REVSTAT – Statistical Journal
Volume 23, Number 1, January 2025, 1–17
https://doi.org/10.57805/revstat.v23i1.492

Control Monitoring Schemes for Percentiles of Generalized Exponential Distribution with Hybrid Censoring

Authors: Shovan Chowdhury
– Quantitative Methods and Operations Management Area,
Indian Institute of Management,
Kozhikode, India
shovanc@iimk.ac.in

Amarjit Kundu ✉
– Department of Mathematics, Raiganj University,
West Bengal, India
bapai_k@yahoo.com

Bidhan Modok
– Department of Mathematics, Raiganj University,
West Bengal, India
bidhanmodok95@gmail.com

Received: June 2021 Revised: December 2022 Accepted: December 2022

Abstract:

- In this article, a parametric bootstrap "control monitoring scheme" equivalently known as "control chart", is proposed for process monitoring of percentiles of the generalized exponential distribution for type-I hybrid censored data assuming in-control parameters to be unknown. Monte Carlo simulations are carried out to evaluate the in-control and out-of-control performance of the proposed scheme in terms of average run lengths. Conventional Shewhart-type scheme is also proposed under the same set-up asymptotically and compared with bootstrap scheme using a skewed data set. Finally, an application of the proposed scheme is shown from clinical practice.

Keywords:

- *average run length; control monitoring scheme; false alarm rate; generalized exponential distribution; hybrid censoring; parametric bootstrap.*

AMS Subject Classification:

- 62P30 (Primary), 62F40 (Secondary), 62N01 (Secondary).

✉ Corresponding author

1. INTRODUCTION

Researchers have often shown interest in developing control charts, also known as control monitoring scheme, for monitoring percentiles of an underlying distribution in reliability studies. Padgett and Spurrier (1990) and Nichols and Padgett (2005) argued in favor of monitoring lower percentile of strength distribution over average to confirm the quality of carbon fiber strength to be in control. Chowdhury *et al.* (2021) emphasized on monitoring the upper (lower) percentile of the proportion of non-conforming (conforming) units, as an upward (downward) shift in the upper (lower) percentile indicates a deterioration in quality. All the available control schemes for monitoring distribution quantiles (see for example, Padgett and Spurrier, 1990; Nichols and Padgett, 2005; Erto and Pallotta, 2007; Lio and Park, 2008, 2010; Lio *et al.* (2014); Erto *et al.* (2015); Chiang *et al.*, 2017, 2018; Leiva *et al.*, 2022; and Ma *et al.*, 2022) used parametric bootstrap method with different distributional assumptions under classical or/and Bayesian set-up. Additionally, Chiang *et al.* (2018) used model selection approaches to choose between competing underlying distributions. Due to the non-availability of closed form expressions of the sampling distribution of the percentiles, computational methods such as parametric bootstrap is used to obtain the control limits. For more discussion on the bootstrap technique and its advantages, one can refer to Efron and Tibshirani (1993), Liu and Tang (1996), Jones and Woodall (1998) and Seppala *et al.* (1996).

The results obtained from the aforementioned papers are useful and valuable, and can be applied to complete data setting only. In practice, reliability data are skewed and censored. Recently Vining *et al.* (2016) emphasized on using censored data in reliability studies as customers expected products and processes to perform with high quality over the entire expected lifetime of the product/process. Most of the available schemes for censored data monitor mean of a process. Few papers are found in the literature for monitoring percentiles of a process using censored data. Haghighi *et al.* (2015) proposed control charts for the quantiles of the Weibull distribution for type-II censored data, based on the distribution of a pivotal quantity conditioned on ancillary statistics. Wang *et al.* (2018) proposed EWMA and CUSUM charts for monitoring the lower Weibull percentiles under complete data and type-II censoring using the same approach as used in Haghighi *et al.* (2015). Encouraged by these findings, in this paper, a control monitoring scheme is proposed based on bootstrap method using hybrid censoring which generalizes control monitoring schemes under type-I and type-II censoring.

In type-I censoring scheme, the experiment is aborted after a pre-decided time $T = x_0$; whereas in type-II censoring, the termination is subject to failure of a pre-fixed number of items r. The hybrid censoring scheme which is popularly known in the literature as type-I hybrid censoring scheme was initially introduced by Epstein (1954) and can be considered as a mixture of type-I and type-II censoring schemes. It can be described briefly as follows: Suppose n identical units are put on an experiment. Now if $X_{1:n}, \ldots, X_{n:n}$ are the ordered lifetimes of the units, then the experiment is aborted either when a pre-chosen number $r(< n)$ out of n items has failed or when a pre-determined time x_0 has elapsed. Hence the life test can be terminated at a random time $X^* = \min\{X_{r:n}, T\}$. One of the following two types of observations can be witnessed under type-I hybrid censoring scheme.

Case I: $\{X_{1:n} < ... < X_{r:n}\}$ if $X_{r:n} < x_0$.

Case II: $\{X_{1:n} < ... < X_{d:n} < T\}$ if $d+1 \le r < n$ and $x_0 \le X_{r:n}$.

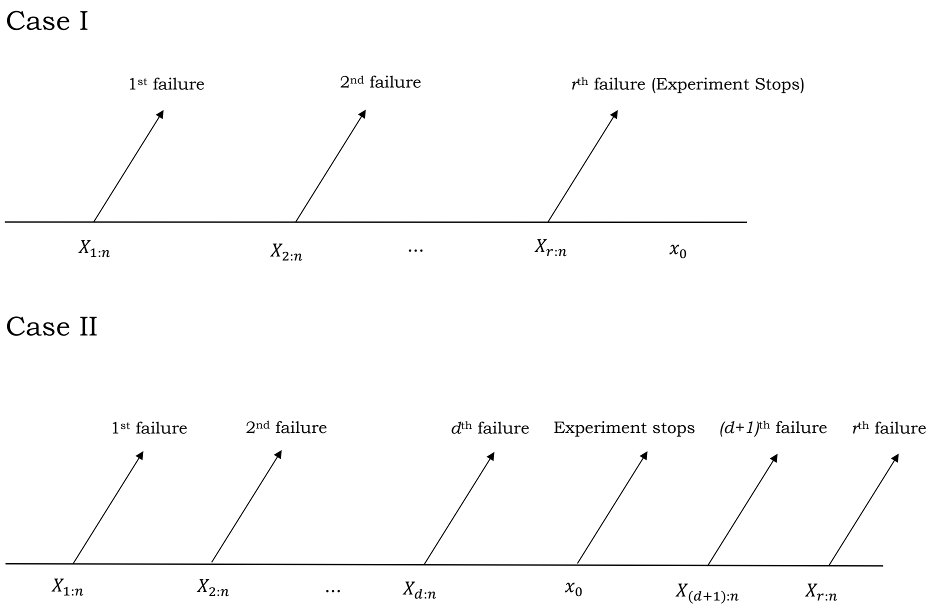


Figure 1: Schematic illustration of type-I hybrid censoring scheme.

In reliability studies, two-parameter Weibull is the most popular distribution to the practitioners. Gupta and Kundu (1999) proposed the two parameter generalized exponential (GE) distribution as an alternative to the Weibull and studied its properties extensively. The scale and shape parameters of the GE distribution bring quite a bit of flexibility in the distribution to analyze any positive real data. Both the Weibull and GE distributions have increasing or decreasing failure rates depending on the shape parameter. Many authors pointed out (see for example, Bain, 1976) that since the hazard function of a GE distribution is bounded above or bounded below as opposed to Weibull which is unbounded, the GE may be more appropriate as a population model when the items in the population are in a regular maintenance environment. The hazard rate may increase initially, but after some times the system reaches a stable condition because of maintenance. Therefore, if it is known that the data are from a regular maintenance environment, it may make more sense to fit the GE distribution over the Weibull. As opposed to Weibull distribution, GE represents a parallel system of independent and identically distributed exponential components. GE has likelihood ratio ordering on the shape parameter indicating the possibility of constructing a uniformly most powerful test for testing a one-sided hypothesis on the shape parameter keeping the scale parameter known. The Weibull distribution doesn't enjoy any such ordering properties and hence no such uniformly most powerful test exists for Weibull. One of the disadvantages of Weibull can be pointed out that the asymptotic convergence to normality for the distribution of the maximum likelihood estimators is very slow (Bain, 1976). Therefore most of the asymptotic inferences may not be very accurate unless the sample size is very large. For a detailed comparison between Weibull and GE, one can refer to Gupta and Kundu (2001). Motivated by these findings, GE is chosen as the underlying distribution to develop a bootstrap control monitoring scheme for hybrid censored data.

The rest of the paper is organized as follows. Section 2 provides the statistical background of the paper. The proposed bootstrap and Shewhart-type control monitoring schemes for GE percentiles with hybrid censored data are introduced in Section 3. Section 4 is devoted to the practical implementation of the schemes including tabulation of the control limits and average run length (ARL). Simulation results of both in-control (IC) and out-of-control (OOC) performance of the bootstrap scheme are presented in Section 4. The effectiveness of the proposed scheme is evaluated in Section 5 using a skewed data set. Bootstrap control monitoring scheme for type-I and type-II censored data are also obtained in Section 5 as a special case and are compared with bootstrap chart under hybrid censoring scheme. An application of the proposed scheme is shown from clinical practice in Section 6. Section 7 concludes the paper.

2. STATISTICAL FRAMEWORK

2.1. Maximum likelihood estimators

Let X be a random variable following two parameter GE distribution with the shape parameter $\theta > 0$ and scale parameter $\lambda > 0$. Then probability density function (pdf) and cumulative distribution function (cdf) of X are given by

$$f(x|\theta,\lambda) = \theta\lambda e^{-\lambda x}(1-e^{-\lambda x})^{\theta-1} \tag{2.1}$$

and

$$F(x|\theta,\lambda) = \Big(1-e^{-\lambda x}\Big)^{\theta}. \tag{2.2}$$

Let ξ_p be the $100p^{\text{th}}$ percentile of the GE distribution and is obtained as

$$\xi_p = -\frac{1}{\lambda}\ln\Big(1-p^{\frac{1}{\theta}}\Big). \tag{2.3}$$

Now, let $x_{i_1}, x_{i_2}, ..., x_{i_n}$ be i^{th} in-control (IC) random subgroup of size n $(i = 1, 2, ..., k)$ drawn from phase I process following GE distribution as in (2.1). On the basis of the observed data and ignoring the additive constant, the log-likelihood function under hybrid censoring (for Case I and II as introduced in Section 1) is given by

$$\begin{aligned} L(\theta,\lambda|data) \;&=\; d\ln\theta + d\ln\lambda - \lambda\sum_{i=1}^{d} x_{i:n} + (\theta-1)\sum_{i=1}^{d}\ln(1-e^{-\lambda x_{i:n}}) \\ &+\; (n-d)\ln\Big(1-(1-e^{-\lambda c})^{\theta}\Big). \end{aligned} \tag{2.4}$$

Note that for Case I, $d = r$ and $c = x_{r:n}$, and for Case II, $0 \le d \le r-1$ and $c = x_0$. Also it can be shown that for $\lambda \to 0$, and for any fixed θ, maximum likelihood estimators (MLE) of θ and λ do not exist when $d = 0$. Assuming d to be positive, the MLEs $\hat{\theta}$ and $\hat{\lambda}$ are obtained by maximizing the log-likelihood function (2.4), and subsequently solving the non-linear equations

$$\frac{\partial L}{\partial\theta} = 0,\ \frac{\partial L}{\partial\lambda} = 0.$$

As closed-form solutions of these two equations are not available, EM algorithm is used to obtain the MLEs. Let the observed data and the censored data be denoted by $\mathbf{X} = (x_{1:n}, ..., x_{d:n})$ and $\mathbf{Y} = (y_1, ..., y_{n-d})$ respectively. Here for given d, $\mathbf{Y}$ is not observable and hence can be thought of as missing data. The combination of $\mathbf{Z} = (\mathbf{X}, \mathbf{Y})$ forms the complete data set. Ignoring the additive constant, the log-likelihood function of the uncensored data set, denoted by $L_c(\theta, \lambda|Z)$ is given by

$$\begin{aligned} L_c(\theta, \lambda|data) &= n \ln \theta + n \ln \lambda - \lambda \left(\sum_{i=1}^{d} x_{i:n} + \sum_{i=1}^{n-d} y_i \right) \\ &+ (\theta - 1) \left(\sum_{i=1}^{d} \ln\left(1 - e^{-\lambda x_{i:n}}\right) + \sum_{i=1}^{n-d} \ln\left(1 - e^{-\lambda y_i}\right) \right). \end{aligned} \tag{2.5}$$

Now for 'E'-step of the EM algorithm, one needs to compute the pseudo log-likelihood function as $L_s(\theta, \lambda|data) = E(L_c(Z; \theta, \lambda)|X)$, obtained as

$$\begin{aligned} L_s(\theta, \lambda|data) &= n \ln \theta + n \ln \lambda - \lambda \sum_{i=1}^{d} x_{i:n} + (\theta - 1) \sum_{i=1}^{d} \ln\left(1 - e^{-\lambda x_{i:n}}\right) \\ &- \lambda \sum_{i=1}^{n-d} E[Y_i|Y_i > c] + (\theta - 1) \sum_{i=1}^{n-d} E\Big[\ln(1 - e^{-\lambda Y_i})|Y_i > c\Big]. \end{aligned} \tag{2.6}$$

Now the 'M'-step involves maximization of the pseudo log-likelihood function given in (2.6). Therefore, if at the k^{th} stage the estimate of (θ, λ) is (θ^k, λ^k), then $(\theta^{(k+1)}, \lambda^{(k+1)})$ can be obtained by maximizing

$$\begin{aligned} g(\theta, \lambda) &= n \ln \theta + n \ln \lambda - \lambda \sum_{i=1}^{d} x_{i:n} + (\theta - 1) \sum_{i=1}^{d} \ln\left(1 - e^{-\lambda x_{i:n}}\right) \\ &- \lambda(n - d) A\left(c, \theta^{(k)}, \lambda^{(k)}\right) + (\theta - 1)(n - d) B\left(c, \theta^{(k)}, \lambda^{(k)}\right), \end{aligned} \tag{2.7}$$

where

$$\begin{aligned} A(c, \theta, \lambda) &= -\frac{\theta}{\lambda(1 - F(c, \theta, \lambda))} u(\lambda c, \theta), \\ B(c, \theta, \lambda) &= \frac{1}{\theta(1 - F(c, \theta, \lambda))} \left[\left(1 - e^{-c\lambda}\right)^{\theta} \left(1 - \theta \ln\left(1 - e^{-c\lambda}\right)\right) - 1 \right]. \end{aligned}$$

The maximization of (2.7) can be performed by using similar technique as of Gupta and Kundu (2001). First, $\lambda^{(k+1)}$ can be obtained by solving a fixed point type equation $h(\lambda) = \lambda$, where the function $h(\lambda)$ is defined as

$$h(\lambda) = \left[\frac{1}{n} \sum_{i=1}^{d} x_{i:n} + \frac{n - d}{n} A - \frac{1}{n} \left(\hat{\theta}(\lambda) - 1\right) \sum_{i=1}^{d} \frac{x_{i:n} e^{-\lambda x_{i:n}}}{1 - e^{-\lambda x_{i:n}}} \right]^{-1},$$

with $A = A\left(c, \theta^{(k)}, \lambda^{(k)}\right)$, $B = B\left(c, \theta^{(k)}, \lambda^{(k)}\right)$ and $\hat{\theta}(\lambda) = -\frac{n}{\sum_{i=1}^{d} \ln\left(1 - e^{-\lambda x_{i:n}}\right) + (n-d)B}$. Once $\lambda^{(k+1)}$ is obtained, $\theta^{(k+1)}$ is obtained by solving the equation $\theta^{(k+1)} = \hat{\theta}\left(\lambda^{(k+1)}\right)$. For more detail on the estimation of GE parameters under hybrid censoring (see Kundu and Pradhan, 2009).

The MLE of the $100p^{\text{th}}$ percentiles, denoted by $\hat{\xi_p}$, is also obtained as

$$\hat{\xi_p} = -\frac{1}{\hat{\lambda}} \ln\left(1 - p^{\frac{1}{\hat{\theta}}}\right). \tag{2.8}$$

2.2. Asymptotic properties

An outline of the Fisher information matrix and asymptotic properties of the estimators are discussed here. For more detail, one may refer to Gupta and Kundu (2001) and Kundu and Pradhan (2009). Using the missing value principle of Louis (1982), it can be written that

$$\text{Observed information} = \text{Complete information} - \text{Missing information}, \tag{2.9}$$

and can be expressed as

$$I_X(\Theta) = I_W(\Theta) - I_{W|X}(\Theta), \tag{2.10}$$

with $\Theta = (\theta, \lambda)$; X = the observed vector; W = the complete data; $I_W(\Theta)$ = the complete information; $I_{W|X}(\Theta)$ = the missing information. The complete information $I_W(\Theta)$ is given by

$$I_W(\Theta) = -E\left[\frac{\partial^2 L_c(W;\Theta)}{\partial\Theta^2}\right],$$

with the Fisher information matrix of the censored observations being written as

$$I_{W|X}(\Theta) = -(n-d)E_{Z|X}\left[\frac{\partial^2 \ln f_Z(z|X,\Theta)}{\partial\Theta^2}\right].$$

The asymptotic variance covariance matrix of $\hat{\Theta}$, the MLE of Θ, can be obtained by inverting $I_X\left(\hat{\Theta}\right)$. The elements of the matrix $I_X(\Theta)$ for the complete data set can be obtained in Kundu and Pradhan (2009).

Let $\hat{\xi}_p\left(\hat{\Theta}_n\right)$ be the value of ξ_p at $\Theta = \hat{\Theta}_n$, obtained from (2.3) and calculated on the basis of n observations. Then as in Chiang *et al.* (2017), it can be shown that $\hat{\xi}_p\left(\hat{\Theta}_n\right)$ follows asymptotic normal distribution with mean $\xi_p(\Theta)$ and variance $\frac{1}{n}\nabla\xi_p^T(\boldsymbol{\Theta})\mathbf{I_Y}^{-1}(\boldsymbol{\Theta})\nabla\xi_p(\boldsymbol{\Theta})$, where $\nabla\xi_p(\Theta)$ is the gradient of $\xi_p(\Theta)$ with respect to Θ. In practice, $\mathbf{I_Y}(\boldsymbol{\Theta})$ is replaced by the observed Fisher Information matrix $\hat{\mathbf{I}}_Y\left(\hat{\boldsymbol{\Theta}}_n\right)$, obtained by substituting the unknown parameters θ and λ by their respective MLEs.

3. CONSTRUCTION OF PROPOSED CONTROL MONITORING SCHEMES

3.1. Charting procedure for the bootstrap hybrid-censored control (BHCC) monitoring scheme

Here, the bootstrap hybrid-censored control (BHCC) monitoring scheme for GE percentiles is developed using the following charting procedure.

Step-1: Collect and establish k reference samples $X_m = (x_{i1}, x_{i2}, ..., x_{im})$ of size m each from an IC process (Phase I process) following GE cdf $F(x|\theta,\lambda)$ as in (2.2).

Step-2: Obtain the MLEs of θ and λ from Step-1 under hybrid censoring following the procedure described in Section 2 and estimate the cdf as $F(x|\hat{\theta}, \hat{\lambda})$.

Step-3: Generate a bootstrap sample of size m, $x_1^*, x_2^*, ..., x_m^*$, from $F(x|\hat{\theta}, \hat{\lambda})$ as obtained in Step-2.

Step-4: Obtain the MLEs of θ and λ under hybrid censoring using the bootstrap sample obtained in Step-3, and denote these as θ^* and λ^*.

Step-5: Using (2.3) and (2.8), compute the bootstrap estimate of the $100p^{\text{th}}$ percentile as

$$\hat{\xi}_p^* = -\frac{1}{\lambda^*} \ln\left(1 - p^{\frac{1}{\theta^*}}\right). \tag{3.1}$$

Step-6: Repeat Steps 3-5 large number of times (B) to obtain bootstrap estimates of $\hat{\xi}_p^*$, denoted by $\hat{\xi}_{1p}^*, \hat{\xi}_{2p}^*, ..., \hat{\xi}_{Bp}^*$.

Step-7: Using B bootstrap estimates as obtained in Step 6, find the $\frac{\nu}{2}^{\text{th}}$ and $(1-\frac{\nu}{2})^{\text{th}}$ empirical percentiles as the lower control limit (LCL) and upper control limit (UCL) respectively to construct a two-sided BHCC chart, where ν is the false alarm rate (FAR) defined as the probability that an observation is considered as out of control (OOC) when the process is actually IC. Here, empirical sample percentiles are obtained following a method proposed by Hyndman and Fan (1996).

Step-8: Sequentially observe the j^{th} phase II (test) sample $Y_{j:m} = (Y_{j1}, Y_{j2}, ..., Y_{jm})$ of size $m, j = 1, 2, ...$

Step-9: Sequentially obtain $\hat{\xi}_{jp}$ using (3.1) after obtaining MLEs of the parameters under hybrid censoring scheme using the j^{th} test sample as described in Step-5.

Step-10: Plot $\hat{\xi}_{jp}$ against LCL and UCL as obtained in Step-7 of the Phase I process.

Step-11: If $\hat{\xi}_{jp}$ falls in between the LCL and UCL, then the process is assumed to be in-control, otherwise, an OOC signal is activated.

3.2. Charting procedure for the Shewhart-type hybrid-censored control (SHCC) monitoring scheme

Shewhart-type control monitoring scheme for the percentiles of GE distribution, named as SHCC scheme is derived in this section following the asymptotic properties of the MLEs obtained in Section 2.2. The steps for designing the SHCC scheme for $100p^{\text{th}}$ percentile of proportion, $\xi_p(\Theta)$, are described as follows.

In phase I, samples are drawn from in-control process following GE distribution in k independent random subgroups of size m each with $n = m \times k$ being the total sample size.

Step-1: As described in Section 2.1, the MLEs $\hat{\Theta}_n = \left(\hat{\theta}_n, \hat{\lambda}_n\right)$ are computed on the basis of n in-control sample values of Phase I. Then the asymptotic standard error of $\hat{\xi}_{p,m}\left(\hat{\Theta}_m\right)$ is computed as

$$SE_{\xi_{p,m}} = \sqrt{\frac{1}{m}\nabla\xi_p^T\left(\hat{\mathbf{\Theta}}_n\right)\mathbf{I}_Y^{-1}\left(\hat{\mathbf{\Theta}}_n\right)\nabla\xi_p\left(\hat{\mathbf{\Theta}}_n\right)}, \tag{3.2}$$

where $\nabla\xi_p\left(\hat{\mathbf{\Theta}}_n\right)$ is the gradient of $\xi_p(\mathbf{\Theta})$ at $\mathbf{\Theta} = \hat{\mathbf{\Theta}}_n$. $\mathbf{I}_Y^{-1}\left(\hat{\mathbf{\Theta}}_n\right)$ is calculated following the procedure as described in Section 2.2.

Step-2: The MLEs $\hat{\mathbf{\Theta}}_m^j$ of Θ and $\xi_p^j\left(\hat{\mathbf{\Theta}}_m^j\right)$ are calculated based on j^{th} $(j = 1, 2, ..., k)$ IC samples of size m each. The sample mean of $\xi_p^j\left(\hat{\mathbf{\Theta}}_m^j\right)$s is calculated as

$$\bar{\xi}_p\left(\hat{\mathbf{\Theta}}_m\right) = \frac{1}{k}\sum_{i=1}^{k}\xi_p^j\left(\hat{\mathbf{\Theta}}_m^j\right).$$

Step-3: The Shewhart-type control monitoring scheme has the center line $CL_{SH} = \bar{\xi}_p\left(\hat{\mathbf{\Theta}}_m\right)$. If ν is the false alarm rate (FAR), then for $0 < \nu < 1$, the upper and lower control limits of the SHCC scheme are found to be

$$UCL_{SH} = \bar{\xi}_p\left(\hat{\mathbf{\Theta}}_m\right) + z_{(1-\nu/2)}SE_{\xi_{p,m}},$$

and

$$LCL_{SH} = \bar{\xi}_p\left(\hat{\mathbf{\Theta}}_m\right) - z_{(1-\nu/2)}SE_{\xi_{p,m}},$$

respectively, where $z_{(1-\nu/2)}$ is the $(1-\nu/2)^{\text{th}}$ quantile of standard normal distribution.

4. SIMULATION STUDY

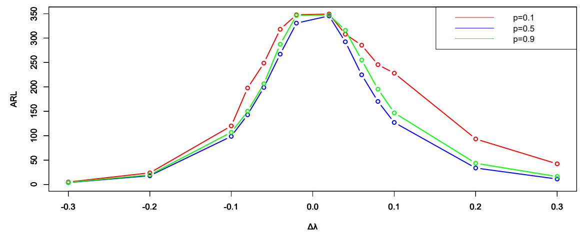
In this section, the IC and OOC performances of the proposed BHCC monitoring scheme are evaluated through a comprehensive simulation study. Numerical computations in R (version 4.0.2) based on Monte-Carlo simulations are used to determine the average UCL and LCL. The MLEs of the parameters θ and λ are obtained for the pair ($\theta = 5.5$, $\lambda = 0.05$). The control limits are obtained based on $B = 5,000$ bootstrap samples. The simulations are carried out with different bootstrap sample sizes m with $k = 20$ subgroups, different percentiles ($p = 0.1, 0.5, 0.9$), different levels of FAR ($\nu = 0.1, 0.005, 0.0027, 0.002, 0.001$) and the following censoring schemes: Scheme $1:$ $m = 25$, $r = 15$, $x_0 = 55$; Scheme $2:$ $m = 25$, $r = 20$, $x_0 = 55$; Scheme $3:$ $m = 40$, $r = 30$, $x_0 = 55$; Scheme $4:$ $m = 40$, $r = 35$, $x_0 = 55$; Scheme $5:$ $m = 25$, $r = 15$, $x_0 = 70$; Scheme $6:$ $m = 25$, $r = 20$, $x_0 = 70$; Scheme $7:$ $m = 40$, $r = 30$, $x_0 = 70$; and Scheme $8:$ $m = 40$, $r = 35$, $x_0 = 70$. The performance of the scheme is assessed by run length, defined as the number of cases required to observe the first OOC signal. For each simulation, the run length is obtained, followed by obtaining the average run length (ARL) and the standard deviation of run length (SDRL) by using $5,000$ simulation runs.

4.1. IC monitoring scheme performance

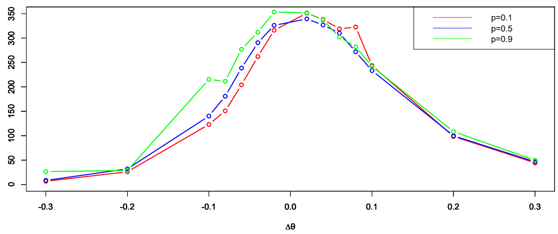
The estimated IC control limits of the BHCC scheme are displayed in Table 1 of the supplementary article, along with the respective ARL and SDRL as the scheme performance measures, denoted by ARL_0 and $SDRL_0$ respectively. It is easy to show that the reciprocal of FAR is same as the nominal (theoretical) ARL, viz. for $\nu = 0.1, 0.005, 0.0027, 0.002$ and 0.001, the nominal ARL should be equal to $10, 200, 370$, 500 and 1000 respectively. In general, the smaller ARLs indicate narrower control limits, while ARLs larger than 370 specifies wider limits that the bootstrap control schemes give fewer false signals. The simulated values of ARL_0 in Table 1 are found to be closer to the theoretical results implying that the BHCC monitoring scheme for percentiles perform well with skewed data. As the bootstrap sample size (m) increases, the estimated control limits get closer together. Moreover, for fixed m, the control limits become farther apart as the percentile (p) increases. Also, $SDRL_0$ is found to be closer to the ARL_0, satisfying the theoretical result of the geometric distribution used as the run length model.

4.2. OOC monitoring scheme performance

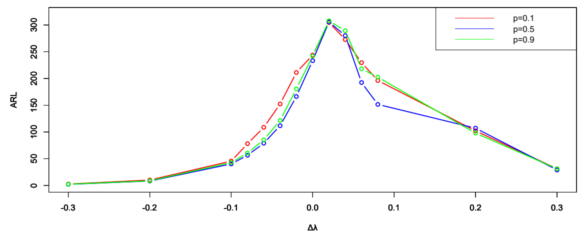
The OOC performance of the BHCC monitoring scheme is investigated by measuring impact of changes in the IC parameter estimates on ARL. In other words, the phase II sample is considered taken from $GE(\theta + \Delta\theta, \lambda + \Delta\lambda)$, while the IC sample comes from $GE(\theta, \lambda)$. The effects of shifts ($\Delta\theta$ and/or $\Delta\lambda$) in the parameters of the GE distribution on ARL of the percentile scheme is examined and exhibited in Table 2 of the supplementary article. In general, the simulation results reveal that for fixed m, r, and x_0, the OOC ARL values (denoted by ARL_1) for the percentiles decrease sharply with both downward and upward small, medium and large shifts in the parameters indicating the effectiveness and usefulness of the scheme. However, the speed of detection varies depending on the type of shifts, the parameters, and the percentile being considered. Except for minor sampling fluctuations, in general, the monitoring scheme detects OOC signal in percentiles faster for downward shifts than the upward shifts (refer Table 2 and Figure 2). In particular, when θ is IC, the ARLs around 50^{th} percentile are smaller than the other percentiles for both upward and downward shifts in λ as is evident from Table 2 and Figure 2(a). For example, for a 4% decrease (increase) in λ when θ is IC ($\Delta\theta = 0$), there is about 27.8% (21%) reduction in the ARL of the 50^{th} percentile. On the other hand, when λ is IC, the ARLs for the lower percentiles (around 10^{th} percentile) is found to be smaller than the other percentiles for downward (upward) shift in θ (refer Table 2 and Figure 2(b)). For example, there is about 44.8% (13.8%) reduction in the ARL of the 10^{th} percentile for a 6% decrease (increase) in θ when λ is IC. From Table 2 and Figure 2(c) it is also clear that for 10% deviation in θ the ARLs around 50^{th} percentile are smaller than the other percentiles for both upward and downward shifts in λ. Again, from Figure 2(d) it can also be observed that, for 10% deviation in λ the ARLs around 50^{th} percentile are smaller than the other percentiles for both upward and downward shifts in θ.



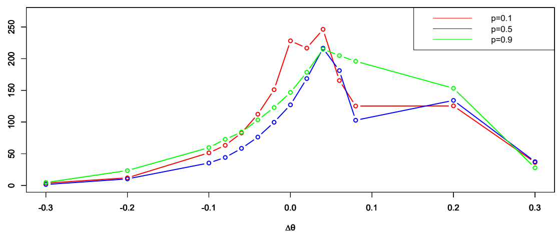
(a) ARL_1 for different choices of $\Delta\lambda$, when $\Delta\theta = 0$



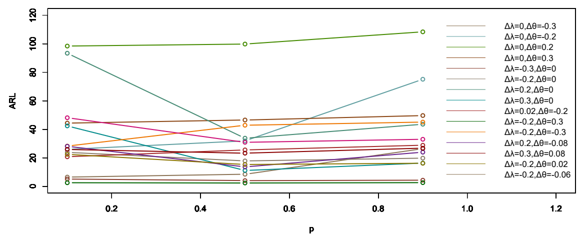
(b) ARL_1 for different choices of $\Delta\theta$, when $\Delta\lambda = 0$



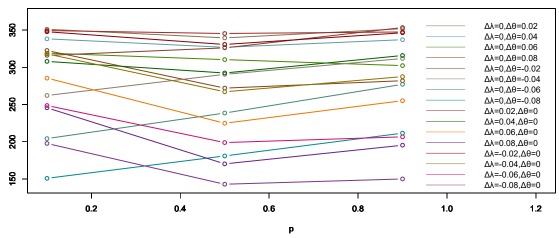
(c) ARL_1 for different choices of $\Delta\lambda$, when $\Delta\theta = 0.1$



(d) ARL_1 for different choices of $\Delta\theta$, when $\Delta\lambda = 0.1$



(e) ARL_1 for different choices of p



(f) ARL_1 for different choices of p

Figure 2: Graphs of ARL_1 for different choices of $\Delta\lambda$, $\Delta\theta$ and p.

5. ILLUSTRATIVE EXAMPLE WITH COMPARISONS

In this section, the BHCC and SHCC monitoring schemes are illustrated by a numerical example which records the waiting times (in minutes) of 100 customers before getting their services (see Ghitany *et al.*, 2008). The BHCC scheme is then compared with bootstrap scheme with Type I and Type II censoring. Various summary measures of the data set can be found below:

Min	5%	10%	25%	50%	75%	90%	95%	Max
0.800	1.895	2.880	4.675	8.100	13.000	19.090	21.955	38.500

First, the Weibull and GE distributions are compared for fitting the data set. For Weibull model, the MLEs of the shape and scale parameters are found to be 1.458 and 10.954 respectively with Kolmogorov-Smirnov test (K-S) statistic value $D = 0.0577$, and p-value, $p = 0.8927$. For GE model, the MLEs are obtained as $\widehat{\theta} = 2.183$ and $\widehat{\lambda} = 0.159$ with $D = 0.0402$ and $p = 0.9970$. The histogram of the data and two fitted densities are provided in Figure 3. The fit results confirm that the GE distribution provides a better fit than Weibull in this case. Moreover, logarithm of the ratio of maximized likelihood (RML), defined as $T = \log L = l_{GE}(\hat{\theta}, \hat{\lambda}) - l_{WE}(\widehat{shape}, \widehat{scale}) = -317.0884 - (-318.7261) = 1.6377 > 0$ indicates to choose GE distribution over Weibull.

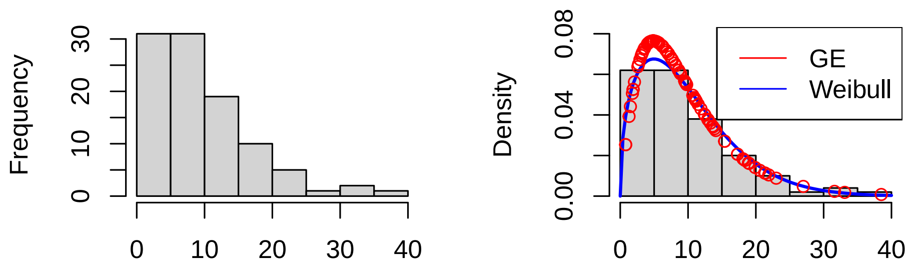


Figure 3: Histogram and density plot of waiting times of 100 patients.

In order to achieve service excellence, the bank may find extreme percentiles of the waiting times worth investigating over the average waiting time. An upward shift in the upper percentile of the waiting times indicates deterioration in the service quality and requires monitoring. In view of this objective, the BHCC chart is constructed for monitoring 90^{th} percentile of the waiting times. The complete data is censored either at the waiting time of the first 60% of the total number of customers ($r = 60$) or at the waiting time of 10 minutes ($x_0 = 10$), whichever occurs earlier. The censoring time $x_0 = 10$ is chosen near to 60^{th} percentile. The complete set of 100 observations is considered as four ($k = 4$) reference samples of size $m = 25$ each. The MLEs of θ and λ under the stated hybrid censoring scheme are obtained as $\widehat{\theta} = 1.760$ and $\widehat{\lambda} = 0.127$ respectively. Using these MLEs, $B = 5,000$ bootstrap samples of size $m = 25$ each are drawn with $r = 15$ (60% of the subgroup size) and $x_0 = 10$. Following the steps $4 - 7$ in subsection 3.1, and using $\nu = 0.0027$ as FAR, the control limits of the BHCC scheme for the 90^{th} percentile are obtained as $UCL = 78.597$, $CL = 26.255$ and $LCL = 12.702$, while the same for the SHCC scheme are computed as $UCL_{SH} = 23.711$, $LCL_{SH} = 19.966$ and $CL_{SH} = 21.839$. It is observed that both schemes provide asymmetric control limits from the respective CL, while the SHCC scheme has narrower interval than the BHCC scheme. Twenty subgroups of size $m = 25$ each are generated from the OOC process under similar hybrid censoring plan having shape parameters $\theta = 2.024$ and $\lambda = 0.108$ (15% increase in θ and 15% decrease in λ).

The OOC performance of the BHCC and SHCC schemes for the 90^{th} percentile are presented in Figure 4 and Figure 5 respectively. The BHCC scheme is able to produce OOC signals quite efficiently with five 90^{th} percentile points falling above the UCL with the first OOC signal being obtained at test sample 2 indicating effectiveness of the scheme in terms of quick detection as well. On the other hand, nine OOC signals are produced by the SHCC scheme with test sample 2 producing the first OOC signal. It is to be noted here that the SHCC scheme grossly underestimates the IC ARL (computation of IC ARL for SHCC scheme is not shown for brevity) due to the narrow band of the control limits which may eventually produce false OOC signals.

Next, bootstrap monitoring scheme is used for type-I (denoted as BT^{I}CC) and type-II (denoted as BT^{II}CC) censored data coming from the GE distribution and their performance is compared with the BHCC monitoring scheme with the same data set and the procedure as used before. The control monitoring schemes for type-I and type-II censored data can be derived as a special case of hybrid censored data for $r = n$ and $T = x_{n:n}$ respectively. The MLEs of θ and λ under type-I censoring with $x_0 = 10$ are obtained as $\widehat{\theta} = 1.803$ and $\widehat{\lambda} = 0.131$ respectively, while the same under type-II censoring with $r = 60$ are found to be

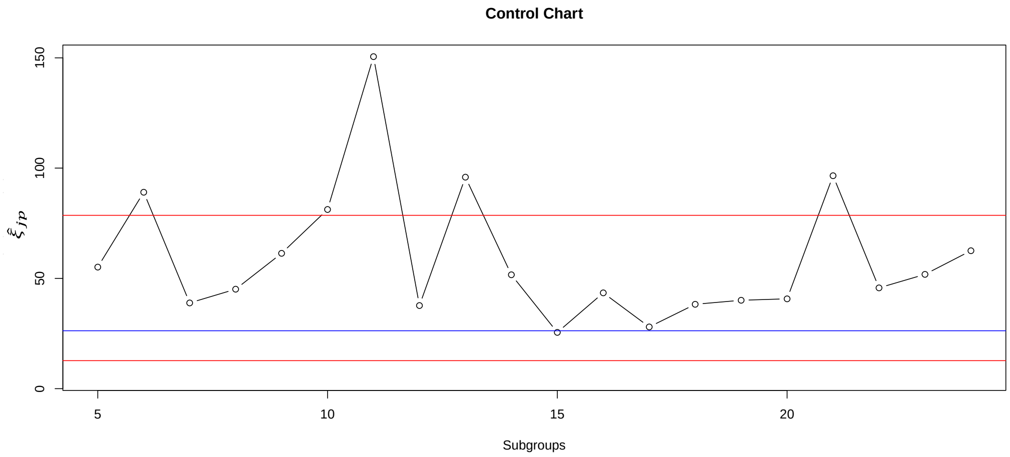


Figure 4: BHCC monitoring scheme for 90$^{\text{th}}$ percentile of the waiting time data with $\Delta\theta = 0.15$, $\Delta\lambda = -0.15$, $UCL = 78.597$, $CL = 26.255$, $LCL = 12.702$.

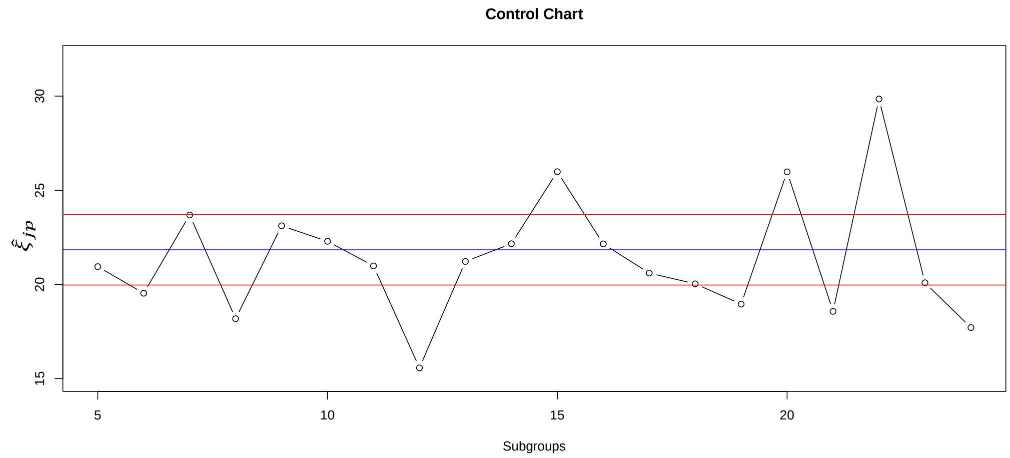


Figure 5: SHCC monitoring scheme for 90$^{\text{th}}$ percentile of the waiting time data with $\Delta\theta = 0.15$, $\Delta\lambda = -0.15$, $UCL_{SH} = 23.711$, $CL_{SH} = 21.839$, $LCL_{SH} = 19.966$; Example 1.

$\widehat{\theta} = 1.766$ and $\widehat{\lambda} = 0.127$ respectively. While the control limits of BT^{I}CC scheme for the 90$^{\text{th}}$ percentile are obtained as $UCL = 76.812, CL = 15.833$ and $LCL = 12.066$, the same for the BT^{II}CC scheme are calculated as $UCL = 44.144, CL = 24.770$, and $LCL = 12.599$. Both the schemes provide asymmetric control limits with the BT^{II}CC scheme having narrower interval than the BT^{I}CC scheme. After the first four IC subgroups, twenty subgroups of size $m = 25$ each are generated from the OOC process having $\theta = 2.073$ and $\lambda = 0.111$ (15% increase in θ and 15% decrease in λ). Figure 6 and Figure 7 provide the OOC performance of the control monitoring schemes for the 90$^{\text{th}}$ percentile. Figure 6 shows that the type-I censored scheme is able to generate three OOC points falling above the UCL with the first being produced at test sample 12. The type-II censored scheme as is shown in Figure 7 produces two OOC signals just above the UCL with test sample 9 providing the first signal. It is evident from the data analysis that the hybrid censored control monitoring scheme performs better than type-I and type-II censored control monitoring schemes in terms of both frequency and speed of detection of OOC signals.

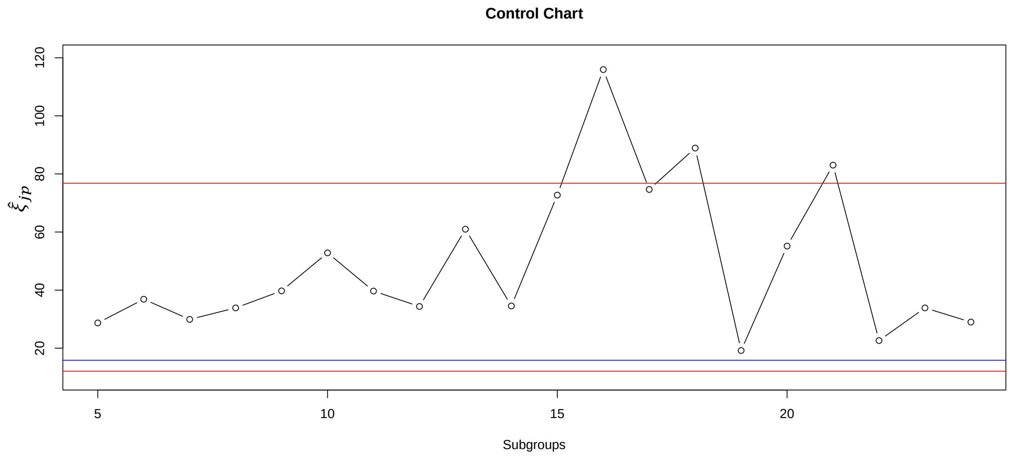


Figure 6: BT^{I}CC monitoring scheme for 90$^{\text{th}}$ percentile of the waiting time data with $\Delta\theta = 0.15$, $\Delta\lambda = -0.15$, $UCL = 76.812$, $CL = 15.833$, $LCL = 12.066$.

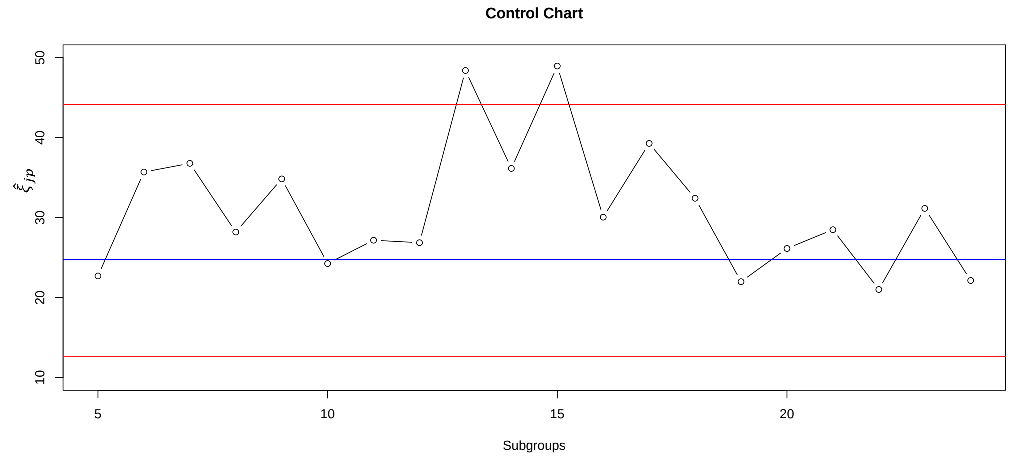


Figure 7: BT^{II}CC monitoring scheme for 90$^{\text{th}}$ percentile of the waiting time data with $\Delta\theta = 0.15$, $\Delta\lambda = -0.15$, $UCL = 44.144$, $CL = 24.770$, $LCL = 12.599$.

6. APPLICATION

This section provides an application of the BHCC monitoring scheme in clinical practice. The scheme is used to monitor the top percentile of the survival times of 120 patients (see Hamedani, 2013) with breast cancer obtained from a large hospital in a period from 1929 to 1938. The histogram of the data set as shown in Figure 8 and the summary measures below suggest the skewed nature of the data set.

Min	5%	10%	25%	50%	75%	90%	95%	Max
0.3	6.585	10.110	17.800	40.000	60.000	105.400	125.050	154.0

The MLEs of θ and λ for the complete data set coming from the GE distribution are found to be $\widehat{\theta} = 1.649$ and $\widehat{\lambda} = 0.029$ respectively. The fitted density is provided in Figure 8. The one sample K-S statistic and corresponding p-value are found to be 0.0717 and 0.5681 respectively.

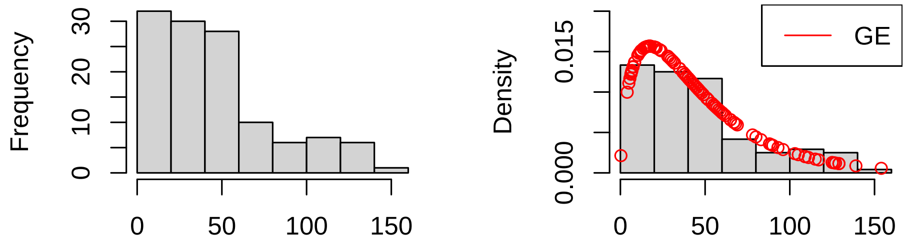


Figure 8: Histogram and density plot of survival times of 120 patients with breast cancer.

The fit results recommend GE distribution to model the survival time data and subsequent development of BHCC scheme. The complete sample data is split into six subgroups of size 20 each. Under hybrid censoring with $r = 72$ and $x_0 = 60$, the estimates of the GE parameters are obtained as $\widehat{\theta} = 1.415$ and $\widehat{\lambda} = 0.024$. Using 5,000 bootstrap samples of size $m = 20$ each with $r = 12$ and $x_0 = 60$, the control limits of the BHCC monitoring scheme for the 90$^{\text{th}}$ percentile are evaluated as $UCL = 322.248, CL = 115.577, LCL = 48.912$. Next, twenty phase II samples of size $m = 20$ each are generated from the process under similar hybrid censoring plan with $\Delta\theta = 0.15$ and $\Delta\lambda = -0.15$ to develop the BHCC monitoring scheme for the 90$^{\text{th}}$ percentile as presented in Figure 9. The scheme has been able to detect OOC signals at 2$^{\text{nd}}$, 5$^{\text{th}}$ and 11$^{\text{th}}$ samples. For the same data set, the control limits for the BT^{I}CC monitoring scheme for 90$^{\text{th}}$ percentile with $x_0 = 60$ are found to be $UCL = 291.965, CL = 106.950, LCL = 51.203$. Figure 10 shows that this scheme has been able to detect only one OOC signal at the 20$^{\text{th}}$ sample. On the other hand, BT^{II}CC monitoring scheme for 90$^{\text{th}}$ percentile with $r = 72$ is presented in Figure 11 with $UCL = 232.296, CL = 115.565, LCL = 48.846$. Figure 11 shows that this scheme also detects only one OOC signal at the 9$^{\text{th}}$ sample. The frequency and speed of detection of OOC signals further justify the use of BHCC monitoring scheme over BT^{I}CC and BT^{II}CC monitoring schemes for the percentiles of survival time in a healthcare set-up.

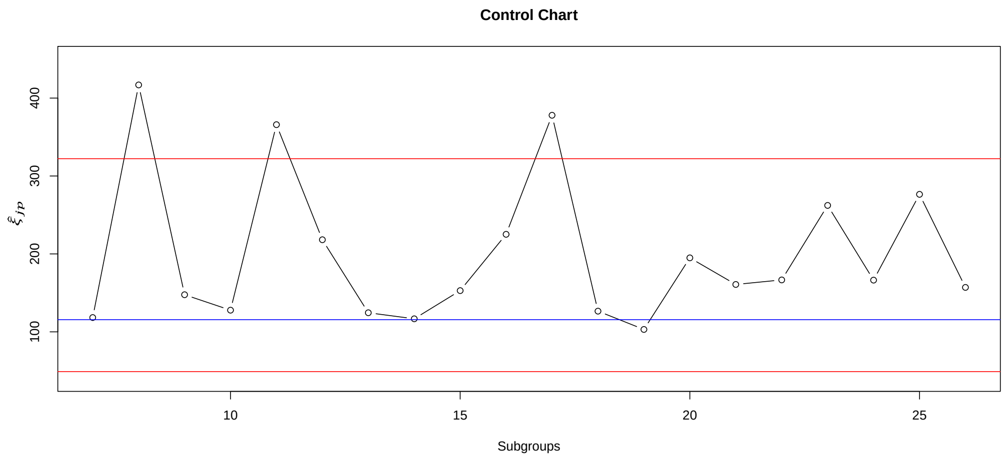


Figure 9: BHCC monitoring scheme for 90$^{\text{th}}$ percentile of the survival time data with $\Delta\theta = 0.15$,$\Delta\lambda = -0.15$, $UCL = 322.248$, $CL = 115.577$, $LCL = 48.912$.

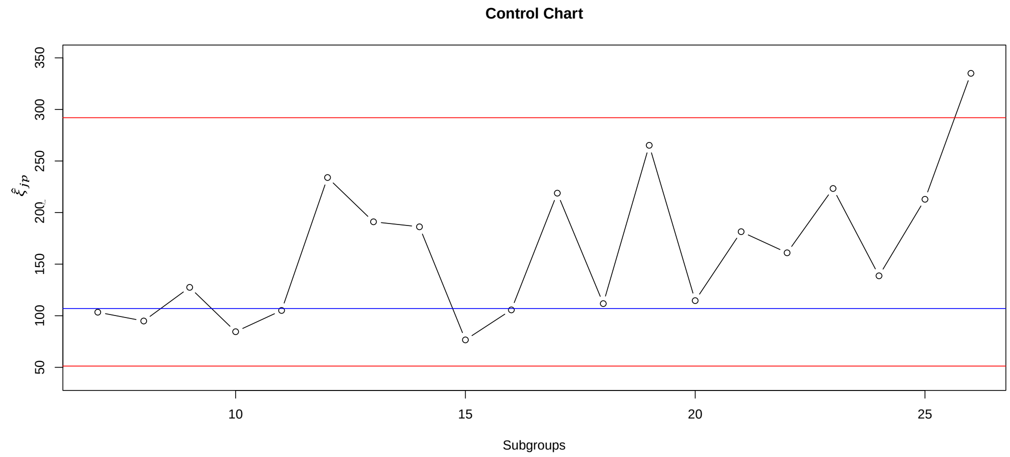


Figure 10: BT^{I}CC monitoring scheme for 90$^{\text{th}}$ percentile of the survival time data with $\Delta\theta = 0.15$, $\Delta\lambda = -0.15$, $UCL = 291.965$, $CL = 106.950$, $LCL = 51.203$.

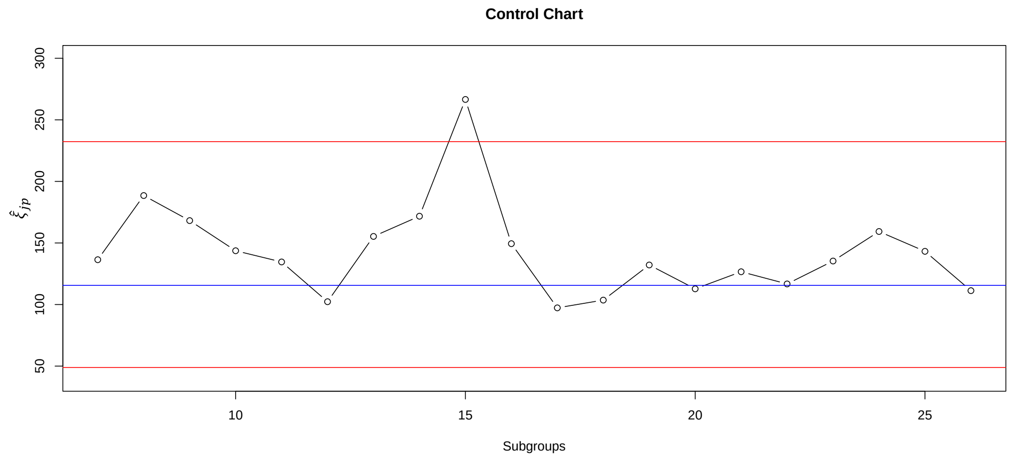


Figure 11: BT^{II}CC monitoring scheme for 90$^{\text{th}}$ percentile of the survival time data with $\Delta\theta = 0.15$, $\Delta\lambda = -0.15$, $UCL = 232.296$, $CL = 115.565$, $LCL = 48.846$.

7. CONCLUDING REMARKS

In this work, hybrid censoring is employed to develop control monitoring schemes for percentiles of GE distribution using bootstrap and asymptotic methods. Bootstrap monitoring schemes for type-I and type-II censored data are also developed under similar set-up as a special case of hybrid censoring plan. An extensive simulation study is conducted to evaluate the IC and OOC performance of the schemes. The hybrid censored schemes are found to be effective in the detection of OOC signals in terms of both magnitude and speed as demonstrated by a real data set. One application from healthcare is also provided to establish the effectiveness of the schemes. In this sense, the present work is the first attempt to apply a new censoring scheme in the process control and generalizes available control monitoring schemes for the GE data. As a scope for future research, hybrid censored schemes may be proposed under Bayesian set-up measuring uncertainty in the parameter(s). One can also think of using progressive censoring scheme to construct such control mechanism. For highly reliable products, accelerated life testing scheme may be employed under various censoring plans for the same purpose.

DATA AVAILABILITY STATEMENT

The data sets used in this manuscript are available in Ghitany *et al.* (2008) and Hamedani (2013).

ACKNOWLEDGMENTS

The authors are thankful to the reviewer, the associate editor and the editor-in-chief for their constructive comments and suggestions which have improved the earlier version of this manuscript.

REFERENCES

Bain, L. J. (1976). *Statistical Analysis of Reliability and Life Testing Model.* Marceland Dekker Inc., New York.

Chiang, J. Y., Jiang, N., Brown, T. N., Tsai, T. R., and Lio, Y. L. (2017). Control charts for generalized exponential distribution percentiles. *Communications in Statistics – Simulation and Computation*, 46(10):7827–7843.

Chiang, J. Y., Lio, Y. L., Ng, H. K. T., Tsai, T. R., and Li, T. (2018). Robust bootstrap control charts for percentiles based on model selection approaches. *Computers & Industrial Engineering*, 123:119–133.

Chowdhury, S., Kundu, A., and Modok, B. (2021). Bootstrap beta control chart for monitoring proportion data. *International Journal of Quality & Reliability Management*, `https://www.emerald.com/insight/content/doi/10.1108/IJQRM-09-2020-0287/full/html`.

Efron, B. and Tibshirani, R. J. (1993). *An Introduction to the Bootstrap.* Chapman & Hall, New York.

Epstein, B. (1954). Truncated life tests in the exponential case. *The Annals of Mathematical Statistics*, 25(3):555–564.

Erto, P. and Pallotta, G. (2007). A new control chart for Weibull technological processes. *Quality Technology & Quantitative Management*, 4(4):553–567.

Erto, P., Pallota, G., and Mastrangelo, C. M. (2015). A semi-empirical Bayesian chart to monitor Weibull percentiles. *Scandinavian Journal of Statistics*, 42(3):701–712.

Ghitany, M. E., Atieh, B., and Nadarajah, S. (2008). Lindley distribution and its application. *Mathematics and Computers in Simulation*, 78(4):493–506.

Gupta, R. D. and Kundu, D. (1999). Generalized exponential distribution. *Australian & New Zealand Journal of Statistics*, 41(2):173–188.

Gupta, R. D. and Kundu, D. (2001). Exponentiated exponential family; an alternative to gamma and Weibull. *Biometrical Journal*, 43:117–130.

Haghighi, F., Pascual, F., and Castagliola, P. (2015). Conditional control charts for Weibull quantiles under type-II censoring. *Quality and Reliability Engineering International*, 31(8):1649–1664.

Hamedani, G. (2013). The Zografos-Balakrishnan log-logistic distribution: Properties and applications. *Journal of Statistical Theory and Applications*, 12(3):225–244.

Hyndman, R. J. and Fan, Y. (1996). Sample quantiles in statistical packages. *American Statistician*, 50:361–365.

Jones, L. A. and Woodall, W. Y. (1998). The performance of bootstrap control charts. *Journal of Quality Technology*, 30:362–375.

Kundu, D. and Pradhan, B. (2009). Estimating the parameters of the generalized exponential distribution in presence of hybrid censoring. *Communications in Statistics – Theory and Methods*, 38(12):2030–2041.

Leiva, V., Santos, R. A. D., Saulo, H., Marchant, C., and Lio, Y. (2022). Bootstrap control charts for quantiles based on log?symmetric distributions with applications to the monitoring of reliability data. *Quality and Reliability Engineering International*, `https://doi.org/10.1002/qre.3072`.

Lio, Y. L. and Park, C. (2008). A bootstrap control chart for Birnbaum-Saunders percentiles. *Quality and Reliability Engineering International*, 24:585–600.

Lio, Y. L. and Park, C. (2010). A bootstrap control chart for inverse Gaussian percentiles. *Journal of Statistical Computation and Simulation*, 80:287–299.

Lio, Y. L., Tsai, T. R., Aslam, M., and Jiang, N. (2014). Control charts for monitoring Burr type-X percentiles. *Communications in Statistics – Simulation and Computation*, 43:761–776.

Liu, R. Y. and Tang, J. (1996). Control charts for dependent and independent measurements based on the bootstrap. *Journal of the American Statistical Association*, 91:1694–1700.

Louis, T. A. (1982). Finding the observed information matrix when using the EM algorithm. *Journal of the Royal Statistical Society, B*, 44(2):226–233.

Ma, Z., Park, C., and Wang, M. (2022). A robust bootstrap control chart for the log-logistic percentiles. *Journal of Statistical Theory and Practice*, 16(1):1–23.

Nichols, M. D. and Padgett, W. J. (2005). A bootstrap control chart for Weibull percentiles. *Quality and Reliability Engineering International*, 22:141–151.

Padgett, W. J. and Spurrier, J. D. (1990). Shewhart-type charts for percentiles of strength distributions. *Journal of Quality Technology*, 22:283–288.

Seppala, T., Moskowitz, H., Plante, R., and Tang, J. (1996). Statistical process control via the subgroup bootstrap. *Journal of Quality Technology*, 27:139–153.

Vining, G., Kuahci, M., and Pedersen, S. (2016). Recent advances and future directions for quality engineering. *Quality and Reliability Engineering International*, 32(3):863–875.

Wang, F. K., Bizuneh, B., and Cheng, X. B. (2018). New control charts for monitoring the Weibull percentiles under complete data and Type-II censoring. *Quality and Reliability Engineering International*, 34(3):403–416.