Process yield analysis for multivariate linear profiles

Fu-Kwun Wang

Department of Industrial Management, National Taiwan University of Science and Technology, Taipei, Taiwan

ABSTRACT

Process capability analysis plays an important role in statistical quality control. We present the process yield index $T_{SpkA}$ to evaluate the process yield for multivariate linear profiles in manufacturing processes. This index provides an exact measure of the process yield. In addition, an approximate confidence interval for $T_{SpkA}$ is constructed. A simulation study is conducted to assess the performance of the proposed method for multivariate linear profiles data under mutually independent normality and multivariate normality. The simulation results confirm that the estimated $T_{SpkA}$ is close to the target value with smaller standard deviation as the sample size increases.

Introduction

Profile monitoring is the use of control charts for cases in which the quality of a process or product can be characterized by a functional relationship between a response variable and one or more explanatory variables (Woodall, 2007). Many researches on profile monitoring considered the case in which the profile can be adequately represented by a simple linear regression model. For example, Kang and Albin (2000) provided a simple linear profile which the process is represented by a linear function of the set points for gas flow in semiconductor manufacturing. In general, two phases are considered in the monitoring of a profile. On one hand, phase I analysis is to estimate the profile parameters and to assess the stability of the process. On the other hand, phase II analysis is to detect any shift in the process parameters. Woodall (2007) reviewed common methods in profile monitoring. A comprehensive review of profile monitoring can be found in (Noorossana, Saghaei, & Amiri, 2011). Many studies have been done by researchers on the monitoring of simple linear profiles, multiple linear profiles, polynomial profiles, non-linear profiles, logistics profiles and circle/cylindrical profiles. When the observations within each profile in continuous manufacturing processes are uncorrelated, the effect of Phase II monitoring of autocorrelated linear and polynomial profiles in a first-order autoregressive AR(1) model using average run length (ARL) criterion can be found in Noorossana, Amiri, and Soleimani (2008),
Soleimani, Noorossana, and Amiri (2009) and Kazemzadeh, Noorossana and Amiri (2010). Multivariate linear profile monitoring was investigated by Amiri, Zou, and Doroudyan (2014), Eyvazian, Noorossana, Saghaei, and Amiri (2011), Noorossana, Eyvazian, and Amiri (2010), Noorossana, Eyvazian, and Vaghefi (2010) and Zou, Ning, and Tsung (2012).

Assuring the process capability in multivariate linear profiles to meet the requirement is also a very important task. Multivariate capability indices for processes with multiple quality characteristics can be found in (Chan, Cheng, & Spiring, 1991; Chen, Pearn, & Lin, 2003; Ebadi & Amiri, 2012; Pearn, Kang, Lee, & Liao, 2009; R Development Core Team, 2013; Shahriari & Sarrafian, 2009; Shahriari, Hubele, & Lawrence, 1995; Taam, Subbaiah, & Liddy, 1993; Wang, 2006; Wang, 2010; Wang & Chu, 2013; Wang & Du, 2000; Wang, Hubele, Lawrence, Miskulin, & Shahriari, 2000). Unfortunately, most multivariate capability indices cannot be used to obtain the process yield. Chen et al. (2003) proposed an index $S_{pk}$ based on the yield index $S_{pk}$ proposed by Boyles (1994) for a mutually independent multivariate normal data. Pearn et al. (2009) investigated this index $S_{pk}$ for a photolithography process in a semiconductor fab that includes the testing process performance of critical dimension (CD), alignment accuracy and photo resist thickness. Wang (2010) proposed a process-yield index $T^S_{pk,PC}$ for a multivariate normal data by using the principal component analysis (PCA) method. This research also provided an approximate lower confidence bound (LCB) for the true process yield. In addition, a comprehensive review of process capability analysis can be found in Kotz and Johnson (2002), Wu, Pearn, and Kotz (2009) and Yum and Kim (2011).

Although in recent years, some univariate capability indices have been proposed for evaluating process capability in simple linear profiles under normality and non-normality (Ebadi & Shahriari, 2013; Hosseinifard & Abbasi, 2012a; Hosseinifard & Abbasi, 2012b; Razavi, Asadzadeh, & Naseri, 2010; Shahriari & Sarrafian, 2009). In general, four different types of multivariate linear profiles are identified: mutually independent normality, multivariate normality, mutually independent non-normality and multivariate non-normality. Recently, Ebadi and Amiri (2012) presented three methods for measuring process capability in multivariate simple linear profiles. However, these methods do not provide the confidence intervals of the process capability indices. More researches on process yield analysis in multivariate profiles are needed.

We present a new method to evaluate the process yield for multivariate linear profiles. The rest of the paper is organized as follows. Process yield index for simple linear profiles section contains the existing process yield index for a simple linear profile. We propose a new method for measuring process yield in multivariate linear profiles with mutually independent normality and multivariate normality. The confidence intervals of the overall process yield is also provided. The lower confidence bound not only be used in the process yield assurance but also be used in process capability testing for decision-making. A simulation study is conducted to evaluate the performance of the proposed method. One real example is used to demonstrate the application of our proposed method. Finally, we offer a conclusion and suggestions for future studies.
**Process yield index for simple linear profiles**

We introduce the following notations for simple linear profiles:

| Symbol | Description |
|--------|-------------|
| $y_{ij}$ | the response at the $i$th level of independent variable from the $j$th profile |
| $e_{ij}$ | error term with $\sim N(0, \sigma^2)$ |
| $S_{pk}$ | the process yield index for a normally distributed data |
| $S_{pki}$ | the process yield index at the $i$th level of independent variable from simple linear profiles with normality |
| $P$ | the overall process yield for simple linear profiles |
| $S_{pkA}$ | the process yield index for simple linear profiles |

A simple linear profile is usually defined by a simple regression model. We assume that $m$ random samples of size $n$ are taken from the process. We also assume that the process is in statistical control. For $n$ fixed values of explanatory variable, $j$th random sample ($j = 1, 2, \ldots, m$) are collected over time. The underlying linear model relating the independent variable $x$ to the response variable $y$ is defined as

$$y_{ij} = \beta_0 + \beta_1 x_i + e_{ij}, \ i = 1, 2, \ldots, n, j = 1, 2, \ldots, m, \quad (1)$$

where the intercept and the slope of the line are called profile coefficients. The predicted value of response variable denoted by $\hat{y}$ can be obtained from $\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$ and the unbiased estimator of $\sigma^2$ is defined as $MSE = \hat{\sigma}^2 = \sum_{i=1}^{n} \sum_{j=1}^{m} e_{ij}^2 / [m(n - 2)]$ (Kunter, Nachtsheim, Neter, & Li, 2005).

For a normally distributed process at the $i$th level of independent variable, the index $S_{pk}$ is used to establish the relationship between the manufacturing specification and the actual process performance (Boyles, 1994). This index is defined as

$$S_{pki} = \Phi^{-1}\left\{ \frac{1}{2} \Phi\left( \frac{USL_i - \mu_i}{\sigma_i} \right) + \frac{1}{2} \Phi\left( \frac{\mu_i - LSL_i}{\sigma_i} \right) \right\} = \frac{1}{3} \Phi^{-1}\left\{ \frac{1}{2} \Phi\left( \frac{1 - C_{dr_i}}{C_{dpi}} \right) + \frac{1}{2} \Phi\left( \frac{1 + C_{dr_i}}{C_{dpi}} \right) \right\}, \quad (2)$$

where $LSL_i$ and $USL_i$ are the lower and upper specification limits, $\mu_i$ and $\sigma_i$ are the mean and the standard deviation, $t_i = (USL_i + LSL_i)/2$, $d_i = (USL_i - LSL_i)/2$, $C_{dr_i} = (\mu_i - t_i)/d_i$, $C_{dpi} = \sigma_i/d_i$, $\Phi$ is the cumulative density function of the standard normal distribution, and $\Phi^{-1}$ is the inverse cumulative density function of the standard normal distribution. The above index provides an exact measure of process yield at the $i$th level of the independent variable, where

$$p_i = \Phi\left( \frac{USL_i - \mu_i}{\sigma_i} \right) - \Phi\left( \frac{LSL_i - \mu_i}{\sigma_i} \right) = \Phi\left( \frac{USL_i - \mu_i}{\sigma_i} \right) + \Phi\left( \frac{\mu_i - LSL_i}{\sigma_i} \right) - 1.$$

We assume that $S_{pki} = c$, where $c$ is a specified value. Using Equation (2), the process yield is obtained as

$$p_i = 2\Phi(3S_{pki}) - 1 = 2\Phi(3c) - 1. \quad (3)$$
Obviously, there exists a one-to-one relationship between $S_{pk_i}$ and the process yield. The overall process yield $P$ for a simple linear profile is defined as the average percentage of conforming parts for all levels of independent variable. Thus, the overall process yield for a simple linear profile is derived by

$$P = \frac{1}{n} \sum_{i=1}^{n} p_i = \frac{1}{n} \sum_{i=1}^{n} \left[ 2\Phi(3S_{pk_i}) - 1 \right].$$

(4)

In addition, the process yield index for a simple linear profile is obtained as

$$S_{pkA} = \frac{1}{3} \Phi^{-1} \left[ \frac{(P + 1)}{2} \right] = \frac{1}{3} \Phi^{-1} \left\{ \sum_{i=1}^{n} \left[ \frac{2\Phi(3S_{pk_i}) - 1}{n} + 1 \right] / 2 \right\}. \quad (5)$$

The above index provides an exact measure of the process yield for a simple linear profile. To construct the confidence interval of the process yield index, we developed the asymptotic distribution of $\hat{S}_{pkA}$ which is shown in the Appendix. Therefore, an approximate 100(1 $- \alpha$)% confidence interval for $S_{pkA}$ is derived by

$$\left( \hat{S}_{pkA} - \frac{1}{6n\sqrt{m\phi(3S_{pkA})}}Z_{\alpha/2}, \hat{S}_{pkA} + \frac{1}{6n\sqrt{m\phi(3S_{pkA})}}Z_{\alpha/2} \right),$$

(6)

where

$$\hat{S}_{pkA} = \frac{1}{3} \Phi^{-1} \left\{ \sum_{i=1}^{n} \left[ 2\Phi(3\hat{S}_{pk_i}) - 1 \right] / n + 1 \right\} / 2,$$

$$\hat{a}_i = d_i / \sqrt{2S_i} \left\{ (1 - \hat{C}_{dri})\phi\left( \frac{1 - \hat{C}_{dri}}{\hat{C}_{dpi}} \right) + (1 + \hat{C}_{dri})\phi\left( \frac{1 + \hat{C}_{dri}}{\hat{C}_{dpi}} \right) \right\},$$

$$\hat{b}_i = \phi\left( \frac{1 - \hat{C}_{dri}}{\hat{C}_{dpi}} \right) - \phi\left( \frac{1 + \hat{C}_{dri}}{\hat{C}_{dpi}} \right),$$

$\phi$ is the probability density function of the standard normal distribution and $Z_{\alpha/2}$ is the upper 100($\alpha/2$)% point of the standard normal distribution. In the Appendix, $C_{dri}$ and $C_{dpi}$ appear in the asymptotic expression of $\hat{S}_{pkA}$ as a consequence of the Taylor expansion used in the asymptotic expansion of $\hat{S}_{pkA}$ around the true values $C_{dri}$ and $C_{dpi}$. It should be noted that $\hat{C}_{dri}$ and $\hat{C}_{dpi}$ are used instead because they will converge to $C_{dri}$ and $C_{dpi}$, respectively (see Lee, Hung, Pearn, & Kueng, 2002). Using the equation $P = 2\Phi(3\hat{S}_{pkA}) - 1$, we can easily obtain an approximate 100(1 $- \alpha$)% confidence interval for $P$. 
Process yield indices for multivariate linear profiles

The notations used in multivariate linear profiles are as follows:

| Symbol | Description |
|--------|-------------|
| $Y_{ijk}$ | the response at the $i$th level of independent variable from the $j$th profile for the $k$th quality characteristic |
| $e_{ijk}$ | error term with $\varepsilon \sim N(0, \sigma^2)$ |
| $S_{pki}$ | the process yield index at the $i$th level of independent variable from multivariate linear profiles with mutually independent normality |
| $P_i^T$ | the overall process yield at the $i$th level of independent variable from multivariate linear profiles |
| $TP_i$ | the overall process yield for multivariate linear profiles with mutually independent normality |
| $TP_{PC}$ | the overall process yield for multivariate linear profiles with multivariate normality |

In this study, we only consider two types of multivariate linear profiles such as mutually independent normality and multivariate normality. For each level of independent variable, we have $v$ corresponding response values or quality characteristics. The multivariate model is defined as

$$
Y_{ijk} = \beta_{0k} + \beta_{ik} x_{ik} + \varepsilon_{ijk}, \quad i = 1, 2, \ldots, n, j = 1, 2, \ldots, m, k = 1, 2, \ldots, v.
$$

(7)

For the $k$th quality characteristic at the $i$th level of independent variable, we assume that $LSL_{ik}$ and $USL_{ik}$ are the lower and upper specification limits, $\mu_{ik}$ and $\sigma_{ik}$ are the mean and the standard deviation, $t_{ik} = (USL_{ik} + LSL_{ik})/2$, $d_{ik} = (USL_{ik} - LSL_{ik})/2$, $C_{d_{ik}} = (\mu_{ik} - t_{ik})/d_{ik}$ and $C_{dp_{ik}} = \sigma_{ik}/d_{ik}$.

The proposed overall process yield indices for the two types of multivariate linear profiles are described as following:

1. Mutually independent normality: We consider the overall process yield index proposed by Chen et al. (1994) for a process with multiple characteristics which are mutually independent normality. For the $i$th level of independent variable, the process yield index for multivariate linear profiles is then defined as

$$
S_{pki}^T = \frac{1}{3} \Phi^{-1} \left\{ \prod_{k=1}^{v} \left( 2\Phi(3S_{pki,k}) - 1 \right) + 1 \right\}/2,
$$

(8)

where $S_{pki,k}$ denotes the value of $S_{pki}$ for the $k$th quality characteristic. Using the one-to-one correspondence between $S_{pki}^T$ and the process yield, the process yield $P_i^T$ can be established as

$$
P_i^T = \prod_{k=1}^{v} \left[ 2\Phi(3S_{pki,k}) - 1 \right] = 2\Phi(3S_{pki}^T) - 1.
$$

(9)

The overall process yield $TP$ for multivariate linear profiles is defined as the average percentage of conforming parts for all levels of independent variable. Thus, the overall process yield $TP$ is derived by

$$
TP = \frac{1}{n} \sum_{i=1}^{n} P_i^T = \frac{1}{n} \sum_{i=1}^{n} \left[ 2\Phi(3S_{pki}^T) - 1 \right].
$$

(10)
In addition, the overall process yield index for multivariate linear profiles is obtained as

\[
T_{SpkA} = \frac{1}{3} \Phi^{-1} \left( \frac{(TP + 1)}{2} \right) = \frac{1}{3} \Phi^{-1} \left\{ \frac{1}{n} \sum_{i=1}^{n} \left[ 2 \Phi(3S_{pki}^T) - 1 \right] + 1 \right\} / 2. \tag{11}
\]

The index \(T_{SpkA}\) provides an exact measure of the overall process yield when multiple characteristics are mutually independent normality. Various \(T_{SpkA}\) values with their corresponding process yield and non-conformities are shown in Table 1. In practice, a process is called inadequate if \(T_{SpkA} < 1.00\), marginally capable if \(1.00 \leq T_{SpkA} < 1.33\), satisfactory if \(1.33 \leq T_{SpkA} < 1.50\), excellent if \(1.50 \leq T_{SpkA} < 2.00\) and super if \(T_{SpkA} \geq 2.00\).

Using the first order of the Taylor expansion and the Central Limit Theorem, the asymptotic distribution of \(\hat{S}_{pki}^T\) can be shown as a normal distribution (Pearn, Wang, & Yen, 2006; Wang, 2010). That is, we have

\[
\hat{S}_{pki}^T \sim N \left( S_{pki}^T, \frac{1}{36m(\Phi(3S_{pki}^T))^2} \left( \sum_{k=1}^{v} \left( a_{ik}^2 + b_{ik}^2 \right) \left[ \prod_{k=1}^{v} \left( 2\Phi(3S_{pki,k}) - 1 \right)^2 \right] \right) \right).
\]

Therefore, an approximate 100\((1 - \alpha)\)% confidence interval for the process yield index \(\hat{S}_{pki}^T\) is derived by

\[
\left( \hat{S}_{pki}^T - \frac{D}{6\sqrt{m(\Phi(3S_{pki}^T))}} Z_{\alpha/2}, \hat{S}_{pki}^T + \frac{D}{6\sqrt{m(\Phi(3S_{pki}^T))}} Z_{\alpha/2} \right), \tag{12}
\]

where

\[
D = \sqrt{\sum_{k=1}^{v} \left( a_{ik}^2 + b_{ik}^2 \right) \left[ \prod_{k=1}^{v} \left( 2\Phi(3S_{pki,k}) - 1 \right)^2 \right] / \left( 2\Phi(3S_{pki,k}) - 1 \right)^2},
\]

\[
\begin{array}{|c|c|c|}
\hline
T_{SpkA} & Process yield & Non-conformities (ppm) \\
\hline
1.00 & 0.997 300 2039 & 2699.796 \\
1.10 & 0.999 003 1517 & 966.848 \\
1.20 & 0.999 681 7828 & 318.217 \\
1.30 & 0.999 903 8073 & 96.193 \\
1.33 & 0.999 933 9267 & 66.073 \\
1.40 & 0.999 973 3085 & 26.691 \\
1.50 & 0.999 993 2047 & 6.795 \\
1.60 & 0.999 998 4133 & 1.587 \\
1.67 & 0.999 999 4557 & 0.544 \\
1.70 & 0.999 999 6603 & 0.340 \\
1.80 & 0.999 999 9334 & 0.067 \\
1.90 & 0.999 999 9880 & 0.012 \\
2.00 & 0.999 999 9980 & 0.002 \\
\hline
\end{array}
\]

ppm represents parts per million.
\[ \hat{a}_{ik} = \frac{d_{ik}}{\sqrt{2S_{ik}}} \left\{ (1 - \hat{C}_{drk})\phi\left(\frac{1 - \hat{C}_{dpk}}{C_{dpk}}\right) + (1 + \hat{C}_{drk})\phi\left(\frac{1 + \hat{C}_{dpk}}{C_{dpk}}\right) \right\}, \]
\[ \hat{b}_{ik} = \phi\left(\frac{1 - \hat{C}_{drk}}{C_{dpk}}\right) - \phi\left(\frac{1 + \hat{C}_{drk}}{C_{dpk}}\right), \]
and \( \forall k = 1, 2, \ldots, v. \) It should be noted that \( \hat{C}_{drk} \) and \( \hat{C}_{dpk} \) are used instead because they will converge to \( C_{drk} \) and \( C_{dpk}, \) respectively.

Similarly, using the first order of the Taylor expansion and the Central Limit Theorem, the asymptotic distribution of \( T_{S_{pkA}} \) can be shown as a normal distribution; that is, we have

\[ T_{S_{pkA}} \sim N\left( T_{S_{pkA}}, \frac{1}{36n^2m[\phi(3T_{S_{pkA}})]^2} \sum_{i=1}^{n} \left\{ \sum_{k=1}^{v} (\hat{a}_{ik}^2 + \hat{b}_{ik}^2) \left[ \prod_{k=1}^{v} \left(2\Phi(3\hat{S}_{pk,i,k}) - 1\right)^2 \right] \right\} \right). \]

Thus, an approximate \( 100(1 - \alpha)\% \) confidence interval for the overall process yield index \( T_{S_{pkA}} \) is derived by

\[ \left( \frac{T_{S_{pkA}} - \frac{C}{6n\sqrt{m}[\phi(3T_{S_{pkA}})]} Z_{\alpha/2}, \frac{T_{S_{pkA}} + \frac{C}{6n\sqrt{m}[\phi(3T_{S_{pkA}})]} Z_{\alpha/2}}{}} \right), \] (13)

where

\[ C = \sqrt{\sum_{i=1}^{n} \left\{ \sum_{k=1}^{v} (\hat{a}_{ik}^2 + \hat{b}_{ik}^2) \left[ \prod_{k=1}^{v} \left(2\Phi(3\hat{S}_{pk,i,k}) - 1\right)^2 \right] \right\} } \]

2. Multivariate normality: Assume that a process with multiple characteristics is a multivariate normal data. In this case, the principal components analysis method can be applied to this kind of data. Consequently, the new variables (principal components) are mutually independent normality. We assume that \( S \) is the sample variance–covariance matrix. By using the spectral decomposition, we can obtain a matrix \( D = U^T S U, \) where \( D \) is a diagonal matrix. The elements of \( D, \lambda_1, \lambda_2, \ldots, \lambda_v, \) are the eigenvalues of \( S \) and the columns of \( U, u_1, u_2, \ldots, u_v \) are the eigenvectors of \( S. \) So the \( k^{\text{th}} \) principal component is expressed as

\[ PC_k = u_k^T Y, \quad k = 1, 2, \ldots, v \] (14)

where \( Y \) is a \( v \times 1 \) vector of original variables. Thus, we have \( PC = U^T Y \) and \( Y = UPC. \) Then, the target values of \( PC_k \)s are derived by \( T_{PC_k} = u_k^T T, \forall k = 1, 2, \ldots, v. \) Also, the specification limits of \( PC_k \)s can be obtained as

\[ \text{Specification region} = \{(PC_1, PC_2, \ldots, PC_v) | LSL \leq UPC \leq LSL\} \] (15)
Similarly, the sample variance and the sample mean of $PC_k$s can be obtained as

\[
\begin{aligned}
S^2_{PC_k} &= \hat{\lambda}_k, \\
\bar{Y}_{PC_k} &= u_k^T\bar{Y} , \quad \forall \ k = 1, 2, \ldots, v
\end{aligned}
\]  

(16)

As we know, the ratio of each eigenvalue to the summation of all eigenvalues is the proportion of variability associated with each principal component variable. That is, we have $\hat{\lambda}_k / \sum_{k=1}^{v} \hat{\lambda}_k, \quad \forall \ k = 1, 2, \ldots, v$. In this study, we proposed the overall process yield index $T_{SpkA;PC}$ for multivariate linear profiles under multivariate normality, which is defined as

\[
T_{SpkA;PC} = \frac{1}{3} \Phi^{-1} \left[ \frac{(TP_{PC} + 1)}{2} \right] = \frac{1}{3} \Phi^{-1} \left\{ \frac{1}{n} \sum_{i=1}^{n} \left[ 2\Phi(3S^T_{pki;PC}) - 1 \right] + 1 \right\}/2,
\]

(17)

where $TP_{PC} = (1/n) \sum_{i=1}^{n} P_{i;PC} = (1/n) \sum_{i=1}^{n} \left[ 2\Phi(3S^T_{pki;PC}) - 1 \right]$, $P_{i;PC} = 2\Phi(3S^T_{pki;PC}) - 1$, $S^T_{pki;PC} = \sum_{k=1}^{v} \left( \hat{\lambda}_k / \sum_{k=1}^{v} \hat{\lambda}_k \right) S_{pki;PC}$, and $S_{pki;PC}$ represents the process yield index for the $k$th principal component at the $i$th level of independent variable. The values of $T_{SpkA;PC}$ with their corresponding process yield and non-conformities are the same as that of $T_{SpkA}$ (see Table 1). The asymptotic distribution of $\hat{T}_{SpkA;PC}$ can be shown as a normal distribution; that is, we have

\[
\hat{T}_{SpkA;PC} \sim N \left( T_{SpkA;PC}, \frac{1}{36n^2m[\phi(3T_{SpkA;PC})]^2} \sum_{i=1}^{n} [\phi(3S^T_{pki;PC})]^2 \left\{ \sum_{k=1}^{v} \frac{\lambda_k^2}{(\sum_{k=1}^{v} \lambda_k)^2} \times \frac{(a^2_{ik;PC} + b^2_{ik;PC})}{[\phi(3S_{pki;PC})]^2} \right\} \right).
\]

Therefore, an approximate 100(1 − $z$)% confidence interval for the overall process yield index $T_{SpkA;PC}$ is derived by

\[
\left( \frac{\hat{T}_{SpkA;PC} - C_{PC}}{6n\sqrt{m}[\phi(3\hat{T}_{SpkA;PC})]} - Z_{\alpha/2}, \frac{\hat{T}_{SpkA;PC} + C_{PC}}{6n\sqrt{m}[\phi(3\hat{T}_{SpkA;PC})]} + Z_{\alpha/2} \right),
\]

(18)

where

\[
C_{PC} = \sqrt{\sum_{i=1}^{n} [\phi(3S^T_{pki;PC})]^2 \left\{ \sum_{k=1}^{v} \frac{\lambda_k^2}{(\sum_{k=1}^{v} \lambda_k)^2} \times \frac{(a^2_{ik;PC} + b^2_{ik;PC})}{[\phi(3S_{pki;PC})]^2} \right\}},
\]

\[
\hat{a}_{ik;PC} = d_{ik;PC} / \sqrt{2S_{ik;PC}} \left\{ (1 - \hat{C}_{dr_{ik;PC}}) \phi \left( \frac{1 - \hat{C}_{dr_{ik;PC}}}{\hat{C}_{dp_{ik;PC}}} \right) + (1 + \hat{C}_{dr_{ik;PC}}) \phi \left( \frac{1 + \hat{C}_{dr_{ik;PC}}}{\hat{C}_{dp_{ik;PC}}} \right) \right\},
\]

\[
\hat{b}_{ik;PC} = \phi \left( \frac{1 - \hat{C}_{dr_{ik;PC}}}{\hat{C}_{dp_{ik;PC}}} \right) - \phi \left( \frac{1 + \hat{C}_{dr_{ik;PC}}}{\hat{C}_{dp_{ik;PC}}} \right).
\]
It should be noted that \( \tilde{C}_{d_{\text{sf},PC}} \) and \( \tilde{C}_{d_{\text{pf},PC}} \) are used instead because they will converge to \( C_{d_{\text{sf},PC}} \) and \( C_{d_{\text{pf},PC}} \), respectively.

In practice, a process is called inadequate if \( TS_{pkA;PC} < 1.00 \), marginally capable if \( 1.00 \leq TS_{pkA;PC} < 1.33 \), satisfactory if \( 1.33 \leq TS_{pkA;PC} < 1.50 \), excellent if \( 1.50 \leq TS_{pkA;PC} < 2.00 \) and super if \( TS_{pkA;PC} \geq 2.00 \).

**Illustrative examples**

Example 1: We consider the profile model from Ebadi and Amiri (2012) to conduct a simulation study. Assuming that the process is in statistical control, the bivariate simple linear profile model is defined as

\[
\begin{align*}
Y_1 & = 2 + 1x + \varepsilon_1 \\
Y_2 & = 3 + 2x + \varepsilon_2
\end{align*}
\]

where the fixed values of \( x \) are defined as \( x = [2, 4, 6, 8] \). We assume that the error term vector \( [\varepsilon_1, \varepsilon_2] \) is a bivariate normal random vector, where

\[
[\varepsilon_1, \varepsilon_2]^T \sim N_2\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_1^2 & \rho \sigma_1 \sigma_2 \\ \rho \sigma_1 \sigma_2 & \sigma_2^2 \end{bmatrix}\right)
\]

For both regression models, we assume that the lower and upper specification limits depend on the independent variable. The values of the specification limits at each level of \( x \) are provided in Table 2. A simulation study for different values of \( \sigma_1^2, \sigma_2^2, \) and \( \rho \) was generated by 10,000 times. A 95% lower limit of the stated nominal for the coverage rate is obtained as \( \left(0.95 - 1.96 \sqrt{0.05 \times 0.95/10000}\right) \times 100\% = 94.57\% \). In each run, we generated 30 and 100 samples for two simple linear profiles at each level of independent variable. When \( \rho = 0 \), the simulated data is mutually independent normality. When \( \rho \neq 0 \), the simulated data is bivariate normality. All simulation programs were run using R (R Development Core Team, 2013).

According to the simulation results in Tables 3 and 4, we observe the following:

i. As the sample size increases, the estimated value of the process yield index is close to the true value.

ii. The standard deviation of the estimated process yield index decrease as the sample size increases.

iii. The width of the estimated confidence interval decreases as the sample size increases.

iv. The coverage rate is greater than the 95% lower limit of the stated nominal (\( = 94.57\% \)) in most cases as the sample size increases.

**Table 2.** Specification limits for the two response variables in each level of independent variable.

| \( i \) | \( x_i \) | \( LSL \) | \( USL \) | \( Target \) | \( LSL \) | \( USL \) | \( Target \) |
|---|---|---|---|---|---|---|---|
| 1 | 2 | 1 | 8 | 4.5 | 2.5 | 10 | 6.25 |
| 2 | 4 | 3 | 10 | 6.5 | 6.85 | 14.35 | 10.60 |
| 3 | 6 | 5.5 | 12.5 | 9 | 11.25 | 18.75 | 15 |
| 4 | 8 | 7.5 | 14.5 | 11 | 16.25 | 23.75 | 20 |
v. The value of $\rho$ slightly affects the true value of the process yield index, the estimated process yield index and the width of the estimated confidence interval.

Example 2: We consider the profile model from Noorossana et al. (2010). Four simple linear profiles are defined as

\[
\begin{align*}
Y_1 &= -8.5 + 0.87x + \varepsilon_1 \\
Y_2 &= -5.8 + 0.95x + \varepsilon_2 \\
Y_3 &= 3.2 + 1.04x + \varepsilon_3 \\
Y_4 &= 13.6 + 1.09x + \varepsilon_4 
\end{align*}
\]

We assume that the fixed values of $x$ are defined as $x = [50, 140, 230, 350]$ and

\[
[\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4]^T \sim N_4 \left( \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \Sigma = \begin{bmatrix} 30 & 0 & 0 & 0 \\ 0 & 32 & 0 & 0 \\ 0 & 0 & 42 & 0 \\ 0 & 0 & 0 & 44 \end{bmatrix} \right)
\]

In this example, 100 samples for the four simple linear profiles at each level of independent variable were generated using R (R Development Core Team, 2013).
The specification limits of all four characteristics are provided in Table 5. This data set is mutually independent normal. Using the proposed method, the estimated process yield is obtained as $\hat{TP} = 0.9995038$. Consequently, the estimated yield index is obtained as $\hat{TS}_{pk/A} = \frac{1}{3}\Phi^{-1}\left[\Phi\left(1 + 0.9995038\right)\right] = 1.1609$. Using the proposed method, an approximate 95% lower confidence bound for the overall process yield index is obtained as 1.0555. We found that the 95% lower confidence bound of $TS_{pk/A}$ is within the range of [1, 1.33]. It should be noted that this process is marginally capable.

Table 4. Simulation results of the bivariate case with mutually dependent normality ($\rho \neq 0$).

| Simulated case | $m$ | True value of $TS_{pk/A}$ | Mean and SD of $TS_{pk/A}$ | 95% confidence interval of $TS_{pk/A}$ | Bias | Coverage rate (%) |
|----------------|-----|---------------------------|-----------------------------|----------------------------------------|------|------------------|
| $\Sigma = \begin{bmatrix} 0.5 & 0.4 \\ 0.4 & 0.5 \end{bmatrix}$ | 30 | 1.6422 | 1.5077, 0.0132 | [1.2520, 1.7634] | -0.1345 | 84.37 |
| | 100 | 1.6422 | 1.5077, 0.0132 | [1.2520, 1.7634] | -0.0585 | 89.12 |
| $\Sigma = \begin{bmatrix} 0.5 & 0.1 \\ 0.1 & 0.5 \end{bmatrix}$ | 30 | 1.6407 | 1.4948, 0.0901 | [1.2365, 1.7531] | -0.1459 | 87.98 |
| | 100 | 1.6407 | 1.4948, 0.0901 | [1.2365, 1.7531] | -0.0615 | 92.28 |
| $\Sigma = \begin{bmatrix} 0.5 & 0.5 \\ 0.5 & 1 \end{bmatrix}$ | 30 | 1.2498 | 1.1828, 0.0791 | [0.9851, 1.3805] | -0.0670 | 93.41 |
| | 100 | 1.2498 | 1.1828, 0.0791 | [0.9851, 1.3805] | -0.0255 | 95.36 |
| $\Sigma = \begin{bmatrix} 0.5 & 0.2 \\ 0.2 & 1 \end{bmatrix}$ | 30 | 1.2497 | 1.1818, 0.0775 | [0.9853, 1.3782] | -0.0679 | 93.47 |
| | 100 | 1.2497 | 1.1818, 0.0775 | [0.9853, 1.3782] | -0.0254 | 95.46 |
| $\Sigma = \begin{bmatrix} 0.5 & 0.8 \\ 0.8 & 2 \end{bmatrix}$ | 30 | 0.8839 | 0.8576, 0.0543 | [0.7253, 0.9900] | -0.0263 | 96.50 |
| | 100 | 0.8839 | 0.8576, 0.0543 | [0.7253, 0.9900] | -0.0088 | 97.29 |
| $\Sigma = \begin{bmatrix} 0.5 & 0.2 \\ 0.2 & 2 \end{bmatrix}$ | 30 | 0.8839 | 0.8576, 0.0544 | [0.7253, 0.9898] | -0.0263 | 96.49 |
| | 100 | 0.8839 | 0.8576, 0.0544 | [0.7253, 0.9898] | -0.0087 | 97.04 |
| $\Sigma = \begin{bmatrix} 0.5 & 1.2 \\ 1.2 & 4 \end{bmatrix}$ | 30 | 0.6250 | 0.6168, 0.0388 | [0.5195, 0.7142] | -0.0082 | 98.58 |
| | 100 | 0.6250 | 0.6168, 0.0388 | [0.5195, 0.7142] | -0.0025 | 98.56 |
| $\Sigma = \begin{bmatrix} 0.5 & 0.2 \\ 0.2 & 4 \end{bmatrix}$ | 30 | 0.6250 | 0.6168, 0.0389 | [0.5194, 0.7141] | -0.0082 | 98.59 |
| | 100 | 0.6250 | 0.6168, 0.0389 | [0.5194, 0.7141] | -0.0024 | 98.48 |

SD: standard deviation; Bias: estimated value-true value.

"Coverage rate is greater than the lower limit of the stated nominal (= 94.57)."

Table 5. Specification limits for the four response variables in each level of independent variable for Example 2.

| $i$ | $x_i$ | $Y_1$ LSL USL | $Y_2$ LSL USL | $Y_3$ LSL USL | $Y_4$ LSL USL |
|-----|-------|----------------|----------------|----------------|----------------|
| 1   | 50    | 5 150          | 5 105          | 5 110          | 5 140          |
| 2   | 140   | 80 150         | 80 180         | 95 200         | 95 230         |
| 3   | 230   | 155 225        | 155 255        | 185 290        | 185 320        |
| 4   | 350   | 255 325        | 255 355        | 305 410        | 305 440        |

The specification limits of all four characteristics are provided in Table 5. This data set is mutually independent normal. Using the proposed method, the estimated process yield is obtained as $\hat{TP} = 0.9995038$. Consequently, the estimated yield index is obtained as $\hat{TS}_{pk/A} = \left(\frac{1}{3}\right)\Phi^{-1}\left[\Phi\left(1 + 0.9995038\right)\right] = 1.1609$. Using the proposed method, an approximate 95% lower confidence bound for the overall process yield index is obtained as $1.0555$. We found that the 95% lower confidence bound of $TS_{pk/A}$ is within the range of [1, 1.33]. It should be noted that this process is marginally capable.
Conclusion

Process yield plays an important role for evaluating the performance of a manufacturing process. Little work has been done to measure the process yield for multivariate linear profiles. The existing process capability indices (Ebadi & Amiri, 2012) only provided the point estimates. However, engineers are often interested in constructing the confidence interval of the process yield. We present a new method to measure the process yield for multivariate linear profiles under mutually independent normality and multivariate normality. The confidence interval of the overall process yield index is constructed. The lower confidence bound of the overall process yield index can be used to determine whether the process performance meets the present yield requirement leading to a reliable decision. Future studies may include more complex profiles such as non-linear profiles or multiple and polynomial regression profiles. Other methods for the other two types of multivariate linear profiles such as mutually independent non-normality and multivariate non-normality can be addressed in the future studies.

Acknowledgements

The author gratefully acknowledges the referee of this paper who helped to clarify and improve the presentation.

Disclosure statement

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of this article.

Notes on contributor

Fu-Kwun Wang, is a Professor in the Department of Industrial Management at the National Taiwan University of Science & Technology, Taiwan. His fields of interest are reliability engineering, quality control and production management.

References

Amiri, A., Zou, C., & Doroudyan, M.H. (2014). Monitoring correlated profile and multivariate quality characteristics. Quality and Reliability Engineering International, 30, 133–142. doi:10.1002/qre.1483
Boyles, R.A. (1994). Process capability with asymmetric tolerances. Communications in Statistics-Simulation and Computation, 23, 615–643. doi:10.1080/03610919408813190
Chan, L.K., Cheng, S.W., & Spiring, F.A. (1991). A multivariate measure of process capability. International Journal of Modeling and Simulation, 11, 1–6.
Chen, H. (1994). A multivariate process capability index over a rectangular solid tolerance zone. Statistica Sinica, 4, 749–758.
Chen, K.S., Pearn, W.L., & Lin, P.C. (2003). Capability measures for processes with multiple characteristics. Quality and Reliability Engineering International, 19, 101–110. doi:10.1002/qre.513
Ebadi, M., & Amiri, A. (2012). Evaluation of process capability in multivariate simple linear profiles. Scientia Iranica, Transaction E: Industrial Engineering, 19, 1960–1968. doi:10.1016/j.scient.2012.09.010
Ebadi, M., & Shahriari, H. (2013). A process capability index for simple linear profile. *International Journal of Advanced Manufacturing Technology*, 64, 857–865. doi:10.1007/s00170-012-4066-7

Eyvazian, M., Noorossana, R., Saghaei, A., & Amiri, A. (2011). Phase II monitoring of multivariate multiple linear regression profiles. *Quality and Reliability Engineering International*, 27, 281–296. doi:10.1002/qre.1119

Hosseinifard, S.Z., & Abbasi, B. (2012a). Evaluation of process capability indices of linear profiles. *International Journal of Quality and Reliability Management*, 29, 161–176. doi:10.1016/j.ijqr.2012.09.010

Hosseinifard, S.Z., & Abbasi, B. (2012b). Process capability analysis in non-normal linear regression profiles. *Communications in Statistics – Simulation and Computation*, 41, 1761–1784. doi:10.1080/03610918.2011.611313

Kang, L., & Albin, S.L. (2000). On-Line monitoring when the process yields a linear profile. *Journal of Quality Technology*, 32, 418–426.

Kazemzadeh, R.B., Noorossana, R., & Amiri, A. (2010). Phase II monitoring of autocorrelated polynomial profiles in AR(1) processes. *Scientia Iranica, Transaction E: Industrial Engineering*, 17, 12–24.

Kotz, S., & Johnson, N.L. (2002). Process capability indices—a review, 1992–2000. *Journal of Quality Technology*, 34, 1–19.

Kunter, M.H., Nachtsheim, C.J., Neter, J., & Li, W. (2005). *Applied linear statistical models*. Boston: McGraw-Hill.

Lee, J.C., Hung, H.N., Pearn, W.L., & Kueng, T.L. (2002). On the distribution of the estimated process yield index $S_{pk}$. *Quality and Reliability Engineering International*, 18, 111–116. doi:10.1002/qre.450

Noorossana, R., Amiri, A., & Soleimani, P. (2008). On monitoring of autocorrelated linear profiles. *Communications in Statistics—Theory and Methods*, 37, 425–442. doi:10.1080/03610920701653136

Noorossana, R., Eyvazian, M., & Amiri, A. (2010). Phase II monitoring of multivariate multiple linear regression profiles. *Quality and Reliability Engineering International*, 26, 291–303. doi:10.1002/qre.1066

Noorossana, R., Eyvazian, M., & Vaghefi, A. (2010). Phase II monitoring of multivariate simple linear profiles. *Computers and Industrial Engineering*, 58, 563–570. doi:10.1016/j.cie.2009.12.003

Noorossana, R., Saghaei, A., & Amiri, A. (2011). *Statistical analysis of profile monitoring*. New York: Wiley.

Pearn, W.L., Kang, H.Y., Lee, A.H.I., & Liao, M.Y. (2009). Photolithography control in wafer fabrication based on process capability indices with multiple characteristics. *IEEE Transactions on Semiconductor Manufacturing*, 22, 351–356. doi:10.1109/TSM.2009.2024851

Pearn, W.L., Wang, F.K., & Yen, C.H. (2006). Measuring production yield for process with multiple quality characteristics. *International Journal of Production Research*, 44, 4649–4661. doi:10.1080/00207540600589119

R Development Core Team. (2013). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Razavi, H., Asadzadeh, A., & Naseri, Z. (2010). Determination of process capability indices for the slope parameter of a single variable linear profile. *International Journal of Industrial Engineering and Production Management*, 21, 33–43.

Shahriari, H., & Sarrafian, M. (2009). *Assessment of process capability in linear profiles*. Proceedings of the 6th International Industrial Engineering Conference, Tehran, Iran (in Farsi).

Shahriari, H., Hubele, N.F., & Lawrence, F.P. (1995). *A multivariate process capability vector*. Proceedings of the 4th Industrial Engineering Research Conference (pp. 304–309), Institute of Industrial Engineers.

Soleimani, P., Noorossana, R., & Amiri, A. (2009). Simple linear profiles monitoring in the presence of within profile autocorrelation. *Computer and Industrial Engineering*, 57, 1015–1021. doi:10.1016/j.cie.2009.04.005
Appendix

Let $\hat{S}_{pkA} = \frac{1}{3} \Phi^{-1}\left(\frac{1}{n} \sum_{i=1}^{n} \left[ 2\Phi(3\hat{S}_{pk}) - 1 \right] + 1 \right) / 2$, where $\hat{S}_{pk}$ denotes the estimator of $S_{pk}$. If $\hat{S}_{pk} \sim N\left(S_{pk}, \frac{a_i^2 + b_i^2}{36m[\phi(3S_{pk})]^2}\right)$ for $i = 1, 2, ..., n$, where

$$a_i = d_i / \sqrt{2\sigma_i}\left\{ (1 - C_{dr})\phi\left(\frac{1 - C_{dr}}{C_{dp}}\right) + (1 + C_{dr})\phi\left(\frac{1 + C_{dr}}{C_{dp}}\right) \right\},$$
$$b_i = \phi\left(\frac{1 - C_{dr}}{C_{dp}}\right) - \phi\left(\frac{1 + C_{dr}}{C_{dp}}\right),$$

$Z_{a/2}$ is the upper $100(a/2)\%$ point of the standard normal distribution and $\phi$ is the probability density function of the standard normal distribution [15], and all $\hat{S}_{pk}, \forall i = 1, 2, ..., n$ are mutually independent. Thus, the asymptotic distribution of $\hat{S}_{pkA}$ is normal. That is, we have

$$\hat{S}_{pkA} \sim N\left(S_{pkA}, \frac{\sum_{i=1}^{n} (a_i^2 + b_i^2)}{36n^2m[\phi(3S_{pkA})]^2}\right).$$

Proof: We take $n = 2$ for example to derive the asymptotic distribution of $\hat{S}_{pkA}$. By taking the first order of the Taylor expansion, we have

$$\hat{S}_{pkA} = f(S_{pk1}, S_{pk2}) + \sum_{i=1}^{2} \frac{\partial f(S_{pk1}, S_{pk2})}{\partial S_{pk1}}(\hat{S}_{pk} - S_{pk}).$$
Since
\[ E(\hat{S}_{pki}) = S_{pki}, \quad V(\hat{S}_{pki}) = \frac{a_i^2 + b_i^2}{36m[\phi(3S_{pki})]^2} \]
and
\[ \frac{\partial f(S_{pki1}, S_{pki2})}{\partial S_{pki}} = \frac{1}{\phi(3S_{pki})}, \quad \forall i = 1, 2, \]
we have
\[
E(\hat{S}_{pKA}) = E(f(S_{pki1}, S_{pki2})) + \sum_{i=1}^{2} E \left( \frac{\partial f(S_{pki1}, S_{pki2})}{\partial S_{pki}} \right) E(\hat{S}_{pki} - S_{pki})
\]
\[ = f(S_{pki1}, S_{pki2}) = S_{pKA} = \frac{1}{3} \Phi^{-1} \left\{ \frac{1}{2} \sum_{i=1}^{2} (2\Phi(3S_{pki}) - 1) + 1 \right\} / 2 \]
and
\[
V(\hat{S}_{pKA}) = V(f(S_{pki1}, S_{pki2})) + \sum_{i=1}^{2} \left( \frac{\partial f(S_{pki1}, S_{pki2})}{\partial S_{pki}} \right)^2 V(\hat{S}_{pki} - S_{pki})
\]
\[ = \left( \frac{1}{3} \phi(3S_{pki}) \right)^2 \times \frac{a_i^2 + b_i^2}{36m[\phi(3S_{pki})]^2} + \left( \frac{1}{3} \phi(3S_{pki}) \right)^2 \times \frac{a_i^2 + b_i^2}{36m[\phi(3S_{pki})]^2}
\]
\[ = \frac{\sum_{i=1}^{2} (a_i^2 + b_i^2)}{144m[\phi(3S_{pki})]^2}. \]

Since \( \hat{S}_{pKA} \) is equal to a linear combination of two independent normal distributed random variables plus one constant. It implies \( \hat{S}_{pKA} \) has normal distribution. Thus, \( \hat{S}_{pKA} \) has the asymptotic normal distribution with the mean \( S_{pKA} \) and the variance 
\[ \left\{ \sum_{i=1}^{2} (a_i^2 + b_i^2) \right\} / \left\{ 144m[\phi(3S_{pki})]^2 \right\} . \]

Similarly, considering \( n \) variables, the asymptotic distribution of \( \hat{S}_{pKA} \) can be obtained as
\[ \hat{S}_{pKA} \sim N \left( S_{pKA}, \frac{\sum_{i=1}^{n} (a_i^2 + b_i^2)}{36n^2m[\phi(3S_{pki})]^2} \right) . \]