is represented by three correlation coefficients (Pearson's R, but mostly Spearman's ρ and Kendall's τ) at the top of both Figures 2 and 3. The methods' accuracy, on the other hand, is measured by MSEs and MAPEs both at the bottom left and center of Figures 2 and 3, respectively. The difference between Figures 2 and 3 lays in the performance results of each sub-table being classified by tracking period AT (Figure 2) or by the value of the Serial-Parallel indicator *SP* (Figure 3). A performance summary table is also included at the bottom right corner of Figures 2 and 3.

< Insert Figure 2 here >

< Insert Figure 3 here >

Interpretations of these two tables are rich but highly varied. We will focus on the most relevant only. Further details presented by projects can be found as *Supplemental material*.

EAC(t) methods stability

Comparatively speaking, the average results of the three correlation coefficients (see summary tables at the bottom right corner of Figures 2 and 3) are not that different. However, Kendall's τ correlation values are significantly lower and produce slightly different results.

As expected, the closer the tracking period is to the end of the project in Figure 2 (AT values closer to 100% of RD), the higher the stability and accuracy of all EAC(t) methods. This agrees with almost all existing research on both project duration and cost forecasting (Henderson and Zwikael, 2008; De Marco *et al.*, 2016; Warburton *et al.*, 2017). It is worth noticing, though, that AT=100%RD does not represent a finished project but a project that has just less than one day left to finish. Otherwise, stability and accuracy results would have been perfectly stable and accurate at AT=100%RD for most methods.

Analogously, but now in Figure 3, the higher the Serial-Parallel indicator (meaning more serial networks), the more stable and accurate all methods become. This echoes the results of Vanhoucke (2010) and Ladeeda and Jeevan (2020). Particularly, R and ρ values are very high for projects with SP>80%. Conversely, for parallel projects with SP<30%, all EAC(t) methods are quite unstable. In fact, among Kendall's τ results, we can even find some negative correlation values. This means EAC(t) methods produced estimates which are reversed in order with the actual RD values. This can only be fully appreciated, though, in the project-by-project results in the *Supplemental material*. Therein, it can be observed how for projects with low SP values, almost all EAC(t) methods have negative correlation values during the early stages of project execution (AT<30%RD). This renders all EAC(t) methods useless at the early stages of a project when the project network is close to parallel. Thus, EAC(t) methods may still be accurate, but lower EAC(t) estimates will point towards longer RD values, whereas higher EAC(t) estimates will end up in shorter project durations.

Regarding projects with SP values between 20% and 80%, we can find moderate correlation coefficients in most cases (R, ρ and τ values between 0.40 and 0.80). But it must be borne in mind that these are *average* values for all tracking periods. That is, if a correlation coefficient is, for example, 0.50, this means that it will probably have been close to 0 at early tracking periods, but surely above 0.80 in later stages. Hence, for projects with moderate SP values, most EAC(t) methods are sufficiently stable, but only once half of the project execution has been exceeded. This is also in agreement with previous studies [e.g. (de Andrade *et al.*, 2019; Batselier and Vanhoucke, 2017a; Wauters and Vanhoucke, 2015)] and can be better appreciated in the AT>50%RD columns presented within the three tables at the top of Figure 2.

EAC(t) methods accuracy

Regarding EAC(t) methods' accuracy (tables at the bottom half of Figures 2 and 3), we can find a similar pattern for both the MSEs and the MAPEs. Particularly, it is only from the middle of the project execution onwards (AT>50%RD in Figure 2) that the errors are relatively small (MSEs<1 and MAPEs<10%) for most EAC(t) methods. This agrees with the findings of Ballesteros-Pérez et al. (2019) who analyzed some EAC(t) methods using MAPEs.

When it comes to the influence of the network topology in Figure 3, most EAC(t) methods perform sufficiently well (MAPEs<10%) for projects with SP>50% (rather serial projects). The same happens with MSEs but only for projects whose SP>60%. This means that, halfway through the project execution, we can still find some sporadic but relevant estimation errors in most EAC(t) methods.

On the other side of the spectrum, Figure 2 shows that most EAC(t) methods seem to be very inaccurate during the first third of the project execution (AT<30%RD), even until much later in projects with SP<30%. This issue had not been identified within the extant literature on EVM, however, the results clearly indicate that, at this execution stage, MAPEs >> 10% and MSEs are usually above unity.

Top-performing EAC(t) methods

In the two tables at the bottom right corner of Figures 2 and 3, we can find a summary of which methods perform well (*) or excel (**) in both stability and accuracy dimensions. Figure 2 depicts an average assessment considering that there are more possible network combinations of projects in the dataset with intermediate values of the SP indicator (as also happens in real projects). In Figure 2, methods 14 and 22 are the top performers, but methods 7, 12, 16, 20 and 25 are also good. It is worth highlighting that almost all of these EAC(t) methods use the configuration $AT+(PD\text{-}ES_x)/PF$ with a Performance Factor $PF=1$, that is, one of the simplest mathematical configurations. Similar findings had also been reported by Vanhoucke (2010) and Ballesteros-Pérez et al. (2019), although the authors compared fewer EAC(t) methods.

However, when analyzing the summary performance results in Figure 3, we find some unexpected results. When averaged by the SP indicator, most methods that are stable are not accurate, and vice versa. This was earlier suggested by Wauters and Vanhoucke (2015) but could not be validated by Batselier and Vanhoucke (2017a) or de Andrade et al. (2019). However, there is one method that is sufficiently stable and highly accurate (method 22). This method uses the configuration $AT+PD\text{-}ES_{min}$, where ES_{min} is the Earned Schedule calculated on the most delayed path during the project execution. Method 22 limits the quadratic errors very early (AT>40%RD) compared to other methods, and from that point onwards we can also expect project duration estimates to differ <10% from the final Real Duration RD.

Finally, it is worth noticing that among methods 22 to 27, there are several which are very stable and others which are very accurate. Since these six methods are quite similar to each other (they use variants of the same metrics ES_{min} and ES_{max} with different Performance Factors), future efforts might look for superior combinations of these metrics. This is left for future research, however, as it is not the intention of this paper to propose new EAC(t) methods.

Discussion

Results have shown that in parallel projects (SP<25%), most EAC(t) methods are very inaccurate and unstable (MSEs>1, MAPEs>15% and correlation coefficients<40%). The opposite happens to serial projects (SP>60%) in which most EAC(t) methods are quite accurate and stable (MSEs<0.5, MAPEs<10% and correlation coefficients>70%). In projects with intermediate topologies (25%<SP<60%), we can expect most EAC(t) methods to become sufficiently accurate and stable too, but only approximately from the middle of the project execution onwards.

A pertinent question is, how can real construction projects benefit from these findings? Inevitably, when a project manager observes that the duration forecast anticipates a project duration significantly longer than that originally envisaged, he or she may need to take action to bring the project back on track. The project manager will then either have to allocate more resources to make the existing project(s) more efficient and/or make the contractors work overtime. However, the specific nature of the systemic project changes triggered by a likely late completion date falls beyond the remit of this work. However, a construction manager might ask 'which are the serial-parallel (SP) values of real construction projects so that we can select the best project duration forecasting method?'

In this regard, serial projects tend to perform the most relevant activities one after the other. Highway projects, pipelines, train tracks, and renewable energy projects are some examples of projects with high SP values (Ballesteros-Pérez, Sanz-Ablanedo, *et al.*, 2020). However, when they have enough resources, these projects' activities can also be partially executed in parallel. Examples of parallel projects with low SP values are usually those that involve several but simple and homogeneous construction units. Examples are residential condominiums, quarry earthworks, off-site prefabrication, etc. (Ballesteros-Pérez, Sanz-Ablanedo, *et al.*, 2020). In between, we can find most construction projects whose schedules boast a mixture of series and parallel paths. Among them, we have most buildings, industrial facilities, water treatment plants, etc.

Hence, when choosing a project duration forecasting method, the project manager will need to consider the project seriality and/or the number of resources he/she counts on to perform the project. If many parallel paths or lines of work are expected to be active simultaneously, most EAC(t) methods will provide very inaccurate and unstable estimates, at least until the end of the project is near. But probably by then, the project might be very difficult to bring back on track if it was delayed and/or was too expensive. For these types of projects, other suitable project control and duration forecasting alternatives would be Schedule Risk Analysis (SRA) or Stochastic Network Analysis (SNA) (Ballesteros-Pérez, Cerezo-Narváez, *et al.*, 2019).

Conversely, if the project has just a few parallel paths and/or few resource teams, earned value management EAC(t) methods constitute a simpler yet accurate and stable alternative for forecasting the project duration. For most construction projects with 30%<SP<100%, the EAC(t) methods described here will also be a good alternative. However, apart from method 22 (which becomes reliable from 40% of the project execution onwards), the others might need to wait until 50-60% of the project has been executed. Many project managers may still find that this point as too late to take corrective actions (Ballesteros-Pérez, Elamrousy, *et al.*, 2019), and it probably is.

Conclusions

Earned Value Management (EVM) is a common project monitoring and control technique for measuring how a project is performing in both time and cost dimensions. Over the last 20 years, many EVM-based project duration forecasting expressions (named EAC(t)) have been proposed. However, their accuracy and stability had not been exhaustively measured. Among the very limited studies comparing EAC(t) methods' accuracy or stability, none has included all recent EAC(t) expressions, most have used *ad hoc* indicators whose results are difficult to generalize, and they did not simulate activities' variability according to real construction projects.

This study has performed a comparison of all deterministic EAC(t) project duration forecasting methods found in the literature (27 methods to date). The methods' stability has been measured using three correlation coefficients (namely, Pearson's, Spearman's, and Kendall's), whereas accuracy has been measured with Mean Squared Errors (MSEs) and Mean Absolute Percentage Errors (MAPEs). Additionally, all of these parameters have been evaluated and analyzed at 10% project progress intervals on a project dataset with 4,100 artificial networks with varied Serial-Parallel (SP) indicator values.

The most relevant results point out that almost all EAC(t) methods produce very inaccurate estimates until at least half of the project is completed. On the contrary, in the last third of the project, most methods are very accurate. Regarding stability, most EAC(t) methods are quite unstable in the early third of the project execution. However, in the last third, all EAC(t) methods become quite stable.

Regarding project topology, the results support that stability and accuracy are inversely correlated for most EAC(t) methods, and that both significantly worsen as project networks become more parallel (mostly for projects with SP<30%).

Among the top-performing EAC(t) methods, method 22: $AT+PD-ES_{min}$ is quite stable and highly accurate, even from the early stages of project progress. It also shows promise for further enhancing its accuracy and stability by combining it with other similar methods that use the ES_{min} and ES_{max} metrics. This is expected to be addressed in future research.

A comparison of current EVM project *cost* forecasting methods is expected to be addressed in future research. This should be relatively straightforward once all cost-related performance indicators have been parametrized and included within the calculation framework developed for this paper. A performance comparison of deterministic *vs* non-deterministic project forecasting methods, incorporating both time and cost dimensions, will also be developed. However, this will involve departing from the EVM framework and implementing alternative calculation approaches (e.g. multivariate regression, statistical learning, fuzzy logic, etc.)

Limitations of this study are mostly related to the size of the projects in the dataset and the number of simulations performed in each project. Regarding the project size, all networks had 30 (non-dummy) activities. Regarding the number of simulations, only 100 simulation runs per project were performed. The size and number of simulations were intentionally restricted to avoid the number of calculations from skyrocketing, especially when it comes to calculating Kendall's rank correlation coefficient. However, considering the analysis encompassed 4,100 different projects, we still expect our results to be sufficiently representative of most real construction projects.

Appendix

This appendix provides detailed explanations on how the activity durations and costs were generated for each of the 4,100 artificial projects. Their activity duration and cost mean values and dispersion values were generated so that they resembled those from real construction project activities. Activity durations and costs were also correlated.

Namely, Lognormally-distributed activity durations (d_i) were generated in this study for each activity i by means of these expressions:

$$d_i \sim e^{Normal\left(mean=\mu_i, st.dev.=\sigma_i\right)} \tag{8}$$

$$\mu_i \sim Normal\left(mean=1, st.dev.=0.25\right) \tag{9}$$

$$\sigma_i \sim \mu_i \cdot Uniform\left(lower\ bound=0.25, upper\ bound=0.75\right) \tag{10}$$

Where e is Euler's number; Normal($\cdot$) and Uniform ($\cdot$) are the Normal and Uniform distributions, respectively; μ_i is the Normal distribution mean calculated according to equation (9); and σ_i is the activity durations standard deviation according to equation (10). The Normal distribution of d_i in equation (8) becomes Lognormal when exponentiated. The

Uniform distribution in equation (10) acts as the Coefficient of Variation ($CV_i = \sigma_i/\mu_i$) whose values are set to resemble those of real construction projects as measured by Ballesteros-Pérez et al. (2020).

Activity costs (c_i) are also randomly generated for each project and activity i, but now introducing a duration-cost correlation coefficient (δ_i):

$$c_i \sim e^{\left(\mu'_i + \sigma'_i \cdot X_i\right)} \tag{11}$$

$$\mu'_i \sim Normal\left(mean = 10, st.dev. = 1\right) \tag{12}$$

$$\sigma'_i \sim \mu'_i \cdot Uniform\left(lower\ bound = 0.25, upper\ bound = 0.75\right) \tag{13}$$

$$X_i \sim \delta_i \cdot \frac{LN\left(d_i\right) - \mu_i}{\sigma_i} + \sqrt{1-\delta_i^2} \cdot Normal\left(mean = 0, st.dev. = 1\right) \tag{14}$$

$$\delta_i \sim Uniform\left(lower\ bound = 0.0, upper\ bound = 1.0\right) \tag{15}$$

Where most variables are analogous to previous equations (8) to (10), but now some of them with an apostrophe (') indicating that they refer to costs instead of durations. X_i refers to an auxiliary correlated Normally-distributed random variable. This variable helps simplifying the mathematical expressions above and allows generating activity costs which are correlated with the activity durations. Namely, the value of X_i is partially conditioned by the d_i value obtained from equation (8), and the rest is randomly generated according to a standard Normal distribution. Again, the values chosen in equations (13) and (15) are representative of real construction activities (Ballesteros-Pérez, Sanz-Ablanedo, *et al.*, 2020).

Each project calculation involves variables μ_i, μ'_i, σ_i, σ'_i and δ_i being randomly generated for each activity at the outset. These variables are forced to remain constant when 100 stochastic simulations are performed for each project. Hence, 100 d_i and c_i values are stochastically generated in these 100 projects simulations while keeping the other variables constant. Only with this approach is it possible to ensure that each activity duration and cost keeps the same average, dispersion, and correlation values across simulations.

References

Acebes, F., Pereda, M., Poza, D., Pajares, J. and Galán, J.M. (2015), “Stochastic earned value analysis using Monte Carlo simulation and statistical learning techniques”, *International Journal of Project Management*, Elsevier Ltd, Vol. 33 No. 7, pp. 1597–1609.

Anbari, F.T. (2004), “Earned value project management method and extensions”, *IEEE Engineering Management Review*, Vol. 32 No. 3, pp. 97–97.

de Andrade, P.A., Martens, A. and Vanhoucke, M. (2019), “Using real project schedule data to compare earned schedule and earned duration management project time forecasting capabilities”, *Automation in Construction*, Elsevier, Vol. 99 No. July 2018, pp. 68–78.

de Andrade, P.A. and Vanhoucke, M. (2017), “Combining EDM and EVM: ap proposed simplification for project time and cost management”, *Journal of Modern Project Management*, No. September, pp. 94–106.

Bai, L., Wang, Z., Wang, H., Huang, N. and Shi, H. (2020), “Prediction of multiproject resource conflict risk via an artificial neural network”, *Engineering, Construction and Architectural Management*, Vol. In press, available at:https://doi.org/10.1108/ECAM-03-2020-0201.

Ballesteros-Pérez, P., Cerezo-Narváez, A., Otero-Mateo, M., Pastor-Fernández, A. and Vanhoucke, M. (2019), “Performance comparison of activity sensitivity metrics in schedule risk analysis”, *Automation in Construction*, Elsevier, Vol. 106 No. February, p. 102906.

Ballesteros-Pérez, P., Cerezo-Narváez, A., Otero-Mateo, M., Pastor-Fernández, A. and Vanhoucke, M. (2020), “Forecasting the Project Duration Average and Standard Deviation from Deterministic Schedule Information”, *Applied Sciences*, Vol. 10 No. 654, available at:https://doi.org/10.3390/app10020654.

Ballesteros-Pérez, P., Elamrousy, K.M. and González-Cruz, M.C. (2019), “Non-linear time-cost trade-off models of activity crashing: Application to construction scheduling and project compression with fast-tracking”, *Automation in Construction*, Vol. 97 No. January, pp. 229–240.

Ballesteros-Pérez, P., Sanz-Ablanedo, E., Mora-Melià, D., González-Cruz, M.C., Fuentes-Bargues, J.L. and Pellicer, E. (2019), “Earned Schedule min-max: Two new EVM metrics for monitoring and controlling projects”, *Automation in Construction*, Elsevier, Vol. 103 No. April, pp. 279–290.

Ballesteros-Pérez, P., Sanz-Ablanedo, E., Soetanto, R., González-Cruz, M.C., Larsen, G.D. and Cerezo-Narváez, A. (2020), “Duration and cost variability of construction activities: an empirical study”, *Journal of Construction Engineering and Management*, Vol. 146 No. 1, p. 04019093.

Batselier, J. and Vanhoucke, M. (2015a), “Construction and evaluation framework for a real-life project database”, *International Journal of Project Management*, Pergamon, Vol. 33 No. 3, pp. 697–710.

Batselier, J. and Vanhoucke, M. (2015b), “Empirical Evaluation of Earned Value Management Forecasting Accuracy for Time and Cost”, *Journal of Construction Engineering and Management*, Vol. 141 No. 11, p. 05015010.

Batselier, J. and Vanhoucke, M. (2015c), “Evaluation of deterministic state-of-the-art forecasting approaches for project duration based on earned value management”, *International Journal of Project Management*, Elsevier Ltd., Vol. 33 No. 7, pp. 1588–1596.

Batselier, J. and Vanhoucke, M. (2017a), “Project regularity: Development and evaluation of a new project characteristic”, *Journal of Systems Science and Systems Engineering*, Vol. 26 No. 1, pp. 100–120.

Batselier, J. and Vanhoucke, M. (2017b), “Improving project forecast accuracy by integrating earned value management with exponential smoothing and reference class forecasting”, *International Journal of Project Management*, Elsevier Ltd and Association for Project Management and the International Project Management Association, Vol. 35 No. 1, pp. 28–43.

Colin, J. and Vanhoucke, M. (2014), “Setting tolerance limits for statistical project control using earned value management”, *Omega*, Pergamon, Vol. 49, pp. 107–122.

Colin, J. and Vanhoucke, M. (2016), “Empirical Perspective on Activity Durations for Project-Management Simulation Studies”, *Journal of Construction Engineering and Management*, Vol. 142 No. 1, p. 04015047.

Demeulemeester, E., Vanhoucke, M. and Herroelen, W. (2003), “RanGen: A random network generator for activity-on-the-node networks”, *Journal of Scheduling*, Vol. 6 No. 1, pp. 17–38.

Elshaer, R. (2013), “Impact of sensitivity information on the prediction of project’s duration using earned schedule method”, *International Journal of Project Management*, Elsevier Ltd and IPMA, Vol. 31 No. 4, pp. 579–588.

Henderson, K. and Zwikael, O. (2008), “Does project performance stability exist? A re-examination of CPI and evaluation of SPI(t) stability”, *CrossTalk*, Vol. 21 No. 4, pp. 7–13.

Hürol, O., Türkakın, E.M. and Arditi, D. (2020), “Delay analysis in construction projects with no updated work schedules”, *Engineering, Construction and Architectural Management*, Vol. 27 No. 10, pp. 2893–2909.

Jacob, D.S. (2003), “Forecasting project schedule completion with earned value metrics”, *The Measurable News*, Vol. 1, pp. 7–9.

Khafri, A.A., Dawson-Edwards, J., Simpson, R. and Abourizk, S. (2018), “Empirical study of the relationship between cost performance index stability and project cost forecast accuracy in industrial construction projects”, *CSCE General Conference 2018, Held as Part of the Canadian Society for Civil Engineering Annual Conference 2018*, pp. 269–276.

Khamooshi, H. and Abdi, A. (2017), “Project Duration Forecasting Using Earned Duration Management with Exponential Smoothing Techniques”, *Journal of Management in Engineering*, Vol. 33 No. 1, p. 04016032.

Khamooshi, H. and Golafshani, H. (2014), “EDM: Earned Duration Management, a new approach to schedule performance management and measurement”, *International Journal of Project Management*, Elsevier Ltd. APM and IPMA, Vol. 32 No. 6, pp. 1019–1041.

Kim, B.C. (2016), “Probabilistic Evaluation of Cost Performance Stability in Earned Value Management”, *Journal of Management in Engineering*, Vol. 32 No. 1, pp. 1–13.

Kim, D.B., White, E.D., Ritschel, J.D. and Millette, C.A. (2019), “Revisiting reliability of estimates at completion for department of defense contracts”, *Journal of Public Procurement*, Vol. 19 No. 3, pp. 186–200.

Koehler, E., Brown, E. and Haneuse, S.J.-P.A. (2009), “On the Assessment of Monte Carlo Error in Simulation-Based Statistical Analyses”, *The American Statistician*, Taylor & Francis, Vol. 63 No. 2, pp. 155–162.

De Koning, P. and Vanhoucke, M. (2016), “Stability of earned value management do project characteristics influence the stability moment of the cost and schedule performance index”, *Journal of Modern Project Management*, Vol. 4 No. 1, pp. 9–25.

Ladeeda, V.H. and Jeevan, J. (2020), “Probabilistic Evaluation of the Stability in Earned Value Management Forecasting on Topologically Diverse Project Networks”, *Lecture Notes in Civil Engineering*, Vol. 46, pp. 713–724.

Lipke, W. (2003), “Schedule is different”, *The Measurable News*, Vol. Summer, pp. 31–34.

Lipke, W. (2011), “Schedule Adherence and Rework”, *PM World Today*, Vol. 13 No. 7, pp. 1–14.

De Marco, A., Rosso, M. and Narbaev, T. (2016), “Nonlinear cost estimates at completion adjusted with risk contingency”, *Journal of Modern Project Management*, Vol. 4 No. 2, pp. 24–33.

Nadaf, S., Moosavirad, S.H. and Ariafar, S. (2019), “Predicting the project time and costs using EVM based on gray numbers”, *Engineering, Construction and Architectural Management*, Vol. 26 No. 9, pp. 2107–2119.

Narbaev, T. and De Marco, A. (2017), “Earned value and cost contingency management: A framework model for risk-adjusted cost forecasting”, *Journal of Modern Project Management*, Vol. 4 No. 3, pp. 12–19.

Petter, J.L., Ritschel, J.D. and White, E.D. (2015), “Stability properties in Department of Defense contracts: Answering the controversy”, *Journal of Public Procurement*, Vol. 15 No. 3, pp. 341–364.

Sato, K., Inoue, J., Yamazaki, A., Kim, J.-H., Maturilli, M., Dethloff, K., Hudson, S.R., *et al.* (2017), “Improved forecasts of winter weather extremes over midlatitudes with extra Arctic observations”, *Journal of Geophysical Research: Oceans*, Vol. 122 No. 2, pp. 775–787.

Trietsch, D., Mazmanyan, L., Gevorgyan, L. and Baker, K.R. (2012), “Modeling activity times by the Parkinson distribution with a lognormal core: Theory and validation”, *European Journal of Operational Research*, Elsevier B.V., Vol. 216 No. 2, pp. 386–396.

Vanhoucke. (2015), “On the use of schedule risk analysis for project management”, *Journal of Modern Project Management*, Vol. 2 No. 3, pp. 108–117.

Vanhoucke, M. (2010), *Measuring Time - Improving Project Performance Using Earned Value Management*, Springer, International Series in Operations Research & Management Science, available at:https://doi.org/10.1007/978-1-4419-1014-1.

Vanhoucke, M. (2011), "On the dynamic use of project performance and schedule risk information during project tracking", *Omega*, Vol. 39 No. 4, pp. 416–426.

Vanhoucke, M. (2013), *Project Management with Dynamic Scheduling*, Springer Berlin Heidelberg, Berlin, Heidelberg, available at:https://doi.org/10.1007/978-3-642-40.

Vanhoucke, M., Coelho, J., Debels, D., Maenhout, B. and Tavares, L. V. (2008), "An evaluation of the adequacy of project network generators with systematically sampled networks", *European Journal of Operational Research*, North-Holland, Vol. 187 No. 2, pp. 511–524.

Warburton, R.D.H., De Marco, A. and Sciuto, F. (2017), "Earned schedule formulation using nonlinear cost estimates at completion", *Journal of Modern Project Management*, No. May/August, pp. 74–81.

Wauters, M. and Vanhoucke, M. (2015), "Study of the stability of earned value management forecasting", *Journal of Construction Engineering and Management*, Vol. 141 No. 4, p. 04014086.

Table I. Project Duration forecasting methods EAC(t) (all variables and mathematical details can be found in the *Supplemental material*)

ID	*Method*	*EAC(t)*	Mathematical expression	Authors	Brief description
1	EVM	*PV1*	$EAC(t)_{EVM\ PV1}=PD(1-(EV-PV)/BAC)$	(Anbari, 2004)	Methods 1-3 adjust the Planned Project Duration (*PD*) by subtracting a Time Variance factor ($TV=SV/PV_{rate}=(EV-PV)\cdot PD/BAC$) in method 1; or by dividing *PD* by the Schedule Performance Index (*SPI*) or the Schedule Critical Index (*SCI*) in methods 2 and 3, respectively.
2	EVM	*PV2*	$EAC(t)_{EVM\ PV2}=PD/SPI$		
3	EVM	*PV3*	$EAC(t)_{EVM\ PV3}=PD/SCI$		
4	EVM	*ED1*	$EAC(t)_{EVM\ ED1}=MAX(PD,AT)+AT(1-SPI)$	(Jacob, 2003)	Methods 4 to 6 were originally expressed as a function of a metric named Earned Duration ($ED'=AT\cdot SPI$). *ED'* has nothing to do with the *ED* metric from methods 14 and 15. These *EAC(t)* expressions, though, have been worked out and skip the use of *ED'*.
5	EVM	*ED2*	$EAC(t)_{EVM\ ED2}=MAX(PD,AT)/SPI$		
6	EVM	*ED3*	$EAC(t)_{EVM\ ED3}=MAX(PD,AT)/SCI+AT(1-(1/CPI))$		
7	EVM	*ES1*	$EAC(t)_{EVM\ ES1}= AT+PD-ES$	(Lipke, 2003)	Methods 7 to 9 follow the formula *AT*+(*PD-ES*)/*PF*, where *ES* is the Earned Schedule and *PF* is a Performance Factor. *PF* equals 1 in method 7; $PF=SPI(t)=ES/AT$ in method 8, or $PF=SCI(t)=SPI(t)\cdot CPI$ in method 9.
8	EVM	*ES2*	$EAC(t)_{EVM\ ES2}= AT+(PD-ES)/SPI(t)$		
9	EVM	*ES3*	$EAC(t)_{EVM\ ES3}= AT+(PD-ES)/SCI(t)$		
10	EDM	*PV1*	$EAC(t)_{EDM\ PV1}= PD(1-(TED-TPD)/BAC(t))$	(Khamooshi and Golafshani, 2014)	Methods 10 to 15 were proposed under the Earned Duration Management (EDM) framework. In EDM, the activity (planned and actual) costs are replaced by activity (planned and actual) durations. Metric names change, but methods 10-15 are equivalent to 1,2, 4, 5, 7 and 8, respectively. There are no counterparts of *CPI* and *SCI* in EDM, that is why methods 3, 6 and 9 do not have equivalent methods in EDM.
11	EDM	*PV2*	$EAC(t)_{EDM\ PV2}= PD/EDI$		
12	EDM	*ED1*	$EAC(t)_{EDM\ ED1}= MAX(PD,AT)+AT(1-EDI)$		
13	EDM	*ED2*	$EAC(t)_{EDM\ ED2}= MAX(PD,AT)/EDI$		
14	EDM	*ES1*	$EAC(t)_{EDM\ ES1}= AT+PD-ED$		
15	EDM	*ES2*	$EAC(t)_{EDM\ ES2}= AT+(PD-ED)/DPI$		
16	ESM	*ESM1*	$EAC(t)_{ESM\ ES1}= AT+PD-ES(e)$	(Lipke, 2011)	Methods 16 to 18 are similar to methods 7 to 9. However, these replace the Earned Schedule (*ES*) with the effective Earned Schedule (*ES(e)*). *ES(e)* represents the fraction of Earned Value (*EV*) performed according to the original schedule, as measured by the p-factor (Lipke 2004).
17	ESM	*ESM2*	$EAC(t)_{ESM\ ES2}= AT+(PD-ES(e))/SPI(t)(e)$		
18	ESM	*ESM3*	$EAC(t)_{ESM\ ES3}= AT+(PD-ES(e))/SCI(t)(e)$		
19	XSM	*XSM1*	$EAC(t)_{XSM\ XSM1}= AT+(PD-ES)/T_{t,SPI(t)}$	(Khamooshi and Abdi, 2017)	Methods 22 to 24 apply exponential smoothing techniques so that they weight not just immediate (current) metrics progress, but also past information to some extent. Namely, methods 19 and 20 are the counterparts of method 7 with different smoothing factors *T*; whereas method 21 is the equivalent of method 14.
20	XSM	*XSM2*	$EAC(t)_{XSM\ XSM2}= AT+(PD-ES)/(T_{t,AT}/T_{t,ES})$	(Batselier and Vanhoucke, 2017)	
21	XSM	*XSM3*	$EAC(t)_{XSM\ XSM2}= AT+(PD-ES)/T_{t,EDI}$	(Khamooshi and Abdi, 2017)	
22	ES_{min}	*ES1*	$EAC(t)_{ESmin\ ES1}= AT+PD-ES_{min}$	(Ballesteros-Pérez et al., 2019)	Methods 22 to 24 are the counterparts of methods 7 to 9, respectively. However, these methods replaced the Earned Schedule metric (*ES*) with the Earned Schedule min metric (ES_{min}). ES_{min} measures the project progress as a function of its most delayed path.
23	ES_{min}	*ES2*	$EAC(t)_{ESmin\ ES2}= AT+(PD-ES_{min})/SPI(t)_{ESmin}$		
24	ES_{min}	*ES3*	$EAC(t)_{ESmin\ ES3}= AT+(PD-ES_{min})/SCI(t)_{ESmin}$		
25	ES_{max}	*ES1*	$EAC(t)_{ESmax\ ES1}= AT+PD-ES_{max}$		Methods 25 to 27 are equivalent to methods 7 to 9, respectively. However, the former replaced the Earned Schedule metric (*ES*) with the Earned Schedule max metric (ES_{max}) while the latter measure the project progress as a function of its most advanced path.
26	ES_{max}	*ES2*	$EAC(t)_{ESmax\ ES2}= AT+(PD-ES_{max})/SPI(t)_{ESmax}$		
27	ES_{max}	*ES3*	$EAC(t)_{ESmax\ ES3}= AT+(PD-ES_{max})/SCI(t)_{ESmax}$		

Table II. Summary of the relevant existing research on deterministic EVM project duration forecasting methods on accuracy and stability

Study (Reference)	**Accuracy** MAPE	MSE	**Stability** Ad-hoc indicator	**Project dataset** Nº simulated projects	Nº real projects	**EAC(t) methods** Methods compared	Top performers
(Vanhoucke 2010)	✓		-	4100	-	1-9	7
(Batselier and Vanhoucke 2015b)	✓		-	-	23	1-9	7
(Batselier and Vanhoucke 2015c)	✓		-	-	23	7, 15 & 16	7
(Khamooshi and Abdi 2017)		✓*		-	19	10-13, 22 & 24	22 & 24
(Batselier and Vanhoucke 2017b)	✓		-	-	23	9 & 20	9
(de Andrade and Vanhoucke 2017)	✓		-	-	14	7 & 15	7
(Ballesteros-Pérez et al. 2019)	✓		-	4100	23	1-22 & 25	22
(Wauters and Vanhoucke 2015)	✓		Mean lags indicator	90	2	1-9**	7
(Batselier and Vanhoucke 2017a)	✓		Regularity index	Not specified	23	1-9	7
(de Andrade et al. 2019)	✓		Regularity index	-	57	7-9, 14 & 15	7 & 14
This study	✓	✓	R, ρ and τ correlations	4100	-***	1-27	22

* Khamooshi and Abdi (2017) actually used the Root Mean Squared Error (RMSE) which is quite an unusual error metric.

** Wauters and Vanhoucke also compared the performance of one non deterministic EVM method proposed by Elshaer (2013).

*** Due to the probabilistic approach adopted in this study to measure the EAC(t) methods stability, it is not possible to resort to real projects.

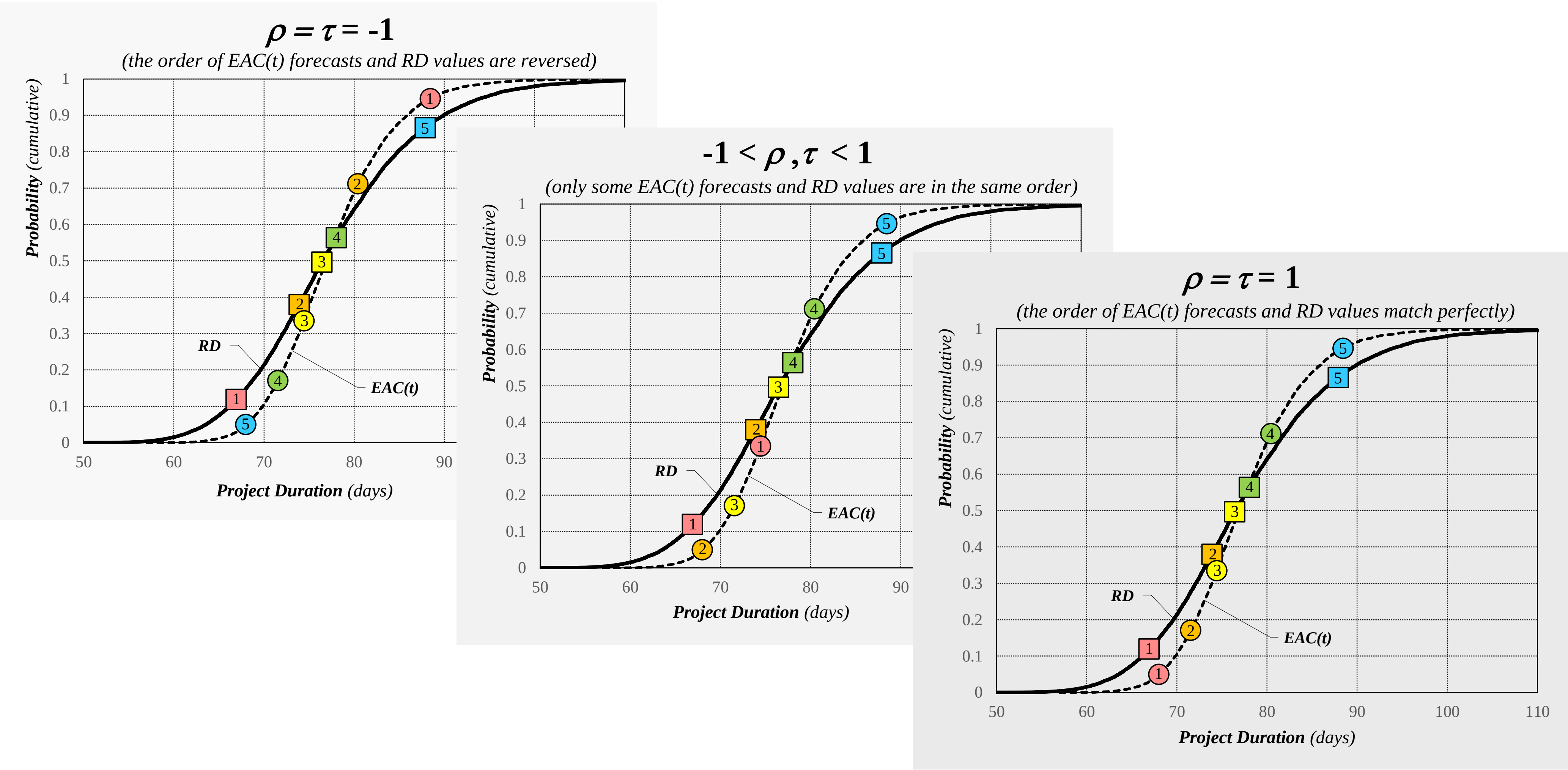

Figure 1: Interpretation of EAC(t) metrics stability through Spearman's ρ and Kendall's τ parametric tests of rank correlation.

Average Pearson's R

ID	Method	EAC(t)	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%	
		Track. per. (%RD) ▶	0.00	0.29	0.40	0.50	0.58	0.66	0.73	0.80	0.86	0.90	0.92	Avg.
1	EVM	PV1	0.00	0.25	0.39	0.50	0.58	0.64	0.69	0.71	0.72	0.70	0.63	0.58
2	EVM	PV2	0.00	0.27	0.37	0.46	0.53	0.59	0.65	0.69	0.71	0.70	0.63	0.56
3	EVM	PV3	0.00	0.20	0.28	0.35	0.41	0.45	0.48	0.49	0.47	0.42	0.35	0.39
4	EVM	ED1	0.00	0.25	0.39	0.51	0.60	0.68	0.75	0.82	0.88	0.94	0.98	0.68
5	EVM	ED2	0.00	0.27	0.37	0.46	0.54	0.61	0.70	0.79	0.87	0.94	0.98	0.65
6	EVM	ED3	0.00	0.21	0.29	0.38	0.46	0.53	0.63	0.73	0.84	0.93	0.98	0.60
7	EVM	ES1	0.00	0.28	0.44	0.56	0.65	0.73	0.80	0.86	0.92	0.97	1.00	0.72
8	EVM	ES2	0.00	0.31	0.40	0.50	0.59	0.67	0.75	0.83	0.90	0.96	1.00	0.69
9	EVM	ES3	0.00	0.23	0.33	0.43	0.52	0.61	0.70	0.79	0.87	0.95	1.00	0.64
10	EDM	PV1	0.00	0.31	0.45	0.55	0.62	0.68	0.73	0.76	0.76	0.74	0.66	0.62
11	EDM	PV2	0.00	0.34	0.42	0.50	0.57	0.63	0.69	0.73	0.75	0.74	0.66	0.60
12	EDM	ED1	0.00	0.31	0.45	0.56	0.64	0.71	0.79	0.85	0.91	0.96	0.99	0.72
13	EDM	ED2	0.00	0.34	0.42	0.50	0.58	0.65	0.74	0.82	0.90	0.96	0.99	0.69
14	EDM	ES1	0.00	0.31	0.46	0.57	0.66	0.74	0.81	0.87	0.93	0.97	1.00	0.73
15	EDM	ES2	0.00	0.35	0.43	0.52	0.60	0.68	0.76	0.83	0.90	0.96	1.00	0.70
16	ESM	ES1	0.00	0.29	0.44	0.55	0.64	0.72	0.80	0.86	0.92	0.97	1.00	0.72
17	ESM	ES2	0.00	0.27	0.38	0.48	0.57	0.66	0.74	0.82	0.89	0.96	1.00	0.68
18	ESM	ES3	0.00	0.21	0.31	0.41	0.51	0.60	0.69	0.78	0.87	0.95	1.00	0.63
19	XSM	XSM1	0.00	0.26	0.41	0.51	0.59	0.67	0.75	0.82	0.89	0.96	1.00	0.69
20	XSM	XSM2	0.00	0.30	0.45	0.55	0.63	0.70	0.76	0.83	0.90	0.96	1.00	0.71
21	XSM	XSM3	0.00	0.29	0.44	0.53	0.61	0.68	0.76	0.84	0.91	0.97	1.00	0.70
22	ES$_{min}$	ES1	0.00	0.34	0.47	0.58	0.67	0.75	0.82	0.88	0.93	0.97	1.00	0.74
23	ES$_{min}$	ES2	0.00	0.32	0.42	0.52	0.60	0.69	0.77	0.84	0.91	0.96	1.00	0.70
24	ES$_{min}$	ES3	0.00	0.28	0.37	0.45	0.54	0.63	0.72	0.80	0.89	0.96	1.00	0.66
25	ES$_{max}$	ES1	0.00	0.24	0.40	0.54	0.64	0.73	0.80	0.87	0.92	0.97	1.00	0.71
26	ES$_{max}$	ES2	0.00	0.38	0.45	0.54	0.62	0.70	0.77	0.84	0.91	0.96	1.00	0.72
27	ES$_{max}$	ES3	0.00	0.30	0.37	0.47	0.55	0.64	0.72	0.81	0.89	0.96	1.00	0.67

Average Spearman's ρ

ID	Method	EAC(t)	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%	
			0.00	0.29	0.41	0.50	0.58	0.66	0.73	0.80	0.86	0.91	0.93	Avg.
1	EVM	PV1	0.00	0.27	0.40	0.50	0.58	0.65	0.70	0.75	0.77	0.77	0.71	0.61
2	EVM	PV2	0.00	0.31	0.40	0.49	0.57	0.63	0.69	0.74	0.77	0.77	0.71	0.61
3	EVM	PV3	0.00	0.23	0.30	0.37	0.42	0.46	0.49	0.51	0.50	0.46	0.37	0.41
4	EVM	ED1	0.00	0.28	0.40	0.50	0.59	0.66	0.73	0.80	0.87	0.95	0.99	0.68
5	EVM	ED2	0.00	0.31	0.40	0.49	0.57	0.64	0.71	0.78	0.87	0.95	0.99	0.67
6	EVM	ED3	0.00	0.24	0.33	0.41	0.49	0.56	0.63	0.72	0.81	0.92	0.98	0.61
7	EVM	ES1	0.00	0.27	0.41	0.52	0.61	0.69	0.76	0.83	0.90	0.96	1.00	0.70
8	EVM	ES2	0.00	0.31	0.42	0.51	0.60	0.67	0.75	0.82	0.89	0.95	0.99	0.69
9	EVM	ES3	0.00	0.26	0.36	0.46	0.54	0.63	0.71	0.79	0.87	0.95	0.99	0.66
10	EDM	PV1	0.00	0.30	0.43	0.52	0.61	0.67	0.73	0.78	0.80	0.80	0.72	0.64
11	EDM	PV2	0.00	0.34	0.43	0.52	0.60	0.66	0.72	0.77	0.80	0.80	0.72	0.64
12	EDM	ED1	0.00	0.30	0.42	0.52	0.61	0.68	0.75	0.83	0.90	0.96	1.00	0.70
13	EDM	ED2	0.00	0.34	0.43	0.52	0.60	0.66	0.74	0.81	0.89	0.96	1.00	0.69
14	EDM	ES1	0.00	0.30	0.43	0.53	0.62	0.70	0.77	0.84	0.91	0.96	1.00	0.71
15	EDM	ES2	0.00	0.35	0.44	0.53	0.61	0.69	0.76	0.83	0.90	0.96	1.00	0.71
16	ESM	ES1	0.00	0.28	0.41	0.51	0.60	0.68	0.76	0.83	0.90	0.96	0.99	0.69
17	ESM	ES2	0.00	0.27	0.39	0.49	0.58	0.66	0.74	0.82	0.89	0.95	0.99	0.68
18	ESM	ES3	0.00	0.24	0.35	0.44	0.53	0.62	0.70	0.79	0.87	0.95	0.99	0.65
19	XSM	XSM1	0.00	0.25	0.40	0.51	0.60	0.67	0.75	0.82	0.89	0.95	0.99	0.68
20	XSM	XSM2	0.00	0.28	0.42	0.52	0.61	0.68	0.75	0.82	0.89	0.95	0.99	0.69
21	XSM	XSM3	0.00	0.27	0.42	0.52	0.61	0.68	0.76	0.83	0.90	0.96	1.00	0.69
22	ES$_{min}$	ES1	0.00	0.32	0.45	0.55	0.64	0.71	0.79	0.85	0.91	0.96	1.00	0.72
23	ES$_{min}$	ES2	0.00	0.34	0.45	0.54	0.63	0.70	0.77	0.84	0.91	0.96	1.00	0.71
24	ES$_{min}$	ES3	0.00	0.31	0.40	0.49	0.58	0.66	0.74	0.81	0.89	0.95	1.00	0.68
25	ES$_{max}$	ES1	0.00	0.26	0.38	0.50	0.60	0.69	0.77	0.84	0.90	0.96	1.00	0.69
26	ES$_{max}$	ES2	0.00	0.38	0.46	0.55	0.63	0.70	0.77	0.83	0.90	0.96	1.00	0.72
27	ES$_{max}$	ES3	0.00	0.32	0.41	0.50	0.58	0.66	0.73	0.81	0.89	0.95	1.00	0.68

Average Kendall's τ

ID	Method	EAC(t)	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%	
			0.00	0.09	0.20	0.28	0.35	0.41	0.47	0.53	0.59	0.65	0.67	Avg.
1	EVM	PV1	0.00	-0.02	0.13	0.23	0.30	0.36	0.40	0.43	0.43	0.36	0.10	0.27
2	EVM	PV2	0.00	0.14	0.21	0.28	0.33	0.38	0.42	0.44	0.44	0.37	0.10	0.31
3	EVM	PV3	0.00	0.11	0.17	0.21	0.25	0.28	0.30	0.31	0.30	0.26	0.20	0.24
4	EVM	ED1	0.00	-0.08	0.09	0.20	0.28	0.34	0.41	0.47	0.55	0.65	0.74	0.36
5	EVM	ED2	0.00	0.14	0.21	0.28	0.34	0.38	0.43	0.49	0.56	0.66	0.75	0.42
6	EVM	ED3	0.00	0.12	0.18	0.24	0.29	0.34	0.39	0.46	0.54	0.65	0.75	0.40
7	EVM	ES1	0.00	-0.03	0.14	0.25	0.34	0.42	0.49	0.56	0.64	0.73	0.80	0.43
8	EVM	ES2	0.00	0.16	0.24	0.31	0.38	0.44	0.50	0.57	0.64	0.73	0.80	0.48
9	EVM	ES3	0.00	0.14	0.21	0.28	0.34	0.40	0.47	0.54	0.63	0.72	0.80	0.45
10	EDM	PV1	0.00	0.01	0.16	0.25	0.33	0.39	0.44	0.47	0.48	0.43	0.18	0.31
11	EDM	PV2	0.00	0.17	0.24	0.30	0.36	0.40	0.45	0.48	0.49	0.43	0.18	0.35
12	EDM	ED1	0.00	-0.06	0.11	0.22	0.31	0.37	0.44	0.51	0.59	0.69	0.77	0.39
13	EDM	ED2	0.00	0.17	0.24	0.30	0.36	0.41	0.46	0.52	0.60	0.69	0.77	0.45
14	EDM	ES1	0.00	-0.03	0.15	0.26	0.35	0.42	0.50	0.57	0.65	0.73	0.81	0.44
15	EDM	ES2	0.00	0.18	0.25	0.32	0.38	0.45	0.51	0.58	0.65	0.73	0.81	0.49
16	ESM	ES1	0.00	-0.03	0.14	0.25	0.33	0.41	0.49	0.56	0.64	0.72	0.80	0.43
17	ESM	ES2	0.00	0.14	0.22	0.30	0.37	0.43	0.50	0.57	0.64	0.73	0.80	0.47
18	ESM	ES3	0.00	0.12	0.20	0.27	0.33	0.40	0.47	0.54	0.62	0.72	0.80	0.45
19	XSM	XSM1	0.00	0.09	0.22	0.30	0.37	0.43	0.50	0.57	0.64	0.72	0.80	0.46
20	XSM	XSM2	0.00	0.06	0.20	0.29	0.37	0.43	0.50	0.57	0.64	0.73	0.80	0.46
21	XSM	XSM3	0.00	0.10	0.22	0.30	0.37	0.43	0.50	0.57	0.64	0.73	0.81	0.47
22	ES$_{min}$	ES1	0.00	0.06	0.19	0.29	0.37	0.44	0.51	0.58	0.66	0.74	0.81	0.46
23	ES$_{min}$	ES2	0.00	0.20	0.27	0.34	0.40	0.46	0.53	0.59	0.66	0.74	0.81	0.50
24	ES$_{min}$	ES3	0.00	0.18	0.25	0.31	0.37	0.43	0.50	0.57	0.64	0.73	0.81	0.48
25	ES$_{max}$	ES1	0.00	-0.01	0.14	0.25	0.34	0.42	0.49	0.57	0.65	0.73	0.81	0.44
26	ES$_{max}$	ES2	0.00	0.19	0.26	0.33	0.40	0.46	0.52	0.58	0.65	0.73	0.81	0.49
27	ES$_{max}$	ES3	0.00	0.17	0.23	0.30	0.36	0.42	0.49	0.56	0.64	0.73	0.81	0.47

Average MSEs

ID	Method	EAC(t)	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%	
		Track. per. (%RD) ▶	1.35	42.37	24.04	6.11	3.25	1.87	1.13	0.75	0.49	0.32	0.25	Avg.
1	EVM	PV1	1.35	1.33	1.21	1.07	0.94	0.83	0.75	0.71	0.74	0.85	1.08	0.95
2	EVM	PV2	1.35	117.7	48.29	14.13	6.96	3.07	1.65	1.17	0.92	0.89	1.08	19.58
3	EVM	PV3	1.35	351.0	204.8	31.47	14.03	7.35	4.16	3.13	2.45	2.04	1.94	62.24
4	EVM	ED1	1.35	1.47	1.30	1.17	1.02	0.87	0.73	0.62	0.48	0.32	0.18	0.82
5	EVM	ED2	1.35	117.7	48.29	14.13	6.96	3.08	1.64	1.09	0.68	0.34	0.13	19.40
6	EVM	ED3	1.35	348.1	201.0	28.82	11.76	5.33	2.33	1.42	0.81	0.37	0.13	60.00
7	EVM	ES1	1.35	1.30	1.14	0.97	0.81	0.65	0.49	0.34	0.21	0.09	0.01	0.60
8	EVM	ES2	1.35	6.84	5.08	3.37	2.26	1.49	0.97	0.60	0.33	0.13	0.01	2.11
9	EVM	ES3	1.35	32.80	21.80	8.67	4.66	2.76	1.60	0.91	0.46	0.16	0.01	7.38
10	EDM	PV1	1.35	1.30	1.14	0.99	0.84	0.72	0.63	0.59	0.61	0.75	1.02	0.86
11	EDM	PV2	1.35	6.70	4.82	3.22	2.10	1.37	0.91	0.68	0.61	0.73	1.02	2.22
12	EDM	ED1	1.35	1.31	1.17	1.02	0.87	0.73	0.59	0.46	0.31	0.17	0.09	0.67
13	EDM	ED2	1.35	6.70	4.82	3.22	2.10	1.36	0.88	0.58	0.34	0.16	0.07	2.02
14	EDM	ES1	1.35	1.29	1.13	0.96	0.80	0.64	0.48	0.33	0.19	0.08	0.01	0.59
15	EDM	ES2	1.35	6.38	4.84	3.23	2.15	1.42	0.91	0.56	0.30	0.11	0.01	1.99
16	ESM	ES1	1.35	1.27	1.12	0.96	0.80	0.64	0.49	0.35	0.21	0.09	0.01	0.59
17	ESM	ES2	1.35	7.95	5.68	3.70	2.46	1.61	1.03	0.64	0.34	0.13	0.01	2.36
18	ESM	ES3	1.35	35.76	23.26	9.34	4.99	2.93	1.68	0.96	0.48	0.17	0.01	7.96
19	XSM	XSM1	1.35	4.19	4.25	3.16	2.18	1.46	0.96	0.60	0.33	0.13	0.01	1.73
20	XSM	XSM2	1.35	1.50	1.55	1.52	1.38	1.13	0.86	0.59	0.36	0.13	0.01	0.90
21	XSM	XSM3	1.35	3.97	3.94	2.89	1.95	1.27	0.80	0.48	0.25	0.09	0.01	1.56
22	ES$_{min}$	ES1	1.35	1.24	1.07	0.90	0.74	0.59	0.44	0.30	0.17	0.07	0.00	0.55
23	ES$_{min}$	ES2	1.35	9.96	6.48	4.07	2.58	1.61	0.98	0.57	0.29	0.10	0.01	2.66
24	ES$_{min}$	ES3	1.35	36.93	23.74	9.63	5.05	2.88	1.62	0.88	0.42	0.13	0.01	8.13
25	ES$_{max}$	ES1	1.35	1.46	1.33	1.15	0.96	0.78	0.59	0.41	0.24	0.09	0.01	0.70
26	ES$_{max}$	ES2	1.35	7.45	5.14	3.34	2.17	1.40	0.91	0.57	0.31	0.11	0.01	2.14
27	ES$_{max}$	ES3	1.35	30.55	20.58	7.99	4.22	2.49	1.45	0.83	0.41	0.14	0.01	6.87

Average MAPEs

ID	Method	EAC(t)	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%	
			0.14	0.25	0.21	0.18	0.15	0.13	0.11	0.09	0.08	0.06	0.03	Avg.
1	EVM	PV1	0.14	0.14	0.13	0.13	0.12	0.11	0.10	0.10	0.10	0.10	0.12	0.12
2	EVM	PV2	0.14	0.29	0.24	0.20	0.17	0.15	0.13	0.11	0.10	0.11	0.12	0.16
3	EVM	PV3	0.14	0.46	0.37	0.30	0.26	0.23	0.20	0.18	0.17	0.16	0.16	0.25
4	EVM	ED1	0.14	0.14	0.14	0.13	0.12	0.11	0.10	0.10	0.08	0.07	0.04	0.10
5	EVM	ED2	0.14	0.29	0.24	0.20	0.17	0.15	0.13	0.11	0.09	0.07	0.03	0.15
6	EVM	ED3	0.14	0.44	0.34	0.26	0.22	0.18	0.14	0.12	0.09	0.07	0.03	0.19
7	EVM	ES1	0.14	0.14	0.13	0.12	0.11	0.10	0.09	0.07	0.06	0.04	0.01	0.09
8	EVM	ES2	0.14	0.25	0.21	0.18	0.15	0.13	0.11	0.09	0.07	0.04	0.01	0.12
9	EVM	ES3	0.14	0.38	0.30	0.24	0.20	0.16	0.13	0.10	0.07	0.04	0.01	0.17
10	EDM	PV1	0.14	0.14	0.13	0.12	0.11	0.10	0.10	0.09	0.09	0.10	0.11	0.11
11	EDM	PV2	0.14	0.22	0.19	0.16	0.14	0.12	0.11	0.10	0.09	0.09	0.11	0.13
12	EDM	ED1	0.14	0.14	0.13	0.12	0.11	0.11	0.10	0.09	0.07	0.05	0.03	0.10
13	EDM	ED2	0.14	0.22	0.19	0.16	0.14	0.12	0.11	0.09	0.07	0.05	0.03	0.12
14	EDM	ES1	0.14	0.14	0.13	0.12	0.11	0.10	0.09	0.07	0.05	0.03	0.01	0.08
15	EDM	ES2	0.14	0.24	0.20	0.17	0.15	0.13	0.11	0.09	0.06	0.04	0.01	0.12
16	ESM	ES1	0.14	0.14	0.13	0.12	0.11	0.10	0.09	0.07	0.06	0.04	0.01	0.09
17	ESM	ES2	0.14	0.26	0.22	0.19	0.16	0.14	0.11	0.09	0.07	0.04	0.01	0.13
18	ESM	ES3	0.14	0.39	0.31	0.25	0.21	0.17	0.14	0.10	0.08	0.04	0.01	0.17
19	XSM	XSM1	0.14	0.21	0.20	0.17	0.15	0.13	0.11	0.09	0.07	0.04	0.01	0.12
20	XSM	XSM2	0.14	0.15	0.14	0.14	0.13	0.12	0.10	0.08	0.06	0.04	0.01	0.10
21	XSM	XSM3	0.14	0.20	0.18	0.16	0.14	0.12	0.10	0.08	0.06	0.04	0.01	0.11
22	ES$_{min}$	ES1	0.14	0.13	0.12	0.11	0.10	0.09	0.08	0.07	0.05	0.03	0.01	0.08
23	ES$_{min}$	ES2	0.14	0.30	0.24	0.20	0.16	0.13	0.11	0.08	0.06	0.04	0.01	0.13
24	ES$_{min}$	ES3	0.14	0.42	0.32	0.26	0.21	0.16	0.13	0.10	0.07	0.04	0.01	0.17
25	ES$_{max}$	ES1	0.14	0.15	0.14	0.13	0.12	0.11	0.10	0.08	0.06	0.04	0.01	0.09
26	ES$_{max}$	ES2	0.14	0.33	0.25	0.20	0.17	0.14	0.11	0.09	0.07	0.04	0.01	0.14
27	ES$_{max}$	ES3	0.14	0.43	0.33	0.26	0.21	0.17	0.13	0.10	0.08	0.04	0.01	0.18

Summary *(top performers for both stability and accuracy)*

		Avg. values ▶	*Stability*				*Accuracy*			*Both*
ID	Method	EAC(t)	*R*	ρ	τ	*(Best)*	*MSE*	*MAPE*	*(Best)*	*(Best)*
1	EVM	PV1	0.58	0.61	0.27		0.95	0.12	*	
2	EVM	PV2	0.56	0.61	0.31		19.58	0.16		
3	EVM	PV3	0.39	0.41	0.24		62.24	0.25		
4	EVM	ED1	0.68	0.68	0.36		0.82	0.10	**	
5	EVM	ED2	0.65	0.67	0.42		19.40	0.15		
6	EVM	ED3	0.60	0.61	0.40		60.00	0.19		
7	EVM	ES1	0.72	0.70	0.43	*	0.60	0.09	**	*
8	EVM	ES2	0.69	0.69	0.48	*	2.11	0.12		
9	EVM	ES3	0.64	0.66	0.45		7.38	0.17		
10	EDM	PV1	0.62	0.64	0.31		0.86	0.11	*	
11	EDM	PV2	0.60	0.64	0.35		2.22	0.13		
12	EDM	ED1	0.72	0.70	0.39	*	0.67	0.10	**	*
13	EDM	ED2	0.69	0.69	0.45	*	2.02	0.12		
14	EDM	ES1	0.73	0.71	0.44	**	0.59	0.08	**	**
15	EDM	ES2	0.70	0.71	0.49	**	1.99	0.12		
16	ESM	ES1	0.72	0.69	0.43	*	0.59	0.09	**	*
17	ESM	ES2	0.68	0.68	0.47	*	2.36	0.13		
18	ESM	ES3	0.63	0.65	0.45		7.96	0.17		
19	XSM	XSM1	0.69	0.68	0.46	*	1.73	0.12		
20	XSM	XSM2	0.71	0.69	0.46	*	0.90	0.10	**	*
21	XSM	XSM3	0.70	0.69	0.47	*	1.56	0.11	*	
22	ES$_{min}$	ES1	0.74	0.72	0.46	**	0.55	0.08	**	**
23	ES$_{min}$	ES2	0.70	0.71	0.50	**	2.66	0.13		
24	ES$_{min}$	ES3	0.66	0.68	0.48	*	8.13	0.17		
25	ES$_{max}$	ES1	0.71	0.69	0.44	*	0.70	0.09	**	*
26	ES$_{max}$	ES2	0.72	0.72	0.49	**	2.14	0.14		
27	ES$_{max}$	ES3	0.67	0.68	0.47	*	6.87	0.18		

Figure 2: EAC(t) estimates' stability and accuracy by percentage of project progress (as %RD) (*: Good performers, **: Top performers).

Average Pearson's R

ID	Method	EAC(t)	7%	17%	28%	38%	48%	59%	69%	79%	90%	
Serial-Par. (%SP) ►			**0.30**	**0.41**	**0.48**	**0.57**	**0.66**	**0.74**	**0.82**	**0.88**	**0.93**	**Avg.**
1	EVM	**PV1**	0.23	0.39	0.47	0.57	0.67	0.73	0.80	0.85	0.86	**0.62**
2	EVM	**PV2**	0.39	0.38	0.42	0.53	0.63	0.71	0.79	0.84	0.86	**0.62**
3	EVM	**PV3**	0.27	0.27	0.32	0.41	0.48	0.55	0.58	0.59	0.57	**0.45**
4	EVM	**ED1**	0.23	0.40	0.48	0.58	0.68	0.73	0.81	0.87	0.93	**0.64**
5	EVM	**ED2**	0.39	0.38	0.42	0.53	0.63	0.71	0.80	0.87	0.93	**0.63**
6	EVM	**ED3**	0.29	0.30	0.35	0.46	0.55	0.65	0.74	0.83	0.92	**0.56**
7	EVM	**ES1**	0.15	0.44	0.53	0.62	0.72	0.79	0.86	0.91	0.96	**0.67**
8	EVM	**ES2**	0.36	0.41	0.46	0.57	0.68	0.76	0.84	0.90	0.96	**0.66**
9	EVM	**ES3**	0.28	0.34	0.42	0.51	0.61	0.71	0.79	0.87	0.95	**0.61**
10	EDM	**PV1**	0.28	0.47	0.52	0.61	0.71	0.77	0.84	0.89	0.89	**0.66**
11	EDM	**PV2**	0.51	0.45	0.47	0.57	0.67	0.75	0.83	0.88	0.89	**0.67**
12	EDM	**ED1**	0.28	0.48	0.53	0.62	0.71	0.77	0.85	0.91	0.96	**0.68**
13	EDM	**ED2**	0.51	0.45	0.47	0.57	0.67	0.75	0.84	0.90	0.96	**0.68**
14	EDM	**ES1**	0.21	0.49	0.56	0.63	0.73	0.79	0.86	0.91	0.96	**0.68**
15	EDM	**ES2**	0.47	0.47	0.50	0.58	0.68	0.76	0.84	0.91	0.96	**0.69**
16	ESM	**ES1**	0.23	0.44	0.52	0.62	0.72	0.79	0.86	0.91	0.96	**0.67**
17	ESM	**ES2**	0.26	0.36	0.43	0.55	0.67	0.76	0.84	0.90	0.96	**0.64**
18	ESM	**ES3**	0.22	0.30	0.39	0.50	0.61	0.70	0.79	0.87	0.95	**0.59**
19	XSM	**XSM1**	0.14	0.43	0.49	0.56	0.67	0.75	0.84	0.90	0.96	**0.64**
20	XSM	**XSM2**	0.18	0.45	0.53	0.61	0.69	0.76	0.83	0.89	0.95	**0.65**
21	XSM	**XSM3**	0.20	0.48	0.52	0.58	0.67	0.76	0.84	0.90	0.96	**0.66**
22	ES_{min}	**ES1**	0.38	0.50	0.58	0.65	0.74	0.80	0.86	0.91	0.96	**0.71**
23	ES_{min}	**ES2**	0.40	0.44	0.50	0.59	0.69	0.77	0.85	0.91	0.96	**0.68**
24	ES_{min}	**ES3**	0.39	0.41	0.46	0.53	0.63	0.71	0.79	0.87	0.95	**0.64**
25	ES_{max}	**ES1**	-0.06	0.33	0.54	0.62	0.72	0.78	0.85	0.91	0.96	**0.63**
26	ES_{max}	**ES2**	0.53	0.49	0.55	0.59	0.69	0.76	0.84	0.91	0.96	**0.70**
27	ES_{max}	**ES3**	0.49	0.42	0.49	0.53	0.62	0.71	0.79	0.87	0.95	**0.65**

Average Spearman's r

ID	Method	EAC(t)	7%	17%	28%	38%	48%	59%	69%	79%	90%	
Serial-Par. (%SP) ►			**0.30**	**0.40**	**0.47**	**0.58**	**0.67**	**0.74**	**0.81**	**0.87**	**0.93**	**Avg.**
1	EVM	**PV1**	0.22	0.38	0.45	0.57	0.67	0.73	0.80	0.87	0.91	**0.62**
2	EVM	**PV2**	0.40	0.38	0.44	0.57	0.66	0.73	0.80	0.87	0.91	**0.64**
3	EVM	**PV3**	0.28	0.29	0.32	0.43	0.49	0.54	0.57	0.60	0.61	**0.46**
4	EVM	**ED1**	0.22	0.38	0.46	0.57	0.67	0.73	0.80	0.87	0.93	**0.63**
5	EVM	**ED2**	0.40	0.38	0.44	0.57	0.66	0.73	0.80	0.87	0.94	**0.64**
6	EVM	**ED3**	0.31	0.32	0.36	0.50	0.58	0.67	0.74	0.82	0.91	**0.58**
7	EVM	**ES1**	0.15	0.40	0.49	0.60	0.69	0.76	0.84	0.90	0.95	**0.64**
8	EVM	**ES2**	0.34	0.40	0.47	0.59	0.69	0.77	0.84	0.90	0.95	**0.66**
9	EVM	**ES3**	0.28	0.35	0.43	0.55	0.64	0.72	0.79	0.87	0.94	**0.62**
10	EDM	**PV1**	0.24	0.44	0.49	0.60	0.68	0.75	0.83	0.89	0.93	**0.65**
11	EDM	**PV2**	0.51	0.44	0.48	0.59	0.68	0.75	0.83	0.89	0.93	**0.68**
12	EDM	**ED1**	0.24	0.44	0.50	0.60	0.68	0.75	0.83	0.89	0.95	**0.65**
13	EDM	**ED2**	0.51	0.44	0.48	0.59	0.68	0.75	0.83	0.89	0.95	**0.68**
14	EDM	**ES1**	0.17	0.45	0.52	0.60	0.69	0.76	0.84	0.90	0.95	**0.65**
15	EDM	**ES2**	0.46	0.45	0.50	0.60	0.69	0.77	0.84	0.90	0.95	**0.69**
16	ESM	**ES1**	0.21	0.40	0.48	0.59	0.69	0.76	0.83	0.90	0.95	**0.65**
17	ESM	**ES2**	0.26	0.35	0.44	0.58	0.68	0.76	0.84	0.90	0.95	**0.64**
18	ESM	**ES3**	0.23	0.32	0.40	0.54	0.64	0.72	0.79	0.87	0.94	**0.61**
19	XSM	**XSM1**	0.13	0.39	0.48	0.59	0.68	0.76	0.83	0.89	0.95	**0.63**
20	XSM	**XSM2**	0.16	0.40	0.49	0.59	0.69	0.76	0.83	0.89	0.95	**0.64**
21	XSM	**XSM3**	0.17	0.45	0.51	0.59	0.68	0.76	0.83	0.89	0.95	**0.65**
22	ES_{min}	**ES1**	0.35	0.47	0.55	0.62	0.71	0.77	0.84	0.90	0.95	**0.68**
23	ES_{min}	**ES2**	0.37	0.46	0.54	0.62	0.70	0.77	0.84	0.90	0.95	**0.68**
24	ES_{min}	**ES3**	0.37	0.43	0.50	0.58	0.65	0.73	0.80	0.87	0.94	**0.65**
25	ES_{max}	**ES1**	0.10	0.31	0.51	0.60	0.69	0.76	0.83	0.90	0.95	**0.63**
26	ES_{max}	**ES2**	0.50	0.49	0.56	0.62	0.70	0.77	0.84	0.90	0.95	**0.70**
27	ES_{max}	**ES3**	0.48	0.44	0.51	0.58	0.65	0.73	0.80	0.87	0.94	**0.67**

Average Kendall's t

ID	Method	EAC(t)	7%	17%	28%	38%	48%	59%	69%	79%	90%	
Serial-Par. (%SP) ►			**-0.16**	**0.13**	**0.24**	**0.35**	**0.44**	**0.51**	**0.58**	**0.66**	**0.74**	**Avg.**
1	EVM	**PV1**	-0.33	0.03	0.17	0.31	0.41	0.48	0.56	0.64	0.68	**0.33**
2	EVM	**PV2**	0.05	0.15	0.22	0.35	0.44	0.50	0.57	0.64	0.69	**0.40**
3	EVM	**PV3**	0.03	0.13	0.17	0.27	0.32	0.36	0.39	0.41	0.42	**0.28**
4	EVM	**ED1**	-0.51	-0.05	0.12	0.29	0.40	0.48	0.55	0.64	0.73	**0.29**
5	EVM	**ED2**	0.05	0.15	0.22	0.35	0.44	0.50	0.57	0.65	0.74	**0.41**
6	EVM	**ED3**	0.04	0.14	0.19	0.31	0.38	0.45	0.52	0.61	0.71	**0.37**
7	EVM	**ES1**	-0.53	0.03	0.22	0.34	0.44	0.52	0.60	0.68	0.77	**0.34**
8	EVM	**ES2**	0.01	0.19	0.28	0.38	0.46	0.54	0.61	0.69	0.77	**0.44**
9	EVM	**ES3**	0.02	0.17	0.25	0.35	0.43	0.50	0.57	0.66	0.76	**0.41**
10	EDM	**PV1**	-0.39	0.08	0.21	0.34	0.43	0.51	0.59	0.67	0.72	**0.35**
11	EDM	**PV2**	0.08	0.19	0.26	0.37	0.45	0.52	0.60	0.67	0.73	**0.43**
12	EDM	**ED1**	-0.57	-0.02	0.17	0.32	0.42	0.50	0.59	0.67	0.77	**0.32**
13	EDM	**ED2**	0.08	0.19	0.26	0.37	0.45	0.52	0.60	0.68	0.77	**0.44**
14	EDM	**ES1**	-0.59	0.04	0.24	0.34	0.45	0.52	0.60	0.68	0.77	**0.34**
15	EDM	**ES2**	0.05	0.22	0.30	0.39	0.47	0.54	0.61	0.69	0.77	**0.45**
16	ESM	**ES1**	-0.51	0.03	0.21	0.34	0.44	0.52	0.60	0.68	0.77	**0.34**
17	ESM	**ES2**	-0.03	0.16	0.26	0.37	0.46	0.54	0.61	0.69	0.77	**0.43**
18	ESM	**ES3**	0.00	0.15	0.24	0.34	0.43	0.50	0.57	0.66	0.76	**0.41**
19	XSM	**XSM1**	-0.28	0.15	0.27	0.37	0.46	0.53	0.61	0.68	0.77	**0.40**
20	XSM	**XSM2**	-0.42	0.11	0.26	0.37	0.46	0.53	0.61	0.68	0.77	**0.37**
21	XSM	**XSM3**	-0.31	0.17	0.28	0.37	0.46	0.53	0.61	0.68	0.77	**0.40**
22	ES_{min}	**ES1**	-0.18	0.14	0.27	0.36	0.46	0.53	0.60	0.68	0.77	**0.40**
23	ES_{min}	**ES2**	0.14	0.26	0.33	0.40	0.48	0.55	0.62	0.69	0.77	**0.47**
24	ES_{min}	**ES3**	0.15	0.24	0.31	0.37	0.44	0.51	0.58	0.66	0.76	**0.45**
25	ES_{max}	**ES1**	-0.40	0.01	0.24	0.34	0.45	0.52	0.60	0.68	0.77	**0.36**
26	ES_{max}	**ES2**	0.06	0.23	0.34	0.40	0.47	0.54	0.61	0.69	0.77	**0.46**
27	ES_{max}	**ES3**	0.07	0.21	0.31	0.37	0.44	0.50	0.57	0.66	0.76	**0.43**

Average MSEs

ID	Method	EAC(t)	7%	17%	28%	38%	48%	59%	69%	79%	90%	
Serial-Par. (%SP) ►			**2.69**	**2.68**	**2.76**	**2.94**	**2.23**	**1.19**	**0.82**	**0.50**	**0.29**	**Avg.**
1	EVM	**PV1**	1.84	1.41	1.00	0.80	0.76	0.66	0.59	0.47	0.43	**0.88**
2	EVM	**PV2**	2.5	3.97	4.46	4.86	4.04	1.94	1.40	0.78	0.54	**2.72**
3	EVM	**PV3**	3.3	7.8	8.78	9.06	10.05	4.87	4.56	3.54	3.01	**6.10**
4	EVM	**ED1**	1.75	1.50	1.03	0.82	0.82	0.80	0.68	0.50	0.36	**0.92**
5	EVM	**ED2**	2.5	3.97	4.46	4.86	4.08	1.95	1.42	0.77	0.45	**2.71**
6	EVM	**ED3**	3.1	7.1	7.77	7.31	7.89	2.69	1.96	1.00	0.52	**4.37**
7	EVM	**ES1**	1.77	1.39	0.93	0.69	0.60	0.46	0.32	0.20	0.09	**0.72**
8	EVM	**ES2**	2.56	1.70	2.63	2.71	1.65	0.95	0.53	0.28	0.11	**1.46**
9	EVM	**ES3**	3.21	3.19	4.28	5.90	3.25	1.60	0.95	0.47	0.16	**2.56**
10	EDM	**PV1**	1.85	1.37	0.94	0.71	0.64	0.53	0.38	0.27	0.25	**0.77**
11	EDM	**PV2**	2.55	1.75	1.56	2.09	1.46	0.99	0.53	0.31	0.23	**1.28**
12	EDM	**ED1**	1.77	1.42	0.99	0.73	0.67	0.54	0.39	0.23	0.10	**0.76**
13	EDM	**ED2**	2.55	1.75	1.56	2.09	1.47	0.99	0.53	0.29	0.11	**1.26**
14	EDM	**ES1**	1.78	1.38	0.91	0.68	0.59	0.46	0.32	0.20	0.09	**0.71**
15	EDM	**ES2**	2.62	1.49	2.29	2.45	1.60	0.95	0.52	0.28	0.11	**1.37**
16	ESM	**ES1**	1.67	1.33	0.93	0.70	0.60	0.46	0.32	0.20	0.09	**0.70**
17	ESM	**ES2**	1.89	1.97	3.45	3.14	1.73	0.97	0.53	0.29	0.11	**1.56**
18	ESM	**ES3**	2.58	3.83	5.50	6.99	3.38	1.63	0.96	0.47	0.16	**2.83**
19	XSM	**XSM1**	2.08	1.49	1.67	2.52	1.76	1.03	0.55	0.29	0.11	**1.28**
20	XSM	**XSM2**	1.82	1.28	0.98	1.13	1.20	0.95	0.65	0.39	0.12	**0.95**
21	XSM	**XSM3**	2.09	1.46	1.23	1.83	1.49	0.92	0.50	0.28	0.10	**1.10**
22	ES_{min}	**ES1**	1.62	1.25	0.84	0.65	0.57	0.44	0.32	0.20	0.09	**0.67**
23	ES_{min}	**ES2**	2.60	3.78	4.17	3.02	1.60	0.93	0.52	0.28	0.11	**1.89**
24	ES_{min}	**ES3**	3.32	5.61	6.22	7.24	3.15	1.58	0.94	0.46	0.15	**3.19**
25	ES_{max}	**ES1**	2.82	1.84	1.06	0.72	0.61	0.46	0.32	0.20	0.09	**0.90**
26	ES_{max}	**ES2**	7.05	3.35	1.92	1.98	1.49	0.93	0.52	0.28	0.11	**1.96**
27	ES_{max}	**ES3**	7.41	4.10	2.84	3.64	2.91	1.54	0.93	0.46	0.15	**2.67**

Average MAPEs

ID	Method	EAC(t)	7%	17%	28%	38%	48%	59%	69%	79%	90%	
Serial-Par. (%SP) ►			**0.34**	**0.22**	**0.18**	**0.15**	**0.12**	**0.09**	**0.07**	**0.05**	**0.04**	**Avg.**
1	EVM	**PV1**	0.27	0.18	0.14	0.11	0.09	0.08	0.07	0.06	0.05	**0.12**
2	EVM	**PV2**	0.34	0.22	0.19	0.17	0.14	0.11	0.09	0.07	0.05	**0.15**
3	EVM	**PV3**	0.41	0.30	0.27	0.26	0.23	0.19	0.18	0.15	0.14	**0.24**
4	EVM	**ED1**	0.26	0.18	0.14	0.11	0.10	0.09	0.08	0.06	0.05	**0.12**
5	EVM	**ED2**	0.34	0.22	0.19	0.17	0.14	0.11	0.09	0.07	0.05	**0.15**
6	EVM	**ED3**	0.40	0.28	0.24	0.21	0.17	0.13	0.11	0.07	0.05	**0.18**
7	EVM	**ES1**	0.26	0.18	0.13	0.10	0.08	0.07	0.05	0.04	0.03	**0.10**
8	EVM	**ES2**	0.35	0.21	0.20	0.16	0.12	0.09	0.07	0.05	0.03	**0.14**
9	EVM	**ES3**	0.40	0.27	0.25	0.21	0.16	0.11	0.08	0.06	0.03	**0.17**
10	EDM	**PV1**	0.27	0.18	0.13	0.10	0.09	0.07	0.06	0.04	0.03	**0.11**
11	EDM	**PV2**	0.35	0.19	0.16	0.15	0.11	0.09	0.07	0.05	0.03	**0.13**
12	EDM	**ED1**	0.26	0.18	0.13	0.10	0.09	0.08	0.06	0.04	0.03	**0.11**
13	EDM	**ED2**	0.35	0.19	0.16	0.15	0.11	0.09	0.07	0.05	0.03	**0.13**
14	EDM	**ES1**	0.26	0.18	0.13	0.10	0.08	0.07	0.05	0.04	0.03	**0.10**
15	EDM	**ES2**	0.36	0.19	0.19	0.16	0.12	0.09	0.07	0.05	0.03	**0.14**
16	ESM	**ES1**	0.25	0.18	0.13	0.10	0.08	0.07	0.05	0.04	0.03	**0.10**
17	ESM	**ES2**	0.28	0.22	0.23	0.17	0.12	0.09	0.07	0.05	0.03	**0.14**
18	ESM	**ES3**	0.34	0.28	0.27	0.22	0.16	0.11	0.08	0.06	0.03	**0.17**
19	XSM	**XSM1**	0.29	0.19	0.17	0.16	0.12	0.09	0.07	0.05	0.03	**0.13**
20	XSM	**XSM2**	0.26	0.17	0.14	0.12	0.11	0.09	0.07	0.05	0.03	**0.12**
21	XSM	**XSM3**	0.29	0.19	0.15	0.14	0.11	0.09	0.06	0.05	0.03	**0.12**
22	ES_{min}	**ES1**	0.25	0.17	0.12	0.10	0.08	0.07	0.05	0.04	0.03	**0.10**
23	ES_{min}	**ES2**	0.33	0.28	0.22	0.16	0.12	0.09	0.06	0.05	0.03	**0.15**
24	ES_{min}	**ES3**	0.37	0.33	0.26	0.21	0.15	0.11	0.08	0.06	0.03	**0.18**
25	ES_{max}	**ES1**	0.36	0.21	0.14	0.10	0.08	0.07	0.05	0.04	0.03	**0.12**
26	ES_{max}	**ES2**	0.67	0.33	0.19	0.15	0.12	0.09	0.07	0.05	0.03	**0.19**
27	ES_{max}	**ES3**	0.69	0.36	0.23	0.20	0.15	0.11	0.08	0.06	0.03	**0.21**

Summary (top performers for both stability and accuracy)

Avg. values ►			Stability				Accuracy			Both
ID	Method	EAC(t)	*R*	r	t	(Best)	*MSE*	*MAPE*	(Best)	(Best)
1	EVM	**PV1**	0.62	0.62	0.33		0.88	0.12	*	
2	EVM	**PV2**	0.62	0.64	0.40		2.72	0.15		
3	EVM	**PV3**	0.45	0.46	0.28		6.10	0.24		
4	EVM	**ED1**	0.64	0.63	0.29		0.92	0.12	*	
5	EVM	**ED2**	0.63	0.64	0.41		2.71	0.15		
6	EVM	**ED3**	0.56	0.58	0.37		4.37	0.18		
7	EVM	**ES1**	0.67	0.64	0.34		0.72	0.10	**	
8	EVM	**ES2**	0.66	0.66	0.44	*	1.46	0.14		
9	EVM	**ES3**	0.61	0.62	0.41		2.56	0.17		
10	EDM	**PV1**	0.66	0.65	0.35		0.77	0.11	**	
11	EDM	**PV2**	0.67	0.68	0.43	*	1.28	0.13		
12	EDM	**ED1**	0.68	0.65	0.32		0.76	0.11	**	
13	EDM	**ED2**	0.68	0.68	0.44	*	1.26	0.13		
14	EDM	**ES1**	0.68	0.65	0.34		0.71	0.10	**	
15	EDM	**ES2**	0.69	0.69	0.45	**	1.37	0.14		
16	ESM	**ES1**	0.67	0.65	0.34		0.70	0.10	**	
17	ESM	**ES2**	0.64	0.64	0.43		1.56	0.14		
18	ESM	**ES3**	0.59	0.61	0.41		2.83	0.17		
19	XSM	**XSM1**	0.64	0.63	0.40		1.28	0.13		
20	XSM	**XSM2**	0.65	0.64	0.37		0.95	0.12	*	
21	XSM	**XSM3**	0.66	0.65	0.40		1.10	0.12	*	
22	ES_{min}	**ES1**	0.71	0.68	0.40	*	0.67	0.10	**	*
23	ES_{min}	**ES2**	0.68	0.68	0.47	**	1.89	0.15		
24	ES_{min}	**ES3**	0.64	0.65	0.45	*	3.19	0.18		
25	ES_{max}	**ES1**	0.63	0.63	0.36		0.90	0.12	**	
26	ES_{max}	**ES2**	0.70	0.70	0.46	**	1.96	0.19		
27	ES_{max}	**ES3**	0.65	0.67	0.43	*	2.67	0.21		

Figure 3: EAC(t) estimates' stability and accuracy by the Serial-Parallel (*SP*) indicator (*SP*=0 Parallel network, *SP*=1 Serial network).

Supplemental online material

Results from the 4100 simulated projects can be downloaded here https://bit.ly/3d8jRIl . The link allows you to download a 87 MB Excel spreadsheet file. Please, be patient when downloading and opening it. A list of all abbreviations, variables and mathematical expressions used in the paper follows in Table S1.

Table S1. List of major EVM-related abbreviations, variables and mathematical expressions used in the paper.

Variable	Unit	Description	Expression
AC	Money	(Project) Actual Cost (at tracking period *AT*)	$AC = \sum_{i \in N} AC_{i,AT}$
$AC_{i, AT}$	Money	Activity *i*'s Actual Cost (at tracking period *AT*)	-
AT	Time	Actual Time (current tracking period). Also named *AD*.	-
AT_t	Time	Ongoing (Project) duration at tracking period *t*	-
BAC	Money	(Project) Budget At Completion (planned total project cost estimate)	-
BAC(t)	Time	(Project) Duration At Completion (sum of all activities' *TPD* or planned d_i at the end of the project)	$BAC(t) = \sum_{i \in N} d_i$
BAC_i	Money	Activity *i*'s total planned cost	-
c_i	Money	Activity *i*'s cost	-
CPI	1	(Project) Cost Performance Index (at tracking period *AT*)	$CPI = \frac{EV}{AC}$
CV_i	1	Activity *i*'s Coefficient of Variation	$CV_i = \frac{\sigma_i}{\mu_i}$
d_i	Time	Activity *i*'s duration	-
d_{mk}	Ranks	Difference between the ranking (order) values of $EAC(t)_{AT}$ and RD_{mk} for project *m*, for the *K* simulation runs at a particular tracking period AT.	$d_{mk} \equiv rank\left(EAC(t)^k_{AT}\right) - rank\left(RD_{mk}\right)$ for *k*=1,2…*K*.
DPI	1	(Project) Duration Performance Index (EVM's *SPI(t)* counterpart in EDM) (at tracking period *AT*)	$DPI = \frac{ED}{AT}$
$EAC(t)_{AT}$	Time	Project Duration forecasting estimate at tracking period AT.	-
$EAC(t)_x$	Time	Project Duration forecasting method *x*	See Table 1
$\overline{EAC(t)_{AT}}$	Time	Average of the EAC(t) estimates at tracking period AT in the K simulation runs.	

Variable	Unit	Description	Expression
ED	Time	(Project) Earned Duration (EVM's *ES* counterpart in EDM) (at tracking period *AT*) as formulated by Khamooshi and Golafshani (2014). Also named *ED(t)*	$ED = t + \frac{TED - TPD_t}{TPD_{t+1} - TPD_t}$
ED'	Time	(Project) Earned Duration (at tracking period AT) as formulated by Jacob (2003)	$ED = AT \cdot SPI$
EDI	1	(Project) Earned Duration Index (EVM's *SPI* counterpart in EDM) (at tracking period *AT*)	$EDI = \frac{TED}{TPD}$
EDI_t	1	(Project) *EDI* at tracking period *t*	$EDI_t = \frac{TED_t}{TPD_t}$
EDM	-	Earned Duration Management	-
ES	Time	(Project) Earned Schedule (at tracking period *AT*)	$ES = t + \frac{EV - PV_t}{PV_{t+1} - PV_t} \cdot (t+1-t)$
ES(e)	Time	(Project) Effective Earned Schedule (ES calculated with *EV(e)*) (at tracking period *AT*)	$ES(e) = t + \frac{EV(e) - PV_t}{PV_{t+1} - PV_t}$
ES_i	Time	Activity *i*'s Earned Schedule (at tracking period *AT*)	$ES_i = SD_i + PC_i \cdot d_i$
ESM	-	Earned Schedule Management	-
ES_{max}	Time	(Project) Maximum Earned Schedule (at tracking period *AT*)	$ES_{max} = MAX\left\{ES_i : PC_i \subset (0,1], i \in N\right\}$
ES_{min}	Time	(Project) Minimum Earned Schedule (at tracking period *AT*)	$ES_{min} = MIN\left\{ES_i + s_i : PC_i \subset [0,1), i \in N\right.$
ES_t	Time	(Project) *ES* at tracking period *t*	-
EV	Money	(Project) Earned Value (at tracking period *AT*)	$EV = \sum_{i \in N} EV_{i,AT}$
EV(e)	Money	(Project) Effective Earned Value (at tracking period *AT*)	$EV(e) = \left[1 - (1-p)\left(1 - PC \cdot e^{-\frac{1-PC}{2}}\right)\right] EV$
$EV_{i,\,AT}$	Money	Activity *i*'s Earned Value (at tracking period *AT*)	-
$EV_{i,AT}$	Money	Activity *i*'s Earned Value at tracking period *AT*	-
EVM	-	Earned Value Management	-
EV_t	Money	(Project) Earned Value at tracking period *t*	$EV_t = \sum_{i \in N} EV_{i,t}$
i	-	Activity *i* (one of the activities of the project schedule with *i*=1,2…*N*)	-
k	-	Each of the simulation runs in the experiments (*k*=1,2, …*K* in the paper)	-
K	Sim. Runs	Number of simulation runs in the experiments (*K*=100 in the paper)	-
m	-	Project *m* (one of the 4100 simulated projects of the dataset)	-
M	-	Total number of projects (4100 in the dataset)	-
$MAPE_{AT}$	-	Mean Absolute Percentage Error at tracking period AT.	See eq. (3) in the paper

Variable	Unit	Description	Expression
MSE_{AT}	-	Mean Squared Error Error at tracking period AT.	See eq. (2) in the paper
n	Activities	Total number of (non-dummy) activities scheduled in the ongoing project (30 in the projects dataset)	-
N	units	Total number of planned tracking periods of a project	-
p	1	(Project) Schedule adherence *p-Factor* (at tracking period *AT*)	$p = \frac{\sum_{i \in N} \min\left(PV_{i,ES}, EV_{i,AT}\right)}{\sum_{i \in N} PV_{i,ES}}$
P_m	Concordant pairs	Number of concordant pairs between the EAC(t) estimates and the RD values for each project *m*.	See eq. (7) in the paper
PC	1	(Project) Percentage of Completion (at tracking period *AT*)	$PC = \frac{EV}{BAC}$
PC_i	1	Activity i's Percentage of Completion (at tracking period *AT*)	$PC_i = \frac{EV_i}{BAC_i}$
PD	Time	(Project) Planned Duration (total project duration estimate)	-
PF	-	Performance Factor that can be 1 or another indicator such as *SPI*, *SPI(t)*, *SPI(t)(e)*, $SPI(t)_{ESmin}$, $SPI(t)_{ESmax}$, *SCI*, *SCI(t)*, *SCI(t)(e)*, $SCI(t)_{ESmin}$, $SCI(t)_{ESmax}$, even an exponential smoothing factor (e.g. $T_{t,SPI(t)}$, $T_{t,AT}/T_{t,ES}$ or $T_{t,EDI}$)	-
PV	Money	(Project) Planned Value (at tracking period *AT*)	$PV = \sum_{i \in N} PV_{i,AT}$
$PV_{i, AT}$	Money	Activity *i*'s Planned Value (at tracking period *AT*)	-
$PV_{i,ES}$	Money	Activity *i*'s Planned Value at tracking period *ES*	-
PV_{rate}	Money/ time	(Project) Planned Value rate (at tracking period *AT*)	$PV_{rate} = BAC/PD$
PV_t	Money	(Project) Planned Value at tracking period *t*	$PV_t = \sum_{i \in N} PV_{i,t}$
q	Activities	Number of activities in the path with the highest number of activities of a project (which is not necessarily the longest in duration)	-
R_{AT}	1	Pearson's linear correlation coefficient at tracking period AT.	See eq. (4) in the paper
RAC	Money	(Project) Real (budget) At Completion (only known once the project is completed)	-
RD	Time	(Project) Real Duration (only known once the project is completed)	-
RD_{mk}	Time	Real (Actual) Duration of project *m* in the simulation run *k*.	-
$\overline{RD_{mk}}$	Time	Average of project *m*'s Real Duration values in the *K* simulation runs.	

Variable	Unit	Description	Expression
SCI	1	(Project) Schedule Cost Index using original EVM metrics (at tracking period AT)	$SCI = SPI \cdot CPI$
$SCI(t)$	1	(Project) Schedule Cost Index using the ES metric (at tracking period AT)	$SCI(t) = SPI(t) \cdot CPI$
$SCI(t)_{ESmax}$	1	(Project) Schedule Cost Index using the ES_{max} metric (at tracking period AT)	$SCI(t)_{ESmax} = SPI(t)_{ESmax} \cdot CPI$
$SCI(t)_{ESmin}$	1	(Project) Schedule Cost Index using the ES_{min} metric (at tracking period AT)	$SCI(t)_{ESmin} = SPI(t)_{ESmin} \cdot CPI$
$SCI(t)(e)$	1	(Project) Effective Schedule Cost Index (at tracking period AT)	$SCI(t)(e) = SPI(t)(e) \cdot CPI$
SD_i	Time	Activity i's (Earliest) Start Date (at tracking period AT)	Critical path calculations (ASAP schedule)
s_i	Time	Activity i's slack or float . Difference between each activity's earliest and latest Finish or Start.	Critical path calculations
SP	1	Serial-Parallel topological indicator	$SP=(q-1)/(n-1)$
SPI	1	(Project) Schedule Performance Index using original EVM metrics (at tracking period AT)	$SPI = \frac{EV}{PV}$
$SPI(t)$	1	(Project) Schedule Performance Index using the ES metric (at tracking period AT)	$SPI(t) = \frac{ES}{AT}$
$SPI(t)_{ESmax}$	1	(Project) Schedule Performance Index using the ES_{max} metric (at tracking period AT)	$SPI(t)_{ESmax} = ES_{max} / AT$
$SPI(t)_{ESmin}$	1	(Project) Schedule Performance Index using the ES_{min} metric (at tracking period AT)	$SPI(t)_{ESmin} = ES_{min} / AT$
$SPI(t)(e)$	1	(Project) Effective Schedule Performance Index (at tracking period AT)	$SPI(t)(e) = \frac{ES(e)}{AT}$
$SPI(t)_t$	1	(Project) $SPI(t)$ at tracking period t	$SPI(t)_t = \frac{ES_t}{t}$
t	Time	Integer tracking period such that: $PV_t \leq EV < PV_{t+1}$ for ES calculations, $TPD_t \leq TED < TPD_{t+1}$ for ED calc., $PV_t \leq EV(e) < PV_{t+1}$ for $ES(e)$ calc., or just t=0,1,2…n for the other calc.	-
TAD	Time	(Project) Total Actual Duration (EVM's AC counterpart in EDM) (at tracking period AT)	$TAD = \sum_{i \in N} TAD_{i,AT}$
$TAD_{i,AT}$	Time	Activity i's Actual Duration (at tracking period AT)	-
TED	Time	(Project) Total Earned Duration (EVM's EV counterpart in EDM) (at tracking period AT)	$TED = \sum_{i \in N} TED_{i,AT}$
$TED_{i,AT}$	Time	Activity i's Earned Duration (at tracking period AT)	-
TED_t	Time	(Project) Total Earned Duration at tracking period t	$TED_t = \sum_{i \in N} TED_{i,t}$
TPD	Time	(Project) Total Planned Duration (EVM's PV counterpart in EDM) (at tracking period AT)	$TPD = \sum_{i \in N} TPD_{i,AT}$
$TPD_{i,AT}$	Time	Activity i's Planned Duration (at tracking period AT)	-
TPD_t	Time	(Project) Total Planned Duration at tracking period t	$TPD_t = \sum_{i \in N} TPD_{i,t}$

Variable	Unit	Description	Expression
$T_{t,AT}$	1	(Project) Trend of *AT* at tracking period *t*	$T_{t,AT} = \gamma\left(AT_t - AT_{t-1}\right) + \left(1-\gamma\right)T_{t-1,AT}$ with $T_{0,AT} = \frac{PD}{n}$
$T_{t,EDI}$	1	(Project) Trend of *EDI* at tracking period *t*	$T_{t,EDI} = \beta \cdot EDI_t + \left(1-\beta\right)T_{t-1,EDI}$ with $T_{0,EDI} = 1$
$T_{t,ES}$	1	(Project) Trend of *ES* at tracking period *t*	$T_{t,ES} = \gamma\left(ES_t - ES_{t-1}\right) + \left(1-\gamma\right)T_{t-1,AT}$ with $T_{0,ES} = \frac{PD}{n}$
$T_{t,SPI(t)}$	1	(Project) Trend of *SPI(t)* at tracking period *t*	$T_{t,SPI(t)} = \beta \cdot SPI(t)_t + \left(1-\beta\right)T_{t-1,SPI(t)}$ with $T_{0,SPI(t)} = 1$
TV	1	(Project) Time Variance (at tracking period *AT*)	$TV = SV/PV_{rate}$
X_i	1	Activity *i*'s correlated Normally-distributed random variable whose variability is conditioned by the d_i from equation (2) and the rest varies according to a st. Normal distrib.	See eq. (14) in the paper
β	1	Exponential smoothing constant (assumed here as 0.25 according to Khamooshi & Abdi (2016))	-
γ	1	Exponential smoothing constant (assumed here as 0.05 according to Batselier & Vanhoucke (2017))	-
δ_i	1	Activity *i*'s duration-cost correlation	See eq. (15) in the paper
μ_i	Log time	Activity *i*'s average log duration.	See eq. (9) in the paper
μ'_i	Log money	Activity *i*'s average log cost.	See eq. (12) in the paper
σ_i	Log time	Activity *i*'s avg. log duration st. deviation.	See eq. (10) in the paper
σ'_i	Log money	Activity *i*'s average log cost st. deviation.	See eq. (13) in the paper
ρ_{AT}	1	Spearman's rank correlation coefficient at tracking period AT.	See eq. (5) in the paper
τ_{AT}	1	Kendall's rank correlation coefficient at tracking period AT.	See eq. (6) in the paper

Acknowledgements (Research funding)

The first author acknowledges the University of Talca for his Doctoral Program Scholarship (RU-056-2019). The second author acknowledges the Spanish Ministry of Science and Innovation for his Ramon y Cajal contract (RYC-2017-22222) co-funded by the European Social Fund.