Available online at www.sciencedirect.com

ScienceDirect

International Journal of Project Management 33 (2015) 697 – 710

International Journal of Project Management

www.elsevier.com/locate/ijproman

Construction and evaluation framework for a real-life project database

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] Department of Management Science and Innovation, University College London, Gower Street, London WC1E 6BT, United Kingdom

Received 16 May 2014; received in revised form 18 August 2014; accepted 9 September 2014
Available online 6 October 2014

Abstract

In this paper, a real-life project database is created, outranking the existing empirical databases from project management literature in both size and diversity. To ensure the quality of the added project data, a database construction and evaluation framework based on the so-called project cards is developed. These project cards incorporate the concepts of dynamic scheduling and introduce two novel evaluation measures for the authenticity of project data. Furthermore, an overview of the constructed database leads to statements on the difference between planned and actual project performance and on the earned value management (EVM) forecasting accuracy. Moreover, the database is publicly available and can thus become the basis for many future studies related to project management, of which a few are suggested in this paper. To further support these studies, the database will continuously be extended utilizing the project cards. Furthermore, the project cards can also serve didactical purposes.

Keywords: Project management; Dynamic scheduling; Earned value management; Empirical database; Database construction framework

1. Introduction

The concept of *dynamic scheduling* was introduced by Uyttewaal (2005) and further extended by Vanhoucke (2012b) and incorporates the three dimensions of project management: baseline schedule, risk analysis, and project control. The *baseline schedule* expresses the planned course of a project and acts as a central point of reference for the two other dimensions. The search for the best possible baseline schedule under certain resource restrictions is known as the resource-constrained project scheduling problem (RCPSP). During the *risk analysis* phase, Monte Carlo simulations are used to generate activity durations and costs that deviate from their baseline values in order to assess the impact of these deviations on the time and cost objectives of the project. This technique is also called schedule risk analysis (SRA) and was first introduced by Hulett (1996). *Project control* then consists of the monitoring of the performance of a project in progress and comparing it to the baseline schedule with the aim of triggering corrective actions when the project's objectives are jeopardized. Earned value management (EVM), with the incorporation of the earned schedule (ES) concept, is the preferred technique for performing project control. This technique will more elaborately be discussed in Section 2.1.4.

In the last few years, there has been extensive high-quality research on each of the three dimensions of dynamic scheduling, which proves their relevance to modern-day project management. Concerning baseline scheduling, Hartmann and Briskorn (2010) provided an overview and classification of the most important extensions of the RCPSP. More recently, several project scheduling researchers focused on one particular extension of the RCPSP, namely the multi-mode RCPSP (MRCPSP), in which multiple execution modes are available for each activity in the project (Coelho and Vanhoucke, 2011; Deblaere et al., 2011; Elloumi and Fortemps, 2010; Van Peteghem and Vanhoucke, 2010, 2014; Wang and Fang, 2011, 2012). Elaborate research into the performance of SRA was conducted by Vanhoucke (2010a), while Trietsch et al. (2012) investigated the use of lognormal distributions for activity durations. Moreover, several new studies

* Corresponding author at: Faculty of Economics and Business Administration, Ghent University, Tweekerkenstraat 2, 9000 Ghent, Belgium. Tel.: +32 9 264 35 69.
E-mail addresses: jordy.batselier@ugent.be (J. Batselier), mario.vanhoucke@ugent.be (M. Vanhoucke).

http://dx.doi.org/10.1016/j.ijproman.2014.09.004
0263-7863/00/

are concerned with risk analysis in general (Hartono et al., 2014; Nasirzadeh et al., 2014; Taroun, 2014; Zhang and Fan, 2014; Zwikael et al., 2014). Research on project control has also been widespread, with an emphasis on the development of probabilistic forecasting techniques based on the EVM methodology (Caron et al., 2013; Kim and Reinschmidt, 2009, 2010, 2011; Lipke et al., 2009; Naeni et al., 2011; Narbaev and De Marco, 2014). Furthermore, several authors proposed to integrate risk analysis techniques with project control approaches (e.g. SRA with EVM) in order to enhance the decision-making process for taking corrective actions (Acebes et al., 2014; Colin and Vanhoucke, 2014; Elshaer, 2013; Pajares and López-Paredes, 2011; Vanhoucke, 2011, 2012a).

This last research trend has inspired the renaming of the concept of dynamic scheduling to *integrated project management and control* (Vanhoucke, 2014). The new term indeed better represents the integrative interaction between the three dimensions, especially between risk analysis and project control. Furthermore, it stresses the importance of project control. Indeed, controlling the project during its progress should be the main concern of a project manager. Therefore, further discussions on integrated project management and control will mainly refer to the project control dimension, and thus to EVM.

Previous studies have conducted research on integrated project management and control based on both generated (Agrawal et al., 1996; Boctor, 1993; Demeulemeester et al., 2003; Kolisch and Sprecher, 1996; Kolisch et al., 1995; Schwindt, 1995; Tavares, 1999; Van Peteghem and Vanhoucke, 2014; Vanhoucke et al., 2008) and real-life project data (Bright and Howard, 1981; Covach et al., 1981; Hecht, 2007; Henderson, 2003, 2005; Lipke, 2009; Riedel and Chance, 1989; Rujirayanyong, 2009; Tzaveas et al., 2010; Vandevoorde and Vanhoucke, 2006; Zwikael et al., 2000). However, in the latter case, the data set was often of insufficient size and always of inadequate diversity for obtaining generalizable conclusions, as Table 1 indicates. Regarding diversity, the data sets consisted of mutually similar projects situated within the same sector and mostly even within the same company. This is, for example, also the case for the study of Riedel and Chance (1989) listed in Table 1, for which a sufficient number of projects was considered, however, all originating from one and the same US Air Force division.

Note that size and diversity can indeed be identified as the two main quality-determining properties of a database (Christensen et al., 1995; Vanhoucke, 2011). Consequently, many authors have expressed the need for performing integrated project management and control research – in particular research on project control – on a large and diverse data set of real-life projects (Henderson, 2003, 2004, 2005; Lipke, 2009, 2013; Tzaveas et al., 2010; Vanhoucke, 2011; Zwikael et al., 2000). However, no existing project database seems to meet these requirements. Therefore, the main goal and contribution of this paper lies in the creation of a large and diverse database consisting of qualitative real-life project data that can become the basis for multiple studies related to one or more of the dynamic scheduling dimensions. Moreover, to ensure that the project data are indeed qualitative and sufficiently diverse, a framework for database construction and evaluation will be introduced.

Table 1
Chronological overview of project control studies using real-life project data.

Paper	# Projects	# Sectors (companies)
Bright and Howard (1981)	11	1 (1)
Covach et al. (1981)	17	1 (1)
Riedel and Chance (1989)	56	1 (1)
Zwikael et al. (2000)	12	1 (1)
Henderson (2003)	6	1 (1)
Henderson (2005)	1	1 (1)
Vandevoorde and Vanhoucke (2006)	3	1 (1)
Hecht (2007)	1	1 (1)
Lipke (2009)	16 [a]	2 (2)
Rujirayanyong (2009)	2	1 (1)
Tzaveas et al. (2010)	1	1 (1)
This paper (2014)	51	5 [b] (47)

[a] These projects are the same as those used by Zwikael et al. (2000) and Henderson (2003).
[b] There are 5 sectors if construction is perceived as one overarching industry. However, if construction is divided into its constituting subsectors (infra Fig. 1a), 9 different sectors can be identified.

Note that none of the earlier data sets mentioned in Table 1 have been made publicly available. For some of the smallest data sets (Rujirayanyong, 2009; Tzaveas et al., 2010; Vandevoorde and Vanhoucke, 2006), somewhat more extensive project data are provided in the paper itself, but never down to the activity level. In contrast, our entire database – including the specific activity data for all projects – can be consulted at www.or-as.be/research/database (OR-AS, 2014a).

Summarized, the contribution of this paper is fourfold:

- A real-life project database is created, outranking the existing empirical databases from project management literature in both size and diversity. Moreover, the database will continuously be extended and is made publicly available, so that it can become the basis for many future studies related to project management.
- A database construction and evaluation framework is developed, based on the so-called project cards (see Section 2.1). Through implementation of these project cards, the quality of the added data can be guaranteed.
- The project cards can also serve didactical purposes. More concretely, they can be used for introducing students to the three dimensions of integrated project management and control – all incorporated in the project cards (see Sections 2.1.2 to 2.1.4) – in a more practical context.
- The performed database evaluation yields interesting initial results (e.g. on the difference between planned and actual project performance and on EVM forecasting accuracy; see Section 3), which can be further explored in a future research.

The specific outline of the rest of the paper is as follows. In Section 2, the framework for database construction and evaluation based on the project cards is introduced. Furthermore, it is presented how these project cards can be applied in order to ensure qualitative database extension. Section 3 then provides an overview of the general characteristics of the constructed database, based on the information that is included in the project cards. Finally, in Section 4, general conclusions are drawn and actions for further research are suggested.

2. Methodology

This section starts with the introduction of a data categorization and evaluation tool called project cards and the definition of the associated metrics of project authenticity and tracking authenticity (Section 2.1). Subsequently, in Section 2.2, it is presented how these project cards can be used as a framework for database construction and evaluation with the aim of ensuring the diversity and authenticity of the real-life project database.

2.1. Introduction of project cards

With the aim of obtaining a clear overview of the projects in the database, the so-called project cards were introduced. A project card summarizes the whole of the project data (i.e. general characteristics, risk information and tracking data) for a certain project in an orderly and structured manner. Its main structure originates from the three dimensions of dynamic scheduling (or integrated project management and control) – baseline scheduling, risk analysis and project control – as presented by Vanhoucke (2012b, 2014). Although every project card possesses self-contained clarity, a brief presentation of the constituting elements is provided in the following subsections. The presentation of the project cards is logically construed according to the three dynamic scheduling dimensions, preceded by an introductory subsection which describes the header of a project card. For illustration, an example project card is included in Appendix A. Moreover, the project cards for all projects in the database are available on the supporting website (OR-AS, 2014a).

2.1.1. Header

The header of a project card comprises information such as the project's name, the sector in which the project is situated, the name of the data submitter, the submission date, and the name of the corresponding data file. Concerning this last aspect, all project data are available in the form of files from the project management software tool ProTrack (OR-AS, 2014b). On the supporting website (OR-AS, 2014a), the project data are also available in this format. Furthermore, the preceding code in the file name reflects the year in which the project data were received and the sequence number of the project. A short description of the concerned project is also included just below the project card header.

It is important to realize that the project cards not only allow an easier categorization of the projects, they also enable the evaluation of the completeness and quality of the received data.

More concretely, the extent to which each of the three dynamic scheduling dimensions was covered by the project data is expressed by a three-level color code which is based on the 'traffic light approach' proposed by Anbari (2003). A green, yellow and orange color (indicated as the respective categories G, Y and O in upcoming figures) respectively indicates full, mediocre and rather poor completeness of data. Materializing the approach for the three dynamic scheduling dimensions individually, full completeness is achieved for the baseline schedule when resources and costs were included, for the risk analysis when non-standard risk distribution profiles for activity durations were defined, and for project control when tracking data originated from user input (instead of from simulations).

Furthermore, the concept of *project authenticity* was introduced. We say that full project authenticity is achieved when activity, resource and (baseline) cost data were all obtained directly from the actual project owner. Full authenticity of data thus implies that the data collector did not make any personal assumptions regarding the relevant data types. Moreover, project authenticity is assessed and represented in the same way as the dynamic scheduling dimensions, that is, using the three-level color code. More specifically, full authenticity corresponds to a green color, and for every mentioned data type that was not obtained from the actual project owner, the color level is reduced by one (first to yellow, then to orange).

2.1.2. Baseline schedule

The baseline schedule section of a project card provides some general information about the number of activities, the planned total duration, and the budgeted total cost of the project. Furthermore, the number of renewable and consumable (or non-renewable) resources identified for the project is mentioned. The topology of the project network is reflected by four indicators, namely activity distribution (AD), length of arcs (LA), topological float (TF), and the serial/parallel-indicator (SP). For an elaborate discussion of these topological indicators, the reader is referred to Tavares et al. (1999) and Vanhoucke et al. (2008).

Nevertheless, the last indicator in the list can be deemed the most important one, as previous research (Vanhoucke and Vandevoorde, 2007) already indicated that the SP has a significant influence on the performance of EVM as a project control technique, whereas similar results could not (yet) be found for the other topological indicators. More concretely, the SP describes a project's network structure in terms of how close the network is to a serial (or parallel) network. The indicator ranges from 0% to 100%, with SP = 0% indicating that all activities are in parallel and SP = 100% signifying that the project network is completely serial. Between these two extreme values, networks can be closer to either a serial or a parallel network. Moreover, a presentation of the SP-related characteristics of the projects in our database will be part of Section 3.

2.1.3. Risk analysis

Risk analysis is based on Monte Carlo simulations. In order to generate activity duration and cost deviations (with respect to the baseline values), risk distribution profiles have to be identified for the activity durations. In this study, the simulations are performed with the project management software tool ProTrack (OR-AS, 2014b) and make use of triangular risk distribution profiles that can be either symmetrical, skewed to the left or skewed to the right, according to the specific activity characteristics. It is assumed that the variable costs of an activity vary uniformly with the corresponding activity duration. In contrast, the fixed cost figures always remain constant. Furthermore, the standard symmetric profiles are assumed when the data providers did not include any information about the specific activity risk distributions. Moreover, 100 Monte Carlo simulation runs are performed for each project.

Recall that the goal of risk analysis is to assess the impact of activity duration and cost deviations on the time and cost objectives of the project. To measure the criticality and/or sensitivity of a project's activities, several sensitivity measures have been proposed, these are the criticality index (CI), the significance index (SI), the schedule sensitivity index (SSI), and three versions of the cruciality index (CRI). An elaborate discussion of these measures can be found in several research papers (Elmaghraby, 2000; PMI, 2004; Vanhoucke, 2010b; Williams, 1992).

2.1.4. Project control

The preferred methodology for performing project control is EVM, a technique that integrates the three critical project management elements of cost, schedule and scope. The reader can consult several works in which a detailed discussion of both the basic and more thoroughgoing concepts of EVM – like the various performance metrics – is performed (Anbari, 2003; Fleming and Koppelman, 2010; PMI, 2005; Vanhoucke, 2010b). The utility of EVM is not limited to the evaluation of the current project schedule and cost performance, moreover, the technique can also be used for predicting the eventual project duration and budget.

The most commonly used EVM time and cost forecasting methods can be identified from Vanhoucke (2012b). More concretely, it concerns nine time and eight cost forecasting methods. For time forecasting, the nine methods can be grouped into three overarching methodologies, being the planned value method (PVM) by Anbari (2003), the earned duration method (EDM) by Jacob and Kane (2004), and the earned schedule method (ESM) by Lipke (2003). All the methods, also those for cost forecasting, differ in the sense that they rely on a specific performance factor (PF) that expresses a certain assumption about the expected performance of future work.

All nine time and eight cost forecasting methods are presented on the project cards. However, they are not only presented, they are also evaluated based on their accuracy. More specifically, the forecasting methods' accuracy is assessed through the mean absolute percentage value (MAPE) and mean percentage error (MPE). For a calculation example of these measures in an EVM setting, the reader is referred to Vanhoucke (2010b). Moreover, the MAPE has been used in multiple other studies on project forecasting accuracy (Elshaer, 2013; Rujirayanyong, 2009; Vanhoucke and Vandevoorde, 2007).

Furthermore, there are two types of forecasting to be identified: simulated forecasting and real forecasting. *Simulated forecasting* is the type of forecasting that is based on the Monte Carlo simulations that were discussed in Section 2.1.3. Since it was mentioned there that 100 simulation runs are performed for each project, the simulated forecasting results presented on the project cards are the average over these 100 runs. Moreover, the results for every individual Monte Carlo run are based on simulated progress data over 20 tracking periods. This implies that the tracking periods for simulation are assigned a length of $\lfloor PD/20 \rfloor$, with PD being the planned duration of the project.

It is important to realize that for projects that have been generated by a project network generator (Agrawal et al., 1996; Demeulemeester et al., 2003; Kolisch et al., 1995; Schwindt, 1995; Tavares, 1999; Vanhoucke et al., 2008) and/or are part of a benchmark dataset (Boctor, 1993; Kolisch and Sprecher, 1996; Van Peteghem and Vanhoucke, 2014), one can only perform the above described simulated forecasting due to the inherent absence of real tracking data for those projects. In contrast, for real-life projects that include tracking information obtained directly from the actual project owner, one can also obtain *real forecasting* results. The accuracy of a real time or cost forecast is assessed by comparing the forecasted values, obtained from a certain forecasting method, with the actual final project duration or cost, and this at predefined times throughout the project (i.e. at the end of each tracking period). This approach for accuracy assessment actually corresponds to the calculation of the MAPE.

One should note that real forecasting accuracy can only be calculated *after* the project has ended. On the other hand, simulated forecasting can already be performed *before* the project starts, meaning that the corresponding accuracy is known prior to project launch. Also notice that simulated forecasting can just as well be conducted for real-life projects, as the only requirement is the definition of risk distribution profiles for the individual activity durations. Hence, previous statements point out the possibility of comparing simulated and real forecasting results for real-life projects with authentic tracking data. This comparison has not yet explicitly been made in literature and can therefore be identified as an interesting topic for future research.

At the end of the previous paragraph, there was an indication of the need for projects with authentic tracking data. Therefore, in order to assess whether or not tracking data are authentic, the concept of *tracking authenticity* was introduced. We say that full tracking authenticity is achieved when the tracking data that were obtained from the project owner include actual activity start dates, durations and costs. Furthermore, tracking authenticity is evaluated according to the same color code-based approach as presented for project authenticity (see Section 2.1.1).

2.2. Application of project cards

In addition to the requirement of size, the aim was to create a diverse database consisting of qualitative (i.e. authentic) real-life project data. To realize this objective, the project cards were applied as a database construction and evaluation framework.

The diversity and authenticity properties of the database are shown in Fig. 1. More specifically, Fig. 1a to c expresses the database diversity with respect to sector, planned duration, and budget at completion, respectively. Fig. 1d displays the level of project authenticity of the projects in the database, using the three-level color code – green (G), yellow (Y) and orange (O) – as described in Section 2.1.1. The tracking authenticity (Fig. 1e) is presented in the same way, only that one category (N/A) is added to represent the projects for which no real tracking data were provided.

Notice that (some of) the bars in the subfigures of Fig. 1 consist of a lighter base and a darker top. The lighter parts represent the initial database of 33 projects which were received during the period 2011–2012, whereas the darker top parts show how the database was extended after the introduction of the project cards. Note that all data were obtained from the master students in Civil Engineering and Business Engineering of Ghent University. More specifically, these students have performed case studies at a wide variety of national (i.e. Belgian) and international companies for the completion of a master's thesis or a project management course, all under the careful supervision of the academic staff.

The project cards enabled the evaluation of the initial database according to the criteria of diversity and authenticity. In that way, gaps in the database characteristics could be identified. In this particular case, an obvious gap was the underrepresentation of projects with real tracking data. Fig. 1e indicates that only five projects comprised tracking information (category G, Y and O together) and that for only two of those the information was completely authentic (category G only). Furthermore, we strived to only extend the database with fully authentic data (category G in Fig. 1d), thus assuring the quality of the database. This specific need for fully authentic projects with complete tracking information was communicated to the data providers (i.e. the master

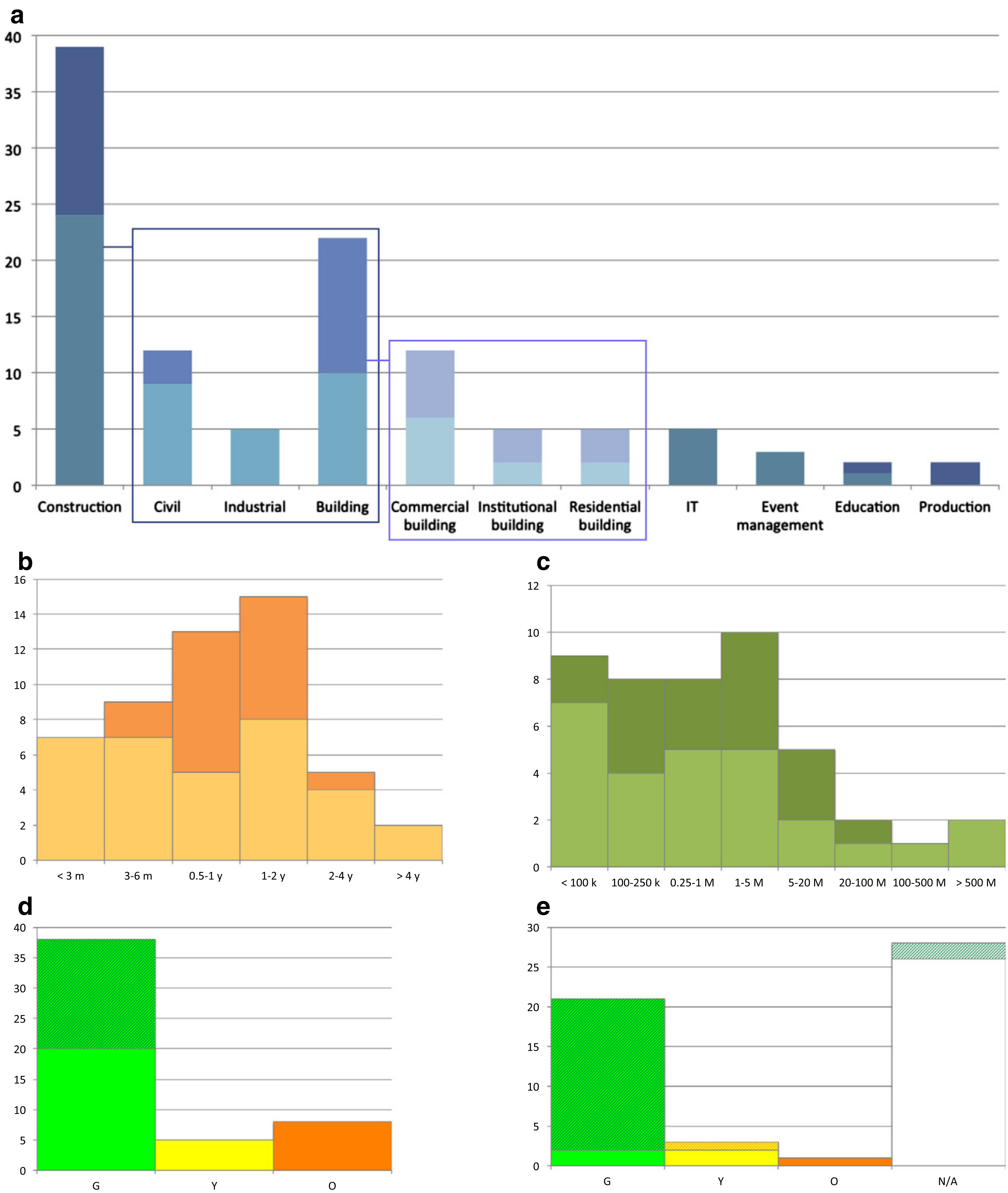

Fig. 1. Diversity and authenticity of the database.

students) and translated into the more visual and comprehensible demand that project control and both project and tracking authenticity should receive a green color on the project card. The respective requirements for 'green' elements are defined on the project card itself (see Appendix A), so that the students were provided with all the information necessary to gather the requested data. In this view, the project cards also exhibit a didactical utility and can be said to reflect the concept of 'education feeding research', and vice versa.

The methodology described in the previous paragraph was applied and proved to be successful. Indeed, only fully authentic projects were added to the database (Fig. 1d), and moreover, all (but one) of them comprised complete tracking information (Fig. 1e). Also notice that two existing projects were 'upgraded' by the data providers (see shaded part of rightmost bar of Fig. 1e) after the latter had received a project card with the requirements for new project data. More specifically, real tracking information was added for both projects, which was in accordance with our primary demand for database extension. This event provides an indication of the benefits of using the project cards for database construction.

From the above discussions, the usefulness of the project card methodology as a framework for qualitative database extension appears. Moreover, database extension can be perceived as an ongoing dynamic process. That is, the requirements for new projects can change over time. In our specific case for example, one could argue that the database diversity is somewhat compromised because of the preponderance of projects situated in the broad construction sector[1] (Fig. 1a). Therefore, we could start a new database extension with an adjusted focus on adding non-construction projects, which of course would still need to be fully authentic and comprise real tracking information. Again, the project cards could be utilized to express and communicate these revised needs. It lies in our intention to effectively implement the dynamical process for database extension as presented here, so that our database can continue to grow in both size and diversity, and increasingly generalizable conclusions could be drawn (Christensen et al., 1995).

3. Overview

This section provides an overview of the global characteristics of the constructed database. This is not only useful for discovering gaps in the current project data so that the relevance and quality of future database extensions can be guaranteed (i.e. a continuation of Section 2.2), but also for the identification of more general research topics related to project management. As already mentioned, the data of all individual projects in the database are available online (OR-AS, 2014a). Moreover, the outcomes discussed here can all be retrieved from the project cards.

Following the structure of the project cards, we start with the assessment of the three dynamic scheduling dimensions for the projects in our database. The results are shown in Fig.2. For the explanation of the applied color-based evaluation methodology, we refer to Section 2.1.1.

Fig. 2a indicates that the baseline schedule of the projects in our database is overall fairly complete, as green (G) is the dominant color. However, still about half of the baseline schedules are not fully complete. This is mainly due to the absence of resource information in the projects. Indeed, for only 27 of the 51 projects in the database, renewable resources were identified. Moreover, only one of these projects also comprises consumable resources. Therefore, we pursue to include resource information for newly added projects.

Concerning the risk analysis dimension, it can be observed that all projects in the database are at least of medium completeness (Fig. 2b). The prevalence of mediocre completeness (i.e. category Y) is due to the absence of customized risk distributions for activity durations in many projects. In these cases, the Monte Carlo simulations are based on the standard symmetric activity duration profiles. In the future, we therefore aim to extend the database with projects containing customized activity duration distributions.

From Fig. 2c, one can observe that the database contains slightly more projects with real tracking data (i.e. category G) than without tracking data (i.e. category O). Note that all projects that were added to the database after the introduction of the project cards indeed comprised real tracking data. This could already be seen from Fig. 1e. Obviously, this evolution will be maintained for future database extensions.

Now consider the size of the database. It was already indicated in Table 1 that the current database comprises 51 projects. The evolution of the database size in the past year and a half is shown in Fig. 3. Moreover, it is our intention to sustain the steady increase in database size in the future.

As the authenticity of the database was already discussed in Section 2.2, we continue this overview with the evaluation of the database diversity. Fig. 1a indicates that the database covers five main sectors: construction; IT; event management; education; and production. Moreover, the last sector was only added after – and thanks to – the introduction of the project cards. Furthermore, it clearly shows that construction is the dominant sector, which was already implied in the last paragraph of Section 2.2. However, the construction industry is very wide and comprises various subdivisions that exhibit mutually different characteristics. As such, we can break down the construction sector into civil, industrial and building construction. Furthermore, building construction can, in its turn, further be split into commercial, institutional and residential building. When considering these different areas of construction, the sector-related diversity of our database proves acceptable.

Fig. 1b and c[2] respectively shows that the projects' planned duration ranges from less than three months to more than four years and that their budget at completion varies from less than € 100,000 to over € 500,000,000 (with a peak project budget of up to 5 billion euros). These ranges indicate that the diversity of our database concerning project duration and budget is also guaranteed. Furthermore, the average project in the database

[1] According to Tzaveas et al. (2010), the prevalence of construction projects should, however, not be regarded as a weakness, as earlier studies on EVM – and project control by extension – were mainly limited to IT and high tech projects.

[2] Note that the total number of projects in Fig. 1c does not sum up to 51, but only to 45. The reason is that for 6 of the projects in the database, no cost data were included.

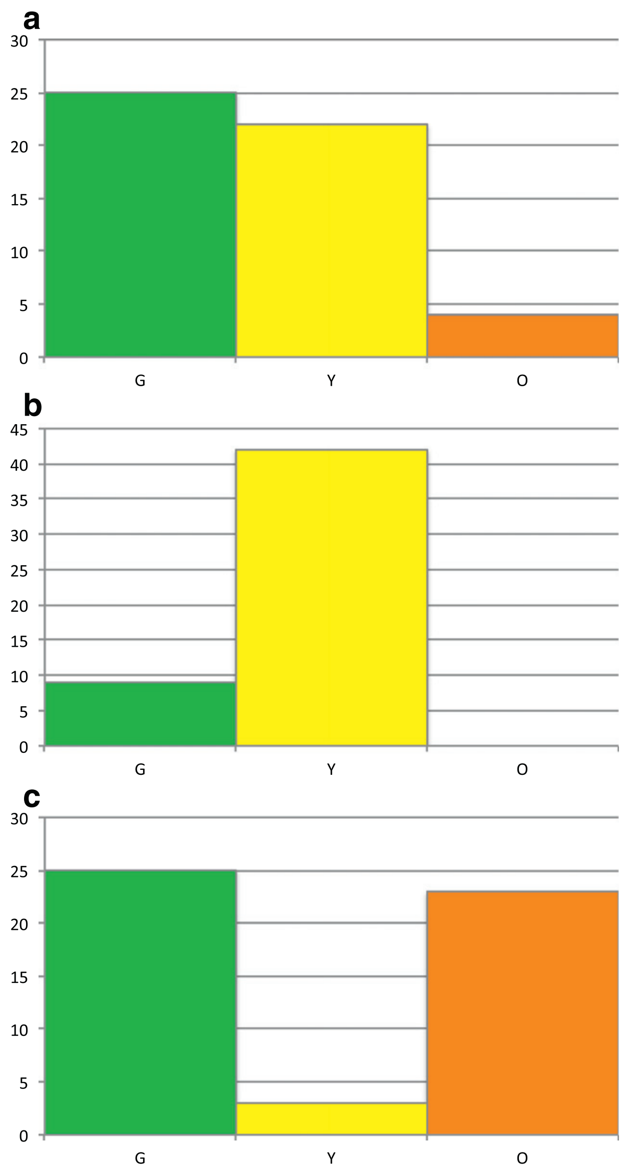

Fig. 2. Completeness of the dynamic scheduling dimensions.

has a planned duration of about ten months and a budget at completion of just under € 130,000,000.

Another characteristic of the average project is that it consists of 85 activities. A distribution of the projects' activity quantum is displayed in Fig. 4a. Note that the number of activities that has been defined for a certain project is not necessarily related to the size of the project (with size expressed in terms of planned duration and/or budget at completion). Indeed, the number of activities within a project primarily depends on the level to which the project has been broken down (this is related to the work breakdown structure or WBS of the project (PMI, 2004)). More specifically, one and the same project can be broken down into hundreds of specific tasks, but it can also be divided into only a dozen very general work packages. Although the chosen division of a project into its constituting activities does not intrinsically change the project, it can influence the project control outcomes and thus the possible management actions. The nature and importance of this effect should be assessed in future studies.

Another factor that can affect the project control results is the chosen tracking frequency, or more specifically, the resulting number of tracking periods (Fig. 4b). Note that the projects presented in Fig. 4b correspond to the 25 projects for which real tracking data were included (see category G in Fig. 2c). For these projects, we observed that project monitoring was performed over an average of 19 tracking periods. Of course, the number of tracking periods is partly related to the planned (and actual) duration of the project. Moreover, tracking periods do not necessarily have to be of equal length (e.g. one week, one month). In our database, for instance, there are 7 projects for which tracking was performed at irregular time intervals. These cases occur, for example, when project performance is assessed after the accomplishment of certain milestones, instead of on a regular weekly or monthly basis. The choice for one monitoring approach or another can indeed influence the tracking results. Therefore, the evaluation of the effect of tracking frequency on project control outcomes is also a subject for further research.

In Section 2.1.2, the serial/parallel-indicator SP was identified as an important topological indicator. Therefore, the SP-related characteristics of our database are presented here. It was already mentioned in Section 2.1.2 that completely parallel projects are indicated by an SP of 0% and completely serial projects by an SP of 100%, with project networks closer to either a parallel or a serial network lying between these two extremes. In order to allow a clear presentation of the SP-related characteristics of our database, the projects are subdivided into three SP-based categories as defined in Table 2.

It is clear from Table 2 that there are considerably more parallel projects in the database than there are serial projects. According to the simulation study of Vanhoucke and Vandevoorde (2007), these parallel projects should exhibit a lower project control performance (or more concretely, a lower EVM time forecasting accuracy) than the serial projects. The testing of this assertion on a real-life project database makes an interesting subject for future research.

Nevertheless, some preliminary results concerning the EVM forecasting accuracy – for both time and cost – can already be presented in the context of this paper. Table 3 displays the number of times that a certain time or cost forecasting method, based on the selection of Vanhoucke (2012b) and included in the project cards (see Appendix A), provides the most accurate predictions (based on MAPE calculations), and this over all projects in the database for which real tracking was performed. Furthermore, when two forecasting methods proved equally accurate for a certain project, both methods received a plus one in Table 3.

Table 3 seems to provide quite a clear first indication that the ES-based method with PF = 1 (i.e. ES-1) is the most accurate method for time forecasting and that performance factors 1 and CPI perform best for cost forecasting. However, keep in mind that these are the results of a fairly basic study. Future research should extend this study so that the preliminary results could be further supported.

For the projects with real tracking information, some interesting observations regarding the difference between the planned performance and the actual performance can be made, and this for both time and cost. We now refer to Table 4 and Fig. 5.

Table 4 and Fig. 5a show that about 60% of the projects finishes late, whereas only a third is completed ahead of schedule.

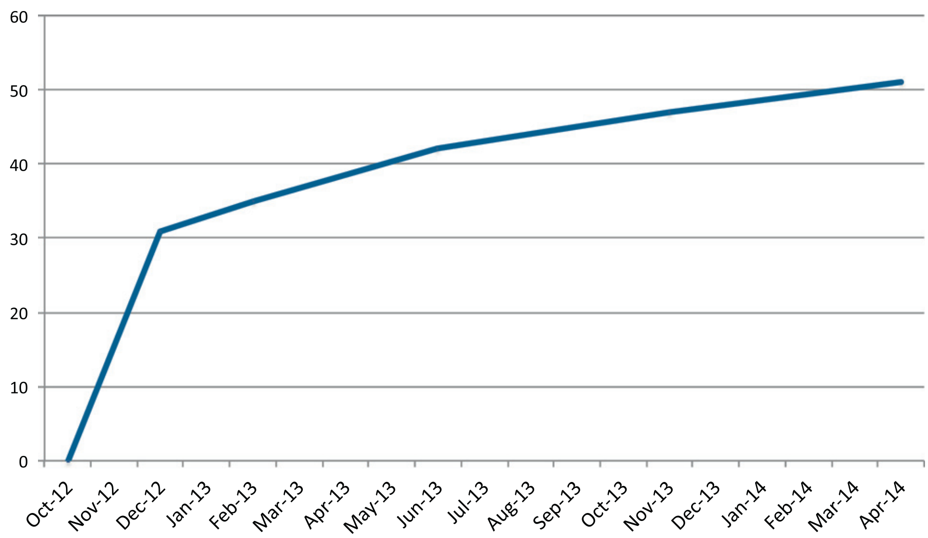

Fig. 3. Evolution of database size.

Moreover, when a project is late, on average, it ends with a delay of 19.1%. The average time savings of an early project are somewhat lower, namely 13.5%.

Regarding cost performance, almost as many projects come in under budget as over budget. However, the average cost reduction for an under budget project (8.9%) is significantly lower than the mean extra cost that is incurred for an over budget project (22.1%).

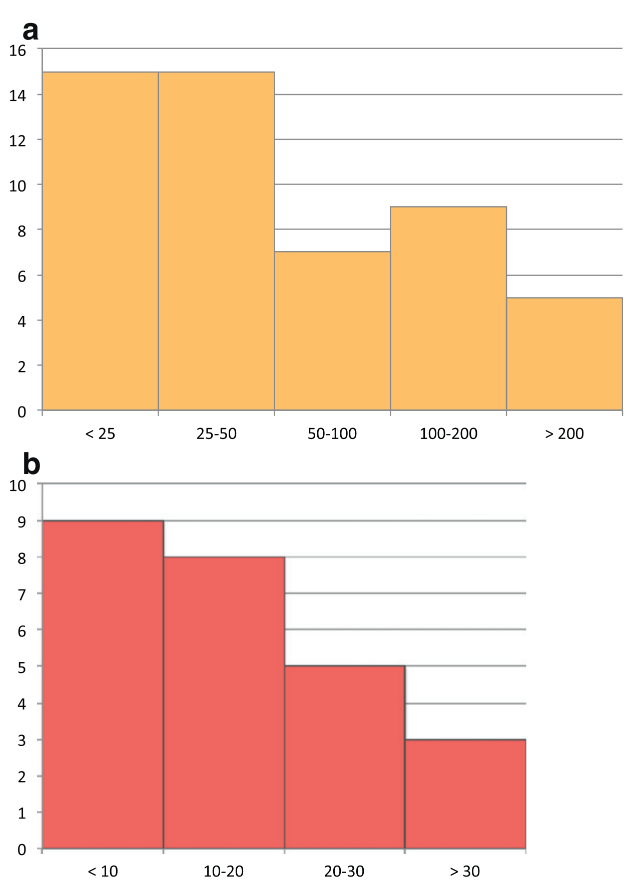

Fig. 4. Characteristics influencing project control results.

Fig. 5b indeed indicates that the magnitude of budget overruns can be much greater than that of cost savings.

Furthermore, the average project in the database is 6.6% late and is 7.3% over budget. In absolute numbers, this is reflected by a delay of just over 18 days (corresponding to the mean SV(t)) and an additional cost of about € 630,000 (corresponding to the mean CV). The fact that the average project is late and over budget can also be derived from the mean SPI(t) and CPI over all projects. These measures have a value of 0.97 and 0.96, respectively, which is both lower than 1 and thus indicating a time and cost performance that is worse than planned.

Just like the average project, a third of the individual projects in the database is also both late and over budget. Moreover, more than 70% of the projects finishes either behind schedule or with additional costs (or both). This implies that less than 30% of the projects do better than (or just as well as) planned. Furthermore, there appears to be a 65% chance for a project to also be late when it is already over budget, and vice versa.

4. Conclusions

The main objective of this paper was to create a large and diverse database consisting of qualitative real-life project data, thereby filling a much debated gap in project management literature. Therefore, a database comprising 51 projects from 47 different companies situated in various sectors was constructed, outranking previous real-life databases from literature in both size and diversity. The entire database is publicly available on the supporting website www.or-as.be/research/database (OR-AS, 2014a).

In order to ensure qualitative database extension regarding diversity and authenticity, a database construction and evaluation

Table 2
SP-based categorization of the projects.

Category	SP	# Projects
Parallel projects	[0%,40%[	26
Serial-parallel projects	[40%,60%]	13
Serial projects	]60%,100%]	12

Table 3
Identification of the most accurate EVM time and cost forecasting methods.

Time									
method-PF	PV-1	PV-SPI	PV-SCI	ED-1	ED-SPI	ED-SCI	ES-1	ES-SPI(t)	ES-SCI(t)
# best	1	0	0	4	0	1	*16*	1	2
Cost									
PF	1	CPI	SPI	SPI(t)	SCI	SCI(t)	0.8CPI+0.2SPI	0.8CPI+0.2SPI(t)	
# best	*8*	*8*	1	1	2	1	3	3	

framework based on the so-called project cards was introduced. A project card describes the specifics of a certain project and provides a tool for categorizing and evaluating these project data. More concretely, the project cards include the three much studied dimensions of dynamic scheduling (Vanhoucke, 2012b) or integrated project management and control (Vanhoucke, 2014) – baseline schedule, risk analysis and project control – and two newly defined characteristics named project authenticity and tracking authenticity. These characteristics form the basis for categorizing the different projects and for evaluating the quality and completeness of the project data. Furthermore, the project cards can also provide assistance for qualitative data acquisition – a merit which was demonstrated in this paper – and through this, can serve a didactical purpose ('education feeds research', and vice versa). Just like the project data itself, the project cards for all projects in the database can be consulted online (OR-AS, 2014a).

Furthermore, some interesting observations could be made from the database overview. First of all, regarding the difference between planned performance and actual performance, it was observed that the average project is both late (by 6.6%) and over budget (by 7.3%). Moreover, more than 70% of the projects finishes either behind schedule or with additional costs (or both), which implies that less than 30% of the projects actually performs according to plan or does better than planned. Furthermore, the chance that a project also ends late when it is already known to be over budget, and vice versa, can be estimated at 65%.

Also from the database overview, some preliminary results concerning the EVM time and cost forecasting accuracy could be obtained. More specifically, for time forecasting, the ES-based method with PF = 1 clearly appears to be the most accurate method. For cost forecasting, the best performing methods are those with performance factors 1 and CPI. However, these results are based on a fairly basic study. Therefore, future research should extend this study to further support the conclusions drawn here.

Other topics for future research were already identified in previous sections and are now summarized. First of all, the constructed real-life project database can form the basis for many future studies related to project management. Therefore, an obvious future activity is the continuous extension of the current real-life project database, so that the validity and generalizability of these future studies could ever increase. The database extension will of course be performed in a qualitative manner through the application of the project cards. More specific suggestions for future studies are the comparison of simulated (i.e. based on Monte Carlo simulations) and real (i.e. based on actual progress data) forecasting results (which is only possible for real-life project data), the assessment of the influence of the chosen number of activities and number of tracking periods on the project control outcomes (and thus on the possible management actions), and the validation of the assertion that the SP influences project control (i.e. EVM) performance (Vanhoucke and Vandevoorde, 2007) on a real-life project database.

Table 4
Actual time and cost performance of the projects with real tracking data.

Time	Early	On time	Late
# projects	8	2	14
avg early/late	13.5%		19.1%
Cost	**Under budget**	**On budget**	**Over budget**
# projects	10	0	11
avg under/over budget	8.9%		22.1%

Acknowledgments

We acknowledge the support provided by the "Bijzonder Onderzoeksfonds" (BOF) for the project with contract number BOF12GOA021.

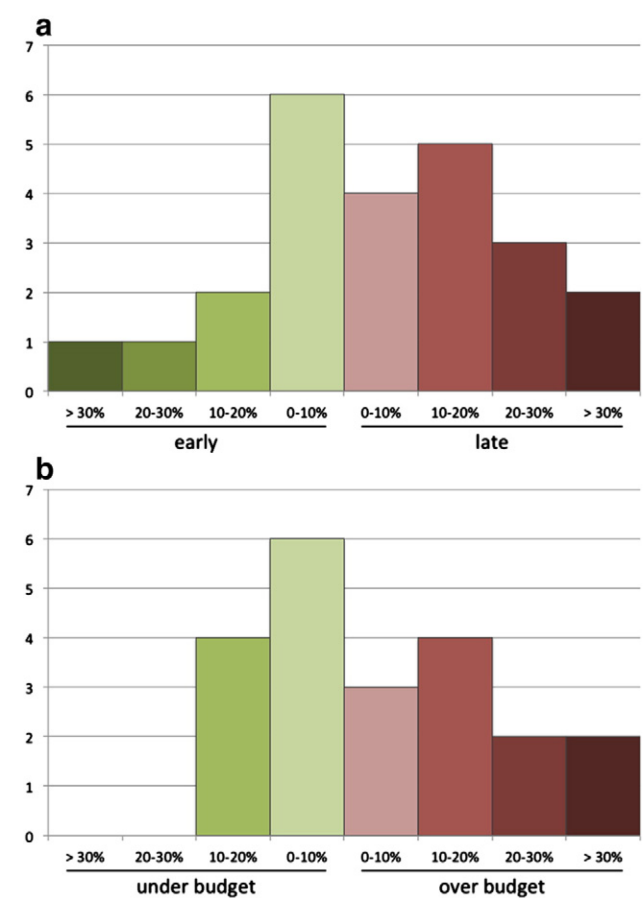

Fig. 5. Actual time and cost performance of the projects with real tracking data.

Appendix A. Example project card

OR-AS Operations Research Applications and Solutions	Case Name: **Pumping Station Jabbeke**	Sector	Construction (Civil)	
	OR-AS Operations Research - Applications and Solutions www.or-as.be info@or-as.be	Baseline Schedule	Schedule with resources	
			Schedule with costs	
		Risk Analysis	Random simulation	
			One of nine std. scenarios	
Submitted by	Niels Crommen		User defined distributions	
Date	December 20, 2012	Project Control	Automatic tracking	
File Name	C2012-13 Pumping Station Jabbeke.p2x		Tracking based on user input	

1. Project description

Project authenticity

The renovation of three pumps used to prevent creeks from overflowing by removing the water from the polder and pumping it into the nearby canal. All main activities are controlled by a general low voltage panel, which also has to be installed. The pumping station is situated in Jabbeke (Belgium).

The project consists of activity, resource and cost data that were obtained directly from the actual project owner.

2. Project properties

2.1. Baseline Schedule

General	
# Activities	74
Planned Duration (PD)	125 days*
Budget At Completion (BAC)	366,410 €
Renewable Resources	2
Consumable Resources	-

* standard eight-hour working days

Network topology	
Serial/Parallel (SP)	64%
Activity Distribution (AD)	59%
Length of Arcs (LA)	3%
Topological Float (TF)	27%

2.2. Risk Analysis

Random simulation by ProTrack was performed using the default symmetric triangular risk distribution profiles.

	Cost sensitivity		
	avg [%]	std dev [%]	skew [-]
CRI-r	10.2	10.6	2.8
CRI-rho	15.2	14.9	1.6
CRI-tau	32.4	24.8	1.3

	Resource sensitivity		
	avg [%]	std dev [%]	skew [-]
CRI-r	54.0	44.0	N/A
CRI-rho	55.0	43.0	N/A
CRI-tau	43.0	43.0	N/A

	Time sensitivity		
	avg [%]	std dev [%]	skew [-]
CI	58.5	48.6	-0.3
SI	64.1	42.6	-0.5
SSI	4.5	10.7	6.0
CRI-r	10.1	12.4	3.8
CRI-rho	14.1	16.1	2.2
CRI-tau	34.7	24.2	1.3

2.3. Project Control

2.3.1. Simulated forecasting accuracy

The accuracy of time and cost forecasting methods has been evaluated based on Monte Carlo simulation runs using the risk profiles described in section “2.2. Risk Analysis”. Based on these risk profiles, the Mean Absolute Percentage Error (MAPE) and Mean Percentage Error (MPE) has been calculated to evaluate the expected accuracy of the time and cost predictions, EAC(t) and EAC, respectively.

Simulated EAC(t) accuracy		
method - PF	MAPE [%]	MPE [%]
PV - 1	14.3	-14.3
PV - SPI	21.2	4.0
PV - SCI	21.2	4.3
ED - 1	21.2	-21.2
ED - SPI	21.2	4.0
ED - SCI	21.2	4.1
ES - 1	12.4	-11.4
ES - SPI(t)	18.6	15.4
ES - SCI(t)	18.7	15.5

Simulated EAC accuracy		
method (PF)	MAPE [%]	MPE [%]
1	0.3	-0.2
CPI	0.3	0.0
SPI	12.2	12.2
SPI(t)	14.6	14.5
SCI	12.3	12.2
SCI(t)	14.6	14.6
0.8 CPI + 0.2 SPI	4.0	3.9
0.8 CPI + 0.2 SPI(t)	5.5	5.5

According to the MAPE values[1] the best performance for time forecasting can be expected from the unweighted Earned Schedule method. For cost forecasting the unweighted and CPI-weighted methods should yield the best results.

2.3.2. Tracking description

Tracking authenticity

Manual tracking was performed over 28 tracking periods with a length of approximately one week. The Real Duration and Real Cost mentioned in section “2.3.3. Earned Value Management” are based on manual user input.

The tracking information obtained from the project owner and introduced in ProTrack includes actual activity start dates, durations and costs.

[1] The MAPE gives the best indication for the forecast accuracy (the lower the MAPE, the more accurate the method) since all deviations from the targeted real duration (real cost) are cumulated, whereas for the MPE underestimates can be compensated by overestimates and vice versa, possibly leading to an overly positive evaluation of a certain method. However, the MPE can provide useful information about the nature of the deviations, i.e. does the method rather underestimate or overestimate the real duration (real cost)?

2.3.3. Earned Value Management

2.3.3.1. Performance metrics

	CV [€]	SV [€]	SV(t) [d]	CPI [-]	SPI [-]	SPI(t) [-]	p-factor [-]
avg	-5,717	6,471	-6.01	0.98	1.00	0.97	0.98
std dev	5,381	17,755	8.89	0.02	0.10	0.12	0.04
final	-14,101	0	-15.00	0.96	1.00	0.89	1.00

2.3.3.2. Time forecasting

PD	125 days	Real Duration	140 days	Late	12.00%

EAC(t)			Real Accuracy	
method - PF	avg [d]	std dev [d]	MAPE [%]	MPE [%]
PV - 1	127.41	6.60	9.0	-9.0
PV - SPI	125.72	11.14	10.5	-10.2
PV - SCI	128.47	13.46	9.6	-8.2
ED - 1	128.50	7.48	8.5	-8.2
ED - SPI	126.83	12.03	10.0	-9.4
ED - SCI	127.46	12.59	9.7	-9.0
ES - 1	131.01	8.89	7.7	-6.4
ES - SPI(t)	130.16	15.09	10.2	-7.0
ES - SCI(t)	130.90	15.73	10.3	-6.5

2.3.3.3. Cost forecasting

BAC	336,410 €	Real Cost	350,511 €	Over Budget	4.19%

EAC			Real Accuracy	
method (PF)	avg [€]	std dev [€]	MAPE [%]	MPE [%]
1	342,127	5,381	2.4	-2.4
CPI	343,266	6,632	2.1	-2.1
SPI	337,585	18,034	4.1	-3.7
SPI(t)	338,319	21,039	4.7	-3.5
SCI	338,860	19,596	4.1	-3.3
SCI(t)	339,646	22,476	5.0	-3.1
0.8 CPI + 0.2 SPI	341,966	9,025	2.5	2.4
0.8 CPI + 0.2 SPI(t)	342,046	9,433	2.5	-2.4

References

Acebes, F., Pajares, J., Galán, J.M., López-Paredes, A., 2014. A new approach for project control under uncertainty. Going back to the basics. Int. J. Proj. Manag. 32, 423–434.

Agrawal, M., Elmaghraby, S., Herroelen, W., 1996. DAGEN: a generator of test sets for project activity nets. Eur. J. Oper. Res. 90, 376–382.

Anbari, F., 2003. Earned value project management method and extensions. Proj. Manag. J. 34, 12–23.

Boctor, F., 1993. Heuristics for scheduling projects with resource restrictions and several resource-duration modes. Int. J. Prod. Res. 31, 2547–2558.

Bright, H., Howard, T., 1981. Weapon System Cost Control: Forecasting Contract Completion Costs. Comptroller/Cost Analysis Division, US Army Missile Command, Redstone Arsenal, Alabama.

Caron, F., Ruggeri, F., Merli, A., 2013. A Bayesian approach to improve estimate at completion in earned value management. Proj. Manag. J. 44, 3–16.

Christensen, D., Antolini, R., McKinney, J., 1995. A review of estimate at completion (EAC) research. Journal of Cost Analysis and Management (Spring) 41–62.

Coelho, J., Vanhoucke, M., 2011. Multi-mode resource-constrained project scheduling using RCPSP and SAT solvers. Eur. J. Oper. Res. 213, 73–82.

Colin, J., Vanhoucke, M., 2014. Setting tolerance limits for statistical project control using earned value management. Omega-Int. J. Manag. Sci. 49, 107–122.

Covach, J., Haydon, J., Reither, R., 1981. A Study to Determine Indicators and Methods to Compute Estimate at Completion (EAC). ManTech International Corporation, Virginia.

Deblaere, F., Demeulemeester, E., Herroelen, W., 2011. Reactive scheduling in the multi-mode RCPSP. Comput. Oper. Res. 38, 63–74.

Demeulemeester, E., Vanhoucke, M., Herroelen, W., 2003. RanGen: a random network generator for activity-on-the-node networks. J. Sched. 6, 17–38.

Elloumi, S., Fortemps, P., 2010. A hybrid rank-based evolutionary algorithm applied to multi-mode resource-constrained project scheduling problem. Eur. J. Oper. Res. 205, 31–41.

Elmaghraby, S., 2000. On criticality and sensitivity in activity networks. Eur. J. Oper. Res. 127, 220–238.

Elshaer, R., 2013. Impact of sensitivity information on the prediction of project's duration using earned schedule method. Int. J. Proj. Manag. 31, 579–588.

Fleming, Q., Koppelman, J., 2010. Earned value project management, Fourth ed. Project Management Institute, Newtown Square, Pennsylvania.

Hartmann, S., Briskorn, D., 2010. A survey of variants and extensions of the resource-constrained project scheduling problem. Eur. J. Oper. Res. 207, 1–15.

Hartono, B., Sulistyo, S.R., Praftiwi, P.P., Hasmoro, D., 2014. Project risk: theoretical concepts and stakeholders' perspectives. Int. J. Proj. Manag. 32, 400–411.

Hecht, L., 2007. Case study of earned schedule to do predictions. The Measurable News (Winter) 16–18.

Henderson, K., 2003. Earned schedule: A breakthrough extension to earned value theory? A retrospective analysis of real project data. The Measurable News (Summer) 13–17, 21–23.

Henderson, K., 2004. Further developments in earned schedule. The Measurable News (Spring) 15–17, 20–22.

Henderson, K., 2005. Earned schedule in action. The Measurable News (Spring) 23–28, 30.

Hulett, D., 1996. Schedule risk analysis simplified. Proj. Manag. Netw. 10, 23–30.

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

Kim, B., Reinschmidt, K., 2009. Probabilistic forecasting of project duration using Bayesian inference and the beta distribution. J. Constr. Eng. Manag. 135, 178–186.

Kim, B., Reinschmidt, K., 2010. Probabilistic forecasting of project duration using Kalman filter and the earned value method. J. Constr. Eng. Manag. 136, 834–843.

Kim, B., Reinschmidt, K., 2011. Combination of project cost forecasts in earned value management. J. Constr. Eng. Manag. 137, 958–966.

Kolisch, R., Sprecher, A., 1996. PSPLIB — A project scheduling problem library. Eur. J. Oper. Res. 96, 205–216.

Kolisch, R., Sprecher, A., Drexl, A., 1995. Characterization and generation of a general class of resource-constrained project scheduling problems. Manag. Sci. 41, 1693–1703.

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

Lipke, W., 2009. Project duration forecasting… a comparison of earned value management methods to earned schedule. The Measurable News (2), 24–31.

Lipke, W., 2013. Earned schedule — Ten years after. The Measurable News (3), 15–21.

Lipke, W., Zwikael, O., Henderson, K., Anbari, F., 2009. Prediction of project outcome: the application of statistical methods to earned value management and earned schedule performance indexes. Int. J. Proj. Manag. 27, 400–407.

Naeni, L., Shadrokh, S., Salehipour, A., 2011. A fuzzy approach to earned value management. Int. J. Proj. Manag. 29, 764–772.

Narbaev, T., De Marco, A., 2014. An earned schedule-based regression model to improve cost estimate at completion. Int. J. Proj. Manag. 32, 1007–1018.

Nasirzadeh, F., Khanzadi, M., Rezaie, M., 2014. Dynamic modeling of the quantitative risk allocation in construction projects. Int. J. Proj. Manag. 32, 442–451.

OR-AS, 2014a. Online consultation of the real-life project database and the corresponding project cards (continuously updated). URL: http://www.or-as.be/research/database.

OR-AS, 2014b. Project management software tool ProTrack. URL: http://www.protrack.be/.

Pajares, J., López-Paredes, A., 2011. An extension of the EVM analysis for project monitoring: the cost control index and the schedule control index. Int. J. Proj. Manag. 29, 615–621.

PMI, 2004. A Guide to the Project Management Body of Knowledge (PMBOK Guide), Third ed. Project Management Institute, Newtown Square, Pennsylvania.

PMI, 2005. Practice Standard for Earned Value Management. Project Management Institute, Newtown Square, Pennsylvania.

Riedel, M., Chance, J., 1989. Estimates at Completion (EAC): A Guide to their Calculation and Application for Aircraft, Avionics, and Engine Programs. Aeronautical Systems Division, Wright-Patterson AFB, Ohio.

Rujirayanyong, T., 2009. A comparison of three completion date predicting methods for construction projects. J. Res. Eng. Technol. 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.

Taroun, A., 2014. Towards a better modelling and assessment of construction risk: insights from a literature review. Int. J. Proj. Manag. 32, 101–115.

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. Eur. J. Oper. Res. 119, 510–537.

Trietsch, D., Mazmanyan, L., Govergyan, L., Baker, K.R., 2012. Modeling activity times by the Parkinson distribution with a lognormal core: theory and validation. Eur. J. Oper. Res. 386–396.

Tzaveas, T., Katsavounis, S., Kalfakakou, G., 2010. Analysis of project performance of a real case study and assessment of earned value and earned schedule techniques for the prediction of project completion date. Proceedings of IPMA Conference, May 2010 — Crete, Greece, pp. 752–759.

Uyttewaal, E., 2005. Dynamic Scheduling With Microsoft Office Project 2003: The Book by and for Professionals. Co-published with International Institute for Learning, Inc.

Van Peteghem, V., Vanhoucke, M., 2010. A genetic algorithm for the preemptive and non-preemptive multi-mode resource-constrained project scheduling problem. Eur. J. Oper. Res. 201, 409–418.

Van Peteghem, V., Vanhoucke, M., 2014. An experimental investigation of metaheuristics for the multi-mode resource-constrained project scheduling problem on new dataset instances. Eur. J. Oper. Res. 235, 62–72.

Vandevoorde, S., Vanhoucke, M., 2006. A comparison of different project duration forecasting methods using earned value metrics. Int. J. Proj. Manag. 24, 289–302.

Vanhoucke, M., 2010a. Using activity sensitivity and network topology information to monitor project time performance. Omega-Int. J. Manag. Sci. 38, 359–370.

Vanhoucke, M., 2010b. Measuring time — Improving project performance using earned value management. International Series in Operations Research and Management Science vol. 136. Springer.

Vanhoucke, M., 2011. On the dynamic use of project performance and schedule risk information during project tracking. Omega-Int. J. Manag. Sci. 39, 416–426.

Vanhoucke, M., 2012a. Measuring the efficiency of project control using fictitious and empirical project data. Int. J. Proj. Manag. 30, 252–263.

Vanhoucke, M., 2012b. Project management with dynamic scheduling: baseline scheduling. Risk Analysis and Project Control vol. XVIII. Springer.

Vanhoucke, M., 2014. Integrated project management and control: first comes the theory, then the practice. Management for Professionals. Springer.

Vanhoucke, M., Vandevoorde, S., 2007. A simulation and evaluation of earned value metrics to forecast the project duration. J. Oper. Res. Soc. 58, 1361–1374.

Vanhoucke, M., Coelho, J., Debels, D., Maenhout, B., Tavares, L., 2008. An evaluation of the adequacy of project network generators with systematically sampled networks. Eur. J. Oper. Res. 187, 511–524.

Wang, L., Fang, C., 2011. An effective shuffled frog-leaping algorithm for multi-mode resource-constrained project scheduling problem. Inf. Sci. 181, 4804–4822.

Wang, L., Fang, C., 2012. An effective estimation of distribution algorithm for the multi-mode resource-constrained project scheduling problem. Comput. Oper. Res. 39, 449–460.

Williams, T., 1992. Criticality in stochastic networks. J. Oper. Res. Soc. 43, 353–357.

Zhang, Y., Fan, Z.P., 2014. An optimization method for selecting project risk response strategies. Int. J. Proj. Manag. 32, 412–422.

Zwikael, O., Globerson, S., Raz, T., 2000. Evaluation of models for forecasting the final cost of a project. Proj. Manag. J. 31, 53–57.

Zwikael, O., Pathak, R.D., Singh, G., Ahmed, S., 2014. The moderating effect of risk on the relationship between planning and success. Int. J. Proj. Manag. 32, 435–441.