Process capability analysis for multivariate simple linear profiles in a multistage process

Saeed Adibfar\textsuperscript{1}, Rassoul Noorossana\textsuperscript{1,2*}, Orod Ahmadi\textsuperscript{3}

\textsuperscript{1}Industrial Engineering Department, Iran University of Science and Technology, Tehran, Iran
\textsuperscript{2}Information Systems and Operations Management Department, College of Business, University of Central Oklahoma, Edmond, OK, 73034, United States
\textsuperscript{3}Industrial Engineering Department, Faculty of Engineering, Kharazmi University, Tehran, Iran

adibfar.saeed@gmail.com; rassoul@iust.ac.ir; orod.ahmadi@khu.ac.ir

Abstract

Process capability indices (PCIs) are developed to assess process performance based on the specification limits (SLs) provided by customer. Sometimes the quality of a process or product is characterized by a regression relationship between a response variable and one or more independent variables referred to as "profile". On the other hand, modern production systems often involve multistage manufacturing processes, in which the output of one stage is the input of the next stage. This property is known as the cascade property. Due to this property, the capability in each stage is dependent on the capability of the preceding stages. This study provides an approach to assess PCIs in a multistage process when the quality characteristics of interest are represented by multivariate linear profiles. Process performance is specified based on profile intercept and slope parameters. In other word, in addition to PCIs of the response variable in each stage, the PCIs of profile parameters are also investigated. By using the SLs of the response variable and considering in-control profile, the SLs for intercept and slope can be obtained. Therefore, PCIs for profile parameters can be computed. The results indicate that the proposed method eliminates the effect of the cascade property for different autocorrelation values. Simulation results reveal satisfactory performance of the proposed method for a two-stage process.

**Keywords:** Process capability index, multivariate simple linear profile, multistage process, cascade property, specification limits

1-Introduction

Process capability indices (PCIs) have become popular and widely used tools in assessing process performance when a process is statistically in-control. PCIs quantify the relationship between the actual process performance and the specification limits (SLs) of the manufactured products. A process is called capable if the product meets customer expectations. The first process capability index introduced by Kane (1986), $C_p$, is defined as

$$C_p = \frac{USL - LSL}{6\sigma}$$

where $\sigma$ is the process standard deviation and USL and LSL are the upper and the lower SLs, respectively which reflect the customer's quality requirements.

*Corresponding author
ISSN: 1735-8272, Copyright c 2022 JISE. All rights reserved

According to the literature, multivariate PCIs may be classified into two groups. PCIs in the first group are defined without using correlation structure between variables. In this case, multivariate PCIs are obtained based on univariate PCIs. An example of these types of PCIs is the multivariate PCI presented by Hubele, Montgomery and Chih (1991). PCIs in the second group are defined based on the assumption that the output quality can be modeled using $p$ correlated quality characteristics. The literature indicates that most of the studies on multivariate PCIs belong to the second group. In the literature, three approaches are applied to compute multivariate PCIs that do take into account the correlation structure between quality characteristics (de-Felipe and Benedito, 2017):

- Multivariate PCIs based on principal component analysis
- Multivariate PCIs based on the relation between tolerance and process regions
- Multivariate PCIs based on the inverse function of the cumulative distribution function

Although in most process capability analysis studies, the quality characteristic of interest is modeled using a continuous or discrete random variable, sometimes the quality characteristics of interest can be modeled by linear relationships between response variables and one or more independent variables. This relationship that is often called ‘profile’ can be represented by a simple linear, a multiple linear, a multivariate linear or a polynomial regression or even by a more complicated relationship such as nonlinear regression. According to Woodall et al. (2004) and others, Phase I and Phase II are two phases for constructing control charts to monitor a process. In Phase I, the main goal is to assess the stability of the process, recognizing and eliminating assignable causes of variation and to estimate the in-control values of the process parameters. The aim of the Phase II is quick detection of any changes in the process parameters. Noorossana, Saghaei and Amiri (2011) addressed the fundamental concepts, methods, and issues related to statistical profile monitoring.

Process capability assessment of linear profiles has been partially studied in recent years. Hosseinfard and Abbasi (2012a) developed a PCI for linear profiles using the proportion of nonconforming items. In another study, Hosseinfard and Abbasi (2012b) investigated and compared five methods to estimate non-normal PCIs for linear profiles. Keshteli et al. (2014) explained a functional approach for measuring PCI for simple linear profiles. Pakzad, Razavi and Sadeghpour Gildeh (2021) proposed a functional approach for a simple linear profile based on fuzzy set theory for the situations in which the specification limits and target values of the response variable are not precisely specified. Pakzad and Basiri (2022) introduced a new functional incapability index for dealing with asymmetric tolerances for simple linear profile. In the study of Mehri et al. (2021), two robust PCIs for multiple linear profiles are proposed. In their study, the process capability is estimated using the M-estimator and the Fast-$\tau$-estimator. For more discussion on this issue, see (Ebadi and Shahriari, 2013; F.-K. Wang, 2014; F. Wang, 2014; and Wang and Tamirat, 2015).

In the area of PCI for multivariate profiles, few studies have been conducted. Ebadi and Amiri (2012) proposed three new methods to measure process capability when process output could be modeled by multivariate simple linear profiles (MVSLP). Wang (2016) presented a new process yield index to evaluate the process yield for multivariate linear profiles in manufacturing processes. Also, Wang and Tamirat (2016) presented two indices to measure the process capability for multivariate linear profiles with one-sided SLs under mutually independent normality. Additionally, they proposed two indices to measure the process capability for multivariate linear profiles with one-sided SLs under multivariate normality assumption. Guevara G and Alejandra López (2022) proposed a two-phase methodology based on the concept of depth to measure the capability of processes characterized by the functional relationship of multivariate nonlinear profile data, treated as multivariate functional observations.

In recent years, studies on other types of profiles have been of interest. See references (Wang and Guo, 2014; Guevara, Vargas and Castagliola, 2016; Rezaye Abbasi Charkhi, Aminnayeri and Amiri, 2016; Mohammad Pour Larimi, Nemati Keshteli and Safaei, 2018; Alevizakos, Koukouvinos and Castagliola, 2019; and Alevizakos and Koukouvinos, 2022) for more details.
In the existing studies, PCIs are generally computed based on response values. However, use of estimated values of profile parameters to measure the capability of process has rarely been studied. Karimi Ghartemanl, Noorossana and Niaki (2016), Wu (2016) and, Chiang, Lio and Tsai (2017) introduced PCIs to measure the process capability for simple linear profiles based on profile intercept and slope. Despite these few studies, we believe that an important issue has been overlooked in the proposed PCIs based on profile intercept and slope. This issue is the lack of using the in-control profiles to obtain accurate SLs for the parameters. It is well known that PCIs are based on predefined SLs. Considering the available literature, it seems that in all studies the SLs for intercept and slope are determined based on the profile SLs, which are not necessarily in-control. Since one of the main assumptions of capability analysis is the stability of the process, by considering profile SLs as well as the in-control profile, accurate SLs for parameters could be obtained.

On the other hand, in practice, manufacturing operations are often involved with multistage processes. In a multistage process, the output of each stage is affected by two main factors: the activities at current stage and the performance of the previous stage(s). This dependence that is referred to as the cascade property is of great importance when a multistage process is monitored. To solve this issue, some approaches such as cause selecting chart (CSC), regression adjusted charts, and state-space models were developed over time.

To the best of the authors’ knowledge, no attempt has been devoted in the literature to establish process capability for MVSLP in multistage processes. Therefore, the main purpose of this study is to propose an approach to assess the process capability for multivariate linear profile in a multistage process. Besides, a new method to compute PCIs for profile parameters is utilized. Note that even when the PCIs related to response values indicate the process is capable; analyzing the PCIs of profile parameters is still recommended.

This paper is organized as follows. Multistage process and profile modeling in multistage processes along with the proposed method for evaluating process capability are presented in sections 2 and 3, respectively. A brief explanation of the multivariate PCI used in this study is provided in section 4. The PCIs of profile parameters is introduced in section 5. A simulation study to evaluate the performance of the proposed method is presented in section 6 and finally, in the last section conclusions and some recommendations for future research are provided.

2-Multistage process

In real world manufacturing systems, many processes consist of several dependent stages. This implies that the output of one stage is the input of its subsequent stage and a change in a quality characteristic may affect some or all output variables in successive stages. This property is called cascade property and is the main feature of multistage processes. As an example, the quality characteristic of interest in a piston machining process is the piston diameter that is measured at different heights from the bottom of the pistons in each of the four series operations involved (Fong and Lawless, 1998). In this multistage process, the authors considered the relationship between the diameter and the height of pistons, as a profile, and analyzed the profile after each stage.

Copious studies about the applications of multistage processes have been done. In recent years, researchers have developed various control charts in multistage processes. However, capability analysis in multistage processes has not been studied as much. Zhang (1990) introduced two kinds of PCIs for multistage processes. The first PCI was total PCI that computes the process capability when the quality characteristic in the present stage is affected by quality characteristics of previous stages. The second one was the specific PCI that indicates the capability of a stage when the effects of precedent stages are excluded. Linn, Au and Tsung (2002) addressed how to prioritize the process variation reduction to enhance the overall process capability in multistage processes. Based on Taguchi loss function, Chen et al. (2012) presented a method to calculate PCI for complex product machining process as a multistage process. Nikzad, Amiri and Abbasi (2017) estimated the process capability of the second stage of two-stage process while the effect of cascade property is removed by using residuals analysis. In another study by Nikzad, Amiri and Amirkhani (2018), the effects of measurement errors on the specific and total PCIs in the second and third stages of a three-stage process are statistically analyzed.

Due to the cascade property, using usual statistical quality control methods in a multistage process may lead to inaccurate results. One way to overcome this problem is the cause-selecting chart (CSC), proposed by Zhang (1984). The advantage of this method is that once a signal is given, it is easy to
determine the stage associated with the signal. Therefore, it is more practical and beneficial for analyzing multistage processes by considering the cascade property. This idea is discussed in section 3.

While most of the studies in the area of multistage processes deal with univariate or multivariate quality characteristic, in some situations, profiles are streamed in the stages of a multistage process. Ghahyazi, Niaki and Soleimani (2014) were the first researchers who considered the quality characteristic in a multistage process as a profile. They proposed an approach to monitor simple linear profile in Phase II in the presence of cascade property. Khedmati and Niaki (2017) addressed the problem of monitoring general linear profiles in multistage processes in Phase I. Khedmati and Niaki (2016) also proposed an approach for monitoring simple linear profiles in multistage processes in Phase II. Bahrami, Niaki and Khedmati (2021) introduced a method to monitor MVSLP in a multistage process in Phase II.

3-Modeling

To model a MVSLP in a multistage process, it is assumed that \( m \) samples of size \( n \) are collected for \( p \) response variables at each of \( k \) stages of a multistage process from historical data. At each stage of the process, for sample \( j \), there are \( n \) fixed values for the explanatory variable, and, the observations \( (x_{ij1}, y_{ij1}, x_{ij2}, ..., y_{ijp}) \), \( i = 1, 2, ..., n, \ j = 1, 2, ..., m \) and \( s = 1, 2, ..., k \) are available. It is also assumed that the explanatory variable is fixed from sample to sample for all stages. Consequently \( x_{ij} = x_i \) for all values of \( j \) and \( s \). The multivariate profile model in a multistage process considering the cascade property can be written as:

\[
Y_{j1} = XA_{j1} + E_{j1} \\
Y_{js} = Y_{(s-1)} \Phi + XA_{js} + E_{js} \quad s > 1
\]

Where \( Y_{j1} \) refers to the response variables, \( A_{j1} \) indicates the parameters, and \( E_{j1} \) refers to the error terms in the first stage, respectively. Also, \( X \) is the matrix of explanatory variables.

For stage \( s \), we can equivalently state:

\[
\begin{bmatrix}
Y_{j1s} \\
Y_{j2s} \\
\vdots \\
Y_{nj2s}
\end{bmatrix}
= \begin{bmatrix}
Y_{1j1(s-1)} & Y_{1j2(s-1)} & \cdots & Y_{1jp(s-1)} \\
Y_{2j1(s-1)} & Y_{2j2(s-1)} & \cdots & Y_{2jp(s-1)} \\
\vdots & \vdots & \ddots & \vdots \\
Y_{nj1(s-1)} & Y_{nj2(s-1)} & \cdots & Y_{njp(s-1)}
\end{bmatrix}
\begin{bmatrix}
\Psi_{11} & \Psi_{12} & \cdots & \Psi_{1p} \\
\Psi_{21} & \Psi_{22} & \cdots & \Psi_{2p} \\
\vdots & \vdots & \ddots & \vdots \\
\Psi_{p1} & \Psi_{p2} & \cdots & \Psi_{pp}
\end{bmatrix}
\begin{bmatrix}
1 \\
x_{11} \\
x_{12} \\
\vdots \\
x_{1n}
\end{bmatrix}
+ \begin{bmatrix}
a_{01js} & a_{02js} & \cdots & a_{0pjs} \\
a_{11js} & a_{12js} & \cdots & a_{1pjs} \\
\vdots & \vdots & \ddots & \vdots \\
a_{nj1s} & a_{nj2s} & \cdots & a_{njps}
\end{bmatrix}
\begin{bmatrix}
\xi_{1j1s} & \xi_{1j2s} & \cdots & \xi_{1jps} \\
\xi_{2j1s} & \xi_{2j2s} & \cdots & \xi_{2jps} \\
\vdots & \vdots & \ddots & \vdots \\
\xi_{nj1s} & \xi_{nj2s} & \cdots & \xi_{njps}
\end{bmatrix}
\]

Where \( Y_{js} \) is a \( n \times p \) matrix of response variables for the \( j^{th} \) sample in \( s^{th} \) stage. \( X = [1 \ x] \) is a \( n \times 2 \) matrix of explanatory variables, in which \( 1 = (1, 1, \ldots, 1)^T \), and \( x = (x_{11}, x_{12}, \ldots, x_{1n})^T \). \( A_{js} = (a_{0js}, a_{1js})^T \) is a \( 2 \times p \) matrix of known parameters, in which \( a_{0js} = (a_{01js}, a_{02js}, \ldots, a_{0pjs})^T \), and \( a_{1js} = (a_{11js}, a_{12js}, \ldots, a_{1pjs})^T \). The autocorrelation values are given by matrix \( \Phi \). To further explain matrix \( \Phi \), we consider a two-variate profile in a two-stage process. For the \( j^{th} \) random sample, quality characteristics in the first stage are defined by \( y_{j11} \) and \( y_{j12} \) (\( s = 1 \)), and quality characteristics in the second stage are defined by \( y_{j12} \) and \( y_{j22} \) (\( s = 2 \)), respectively. Based on the nature of a process, \( y_{j12} \)
can be correlated with only $y_{j11}$ or both $y_{j11}$ and $y_{j21}$. Similarly, $y_{j22}$ can be correlated with only $y_{j21}$ or both $y_{j11}$ and $y_{j21}$. So, a general form of matrix $\Phi$ is considered in equation (3). Also $E_{js}$ is a $n \times p$ matrix of error terms. It is assumed that in each stage, the vector of error terms follows multivariate normal distribution with mean vector zero and known covariance matrix $\Sigma$, which can be shown by

$$
\Sigma = \begin{bmatrix}
\sigma_{11} & \sigma_{12} & \cdots & \sigma_{1p} \\
\sigma_{21} & \sigma_{22} & \cdots & \sigma_{2p} \\
\vdots & \vdots & \ddots & \vdots \\
\sigma_{p1} & \sigma_{p2} & \cdots & \sigma_{pp}
\end{bmatrix}
$$

(4)

Where $\sigma_{hi}$ indicates the covariance between $h^{th}$ and $i^{th}$ ($h = 1, 2, \ldots, p$, $i = 1, 2, \ldots, p$) error terms at each observation. Figure 1 presents a graphical display of the proposed multistage model.

![A graphical representation of a multistage process](Fig 1)

After specifying the profile model in a multistage process, SLs are defined. Generally, SLs can be considered as fixed values or a function of explanatory variables. For each variable and stage, it is assumed that the SLs associated with the response variables are linear functions of the explanatory variable as shown in equation (5)

$$
USL_{ihs} = a_{0hs} + a_{1hs}X_i \\
LSL_{ihs} = a_{0hs} + a_{1hs}X_i
$$

(5)

Where $USL_{ihs}$ and $LSL_{ihs}$ are the upper and the lower SLs for the $i^{th}$ level of $h^{th}$ response variable in $s^{th}$ stage. Also, $a_{0hs}$, $a_{1hs}$, $a_{0hs}'$, and $a_{1hs}'$ are the intercepts and slopes for $USL_{ihs}$ and $LSL_{ihs}$, respectively. Note that the SLs are not necessarily parallel to each other as well as to the profile line. However, in this article it is assumed that the SLs are parallel.

Due to the cascade effect, using common PCIs to assess the capability of intermediate stages ($s > 1$) may lead to misleading results. To deal with this issue, PCI for the residuals is considered. Residual analysis is the main idea of CSC, as one of the most popular approaches in multistage studies. Under this condition, residuals are not affected by previous stages. Thus, the PCIs for the residuals indicate the specific process capability of the process in each stage. The residual for the $i^{th}$ level of $h^{th}$ response variable is computed as

$$
e_{ijhs} = Y_{ijhs} - \hat{Y}_{ijhs}
$$

(6)

Where $\hat{Y}_{ijhs}$ is the fitted value for $Y_{ijhs}$. The fitted value $\hat{Y}_{ijhs}$ can be obtained using equations (7)

$$
\hat{Y}_{ijhs} = Y_{ijh(s-1)} + XA_{ijhs}
$$

(7)
The variance of the residuals is calculated by

$$\sigma^2_{eihs} = \frac{\sum_{j=1}^{m}(Y_{ijhs} - \hat{Y}_{ijhs})^2}{m - 2} = \frac{\sum_{j=1}^{m}e_{ijhs}^2}{m - 2}$$  \hspace{1cm} (8)

To assess the PCI for residuals, SLs of residuals have to be obtained. Nikzad, Amiri and Abbasi (2017) proposed a method to calculate the SLs for residuals using process yield. According to F.-K. Wang (2014), process yield has been recognized as a common criterion for measuring process performance. It measures the performance of process by computing the percentage of conforming items based on the SLs of process. Under the assumptions that the mean of residuals is equal to zero, which is considered as the target value, then the process yield of residuals is 0.9973. Thus, the SLs of residuals are acquired as follows.

$$\text{Process Yield} = P\{LSL_e \leq \varepsilon \leq USL_e\} = 0.9973$$  \hspace{1cm} (9)

$$P\left\{\frac{LSL_e - \mu_e}{\sigma_e} \leq z \leq \frac{USL_e - \mu_e}{\sigma_e}\right\} = 0.9973 \hspace{0.5cm}, \hspace{0.5cm} \mu_e = 0$$  \hspace{1cm} (10)

$$USL_e = \sigma_e \phi^{-1}(0.99865) \hspace{0.5cm}, \hspace{0.5cm} LSL_e = \sigma_e \phi^{-1}(0.00135)$$  \hspace{1cm} (11)

Where $\mu_e$ is the residuals mean, $\sigma_e$ is the residuals standard deviation, $USL_e$ and $LSL_e$ are the upper and the lower SLs of the residuals, respectively, and $\phi^{-1}(.)$ is the inverse cumulative distribution function of standard normal distribution.

4-Multivariate PCI

Many authors have introduced multivariate PCIs under different assumptions over past few years. Among the introduced approaches to assess multivariate PCIs, the ratio of tolerance region to process region has been of interest to some researchers. Chan, Cheng and Spiring (1991), Taam, Subbaiah and Liddy (1993), Shahriari, Hubele and Lawrence (1995), Grau (2007), Pan and Lee (2010), Niavarani, Noorossana and Abbasi (2012), Wang et al. (2013), Ciupke (2015), Pan, Li and Shih (2015), Abbasi Ganji and Sadeghpour Gildeh (2016, 2017), Abbasi Ganji (2019), and Govinda Khadse and Kailas Khadse (2020) are some of the authors who used the relation between tolerance and process regions to compute the preferred PCIs.

In this study, we consider the multivariate PCI referred to as $NMC_{PM}$ which was introduced by Niavarani, Noorossana and Abbasi (2012). This process capability index is a modified version of $MC_{PM}$ index proposed by Taam, Subbaiah and Liddy (1993) and it measures the ratio of modified tolerance region to a scaled 99.73% process region. Process region and modified tolerance region in two dimensions is shown in figure 2.

![Fig 2. Modified tolerance region vs. tolerance region](image-url)
Based on figure 2, Niavarani, Noorossana and Abbasi (2012) declared that if the modified tolerance region is ellipse, then the area of the original tolerance region is underestimated by the modified tolerance region; and consequently \( MC_{PM} \) would be underestimated. This concept can be expanded to higher dimensions as well. To eliminate the error in the estimation of the tolerance region, a new \( MC_{PM} \) referred to as \( NMC_{PM} \) is proposed as follows.

\[
NMC_{PM} = \frac{NC_P}{D}
\]

Where

\[
NC_P = \frac{Vol\left(\text{original tolerance region}\right)}{Vol\left(\text{estimated 99.73\% process region}\right)} = \frac{\prod_{i=1}^{p}[USL_h - LSL_h]}{|S|^{0.5}(\pi R)^{p/2}[\Gamma(\frac{P}{2}+1)]^{-1}}
\]

and

\[
D = [1 + \frac{m}{m+1}(\bar{Y} - T_0)'S^{-1}(\bar{Y} - T_0)]^{1/2}
\]

In above equations, \( Vol(.) \) is the volume of the region, \( p \) is the number of response variables, \( S \) contains the unbiased sample variance-covariance of the observations, \( R \) is the 99.73\% quantile of a \( \chi^2 \) distribution with \( p \) degrees of freedom, \( \Gamma \) denotes gamma function, \(||\) denotes the determinant, \( m \) is the number of observations, and \( T_0 \) denotes the \( p \)-vector target values for the \( p \) response variables. The quantity \( NC_P \) focuses on variation. If \( NC_P \) is larger than 1, it indicates that the process variation is lower than the acceptance variation criteria. The quantity \( \frac{1}{D} \) measures the closeness between the process mean and the target and a larger \( \frac{1}{D} \) indicates that the mean is closer to target. By determining the SLs of residuals, considering the profile model in multistage processes, and employing \( NMC_{PM} \) as the multivariate PCI, the PCI for each stage can be assessed.

5-PCIs for parameters

Along with the PCIs for each stage that is calculated based on the response variable, the performance of the stages can be inspected through the PCIs for the parameters. This helps to detect the parameter which contributes to low performance of the process. To assess the PCIs for profile parameters, it is required to determine the SLs for the intercept and slope. Pakzad (2021) provided a new method to measure PCI for a SLP based on its parameters. She considered profile SLs as well as the in-control profile to obtain accurate SLs for parameters. To assess the in-control profile, control chart limits for monitoring each parameter was considered. Her method is based on Kim, Mahmoud and Woodall (2003) study, which uses coded \( X \)-values to make the intercept estimator and the slope estimator of each profile independent.

Kim, Mahmoud and Woodall (2003) introduced a method for Phase I profile monitoring that monitors the intercept and slope individually by using a separate control chart. In this method, coded \( X \)-values are used which make the intercept estimator and the slope estimator of each profile independent. The transformed form of the model is obtained by equation (15).

\[
y_{ij} = b_{0j} + b_{1j}x_i^j + \varepsilon_{ij} \quad i = 1,2,...,n, \quad j = 1,2,...,m.
\]

164
Where $b_{0j} = a_{0j} + a_{1j} \bar{x}$, $b_{1j} = a_{1j}$ and $x_i' = x_i - \bar{x}$. In this situation the least-square estimators of coefficients are calculated by $b_{0j} = \bar{y}_j$ and $b_{1j} = a_{1j} = \frac{s_{xy(j)}}{s_{xx}}$. Also, $\bar{y}_j = \frac{\sum_{i=1}^{n} x_i y_{ij}}{n}$, $s_{xy(j)} = \sum_{i=1}^{n} (x_i - \bar{x})y_{ij}$, and $s_{xx} = \sum_{i=1}^{n} (x_i - \bar{x})^2$.

It is well known that when process is in-control, $b_{0j}$ and $b_{1j}$ are independent and follow normal distributions as $b_{0j} \sim N(B_{0j}, \frac{a_2^2}{n})$ and $b_{1j} \sim N(B_{1j}, \frac{a_2^2}{s_{xx}})$. A separate Shewhart control chart for monitoring intercept and slope are given in equations (16) to (19).

$$LCL_{b_0} = \bar{b}_0 - t_{m(n-2), \frac{a_2^2}{2}} \sqrt{\frac{(m-1)MSE}{mn}}$$

$$UCL_{b_0} = \bar{b}_0 + t_{m(n-2), \frac{a_2^2}{2}} \sqrt{\frac{(m-1)MSE}{mn}}$$

$$LCL_{b_1} = \bar{b}_1 - t_{m(n-2), \frac{a_2^2}{2}} \sqrt{\frac{(m-1)MSE}{mS_{XX}}}$$

$$UCL_{b_1} = \bar{b}_1 + t_{m(n-2), \frac{a_2^2}{2}} \sqrt{\frac{(m-1)MSE}{mS_{XX}}}$$

Where $\bar{b}_0 = \frac{\sum_{j=1}^{m} b_{0j}}{m}$, $\bar{b}_1 = \frac{\sum_{j=1}^{m} b_{1j}}{m}$, $MSE = \frac{\sum_{i=1}^{m} MSE_j}{m}$ and $t_{m(n-2), \frac{a_2^2}{2}}$ is a $100(1 - \frac{a_2^2}{2})$ percentile of $t$ distribution with $(m(n-2))$ degrees of freedom. Note that $\alpha_2 = n\sqrt{(1 - \alpha_1)}$ is the marginal probability of signal for each control chart and $\alpha_1 = \frac{3}{\sqrt{1 - \alpha}}$ specifies the overall probability of false alarm by each chart.

Pakzad (2021) assumed that the SLs for response variable for each level of the explanatory variable ($i = 1, 2, ..., n$) are linear functions of the explanatory variable as was stated in equation (9). According to separate control chart method, the transformed model for the SLs of $h^{th}$ profile in $s^{th}$ stage can be written as

$$USL_{ih} = b_{0hs} + b_{1hs}'x_i'$$

$$LSL_{ih} = b_{0hs} + b_{1hs}'x_i'$$

(20)

Where $b_{0hs}'$, $b_{1hs}'$, $b_{0hs}$, and $b_{1hs}$ are the intercepts and slopes for $USL_{ih}$ and $LSL_{ih}$, respectively. Note that the SLs are not necessarily parallel to each other as well as to the profile line. However, in this study, it is assumed that the SLs are parallel, so $b_{1hs}' = b_{1hs} = b$. A process is called “capable” if the response variable falls within the profile SLs. Hence, equation (21) may be stated as

$$b_{0hs}' + b_{1hs}'x_i' \leq b_{0hs} + b_{1hs}x_i' \leq b_{0hs}' + b_{1hs}'x_i'$$

(21)

According to Pakzad (2021), SLs for the intercept and slope can be calculated using equations (22) and (23).

$$b_{0hs}' + (b - b_{1hs})x_i' \leq b_{0hs} \leq b_{0hs}' + (b - b_{1hs})x_i'$$

(22)
\[
\begin{cases}
  b + \frac{(b_{0hs} - b_{0hs})}{x_i'} \leq b_{1hs} \leq b + \frac{(b_{0hs} - b_{0hs})}{x_i'}, & x_i' > 0 \\
  b + \frac{(b_{0hs} - b_{0hs})}{x_i'} \leq b_{1hs} \leq b + \frac{(b_{0hs} - b_{0hs})}{x_i'}, & x_i' < 0
\end{cases}
\]

(23)

It must be noted that although all profiles in equation (21) are within the SLs of the response variable, they are not necessarily in-control. Thus, all intercepts and slopes in equations (22) and (23) are not necessarily in-control either. To determine correct SLs for profile parameters, (Pakzad, 2021) considered both conforming and statistically in-control profiles in equation (21). As a result, the SLs for the intercept and slope parameters are given by equations (24) and (25).

\[
b_0^* + (b - LCL_{b_1})x_L' \leq b_0 \leq b_0^* + (b - UCL_{b_1})x_L'
\]

(24)

\[
\text{Min } \{\text{conforming and in-control slopes}\} \leq b_1 \leq \text{Max } \{\text{conforming and in-control slopes}\}
\]

(25)

Where \(X_i'\) is the minimum value of \(X_i's\), \(X_U'\) is the maximum value of \(X_i's\), and \(b + \frac{(b_0^* - LCL_{b_0})}{x_U'}, b + \frac{(b_0^* - UCL_{b_0})}{x_L'}\) and \(b + \frac{(b_0^* - LCL_{b_0})}{x_L'}\) are all conforming and in-control slopes.

Once the SLs for intercept and slope parameters are determined using equations (24) and (25), a univariate PCI for each parameter should be applied. Among univariate PCIs, \(C_{pmk}\) is considered in this study. This index provides indications of both process variability and proximity to the target. The index \(C_{pmk}\) for \(b_{0hs}\) and \(b_{1hs}\) can be written as in equations (26) and (27).

\[
C_{pmk_{b_{0hs}}} = \min \left\{ \frac{USL_{b_{0hs}} - \mu_{b_{0hs}}}{3\sqrt{\sigma_{b_{0hs}}^2 + (\mu_{b_{0hs}} - T_{b_{0hs}})^2}}, \frac{\mu_{b_{0hs}} - LSL_{b_{0hs}}}{3\sqrt{\sigma_{b_{0hs}}^2 + (\mu_{b_{0hs}} - T_{b_{0hs}})^2}} \right\}
\]

(26)

\[
C_{pmk_{b_{1hs}}} = \min \left\{ \frac{USL_{b_{1hs}} - \mu_{b_{1hs}}}{3\sqrt{\sigma_{b_{1hs}}^2 + (\mu_{b_{1hs}} - T_{b_{1hs}})^2}}, \frac{\mu_{b_{1hs}} - LSL_{b_{1hs}}}{3\sqrt{\sigma_{b_{1hs}}^2 + (\mu_{b_{1hs}} - T_{b_{1hs}})^2}} \right\}
\]

(27)

Where \(\mu_{b_{0hs}}\) and \(\sigma_{b_{0hs}}^2\) are the mean and variance of the sample mean of intercept of \(h^{th}\) profile in \(s^{th}\) stage, which follows \(N(B_0, \frac{\sigma_{b_{0hs}}^2}{mX_s})\). Similarly, \(\mu_{b_{1hs}}\) and \(\sigma_{b_{1hs}}^2\) are the mean and variance of the sample mean of slope of \(h^{th}\) profile in the \(s^{th}\) stage, which follows \(N(B_1, \frac{\sigma_{b_{1hs}}^2}{mS_X})\). The specifications \(USL_{b_{0hs}}, LSL_{b_{0hs}}, USL_{b_{1hs}}, \) and \(LSL_{b_{1hs}}\) are the upper and lower SLs of intercept and slope, respectively. Also, \(T_{b_{0hs}}\) and \(T_{b_{1hs}}\) are assumed to be the target value of the intercept and slope, respectively.

Using the SLs of response variable and considering the in-control profile, the SLs for intercept and slope can be obtained. Therefore, \(C_{pmk}\) index for profile parameters can be computed. It is worth mentioning that both indices \(C_{pmk_{b_0}}\) and \(C_{pmk_{b_1}}\) are used simultaneously and the process is deemed “incapable” if at least one of the indices indicates a low process performance.
6-Simulation study

In this section, a simulation study in MATLAB environment is carried out to investigate the performance of the proposed method. The MVSLP is considered which was stated in the study of Bahrami, Niaki and Khedmati (2021). The underlying model in the first and the second stage of a two-stage process, respectively, is defined as

\[
\begin{bmatrix}
    y_{ij11} \\
    y_{ij21}
\end{bmatrix} = \begin{bmatrix} 1 & x_i \end{bmatrix} \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} + \begin{bmatrix} \varepsilon_{ij11} \\ \varepsilon_{ij21} \end{bmatrix}
\]

(34)

\[
\begin{bmatrix}
    y_{ij12} \\
    y_{ij22}
\end{bmatrix} = \begin{bmatrix} \varphi_{11} & 0 \\ 0 & \varphi_{22} \end{bmatrix} \begin{bmatrix} 1 \\ x_i \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} + \begin{bmatrix} \varepsilon_{ij12} \\ \varepsilon_{ij22} \end{bmatrix}
\]

(35)

Where \( n = 4 \), \( \varepsilon_{ijp1} \sim \text{MVN} \left( \begin{bmatrix} 0 \\ 1 \\ 0.5 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0.5 & 0 \\ 1 & 0 & 0.5 & 0 \\ 0.5 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \right) \), and \( \varepsilon_{ijp2} \sim \text{MVN} \left( \begin{bmatrix} 0 \\ 1 \\ 0.5 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0.5 & 0 \\ 1 & 0 & 0.5 & 0 \\ 0.5 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \right) \). The explanatory variable with four fixed \( X_i \)-values of 2, 4, 6, and 8 is used in the simulation study. In the proposed method, by coding \( X_i \)-values, the transformed model is obtained as

\[
\begin{bmatrix}
    y_{ij11} \\
    y_{ij21}
\end{bmatrix} = \begin{bmatrix} 1 & x_i' \end{bmatrix} \begin{bmatrix} 13 \\ 9 \\ 2 \\ 1 \end{bmatrix} + \begin{bmatrix} \varepsilon_{ij11} \\ \varepsilon_{ij21} \end{bmatrix}
\]

(36)

\[
\begin{bmatrix}
    y_{ij12} \\
    y_{ij22}
\end{bmatrix} = \begin{bmatrix} \varphi_{11} & 0 \\ 0 & \varphi_{22} \end{bmatrix} \begin{bmatrix} 1 \\ x_i' \end{bmatrix} \begin{bmatrix} 7 \\ 16 \\ 1 \\ 2 \end{bmatrix} + \begin{bmatrix} \varepsilon_{ij12} \\ \varepsilon_{ij22} \end{bmatrix}
\]

(37)

Where \( X_i' \)-values are -3, -1, 1 and 3.

The regression lines associated with the SLs in each stage for the transformed model are provided in Table 1.

<table>
<thead>
<tr>
<th>Table 1. SLs of response variable in each stage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stage</td>
</tr>
<tr>
<td>-------</td>
</tr>
<tr>
<td>Stage 1</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Stage 2</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

Now we investigate the PCIs associated with the response variables and parameters. As mentioned earlier, it is better to evaluate the PCIs for profile parameters. It must be noted again that the specified PCI for stage 2 is calculated based on the residuals. Thus, the performance of stage 1 is related to the parameters of the profiles in the first stage, while total performance is affected by the parameters of profiles in all stages.
The effect of different values of sample size on the capability of each stage using 10,000 simulation replications for both weak and strong autocorrelation coefficients (\( \varphi_{11} = \varphi_{22} = 0.1, 0.9 \)) is presented in table 2. It must be noted that in the following tables, \( b_{0-hs} \) and \( b_{1-hs} \) \((h = 1, 2 \text{ and } s = 1, 2)\) refer to intercept and slope of the \( h^{th} \) profile in stage \( s \), respectively.

| Table 2. Capability of each stage and parameter under different values of \( \varphi \) and \( m \) |
|-----------------|-----------------|-----------------|-----------------|-----------------|
| \( \varphi = 0.9 \) | \( \text{Stage 1} \) | \( \text{Stage 2} \) | \( \text{Total} \) | \( \text{Stage 1} \) | \( \text{Stage 2} \) | \( \text{Total} \) | \( \text{Stage 1} \) | \( \text{Stage 2} \) | \( \text{Total} \) |
| \( \text{Stage 1} \) | \( \text{Stage 2} \) | \( \text{Total} \) | \( \text{Stage 1} \) | \( \text{Stage 2} \) | \( \text{Total} \) | \( \text{Stage 1} \) | \( \text{Stage 2} \) | \( \text{Total} \) |
| \( \varphi = 0.9 \) | \( \text{Stage 1} \) | \( \text{Stage 2} \) | \( \text{Total} \) | \( \text{Stage 1} \) | \( \text{Stage 2} \) | \( \text{Total} \) | \( \text{Stage 1} \) | \( \text{Stage 2} \) | \( \text{Total} \) |
| \( \text{Stage 1} \) | \( \text{Stage 2} \) | \( \text{Total} \) | \( \text{Stage 1} \) | \( \text{Stage 2} \) | \( \text{Total} \) | \( \text{Stage 1} \) | \( \text{Stage 2} \) | \( \text{Total} \) |
| \( \varphi = 0.9 \) | \( \text{Stage 1} \) | \( \text{Stage 2} \) | \( \text{Total} \) | \( \text{Stage 1} \) | \( \text{Stage 2} \) | \( \text{Total} \) | \( \text{Stage 1} \) | \( \text{Stage 2} \) | \( \text{Total} \) |
| \( \varphi = 0.9 \) | \( \text{Stage 1} \) | \( \text{Stage 2} \) | \( \text{Total} \) | \( \text{Stage 1} \) | \( \text{Stage 2} \) | \( \text{Total} \) | \( \text{Stage 1} \) | \( \text{Stage 2} \) | \( \text{Total} \) |

From table 2, different values of the autocorrelation coefficient do not affect the process performance in stage 1 and stage 2 (when the cascade property is removed). However, the total index is strongly correlated to this coefficient. Higher values of autocorrelation coefficients result in higher values of the total \( \text{NMC}_{PM} \). The performance of parameters in stage 1 is not correlated to the autocorrelation coefficient either. As can be seen in table 2, the values of \( \text{C}_{pmk} \) for \( b_{0-h1} \) and \( b_{1-h1} \) \((h = 1, 2)\) are not affected by different values of \( \varphi \); while the capability of parameters in stage 2 are highly affected by the value of \( \varphi \). On the other hand, it is clear from table 2 that different values of sample size do not have a noticeable effect on the capability values.

The other issue which should be investigated is the effect of variance of error terms on the capability values. It should be noted that the index \( \text{NMC}_{PM} \) considers the correlation between variables for computing process capability. Besides, in this example, we deal with four variances of error terms which are relating to the four mentioned profiles. \( \sigma_{11}^2 \) and \( \sigma_{21}^2 \) refer to the profiles in the first stage and \( \sigma_{12}^2 \) and \( \sigma_{22}^2 \) refer to the profiles in the second stage, respectively. To investigate the effect of variances of error terms along with the correlation coefficient, there will be many different cases. In table 3, we present the effect of different variances of the error terms relating to the profiles in stage 1 (\( \sigma_{11}^2 \) and \( \sigma_{21}^2 \) and
the correlation coefficient ($\rho$) on capability values in each stage and related parameters, while $\phi = 0.9$ and the sample size equals 25.

<table>
<thead>
<tr>
<th>$\sigma^2_{11}$</th>
<th>Capability</th>
<th>$\sigma^2_{11} = 0.7$</th>
<th>$\sigma^2_{11} = 1$</th>
<th>$\sigma^2_{11} = 1.3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.7</td>
<td>$NMC_{\text{PM}}$</td>
<td>Stage1</td>
<td>2.0389</td>
<td>2.3282</td>
</tr>
<tr>
<td>0.7</td>
<td>$NMC_{\text{PM}}$</td>
<td>Stage2</td>
<td>1.0375</td>
<td>1.1307</td>
</tr>
<tr>
<td>0.7</td>
<td>$NMC_{\text{PM}}$</td>
<td>Total</td>
<td>1.2857</td>
<td>1.4820</td>
</tr>
<tr>
<td>0.7</td>
<td>$C_{pmk-b_{0-11}}$</td>
<td>3.9751</td>
<td>3.9751</td>
<td>3.9751</td>
</tr>
<tr>
<td>0.7</td>
<td>$C_{pmk-b_{1-11}}$</td>
<td>0.9120</td>
<td>0.9120</td>
<td>0.9120</td>
</tr>
<tr>
<td>0.7</td>
<td>$C_{pmk-b_{0-21}}$</td>
<td>3.9523</td>
<td>3.9523</td>
<td>3.9523</td>
</tr>
<tr>
<td>0.7</td>
<td>$C_{pmk-b_{1-21}}$</td>
<td>0.8942</td>
<td>0.8942</td>
<td>0.8942</td>
</tr>
<tr>
<td>0.7</td>
<td>$C_{pmk-b_{0-12}}$</td>
<td>3.7260</td>
<td>3.7260</td>
<td>3.7260</td>
</tr>
<tr>
<td>0.7</td>
<td>$C_{pmk-b_{1-12}}$</td>
<td>0.7181</td>
<td>0.7181</td>
<td>0.7181</td>
</tr>
<tr>
<td>0.7</td>
<td>$C_{pmk-b_{0-22}}$</td>
<td>3.7190</td>
<td>3.7190</td>
<td>3.7190</td>
</tr>
<tr>
<td>0.7</td>
<td>$C_{pmk-b_{1-22}}$</td>
<td>0.7278</td>
<td>0.7278</td>
<td>0.7278</td>
</tr>
<tr>
<td>1.3</td>
<td>$NMC_{\text{PM}}$</td>
<td>Stage1</td>
<td>1.4157</td>
<td>1.6331</td>
</tr>
<tr>
<td>1.3</td>
<td>$NMC_{\text{PM}}$</td>
<td>Stage2</td>
<td>1.0359</td>
<td>1.1007</td>
</tr>
<tr>
<td>1.3</td>
<td>$NMC_{\text{PM}}$</td>
<td>Total</td>
<td>1.1196</td>
<td>1.2930</td>
</tr>
<tr>
<td>1.3</td>
<td>$C_{pmk-b_{1-11}}$</td>
<td>0.9120</td>
<td>0.9120</td>
<td>0.9120</td>
</tr>
<tr>
<td>1.3</td>
<td>$C_{pmk-b_{0-21}}$</td>
<td>3.5661</td>
<td>3.5661</td>
<td>3.5661</td>
</tr>
<tr>
<td>1.3</td>
<td>$C_{pmk-b_{1-21}}$</td>
<td>0.7452</td>
<td>0.7452</td>
<td>0.7452</td>
</tr>
<tr>
<td>1.3</td>
<td>$C_{pmk-b_{0-12}}$</td>
<td>3.7260</td>
<td>3.7260</td>
<td>3.7260</td>
</tr>
<tr>
<td>1.3</td>
<td>$C_{pmk-b_{1-12}}$</td>
<td>0.7181</td>
<td>0.7181</td>
<td>0.7181</td>
</tr>
<tr>
<td>1.3</td>
<td>$C_{pmk-b_{1-22}}$</td>
<td>0.6768</td>
<td>0.6768</td>
<td>0.6768</td>
</tr>
</tbody>
</table>

It is inferred from Table 3 that variances of error terms and the correlation coefficient have a remarkable effect on the capability values. Generally, as the correlation coefficient increases, the capability of each stage and the total capability also increase. However, it does not affect the capability of the parameters when $\sigma^2_{11}$ and $\sigma^2_{21}$ are fixed. On the other hand, different values of $\sigma^2_{11}$ and $\sigma^2_{21}$ results
in different values of capability for all stages and parameters. As we find in table 3, when $\sigma^2_{11} = \sigma^2_{21} = 0.7$ we achieve better capabilities in comparison to the case when $\sigma^2_{11}$ and $\sigma^2_{21}$ are equal to $1.3$. Similarly, simultaneous changes of all the variance terms can be explained too. For example, when $\sigma^2_{hi} = 0.7 (h = 1, 2$ and $s = 1, 2)$, the capability values of Stage1, Stage2, total, $b_{0-11}$, $b_{1-11}$, $b_{0-21}$, $b_{1-21}$, $b_{0-12}$, $b_{1-12}$, $b_{0-22}$, and $b_{1-22}$ are obtained as $2.3264$, $1.1260$, $2.3224$, $3.9751$, $0.9120$, $3.9523$, $0.8942$, $3.9657$, $0.8975$, $3.9579$ and $0.9157$, respectively which shows better performance. As expected, the lower values of the variance terms result in better performance of the process. The effect of different values of variances of the error terms relating to the profiles in Stage 2 ($\sigma^2_{12}$ and $\sigma^2_{22}$) on capability values in each stage and parameters can be analyzed in the same way.

7-Conclusion

In this study, an approach was presented to assess process capability in a multistage process when quality outputs are characterized by a MVSLP. Moreover, a method was developed to specify the performance of a profile based on its parameters. The capability of an in-control process was evaluated by new independent PCIs for profile intercept and slope. The SLs of profile parameters were obtained based on SLs of the response variable by considering the in-control profiles. The results of a two-stage process showed that total capability is strongly correlated to the autocorrelation coefficient, while different values of this coefficient do not affect the process performance in Stage 1 and Stage 2. The other important result was the remarkable effect of variance of error terms on the capability values. Generally, the lower values of the variance terms result in better performance of the process. This study focused on evaluating the performance of a MVSLP process based on profile parameters. The proposed approach can be extended for more complex profile models such as polynomial and nonlinear. Also, future studies may include calculation of PCIs in the presence of contamination.

References


*International Journal of Advanced Manufacturing Technology*, 64(5–8), pp. 857–865. doi: 

multivariate measurements’, *Statistica Sinica*, pp. 151–164.


Quality Management & Business Excellence*, 30(5-6), pp.525-536.


profiles’, *Communications in Statistics-Simulation and Computation*, 41(10), pp. 1761–1784. doi: 


https://doi.org/10.1080/00224065.1986.11978984.

analysis of processes monitored by a simple linear regression profile’, *Quality and Reliability 


a MaxEWMA control chart’, *Computers & Industrial Engineering*, 98, pp. 125–143. doi:


