Evaluation of process capability in multivariate simple linear profiles

Mohsen Ebadi*, Amirhossein Amiri

* Department of Industrial Engineering, K.N. Toosi University of Technology, Tehran, P.O. Box 15875-4416, Iran
Department of Industrial Engineering, Faculty of Engineering, Shahed University, Tehran, P.O. Box 18151-159, Iran

Received 7 January 2012; revised 29 May 2012; accepted 27 June 2012

Abstract In some situations, the quality of a process or product is characterized by a linear regression model between two or more variables which is called a linear regression profile. Moreover, in some cases, several correlated quality characteristics can be modeled as a set of linear functions of one explanatory variable which is typically referred to as multivariate simple linear profiles structure. On the other hand, process capability index is an important concept in statistical process control and measures the ability of the process to provide products that meet certain specifications. Little work, however, is done to evaluate the capability of a process with profile quality characteristic. This paper proposes three new methods for measuring process capability in multivariate simple linear profiles. Performance of the proposed methods is evaluated through simulation studies. In addition, the applicability of the proposed methods is illustrated using a real case of calibration application.

© 2012 Sharif University of Technology. Production and hosting by Elsevier B.V. All rights reserved.

1. Introduction

In many applications of Statistical Process Control (SPC), use of a single or several distinct quality characteristics is sufficient to monitor the quality of the process or product. However, sometimes the quality of the process or product is characterized by a relationship between response variable and one or more independent variables, which is referred to as profile. This relationship may be represented by a linear regression model or a more complicated model, such as nonlinear model. Kang and Albin [1] introduced the concept of linear profiles for the first time and proposed two methods for Phase II monitoring of simple linear profiles. In the first method, they applied a bivariate $T^2$ control chart to monitor the parameters of a simple linear profile. The second method uses an Exponentially Weighted Moving Average (EWMA) chart and an R-chart to monitor the regression residuals mean and the standard deviation, respectively. Kang and Albin [1] also investigated a usual application of simple linear profiles in the calibration industries. Kim et al. [2] proposed three separate univariate control charts for monitoring the intercept, and the slope of a profile model and the error standard deviation independently. Stover and Brill [3], Mahmoud and Woodall [4], Woodall et al. [5], Wang and Tsung [6], Zou et al. [7], Gupta et al. [8], Zhang et al. [9], Saghaei et al. [10], and Mahmoud et al. [11] investigated monitoring of simple linear profiles in Phases I and II.

Sometimes we should apply more complicated models than simple linear regression model to describe a profile. Kazemzadeh et al. [12,13] considered and developed some methods for monitoring polynomial profiles in Phases I and II. Mahmoud [14] developed a reduction parameter technique and extended some of the simple linear regression profile methods for the analysis of multiple linear regression profiles in Phase I. Furthermore, in some applications, the quality of the process or a product can be characterized by a multivariate regression model. In this case, there are some dependent quality characteristics as response variables, which are modeled as functions of one or more explanatory variables. Noorossana et al. [15] investigated multivariate simple linear regression profiles and developed some control charts to monitor such profiles in Phase II. Noorossana et al. [16] proposed four methods including likelihood ratio, Wilk’s lambda, $T^2$, and principal components to monitor multivariate multiple linear regression profiles in Phase I. The performance of these
methods is compared through simulation studies in terms of probability of a signal. Eyvazian et al. [17] developed four methods for Phase II monitoring of multivariate multiple linear regression profiles, and compared the performance of developed methods through simulation studies via Average Run Length (ARL) criterion.

On the other hand, various quality measures have been proposed to evaluate a process performance. One of the most important issues which must be considered in assessing product quality is the process capability analysis. A Process Capability Index (PCI) is a numerical summary that compares the actual process performance related to engineering specifications. So, this concept will be acceptable for both consumer and manufacturer. As mentioned in Montgomery [18], an important consideration for measuring the process capability is to survey whether the process is in statistical control or not. During last years, many authors have investigated the use of various univariate process capability indices which could be found in Kotz and Johnson [19]. Furthermore, in some processes, multiple quality characteristics may be considered for measuring process capability. When these variables are correlated, a multivariate statistical technique must be used to analyze the process capability. In [20–31] multivariate capability indices have been developed and presented for assessing process capability. Some of multivariate process capability indices are defined based on the ratio of a tolerance region to a process region, such as the method proposed by Shahriari et al. [25], while some authors have used principal components analysis or the probability of producing nonconforming items, such as Wang and Chen [32] and Polansky [33], respectively.

A general framework for research topics about profile monitoring could be found in a review article which is presented by Woodall [34]. As mentioned by Woodall [34], there is no research on assessing the process capability in linear or nonlinear profile up to the year 2007. After that, Shahriari and Sarrafian [35] proposed a basic method to measure process capability in simple linear profile. Razavi et al. [36] proposed a method to determine the capability for the intercept and slope of simple linear profile, independently. Hosseinfard and Abbasi [37] employed the proportion of the non-conformance criteria to estimate the process capability index in linear profiles. Ebadi and Shahriari [38] also investigated measuring process capability in simple linear profiles. Hosseinfard and Abbasi [39] considered five methods to estimate process capability index in non normal linear profiles.

Although in recent years some approaches have been proposed for determination of process capability in simple linear profiles, to the best of our knowledge, there is no research on the evaluation of process capability in multivariate simple linear profiles. However, there are several applications of multivariate simple linear profiles in industry introduced by researchers, such as Noorossana et al. [15,16] and Eyvazian et al. [17].

In this paper, we develop three methods to handle process capability analysis in multivariate simple linear profiles. In addition, the performance of the proposed methods is evaluated through simulation studies and a real case. The results show the suitable performance of the proposed methods. The structure of the paper is as follows: Section 2 contains the general framework of multivariate simple linear profile model. We explain our proposed methods in Section 3. In Section 4, the performance of the proposed process capability indices is evaluated through simulation studies. The applicability of the proposed methods is illustrated through a real dataset given from the 1600-ton press machine case in Section 5. The last section contains our concluding remarks.

### 2. Model and assumptions

As mentioned in the introduction section, a multivariate simple linear profile is usually modeled by a multivariate simple linear regression model. In this case, there are relationships between some dependent response variables and one explanatory variable. Each of response variables $y_1, y_2, \ldots, y_p$ is predicted by the explanatory variables $x$. It is assumed that $m$ random samples of size $n$ are taken from the process. For the $k$th random sample collected over time, we have $n$ fixed values for explanatory variable (sample size), and for each value of explanatory variable, we have $p$ corresponding response values. We specify the observations in each sample collected over time by $(x_{ik}, y_{1ik}, y_{2ik}, \ldots, y_{pik})$, $i = 1, 2, \ldots, n$, $k = 1, 2, \ldots, m$. Thus, we have a column of $\beta$'s for each column of $Y$, and these columns form a matrix $B = (\beta_1, \beta_2, \ldots, \beta_p)$. Hence, our multivariate model could be defined as:

$$y_k = X\hat{\beta}_k + e_k, \quad k = 1, 2, \ldots, m,$$  \hspace{1cm} (1)

or equivalently:

$$
\begin{bmatrix}
    y_{11k} & y_{12k} & \cdots & y_{1pk} \\
    y_{21k} & y_{22k} & \cdots & y_{2pk} \\
    \vdots & \vdots & \ddots & \vdots \\
    y_{n1k} & y_{n2k} & \cdots & y_{npk}
\end{bmatrix}
= 
\begin{bmatrix}
    1 & x_{1k} & \cdots & \epsilon_{11k} & \cdots & \epsilon_{1pk} \\
    1 & x_{2k} & \vdots & \beta_{11k} & \cdots & \beta_{1pk} \\
    \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\
    1 & x_{nk} & \epsilon_{n1k} & \cdots & \epsilon_{n2k} & \cdots & \epsilon_{npk}
\end{bmatrix},
\hspace{1cm} (2)
$$

where $Y_k = (y_{1k}, y_{2k}, \ldots, y_{nk})^T$ is a $n \times p$ matrix of response variables for the $k$th sample, and $X = [1 \ x]$ is a $n \times 2$ matrix of explanatory variable. For simplicity, it is assumed that the $x$-values are fixed and take the same set of values for each sample. It is assumed that the vector of error terms has a known covariance matrix $\Sigma$, which can be obtained from:

$$
\Sigma = 
\begin{bmatrix}
    \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1p} \\
    \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2p} \\
    \vdots & \vdots & \ddots & \vdots \\
    \sigma_{p1} & \sigma_{p2} & \cdots & \sigma_{pp}
\end{bmatrix},
\hspace{1cm} (3)
$$

where $\sigma_{ij}$ represents the covariance between $i$th and $j$th error vector terms at each observation. Covariance matrix between response variables can be also given by Eq. (3).

Based on Rencher [40], the least squares estimators of the matrix $\hat{\beta}_k$ can be now defined as:

$$\hat{\beta}_k = (X^TX)^{-1}X^TY_k. \hspace{1cm} (4)$$

An unbiased estimator of $\text{cov}(y) = \Sigma$ is given by:

$$S_e = \frac{(Y - \hat{X}\hat{\beta})^T(Y - \hat{X}\hat{\beta})}{n - 2} = \frac{Y^TY - \hat{B}^TX^TY}{n - 2}. \hspace{1cm} (5)$$
3. Proposed methods

In this section, three methods are proposed for evaluating the capability of a multivariate simple linear profile. The first method is based on the percentage of nonconforming parts produced at each response variable. The second method is a multivariate capability vector and the third one applies the principal component analysis to measure process capability. These methods will be fully explained in the coming sections.

3.1. Method A: process capability estimation based on nonconforming percentage

In this method, we first determine the average percentage of nonconforming parts for each response variables $y_1, y_2, \ldots, y_p$, by computing the percentage of the parts falling outside the specification limits. For the $j$th response variable, the nonconforming percentage can be calculated as follows:

$$P_j = P_{U,j} + P_{L,j} = \frac{1}{n} \sum_{i=1}^{n} P_{U_{i,j}} \sum_{i=1}^{n} P_{L_{i,j}},$$

where $P_{U_{i,j}}$ is the percentage of nonconforming parts below the USL$_j$, and $P_{L_{i,j}}$ is the percentage of nonconforming parts above the USL$_j$ and they are defined as follows, respectively:

$$P_{U_{i,j}} = P(y_i < \text{USL}_j) = P \left( Z < \frac{\text{USL}_j - \mu_{y_i}}{\sigma_{y_i}} \right),$$

$$P_{L_{i,j}} = P(y_i > \text{USL}_j) = P \left( Z > \frac{\text{USL}_j - \mu_{y_i}}{\sigma_{y_i}} \right),$$

where $\mu_{y_i}$ and $\sigma_{y_i}$ are the mean and standard deviation for the $i$th level of $j$th response variable, respectively. Then, we use the univariate index $S_{pk}$ to calculate the process capability for each of $y_1, y_2, \ldots, y_p$ responses. It is well known that there is a relation between the process yield and $S_{pk}$ which can be expressed for each response as [41]:

$$\%\text{yield} = 2\Phi(3S_{pk}) - 1, \quad j = 1, 2, \ldots, p,$$

where $\%\text{yield} = 1 - P_j$ and $P_j = 1.2, \ldots, p$ is obtained from Eq. (6). So, the value of $S_{pk}$ for each response variable can be computed by Eq. (9). Obviously, there is a one-by-one relationship between $S_{pk}$ and the process yield. Thus, $S_{pk}$ provides an exact (rather than approximate) measure of the process yield. After computing the value of $S_{pk}$ for all response variables separately, we use multivariate capability index, $S^T_{pk}$, proposed by Chen et al. [30] to estimate the overall capability ($S_{pk}^T$) of a process described by a multivariate simple linear profile:

$$S_{pk}^T = \frac{1}{3} 2\Phi^{-1} \left[ \left( \prod_{j=1}^{p} (2\Phi(3S_{pk}) - 1) + 1 \right)/2 \right].$$

A one-by-one relationship between the index $S_{pk}^T$ and the overall process yield $1 - P$ can be established as:

$$1 - P = \prod_{j=1}^{p} (1 - P_j) = \prod_{j=1}^{p} [2\Phi(3S_{pk}^T) - 1] = 2\Phi(3S_{pk}^T) - 1.$$

For a process with $p$ response variables, the requirement of $C_1 \leq S_{pk} \leq C_2$ for the overall process capability would be satisfied, if the capability of the $j$th response satisfies $S_j \leq S_{pk} \leq S_j$ for all $j = 1, 2, \ldots, p$. It can be shown that the lower and upper bounds, respectively, for each $S_{pk}$ can be determined as follows:

$$S_{pk} = \frac{1}{3} 2\Phi^{-1} \left( \frac{\sqrt{2\Phi(3C_1)} - 1}{2} \right),$$

$$S_{pk} = \frac{1}{3} 2\Phi^{-1} \left( \frac{\sqrt{2\Phi(3C_2)} - 1}{2} \right).$$

A drawback of this approach is that it measures the overall process yield, assuming the quality characteristics are independent, and ignores the correlation between response variables in estimating the overall process capability. This problem would be more explained in Section 4.

3.2. Method B: a multivariate capability vector

In this section, we propose a method for measuring process capability of a multivariate simple linear profile based on the method proposed by Shahriari et al. [25]. They developed a three-component multivariate process capability vector $C_{pm}$, PV, UL]. Considering their proposed method, the first component, labeled $C_{pm}$, is defined for multivariate simple linear profiles as follows:

$$\hat{C}_{pm} = \left[ \frac{\text{Vol of tolerance region}}{\text{Vol of modified process region}} \right]^\frac{1}{m}.$$

In analyzing the value of this component, values higher than 1 indicate that the volume of modified process region is smaller than the volume of the tolerance box, namely the part is acceptable. USL$_j$ and LSL$_j$ are upper and lower specification limits for the $j$th level of $j$th response variable, $i = 1, 2, \ldots, n, \quad j = 1, 2, \ldots, p$.

This approach uses a modified process region whose limits are simplified by Hardle and Simar [42] as follows:

$$\text{UPL}_j = \mu_j + \sqrt{\chi^2_{\alpha, p}} \cdot \sigma_j,$$

$$\text{LPL}_j = \mu_j - \sqrt{\chi^2_{\alpha, p}} \cdot \sigma_j,$$

where $\mu_j$ and $\sigma_j$ are the mean and variance for the $j$th level of $j$th response variable, respectively, and $\chi^2_{\alpha, p}$ is the upper 100 $\alpha$% of a chi-square distribution with $p$ degrees of freedom. So, the first component could be rewritten as:

$$\hat{C}_{pm} = \left[ \prod_{j=1}^{p} \prod_{i=1}^{n} (\text{USL}_j - \text{LSL}_j) \right]^\frac{1}{m}.$$

It should be noted that in multivariate simple linear profile, for the $j$th response variable, the variance of response is the same at levels of independent variable. So, $\sigma_j^2$ for each response is equivalent to $\sigma^2_j$ and could be obtained from Eq. (3). The first component represents the adequacy of our process dispersion.
The second component measures the significance level of the observed value of responses based on the Hotelling’s $T^2$ statistic, which could be defined for multivariate simple linear profiles as:

$$T_i^2 = m\langle \overline{y}_i - \overline{T}_i \rangle^2 S^{-1}\langle \overline{y}_i - \overline{T}_i \rangle,$$

where $\overline{y}_i$ and $\overline{T}_i$ are the sample mean vector and the target vector of response variables at $i$th level of $x$, respectively, and $S$ is the sample variance–covariance matrix. We consider $T_i = \overline{y}_i - \overline{T}_i$ in Eq. (16) when it is used for multivariate simple linear profile, and $m$ is the number of samples used to estimate the parameters. So, this test measures the similarity of $\overline{y}_i$ to the center of specification limits of responses, which is equivalent to evaluation of the centrality of the process. Then the $P$-Value may be computed from [43]:

$$PV_i = P\left(F_{p,m-p} > \frac{m - p}{p(m - 1)} T^2_i \right),$$

where $F_{p,m-p}$ is the $F$-distribution with $p$ and $m - p$ degrees of freedom. The $P$-Value will never exceed 1 and the center of process will be far from the engineering target value when the $P$-Value is close to zero. For assessing capability with respect to this component, a level of significance $\alpha$ can be chosen. Typically, $\alpha$ is set equal to 0.05.

Because the response values at each level of independent variable have multivariate normal distribution, we define the second component for each level, separately. So, in the case of multivariate simple linear profile, we should have different $P$-Values at each level of independent variable, and we can determine the centrality of process at each level. Then, the variable for process mean would be taken at levels with low $P$-Values.

The third component (LI) compares the location of the process region, and the tolerance region, and is defined as:

$$LI = \begin{cases} 1 & \text{if modified process region is located within tolerance region} \\ 0 & \text{otherwise}. \end{cases}$$

In the case of multivariate simple linear profile, the third component is equal to 1 when the modified process region of all the regression models is contained within tolerance region.

Consequently, the multivariate capability vector has three components, $C_{pv}, PV, LI$, which provide a comparison of the values of the regions, locations of the centers and location of the regions.

By using this method, we do not need to either compute the process capability or to find the proportion of nonconforming items at each level of the predictor, and we compute the three components of the process capability vector. The gain in using this index is that when the process is incapable, one can easily recognize that it is due to the location or variability. Furthermore, we can determine non-centered levels with respect to low $P$-Values. This method also provides comparison between two or more than two multivariate simple linear profiles in different aspects of process dispersion or process centrality.

3.3. Method C. process capability estimation using principal component analysis

Wang and Du [27] proposed a useful method using the Principal Component Analysis (PCA), for describing the process performance of a multivariate process.

A set of few PCs normally comprises 80%–90% of the process variability. By using this subset, the multivariate quality characteristic problem can be reduced in dimensionality. To determine the principal components of response variables, we use the method proposed by Norossana et al. [16] for monitoring multivariate multiple linear profiles. In this method, they used principal components analysis to make the correlated response variables independent, and to reduce the number of response variables. We first pool all the $m$ samples into one sample of size $mn$ and compute the principal components of the response variables, which explain the most variations in the response variables. These new components are linear combinations of the $p$ response variables and follow normal distributions. Then, we calculate the scores of the principal components of the $p$ response variables for each observation and choose the first few principal components, which explain most of the variation. These components are considered as new response variables and, we model the relationship between the new response variables and independent variable.

Considering the eigenvalues and eigenvectors of variance–covariance matrix $\Sigma$ in Eq. (3) as $\lambda_1, \lambda_2, \ldots, \lambda_p$ and $e_1, e_2, \ldots, e_p$, respectively, the overall capability index for the process is defined as:

$$M_{cp}^\ast = \left( \prod_{i=1}^{q} \prod_{j=1}^{n} \hat{C}_{p,ij} \right)^{-\frac{1}{mn}},$$

where:

$$\hat{C}_{p,ij} = \frac{USL_{p,ij} - LSL_{p,ij}}{\theta S_{p,ij}},$$

$$USL_{p,ij} = e^T_j USL_i,$n, \quad J = 1, 2, \ldots q.$

$q$ is the number of principal components comprising much of the process variability; $USL_i$ and $LSL_i$ are the vectors of upper and lower specification limits at $i$th level of $x$, respectively, and $S$ is the sample variance–covariance matrix. $\hat{C}_{p,ij}$ represents the univariate measure of potential process capability for the $i$th level of $j$th PC. Similarly, we can define $M_{cpk}, M_{cpm}$ and $M_{cppk}$ by replacing $\hat{C}_{p,ij}$ with $\hat{C}_{pk,ij}, \hat{C}_{pm,ij}$ and $\hat{C}_{ppk,ij}$, respectively, for $i = 1, 2, \ldots, n, j = 1, 2, \ldots, q$. Furthermore, since the new response variables (principal components) are normally distributed, an approximate $(1 - \alpha)100\%$ confidence interval for $M_{cp}^\ast$ is obtained from [27]:

$$\left[ \prod_{i=1}^{q} \prod_{j=1}^{n} \hat{C}_{p,ij} \right]^{\frac{1}{mn}} \pm \delta M_{cp},$$

where $m$ is the number of random samples.

Table 1 compares the advantages and disadvantages of the proposed methods.

4. Performance comparisons

In this section, the proposed methods are evaluated via simulation studies. Although analyses can be performed for different values of $p$, without loss of generality, we consider the
bivariate case \((p = 2)\). Assuming the process is in statistical control, the underlying multivariate simple linear profile model is used for simulation:

\[
Y_1 = 2 + 1x + \varepsilon_1, \quad Y_2 = 3 + 2x + \varepsilon_2.
\]

The fixed values of \(x\) are defined as \(x = [2, 4, 6, 8]\) with mean \(\bar{x} = 5\). The error terms vector \((\varepsilon_1, \varepsilon_2)\) is a bivariate normal random vector with mean vector zero and known covariance matrix \(\Sigma = \begin{bmatrix} \sigma_1^2 & \rho \sigma_1 \sigma_2 \\ \rho \sigma_1 \sigma_2 & \sigma_2^2 \end{bmatrix}\). Figure 1 illustrates the effect of correlation between error terms which leads to correlation between response variables for \(\rho = 0.1, 0.5\) and 0.9.

For the first regression model, we considered USL = 25 and LSL = 0. For the second one, it is assumed that the upper and lower specification limits depend on the independent variable. The values of the specification limits at each level of \(x\) are defined in Table 2. The values of USL and LSL could be obtained by fitting regression lines such as USL = \(b_0^U + b_1^U x_i\) and LSL = \(b_0^L + b_1^L x_i\), respectively. These lines are illustrated in Figure 2 and are obtained by 10,000 simulation runs. In each run, we generated 20 samples including two simple linear profiles. The scatter plot and the fitted line for the two regression models are also shown in Figure 2.

Table 3 shows the simulation results of the first method for different values of \(\sigma_1^2\) and \(\sigma_2^2\). Through the simulation studies, we concluded that the value of \(\rho\) does not affect the estimated process capability index \(S_{pk}^T\). However, our simulation studies (not reported here) showed that the actual nonconforming percentage of process does not differ significantly when the error term’s variances are constant and the correlation between error terms remains unchanged. So, for the sake of simplicity, we assumed \(\rho = 0.5\) in Table 3. \(S_{pk1}\) and \(S_{pk2}\) are the estimated process capability for the first and second regression models, respectively. \(N_i\) is the total number of parts falling outside the specification limits, which is computed for the two regression models through 10,000 simulation runs.

As mentioned, there is a relation between the process yield and \(S_{pk}\). Furthermore, it can be concluded from Table 3, that there is proportionality between the real total number of nonconforming parts and estimated process capability. Increasing the total number of nonconforming parts leads to a decrease in the overall process capability.

Table 4 displays the lower bound \(S_{Lk}\) and upper bound \(S_{Uk}\) on \(S_{pk}\), if the requirement of the overall process capability is 1.000 \(\leq S_{pk}^T \leq 1.333\) and 1.50 \(\leq S_{pk}^T \leq 2\). Based on computed \(S_1\) and \(S_2\), we determined the minimum and maximum value of process variance corresponding to the maximum and minimum yield percent, respectively, by using Eqs. (7)–(9).

Figure 3 illustrates the performance of the first proposed index under different variance values. As it is obvious from Figure 3, the estimated process capability index \(S_{pk}^T\) deteriorates when the error terms variance increases continuously. Beside the aforementioned reasons, this could confirm the usefulness and proper performance of this proposed method.

Table 5 shows the simulation results of the estimated process capability by using the multivariate capability vector and PCA method for different values of \(\sigma_1^2\) and \(\sigma_2^2\).

For different values of \(\sigma_1^2\) and \(\sigma_2^2\), we determined the principal components of responses to estimate \(M\hat{C}_{pc}\). The \(P\)-Values in Table 5 are close to zero because the vector of mean samples is far from the target vector, especially in the first linear model. These low \(P\)-Values lead to a decrease in the overall process capability.

By comparison of the computed multivariate process capability vectors which are presented in Table 5 with estimated values of \(S_{pk}^T\) in Table 3, we can conclude that potential process capability is high, and correcting the process centrality or changing the center of defined specification limits will improve the overall process capability significantly.

It should be noted that the index \(M\hat{C}_{pc}\) considers the correlation between variables for computing process capability. However, the value of index \(C_{pm}\) in the multivariate capability vector does not depend on the correlation between variables. In the second proposed method, this correlation affects the calculated \(P\)-Values. That is the reason of significant differences between estimated \(M\hat{C}_{pc}\) and \(C_{pm}\) in Table 5 for some values of \(\sigma_1^2\) and \(\sigma_2^2\). However, as our simulation studies showed and it is obvious from Table 5, \(M\hat{C}_{pc}\) is close to \(C_{pm}\) when \(\sigma_1^2 = \sigma_2^2\), and the defined confidence interval for \(M\hat{C}_{pc}\) would be also more reliable for this bivariate profile.

5. Example

In this section, we illustrate how the proposed methods can be applied to a calibration case. Noorossana et al. [15] considered a real example of calibration between the desired force and the real force produced by a 1600-ton hydraulic press machine at Body Shop of an automotive industrial group. Similar to all hydraulic machines, the 1600-ton hydraulic press machine consists of important components such as pistons, cylinders, and hydraulic pipe. A main input value in this machine is the nominal force exerted by cylinders. In this case, for each value of nominal force (explanatory variable), there are four real forces measured in the four cylinders (response variables) of the machine. The purpose is to evaluate whether...
the nominal force \( x \) and the real forces \( y_1, y_2, y_3 \) and \( y_4 \) are close enough or not. Noticeable differences between the nominal force and the real forces, which could be due to oscillation in oil temperature, variation in oil volume, etc., might lead to defective or low quality parts. Because in this situation there is a correlation between responses, the process capability

![Image](https://via.placeholder.com/150)

**Figure 1:** Distribution of error terms for (a) \( \rho = 0.1 \), (b) \( \rho = 0.5 \), and (c) \( \rho = 0.9 \) in a bivariate profile.

![Image](https://via.placeholder.com/150)

**Figure 2:** Fitted line and specification limits for linear profile. (a) \( Y_{ij} = 2 + 1X_i \), and (b) \( Y_{ij} = 3 + 2X_i + \varepsilon_{ij}, \varepsilon_{ij} \sim N(0,1) \).

| Requirement | \( S_L \) | \( S_U \) | \( \%\text{yield}_L \) | \( \%\text{yield}_U \) | \( \sigma^2_{\text{US}} \) | \( \sigma^2_{\text{DS}} \) | \( \sigma^2_{\text{TD}} \) |
|-------------|------|------|----------------|----------------|----------------|----------------|----------------|
| 1.00 \( \leq S^2_{\text{US}} \leq 1.33 \) | 1.068 | 1.383 | 98.8649 | 99.9967 | 0.5488 | 1.0205 | 1.202 |
| 1.50 \( \leq S^2_{\text{US}} \leq 2 \) | 1.548 | 2.037 | 99.9997 | 1 | 0.2266 | 0.422 | 0.4805 |

Table 4: Minimum and maximum value of process variance for the determined overall process capability.
Figure 3: Process capability estimation using $S_{pm}^2$ index for different values of error variance.

of response variables cannot be investigated separately, and the overall process capability should be estimated using multivariate approaches.

Noorossana et al. [15] considered the relationships between the nominal force (x) and the real forces measured in four cylinders ($y_1$, $y_2$, $y_3$, and $y_4$) as key measurable quality characteristics, which should be kept in control and should be monitored over time. Their Phase I analysis of a set of historical data showed that the stable model between the response variables and independent variable is as follows:

$$Y_1 = -8.5 + 0.87x + \varepsilon_1,$$
$$Y_2 = -5.8 + 0.95x + \varepsilon_2,$$
$$Y_3 = 3.2 + 1.04x + \varepsilon_3,$$
$$Y_4 = 13.6 + 1.09x + \varepsilon_4,$$

where $(\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4)$ is normal random vector with mean vector of zero and covariance matrix of:

$$\Sigma = \begin{bmatrix}
80.0 & 89.6 & 45.1 & 25.3 \\
89.6 & 122.1 & 71.5 & 29.1 \\
45.1 & 71.5 & 189.0 & -28.8 \\
25.3 & 29.1 & -28.8 & 84.4
\end{bmatrix}.$$

In this case, 15 in-control samples, where each sample consists of 11 values for nominal force as 50, 80, 110, 140, 170, 200, 230, 260, 290, 320 and 350, are collected and the corresponding output force associated with each cylinder is measured. The dataset is available by authors upon request. We defined the specification limits of $y_1$, $y_2$, $y_3$ and $4\alpha$ as:

- LSL$_1$ : 5(25)/255, USL$_1$ : 75(25)/325,
- LSL$_2$ : 5(25)/255, USL$_2$ : 105(25)/355,
- LSL$_3$ : 5(30)/305, USL$_3$ : 110(30)/410,
- LSL$_4$ : 5(30)/305, USL$_4$ : 140(30)/440.

Note that in the above specification limits, for example, 5(25)/255 implies that the LSL$_1$ increases in amount of 25 in each level from the LSL$_1$ = 5 for the first x-level. We considered significance level of $\alpha = 5\%$ to compute the process capability. Table 6 shows the calculated values of modified process regions for the response variables which are used for computing multivariate capability vector. It can be concluded from Table 6 that the modified process region for the response $Y_2$ is not located within tolerance region at levels $i = 10, 11$. So, LI is equal to zero. Table 7 shows the process capability estimation for these calibration application data using the proposed methods.

Explained percent of total variability for $y_1$, $y_2$, $y_3$ and $y_4$ are $%1.77$, $%9.85$, $%30.42$ and $%57.95$, respectively. Thus, for the PCA approach, we used the first and the second principal components as new response variables to evaluate the capability of this process at 88.37% total variability. An approximate 95% confidence interval of $M^P$ is $[1.119, 1.642]$. As it is obvious from Table 7, the estimated $C_{pm}$ by using second method is in the confidence interval and it confirms the proper performance of calculated confidence interval.

It can be concluded from Table 7 that the estimated values of $M^P$, $C_{pm}$ and $C_{pm}$ are close to that of estimated for $S_{pm}^2$, because all of these indices consider the mean and the standard deviation of the process simultaneously, in determining the process performance.

Using the proposed methods altogether leads to a comprehensive analyze about the process capability.

The overall process capability would be acceptable due to estimated $S_{pm}^2$ (it is greater than 1). However, improving the capability of $y_2$ which has the minimum estimated process capability will increase the overall process capability.

Based on the second method, the centrality of the process should be corrected at levels $i = 1, 7, 8, 9, 10, 11$ and the process centrality at levels $i = 2, 3, 4, 5, 6$ is acceptable at significance level $\alpha = 5\%$. It can be concluded from the comparison of estimated $S_{pm}^2$ and $C_{pm}$ that improving the centrality of non-centered levels leads to better overall process capability.

Based on the third method, the process capability estimation could be reduced to determine the capability of two independent variables. Furthermore, the potential capability of the process remains between 1.119 and 1.642 at the 95% confidence level.

6. Conclusions

In this paper, three new methods for measuring the process capability in multivariate simple linear profiles were proposed.
and their performance was investigated through simulation studies. The first method uses the average percentage of nonconforming parts to transform it into a process capability measure. The advantage of this method is that there is a one-to-one relationship between process capability and the process yield. So, this new index is useful to the engineers or practitioners in measuring the yield. The second method is a multivariate capability vector which separates the process dispersion and its centrality to measure process capability and the third is based on principal component analysis. Simulation results indicated that using the proposed methods simultaneously gives comprehensive information about the capability of multivariate simple linear profiles. We also illustrated the use of our proposed methods under a calibration application in a hydraulic press at Body Shop of an automotive company. The approaches presented in this paper can be investigated to develop the process capability indices for more complicated models such as multivariate multiple linear profiles.

References

[1] Kang, L. and Albin, S.L. “On-line monitoring when the process yields a linear profile”, Journal of Quality Technology, 32, pp. 418–426 (2000).

[2] Kim, K., Mahmoud, M.A. and Woodall, W.H. “On the monitoring of linear profiles”, Journal of Quality Technology, 35, pp. 317–328 (2003).

[3] Stover, F.S. and Brill, R.V. “Statistical quality control applied to ion chromatography calibrations”, Journal of Chromatography, 804, pp. 37–43 (1998).

[4] Mahmoud, M.A. and Woodall, W.H. “Phase I monitoring of linear profiles with calibration application”, Technometrics, 46, pp. 380–391 (2004).

[5] Woodall, W.H., Spitzner, D.J., Montgomery, D.C. and Gupta, S. “Using control charts to monitor process and product quality profiles”, Journal of Quality Technology, 36, pp. 309–320 (2004).

[6] Wang, K. and Tsung, F. “Using profile monitoring techniques for a data-rich environment with huge sample size”, Quality and Reliability Engineering International, 21, pp. 677–688 (2005).

[7] Zou, C., Zhang, Y. and Wang, Z. “Control chart based on change-point model for monitoring linear profiles”, IIE Transactions, 38, pp. 1093–1103 (2006).

[8] Gupta, S., Montgomery, D.C. and Woodall, W.H. “Performance evaluation of two methods for online monitoring of linear calibration profiles”, International Journal of Production Research, 44, pp. 1927–1942 (2006).

[9] Zhang, J., Li, Z. and Wang, Z. “Control chart based on likelihood ratio for monitoring linear profiles”, Computational Statistics and Data Analysis, 53, pp. 1486–1495 (2009).

[10] Saghaei, A., Mehrjoo, M. and Amiri, A. “A CUSUM-based method for monitoring linear profiles”, The International Journal of Advanced Manufacturing Technology, 45, pp. 1252–1260 (2009).

[11] Mahmoud, M.A., Morgan, J.P. and Woodall, W.H. “The monitoring of simple linear regression profiles with two observations per sample”, Journal of Applied Statistics, 37, pp. 1243–1261 (2010).

[12] Kazemzadeh, R.B., Noorossana, R. and Amiri, A. “Phase I monitoring of polynomial profiles”, Communications in Statistics-Theory and Methods, 37, pp. 1671–1686 (2008).

[13] Kazemzadeh, R.B., Noorossana, R. and Amiri, A. “Monitoring polynomial profiles in quality control applications”, The International Journal of Advanced Manufacturing Technology, 42, pp. 703–712 (2009).

[14] Mahmoud, M.A. “Phase I analysis of multiple linear regression profiles”, Communications in Statistics-Simulation and Computation, 37, pp. 2106–2130 (2008).

[15] Noorossana, R., Eyvazian, M. and Vaghefi, A. “Phase II monitoring of multivariate simple linear profiles”, Computers and Industrial Engineering, 58, pp. 563–570 (2010).

[16] Noorossana, R., Eyvazian, M., Amiri, A. and Mahmoud, M.A. “Statistical monitoring of multivariate multiple linear regression profiles in Phase I with calibration application”, Quality and Reliability Engineering International, 26, pp. 291–303 (2010).

[17] Eyvazian, M., Noorossana, R., Saghaei, A. and Amiri, A. “Phase II monitoring of multivariate multiple linear regression profiles”, Quality and Reliability Engineering International, 27, pp. 281–296 (2011).

[18] Montgomery, D.C., Introduction to Statistical Quality Control, John Wiley & Sons, New York (2005).

[19] Kotz, S. and Johnson, N.L. “Process capability indices—a review: 1992–2000”, Journal of Quality Technology, 34, pp. 3–19 (2000).

[20] Chan, L.K., Cheng, S.W. and Spring, F.A. “A multivariate measure of process capability”, International Journal of Modelling and Simulation, 11, pp. 6–10 (1991).

[21] Taam, W., Subbaiah, P. and Liddy, J.W. “A note on multivariate capability indices”, Journal of Applied Statistics, 20, pp. 339–351 (1993).

[22] Pearn, W.L., Kotz, S. and Johnson, N.L. “Distributional and inferential properties of PCI”, Journal of Quality Technology, 24, pp. 216–231 (1992).

[23] Chen, H.F. “A multivariate process capability index over a rectangular solid tolerance zone”, Statistica Sinica, 4, pp. 749–758 (1994).

[24] Karl, D.P., Mosiette, J. and Taam, W. “Some applications of a multivariate capability index in geometric dimensioning and tolerancing”, Quality Engineering, 6, pp. 649–665 (1994).
[25] Shahriari, H., Hubele, N.F. and Lawrence, F.P. “A multivariate process capability vector”, In Proceeding of The 4th Industrial Engineering Research Conference, pp. 304–309, Institute of Industrial Engineers (1995).

[26] Boyles, R.A. “Exploratory capability analysis”, Journal of Quality Technology, 28, pp. 91–98 (1996).

[27] Wang, F.K. and Du, T.C.T. “Using principal component analysis in process performance for multivariate data”, Omega—The International Journal of Management Science, 28, pp. 185–194 (2000).

[28] Wang, F.K., Hubele, N.F., Lawrence, P., Miskulin, J.D. and Shahriari, H. “Comparison of three multivariate process capability indices”, Journal of Quality Technology, 32, pp. 263–275 (2000).

[29] Castagliola, P. “New process capability indices for two quality characteristics”, 5th International Industrial Engineering Conference, Quebec, Canada (2003).

[30] Chen, K.S., Pearn, W.L. and Lin, P.C. “Capability measures for process with multiple characteristics”, Quality and Reliability Engineering International, 11, pp. 101–110 (2003).

[31] Khadse, K.G. and Shinde, R.L. “Probability-based process capability indices”, Communications in Statistics-Simulation and Computation, 38, pp. 884–904 (2009).

[32] Wang, F.K. and Chen, J.C. “Capability index using principal component analysis”, Quality Engineering, 11, pp. 21–27 (1999).

[33] Polansky, A.M. “A smooth nonparametric approach to multivariate process capability”, Technometrics, 43, pp. 199–211 (2001).

[34] Woodall, W.H. “Current research on profile monitoring”, Rev Producao, 17, pp. 420–625 (2007).

[35] Shahriari, H. and Sarrafi, M. “Assessment of process capability in linear profiles”, Proceedings of the 6th International Industrial Engineering Conference, Tehran, Iran (in Farsi) (2009).

[36] Ehtesabi, H., Asadzadeh, A. and Naseer, Z. “Determination of process capability indices for the slope parameter of a single variable linear profile”, International Journal of Industrial Engineering and Production Management, 21, pp. 33–43 (2010).

[37] Hosseinifard, S.Z. and Abbasi, B. “Evaluation of process capability indices of linear profiles”, International Journal of Quality & Reliability Management, 29, pp. 162–176 (2012).

[38] Ebadi, M. and Shahriari, H. “A process capability index for simple linear profile”, The International Journal of Advanced Manufacturing Technology, (2012), http://dx.doi.org/10.1007/s00170-012-4066-7.

[39] Hosseinifard, S.Z. and Abbasi, B. “Process capability analysis in non normal linear regression profiles”, Communications in Statistics-Simulation and Computation, 41, pp. 1761–1784 (2012).

[40] Bencher, A.C. Methods of Multivariate Analysis, Wiley, New York (2002).

[41] Boyles, R.A. “Process capability with asymmetric tolerances”, Communications in Statistics: Computation and Simulation, 23, pp. 615–643 (1994).

[42] Hardle, W. and Simar, L., Applied Multivariate Statistical Analysis, Springer Verlag, Berlin (2007).

[43] Pearn, W.L. and Kotz, S., Encyclopedia and Handbook of Process Capability Indices, A Comprehensive Exposition of Quality Control Measures, World Scientific Publishing, New Jersey (2006).

Mohsen Ebadi received his B.S. and M.S. degrees in Industrial Engineering from Iran University of Science and Technology and Khajeh Nasir University of Technology, respectively. His research interests are in the areas of statistical quality control, robust statistics and process capability analysis.

Amirhossein Amiri is Assistant Professor at Shahed University, in Iran. He holds B.S., M.S., and Ph.D. degrees in Industrial Engineering from Khajeh Nasir University of Technology, Iran University of Science and Technology, and Tarbiat Modares University, respectively, and is a member of the Iranian Statistical Association. His research interests are statistical quality control, profile monitoring and Six Sigma.