Project regularity: Development and evaluation of a new project characteristic

Jordy Batselier[a], Mario Vanhoucke[a,b,c,*]

[a] Faculty of Economics and Business Administration, Ghent University, Tweekerkenstraat 2, 9000 Ghent, Belgium
[b] Technology and Operations Management Area, Vlerick Business School, Reep 1, 9000 Ghent, Belgium
[c] UCL School of Management, University College London, 1 Canada Square, London E14 5AA, United Kingdom

Abstract

The ability to accurately characterize projects is essential to good project management. Therefore, a novel project characteristic is developed that reflects the value accrue within a project. This characteristic, called project regularity, is expressed in terms of the newly introduced regular/irregular-indicator RI. The widely accepted management system of earned value management (EVM) forms the basis for evaluation of the new characteristic. More concretely, the influence of project regularity on EVM forecasting accuracy is assessed, and is shown to be significant for both time and cost forecasting. Moreover, this effect appears to be stronger than that of the widely used characteristic of project seriality expressed by the serial/parallel-indicator SP. Therefore, project regularity could also be useful as an input parameter for project network generators. Furthermore, the introduction of project regularity can provide project managers with a more accurate indication of the time and cost forecasting accuracy that is to be expected for a certain project and, correspondingly, of how a project should be built up in order to obtain more reliable forecasts during project control.

Keywords: Project management, earned value management, time and cost forecasting, empirical database, Monte Carlo simulation, project control system

Cite as: Batselier, J. and Vanhoucke, M., 2017, "Project regularity: Development and evaluation of a new project characteristic", Journal of Systems Science and Systems Engineering, 26(1), 100-120 (doi:10.1007/ s11518-016-5312-6).

1. Introduction

Project managers need reliable estimates of the performance that is required for the remaining work in a project so that the project's schedule and budget objectives could be met (Cioffi, 2006). To this end, accurate final cost and duration forecasts of a project in progress are essential. Since we consider projects in progress, this study is situated in the project control phase. Earned value management (EVM) is a widely accepted technique for performing project control of which the benefits have been amply demonstrated (Christensen, 1998; Anbari, 2003; Kim et al., 2003; Henderson, 2007; Marshall, 2007; Van De Velde, 2007; Egnot, 2011). Furthermore, Hall (2012) also identified EVM as a management system that exhibits multiple research opportunities.

Conceptually, EVM integrates the three critical project management elements of cost, schedule and scope. More concretely, through the application of EVM, the project manager can monitor the performance of the project during execution and receive warning signals for taking corrective actions that are needed to get the project back on track. A overview of EVM's key metrics and formulas can be found in Anbari (2003), PMI (2008), Fleming & Koppelman (2010), Vanhoucke (2010a) and Vanhoucke (2014). Nevertheless, a summary table (table 1) is included to ensure the standalone comprehensibility of this paper.

Furthermore, the EVM technique can also be used to produce project forecasts, which is the focus of the current study. The most common EVM time and cost forecasting methods were identified by Vanhoucke (2012b) and are presented in table 2.

*Corresponding author. Tel.: +32 92643569.
Email addresses: jordy.batselier@ugent.be (Jordy Batselier), mario.vanhoucke@ugent.be (Mario Vanhoucke)

Table 1: Definitions of the EVM key metrics

Metric	Definition
PD	Planned duration, the planned total duration of the project
BAC	Budget at completion, the budgeted total cost of the project
AT	Actual time
PV	Planned value, the value[a] that was planned to be earned at AT
EV	Earned value, the value that has actually been earned at AT
AC	Actual cost, the costs that have actually been incurred at AT
ES	Earned schedule, the time at which the EV should have been earned according to plan, $ES = t + \frac{EV - PV_t}{PV_{t+1} - PV_t}$ with t the (integer) point in time for which $EV \geq PV_t$ and $EV < PV_{t+1}$
RD	Real duration, the actual total duration of the project
RC	Real cost, the actual total cost of the project
$EAC(t)$	Estimated duration at completion, the prediction of RD made at AT
EAC	Estimated cost at completion, the prediction of RC made at AT
SV	Schedule variance, $SV = EV - PV$
SPI	Schedule performance index, $SPI = \frac{EV}{PV}$
$SV(t)$	Schedule variance (time), $SV(t) = ES - AT$
$SPI(t)$	Schedule performance index (time), $SPI(t) = \frac{ES}{AT}$
CV	Cost variance, $CV = EV - AC$
CPI	Cost performance index, $CPI = \frac{EV}{AC}$
SCI	Schedule cost index, $SCI = SPI * CPI$
$SCI(t)$	Schedule cost index (time), $SCI(t) = SPI(t) * CPI$

[a] In these definitions, *value* always alludes to the cumulative value over all activities up to a certain point in time.

Table 2: EVM time and cost forecasting methods

According to	plan	current time performance	current cost performance	current time/cost performance	weighted time/cost performance
PF =	1	SPI SPI(t)	CPI	SCI SCI(t)	0.8CPI+0.2SPI 0.8CPI+0.2SPI(t)
Time					
PVM	**PVM-1**	**PVM-SPI**		**PVM-SCI**	
$EAC(t) =$	$PD - \frac{SV*PD}{BAC}$	$\frac{PD}{SPI}$		$\frac{PD}{SCI}$	
EDM	**EDM-1**	**EDM-SPI**		**EDM-SCI**	
$EAC(t) =$	$PD + AT * (1 - SPI)$	$\frac{PD}{SPI}$		$\frac{PD}{SCI} + AT * (1 - \frac{1}{CPI})$	
ESM	**ESM-1**	**ESM-SPI(t)**		**ESM-SCI(t)**	
$EAC(t) =$	$AT + PD - ES$	$AT + \frac{PD-ES}{SPI(t)}$		$AT + \frac{PD-ES}{SCI(t)}$	
Cost					
EAC	**EAC-1**	**EAC-SPI**	**EAC-CPI**	**EAC-SCI**	**EAC-0.8CPI+0.2SPI**
$EAC =$	$AC + BAC - EV$	$AC + \frac{BAC-EV}{SPI}$	$AC + \frac{BAC-EV}{CPI}$	$AC + \frac{BAC-EV}{SCI}$	$AC + \frac{BAC-EV}{0.8CPI+0.2SPI}$
		EAC-SPI(t)		**EAC-SCI(t)**	**EAC-0.8CPI+0.2SPI(t)**
$EAC =$		$AC + \frac{BAC-EV}{SPI(t)}$		$AC + \frac{BAC-EV}{SCI(t)}$	$AC + \frac{BAC-EV}{0.8CPI+0.2SPI(t)}$

Here, the performance factor (PF) reflects the assumption that has been made to produce a certain forecast. The different possible assumptions about the expected performance of future work are presented on the first row of table 2, while the second row shows the corresponding performance factors. We can identify nine methods for time forecasting, grouped into three overarching methodologies: the planned value method (PVM) by Anbari (2003); the earned duration method (EDM) by Jacob & Kane (2004); and the earned schedule method (ESM) by Lipke (2003). Furthermore, eight different EVM-based methods are considered for forecasting project cost.

In order to further clarify the nature and use of the methods summarized in table 2, we now present an example. Assume that it is our goal to forecast a project's duration by taking into account both its current time and cost performance. We see from the first two rows of table 2 that if we want to do this, we have to apply a PF equal to SCI or SCI(t) (fifth column). From table 1 we learn that the difference between SCI and SCI(t) is that the first metric is calculated with the traditional SPI, whereas the second one uses the more novel ES-based SPI(t) for multiplication with CPI. Assume we choose to apply SCI(t) as PF. From the fifth column of table 2, we can identify ESM-SCI(t) as the only EVM-based time forecasting method for doing so, since both PVM-SCI and EDM-SCI - as their naming suggests - use SCI. Below the method's name, the formula for applying it is provided. For ESM-SCI(t), this comes down to a project duration forecast or EAC(t) that can be calculated as $AT + (PD - ES)/SCI(t)$, with all the necessary metrics defined in table 1. This procedure can be followed for every other forecasting method included in table 2. This is also true for cost forecasting, for which there are no subdividing methodologies (like PVM, EDM and ESM for time forecasting), but for which two extra assumptions about the expected future performance can be made (see fourth and sixth column of table 2).

Batselier & Vanhoucke (2015b) evaluated the overall accuracy of these time and cost forecasting methods on a real-life project database constructed by Batselier & Vanhoucke (2015a). As a continuation of the study of Vanhoucke & Vandevoorde (2007), the authors also assessed the influence of project seriality, a characteristic which describes the network structure of a project (see section 2.1 for further definition), on the methods' forecasting accuracy. They concluded that project seriality indeed seemed to influence time forecasting accuracy, supporting the simulation results obtained by Vanhoucke & Vandevoorde (2007). However, the said characteristic did not show any effect on cost forecasting accuracy.

In this paper, a novel project characteristic is developed that influences both time and cost forecasting accuracy, moreover, in a more significant manner than project seriality. This is an important result, as project seriality can be regarded as the most influential input parameter (topological indicator) for project network generators up to now (Vanhoucke et al., 2008). The newly introduced characteristic, called project regularity, reflects the value accrue within a project and is able to provide a better indication of the expected accuracy of a certain EVM forecasting method. More generally, the aims of introducing the concept of project regularity are to:

- increase the potential to adequately express a project's schedule and cost characteristics;
- more accurately indicate to project managers what level of time and cost forecasting accuracy is to be expected for a certain project;
- provide an indication to project managers of how a project should be built up in order to obtain more accurate forecasts during project control;
- define a new meaningful project characteristic (or topological indicator) that can be used as an input parameter for project network generators (Kolisch et al., 1995; Schwindt, 1995; Agrawal et al., 1996; Tavares, 1999; Demeulemeester et al., 2003; Vanhoucke et al., 2008).

The remainder of this paper is organized as follows. In section 2, the existing characteristic of project seriality is described and the novel characteristic of project regularity is introduced, together with their corresponding indicators. Section 3 then assesses the influence of the latter characteristic on EVM forecasting accuracy and compares the outcomes with those obtained for project seriality by Batselier & Vanhoucke (2015b). Moreover, all forecasting accuracy evaluations are based on the real-life project database from Batselier & Vanhoucke (2015a) and are also supported by simulation results. Section 4 further expands the study by investigating the combined influence of the two mentioned project characteristics on EVM forecasting accuracy. Finally, general conclusions and future research topics are presented in section 5.

2. Project characteristics

This section starts with the description of the existing characteristic of project seriality (subsection 2.1) and then proceeds to the introduction of the novel characteristic of project regularity (subsection 2.2).

2.1. Project seriality

The characteristic of project seriality describes the network structure of a project in terms of how close the network is to a serial (or parallel) network and can be represented by the so-called serial/parallel-indicator (SP). The SP can lie anywhere between 0 and 1, with the boundary values expressing a project for which all activities are in parallel or for which the network is completely serial, respectively (Tavares et al., 1999; Vanhoucke et al., 2008). All SP-values between these two extremes represent networks that are either closer to a serial or a parallel network.

Vanhoucke et al. (2008) mention that completely parallel and completely serial projects can also be indicated by the order strength (OS) measure (Mastor, 1970), that is, by the values OS = 0 and OS =

1, respectively. However, given the definition of OS (i.e. the number of precedence relations - including the transitive ones - divided by the theoretical maximum number of precedence relations), this measure better represents the density of a project network (Kao & Queyranne, 1982) rather than its seriality. Therefore, the SP is selected as the most appropriate measure for expressing project seriality. The formula for the SP is as follows.

$$SP = \frac{n_s - 1}{n_t - 1} \tag{1}$$

In this formula, n_s is the maximum number of subsequent activities in the network (or the maximum progressive level) and n_t is the total number of activities. Note that for a project of only one activity, SP is per definition equal to unity (completely serial project).

For accuracy evaluation purposes, the same SP-categorization as proposed by Batselier & Vanhoucke (2015a,b) is utilized. All projects in the considered database (see section 3.1) can thus be subdivided based on their SP-value according to the three following categories:

- 0% ≤ SP < 40%: parallel projects
- 40% ≤ SP ≤ 60%: serial-parallel projects
- 60% < SP ≤ 100%: serial projects

2.2. Project regularity

The characteristic of project regularity is completely new to literature. Therefore, a brief situating of the concept's feeding ground might be appropriate. Similar to what was done in Batselier & Vanhoucke (2015b), Jacob & Kane (2004) compared the ESM, EDM and PVM time forecasting approaches and came to the conclusion that "as long as the planned value (PV) is linear, all formulas will always yield exact results and if the PV were non-linear, it is possible errors or discrepancies could be introduced." Consequently, a project with a perfectly linear PV-curve can be identified as a completely regular project with minimal potential forecasting errors (Vandevoorde & Vanhoucke (2006) also endorse this reasoning). Thus, irregularities can be defined as deviations of the actual PV-curve from a perfectly linear one.

We now introduce a novel metric for expressing the level of regularity of a project, called the regular/irregular-indicator (RI). For the purpose of uniformity and comprehensibility, the regular/irregular-indicator is conceived strongly similar to the serial/parallel-indicator. Just as a completely serial project has an SP of 1, a perfectly regular project - that is a project with a perfectly linear PV-curve - is characterized by an RI of 1. At the opposite end of the regularity spectrum, a maximally irregular project is represented by RI = 0 and occurs when the earned value (EV) is zero throughout the entire course of the project and shoots up to the budget at completion (BAC) only at the very end. Just like for project seriality, projects with different degrees of regularity are situated between these two extreme cases. For those projects, the RI is calculated based on the following formula.

$$RI = \frac{\sum_{i=1}^{r} m_i - \sum_{i=1}^{r} a_i}{\sum_{i=1}^{r} m_i} \tag{2}$$

Here, m_i and a_i are, respectively, the maximal possible deviation and the actual deviation of the project's PV-curve from a perfectly linear curve at time instance $i = 1, ..., r$, with r the number of equidistant evaluation points. We want to stress that r is not related to the chosen number of tracking periods (which will be denoted by p in formula (4)), but is primarily dependent on the shape of the PV-curve. More concretely, for projects where the PV follows a more capricious course, r is augmented in order to guarantee an accurate calculation of RI. Remark that for many projects, the number of tracking periods is not large enough to allow a sufficiently detailed characterization of the PV-curve, and thus, a precise calculation of RI, which is the reason for not simply using p in formula (2). For the various projects considered in this paper, r ranges from at least 50 to about 120. The optimal value of r is determined by the project management software tool ProTrack (www.protrack.be). An elaboration on the concrete calculations of r, however, is outside the scope of this paper. Figure 1 clarifies the calculation of RI for one of the projects in the utilized database.

The RI for the example project of figure 1, which corresponds to project C2011-12 from the database of Batselier & Vanhoucke (2015a), is 0.75 or 75%. It is important to realize that project regularity, just

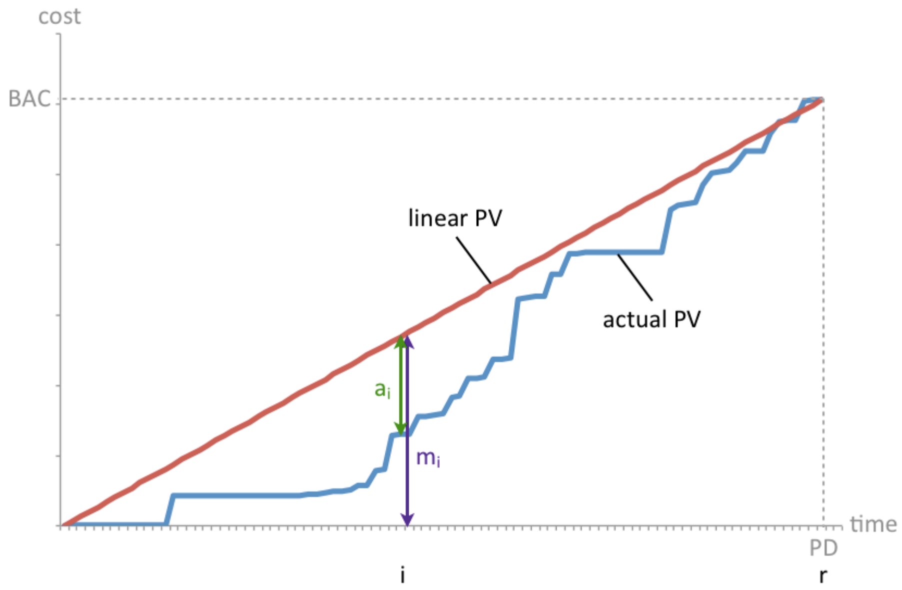


Figure 1: Example for calculation of RI

as project seriality, defines an intrinsic characteristic of a project. More specifically, the RI-value does not change during project execution, since its calculation is based on the course of the PV-curve, which represents the pre-project planning that forms the unchangeable point of reference for performing project control (through EVM). This is also true when simulating project progress (i.e. simulated forecasting; see section 3.1); the RI of the project will remain the same for every simulation run.

Furthermore, for enabling the assessment of the effect of project regularity on forecasting accuracy (section 3), the projects in the considered database have to be subdivided into categories based on their RI-value. In first instance, we are interested in the RI-value that defines the boundary between projects that can be regarded as 'regular' and projects that should be identified as 'irregular'. To this end, we consider the so-called perfect S-curve displayed in figure 2 together with the linear curve. The perfect S-curve is identified as the familiar textbook S-curve which assumes symmetrical spending in time (Meredith & Mantel, 2003). The mathematical expression of such an S-curve was proposed by Cioffi (2005) and is given by next formula.

$$y(\beta) = 1.018 \, \frac{1 - exp(-8\beta)}{1 + 54.6 \; exp(-8\beta)} \tag{3}$$

Here, β is a variable ranging from 0 to 1 which expresses time as a fraction of the total project duration. In other words, β could be calculated as AT/PD. Remark that the variable β merely reflects the normalized advancement of time in a project and therefore does not influence the shape of the S-curve. In order to make the S-curve more/less concave/convex, the numbers in formula (3) should be modified. For the discussion of the necessary modifications we refer to Cioffi (2005), as this is outside the scope of the current study. Furthermore, the $y(\beta)$ in formula (3) represents the project cost at the point in time corresponding to β. Based on the expression of formula (3), the RI for the perfect S-curve is calculated to be 81% according to formula (2).

Although the perfect S-curve displays noticeable deviations from the linear curve (see figure 2), it is suggested that a project for which the PV follows the trajectory of the perfect S-curve should still be considered a regular project. In other words, a project that is *typical* (i.e. for which the value accrue follows a perfect S-curve (Meredith & Mantel, 2003)) shows forecasting discrepancies that are *typically* to be expected and should therefore still be deemed *regular*. However, when a project would, for example, exhibit a smaller number of activities to be completed at the beginning and would consequently show a slower value accrue than the perfect S-curve in the early stages, it would presumably be prone to greater forecasting errors than the *typical* project. This example project would display more irregularities than what is *typical*, and could thus be characterized as *irregular*. Obviously, when a project's PV-curve would draw nearer to a linear curve than the perfect S-curve, the project becomes increasingly regular per definition (see formula (2)).

Given the RI of 81% for the perfect S-curve, the RI-value defining the boundary between the categories

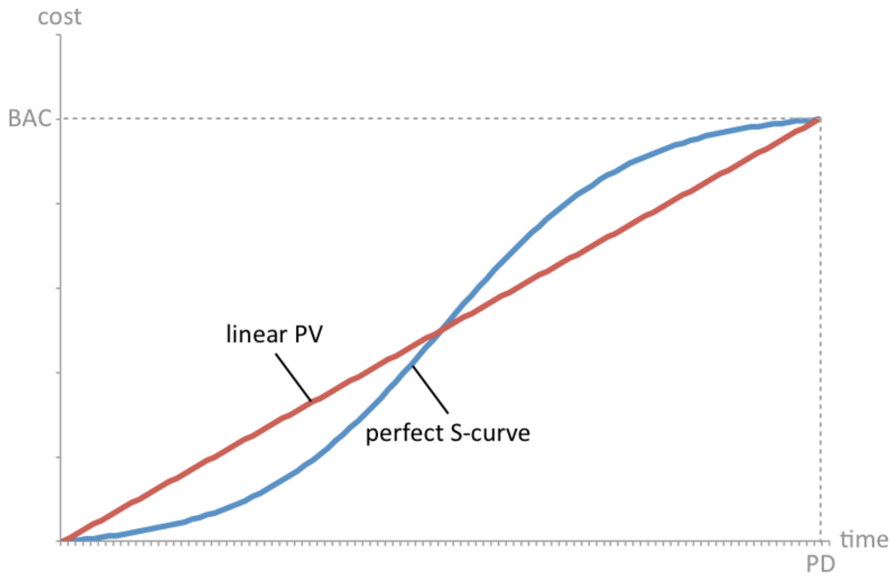


Figure 2: The perfect S-curve

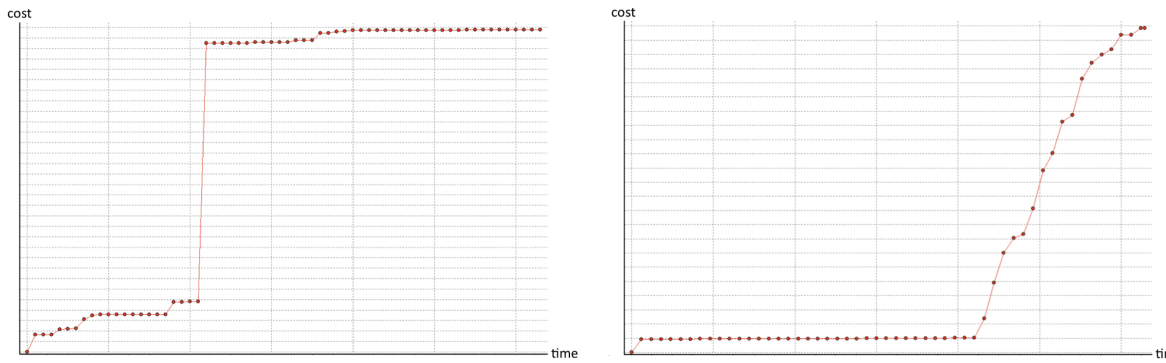
of regular and irregular projects is set to 80%. Consequently, projects that show a PV-curve that deviates from the perfect S-curve more towards the linear curve are categorized as regular (RI $\geq$ 80%), while projects for which those deviations bring the PV-curve further away from the linear curve are classified as irregular (RI $<$ 80%). Note, however, that RI-calculations are always based on deviations from the linear curve, and not from the perfect S-curve (see figure 1).

Although we will predominantly use the above described two-part categorization to allow a clear evaluation of the effect of project regularity on forecasting accuracy (section 3), the more extended research into the combined influence of project regularity and seriality on EVM forecasting accuracy (section 4) requires the definition of three regularity categories, just as for project seriality. Therefore, the category of irregular projects is further subdivided into mildly irregular projects and strongly irregular projects as follows:

- 0% $\leq$ RI $<$ 60%: strongly irregular projects
- 60% $\leq$ RI $<$ 80%: mildly irregular projects
- 80% $\leq$ RI $\leq$ 100%: regular projects

The additional interval boundaries are chosen in such a way that the number of considered projects in each category does not differ too much, as was also done for the SP-categorization (Batselier & Vanhoucke, 2015b). However, note that the boundary values for RI are not distributed symmetrically around 50%, as was the case for the project seriality categorization. Indeed, the distribution of the occurring RI-values is inclined more towards the higher end of the spectrum, which is an intrinsic consequence of the definition of RI by formula (2).

Since project regularity is a new concept, it might be appropriate to elaborate on what properties exactly make a project irregular and how one can spot such irregularities. Recall that we defined a completely regular project as a project for which the PV-curve is perfectly linear. For a project to become more irregular, its PV-curve should thus show some deviations from the perfectly linear curve. This implies that the PV-curve will be flatter than the linear curve in some sections, and since the BAC has to be reached at the planned duration (PD), steeper than the linear curve in other sections. This is exactly what we see for the perfect S-curve (figure 2): it is flatter than the linear curve in the beginning and at the end, and steeper in the middle of the project. In more extreme cases, the flatter parts of the PV-curve can become completely horizontal - resulting from a period of no value accrue - and very long, while the steeper parts can grow almost vertical and resemble very high ‘jumps’ in the PV-curve - which are generally induced by a fixed activity cost constituting a considerable part of the BAC. Strongly irregular projects are therefore easily recognizable - even at first glance - by two types of irregularities: long (almost) horizontal sections and/or high jumps in the PV-curve. Two clear examples of strongly irregular projects from the used database are displayed in figures 3a and 3b. More specifically, it concerns the projects with codes C2012-11 and C2013-10, respectively (see www.or-as.be/research/database). These projects show both types of irregularities: a high jump in PV in figure 3a, and a long period of no value accrue in figure 3b. The RI for both cases is 54% and 41%, respectively, so indeed these are strongly irregular projects.



(a) Jump in PV-curve

(b) Long period of no value accrue

Figure 3: Example PV-curves of strongly irregular projects

3. Influence of project regularity on EVM forecasting accuracy

In section 3.1, the applied approach for the evaluation of the influence of project regularity on EVM forecasting accuracy is described. More specifically, we discuss the utilized empirical project database, the evaluation measure used to express forecasting accuracy, and the important difference between simulated and real forecasting. Section 3.2 then presents the obtained results and compares the outcomes with those for the influence of project seriality on EVM forecasting accuracy. Time and cost forecasting are considered separately.

3.1. Evaluation approach

As a basis for this study, the empirical database constructed by Batselier & Vanhoucke (2015a) is utilized, which - at the time of this study - consisted of 51 real-life projects exhibiting wide ranges of project duration and budget. Moreover, the constituting projects originate from many different companies situated in various sectors (e.g. construction, event management, IT, production, education). This diversity - as well as the completeness and authenticity - of the project data is guaranteed through the application of so-called project cards, which summarize the most important properties of a certain project and enable their categorization and evaluation. More information on these project cards and the resulting database construction and evaluation framework can be found in the originating paper (Batselier & Vanhoucke, 2015a). Furthermore, the complete project database - including the project cards - is publicly available at www.or-as.be/research/database. All project data can be consulted in both ProTrack and MS Excel format thanks to the novel software tool PMConverter (available on the website), which also enables a more detailed presentation and evaluation of the considered projects.

The EVM forecasting accuracies are expressed in terms of mean absolute percentage error (MAPE). The MAPE is a well-known forecast error measure that is also frequently used in EVM accuracy studies (Vanhoucke & Vandevoorde, 2007; Rujirayanyong, 2009; Vanhoucke, 2010a; Elshaer, 2012; Batselier & Vanhoucke, 2015b). The measure is calculated as follows.

$$MAPE = \frac{1}{p}\sum_{t=1}^{p}\left|\frac{A - F_t}{A}\right| \tag{4}$$

In this formula, A represents the actual final value and F_t the forecasted value at time t. The time instances $t = 1, ..., p$ correspond to the p tracking periods that were chosen for the considered project. Furthermore, formula (4) can be particularized for time forecasting by substituting A and F_t by the real duration (RD) and the estimated duration at completion (EAC(t)), respectively. For cost forecasting,

A and F_t should be substituted by the real cost (RC) and the estimated cost at completion (EAC), respectively. The abbreviations used here are also defined in table 1. Obviously, a lower MAPE signifies a higher forecasting accuracy for a particular forecasting method.

The MAPE is the basis for both simulated and real forecasting accuracy evaluation. It is important to understand the essential difference between both types of forecasting outcomes. On the one hand, simulated forecasting results are based on Monte Carlo simulation runs for which appropriate risk distribution profiles for the individual activity durations are used as input, which means that the RD and RC required to calculate the MAPE from formula (4) originate from simulation outcomes in this case. For real forecasting, on the other hand, this RD and RC represent the *actual* final project outcomes, which can be obtained from the project owner.

We now list some specifics of the applied Monte Carlo simulation approach:

- Number of simulation runs for each project: 100
- Number of tracking periods for each simulation run: 20
- Risk distribution profiles for activity durations: triangular (symmetrical, skewed to the left, or skewed to the right)
- Cost assumptions: variable costs of an activity vary uniformly with the corresponding activity duration, fixed costs always remain constant and are fully incurred at the start of an activity

The simulations are performed with the project management software tool ProTrack. Note that for a number of 100 simulation runs, the forecasting accuracy results converge for every project in the employed database. A standard risk distribution profile is symmetrical with a best case and worst case activity duration of 80% and 120% of the most likely duration, respectively. However, the project owner can specify the best case and worst case durations in order to better reflect the risk of a certain activity, which yields specific triangular distributions that are mostly skewed. Furthermore, the fixed and variable costs (i.e. cost per hour) that are used in the Monte Carlo simulation are all obtained directly from the project owner as well. Note, however, that the simulation outcomes primarily have a supportive function for the real forecasting results in this paper.

3.2. Results

In this section, the results for the evaluation of the influence of project regularity on EVM forecasting accuracy are presented and compared with the corresponding outcomes for project seriality. First, time forecasting will be considered (section 3.2.1), then cost forecasting (section 3.2.2). Recall that the notations and formulas of the EVM forecasting methods considered here are included in table 2.

3.2.1. Time forecasting

As was mentioned in section 2.2, we focus on the presented two-part categorization (i.e. regular versus irregular projects) for the evaluation of the effect of project regularity on forecasting accuracy. The specific results for time forecasting are shown in table 3.

Table 3: Accuracies of EVM time forecasting methods w.r.t. project regularity

	PVM			**EDM**			**ESM**			
MAPE [%]	1	SPI	SCI	1	SPI	SCI	1	SPI(t)	SCI(t)	
simulated										
regular	13.3	19.5	20.2	16.8	19.3	19.7	11.6	16.5	16.9	
irregular	30.5	42.0	42.5	M	41.9	42.1	16.9	24.1	24.4	*avg*
diff.	17.2	22.5	22.3	M	22.6	22.4	5.3	7.6	7.5	*15.9*
real										
regular	11.1	19.0	33.2	8.7	16.7	25.1	8.2	15.0	23.2	
irregular	13.5	39.7	42.2	15.8	38.5	37.4	8.8	18.5	18.7	*avg*
diff.	2.4	20.7	9.0	7.1	21.8	12.3	0.6	3.5	-4.5	*8.1*

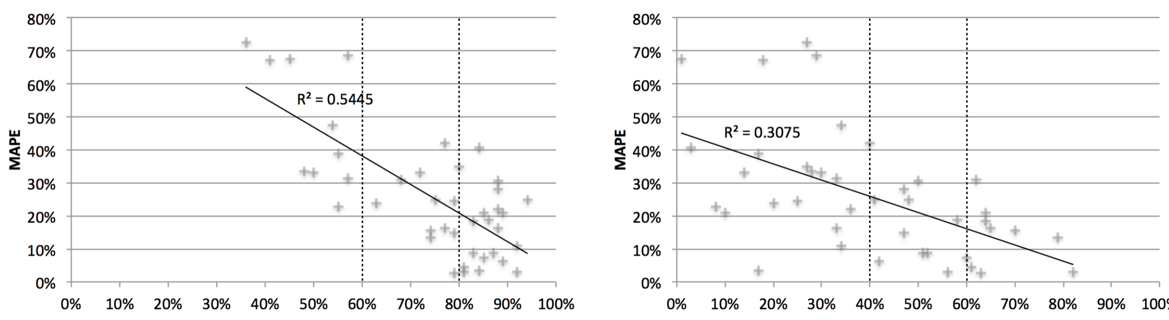
First of all, notice that the EDM-1 results appear to be biased for irregular projects (unrealistically high values, indicated by a big M). This bias was also observed by Batselier & Vanhoucke (2015b), who eliminated the projects with an initial period of very low PV-accrue in order to remove the biased results for EDM-1 and thus allow a correct assessment of the influence of project seriality on time forecasting accuracy. Since a project with an initial period of very low PV-accrue is an irregular project (see section 2.2 and more specifically figure 3b), the existence of an at that time still undefined but apparently influential project characteristic - introduced in this paper as project regularity - was already supported by Batselier & Vanhoucke (2015b).

Again considering table 3, it can be observed that for both simulated and real forecasting and for each method (except for ESM-SCI(t)), the average MAPE of the irregular projects is higher than that of the regular projects. This indicates that project irregularity appears to have a clear adverse effect on time forecasting accuracy. To further support this assertion, the difference between the regular and irregular projects' MAPEs was calculated for each method. Further calculations reveal that the average accuracy decrease over all methods is respectively 15.9% and 8.1% according to simulated forecasting (without EDM-1) and real forecasting. Focusing on the more important latter result, a decrease of 8.1% can indeed be deemed significant as it is of the exact same magnitude as the real forecasting accuracy of the best performing EVM time forecasting method (i.e. ESM-1) for regular projects identified by Batselier & Vanhoucke (2015b), which was 8.2% MAPE.

Moreover, it can be observed that, in reality, the methods based on SCI and SCI(t) perform relatively worse compared to the other methods than what would have been expected from simulations, even - and especially - for regular projects. This leads to the idea that SCI and SCI(t) are not particularly fit as a performance factor for time forecasting. Indeed, one might see the possible bias in weighing the current cost performance as much as the current schedule performance when forecasting project duration, where common sense would probably add more weight to the latter.

It is also remarkable that the ESMs seem to suffer less from PV-irregularities than the other methods. Real forecasting results even indicate that the ESMs feel virtually no effect at all of project irregularity. A qualitative explanation is that PVMs and EDMs weigh EV against PV (i.e. focus on the cost dimension), whereas ESMs compare ES and AT (i.e. focus on the time dimension). For example, consider figure 3a, which presents a strongly irregular project (RI = 54%) from the considered database. It becomes apparent that, if this project were to fall behind schedule, the cost dimension could be subject to much larger deviations than the time dimension due to the existing jump in the PV-curve. These possible excessive deviations cause PVM and EDM forecasts to be significantly less accurate for irregular projects.

Vanhoucke & Vandevoorde (2007) proved that project seriality has an influence on EVM time forecasting accuracy based on a simulation study. Moreover, this statement was confirmed by the empirical study of Batselier & Vanhoucke (2015b). The fact that irregular projects had to be eliminated in the latter study in order to allow a correct assessment of the influence of project seriality on time forecasting accuracy suggests that project regularity has a more significant effect on the said accuracy than project seriality. Indeed, figure 4 shows that the relation between RI and time forecasting accuracy (i.e. a higher RI yields a lower MAPE and thus a higher accuracy) is stronger than that between SP and time forecasting accuracy (i.e. a higher SP yields a lower MAPE and thus a higher accuracy), as indicated by the respective R-squared values of 54.5% (figure 4a) and 30.1% (figure 4b). Note that the MAPE data points in figure 4 represent the average simulated forecasting outcomes over all time forecasting methods (with exception of the biased EDM-1) for all projects in the utilized database. The simulated forecasting results are considered here as they are more extensive than the real forecasting results (51 relevant projects compared to 23 projects) and are therefore better fit for the performed comparison.



(a) Project regularity (b) Project seriality

Figure 4: Comparison of the influence of project regularity and project seriality on EVM time forecasting accuracy

To conclude, the outcomes of the last paragraph further support the relevance and validity of project regularity as a metric for project characterization and indicate the strength of the novel project characteristic to provide a reliable indication of the expected accuracy of a certain EVM time forecasting method.

3.2.2. Cost forecasting

In previous section, project regularity was shown to have a significant impact on time forecasting accuracy. Since the concept of project regularity is related to the shape of the PV-curve and PV covers the cost dimension just as well as the time dimension, it could be suspected that EVM cost forecasting also experiences an influence of project irregularities. This would be opposite to what was found for project seriality, which was shown to have no apparent effect on cost forecasting accuracy (Batselier & Vanhoucke, 2015b). Therefore, proving that project regularity does exhibit such an effect would further endorse the validity and predictive strength of the novel project characteristic, also for cost forecasting. The required assessment can be performed by means of the results in table 4. Just like in section 3.2.1, we utilize the two-part categorization of regular versus irregular projects.

Table 4: Accuracies of EVM cost forecasting methods w.r.t. project regularity

	EAC								
MAPE [%]	1	CPI	SPI	SPI(t)	SCI	SCI(t)	0.8CPI+ 0.2SPI	0.8CPI+ 0.2SPI(t)	
simulated									
regular	0.8	0.9	9.7	10.0	10.5	10.8	4.5	3.7	
irregular	1.1	0.9	15.7	13.1	15.8	13.2	8.7	5.7	*avg*
diff.	0.3	0.0	6.0	3.1	5.3	2.4	4.2	2.0	*2.9*
real									
regular	6.5	5.7	13.4	12.3	16.8	14.9	5.8	5.9	
irregular	10.2	14.6	37.5	15.0	38.8	16.6	14.5	11.5	*avg*
diff.	3.7	8.9	24.1	2.7	21.9	1.7	8.7	5.6	*9.7*

For both simulated and real forecasting, the accuracy of the cost forecasts is always higher for the regular projects than for the irregular projects, and this for each method. More generally, for simulations, an average accuracy decrease of 2.9% for irregular projects with respect to regular projects is witnessed. However, this outcome might underestimate the actual effect of irregularities as the somewhat unrealistic simulated MAPEs for EAC-1 and EAC-CPI lower the result. Indeed, when considering the real forecasting results, we observe an enlarged accuracy drop of 9.7% over all forecasting methods, which is even more substantial than the 8.1% accuracy decrease witnessed for time forecasting (see section 3.2.1).

However, the real forecasting accuracy differences for EAC-SPI and EAC-SCI are quite large and, just like EAC-SPI(t) and EAC-SCI(t), these methods are not really relevant to our evaluation since they would probably never be recommended. Therefore, we only take into account the top 4 cost forecasting methods as identified by Batselier & Vanhoucke (2015b) (i.e. EAC-1, EAC-CPI, EAC-0.8CPI+0.2SPI, and EAC-0.8CPI+0.2SPI(t)) and obtain a more adequate average accuracy decrease of 6.7%. This is still a considerable decrease, as the average MAPE of the top 4 methods, when used for regular projects, is only 6% (Batselier & Vanhoucke, 2015b). Therefore, it can be stated that, just as for time forecasting, project regularity appears to have a significant influence on cost forecasting accuracy.

4. Combining project regularity and project seriality

In section 3.2.1, it was indicated that the regularity of a project could influence the observations concerning the impact of project seriality on time forecasting accuracy. Obviously, the same applies in the opposite direction, that is, the seriality of a project could affect the perceived influence of project regularity on forecasting accuracy. Therefore, the combined influence of project regularity and project seriality on time forecasting accuracy is assessed and presented in table 5.

The table consists of four subtables; two for the simulation results and two for the real forecasting outcomes. For all subtables, RI and SP are plotted against each other, both divided into three levels - low (L), medium (M) and high (H) - of project regularity and project seriality, respectively. Since a two-part categorization for project regularity (as was applied in section 3.2.1) would not be sufficient to identify a trend in the outcomes and would not allow the necessary correspondence with the project seriality

Table 5: Accuracies of EVM time forecasting methods w.r.t. project regularity and project seriality

simulated

		SP		
		L	M	H
RI	L	46.5	N/A	N/A
	M	24.5	22.0	15.7
	H	24.1	15.2	11.8

mean MAPE[%]*

		L	M	H
RI	L	35.4	-	-
	M	-	19.9	-
	H	-	-	15.7

mean MAPE[%]*

real

		SP		
		L	M	H
	L	23.4	N/A	N/A
	M	27.4	17.9	41.8
	H	22.5	12.6	11.9

mean MAPE[%]

		L	M	H
	L	24.2	-	-
	M	-	23.5	-
	H	-	-	20.1

mean MAPE[%]

* With exclusion of the EDM-1 results to avoid biasing effects.

categorization, project regularity is subdivided into three categories here, as was already announced in section 2.2. More specifically, for project regularity, L corresponds to highly irregular, M to mildly irregular, and H to regular projects. Similarly, for project seriality, L, M and H refer to parallel, serial-parallel and serial projects, respectively. Recall that the limit values for the project seriality and project regularity categorizations were defined in the respective sections 2.1 and 2.2, and moreover, that they were represented by vertical dotted lines in figure 4. Furthermore, figure 5 is introduced to visualize the distribution of the considered projects according to the combined regularity/seriality categorization.

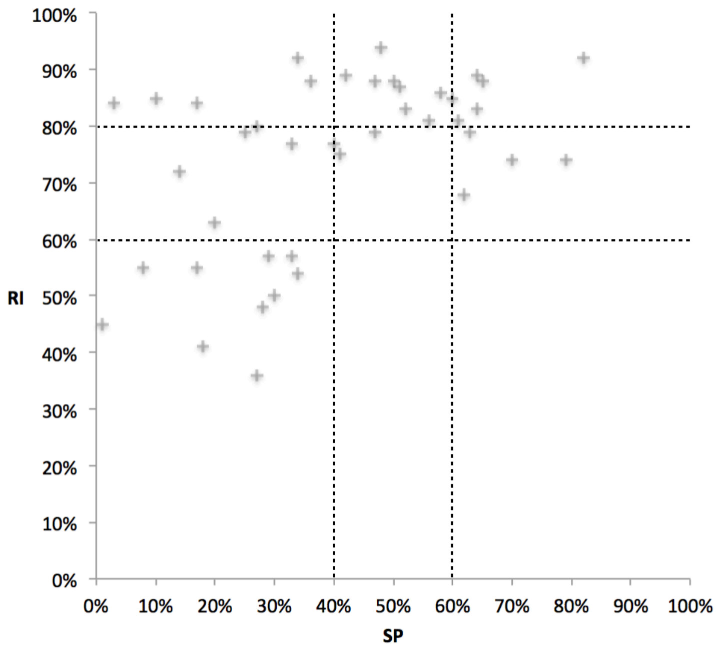


Figure 5: Combined project regularity/seriality categorization of the considered projects

For both simulated and real forecasting, the displayed MAPEs in table 5 are the means over all time forecasting methods. As such, we can evaluate the average EVM time forecasting accuracy for a variety of projects with different regularity and seriality characteristics. Note that for simulated forecasting, we do not include the results for EDM-1, again, to avoid biasing effects. Initially, we thus obtain the two upper subtables of table 5 - one for simulation and one for real forecasting - showing the mean time forecasting accuracies for nine categories of projects with different combinations of project regularity and project seriality levels.

Now consider these two upper subtables. For both simulated and real results, we observe a tendency of increasing forecasting accuracy when fixing the level of one characteristic and letting the other characteristic vary from low to high. For simulated forecasting, this applies for each row and each column. For real forecasting, however, not every combination yields equally nice results (e.g. when setting RI to medium). The main reason is that the utilized project database is divided into nine categories, thus creating strongly reduced samples for each category, even to the extent that two categories do not contain any projects at all. This effect is more tangible for real forecasting since only projects containing real tracking data are considered there.

In order to partially overcome the biasing effect of the smaller sample sizes and to eliminate the impact of outlier values to some extent, the results of the nine categories are grouped on the downward diagonal. This is done for both real forecasting - where it is most needed - and simulated forecasting, thus yielding the two bottom subtables of table 5. More specifically, each outcome on the downward diagonal of a bottom subtable is calculated as the average of all results in the corresponding row and column of the subtable directly above it, with a double counting of the result exactly on the diagonal. From the two bottom subtables, the combined influence of project regularity and project seriality on time forecasting accuracy indeed appears, for both simulated and real forecasting.

Referring back to the two upper subtables of table 5, another interesting observation can be made. That is, for both simulated and real forecasting, the values on the respective downward diagonals are already perfectly ordered and even show a clearer mutual difference than was to be observed for the bottom subtables. Furthermore, the apparently stronger forecasting accuracy relations for projects with regularity and seriality characteristics of the same level (i.e. RI and SP both low, both medium, or both high) could be conceived as an indication that combining RI and SP might yield a new indicator that produces better indications of the expected time forecasting accuracy than the individual indicators. Or alternatively stated, the combined RI/SP-indicator could display a stronger relationship with time forecasting accuracy than the individual indicators do. The explicit development of such a combined indicator is left to future research. Nevertheless, preliminary studies suggest that, for example, a combination of 0.8 RI + 0.2 SP should yield improved results. An evaluation similar to that of figure 4 indeed shows an increased R-squared value of 59.2% for 0.8 RI + 0.2 SP compared to the 54.5% for RI. Consequently, previous statements indicate that there is a combined and mutual influence of project regularity and project seriality on time forecasting accuracy.

Notice that the above discussion only concerned time forecasting. Indeed, since project seriality as an individual characteristic was shown to have no apparent impact on cost forecasting accuracy, it is redundant to assess its influence in combination with project regularity for the cost dimension. This statements again suggests the dominance of project regularity over project seriality with respect to project characterization potential.

5. Conclusions

In this paper, a novel project characteristic is introduced which can accurately characterize a project and provide reliable indications of the expected EVM forecasting accuracy for that project. This novel characteristic, called project regularity, reflects the value accrue within a project and is therefore related to the course of the PV-curve. More specifically, the level of project regularity can be expressed by the regular/irregular-indicator RI, which measures the deviation of the project's PV-curve from a perfectly linear curve (which represents a perfectly regular project; RI = 100%). Moreover, the RI is constructed completely similar to the well-known serial/parallel-indicator SP that is used to express project seriality (Tavares et al., 1999; Vanhoucke et al., 2008), which facilitates its implementation.

An evaluation of the influence of project regularity on EVM forecasting accuracy was performed, based on the most commonly used methods for both time and cost forecasting as identified by Vanhoucke (2012b) and making use of the real-life project database of Batselier & Vanhoucke (2015a). For both forecasting dimensions, a decreasing level of project regularity (i.e. a PV-curve that shows larger deviations from a perfectly linear curve, thus a lower RI) appears to have a significant negative impact on the forecasting accuracy. This implies that project managers could try to build up their project in a more 'regular' way (i.e. strive for a more linear PV-curve) in order to increase the expected accuracy of the time and cost forecasts performed during project execution. However, for time forecasting, the ES-based approach (i.e. ESM) already seems to cope quite well with irregular projects - certainly better than the alternative methods (i.e. PVM and EDM) - and is thus recommended for projects with a low RI.

Also for time forecasting, project regularity exhibits a stronger influence on forecasting accuracy than project seriality. Furthermore, for cost forecasting, project seriality has no apparent effect on the forecasting accuracy (Batselier & Vanhoucke, 2015b), whereas project regularity does. These results are important, as project seriality can be regarded as the most influential input parameter (topological indicator) for project network generators up to now (Vanhoucke et al., 2008). It can thus be concluded that project regularity can provide a more reliable indication of the expected accuracy of a certain EVM forecasting method than the widely used characteristic of project seriality. Moreover, this implies that project regularity can also be useful as an input parameter for project network generators, certainly as it characterizes both the schedule and cost aspects of a project (whereas project seriality only considers the schedule aspects).

Furthermore, the combined influence of project regularity and project seriality on time forecasting accuracy was assessed, as both characteristics showed an individual influence in the time forecasting case and could thus affect one another. Such a combined impact was indeed observed, suggesting that a combined RI/SP-indicator could produce better prognoses of the expected time forecasting accuracy than the individual indicators. The development of such a combined indicator, however, is left to future research. Furthermore, the apparent influence of project regularity on time and cost forecasting accuracy could further be validated by a large and more detailed simulation study (e.g. with incorporation of possible management actions as in Vanhoucke (2010b, 2011, 2012a)). Finally, we could also investigate the impact of project regularity on the performance of the schedule risk analysis (SRA) project control system. This research would be in line with the studies performed by Vanhoucke (2010b, 2011, 2012a), which considered project seriality and assessed its effect on the performance of both SRA and EVM. By applying a similar approach for project regularity, a guideline on which control technique - SRA or EVM - best to use under certain project regularity conditions could be provided to project managers.

Acknowledgements

We acknowledge the support provided by the "Nationale Bank van België" (NBB) and by the "Bijzonder Onderzoeksfonds" (BOF) for the project with contract number BOF12GOA021. Furthermore, we would also like to thank Gilles Bonne, Eveline Hoogstoel and Gilles Vandewiele for their efforts in developing PMConverter.

References

Agrawal, M., Elmaghraby, S., & Herroelen, W. (1996). DAGEN: A generator of testsets for project activity nets. *European Journal of Operational Research*, *90*, 376–382.

Anbari, F. (2003). Earned value project management method and extensions. *Project Management Journal*, *34*, 12–23.

Batselier, J., & Vanhoucke, M. (2015a). Construction and evaluation framework for a real-life project database. *International Journal of Project Management*, *33*, 697–710.

Batselier, J., & Vanhoucke, M. (2015b). Empirical evaluation of earned value management forecasting accuracy for time and cost. *Journal of Construction Engineering and Management*, *141*, 05015010.

Christensen, D. (1998). The costs and benefits of the earned value management process. *Acquisition Review Quarterly*, *Fall*, 373–386.

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

Cioffi, D. (2006). Completing projects according to plans: An earned-value improvement index. *Journal of the Operational Research Society*, *57*, 290–295.

Demeulemeester, E., Vanhoucke, M., & Herroelen, W. (2003). RanGen: A random network generator for activity-on-the-node networks. *Journal of Scheduling*, *6*, 17–38.

Egnot, J. (2011). Earned progress management - A unified theory of earned value & earned schedule concepts. *The Measurable News*, *Issue 4*, 8–24.

Elshaer, R. (2012). Impact of sensitivity information on the prediction of project's duration using earned schedule method. *International Journal of Project Management*, *31*, 579–588.

Fleming, Q., & Koppelman, J. (2010). *Earned value project management*. (4th ed.). Project Management Institute, Newtown Square, Pennsylvania.

Hall, N. (2012). Project management: Recent developments and research opportunities. *Journal of Systems Science and Systems Engineering*, *21*, 129–143.

Henderson, K. (2007). Earned schedule: A breakthrough extension to earned value management. In *PMI Asia Pacific Global Congress Proceedings - Hong Kong*.

Jacob, D., & Kane, M. (2004). Forecasting schedule completion using earned value metrics? Revisited. *The Measurable News*, *Summer*, 1, 11–17.

Kao, E., & Queyranne, M. (1982). On dynamic programming methods for assembly line balancing. *Operations Research*, *30*, 375–390.

Kim, E., Wells, W., & Duffey, M. (2003). A model for effective implementation of earned value management methodology. *International Journal of Project Management*, *21*, 375–382.

Kolisch, R., Sprecher, A., & Drexl, A. (1995). Characterization and generation of a general class of resource-constrained project scheduling problems. *Management Science*, *41*, 1693–1703.

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

Marshall, R. (2007). The contribution of earned value management to project success on contracted efforts. *Journal of Contract Management*, *Summer*, 21–33.

Mastor, A. (1970). An experimental and comparative evaluation of production line balancing techniques. *Management Science*, *16*, 728–746.

Meredith, J., & Mantel, S. (2003). *Project management: A managerial approach*. (5th ed.). Wiley, New York.

PMI (2008). *A guide to the Project Management Body of Knowledge (PMBOK guide). 3rd Edition*. Newtown Square, PA: Project Management Institute.

Rujirayanyong, T. (2009). A comparison of three completion date predicting methods for construction projects. *Journal of Research in Engineering and Technology*, *6*, 305–318.

Schwindt, C. (1995). A new problem generator for different resource-constrained project scheduling problems with minimal and maximal time lags. *WIOR-Report-449. Institut für Wirtschaftstheorie und Operations Research, University of Karlsruhe*, .

Tavares, L. (1999). *Advanced models for project management*. Kluwer Academic Publishers, Dordrecht.

Tavares, L., Ferreira, J., & Coelho, J. (1999). The risk of delay of a project in terms of the morphology of its network. *European Journal of Operational Research*, *119*, 510–537.

Van De Velde, R. (2007). Time is up: Assessing schedule performance with earned value. *PM World Today*, *9*, 1–10.

Vandevoorde, S., & Vanhoucke, M. (2006). A comparison of different project duration forecasting methods using earned value metrics. *International Journal of Project Management*, *24*, 289–302.

Vanhoucke, M. (2010a). *Measuring Time - Improving Project Performance using Earned Value Management* volume 136 of *International Series in Operations Research and Management Science*. Springer.

Vanhoucke, M. (2010b). Using activity sensitivity and network topology information to monitor project time performance. *Omega The International Journal of Management Science*, *38*, 359–370.

Vanhoucke, M. (2011). On the dynamic use of project performance and schedule risk information during project tracking. *Omega The International Journal of Management Science*, *39*, 416–426.

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

Vanhoucke, M. (2012b). *Project Management with Dynamic Scheduling: Baseline Scheduling, Risk Analysis and Project Control* volume XVIII. Springer.

Vanhoucke, M. (2014). *Integrated Project Management and Control: First comes the theory, then the practice*. Management for Professionals. Springer.

Vanhoucke, M., Coelho, J., Debels, D., Maenhout, B., & Tavares, L. (2008). An evaluation of the adequacy of project network generators with systematically sampled networks. *European Journal of Operational Research*, *187*, 511–524.

Vanhoucke, M., & Vandevoorde, S. (2007). A simulation and evaluation of earned value metrics to forecast the project duration. *Journal of the Operational Research Society*, *58*, 1361–1374.