

Evaluating capability of a bivariate zero-inflated poisson process

Surajit Pal¹, Susanta Kumar Gauri^{2*}

¹ SQC & OR Unit, Indian Statistical Institute, 110, N. Manickam Road, Chennai-600029, INDIA

² SQC & OR Unit, Indian Statistical Institute, 203, B. T. Road, Kolkata-700108, INDIA

*Corresponding Author: e-mail: susantagauri@hotmail.com, Tel. No.: 091-033-2575-3350; Fax No.: 091-033-2577-6042
ORCID iDs: <https://orcid.org/0000-0001-6066-5782> (Pal), <https://orcid.org/0000-0003-4654-7696> (Gauri)

Abstract

A zero-inflated Poisson (ZIP) distribution is commonly used for modelling zero-inflated process data with single type of defect, and for developing appropriate tools for instituting statistical process control of manufacturing processes. However, in reality, such manufacturing scenarios are very common where more than one type of defect can occur. For example, occurrences of defects like solder short circuits (shorts) and absence of solder (skips) are very common on printed circuit boards. In literature, different forms of bivariate zero-inflated Poisson (BZIP) distributions are proposed, which can be used for modelling the manufacturing scenarios where two types of defects can occur. Control charts are designed for monitoring for such processes using BZIP models. Although evaluation of capability is an integral part of statistical process control of a manufacturing process, researchers have given very little effort on this aspect of zero-inflated processes. Only a few articles attempted to evaluate the capability of a univariate zero-inflated process and no work is reported on evaluating capability of a bivariate zero-inflated process. In this paper, a methodology for measuring capability of a bivariate zero-inflated process is presented. The proposed methodology is illustrated using two case studies.

Keywords: Count data; process capability index; zero-inflated process; bivariate zero-inflated count data; bivariate zero-inflated poisson (BZIP) distribution.

DOI: <http://dx.doi.org/10.4314/ijest.v14i1.2>

Cite this article as:

Pal S., Gauri S.K. 2022. Evaluating capability of a bivariate zero-inflated poisson process. *International Journal of Engineering, Science and Technology*, Vol. 14, No. 1, pp. 10-20. doi: 10.4314/ijest.v14i1.2

Received: February 27, 2022; Accepted: March 7, 2022; Final acceptance in revised form: March 7, 2022

1. Introduction

Count data with excessive number of zero counts are quite common in many manufacturing processes as well as in many other non-manufacturing scenarios. For modelling such zero-inflated count data, several zero-inflated models such as zero-inflated Poisson (ZIP) distribution (Xie and Goh, 1993; Xie et al., 2001), zero-inflated negative binomial (ZINB) distribution (Chaney et al., 2013), zero-inflated generalized Poisson (ZIGP) distribution (Wagh and Kamalja, 2018) are developed by researchers over the last three decades. Among them, the most popular zero-inflated model is zero-inflated Poisson (ZIP) distribution where it is assumed that the defects occur due to random shock with certain probability Ω and those defects follow a Poisson distribution with parameter λ (Xie and Goh, 1993). More detailed information on the models and monitoring of univariate zero-inflated processes are available in Mahmood and Xie (2019) and Wagh and Kamalja (2018).

There are many high quality manufacturing processes where more than one type of defects may occur due to several reasons. For example, occurrences of defects like solder short circuits (shorts) and absence of solder (skips) are common on printed circuit boards (PCBs) (Li et al.; 1999); occurrences of LED mounting errors and soldering errors are observed in the soldering process of

LED mounting on the PCBs (He et al.; 2012). In a non-manufacturing scenario, Fatahi et al. (2012) have investigated the presence of number of particles and number of microorganisms in the environmental air of a sterilization process in a pharmaceutical factory.

When multiple types of defects are present in a process, it is essential to utilize multivariate distributions for modelling multivariate count data instead of using univariate distributions for modelling each type of defects separately. If the different types of defects are correlated with each other, it becomes more important to model such multivariate count data with multivariate distributions. Bivariate Poisson distribution, first introduced by Campbell (1938) and later developed by Holgate (1964), is often used for modelling correlated bivariate count data. The multivariate Poisson distribution was developed by extending the form of bivariate Poisson distribution for modelling multivariate count data (Krishnamoorthy, 1951). Famoye and Consul (1995) presented a bivariate generalized Poisson distribution with some applications.

For zero-inflated bivariate count data, Li et al. (1999) proposed two different types of BZIP models and extended them further for multivariate zero-inflated Poisson (MZIP) models. The BZIP model, proposed by Li et al. (1999), is a mixture of a bivariate Poisson, two univariate Poisson and a point mass at (0,0). Fatahi et al. (2012) applied the copula function approach to develop the joint distribution of two correlated ZIP distributions. Liu and Tian (2015) proposed a Type I multivariate ZIP distribution to model correlated multivariate count data with extra zeroes. They also described methods for obtaining maximum likelihood estimates of the proposed Type I MZIP distribution. Zhang et al. (2015) proposed two new bivariate zero-inflated generalized Poisson (BZIPG) distribution with a flexible correlation structure by incorporating a multiplicative parameter. Wang (2003) developed a bivariate zero-inflated negative binomial regression model for bivariate count data with excess zeroes. Faroughi and Ismail (2017) have used three different forms of bivariate zero-inflated negative binomial (BZINB) distribution for modelling bivariate zero-inflated count data. Cho et al. (2020) proposed a bivariate ZINB model constructed using a bivariate Poisson-Gamma mixture with application to single cell RNA sequencing data. Purhadi et al. (2021) have discussed about the development of geographically weighted bivariate zero-inflated generalized Poisson regression model and applied the same for modelling the number of pregnant maternal mortality and postpartum maternal mortality. Young et al. (2020^a, 2020^b) have discussed about various univariate zero-inflated models, multivariate models and zero-inflated models for complex data structures.

Capability evaluation and process monitoring with the help of control charts are integral part of statistical process control of a manufacturing process. There are a considerable amount of research works on monitoring of multivariate processes for continuous data as well as for multi-attribute processes for count data. Patel (1973) proposed a Hotelling-type χ^2 chart to monitor observations from multivariate Binomial or multivariate Poisson distribution. Marruci (1985) applied a Multinomial distribution to develop a control chart for multi-attribute processes. Lu et al. (1998) proposed a multivariate Shewhart np -chart (Mnp chart) for combined number of defective items. Jolayemi (2000) developed a model for an optimal design of multi-attribute control charts for processes with multiple assignable causes. Niaki and Abbasi (2007) proposed a methodology to monitor correlated multi-attribute high quality processes. As per the methodology, the multi-attribute data is first transformed, then the transformed mean and covariance matrix is estimated, and finally the well-known χ^2 control chart is applied. Topalidou and Psarakis (2009) presented a review of multinomial and multi-attribute quality control charts. Aebtarm and Bouguila (2010) proposed an optimal bivariate Poisson field chart to monitor two correlated characteristics of count data for both manufacturing and non-manufacturing processes. Albers (2012) proposed a non-parametric control chart for high quality bivariate process. Fatahi et al. (2012) applied copula function approach to achieve the joint distribution of two correlated ZIP distributions for developing a bivariate control chart which can be used for monitoring rare events. He et al. (2012) proposed a CUSUM based procedure to monitor bivariate zero-inflated Poisson processes with an application in LED packaging Industry. The methodology uses the combination of two CUSUM based control procedures for detecting shifts in the two sets of parameters in a BZIP process.

Although evaluation of capability is an integral part of statistical process control of a manufacturing process, researchers have taken very little interest on this aspect of zero-inflated processes. Only Patil and Shirke (2012) and Pal and Gauri (2021) have attempted to measure capability of the univariate zero-inflated Poisson processes. Patil and Shirke (2012) have modified the Perakis and Xekalaki (2005) proposed C_{pcu} index by incorporating the inflation of zero parameter into C_{pcu} index, and applied it for measuring the capability of a zero-inflated Poisson process. But it fails to represent the true capabilities of zero-inflated processes consistently. On the other hand, Pal and Gauri (2021) have applied the concept of Borges and Ho (2001) for measuring the capability of a zero-inflated Poisson process. Pal and Gauri (2021) proposed approach overcomes the limitation of Patil and Shirke's (2012) approach, and can reveal the true capabilities of zero-inflated processes consistently. However, to the best of our knowledge, no work is reported in literature on the measurement of process capability index of a bivariate zero-inflated process. In this paper, a methodology for evaluating capability of a bivariate zero-inflated process is presented. In the proposed approach, a BZIP model is first fitted to sample data and then the expected nonconformance in the process is estimated which is finally converted into a process capability index by using a transformation.

The article is organized as follows: in Section 2, we describe two different sets of BZIP distributions derived from Li et al. (1999) MZIP distribution and Liu and Tian (2015) Type I MZIP distribution for modelling bivariate zero-inflated count data. In section 3, we propose our approach for computing process capability index C_u^{BZIP} for the bivariate zero-inflated Poisson process.

We present two case studies in Section 4 for illustration of computational methods using the proposed approach. Finally, we conclude in Section 5.

2. Bivariate Zero-Inflated Poisson (BZIP) Models and Parameter Estimation

There are several possible ways, proposed by various researchers, to develop bivariate zero-inflated Poisson distributions. Li et al. (1999) proposed two different types of BZIP models and extended them further for multivariate zero-inflated Poisson (MZIP) models. They also investigated the distributional properties of BZIP and MZIP models. The BZIP model, proposed by Li et al. (1999), is a mixture of a bivariate Poisson, two univariate Poisson and a point mass at (0,0). Fatahi et al. (2012) applied the copula function approach to develop the joint distribution of two correlated ZIP distributions. Then, using the joint distribution, they developed a bivariate control chart which can be used for monitoring correlated rare events in a BZIP process. Liu and Tian (2015) proposed a Type I multivariate ZIP distribution to model correlated multivariate count data with extra zeroes. The bivariate ZIP distribution having only three parameters, derived from the proposed MZIP distribution, is much simpler in comparison with Li et al. (1999) proposed BZIP distribution and copula-based BZIP distribution proposed by Fatahi et al. (2012). Hence, in this article, we consider two BZIP models: the first one proposed by Li et al. (1999) and the Type I BZIP distribution, proposed by Liu and Tian (2015). Let us first describe an univariate ZIP model and then we describe about both the BZIP models.

2.1. Univariate Zero-inflated Poisson (ZIP) distribution and parameter estimation

Here, we assume that only one type of random shock occurs in the zero-inflated process and the probability of occurrence of the random shock is Ω . When the random shock occurs, defects (nonconformities) appear in the produced items according to Poisson distribution with parameter λ . The probability of occurrence of the other stationary state is $1 - \Omega$ and in this state, only zero defect items are produced. Thus, if Y is an independent random variable that follows ZIP distribution, the probability mass function (pmf) for ZIP model can be written as

$$f(y; \Omega, \lambda) = \begin{cases} (1 - \Omega) + \Omega e^{-\lambda} & \text{for } y = 0 \\ \Omega \frac{\lambda^y e^{-\lambda}}{y!} & \text{for } y = 1, 2, 3, .. \end{cases} \tag{1}$$

for some $0 \leq \Omega \leq 1$ and $\lambda > 0$. The mean and variance of the underlying Poisson distribution is λ , and the mean and variance of the ZIP distribution are $E(Y) = \Omega\lambda$ and $Var(Y) = \Omega\lambda[1 + (1 - \Omega)\lambda]$ respectively.

The unknown parameters Ω and λ of the ZIP distribution can be estimated from a random sample of size n using the maximum likelihood estimation method (Xie and Goh 1993). An enumerative search procedure can be performed using ‘Solver’ tool of Microsoft Excel for obtaining parameter estimates by solving two likelihood equations and maximizing the log-likelihood function (Pal and Gauri 2021). A Chi-square goodness-of-fit must be performed to check the adequacy of the fitted model for the sample data.

2.2. Bivariate Zero-inflated Poisson (BZIP) model (Li et al., 1999)

A BZIP model can be constructed as a combination of two univariate Poisson models, one bivariate Poisson model and a point mass at (0,0) for stationary state with probability p_{00} . Let Y_1 and Y_2 are two random variables representing two different types of defects. Then,

$$\begin{aligned} (Y_1, Y_2) &\sim (0, 0) \text{ with probability } p_{00} \\ &\sim (\text{Poisson}(\lambda_1), 0) \text{ with probability } p_{10} \\ &\sim (0, \text{Poisson}(\lambda_2)) \text{ with probability } p_{01} \\ &\sim \text{bivariate Poisson}(\lambda_{10}, \lambda_{20}, \lambda_{00}) \text{ with probability } p_{11} \end{aligned} \tag{2}$$

where each $p_{00}, p_{10}, p_{01}, p_{11} > 0$, and $p_{00} + p_{10} + p_{01} + p_{11} = 1$. A bivariate Poisson distribution (X_1, X_2) with parameters $(\lambda_{10}, \lambda_{20}, \lambda_{00})$ is represented as follows (Marshall and Olkin, 1995):

$$X_1 = U_1 + Z \text{ and } X_2 = U_2 + Z$$

where U_1, U_2 , and Z are independent and have univariate Poisson distributions with respective means $\lambda_{10}, \lambda_{20}$, and λ_{00} , ($eac \square > 0$).

The probability mass function of the BZIP is given by:

$$\begin{aligned} P(Y_1 = 0, Y_2 = 0) &= p_{00} + p_{10} \exp(-\lambda_1) + p_{01} \exp(-\lambda_2) + p_{11} \exp(-\lambda), \\ P(Y_1 = y_1, Y_2 = 0) &= \frac{1}{y_1!} [p_{10} \lambda_1^{y_1} \exp(-\lambda_1) + p_{11} \lambda_{10}^{y_1} \exp(-\lambda)] \\ P(Y_1 = 0, Y_2 = y_2) &= \frac{1}{y_2!} [p_{01} \lambda_2^{y_2} \exp(-\lambda_2) + p_{11} \lambda_{20}^{y_2} \exp(-\lambda)] \\ P(Y_1 = y_1, Y_2 = y_2) &= p_{11} \left\{ \sum_{j=0}^{\min(y_1, y_2)} \left[\frac{\lambda_{10}^{y_1-j} \lambda_{20}^{y_2-j} \lambda_{00}^j \exp(-\lambda)}{(y_1-j)!(y_2-j)!j!} \right] \right\} \end{aligned} \tag{3}$$

for $y_1, y_2 = 1, 2, \dots$, and $\lambda = \lambda_{10} + \lambda_{20} + \lambda_{00}$. The authors also assumed that $\lambda_1 = \lambda_{10} + \lambda_{00}$ and $\lambda_2 = \lambda_{20} + \lambda_{00}$ (see Li et al. 1999), for simplification of the BZIP model and to let the marginal distributions become univariate ZIP distributions.

Estimation of Parameters of BZIP Model of Li et al. (1999)

It can be observed from the above BZIP model that there are total 10 unknown parameters, namely $p_{00}, p_{10}, p_{01}, p_{11}, \lambda, \lambda_1, \lambda_2, \lambda_{10}, \lambda_{20}$, and λ_{00} . Among them, three parameters namely, λ, λ_1 , and λ_2 can be obtained from $\lambda_{10}, \lambda_{20}$, and λ_{00} values, and parameter p_{11} can be automatically obtained since the constraint $p_{00} + p_{10} + p_{01} + p_{11} = 1$ must be satisfied.. Therefore, we need to estimate only six unknown parameters, namely $p_{00}, p_{10}, p_{01}, \lambda_{10}, \lambda_{20}$, and λ_{00} . The unknown six parameters can be estimated from a random sample of size n using the maximum likelihood method (Li et al. 1999). The maximum likelihood procedure is quite complicated and may not be easily implementable for many practitioners.

Let there are n pairs of values of (y_1, y_2) from a BZIP process where y_1 gives the number of first type of defects and y_2 gives the number of second type of defects in a single item. For a high quality BZIP process, most of the values will be $(0,0)$. Let there are n_{00} number of such $(0,0)$ pairs. Let us denote the maximum number of type 1 defects and type 2 defects in an item by u and v respectively. So, there can be a maximum of $(u+1)(v+1)$ number of (i,j) pairs where $i=0,1,2,\dots,u$, and $j = 0,1,2,\dots,v$. Let the frequency of occurrence of $(i,j)^{th}$ pair is denoted by n_{ij} . Therefore, $\sum_{i=0}^u \sum_{j=0}^v n_{ij} = n$. In the sample data, there can be many (i,j) pairs whose frequencies may be equal to zero.

The log likelihood function can be written as

$$\ln L = \ln \prod_{i=0}^u \prod_{j=0}^v P(Y_1 = i, Y_2 = j)^{n_{ij}} = \sum_{i=0}^u \sum_{j=0}^v n_{ij} \times \ln P(Y_1 = i, Y_2 = j) \tag{4}$$

This log likelihood function is to be maximized by changing those six unknown parameters ($p_{00}, p_{10}, p_{01}, \lambda_{10}, \lambda_{20}$, and λ_{00}) subjecting to few constraints. The constraints are: i) all parameters are positive, and ii) $p_{00} + p_{10} + p_{01} + p_{11} = 1$. Using enumerative search procedures, the estimates of the unknown parameters can be obtained. While performing this search procedure, one has to check about the expected proportion value computed from the fitted BZIP model. The expected probability $P(Y_1 = i, Y_2 = j)$ value for $i=0,1,2,\dots,u$ and $j=0,1,2,\dots,v$, should be close to (n_{ij}/n) value obtained from sample data. The analyst must perform Chi-square goodness-of-fit test for checking the adequacy of the fitted model.

2.3. Type I Bivariate ZIP Distribution of Liu and Tian (2015)

Liu and Tian (2015) proposed a Type I multivariate ZIP distribution to model correlated multivariate count data with extra zeroes. The proposed BZIP distribution can be thought of as an extension of univariate ZIP distribution with an extra parameter for accounting the second type of defect data.

The Type I BZIP distribution is represented as $(Y_1, Y_2) \sim BZIP^{(1)}(\Omega, \lambda_1, \lambda_2)$ and the joint probability mass function of (Y_1, Y_2) is given as

$$f((y_1, y_2) | \Omega, \lambda_1, \lambda_2) = \begin{cases} (1 - \Omega) + \Omega e^{-(\lambda_1 + \lambda_2)} & \text{where } (y_1, y_2) = (0,0) \\ \Omega e^{-(\lambda_1 + \lambda_2)} \frac{\lambda_1^{y_1} \lambda_2^{y_2}}{y_1! y_2!} & \text{where } (y_1, y_2) \neq (0,0) \end{cases} \tag{5}$$

where $0 < \Omega < 1$ and $\lambda_1, \lambda_2 > 0$.

Let us assume that a sample of size n is collected from a BZIP process and there are n_{00} number of defect free items. The observed-data likelihood function for the Type I BZIP model can be written as

$$\begin{aligned} \ln L &= \ln \{ [(1 - \Omega) + \Omega e^{-(\lambda_1 + \lambda_2)}]^{n_{00}} * \Omega^{n - n_{00}} * e^{-(n - n_{00})(\lambda_1 + \lambda_2)} * \prod_{i,j} \left(\frac{\lambda_1^{y_{1i}} \lambda_2^{y_{2j}}}{y_{1i}! y_{2j}!} \right)^{n_{ij}} \} \\ &= n_{00} \ln [(1 - \Omega) + \Omega e^{-(\lambda_1 + \lambda_2)}] + (n - n_{00}) [\ln \Omega - (\lambda_1 + \lambda_2)] + \sum_{i,j} n_{ij} \ln \left(\frac{\lambda_1^{y_{1i}} \lambda_2^{y_{2j}}}{y_{1i}! y_{2j}!} \right) \end{aligned} \tag{6}$$

This log likelihood function is to be maximized by changing three unknown parameters $(\Omega, \lambda_1, \lambda_2)$. Using enumerative search procedures, the estimates of the unknown parameters can be obtained. This search procedure can be performed using Solver tool of Excel package. While performing this search procedure, one has to check about the expected proportion value computed from the fitted Type I BZIP model. The expected probability $P(Y_1 = i, Y_2 = j)$ value for $i=0,1,2,\dots,u$ and $j=0,1,2,\dots,v$, should be close to (n_{ij}/n) value obtained from sample data. The analyst must perform Chi-square goodness-of-fit test for checking the adequacy of the fitted model.

2.4. Selection of the most appropriate model

To a given sample dataset, both BZIP distribution of Li et al. (1999) and Type I BZIP distribution of Liu and Tian (2015) can be fitted. However, one should use the most appropriate model for the purpose of statistical process control. For identification of the most appropriate model for a zero-inflated process generally the following two information criterion are used: i) Akaike Information Criterion (AIC) and ii) Bayesian Information Criterion (BIC). The AIC and BIC are formally defined as

$$AIC = 2K - 2 \times \ln L \tag{7}$$

$$BIC = K \ln(n) - 2 \times \ln L \tag{8}$$

where K is the number of estimated parameters in the model, $\ln L$ is the log-likelihood function for the model, and n is the sample size. A smaller value of AIC (or BIC) implies that the existing discrepancy between the fitted model and the data is less, and thus the fitted model that results in the minimum value of AIC (or BIC) can be considered as the most appropriate one.

It is important to note that AIC (or BIC) value reveals the relative goodness of two or more fitted models. But it does not say anything about the adequacy of the fitted model. Therefore, the primary criterion for considering a fitted model as the candidate model for the comparison is that it must pass the Chi-square goodness-of-fit.

3. Proposed Approach for Computation of PCI of a Bivariate ZIP process

The standard process capability indices (Cp, Cpk, Cpm, Cpmk) are developed for univariate processes where the process data are continuous in nature and follow a normal distribution. Subsequently, generalized PCIs have been developed for continuous process data that do not follow a normal distribution. Bivariate PCIs are also developed for two continuous characteristics from a process with the assumption that the bivariate process data jointly follow a bivariate normal distribution. For detailed information on process capability indices for continuous data, refer Kotz and Johnson (2002), Polhemus (2018).

Defects data are discrete in nature (also known as count data) and generally follows a Poisson distribution or negative binomial distribution. Various PCIs have been proposed by several researchers for attribute process data (number of defects, proportion of defective items). For example, Yeh and Bhattacharya (1998) proposed C_f index, Borges and Ho (2001) proposed C_u index, Perakis and Xekalaki (2005) presented C_{pc} index and Maiti et al. (2010) suggested C_{py} index for assessment of process capability of attribute process data. Pal and Gauri (2020) assessed relative goodness of these indices and suggested about the best measure of PCI for Poisson process data.

For high quality zero-inflated process where most of the items are defect free and only a few items contain defects due to the occurrence of random shock, the defect data are assumed to follow a zero-inflated Poisson (ZIP) distribution or a zero-inflated negative binomial (ZINB) distribution (Pal and Gauri 2021). Pal and Gauri (2021) have proposed a method for computation of PCI of univariate zero-inflated processes. As per the proposed methodology, a ZIP or ZINB distribution is utilized for modelling the sample zero-inflated process data and then the expected proportion of conformance is estimated using the fitted model. Then, adopting the approach of Borges and Ho (2001) for computation of PCI, the estimated proportion of conformance is converted into PCI C_u -index by using inverse transformation of standard normal distribution. We adopt the similar methodology to compute the PCI for a bivariate process where two different types of defects are assumed to follow a bivariate zero-inflated Poisson (BZIP) distribution.

In this method, a BZIP distribution is first fitted to the sample data using maximum likelihood method. Then, the expected proportion of nonconforming items having combined number of defects (arising from two types of defects) more than the specified upper limit of defects in an item is estimated using the fitted BZIP model. The expected proportion of non-conformance above USL is mapped to the Z-score in the right side of standard normal distribution, and 1/3rd of this Z-score is considered as the measure of the process capability with respect to USL and it is denoted as C_u^{BZIP} . The C_u^{BZIP} index can be evaluated from a bivariate zero-inflated Poisson process using the following procedures:

- 1) Collect a sample of n units from the concerned zero-inflated process and observe the numbers of two different types of defects present in each of the sample items. Let the random variables (Y_1, Y_2) represent the number of two types of defects present in an item.
- 2) Fit appropriate bivariate zero-inflated Poisson distribution to the observed count data.
- 3) Estimate the expected proportion of nonconforming items (PNU_{USL}) with respect to specified USL for defects in an item of the concerned zero-inflated process. The specified USL can be on the combined number of defects in an item or two individual USL values for two different types of defects in an item.

Let c^{USL} be the USL specified by the manufacturer on the combined number of defects in a unit. A unit will be considered nonconforming if the number of defects in it is more than c^{USL} . Then the expected proportion of non-conforming items PNU_{USL} can be estimated as follows:

$$\overline{PNU}_{USL} = P((Y_1 + Y_2) > c^{USL}) = 1 - P((Y_1 + Y_2) \leq c^{USL}) \tag{9}$$

$$P((Y_1 + Y_2) \leq c^{USL}) = \sum_{i=0}^{c^{USL}-j} \sum_{j=0}^{c^{USL}} P(Y_1 = i, Y_2 = j) \tag{10}$$

The probability $P((Y_1 + Y_2) \leq c^{USL})$ gives the proportion of conforming items having combined number of defects less than equal to the specified USL.

Suppose (c_1^{USL}, c_2^{USL}) are the individual USL values for two different types of defects specified by the manufacturer. A unit will be considered nonconforming if the number of first type of defects in it is more than c_1^{USL} or the number of second type of defects is more than c_2^{USL} or both. Then PNU_{USL} can be estimated as follows:

For a BZIP process, the proportion PNU_{USL} can be estimated as

$$\begin{aligned} \overline{PNU}_{USL} &= 1 - P(Y_1 \leq c_1^{USL}, Y_2 \leq c_2^{USL}) \\ &= 1 - \sum_{i=0}^{c_1^{USL}} \sum_{j=0}^{c_2^{USL}} P(Y_1 = i, Y_2 = j) \end{aligned} \tag{11}$$

The probability $P(Y_1 \leq c_1^{USL}, Y_2 \leq c_2^{USL})$ gives the proportion of conforming items having number of two types of defects less than equal to the respective specified USL values.

- 4) Determine the Z-value in the right side of the standard normal distribution that results in probability area equal to \overline{PNU}_{USL} value. In other words, map the computed \overline{PNU}_{USL} value to the Z-score in the right side of standard normal distribution. Let Z_U is the value of Z that results in probability area \overline{PNU}_{USL} above it. The Z_U value can be obtained by using inverse cumulative probability of the standard normal distribution function as follows:

$$Z_U = \begin{cases} 0 & \overline{PNU}_{USL} \geq 0.5 \\ \Phi^{-1}(1 - \overline{PNU}_{USL}) & 0.0 < \overline{PNU}_{USL} < 0.5 \\ 4 & \overline{PNU}_{USL} = 0 \end{cases} \tag{12}$$

where, $\Phi(\cdot)$ denotes the standard normal cumulative distribution function.

- 5) Finally, obtain the estimate of the process capability index C_u^{BZIP} of the concerned zero-inflated process as follows:

$$\hat{C}_u^{BZIP} = (1/3) \times Z_U \tag{13}$$

If the value of estimated index \hat{C}_u^{BZIP} is greater than 1, then the capability of the concerned zero-inflated process can be considered good. In this case, the process is capable of producing more than 99.865% conforming items, i.e. more than 99.865% of produced items will have total number of defects less than equal to c^{USL} (the specified USL).

It is important to mention that if the value of proportion \overline{PNU}_{USL} is more than or equal to 0.5, then Z_U is considered as zero instead of taking the negative values. Consequently, the index C_u^{BZIP} becomes equal to zero. When \overline{PNU}_{USL} is more than 0.5, it means that more than 50% of produced items are nonconforming, i.e. more than 50% of produced items will have combined number of defects more than c^{USL} (the specified USL). Thus, it is concluded that the corresponding manufacturing process is not capable at all.

3.1 Estimation of confidence interval of C_u^{BZIP}

Since \hat{C}_u^{BZIP} is a point estimate obtained from sample data, it is necessary to construct confidence interval (CI) of the capability index C_u^{BZIP} for inference purpose, especially when the sample size is relatively small. However, construction of CI using the sampling distribution of the estimated \hat{C}_u^{BZIP} is found to be quite difficult. Hence, we use Nagata and Nagahata (1994) proposed generalized approximation formula for construction of two-sided CI of C_u^{BZIP} . According to Nagata and Nagahata (1994),

$$(1 - \alpha)\% \text{ two-sided CI of } C_u^{BZIP} = \left(\hat{C}_u^{BZIP} - Z_{1-\frac{\alpha}{2}} \sqrt{\frac{1}{9n} + \frac{(\hat{C}_u^{BZIP})^2}{2(n-1)}}, \hat{C}_u^{BZIP} + Z_{1-\frac{\alpha}{2}} \sqrt{\frac{1}{9n} + \frac{(\hat{C}_u^{BZIP})^2}{2(n-1)}} \right) \tag{14}$$

where, α is the level of significance and $(1 - \alpha)$ is the confidence coefficient.

4. Analysis and Results

For the purpose of illustrations of computations of process capability indices using the proposed approach and assessing its effectiveness, two case studies are presented here.

4.1 Case Study 1

This case study from a LED packaging process was presented by He et al. (2012). In LED packaging process, LEDs are placed onto a printing circuit board (PCB) and then a soldering process is done that connects the LEDs and PCB via golden wires. There are two types of defects in the LEDs on a manufactured PCB, namely – LED mounting errors (defect 1) and soldering errors (defect 2). They have selected 100 PCBs from the packaging process when the process is under an in-control situation and then counted the number of two types of defects in them. He et al. (2012) used this data to estimate the process parameters and develop control limits for CUSUM control charts for monitoring the packaging process. The raw data is given in the article (He et al. 2012). Here, the frequency distribution of two types of defects from the sample data is presented in the following Table 1.

Table 1. Frequency distribution of two types of defects

(y_1, y_2)	Frequency	(y_1, y_2)	Frequency	(y_1, y_2)	Frequency
(0, 0)	62	(0, 1)	1	(0, 8)	2
(1, 0)	2	(0, 2)	1	(0, 9)	1
(2, 0)	1	(0, 3)	3	(2, 3)	2
(3, 0)	2	(0, 4)	4	(3, 8)	2
(4, 0)	3	(0, 5)	2	(7, 5)	1
(5, 0)	1	(0, 6)	1	(7, 8)	1
(7, 0)	2	(0, 7)	6		

Following the notations described earlier, it can be written that:

$$n_{00} = 62, \quad \sum_{i=1}^7 n_{i0} = 11, \quad \sum_{j=1}^9 n_{0j} = 21, \quad \sum_{i=1}^7 \sum_{j=1}^9 n_{ij} = 6, \quad n = 100$$

Also, it can be computed from the sample data that the total number of type 1 defects is 65, total number of type 2 defects is 146 and combined number of defects is 211.

We now apply the maximum likelihood procedure to find the estimates of the unknown six parameters ($p_{00}, p_{10}, p_{01}, \lambda_{10}, \lambda_{20},$ and λ_{00}) of BZIP distribution, proposed by Li et al. (1999). The iterative procedure for finding the estimates by maximizing the log-likelihood function is carried out using Solver tool of Microsoft Excel. A starting solution of ($p_{00}, p_{10}, p_{01},$ and p_{11}) for the iterative procedure can be considered as $(\frac{n_{00}}{n}, \frac{\sum_i n_{i0}}{n}, \frac{\sum_j n_{0j}}{n}, \frac{\sum_{i,j>1} n_{ij}}{n})$ values respectively. The procedure of computing expected proportions and likelihood function is slightly complicated and will be difficult for many practitioners. The maximum likelihood estimates of the parameters of the fitted BZIP distribution are obtained as:

$$\hat{p}_{00} = 0.6162, \hat{p}_{10} = 0.1125, \hat{p}_{01} = 0.2094, \hat{p}_{11} = 0.0619$$

$$\hat{\lambda}_{00} = 1.3543, \hat{\lambda}_{10} = 2.3426, \hat{\lambda}_{20} = 4.0187, \hat{\lambda}_1 = 3.6969, \hat{\lambda}_2 = 5.3730 \text{ and } \hat{\lambda} = 7.7156$$

with the log-likelihood value as $\ln L = -197.231$. Although, the Chi-square goodness-of-fit must be performed to check the adequacy of the fitted distribution, it could not be done for this small size sample data because of very small frequency values of the observed pairs.

We then try to find the maximum likelihood estimates of Type I BZIP distribution, proposed by Liu and Tian (2015). The maximum likelihood estimates of the parameters of the fitted BZIP distribution are obtained as:

$$\hat{\Omega} = 0.3815, \hat{\lambda}_1 = 1.7037, \hat{\lambda}_2 = 3.8269 \text{ with } \ln L = -263.8$$

The expected frequencies computed using this fitted Type I BZIP distribution of Liu and Tian (2015) are found as highly deviated from the observed frequencies in the sample data. Also, the log-likelihood value of the fitted Type I BZIP distribution is much less compared to the fitted BZIP distribution of Li et al. (1999). Hence, the fitted BZIP model of Li et al. (1999) is chosen for further computation of non-conformances and process capability index.

For computation of process capability index and expected non-conformance, we need a specified upper limit for the combined number of defects or individual upper limits for both type of defects in a PCB. He et al. (2012) have not specified USL value either for the combined number of defects or for individual USL values for two types of defects. Obviously, the desirable value for number of defects in a PCB is zero. However, for computation of PCI, we need to have positive integer values as USL for the number of defects.

In the first case, let us assume that the specified USL for the combined number of defects is $c^{USL} = 7$, that means, a PCB will be considered as non-conforming if the combined number of defects in it exceeds 7. Then, the expected proportion of non-conforming PCBs can be computed as:

$$\text{Expected Conformance } P((Y_1 + Y_2) \leq c^{USL}) = \sum_{i=0}^{7-j} \sum_{j=0}^7 P(Y_1 = i, Y_2 = j) = 0.9168$$

$$\text{Expected Nonconformance } \widehat{PNU}_{USL} = 1 - P((Y_1 + Y_2) \leq c^{USL}) = 0.0832$$

The process capability index C_u^{BZIP} can be computed as

$$\hat{C}_u^{BZIP} = \frac{1}{3} \Phi^{-1}(1 - \widehat{PNU}_{USL}) = \frac{1}{3} \Phi^{-1}(0.9168) = 0.461$$

Since, the estimated PCI \hat{C}_u^{BZIP} is much less than 1.0, it can be concluded that the LED packaging process is not capable. In fact, around 8.32% of manufactured PCBs will have the combined number of defects more than the specified USL of 7.

The zero-defect proportion $P(Y_1 = 0, Y_2 = 0)$ is computed from the fitted BZIP model as 0.62. This means, 62% of PCBs are expected not to have any type of defect and around 38% of manufactured PCBs will have either LED mounting defect or soldering defect or both.

In the second case, let us assume that the specified USL values for the two types of defects in a PCB are given as ($c_1^{USL} = 3, c_2^{USL} = 4$). This means, a PCB will be considered as non-conforming if the number of first type of defects in it exceeds 3 or the number of second type of defects exceeds 4. Then, the expected proportion of non-conforming PCBs can be computed as:

$$\text{Expected Conformance } P(Y_1 \leq c_1^{USL}, Y_2 \leq c_2^{USL}) = \sum_{i=0}^{c_1^{USL}} \sum_{j=0}^{c_2^{USL}} P(Y_1 = i, Y_2 = j) = 0.7636$$

$$\text{Expected Nonconformance } \widehat{PNU}_{USL} = 1 - P(Y_1 \leq c_1^{USL}, Y_2 \leq c_2^{USL}) = 0.2364$$

The process capability index C_u^{BZIP} can be computed as

$$\hat{C}_u^{BZIP} = \frac{1}{3} \Phi^{-1}(1 - \widehat{PNU}_{USL}) = \frac{1}{3} \Phi^{-1}(0.7636) = 0.239$$

Since, the estimated PCI \hat{C}_u^{BZIP} is much less than 1.0, it can be concluded that the LED packaging process is not capable. In fact, around 23.6% of manufactured PCBs will either have the first type of defects more than the specified USL of 3 or the second type of defects more than the specified USL of 4 or both.

It may be noted here that the second type of specified limits (USL for each type of defects) is more stringent than first type of specified USL for the combined number of defects. This is because of many defect combinations like [(4,0), (4,1), (4,2), (4,3), (1,5), (1,6), (2,5),.. etc] which are considered as non-conforming PCBs under the second type of specification limits, will be considered as conforming PCBs under the first type of USL for combined number of defects. This becomes clear from the expected proportion of conforming PCBs estimated above.

4.2 Case Study 2

The following study was carried out in an automobile industry for one critical component used in motor cycles. After receiving the input components from an authorized vendor, these components are processed for a smooth outer surface finish using a special purpose grinding machine. The defects are detected only after removal of a fixed amount of material from component outer surface after the grinding process. There are mainly two types of defects, namely, holes and uncleaned surface, either one or both occurring in one or more locations. Although around 90% components are observed as defect free, there are around 10% components having either one or both type of defects in one or more locations along the component surface. It is desirable not to have any kind of defects in any component. However, looking at the complex structure of the component, it was decided that a maximum of total 4 defects ($c^{USL} = 4$) in any component will be treated as acceptable and any component containing more than 4 defects will be rejected and will be sent back to the supplier at his expenses. A sample of 1000 components is selected randomly from a lot and all those components are inspected after the grinding process. The following Table 2 gives the number of defect occurrences in a component and their frequencies in the entire sample of size 1000.

Table 2. Frequency distribution of two types of defects in components after grinding

(y_1, y_2)	Frequency	(y_1, y_2)	Frequency	(y_1, y_2)	Frequency
(0, 0)	893	(0, 3)	5	(2, 1)	6
(1, 0)	16	(0, 4)	2	(2, 2)	7
(2, 0)	9	(0, 5)	1	(2, 3)	5
(3, 0)	4	(1, 1)	8	(3, 1)	3
(4, 0)	2	(1, 2)	5	(3, 2)	4
(0, 1)	15	(1, 3)	3	(4, 1)	1
(0, 2)	9	(1, 4)	2		

Following the notations described earlier, it can be written that:

$$n_{00} = 893, \quad \sum_{i=1}^4 n_{i0} = 31, \quad \sum_{j=1}^5 n_{0j} = 32, \quad \sum_{i=1}^4 \sum_{j=1}^5 n_{ij} = 44, \quad n = 1000$$

Also, it can be computed from the sample data that the total number of type 1 defects is 133, total number of type 2 defects is 143 and combined number of defects is 276.

We now apply the maximum likelihood procedure to find the estimates of the unknown six parameters ($p_{00}, p_{10}, p_{01}, \lambda_{10}, \lambda_{20}$, and λ_{00}) of BZIP distribution, proposed by Li et al. (1999). The maximum likelihood estimates of the parameters of the fitted BZIP distribution are obtained as:

$$\hat{p}_{00} = 0.8697, \hat{p}_{10} = 0.0312, \hat{p}_{01} = 0.0295, \hat{p}_{11} = 0.0696$$

$$\hat{\lambda}_{00} = 0.6432, \hat{\lambda}_{10} = 0.6929, \hat{\lambda}_{20} = 0.8049, \hat{\lambda}_1 = 1.3361, \hat{\lambda}_2 = 1.4481 \text{ and } \hat{\lambda} = 2.1410$$

with the log-likelihood value as $\ln L = -637.397$. Since, six unknown parameters are estimated from the sample data of size $n = 1000$, the AIC and BIC values are computed as

$$AIC = 2 * 6 - 2\ln L = 1286.794 \text{ and } BIC = 6 * \ln(n) - 2\ln L = 1316.24$$

We then try to find the maximum likelihood estimates of Type I BZIP distribution, proposed by Liu and Tian (2015). The maximum likelihood estimates of the parameters of the fitted BZIP distribution and the corresponding AIC, BIC values are obtained as:

$$\hat{\Omega} = 0.1186, \hat{\lambda}_1 = 1.1218, \hat{\lambda}_2 = 1.2062 \text{ with } \ln L = -642.302$$

$$AIC = 2 * 3 - 2\ln L = 1290.60 \text{ and } BIC = 3 * \ln(n) - 2\ln L = 1305.33$$

Following the model selection criteria of minimum AIC value and maximum log-likelihood value, it is decided to select the Li et al. (1999) BZIP distribution with 6 estimated parameters as the fitted model for this sample data. Using those 6 estimated model parameters, the expected frequencies of defects pairs are computed which is found quite similar with the observed frequencies. However, Liu and Tian (2015) Type I BZIP distribution is also found as a good model for the sample data and can be utilized for future estimation purposes.

The specified USL for the combined number of defects is given as $c^{USL} = 4$. The expected proportion of non-conforming components is computed using Li et al. (1999) BZIP distribution as:

Expected Conformance $P((Y_1 + Y_2) \leq c^{USL}) = \sum_{i=0}^{4-j} \sum_{j=0}^4 P(Y_1 = i, Y_2 = j) = 0.9879$

Expected Nonconformance $\widehat{PNU}_{USL} = 1 - P((Y_1 + Y_2) \leq c^{USL}) = 0.0121$

The process capability index C_u^{BZIP} can be computed as

$$\hat{C}_u^{BZIP} = \frac{1}{3} \Phi^{-1}(1 - \widehat{PNU}_{USL}) = \frac{1}{3} \Phi^{-1}(0.9879) = 0.751$$

The 95% confidence interval of process capability index C_u^{BZIP} is computed using equation no. (14) as

$$95\% \text{ Confidence interval of } C_u^{BZIP} \text{ is } (0.708, 0.794)$$

Since, the estimated PCI \hat{C}_u^{BZIP} is less than 1.0, it can be concluded that the supplied materials are not conforming to permissible limit with respect to total number of defects. In fact, around 1.2% components will be rejected due to containing the combined number of defects more than the specified USL of 4.

If we use the fitted Type I BZIP distribution with 3 parameters of Liu and Tian (2015), the expected proportion of non-conformance \widehat{PNU}_{USL} is estimated as 0.0103 and accordingly, the process capability index C_u^{BZIP} can be computed as 0.771. The expected proportion of non-conformance and process capability index values estimated using both BZIP models are observed as quite close to each other. In this particular case, Liu and Tian's Type I BZIP distribution may be preferred for future prediction purposes because of its simplicity and easier computational procedure of expected proportions in comparison with Li et al. (1999) proposed BZIP distribution.

5. Conclusions and Future Research

In high quality manufacturing processes, most of the items produced are defect free and only a few items contain one or more number of single or multiple types of defects. Such processes are referred to as zero-inflated processes with random shocks. A zero-inflated Poisson (ZIP) distribution is commonly used for modelling zero-inflated process data with single type of defect and a bivariate zero-inflated Poisson (BZIP) distribution is used for modelling zero-inflated process data with two types of defects and a multivariate zero-inflated Poisson (MZIP) distribution is used for modelling zero-inflated process data with more than two types of defects. Often evaluation of capabilities of such zero-inflated processes becomes necessary for their assessment, comparison and decision making for improvement. Although some works on evaluation of process capability of univariate zero-inflated Poisson (ZIP) processes are available in literature, no work is reported on measuring the capabilities of BZIP or MZIP processes. This paper presents a methodology for measuring capability of a BZIP process. In the proposed approach, a BZIP model is first fitted to sample data and then the expected nonconformance in the process is estimated which is finally converted into a process capability index by using a transformation. The proposed methodology is illustrated using two case studies and the results reveal that the true capabilities of these processes are well represented by the measured values of the proposed process capability index.

In this article, the proposed methodology is developed considering only two different forms of BZIP distribution. In literature, some other forms of BZIP distributions are reported. It will be an interesting future work to study how robust is the proposed methodology when other forms of BZIP distribution are used for modelling the process data. Sometimes bivariate zero-inflated negative binomial (BZINB) distribution and bivariate zero-inflated generalized Poisson (BZIGP) distributions are used for modelling bivariate count data, especially when there is over dispersion in the count data. Future studies are required to examine the effectiveness of the proposed methodology when BZINB or BZIGP distribution is used for modelling process data. Future studies may also be aimed at evaluating process capabilities of multivariate zero-inflated processes.

Acknowledgment

Thanks to the referees for their comments that have enhanced the quality of this article.

References

- Aebtarm, S., and Bouguila, N., 2010. An optimal bivariate Poisson field chart for controlling high-quality manufacturing processes. *Expert Systems with Applications*, Vol. 37, pp. 5498-5506. <https://doi.org/10.1016/j.eswa.2010.02.060>
- Albers, W., 2012. Nonparametric control charts for bivariate high quality processes. *International Journal of Pure and Applied Mathematics*, Vol. 79, pp. 139-154.
- Borges, W. and Ho, L. L., 2001. A fraction defective based capability index. *Quality and Reliability Engineering International*, Vol. 17, pp. 447-458. <https://doi.org/10.1002/qre.438>
- Campbell, J.T. (1938). The Poisson correlation function, *Proceedings of the Edinburgh Mathematical Society (Series 2)*, Vol.4, pp. 18-26.
- Cho, H., Liu, C., Preisser, J.S., and Wu, D., 2020. A bivariate zero-inflated negative binomial model for identifying underlying dependence with application to single cell RNA sequencing data. <https://doi.org/10.1101/2020.03.06.977728>.
- Famoye, F., and Consul, P.C., 1995. Bivariate generalized Poisson distribution with some applications. *Metrika*, Vol. 42, pp. 127-138. <https://doi.org/10.1007/BF01894293>

- Faroughi, P., and Ismail, N., 2017. Bivariate zero-inflated negative binomial regression model with applications. *Journal of Statistical Computation and Simulation*, Vol. 87, pp. 457-477. <https://doi.org/10.1080/00949655.2016.1213843>
- Fatahi, A.A., Noorossana, R., Dokouhaki, P., and Moghaddam, B.F., 2012. Copula-based bivariate ZIP control chart for monitoring rare events. *Communication in Statistics - Theory and Methods*, Vol. 41, pp.2699-2716. <https://doi.org/10.1080/03610926.2011.556296>
- He, S., He, Z., Wang, G.A., 2012. CUSUM charts for monitoring bivariate zero-inflated Poisson processes with an application in the LED packaging industry. *IEEE Transactions on Components, Packaging and Manufacturing Technology*, Vol. 2, pp. 169-180. <https://doi.org/10.1109/TCPMT.2011.2176948>
- Holgate, P., 1964. Estimation for the bivariate Poisson distribution, *Biometrika*, Vol. 51, pp.241-245. <https://doi.org/10.2307/2334210>
- Jolayemi, J.K., 2000. An optimal design of multiattribute control charts for processes subject to a multiplicity of assignable causes. *Applied Mathematics and Computation*, Vol. 114, pp.187-203. [https://doi.org/10.1016/S0096-3003\(99\)00111-3](https://doi.org/10.1016/S0096-3003(99)00111-3)
- Kotz, S. and Johnson, N. L., 2002. Process capability indices - a review, 1992-2000. *Journal of Quality Technology*, Vol. 34, pp 2-19. <https://doi.org/10.1080/00224065.2002.11980119>
- Krishnamoorthy, A.S., 1951. Multivariate Binomial and Poisson distributions. *Sankhya*, Vol. 11, pp. 117-124. <https://www.jstor.org/stable/25048072>
- Li, C.S., Lu, J.C., Park, J., Kim, K., Brinkley, P.A., and Peterson, J.P. 1999. Multivariate zero-inflated Poisson models and their applications. *Technometrics*, Vol. 41, pp. 29-38. <https://doi.org/10.2307/1270992>
- Lu, X.S., Xie, M., Goh, T.N., and Lai, C.D., 1998. Control chart for multivariate attribute processes. *International Journal of Production Research*, Vol. 36, No. 12, pp. 3477-3489. <https://doi.org/10.1080/002075498192166>
- Mahmood, T. and Xie, M., 2019. Models and monitoring of zero-inflated processes: The past and current trends. *Quality and Reliability Engineering International*, Vol. 35, pp. 2540-2557. <https://doi.org/10.1002/qre.2547>
- Maiti, S.S., Saha, M. and Nanda, A.K., 2010. On generalising process capability indices. *Quality Technology & Quantitative Management*, Vol. 7, No. 3, pp. 279-300. <https://doi.org/10.1080/16843703.2010.11673233>
- Marshall, A.W., and Olkin, I., 1985. A family of bivariate distributions generated by the bivariate Bernoulli distribution. *Journal of the American Statistical Association*, Vol. 80, pp. 332-338.
- Nagata, Y. and Nagahata, H., 1994. Approximation formulas for the confidence intervals of process capability indices. *Okayama Economic Review*, Vol. 25, pp. 301-314.
- Niaki, S.T.A., and Abbasi, B., 2007. On the monitoring of multi-attribute high-quality production processes. *Metrika*, Vol. 66, pp. 373-388. <https://doi.org/10.1007/s00184-006-0117-0>
- Pal, S., and Gauri, S.K., 2020. Measuring capability of a Poisson process: relative goodness of the estimates obtained by different approaches. *International Journal of Engineering, Science and Technology*, Vol. 12, pp. 1-13. <https://doi.org/10.4314/ijest.v12i4.1>
- Pal, S., and Gauri, S.K., 2021. Measuring capabilities of zero-inflated processes. *International Journal of Engineering, Science and Technology*, Vol. 13, No. 3, pp. 37-48. <https://doi.org/10.4314/ijest.v13i3.4>
- Patil, M. K. and Shirke, D.T., 2012. Process capability index for zero-inflated Poisson process. *International Journal of Mathematical Sciences & Engineering Applications*, Vol. 6, pp. 87-97.
- Perakis, M. and Xekalaki, E., 2005. A process capability index for discrete processes. *Journal of Statistical Computation and Simulation*, Vol. 75, No. 3, pp. 175-187. <https://doi.org/10.1080/00949650410001687244>
- Polhemus, N.W., 2018. *Process Capability Analysis: Estimating Quality*. CRC press, Taylor & Francis Group, Boca Ration, Florida.
- Purhadi, Sari, D.N., Aini, Q., and Irhamah, 2021. Geographically weighted bivariate zero-inflated Poisson regression model and its application. *Heliyon*, 7, e07491. <https://doi.org/10.1016/j.heliyon.2021.e07491>
- Tian, W., You, H., Zhang, C., Kang, S., Jia, X. and Chien, W.K., 2019. Statistical process control for monitoring the particles with excess zero counts in semiconductor manufacturing. *IEEE Transactions on Semiconductor Manufacturing*, Vol. 32, pp.93-103. <https://doi.org/10.1109/TSM.2018.2882862>
- Topalidou, E., and Perakis, S., 2009. Review of multinomial and multiattribute quality control charts. *Quality and Reliability Engineering International*, Vol. 25, pp. 773-804. <https://doi.org/10.1002/qre.999>
- Wagh, Y.S., and Kamalja, K.K., 2018. Zero-inflated models and estimation in zero-inflated Poisson distribution. *Communications in Statistics - Simulations and Computation*. Vol. 47, pp. 2248-2265. <https://doi.org/10.1080/03610918.2017.1341526>
- Wang, P., 2003. A bivariate zero-inflated negative binomial regression model for count data with excess zeroes. *Economic Letters*, Vol. 78, pp. 373-378. [https://doi.org/10.1016/S0165-1765\(02\)00262-8](https://doi.org/10.1016/S0165-1765(02)00262-8)
- Xie M, Goh T., 1993. Spc of a near zero-defect process subject to random shocks. *Quality and Reliability Engineering International*. Vol. 9, pp. 89-93. <https://doi.org/10.1002/qre.4680090205>
- Xie, M. He, B. and Goh, T. N., 2001. Zero-inflated Poisson model in statistical process control. *Computational Statistics & Data Analysis*, Vol. 38, pp. 191-201. [https://doi.org/10.1016/S0167-9473\(01\)00033-0](https://doi.org/10.1016/S0167-9473(01)00033-0)

- Yeh, F.B. and Bhattacharya, S., 1998. A robust process capability index. *Communications in Statistics - Simulation and Computation*, Vol. 27, No. 2, pp. 565-589. <https://doi.org/10.1080/03610919808813495>
- Young, D.S., Roemmele, E.S., and Yeh P., 2020^a. Zero-inflated modelling Part I: Traditional zero-inflated count regression models, their applications, and computational tools. *WIREs Computational Statistics*. <https://doi.org/10.1002/wics.1541>.
- Young, D.S., Roemmele, E.S., and Yeh P., 2020^b. Zero-inflated modelling Part II: Zero-inflated models for complex data structures. *WIREs Computational Statistics*. <https://doi.org/10.1002/wics.1540>.
- Zhang, C., Tian, G., and Huang, X. 2015. Two new bivariate zero-inflated generalized Poisson distribution with a flexible correlation structure. *Statistics, Optimization and Information Computing*, Vol. 3, pp. 105–137. <https://doi.org/10.19139/soic.v3i2.133>

Biographical notes

Dr. Surajit Pal is a Faculty Member in the Statistical Quality Control and Operations Research Unit of the Indian Statistical Institute, Chennai Centre, India. His fields of interest are quality engineering, process optimization, statistical quality control and multiple response optimization. He has published about forty five papers in different National and International journals.

Dr. Susanta Kumar Gauri is a Faculty Member in the Statistical Quality Control and Operations Research Unit of the Indian Statistical Institute, Kolkata, India. His fields of interest are quality engineering, process optimization, statistical quality control and multiple response optimization. He has published about sixty papers in different National and International journals.