Noname manuscript No.
(will be inserted by the editor)

A heuristic approach of the estimation of process capability indices for non-normal process data using the Burr XII distribution

Andrea Molina-Alonso ·
Myladis Rocio Cogollo-Flórez

Received: date / Accepted: date

Abstract The Burr type XII distribution plays an important role in a variety of applied mathematics contexts (Watkins, 1999). One of them is the process capability analysis (Ahmad et al., 2009) and the estimation of the distribution parameters is essential for its applications. The estimation with tabulated values is a common method. This paper implement three heuristics to find good solutions for the estimation of the parameters: Particle Swarm Optimization, Median-oriented Particle Swarm Optimization and Artificial Bee Colony. A comparison between the solution given by these methods and other proposed in literature is presented. Finally, the heuristic methods are implemented to estimate the Process Capability Index.

Keywords Burr XII distribution · Parameter estimation · Heuristic methods · Process Capability Index

1 Introduction

Process capability analysis has been an issue addressed in several studies because of its importance in the field of monitoring and quality control of industrial processes. One known method to estimate risk or variability of the processes is the Process Capability Indices (PCIs). PCIs measure how much variation a process experiences according to its specification limits (Miao et al., 2011) and they are generally defined based on three basic assumptions (Liu and Chen, 2006): (i) the system determining which data are collected is under control, (ii) the collected process data are independent and identically

A. Molina-Alonso
Mathematical Engineering, Department of Mathematical Sciences, EAFIT University
E-mail: amolin13@eafit.edu.co

M.R. Cogollo-Flórez, PhD
Department of Mathematical Sciences, EAFIT University
E-mail: mcogollo@eafit.edu.co

distributed, (iii) the collected process data are normally distributed. In practice, industrial production involves many non-normal processes, implying the need of estimation of PCIs where the process output does not follow a normal distribution. Different process capability indices such as Clements percentile method, data transformation method and the Burr based percentile method have been proposed to deal with the non-normal situation (Ahmad et al., 2008).

Liu and Chen (2006) modified Clements method for the estimation of PCIs that follows a normal distribution and uses the Burr XII distribution for estimating the indices for non-normal processes. The authors consider the two parameter Burr distribution, called Burr type XII distribution and introduced by Burr (1942). The specific forms of the cumulative distribution function and the probability density function are

$$F(x) = 1 - \frac{1}{(1+x^c)^k}; \quad x > 0,\ c > 0,\ k > 0 \tag{1}$$

$$f(x) = \frac{k\ c\ x^{c-1}}{(1+x^c)^{k+1}}; \quad x > 0,\ c > 0,\ k > 0 \tag{2}$$

where k is shape parameter and c is scale parameter. Note that changing the values of these parameters, it can be obtained several distributions commonly used such as gamma, lognormal, loglogistic and beta.

Liu and Chen (2006) proposed the following procedure of the estimation of the PCIs using the Burr type XII distribution:

1. Estimate the mean ($\bar{x}$), standard deviation (s), skewness (a_3), and kurtosis (a_4) from the process data.
2. Select the parameter (c) and (k) based on the estimations of skewness and kurtosis coefficients, using the Burr XII distribution table.
3. With those obtained parameters (c) and (k), determine $Z^*_{0.00135}$, $Z^*_{0.5}$ and $Z^*_{0.99865}$ (standardized lower percentile, standardized median and standardized upper percentile respectively) using the table of standardized tails of the Burr XII distribution.
4. Calculate estimated percentiles:
   - $L_p = \bar{x} + Z^*_{0.00135} \cdot s$
   - $M = \bar{x} + Z^*_{0.5} \cdot s$
   - $U_p = \bar{x} + Z^*_{0.99865} \cdot s$
5. Estimate process capability indices using:

$$C_{pu} = \frac{U_t - M}{U_p - M} \tag{3}$$

$$C_{pl} = \frac{M - L_t}{M - L_p} \tag{4}$$

$$C_{pk} = \min\{C_{pu}, C_{pl}\} \tag{5}$$

where U_t is the upper specification limit and L_t is the lower specification limit.

Ahmad et al. (2008) show that the Burr method is better in accuracy than the method of Clements. However the estimation of the parameters of the Burr type XII distribution, in the second step, is not directly but through tabulated values; which leads to problems of underestimation and overestimation of the parameters. In this regard, it is noteworthy that some of the values reported in these tables violate the assumption that the shape and scale parameters of the distribution are strictly positive.

Moreover, in the literature there are several proposals to address the problem of estimating the parameters of Burr type XII distribution. Abbasi et al. (2010) uses a neural network type Multilayer perceptron (MLP) to estimate the distribution parameters, however is not presented a systematic procedure for the construction of model. Malinowska et al. (2006) presents the theoretical development of obtaining the minimum variance linear unbiased estimators (MVLUE), the best linear invariant estimators (BLIE) and the maximum likelihood estimators (MLE) using generalized order statistics for the parameters of the Burr XII distribution. Wang and Lee (2014) use the least squares (LS) method and the M-estimator to estimate the parameters based on the quantile function for complete data with outliers.Watkins (1999) uses the MLE, exploiting the link between the Burr XII and the two parameter Weibull distribution. However these proposals have been made in isolation to the calculus of PCIs and also have a high computational cost due to the use of numerical methods.

The goal of this paper is to propose a new Clements method, using Burr type XII distribution and metaheuristics, that does not require the use of tabulated values and allow accurate estimates. The originality and importance of this paper is framed as follows:

- Although the method of Clements using Burr type XII distribution is used in the industry to monitor processes, the problem of underestimation and overestimation associated with the use of tables in your estimation process can lead to erroneous results. This makes it necessary to explore new methods of estimation that capture the actual behavior of the processes.
- In the literature there is no evidence of the use of metaheuristic techniques in the estimation process of PCIs.

This paper is organized in the following way. The heuristic approach for estimating parameters Burr type XII distribution is presented in Section 2. The new Burr percentile method using metaheuristic techniques is discussed in Section 3. For illustrative purposes, the method proposed in this paper is compared with the conventional method using a simulation study presented in Section 4. Also in this section the properties of the estimators proposed in this paper are discussed. Finally, conclusions are given in Section 5.

2 Parameter estimation of Burr type XII distribution using metaheuristics

2.1 Optimization problem

Given a random sample, to estimate the parameters c and k in (2), conventional methods of estimation: Maximum Likelihood (MLE) and Least Squares (LS) have been considered (Abbasi et al., 2010; Malinowska et al., 2006; Wang and Lee, 2014; Watkins, 1999). Under these methods, the best estimate is given by

$$\theta_{opt} = \arg \min_{\theta \in \mathbf{R}^+ \times \mathbf{R}^+} L_T(\theta), \tag{6}$$

where $\theta = (c, k)$ and

$$L_T(\theta) = n(\ln(c) + \ln(k)) + (c-1)\sum_{i=1}^{n} \ln(x_i) - (k+1)\left(\sum_{i=1}^{n} \ln(1 + x_i^c)\right) \tag{7}$$

is the logarithm of the likelihood function for the Burr XII distribution. In the process of optimization the following equations are obtained:

$$\frac{n}{c} + \sum_{i=1}^{n} \ln(x_i) - (k+1)\sum_{i=1}^{n} \frac{x_i^c \ln(x_i)}{1 + x_i^c} = 0 \tag{8}$$

$$\frac{n}{k} - \left(\sum_{i=1}^{n} \ln(1 + x_i^c)\right) = 0 \tag{9}$$

The set of Equations, (8) and (9), are intractable analytically, and numerical methods must be used to resolve them. The difficulty of mathematical manipulation, negatively impacts the development of statistics asymptotic properties of the estimators (Malinowska et al., 2006; Watkins, 1999). Furthermore, depending on the numerical method used the process of estimating c and k can be very costly in terms of computational time, which limits the use of Burr type XII distribution to build process capability indices of non-normal data in real time.

2.2 Metaheuristic methods

According to Voss (2001), a heuristic is a technique (consisting of a rule or a set of rules) which seeks (and hopefully finds) good solutions at a reasonable computational cost. Meta-heuristic refers to a master strategy that guides and modifies other heuristics to produce solutions beyond those that are normally generated in a quest for local optimality.

For the estimation of the Burr type XII distribution, were implemented three metaheuristic algorithms: Particle Swarm Optimization (PSO) (Marini and Walczak, 2015), Median-oriented Particle Swarm Optimization (MPSO) (Beheshti et al., 2013) and Artificial Bee Colony (ABC) (Karaboga and Basturk, 2008). These algorithms are inspired in the behavior of social animals and are known as optimizers and population-based techniques that are not affected by the size and non-linearity of the problem (Beheshti et al., 2013; Gao et al., 2015; Kennedy, 2010; Marini and Walczak, 2015).

Algorithm 1 PSO pseudocode

Initialize the position $x_i(0) \quad \forall i \in 1 : N$
Initialize the particle's best position $p_i(0) = x_i(0)$
Calculate the fitness of each particle and if $f(x_j(0)) \geq f(x_i(0)) \quad \forall i \neq j$ initialize the global best as $g = x_j(0)$
Until a stopping criterion is met, repeat:
for $i = 1$ to N **do**
  $v_i(t+1) = v_i(t) + c_1(p_i - x_i(t))R_1 + c_2(g - x_i(t))R_2$
  $x_i(t+1) = x_i(t) + v_i(t+1)$
  Evaluate the fitness of the particle $f(x_i(t+1))$
  if $f(x_i(t+1)) \geq f(p_i)$ **then**
    $p_i = x_i(t+1)$
  end if
  if $f(x_i(t+1)) \geq f(g)$ **then**
    $g = x_i(t+1)$
  end if
end for
return g

In Particle Swarm Optimization (PSO) the set of candidate solutions to the optimization problem is defined as a swarm of particles which may flow through the parameter space defining trajectories which are driven by their own and neighbors' best performances (Marini and Walczak, 2015).

The MPSO has the same behavior of the PSO. The difference between them is the way the velocity equation is calculated.

The ABC algorithm simulate the foraging behavior of a honeybee colony. A typical honeybee swarm consists of three fundamental components: employed, onlookers and scout (bees). Onlookers are the bees that currently exploiting, a certain food source. They carry information about the (distance, direction and the profitability) of the food source and communicate the information with other bees waiting at the hive. The onlooker tries to find a food source by means of the information given by an employed bee; while the scout randomly

Algorithm 2 MPSO pseudocode

Initialize the position $x_i(0) \quad \forall i \in 1 : N$
Initialize the particle's best position $p_i(0) = x_i(0)$
Calculate the fitness of each particle and if $f(x_j(0)) \geq f(x_i(0)) \quad \forall i \neq j$ initialize the global best as $g = x_j(0)$
Until a stopping criterion is met, repeat:
for $i = 1$ to N **do**
  if $f(x_{id}(t)) \geq f(p_{id}(t))$ **then**
    $p_{id}(t) = x_{id}(t)$
  end if
  if $f(x_{id}(t)) \geq f(g)$ **then**
    $g = x_{id}(t)$
  end if
  $A_i(t) = \dfrac{fit_i(t) - Maxfit(t)}{Medfit(t) - Maxfit(t)}$
end for
for $i = 1$ to N **do**
  $a_i(t) = \dfrac{A_i(t)}{\sum_{j=1}^{N} A_j(t)}$
  $M_{id}(t) = a_i(t)[rand*(p_{id}(t) - p_{md}(t) - x_{id}(t)) + rand*(g - p_{md}(t) - x_{id}(t))]$
  $v_{id}(t+1) = v_i(t) + M_{id}(t)$
  $x_{id}(t+1) = x_{id}(t) + v_{id}(t+1) + \frac{1}{2} * [rand * (p_{id}(t) - x_{id}(t)) + rand * (g - x_{id}(t))]$
end for
return g

Algorithm 3 ABC pseudocode

Randomly generate SN points in the search space to find an initial population
Evaluate the objective function values of the population
Until a stopping criterion is met, repeat:
Move the employed bees onto their food sources and determine their nectar amounts
Move the onlookers onto the food sources and determine their nectar amounts
Move the scouts for searching new food sources
Memorize the best food source found so far
return best food source

searches the environment to find a new (better) food source. Presumably, an employed bee whose food source is depleted becomes a scout bee, and starts to search for a new food source. Furthermore, it assumes the number of employed bees in the colony to be equal to the number of food sources. Conceivably, the position of a food source represents a possible solution to the optimization problem; whereas the amount of a food source corresponds to the quality (fitness) of the associated solution (Gao et al., 2015; Karaboga and Basturk, 2008)

Algorithms 1, 2 and 3 find good solutions at a reasonable computational cost and they also become simple and easy to implement. Reason for which they were selected in this paper.

3 New Burr percentile method using metaheuristic techniques

The fact that the Burr type XII distribution can be used to describe data that arise in the real world (Liu and Chen, 2006), and its cumulative distribution function (Equation 1) is invertible, makes it a good candidate for process capability analysis. The method to estimate process capability indices under a heuristic approach is as follows:

Step 1. Input the process data (Y), and estimate the mean ($\bar{y}$) and standard deviation (s_Y).

Step 2. Estimate the parameters (c) and (k) using the methaheuristic method.

Step 3. With those estimates obtained ($\widehat{c}$ and $\widehat{k}$), determine the estimated mean and variance of the random variable Burr, given by:

$$\widehat{\mu} = \widehat{k}\, B\left(\widehat{k} - \frac{1}{\widehat{c}},\, 1 + \frac{1}{\widehat{c}}\right)$$
$$\widehat{\sigma}^2 = \widehat{k}\, B\left(\frac{\widehat{ck} - 2}{\widehat{c}},\, \frac{\widehat{c} + 2}{\widehat{c}}\right) - \widehat{\mu}^2$$

where $B(\alpha,\, \beta) = \int_0^1 t^{\alpha-1}(1-t)^{\beta-1}\,\mathrm{d}t, \quad \alpha,\, \beta > 0$, is the beta function.

Step 4. Calculate estimated percentiles:

$$L_p = \bar{y} - \frac{s_Y}{\widehat{\sigma}}\widehat{\mu} + \frac{s_Y}{\widehat{\sigma}} Q(0.00135; \widehat{c}, \widehat{k})$$
$$M = \bar{y} - \frac{s_Y}{\widehat{\sigma}}\widehat{\mu} + \frac{s_Y}{\widehat{\sigma}} Q(0.5; \widehat{c}, \widehat{k})$$
$$U_p = \bar{y} - \frac{s_Y}{\widehat{\sigma}}\widehat{\mu} + \frac{s_Y}{\widehat{\sigma}} Q(0.99865; \widehat{c}, \widehat{k})$$

where $Q(p; \widehat{c}, \widehat{k}) = \{[(1-p)^{(-1/\widehat{k})}] - 1\}^{(1/\widehat{c})}$, with $0 < p < 1$, is the p-th estimated percentile of the distribution Burr.

Step 5. Estimate process capability indices using (3-5).

4 Simulation study

4.1 Properties of the estimators found with metaheuristics

To determine if the results obtained with the algorithms were consistent, a series of simulations were implemented with sample sizes of $n = 30, 50, 100, 500$

and 1000. Each run was replicated 1000 times to yield the average of 1000 ($\hat{c}$) and ($\hat{k}$) values. The simulation procedure is as follows:

1. Choose the targeted c and k values. For this simulation were used the values $c = 2$ and $k = 1$.
2. Generate $n = 30, 50, 100, 500$ or 1000 sample data points that follow a Burr XII distribution with the selected c and k values. For this, it was used the inverse cumulative function
$$x_i = \left[\left(\frac{1}{1-y} \right)^{1/k} - 1 \right]^{1/c}$$
where $y \sim Uniform(0, 1)$.
3. Estimate the parameters (c) and (k) using the methaheuristic methods. Repeat the procedure with each of the estimation methods.

With this procedure, Figures 1-3 were obtained. The figures presents the box plot of $\hat{c}$ and $\hat{k}$ by applying the PSO, MPSO and ABC metaheuristic. Clearly, the average of the estimated values are considerably close to the target values while the sample size increase. This indicates that the values of the estimates are consistent, except in Figure 3 where consistency with this algorithm is affected by outliers.

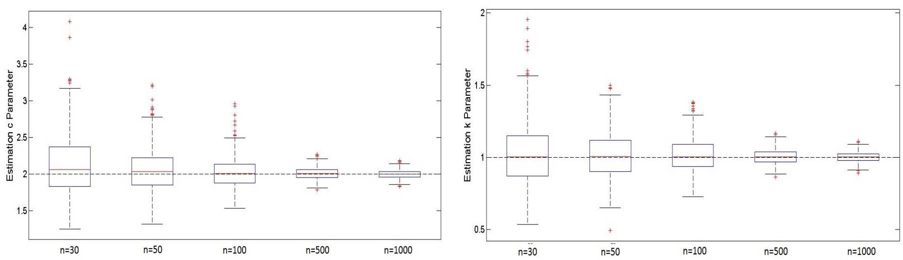

Fig. 1: Box plots of the PSO algorithm with target values $c = 2$ and $k = 1$ with different sample sizes

Moreover, Table 1 presents a comparison of the results obtained with the three algorithms implemented with the ones obtained with Neural Networks by Abbasi et al. (2010) and other obtained with tabulated values (Liu and Chen, 2006). Results obtained with the PSO, MPSO and ABC algorithms, in most of the cases, are more accurate and close to the real value than the others obtained with different methods. Additionally, a larger sample size n yields better estimates in each of the estimation methods, which indicates consistence on the methods.

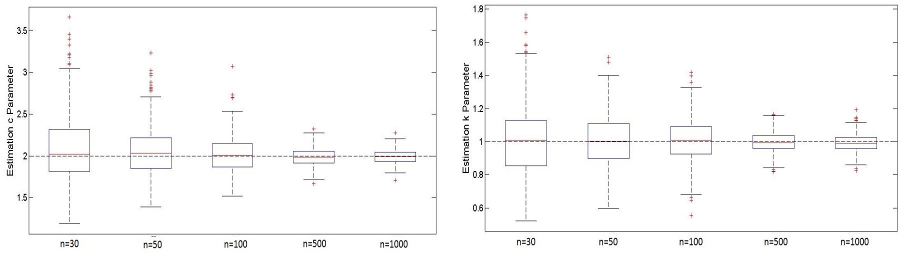

Fig. 2: Box plots of the MPSO algorithm with target values $c = 2$ and $k = 1$ with different sample sizes

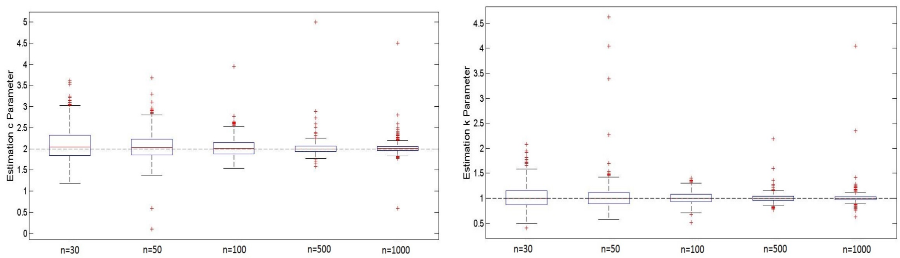

Fig. 3: Box plots of the ABC algorithm with target values $c = 2$ and $k = 1$ with different sample sizes

4.2 PCI with new method

Since Burr type XII distribution and its variants assess processes whose data have a positive asymmetry, to examine the effect of using metaheuristics techniques in the calculation of PCI, it is considered the upper one sided capability index C_{pu}. In this case, it was just considered an upper specification limit, which in this simulation study was assumed as $U_t = 2.2943$.
To calculate it were considered Steps 1-5 in Section 3, and the estimations obtained in Table 1 with the parameter values of $c = 2$ and $k = 5$ and the four sample sizes. Results are presented in Table 2.

According to the results obtained in this section, the estimations obtained with the PSO and the MPSO algorithms are closer to the real value showed in the second column of Table 2. Additionally, with the estimation method of tabulated values (Liu and Chen, 2006), was obtained a negative estimation of the C_{pu}, having no sense for a positive asymmetric function in real world.

Table 1: Comparison of the results of methods with other proposed in the literature

Parameters	n	PSO	MPSO	ABC	Abbasi et al. (2010)	Liu and Chen (2006)
c=4,14224 k=9,13497	100	c=4,1007 k=8,3162	c=4,0761 k=8,2304	c=4,1265 k=8,2988	c=3,9113 k=10,1661	c=4,8737 k=6,1575
	1000	c=4,1620 k=9,4552	c=4,1625 k=9,4310	c=4,1620 k=9,4458	c=4,0531 k=9,0436	c=4,87371 k=6,15756
	2500	c=4,1422 k=9,3333	c=3,8042 k=8,0458	c=4,1270 k=9,2339	c=4,0758 k=9,0610	c=4,87371 k=6,15756
	10000	c=4,1264 k=9,1480	c=4,0712 k=8,6973	c=4,1278 k=9,1531	c=4,0998 k=9,0781	c=3,93893 k=19,864823
c=2 k=5	100	c=1,8809 k=4,6006	c=1,8442 k=4,6556	c=2,0498 k=5,5355	c=1,6068 k=4,2142	c=-8,83754 k=0,09995
	1000	c=2,0456 k=5,1508	c=2,0242 k=5,1657	c=2,2792 k=5,7089	c=1,8145 k=4,5983	c=2,55182 k=3,68994
	2500	c=1,9975 k=5,2704	c=1,9681 k=5,0915	c=2,0669 k=5,6685	c=1,8841 k=4,7489	c=3,58714 k=2,19903
	10000	c=1,9931 k=4,9063	c=1,9878 k=4,9445	c=2,0829 k=5,7299	c=1,9586 k=4,9288	c=2,770828 k=3,17098
c=3 k=4	100	c=2,7973 k=3,7460	c=2,7887 k=3,7531	c=2,9279 k=4,4239	c=2,5043 k=3,5291	c=3,22768 k=2,46376
	1000	c=2,9872 k=4,2007	c=2,9919 k=4,3154	c=3,1187 k=5,2343	c=2,8966 k=3,8751	c=2,53777 k=12,52234
	2500	c=3,0094 k=4,0019	c=3,0129 k=3,9275	c=3,0707 k=4,2874	c=2,8988 k=3,8987	c=-7,17600 k=0,07865
	10000	c=2,9870 k=4,0126	c=3,0044 k=3,8206	c=2,9870 k=4,0126	c=2,9576 k=3,9662	c=-7,17600 k=0,07865

Table 2: Comparison of the real C_{pu} with the estimations obtained using the metaheuristics methods and the other conventional methods of estimation

n	Real	PSO	MPSO	ABC	Abbasi et al. (2010)	Liu and Chen (2006)
100	1,5519	1,4596	1,4468	1,6166	1,2823	-5,4809
1000	1,4546	1,4859	1,4778	1,6211	1,3376	1,5253
2500	1,5250	1,5452	1,5177	1,6055	1,4493	1,5461
10000	1,4938	1,4830	1,4838	1,5837	1,4693	1,5515

5 Conclusions

We proposed a new method of type Burr percentile using metaheuristic techniques that provide more accurate estimates than those found with the traditional method using tabulated values. The proposed method is characterized allowed to work in real time with no normal data, and therefore better capture the behavior of the real processes.

References

Abbasi, B., Hosseinifard, S. Z., and Coit, D. W. (2010). A neural network applied to estimate Burr XII distribution parameters. *Reliability Engineering & System Safety*, 95(6):647–654.

Ahmad, S., Abdollahian, M., and Zeephongsekul, P. (2008). Process capability estimation for non-normal quality charactersitics using Clement, Burr and Box–Cox methods. *ANZIAM Journal*, 49:642–665.

Ahmad, S., Abdollahian, M., Zeephongsekul, P., and Abbasi, B. (2009). Multivariate nonnormal process capability analysis. *The International Journal of Advanced Manufacturing Technology*, 44(7-8):757–765.

Beheshti, Z., Shamsuddin, S. M. H., and Hasan, S. (2013). MPSO: Median-oriented Particle Swarm Optimization. *Applied Mathematics and Computation*, 219(11):5817–5836.

Burr, I. W. (1942). Cumulative frequency functions. *The Annals of Mathematical Statistics*, 13(2):215–232.

Gao, W., Chan, F. T., Huang, L., and Liu, S. (2015). Bare bones artificial bee colony algorithm with parameter adaptation and fitness-based neighborhood. *Information Sciences*, 316:180–200.

Karaboga, D. and Basturk, B. (2008). On the performance of artificial bee colony (ABC) algorithm. *Applied soft computing*, 8(1):687–697.

Kennedy, J. (2010). Particle swarm optimization. In *Encyclopedia of Machine Learning*, pages 760–766. Springer.

Liu, P.-H. and Chen, F.-L. (2006). Process capability analysis of non-normal process data using the Burr XII distribution. *The International Journal of Advanced Manufacturing Technology*, 27(9-10):975–984.

Malinowska, I., Pawlas, P., and Szynal, D. (2006). Estimation of location and scale parameters for the Burr XII distribution using generalized order statistics. *Linear algebra and its applications*, 417(1):150–162.

Marini, F. and Walczak, B. (2015). Particle swarm optimization (PSO). A tutorial. *Chemometrics and Intelligent Laboratory Systems.*

Miao, R., Zhang, X., Yang, D., Zhao, Y., and Jiang, Z. (2011). A conjugate Bayesian approach for calculating process capability indices. *Expert Systems with Applications*, 38(7):8099–8104.

Voss, S. (2001). Meta-heuristics: The state of the art. In *Local Search for Planning and Scheduling*, volume 2148 of *Lecture Notes in Computer Science*, pages 1–23. Springer Berlin Heidelberg.

Wang, F.-K. and Lee, C.-W. (2014). M-estimator for estimating the Burr type III parameters with outliers. *Mathematics and Computers in Simulation*, 105:144–159.

Watkins, A. J. (1999). An algorithm for maximum likelihood estimation in the three parameter Burr XII distribution. *Computational statistics & data analysis*, 32(1):19–27.