Automation in Construction 89 (2018) 71–85

Contents lists available at ScienceDirect

Automation in Construction

journal homepage: www.elsevier.com/locate/autcon

An empirical validation of the performance of project control tolerance limits

Annelies Martens[a], Mario Vanhoucke[a,b,c,*]

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

ARTICLE INFO

Keywords:
Project management
Schedule control
Earned Value Management
Empirical data

ABSTRACT

The goal of project control is monitoring the project progress during project execution to detect potential problems and taking corrective actions when necessary. Tolerance limits are a tool to assess whether the project progress is acceptable or not, and generate warnings signals that act as triggers for corrective action to the project manager. In this paper, three distinct types of tolerance limits that have been proposed in literature are validated on a large and diverse set of real-life projects mainly situated in the construction sector. Moreover, a novel approach to construct tolerance limits that integrate the project risk information into the monitoring process is introduced. The results of the empirical experiment have shown that integrating project-specific information into the construction of the tolerance limits results in a higher efficiency of the monitoring process. More specifically, while including cost information increases the efficiency only marginally, incorporating the available resource information substantially improves the efficiency of the monitoring process. Furthermore, when projects are not restricted by scarce resources, the efficiency can be enhanced by integrating the available project risk information.

1. Introduction

An important factor of project success is the timely completion of projects. In this paper, three control methodologies proposed in recent literature to control the schedule progress of projects are empirically compared and validated on the large and diverse real-life project database of Batselier and Vanhoucke [1]. From this database, 93 projects have been selected, of which 71 are situated in the broad construction sector. More specifically, the tolerance limits have been evaluated for commercial, residential and institutional building projects and for civil and industrial construction projects. Moreover, a novel control methodology that integrates the activity risk information into the project control phase is introduced.

Project control is, together with baseline scheduling and risk analysis, one of the three major components of Integrated Project Management and Control [2]. While the scheduling and risk analysis phases are performed before the project execution is started, the project control phase is conducted during project execution. The goal of this phase is to identify potential problems or opportunities during project execution, and to take corrective actions to get the project back on track if necessary. During project execution, the actual project progress is monitored and evaluated by comparing it to the baseline schedule. A well-known methodology to monitor the project progress is Earned Value Management (EVM), which originated in the 1960s at the US Department of Defense [3]. While EVM provides simple metrics to measure the current performance of a project, they should be used in conjunction with tolerance limits to assess this performance. These tolerance limits for project control have been established as a tool to support the project manager in deciding whether corrective actions should be taken to get the project back on track. Hence, the goal of these tolerance limits is generating warning signals when the monitored project progress is below a certain threshold, indicating that it is likely that the project will exceed its deadline. These warning signals thus act as a trigger for corrective action for the project manager. The control methodologies validated in this paper are analytical tolerance limits for schedule control using EVM metrics. This type of tolerance limits sets threshold values for the schedule progress at each project phase based on project-specific characteristics. Moreover, each of the tolerance limits evaluated in this paper are constructed for projects with a project buffer. Hence, these tolerance limits generate warning signals when it is expected that the project buffer will be consumed entirely before the project is completed, resulting in a project exceeding its deadline.

---

[*] Corresponding author at: Faculty of Economics and Business Administration, Ghent University, Tweekerkenstraat 2, Gent 9000, Belgium.
E-mail addresses: annelies.martens@ugent.be (A. Martens), mario.vanhoucke@ugent.be (M. Vanhoucke).

https://doi.org/10.1016/j.autcon.2018.01.002
Received 2 August 2017; Received in revised form 14 December 2017; Accepted 9 January 2018

The contribution of this paper is twofold. First, the analytical tolerance limits proposed in [4,5] and [6] are empirically validated using the real-life project database of Batselier and Vanhoucke [1]. A distinctive characteristic of these tolerance limits is their ease of implementation, since they are constructed using only project-specific information and do not require any historical or simulated data. However, since the performance of these limits has been validated using large simulation studies only, the ease of implementation in a real-life context has not been verified yet. Therefore, in this paper, both the ease of implementation of the tolerance limits for real-life data and their performance will be reviewed, using empirical data rather than artificial simulations. These empirical data comprise of 71 real-life construction projects and 22 real-life projects in the education, event management, engineering and IT management sector. Second, a new approach is proposed, which integrates the available activity risk information into the construction of the tolerance limits. Based on the results of the empirical study, guidelines are proposed for when to use which type of tolerance limit.

The outline of this paper is as follows. In Section 2, the literature on Integrated Project Management and Control and on project monitoring is briefly introduced. Subsequently, the empirical experiment is described in Section 3. Further, the results of this experiment are reviewed in Section 4. Finally, Section 5 discusses the conclusion of this paper.

2. Literature review

Since project control is one of the three major components of Integrated Project Management and Control, these components are briefly discussed in Section 2.1. Subsequently, in Sections 2.2 and 2.3, we elaborate on the project monitoring process and the construction of tolerance limits for project control.

2.1. Integrated Project Management and Control

In this section, the three components of Integrated Project Management and Control, namely baseline scheduling, risk analysis and project control, are introduced. Further, research efforts in recent literature are highlighted for each component.

Baseline scheduling

During the baseline scheduling phase, a feasible baseline schedule is constructed that acts as a point of reference during the risk analysis and project control phase. While the project scheduling problem has initially been addressed in absence of resource restrictions by the critical path method (CPM, [7]) and the program evaluation and research technique (PERT, [8]), the resource-constrained project scheduling problem (RCPSP) has been defined to explicitly incorporate resource restrictions. The aim of the RCPSP is minimising the project makespan when limited resources are available. For an overview of the variants and extensions of this problem that are explored in literature, the reader is referred to [9–11]. Moreover, extensions and novel approaches to solve this problem are still being developed, e.g. by [12–14].

Schedule risk analysis

In this phase, a Schedule Risk Analysis (SRA) is performed to connect the activity risk information to the baseline schedule such that the sensitivity and cruciality of the project activities can be measured. The resulting SRA measures indicate the impact of the activities on the final project duration and can be used by the project manager to decide which activities require managerial effort during project execution [15,16]. Further, the merits and pitfalls of these SRA measures are reviewed by Elmaghraby [17].

Project control

During the project control phase, deviations from the baseline schedule are measured during project execution such that corrective actions can be taken by the project manager when problems are detected. A wide variety of project control problems has been studied in literature, e.g. determining the timing of control points [18–20], forecasting the final project duration [21,22], optimal buffer sizing [23–25] and corrective action taking [16,26]. In this paper, we focus on constructing tolerance limits for the schedule progress of projects. The aim of these tolerance limits for project control is to support the project manager in deciding whether corrective actions should be taken by comparing the monitored progress to a certain threshold. Therefore, in Section 2.2, the monitoring process is briefly discussed. Subsequently, the literature on tolerance limits for project control is reviewed in Section 2.3.

2.2. Progress monitoring: Earned Value Management

A well-known technique to monitor project progress is Earned Value Management (EVM, [3]). This methodology measures the project schedule and cost progress in terms of Earned Value (EV) and Actual Cost (AC). The cost progress is evaluated by comparing the AC with the EV using two EVM cost performance metrics, the Cost Performance Index $\left(\text{CPI} = \frac{\text{EV}}{\text{AC}}\right)$ and the Cost Variance (CV = EV − AC). Further, the schedule progress is examined by comparing the EV to the Planned Value (PV), which is the value that is planned to be earned according to the baseline schedule. Based on these key metrics, two EVM schedule performance metrics are constructed, namely the Schedule Performance Index $\left(\text{SPI} = \frac{\text{EV}}{\text{PV}}\right)$ and the Schedule Variance (SV = EV − PV). However, it is known that, since the PV and EV are both cost-based rather than time-based metrics, the SPI behaves unreliably towards the end of the project. In order to overcome this drawback of EVM, the Earned Schedule (ES) concept has been introduced by Lipke [27] as an extension of EVM. The ES is a time-based metric for the schedule progress, and is determined as follows:

$$\text{ES} = t + \frac{\text{EV} - \text{PV}_t}{\text{PV}_{t+1} - \text{PV}_t} \tag{1}$$

with t such that $\text{EV} \geq \text{PV}_t$ and $\text{EV} < \text{PV}_{t+1}$. Consequently, the schedule performance of projects can be measured using the SPI(t) $\left(=\frac{\text{ES}}{\text{AT}}\right)$, which is a more reliable metric than the SPI. Therefore, the SPI(t) will be used in this experiment to measure the schedule progress of the projects. In the remainder of this section, the integrated earned value/earned schedule method will be referred to as EVM/ES. For an extensive introduction to EVM/ES, the reader is referred to [28] and [27]. Further, a recent comprehensive overview of the applications and extensions of EVM/ES is given in [29].

2.3. Tolerance limits for project control

In order to evaluate the monitored project progress, tolerance limits for project control are constructed. For each phase of the project, a threshold value for the progress is determined. When the actual progress is below this threshold at a certain time during execution, a warning signal is generated. This signal indicates that, given the current progress, the project is likely to exceed its deadline and corrective actions should be taken in order to get the project back on track.

In recent literature, tolerance limits have been proposed that can be classified in three groups; static tolerance limits, statistical tolerance limits and analytical tolerance limits. *Static tolerance limits* are constant throughout the entire project life cycle and are determined using rules of thumb. Since these limits do not consider any project-specific information or information from historical data, they are not always very accurate or fail to dynamically take project progress features into account. This type of tolerance limits has been introduced by Goldratt [30] and Leach [31]. Further, *statistical tolerance limits* have been proposed. This type of limits applies concepts of Statistical Process Control (SPC, [32]) and requires historical data or Monte Carlo simulations to determine the desired state of the project's progress at each time during the execution. In literature, statistical tolerance limits have been proposed and validated using empirical data [33–37] or simulation studies [4,38,39]. Finally, *analytical tolerance limits* use project-specific information to determine the progress threshold at each project phase. Since this type of limits does not require historical data or monte carlo simulations, but only uses information which is readily available during the scheduling phase, this type of limits is easier to implement than statistical tolerance limits. Moreover, by including project-specific information, they are often more accurate than static tolerance limits. Analytical tolerance limits for projects with a project buffer have been introduced by [4-6,40]. Since the limits used in [40] employ the same principle as those in [4], but do not translate their threshold values into EVM performance metrics, they are not further considered in this research. The analytical tolerance limits introduced in [4,5] and [6] are discussed in more detail in the remainder of this section and empirically validated in Section 3.

More specifically, both the analytical tolerance limits for projects with a project buffer introduced in [5] and [6] as well as a novel approach to construct analytical tolerance limits will be compared to the basic analytical tolerance limits proposed in [4]. Each of these analytical tolerance limits focuses on a different perspective to determine the threshold values for the project's progress. First, the limits introduced in [4] consider the time perspective. In [5], these limits are referred to as *linear limits*, since they are established by imposing that the project buffer is allowed to be consumed linearly with the time during project execution. As a consequence, these limits are equally strict for each project phase. Since these linear limits are the most straightforward limits that can be set, they will be used as a benchmark to assess the improvements of the other more advanced analytical tolerance limits. The first more advanced type of analytical tolerance limits has been proposed in [5]. More specifically, since the planned monetary value of the project generally does not accrue linearly, but often follows an S-curve, Martens and Vanhoucke [5] argue that the allowed buffer consumption of each project phase should be determined based on the planned value of that phase rather than just following a linear shape. By employing a so-called cost perspective, the *cost limits* that are constructed are more severe for delays of project phases that deliver limited (monetary) value to the project. Further, rather than following a cost approach to construct the tolerance limits, Martens and Vanhoucke[6] incorporated resource information into the construction of the limits. In particular, since the limited availability of renewable resources is a frequent cause of project delays, a resource perspective is considered in [6]. The proposed *resource limits* are based on the resource availability and resource requirements of the project activities. These resource limits are constructed such that they are more strict for project phases in which resource conflicts leading to project delays are more likely. Finally, the novel approach that has not been used earlier in literature focuses on the risk perspective, by considering the risk level of all project activities. In this approach, the most risky project phases in terms of activity duration variation are allowed to consume a larger portion of the project buffer than the less risky phases. Consequently, the *risk limits* are less strict for the most risky project phases. The procedure to construct these analytical tolerance limits is discussed in more detail in Section 3.2.

3. Empirical experiment

In this section, the empirical experiment is discussed. First, the data selection procedure is described in Section 3.1. Subsequently, the procedure to construct the tolerance limits is defined in Section 3.2. Both the construction of the limits and the implications of using real-life data are discussed in this section. Further, the design of the experiment is discussed in Section 3.3. Finally, the performance measurement process is specified in Section 3.4.

3.1. Data selection

The real-life projects used in this paper are part of the large and diverse real-life project database of Batselier and Vanhoucke [1], which currently consists of 125 projects situated within 5 distinct sectors and which is publicly available at [41]. This project database has been used in several studies to evaluate project control techniques empirically, e.g. in [42,43] and [44]. Each project in this database has a project card, in which the project data are summarised and the authenticity and completeness of the data are indicated. The authenticity of the project data is reviewed using two concepts, namely the *project authenticity* and the *tracking authenticity*. Full project authenticity implies that all activity, resource and baseline cost data were collected from the actual project owner, and no assumptions were made by the data collector. Similarly, full tracking authenticity means that the actual activity start dates, durations and costs were all obtained from the actual project owner. As a result, each project with full tracking authenticity can be monitored using the EVM/ES methodology. Further, the project cards represent the completeness of the project data for the three dimensions of Integrated Project Management and Control (baseline schedule, risk analysis and project control).

Using the information provided by the project cards, we have selected the projects suitable for this research as follows. First of all, we performed an *authenticity selection* in order to only retain those projects with a full project and tracking authenticity. Out of the 125 projects in the database, 111 projects have a full project authenticity. 93 of these projects have a full tracking authenticity as well. Consequently, these 93 projects were retained and subjected to a *completeness selection*, in which we verified whether the required baseline, risk and project control information is included. Since each type of analytical tolerance limit which is empirically validated in this paper focuses on a different project perspective, as depicted in Table 1, the required project information differs for each type. First, the *linear limits* do not consider any project-specific information. Therefore, only the planned duration of the project and the time of the tracking periods are required. For each of the 93 fully authentic projects, these limits can thus be constructed. Further, the *cost limits*, which deploy a cost perspective, require that the planned activity start times, duration and cost are known. Since these requirements have to be met in order to be fully authentic, the *cost limits* can be constructed for all 93 projects as well. The resource limits proposed by Martens and Vanhoucke [6], however, require additional information on the resource availability and resource requirements of the activities. This information is only provided for 21 of the 93 projects. Finally, we observed that, from the 93 projects, 32 provide additional information on the activity risk profiles. More specifically, the three point estimates and standard deviation of the activity durations are specified. Therefore, we propose an approach to include this information into the construction of tolerance limits in order to improve their performance. These performance limits will be referred to as *risk limits* and are introduced in Section 3.2.2. An overview of the number of selected projects for each type of limits is given in Fig. 1. Further, the project codes of the selected projects for each limit are listed in Appendix A. Using these codes, the project cards with detailed information on the project characteristics can be found at [41].

Table 1
Overview of limits.

	Perspective	Reference	Number of projects
Linear limits	Time	[4]	93
Cost limits	Cost	[5]	93
Resource limits	Resource	[6]	21
Risk limits	Risk	–	32

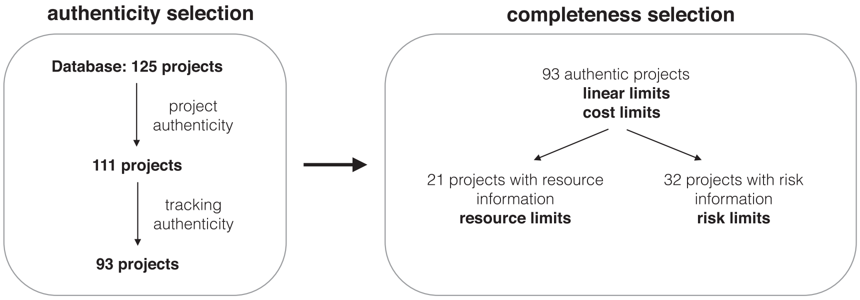

Fig. 1. Dataset.

3.2. Construction of limits

In this section, the construction of the analytical tolerance limits for projects with a project buffer will be discussed. First, the general procedure to construct these limits is briefly introduced in Section 3.2.1. Subsequently, the aspects which are specific to each type of limit are discussed in Section 3.2.2. Finally, the specific procedures are illustrated using a toy example in Section 3.2.3.

3.2.1. General procedure

The procedure proposed in this paper constructs EVM tolerance limits for projects with a project buffer, by determining the maximal amount of the buffer that is allowed to be consumed at each project phase. In the empirical experiment, we determined this allowable buffer consumption for each time unit from the start until the planned duration of the projects. The three steps required to construct these limits are depicted in Fig. 2. In this figure, the abbreviations PB, PD, DL and BAC are used to denote the project buffer, planned duration, deadline and budget at completion respectively.

Step 1 *Buffer assignment.* In the first step, the allowable buffer consumption is determined for each project phase. The allowable buffer consumption reflects the portion of the project buffer that is allowed to be consumed by a project phase without endangering the project deadline. The left pane of Fig. 2 illustrates that different project phases are not necessarily assigned an equal portion of the project buffer. In particular, this portion is determined for each project phase, based on project specific characteristics. Each limit discussed in this paper focuses on a different characteristic, namely time, cost, resources and risk, resulting in a different accrue of allowable buffer consumption.

Step 2 *Buffered planned progress.* In this step, the buffered planned progress curve (BPP-curve) is constructed. This curve reflects the project progress in case each project phase consumes the entire allowable portion of the project buffer. As depicted in the middle pane of Fig. 2, the BPP curve can be determined by adding the allowable buffer consumption of each project phase to the planned progress of that phase. Accordingly, the BPP-curve reflects the minimal progress of the project in order to meet its deadline.

Step 3 *Threshold values.* In the final step, the tolerance limits are expressed in terms of SPI(t) based on the minimally required progress (reflected by the BPP-curve) and the planned progress. At each time, the BPP-curve is compared to the planned progress in order to determine at which time the buffered planned progress was scheduled to be completed. This time is referred to as the minimally required earned schedule at time *t* ($ES_{R,t}$). In the right pane of Fig. 2, this step is illustrated for time t_3. Finally, by dividing the $ES_{R,t}$ by time *t*, the SPI(t) threshold value for time *t* is found. This SPI(t) threshold value represents the minimum schedule performance that guarantees a timely project completion. Hence, when the performance drops below this threshold, corrective actions are required.

For a more detailed discussion on this general procedure, the reader is referred to [5]. Since each type of tolerance limit used in this research focuses on a different project-specific characteristic in order to determine the allowable buffer consumption during project execution, the first step of this procedure is specific to each type of tolerance limit. The second and third steps, however, are identical for all types.

3.2.2. Specific procedure: buffer assignment

In this section, the determination of the allowable buffer consumption for each project phase is described for each type of analytical tolerance limit listed in Table 1. Further, the required data and possible adjustments to the construction of the limits are discussed for each of these limits.

Linear limits. For the linear limits, introduced by Colin and Vanhoucke [4], the allowable buffer consumption is assumed to increase linearly with the time. Specifically, at *x*% of the project makespan, *x*% of the project buffer is allowed to be consumed. In other words, the BPP-curve is established by stretching the PV-curve

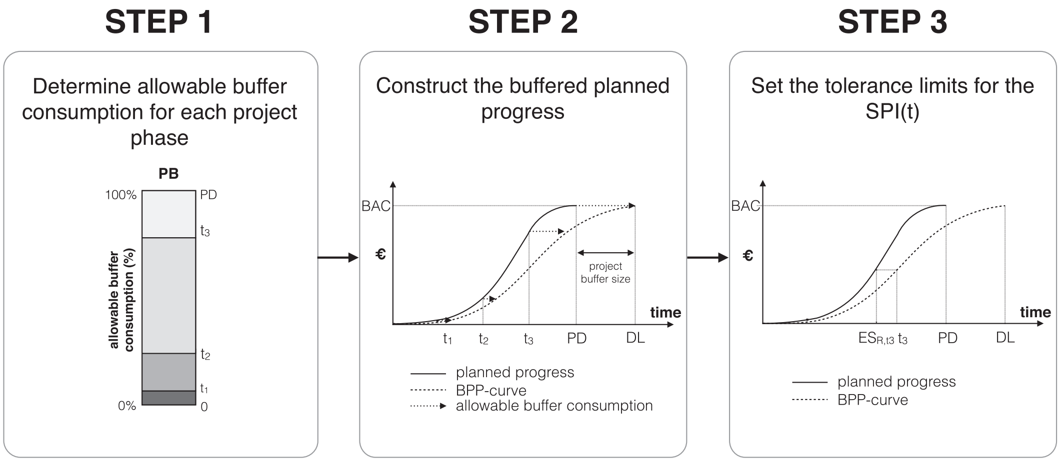

Fig. 2. General procedure.

horizontally with a factor $\frac{DL}{PD}$. Consequently, the only information that is required is the time of each tracking period for which tolerance limits will be set, and the total project duration. Since only a limited amount of data is required, the linear limits can be constructed for each project in our real-life project data set. Moreover, no adaptations to the data are required.

Cost limits. In [5], a cost perspective is deployed to determine the allowable buffer consumption throughout the project execution. More precisely, the cost information of the activities is used to determine the aggregated PV accrue of the baseline schedule. Accordingly, the allowable buffer consumption of a project phase is set proportionally with the PV accrue. Thus, at each project phase, the allowable buffer consumption is equal to the percentage completed (PC) of that phase. Specifically, when x% of the project is completed, x% of the project buffer is allowed to be consumed. Therefore, the PC of each project phase should be known in order to construct the cost limits. Since the projects in our dataset are selected such that they all have full project authenticity, this information is available for each project in the dataset. Hence, the cost limits can be constructed for all 93 projects.

Resource limits The limits proposed in [6] consider a resource perspective. This perspective is deployed at two levels, the activity level and the project level. At the *activity level*, the project progress is determined using the work content (WC) of the project activities rather than the activity cost. The WC of each activity is defined as the planned duration of the activity multiplied by its resource requirement per time unit, and is used to establish the aggregated WC accrue. At the *project level*, the so-called shiftability of the schedule, which reflects the ability to shift activities to the right without endangering the PD of the project, is considered. The higher this shiftability, the less likely the probability that resource conflicts will cause further delays that endanger the PD of the project.

Hence, in order to determine the allowable buffer consumption using the resource perspective, the following procedure should be followed. First, at the activity level, a planned WC-curve should be constructed which reflects the amount of work content that should be performed according to the baseline schedule. Second, at the project level, the cumulative shiftability should be determined by comparing the planned WC-curve of the baseline schedule to the planned WC-curve of the resource-feasible latest start schedule (LSS). This schedule is determined by scheduling the project activities as late as possible without violating the precedence relations and resource constraints using the Latest Finish Time (LFT) priority rule [45]. Hence, the shiftability of each time t until the PD is defined as the difference between the planned WC of the baseline schedule and the planned WC of the LSS at time t. Further, the cumulative shiftability at time t is determined by summing the shiftability of each time from 0 until t. Finally, since both the aggregated WC accrue and the cumulative shiftability are considered to determine the allowable buffer consumption, the project buffer is divided in two parts that are dedicated to these two aspects. The larger the shiftability of the project, the higher the focus on this aspect should be. Similarly, the lower the project shiftability, the higher the focus on the WC accrue. To determine the size of the project buffer that should be dedicated to the shiftability (implying that the remaining part of the buffer is dedicated to the WC accrue), the *degree of shiftability* has been introduced. This degree of shiftability is defined by two components, namely the absolute shiftability of the project and the maximal shiftability of the project. The absolute shiftability represents the total cumulative shiftability at the PD of the project. Further, the maximal shiftability is determined by comparing the baseline schedule to the maximal shiftability schedule, which is defined as the latest start schedule when the precedence relations of the project are not considered and a fixed work content is assumed for the activities. Accordingly, the degree of shiftability is calculated by dividing the absolute shiftability by the maximal shiftability. For a more elaborate description of this procedure, the reader is referred to [6]. Moreover, this procedure is illustrated on a toy example in Section 3.2.3.

Contrarily to the linear and cost limits, some adjustments to the data and the theoretical procedure are required to construct the resource limits. At the activity level, since the planned and actual progress of the projects in the database is provided in monetary units rather than in terms of work content, the planned and earned WC-curves have to be constructed using the provided project data. While in [6] only one type of renewable resources is assumed to be used, most real-life projects deploy several distinct renewable resources. Therefore, the availability and requirements of all required resources should be considered. In order to avoid biases due to different measure units used for the different types of resources, we normalised the resource availability of each resource to 1. Hence, the resource requirements of the project activities are defined as a portion of the total resource availability. Accordingly, the work content of each activity is defined as its planned duration multiplied with the sum of all proportional resource requirements per time unit. Since complete information on the planned and actual activity start times and resource requirements is available for the 21 selected projects, constructing these curves did not pose any problems. Moreover, this additional step was required in this experiment due to the fact that the presented projects all use the EVM/ES methodology to monitor their progress. When the resource limits would be used for a new real-life project, the progress can readily be expressed and measured in terms of work content without these additional calculations. At the project level, additional calculations are required as well. First, to determine the shiftability of the baseline schedule, the LSS of the project should be established. While this LSS is not included in the project database, it can be determined using the software tool ProTrack [46]. Using this information, all calculations to determine the resource limits can be made. Consequently, while the construction of the resource limits requires more calculations than the linear or cost limits, these calculations can be easily done with the information available at the scheduling phase.

Risk limits. While the resource limits have been proposed in order to improve the performance of analytical tolerance limits by including additional project information in the construction of the limits, most projects in our dataset (93 − 21 = 72) do not provide the required resource information. However, 32 of these projects provide additional project information in terms of risk. More specifically, for these projects, the three point estimates and standard deviation of the activity durations are listed. Therefore, we propose an approach to include this information in the determination of the allowable buffer consumption of each project phase.

The allowable buffer consumption according to the risk limits should be proportional with the risk accrue of the project. Two steps are required to determine the allowable buffer consumption. First, the individual risk value of each activity is determined. Second, based on the risk value of all activities, the aggregated risk curve can be constructed. The allowable buffer consumption of each phase is then set proportional to the aggregated risk at that phase.

Similar to the existing analytical tolerance limits, the individual activity risk value is determined using information that is readily available during the scheduling phase, without requiring historical data or Monte Carlo simulations. In order to determine the individual activity risk value, we have considered two aspects that affect the activity risk and that are known before project execution, namely the activity variability and the position of the activity in the project schedule. The variability of the activity duration is represented by its standard deviation, which is an assessment made by the project manager of the uncertainty related to the estimated activity duration. Further, we identified three situations in which early activities have a larger potential impact on the project outcome. First, activities that are scheduled early in the project can have more total successors that can suffer from a delayed start time than activities at the end of the project. Second, when limited resources are available, activities early in the project can cause more resource conflicts than activities at the end of the project. Finally, activity duration estimates are known to be stochastically dependent [47]. For early activities, there can be more remaining activities that suffer from the same estimation bias. Hence, rather than analysing for each activity the number of activities that could be affected due to the above reasons, which would require additional information on the dependency between activities, we assume that all activities with a start time later than the planned finish of an activity can be affected by delays of that activity. Therefore, we defined the individual activity risk value as follows:

$$risk_i = \sigma_i \times \#act_{i,f} \tag{2}$$

with $risk_i$ and σ_i the risk value and standard deviation of activity i and $\#act_{i,f}$ the number of activities following activity i, including the dummy end activity. Similarly to the PV-curve, the aggregated risk accrue curve can be constructed based on the activity start time, duration and risk value, and thus reflects the cumulative increase in risk value of the schedule. Finally, the allowable buffer consumption according to the risk limits is proportional with the aggregated risk value of each project phase.

3.2.3. Toy example

In this section, a toy project, for which the project network and baseline schedule are depicted in Fig. 3, is used to illustrate how each of the limits presented in Section 3.2.2 are constructed. First, the allowable buffer consumption is determined for each time period of the project. Each tolerance limit uses a different perspective to determine the allowable buffer consumption. More specifically, the allowable buffer consumption is set proportionally with the planned progress of the project, which is expressed using a time, cost, resource and risk perspective. Subsequently, the threshold values for the SPI(t) of each phase are set based on the allowable buffer consumption of that phase.

In Table 2, the planned progress according to the time, cost, resource and risk perspectives are listed for each time period of the project. For the linear limits, the planned progress per time period is equal to the time of that period. For the cost limits, the planned progress is defined by the PV-curve. At time t_3 $(t = 3)$, for instance, the progress is equal to 76.67 $\left(=cost_{act1} + cost_{act2} + \frac{cost_{act3}}{3} \times 2 = 20 + 30 + \frac{40}{3} \times 2\right)$. Further, the allowable buffer consumption of the resource limits depends on two aspects, namely the progress in terms of WC and the progress in terms of shiftability. At time t_3, the progress in terms of WC is equal to 9 $(=3 + 4 + \frac{3}{3} \times 2)$. In terms of shiftability, the progress is equal to 6 $(= shift_{t1} + shift_{t2} + shift_{t3} = 0 + (WC_{BS,t_2} - WC_{LS,t_2}) + (WC_{BS,t_3} - WC_{LS,t_3}) = 0 + (6 - 4) + (9 - 5))$. Finally, the progress in terms of risk depends on the standard deviation and the number of ensuing activities (including the dummy end activity) of each activity. For real-life projects, the standard deviation of the project activities is assessed by the project manager. In this example, we assume that activity 1 is the least risky and activity 3 the most risky activity of the project. This is reflected by their standard deviation, which is 0.01, 0.1, 0.3 and 0.1 for activities 1, 2, 3 and 4 respectively. Moreover, the activities have respectively 4, 2, 2 and 1 ensuing activities. The individual risk levels of the activities, which will be used to determine the project risk progress, are thus equal to 0.04, 0.2, 0.6 and 0.1. Consequently, at time t_3, the risk progress is equal to 0.64 $(= 0.04 + 0.2 + \frac{0.6}{3} \times 2)$.

Table 2
Planned progress for four perspectives.

	t_0	t_1	t_2	t_3	t_4	t_5	t_6
Linear limits	0	1	2	3	4	5	6
Cost limits	0	20	48.33	76.67	90	105	120
Resource limits							
WC	0	3	6	9	10	11	12
Shiftability	0	0	2	6	10	12	12
Maximal shiftability	0	3	9	15	19	21	21
Risk limits	0	0.04	0.34	0.64	0.84	0.89	0.94

Table 3
Allowable buffer consumption for four perspectives (in %).

	t_0	t_1	t_2	t_3	t_4	t_5	t_6
Linear limits	0	17	33	50	67	83	100
Cost limits	0	17	40	64	75	88	100
Resource limits	0	11	31	61	83	96	100
WC	0	25	50	75	83	92	100
Shiftability	0	0	17	50	83	100	100
Risk limits	0	4	36	68	89	95	100

Table 3 shows the allowable buffer consumption of each time period for all activities. For the linear, cost and risk limits, these proportion can be readily determined by dividing the progress at each time period by the total progress at the planned duration. Hence, for these limits, Table 3 depicts the normalised values of Table 2. For the resource limits, however, Table 3 shows the normalised values for the WC and shiftability. Accordingly, the allowable buffer consumption for the resource limits is determined as the weighted sum of the WC and shiftability, with weights (1 − *degreeof shiftability*) and *degree of shiftability*, respectively.

For instance, the allowable buffer consumption at time t_3 is 50% $\left(=\frac{3}{6}\right)$ for the linear limits, 64% $\left(=\frac{76.67}{120}\right)$ for the cost limits and 68% $\left(=\frac{0.64}{0.94}\right)$ for the risk limits. For the resource limits, however, an intermediate step is required, since the allowable buffer consumption depends on both the progress in terms of WC and the progress in terms of shiftability. First, the portion of the project buffer that is assigned to the shiftability progress is determined by dividing the actual shiftability of the schedule by the maximal shiftability. As shown in Table 2, the actual and maximal shiftability of the project are equal to 12 and 21 respectively. Therefore, 57% $\left(=\frac{12}{21}\right)$ of the project buffer is divided over the project phases proportionally with the shiftability. Similarly, 43% of the buffer is divided over the project phases proportionally with the WC. At time t_3, the proportion of the total WC is 75% $\left(=\frac{9}{12}\right)$. The proportion of the shiftability is equal to 50% $\left(=\frac{6}{12}\right)$. Consequently, the allowable buffer consumption at time t_3 according to the resource limits is equal to 61% $(= 0.43 \times 0.75 + 0.57 \times 0.5)$.

In Fig. 4, the allowable buffer consumption for the cost, resource and risk limits are graphically compared to the linear limits. As shown in the left pane, the cost limits allow a larger buffer consumption at each time period, due to the fact that a considerable part of the project

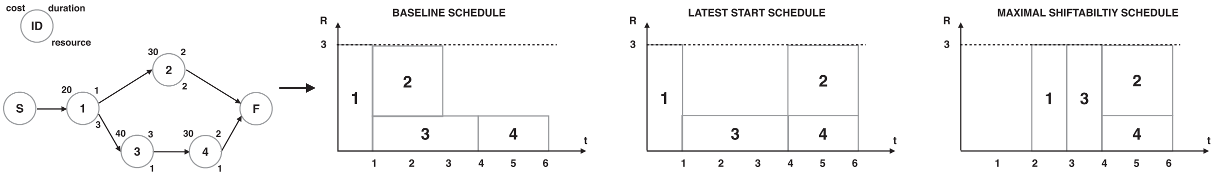

Fig. 3. Toy example.

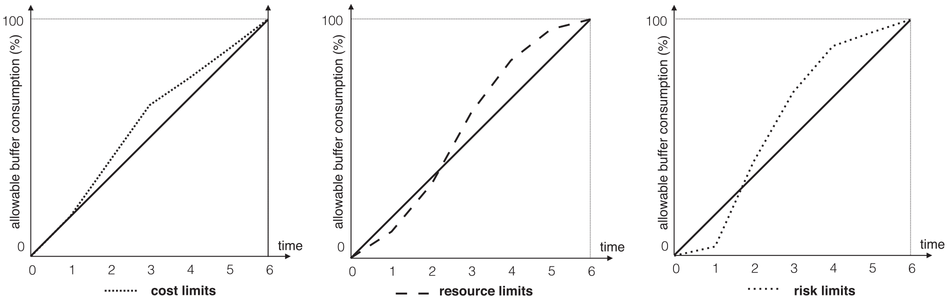

Fig. 4. Allowable buffer consumption.

is scheduled in the early phases. The middle pane shows that the resource limits are more severe in the early phase and less strict in the final phases, due to the limited shiftability in the early phases and the high shiftability in the middle phases. Further, since the first activity is very unrisky, the risk limits allow a very limited portion of the project buffer to be consumed in the early phases. Accordingly, a larger portion of the buffer is available for the middle phases, in which the most risky activity is scheduled.

Finally, the progress of two fictitious executions will be monitored and evaluated at time t_3 to illustrate how the actual SPI(t) limits are calculated. In Fig. 5, these fictitious executions are depicted. For both executions, a project buffer of 25% of the PD is assumed (= 1.5 days). During execution A, activity 3 has a delay of 50%. During execution B, both activity 2 and 3 suffer from a delay of 50%. While the delay of activity 3 affects the final project duration, both executions are completed on time due to the provided project buffer. Moreover, Fig. 5 shows that the delay in activity 2 has no impact on the final project duration.

In Fig. 6, the EV of execution A and the PV of the baseline schedule are depicted. Further, this figure shows how the ES can be determined graphically using the EV- and PV-curve. By dividing this ES by the actual time t_3, the SPI(t) at time t_3 is determined. Further, the BPP, which is required to determine the SPI(t) tolerance limit, is depicted for the cost perspective. As shown in Fig. 6, the appropriate BPP value to determine the tolerance limit for the project progress at t_3 is established as the ES of t_3 increased with the portion of the project buffer that is assigned to time ES. The tolerance limit is hence established by dividing the ES by the BPP. Since the BPP is higher than t_3 for execution A, no warning signal will be generated by the cost limits for execution A at time t_3. In the remainder of this section, the calculations to determine the schedule performance and tolerance limits will be demonstrated.

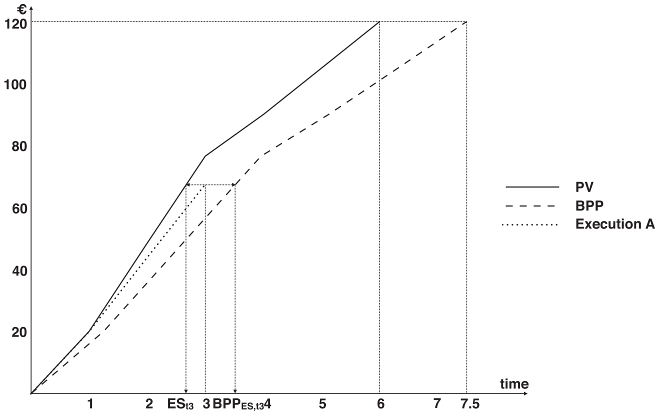

Fig. 6. Determine tolerance limit.

Table 4
Progress and performance metrics.

	Progress monitoring				Performance metrics	
	EV	Earned WC	ES	ES_{WC}	SPI(t)	$SPI(t)_{WC}$
Execution A	67.78	8.33	2.69	2.78	0.90	0.93
Execution B	57.78	7	2.33	2.33	0.78	0.78

Table 5
Threshold values at t_3.

	Allowable BC (%)		BPP		Tolerance limit	
	A	B	A	B	A	B
Time	45	39	3.36	2.91	0.80	**0.80**
Cost	56	48	3.53	3.06	0.76	0.76
Resource	69	58	3.82	3.21	0.73	0.73
Risk	54	47	3.49	3.04	0.77	0.77

In Table 4, the monitored progress and schedule performance of both executions is listed. In case the time, cost or risk perspectives are deployed to construct tolerance limits, the well-known EVM/ES performance metric SPI(t) is used to monitor and evaluate the progress. However, in the case the resource perspective is used, the progress is monitored and evaluated in terms of earned WC using the $SPI(t)_{WC}$ metric. This corresponds with the approach introduced in [6]. Therefore, Table 4 shows the progress for both the EV and earned WC. Using the information provided in Fig. 3 and Table 2, these values can be determined. For instance, the EV of execution A at time t_3 is equal to 67.78 $\left(=20+30+\frac{40}{4.5}\times 2\right)$. Further, the ES equals 2.69 $\left(=2+\frac{67.78-48.33}{76.67-48.33}\right)$. As a result, execution A has an SPI(t) of 0.90 $\left(=\frac{2.69}{3}\right)$ at time t_3.

In Table 5, the SPI(t) tolerance limits are determined for each perspective and each execution. First, the allowable buffer consumption is determined based on the perspective that is deployed. For instance, since the ES of execution A is equal to 2.69, 45% $\left(=\frac{2.69}{6}\right)$ of the PB is

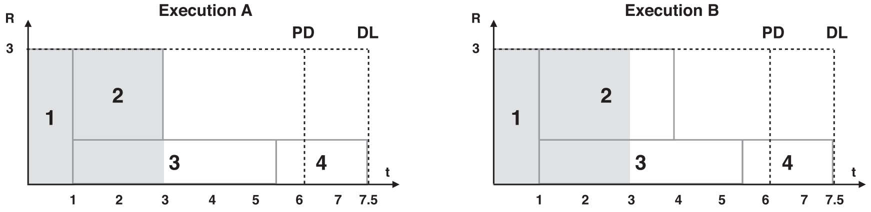

Fig. 5. Fictitious project executions.

allowed to be consumed at t_3 by execution A. Subsequently, the BPP is established by adding the allowable buffer consumption portion of the PB to the ES of the execution. For the time, cost and risk perspective, the traditional ES is used. The resource perspective, however, uses the ES_{WC}. Hence, the BPP for execution A at t_3 for the time perspective equals 3.36 ($= 2.69 + 0.45 \times 1.5$). Finally, the tolerance limits are determined by dividing the ES (or ES_{WC}) by the BPP. For the time perspective, this results in an SPI(t) tolerance limit of 0.80 $\left(=\frac{2.69}{3.36}\right)$.

Finally, when the SPI(t) of an execution is lower than the SPI(t) limit, a warning signal is generated. For the time, cost and risk perspective, the traditional SPI(t) is compared to the tolerance limits. For the resource perspective, the $SPI(t)_{WC}$ is evaluated. In Table 5, the limits that generate a warning signal are shown in bold. For execution A, no warning signals are generated by the four limits. For execution B, however, the time limits generate a warning signal due to the simultaneous delay of activities 2 and 3. Since the delay of activity 2 does not affect the project duration and the project finishes on time, this signal is considered a false warning signal.

3.3. *Design of experiment*

In the empirical experiment, the performance of the cost, resource and risk limits will be evaluated and compared to the linear benchmark using the performance metrics discussed in Section 3.4. For each tracking period at which the projects have been monitored, the tolerance limits are determined to evaluate the project progress. Since not all projects provide the necessary information to evaluate the resource and risk limits, this experiment consists of three parts, in which the cost, resource and risk limits will be compared to the linear benchmark using the appropriate projects. Therefore, the specific settings of each part of the experiment are discussed in the following paragraphs.

3.3.1. *Evaluation of cost limits*

In several simulation studies on the performance of various project control techniques, the *project seriality* has shown to have an important impact (e.g. in [15,48]). The project seriality is represented by the serial/parallel indicator SP, which indicates how close a project network is to a completely serial network (SP = 1) or a completely parallel network (SP = 0) [49].

The simulation studies performed to evaluate the performance of the cost and resource limits showed that the SP-indicator has an important impact on the efficiency of the limits [5,6]. However, [50] recently introduced a new project characteristic, the *project regularity*, which reflects the value accrue of the project. The authors define a project as completely regular when the PV-curve of the project is perfectly linear. In order to express the level of project regularity of projects, the regular/irregular indicator (*RI*) has been introduced (Eq. (3)), with m_i the maximal possible deviation of the PV-curve from a perfectly linear curve, a_i the actual deviation of the PV curve from a perfectly linear curve and r the number of equidistant evaluation points.

$$RI = \frac{\sum_{i=1}^{r}(m_i - a_i)}{\sum_{i=1}^{r} m_i} \tag{3}$$

For completely regular projects, the actual deviation from a perfectly linear curve will be zero at all evaluation points, resulting in an *RI* of 1. Further, a maximally regular project occurs when the PV of the project is zero throughout the entire makespan of the project, and reaches the BAC at the end of the project. Consequently, the *RI* of a maximally irregular project is equal to zero. In order to categorise projects as regular or irregular, an RI cut-off value of 80% is proposed by [50]. This value has been determined based on the RI of a perfect S-curve (81%). Consequently, projects with an RI greater than 80% are classified as regular, while projects with an RI below 80% are labelled as irregular.

As an illustrative example, the RI of the toy example introduced in Fig. 3 is determined. For this example project, there are 5 equidistant evaluation points, namely $t_1, t_2, \ldots, t_5$. Further, for a perfectly linear PV curve, the PV would increase with 20 units at each evaluation point. Hence, at t_3, m_3 is equal to 60 ($= 60 - 0$). With an actual PV of 76.67, a_3 is equal to 16.67 ($= |60 - 76.67|$). The RI of the project can be determined by calculating m_i and a_i for $i = 1,2,\ldots,5$ and is equal to 87% $\left(=\frac{260}{300}\right)$. Consequently, this example project would be classified as a regular project.

In this experiment, we have decided to examine the impact of the project regularity on the performance of the tolerance limits, rather than the project seriality, for the following two reasons. First, [50] showed that the project regularity has a stronger effect on the EVM forecasting accuracy than the widely used SP-indicator, due to its direct link with the value accrue of the project. Second, for completely regular projects, the PV-curve is completely linear. As a result, the cost limits will be identical to the linear benchmark limits in this case. Accordingly, it is expected that the cost limits for regular projects will not result in substantially different results compared to the linear limits. For irregular projects, however, the cost limits are expected to substantially differ from the linear benchmark. Therefore, the dataset of 93 projects has been partitioned in a set of 48 regular and 45 irregular projects. For both sets, the performance of the cost limits has been compared to the linear benchmark limits.

3.3.2. *Evaluation of resource limits*

Further, the real-life project data set can be divided into projects with or without available resource information. For the 21 projects with available resource information, the resource limits are constructed and compared to the linear limits.

3.3.3. *Evaluation of risk limits*

Similarly, the project dataset consists of 32 projects with complete risk information. In order to evaluate the performance of the proposed risk limits, they are compared to the linear benchmark limits.

3.4. *Performance measurement*

The performance of project control tolerance limits is defined by two characteristics. First of all, the capability of generating warning signals for projects that exceed their deadline should be as high as possible. Second, for projects that finish on time, the number of generated (false) warning signals should be as low as possible. These two characteristics are evaluated by the *efficiency* and the *reliability*, respectively [5]. Further, the performance of the tolerance limits can be evaluated at two levels, the project level and the tracking period level. At the *project level*, focus lies on reviewing whether or not one or more warning signals are generated during the monitoring process. At the *tracking period level*, the correctness of generated or absent warning signals is analysed for individual tracking periods. Finally, while simulation studies can provide multiple observations of project executions per project, using real-life data implies that only one observation of a project execution is available per project. Consequently, the performance metrics used for this experiment differ slightly from those used in [5]. Therefore, in the remainder of this section, the interpretation and definition of the applied performance metrics and the differences with the performance metrics used for simulation studies are discussed at the project and the tracking period level.

3.4.1. *Project level metrics*

The project level efficiency expresses the probability that a project will exceed its deadline when 1 or more warning signals are generated by the tolerance limits. Similarly, the project level reliability indicates the probability that a project will be completed before or at its deadline, when no warning signals are generated throughout the monitoring process. Both the efficiency and reliability express probabilities, indicating that multiple observations are required to calculate these metrics. Since simulation studies enable having multiple observations per project, the efficiency and reliability can be calculated per project when Monte Carlo simulations were used to generate fictitious project executions. Subsequently, the efficiency and reliability of the tolerance limits for the entire project dataset are defined as the average of all projects' efficiency and reliability. Consequently, for a given project dataset, the variability of the efficiency and reliability over the different projects can be determined as well. When real-life data are used, however, each project execution represents a single observation per project. As a result, the efficiency and reliability of individual projects cannot be determined. Consequently, the efficiency and reliability of the project dataset are determined by considering the single observations of all projects in the dataset. Therefore, the efficiency and reliability of a real-life project dataset are both a single value, rather than an average value over all projects. The project level efficiency and reliability are thus defined as follows:

$$efficiency = \mathrm{P}[Late|S_P] = \frac{\mathrm{P}(S_P|Late) \times \mathrm{P}(Late)}{\mathrm{P}(S_P)} \tag{4}$$

$$reliability = \mathrm{P}[OT|noS_P] = \frac{\mathrm{P}(noS_P|OT) \times \mathrm{P}(OT)}{\mathrm{P}(noS_P)} \tag{5}$$

with S_P indicating that one or more warning signals were generated for project P, noS_P indicating that no warning signals were generated for project P and $Late$ and OT representing projects exceeding their deadline or finishing before or at their deadline, respectively. For instance, when 500 out of 1000 project executions generated one or more warning signals of which 250 are late projects and 250 are on time projects, while only 300 of the 1000 executions are late projects, the efficiency and reliability of the monitoring system are respectively 50% $\left(=\frac{250/300 \times 300/1000}{\frac{500}{1000}}\right)$ and 90% $\left(=\frac{(700-250)/700 \times 700/1000}{500/1000}\right)$. Thus, while the efficiency indicates the probability that a project of the dataset will be late when one or more warnings signals are generated, the reliability reflects the probability that a project of the dataset will be completed on time when no warning signals are generated.

3.4.2. Tracking period level metrics

In [5], the signal efficiency and signal reliability have been defined as the relative number of correct warning signals and the relative number of times signals were correctly absent, respectively. However, contrarily to the project executions in the simulation study, the real-life projects in our data set have varying numbers of tracking periods. Consequently, in order to ensure a fair comparison, the signal efficiency and reliability in this study have been slightly adapted. Instead of considering the absolute number of signals generated per project execution, the proportion of tracking periods that generated a warning signal for late projects ($\%S_{TP,Late}$) and on time projects ($\%S_{TP,OT}$) are considered. As a result, each project has the same weight in the calculation of this metric. The adapted signal efficiency and signal reliability are defined in Eqs. (6) and (7). While the signal efficiency expresses the probability that a project will be late when a specific tracking period of that project generates a warning signal, the signal reliability reflects the probability that a project will finish on time when a specific tracking period has not generated a warning signal.

$$signal\ efficiency = \mathrm{P}[Late|S_{TP}] = \frac{\%S_{TP,Late} \times \mathrm{P}(Late)}{\%S_{TP}} \tag{6}$$

$$signal\ reliability = \mathrm{P}[OT|noS_{TP}] = \frac{(1-\%S_{TP,OT}) \times \mathrm{P}(OT)}{(1-\%S_{TP})} \tag{7}$$

4. Results and discussion

In this section, the performance of the analytical tolerance limits for empirical projects is evaluated and discussed using the performance metrics discussed in Section 3.4. In Section 4.1, the performance of the cost limits is compared to the linear benchmark for 93 real-life projects. The performance of the resource limits is evaluated on a set of 21 real-life projects in Section 4.2. Finally, the risk limits are validated on 32 real-life projects in Section 4.3.

Since the performance metrics described in this section are Bayesian metrics, the proportion of late and on time projects in the data set has an impact on their value. Thus, in order to ensure a fair comparison between the different experiments, the buffer size for each part of the experiment will be set such that an identical proportion of the projects in the test set will be late. More specifically, in each part of the experiment, the limits have been evaluated for buffer sizes such that the probability of late projects is equal to 10, 20 and 30%. Further, while Marten and Vanhoucke [5] showed that the efficiency of the linear and cost limits decreases for increasing buffer sizes, we did not empirically analyse the impact of the buffer size for the following two reasons. First, by fixing the proportional project buffer size for the three experiments (for instance at 5, 10, 15 and 20% of the PD), the performance metrics will be affected by the substantial differences in the number of late projects in each experiment. For instance, for a project buffer of 5%, the portion of late projects in the cost, resource and risk experiment is equal to 52, 33 and 47% respectively. Further, for a project buffer of 20%, these portions reduce to 17, 5 and 3%. Hence, due to these substantial discrepancies, the result of the three experiments would not be comparable. Second, due to the small number of projects with a delay of 20% or more in the resource and risk group (i.e. only 1 project in each group), we believe that the impact of the buffer size cannot be evaluated adequately given the available project dataset.

4.1. Cost limits

As discussed in Section 3.3, the 93 projects of our dataset are divided into regular and irregular projects to evaluate the performance of the cost limits. In the following paragraphs, the results of this experiment are discussed for the group of regular projects and the group of irregular projects.

4.1.1. Regular projects

The linear and cost limits have been evaluated for 48 regular projects. The results have shown that, as conjectured, the difference between both methods is negligible for regular projects. Consequently, employing the cost limits for regular projects does not provide any advantage over using the linear benchmark limits.

4.1.2. Irregular projects

The cost limits have been evaluated for the group of 45 irregular projects as well. In Fig. 7, the performance of the linear and cost limits for the irregular projects in our dataset are depicted. The upper pane shows the project level efficiency and reliability. Further, the signal efficiency and signal reliability are presented in the lower pane. For each performance metric, the results are listed for a project buffer resulting in 10, 20 and 30% of the projects being late. Further, the results for the linear limits are shown by the light grey bar, while the dark grey bars depict the performance of the cost limits.

At the project level, the reliability of both limits is equal to 1 for all project buffer settings. Since the reliability indicates the probability that a project will be completed on time when no warning signals are generated, this implies that all projects that finish late are detected by the tolerance limits. Further, the efficiency of the cost limits is slightly higher than the efficiency of the linear limits, indicating that less false warnings are generated. This difference, however, is marginal. At the tracking period level, both the signal efficiency and signal reliability of the cost limits are higher than the linear benchmark.

Compared to the results of the simulation study of Marten and Vanhoucke [5], the difference in performance between the cost limits

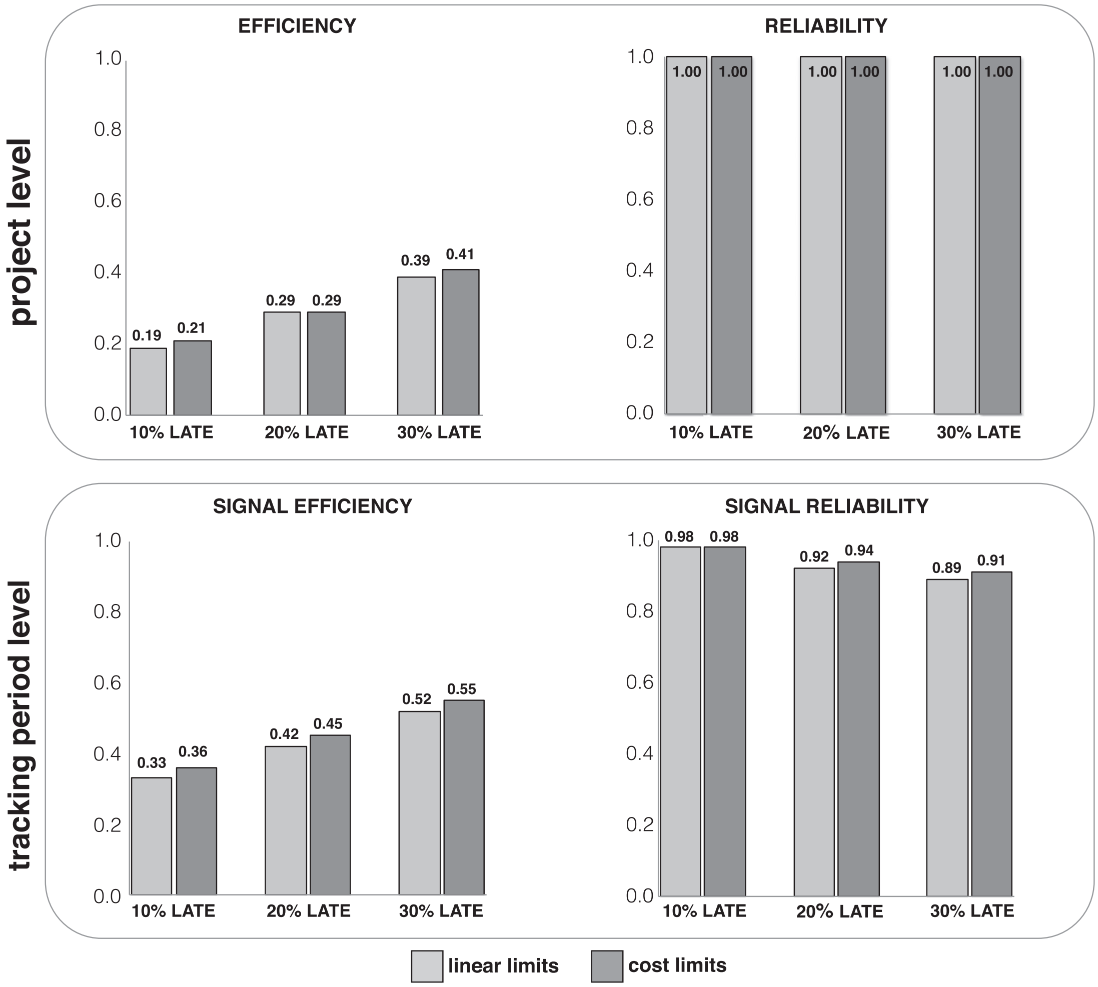

Fig. 7. Performance of cost limits (irregular projects).

and linear limits is less pronounced. This indicates that in real-life situations the relation between the activity duration and activity cost is less explicit. An analysis of the project database showed that, opposed to simulated projects, for real-life projects the portion of fixed activity costs is indeed higher compared to the variable activity costs. However, while the performance improvement of the cost limits is only marginal compared to the linear limits, the additional effort to construct these limits is limited as well. Therefore, when corrective actions are expensive and should thus be avoided when not necessary, the cost limits are preferred over the linear limits.

4.2. Resource limits

The performance of the linear and resource limits is depicted in Fig. 8. At the project level, the reliability of the resource and linear limits is equal to 1 for each project buffer setting. Further, Fig. 8 shows that the efficiency of the linear limits is rather low for the 21 projects with sufficient resource information. Moreover, the efficiency of the resource limits is substantially higher for each project buffer setting. Further, the enhanced efficiency of the resource limits is confirmed at the tracking period level. In terms of signal reliability, however, the resource limits performance slightly less compared to the linear benchmark.

This empirical experiment confirms the conclusion of [6] that adding the available resource information into the monitoring process significantly improves the efficiency of this process. It is shown that the linear limits are not efficient for projects with scarce resources. Further, while constructing resource limits requires some additional effort, the using resource limits increases the efficiency of the monitoring process considerably. Consequently, when the required resource information is available, the resource limits are preferred over the linear limits.

4.3. Risk limits

Fig. 9 shows the performance of the risk limits compared to the linear benchmark. The project reliability for both limits is again equal to 1. Further, the project efficiency of the risk limits outperforms the linear benchmark for all project buffer settings. At the tracking period level, the signal efficiency of the resource limits is higher than the linear benchmark. However, for an increasing portion of late projects in the dataset, the difference between both limits decreases. The signal reliability shows an opposite trend. The signal reliability of the risk limits is lower than the linear benchmark, and this difference increases with the portion of late projects in the dataset.

The results at the project level imply that the risk limits generate less false warning signals than the linear limits, and are thus more efficiency. At the tracking period level, this observation is confirmed by the increased signal efficiency of the risk limits. At the same time, however, the reduced signal reliability indicates that less correct warning signals are generated per project. Consequently, for projects where corrective actions are expensive and risk information on the activity duration is available, the risk limits could be deployed to reduce the number of false warning signals and to detect late projects efficiently.

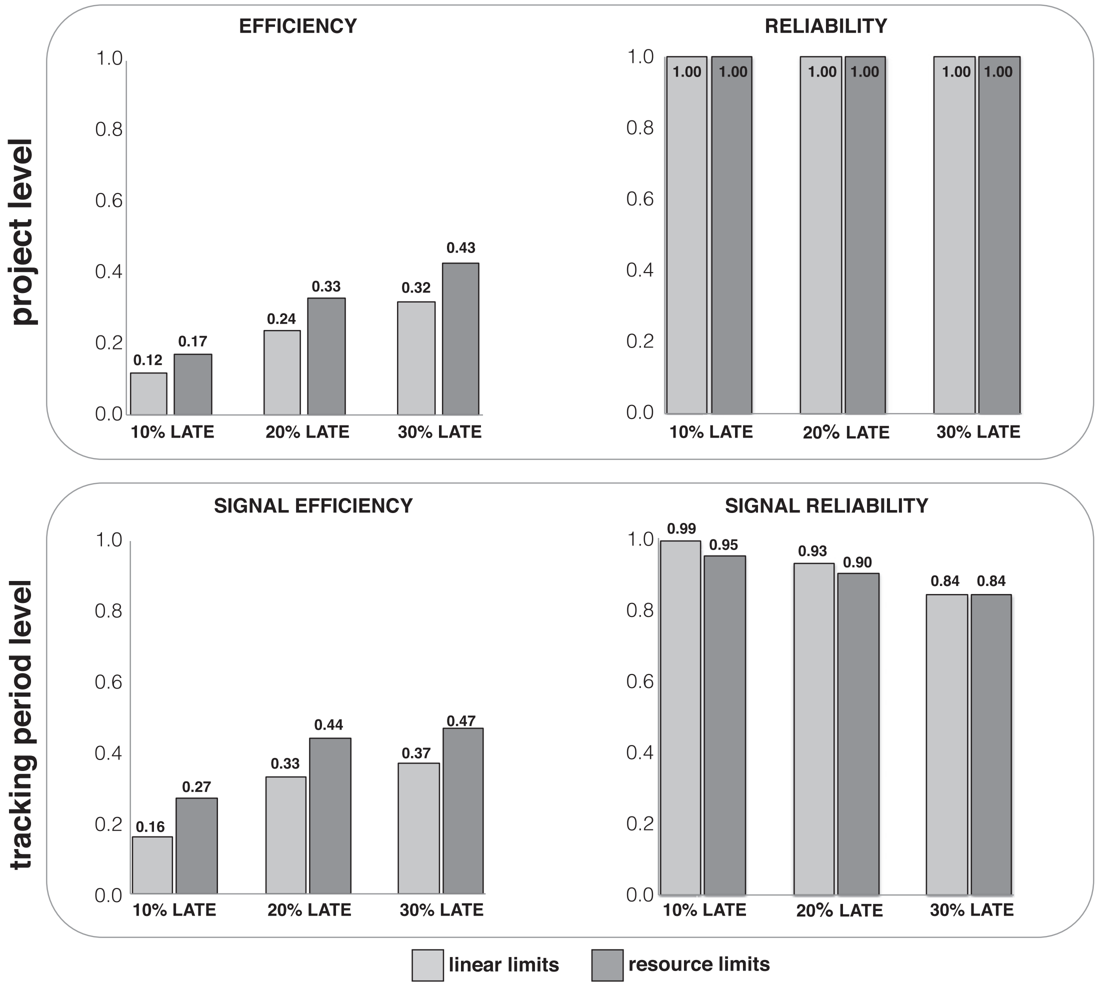

Fig. 8. Performance of resource limits.

5. Conclusion

In this paper, three types of analytical tolerance limits proposed in literature have been empirically validated on a large set of real-life project data, mainly situated in the construction sector. Further, a novel approach to construct analytical tolerance limits that incorporate activity risk information are introduced and empirically validated.

Three experiments have been conducted in which the cost, resource and risk limits have been compared to the linear benchmark limits. First, the cost limits established in [5] are evaluated on a set of 93 real-life projects. These cost limits adopt a cost perspective by assigning a portion of the project buffer to each project phase proportionally with the planned cost of that phase. The performance of the cost limits is compared to the linear limits for 48 regular projects and 45 irregular projects. Second, the resource limits have been implemented for the 21 projects in the data set that contained the required information. These limits consider a resource perspective for two reasons. Firstly, expressing the planned and earned progress in terms of value is, as demonstrated in the first experiment, not necessarily accurate for real-life projects. Therefore, [6] propose to express and monitor the planned and earned progress in terms of work content units instead of costs. Secondly, when renewable resources are limitedly available, resource conflicts are an important cause of project delays. Therefore, the shiftability of each project phase, defined as the ability to shift activities to the right without endangering the project makespan, is considered in the construction of the resource limits. In the third experiment, the risk limits have been evaluated on a set of 32 projects that provided the required activity risk information. These limits have been introduced to improve the performance of analytical tolerance limits by considering the activity duration variability.

In terms of reliability, each of the tolerance limits has a maximal performance of 1, and thus no difference between the three experiments has been observed. Further, the main distinctive findings of the experiments can be summarised as follows. First, as conjectured, the performance of the cost limits is almost identical to the linear benchmark for regular projects. For irregular projects, a slight improvement of the efficiency compared to the linear benchmark has been observed. This improvement, however, is lower than in the simulation study conducted in [5], indicating that the relation between the activity duration and activity cost is less straightforward for real-life projects. Second, the experiment showed that the linear limits did not suffice to efficiently monitor the progress of projects that are limited by scarce resources. The resource limits, however, substantially increase the efficiency of the monitoring process. Consequently, while additional effort is required to deploy the resource limits, they are preferred over the linear limits due to their enhanced efficiency. Finally, the risk limits generate in general less warning signals than the linear benchmark limits. At the project level, this results in a considerably higher efficiency compared to the benchmark. At the tracking period level, an increased signal efficiency but decreased signal reliability has been observed. This lower signal reliability indicates that less warning signals are generated per late project. However, the project reliability, which is equal to 1, indicates that all late projects are still detected. Consequently, when no resource information is available or when the

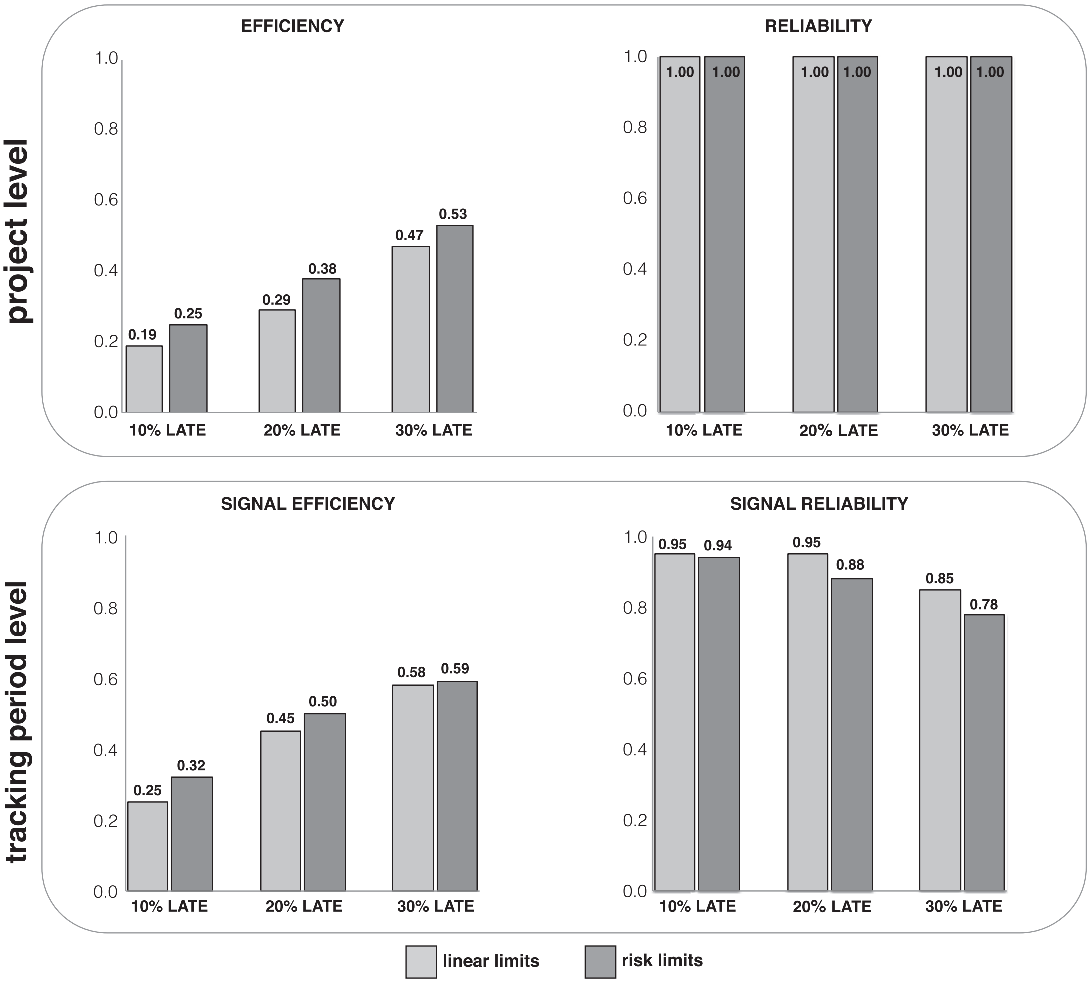

Fig. 9. Performance of risk limits.

project execution is not limited by scarce resources, risk limits are an adequate alternative to improve the efficiency of the monitoring process.

Consequently, while the cost, resource and risk limits are as reliable as the linear limits, their efficiency is higher. In general, while the resource limits require the most additional effort to be constructed, they increase the efficiency of the monitoring process the most as well. Further, when the project is not limited by scarce renewable resources, but the activity durations are known to vary to a certain extent, the novel risk limits are an appropriate alternative to increase the monitoring efficiency. Furthermore, the experiment has shown that the cost limits are easy to implement and are applicable on many projects. Therefore, when no information on the availability and requirements of the renewable resources and on the activity duration variability is known, implementing the cost limits can increase the efficiency of the monitoring process for irregular projects. Finally, the linear limits can be used to monitor regular projects for which no resource and risk information is available.

Appendix A. List of projects

Table A.6
List of used projects.

Project code	Early/late (in %)	Cost limits	Resource limits	Risk limits
C2011-13	14.3	x	–	–
C2012-13	12.0	x	x	–
C2013-01	0.0	x	–	–
C2013-02	0.0	x	–	–
C2013-03	0.2	x	–	–
C2013-04	36.0	x	–	–
C2013-06	0.0	x	–	–
C2013-07	17.6	x	–	–
C2013-08	8.8	x	–	–
C2013-09	23.7	x	–	–
C2013-10	1.0	x	–	–

C2013-11	−7.8	x	–	–
C2013-12	72.2	x	–	–
C2013-13	−8.1	x	–	–
C2013-14	23.8	x	–	–
C2013-15	−29.8	x	–	–
C2013-16	−33.2	x	–	–
C2013-17	−18.0	x	–	–
C2014-01	0.0	x	–	–
C2014-02	17.7	x	–	–
C2014-03	13.5	x	x	–
C2014-05	20.2	x	–	–
C2014-06	11.7	x	–	–
C2014-07	14.4	x	–	–
C2014-08	18.0	x	–	–
C2015-01	60.3	x	–	–
C2015-02	20.1	x	–	–
C2015-03	9.7	x	–	–
C2015-04	34.4	x	–	–
C2015-05	9.2	x	–	–
C2015-06	13.5	x	–	–
C2015-07	21.2	x	–	–
C2015-08	4.7	x	–	–
C2015-09	61.0	x	–	–
C2015-10	0.0	x	x	–
C2015-11	0.0	x	x	–
C2015-12	0.0	x	x	–
C2015-13	0.0	x	x	–
C2015-14	0.0	x	x	–
C2015-15	−1.8	x	x	–
C2015-16	0.0	x	x	–
C2015-17	13.4	x	x	–
C2015-18	0.0	x	x	–
C2015-19	0.5	x	x	–
C2015-20	0.0	x	x	–
C2015-21	0.0	x	x	–
C2015-22	0.0	x	x	–
C2015-23	−1.6	x	x	–
C2015-24	−0.4	x	x	–
C2015-25	7.1	x	x	–
C2015-26	5.4	x	x	–
C2015-27	20.6	x	x	–
C2015-28	10.4	x	x	–
C2015-29	5.3	x	–	–
C2015-30	4.1	x	–	–
C2015-31	36.2	x	–	–
C2015-32	−25.7	x	–	–
C2015-33	100.0	x	–	–
C2015-34	98.3	x	–	–
C2015-35	10.0	x	–	–
C2016-01	27.6	x	–	x
C2016-02	10.5	x	–	x
C2016-03	13.3	x	–	x
C2016-04	1.6	x	–	x
C2016-05	8.7	x	–	x
C2016-06	7.9	x	–	x
C2016-07	48.2	x	–	–
C2016-08	2.1	x	–	x
C2016-09	−1.0	x	–	x
C2016-11	5.4	x	–	x
C2016-12	0.0	x	–	x
C2016-13	7.8	x	–	x
C2016-14	0.0	x	–	x
C2016-15	3.2	x	–	x
C2016-16	0.0	x	–	x
C2016-17	0.0	x	–	x
C2016-18	0.0	x	–	x

C2016-19	0.0	x	–	x
C2016-20	0.0	x	–	x
C2016-21	0.0	x	–	x
C2016-22	0.0	x	–	x
C2016-23	0.0	x	–	x
C2016-24	0.0	x	–	x
C2016-25	0.0	x	–	x
C2016-26	0.0	x	–	x
C2016-27	12.8	x	–	x
C2016-28	11.3	x	–	x
C2016-29	14.7	x	–	x
C2016-30	12.9	x	–	x
C2016-31	11.4	x	–	x
C2016-32	9.0	x	–	x
C2016-33	11.2	x	–	x
C2016-34	10.8	x	–	x

References

[1] J. Batselier, M. Vanhoucke, Construction and evaluation framework for a real-life project database, Int. J. Proj. Manag. 0263-7863, 33 (2015) 697–710, http://dx.doi.org/10.1016/j.ijproman.2014.09.004.

[2] M. Vanhoucke, Integrated Project Management and Control: First Come the Theory, Then the Practice, Springer, 978-3-319-04331-9, 2014.

[3] Q. Fleming, J. Koppelman, Earned Value Project Management, 3rd edition, Project Management Institute, Newton Square, Pennsylvania, 9781935589082, 2010.

[4] J. Colin, M. Vanhoucke, A comparison of the performance of various project control methods using earned value management systems, Expert Syst. Appl. 0957-4174, 42 (2015) 3159–3175, http://dx.doi.org/10.1016/j.eswa.2014.12.007.

[5] A. Martens, M. Vanhoucke, A buffer control method for top-down project control, Eur. J. Oper. Res. 0377-2217, 262 (2017) 274–286, http://dx.doi.org/10.1016/j.ejor.2017.03.034.

[6] A. Martens, M. Vanhoucke, The integration of constrained resources into top-down project control, Comput. Industr. Eng. 0360-8352, 110 (2017) 277–288, http://dx.doi.org/10.1016/j.cie.2017.05.020.

[7] J. Kelley, M. Walker, Critical Path Planning and Scheduling: An Introduction, Mauchly Associates, Ambler, PA, 1959.

[8] W. Fazar, Federal statistical activities, Am. Stat. 0003-1305, 13 (1959) 9–12 http://www.jstor.org/stable/2681377 last visited 09/01/2017.

[9] S. Hartmann, D. Briskorn, A survey of variants and extensions of the resource-constrained project scheduling problem, Eur. J. Oper. Res. 0377-2217, 207 (2010) 1–15, http://dx.doi.org/10.1016/j.ejor.2009.11.005.

[10] W. Herroelen, B. De Reyck, E. Demeulemeester, Resource-constrained project scheduling: a survey of recent developments, Comput. Oper. Res. 0305-0548, 25 (1998) 279–302, http://dx.doi.org/10.1016/S0305-0548(97)00055-5.

[11] R. Kolisch, R. Padman, An integrated survey of deterministic project scheduling, OMEGA Int. J. Manag. Sci. 0305-0483, 29 (3) (2001) 249–272, http://dx.doi.org/10.1016/S0305-0483(00)00046-3.

[12] M. Bruni, L. Di Puglia Pugliese, P. Beraldi, F. Guerriero, An adjustable robust optimization model for the resource-constrained project scheduling problem with uncertain activity durations, OMEGA Int. J. Manag. Sci. 0305-0483, 71 (2017) 66–84, http://dx.doi.org/10.1016/j.omega.2016.09.009.

[13] A. Laurent, L. Deroussi, N. Grangeon, S. Norre, A new extension of the RCPSP in a multi-site context: mathematical model and metaheuristics, Comput. Indust. Eng. 0360-8352, 112 (2017) 634–644, http://dx.doi.org/10.1016/j.cie.2017.07.028.

[14] M. Tritschler, A. Naber, R. Kolish, A hybrid metaheuristic for resource-constrained project scheduling with flexible resource profiles, Eur. J. Oper. Res. 0377-2217, 262 (1) (2017) 262–273, http://dx.doi.org/10.1016/j.ejor.2017.03.006.

[15] M. Vanhoucke, Using activity sensitivity and network topology information to monitor project time performance, OMEGA Int. J. Manag. Sci. 0305-0483, 38 (2010) 359–370, http://dx.doi.org/10.1016/j.omega.2009.10.001.

[16] M. Vanhoucke, On the dynamic use of project performance and schedule risk information during project tracking, OMEGA Int. J. Manag. Sci. 0305-0483, 39 (2011) 416–426, http://dx.doi.org/10.1016/j.omega.2010.09.006.

[17] S. Elmaghraby, On criticality and sensitivity in activity networks, Eur. J. Oper. Res. 0377-2217, 127 (2000) 220–238, http://dx.doi.org/10.1016/S0377-2217(99)00483-X.

[18] F.Y. Partovi, J. Burton, Timing of monitoring and control of CPM projects, IEEE Trans. Eng. Manag. 0018-9391, 40 (1) (1993) 68–75, http://dx.doi.org/10.1109/17.206653.

[19] T. Raz, E. Erel, Optimal timing of project control points, Eur. J. Oper. Res. 0377-2217, 127 (2) (2000) 252–261, http://dx.doi.org/10.1016/S0377-2217(99)00492-0.

[20] N. Sabeghi, H. Tareghian, E. Demeulemeester, H. Taheri, Determining the timing of project control points using a facility location model and simulation, Comp. Oper. Res. 0305-0548, 61 (2015) 69–80, http://dx.doi.org/10.1016/j.cor.2015.03.006.

[21] R. Elshaer, Impact of sensitivity information on the prediction of project's duration using earned schedule method, Int. J. Proj. Manag. 0263-7863, 31 (2013) 579–588, http://dx.doi.org/10.1016/j.ijproman.2012.10.006.

[22] M. Wauters, M. Vanhoucke, A nearest neighbour extension to project duration forecasting with artificial intelligence, Eur. J. Oper. Res. 0377-2217 0377-2217 0377-2217, 259 (3) (2017) 1097–1111, http://dx.doi.org/10.1016/j.ejor.2016.11.018.

[23] O.I. Tukel, W.O. Rom, S.D. Eksioglu, An investigation of buffer sizing techniques in critical chain scheduling, Eur. J. Oper. Res. 0377-2217, 172 (2) (2006) 401–416, http://dx.doi.org/10.1016/j.ejor.2004.10.019.

[24] H. Iranmanesh, F. Mansourian, S. Kouchaki, Critical chain scheduling: a new approach for feeding buffer sizing, Int. J. Oper. Res. 1745-7645, 25 (1) (2016) 114–130, http://dx.doi.org/10.1504/IJOR.2016.073254.

[25] J. Zhang, X. Song, E. Diaz, Project buffer sizing of a critical chain based on comprehensive resource tightness, Eur. J. Oper. Res. 0377-2217, 248 (1) (2016) 174–182, http://dx.doi.org/10.1016/j.ejor.2015.07.009.

[26] M. Madadi, H. Iranmanesh, A management oriented approach to reduce a project duration and its risk (variability), Eur. J. Oper. Res. 0377-2217, 219 (3) (2012) 751–761, http://dx.doi.org/10.1016/j.ejor.2012.01.006.

[27] W. Lipke, Schedule is different, The Measurable News Summer, (2003), pp. 31–34 http://earnedschedule.com/Papers.shtml last visited 09/01/2017.

[28] M. Vanhoucke, Measuring Time - Improving Project Performance Using Earned Value Management, International Series in Operations Research and Management Science, vol. 136, Springer, 978-1-4419-1014-1, 2010.

[29] L. Willems, M. Vanhoucke, Classification of articles and journals on project control and earned value management, Int. J. Proj. Manag. 0263-7863, 33 (7) (2015) 1610–1634, http://dx.doi.org/10.1016/j.ijproman.2015.06.003.

[30] E. Goldratt, Critical Chain, North River Press, Great Barrington, MA, 9780884271536, 1997.

[31] L.P. Leach, Critical Chain Project Management, vol. 2nd, Artech House, 9781580534192, 2005.

[32] W.A. Shewhart, Economic Control of Quality of Manufactured Product, vol. 509, ASQ Quality Press, 9780873890762, 1931.

[33] R. Aliverdi, L. Moslemi Naeni, A. Salehipour, Monitoring project duration and cost in a construction project by applying statistical quality control charts, Int. J. Proj. Manag. 0263-7863, 31 (3) (2013) 411–423, http://dx.doi.org/10.1016/j.ijproman.2012.08.005.

[34] G.T. Bauch, C.A. Chung, A statistical project control tool for engineering managers, Proj. Manag. J. 1938-9507, 32 (2001) 37–44 https://www.pmi.org/learning/library/statistical-_project-_control-_tool-_engineering-_managers-_5327 last visited 09/01/2017.

[35] S.S. Leu, Y.C. Lin, Project performance evaluation based on statistical process control techniques, J. Constr. Eng. Manag. 0733-9364, 134 (2008) 813–819, http://dx.doi.org/10.1061/(ASCE)0733-9364(2008)134:10(813).

[36] W. Lipke, J. Vaughn, Statistical process control meets earned value, CrossTalk: J. Def. Softw. Eng. 2160-1577(June, 2000), http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.639.8551&rep=rep1&type=pdf 16–20, 28–29 last visited 09/01/2017.

[37] Q. Wang, N. Jiang, L. Gou, M. Che, R. Zhang, Practical experiences of cost/schedule measure through earned value management and statistical process control, Lect. Notes Comput. Sci 0302-9743, 3966 (2006) 348–354, http://dx.doi.org/10.1007/11754305_38.

[38] J. Colin, M. Vanhoucke, Setting tolerance limits for statistical project control using earned value management, OMEGA Int. J. Manag. Sci. 0305-0483, 49 (2014) 107–122, http://dx.doi.org/10.1016/j.omega.2014.06.001.

[39] J. Colin, A. Martens, M. Vanhoucke, M. Wauters, A multivariate approach for top-down project control using earned value management, Dec. Supp. Syst. 0167-9236, 79 (2015) 65–76, http://dx.doi.org/10.1016/j.dss.2015.08.002.

[40] X. Hu, N. Cui, E. Demeulemeester, L. Bie, Incorporation of activity sensitivity measures into buffer management to manage project schedule risk, Eur. J. Oper. Res. 0377-2217, 249 (2015) 717–727, http://dx.doi.org/10.1016/j.ejor.2015.08.066.

[41] OR-AS, Online consultation of the real-life project database, URL: http://www.projectmanagement.ugent.be/research/data/realdata, last visited 09/01/2017, 2017.

[42] J. Batselier, M. Vanhoucke, Evaluation of deterministic state-of-the-art forecasting approaches for project duration based on earned value management, Int. J. Proj. Manag. 0263-7863, 33 (7) (2015) 1558–1596, http://dx.doi.org/10.1016/j.ijproman.2015.04.003.

[43] J. Batselier, M. Vanhoucke, Improving project forecast accuracy by integrating earned value management with exponential smoothing and reference class forecasting, Int. J. Proj. Manag. 0263-7863, 35 (1) (2017) 28–43, http://dx.doi.org/10.1016/j.ijproman.2016.10.003.

[44] J. Colin, M. Vanhoucke, Empirical Perspective on Activity Durations for Project-Management Simulation Studies, J. Constr. Eng. Manag. 0733-9364, 142 (1) (2016), http://dx.doi.org/10.1061/(ASCE)CO.1943-7862.0001022.

[45] R. Kolisch, Efficient priority rules for the resource-constrained project scheduling problem, J. Oper. Manag. 0272-6963, 14 (1996) 179–192, http://dx.doi.org/10.1016/0272-6963(95)00032-1.

[46] OR-AS, Project management software tool ProTrack url: www.protrack.be, last visited 09/01/2017, 2017.

[47] D. Trietsch, The effect of systemic errors on optimal project buffers, Int. J. Proj. Manag. 0263-7863, 23 (4) (2005) 267–274, http://dx.doi.org/10.1016/j.ijproman.2004.12.004.

[48] M. Wauters, M. Vanhoucke, A comparative study of Artificial Intelligence methods for project duration forecasting, Expert Syst. Appl. 0957-4174, 46 (2016) 249–261, http://dx.doi.org/10.1016/j.eswa.2015.10.008.

[49] M. Vanhoucke, J. Coelho, D. Debels, B. Maenhout, L. Tavares, An evaluation of the adequacy of project network generators with systematically sampled networks, Eur. J. Oper. Res. 0377-2217, 187 (2008) 511–524, http://dx.doi.org/10.1016/j.ejor.2007.03.032.

[50] J. Batselier, M. Vanhoucke, Project regularity: development and evaluation of a new project characteristic, J. Syst. Sci. Syst. Eng. 1004-3756, 26 (1) (2017) 100–120, http://dx.doi.org/10.1007/s11518-016-5312-6.