Manuscript in submission

Research undertaken on behalf of Ghent University, Faculty of Economics and Business Administration

WORKING PAPER

A multivariate approach for top-down project control using earned value management

Jeroen Colin and Mario Vanhoucke

November 2014

A multivariate approach for top-down project control using earned value management

Jeroen Colin[a], Mario Vanhoucke[a,b,c,*]

[a] *Faculty of Economics and Business Administration, Ghent University, Tweekerkenstraat 2, 9000 Gent (Belgium)*

[b] *Technology and Operations Management, Vlerick Business School, Reep 1, 9000 Gent (Belgium)*

[c] *Department of Management Science and Innovation, University College London, Gower Street, London WC1E 6BT (United Kingdom)*

Abstract

The application of a top-down project schedule control process requires an advanced decision support system (DSS), since aggregated performance measures to control a project's makespan are likely to obscure details on the activity level schedule progress. Earned value management/earned schedule (EVM/ES) is typically applied in a DSS for top-down project schedule control. However, traditional models do not correctly account for the multivariate nature of the EVM/ES measurement system. We therefore propose a multivariate model for EVM/ES, which implements a principal component analysis (PCA) on a simulated schedule control reference. During project progress, the real EVM/ES observations can then be projected onto these principal components. This allows for two new multivariate schedule control metrics (T^2 and SPE) to be calculated, which can be dynamically monitored on project control charts. Using a computational experiment, we show that these multivariate schedule control metrics lead to enhanced decision making in top-down schedule control, when compared to traditional univariate EVM/ES models.

Keywords: project management, schedule control, earned value management (EVM), simulation, principal component analysis (PCA)

*Corresponding author

Email addresses: `jeroen.colin@ugent.be` (Jeroen Colin), `mario.vanhoucke@ugent.be` (Mario Vanhoucke)

1. Introduction

The focus of this research is on the development of a decision support system (DSS) for project schedule control. Although cost and scope were also identified as important dimensions of project management success [1], we will only focus on the control of the schedule performance of a project. If the project manager has a detailed overview of the progress of all the individual activities, at any given time in the project, a decision support framework would be a relatively straightforward one. However, it has been argued in literature that the activity level monitoring of all activities would be a cumbersome and often disruptive task for the project manager [2, 3]. We therefore focus on a DSS that is implemented on a higher level of the work breakdown structure (WBS) and which can be situated within the risk monitoring and control phase of a classical project risk management (PRM, [4]) process. If the project is monitored at these high levels of the WBS, the need for an advanced DSS becomes apparent. In this case, only aggregated measures of the project schedule performance are monitored regularly and details on the underlying activity level performance might be obscured. The DSS that we present here, aims to help the project manager in making the decision whether he/she should invest time and effort to obtain a more detailed (activity level) view of the project schedule performance, in order to take further actions to bring the project back on track. This drill-down of the WBS in order to obtain a more detailed activity look is known as top-down control [3, 5]. In our research, we focus on the development of such a DSS and we will not incorporate possible actions that can be taken to bring the project back on track. Earned value management/Earned schedule (EVM/ES) is often implemented in a DSS for top-down project schedule control, and we will discuss its use in this paper due to its popularity in practice. The subject of EVM/ES has received a lot of attention from academics in recent years, as can be seen from the numerous publications on case studies [6, 7], consolidated research experiments [3, 5] and possible extensions [2, 8–10].

We will briefly introduce the EVM/ES methodology for project schedule control using the three fundamental components of any DSS architecture, i.e. the database, the model and the user interface [11].

Database. The data that are used in an EVM/ES system are aggregated performance metrics calculated at the project level of the WBS from the earned value (EV) and the planned value (PV). These metrics compare the preliminary planning variable (PV) with the actual performance (EV) in order to asses the schedule progress of the project and to indicate the direction of changes with respect to the baseline schedule. The EV is the budgeted value of the work that is performed in the project, whereas the PV is the budgeted value of the work that was scheduled to be performed [8]. The schedule variance (SV=EV-PV) and schedule performance index (SPI=EV/PV) can be directly calculated from the EV and PV, which are expressed in monetary units. The earned schedule (ES) translates the EV into time units and expresses the time since the beginning of the project at which the current EV should have been accrued, according to the baseline schedule. When compared to the actual time (AT) since the beginning of the project, the ES gives rise to the schedule variance using earned schedule (SV(t)) and the schedule performance index using earned schedule (SPI(t)).

Model. The model to interpret the schedule control metrics, provided by the EVM/ES system, has been subject to some debate in the literature. Straightforward control limits for these metrics where proposed [12] as well as implementations of statical process control charts [13–15] and statistical tolerance limits [5]. However, these methods do not account for the project-specific dynamics of the EVM/ES metrics [5] or fail to accurately incorporate the multivariate nature of the EVM/ES system. As a consequence, the use of EVM/ES in practice is typically characterised by decision making based on experience, arbitrary rules-of-thumb and anecdotal evidence. In order to aid the project manager in his/her decision making process, the EVM/ES system is often implemented in software. Commercial products have emerged on the market (for example, Decision Edge [16] to go along with Microsoft Project, or the stand-alone applications ProTrack [17] and P2 Engine [18]) to present the project manager with the schedule performance metrics throughout the execution of the project.

User interface. The user interface is mostly comprised of a set of charts that depict the time-series of the different schedule performance metrics. The interpretation of these charts can be impeded by the shear overload of observations presented to the project manager, the redundancy in these observations or the noise that is inherent to the EVM/ES system. Ideally, the model that is applied to the EVM/ES data should handle these characteristics, in order for the user interface to be effective in driving the decision process.

In this paper, we present a multivariate model to be applied in a DSS using EVM/ES for top-down project schedule control, that correctly accounts for the characteristics that were briefly introduced above. We will elaborate on these characteristics in section 2. We account for these characteristics by analysing the correlation structure of EVM/ES measurements over the course of the execution of the project in a simulated schedule control reference. From this correlation structure, we calculate latent variables that reflect the underlying dynamics of the EVM/ES measurement system. These latent variables are identified as principal components, calculated by a principal component analysis (PCA). We will introduce PCA in section 3. The principal components represent the basis for a new coordinate system onto which the EVM/ES observations during the project execution can be projected. Using this projection, we will show in section 4 how two new performance metrics can be calculated, which accurately incorporate the multivariate nature of the EVM/ES system. These new multivariate control metrics can be presented on project control charts in order to steer the decision process in top-down project schedule control. We will demonstrate the enhanced decision making that results from these two new metrics using a large computational experiment, outlined in section 5. Section 6 provides the results and overall conclusions are drawn in section 7.

Four appendices to this paper are available online. They can be freely accessed on the `statistical project control` research page on `www.projectmanagement.ugent.be`. In appendix A, a more detailed overview is given for the mathematics involved in calculating a PCA. Appendix B contains a geometrical interpretation for a PCA. In appendix B, we introduces the theoretical distributions that are used in the literature on batch process control to derive control limits for the T^2 and the SPE metrics. Finally, in appendix D a numerical example for a project schedule control process using the multivariate T^2 and SPE metrics is presented.

2. Multivariate nature of schedule control

In this section, we first present a formal characterisation of the top-down project schedule control process. This characterisation will be used in this paper to quantify the performance of the decision process that follows from a DSS using EVM/ES for top-down project schedule control. Second, we will discuss the multivariate nature of the EVM/ES system and we will show which aspects need to be considered in order to improve the top-down schedule control process. Finally, we will briefly introduce the body of literature that has been developed to handle these aspects in a batch process control context.

In order to present a formal characterisation of the top-down schedule control process, let us consider a vector $\boldsymbol{x}$ of EVM/ES measurements along the lifetime of the project as an observation for a multivariate random variable X.

This random variable X represents the schedule performance measurements of a project as observed at the top WBS level (SV, SPI, SV(t), SPI(t)). In EVM/ES, X is a function of the underlying activity level performance, which can be expressed as the multivariate random variable D, containing the real durations for all activities in the project. From a computational perspective, it is not opportune to derive a closed-form analytical expression for the function $X = f(D)$ for projects with a large number of activities. Rather, we wish to deduce the activity level performance $\boldsymbol{d}$ from the vector of measurements $\boldsymbol{x}$ during the execution of the project. In this research, we do not intend to calculate the activity level schedule performance explicitly, but we would like to infer *a state of schedule control*, i.e. whether the activity durations conform a pre-defined state of control.

This implicit inference process will be used to quantify the performance of the decision making that follow from our multivariate model for a DSS using EVM/ES in section 5. The end user of our DSS (i.e. the project manager) will be presented with control charts and corresponding control limits. If the control chart produces a signal (a control limit that is exceeded), the project manager is provided with an indication that the underlying activity durations do not conform with the pre-defined state of control. The project manager will then likely invest time and resources to drill down the WBS of the project to find the activities that cause this departure from the pre-defined state of control. If, at the activity level of the project, the departure from the pre-defined state of control can not be confirmed, the efforts spent by the project manager were in vain. The signal, produced by the DSS is then said to be false. If however, the departure can be confirmed at the activity level, the signal can be said to be true. In section 5, we will provide more details on how these true and false signals are used to quantify the performance of the decision making process.

With respect to the multivariate nature of EVM/ES observations, and improved DSS using EVM/ES should consider the following aspects:

- **Data overload:** The project manager using an EVM/ES system is provided with an abundance of data over the lifetime of a project. The choice of which of the performance metrics (SV, SV(t), SPI or SPI(t)) is likely to result in an as good as possible inference of activity level schedule performance, is a question that is often answered based on subjective arguments.

- **Redundancy:** Even in the case that the vector $\boldsymbol{x}$ is composed out of observations for a single EVM/ES performance metric along the lifetime of the project, $\boldsymbol{x}$ can still be subject to redundancy. Redundancy, or collinearity, is a problem that is found when different observed variables are influenced by a common factor. Redundancy should be addressed when inference is made from a multivariate variable [19]. If we suppose that SPI(t) is measured along the lifetime of the project, it is likely that some of the observations for SPI(t) are influenced by the same couple of activities, hence collinearity will exist between these variables.

- **Noise:** Multivariate systems can be subject to noise. It has been shown that SPI behaves unreliable towards the end of late projects, since for any execution of a project that experiences delays, SPI will equal 1 at the end of the project [2]. The use of SPI for duration forecasts in the project is therefore not recommended. One could argue that SPI is erroneous, and should therefore not be considered for schedule control. However, in early stages of the project, the information contained in SPI can be valuable for schedule control. In late projects, even the rate at which SPI converges towards 1 and the timing of the point at which the increase towards 1 takes place can provide the project manager with effective information, with respect to the underlying activity level performance. What is regarded by one as noise in the system, could be usefully employed by another. We therefore reason that a complex multivariate system such as EVM/ES could benefit from an automatic mechanism to identify and handle noise.

The aspects discussed here are common to many multivariate measurement systems. They have previously been addressed for the control of batch processes by MacGregor [20], to whom we refer for an illustration on how a principal component analysis overcomes these problems. Multivariate projection methods have been used to overcome the problems associated with the multivariate nature of the measured quality of a system since the first application of Hotelling's [21] multivariate t-test measure, in combination with PCA. Ever since, strong interest has been shown for this field of study, which resulted in related multivariate projection methods. The reader is referred to Bersimis et al. (2006) [22] for a recent overview and a comparison of these techniques. For the scope of this paper we will refrain from these more recent procedures. We believe that for this introduction on how a multivariate projection method can be used in a DSS using EVM/ES for top-down project schedule control, PCA is an appropriate choice. PCA is algorithmically easier, and provides some insight in how a new basis for comparison of project schedule performance is chosen, based on a reference set of project executions.

3. Principal component analysis of EVM/ES schedule performance metrics

In this section, we will first introduce the basics of a principal component analysis and how it can be performed. Second, we will discuss how the schedule control reference matrix should be composed.

3.1. Calculating principal components

In the very early years of the 20th century PCA was developed by Pearson [23]. Ever since, PCA has been a popular procedure to reduce the dimensionality of a variable space. A full coverage on the basics of linear algebra, fundamental to PCA, lies outside the scope of this paper and the reader is referred to the recent book of Jolliffe [24].

Instead, we will give a brief overview of the matrix calculus that is required for a PCA.

PCA assumes that the true rank of a matrix of observations X is less than the number of observations which are made. Consequently, it conjectures that the observations can be projected onto a new set of coordinate axes, thereby removing redundancy and noise from the system. The PCA decomposition method first calculates the principal components of the observation space, i.e. the directions that will make up the new coordinate axes. These principal components are often also named the latent variables, since they represent the underlying (unobservable) factors really influencing the system dynamics. Consider a $(n \times P)$ matrix X that is a collection of n measurements for a P-variate randomly distributed variable $\mathbf{x}$. The first principal component of $\mathbf{x}$ is defined as the vector of coefficients $\mathbf{p}_1$ for which the linear combination $t_1 = \mathbf{x}\mathbf{p}_1$ captures as much as possible of the variance contained in X, subject to $|\mathbf{p}_1| = 1$. The second principal component is then the vector of coefficients $\mathbf{p}_2$ for which the linear combination $t_2 = \mathbf{x}\mathbf{p}_2$ contains as much of the variance from X that is not captured within t_1. Additional principal components up to P are similarly defined.

In practice, a PCA is always performed using the computationally efficient singular value decomposition (SVD, [25]) of X ($X = \mathrm{ULA}^T$). A PCA decomposition of the matrix X can be written as:

$$X = \mathrm{TP}^T \tag{1}$$

Correspondingly, the standard deviation of the i^{th} principal component:

$$s_{\mathbf{t_i}} = \sqrt{\frac{l_i^2}{n-1}} \tag{2}$$

can be obtained from the singular values on the diagonal of L, since these are equal to the square roots of the eigenvalues ($\sqrt{l_i^2}, \forall i \in \{1, \ldots, P\}$) of $X^T X$.

The score-loading nomenclature is very common in PCA literature and is therefore adopted here. The matrix of loadings $\mathrm{P} = \begin{bmatrix} \mathbf{p}_1 & \mathbf{p}_2 & \dots & \mathbf{p}_P \end{bmatrix}$ can be seen as a $(1 \times P)$ collection of $(P \times 1)$ vectors of coefficients for the linear combination $t_i = \mathbf{x}\mathbf{p}_i$ that defines the i^{th} principal component. A related term "matrix of rotations" expresses the geometrical interpretation of the loadings, as they represent a new coordinate space onto which $\mathbf{x}$ is projected. The $(n \times P)$ matrix of scores T can then be interpreted as the values in the new coordinate space for the collection of n measurements of $\mathbf{x}$.

With respect to the ability of a PCA to remove redundancy and noise from the system, we need to discuss the relative importance of the different principal components, as expressed by their standard deviations. We will now explore this further, while discussing whether or not all principal components should be retained in the analysis.

If all of the P principal components are retained, for a further analysis of the data, an observation for the P-variate vector of observations $\mathbf{x}$ can be reconstructed from its $(1 \times P)$ vector of scores $\mathbf{t}$.

$$\mathbf{x} = \mathbf{t}\mathrm{P}^T = \sum_{i=1}^{P} t_i \mathbf{p}_i \tag{3}$$

However, since PCA is used to find a solution to the problems associated with the multivariate nature of project data, it will try to reduce the dimensionality of the problem in a structured manner, without losing valuable information. We assume that an integer $0 < k \leq P$ exists such that the last $P - k$ principal components do not represent valuable information for our system. By definition, each principal component will only explain a very small part of the original variation contained in X. For now we state that if only k principal components are retained, the original observation for the P-variate vector of observations $\mathbf{x}$ can be estimated as $\hat{\mathbf{x}}$ from its $(1 \times k)$ vector of scores $\mathbf{t}$.

$$\hat{\mathbf{x}} = \mathbf{t}\mathrm{P}^T = \sum_{i=1}^{k} t_i \mathbf{p}_i \tag{4}$$

For the collection of n observations for $\mathbf{x}$ in X, the matrix form is:

$$X = \sum_{i=1}^{k} \mathrm{T}_k \mathrm{P}_k^T + \mathrm{E}_k \tag{5}$$

where T_k is the $(n \times k)$ matrix of scores, P_k is the $(P \times k)$ matrix of loadings when only k principal components are retained and E_k is the $(n \times P)$ error matrix. E_k can be seen as the collection of error vectors $\mathbf{e}$, each corresponding to the vector $\mathbf{x}$ when only k principal components are retained:

$$\mathbf{e} = \mathbf{x} - \hat{\mathbf{x}} = \sum_{i=k+1}^{P} t_i \mathbf{p}_i \tag{6}$$

3.2. PCA model for a DSS using EVM/ES for project schedule control

When PCA is implemented on a simulated schedule control reference, it is essential that this matrix X is in an appropriate format. This aspect has been dealt with in numerous publications for the continuous monitoring of batch processes. An EVM/ES schedule control matrix can be considered to resemble the process matrix of a batch process that can be unfolded into a flat structure according to MacGregor and Kourti [26, 27]. Table 1 introduces the symbols and variables that are used in this section.

During the execution of a project, a functional EVM/ES system will require periodic measurements of the performance metrics displayed in table 1. In practice this might be done using software [16–18] and
as specified within a contractual agreement or scheduled at distinct time interval along the project lifetime. The schedule control reference, from which principal components will be calculated, are fictitious project executions that are produced by a Monte Carlo simulation on activity durations. Different fictitious project executions can lead to different numbers of performance variables recorded for each execution. To have an equal amount of observations for each execution, equally spaced over the lifespan of a project, it is not appropriate to use a time index increasing from the start to the end of the project. Instead, our simulation model will use the project percentage complete (PC_t) as a monotonically rising, scaled time-indicator ranging from 0% to 100%. Without loss of generality, we will assume that after each ΔPC percentage of the work performed, the project schedule performance metrics are recorded.

We present our methodology in a general form, and proceed with a vector of observations $\mathbf{x}_{\kappa,1:J}$ with length J reported at each review period $\kappa \in \{1, \ldots, K\}$, when $PC = \kappa \Delta PC$ percentage of the work in the project is completed. In the experimental results sections of

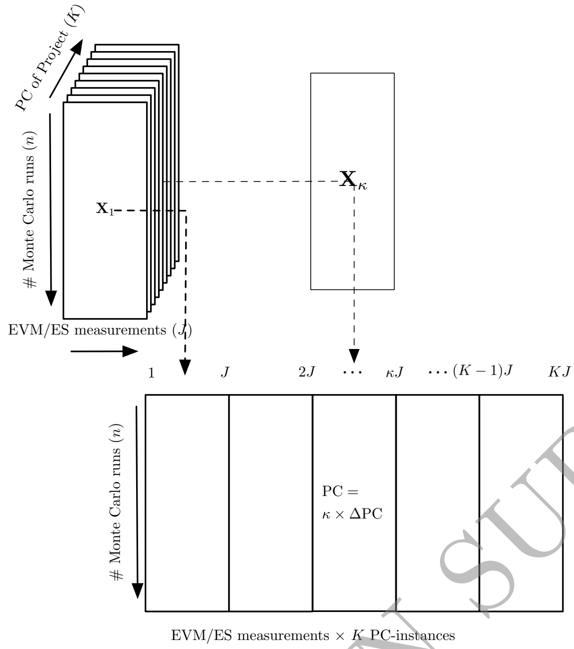

Figure 1: Unfolding of the three dimensional project data matrix ([28])

this paper, $\mathbf{x}_{\kappa,j}$ will be equal to $\left[\mathrm{SV}, \mathrm{SPI}, \mathrm{SV(t)}, \mathrm{SPI(t)}\right]$ for $j \in \{1, \ldots, 4\}$ as the periodically reviewed EVM/ES schedule control metrics. The outcome of a Monte Carlo simulation on the duration of the activities results in a set of fictitious project executions, for which the EVM/ES observations are structured as the three dimensional matrix presented at the top of figure 1. For each periodic review period $\kappa \in \{1, \ldots, K\}$, J EVM/ES schedule control metrics are observed for n Monte Carlo runs. In order to use equation 1 to perform a PCA decomposition, the three dimensional matrix consisting of K times $n \times J$ matrices has to be unfolded into a big $n \times KJ$ (with $P = KJ$) two dimensional matrix as presented in figure 1.

Since we will combine SV and SV(t), which are expressed in monetary and time units respectively, with dimensionless indices SPI and SPI(t), we need to scale and center each column of the matrix X before the PCA is performed. The SV and SV(t) values will be much larger than the SPI and SPI(t) values. This could result in relatively higher weights assigned to the SV and SV(t) observations in the loading vectors found by PCA, which might then lead to wrong interpretations. We will adopt the normalisation per column which is most common in literature, but other weighted PCA examples can also be found [22].

4. Multivariate schedule control in a DSS using EVM/ES

In this section, we will apply the PCA decomposition discussed in section 3 to produce two multivariate schedule control metrics (Hotelling's T^2 and SPE, section 4.1). In order to benefit from these metrics during the execution of a project

, so that they can lead to enhanced decision making

, we need to be able to apply them dynamically on project progress data. Section 4.2 introduces how missing observations are filled in to overcome the problems associated with the dynamic use of the T^2 and SPE metrics. These two newly proposed schedule metrics should improve the top-down schedule control process by generating correct warning signals. A signal is said to be produced when a tolerance limit for either T^2 or SPE is exceeded. Section 4.3 introduces how tolerance limits can be calculated prior to the execution of a project.

4.1. Two new schedule control metrics

Let us assume that we have a simulated schedule control reference in the appropriate matrix format X. This matrix contains observations for all J EVM/ES schedule performance metrics, for all K review periods and for all n fictitious project executions produced by a Monte Carlo simulation. The PCA decomposition of X results in a matrix of loadings P and a matrix of scores T. Let us assume that the project is now executed in real life, which results in a vector of observations $\boldsymbol{x}_{new}$. This P-variate vector of observations will now be referenced against the PCA model, producing two schedule control metrics (T^2, section 4.1.1 and SPE, section 4.1.2), in order to

enhance the decision making process in top-down schedule control.

4.1.1. Hotelling's T^2

Hotteling [21] proposed the T^2 measure as a multivariate extension of the t-statistic, frequently used in statistical hypothesis testing. If the population covariance Σ for the P-variate variable $\mathbf{x}$ is not known, it can be estimated from its sample covariance matrix S using n samples,

$$\mathrm{S} = \frac{1}{n-1}\sum_{i=1}^{n}(\mathbf{x}_i - \bar{\mathbf{x}})^T(\mathbf{x}_i - \bar{\mathbf{x}}) \tag{7}$$

where $\bar{\mathbf{x}}$ is used to represent the P-variate mean of $\mathbf{x}$. Hotellings T^2 statistic can then be used to measure the weighted multivariate distance of $\boldsymbol{x}_{new}$ from this mean $\bar{\mathbf{x}}$:

$$T^2 = (\mathbf{x}_{new} - \bar{\mathbf{x}})\mathrm{S}^{-1}(\mathbf{x}_{new} - \bar{\mathbf{x}})^T \tag{8}$$

It should be noted that, in order to calculate T^2 directly from $\mathbf{x}_{new}$, the inverse of the covariance matrix S needs to be calculated. As discussed earlier, collinearity and noise in the multivariate variable severely impede the accurate calculation of this inverse. Calculating T^2 from the scores however, resulting from a prior PCA decomposition, is much more computationally stable. Principal components are mutually independent and thus, the covariance matrix is reduced to a diagonal matrix, for which calculating the inverse is trivial. The scores for the new vector of observations $\mathbf{t}_{new}$ can be found by projecting the P-variate EVM/ES observation vector $\mathbf{x}_{new}$ onto the k-dimensional principle component space.

$$\mathbf{t}_{new} = \mathbf{x}_{new}\mathrm{P} \tag{9}$$

Hotellings T^2 can then be calculated as:

$$T_k^2 = \sum_{i=1}^{k} \left(\frac{(\mathbf{t}_{new})_i}{s_{\mathbf{t}_i}} \right)^2 \tag{10}$$

with $(\mathbf{t}_{new})_i$ the i^{th} element of the vector of scores $\mathbf{t}_{new}$ and $s^2_{\mathbf{t}_i}$ the estimated variance of the i^{th} principal component.

From equation 10, it is apparent that the last $P - k$ principal components should not be retained in the calculation. These explain very little of the variation in X and generally represent random noise or errors introduced by the measurement system. Calculating T^2 on the scores results in the summation of terms $((\mathbf{t}_{new})_i / s_{\mathbf{t}_i})^2$. Due to the small variances $s^2_{\mathbf{t}_i}$ explained by the last principal components, the slightest deviation (for principal components that have almost no effect on X) would lead to very large weighted distances. Therefore, retaining only the first k principal components is of importance to ensure that only the principal component that have the greatest influence on the EVM/ES schedule control vector $\mathbf{x}$ are expressed in this distance.

4.1.2. The squared prediction error SPE

Hotelling's T^2 statistic can be used to monitor the weighted distance of the vector of schedule performance observations as projected onto a reference defined by the principal components.

However, this metric will only detect whether or not the observed variation is greater than what was contained within the reference data matrix X. When this variation on the activity level is significantly different than the reference variation, the basis formed by the principal components might no longer be representative. In order to monitor whether the applied principal component analysis transformation is still representative for the new vector of observations $\boldsymbol{x}_{new}$, the squared prediction error (SPE) should be calculated. SPE represents the squared perpendicular distance of the new P-variate observation from the projection space defined by the principal components and can be calculated as:

$$SPE = \mathbf{e}\mathbf{e}^T = \sum_{i=1}^{P}(\mathbf{x}_{new,i} - \hat{\mathbf{x}}_{new,i})^2 \tag{11}$$

where $\mathbf{e}$ is defined for $\mathbf{x}_{new}$ using equation 6 if k principal component are retained.

4.2. Dynamic use of the T^2 and SPE schedule control metrics

We want to apply the proposed T^2 and SPE schedule measures dynamically during the project execution phase. The new vector of observations $\mathbf{x}_{new}$ needs to be projected onto the principal component space produced by the PCA decomposition of the schedule control reference X. Moreover, in order to calculate the scores $\mathbf{t}_{new}$ with equation 9, the $(1 \times P)$ vector $\mathbf{x}_{new}$ needs to be complete. However, at a review period $\kappa \in \{1, \ldots, K\}$ of the project in progress, only κJ of the P observations will be available. To come up with an estimate for the scores $\mathbf{t}_{new}$, we will need to deal with the problem of missing observations [29]. For notation purposes, we divide the vector of observations $\mathbf{x}$ into a $(1 \times \kappa J)$ vector $\mathbf{x}^*$ for which values are already recorded and a $(1 \times (K - \kappa)J)$ vector $\mathbf{x}^{\#}$ of missing measurements ($\mathbf{x} = \begin{bmatrix}\mathbf{x}^* & \mathbf{x}^{\#}\end{bmatrix}$).

The problem of producing an estimate for $\mathbf{t}_{new}$ when there are only observations available for $\mathbf{x}^*$ is described in the literature as dealing with missing data for on-line process monitoring. Different procedures have been proposed to deal with missing data in the literature and a comprehensive comparison is given by Nelson [30]. Conditional Mean Replacement (CMR) has been found to perform well by this study and is implemented here to produce an estimate $\hat{\boldsymbol{\tau}}$ for $\mathbf{t}_{new}$ during the dynamic project control process.

In CMR, the expected values for the conditional multivariate distribution are used to estimate $\mathbf{x}^{\#}$, given the present data and the most accurate estimate for the mean $\bar{\mathbf{x}}$ and covariance matrix S ($\hat{\mathbf{x}}^{\#} = \mathbb{E}(\mathbf{x}^{\#}|\mathbf{x}^{*}, \bar{\mathbf{x}}, \mathrm{S})$) . CMR allows us to estimate the score $\hat{\boldsymbol{\tau}}$ as:

$$\hat{\boldsymbol{\tau}} = \mathbf{x}^{*}\mathrm{S}_{22}^{-1}\mathrm{S}_{21}\mathrm{P}^{\#} + \mathbf{x}^{*}\mathrm{P}^{*} \tag{12}$$

with the covariance matrix S rewritten as:

$$\mathrm{S} = \begin{bmatrix} \mathrm{P}^{\#}\Theta\mathrm{P}^{\#T} & \mathrm{P}^{\#}\Theta\mathrm{P}^{*T} \\ \mathrm{P}^{*}\Theta\mathrm{P}^{\#T} & \mathrm{P}^{*}\Theta\mathrm{P}^{*T} \end{bmatrix} = \begin{bmatrix} \mathrm{S}_{11} & \mathrm{S}_{12} \\ \mathrm{S}_{21} & \mathrm{S}_{22} \end{bmatrix} \tag{13}$$

where $\Theta = \mathrm{T}^{T}\mathrm{T}/(n-1)$ represents a diagonal matrix with the variances explained by each principal component on its diagonal and the matrix of loadings is restructured as $\mathrm{P} = \begin{bmatrix} \mathrm{P}^{\#} \\ \mathrm{P}^{*} \end{bmatrix}$.

4.3. Tolerance limits for the T^2 and SPE schedule control metrics

During the execution of a project, the T^2 and SPE metrics, calculated at the top WBS level, should be interpreted to infer a state of activity level schedule control. The most straightforward interpretation follows when a warning signals arises at the top WBS level. If either the T^2 or the SPE metric exceeds a tolerance limit, a warning signal is said to be produced. This warning signal should then indicate that the activity level performance does not longer conform a pre-defined state of schedule control.

For independence with respect to distributional assumptions, we propose the use of the empirical cumulative distribution function (ecdf) for T^2 and SPE. In practice, this requires the calculation of the T^2 and SPE schedule control metric at each review period ($\forall 1 \leq \kappa \leq K$) and for each fictitious project execution in our schedule control reference X. Obviously, the estimate $\hat{\boldsymbol{\tau}}$, produced using CMR, now replaces the score vector $\mathbf{t}_{new}$ in calculating T^2 and SPE for all $\kappa < K$. The tolerance limits at a review period κ for T^2 and SPE can then be calculated in accordance with a tolerance level α as the α^{th} sample quantile. We refer to Hyndman and Fan [31] for more detail on the calculations of the sample quantiles.

5. Experimental test design

In this section the design of the experiments to characterise the performance of the T^2 and the SPE metrics for project control is outlined. Section 5.1 introduces the data generation process, to create a project benchmark set. In section 5.2 the Monte Carlo simulations are described in detail. Section 5.3 provides the measures to quantify the performance of the decision process using the T^2 and the SPE schedule control metrics.

5.1. Data generation

We test the performance of the T^2 and SPE schedule control metrics on a set of 900 projects generated by the project network generator RanGen [32], recommended by Herroelen [33] to protect the test ensemble from a possible bias in network structure. Projects from this set were extensively used in previous research on project control [3, 5, 10, 34].

The baseline durations are randomly assigned to the 30 activities in the projects. They are sampled from a uniform distribution between 8 and 56 days. The baseline duration (estimate) for an activity i will be denoted as $\hat{d}_i$. The fixed cost for each activity is sampled uniformly between €0 and €500 and the variable cost is sampled uniformly between €700 and €1500. The schedule, referred to as the baseline during the execution of the project, is the earliest start schedule obtained from a single forward pass of the critical path method.

5.2. Monte Carlo simulations

In order to quantify the use of the two newly proposed schedule control metrics, we ran an extensive set of two-phased Monte Carlo simulation experiments. Section 5.2.1 describes the dynamic project progress model with which EVM/ES measurements are produced for fictitious project executions. Section 5.2.2 recounts how Monte Carlo simulations serve a double purpose in this paper and section 5.2.3 describes the combined use of variation and risk input modelling for activity durations.

5.2.1. EVM/ES model

Fictitious project executions are simulated using P2 Engine [18] to generate EVM/ES data at $K = 19$ distinct PC intervals, with $\Delta\text{PC} = 5\%$, from 5% to 95%. At any review period $\kappa \in 1, .., K$ we chose $\mathbf{x}_{\kappa,1:J} = \left[\text{SV}, \text{SPI}, \text{SV(t)}, \text{SPI(t)}\right]$ for $j \in \{1, \ldots, 4\}$ to be the periodically reviewed EVM/ES schedule control metrics. A total of $P = 19 \times 4 = 76$ original variables is recorded for the matrix X.

Our model assumes that the earned value (EV) for a single activity follows a linear accrue, starting from its actual start up to its budget at completion (BAC) when it is finished [3]. Planned value (PV) follows this same linear accrue from the planned start up to the planned finish time of the activity. EV and PV are calculated in P2 Engine at the project level and are compared to produce the SV, SPI, SV(t) and SPI(t) schedule performance metrics at each review period κ.

The calculations for the PCA decomposition and the CMR procedure to produce the T^2 and SPE schedule control metrics, along with the analysis of the results presented in section 6, were implemented in the statistical programming language R [35].

5.2.2. Two-phased experiment

Monte Carlo simulations serve a dual purpose in our experiments. In the *first phase* the outcome of a large simulation (10,000 runs) is used to build a schedule control reference set for each of the 900 projects,

which is then used to perform the PCA decomposition on for our enhanced DSS using EVM/ES.

The preferred state of schedule control will then determine which fictitious project executions end up in the reference data matrix X. We will provide additional detail on how we define this state of schedule control in section 5.3.

In the *second phase* of the Monte Carlo experiment, project progress situations are simulated (now with 1,000 runs per project), where the activity level performance might not conform with the pre-defined state of schedule control.

This set is used to quantify the performance of the decision process that follows from our DSS using EVM/ES for project schedule control. The warning signals that are generated by the T^2 and SPE control charts should accurately indicate whether or not the pre-defined state of control can be confirmed at the activity level. We will provide details on how the performance of the decision process for top-down schedule control is quantified in section 5.3

5.2.3. Activity duration input modelling

The Monte Carlo simulations in this paper produce fictitious executions to generate project progress data. To that purpose, we need an appropriate model to accurately represent the uncertainty experienced at the activity level of the project. We opted for a combined input modelling where both risk and variation, often considered as separate sources of uncertainty in project management literature [36], are represented using separate probability distributions. We have chosen to implement probability functions from the family of generalised beta distributions, which have long been used in project management [3, 34, 37] due to their ability to accurately mimic the behaviour of random input processes driving the system [38], and their association to PERT-style three point estimates [39]. The probability density function for a generalised beta random variable $\mathbf{D}$ can be stated as:

$$f_{\mathbf{D}}(d|a,b,\theta_1,\theta_2) = \begin{cases} \frac{\Gamma(\theta_1+\theta_2)}{\Gamma(\theta_1)\Gamma(\theta_2)} \frac{(d-a)^{\theta_1-1}(b-d)^{\theta_2-1}}{(b-d)^{\theta_1+\theta_2-1}} & \text{if } a \leq d \leq b \\ 0 & \text{if } d < a \vee b < d \end{cases} \tag{14}$$

where $\Gamma()$ denotes the gamma function and θ_1 and θ_2 represent shape parameters. We propose the use of a parameter vector $\boldsymbol{\omega} = (a, b, m, \mu)$ with estimates for the minimum (a), the maximum (b), the mode (m) and the mean (μ) of the distribution, expressed as fractions of the baseline estimate duration $\hat{d}_i$ for activity $i \in \mathcal{N}$. In doing so, we can sample for all activities i, using only a limited number of distributions, represented by their parameter vector $\boldsymbol{\omega}$ (for which settings are presented in table 2). From $\boldsymbol{\omega}$ the shape parameters can be found using:

$$\begin{cases} \theta_1(\boldsymbol{\omega}) &= -\frac{(b+a-2m)(a-\mu)}{(m-\mu)(a-b)} \\ \theta_2(\boldsymbol{\omega}) &= \frac{(b+a-2m)(b-\mu)}{(m-\mu)(a-b)} \end{cases} \tag{15}$$

In each Monte Carlo simulation, *risk* is first modelled using the concept of linear association [5, 40] and a parameter vector $\boldsymbol{\omega}_R$ is assigned. Due to the absence of a unified methodology to test the impact of riskful events or dependencies between activities in the project management literature, Trietsch et al. [40] suggest the use of a positive random variable B to act as a bias term in simulations of project executions. B is most easily perceived as a consistent over- or underestimation of activity durations, or a project-wide effect of an uncertain event. In our experiments, we will assign a parameter vector $\boldsymbol{\omega}_R$, to represent this risk factor. Consequently, for each fictitious project execution in the simulation, a bias term is sampled from the generalised beta distribution with parameter vector $\boldsymbol{\omega}_R$. Subsequently, for all activities in the project, variation is added using an parameter vector $\boldsymbol{\omega}_V$, which can be chosen independently from $\boldsymbol{\omega}_R$. The duration d_i of an activity i in a fictitious project execution can then ultimately be considered as a sample from a generalised beta distribution (with parameter vector $\boldsymbol{\omega}_V$) multiplied with a biased baseline estimate $B\hat{d}_i$.

In the experiments, for which results will be shown in section 6, $\boldsymbol{\omega}_R$ and $\boldsymbol{\omega}_V$ are chosen separately from the set of parameter vectors presented in table 2. In the *general performance experiment*, the parameter vectors are chosen such that they represent realistic activity duration distributions for activities that either finish early ($\boldsymbol{\omega}_1$), on time ($\boldsymbol{\omega}_2$) or late ($\boldsymbol{\omega}_3$), on average. From the distribution for which activities end on time, we derived two distinct sets of parameter vectors ($\boldsymbol{\omega}_{2\mu_1}$ to $\boldsymbol{\omega}_{2\mu_5}$ and $\boldsymbol{\omega}_{2s_1}$ to $\boldsymbol{\omega}_{2s_6}$) to model changes in respectively the mean and the standard deviation. These can be used in the *sensitivity experiment* to test the robustness to over- or underestimation, when assumptions are made with respect to distributional characteristics of the activity durations.

5.3. Quantifying the performance of the decision making in a DSS using EVM/ES

In order to show the enhanced decision making that follows from our multivariate schedule control metrics in a DSS using EVM/ES for top-down schedule control, we will compare it to the schedule control process using statistical tolerance limits. This univariate procedure was presented by Colin and Vanhoucke (2014) [5] and uses statistical tolerance limits for SV, SPI, SV(t) and SPI(t) to steer the decision making in top-down schedule control using EVM/ES. The schedule control metrics were applied either directly (X chart) or the difference between two consecutive measurements was monitored (R chart). The X and R charts were shown to outperform the decision making based on rules-of-thumb that is omnipresent in EVM/ES schedule control in practice.

Moreover, we were able to obtain a fair comparison in our experiments between the newly proposed multivariate schedule control metrics and the univariate schedule control process using EVM/ES. Since, the same fictitious project executions were used to calculate tolerance limits for both the univariate and multivariate EVM/ES schedule control metrics. Subsequently, we compared the warning signals generated by all performance metrics with respect

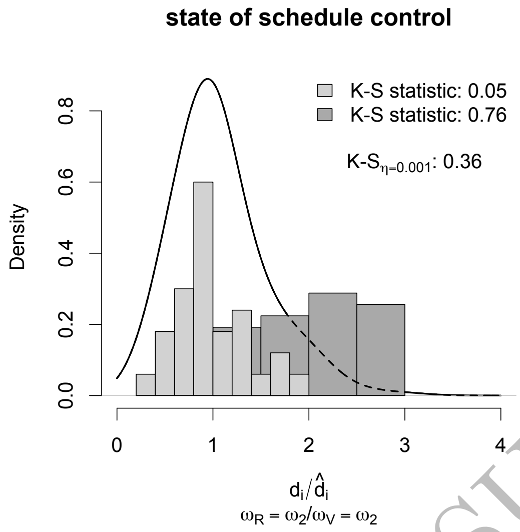

Figure 2: Activity level test for schedule control: Kolmogorov-Smirnov

to the underlying activity level of the project for additional fictitious project executions in the second phase,

in order to quantify the performance of the decision making process.

In order to demonstrate the enhanced decision process that follows from the newly proposed multivariate control metrics, we first needed to define a *state of schedule control*, as established at the activity level of the project. The performance of the decision process could then be quantified by whether or not this state of control can be correctly inferred from the observations made at the highest WBS level. We measured this performance using two descriptive quantities, the *detection performance* and the *probability of overreactions*. These can be combined into a single measure, known in the classification techniques literature as the *area under the curve*.

Defining a state of schedule control. The state of schedule control is primarily defined by the fictitious project executions that are included in the schedule control reference matrix X. This reference is then used to compare with future executions and to calculate tolerance limits. We will illustrate along the following lines how a set of fictitious project executions, generated in the first phase of the simulation model, can be chosen to constitute the project schedule control reference.

In the first phase of the Monte Carlo simulation, specific values are chosen for the parameter vectors $\boldsymbol{\omega}_R$ and $\boldsymbol{\omega}_V$, which results in a probability distribution for the simulated real activity durations of a project. Figure 2 shows the approximated density of the scaled activity durations $(d_i/\hat{d}_i)$ from a sample of 10,000 fictitious project executions, where $\boldsymbol{\omega}_R = \boldsymbol{\omega}_2$ and $\boldsymbol{\omega}_V = \boldsymbol{\omega}_2$. The histograms in lightgrey and darkgrey in figure 2 represent two fictitious project executions, from this sample of 10,000. Even though these fictitious executions where simulated from the same distribution, they are very dissimilar in terms of their appearance on the activity duration scale.

Figure 2 illustrates how it can be difficult to define a state of schedule control due to random variation. We will therefore employ a metric to describe how close a fictitious project execution lies to the empirical density, approximated from all 10,000 samples. We test for each execution whether it is likely to have been sampled from the empirical distribution function, which is based on $\boldsymbol{\omega}_R$ and $\boldsymbol{\omega}_V$. We do this using the two-sample Kolmogorov-Smirnov (K-S) statistic, a well-established non-parametric statistic to test the equality of two distributions [41]. Corresponding to a significance level of $\eta = 0.001$, a critical value (K-S$_\eta = 0.36$) for the K-S statistic can be found [42]. Thus, using the K-S statistic and its critical value, we will only include those fictitious project executions, in our schedule control reference, for which the sample of activity durations is likely to have been drawn from the distribution defined by $\boldsymbol{\omega}_R$ and $\boldsymbol{\omega}_V$ (K-S $<$ K-S$_\eta$).

Detection performance. In the second phase of the simulation experiment, we test whether the signals generated at the highest WBS level correspond to the underlying activity level schedule performance. In other words, for a project execution that does not conform with the defined state of schedule control (K-S $>$ K-S$_\eta$),
the project manager should be given a warning during the lifetime of the project, using the control charts that act as the user interface of the DSS.
The *detection performance* is thereby calculated for all runs in the second phase that do not comply to the state of schedule control, as the ratio of the number of those runs that generate a signal to the total number of those runs.

Probability of overreactions. In conjunction with the detection performance, a dual measure needs to be formulated. The *probability of overreactions* measures whether the

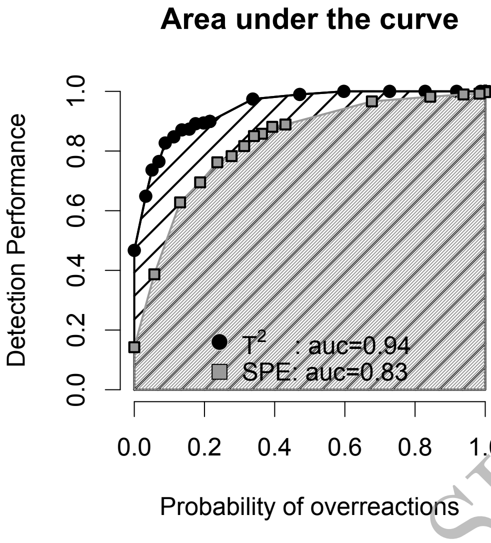

Figure 3: Illustration of the duality of the detection performance and the probability of overreactions. The area under the curve allows to combine both measures into one descriptive value

control charts provide a false warning signal to the project manager, who is the end user of the DSS

. If a warning signal is produced at the highest WBS level, we do not want the project manager's effort spent in drilling down the WBS to be in vain. The probability of overreactions can be calculated from all runs in the second phase of the simulation experiment for which the state of schedule control can still be assumed to be representative (K-S < K-S_η).

Area under the curve. Figure 3 shows an illustration of the probability of overreactions and the detection performance for the multivariate schedule control metrics proposed in this paper. The probability of overreactions is denoted along the x-axis and the detection performance along the y-axis. An ideal control metric should have an as high as possible detection performance, whilst keeping the probability of overreactions close to 0. The points presented on the graph represent different probability of overreaction/detection performance couples for varying values of the level α. This α represents the α^{th} quantile used for the statistical tolerance limits. In order to capture the dynamics of both the probability of overreactions and the detection performance into a single descriptive measure, we propose the use of the *area under the curve*, which should be maximised for

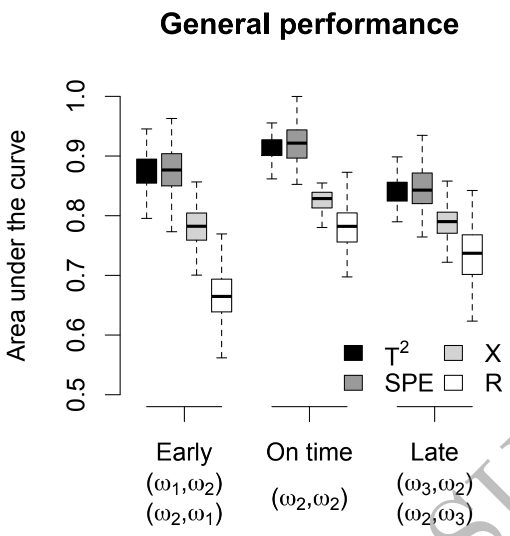

Figure 4: General performance comparison of the models for a DSS using EVM/ES for top-down schedule control

as good as possible decision making from the DSS using EVM/ES for top-down schedule control

. This measure is widely used in classification testing and machine learning [43] and can be obtained by Riemann integration of the curves produced by the probability of overreaction/detection performance couples, as displayed in figure 3.

6. Results and Discussion

In order to demonstrate the improved decision making that follows from our multivariate model for a DSS using EVM/ES, we conducted the following experiments.

Using the parameter vectors introduced in table 2 of section 5.2.3, we conducted two large simulation experiments. Using the parameter vectors outlined under *general performance experiment*, we will discuss the overall improvement of multivariate schedule control over the univariate use of EVM/ES metrics in section 6.1. While using these results, we explore how many principal components should be retained in the analysis for schedule control. The *sensitivity experiment* was conducted using the parameter vectors for the generalised beta family of distributions where either the mean ($\boldsymbol{\omega}_{2\mu_1}$ to $\boldsymbol{\omega}_{2\mu_5}$) or the standard deviation is varied ($\boldsymbol{\omega}_{2s_1}$ to $\boldsymbol{\omega}_{2s_6}$). In section 6.2, we present the effect of an under- or overestimation of the mean or standard deviation on the performance of the decision making process. In section 6.3, we will add discussion to the results that are presented here.

6.1. General performance experiment

In order to compare the general performance of the multivariate and univariate schedule control metrics described in this paper, a simulation experiment was carried out with scenarios where activities are either early, on time or late, on average. Figure 4 displays the assignments made to the couple ($\boldsymbol{\omega}_R$,$\boldsymbol{\omega}_V$) to obtain these scenarios, along its x-axis.

The boxplots of figure 4 present the recorded area under the curve for the multivariate T^2 and SPE metrics, and the univariate schedule control metrics X and R from Colin and Vanhoucke (2014) [5]. Overall, figure 4 shows that the multivariate metrics outperform the traditional univariate use of EVM/ES. The difference between the T^2 metric and the squared prediction error SPE is small, with the latter outperforming the former slightly. The difference between the multivariate and the univariate control metrics becomes smaller in the situations where activities are on time or late on average, but the multivariate approaches still show to be significantly better.

In conclusion, our multivariate approach is likely to lead to enhanced decision making in a DSS using EVM/ES for top-down schedule control.

Number of principal components to retain. The results of figure 4 are restructured to analyse the optimal number of principal components that needs to be retained for building an as accurate as possible PCA model. Figure 5 shows the recorded area under the curve for the T^2 and SPE metrics in function of the number of principal components retained in the model k. It shows that, in general, the T^2 metric benefits from a very low number of principal components (preferably only 1), while the SPE metric performs better when more principal components are included in the PCA model.

Analogously, we investigated which EVM/ES control metric (SV, SPI, SV(t), SPI(t)) should be used in the schedule control procedures X and R in order to deliver the best performance. We present these findings in figure 6, where a birds-eye perspective is given on six histograms

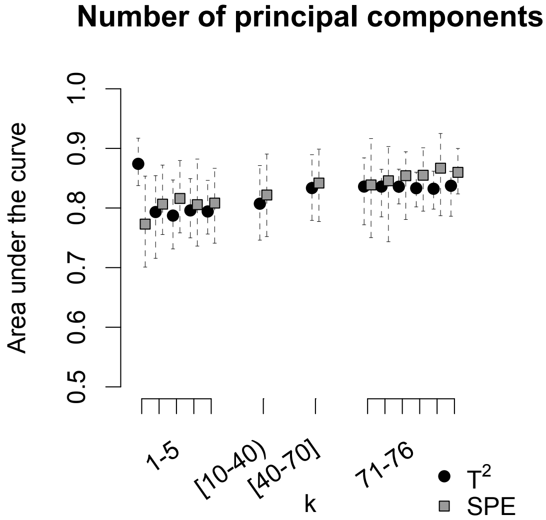

Figure 5: Optimal number of principal components for the multivariate schedule control metrics

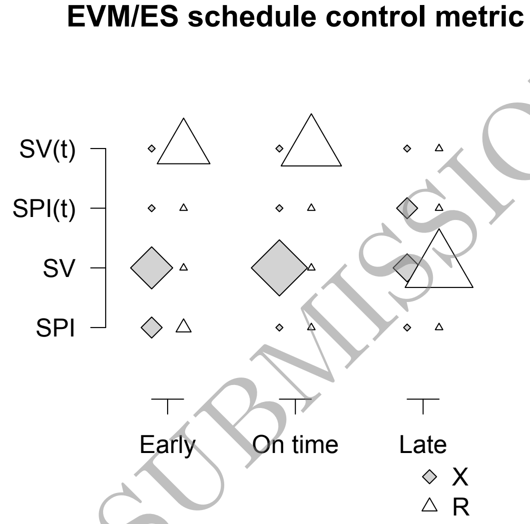

Figure 6: Optimal schedule control metrics

in which the relative size of the points represents how often a certain metric is found to be performing best. A single EVM/ES metric that outperforms all other could not be found. Therefore a combined use of all EVM/ES schedule control metrics is advised, where none is neglected during the execution of the project. This can been seen as a confirmation of the problems discussed in section 2, which inspired our multivariate model for a DSS using EVM/ES for top-down project schedule control.

6.2. Sensitivity experiment

The statistical tolerance limits for the schedule control metrics discussed in this paper are generated from a simulated set of fictitious project executions. These are generated using a Monte Carlo simulation, for which probability distributions serve as input to model the real activity durations. In a realistic project environment, it is not always possible to produce probability distributions that accurately reflect the stochastic nature of the input processes. Consequently, an under or over-estimation of one of the distributional characteristics can not always be avoided.

We model this situation using the parameter vectors $\boldsymbol{\omega}_{2\mu_1}$ to $\boldsymbol{\omega}_{2\mu_5}$ to test a deviation in

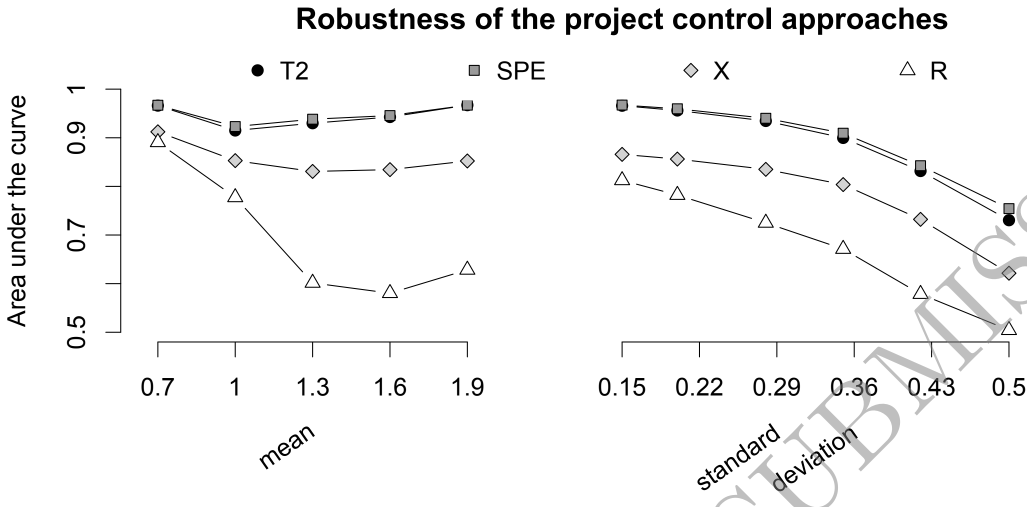

Figure 7: Influence of a change in the mean and the standard deviation on the project schedule control performance

the mean and $\boldsymbol{\omega}_{2s_1}$ to $\boldsymbol{\omega}_{2s_6}$ to test a deviation of the standard deviation. In the sensitivity experiment, the statistical tolerance limits are produced from a state of schedule control defined by $\boldsymbol{\omega}_R = \boldsymbol{\omega}_V = \boldsymbol{\omega}_2$. Additional fictitious project executions are generated with alternately one of the parameter vectors $\boldsymbol{\omega}_{2\mu_1} \ldots \boldsymbol{\omega}_{2\mu_5}$ or $\boldsymbol{\omega}_{2s_1} \ldots \boldsymbol{\omega}_{2s_6}$ assigned to either $\boldsymbol{\omega}_R$ or $\boldsymbol{\omega}_V$, in order to test the effect of a change in the mean or standard deviation.

Figure 7 shows the calculated area under the curve for the univariate and multivariate procedures that are used in a DSS for top-down schedule control, as described in this paper . On the left x-axis, the average activity duration for the additional runs, from which the detection performance and probability of overreactions are calculated, is presented. The standard deviation is presented on the right x-axis. The pre-defined state of schedule control ($\boldsymbol{\omega}_R = \boldsymbol{\omega}_V = \boldsymbol{\omega}_2$) has mean 1 and standard deviation 0.3.

From figure 7, we can conclude that the performance of the decision making process that follows from the R schedule control procedure decreases drastically for projects for which the average activity duration increases. The decision making processes based on T^2, the SPE and the X schedule control procedure seem to be much more robust to under- or overestimates of the mean activity duration.

With respect to the sensitivity to changes in the standard deviation of the underlying activity durations, all of the investigated schedule control procedures exhibit the same behaviour. When the standard deviation is larger than what was assumed for calculating the schedule control reference, the performance of the decision making process decreases significantly, as shown on the right of figure 7.

6.3. Discussion

In this section, we will add some discussion to the results that are presented in this section. Although we have shown how our multivariate model can lead to enhanced decision making for a top-down schedule control process, some aspects still need to be addressed.

The newly proposed schedule metrics T^2 and SPE outperform the current best-known choices for schedule control using EVM/ES with respect to the decision making by the project manager of whether or not he/she should invest time and effort in drilling down the project schedule. Despite its improved performance compared to the traditional univariate approach, the area under the curve is still mostly lower than 1, which means that some portion of the activity level performance is still obscured by the aggregated EVM/ES observations at the high WBS level of the project. This issue is intrinsically connected to a top-down schedule control process, but our research shows that considerable improvements can be made with the application of a multivariate model to the EVM/ES measurement system.

The practicality of the DSS for the end user is also significantly improved. Whereas all schedule control metrics (SV, SPI, SV(t), SPI(t)) should be monitored in a univariate model to ensure the best possible decision making, we show that the use of a single multivariate metric (T^2 or SPE) can lead to considerable improvements. The data overload from which the univariate EVM/ES models suffer is thereby reduced. When only a single multivariate metric should be monitored by the end user of the DSS, this is likely to improve the accuracy and the timeliness of the decision making. In addition, a single multivariate metric is also useful from a reporting and communications point of view. Moreover, due to the application of a PCA, redundancies and noise are removed from the inference process of underlying activity level schedule performance.

7. Conclusion

In this paper, we present a multivariate model for a DSS using EVM/ES for top-down project schedule control

. The need for a multivariate approach is dictated by the very nature of the EVM/ES measurement system. Observations for the EVM/ES variables are characterised by a data-overload, noise and redundancy. Based on the principle of batch process control from the chemometrics literature, we have initiated a multivariate

model to be used in a DSS using EVM/ES for top-down project schedule control

. In the resulting DSS, all the available EVM/ES observations are used in a structured manner, through analysis of their correlation structure. Two multivariate schedule control metrics (T^2 and SPE) are proposed,

that can be presented to the end user of the DSS on control charts.

We tested our multivariate model on a large simulation experiment that is in line with previous project control research. The performance of the T^2 and the SPE schedule control metrics was quantified and compared to the univariate use of EVM/ES observations

in a DSS for top-down schedule control

. The T^2 and the SPE metrics outperform the univariate EVM/ES approaches in terms of their ability to infer underlying activity level schedule performance more accurately,

which ultimately leads to enhanced decision making

. Moreover, the use of the multivariate control metrics leaves a project manager with only a single metric to monitor over the lifetime of a project. Traditional use of the EVM/ES metrics would require monitoring all the EVM/ES schedule control metrics, since their performance might vary from situation to situation. The SPE metric slightly outperforms the T^2 metric, but both these multivariate schedule control metrics are likely to improve not only the accuracy and the timeliness of the inference process of the underlying state of control at the activity level, but are also useful from a reporting and communications point of view.

Further work on the topic of multivariate metrics for top-down schedule control will expand on the results presented here. Other projection methods or transformations might improve the inference process even more,

which will consequently lead to more improvements in the decision making process

. PCA should currently be considered as the current best option for

a model to handle EVM/ES measurements in a DSS for top-down project schedule control.

References

[1] PMBOK, A Guide to the Project Management Body of Knowledge, Third Edition, Newtown Square, Pa.: Project Management Institute, Inc., 2004.

[2] W. Lipke, O. Zwikael, K. Henderson, F. Anbari, Prediction of project outcome: The application of statistical methods to earned value management and earned schedule performance indexes, International Journal of Project Management 27 (2009) 400–407.

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

[4] C. Fang, F. Marle, A simulation-based risk network model for decision support in project risk management, Decision Support Systems 52 (3) (2012) 635 – 644. doi:http://dx.doi.org/10.1016/j.dss.2011.10.021.
URL http://www.sciencedirect.com/science/article/pii/S0167923611002016

[5] J. Colin, M. Vanhoucke, Setting tolerance limits for statistical project control using earned value management, Omega The International Journal of Management Science 49 (2014) 107–122.

[6] H. Hunter, R. Fitzgerald, D. Barlow, Improved cost monitoring and control through the earned value management system, Acta Astronautica 93 (0) (2014) 497 – 500. doi:http://dx.doi.org/10.1016/j.actaastro.2012.09.010.
URL http://www.sciencedirect.com/science/article/pii/S0094576512003670

[7] H. L. Chen, Improving forecasting accuracy of project earned value metrics: Linear modeling approach, Journal of Management in Engineering 30 (2) (2014) 135–145.

[8] M. Plaza, O. Turetken, A model-based DSS for integrating the impact of learning in project control, Decision Support System 47 (2009) 488–499.

[9] H. Khamooshi, H. Golafshani, EDM: Earned duration management, a new approach to schedule performance management and measurement, International Journal of Project Management 32 (6) (2014) 1019–1041.

[10] R. Elshaer, Impact of sensitivity information on the prediction of project's duration using earned schedule method, International Journal of Project Management 31 (2013) 579–588.

[11] G. M. Marakas, Decision support systems in the 21st century, Vol. 2, Pearson Education, Prentice Hall, New Jersey, 2003.

[12] F. Anbari, Earned value project management method and extensions., Project Management Journal 34(4) (2003) 12–23.

[13] Q. Wang, N. Jiang, L. Gou, M. Che, R. Zhang, Practical experiences of cost/schedule measure through earned value management and statistical process control, Lecture Notes in Computer Science 3966 (2006) 348–354.

[14] S. S. Leu, Y. C. Lin, Project performance evaluation based on statistical process control techniques, Journal of Construction Engineering and Management 134 (2008) 813–819.

[15] R. Aliverdi, L. Moslemi Naeni, A. Salehipour, Monitoring project duration and cost in a construction project by applying statistical quality control charts, International Journal of Project Management 31 (3) (2013) 411–423.

[16] DecisionEdge, `http://www.decisionedge.com/` (2014).

[17] M. Vanhoucke, Dynamic Scheduling: Integrating Schedule Risk Analysis with Earned Value Management, The Measurable News 2 (2012) 11–13.

[18] P2 Engine, Visit the P2 Engine website at www.p2engine.com (Jan. 2013).

[19] G. W. Stewart, et al., Collinearity and least squares regression, Statistical Science 2 (1) (1987) 68–84.

[20] J. F. MacGregor, Using on-line process data to improve quality: Challenges for statisticians*, International Statistical Review 65 (3) (1997) 309–323.

[21] H. Hotelling, A generalized t test and measure of multivariate dispersion, Proc. Second Berkeley Symp. on Math. Statist. and Prob. (1951) 23–41.

[22] S. Bersimis, S. Psarakis, J. Panaretos, Multivariate statistical process control charts: an overview, Quality and Reliability Engineering International 23 (5) (2006) 517–543.

[23] K. Pearson, Liii. on lines and planes of closest fit to systems of points in space, The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science 2 (11) (1901) 559–572.

[24] I. Jolliffe, Principal component analysis, Wiley Online Library, 2005.

[25] E. Anderson, Z. Bai, C. Bischof, S. Blackford, J. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammerling, A. McKenney, et al., LAPACK Users' guide, Vol. 9, Siam, 1999.

[26] J. MacGregor, T. Kourti, Statistical process control of multivariate processes, Control Engineering Practice 3 (3) (1995) 403–414.

[27] T. Kourti, J. F. MacGregor, Process analysis, monitoring and diagnosis, using multivariate projection methods, Chemometrics and intelligent laboratory systems 28 (1) (1995) 3–21.

[28] D. S. Lee, J. M. Park, P. A. Vanrolleghem, Adaptive multiscale principal component analysis for on-line monitoring of a sequencing batch reactor, Journal of Biotechnology 116 (2) (2005) 195.

[29] P. Nomikos, J. F. MacGregor, Multivariate SPC charts for monitoring batch processes, Technometrics 37 (1) (1995) 41–59.

[30] P. R. Nelson, J. F. MacGregor, P. A. Taylor, The impact of missing measurements on pca and pls prediction and monitoring applications, Chemometrics and intelligent laboratory systems 80 (1) (2006) 1–12.

[31] R. J. Hyndman, Y. Fan, Sample quantiles in statistical packages, The American Statistician 50 (4) (1996) 361–365.

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

[33] W. Herroelen, Project scheduling — theory and practice, Production and Operations Management 14 (4) (2005) 413–432.

[34] M. Wauters, M. Vanhoucke, Support vector machine regression for project control forecasting, Automation in Construction 47 (2014) 92–106.

[35] R Core Team, R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria (2013).
URL http://www.R-project.org

[36] C. Loch, A. De Meyer, M. Pich, Managing the unknown: A new approach to managing high uncertainty and risk in project, John Wiley and Sons, Inc., New Jersey., 2006.

[37] S. AbouRizk, D. Halpin, J. Wilson, Fitting beta distributions based on sample data, Journal of Construction Engineering and Management 120 (1994) 288–305.

[38] M. E. Kuhl, E. K. Lada, N. M. Steiger, M. A. Wagner, J. R. Wilson, Introduction to modeling and generating probabilistic input processes for simulation, in: S. Henderson, B. Biller, M. Hsieh, J. Shortle, J. Tew, R. Barton (Eds.), Proceedings of the 2007 Winter Simulation Conference, New Jersey: Institute of Electrical and Electronics Engineers, 2007, pp. 63–76.

[39] W. J. McBride, C. W. McClelland, Pert and the beta distribution, Engineering Management, IEEE Transactions on (4) (1967) 166–169.

[40] D. Trietsch, L. Mazmanyan, L. Govergyan, K. R. Baker, Modeling activity times by the Parkinson distribution with a lognormal core: Theory and validation, European Journal of Operational Research 216 (2012) 386–396.

[41] A. N. Pettitt, M. A. Stephens, The kolmogorov-smirnov goodness-of-fit statistic with discrete and grouped data, Technometrics 19 (2) (1977) 205–210.

[42] G. Marsaglia, W. W. Tsang, J. Wang, Evaluating kolmogorov's distribution, Journal of Statistical Software 8 (18) (2003) 1–4.

[43] T. Hastie, R. Tibshirani, J. Friedman, T. Hastie, J. Friedman, R. Tibshirani, The elements of statistical learning, Vol. 2, Springer, 2009.

Table 1: Symbols and abbreviations used in PCA for EVM/ES schedule control

Earned Value/Earned Schedule	
Project EVM/ES key metrics	
BAC	Budget at completion for the project
PV_t	Planned value for the project at period t
EV_t	Earned value for the project at period t
ES_t	Earned schedule for the project at period t
PC_t	The percentage complete for a project at period t: $\mathrm{PC}_t = \mathrm{EV}_t/\mathrm{BAC}$
Project EVM/ES performance metrics	
$\mathrm{SV}_t = \mathrm{EV}_t - \mathrm{PV}_t$	Schedule variance for the project at period t
$\mathrm{SPI}_t = \mathrm{EV}_t/\mathrm{PV}_t$	Schedule performance index for the project at period t
$\mathrm{SV(t)}_t = \mathrm{ES}_t - t$	Schedule variance using earned schedule for the project at period t
$\mathrm{SPI(t)}_t = \mathrm{ES}_t/t$	Schedule performance index using earned schedule for the project at period t
with:	
$t = 1, \ldots, T$	Current time period (otherwise denoted as AT)
T	Total duration of the project
Schedule control observations	
$\boldsymbol{x}_{\kappa,1:J}$	Vector of observations at review period κ, including J performance metrics
κ	The index for the review period ($\kappa \in 1 : K$)
K	The total number of review periods for which each project execution in the reference data has observations: $K = 100/\Delta\mathrm{PC}$
ΔPC	The percentage complete increment between two review periods
J	The number of performance metrics included in the schedule analysis.
$P = J * K$	The total number of observations for each project execution.

Table 2: Parameter vectors for the generalised beta distribution

General performance experiment			**Sensitivity experiment**					
			Mean			*Standard deviation*		
$\boldsymbol{\omega}$	(a, b, m, μ)	σ	$\boldsymbol{\omega}$	(a, b, m, μ)	σ	$\boldsymbol{\omega}$	(a, b, m, μ)	σ
$\boldsymbol{\omega}_1$	$(0.1, 1.2, 0.7, 0.6)$	0.38						
			$\boldsymbol{\omega}_{2\mu_1}$	$(0.2, 4, 0.51, 0.7)$	0.3	$\boldsymbol{\omega}_{2s_1}$	$(0.2, 4, 0.40, 1)$	0.50
			$\boldsymbol{\omega}_{2\mu_2}$	$(0.2, 4, 0.90, 1.0)$	0.3	$\boldsymbol{\omega}_{2s_2}$	$(0.2, 4, 0.70, 1)$	0.42
$\boldsymbol{\omega}_2$	$(0.2, 4, 0.9, 1)$	0.30	$\boldsymbol{\omega}_{2\mu_3}$	$(0.2, 4, 1.22, 1.3)$	0.3	$\boldsymbol{\omega}_{2s_3}$	$(0.2, 4, 0.85, 1)$	0.35
			$\boldsymbol{\omega}_{2\mu_4}$	$(0.2, 4, 1.57, 1.6)$	0.3	$\boldsymbol{\omega}_{2s_4}$	$(0.2, 4, 0.92, 1)$	0.28
			$\boldsymbol{\omega}_{2\mu_5}$	$(0.2, 4, 1.89, 1.9)$	0.3	$\boldsymbol{\omega}_{2s_5}$	$(0.2, 4, 0.96, 1)$	0.20
						$\boldsymbol{\omega}_{2s_6}$	$(0.2, 4, 0.98, 1)$	0.15
$\boldsymbol{\omega}_3$	$(0.9, 4, 1.3, 1.4)$	0.38						