"This is the peer reviewed version of the following article: Zwetsloot I M, Mahmood T, Taiwo F M, Wang Z. A real-time monitoring approach for bivariate event data. Applied Stochastic Models in Business and Industry. 2024;39 (6):pp. 789-817, which has been published in final form at https://doi.org/10.1002/asmb.2800. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Use of Self-Archived Versions. This article may not be enhanced, enriched or otherwise transformed into a derivative work, without express permission from Wiley or by statutory rights under applicable legislation. Copyright notices must not be removed, obscured or modified. The article must be linked to Wiley's version of record on Wiley Online Library and any embedding, framing or otherwise making available the article or pages thereof by third parties from platforms, services and websites other than Wiley Online Library must be prohibited."
A Real Time Monitoring Approach for Bivariate Event Data

Inez Maria Zwetsloot\textsuperscript{a,b} and Tahir Mahmood\textsuperscript{c,d} and Funmilola Mary Taiwo\textsuperscript{e} and Zezhong Wang \textsuperscript{a}

\textsuperscript{a}Department of Advanced Design and Systems Engineering, City University of Hong Kong, Tat Chee Avenue, Kowloon, Hong Kong; \textsuperscript{b}School of Data Science, City University of Hong Kong, Tat Chee Avenue, Kowloon, Hong Kong; \textsuperscript{c}Industrial and Systems Engineering Department, King Fahd University of Petroleum and Minerals, Dhahran, 31261, Saudi Arabia; \textsuperscript{d}Interdisciplinary Research Centre for Smart Mobility and Logistics, King Fahd University of Petroleum and Minerals, Dhahran, 31261, Saudi Arabia; \textsuperscript{e}Department of Statistics, University of Manitoba, Winnipeg, MB R3T 2N2 Canada

ARTICLE HISTORY
Compiled July 4, 2023

Abstract
Early detection of changes in the frequency of events is an important task in many fields, such as disease surveillance, monitoring of high-quality processes, reliability monitoring, and public health. This article focuses on detecting changes in multivariate event data by monitoring the time-between-events (TBE). Existing multivariate TBE charts are limited because they only signal after an event occurred for each of the individual processes. This results in delays (i.e., long time to signal), especially when we are interested in detecting a change in one or a few processes with different rates. We propose a bivariate TBE (BTBE) chart, which can signal in real-time. We derive analytical expressions for the control limits and average time-to-signal performance, conduct a performance evaluation and compare our chart to an existing method. Our findings showed that our method is an effective approach for monitoring bivariate time-between-events data and has better detection ability than the existing method under transient shifts and is more generally applicable. A significant benefit of our method is that it signals in real-time and that the control limits are based on analytical expressions. The proposed method is implemented on two real-life datasets from reliability and health surveillance.

Keywords: early event detection; lifetime expectancy; multivariate control chart; statistical process monitoring; time-between-events; real-time monitoring; superimposed process.

1. Introduction

Detecting changes in time-between-events data has many applications. In healthcare, examples are: monitoring of times to blindness in the eyes (Huster et al., 1989; Li et al., 2012), response time to different treatments (Gross and Lam, 1981; Lu and Bhattacharyya, 1991), or recurrence time after (cancer) treatment (Byar, 1980; Chiou et al., 2018). In manufacturing, time-between-events data are: the production times of different batches and the failure time of subsystems (Flury and Quaglino, 2018; Nelson, 1982).
In public surveillance, time-between-events data are: time to disease outbreaks (Sparks et al., 2019), time between polio cases, and the time between crimes (Qiu et al., 2019; Wang and Qiu, 2018).

Event data are discrete observations occurring in continuous time, where the time interval between two events carry substantial information about the underlying systems. Hence, it is of interest to study the time period between the events, referred to as time-between-events data. Two groups of methods have been proposed in the literature for monitoring event data. One category of methods is the monitoring of count data. Count data can be obtained by counting the number of events in pre-specified periods. Reviews of the monitoring methods for count data can be found in Saghir and Lin (2015); Ali et al. (2016) and Mahmood and Xie (2019). However, the count-based approach is not real-time, as one needs to wait until the end of each time period, e.g., a week, before changes can be detected (Zwetsloot and Woodall, 2021). In addition, the selection of aggregation window length is always somewhat arbitrary.

The second category of methods are Time-Between-Events (TBE) control charts. With a TBE control chart, we monitor the length of time between events. For recent studies on univariate TBE control charts, the reader is referred to Sparks et al. (2019, 2020). Methods for multivariate TBE data are categorized into two types; methods for (a) vector-based event data and (b) point-process data (Zwetsloot et al., 2021). Vector-based event data occur when by construction, each subprocess can signal once only and hence we can create a vector of event times. Consider for example time to blindness, we may observe different times for the left and the right eye but by the nature of the system each will yield only one event time. Another example of vector-based event data is breakdown and repair times of a system. In Section 6, we will implement our method on a case study considering the monitoring of breakdown and repair times of escalators. For each event—breakdown and repair, we record the time elapsed since the previous breakdown (in days) as well as the time to repair (in hours). Hence, we will have bivariate vectors consisting of time to break down ($X_1$) and time to repair ($X_2$). For the escalator management company, it is of interest to know if $X_1$ is decreasing, indicating possible unexpected rapid deterioration of the escalator, as well as increase in $X_2$, indicating a possible breach of maintenance contract by the subcontractor responsible for the maintenance and repair.

The other type of multivariate TBE data is point-process data. In point-process data, one event may occur several times before another shows up. For example, consider manufactured items that can fail in several ways and are repaired after a failure. One may observe a failure type A twice before observing failure type B or C. In this paper, we focus on vector-based event data only.

As far as we know, all literature on Multivariate Time-Between-Events (MTBE) control charts focuses on vector-based event data. The most well-known method is developed by Xie et al. (2011). The authors considered Gumbel’s Bivariate Exponential (GBE) distributed data and proposed a vector-based Multivariate Exponentially Weighted Moving Average (MEWMA) control chart. We will provide more details on the existing MTBE control chart literature in Section 5. Noteworthy is that all current MTBE control charts have a built-in detection delay, which requires that one event is available for each of the $p$ variables under consideration. Hence, changes can only be detected when we have an observed event on each variable. As these events happen asynchronously in time, the methods have a built-in delay until the vector of event data is completely observed. Figure 1 illustrates these delays. For the bivariate event data, existing methods can only signal when the observation of both events is available. Delays are undesirable when we wish to detect changes in the process as
quickly as possible. Furthermore, it is easy to see how an extension from a bivariate approach to a multivariate approach will result in longer delays.

| Illustration of bivariate event data |
|--------------------------------------|
| **Process A** | **System 1** | **Process B** | **System 2** | **System 3** |
| Time          | Time         | Time          |

| Our approach: monitor data in real time |
|------------------------------------------|
| Monitoring points                        |

| Existing approaches: monitoring complete vectors |
|-------------------------------------------------|
| Monitoring points                              |
| Delay                                           |
| No Delay                                        |
| Delay                                           |

**Figure 1.** Illustration of delay when monitoring complete vectors of event data

Therefore, in this article, we propose a novel and effective method for *real-time bivariate event-based monitoring*, named the Bivariate Time-Between-Events (BTBE) control chart. This method is designed for bivariate event data of the vector-based type. Our proposed method has four main features. First, it has real-time detection power and, unlike all existing methods, does not have a built-in detection delay. Second, it is applicable to any type of bivariate event data, irrespective of the distribution. Third, we derive analytical expressions for the control limits and the average time-to-signal (ATS), making simulations unnecessary in most applications. Fourth, it provides exact information about the root cause behind an out-of-control signal.

The remainder of this article is organized as follows. The proposed method is introduced in Section 2. Analytical expressions for the theoretical performance of our method are given in Section 3. The performance of our proposed method under different distributional environments is discussed in Section 4, and a comparison with the existing method by Xie et al. (2011) is presented in Section 5. Implementation of the proposed method on two case studies is discussed in Section 6. Finally, the article is summarized in Section 7. Moreover, mathematical proofs and other details are provided in the *Appendices* A-C and the supplementary material. Codes for implementing our method and reproducing all results in this paper are provided on https://github.com/tmahmood5/Codes-BTBE-Monitoring-Method.

2. Proposed Method

This section presents our proposed BTBE chart to monitor bivariate vector-based event data. This method can signal changes as data comes in, and unlike other existing
methods, there is no need to wait until we observe a complete vector of events (refer to Figure 1). We first present monitoring statistics in Section 2.1, followed by the control limits in Section 2.2 and a discussion on the implementation of our chart in Section 2.3.

2.1. Monitoring statistic

Consider $X = (X_1, X_2)$ as a vector of bivariate lifetimes, where $X_1$ indicates the time to an event in the first subprocess and $X_2$ indicates the time to an event in the second subprocess. We assume that $X$ is drawn from a bivariate probability density function $f(x|\theta)$ where $\theta$ is the parameter vector. We will discuss some typical choices for event time distributions $f()$ in Section 2.3. We denote the corresponding cumulative joint distribution function, the joint survival function, and the partial survival functions by

$$F(x_1, x_2) = \Pr[X_1 \leq x_1, X_2 \leq x_2]$$
$$S(x_1, x_2) = \Pr[X_1 > x_1, X_2 > x_2]$$
$$S_1(x_1, x_2) = \frac{\partial}{\partial x_1} S(x_1, x_2)$$
$$S_2(x_1, x_2) = \frac{\partial}{\partial x_2} S(x_1, x_2)$$

As $X = (X_1, X_2)$ denotes event times, one of the two is observed first. In order to model the data in real-time, we define order statistics $X_{(1)}$ as the first observed event time and $X_{(2)}$ as the second observed event time:

$$X_{(1)} = \min(X_1, X_2), \quad X_{(2)} = \max(X_1, X_2)$$

For our method, we create a superimposed data stream of the order statistics and hence plot each event consecutively on our chart. As example, consider the three observations in Figure 1. These could be time to blindness for the left eye (process A) and the right eye (process B) after treatment. Or time to failure and time to repair (see our first case study in Section 6.1 for details) or time-to-infection and time to diagnosis (see our second case study in Section 6.2 for details). The data in Figure 1 are:

$$X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = \begin{bmatrix} 0.21 \\ 1.33 \\ 1.40 \\ 2.00 \end{bmatrix}.$$  

The superimposed process consists of the events as they are observed and is given by

$$[0.21, 1.33, 1.40, 0.50, 2.00].$$

Assume that these time-between-event data are measured in days. For our method, we plot a monitoring statistic after 5 hours (0.21 day) and after 1 day and 8 hours (1.33 day) for the first vector. Whereas, existing methods wait for 1 day and 8 hours until they observed both 0.21 and 1.33 before plotting the first monitoring statistic. Similarly for the third vector, we first plot $X_2$ after 12 hours and then plot $X_1$ 1.5 days later. Whereas, existing methods wait until they observe both, $X_1$ and $X_2$, after
2 days before plotting. Note that in the second case, we observe both processes at the same time and our method and existing methods would signal at the same moment.

2.2. Control limits

Our control chart consists of the plotted data \((X)\) and its corresponding upper and lower control limits. For our method, the plotted data is the superimposed process as illustrated in Equation (3). The control limits are obtained using the following theorem:

**Theorem 1.** Let \(X = (X_1, X_2)\) be a bivariate event vector distributed according to \(f(x_1, x_2)\). Define the order statistics \(X^{(1)}_1\) and \(X^{(2)}_1\) as in Equation (2). Define \(V\) as a random variable indicating which event is observed first, where \(V = 0\) if \(x_1 < x_2\), \(V = 1\) if \(x_1 > x_2\), and \(V = 2\) if \(x_1 = x_2\). For the first observed event \(X^{(1)}_1\) we define the control limits as

\[
LCL^{(1)} = F_{X^{(1)}_1}^{-1} \left( \frac{\alpha}{2}, v \right)
\]

\[
UCL^{(1)} = S_{X^{(1)}_1}^{-1} \left( \frac{\alpha}{2}, v \right)
\]

and for the second event we defined them as

\[
LCL^{(2)} = F_{X^{(2)}_1|X^{(1)}_1}^{-1} \left( \frac{\alpha}{2}, x^{(1)}_1, v \right)
\]

\[
UCL^{(2)} = S_{X^{(2)}_1|X^{(1)}_1}^{-1} \left( \frac{\alpha}{2}, x^{(1)}_1, v \right)
\]

where \(S^{-1}\) and \(F^{-1}\) are the inverses of the conditional survival and distribution functions as defined in Definitions 1 and 2.

**Proof.** For a univariate data point \(X\), control limits are set such that the exceeded probability for in-control data is equal to a pre-specified value \(\alpha\). Mathematically, this is equal to setting the control limits such that \(F_X(LCL) = \alpha/2\) and \(S_X(UCL) = \alpha/2\) for the lower (LCL) and upper (UCL) control limits, respectively. Here, \(F_X()\) and \(S_X()\) are the cumulative distribution and survival function. In our superimposed process, a data point is either a first observed event \(X^{(1)}_1\) or a second observed event \(X^{(2)}_1\). Accordingly, we take the inverse of cumulative distribution and survival function to set the control limits.

The expressions for the cumulative distribution and survival function in Theorem 1 will differ depending on whether the data point is the first observed event \((X^{(1)}_1)\) or the second observed event \((X^{(2)}_1)\) and whether the data point is from the first subprocess \((X_1)\) or the second subprocess \((X_2)\). We derive control limits for all these situations in the following two definitions:

**Definition 1.** Assume \(X = (X_1, X_2) \sim f(x_1, x_2)\) and let \(X^{(1)}_1\) and \(X^{(2)}_1\) be defined as in Equation (2). Define \(V\) as a random variable indicating which event is observed first, where \(V = 0\) if \(x_1 < x_2\), \(V = 1\) if \(x_1 > x_2\), and \(V = 2\) if \(x_1 = x_2\). Then the conditional cumulative distribution function of \(X^{(1)}_1\) conditional on \(V = v\), is defined
as $F_{X(1)}(x(1)|v) = P[X(1) \leq x(1), V = v] / P[V = v]$ and equal to

$$F_{X(1)}(x(1)|v) = \begin{cases} \int_0^{x(1)} S_1(x_1, x_1)dx_1 & \text{if } v = 0 \\ \int_0^{x(1)} S_1(x_1, x_1)dx_1 + \int_{x(1)}^{\infty} S_2(x_2, x_2)dx_2 & \text{if } v = 1 \\ \int_0^{\infty} S_2(x_2, x_2)dx_2 + \int_0^{x(1)} f(x_1, x_1)dx_1 & \text{if } v = 2 \end{cases}$$

(4)

where $S_1$ and $S_2$ are defined as in Equation (1). Equivalently, we define the conditional survival function of $X(1)$ conditional on $V = v$, as $S_{X(1)}(x(1)|v) = P[X(1) > x(1), V = v] / P[V = v]$ is equal to

$$S_{X(1)}(x(1)|v) = \begin{cases} \int_{x(1)}^{\infty} S_1(x_1, x_1)dx_1 & \text{if } v = 0 \\ \int_0^{x(1)} S_1(x_1, x_1)dx_1 + \int_{x(1)}^{\infty} S_2(x_2, x_2)dx_2 & \text{if } v = 1 \\ \int_0^{\infty} S_2(x_2, x_2)dx_2 + \int_0^{x(1)} f(x_1, x_1)dx_1 & \text{if } v = 2 \end{cases}$$

(5)

Detailed derivations of Equations (4) and (5) can be found in Appendix A.1.

We note that this definition allows for the case $X_1 = X_2$, which is unusual when the distribution function $f(\cdot)$ is assumed to be continuous. We explicitly include this case because the failure mechanism is based on an external random shock model for some bivariate lifetime distributions. Some shocks can affect the lifetime of both components simultaneously, where lifetime $X_1$ and $X_2$ will be equal.

**Definition 2.** Assume $X = (X_1, X_2) \sim f(x_1, x_2)$ and let $X(1)$ and $X(2)$ be defined as in Equation (2). Define $V$ as a binary random variable indicating which event is observed first, where $V = 0$ if $x_1 < x_2$ and $V = 1$ if $x_1 > x_2$. Then the conditional cumulative distribution function $F_{X(2)|X(1), V}(x(2)|x(1), v) = P[X(2) \leq x(2)|X(1) = x(1), V = v]$ is defined as

$$F_{X(2)|X(1), V}(x(2)|x(1), v) = \begin{cases} 1 - \frac{S_1(x_1, x_2)}{S_1(x_1, x_1)} & \text{if } v = 0 \\ 1 - \frac{S_2(x_1, x_2)}{S_2(x_2, x_2)} & \text{if } v = 1 \end{cases}$$

(6)

Similarly, the conditional survival function $S_{X(2)|X(1), V}(x(2)|x(1), v) = P[X(2) > x(2)|X(1) = x(1), V = v]$ is defined as

$$S_{X(2)|X(1), V}(x(2)|x(1), v) = \begin{cases} \frac{S_1(x_1, x_2)}{S_1(x_1, x_1)} & \text{if } v = 0 \\ \frac{S_2(x_1, x_2)}{S_2(x_2, x_2)} & \text{if } v = 1 \end{cases}$$

(7)

Detailed derivations of Equations (6) and (7) can be found in Appendix A.2.
We note that the case $V = 2$, i.e., $x_1 = x_2$, is not included in Definition 2, because both event times were observed simultaneously. Hence, conditioning on the first observed event time is irrelevant.

Control limits are usually obtained by setting $\alpha = \frac{1}{ARL_0}$, where the in-control average run length ($ARL_0$) is assigned a pre-defined value. However, when the data represent event times, it is common to design the chart using the average time-to-signal ($ATS$) rather than the $ARL$ (Zwetsloot et al., 2021). We can relate the $ARL$ and $ATS$ by $ATS = ARL \cdot E[TBE]$, and as $ARL = 1/\alpha$, it holds that:

$$\alpha = \frac{E[TBE]}{ATS_0}$$  \hspace{1cm} (8)

where $E[TBE]$ is the expected time between two events (for details and derivations of $E[TBE]$ see Section 3) and, $ATS_0$ is the in-control desired $ATS$ value (we assume known parameters for designing the control limits, using $ARL = 1/\alpha$).

2.3. Implementation of our proposed BTBE chart

Control charts are generally implemented in two phases: Phase I for determining and estimating a distribution model for the data and Phase II for prospective monitoring (Montgomery, 2017; Jones-Farmer et al., 2014).

In Phase I, a stable and in-control dataset should be gathered. We assume that we have such a dataset and denote it by $X^i$ for $i = 1, 2, ..., n$. Where $X^i = (X^i_1, X^i_2)$ are the bivariate vectors of events’ time collected asynchronously in time. Next, a distribution function $f(x_1, x_2|\theta)$ should be selected and fitted. A variety of possible models $f()$ have been proposed in the literature (Kotz et al., 2004). We introduce three well-known models:

- The Gumbel’s Bivariate Exponential (GBE) distribution is based on a random external stress factor and was introduced by Gumbel (1960) and further developed by Hougaard (1986). The GBE model’s survival function is $S(x_1, x_2) = \exp((x_1/\theta_1)^\delta + (x_2/\theta_2)^\delta)^{1/\delta}$ where $\theta_1$ and $\theta_2$ are the expected values of $x_1$ and $x_2$ and, $\delta \in (0, 1]$ is the dependence parameter. When $\delta = 1$ the model generates independent data.

- The Marshall Olkin Bivariate Exponential (MOBE) distribution (Marshall and Olkin, 1967) is based on the assumption that the lifetime of the main system does not depend on the failure time of the two components but is affected by external factors. In this model, random shocks to the system appear as a homogeneous Poisson process. The MOBE model’s survival function is $S(x_1, x_2) = \exp(-\lambda_1 x_1 - \lambda_2 x_2 - \lambda_{12} \max(x_1, x_2))$ where $\lambda_1$ is related to the rate of $X_1$, $\lambda_2$ to the rate of $X_2$ and $\lambda_{12}$ determines dependence (when $\lambda_{12} = 0$, the two variables are independent).

- The Marshal Olkin Bivariate Weibull (MOBW) distribution was also introduced by Marshall and Olkin (1967). It is based on the MOBE model but allows for a more flexible shock process modeled using a non-homogeneous Poisson process. The MOBW model’s survival function is $S(x_1, x_2) = \exp(-\lambda_1 x_1^\eta - \lambda_2 x_2^\eta - \lambda_{12} \max(x_1, x_2)^\eta)$ where $\lambda_1$ is related to the rate of $X_1$, $\lambda_2$ to the rate of $X_2$, $\lambda_{12}$ determines dependence (when $\lambda_{12} = 0$, the two variables are independent) and, $\eta$ governs the shape of the marginal distributions. When $\eta = 1$ the MOBW model is equal to the MOBE model.
More details regarding each of these three models, including the maximum likelihood estimation equations, can be found in Appendix B and properties are given Tables 6, 7 and 8.

Selecting the appropriate model is an important task that requires a combination of process knowledge and statistical knowledge. We recommend that the practitioner carefully considers the failure mechanism of the underlying process and uses statistical tests to validate the fit of the select model to the data. One could consider for example the package ‘fitdistrplus’ (Delignette-Muller et al., 2015) in R for fitting the data distribution. We have used this for our case study (see Section 6 for more details.) Simultaneously, the parameters of the model have to be estimated.

Here, we have introduced three possible models for the data. However, many other bivariate distributional models exist (Kotz et al., 2004). One strength of our suggested approach is that it is easy to modify the method to any of these distributions. Hence, the practitioner is free to select any other model. Our method works for any distribution chosen, and the only thing that needs to be done is to derive the control limits in accordance with Theorem 1 using the distribution function of the selected data model. This is a great benefit of our method compared to the existing methods, which are only developed for one specific data model.

Once the model is established and the parameters are estimated, the control limits for our proposed BTBE monitoring method should be derived according to Theorem 1. For the three discussed bivariate lifetime models, we provide the control limits in Table 1, and the derivations for these limits are provided in the supplementary material. Before setting the control limits, we need to assign a desirable value for the \( ATS_0 \) and set \( \alpha \) according to Equation (8).
### Table 1. Control limits for the BTBE chart for various lifetime distributions

| Distribution | \(X\) | \(LCL\) | \(UCL\) |
|--------------|-------|-------|-------|
| GBE \(X_1\) | - \((\frac{\lambda_1}{\alpha/2} + \frac{\lambda_2}{\alpha/2})^{1/\delta}\) \ln(1 - \alpha/2) | - \((\frac{\lambda_1}{\alpha/2} + \frac{\lambda_2}{\alpha/2})^{-\delta}\) \ln(\alpha/2) |
| \(X_2\) \(\delta = 1\) | \(x_1 - \theta_2 \ln(1 - \alpha/2)\) \(x_1 < x_2\) | \(x_2 - \theta_1 \ln(1 - \alpha/2)\) \(x_1 > x_2\) |
| \(X_1\) | \(x_1 - \theta_2 \ln(\alpha/2)\) \(x_1 < x_2\) | \(x_2 - \theta_1 \ln(\alpha/2)\) \(x_1 > x_2\) |
| 0 < \(\delta < 1\) \(X_2\) | \(\left\{ \begin{array}{l} \left(\frac{1 - \frac{\lambda_2}{\alpha/2} W_0(G_1)}{1/\delta} - \theta_2 x_1 \right)^{1/\delta} \quad x_1 < x_2 \\ \left(\frac{1 - \frac{\lambda_1}{\alpha/2} W_0(G_1)}{1/\delta} - \theta_1 x_2 \right)^{1/\delta} \quad x_1 > x_2 \end{array} \right. | \(W_0(.)\) is the Lambert function* |
| \(X_2\) | \(\left\{ \begin{array}{l} \left(\frac{1 - \frac{\lambda_2}{\alpha/2} W_0(G_2)}{1/\delta} - \theta_2 x_1 \right)^{1/\delta} \quad x_1 < x_2 \\ \left(\frac{1 - \frac{\lambda_1}{\alpha/2} W_0(G_2)}{1/\delta} - \theta_1 x_2 \right)^{1/\delta} \quad x_1 > x_2 \end{array} \right. | \(W_0(.)\) is the Lambert function* |
| MOBE \(X_1\) | - \((\lambda_1 + \lambda_2 + \lambda_{12})^{-1}\) \ln(1 - \alpha/2) | - \((\lambda_1 + \lambda_2 + \lambda_{12})^{-1}\) \ln(\alpha/2) |
| \(X_2\) | \(x_1 - (\lambda_2 + \lambda_{12})^{-1}\) \ln(1 - \alpha/2) \(x_1 < x_2\) | \(x_2 - (\lambda_1 + \lambda_{12})^{-1}\) \ln(1 - \alpha/2) \(x_1 > x_2\) |
| \(X_1\) | \(x_1 - (\lambda_2 + \lambda_{12})^{-1}\) \ln(\alpha/2) \(x_1 < x_2\) | \(x_2 - (\lambda_1 + \lambda_{12})^{-1}\) \ln(\alpha/2) \(x_1 > x_2\) |
| MOBW \(X_1\) | - \((\lambda_1 + \lambda_2 + \lambda_{12})^{-1}\) \ln(1 - \alpha/2) \(1/\eta\) | - \((\lambda_1 + \lambda_2 + \lambda_{12})^{-1}\) \ln(\alpha/2) \(1/\eta\) |
| \(X_2\) | \(x_1 - (\lambda_2 + \lambda_{12})^{-1}\) \ln(1 - \alpha/2) \(1/\eta\) \(x_1 < x_2\) | \(x_2 - (\lambda_1 + \lambda_{12})^{-1}\) \ln(1 - \alpha/2) \(1/\eta\) \(x_1 > x_2\) |
| \(X_1\) | \(x_1 - (\lambda_2 + \lambda_{12})^{-1}\) \ln(\alpha/2) \(1/\eta\) \(x_1 < x_2\) | \(x_2 - (\lambda_1 + \lambda_{12})^{-1}\) \ln(\alpha/2) \(1/\eta\) \(x_1 > x_2\) |

* The Lambert function \(y = W_0(z)\) gives a solution for the equation \(y \exp(y) = z\).

For details see Lambert (1758), Euler (1779), and the supplementary material.

## 3. Theoretical performance

In this section, we will derive performance equations for the control charts in terms of the Average Time to Signal (ATS). In Section 3.1, we will derive the general results followed in Section 3.2 by the results for the three introduced data models (see Section 2.3).
3.1. **Generic ATS equations**

Given a series of observed bivariate event data \(X^i = (X_1^i, X_2^i)\) for \(i = 1, 2, \ldots\), the data model is assumed as follows:

\[
\begin{align*}
X^i &\sim f(x|\theta) \quad i \leq \tau, \\
X^i &\sim f(x|\theta^*) \quad i > \tau,
\end{align*}
\]

(9)

Given this model, the objective of a control chart is to signal as quickly as possible after \(\tau\). Usually, a control chart’s performance is quantified using the \(\text{ARL}\), defined as the average number of points after \(\tau\) after which a signal is triggered. Therefore, we use the \(\text{ATS}\) because the time between plotted statistics varies when monitoring event data. However, Zwetsloot et al. (2021) highlighted that it is more appropriate to quantify the performance of methods for event data based on the time-to-signal performance metric because the time between plotted statistics varies when monitoring event data. Therefore, we use the \(\text{ATS}\) as the performance metric, defined as the average time elapsed after \(\tau\) until an out-of-control point is signaled. The \(\text{ATS}\) can be computed as \(\text{ATS} = \text{ARL} \times \text{E}[\text{TBE}]\), where

\[
\text{E}[\text{TBE}] = P[X_1 = X_2]E[X(1)|X_1 = X_2] + 0.5 P[X_1 \neq X_2]E[X(1)|X_1 \neq X_2] + 0.5 P[X_1 \neq X_2]E[X(2) - X(1)|X_1 \neq X_2].
\]

(10)

With probability \(P[X_1 = X_2]\), we can observe both events simultaneously, where \(E[X(1)|X_1 = X_2]\) is the expected time for this simultaneous event. With probability \(P[X_1 \neq X_2]\), we can observe either \(X(1)\) or \(X(2)\) since both occur with probability 0.5 and they are multiplied by the expected time for \(X(1)\) and \(X(2)\), respectively.

Since we used a superimposed data stream, not all data points are independent. Thus, the computation of \(\text{ARL}\) and consequently the \(\text{ATS}\), is subtle. Taking this into account, we derive the analytical expression for the \(\text{ATS}\) in Theorem 2. Proof of Theorem 2 can be found in Appendix C.

**Theorem 2.** Assume \(X = (X_1, X_2) \sim f(x|\theta^*)\) as in Equation 9. Let the control limits be defined as in Theorem 1, and obtained from an in-control dataset \(X \sim f(x|\theta)\). The proposed BTBE chart will have an ATS performance equal to:

\[
\text{ATS} = E_x[\text{TBE}] \frac{1 + P_s[NS_1, \neq]}{P_s[S_1] + P_s[NS_1, S_2, \neq]},
\]

(11)

where \(P_s\) and \(E_x\) denote the probability and expectation over \(f(x|\theta^*)\) respectively. \(S_1\) and \(NS_1\) indicate a signal or no signal on the first event, while \(S_2\) indicates a signal on the second event, and \(\neq\) is an abbreviation for \(X_1 \neq X_2\).

Formally, the probabilities are defined as:

\[
\begin{align*}
P_s[NS_1, \neq] & = P_s[LCL(1) < X(1) \leq UCL(1), X_1 \neq X_2], \\
P_s[S_1] & = P_s[X(1) \leq LCL(1)] + P_s[X(1) > UCL(1)], \\
P_s[NS_1, S_2, \neq] & = P_s[LCL(1) < X(1) \leq UCL(1), X(2) \leq LCL(2), X_1 \neq X_2], \\
& \quad + P_s[LCL(1) < X(1) \leq UCL(1), X(2) > UCL(2), X_1 \neq X_2].
\end{align*}
\]

(12)

In Equation (11), the fraction is related to \(1/\alpha\) (used for a univariate Shewhart...
control chart). The ‘α’ in the denominator, is the sum of the probability to observe a signal at the first observed event plus the probability to observe a signal at the second event. The numerator is larger than 1 because we observe two events in the later case. The ARL performance (rather than the ATS performance) is equal to this fraction.

3.2. ATS equations for GBE, MOBE and MOBW models

Next, we compute equation (11) for the GBE, MOBE, and MOBW models in Corollary 1-3. Detailed derivations to obtain the results in Corollary 1-3 can be found in the supplementary material.

Corollary 1. Assume \( X \sim f(x; \theta) \) as in Equation 9, where \( f() \) is the GBE distribution function with \( \theta = (\theta_1, \theta_2, \delta) \) for in-control data, and \( \theta* = (\theta_1^*, \theta_2^*, \delta^*) \) for out-of-control data. The ATS performance of our proposed method when \( \delta = 1 \), is equal to

\[
\text{ATS} = E_\gamma[TBE] \frac{1 + (1 - \alpha^*)}{\alpha^* + (1 - \alpha^*)} \left[ \frac{\theta^*}{\theta^* - \alpha^*} \left( 1 - \left( 1 - \frac{\theta^*}{\theta^* + \alpha^*} \right) \right) \right]
\]

(13)

where \( E_\gamma[TBE] = 0.5(\theta_1^* + \theta_2^*) - C(1,1)^{-1} \), \( C = C^*(1,1)/C(1,1) \) and, \( C(1,1) = (1/\theta_1)^{1/\delta} + (1/\theta_2)^{1/\delta} \), \( \alpha^* = \alpha_L^* + \alpha_U^* = 1 - \left( 1 - \frac{\theta^*}{\theta^* + \alpha^*} \right)^C \), \( \alpha_L^* = \left( \frac{\theta^*}{\theta^* + \alpha^*} \right)^C \).

We note that the analytical expression for the ATS when \( \delta < 1 \) is too complex; thus, we use simulation to obtain the ATS if \( \delta < 1 \).

Corollary 2. Assume \( X \sim f(x; \theta) \) as in Equation 9, where \( f() \) is the MOBE distribution function with \( \theta = (\lambda_1, \lambda_2, \lambda_1^*) \) for in-control data and \( \theta* = (\lambda_1^*, \lambda_2^*, \lambda_1^* \lambda_1^*) \) for out-of-control data. The ATS performance of our proposed method, is equal to

\[
\text{ATS} = E_\gamma[TBE] \frac{1 + \frac{\lambda_1^* + \lambda_2^*}{\Lambda^*} (1 - \alpha^*)}{\alpha^* + (1 - \alpha^*)} \left[ \frac{\Lambda^*}{\Lambda^*} \left( 1 - \left( 1 - \frac{\lambda_1^* + \lambda_2^*}{\lambda_1^* + \lambda_2^*} \right) \right) \right]
\]

(14)

where \( \alpha_L^* = 1 - \left( 1 - \frac{\lambda_1^*}{\lambda_2^*} \right)^\Lambda^*, \alpha_U^* = \left( \frac{\lambda_2^*}{\lambda_1^*} \right)^\Lambda^*, \alpha^* = \alpha_L^* + \alpha_U^*, \Lambda = \lambda_1 + \lambda_2 + \lambda_1^*, \) and

\[
E_\gamma[TBE] = 0.5 \left( \frac{\lambda_1^* + \lambda_2^*}{\Lambda^*} \right)^\Lambda^* + \frac{\lambda_2^*}{\Lambda^*(\lambda_1^* + \lambda_1^*)} + \frac{\lambda_1^*}{\Lambda^*(\lambda_2^* + \lambda_1^*)} = \frac{\lambda_1^* + \lambda_2^*}{\Lambda^*}
\]

Corollary 3. Assume \( X \sim f(x; \theta) \) as in Equation 9, where \( f() \) is the MOBW distribution function, with \( \theta = (\lambda_1, \lambda_2, \lambda_1^*, \eta) \) for in-control data, and \( \theta* = (\lambda_1^*, \lambda_2^*, \lambda_1^* \lambda_1^*, \eta) \) for out-of-control data. The ATS performance of our proposed method, is equal to

\[
\text{ATS} = E_\gamma[TBE] \frac{1 + \frac{\lambda_1^* + \lambda_2^*}{\Lambda^*} (1 - \alpha^*)}{\alpha^* + (1 - \alpha^*)} \left[ \frac{\Lambda^*}{\Lambda^*} \left( 1 - \left( 1 - \frac{\lambda_1^* + \lambda_2^*}{\lambda_1^* + \lambda_2^*} \right) \right) \right]
\]

(15)
where \( \alpha^*_L, \alpha^*_U, \alpha^*, \) and \( \Lambda \) have the same definition as Corollary 2. And

\[
E_*[TBE] = 0.5 E \left( \frac{1}{\eta} \right) \left( \frac{1}{(\lambda^*_2 + \lambda^*_2)^{1/n}} \right) + \frac{1}{(\lambda^*_1 + \lambda^*_1)^{1/n}} - \frac{\lambda^*_1 + \lambda^*_2}{\lambda^* + 1 + 1/\eta} + 2 \frac{\lambda^*_2}{\lambda^* + 1 + 1/\eta}.
\]

The equations in Corollary 1-3 allow us to do performance evaluations of our proposed chart without simulations. Whereas most existing methods for monitoring event data rely on simulations to obtain ATS values.

4. Performance of proposed method

In this section, we evaluate the performance of our proposed BTBE chart. First, we describe our experiments, followed by a performance evaluation.

4.1. Synthetic data experiments

We consider data from either of the three lifetime distributions introduced in Section 2.3. We select four in-control models to evaluate the performance of our method.

- **Scenario 1:** equal expectations for both event times; \( E[X_1] = E[X_2] = 5 \), with \( X_1 \) and \( X_2 \) modeled to be independent.
- **Scenario 2:** equal expectations for both event times; \( E[X_1] = E[X_2] = 5 \), with \( X_1 \) and \( X_2 \) modeled to be dependent.
- **Scenario 3:** unequal expectations for the event times; \( E[X_1] = 5 \) and \( E[X_2] = 15 \), with \( X_1 \) and \( X_2 \) modeled to be independent.
- **Scenario 4:** unequal expectations for the event times; \( E[X_1] = 5 \) and \( E[X_2] = 15 \), with \( X_1 \) and \( X_2 \) modeled to be dependent.

For the performance assessment of the BTBE chart, four different types of shifts are considered for each of the four in-control scenarios:

- **Shift type I1:** Positive shift (increase) in only \( E[X_1] \) by 50 and 100 percent.
- **Shift type I2:** Positive shift (increase) in both \( E[X_1] \) and \( E[X_2] \) by 50 and 100 percent.
- **Shift type D1:** Negative shift (decrease) in \( E[X_2] \) from 30 to 80 percent.
- **Shift type D2:** Negative shift (decrease) in both \( E[X_1] \) and \( E[X_2] \) from 30 to 80 percent.

For each of the three distributions, we have selected parameters (\( \theta_1, \theta_2, \delta \) for GBE, \( \lambda_1, \lambda_2, \lambda_{12} \) for MOBE, and \( \lambda_1, \lambda_2, \lambda_{12}, \eta \) for MOBW) such that the above mentioned expectations are met under in-control and out-of-control scenarios. Table 9 in Appendix B provides the parameter values.

Note that dependence has a careful interpretation when the data is MOBE or MOBW distributed since dependence is related to the probability that \( X_1 = X_2 \). The nature of the MOBE and MOBW models is that \( X_1 \) can be equal to \( X_2 \) even though they are continuous random variables. The rationale is that an external shock could influence both components in a system, and hence they have an equal event time (\( X_1 = X_2 \)). Therefore, to model independence in scenarios 1 and 3, we set \( P[X_1 = X_2] = 0 \) and to model dependence in scenarios 2 and 4, we set \( P[X_1 = X_2] = 0.1 \).

We study two types of shifts. First, we consider the classic scenario of sustained
shifts in the process—once the process changes, it stays that way until it is detected and corrected. The second shift type we study is that of transient shifts. A transient shift appears and disappears again after a certain number of periods.

### 4.2. Performance evaluation for sustained shifts

Table 2 shows the ATS performance of our proposed BTBE chart for the selected experiments and the three distributions. Results in Table 2 were obtained using Corollary 1-3, except for the GBE results in scenarios 2 and 4 (the dependent scenarios \(\delta < 1\)). All codes for replicating the results are available on [https://github.com/tmahmood5/Codes-BTBE-Monitoring-Method](https://github.com/tmahmood5/Codes-BTBE-Monitoring-Method).

| In-control scenario | Shift type | \(E[X_1]\) | \(E[X_2]\) | GBE | MOBE | MOBW |
|---------------------|------------|-------------|-------------|-----|------|------|
| 1. Equal expectations, independence | IC | 5 | 5 | 200.0 | 200.0 | 200.0 |
| OC-I | 7.5 | 5 | 143.8 | 143.9 | 67.0 |
| OC-I | 10 | 5 | 105.7 | 105.7 | 35.9 |
| OC-I | 7.5 | 7.5 | 110.9 | 110.9 | 40.0 |
| OC-I | 10 | 10 | 73.9 | 73.9 | 21.4 |
| OC-D | 5 | 2.5 | 173.3 | 173.3 | 133.6 |
| OC-D | 5 | 1 | 100.9 | 100.9 | 39.6 |
| OC-D | 2.5 | 2.5 | 100.0 | 100.0 | 50.6 |
| OC-D | 1 | 1 | 16.3 | 16.3 | 3.5 |
| 2. Equal expectations, dependence | IC | 5 | 5 | 199.9 | 200.0 | 200.0 |
| OC-I | 7.5 | 5 | 151.0 | 143.6 | 66.9 |
| OC-I | 10 | 5 | 109.0 | 104.8 | 35.4 |
| OC-I | 7.5 | 7.5 | 125.1 | 111.0 | 40.7 |
| OC-I | 10 | 10 | 86.2 | 73.9 | 21.9 |
| OC-D | 5 | 2.5 | 162.5 | 174.2 | 135.9 |
| OC-D | 5 | 1 | 66.3 | 104.0 | 106.5 |
| OC-D | 2.5 | 2.5 | 119.1 | 100.0 | 50.6 |
| OC-D | 1 | 1 | 24.7 | 16.3 | 3.5 |
| 3. Unequal expectations, independence | IC | 5 | 15 | 200.0 | 200.0 | 200.0 |
| OC-I | 7.5 | 15 | 138.9 | 138.9 | 71.5 |
| OC-I | 10 | 15 | 101.8 | 101.8 | 37.3 |
| OC-I | 7.5 | 22.5 | 138.4 | 138.4 | 63.4 |
| OC-I | 10 | 30 | 106.4 | 106.4 | 40.5 |
| OC-D | 5 | 10.5 | 170.3 | 170.3 | 151.9 |
| OC-D | 5 | 7.5 | 129.3 | 129.3 | 89.6 |
| OC-D | 3.5 | 7.5 | 119.7 | 119.7 | 71.5 |
| OC-D | 2.5 | 7.5 | 100.0 | 100.0 | 51.5 |
| 4. Unequal expectations, dependence | IC | 5 | 15 | 197.7 | 200.0 | 200.0 |
| OC-I | 7.5 | 15 | 124.8 | 140.1 | 73.8 |
| OC-I | 10 | 15 | 87.1 | 102.8 | 38.4 |
| OC-I | 7.5 | 22.5 | 147.7 | 138.5 | 63.9 |
| OC-I | 10 | 30 | 114.0 | 106.5 | 41.1 |
| OC-D | 5 | 10.5 | 145.9 | 171.0 | 153.2 |
| OC-D | 5 | 7.5 | 97.2 | 130.0 | 90.8 |
| OC-D | 3.5 | 7.5 | 113.5 | 119.6 | 71.2 |
| OC-D | 2.5 | 7.5 | 117.6 | 100.0 | 51.5 |

From Table 2, we conclude that our method can signal shifts for all three data distributions. The method is slightly slower when the data are dependent (scenarios 2 and 4) compared to similar independent scenarios (1 and 3), but the difference is
negligible. Next, we compare the results of our BTBE method for observations with equal and unequal expectations (scenarios 1 and 2 versus scenarios 3 and 4). ATS values for scenarios 3 and 4 are larger compared to scenarios 1 and 2. This is a direct result that the expectation of $X_2$ is larger in scenarios 3 and 4.

The two-sided chart applied with the MOBW data has good performance for detecting decreases in both variables. The charts are all designed to have an exact in-control ATS of 200, obtained using the analytical expression. The simulation results for GBE with $\delta < 1$ (scenarios 2 and 4) show a little bit of simulation error. We note that the ATS values for GBE and MOBE under the independent scenarios (1 and 3) are equal when the data are independent, as the models are equal.

4.3. Performance evaluation for transient shifts

In this section, we look at transient shifts, i.e. shifts that come and disappear again. These types of shifts are particularly of interest in application fields where it is difficult to stop a process after a signal has occurred. Transient shifts are common in disease surveillance, as Fricker Jr and Burkom (2021) point out "eventually a disease outbreak subsides and the incidence of the disease returns to some sort of normal background rate". This is the case in Wang and Zwetsloot (2021) where a dengue outbreak lasts one month. Or consider Mahmood et al. (2019) who develop a method for health surveillance that signals outliers, i.e. elder patients with abnormal vital signs. Another field with transient shifts is network monitoring. For example, Yu et al. (2022) study shifts of length 50. Social media data often display transient shifts as illustrated in, for example, Zwetsloot and Woodall (2021); Sparks and Paris (2018); Sparks et al. (2019).

Overall, transient shifts occur in many fields and are important to detect. The anticipated length of any important transient shift in the process has an impact on the performance of any monitoring method. The length of the transient shift of interest will vary from application to application. Very short transient shifts can be modeled as outliers and in the limit, a very long transient shift will be equal to a sustained shift. For conciseness, we only study these two extremes and point out that the detection performance of our methods for transient shifts of moderate length will be somewhere in between the performance for these two extremes presented in Section 4.2 (sustained shifts) and this section (short transient shifts equal to outliers). Hence, for our purpose, we model the transient shift as an outlier. The shift influenced one event vector and the process returns to its in-control state after this event vector.

To study the performance of our methods, we consider the Detection Probability (DP), which is defined as the probability that the chart will signal an out-of-control process at the time point of the transient shift. The Detection Probability is defined as $DP_\ast = E_\ast[TBE]/ATS_\ast$ where we take ATS from Corollary 1-3, and the star denotes that these are computed using the out-of-control process parameters.

Figure 2 shows the DP of our proposed chart for the three considered data models for each of the 4 scenarios and for a range $E[X_1]$ and $E[X_2]$ values.

The in-control detection rate is set at 5 percent, and we see that all charts have a detection rate of 5 percent when $E[X_1] = 5$ and $E[X_2] = 5$, for scenarios 1 and 2 and $E[X_2] = 15$ for scenarios 3 and 4. The detection power of the transient shifts increases as the shift size increases. In particular, for the MOBW model, the detection probability increases rapidly. Decreases in $E[X_1]$ are easier to detect for the dependent models (scenarios 2 and 4).
Figure 2. DP values for our BTBE chart under transient shifts for various data distributions and scenarios

5. Comparative Analysis

In this section, we compare our proposed BTBE chart with the existing method by Xie et al. (2011). We select this method as it is the most well-known and widely studied method. Other options would be the more recent method proposed by Xie et al. (2021),
who extended the idea of Xie et al. (2011) by using the Multivariate Cumulative Sum (MCUSUM) control chart. Koutras and Sofikitou (2017) and Triantafyllou and Panayiotou (2020) used control charts based on the order statistic to monitor the bivariate vector-based data. A two-level multivariate Bayesian control chart based on the Marshall-Olkin Bivariate Exponential (MOBE) distributed data was proposed by Duan et al. (2020). For bivariate vector-based event data, copula based MEWMA, multivariate double EWMA and MCUSUM charts were proposed by Kuvattana and Sukparungsee (2015), Sasiwannapong et al. (2022), and Sukparungsee et al. (2021), and the Hotelling’s $T^2$ chart based on the different type of copulas was discussed by Sukparungsee et al. (2018). For the multivariate vector-based event data, copula based MCUSUM chart was proposed by Sukparungsee et al. (2017) and the MEWMA charts based on transformed exponential data and asymmetric gamma distributions were discussed by Khan et al. (2018) and Flury and Quaglino (2018), respectively.

Table 3. Limit $h$ for the MEWMA chart against a fixed $ATS_0 = 200$

| Scenario | 1  | 2  | 3  | 4  |
|----------|----|----|----|----|
| $\lambda = 0.1$ | 3.60 | 3.87 | 2.09 | 2.12 |
| $\lambda = 1$  | 9.51 | 11.40 | 5.33 | 5.86 |

The method by Xie et al. (2011) is a Multivariate EWMA (MEWMA) chart designed for GBE distributed data only. The MEWMA statistic is defined as: $z_i = r(X_i - \mu_X) + (1 - r)z_{i-1}$ and the charting statistics is equal to $E_i = z_i^T \Sigma_{Z_i}^{-1} z_i$ where $r \in (0, 1]$ is the EWMA smoothing parameter, $\Sigma_{Z_i} = r/(2 - r) \Sigma_X$ and, $\mu_X$ and $\Sigma_X$ are the mean vector and covariance matrix, respectively (see Xie et al. (2011) for more details). The MEWMA chart signals when $E_i > h$ and this chart converts to a Hotelling’s $T^2$ chart when $r = 1$. We run the MEWMA chart with $r = 0.1$ and $r = 1$.

Table 4 gives the $ATS$ values for our chart and the MEWMA chart when the data are drawn from a GBE distribution with sustained shift with scenarios similar to those considered in Section 4. Results for our method when $\delta = 1$, are obtained using the $ATS$ expression in Corollary 1 while the results for $\delta < 1$ and the MEWMA chart are obtained using 10,000 Monte Carlo simulations.

5.1. Performance comparison for sustained shifts

We design our method and the MEWMA chart with an in-control overall $ATS_0 = 200$. The control limits for our BTBE method are taken from Table 1. For the MEWMA chart, we obtain the control limits using simulation. The limits are reported in Table 3. Note that these limits are different compared to the limits reported by Xie et al. (2011) because we use the $ATS$ as a performance measure, and they used the $ARL$.

We compare the performance for two types of scenarios: 1) for detecting sustained shifts in the process parameters and 2) for detecting transient shifts in the process. In Section 5.1, we present the performance comparisons for sustained shifts, and in Section 5.2, we present the performance comparison for transient shifts and outliers.
The results in Table 4 show that our method has better detection power for decreasing shifts compared to MEWMA with \( r = 1 \). This is a desirable property as usually decreases in the expected time between events indicate a deterioration of the system under consideration. Note that for many decreasing scenarios, the MEWMA chart has no detection power (the ATS values are larger than the in-control value of 200). When we compare our methods to the MEWMA with \( r = 0.1 \), we see that it is a bit slower in detecting small sustained decreasing shifts. This can be expected as we do not incorporate any memory into our monitoring statistics. There is a potential to improve our method by using a time-weighted scheme on top of our proposed monitoring statistics.

When we consider increasing shifts, we see that in scenarios 1 and 2, the equal expectation scenarios, the MEWMA chart is quicker in detecting the change. For the more realistic scenarios of unequal expectations (scenarios 3 and 4), our method’s performance is a little slower than the MEWMA method. We believe scenarios 3 and 4 to be more realistic than scenarios 1 and 2 as seldom multiple components of one system will have identical expected lifetimes.

5.2. Performance comparison for transient shifts

Next, we compare the performance in detecting transient shifts for our method with the MEWMA methods based on \( r = 0.1 \) and \( r = 1 \) for each of the four scenarios. Figure 3 shows the results. The y-axis shows the percent point difference between the DP for the BTBE chart and the MEWMA chart. So a value above 0 means that our chart has a larger detection probability, and a value below 0 means that the MEWMA chart has a better chance of detecting the transient shift.

Unsurprisingly, the MEWMA chart based on \( r = 0.1 \) (left column in figure 3) does not have strong performance for detecting transient shifts. The DP is a lot higher for our method, especially for decreasing shifts (\( \theta_1 < 5 \) and \( \theta_2 < 5 \) for scenarios 1 and 2, and \( \theta_2 < 15 \) for scenarios 3 and 4). For scenarios 3 and 4, the MEWMA has a slightly larger DP in the range of \( \theta_1 \) between 2.5 and 7.5 but the difference with our method is very small.

The MEWMA chart based on \( r = 1 \) performs better than the MEWMA chart based on \( r = 0.1 \) in detecting transient shifts, which is as expected. The performance of the MEWMA with \( r = 1 \) compared to our BTBE methods depends on the type of shift. For decreasing shifts with unequal expectations (scenarios 3 and 4), the MEWMA is strongly outperformed by our proposed methods. For larger increasing shifts in either \( \theta_1 \) or \( \theta_2 \), the MEWMA has slightly better performance, but the difference is small with our method.

Overall, our method performs well for detecting decreasing and increasing shifts when the expected lifetimes are unequal among components. Though performance is similar to the MEWMA chart for the GBE data with sustained shifts, our method is more generally applicable (to any distribution) and detects transient shifts quicker. Hence, we recommend using our method, unless the practitioner is interested in monitoring GBE data only and is confident all shifts will always be sustained in nature.
Figure 3. Difference in DP values for our BTBE chart versus the MEWMA chart under transient shifts for the 4 scenarios

6. Application to two case studies

6.1. Application to failure and repair time of a system

In this case study, we use fault data to monitor the health condition of escalators in order to help decision-making regarding maintenance. The project is a collaboration research project between City University of Hong Kong and a Hong Kong based company that operates many escalators in various buildings throughout Hong Kong. Due to confidentiality, many project details cannot be shared.

A typical escalator has a projected lifetime of 30 years, can transport up to 8000 passengers per hour, and usually operates about 20 hours per day. Regular examinations and preventive maintenance are necessary for escalators to ensure safe operation. An escalator may break down in its service lifetime, and we regard these breakdowns as faults. Faults occur because of different reasons. Firstly, escalators are engineering systems with multiple components and parts, which have risks of breaking down and causing system faults. Secondly, unpredictable external factors can cause faults, including passenger accidents, jammed objects, etc.

Instead of quantifying each fault, the time between faults can be used to monitor the escalator’s health status. A long time between two faults may indicate a good
health condition of the escalator. On the contrary, a short time interval between two faults indicates a bad health status. Repairing a fault after its occurrence is necessary for the maintenance work and should be finished as soon as possible. Hence, the time to repair, which is the time cost to check and fix the fault, is also of interest. A short repair time is preferable, indicating less severe faults, efficient maintenance work, and good service.

The data we use in this case study is the 7-year fault record of an escalator (2013-2020). For each fault, we record the time since its previous fault in days ($X_1$) and the time to repair it in hours ($X_2$) (cf. Figure 4). Though the fault is observed earlier than its repair, the number of hours can be smaller than the number of days. Therefore, we include the setting $X_1 > X_2$. The data fits our method as the data are vector-based time-between-events data. This project aims to develop a monitoring framework that can signal significant changes in the time to break down ($X_1$), the time to repair ($X_2$), or both. Decreases in $X_1$ and increases in $X_2$ are of our concern. A shorter time between two faults indicates an increasing frequency of accidents related to the escalator’s bad status. A longer repair time implies more maintenance costs and significant effects on service quality.

To implement our BTBE chart, we use the first 100 faults records as in-control data, where a record indicates a vector with time to fault and time to repair. The remaining 145 records are used for prospective monitoring. For the in-control data, we fit several distributions using the R-package fitdistrplus (Delignette-Muller et al., 2015) and conclude that $X_1$ and $X_2$ follow a Weibull distribution. Marshall and Olkin (1967) stated that a bivariate dataset fits the MOBW distribution if (i) the marginal distribution of each variable follows a Weibull distribution, and (ii) $\min(X_1, X_2)$ fol-
lows a Weibull distribution. From the above analysis, it is concluded that a MOBW model best fits our data.

We estimate the MOBW parameters for the in-control data using the EM algorithm proposed by Kundu and Dey (2009) and obtained $\eta = 1.1677$, $\lambda_1 = 0.0435$ $\lambda_2 = 0.0105$ and $\lambda_{12} = 5.78e^{-0.8}$. With the estimated parameters and the Corollary 3, the expected time between events is estimated as 24.68, and $ARL$ is set at 50. Hence, the control limits given in Table 1 are set on the basis of $ATS_0 = 1233.75$ and yield as: $LCL(1) = 0.2369$, $UCL(1) = 45.0326$ for the first event time and $LCL(2) = (x_1^{1.1677} + 0.9595)^{0.8564}$, $UCL(2) = (x_1^{1.1677} + 439.6717)^{0.8564}$ for the second event time when $X_1 < X_2$. While for $X_1 > X_2$, control limits for the second event time are $LCL(2) = (x_1^{1.1677} + 0.2309)^{0.8564}$ and $UCL(2) = (x_1^{1.1677} + 105.7797)^{0.8564}$.

![Original Plot](image1.png)

![Zoom Plot](image2.png)

**Figure 5.** The BTBE chart for the escalator’s failure case study.

The proposed BTBE chart for the escalator’s failure dataset is plotted in Figure 5. However, a zoomed version of the original plot is also plotted in Figure 5 for further enhancement of the main plot’s readability. The chart plots the univariate superimposed data stream: the time to faults and the time to repair for the first fault, followed by the time to faults and the repair time for the second fault. The order of time to fault and time to repair for each fault depends on their numerical values. The chart showed seven signals in total, three out of them due to the decrease of time to faults (i.e., indicated small “1”) and remaining are caused by the increase of time to repair—second event signals—(i.e., marked with a small “2”). This case shows that the proposed method can identify the outliers in the dataset. In particular, the second and third signals indicate a very long repair time which could be investigated. The fourth signal is a decrease in time to failure, which should be monitored to ensure it does not become a persistent change.
6.2. Application to AIDS Data

In this section, we implement the proposed BTBE monitoring method to the AIDS dataset obtained from the Centers for Disease Control (CDC) in Atlanta, Georgia, which is also available in R-package SurvTrunc (Rennert, 2018). In the data, we have a total of 295 people, among which 258 are adults and 37 are children. All people in the sample got AIDS through contaminated blood transfusion. For each person, we have a time to HIV infection (the first event $X(1)$ referred to as infection time), which is the time between blood transfusion and infection. We have the total time to AIDS diagnosis (the second event $X(2)$ referred to as total incubation time), which is the time between blood transfusion and AIDS diagnosis. Therefore, the total incubation time is always larger than the infection time for each patient. We have excluded one person from the data whose event time equals zero and most likely contracted AIDS before the blood transfusion.

The data is summarized in Table 5 and visualized in Figure 6. The infection time ($X(1)$) is significantly higher for children than for adults (at a 5% significance level). The total incubation time ($X(2)$) is shorter for children than for adults (at a 10% significance level). These results align with Hu et al. (2014), who concluded that children, compared to adults, have shorter HIV incubation times. Note that we will work with transformed data (division by 100) to have shape and scale parameters of the same size (Kundu and Dey, 2009).

![AIDS data for each subject.](image)

To illustrate our BTBE chart, we use the adult’s data as in-control data and the children’s data as shifted data (upward shifted for $X(1)$ and downward shifted for $X(2)$).
In this dataset, $X_1 < X_2$ for all observations/people, hence it follows that $X_{(1)} = X_1$ and $X_{(2)} = X_2$ for all observations/people. We first fit a distribution to the data in a Phase I analysis to implement our chart. We used the R-package `fitdistrplus` (Delignette-Muller et al., 2015) to evaluate various distributions. Figure 7 shows the Q-Q plots for $X_{(1)}$ and $X_{(2)}$, which concludes that a MOBW model best fits our data. The estimated MOBW parameters of the in-control (adults) data are $\eta = 4.31$, $\lambda_1 = 0.574$, $\lambda_2 = 0.905$, and $\lambda_{12} = 1.12$.

Figure 7. QQ-plot; (a) for $X_{(1)}$, and (b) for $X_{(2)}$.

Using these estimates, the control limit formulas in Table 1 yield: $LCL_{(1)} = 0.180$, $UCL_{(1)} = 0.794$ for the first event time and $LCL_{(2)} = \left(x_1^{4.311} + 0.00374\right)^{0.232}$, $UCL_{(2)} = \left(x_1^{4.311} + 2.247\right)^{0.232}$ for the second event time. We have set $AT_{S0} = 25$ because the expected time between events is approximately 0.4, yielding about one false alarm for approximately every 60 events, i.e., 30 observations/people.

The proposed BTBE chart for the AIDS dataset is plotted in Figure 8. The chart plots the univariate superimposed data stream: the infection time $X_{(1)}$ and total incubation time $X_{(2)}$ of the first person, followed by the infection time and total incubation time of the second person, etc. The chart shows that the infection time signals five times (the small “1” indicates that the signal is related to a first event). This is in line with the results of Table 5.
7. Concluding Remarks and Recommendations

This article proposed a novel method, named BTBE chart, to monitor sustained or transient shifts in multivariate time-between-events data. The strengths of the proposed methods are that 1) it has real-time detection power and does not have a built-in delay like all existing multivariate time-between-event methods; 2) it can be used for any bivariate distribution of the underlying data, and it only needs the CDF and survival functions for control limit computation; 3) it has analytical expressions for the control limits and performance metrics making it quick to run; 4) it shows reasonable performance compared to other methods when data are GBE distributed (especially for transient shifts); 5) it gives insight into the underlying variable that shifted. Overall, the method is more generally applicable than the existing MEWMA method and performs excellently, especially when the observations have unequal time-between-events and/or are transient in nature, which we consider a more realistic scenario than equal expectations. Under equal expectations and sustained shifts, the MEWMA outperforms for GBE data only.

Future interesting work would be to extend the method to point-process data. Our method is only applicable for the monitoring of vector-based event data. Also, adapting our bivariate chart to multivariate data will be of interest. In the current methodology, the history of the events is not accumulated, so a method based on EWMA and CUSUM type structures will also be an interesting issue for future research. A potential to fine-tune our method is to have a signal once time-between-events gets bigger than the UCL rather than to wait until we observe time-between-events.
Data Availability Statement

Data available on request from the authors.

Acknowledgement

The work of Inez M. Zwetsloot described in this paper was partially supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China (Project No. CityU 21215319) and by a grant from City University of Hong Kong (Project No. 7005567). Further, the author Tahir Mahmood at KFUPM acknowledges the Interdisciplinary Research Center for Smart Mobility and Logistics for the support received under Grant no. INML2205.

Conflict of Interest Disclosure

The authors have no conflicts of interest to declare. All co-authors have seen and agree with the contents of the manuscript and there is no financial interest to report. We certify that the submission is original work and is not under review at any other publication.

ORCID

Inez M. Zwetsloot: https://orcid.org/0000-0002-6144-4188
Tahir Mahmood: https://orcid.org/0000-0002-8748-5949
Funmilola Mary Taiwo: https://orcid.org/0000-0002-4025-0878
Zezhong Wang: https://orcid.org/0000-0002-7147-1489

Supplementary Material

The supplementary material of this manuscript, which consists of some proofs is also provided in form of pdf.

References

Ali, S., A. Pievatolo, and R. Göb (2016). An overview of control charts for high-quality processes. Quality and Reliability Engineering International 32(7), 2171–2189.
Bemis, B. M., L. J. Bain, and J. J. Higgins (1972). Estimation and hypothesis testing for the parameters of a bivariate exponential distribution. Journal of the American Statistical Association 67(343), 704–706.
Bhattacharyya, G. and R. A. Johnson (1973). On a test of independence in a bivariate exponential distribution. Journal of the American Statistical Association 68(343), 704–706.
Byar, D. (1980). The veterans administration study of chemoprophylaxis for recurrent
stage i bladder tumours: comparisons of placebo, pyridoxine and topical thiotepa. In *Bladder tumors and other topics in urological oncology*, pp. 363–370. Springer.

Chiou, S. H., G. Xu, J. Yan, and C.-Y. Huang (2018). Semiparametric estimation of the accelerated mean model with panel count data under informative examination times. *Biometrics* 74(3), 944–953.

Delignette-Muller, M. L., C. Dutang, et al. (2015). fitdistrplus: An r package for fitting distributions. *Journal of Statistical Software* 64(4), 1–34.

Delignette-Muller, M. L., C. Dutang, R. Pouillot, J.-B. Denis, and M. M. L. Delignette-Muller (2015). Package ‘fitdistrplus’. *Journal of Statistical Software* 64(4), 1–34.

Duan, C., V. Makis, and C. Deng (2020). A two-level bayesian early fault detection for mechanical equipment subject to dependent failure modes. *Reliability Engineering & System Safety* 193, 106676.

Euler, L. (1779). *De serie lambertina plurimis queeius insignibus proprietatibus* (on the remarkable properties of a series of lambert and others). *Opera Omnia* (Series 1) 6, 350–369.

Feizjavadian, S. and R. Hashemi (2015). Analysis of dependent competing risks in the presence of progressive hybrid censoring using marshall–olkin bivariate weibull distribution. *Computational Statistics & Data Analysis* 82, 19–34.

Flury, M. I. and M. B. Quaglino (2018). Multivariate ewma control chart with highly asymmetric gamma distributions. *Quality Technology & Quantitative Management* 15(2), 230–252.

Fricker Jr, R. D. and H. S. Burkom (2021). Data aggregation in disease surveillance. *Journal of Quality Technology* 53(1), 38–43.

Gross, A. J. and C. F. Lam (1981). Paired observations from a survival distribution. *Biometrics* 37(3), 505–511.

Gumbel, E. (1960). Bivariate exponential distributions. *Journal of the American Statistical Association* 55(292), 698–707.

Hougaard, P. (1986). A class of multivariate failure time distributions. *Biometrika* 73(3), 671–678.

Hu, T., X. Lin, and B. Nan (2014). Cross-ratio estimation for bivariate failure times with left truncation. *Lifetime Data Analysis* 20(1), 23–37.

Huster, W. J., R. Brookmeyer, and S. G. Self (1989). Modelling paired survival data with covariates. *Biometrics* 45(1), 145–156.

Jones-Farmer, L. A., W. H. Woodall, S. H. Steiner, and C. W. Champ (2014). An overview of phase i analysis for process improvement and monitoring. *Journal of Quality Technology* 46(3), 265–280.

Khan, N., M. Aslam, M. S. Aldosari, and C.-H. Jun (2018). A multivariate control chart for monitoring several exponential quality characteristics using ewma. *IEEE Access* 6, 70349–70358.

Kotz, S., N. Balakrishnan, and N. L. Johnson (2004). *Continuous multivariate distributions, Models and applications*, Volume 1. John Wiley & Sons.

Koutras, M. and E. Sofikitou (2017). A new bivariate semiparametric control chart based on order statistics and concomitants. *Statistics & Probability Letters* 129, 340–347.

Kundu, D. and A. K. Dey (2009). Estimating the parameters of the marshall–olkin bivariate weibull distribution by em algorithm. *Computational Statistics & Data Analysis* 53(4), 956–965.

Kuvattana, S. and S. Sukparungsee (2015). Comparative the performance of control charts based on copulas. In *The World Congress on Engineering and Computer Science*, pp. 47–58. Springer.

25
Lai, C.-D., G. Dong Lin, K. Govindaraju, and S. Pirikahu (2017). A simulation study on the correlation structure of Marshall–Olkin bivariate Weibull distribution. *Journal of Statistical Computation and Simulation* 87(1), 156–170.

Lambert, J. H. (1758). Observationes variae in mathematicam puram. *Acta Helvetica* 3(1), 128–168.

Li, Y., J. Sun, and S. Song (2012). Statistical analysis of bivariate failure time data with Marshall–Olkin Weibull models. *Computational Statistics & Data Analysis* 56(6), 2041–2050.

Lu, J. and G. Bhattacharyya (1991). Inference procedure for bivariate exponential distributions. *Statistics & Probability Letters* 12(1), 37–50.

Mahmood, T., P. Wittenberg, I. M. Zwetsloot, H. Wang, and K. L. Tsui (2019). Monitoring data quality for telehealth systems in the presence of missing data. *International Journal of Medical Informatics* 126, 156–163.

Mahmood, T. and M. Xie (2019). Models and monitoring of zero-inflated processes: The past and current trends. *Quality and Reliability Engineering International* 35(8), 2540–2557.

Marshall, A. W. and I. Olkin (1967). A multivariate exponential distribution. *Journal of the American Statistical Association* 62(317), 30–44.

Montgomery, D. C. (2017). *Design and analysis of experiments* (9-th ed.). John Wiley & Sons.

Nelson, W. (1982). *Applied life data analysis, 1982*. John Wiley & Sons.

Proschan, F. and P. Sullo (1976). Estimating the parameters of a multivariate exponential distribution. *Journal of the American Statistical Association* 71(354), 465–472.

Qiu, P., Z. He, and Z. Wang (2019). Nonparametric monitoring of multiple count data. *IISE Transactions* 51(9), 972–984.

Rennert, L. (2018). *SurvTrunc: Analysis of Doubly Truncated Data*. R package version 0.1.0.

Saghir, A. and Z. Lin (2015). Control charts for dispersed count data: an overview. *Quality and Reliability Engineering International* 31(5), 725–739.

Sasiwannapong, S., S. Sukparungsee, P. Busababodhin, and Y. Areepong (2022). The efficiency of constructed bivariate copulas for newma and hotelling’s $t^2$ control charts. *Communications in Statistics-Simulation and Computation* 51(4), 1837–1851.

Sparks, R., B. Jin, S. Karimi, C. Paris, and C. MacIntyre (2019). Real-time monitoring of events applied to syndromic surveillance. *Quality Engineering* 31(1), 73–90.

Sparks, R., A. Joshi, C. Paris, S. Karimi, and C. R. MacIntyre (2020). Monitoring events with application to syndromic surveillance using social media data. *Engineering Reports* 2(5), e12152.

Sparks, R. and C. Paris (2018). Sentiment monitoring of social media from oceania. *Journal of Health & Medical Informatics* 9(2), 100312.

Sukparungsee, S., S. Kuvattana, P. Busababodhin, and Y. Areepong (2017). Multivariate copulas on the mcusum control chart. *Cogent Mathematics* 4(1), 1342318.

Sukparungsee, S., S. Kuvattana, P. Busababodhin, and Y. Areepong (2018). Bivariate copulas on the hotelling’s $t^2$ control chart. *Communications in Statistics-Simulation and Computation* 47(2), 413–419.

Sukparungsee, S., S. Sasiwannapong, P. Busababodhin, and Y. Areepong (2021). The effects of constructed bivariate copulas on multivariate control charts effectiveness. *Quality and Reliability Engineering International* 37(5), 2156–2168.

Triantafyllou, I. S. and N. I. Panayiotou (2020). Distribution-free monitoring schemes
based on order statistics: a general approach. *Journal of Applied Statistics* **47**(12), 2230–2257.

Wang, Z. and P. Qiu (2018). Count data monitoring: Parametric or nonparametric? *Quality and Reliability Engineering International* **34**(8), 1763–1774.

Wang, Z. and I. M. Zwetsloot (2021). Exploring the usefulness of functional data analysis for health surveillance. In *Frontiers in Statistical Quality Control* **13**, pp. 247–264. Springer.

Wienke, A. (2010). *Frailty models in survival analysis*. Chapman and Hall/CRC.

Xie, F., J. Sun, P. Castagliola, X. Hu, and A. Tang (2021). A multivariate cusum control chart for monitoring gumbel’s bivariate exponential data. *Quality and Reliability Engineering International* **37**(1), 10–33.

Xie, Y., M. Xie, and T. N. Goh (2011). Two mewma charts for gumbel’s bivariate exponential distribution. *Journal of Quality Technology* **43**(1), 50–65.

Yu, L., I. M. Zwetsloot, N. T. Stevens, J. D. Wilson, and K. L. Tsui (2022). Monitoring dynamic networks: A simulation-based strategy for comparing monitoring methods and a comparative study. *Quality and Reliability Engineering International* **38**(3), 1226–1250.

Zwetsloot, I. M., T. Mahmood, and W. H. Woodall (2021). Multivariate time-between-events monitoring–an overview and some (overlooked) underlying complexities. *Quality Engineering* **33**(1), 13–25.

Zwetsloot, I. M. and W. H. Woodall (2021). A review of some sampling and aggregation strategies for basic statistical process monitoring. *Journal of Quality Technology* **53**(1), 1–16.
Appendix A

This appendix provided the derivations for definition 1 and 2.

A.1 Derivations for Definition 1

This proof is based on censored data modeling. We follow similar steps to Wienke (2010).

Proof of Definition 1. We derive the conditional cumulative distribution function for the first observed event time $X_{(1)}$ for each case $V = 0$, $V = 1$, and $V = 2$, separately.

When $V = 0$, the conditional CDF is defined as $F_{X_{(1)}}(x_{(1)}|v) = P[X_1 \leq x_{(1)}, V = 0]/P[V = 0]$, for which

$$P[V = 0] = \int_0^\infty \int_{x_1}^\infty f(x_1, x_2) \, dx_2 \, dx_1 = \int_0^\infty \int_{x_1}^\infty \frac{\partial}{\partial x_2} S_1(x_1, x_2) \, dx_2 \, dx_1$$

$$= \int_0^\infty S_1(x_1, x_2) \bigg|_{x_2=x_1} \, dx_1 = - \int_0^\infty S_1(x_1, x_1) \, dx_1$$

where $f(x_1, x_2) = \frac{\partial^2}{\partial x_1 \partial x_2} S_1(x_1, x_2) = \frac{\partial}{\partial x_2} S_1(x_1, x_2)$ and $S_1(x_1, \infty) = 0$. We also obtain

$$P[X_1 \leq x_{(1)}, V = 0] = \int_0^{x_{(1)}} \int_{x_1}^\infty f(x_1, x_2) \, dx_2 \, dx_1 = \int_0^{x_{(1)}} \int_{x_1}^\infty \frac{\partial}{\partial x_2} S_1(x_1, x_2) \, dx_2 \, dx_1$$

$$= \int_0^{x_{(1)}} S_1(x_1, x_2) \bigg|_{x_2=x_1} \, dx_1 = - \int_0^{x_{(1)}} S_1(x_1, x_1) \, dx_1.$$

Therefore, $F_{X_{(1)}}(x_{(1)}|v) = \int_0^{x_{(1)}} S_1(x_1, x_1) \, dx_1$ if $v = 0$. A symmetric argument gives $F_{X_{(1)}}(x_{(1)}|v) = \int_0^{x_{(1)}} S_2(x_2, x_2) \, dx_2$ if $v = 1$.

When $V = 2$, the conditional CDF is defined as $F_{X_{(1)}}(x_{(1)}|v) = P[X_1 \leq x_{(1)}, V = 2]/P[V = 2]$, where $P[X_1 \leq x_{(1)}, V = 2] = \int_0^{x_{(1)}} f(x_1, x_1) \, dx_1$ and $P[V = 2] = \int_0^\infty f(x_1, x_1) \, dx_1$.

Similarly, the conditional survival function is defined as $S_{X_{(1)}}(x_{(1)}|v) = P[X_1 > x_{(1)}, V = 0]/P[V = 0]$ if $v = 0$, for which

$$P[X_1 > x_{(1)}, V = 0] = \int_{x_{(1)}}^\infty \int_{x_1}^\infty f(x_1, x_2) \, dx_2 \, dx_1 = \int_{x_{(1)}}^\infty \int_{x_1}^\infty \frac{\partial}{\partial x_2} S_1(x_1, x_2) \, dx_2 \, dx_1$$

$$= \int_{x_{(1)}}^\infty S_1(x_1, x_2) \bigg|_{x_2=x_1} \, dx_1 = - \int_{x_{(1)}}^\infty S_1(x_1, x_1) \, dx_1.$$

So it follows that $S_{X_{(1)}}(x_{(1)}|v) = \int_0^{x_{(1)}} \frac{S_1(x_1, x_1) \, dx_1}{S_1(x_1, x_1)}$ if $v = 0$. Equivalently, we obtain
Lemma 1. The probability density function of $X(1)$ is given as

$$f_{X(1)}(x(1)|v) = \begin{cases} -S_1(x(1), x(1)) & \text{if } v = 0 \\ -S_2(x(2), x(2)) & \text{if } v = 1. \end{cases}$$ \hspace{1cm} (16)$$

Proof of Lemma 1. From Definition 1, we know that $P[X(1) \leq x(1), V = 0] = -\int_0^{x(1)} S_1(x(1), x(1)) \, dx$. The probability density function of $X(1)$ when $V = 0$ is

$$f_{X(1)}(x(1)|V = 0) = \frac{d}{dx} \int_0^{x(1)} -S_1(x, x) \, dx = -S_1(x(1), x(1)).$$

Similarly, for $V = 1$, we obtained $f_{X(1)}(x(2)|V = 1) = -S_2(x(2), x(2))$. This completes the proof.

Lemma 2. The partial distribution function of $(X(1), X(2))$ with respect to $X(1)$ is defined as $F_{X(1)}(x(1), x(2)|v) = P[X(1) = x(1), X(2) \leq x(2)|V = v]$ which is equal to

$$F_{X(1)}(x(1), x(2)|v) = \begin{cases} S_1(x(1), x(2)) - S_1(x(1), x(1)) & \text{if } v = 0 \\ S_2(x(1), x(2)) - S_2(x(2), x(2)) & \text{if } v = 1. \end{cases}$$ \hspace{1cm} (17)$$

Proof of Lemma 2. For $v = 0$, it follows that

$$F_{X(1)}(x(1), x(2)|V = 0) = P[X_1 = x_1, X_2 \leq x_2|V = 0] = P[X_1 = x_1, x_1 < X_2 \leq x_2] = \int_{x_1}^{x_2} f(x_1, x_2) \, dx_2 = \int_{x_1}^{x_2} \frac{\partial}{\partial x_2} S_1(x_1, x(2)) \, dx_2 = S_1(x_1, x(2)) \bigg|_{x_2 = x_2} - S_1(x_1, x(1)).$$

Equivalently, when $v = 1$, it follows that $F_{X(1)}(x(1), x(2)|V = 1) = S_2(x_1, x_2) - S_2(x_2, x_2)$. This completes the proof.

Lemma 3. The partial survival function of $(X(1), X(2))$ with respect to $X(1)$ is defined as $S_{X(1)}(x(1), x(2)|v) = P[X(1) = x(1), X(2) > x(2)|V = v]$ and is equal to

$$S_{X(1)}(x(1), x(2)|v) = \begin{cases} -S_1(x(1), x(2)) & \text{if } v = 0 \\ -S_2(x(1), x(2)) & \text{if } v = 1. \end{cases}$$ \hspace{1cm} (18)$$
Proof of Lemma 3. For \( v = 0 \) it follows that

\[
S_{X(1)}(x(1), x(2)|V = 0) = P[X_1 = x_1, X_2 > x_2, V = 0] = \int_{x_2}^{\infty} f(x_1, x_2)dx_2 = \int_{x_2}^{\infty} \frac{\partial}{\partial x_2}S_1(x_1, x_2)dx_2 = S_1(x_1, x_2)\Bigg|_{x_2=\infty} = -S_1(x_1, x_2).
\]

Equivalently, when \( v = 1 \), it follows that \( S_{X(1)}(x(1), x(2)|V = 1) = -S_2(x_1, x_2) \). This completes the proof.

Proof of Definition 2. By definition, the conditional distribution function is

\[
F_{X(2)|X(1), V}(x(2)|x(1), v) = P[X(2) \leq x(2)|X(1) = x(1), V = v] = \frac{P[X(1) = x(1), X(2) \leq x(2)|V = v]}{P[X(1) = x(1)|V = v]} = \frac{F_{X(1)}(x(1), x(2)|v)}{f_{X(1)}(x(1)|v)}.
\]

From lemmas 1 and 2 it follows that

\[
F_{X(2)|X(1), V}(x(2)|x(1), v) = \begin{cases} 1 - \frac{S_1(x_1, x_2)}{S_1(x_1, x_1)} & \text{if } v = 0 \\ 1 - \frac{S_2(x_1, x_2)}{S_2(x_2, x_2)} & \text{if } v = 1. \end{cases} \tag{19}
\]

Similarly, the conditional survival function is

\[
S_{X(2)|X(1), V}(x(2)|x(1), v) = \frac{P[X_1 = x_1, X_2 > x_2|V = v]}{P[X_1 = x_1|V = v]} = \frac{S_{X(1)}(x(1), x(2))}{f_{X(1)}(x(1))}.
\]

From lemmas 1 and 3 it follows that

\[
S_{X(2)|X(1), V}(x(2)|x(1), v) = \begin{cases} S_1(x_1, x_2) & \text{if } v = 0 \\ S_2(x_1, x_2) & \text{if } v = 1. \end{cases} \tag{20}
\]

This completes the proof.

Appendix B

In this Appendix, we provide more details on the selected bivariate lifetime distributions: the GBE, MOBE, and MOBW models.

B.1 Gumbel’s Bivariate Exponential distribution

The Gumbel’s Bivariate Exponential (GBE) distribution is the most well-known model, which was first introduced by Gumbel (1960). The GBE model assumed a
failure mechanism driven by a random external stress factor. Gumbel (1960) provided two types of GBE models while Hougaard (1986) extended GBE type B model.

For parameter estimation of the GBE model parameters, one can derive the maximum likelihood estimators as: \( \hat{\theta}_1 = \bar{x}_1 = n^{-1} \sum_{i=1}^{n} x_{1i} \), \( \hat{\theta}_2 = \bar{x}_2 = n^{-1} \sum_{i=1}^{n} x_{2i} \), and \( \hat{\delta} = -(\log 2)^{-1} n^{-1} \sum_{i=1}^{n} \min\{x_{1i}/\bar{x}_1, x_{2i}/\bar{x}_2\} \). For more details on deriving these estimates, the reader is referred to Lu and Bhattacharyya (1991).

**B.2 Marshall Olkin Bivariate Exponential distribution**

The MOBE model was built to model the lifetime of a system with two-components which is affected by external shocks.

For parameter estimation of the MOBE model, one can derive the maximum likelihood estimations by solving the following maximum likelihood equations:

\[
\frac{n_1}{\hat{\lambda}_1} + \frac{n_2}{(\hat{\lambda}_1 + \hat{\lambda}_{12})} = \sum x_{1,i}, \quad \frac{n_1}{(\hat{\lambda}_2 + \hat{\lambda}_{12})} + \frac{n_2}{\hat{\lambda}_2} = \sum x_{2,i} \quad \text{and} \quad \frac{n_1}{(\hat{\lambda}_2 + \hat{\lambda}_{12})} + \frac{n_2}{\hat{\lambda}_2} + \frac{n_3}{\hat{\lambda}_{12}} = \sum \max(x_{1,i}, x_{2,i})
\]

where \( n_1, n_2, n_3 \) are the number of observations in the regions \( X_1 < X_2 \), \( X_1 > X_2 \), and \( X_1 = X_2 \), respectively. For more details on estimation of the MOBE model, the reader is referred to Bemis et al. (1972); Bhattacharyya and Johnson (1973); Proschan and Sullo (1976).

**B.3 Marshall Olkin Bivariate Weibull distribution**

In the MOBW model, random shocks affect the system and they are modelled as a non-homogeneous Poisson process. The MOBW model was developed by Marshall and Olkin (1967).

For parameter estimation of the MOBW model parameters, one can derive the maximum likelihood estimators by solving the following equations: \( \hat{\lambda}_1(\eta) = \frac{n_1}{\sum_{i=1}^{n_1} (r_i+1)\eta^i} \), \( \hat{\lambda}_2(\eta) = \frac{n_2}{\sum_{i=1}^{n_2} (r_i+1)\eta^i} \), \( \hat{\lambda}_{12}(\eta) = \frac{n_3}{\sum_{i=1}^{n_3} (r_i+1)\eta^i} \) and \( \hat{\eta} = h(\eta) \). For more details on deriving these estimates and more explanation of the symbols, the reader is referred to Feizjavadian and Hashemi (2015). For obtaining estimates by EM algorithm the reader is referred to Kundu and Dey (2009).

**Appendix C**

This section provides the proof of Theorem 2. For the proof, we need the following geometric series:

**Lemma 4.** The following holds when \(|r| < 1\)

\[
\sum_{t=k}^{\infty} \frac{t!}{(t-k)!} r^{t-k} = \frac{k!}{(1-r)^{k+1}} \quad \text{for} \quad k = 0, 1, 2, 3, \ldots \tag{21}
\]

Some mathematical manipulation of Equation (21) results in the following series,
which holds when \(|r| < 1\):

\[
\begin{align*}
\sum_{t=k}^{\infty} t \frac{(t-k+1)!}{(t-k)!} r^{t-k} &= \left(\frac{k-1}{2}\right)! \frac{k-1}{2} + \frac{(k+1)!}{(1-r)^{k+1}} + \frac{(k+1)!}{(1-r)^{k+1}} \\
&\quad \text{for } k = 1, 3, 5, 7, \ldots
\end{align*}
\]

\[
\sum_{t=k}^{\infty} i \frac{(t-k-2)!}{(t-k)!} r^{t-k} = \left(\frac{k}{2}\right)! \frac{k}{2} + \frac{(k+1)!}{(1-r)^{k+1}} \\
&\quad \text{for } k = 2, 4, 6, \ldots
\]

**Proof of Theorem 2.** We use the shorthand notation, = and \(\neq\), as defined in Equation (12). Equation (10) gives the expressions for \(E[TBE]\) so we only need to derive an expression for the ARL. By definition, \(ARL = \sum_{i=1}^{\infty} i \ast P_s[\text{Signal at event } i]\), the signal can be caused by either the first event or the second event. Hence we can split our ARL into:

\[
ARL = \sum_{i=1}^{\infty} i \times P_s[\text{No signal for all events upto } i - 1] P_s[S_1] \\
+ \sum_{i=2}^{\infty} i \times P_s[\text{No Signal for all events upto } i - 2] P_s[NS_1, S_2, \neq]
\]

(23)

There are two signal scenarios (\(S_1\) or \(S_2\)), first we focus on the probability of observing a signal at the first event time \(P_s[S_1]\) (Equation (23)). We need the probability that we did not observed a signal up to event \(i - 1\). There are two no-signal scenarios: either we do not observe a signal with probability \(P_s[NS_1, =]\) when \(X_1 = X_2\), or with probability \(P_s[NS_1, NS_2, \neq]\) when \(X_1 \neq X_2\). We include all possible combinations of these two to obtain \(i - 1\) events without a signal.

\[
\sum_{i=1}^{\infty} i \times P_s[\text{No signal upto } i - 1] P_s[S_1] = \sum_{i=1}^{\infty} i \times P_s[S_1] P_s[NS_1, =]^{i-1} \\
+ \sum_{i=3}^{\infty} i(i-2)P_s[S_1]P_s[NS_1, NS_2, \neq]P_s[NS_1, =]^{i-3} \\
+ \sum_{i=5}^{\infty} i \frac{(i-3)!}{2(i-5)!} P_s[S_1]P_s[NS_1, NS_2, \neq]^2P_s[NS_1, =]^{i-5} \\
+ \ldots...
\]

Next, we focus on a signal at the second event time \(P_s[NS_1, S_2, \neq]\) in Equation (23).
We include all combinations of signal and no-signal scenarios and we get

\[
\sum_{i=2}^{\infty} i \times P_x[\text{No signal up to } i-2] P_x[NS_1, S_2, X_1 \neq X_2] = \sum_{i=2}^{\infty} i \times P_x[NS_1, S_2, \neq] P_x[NS_1, =]^{i-2} \\
+ \sum_{i=4}^{\infty} i(i-3) P_x[NS_1, S_2, \neq] P_x[NS_1, NS_2, \neq] P_x[NS_1, =]^{i-4} \\
+ \sum_{i=6}^{\infty} i(i-4)! P_x[NS_1, S_2, \neq]^2 P_x[NS_1, NS_2, \neq]^2 P_x[NS_1, =]^{i-6} \\
+ \ldots.,
\]

Adding up these two components, using Lemma 4 the ARL is equal to

\[
ARL = P_x[S_1] \left( \frac{0}{1 - P_x[NS_1, =]} + \frac{1}{(1 - P_x[NS_1, =])^2} \right) \\
+ P_x[S_1] P_x[NS_1, NS_2, \neq] \left( \frac{1}{(1 - P_x[NS_1, =])^2} + \frac{2}{(1 - P_x[NS_1, =])^3} \right) \\
+ P_x[S_1] P_x[NS_1, NS_2, \neq] \left( \frac{2}{(1 - P_x[NS_1, =])^3} + \frac{3}{(1 - P_x[NS_1, =])^4} \right) \\
+ \ldots.,
\]

This can also be written as

\[
ARL = \sum_{i=0}^{\infty} P_x[S_1] P_x[NS_1, NS_2, \neq]^i \left( \frac{i}{(1 - P_x[NS_1, =])^{i+1}} + \frac{i + 1}{(1 - P_x[NS_1, =])^{i+2}} \right) \\
+ \sum_{i=0}^{\infty} P_x[NS_1, S_2, \neq] P_x[NS_1, NS_2, \neq]^i \left( \frac{i + 1}{(1 - P_x[NS_1, =])^{i+1}} + \frac{i + 1}{(1 - P_x[NS_1, =])^{i+2}} \right).
\]

By carrying out some mathematical manipulations and applying Equation (21) for \( k = 0, 1 \) (the traditional geometric series), we get the following expression for the
ARL:

\[ ARL = \frac{P_s[S_1]}{1 - P_s[NS_1,=]} \sum_{i=0}^{\infty} i \left( \frac{P_s[NS_1,NS_2,\neq]}{1 - P_s[NS_1,=]} \right)^i \]

\[ + \frac{P_s[S_1]}{(1 - P_s[NS_1,=])^2} \sum_{i=0}^{\infty} (i + 1) \left( \frac{P_s[NS_1,NS_2,\neq]}{1 - P_s[NS_1,=]} \right)^i \]

\[ + \frac{P_s[NS_1,S_2,\neq]}{(1 - P_s[NS_1,=])^2} \sum_{i=0}^{\infty} (i + 1) \left( \frac{P_s[NS_1,NS_2,\neq]}{1 - P_s[NS_1,=]} \right)^i \]

\[ = \frac{P_s[S_1]}{1 - P_s[NS_1,=]} \sum_{i=1}^{\infty} (i - 1) \left( \frac{P_s[NS_1,NS_2,\neq]}{1 - P_s[NS_1,=]} \right)^{i-1} \]

\[ + \frac{P_s[S_1]}{(1 - P_s[NS_1,=])^2} \sum_{i=1}^{\infty} i \left( \frac{P_s[NS_1,NS_2,\neq]}{1 - P_s[NS_1,=]} \right)^{i-1} \]

\[ + \left( \frac{P_s[NS_1,S_2,\neq]}{1 - P_s[NS_1,=]} + \frac{P_s[NS_1,S_2,\neq]}{(1 - P_s[NS_1,=])^2} \right) \sum_{i=1}^{\infty} (i - 1) \left( \frac{P_s[NS_1,NS_2,\neq]}{1 - P_s[NS_1,=]} \right)^{i-1} \]

\[ \frac{P_s[S_1]P_s[NS_1,NS_2,\neq] + P_s[S_1] + (2 - P_s[NS_1,=])P_s[NS_1,NS_2,\neq]}{(P_s[S_1] + P_s[NS_1,=])^2} \]

\[ = \frac{1 + P_s[NS_1,\neq]}{P_s[S_1] + P_s[NS_1,=]}. \]

Therefore,

\[ ATS = E_s[TBE] \frac{1 + P_s[NS_1,\neq]}{P_s[S_1] + P_s[NS_1,=]}. \]

This completes the proof.
Table 4. Comparison of ATS performance for our BTBE chart and the MEWMA chart

| In-control scenario | Shift type | $\theta_1$ | $\theta_2$ | $\delta$ | BTBE | MEWMA $r = 0.1$ | MEWMA $r = 1$ |
|---------------------|------------|------------|------------|---------|------|----------------|----------------|
| 1. Equal expectations, independence | IC | 5 | 5 | 1 | 199.7 | 200.1 | 200.2 |
| | OC-I1 | 7.5 | 5 | 1 | 143.1 | 104.0 | 106.8 |
| | OC-I1 | 10 | 5 | 1 | 104.8 | 72.7 | 76.5 |
| | OC-I2 | 7.5 | 7.5 | 1 | 110.0 | 81.2 | 79.8 |
| | OC-I2 | 10 | 10 | 1 | 72.6 | 59.5 | 56.9 |
| | OC-D1 | 5 | 2.5 | 1 | 172.5 | 71.7 | 251.2 |
| | OC-D1 | 5 | 1 | 1 | 99.9 | 35.9 | 211.1 |
| | OC-D2 | 2.5 | 2.5 | 1 | 99.5 | 34.6 | 2054.0 |
| | OC-D2 | 1 | 1 | 1 | 16.2 | 7.7 | >10000 |
| 2. Equal expectations, dependence | IC | 5 | 5 | 0.5 | 199.9 | 199.9 | 200.1 |
| | OC-I1 | 7.5 | 5 | 0.5 | 151.0 | 85.8 | 105.3 |
| | OC-I1 | 10 | 5 | 0.5 | 109.0 | 59.5 | 70.9 |
| | OC-I2 | 7.5 | 7.5 | 0.5 | 125.1 | 86.9 | 87.9 |
| | OC-I2 | 10 | 10 | 0.5 | 86.2 | 64.4 | 63.2 |
| | OC-D1 | 5 | 2.5 | 0.5 | 162.5 | 57.0 | 224.2 |
| | OC-D1 | 5 | 1 | 0.5 | 66.3 | 30.7 | 150.1 |
| | OC-D2 | 2.5 | 2.5 | 0.5 | 119.1 | 44.1 | 3820.1 |
| | OC-D2 | 1 | 1 | 0.5 | 24.7 | 9.1 | >10000 |
| 3. Unequal expectations, independence | IC | 5 | 15 | 1 | 197.9 | 199.9 | 200.8 |
| | OC-I1 | 7.5 | 15 | 1 | 136.8 | 126.1 | 112.0 |
| | OC-I1 | 10 | 15 | 1 | 99.9 | 90.3 | 79.2 |
| | OC-I2 | 7.5 | 22.5 | 1 | 134.9 | 128.4 | 108.4 |
| | OC-I2 | 10 | 30 | 1 | 101.6 | 102.8 | 87.6 |
| | OC-D1 | 5 | 10.5 | 1 | 168.8 | 129.9 | 238.1 |
| | OC-D1 | 5 | 7.5 | 1 | 128.5 | 72.4 | 225.1 |
| | OC-D2 | 3.5 | 7.5 | 1 | 118.6 | 64.3 | 718.0 |
| | OC-D2 | 2.5 | 7.5 | 1 | 98.8 | 50.6 | 2399.0 |
| 4. Unequal expectations, dependence | IC | 5 | 15 | 0.5 | 197.7 | 199.2 | 200.1 |
| | OC-I1 | 7.5 | 15 | 0.5 | 124.8 | 107.1 | 106.1 |
| | OC-I1 | 10 | 15 | 0.5 | 87.1 | 70.0 | 71.7 |
| | OC-I2 | 7.5 | 22.5 | 0.5 | 147.7 | 139.0 | 119.9 |
| | OC-I2 | 10 | 30 | 0.5 | 114.0 | 116.3 | 100.9 |
| | OC-D1 | 5 | 10.5 | 0.5 | 145.9 | 111.0 | 193.7 |
| | OC-D1 | 5 | 7.5 | 0.5 | 97.2 | 56.4 | 147.2 |
| | OC-D2 | 3.5 | 7.5 | 0.5 | 113.5 | 69.4 | 480.5 |
| | OC-D2 | 2.5 | 7.5 | 0.5 | 117.6 | 63.8 | 1586.1 |

Table 5. Descriptive statistics for AIDS dataset

| | Mean Adults | Mean Children | p-value |
|---------------------|--------------|--------------|---------|
| Infection time (in months) $X_{(1)}$ | 48.7 | 56.9 | 0.0127 |
| Total incubation time (in months) $X_{(2)}$ | 81.2 | 76.2 | 0.0514 |
| Transformed $X_{(1)}$ | 0.487 | 0.569 |
| Transformed $X_{(2)}$ | 0.812 | 0.762 |
Table 6. Characteristics of the GBE distribution

| pdf | \( f(x_1, x_2) = \left( \frac{\theta_1}{\theta_2^2} \right)^{1/\delta} \left( \frac{\theta_2}{\theta_1} \right)^{1/\delta - 1} C(x_1, x_2)^{\delta - 2} \left( C(x_1, x_2)^{\delta} + \frac{1}{\delta} - 1 \right) \exp(-C(x_1, x_2)^{\delta}) \) |
|-----|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| \( C(x_1, x_2) = \left( \frac{\theta_1}{\theta_2} \right)^{1/\delta} + \left( \frac{x_2}{\theta_2} \right)^{1/\delta} \) |
| Survival function | \( S(x_1, x_2) = \exp(-C(x_1, x_2)^{\delta}) \) |
| Expectations | \( E[X_1] = \theta_1 \) |
| \( E[X_2] = \theta_2 \) |
| \( E[X(1)] = C(1, 1)^{-\delta} \) |
| \( E[X(2)] = \theta_1 + \theta_2 - C(1, 1)^{-\delta} \) |
| \( E[TBE] = 0.5(\theta_1 + \theta_2 - C(1, 1)^{-\delta}) \) |
| Probabilities | \( P[X < Y] = \frac{\theta_1^{-1/\delta}}{C(1, 1)} \) |
| \( P[X > Y] = \frac{\theta_2^{-1/\delta}}{C(1, 1)} \) |

Table 7. Characteristics of the MOBE distribution

| pdf | \( f(x_1, x_2) = \begin{cases} \frac{\lambda_1}{\lambda_2 + \lambda_12} \exp(-\lambda_1 x_1 - (\lambda_2 + \lambda_12) x_2), & x_1 < x_2 \\ \frac{\lambda_2}{\lambda_1 + \lambda_12} \exp(-\lambda_1 + \lambda_12) x_1 - \lambda_2 x_2), & x_1 > x_2 \\ \lambda_1 \exp(-\lambda_1 x_1) \end{cases} \) |
|-----|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Survival function | \( S(x_1, x_2) = \exp(-\lambda_1 x_1 - \lambda_2 x_2 - \lambda_12 \max(x_1, x_2)) \) |
| Expectations | \( E[X_1] = \frac{\lambda_1}{\lambda_2 + \lambda_12} \) |
| \( E[X_2] = \frac{\lambda_2}{\lambda_1 + \lambda_12} \) |
| \( E[X(1)\mid X_1 = X_2] = \frac{\lambda_1 \lambda_2}{\lambda_1 + \lambda_12} \lambda_1 \) |
| \( E[X(2)\mid X_1 \neq X_2] = \frac{\lambda_1 \lambda_2}{\lambda_2 + \lambda_12} \lambda_1 \) |
| \( E[TBE] = 0.5(\frac{\lambda_2}{\lambda_2 + \lambda_12} + \frac{\lambda_1 \lambda_2}{\lambda_1 + \lambda_12} + \frac{\lambda_2}{\lambda_2 + \lambda_12}) \) |
| Probabilities | \( P[X < Y] = \frac{\lambda_1}{\lambda_2 + \lambda_12} \) |
| \( P[X > Y] = \frac{\lambda_2}{\lambda_1 + \lambda_12} \) |
| \( P[X = Y] = \frac{\lambda_1 \lambda_2}{\lambda_2 + \lambda_12} \) |

Table 8. Characteristics of the MOBD distribution

| pdf | \( f(x_1, x_2) = \begin{cases} \eta^2 \lambda_1 x_1^{\eta-1} x_2^{\eta-1} \exp(-\lambda_1 x_1 - (\lambda_2 + \lambda_12) x_2) x_1 < x_2 \\ \eta^2 \lambda_2 x_1^{\eta-1} x_2^{\eta-1} \exp(-\lambda_1 + \lambda_12) x_1 - \lambda_2 x_2), x_1 > x_2 \\ \eta^2 \lambda_12 x_1^{\eta-1} \exp(-\lambda_1 + \lambda_12) x_1) x_1 = x_2 \end{cases} \) |
|-----|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Survival function | \( S(x_1, x_2) = \exp(-\lambda_1 x_1^\eta - \lambda_2 x_2^\eta - \lambda_12 \max(x_1, x_2)^\eta), \ x_1, x_2 > 0 \) |
| Expectations | \( E[X_1] = \Gamma(1 + \frac{1}{\eta}) (\lambda_1 + \lambda_12)^{\eta}/\eta \) |
| \( E[X_2] = \Gamma(1 + \frac{1}{\eta}) (\lambda_2 + \lambda_12)^{\eta}/\eta \) |
| \( E[X(1)\mid X_1 = X_2] = \Gamma(1 + \frac{1}{\eta}) (\lambda_1 + \lambda_12)^{\eta}/\eta \) |
| \( E[X(2)\mid X_1 \neq X_2] = \Gamma(1 + \frac{1}{\eta}) (\lambda_2 + \lambda_12)^{\eta}/\eta \) |
| \( E[TBE] = 0.5\Gamma(1 + \frac{1}{\eta}) (\lambda_2 + \lambda_12)^{\eta}/\eta - \frac{\lambda_2 + \lambda_12}{\lambda_1 + \lambda_12} \) |
| Probabilities | \( P[X < Y] = \frac{\lambda_1}{\lambda_2 + \lambda_12} \) |
| \( P[X > Y] = \frac{\lambda_2}{\lambda_1 + \lambda_12} \) |
| \( P[X = Y] = \frac{\lambda_1 \lambda_2}{\lambda_2 + \lambda_12} \) |
| Scene | Shift | $E[X_1]$ | $E[X_2]$ | GBE | MOBE | MOBW |
|-------|-------|----------|----------|-----|------|------|
| 1. IC | 5 | 5 | 5 | 5 | 0.2 | 0.2 | 0 | 0.0314 | 0.0314 | 0 | 2 |
| OC-I | 7.5 | 5 | 7.5 | 5 | 1 | 0.133 | 0.2 | 0 | 0.0140 | 0.0314 | 0 | 2 |
| OC-I | 10 | 5 | 10 | 5 | 1 | 0.1 | 0.2 | 0 | 0.0079 | 0.0314 | 0 | 2 |
| OC-D | 7.5 | 7.5 | 7.5 | 1 | 0.133 | 0.133 | 0 | 0.0140 | 0.0140 | 0 | 2 |
| OC-D | 10 | 10 | 10 | 1 | 0.1 | 0.1 | 0 | 0.0079 | 0.0079 | 0 | 2 |
| OC-D | 5 | 2.5 | 5 | 2.5 | 1 | 0.2 | 0.2 | 0 | 0.0314 | 0.1257 | 0 | 2 |
| OC-D | 2.5 | 2.5 | 2.5 | 2.5 | 1 | 0.4 | 0.4 | 0 | 0.1257 | 0.1257 | 0 | 2 |
| 2. IC | 5 | 5 | 5 | 5 | 0.5 | 0.164 | 0.164 | 0.036 | 0.0257 | 0.0257 | 0.0057 | 2 |
| OC-I | 7.5 | 5 | 7.5 | 5 | 0.5 | 0.103 | 0.170 | 0.030 | 0.0098 | 0.0273 | 0.0041 | 2 |
| OC-I | 10 | 5 | 10 | 5 | 0.5 | 0.073 | 0.173 | 0.027 | 0.0043 | 0.0278 | 0.0036 | 2 |
| OC-D | 7.5 | 7.5 | 7.5 | 7.5 | 0.5 | 0.109 | 0.109 | 0.024 | 0.0114 | 0.0114 | 0.0025 | 2 |
| OC-D | 10 | 10 | 10 | 10 | 0.5 | 0.081 | 0.081 | 0.018 | 0.0064 | 0.0064 | 0.0014 | 2 |
| OC-D | 5 | 2.5 | 5 | 2.5 | 0.5 | 0.145 | 0.345 | 0.055 | 0.0171 | 0.1114 | 0.0143 | 2 |
| OC-D | 2.5 | 2.5 | 2.5 | 2.5 | 0.5 | 0.327 | 0.327 | 0.073 | 0.1028 | 0.1028 | 0.0228 | 2 |
| 3. IC | 5 | 5 | 5 | 5 | 0.5 | 0.164 | 0.164 | 0.036 | 0.0257 | 0.0257 | 0.0057 | 2 |
| OC-I | 7.5 | 15 | 7.5 | 15 | 1 | 0.133 | 0.067 | 0 | 0.0314 | 0.0035 | 0 | 2 |
| OC-I | 10 | 15 | 10 | 15 | 1 | 0.1 | 0.067 | 0 | 0.0079 | 0.0035 | 0 | 2 |
| OC-D | 7.5 | 22.5 | 7.5 | 22.5 | 1 | 0.133 | 0.044 | 0 | 0.0140 | 0.0016 | 0 | 2 |
| OC-D | 10 | 30 | 10 | 30 | 1 | 0.1 | 0.033 | 0 | 0.0079 | 0.0009 | 0 | 2 |
| OC-D | 5 | 10.5 | 5 | 10.5 | 1 | 0.2 | 0 | 0.0314 | 0.0071 | 0 | 2 |
| OC-D | 2.5 | 7.5 | 2.5 | 7.5 | 1 | 0.2 | 0 | 0.0641 | 0.0140 | 0 | 2 |
| OC-D | 2.5 | 7.5 | 2.5 | 7.5 | 1 | 0.4 | 0 | 0.1257 | 0.1040 | 0 | 2 |
| 4. IC | 5 | 5 | 5 | 5 | 0.5 | 0.176 | 0.042 | 0.024 | 0.0282 | 3.17e-04 | 0.0032 | 2 |
| OC-I | 7.5 | 15 | 7.5 | 15 | 1 | 0.133 | 0.167 | 0 | 0.0140 | 0.0035 | 0 | 2 |
| OC-I | 10 | 15 | 10 | 15 | 1 | 0.1 | 0.167 | 0 | 0.0079 | 0.0035 | 0 | 2 |
| OC-D | 7.5 | 22.5 | 7.5 | 22.5 | 1 | 0.133 | 0.044 | 0 | 0.0140 | 0.0016 | 0 | 2 |
| OC-D | 10 | 30 | 10 | 30 | 1 | 0.1 | 0.033 | 0 | 0.0079 | 0.0009 | 0 | 2 |
| OC-D | 5 | 10.5 | 5 | 10.5 | 1 | 0.2 | 0 | 0.0314 | 0.0071 | 0 | 2 |
| OC-D | 3.5 | 7.5 | 3.5 | 7.5 | 1 | 0.286 | 0.133 | 0 | 0.0641 | 0.0140 | 0 | 2 |
| OC-D | 2.5 | 7.5 | 2.5 | 7.5 | 1 | 0.4 | 0 | 0.1257 | 0.1040 | 0 | 2 |
Supplementary Material for “A Real Time Monitoring Approach for Bivariate Event Data”

In this supplementary material, we provide the mathematical derivations for the control limits displayed in Table 1. We derive the control limits for each of the three life-time distributions in Sections 1-3 and in addition, we provide the mathematical derivations of the ATS results provided in Corollary 3.1 - 3.3 in Sections 4-6.

1. Control limits for GBE distributed data

In this section, we derive the control limits for our control chart when the data can be modelled using a Gumbel’s Bivariate Exponential (GBE) distribution. The joint survival function \( S(x_1, x_2) \) is defined as,

\[
S(x_1, x_2) = \exp \left( -C(x_1, x_2)^\delta \right)
\]

where

\[
C(x_1, x_2) = \left( \frac{x_1}{\theta_1} \right)^{\frac{1}{\delta}} + \left( \frac{x_2}{\theta_2} \right)^{\frac{1}{\delta}}.
\]

For setting up the multivariate TBE chart, we will need the following:

\[
S_1(x_1, x_2) = \frac{\partial}{\partial x_1} S(x_1, x_2) = -\frac{1}{\theta_1} \left( \frac{x_1}{\theta_1} \right)^{\frac{1}{\delta}-1} C(x_1, x_2)^{\delta-1} S(x_1, x_2) \tag{1}
\]

\[
S_2(x_1, x_2) = \frac{\partial}{\partial x_2} S(x_1, x_2) = -\frac{1}{\theta_2} \left( \frac{x_2}{\theta_2} \right)^{\frac{1}{\delta}-1} C(x_1, x_2)^{\delta-1} S(x_1, x_2). \tag{2}
\]

To set the upper control limits for the first event time \((UCL_{(1)})\), we need to solve \(\alpha/2 = S(X_{(1)}|X_{(1)}, UCL_{(1)})\). Simple derivations yield \(UCL_{(1)} = -\ln(\alpha/2) C(1,1)^{-\delta}\).

To set the upper control limit for the second event time \((UCL_{(2)})\), we need to solve \(\alpha/2 = S_{X_{(2)}|X_{(1)}, V}(UCL_{(2)}|x_{(1)}, v)\) for \(UCL_{(2)}\). For \(V = 0\), we follow these
Since

\[ S_1(x_1, UCL(2)) = (\alpha/2) S_1(x_1, x_1) \]

\[ -\frac{1}{\theta_1} \left( \frac{x_1}{\theta_1} \right)^{1-\delta} \exp(-C(x_1, UCL) \delta) = (\alpha/2) S_1(x_1, x_1) \]

\[ C(x_1, UCL)^{\delta-1} \exp(-C(x_1, UCL) \delta) = -\theta_1 \left( \frac{x_1}{\theta_1} \right)^{1-\frac{1}{\delta}} (\alpha/2) S_1(x_1, x_1) \]

\[ C(x_1, UCL)^{1-\delta} \exp(C(x_1, UCL) \delta) = \left( -\theta_1 \left( \frac{x_1}{\theta_1} \right)^{1-\frac{1}{\delta}} (\alpha/2) S_1(x_1, x_1) \right)^{-1} \]

\[ \frac{\delta}{1 - \delta} C(x_1, UCL) \delta \exp \left( \frac{\delta}{1 - \delta} C(x_1, UCL) \delta \right) = \frac{\delta}{1 - \delta} \left( -\theta_1 \left( \frac{x_1}{\theta_1} \right)^{1-\frac{1}{\delta}} (\alpha/2) S_1(x_1, x_1) \right)^{-1} \]

We define \( H = \frac{\delta}{1 - \delta} C(x_1, UCL) \delta \) and \( G = \frac{\delta}{1 - \delta} \left( -\theta_1 \left( \frac{x_1}{\theta_1} \right)^{1-\frac{1}{\delta}} (\alpha/2) S_1(x_1, x_1) \right)^{-1} \), where some computations give the following form for \( G \):

\[ G = \frac{\delta}{1 - \delta} \exp \left( \frac{\alpha}{2} (1, 1) \delta \right) \left( -\theta_1 \left( \frac{x_1}{\theta_1} \right)^{1-\frac{1}{\delta}} \right) \]

Since \( H \) and \( G \) are real numbers and as \( G > 0 \), Definition 1 gives use the following solution:

\[ H = W_0(G) \]

\[ \left( \frac{x_1}{\theta_1} \right)^{\frac{1}{\delta}} + \left( \frac{UCL(2)}{\theta_2} \right)^{\frac{1}{\delta}} = \left( 1 - \frac{1}{\delta} W(G) \right)^{1/\delta} \]

\[ UCL = \theta_2 \left( \left( 1 - \frac{1}{\delta} W(G) \right)^{\frac{1}{\delta}} - \left( \frac{x_1}{\theta_1} \right)^{\frac{1}{\delta}} \right)^{\delta} \]

\[ UCL = \left( \theta_2 \left( 1 - \frac{1}{\delta} W(G) \right)^{\frac{1}{\delta}} - \left( \frac{\theta_2}{\theta_1} x_1 \right)^{\frac{1}{\delta}} \right)^{\delta}. \]

2. Control limits for MOBE distributed data

In this section, we derive the control limits when the data can be modeled by a MOBE distribution. The MOBE joint survival function \( S(x_1, x_2) \) is defined as

\[ S(x_1, x_2) = \exp(-\lambda_1 x_1 - \lambda_2 x_2 - \lambda_{12} \max(x_1, x_2)). \]

To setup the control limits, we need the following partial derivatives:

\[ S_1(x_1, x_2) = -\lambda_1 \exp(-\lambda_1 x_1 - \lambda_2 x_2 - \lambda_{12} x_2) \text{ when } 0 < x_1 < x_2 \]  \hspace{1cm} (3)

\[ S_2(x_1, x_2) = -\lambda_2 \exp(-\lambda_1 x_1 - \lambda_2 x_2 - \lambda_{12} x_2) \text{ when } x_1 > x_2 > 0 \]  \hspace{1cm} (4)
According to Theorem 1, we get the upper control limit for the first event time \((UCL_{(1)})\) by solving \(\alpha/2 = S(UCL_{(1)}, UCL_{(1)})\). Simple derivations yield \(UCL_{(1)} = - (\lambda_1 + \lambda_2 + \lambda_{12})^{-1} \ln(\alpha/2)\).

To set lower control limit for the first event time \((LCL_{(1)})\), we need to solve from Theorem 1. For \(X_1 < X_2\), \(\alpha/2 = \int_{0}^{LCL_{(1)}} \frac{S_1(x_1, x_1) dx_1}{S_1(x_1, x_1)}\), where

\[
\int_{0}^{LCL_{(1)}} S_1(x_1, x_1) dx_1 = \int_{0}^{LCL_{(1)}} \lambda_1 \exp\left(- (\lambda_1 + \lambda_2 + \lambda_{12}) x_1\right) dx_1 \\
= \frac{-\lambda_1}{\lambda_1 + \lambda_2 + \lambda_{12}} \left[ \exp\left(- (\lambda_1 + \lambda_2 + \lambda_{12}) LCL_{(1)}\right) - 1 \right]
\]

and

\[
\int_{0}^{\infty} S_1(x_1, x_1) dx_1 = \int_{0}^{\infty} \lambda_1 x_1 \exp\left(- (\lambda_1 + \lambda_2 + \lambda_{12}) x_1\right) dx_1 \\
= \frac{\lambda_1}{\lambda_1 + \lambda_2 + \lambda_{12}}.
\]

By substituting in \(\alpha/2\) above, \(LCL_{(1)} = - (\lambda_1 + \lambda_2 + \lambda_{12})^{-1} \ln(1 - \alpha/2)\). The lower control limit for \(X_1 > X_2\) follows equivalently.

To set the upper control limit for the second event time \((UCL_{(2)})\), we need to solve \(\alpha/2 = S_{X_2|X_1}(UCL_{(2)}|x_{(1)})\). For \(X_1 < X_2\), we solve \(\alpha/2 = \frac{S_1(x_1, UCL_{(2)})}{S_1(x_1, x_1)}\). Straightforward computations yield \(UCL_{(2)} = x_1 - (\lambda_2 + \lambda_{12})^{-1} \ln(\alpha/2)\). The \(UCL_{(2)}\) for the case \(X_1 > X_2\) follows equivalently.

When \(X_1 < X_2\), to set the lower control limit for the second event time \((LCL_{(2)})\), we solve \(\alpha/2 = F_{X_2|X_1}(LCL_{(2)}|x_{(1)})\). Which is \(\alpha/2 = 1 - \frac{S_1(x_1, LCL_{(2)})}{S_1(x_1, x_1)}\). Straightforward computations yield \(LCL_{(2)} = x_1 - (\lambda_2 + \lambda_{12})^{-1} \ln(1 - \alpha/2)\). The \(LCL_{(2)}\) for the case \(X_1 > X_2\) follows equivalently.

### 3. Control limits for MOBW distributed data

In this section, we derive the control limits for our control chart when the data can be modeled using a MOBW distribution. The MOBW joint survival function \(S(x_1, x_2)\) is defined as

\[
S(x_1, x_2) = \exp\left(-\lambda_1 x_1^\eta - \lambda_2 x_2^\eta - \lambda_{12} \max(x_1, x_2)^\eta\right).
\]

To derive the control limits, we need the following partial derivatives:

\[
S_1(x_1, x_2) = -\eta \lambda_1 x_1^{\eta-1} \exp\left(-\lambda_1 x_1^\eta - (\lambda_2 + \lambda_{12}) x_2^\eta\right) \text{ when } 0 < x_1 < x_2 \quad (5)
\]

\[
S_2(x_1, x_2) = -\eta \lambda_2 x_2^{\eta-1} \exp\left(- (\lambda_1 + \lambda_{12}) x_1^\eta - \lambda_2 x_2^\eta\right) \text{ when } x_1 > x_2 > 0 \quad (6)
\]

To set upper control limit for the first event time \((UCL_{(1)})\), we need to solve from Theorem 1, \(\alpha/2 = S(UCL_{(1)}, UCL_{(1)})\). Simple derivations yield \(UCL_{(1)} = - (\lambda_1 + \lambda_2 + \lambda_{12})^{-1} \ln(\alpha/2)^{1/\eta}\).
To set lower control limit for the first event time ($LCL(1)$), we need to solve from Theorem 1. For $X_1 < X_2$, $\alpha/2 = \int_{0}^{\infty} S_i(x_1,x_1)dx_1$, where

$$\int_{0}^{LCL(1)} S_1(x_1,x_1)dx = \int_{0}^{LCL(1)} \eta \lambda_1 x^{\eta-1} \exp \left( - (\lambda_1 + \lambda_2 + \lambda_12) x^{\eta} \right) dx.$$  

$$= \frac{-\lambda_1}{(\lambda_1 + \lambda_2 + \lambda_12)} \left[ \exp \left( - (\lambda_1 + \lambda_2 + \lambda_12) LCL(1)^{\eta} \right) - 1 \right]$$

and

$$\int_{0}^{\infty} S_1(x_1,x_1)dx = \int_{0}^{\infty} \eta \lambda_1 x^{\eta-1} \exp \left( - (\lambda_1 + \lambda_2 + \lambda_12) x^{\eta} \right) dx.$$  

$$= \frac{\lambda_1}{\lambda_1 + \lambda_2 + \lambda_12}.$$  

By substituting in $\alpha/2$ above, $LCL(1) = \left( - (\lambda_1 + \lambda_2 + \lambda_12)^{-1} \ln(1 - \alpha/2) \right)^{1/\eta}$. The lower control limit for $X_1 > X_2$ follows equivalently.

To set the upper control limit for the second event time ($UCL(2)$), we need to solve $\alpha/2 = S_{X(2)|X(1)}(UCL(2)|x(1))$. For $X_1 < X_2$, we solve $\alpha/2 = \frac{S_1(x_1,UCL(2))}{S_1(x_1,x_1)}$. Straightforward computations yield $UCL(2) = \left( x_1^{\eta} - (\lambda_2 + \lambda_12)^{-1} \ln(\alpha/2) \right)^{1/\eta}$. The $UCL(2)$ for the case $X_1 > X_2$ follows equivalently.

To set the lower control limit for the second event time ($LCL(2)$), we need to solve from Theorem 2, $\alpha/2 = F_{X(2)|X(1)}(LCL(2)|x(1))$. For $X_1 < X_2$ we solve $\alpha/2 = 1 - \frac{S_1(x_1,LCL(2))}{S_1(x_1,x_1)}$. Straightforward computations yield $LCL(2) = \left( x_1^{\eta} - (\lambda_2 + \lambda_12)^{-1} \ln(1 - \alpha/2) \right)^{1/\eta}$. The $LCL(2)$ for the case $X_1 > X_2$ follows equivalently.

### 4. Proof for Corollary 3.1

Here, we provide the derivation to obtain the results of Corollary 3.1, which gives an expression for the ATS of the GBE model. We begin by restating Equation 12 from the main paper:

$$ATS = E_*[TBE] \frac{2 - P_*[S_1]}{P_*[S_1] + P_*[NS_1,S_2]}.$$  

(7)

We then obtain expressions for each of the elements in the above ATS equation:

$$E_*[TBE] = 0.5(\theta_1^* + \theta_2^* - C(1,1)^{-\delta}).$$

Next $P_*[S_1]$, and $P_*[NS_1,S_2]$ are computed as
\[ P_\ast \{S_1 \} = P_\ast \{X_1 \leq LCL_{(1)}, X_1 < X_2 \} + P_\ast \{X_1 > UCL_{(1)}, X_1 < X_2 \} \\
+ P_\ast \{X_2 \leq LCL_{(1)}, X_1 > X_2 \} + P_\ast \{X_2 > UCL_{(1)}, X_1 > X_2 \} \\
= \int_0^{LCL_{(1)}} -S_1(x_1, x_1) dx_1 + \int_0^{LCL_{(1)}} -S_2(x_2, x_2) dx_2 + \int_{UCL_{(1)}}^\infty -S_1(x_1, x_1) dx_1 \\
+ \int_{UCL_{(1)}}^\infty -S_2(x_2, x_2) dx_2 \\
= \int_0^{LCL_{(1)}} \frac{1}{\theta_1^*} \exp(-C_{x_1, x_1}) dx_1 + \int_0^{LCL_{(1)}} \frac{1}{\theta_2^*} \exp(-C_{x_2, x_2}) dx_2 + \int_{UCL_{(1)}}^\infty \frac{1}{\theta_1^*} \exp(-C_{x_1, x_1}) dx_1 \\
+ \int_{UCL_{(1)}}^\infty \frac{1}{\theta_2^*} \exp(-C_{x_2, x_2}) dx_2 \\
= \int_0^{LCL_{(1)}} \frac{1}{\theta_1^*} \exp(-x_1 C_{11}^*) dx_1 + \int_0^{LCL_{(1)}} \frac{1}{\theta_2^*} \exp(-x_2 C_{11}^*) dx_2 + \int_{UCL_{(1)}}^\infty \frac{1}{\theta_1^*} \exp(-x_1 C_{11}^*) dx_1 \\
+ \int_{UCL_{(1)}}^\infty \frac{1}{\theta_2^*} \exp(-x_2 C_{11}^*) dx_2 \\
= -\frac{1}{\theta_1^* C_{11}^*} \left[ e^{-LCL_{(1)} C_{11}^*} - 1 \right] - \frac{1}{\theta_2^* C_{11}^*} \left[ e^{-LCL_{(1)} C_{11}^*} - 1 \right] + \frac{1}{\theta_1^* C_{11}^*} \left[ e^{-UCL_{(1)} C_{11}^*} - 1 \right] + \frac{1}{\theta_2^* C_{11}^*} \left[ e^{-UCL_{(1)} C_{11}^*} - 1 \right] \\
= -\frac{1}{\theta_1^* C_{11}^*} \left[ \exp \left( C_{11}^- \ln(1 - \alpha/2) C_{11}^* \right) - 1 \right] - \frac{1}{\theta_2^* C_{11}^*} \left[ \exp \left( C_{11}^- \ln(1 - \alpha/2) C_{11}^* \right) - 1 \right] \\
+ \frac{1}{\theta_1^* C_{11}^*} \left[ \exp \left( C_{11}^- \ln(\alpha/2) C_{11}^* \right) \right] + \frac{1}{\theta_2^* C_{11}^*} \left[ \exp \left( C_{11}^- \ln(\alpha/2) C_{11}^* \right) \right] \\
= \frac{1}{\theta_1^* C_{11}^*} \left[ 1 - (1 - \frac{\alpha}{2}) \frac{C_{11}^*}{C_{11}^-} \right] + \frac{1}{\theta_2^* C_{11}^*} \left[ 1 - (1 - \frac{\alpha}{2}) \frac{C_{11}^*}{C_{11}^-} \right] + \frac{1}{\theta_1^* C_{11}^*} \left[ \frac{\alpha}{2} \frac{C_{11}^*}{C_{11}^-} \right] + \frac{1}{\theta_2^* C_{11}^*} \left[ \frac{\alpha}{2} \frac{C_{11}^*}{C_{11}^-} \right] \\
= 1 - \left( 1 - \frac{\alpha}{2} \right) \frac{C_{11}^*}{C_{11}^-} + \left( \frac{\alpha}{2} \right) \frac{C_{11}^*}{C_{11}^-} \\
\]

where we used \( LCL_{(1)} \) and \( UCL_{(1)} \) as in Table 1 and, \( S_1 \) and \( S_2 \) from Equations 1 and 2.
\[ P_s[NS_1, S_2] = P_s[LCL_{(1)} < X_{(1)} \leq UCL_{(1)}], X_{(2)} \leq LCL_{(2)}, X_1 < X_2] \\
+ P_s[LCL_{(1)} < X_{(2)} \leq UCL_{(1)}, X_{(1)} \leq LCL_{(2)}, X_1 > X_2] \\
+ P_s[LCL_{(1)} < X_{(1)} \leq UCL_{(1)}, X_{(2)} > UCL_{(2)}, X_1 < X_2] \\
+ P_s[LCL_{(1)} < X_{(2)} \leq UCL_{(1)}, X_{(1)} > UCL_{(2)}, X_1 > X_2] \\
= \int_{LCL_{(1)}}^{UCL_{(1)}} \int_{LCL_{(2)}}^{UCL_{(2)}} S_1(x_1, x_2) dx_2 dx_1 + \int_{LCL_{(1)}}^{UCL_{(1)}} \int_{x_2}^{UCL_{(2)}} S_2(x_1, x_2) dx_1 dx_2 \\
\quad + \int_{LCL_{(1)}}^{UCL_{(1)}} \int_{UCL_{(2)}}^{\infty} S_1(x_1, x_2) dx_2 dx_1 + \int_{LCL_{(1)}}^{UCL_{(1)}} \int_{LCL_{(2)}}^{\infty} S_2(x_1, x_2) dx_1 dx_2 \\
= \int_{LCL_{(1)}}^{UCL_{(1)}} S_1(x_1, LCL_{(2)}) - S_1(x_1, x_1) dx_1 + \int_{LCL_{(1)}}^{UCL_{(1)}} S_2(LCL_{(2)}, x_2) - S_2(x_2, x_2) dx_2 \\
\quad + \int_{LCL_{(1)}}^{\infty} S_1(x_1, \infty) - S_1(x_1, UCL_{(2)}) dx_1 + \int_{LCL_{(1)}}^{\infty} S_2(\infty, x_2) - S_2(UCL_{(2)}, x_2) dx_2 \\
= \int_{LCL_{(1)}}^{UCL_{(1)}} \frac{-1}{\theta_1^*} \exp \left( \frac{-x_1}{\theta_1^*} - \frac{(x_1 - \theta_2 \ln(1 - \frac{\alpha}{2}))}{\theta_2^*} \right) + \frac{1}{\theta_1^*} \exp \left( -x_1 C_{11}^* \right) dx_1 \\
\quad + \int_{LCL_{(1)}}^{UCL_{(1)}} \frac{-1}{\theta_2^*} \exp \left( \frac{-(x_2 - \theta_1 \ln(1 - \frac{\alpha}{2}))}{\theta_1^*} - \frac{x_2}{\theta_2^*} \right) + \frac{1}{\theta_2^*} \exp \left( -x_2 C_{11}^* \right) dx_2 \\
\quad + \int_{LCL_{(1)}}^{UCL_{(1)}} \frac{1}{\theta_1^*} \exp \left( \frac{-x_1}{\theta_1^*} - \frac{(x_1 - \theta_2 \ln(\frac{\alpha}{2}))}{\theta_2^*} \right) dx_1 + \int_{LCL_{(1)}}^{UCL_{(1)}} \frac{1}{\theta_2^*} \exp \left( \frac{-(x_2 - \theta_1 \ln(\frac{\alpha}{2}))}{\theta_1^*} - \frac{x_2}{\theta_2^*} \right) dx_2 \\
= \left( \frac{1 - \frac{\alpha}{2}}{\theta_1^* C_{11}^*} \right)^{\frac{\theta_2^*}{\theta_1^*}} \left[ e^{-UCL_{(1)} C_{11}^*} - e^{-LCL_{(1)} C_{11}^*} \right] - \frac{1}{\theta_1^* C_{11}^*} \left[ e^{-UCL_{(1)} C_{11}^*} - e^{-LCL_{(1)} C_{11}^*} \right]
\quad + \left( \frac{1 - \frac{\alpha}{2}}{\theta_2^* C_{11}^*} \right)^{\frac{\theta_1^*}{\theta_2^*}} \left[ e^{-UCL_{(1)} C_{11}^*} - e^{-LCL_{(1)} C_{11}^*} \right] - \frac{1}{\theta_2^* C_{11}^*} \left[ e^{-UCL_{(1)} C_{11}^*} - e^{-LCL_{(1)} C_{11}^*} \right]
\quad - \frac{\theta_1^*}{\theta_1^* C_{11}^*} \left[ e^{-UCL_{(1)} C_{11}^*} - e^{-LCL_{(1)} C_{11}^*} \right] - \frac{\theta_2^*}{\theta_2^* C_{11}^*} \left[ e^{-UCL_{(1)} C_{11}^*} - e^{-LCL_{(1)} C_{11}^*} \right]
\quad \frac{\left( \frac{\alpha}{2} \right)^C - (1 - \frac{\alpha}{2})^C}{\theta_1^* C_{11}^*} \left( \left( \frac{\alpha}{2} \right)^C - (1 - \frac{\alpha}{2})^C \right) + \frac{\left( \frac{\alpha}{2} \right)^C - (1 - \frac{\alpha}{2})^C}{\theta_2^* C_{11}^*} \left( \left( \frac{\alpha}{2} \right)^C - (1 - \frac{\alpha}{2})^C \right)
\quad \frac{	heta_1^*}{\theta_1^* C_{11}^*} - \frac{\theta_2^*}{\theta_2^* C_{11}^*} - 1\right]

By using \( LCL_{(1)} \) and \( UCL_{(1)} \) as in Table 1 and writing \( C_{11}^*/C_{11} \) as \( C \), we complete the derivation of Corollary 3.1.

5. Proof for Corollary 3.2

Here, we provide the derivation to obtain the results of Corollary 3.2, which gives an expression for the ATS of the MOBE model. We obtain expressions for each of the elements in the ATS equation in Theorem 2:

\[ ATS = E_s[TBE] \frac{1 + P_s[NS_1, X_1 \neq X_2]}{P_s[S_1] + P_s[NS_1, S_2, X_1 \neq X_2]} \]
\[ E_s[TBE] = P_s[X_1 = X_2]E_s[X(1)|X_1 = X_2] + 0.5P_s[X_1 \neq X_2]E_s[X(2)|X_1 \neq X_2] \]
\[ = \frac{\lambda_1^2}{\Lambda^2} + 0.5 \left( \frac{\lambda_1^*}{\Lambda} + \frac{\lambda_2^*}{\Lambda} + \frac{\lambda_1^*}{\Lambda^*} \right) \]

Then, \( P_s[NS_1, X_1 \neq X_2] \), \( P_s[S_1] \), and \( P_s[NS_1, S_2, X_1 \neq X_2] \) are computed as:

\[ P_s[NS_1, X_1 \neq X_2] = P_s[LCL(1) < X_1 < UCL(1), X_1 < X_2] + P_s[LCL(1) < X_2 < UCL(1), X_1 > X_2] \]
\[ = \int_{LCL(1)}^{UCL(1)} -S_1(x_1, x_1)dx_1 + \int_{LCL(1)}^{UCL(1)} -S_2(x_2, x_2)dx_2 \]
\[ = \int_{LCL(1)}^{UCL(1)} \lambda_1^* \exp(-\Lambda^* x_1)dx_1 + \int_{LCL(1)}^{UCL(1)} \lambda_2^* \exp(-\Lambda^* x_2)dx_2 \]
\[ = \frac{\lambda_1^* + \lambda_2^*}{\Lambda^*} \left( \exp(-\Lambda^* LCL(1)) - \exp(-\Lambda^* UCL(1)) \right) \]
\[ = \frac{\lambda_1^* + \lambda_2^*}{\Lambda^*} \left[ (1 - \alpha/2)^{\frac{\Lambda^*}{\Lambda}} - (\alpha/2)^{\frac{\Lambda^*}{\Lambda}} \right] \]

where we used \( LCL(1) = -\Lambda^{-1} \ln(1 - \alpha/2) \) and \( UCL(1) = -\Lambda^{-1} \ln(\alpha/2) \) from Table 1, and \( S_1, S_2 \) from Equations 3 and 4.

\[ P_s[S_1] = P_s[X_1 \geq UCL(1), X_1 < X_2] + P_s[X_1 \leq LCL(1), X_1 < X_2] \]
\[ + P_s[X_2 \geq UCL(1), X_1 > X_2] + P_s[X_2 \leq LCL(1), X_1 > X_2] \]
\[ + P_s[X_1 \geq UCL(1), X_1 = X_2] + P_s[X_1 \leq LCL(1), X_1 = X_2] \]
\[ = \int_{0}^{\infty} -S_1(x_1, x_1)dx_1 + \int_{0}^{LCL(1)} -S_1(x_1, x_1)dx_1 \]
\[ + \int_{0}^{\infty} -S_2(x_2, x_2)dx_2 + \int_{0}^{LCL(1)} -S_2(x_2, x_2)dx_2 \]
\[ + \int_{0}^{\infty} \lambda_1^* \exp(-\Lambda^* x_1)dx_1 + \int_{0}^{LCL(1)} \lambda_1^* \exp(-\Lambda^* x_1)dx_1 \]
\[ = \exp(-\Lambda^* UCL(1)) + 1 - \exp(-\Lambda^* LCL(1)) \]
\[ = (\alpha/2)^{\frac{\Lambda^*}{\Lambda}} + 1 - (1 - \alpha/2)^{\frac{\Lambda^*}{\Lambda}}. \]
Also,

\[ P_\epsilon[N_S, S, X_1 \neq X_2] = P_\epsilon[LCL_{(1)} < X_1 < UCL_{(1)}, X_2 < UCL_{(2)} < X_2] + P_\epsilon[LCL_{(1)} < X_1 < UCL_{(1)}, X_2 < UCL_{(2)}] + P_\epsilon[LCL_{(1)} < X_2 < UCL_{(1)}, X_1 < X_2] + P_\epsilon[LCL_{(1)} < X_2 < UCL_{(1)}, X_1 > X_2] \]

\[ = \int_{LCL_{(1)}}^{UCL_{(1)}} -S_1(x_1, UCL_{(2)})dx_1 + \int_{LCL_{(1)}}^{UCL_{(1)}} -S_1(x_1, x_1) + S_1(x_1, LCL_{(2)})dx_1 \]

\[ + \int_{LCL_{(1)}}^{UCL_{(1)}} -S_2(UCL_{(2)}, x_2)dx_2 + \int_{LCL_{(1)}}^{UCL_{(1)}} -S_2(x_2, x_2) + S_2(LCL_{(2)}, x_2)dx_2 \]

\[ = \left[ -\left(\frac{\alpha}{2}\right)^{\frac{\lambda_1}{\Lambda^*}} + (1 - \frac{\alpha}{2})^{\frac{\lambda_1}{\Lambda^*}} \right] \left[ \frac{\lambda_1^*}{\Lambda^*} \left( 1 - (1 - \frac{\alpha}{2})^{\frac{\lambda_1^* + \lambda_1^*}{\lambda_1^* + \lambda_1^*}} + \left(\frac{\alpha}{2}\right)^{\frac{\lambda_1^* + \lambda_1^*}{\lambda_1^* + \lambda_1^*}} \right) + \frac{\lambda_2^*}{\Lambda^*} \left( 1 - (1 - \frac{\alpha}{2})^{\frac{\lambda_1^* + \lambda_1^*}{\lambda_1^* + \lambda_1^*}} + \left(\frac{\alpha}{2}\right)^{\frac{\lambda_1^* + \lambda_1^*}{\lambda_1^* + \lambda_1^*}} \right) \right] \]

This completes the derivation of Corollary 3.2.

6. Proof for Corollary 3.3

Here we provide the derivation to obtain the results of Corollary 3.3, which gives an expression for the ATS of the MOBW model. We obtain expressions for each of the elements in the ATS equation 8:

\[ E_\epsilon[TBE] = 0.5 \left( 1 + \frac{1}{\eta} \right) \left( \frac{1}{(\lambda_2^* + \lambda_1^*)/\eta} - \frac{\lambda_2^* + \lambda_1^*}{\lambda_1^* + \lambda_1^*/\eta} + \frac{1}{(\lambda_1^* + \lambda_1^*)/\eta} - \frac{\lambda_1^* + \lambda_1^*}{\lambda_1^* + \lambda_1^*/\eta} + 2 \frac{\lambda_1^*}{\lambda_1^* + \lambda_1^*/\eta} \right) \]

\[ P_\epsilon[N_S, X_1 \neq X_2] = \int_{LCL_{(1)}}^{UCL_{(1)}} \lambda_1^* \eta x_1^{\eta - 1} \exp(-\Lambda^* x_1^{\eta})dx_1 + \int_{LCL_{(1)}}^{UCL_{(1)}} \lambda_2^* \eta x_2^{\eta - 1} \exp(-\Lambda^* x_2^{\eta})dx_2 \]

\[ = \frac{\lambda_1^* + \lambda_2^*}{\Lambda^*} \left[ \exp(-\Lambda^* LCL_{(1)}) - \exp(-\Lambda^* UCL_{(1)}) \right] \]

\[ = \frac{\lambda_1^* + \lambda_2^*}{\Lambda^*} \left[ (1 - \frac{\alpha}{2})^{\frac{\lambda_1^*}{\Lambda^*}} - (\frac{\alpha}{2})^{\frac{\lambda_1^*}{\Lambda^*}} \right] \]

where \( LCL_{(1)} \) and \( UCL_{(1)} \) are as in Table 1. Then,

\[ P_\epsilon[S_1] = \int_{UCL_{(1)}}^{\infty} -S_1(x_1, x_1)dx_1 + \int_{0}^{LCL_{(1)}} -S_1(x_1, x_1)dx_1 \]

\[ + \int_{UCL_{(1)}}^{\infty} -S_2(x_2, x_2)dx_2 + \int_{0}^{LCL_{(1)}} -S_2(x_2, x_2)dx_2 \]

\[ + \int_{UCL_{(1)}}^{\infty} \eta \lambda_1 x_1^{\eta - 1} \exp(-\Lambda^* x_1^{\eta})dx_1 + \int_{0}^{LCL_{(1)}} \eta \lambda_1 x_1^{\eta - 1} \exp(-\Lambda^* x_1^{\eta})dx_1 \]

\[ = \exp(-\Lambda^* UCL_{(1)}) + 1 - \exp(-\Lambda^* LCL_{(1)}) \]

\[ = \left(\frac{\alpha}{2}\right)^{\frac{\lambda_1}{\Lambda^*}} + 1 - (1 - \frac{\alpha}{2})^{\frac{\lambda_1}{\Lambda^*}}. \]
Also,

\[ P_x[NS_1, S_2, X_1 \neq X_2] \]

\[ = \int_{LCL_{(1)}}^{UCL_{(1)}} -S_1(x_1, UCL_{(2)})dx_1 + \int_{LCL_{(1)}}^{UCL_{(1)}} -S_1(x_1, x_1) + S_1(x_1, LCL_{(2)})dx_1 \]

\[ + \int_{LCL_{(1)}}^{UCL_{(1)}} -S_2(UCL_{(2)}, x_2)dx_2 + \int_{LCL_{(1)}}^{UCL_{(1)}} -S_2(x_2, x_2) + S_2(LCL_{(2)}, x_2)dx_2 \]

\[ = \left[ -\left(\frac{\alpha}{2}\right)^\lambda + (1 - \frac{\alpha}{2})^\lambda \right] \left[ \frac{\lambda_1^s}{\Lambda^s} \left( 1 - \left(1 - \frac{\alpha}{2}\right)^{\lambda_1^{12} + \lambda_2^{12}} + \left(\frac{\alpha}{2}\right)^{\lambda_1^{12} + \lambda_2^{12}} \right) + \frac{\lambda_2^s}{\Lambda^s} \left( 1 - \left(1 - \frac{\alpha}{2}\right)^{\lambda_1^{12} + \lambda_2^{12}} + \left(\frac{\alpha}{2}\right)^{\lambda_1^{12} + \lambda_2^{12}} \right) \right] \]

This completes the derivation of Corollary 3.3.

References

Euler, L. (1779). De serie lambertina plurimis queeius insignibus proprietatibus (on the remarkable properties of a series of lambert and others). *Opera Omnia (Series 1)* 6, 350–369.

Lambert, J. H. (1758). Observationes variae in mathesin puram. *Acta Helvetica* 3(1), 128–168.