

ESTIMATION OF MULTIVARIATE PROCESS YIELD INDEX S_{pk}^T FOR VARIOUS NON-NORMAL DISTRIBUTIONS USING ANN AND BURR XII DISTRIBUTION

Babar Ali^{*1}, Suboohi Safdar²

^{*1,2}Department of Statistics, University of Karachi, Pakistan

¹babarali.sqc@gmail.com, ²suboohi@uok.edu.pk

DOI: <http://doi.org/10.5281/zenodo.22092163>

Keywords

Multivariate non-normal capability index, Yield index, Burr XII distribution, Artificial Neural Network.

Article History

Received: 05 January 2026

Accepted: 18 February 2026

Published: 04 March 2026

Copyright @Author

Corresponding Author: *

Babar Ali

Abstract

Process capability indices have been used in manufacturing industry to evaluate the ability of processes to meet consumer's expectations. Traditionally, frequently used univariate and multivariate capability indices rely over the assumption of normality and multivariate normality respectively. Practically the underlying process distribution is not always univariate and normal. This gives rise to define capability indices for multivariate processes under non-normal scenario. Several multivariate process capability indices have been defined so far based on normality and non-normality. Among multivariate process capability indices based on normality, S_{pk}^T is a multivariate overall yield index that needs normality of each independent quality characteristic. This paper chooses S_{pk}^T and utilizes artificial neural network (ANN) and the Burr XII distribution approach to estimate S_{pk}^T under multivariate non-normal situations. The method was applied to the data sets under non-normal scenarios. The estimates obtained from proposed method were compared with the results of existing traditional indices under normality by translating the non-normal data sets to normality using suitable transformations. The results suggest that the proposed method successfully estimates S_{pk}^T under multivariate non-normal environments.

1. Introduction

Capability analysis evaluates how well manufacturing processes fulfill their pre-defined requirements. These unitless indices measure the ability of a process to produce output in accordance with the customers' demands. To achieve this goal, the use of process capability indices (PCIs) is one of the extensively utilized tools in capability analysis (Spiring, 1995). Traditionally, PCIs were designed on the assumption of normality. But in practice,

manufacturing processes are not always univariate and follow normal distribution. Therefore, the need of multivariate process capability indices (MPCIs) for non-normal output gave rise to a new stream in literature. An outnumbered PCIs are found in the literature that can be categorize with different criteria, viz, the purpose, number of quality characteristics, normality, tolerance etc. The PCIs defined for normal and non-normal distributions are shortly discussed below:

1.1. Capability indices under normality

A large number of indices have been formulated so far. The, frequently used PCIs for normal outputs are C_p , C_{pk} , C_{pm} and C_{pmk} (Kotz & Johnson, 2002). The first-generation index C_p , proposed by Kane (1986) defined as follows:

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

where USL and LSL are the upper and lower process specification limits respectively and σ is the process standard deviation. C_p is the simplest, only takes the process spread into account but not the shift of the process means. Kane (1986) introduced C_{pk} as a remedy to include the deviation of process from the process mean. C_{pk} can be calculated as:

$$C_{pk} = \min\left(\frac{USL - \mu}{3\sigma}, \frac{\mu - LSL}{3\sigma}\right)$$

where μ is the process mean. C_{pk} was still independent of process target value, therefore Chan *et al.* (1988) developed a more refined version C_{pm} , that takes the process targeting into consideration. C_{pm} is defined as:

$$C_{pm} = \frac{USL - LSL}{6\sqrt{\sigma^2 + (\mu - T)^2}}$$

where T is the process target value. C_{pk} is the modified form of C_p , whereas C_{pm} is that of C_{pk} . Pearn *et al.* (1992) combined these two modifications to propose C_{pmk} , an index more sensitive to departure from target value. It is defined as follows:

$$C_{pmk} = \min\left(\frac{USL - \mu}{3\sqrt{\sigma^2 + (\mu - T)^2}}, \frac{\mu - LSL}{3\sqrt{\sigma^2 + (\mu - T)^2}}\right)$$

Vännman (1995) connected these indices by the following superstructure for univariate normal processes.

$$C_p(u, v) = \frac{d - u|\mu - M|}{3\sqrt{\sigma^2 + v(\mu - T)^2}}; u, v \leq 0$$

where $d = (USL - LSL) \div 2$, $M = (USL + LSL) \div 2$. These classical indices are $C_p(0,0) = C_p$, $C_p(1,0) = C_{pk}$, $C_p(0,1) = C_{pm}$ and $C_p(1,1) = C_{pmk}$.

1.1.1. Process yield index S_{pk}

The indices C_p , C_{pk} , C_{pm} and C_{pmk} are potential indicators, nevertheless they do not provide an exact measure of process yield. Boyles (1994) fixed this problem by suggesting a process yield measure S_{pk} for a univariate, normally distributed, bilateral specification process with μ and σ as true process parameters. It is defined as follows.

$$S_{pk} = \frac{1}{3} \Phi^{-1} \left\{ \frac{1}{2} \Phi \left(\frac{USL - \mu}{\sigma} \right) + \frac{1}{2} \Phi \left(\frac{\mu - LSL}{\sigma} \right) \right\} \quad (1)$$

where USL and LSL respectively be the lower and upper specification limits. $\Phi(\cdot)$ and $\Phi^{-1}(\cdot)$ represent cumulative distribution function and inverse cumulative distribution function of standard normal distribution respectively. The natural estimator of S_{pk} was given by Lee *et al.* (2002) who replaced the parameters by their sample estimates.

$$\hat{S}_{pk} = \frac{1}{3} \Phi^{-1} \left\{ \frac{1}{2} \Phi \left(\frac{USL - \bar{x}}{s} \right) + \frac{1}{2} \Phi \left(\frac{\bar{x} - LSL}{s} \right) \right\}$$

where $\bar{x}$ and s are the sample mean and sample standard deviation respectively.

Lee *et al.* (2002) also evaluated an approximate distribution of $\hat{S}_{pk}$ by Taylor expansion technique. It is defined as:

$$\hat{S}_{pk} = S_{pk} + \frac{1}{6\sqrt{n}(\phi(3S_{pk}))} W$$

$$W = \frac{\sqrt{n}(s^2 - \sigma^2)}{\sigma^2} \left\{ \frac{USL - \mu}{2\sigma} \phi\left(\frac{USL - \mu}{\sigma}\right) + \frac{\mu - LSL}{2\sigma} \phi\left(\frac{\mu - LSL}{\sigma}\right) \right\} \\ - \frac{\sqrt{n}(\bar{x} - \mu)}{\sigma} \left\{ \phi\left(\frac{USL - \mu}{\sigma}\right) - \phi\left(\frac{\mu - LSL}{\sigma}\right) \right\}$$

where $\phi(\cdot)$ is the probability density function of the standard normal distribution. The statistic W is asymptotically normally distributed with mean zero and variance $a^2 + b^2$ where a and b can be represented as the functions of C_p and C_a , the precision and accuracy indices successively.

$$a = \{3C_p(2 - C_a)\phi(3C_p(2 - C_a)) + 3C_pC_a\phi(3C_pC_a)\}/\sqrt{2}$$

$$b = \phi(3C_p(2 - C_a)) - \phi(3C_pC_a)$$

where $C_a = 1 - |\mu - m|/d$ and $d = (USL - LSL)/2$.

Thus natural estimator $\hat{S}_{pk}$ is asymptotically distributed with mean S_{pk} and variance $(a^2 + b^2)/36n[\phi(3S_{pk})]^2$.

For a normally distributed quality characteristic X , S_{pk} and respective process yield are connected with a one-to-one correspondence which is given in equation (2).

$$\text{Process Yield} = p = P(LSL \leq x \leq USL) = 2\Phi(3S_{pk}) - 1 \quad (2)$$

Some values of S_{pk} with their corresponding yields are tabulated in table 1.

Table 1: S_{pk} values and their corresponding process yields.

S_{pk}	Process yield
1.00	0.997 300 204
1.24	0.999 800 777
1.35	0.999 933 927
1.50	0.999 993 205
1.67	0.999 999 456
2.00	0.999 999 998

1.1.2. The overall yield index S_{pk}^T

The overall yield index S_{pk}^T is an extension of S_{pk} in a multivariate scenario. Chen *et al.* (2003) defined an overall yield index S_{pk}^T for q independent normally distributed quality characteristics as:

$$S_{pk}^T = \frac{1}{3} \Phi^{-1} \left[\frac{1}{2} \left\{ \prod_{j=1}^q (2\Phi(3S_{pkj}) - 1) + 1 \right\} \right]; j = 1, 2, \dots, q \quad (3)$$

where S_{pkj} is the process yield index of j^{th} quality characteristic.

1.2. Process capability indices under non-normality

There are various approaches being used to estimate PCIs for non-normal processes. Probability plot (Bissell, 1994) is a commonly accepted approach that takes into count the upper and lower 0.135% points of a non-normal process output. The use of transformation is another approach. In this approach classical PCIs are applied after transforming the process data into normality using suitable transformation.

Some of these transformations include Johnson transformation (Johnson, 1949), Box and Cox transformation (Box & Cox, 1964), square root transformation (Somerville and Montgomery, 1996), best root transformation (Niaki and Abbasi, 2007). Another approach is the use of quantiles of non-normal outputs. Clement (1989) used this approach to propose C_p and C_{pk} for non-normal processes as below:

$$C_p = \frac{USL - LSL}{X_{99.865} - X_{0.135}}$$

$$C_{pk} = \min\left(\frac{USL - X_{50}}{X_{99.865} - X_{50}}, \frac{X_{50} - LSL}{X_{50} - X_{0.135}}\right)$$

where X_p is the $p\%$ percentile of the distribution of quality characteristic x . Clement's method was used by Pearn and Kotz (1994) to reform C_{pm} and C_{pmk} for non-normal outputs.

These reformed indices are:

$$C_{pm} = \frac{USL - LSL}{6 \cdot \sqrt{\left(\frac{x_{99.865} - x_{0.135}}{6}\right)^2 + (x_{50} - T)^2}}$$

$$C_{pmk} = \min\left(\frac{USL - x_{50}}{3 \cdot \sqrt{\left(\frac{x_{99.865} - x_{50}}{3}\right)^2 + (x_{50} - T)^2}}, \frac{x_{50} - LSL}{3 \cdot \sqrt{\left(\frac{x_{50} - x_{0.135}}{3}\right)^2 + (x_{50} - T)^2}}\right)$$

Liu and Chen (2006) revised Clement's method to proposed PCIs for non-normal distributions. Another method is proportion of nonconforming items (PNC). Castagliola (1996) established PNC method by relating proportion of nonconforming units with process capability. Bai and Choi (1997) used weighted variance method to use non-normal PCIs. As far as multivariate processes are concerned, Wang and Du (2000) were the pioneer to propose first multivariate capability index (MPCI) for multivariate non-normal processes using principal component analysis. Wang (2006) used distance approach for non-normal MPCI. Ahmad *et al.* (2009) have defined MPCI using Mahalanobis distance and Das and Dwivedi (2014) utilized multivariate g and h distribution. Extended Burr XII distribution was used by Takahashi (1965) to define MPCI. Pan *et al.* (2016) used multivariate weighted standard deviation approach (MWSD). Yeh and Chen (2001) and Polansky (2001) defined distributional assumption free MPCIs called nonparametric MPCIs. Bernardo and Irony (1996) adopted Bayesian approach to evaluate

capability of univariate and multivariate non-normal processes.

1.3. The Burr distribution

In 1942, Burr (1942) developed a number of useful cumulative frequency functions. The Burr XII distribution is one of them and is widely used in quality and reliability. It encompasses a variety of skewness and kurtosis, that can be utilized to fit almost any unimodal data. Burr XII distribution has the relationship with many other statistical distributions with selection of various combinations of c and k (Zimmer *et al.*, 1963). These parameters can be set to fit the Normal, Gamma, Beta, Weibull, Log-normal, and Extreme value type-I distributions. For example, Burr with $c = 4.85437$ and $k = 6.22665$ estimates normal distribution. An extensive discussion for the usage of Burr distribution can be found in Wang *et al.* (1996), Zimmer *et al.* (1998) and Mousa & Jaheen (2002).

The cumulative distribution function *cdf* of Burr distribution is

$$F(x) = \begin{cases} 1 - (1 - x^c)^{-k} & ; x \geq 0 \quad c, k > 0 \\ 0 & ; x < 0 \end{cases}$$

Burr (1942;1973) has constructed several tables of expected mean values, standard deviations, skewness and kurtosis coefficients for the Burr XII distribution (See tables of Burr (Chou *et al.*, 1998; Castagliola, 1996)). These tables enable

users to make a standardized transformation between a Burr variate and another random variate.

Suppose a random variable Y follows a certain distribution with mean μ_Y and standard

deviation σ_T . Suppose a random variable B follows Burr distribution with mean μ_B , and standard deviation, σ_B . We want to approximate

$$\frac{B - \mu_B}{\sigma_B} = \frac{Y - \mu_Y}{\sigma_Y}$$

or equivalently

$$Y = \sigma_Y \left(\frac{B - \mu_B}{\sigma_B} \right) + \mu_Y$$

The *cdf* of the Burr distribution in terms of Y can be written as:

$$F(y) = P\{Y \leq y\} = 1 - \left[1 + \left(\frac{y-M}{S} \right)^c \right]^{-k}; M \leq y < \infty, M, S, c, k > 0 \quad (4)$$

where $S = \sigma_Y/\sigma_B$ and $M = \mu_Y - S\mu_B$.

1.4. Artificial Neural Network (ANN)

Since the 1980s, artificial neural networks (ANNs) have been a widely explored topic in artificial intelligence, originated in 1943 with an outclass NN. The magnitude of the drift of variance was

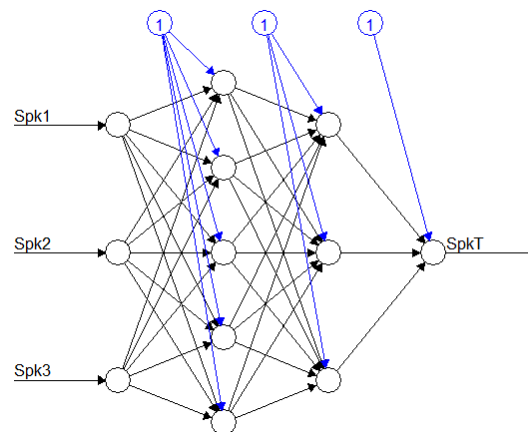


Figure 1: Trained NN to estimate S_{pk}^T using $S_{pkj}; j = 1, 2, 3$.

work of McCulloch & Pitts (1943). ANNs are extensively parallel computing mechanisms mimicking human brain. Like biological neurons, NNs model and solve a variety of problems involving prediction, approximation, pattern recognition, clustering, and categorization. A NN consists of an architecture, weights, and transfer functions. The combination of these components determines the overall output for a set of given input variables. ANNs are also spreading their roots in statistical process control (SPC). A NN was used to indicate the deviation of the process mean for univariate processes (Pugh, 1991). Hwang and Hubele (1993) used a NN to indicate unnatural patterns on control charts. In multivariate processes, out of control causes were studied by Martin and Morris (1995) using fuzzy

the distribution of Y by the distribution of B . This is done by equating both standardized random variables as below.

detected by Chang and Ho (1999) using a NN. A multilayer perceptron was used for classification of pattern in case of a signal in control charts for multivariate processes (Niaki *et al.*, 2005). A NN was used in capability analysis for non-normal outputs by Abbasi (2009). Zan *et al.* (2020) used convolutional NNs for identifying complex patterns automatically. Long short-term memory neural networks were used for fault detection and prediction (Jiang *et al.* 2019). In 2025, Abdolahi *et al.* (2025) used machine learning approach to calculate capabilities of additive manufacturing processes.

2. Proposed methodology

2.1. Formation of neural network

2.1.1. Simulating training and testing data sets

Suppose a process has q non-normal independent quality characteristics with individual yield indices $S_{pkj}; j = 1, 2, \dots, q$ and their overall yield index S_{pk}^T . We want to train and test a NN to estimate the output variable S_{pk}^T using q input variables $S_{pkj}; j = 1, 2, \dots, q$. To simulate data set (will be split into training and testing data sets later), we generated $S_{pkj}; j = 1, 2, \dots, q$ between 0 and 2 using uniform distribution and their corresponding S_{pk}^T was recorded using equation (3). The process was repeated over and over to have N such instances. This data set would have a large variety of processes having an outnumber combinations of $S_{pkj}; j = 1, 2, \dots, q$ and S_{pk}^T . In narrow this study we used $q = 3$ and $N = 1000$. The entire data set was split into 80% and 20% instances dedicated to training and testing data respectively.

2.1.2. Training and testing NN

After the selection of input and output variables, the next most important step in training NN is to decide number of hidden layers and number of neurons in each layer. Unfortunately, there is no exact solution to this question (Mas & Flores, 2008; Stathakis 2009). Traditionally this topology is done by trial and error or sometimes by heuristics followed by trial and error, but still, none of these methods guarantees the best solution. To decide these numbers, we used hyperparameter tuning for optimizing NN performance. This tuning finds the best combination of layer sizes, activation functions, and learning parameters. This was carried out in R language (R version 4.5.1) using "train()" function in "caret" package, providing a feasible systematic evaluation of various hyperparameter settings. For a grid of 3 layers with 0 to 5 neurons in each, we obtained the optimal tuning at 2 numbers of layers with 5 and 3 neurons respectively. The trained model is shown in figure 1.

The testing phase of trained NN gave evaluation metrices as MSE = 0.000748, MAE = 0.022286

and $R^2 = 99.4781\%$. Suppose a process has true $S_{pk1} = S_{pk2} = S_{pk3} = 1$ and $S_{pk}^T = 0.88$, the trained NN estimates $\hat{S}_{pk}^T = 0.85384$ for the process.

2.2. Estimating individual yield indices S_{pkj}

The first step in estimating S_{pk}^T for a process using NN is to estimate S_{pkj} for its non-normal quality characteristics. This would be done by using Burr XII distribution. The steps 1 to 4, given below, explain simulating a non-normal sample, setting its true process yield, setting specification limits and finding true S_{pk} . Whereas, the steps 5 to 9 show how this true S_{pk} is estimated using Burr XII distribution.

1. Randomly selected a process yield p from 0 to 1.

2. Using equation (2), calculate $S_{pk} = \frac{1}{3} \Phi^{-1} \left(\frac{p+1}{2} \right)$

3. Draw a random sample x of size n from a non-normal distribution.

4. To ensure process yield p generated in step 1, calculate process specification limits as follows:

$LSL = \left(\frac{1-p}{2} \right)$ percentile of distribution of x

$USL = \left(1 - \frac{1-p}{2} \right)$ percentile of distribution of x

5. From the generated sample, calculate $\bar{x}$, s , sample skewness a_3 and sample kurtosis a_4 .

6. With reference to a_3 and a_4 , find Burr parameters c, k, μ and σ using Burr table (Burr, 1973).

7. Find $F(USL)$ and $F(LSL)$ using equation (4).

8. Find $\hat{p} = F(USL) - F(LSL)$

9. Estimate $\hat{S}_{pk} = \frac{1}{3} \Phi^{-1} \left(\frac{\hat{p}+1}{2} \right)$

2.3. Estimating overall yield index S_{pk}^T

To estimate overall yield index S_{pk}^T using NN, first confirm the independence among q quality characteristics. If independence is not met, regenerate sample (Repeat all 9 steps given) as many times as to have q independent samples. Finally feed $\hat{S}_{pkj}; j = 1, 2, \dots, q$ into NN (trained

in section 2.1.2) to get an estimated overall yield index $\hat{S}_{pk}^T$.

3. Illustrative examples

This section contains three examples. Examples 1 and 2 use two real world data sets to estimate univariate S_{pk} using Burr method. In Example 3, a bivariate non-normal process and a tri-variate non-normal process were simulated for the estimation of overall yield index S_{pk}^T for each process using Burr distribution and NN

Example 1.

A firm makes and delivers various kinds of rubber edges to speaker driver units making factories. The rubber edge is one of the main components of a standard (woofer) driver unit that affects the quality of sound of the unit. The quality characteristic of our interest is the weight of the rubber edge x in grams. The specification limits of the quality characteristic are $(USL, LSL) = (8.46, 8.94)$. For acceptance of edges by the manufacturers, the weights should lie within the specification limits. This data set was used by K.S. Chen & W.L. Pearn (1997) and tabulated in table 2.

Table 2: The weights of 100 rubber edges in grams.

8.61	8.81	8.72	8.69	8.65	8.64	8.68	8.74	8.68	8.67	8.73	8.61	8.84
8.64	8.68	8.98	8.70	8.74	8.75	8.66	9.00	8.64	8.70	8.71	8.75	8.67
8.53	8.74	8.59	8.69	8.70	9.03	8.83	8.87	8.79	8.68	8.81	8.60	8.64
8.76	8.71	8.71	8.67	8.67	8.68	8.69	8.74	8.80	8.59	8.71	8.53	8.74
8.68	8.55	8.73	8.67	8.71	8.73	8.67	8.68	8.69	8.74	8.80	8.70	8.59
8.55	8.71	8.74	8.70	8.62	8.61	8.79	8.69	8.68	8.77	8.54	8.71	8.69
8.66	8.72	8.81	8.63	8.78	8.64	8.66	8.63	8.71	8.99	8.76	8.64	8.81
8.67	8.71	8.63	8.74	8.67	8.69	8.69	8.68	8.70				

Shapiro-Wilk normality test for the data shows $W = 0.90824$ with p -value of 3.526×10^{-06} , resulting in failure of normality. Now traditionally transforming the data to normality using Johnson transformation. By fitting the best transformation type SU, the transformed variable z is

$$z = -0.229663 + 0.901228 * \operatorname{asinh}((x - 8.68123)/0.0476708)$$

Using equation (1) for z , we get $S_{pk} = 0.701$. The sample data yield $skewness = 1.172046$ and $kurtosis = 5.794697$, that correspond to

$c = 430.089604$, $k = 0.457211$, $\mu_B = 1.003780$ and $\sigma_B = 0.006369$. After executing steps from 5 to 9 stated above, we get $\hat{S}_{pk} = 0.772$ for the rubber edges making process.

Example 2.

The sample in table 3 shows the sizes of ball wire bonding for connection from the integrated circuit apparatus to the lead frame (Leiva *et al.*, 2014).

Table 3: Sizes (in millimeters) of ball wire bonding.

2.891	4.035	4.495	2.890	2.312	3.158	5.228	3.334	5.896	5.639
3.842	1.590	1.954	1.842	0.680	2.752	1.301	2.260	0.889	2.381
0.619	2.788	1.050	3.750	3.508	6.123	6.549	5.954	2.207	4.417
4.805	1.516	2.227	2.797	1.636	1.066	0.940	4.101	4.542	1.295
1.770	3.492	5.706	3.722	6.644	2.472	1.383	4.494	1.694	2.892
2.111	3.591	2.093	3.222	2.891	2.582	0.665	3.234	1.102	1.083
1.508	1.811	2.803	6.659	0.923	6.229	3.177	2.333	1.311	4.419
2.495	0.921	4.061	9.725	1.600	4.281	3.360	1.131	1.618	4.489

3.696	1.982	2.413	5.480	1.992	2.573	1.845	4.620	6.221	1.694
4.882	1.380	3.982	2.260	2.366	2.899	3.782	2.336	1.175	3.055

The process specifications are $LSL = 0.5$ mil and $USL = 8.0$ mil ($1\text{mil} = 1/1000\text{in} = 0.0254\text{mm}$). The normality was rejected by Anderson Darling test producing a p -value of 0.000291. Using Burr distribution with $c = 1.600580$, $k = 75.150186$, $\mu_B = 0.060742$ and $\sigma_B = 0.039208$, the estimated yield index $\hat{S}_{pk}$ is found to be 0.809. For result validation, the data set was transformed to normality using Johnson transformation that fits SB distribution. The transformed variable z is given below:

$$z = 1.50948 + 1.12161 * \log((x - 0.209985)/(12.2927 - x))$$

Applying equation (1) for z , we get $S_{pk} = 0.790$.

Example 3.

Example 3 discusses two processes. The first process contains three quality characteristics,

$x_1 \sim \text{Gamma}(3,1)$, $x_2 \sim \text{Gamma}(16,3)$ and $x_3 \sim \text{Gamma}(9,4)$ having true individual yield indices as $S_{pk1} = 2/3$, $S_{pk2} = 1$ and $S_{pk3} = 1/3$ respectively. The actual overall process yield index of the process is $S_{pk}^T = 0.311$ using equation (3). To estimate overall yield index, we begin with drawing random samples of size 1000 each from Gamma distribution with the given parameters. After simulation, first checking mutually independence assumption among quality characteristics according to Chen *et al.* (2003). Table 4 shows Pearson correlation test result for each pair of quality characteristics.

Table 4: Pearson correlation test results for pairs of quality characteristics in example 3.

Pair	Correlation	P-value	Interpretation
x_1 vs x_2	-0.0146	0.6450680	Independent
x_1 vs x_3	-0.0496	0.1170202	Independent
x_2 vs x_3	-0.0041	0.8979120	Independent
y_2 vs y_2	-0.1259	0.2118	Independent

The p -value for each pair exceeds 0.05, resulting in mutually independence of all quality characteristics. Using Burr method, the estimated individual process yields are $\hat{S}_{pk1} = 0.705$, $\hat{S}_{pk2} = 1.077$ and $\hat{S}_{pk3} = 0.345$ (see table 5). While these estimates were fed into the neural

network, estimated overall yield index was found to be $\hat{S}_{pk}^T = 0.318$.

The second process is bivariate having independent quality characteristics following Weibull distribution such that $y_1 \sim W(1.5,2)$ and $y_1 \sim W(5,6)$ (See table 2). The actual overall yield index 0.321, whereas using proposed method, it is estimated as 0.335 (See table 5).

Table 5: Summary of processes in example 3.

	x_1	x_2	x_3	y_1	y_2
LSL	0.5962877	2.199725	1.514135	0.5666847	2.74937
USL	7.348429	10.22959	2.986285	2.99506	7.711297
p	0.9544997	0.9973002	0.6826895	0.7	0.95
S_{pkj}	0.67	1.00	0.33333	0.3454778	0.6533213
Skewness	1.148517	0.4670957	0.6705799	1.392258	-0.1989633
Kurtosis	4.831017	3.361826	3.768637	6.697235	2.979756
c	1.723358	3	2.501004	1.845063	3.704710

	x_1	x_2	x_3	y_1	y_2
k	11.539059	6	6.355901	4.984891	7.243856
μ_B	0.224684	0.501498	0.443594	0.407011	0.541973
σ_B	0.142462	0.201715	0.207155	0.263012	0.173700
$\hat{S}_{pkj}$	0.7054063	1.077355	0.344678	0.3664854	0.6590968
S_{pk}^T	0.311	*	*	0.321	*
$\widehat{S}_{pk}^T$	0.318	*	*	0.335	*

4. Result

Process capability indices play a vital role in evaluating process performance. Initially PCIs were designed for univariate normal processes. Later these were extended to univariate non-normal, multivariate normal and multivariate non-normal processes. S_{pk}^T is a normality based overall yield index and this study provides a comparative study to estimate this index under multivariate non-normal scenario using Burr XII distribution and artificial neural network. Using Burr distribution, individual yield index for each quality characteristic is estimated and NN combined these indices to estimate overall yield index. The method was elaborated using examples and their results were compared with existing indices. The results show that the combination of Burr XII distribution and ANN successfully estimated overall yield index while the processes are multivariate non-normal.

5. Recommendations

The overall yield index estimated in this study is conventionally based on normality of each individual quality characteristic. The quality characteristics are normal but mutually independent. The approach used in this study may be extended to multivariate non-normal processes having covariance structure among quality characteristics. Assessing capabilities of multivariate non-normal processes using NN will be a milestone in the literature.

Code availability

The custom R scripts developed for this study are not publicly archived but may be made available by the corresponding author upon reasonable request.

REFERENCES

- Abbasi, B. (2009). A neural network applied to estimate process capability of non-normal processes. *Expert Systems with Applications*, 36(2), 3093–3100.
- Abdollahi, A., Soroush, H., & Khodaygan, S. (2025). Process capability analysis of additive manufacturing process: A machine learning-based predictive model. *Rapid Prototyping Journal*, 31(4), 724–741.
- Ahmad, S., Abdollahian, M., Zeephongsekul, P., & Abbasi, B. (2009). Multivariate non-normal process capability analysis. *International Journal of Advanced Manufacturing Technology*, 44(7–8), 757–765.
- Bai, D. S., & Choi, S. S. (1997). *Process capability indices for skewed populations* [Master's thesis]. Department of Industrial Engineering, Korea Advanced Institute of Science and Technology.
- Bernardo, J. M., & Irony, T. Z. (1996). A general multivariate Bayesian process capability index. *The Statistician*, 45(3), 487–502.
- Bissell, D. (1994). *Statistical methods for SPC and TQM*. Chapman and Hall.
- Box, G. E. P., & Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society: Series B (Methodological)*, 26(2), 211–243.
- Boyles, R. A. (1994). Process capability with asymmetric tolerances. *Communications in Statistics—Simulation and Computation*, 23(3), 615–635.
- Burr, I. W. (1942). Cumulative frequency functions. *Annals of Mathematical Statistics*, 13(2), 215–232.

- Burr, I. W. (1973). Parameters for a general system of distributions to match a grid of α_3 and α_4 . *Communications in Statistics*, 2(1), 1-21.
- Castagliola, P. (1996). Evaluation of non-normal process capability indices using Burr's distributions. *Quality Engineering*, 8(4), 587-593.
- Chan, L. K., Spiring, F. A., & Xiao, H. (1988). An OC curve approach for analyzing the process capability index C_{pm} (Technical Report). Department of Statistics, University of Manitoba.
- Chang, S. I., & Ho, E. S. (1999). A two-stage network approach for process variance change detection and classification. *International Journal of Production Research*, 37(8), 1581-1599.
- Chen, K. S., & Pearn, W. L. (1997). An application of non-normal process capability indices. *Quality and Reliability Engineering International*, 13(6), 355-360.
- Chen, K. S., Pearn, W. L., & Lin, P. C. (2003). Capability measures for processes with multiple characteristics. *Quality and Reliability Engineering International*, 19(2), 101-110.
- Chou, Y. M., Polansky, A. M., & Mason, R. L. (1998). Transforming non-normal data to normality in statistical process control. *Journal of Quality Technology*, 30(2), 133-141.
- Clements, J. A. (1989). Process capability calculations for non-normal distributions. *Quality Progress*, 22(12), 95-100.
- Das, N., & Dwivedi, P. S. (2014). Multivariate process capability index: A review and some results. *Economic Quality Control*, 28(2), 151-166.
- Hwarng, H. B., & Hubele, N. F. (1993). X-bar control chart pattern identification through efficient off-line neural network training. *IIE Transactions*, 25(3), 27-40.
- Jiang, C., Jin, P., Kang, Y., & Ye, C. (2019). Fault diagnosis and prediction method of SPC for engine block based on LSTM neural network. In *Intelligent robotics and applications* (pp. 483-492). Springer, Cham.
- Johnson, N. L. (1949). Systems of frequency curves generated by methods of translation. *Biometrika*, 36(1-2), 149-176.
- Kane, V. E. (1986). Process capability indices. *Journal of Quality Technology*, 18(1), 41-52.
- Kotz, S., & Johnson, N. L. (2002). Process capability indices—A review, 1992-2000. *Journal of Quality Technology*, 34(1), 2-19.
- Lee, J. C., Hung, H. N., Pearn, W. L., & Kueng, T. L. (2002). On the distribution of the estimated process yield index Sp_k . *Quality and Reliability Engineering International*, 18(2), 111-116.
- Leiva, V., Marchant, C., Saulo, H., Aslam, M., & Rojas, F. (2014). Capability indices for Birnbaum-Saunders processes applied to electronic and food industries. *Journal of Applied Statistics*, 41(9), 1881-1902.
- Liu, P. H., & Chen, F. L. (2006). Process capability analysis of non-normal process data using the Burr XII distribution. *International Journal of Advanced Manufacturing Technology*, 27(9-10), 975-984.
- Martin, E. B., & Morris, A. J. (1995). Multivariate statistics and neural networks in process fault detection. *IEE Colloquium (Digest)*, 1995(79), 1/1-1/8.
- Mas, J. F., & Flores, J. J. (2008). The application of artificial neural networks to the analysis of remotely sensed data. *International Journal of Remote Sensing*, 29(3), 617-663.
- McCulloch, W. S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, 5(4), 115-133.

- Mousa, M. A., & Jaheen, Z. F. (2002). Statistical inference for the Burr model based on progressively censored data. *Computers & Mathematics with Applications*, 43(10-11), 1441-1449.
- Niaki, S. T. A., & Abbasi, B. (2005). Fault diagnosis in multivariate control charts using artificial neural networks. *Quality and Reliability Engineering International*, 21(7), 825-840.
- Niaki, S. T. A., & Abbasi, B. (2007). Skewness reduction approach in multi-attribute process monitoring. *Communications in Statistics-Theory and Methods*, 36(12), 2313-2325.
- Pan, J. N., Li, C. I., & Shih, W. C. (2016). New multivariate process capability indices for measuring the performance of multivariate processes subject to non-normal distributions. *International Journal of Quality & Reliability Management*, 33(1), 42-61.
- Pearn, W. L., & Kotz, S. (1994). Application of Clements' method for calculating second and third generation process capability indices for non-normal Pearsonian populations. *Quality Engineering*, 7(1), 139-145.
- Pearn, W. L., Kotz, S., & Johnson, N. L. (1992). Distributional and inferential properties of process capability indices. *Journal of Quality Technology*, 24(4), 216-231.
- Polansky, A. M. (2001). A smooth nonparametric approach to multivariate process capability. *Technometrics*, 43(2), 199-211.
- Pugh, G. A. (1991). A comparison of neural networks to SPC charts. *International Journal of Production Research*, 29(2), 253-255.
- Somerville, S., & Montgomery, D. C. (1996). Process capability indices and non-normal distributions. *Quality Engineering*, 19(2), 305-316.
- Spiring, F. A. (1995). Process capability: A total quality management tool. *Total Quality Management*, 6(4), 377-387.
- Stathakis, D. (2009). How many hidden layers and nodes? *International Journal of Remote Sensing*, 30(8), 2133-2147.
- Takahashi, K. (1965). Note on the multivariate Burr's distribution. *Annals of the Institute of Statistical Mathematics*, 17(1), 257-260.
- Vännman, K. (1995). A unified approach to capability indices. *Statistica Sinica*, 5(2), 805-820.
- Wang, F. K. (2006). Quality evaluation of a manufactured product with multiple characteristics. *Quality and Reliability Engineering International*, 22(2), 225-236.
- Wang, F. K., & Du, T. C. T. (2000). Using principal component analysis in process performance for multivariate data. *Omega*, 28(2), 185-194.
- Wang, F. K., Keats, J. B., & Zimmer, W. J. (1996). The maximum likelihood estimation of the Burr XII parameters with censored and uncensored data. *Microelectronics Reliability*, 36(3), 359-362.
- Yeh, A. B., & Chen, H. (2001). A nonparametric multivariate process capability index. *International Journal of Modelling and Simulation*, 21(4), 218-223.
- Zan, T., Liu, Z., Wang, H., & others. (2020). Control chart pattern recognition using the convolutional neural network. *Journal of Intelligent Manufacturing*, 31(3), 703-716. <https://doi.org/10.1007/s10845-019-01473-0>
- Zimmer, W. J., & Burr, I. W. (1963). Variables sampling plans based on non-normal populations. *Industrial Quality Control*, 19(6), 18-36.
- Zimmer, W. J., Keats, J. B., & Wang, F. K. (1998). The Burr XII distribution in reliability analysis. *Journal of Quality Technology*, 30(4), 2-19.