2021

Parametric, Nonparametric, and Semiparametric Linear Regression in Classical and Bayesian Statistical Quality Control

Chelsea L. Jones
Virginia Commonwealth University

Parametric, Nonparametric, and Semiparametric Linear Regression in Classical and Bayesian Statistical Quality Control

Chelsea L. Jones

Dissertation submitted to the Faculty of the
Virginia Commonwealth University
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Systems Modeling and Analysis

Co-Director: Dr. D'Arcy Mays, Associate Professor of Statistics
Statistical Sciences and Operations Research, Virginia Commonwealth University

Co-Director: Dr. Abdel-Salam Abdel-Salam, Associate Professor of Statistics
Mathematics, Statistics and Physics, Qatar University

May 2021
Richmond, Virginia

Keywords: statistical process control; profile monitoring; EWMA; CUSUM; mCUSUM; mEWMA; Hotelling's T^2; nonparametric; semiparametric; Bayesian; loss functions

Abstract

Parametric, Nonparametric, and Semiparametric Linear Regression in Classical and Bayesian Statistical Quality Control

By:
Chelsea L. Jones

Statistical process control (SPC) is used in many fields to understand and monitor desired processes, such as manufacturing, public health, and network traffic. SPC is categorized into two phases; in Phase I historical data is used to inform parameter estimates for a statistical model and Phase II implements this statistical model to monitor a live ongoing process. Within both phases, profile monitoring is a method to understand the functional relationship between response and explanatory variables by estimating and tracking its parameters. In profile monitoring, control charts are often used as graphical tools to visually observe process behaviors. We construct a practitioner's guide to provide a step-by-step application for parametric, nonparametric, and semiparametric methods in profile monitoring, creating an in-depth guideline for novice practitioners. We then consider the commonly used cumulative sum (CUSUM), multivariate CUSUM (mCUSUM), exponentially weighted moving average (EWMA), multivariate EWMA (mEWMA) charts under a Bayesian framework for monitoring respiratory disease related hospitalizations and global suicide rates with parametric, nonparametric, and semiparametric linear models.

Dedication

I dedicate this milestone to

- loved ones that I have lost to their battles with mental health and those still fighting.
- Black women who struggle to find their place in a world that doesn't provide a space for us to exist unapologetically.

Acknowledgments

Thank you to my co-advisors D'Arcy Mays and Abdel-Salam G. Abdel-Salam for being the guidance, support, and encouragement that I needed throughout this journey. Your advice, affirming words, and belief in me are a large reason I have made it to this point and you have made an everlasting impact on my life. I strive to be the kind of mentors you were to me for others.

Thank you to David Edwards and Yanjun Qian for joining me on this journey as my committee members. I am thankful for your support and advice, they have made me a better researcher and academic.

To my husband Jemel: thank you for being crazy enough to marry a PhD student in the middle of her program. It's been a roller coaster and I'm glad you've been by my side to ride out the lows, rejoice in the peaks, and to finally make it to the end.

To my family and friends: thank you for supporting and encouraging me when I needed it the most. I hope that I've made you proud.

To my Virginia State University family: Dr. Wynn thank you for your never-ending kindness and always on-time advice, Dr. Toni thank you for giving me the push I needed to start this journey, and Dr. A thank you for all the opportunities you've presented to me.

And thank you to the life-long friends I've made in this program.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Within Statistical process control (SPC), profile monitoring is an essential step for many applications. In this dissertation we will explore the profile monitoring aspect of SPC under 3 different distributional assumptions using a classical approach and a Bayesian approach.

Our objective is to lay the foundation for novice practitioners in Chapter 2 by constructing an in-depth review of key literature and providing guided examples. In Chapters 3 and 4 we examine the current research pertaining to Bayesian-based control charts under several loss functions. We compile and analyze simulation and real-data application results to endorse the Bayesian CUSUM and EWMA control charts in chapter 3 and 4 respectively, under different loss functions, priors and likelihood distributions. Chapter 5 extends the Bayesian control charts obtained in the previous chapters onto nonparametric and semiparametric multivariate regression models. Here, we intend to show the capabilities of the Bayesian control chart framework and generalize its use for data under any distributional assumption.

In Chapter 2, we complete a comprehensive review on current methodology in profile monitoring under different nonparametric and semiparametric regression techniques. The objective of this chapter is to create an instrument for practitioners to easily learn the key concepts of SPC and a guide on how to implement profile monitoring methods. An overview of essential in-

formation pertaining to SPC and profile monitoring is introduced along with general practices when engaging in the quality control. After a foundation is laid, commonly used univariate and multivariate control charts are defined. We use data from the European Network for Business and Industrial Statistics (ENBIS) to show how the multivariate CUSUM, multivariate mEWMA, and Hotelling's T^2 charts are computed and how they differ. Following the example of multivariate control charts, we focus in on the five different variations of the Hotelling's T^2 chart for a comparative analysis using linear data from Mahmoud et al., 2007 and Croarkin et al., 2006. We then define three nonparametric regression techniques and conduct a comparison. The comparison uses parametric, kernel, p-spline, and wavelet regression methods to estimate a linear model then charts them in an EWMA, CUSUM, and T^2 control chart respectively. The same analysis is done using the Model Robust Regression 1 (MRR1) semiparametric method (Abdel-Salam, 2009). The previously stated nonparametric techniques are used to inform the nonparametric aspect of the MRR1 method. In both the nonparametric and semiparametric examples, the mean squared error (MSE) and the probability of signal (PoS) is used to measure the capability of the regression method and the control chart respectively.

For chapters 3 and 4, we implement Bayesian methods under different loss functions to assess the Bayesian EWMA chart and Bayesian CUSUM chart capabilities. We compile and analyze simulation results to endorse the Bayesian EWMA and CUSUM control charts under different prior distributions, likelihood distributions, and loss functions. A sensitivity analysis of the hyper-parameters and sample size of each chart is conducted, using average run length (ARL), standard deviation of the run length (SDRL), average time to signal (ATS), and standard deviation of time to signal (SDTS) as measuring tools for performance. We support our simulation work with hospitalization count data.

Chapter 5 considers the framework of the Bayesian control charts recounted in chapters 3 and 4 and applies them to simulated data obtained using nonparametric and semiparametric

regression models. The objective of this chapter is to generalize the use of the Bayesian multivariate CUSUM (mCUSUM) and multivariate EWMA (mEWMA) control charts onto all data despite distributional assumptions. Our charting methods are assessed using the ARL, SDRL, ATS, and SDTS while a comparison of the regression methods used is done using mean squared error (MSE), Akaike information criterion (AIC), and Bayesian information criterion (BIC).

Chapter 2

Practitioners Guide on Parametric, Nonparametric, and Semiparametric Profile Monitoring

Abstract

Profile monitoring is one of the methods used in statistical process control (SPC) to understand the functional relationship between response and explanatory variables by tracking this relationship and estimating parameters. SPC is done in two phases: In Phase I a statistical model is created and its parameters estimated using historical data. Phase II implements the statistical model and monitors the live ongoing process. Control charts are graphical tools used to monitor these functional relationships over time in both Phase I and Phase II. This study provides a step-by-step application for parametric, nonparametric, and semiparametric methods in profile monitoring and creates an in-depth guideline with comparative analysis studies for novice practitioners. A comparative analysis under each distributional assumption is con-

ducted for various control charts.

Keywords: statistical process control; profile monitoring; EWMA; CUSUM; Hotelling's T^2; nonparametric; semiparametric

2.1. Introduction

Statistical Process Control (SPC) (or Statistical Process Monitoring (SPM)) is widely applied to monitor quality performance within industry settings to track and improve product quality. It is important to understand which components of the manufacturing process are important in capturing variability that may be between or within manufacturing processes. More can be seen in review papers Woodall and Montgomery, 2014 and Woodall, 2007. Once the product characteristics are determined and SPC methods are ready to be used, we move into the profile monitoring aspect.

2.1.1. Profiling Monitoring

Profile monitoring is one of the tools used in SPC to understand the functional relationship between a response variable and explanatory variables through observing/tracking this relationship and estimating parameters. A profile is a vector of measurement values associated with a single unit/product that when plotted over a range, can take the shape of a curve. The process to be monitored is estimated based on known or historical data, then the actual process is ran and its curve plotted. An example of what a profile can look like is given in figure 2.1. The example shows the revolutions per minute (RPM) versus the torque of a given engine (Abdel-Salam et al., 2013). Next we will look at control charts, which are a tool that helps track changes in these profiles. In this guide we consider and use control charts as an essential tool

within profile monitoring.

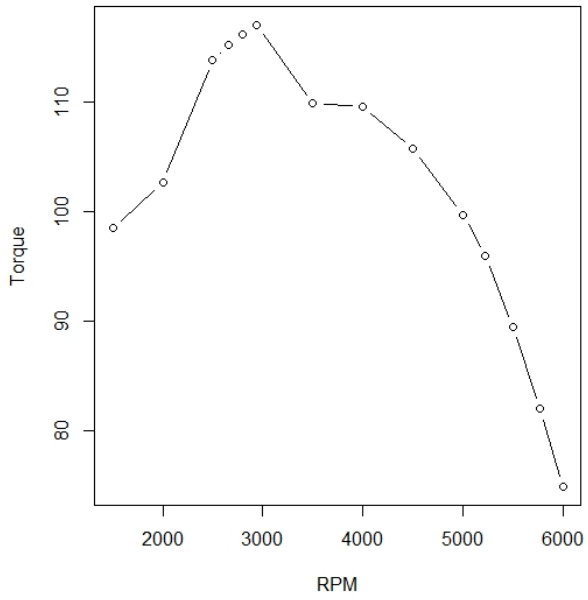


Figure 2.1: Profile Example

2.1.2. Control Charts

The overall purpose of *SPC* is to improve a desired process, and control charts are one tool utilized to do so. They are graphs of a quality characteristic that is measured or computed from a sample vs sample number (or time). If a process is in-control this implies that it is operating with only chance causes (an inherent part of the process) of variation being present. A process is out-of-control if it is operating in the presence of assignable causes (variability stemming from incorrect machine calibration, operator error, or defective raw materials). In subsequent sections we will explain how to identify an out-of-control system, define some of the most commonly used control charts, and highlight current research pertaining to the matter.

The general rule for determining if a process is out-of-control is if one or more of the cal-

culated statistics used to characterize the profile falls above the upper control limit (UCL) or below the lower control limit (LCL). However, this may not be the best approach for all profiles, thus sensitizing rules were established to make the process more versatile for capturing out-of-control conditions (Montgomery, 1990). The sensitivity criteria are useful in its application, but the tradeoff in using multiple of them simultaneously causes an increase in Type I error (rejecting a true hypothesis) probability.

2.1.3. Phase I

There are two phases that take place within process monitoring: Phase I and Phase II. Phase I is where preparation for the profile monitoring and gaining understanding of the process takes place. Within this phase, time-ordered samples are collected, observed, and used to assess tendencies of the process data. Once this is completed, the stability of the process is assessed, and an in-control model is chosen that best fits the results from the process sample. Step shifts and drifts can occur within Phase I while monitoring a profile and have a residual impact when estimating parameters in Phase II. Probability of signal (PoS) is a measurement tool used to assess the likelihood of a control chart signaling out of control. The PoS tracks the proportion of a chart signaling out-of-control within Phase I and is given in eq. 2.1, where q is the number of out-of-control profiles and m is the total number of profiles. Phase I is complete and moving to Phase II is ideal after the in-control model and parameters are chosen.

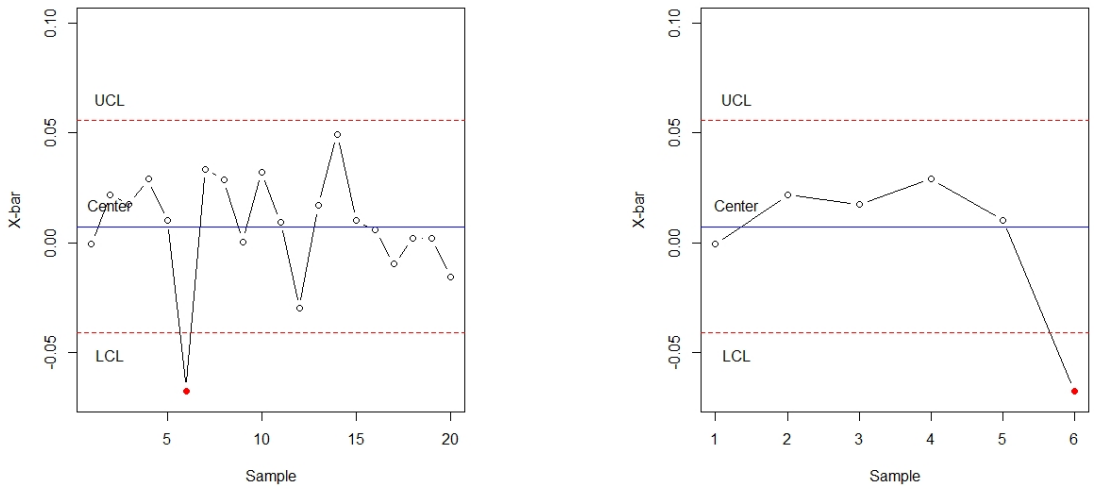
$$PoS = \frac{q}{m} \tag{2.1}$$

2.1.4. Phase II

After the model is chosen, the parameters of the model are estimated and the design parameters of the monitoring method can be determined. Phase *II* consists of live data collected through time and fit into a profile that is monitored and analyzed against the selected in-control model. The Average Run Length (*ARL*) is the average number of sample points that must be plotted before a point is indicated as out-of-control and given in eq. 2.2, where p is the probability that a single point exceeds in-control limits Montgomery, 2013.

$$ARL = \frac{1}{p} \tag{2.2}$$

A typical example of a Shewhart $\bar{x}$ univariate control chart is illustrated in figure 2.2 with the UCL and LCL represented as red lines. Figure 2.2a represents Phase *I* of a process, note the out-of-control point at observation 6. Since this has happened in Phase *I*, observation 6 should be removed and the control chart recalculated to monitor an in-control process. However, if this data were in Phase *II* then the chart would stop at observation 6 because it breaches one of the control limits, as in figure 2.2b.



(a) A basic control chart (Phase I)

(b) A basic control chart (Phase II)

Figure 2.2: Shewhart $\bar{x}$ Control Chart

2.2. Control Charts

2.2.1. Univariate Control Charts

The simplest case of profile monitoring is tracking only one characteristic over several different profiles. The Shewhart, Cumulative Sum (CUSUM) and Exponentially Weighted Moving Average (EWMA) control charts are often used for monitoring data with one variable, but the Hotelling's T^2 can also be used. Shewhart control charts are the simplest to construct since they solely use point estimates from the data. The CUSUM and EWMA control charts are known as memory-based charts since they calculate the current profile's statistic using information from the previous profiles, while the Shewhart and T^2 charts are memoryless.

2.2.1.1. Shewhart Control Chart

The Shewhart control chart was created by Walter Shewhart, introduced in Shewhart, 1926. It is the simplest control chart method as it monitors sample statistics directly. It is commonly used to monitor the mean ($\bar{x}$), range (R), or standard deviation (s) of a sample. Figure 2.2 under section 2.1.2 gives a simple simulated example of what the Shewhart $\bar{x}$ chart looks like under Phase I and Phase II analysis. For the control limits, μ_x is the population mean, σ_x is the population standard deviation, and k is some constant that represents a number of standard deviations away from the mean.

Algorithm 1 $\bar{x}$ chart algorithm

1: **for** $i \leq m$ **do** ▷ m is the number of profiles
2: Calculate $\bar{x} = \frac{\sum_{j=1}^{n} x_j}{n}$ ▷ n is the number of observations in m^{th} profile
Require: $\hat{\mu}_0 = \bar{\bar{x}}$ ▷ $\bar{\bar{x}}$ is the average of all profile means ($\hat{\mu}_0$ is the in-control mean)
3: **Calculate** $UCL = \mu_{\bar{x}} + k\sigma_{\bar{x}}$ ▷ k is typically 3
4: **Calculate** $LCL = \mu_{\bar{x}} - k\sigma_{\bar{x}}$
5: **Plot** profile numbers versus $\bar{x}$ statistics with the control limits

2.2.1.2. CUSUM Control Chart

The CUSUM control chart was initially introduced in Page, 1954 and Page, 1961. It is known for its ability to detect gradual changes in the data by keeping a memory of the previous samples, essentially increasing the sample size. This method uses the previous and the current distances from the target mean to determine the statistic (see eq. 2.3).

$$\begin{aligned} c_i &= \sum_{i=1}^{m} (\bar{x}_i - \hat{\mu}_0), \text{ for } i = 1, 2, \ldots, m = \text{ \# of profiles,} \\ &\text{where } \bar{x} = \frac{\sum_{j=1}^{n} x_j}{n}, j = 1, 2, \ldots, n = \text{ \# of observations in the current profile} \end{aligned} \tag{2.3}$$

Algorithm 2 CUSUM algorithm

Require: $\hat{\mu}_0 = \bar{\bar{x}} = \frac{\sum_{i=1}^{m} \bar{x}_i}{m}$ ▷ $\bar{\bar{x}}$ is the average of all profile means ($\hat{\mu}_0$ is the in-control mean)

1: **for** $i \leq m$ **do** ▷ let $c_0 = 0$, m is the number of profiles
2: $c_i = \bar{x}_i - \hat{\mu}_0$
3: $c_i = c_i + c_{i-1}$
4: **Calculate** control limit ($h = 4\sigma$) ▷ σ is the standard deviation
5: **Plot** profile numbers versus c_i statistics with the control limits

2.2.1.3. EWMA Control Chart

The EWMA chart was first introduced in Roberts, 1959 and is an alternative to the CUSUM. The idea for this chart is to assign a weight value, $(1-\lambda)$, to the previous statistic which controls the length of memory for the chart (see eq. 2.4). A λ value of 1 produces the Shewhart control chart and a λ value closest to 0 returns a control chart with a longer memory.

$$z_i = \lambda \bar{x}_i + (1-\lambda) z_{i-1}, \text{ for } i = 1, 2, \ldots, m = \text{ \# of profiles,} \quad \text{where } z_0 = \hat{\mu}_0 \tag{2.4}$$

Algorithm 3 EWMA algorithm

Require: $\hat{\mu}_0 = \bar{\bar{x}} = \frac{\sum_{i=1}^{m} \bar{x}_i}{m}$ ▷ $\bar{\bar{x}}$ is the average of all profile means ($\hat{\mu}_0$ is the in-control mean)

1: **Choose** $\lambda \in (0, 1]$ ▷ Recommended as $\lambda \in [0.05, 0.25]$ Montgomery2012
2: **for** $i \leq m$ **do** ▷ m is the number of profiles
3: $\bar{x}_i = \frac{\sum_{j=1}^{n} x_j}{n}$ ▷ n is the number of observations in the m^{th} profile
4: $z_i = \lambda \bar{x}_i + (1-\lambda) z_{i-1}$
5: **Calculate** $UCL = \hat{\mu}_0 + L\sigma\sqrt{\frac{\lambda}{2-\lambda}(1-(1-\lambda)^{2i})}$ ▷ $L = 3$ and σ is the standard deviation
6:
7: **Calculate** $LCL = \hat{\mu}_0 - L\sigma\sqrt{\frac{\lambda}{2-\lambda}(1-(1-\lambda)^{2i})}$
8:
9: **Plot** profile numbers versus z_i statistics with the control limits

2.2.2. Multivariate Control Charts

For data that has multiple characteristics to be monitored, it's best to use multivariate control charts rather than using multiple univariate control charts. These characteristics are often correlated and conducting univariate tests for each of the characteristics fails to capture changes in the correlations. Using multiple univariate charts can also lead to inflated Type I error, especially with a large number of variables. In this section different multivariate control charts will be introduced, followed by an application with comparison in the successive section.

2.2.2.1. Hotelling's T^2

Hotelling's T^2 method is briefly mentioned in this section as a multivariate control chart, however the topic is visited in depth in section 2.3.1. The T^2 statistic is calculated as the squared statistical distance, it measures the likelihood of obtaining the observation from a given population (Hotelling et al., 1951). This method is effective in detecting large shifts in the data. The general form of the T^2 statistic is given in eq. 2.5, where $\boldsymbol{\mu_0}$ is the in-control mean and $\boldsymbol{\Sigma_0}$ is the variacne-covariance matrix.

$$T^2 = (\boldsymbol{x_i} - \boldsymbol{\mu_0})' \boldsymbol{\Sigma_0^{-1}} (\boldsymbol{x_i} - \boldsymbol{\mu_0}) \tag{2.5}$$

2.2.2.2. mCUSUM

The multivariate Cumulative Sum (mCUSUM) is an extension of the univariate CUSUM. The concept for the mCUSUM is the same as the univariate case, it sums the previous statistic calculations together to obtain the statistic at the current location (Woodall and Ncube, 1985). The equation begins with setting $\mu_1 = \mu_0 + 1$ (μ_0 is the in-control mean and μ_1 is the out-of-

control mean) and obtaining the variance-covariance matrix, $\mathbf{\Sigma_0}$. Equation 5.4 shows how to obtain the mCUSUM statistic along with necessary variable definitions. Figure 2.3 shows the implementation of this mCUSUM equation in R.

$$MC_i = max\{0, (\boldsymbol{C_i^T \Sigma_0^{-1} C_i})^{\frac{1}{2}} - kn_i\},$$

$$\text{where } \boldsymbol{C_i} = \sum_{j=i-n_i+1}^{i} (\boldsymbol{x_j} - \boldsymbol{\mu_0}), \text{such that } n_i = \begin{cases} n_{i-1} + 1 \text{ if } MC_i > 0, \\ 1, otherwise \end{cases} \tag{2.6}$$

$$\text{and } k = \frac{1}{2}\sqrt{(\mu_1 - \mu_0)^T \mathbf{\Sigma_0} (\mu_1 - \mu_0)} \text{ for } i = 1, 2, \ldots, m = \text{number of profiles}$$

```
for(i in 1:m2){
  if(i == 1){
    n_i[i] <- 1
  }else{
    if(mc_i[i] > 0){
      n_i[i] <- n_i[i-1] + 1
    }else{
      n_i[i] <- 1
    }
  }
  j <- i - n_i[i] + 1
  c_i[, i] <- sum( diff1[, j:i], 2 )
  mc_i[i] <- max(0, sqrt( t(c_i[, i])%*%Ginv(s2)%*%c_i[, i]) - ( k1 * n_i[i] ) )
}
plot(1:m2, mc_i, type = 'b', xlab = 'Sample', ylab = 'Statistic')
abline(h = cUCL, col = 'red')
```

Figure 2.3: R Code for mCUSUM Control Chart

2.2.2.3. mEWMA

Multivariate Exponentially Weighted Moving Average (mEWMA) is an alternative control chart with memory to the mCUSUM . The method relies on a variable r which controls the length of memory for the chart. A r value of 1 produces the T^2 method which is memoryless, and a smaller r value returns a control chart with a longer memory. Another variable, γ, in-

dicates the magnitude of the out-of-control mean shift. A UCL and r value can be found after defining the γ and p values using techniques found in Lowry et al., 1992. The mEWMA equation is given in eq. 5.5 and figure 2.4 shows the implementation of the mEWMA equation in R.

$$\begin{gathered} \boldsymbol{z_i} = r \cdot \boldsymbol{x_i} + (1-r)\boldsymbol{z_{i-1}}, \text{ for } i = 1,2,\ldots,m = \text{number of profiles}, \\ \text{such that } T_i^2 = (\boldsymbol{z_i} - \boldsymbol{\mu_0})^T \Sigma_{z_i}^{-1} (\boldsymbol{z_i} - \boldsymbol{\mu_0}) \\ \text{where } \boldsymbol{z_0} = \mu_0 \text{ and } \boldsymbol{\Sigma_{z_i}} = cov(z_i) = \frac{r\left[1-(1-r)^{2i}\right]}{2-r}\boldsymbol{\Sigma_0} \end{gathered} \tag{2.7}$$

```
for(i in 1:m2){
  if(i == 1){
    z_i[, i] <- ( r1 * data3[, i] ) + ( (1-r1)*mu0 )
  }else{
    z_i[, i] <- ( r1*data3[, i] ) + ( ( 1-r1 ) * z_i[, (i-1)] )
  }
  invSig <- Ginv(s2) * ( (2-r1) / (r1 * ( 1 - ( ( 1- r1)^(2*i) ) ) ) )
  t2_i[i] <- t( z_i[, i] - mu0 )%*%invSig%*%( z_i[, i] - mu0 )
}
plot(1:m2, t2_i, type = 'b', xlab = 'Sample', ylab = 'Statistic')
abline(h = h4, col = 'red')
```

Figure 2.4: R Code for mEWMA Control Chart

2.2.3. Multivariate Control Charts Example

The data used in this section was obtained from the European Network for Business and Industrial Statistics (ENBIS) 2014 challenge. The challenge implored participants to use data describing consumer confidence to understand when the 2008 financial crisis began. The data is comprised of 97 months (May 2004 to May 2012) with 4 index category responses within each month. These indices are National Consumer Confidence Index, Present Situation Index, Expectation Index, and Spending Index. We consider each month-year combination to be a profile and the 4 indices to be observations within the profile (i.e January 2009 is a profile and it con-

tains a measure for each index). The data and information about the ENBIS 2014 challenge can be found at https://www.enbis.org/news/383. Figure 2.5 gives a visual of the steps taken when operating between Phase I and Phase II for this data along with the list below. In general, the entirety of the historical data that is provided should be used for Phase I analysis, but since our example does not have a live process to monitor for Phase II we partitioned the data to validate our model. We use the T^2 chart to determine if the data is in-control, but the practitioner can use the chart that best suits their process.

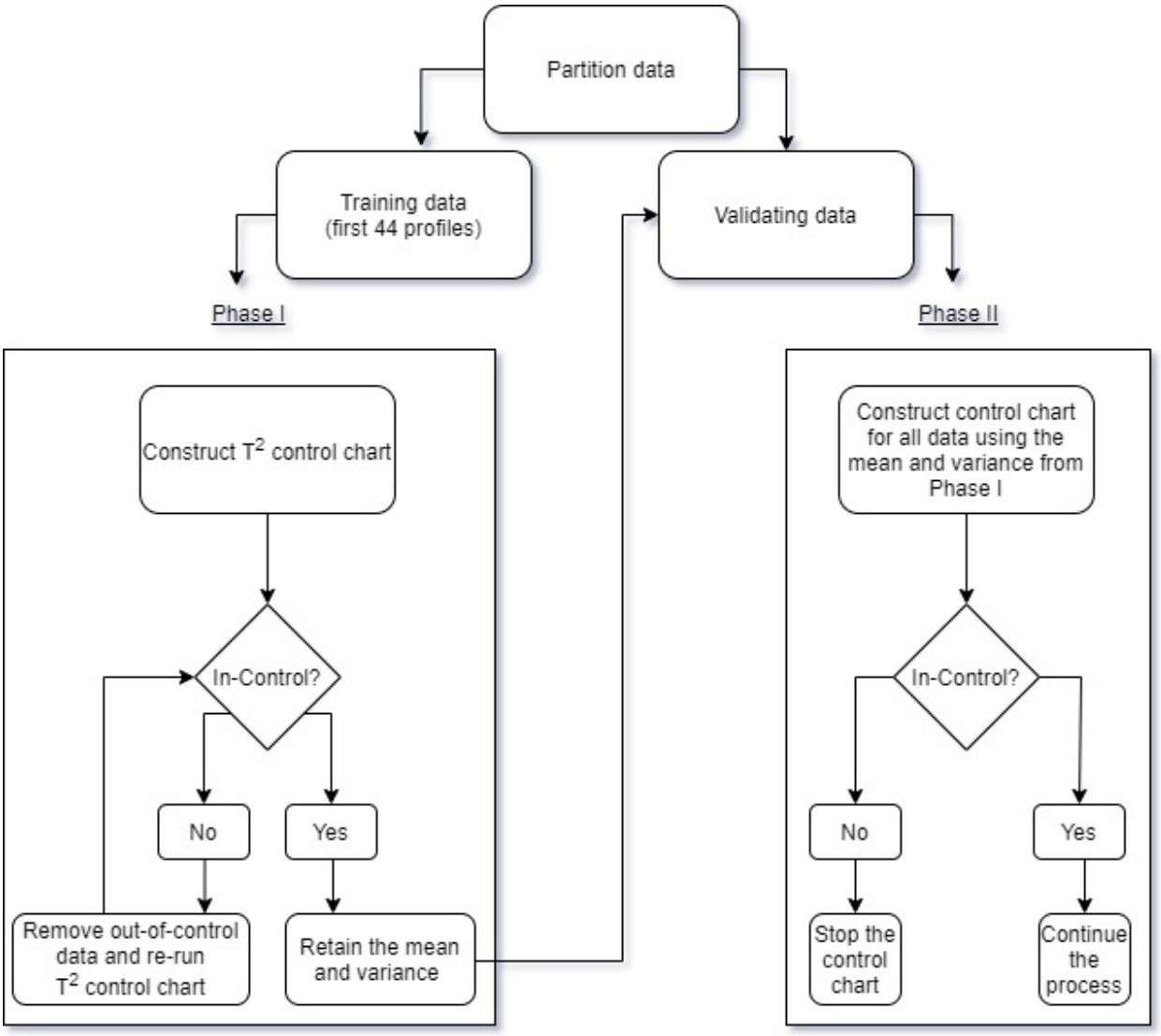


Figure 2.5: Phase I & Phase II Process Flowchart

(1) Partition data into 2 sets: training data and data to monitor. You can use prior knowledge to do this, we took the first 44 profiles (May 2004-December 2007) as our training data since the economic crisis began in 2008.

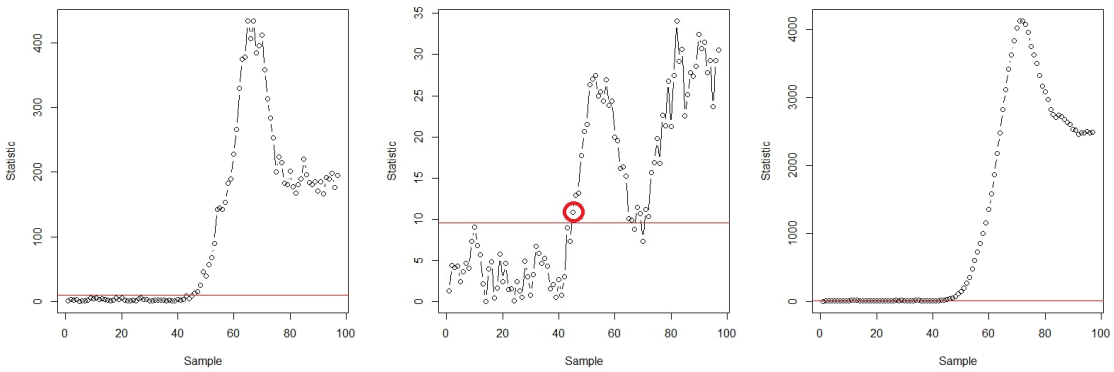
(2) Set control limits for your chart of choice. For the T^2 chart our control limit was $\chi^2_{\alpha,p} = 9.488$, where $\alpha = 0.05$ and $p = 4$.

(3) Determine if there are any out-of-control points in the training data, if there are then remove them and re-run the control chart. In this data there were no outlying points, so our training data was in-control.

(4) Use the $\bar{x}$ and s^2 from Phase I to calculate the T^2 statistics for the remaining 52 profiles.

(5) Run the control chart for all of the data (the training data + the data to monitor). Use this chart to conclude if/when the data is out of the acceptable bounds. The ENBIS data breeches the control limit at sample 46 (figure 2.6a) and suggests that the economic crisis began February 2008.

To enhance sensitivity to small change detection, we performed the mCUSUM and mEWMA methods. The point estimates found in Phase I using the T^2 method are also used here to calculate the statistics for both methods. The UCL_C for the mCUSUM method was found via interpolation as $UCL_C = 0.5857p + 7.2371$, where $p = 4 \implies UCL_C = 9.5779$ using simulation data found in Pignatiello Jr and Runger, 1990. For the mEWMA chart, we let $\lambda = 1$ and $p = 4$ for the UCL (h_4) and r ($UCL_E = 13.34$ and $r = 0.14$) from techniques found in Lowry et al., 1992.



(a) Hotelling's T^2 (b) $mCUSUM$ (c) $mEWMA$

Figure 2.6: Multivariate Control Charts Example

The mCUSUM chart signaled at January 2008 (sample 45) (figure 2.6b) while the mEWMA

chart signaled at March 2005 (sample 11) and continued to fluctuate until November 2007 (sample 43) (figure 2.6c), then it remained out-of-control. Since the mCUSUM and mEWMA charts detect small changes that have accumulated over time it is logical that they signaled before the T^2. Hotelling's T^2 chart is best at detecting large sudden changes and led to the later signaling in this instance. We recommend the use of the mCUSUM chart when immediate response information is needed but recommend the Hotelling's T^2 chart when overall detection is needed as it is less computationally intense. The early signaling of the mEWMA control chart results from the small r value, which causes the chart to have a longer memory. For applications where less sensitivity of the control charts is necessary (i.e. manufacturing), the mEWMA control chart is not ideal compared to the T^2 and mCUSUM charts. Using the mEWMA chart for the ENBIS data resulted in early signaling which can be considered a false alarm, in real-time applications this can have detrimental repercussions.

2.3. Parametric Methods for Profile Monitoring in Regression

An impressive amount of research has been done with profiles represented as a simple linear regression model ($y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$, $i = 1, 2, \cdots, n$). Literature for simple regression profile monitoring often assumes that the in-control model parameters are known and often considers Phase II applications. This section focuses on profile monitoring methods and estimating parameters using parametric techniques in Phase I.

2.3.1. Hotelling's T^2

The Hotelling's T^2 control chart is a well-known charting method characterized by a single UCL, often using the 95^{th} percentile, that is the χ^2 statistic or F statistic (see eq. 2.8) for uni-

variate or multivariate data, respectively. There are five different methods used to estimate the statistic: sample average and intra-profiling (T_I^2), sample average and covariance matrix (T_H^2), sample average and moving ranges (T_R^2), Minimum Volume Ellipsoid (T_{MVE}^2), and Minimum Covariance Determinant (T_{MCD}^2). The last two were initially proposed in Rousseeuw, 1984 and further detailed in Vargas, 2003. All methods have the same form for the statistic calculation which is defined by eq. 2.5, but differ in how the variance-covariance matrix is calculated (table 2.1). T_I^2, T_H^2, and T_R^2 are known to deteriorate in performance due to the masking effect (loss in power), thus inspiring researchers to create both the T_{MVE}^2 and T_{MCD}^2 methods to incorporate estimators robust to outliers. Figure 2.7 shows the implementation of the general Hotelling's T^2 equation in R. The s2 variable seen in the figure is defined based on which T^2 method is being used.

Method	Variance-Covariance Matrix
T_I^2	$S_I = \frac{1}{m}\sum_{i=1}^{m} var(\hat{\boldsymbol{\beta}}_{\boldsymbol{i}}) = (X'W_iX)^{-1}$
T_H^2	$S_H = \frac{1}{m-1}\sum_{i=1}^{m}(\hat{\boldsymbol{\beta}}_{\boldsymbol{i}} - \bar{\boldsymbol{\beta}})(\hat{\boldsymbol{\beta}}_{\boldsymbol{i}} - \bar{\boldsymbol{\beta}})'$
T_R^2	$S_R = \frac{1}{2(m-1)}\sum_{i=1}^{m-1}(\hat{\boldsymbol{\beta}}_{\boldsymbol{i+1}} - \hat{\boldsymbol{\beta}}_{\boldsymbol{i}})(\hat{\boldsymbol{\beta}}_{\boldsymbol{i+1}} - \hat{\boldsymbol{\beta}}_{\boldsymbol{i}})'$
T_{MVE}^2	see section 2.3.1.1
T_{MCD}^2	see section 2.3.1.2

Table 2.1: Variance-Covariance Matrix for Hotelling's T^2 Methods

$$\frac{(m-1)(n-1)}{m(n-1)+1-p}F_{1-\alpha}(p, m(n-1)+1-p), \text{ where}$$
$$m = \#\text{ of profiles},$$
$$n = \#\text{ of observations},$$
$$p = \#\text{ of columns in the x matrix} \tag{2.8}$$

```
for(j in 1:m2){
  t2 <- t( data2[j,] - xbar )%*%Ginv(s2)%*%( data2[j,] - xbar )
  t2all <- rbind(t2all, t2)
}
plot(1:m2, t2all, type = 'b', xlab = "Sample", ylab = "Statistic")
abline(h = ucl, col='red')
```

Figure 2.7: R Code for T^2 Control Chart

2.3.1.1. T^2_{MVE} Algorithm

The Minimum Volume Ellipsoid algorithm iteratively constructs ellipsoids from bootstrapped sub-samples taken from the original dataset, X_N. The volumes for these ellipsoids are calculated, and the mean and covariance matrix corresponding to the subsample with the minimal volume is used to calculate the T^2 statistic. By minimizing the volume of the ellipsoids (calculated by Mahalnobis distance), T^2_{MVE} is robust to outliers. In the algorithm we define p as the number of parameters in X_N (also can be considered the number of columns in a dataset X_N, including the intercept) and m as the number of samples. It is important to note that in this research we consider a sample to be a single profile and the terms will be interchangeable. The `MVE` call in `SAS` or `CovMve` function in `R` can be used to run the algorithm. More can be seen on the algorithm and its usage in Rousseeuw, 1984, Vargas, 2003, Rousseeuw and Van Zomeren, 1990, Van Aelst and Rousseeuw, 2009, and Woodruff and Rocke, 1993.

Algorithm 4 T^2_{MVE} algorithm

Require: $1-(1-(1-\epsilon)^{p+1})^n \geq p_0$ ▷ where $\epsilon = 0.5$ and p_0 is near unity (i.e $p_0 = 0.95$)

for $i \leq n$ **do**

Sample (p+1) observations from a dataset X_N to form a subset,

$$\text{J} = \{i_1, i_2, \cdots i_{p+1}\} \subset \{1, 2, \cdots N\}$$

Compute the sample mean and sample covariance matrix for the subset using:

$$\bar{x}_J = \frac{1}{p+1}\sum_{k=1}^{p+1} x_k \text{ and } S_J = \frac{1}{p}\sum_{k=1}^{p+1}(x_k - \bar{x}_J)(x_k - \bar{x}_J)^T$$

Calculate the distance for each profile using the sub-sample mean and covariance matrix:

$$d^2_J(i) = (x_i - \bar{x}_J)S_J^{-1}(x_i - \bar{x}_J)$$

Set the h^{th} order statistic of $d^2_J = D^2_J$, define h = [a] and $a = \frac{1}{2}(m+p+1)$ where h represents the integer part of a

Calculate the volume of the corresponding ellipsoid using:

$$V_J = (D^2_J)^p det(S_J)$$

Take the minimal V_J over all sub-samples and set $J^* = J$, such that J is the sub-sample index that $min(V_J)$

Set MVE estimators as: $\bar{x}_{MVE} = \bar{x}_{J^*}$ and $S_{MVE} = c^2_{m,p}(\chi^2_{p,0.5})^{-1} D^2_{J^*} S_{J^*}$

where $\chi^2_{p,0.5}$ is the median of a chi-squared distribution with p degrees of freedom and $c^2_{m,p}$ is a correction factor for a small sample size, m, given by Vargas, 2003 as: $c^2_{m,p} = \left(1 + \frac{15}{m-p}\right)$

2.3.1.2. T^2_{MCD} Algorithm

The Minimum Covariance Determinant method finds a subset of the data with the smallest variance-covariance determinant, using its mean and variance to calculate T^2_{MCD}. Minimizing the determinant of the variance-covarince matrix makes this method robust to location and scatter. It is asymptotically normal, making its statistical efficiency better than the MVE method Butler et al., 1993, but in Davies et al., 1992 it is noted that the latter has a lower convergence rate. T^2_{MCD} exposes outliers better as it produces more precise robust distances over the T^2_{MVE}.

However, T^2_{MCD} is more difficult to compute, so the MVE method is used more often. A modified MCD algorithm is introduced in Rousseeuw and Driessen, 1999 (also see Hardin and Rocke, 2004) to improve computational efficiency.

Algorithm 5 T^2_{MCD} algorithm

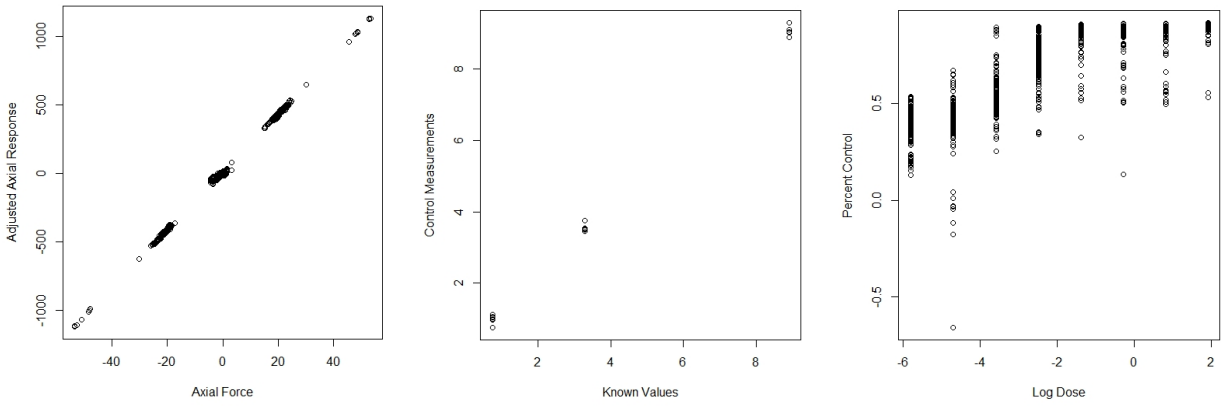
Require: Randomly choose h profiles where we again define h = [a] and $a = \frac{1}{2}(m+p+1)$ where h represents the integer part of a. Such that a subset $|H_0| = h$.

repeat
  Calculate $\bar{x}_0$ and S_0
  if $\det(S_0) \neq 0$ **then** add another profile to the sample and recheck condition
until $\det(S_0) = 0$ ▷ Let $H_{old} = H_0$, $\bar{x}_{old} = \bar{x}_0$, and $S_{old} = S_0$
repeat Calculate for $i = 1$
  Calculate the distances using the mean and variance from the current subset
  Order the distances and select the profile indices for the h smallest distances and set as H_{new}
  Calculate $\bar{x}_{new}$ and S_{new}
  if $\det(S_{new}) = 0$ *or* $\det(S_{new}) = \det(S_{old})$ **then** Set $\bar{x}_{MCD} = \bar{x}_{new}$ and $S_{MCD} = S_{new}$
  else Set $\bar{x}_{old} = \bar{x}_{new}$ and $S_{old} = S_{new}$
until $\det(S_{new}) = 0$ *or* $\det(S_{new}) = \det(S_{old})$

2.3.2. Parametric Profile Monitoring Example

The first data observed in this section is from the NASA Langley Research Center, seen in Mahmoud et al., 2007. The objective was to investigate replicated calibrations of a force balance within wind tunnel experiments. The data has 11 independent samples totaling 723 observations, we consider the axial force as our regressor variable with an adjusted axial response. Experiment details are in Parker et al., 2001 and the data is plotted in figure 2.8a and will be referred to as the NASA data. We use the NASA data to exemplify how the profile monitoring techniques can be applied to industry-type processes. This data contains a good number of observations for us be able to engage in using all of the regression methods. The second data used in this section, found in Croarkin et al., 2006, is plotted in figure 2.8b and will be referred to

as the NIST data. Here, the line widths of three photomask reference standards were measured over six days via an optical imaging system at the low, middle, and high ends of the calibration line. The regressor variable for this is the known value of the photomask reference and the response is the control measurement (what the optical imaging system observes). We consider each day as a profile containing 3 observations (low, middle, and high) each, totaling 18 observations. We chose to use the NIST data because of its small size, allowing us to compare profile monitoring methods for smaller and larger datasets. The third dataset is from HDS from the DuPont Crop Protection and has been used in Williams, 2011 and Gomaa and Birch, 2019, its data is plotted in figure 2.8c and will be referred to as the DuPont data. The DuPont data consists of 44 weeks of bioassay experiments with 8 dose levels of a diluted commercial growth-inhibiting herbicide compound which is replicated 4 times. That is, we have 44 profiles each containing $8 * 4 = 32$ observations for a total of 1408 observations. The DuPont data is desirable for us to show how profile monitoring techniques can be used on biological data. The NASA and NIST datasets have a linear behavior so will be fit with a linear model. The DuPont data is logistic in nature and will follow a 4-parameter logistic model as seen in Williams et al., 2007 and Jensen and Birch, 2009.



(a) NASA Data (b) NIST Data (c) DuPont Data

Figure 2.8: Raw Data

For the NASA and NIST data we fit a linear model to each profile and obtain a column vector, $\hat{\boldsymbol{\beta}}_{\boldsymbol{i}}(i = 1,2,\ldots,m)$, containing an estimate of β_0 and β_1. The estimates of β_0 and β_1 are our observations, meaning we have 11 samples with 2 observations each ($m = 11$ and $p = 2$) and 6 samples with 2 obersavions each ($m = 6$ and $p = 2$) for the NASA and NIST data respectively. We take the average over all the profiles to obtain $\bar{\boldsymbol{\beta}}$ for the respective datasets. Since $T^2 \dot{\sim} \chi^2$ because the population mean and standard deviation are unknown, the UCL for this data is $\chi^2_{0.05,p=2} = 5.991$.

The DuPont data will follow the 4-parameter logistic model given as: $y_i = a_i + \frac{d_i - a_i}{1+10^{b_i(c_i - x_i)}} + \epsilon_i$ for $i = 1,2,\ldots,m$. In the model, a_i is the maximum response parameter, d_i is the minimum response parameter, c_i is the point of inflection or the ED_{50} (elicits 50% response), and b_i is the rate parameter (how quickly the response changes from minimum to maximum response or slope at point c_i). This model will be fit for each profile and the parameters estimated to obtain a column vector, $\hat{\boldsymbol{\beta}}_{\boldsymbol{i}} = \begin{bmatrix} \hat{a} \\ \hat{d} \\ \hat{c} \\ \hat{b} \end{bmatrix} (i = 1,2,\ldots,m)$ and the UCL for this data is $\chi^2_{0.05,p=4} = 9.488$. Again we will obtain $\bar{\boldsymbol{\beta}}$ by averaging the $\hat{\boldsymbol{\beta}}'s$ over all profiles. Figure 2.9 provides visual guidance when conducting analysis using the T^2 chart. The calculation for the variances shown in this figure can be found in table 2.1.

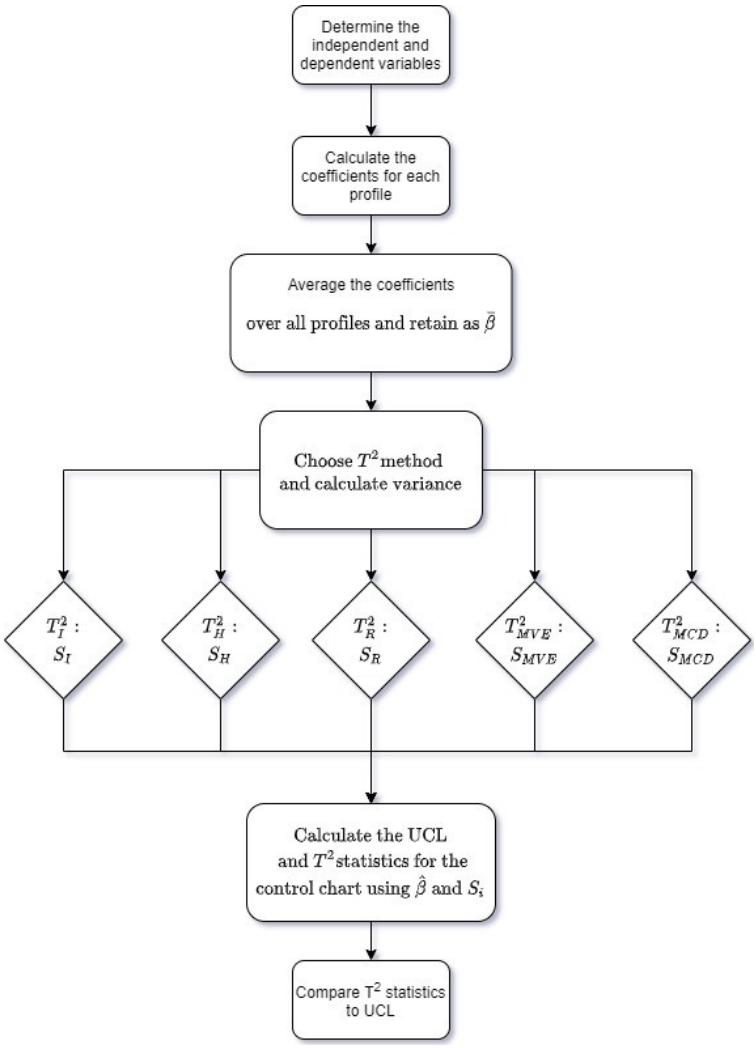
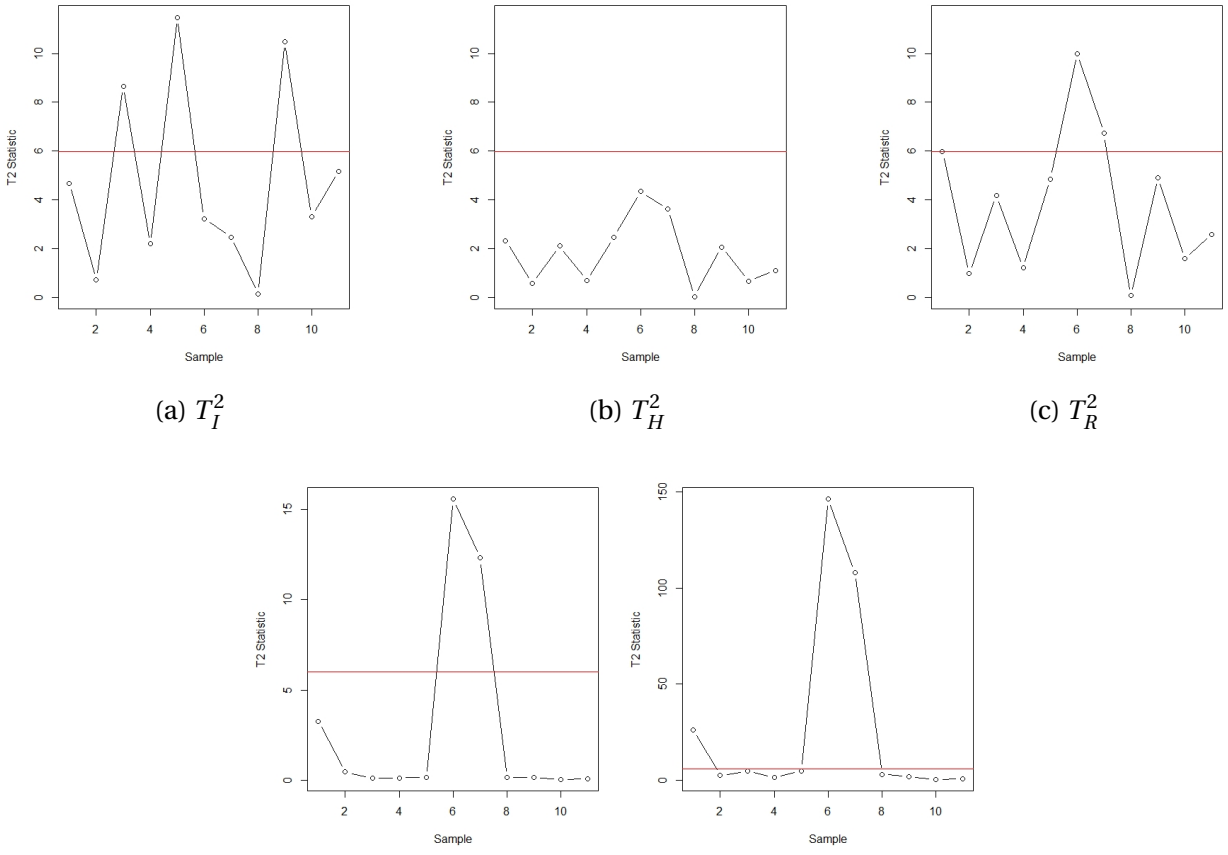


Figure 2.9: T^2 Method Flowchart

For the NASA data using traditional methods, the T_I^2 chart signaled at profiles {3, 5, 9}, the T_H^2 chart did not signal, and the T_R^2 chart signaled at samples {6, 7}, but nearly signaled at sample 1 (see figures 2.10a, 2.10b, 2.10c). For the robust methods, the T_{MVE}^2 chart signaled at samples {6, 7} (figure 2.10d) and the T_{MCD}^2 chart signaled at samples {1, 6, 7} (figure 2.10e), showing that there is more consistency among the robust charts. The PoS of the methods are: T_I^2 = 0.2727, T_H^2 = 0, T_R^2 = 0.1818, T_{MVE}^2 = 0.1818, and T_{MCD}^2 = 0.2727. The T_I^2 and T_{MCD}^2 return the same PoS, but the out-of-control profiles are not exact. However, the T_R^2 and T_{MVE}^2 PoS and out-of-

control profiles were the same. We recommend using either the T^2_{MVE} or T^2_{MCD} chart because of their robust features or the T^2_R chart because of the similar results it produced compared to the robust charts. Since this analysis is done within Phase I and historical information is not provided about the data it is difficult to make any further recommendation. Note, comparisons for the MVE method may not be exact due to the bootstrap sub-sampling and yields different numerical results each time the method is run. The pattern for which profiles are outliers and *PoS* should remain the same.



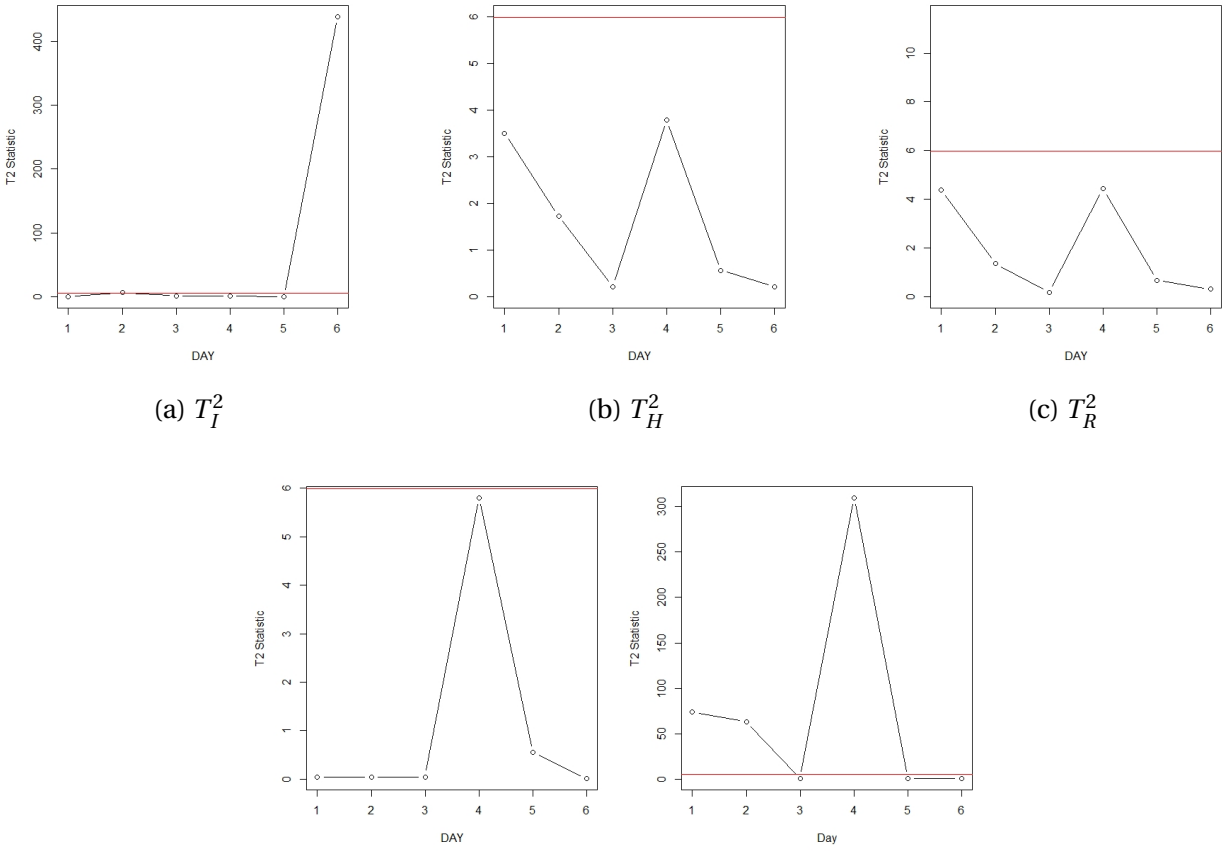
(a) T^2_I (b) T^2_H (c) T^2_R

(d) T^2_{MVE} (e) T^2_{MCD}

Figure 2.10: Hotelling's T^2 Control Charts for NASA Data

For the NIST data using the traditional methods, the T^2_I chart signaled at profiles {2, 6},

the T^2_H and T^2_R charts did not signal (see figures 2.11a, 2.11b, 2.11c). For the robust methods, the T^2_{MVE} chart did not signal (figure 2.11d) and the T^2_{MCD} chart signaled at samples {1, 2, 4} (figure 2.11e).The PoS of the methods are: $T^2_I = 0.3333$, $T^2_H = T^2_R = T^2_{MVE} = 0$, and $T^2_{MCD} = 0.5$. There is some consistency in the shape of the charts for the robust methods, but they do not indicate the same thing. Three of the methods show that the data in Phase I was in-control and no profiles needed to be removed. However, T^2_I and T^2_{MCD} suggest that profiles be removed because of their signaling. In Montgomery, 2013 the authors look at this data using separate Shewhart charts for the intercept, slope, and error variance. In both Croarkin et al., 2006 and Montgomery, 2013 they conclude that day 4 was out-of-control. Using the multivariate chart, we see that day 4 consistently shows a spiking pattern (except when using T^2_I) and breaching the control limit in the T^2_{MCD} chart.



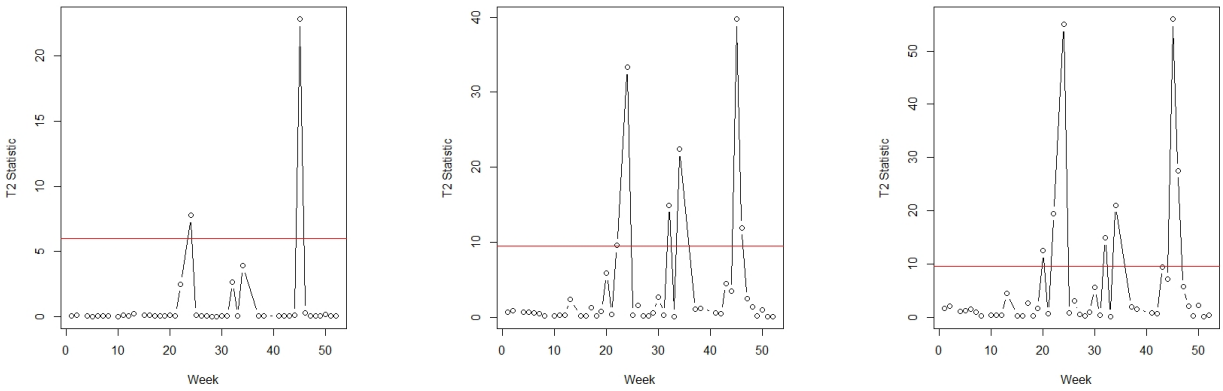
(a) T_I^2 (b) T_H^2 (c) T_R^2

(d) T_{MVE}^2 (e) T_{MCD}^2

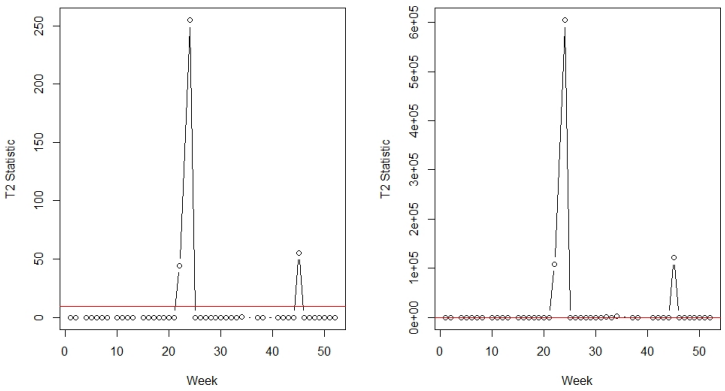
Figure 2.11: Hotelling's T^2 Control Charts for NIST Data

For the DuPont data using the traditional methods, the T_I^2 chart signaled at profiles {20, 37}, the T_H^2 chart signaled at {17, 19, 20, 28, 30, 37, 38}, and T_R^2 signaling at {19, 20, 28, 30, 37, 38} (see figures 2.12a, 2.12b, 2.12c). For the robust methods, the T_{MVE}^2 chart signaled at {19, 20, 37} (figure 2.12d) and the T_{MCD}^2 chart signaled at nearly half of the samples (figure 2.12e).The PoS of the methods are: T_I^2 = 0.04545455, T_H^2 = 0.1363636, T_R^2 = 0.1590909, T_{MVE}^2 = 0.06818182, and T_{MCD}^2 = 0.4772727. We see again that there is some consistency in the shape of the charts for the robust methods, but they do not indicate the same thing. All of the methods show that profiles 20 and 37 are outliers in Phase I and need to be removed. This differs from results

obtained in Gomaa and Birch, 2019, whose T^2 chart using parametric methods did not detect any outliers. But their findings for their nonparametric and semiparametric approaches with the T^2 chart found that profiles {17, 19, 20, 30, 37, 38} were outliers, which match many of the outlier profiles in the charts observed in figure 2.12.



(a) T^2_I (b) T^2_H (c) T^2_R



(d) T^2_{MVE} (e) T^2_{MCD}

Figure 2.12: Hotelling's T^2 Control Charts for DuPont Data

In the following section we will examine methods in nonparametric statistics as they relate to profile monitoring. Three regression techniques will be introduced and used on real data and a comparative analysis conducted with recommendations.

2.4. Nonparametric Methods for Profile Monitoring in Regression

Nonparametric statistics refers to a branch of statistics where the underlying distribution is unknown, or a distribution is specified but the parameters of the distribution are unknown. These methods are often used on ordinal/ranked datasets and aim to make as few assumptions on the data as possible. This leads to more robust estimation results and a broader application compared to parametric statistics. However, if nonparametric tests are used where it is more appropriate to use parametric tests, the nonparametric test will achieve a decreased power. We will discuss Gaussian kernels, B-splines, P-splines, and Haar wavelets used as curve smoothing tools.

2.4.1. Kernel Regression

Kernel regression is a heavily investigated nonparametric method as it is known for its computational simplicity. In general, kernel regression puts heavier weights on points close to desired prediction points and less weight on further away points. This returns a normal-curve shaped kernel density function for each prediction point. The method of using a kernel as a weighting function was proposed in Nadaraya, 1964 and Watson, 1964, their combined efforts produced the Nadaraya-Watson estimator, more can be seen in Bierens, 1988.

In eq. 2.9, $K(\cdot)$ is the chosen kernel function. The rate the weight decreases, the further the point is from prediction which determines the smoothness and is controlled by the bandwidth, h. Choice of bandwidth is vital to obtaining a proper balance between the variance and bias of the model. If h chosen too close to zero (very small h), less points are included in prediction and nearly all the weight is on the prediction point. This produces an overfit model, resulting in large

variance (like connecting the dots). Contrarily, h chosen too close to the range of the x-values (very large h), the weights are distributed evenly across all observations and gives an underfit model high in bias (like fitting the mean). There are different kernel functions, but the commonly mentioned are Boxcar, Epanechinkov, Gaussian, and Tricube which can be seen in Mays et al., 2001. It is more important which bandwidth is chosen than the kernel function used, so most research uses the Gaussian kernel ($(2\pi)^{-\frac{1}{2}} e^{\frac{-u^2}{2}}$) because its simplicity and properties. We will use the Gaussian kernel function for our examples. While `R` offers a `ksmooth` function for kernel regression under the `stats` package, we show in figure 2.13 how the Gaussian kernel can be written in `R`.

$$\hat{y}_j = \sum_{j=1}^{n} y_j \gamma_j, \text{ for } j = 1,2,\ldots,n = \#\text{ of observations}$$
$$\text{where } \gamma_j = \frac{K\left(\frac{X_j - x}{h}\right)}{\sum_{j=1}^{n} K\left(\frac{X_j - x}{h}\right)}, h > 0 \tag{2.9}$$

```
for(i in 1:n1){
  for(k in 1:m1){

    top2 <- cbind()
    for(j in 1:m1){
      dtemp1 <- (xtemp[k] - xtemp[j])/h1
      dtemp2 <- kernfunc(dtemp1)
      top2 <- c(top2, dtemp2)
    }
    s2 <- sum(top2)
    tempWts <- as.matrix(top2/s2)
    wts2 <- cbind(wts2, tempWts)
  }
  yhtemp <- wts2%*%ytemp
  yh_kern <- qpcR:::cbind.na(yh_kern, yhtemp)
  yh2_k <- rbind(yh2_k, yhtemp)
  meanTemp <- mean(yhtemp)
  means2 <- cbind(means2, meanTemp)
}
for(e in 1:n1){

  tyH1 <- yh_kern[, e]
  tyH2 <- tyH1[!is.na(tyH1)]
  tx1 <- x1[, e]
  tx2 <- tx1[!is.na(tx1)]
  xmat1 <- cbind( rep(1, (m2[e]) ), tx2)
  bHat2 <- solve(t(xmat1)%*%xmat1)%*%t(xmat1)%*%tyH2
  bhatAll_k <- cbind(bhatAll_k, bHat2)
}
betabar_k <- rbind(mean(bhatAll_k[1 ,]), mean(bhatAll_k[2 ,]))
```

Figure 2.13: R Code for Kernel Regression

2.4.2. Splines

Statistical splines are used to minimize oscillating behavior in the data and are a collection of generally different approximated functions of sub-intervals divided over a finite interval Burden and Faires, 2010. This is done by partitioning the data at locations, called knots, fitting a

polynomial function to the data within the partitioned intervals, and smoothly joining these segments together. Knots are endpoints of the sub-intervals that tie the functions together and are either prespecified by an expert who is knowledgeable on the behavior of the data or chosen as coordinate locations where the plot appears to have a steep slope and initiates a change in direction. Various methods for knot selection were suggested in Kauermann et al., 2009 and Ruppert et al., 2009, with the latter concluding the number of knots and their position do not make an extreme impact on the results if the range of values is covered well. Some literature pertaining to knot selection and position include: Mosteller and Wallace, 1963, Stone, 1974, Geisser, 2017 (leave-one-out Cross-Validation methods (CV)), Friedman and Silverman, 1989 (TURBO method), Breiman, 1993 (Delete-Knot/Cross-Validation method (DKCV)), Denison et al., 1998 (reversible jump Markov chain Monte Carlo method), Molinari et al., 2004(Akaike Information Criterion (AIC) or Bayesian Information Criterion (BIC)), and Lolive et al., 2006 (Simulated-Annealing strategy). Different types of splines are available for different data needs. Splines are recommended for handling sparse data while polynomial smoothers are preferred for handling dense designs Piri et al., 2019. More on splines can be found in Ruppert et al., 2009, O'Sullivan, 1986, Abdel-Salam, 2009, and De Boor et al., 1978. Keep in mind that spline functions are used to fit the data and aim to capture its behavior, this creates a limitation that if error is present within the data the values obtained using the spline function may vary from the expected values.

Let S be the spline over the [a, b] interval that maps to the real numbers: $S : [a, b] \rightarrow \mathbb{R}$. The [a, b] interval is partitioned into k sub-intervals, where there are k knots: $[a, b] = [t_0, t_1] \cup [t_1, t_2] \cup \cdots \cup [t_{k-2}, t_{k-1}] \cup [t_{k-1}, t_k]$, such that $a = t_0 \leq t_1 \leq \cdots \leq t_{k-1} \leq t_k = b$. Let ρ_i represent a function associated with the sub-interval [i, i+1]: $\rho_i : [t_i, t_{i+1}]$. Finally, define the spline function as the sum of the k sub-interval functions with an averaged weight associated with the sub-interval: $S(t) = \sum_i^k \alpha_i \rho_{i,n}(x)$.

We define i to be the number of knots (0 to k), and the number of sub-intervals is from 1 to *k-1*.

(1) $\rho_i(t_{i+1}) = \rho_{i+1}(t_{i+1})$: The i^{th} function at point t_k is equal to the $(i+1)^{th}$ function at point t_k. Meaning there exists continuity in the spline where the intervals connect at the knots.

(2) $\rho_i'(t_{i+1}) = \rho_{i+1}'(t_{i+1})$: The derivative of the i^{th} function at point t_k is equal to the derivative of the $(i+1)^{th}$ function at point t_k. That is, if the derivative of the left side (ρ_i) is the same as the derivative of the right side (ρ_{i+1}) then there is no relative extremum at the point (t_k). This ensures smoothness at these continuous points.

(3) $\rho_i''(t_{i+1}) = \rho_{i+1}''(t_{i+1})$: The second derivative of the i^{th} function at point t_k is equal to the second derivative of the $(i+1)^{th}$ function at point t_k. That is, if the second derivative of the left side (ρ_i) is the same as the second derivative of the right side (ρ_{i+1}) then this results in either a graph that is concave up if both are positive or concave down if both are negative at the point (t_k). This ensures the same concavity at the knot. Where,

   (a) $f(t_j) = \rho_i(t_i)$: The response values for the original function and for the polynomial functions are the same.

   (b) $f'(t_i) \neq \rho_i'(t_i)$ and $f'(t_k) = \rho_{k-1}'(t_k)$: The derivatives for the original function and for the polynomial functions are not the same except at the end point.

Conditions imposed on the endpoints of the spline, such as the derivatives must be equal to zero, indicates a spline clamped spline. Otherwise the spline is natural/free.

2.4.2.1. B-Splines

A b-spline with degree ω is made of ω+1 polynomial pieces and each piece is of degree ω. These pieces overlap with 2ω of its polynomial neighbors, except at the spline boundaries. Each polynomial piece is joined at ω inner knots and at these joining locations, their derivatives are continuous up to the ω-1 order (i.e if ω=3, the derivative of the joining knots is continuous up to the second derivative). The b-spline is positive within the range of ω+2 knots, zero elsewhere, and at a given x value, (ω+1) b-splines are nonzero.

The B-Spline is defined as

$$S(t) = \sum_{i}^{k} \alpha_i B_{i,n}(x)$$

$$\text{where, } B_{i,0}(x) = \begin{cases} 1, \text{ if } t_i \leq x < t_{i+1} \\ 0, otherwise \end{cases} \tag{2.10}$$

For $m \leq M$,

$$B_{i,m}(x) = \frac{x - t_i}{t_{i+m} - t_i} B_{i,m-1} \frac{t_{i+m+1} - x}{t_{i+m+1} - t_{i+1}} B_{i+1,m-1}(x)$$

The goal of the b-spline is to minimize the sums of squares ($S = \sum_i (y_i - \sum_j \alpha_j B_j(x_i))^2$) over all functions. However, by simply minimizing this estimate the least squared error is obtained and isn't accurate for fitting a smooth spline to the data. As a solution to this problem adding a discrete penalty to acquire greater smoothness of the spline and measured the roughness of a curve was suggested in O'Sullivan, 1986 using equation $R = \int_l^u [f''(x)]^2 dx$, l and u are the lower and upper bounds of the domain.

2.4.2.2. P-Splines

Penalized splines (P-splines) were introduced in O'Sullivan, 1986 and brought to the forefront after discussion paper Eilers and Marx, 1996 was published. P-splines are the same as b-splines, but impose a discrete penalty factor to achieve a smoother fit. An extensive review of P-splines from 1995 to 2015 can be seen in Eilers et al., 2015. The $\hat{y}$ values are calculated in eq 5.1 where D is a lower diagonal matrix corresponding to the number of knots, Λ is the penalizing parameter, and p is the number of columns in $\boldsymbol{X}$. `R` offers a `smooth.spline` function for p-spline regression under the `pspline` package, in figure 2.14 we give an example of how the p-spline can be written in `R`.

$$\hat{y} = \boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{X} + \Lambda^{2p}D)^{-1}\boldsymbol{X}^T y \tag{2.11}$$

```
for(i in 1:n1){
  mid1 <- ( t(xMat1)%*%xMat1 + ( lam1^(2*p1) ) *D1 )
  mid2 <- ifelse( solve(mid1) == 0, Ginv(mid1), solve(mid1) )
  yHat1 = xMat1%*%mid2%*%t(xMat1)%*%ytemp
  yh_Splines <- qpcR:::cbind.na(yh_Splines, yHat1)
  yh2_sp <- rbind(yh2_sp, yHat1)
  meanTemp <- mean(yHat1)
  means1 <- cbind(means1, meanTemp)
  bhat1 <- mid2%*%t(xMat1)%*%yHat1
  bhatAll_s <- cbind(bhatAll_s, bhat1)
}
betabar_s <- rbind()
n2 <- nrow(bhatAll_s)
for(i in 1:n2){
  betabar_s <- rbind(betabar_s, sum(bhatAll_s[i, ])/n1)
}
```

Figure 2.14: R Code for P-Spline Regression

2.4.3. Wavelets

Wavelets, like splines, use a basis function to represent the underlying function. They are comparable to Fourier series which use sine and cosine waves to represent a function. The disadvantage of the Fourier series is it solely localizes in frequency, meaning it has a one-dimensional function mapping to a one-dimensional sequence. Wavelets are more informative as they map the one-dimensional function into a two-dimensional array, representing the location in time and frequency. Wavelets are often preferred over splines when the profile exhibits sudden changes/spikes. It is usually recommended when the shape of the profile is too complicated to be modeled by a linear or nonlinear model Woodall, 2007.

A wavelet basis is made from a mother wavelet ($\psi(x)$) and a father wavelet ($\phi(x)$). The mother wavelet gives the essential behavior of the basis and the father wavelet is a scaling function. Reis and Saraiva, 2006, Jeong et al., 2006, Zou et al., 2007 Chicken et al., 2009, Chang and Yadama, 2010, and Nikoo and Noorossana, 2013 use the nonparametric wavelet approach in some capacity to help inform a model to use in constructing control charts. A more in-depth understanding of wavelets can be acquired from Daubechies, 1992, Ogden, 2012, and Burrus et al., 1998.

In this work we use the Haar wavelet to estimate model parameters, which is said to generate the simplest possible orthogonal wavelet system Burrus et al., 1998. Valuable insight of wavelet use in SPC is offered in Piri et al., 2019 by exploring the use of the Haar basis function on the poisson distribution for profile monitoring. Before using the Haar wavelet two conditions must be met. First, the number of observations in the data must be a power of two, and the second is the function is defined on the interval $[0,1)$. In eq. 2.12, c_k represents the father wavelet coefficients and $d_{j,k}$ is are the coefficients for the mother wavelet. We choose to use the built-in `wd` function offered in `R` under the `wavethresh` package to perform our regression

under the Haar wavelet, this can be seen in figure 2.15.

$$f(x) = \sum_{k=-\infty}^{\infty} c_k \phi(x-k) + \sum_{j=0}^{\infty} \sum_{k=-\infty}^{\infty} d_{j,k} \psi(2^j x - k) \text{ such that,}$$

$$\psi(x) = \begin{cases} 1, \text{if } 0 \le x < \frac{1}{2} \\ -1, \text{if } \frac{1}{2} \le x < 1 \\ 0, \text{otherwise} \end{cases} \quad \text{and } \phi(x) = \begin{cases} 1, \text{if } 0 \le x < 1 \\ 0, \text{otherwise} \end{cases} \tag{2.12}$$

$$\text{where, } \begin{cases} \phi(x)_{j,k} = 2^{j/2} \phi(2^j x - k) \\ \psi(x)_{j,k} = 2^{j/2} \psi(2^j x - k) \end{cases}$$

```
for(i in 1:n1){
  jtemp <- as.integer(log2(m1))
  j1 <- cbind(j1, jtemp)
  jMax <- jtemp - 1
  n2 <- 2^jtemp
  wd_temp <- wd(ytemp2, filter.number = 1, family = "DaubExPhase")
  yh_temp1 <- as.matrix(wr(nullevels(wd_temp, levelstonull = jMax)))
  yh_wave <- qpcR:::cbind.na(yh_wave, yh_temp1)
  yh2_w <- rbind(yh2_w, yh_temp1)
}
for(e in 1:n1){
  tyH1 <- yh_wave[, e]
  tyH2 <- tyH1[!is.na(tyH1)]
  tx1 <- xnew2[, e]
  tx2 <- tx1[!is.na(tx1)]
  xmat1 <- cbind( rep(1, (m3[e]) ), tx2)
  bHat1 <- solve(t(xmat1)%*%xmat1)%*%t(xmat1)%*%tyH2
  bhatAll_W <- cbind(bhatAll_W, bHat1)
}
betabar_W <- rbind(mean(bhatAll_W[1 ,]), mean(bhatAll_W[2 ,]))
```

Figure 2.15: R Code for Wavelet Regression

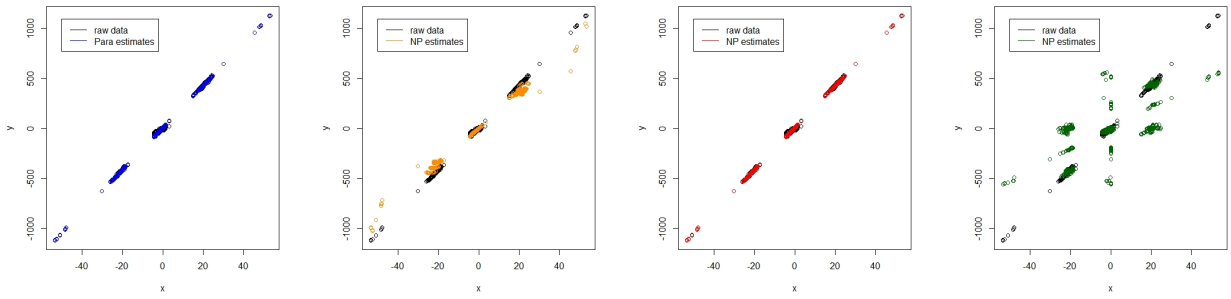
2.4.4. Nonparametric Profile Monitoring Example

This section revisits data from section 2.3.2 and applies nonparametric methods under linear ($\hat{\boldsymbol{\beta}} = \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{bmatrix}$) and nonlinear ($\hat{\boldsymbol{\beta}} = \begin{bmatrix} \hat{a} \\ \hat{d} \\ \hat{c} \\ \hat{b} \end{bmatrix}$) models. $\hat{\boldsymbol{\beta}}_{\boldsymbol{i}}$ $(i = 1, 2, \ldots, m)$ is calculated using estimated $\hat{y}$ values for each profile and used to construct the control charts. The wavelet approach was not used on the NIST data because the condition that the number of observations must be of power two was not met. In figures 2.16, 2.17, and 2.18 the $\hat{y}$ values are plotted for each method against the raw NASA data, NIST data, and DuPont data respectively (parametic in 2.16a, 2.17a, and 2.18a, kernel in 2.16b, 2.17b, and 2.18b, p-spline in 2.16c, 2.17c, and 2.18c, and wavelet in 2.16d, 2.18d). For the Gaussian kernel we use a bandwidth of h=10, 3, 15 for the NASA, NIST, and DuPont data respectively, which is chosen to be a fair balance for each dataset. Since the original data lacks curvature, only two knots are chosen for the P-spline for the linear data and are taken to be the ends of the data for both data sets, while the DuPont data takes four evenly spaced knots under the P-spline method. For the wavelet approach on the NASA and DuPont data we choose to use resolution $J = 4$.

The $\hat{y}$ values using the kernel method in figures 2.16b, 2.17b, and 2.18b vary from the true data because the method uses the nearby points to estimate the current point, causing prediction error near the ends. It appears to be an almost exact fit in figures 2.16c and 2.17c because of the choice of knots and the lack of curvature. The estimates vary widely for the linear data in figure 2.16d, but prove to be a better fit for the nonlinear data in figure 2.18d. Mean Squared Error (MSE) values in table 2.2 reinforce what is seen in figures 2.16, 2.17, and 2.18.

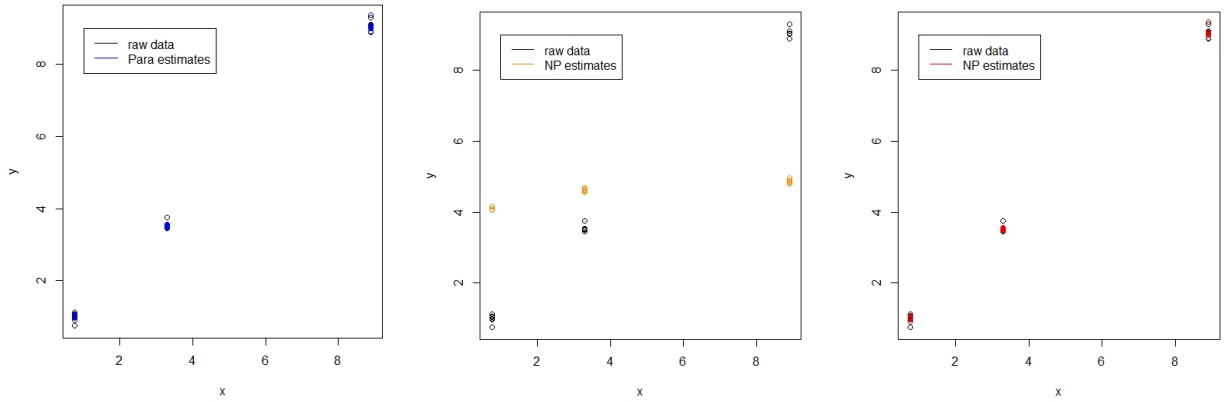
The EWMA chart uses $\lambda = 0.3$, allowing for longer memory. In both the CUSUM and EWMA charts we use a bootstrap technique of in-control profiles (reference section 2.3.2) to calculate the UCL's. This is done since we do not know the underlying truth for in-control profiles for the data. The UCL for the T^2 chart is the same as in section 2.3.2, except when using splines, it is

$\chi^2_{0.05,4} = 9.488$ for the linear data (p = 2 knot estimates plus 2 $\hat{\beta}$ estimates) and $\chi^2_{0.05,8} = 15.507$ for the nonlinear data (p = 4 knot estimates plus 4 $\hat{\beta}$ estimates).



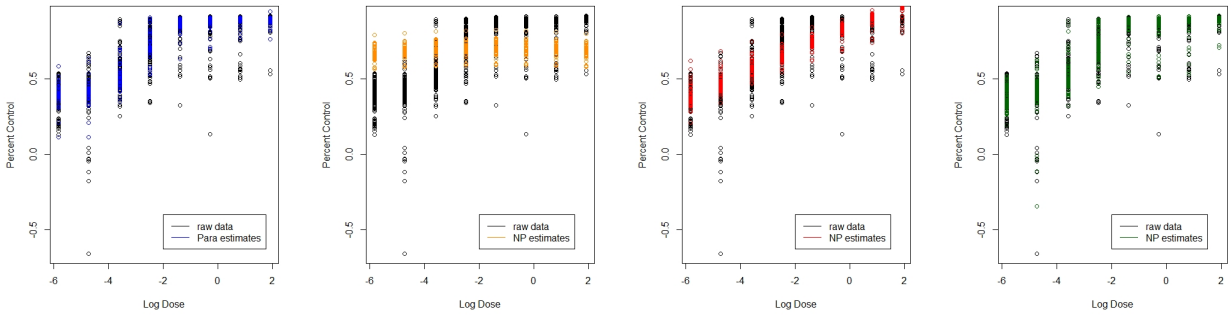
(a) $\hat{y}_p$ (b) $\hat{y}_k$ (c) $\hat{y}_s$ (d) $\hat{y}_w$

Figure 2.16: NASA Data $\hat{y}$ values using parametric and nonparametric methods



(a) $\hat{y}_p$ (b) $\hat{y}_k$ (c) $\hat{y}_s$

Figure 2.17: NIST Data $\hat{y}$ values using parametric and nonparametric methods



(a) $\hat{y}_p$ (b) $\hat{y}_k$ (c) $\hat{y}_s$ (d) $\hat{y}_w$

Figure 2.18: DuPont Data $\hat{y}$ values using parametric and nonparametric methods

Method	NASA MSE	NIST MSE	DuPont MSE
Parametric	104.12	0.0049	0.0044
Gaussian Kernel	4296.72	0.8663	0.0457
P-Splines	104.12	0.0049	0.0104
Haar Wavelets	67363.67	n/a	0.0016

Table 2.2: Mean Squared Error Results for nonparametric methods

	NASA			**PoS NIST**			**DuPont**		
Method	EWMA	CUSUM	T^2	EWMA	CUSUM	T^2	EWMA	CUSUM	T^2
Gaussian Kernel	0.9091	0.8182	0.2727	0.6667	0.8333	0	0.8636	0.9545	0.7727
P-Splines	0.9091	0.8182	0.5455	0.6667	0.6667	0.5	0.5227	0.9545	0.7727
Haar Wavelets	0.7273	0.8182	0.0909	n/a	n/a	n/a	0.8636	0.9545	0.7727

Table 2.3: Probability of Signal Results for control charts using nonparametric methods

From table 2.2 and figures 2.16, 2.17, and 2.18 we recommend the use of P-splines over the Gaussian kernel and Haar wavelet for simple linear data but recommend the Haar wavelet for nonlinear data. The P-spline was simple to implement and gave the lowest MSE for the simple linear regression cases, but when applied to the nonlinear data it did not perform as well. The Haar wavelet had the lowest MSE for the nonlinear data and fit the ends of the data better than

the other methods used. We recommend the use of the CUSUM chart with our nonlinear data when detecting drift shifts because it consistently had the largest PoS between the EWMA and itself (seen in table 2.3), however the choice of λ is imperative to PoS results. The EWMA chart out-performed the CUSUM chart overall for the NASA data, but the CUSUM chart remained consistent in its detection ability. For the NIST data both charts performed almost the same. The T^2 chart proved to have the lowest PoS overall, but this is because it is meant to detect large/sudden changes whereas there did not seem to be any occurrences in the data used.

2.5. Semiparametric Methods in Regression

If there is information known about the data distribution or its parameters, then parametric methods are applied, otherwise nonparametric methods must be employed to inform a model. However, solely using parametric methods for estimating a model often leads to model misspecification since we don't know the true model, which causes uncertainty resulting in high bias. Whereas estimating a curve solely from nonparametric methods can fit the data too closely, introducing a high amount of variance into the model. But what if there was a way to mediate the consequences of relying exclusively on the parametric method or the nonparametric method? Olkin and Spiegelman, 1987 used a mixing proportion to create a parametric and nonparametric estimate in an approach to estimate density functions, utilizing maximum likelihood to estimate the parameter of the parametric model. Meanwhile, Ke and Wang, 2001 approximated the log-likelihood function by way of the Laplace method. In regression both approaches are considered as methods of model robust regression (MRR), and the concept was extended into phase I analysis and introduced as model robust profile monitoring (MRPM) in Abdel-Salam, 2009. These methods combine information obtained from parametric and nonparametric techniques via a mixing parameter (Λ), known as a semiparametric method. This

allows for retention of known information about the data by including part of a parametric fit while including the functional flexibility gained from a nonparametric fit.

Under the normality assumption, Einsporn, 1987 and Einsporn and Birch, 1993 introduced their model robust regression 1 (MRR1) semiparametric method for modeling mean response. MRR1 is an umbrella term encompassing the use of original data to estimate a model using a convex combination of the parametric and nonparametric fits via a mixing parameter, $\eta \in [0,1]$. When η is chosen to be 0 only the parametric method is being used, while if η is taken to be 1 the nonparametric method is the sole contributor. The MRR1 concept was developed in both Burman and Chaudhuri, 2012 and Mays et al., 2000, termed by the latter . Model robust regression 2 (MRR2), seen in Mays et al., 2000, uses a combination of the residuals from the parametric and nonparametric fits rather than their $\hat{y}$ values.

Nottingham and Birch, 2000 strayed from the frequented logistic regression methods and extend semiparametric methods to analyze dose-response data. Waterman et al., 2015 compare parametric, nonparametric, and mixed model robust methods via their approximated integrated mean square errors then apply the methods to two data sets. Gomaa and Birch, 2019 introduces and evaluates a nonlinear mixed robust profile monitoring (NMRPM) method against its parametric counterpart, concluding their method is capable of easily computing control charts. The mixed model robust residuals profile monitoring (MMRRPM) method was proposed in Siddiqui and Abdel-Salam, 2019, which implements the MRR2 technique as an expansion of the mixed model robust profile monitoring MMRPM (Abdel-Salam, 2009) method. In this paper we focus on MRR1 as a semiparametric approach, defined as the following:

$$\hat{y}^{MRR1} = (1-\eta)\,\hat{y}^{P} + \eta\,\hat{y}^{NP} \tag{2.13}$$

Figure 2.19 gives R code implementing the MRR1 equation, where *lam2* represents the mix-

ing parameter η.

```
for(i in 1:n3){
  spTemp <- ((1 - lam2) * yh_para[i]) + (lam2 * yh2_sp[i])
  yhm_s <- rbind(yhm_s, spTemp)
}
```

Figure 2.19: R Code for Semiparametric Regression

2.5.1. Semi-Parametric Profile Monitoring Example

We implement the semiparametric method mixing the parametric results with each of the nonparametric methods from section 2.4.4. The $\hat{y}$ values from the parametric and the nonparametic methods are used in the MRR1 equation and η is taken to be 0.3 since the parametric results appear to fit the data best. Figure 2.20 provides visual guidance for using the nonparamteric and semiparametric regression methods described in sections 2.5 and 2.4.

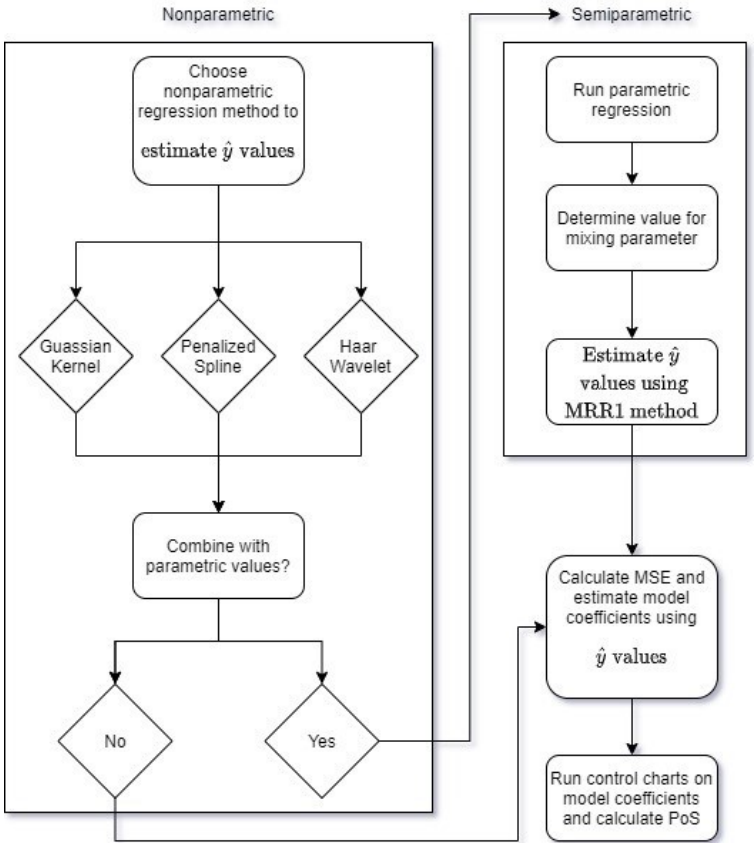


Figure 2.20: Nonparametric and Semiparametric Regression Flowchart

The $\hat{y}$ values from the semiparametric method using each nonparametric method are compared via MSE. The kernel method fits similar to figures 2.16b, 2.17b, and 2.18b, respective of the data used, but has a better fit on the ends. The P-spline method yields the same fit as before for the linear data, but has a better fit for the DuPont data than the nonparametric by itself. While the Haar wavelet approach for the NASA data fits closer to the original data it still has a high prediction error, for the DuPont data it fits the same as solely using the nonparametric method. In table 2.4, the MSE in both linear data sets for the kernel and wavelet decrease drastically from the nonparametric approach. Notice the MSE for the parametric method is the same as for the P-splines in both cases. And for the nonlinear data the kernel and spline methods show a decrease in MSE from the nonparametric method as the wavelet method remains

constant. We conclude that using the P-spline approach for our linear data yielded comparable results to the parametric approach, but the Haar wavelet was best in our nonlinear data case.

The EWMA, CUSUM, and T^2 control charts are applied to the semiparametric results. We use the same conditions as stated in section 2.4.4, besides the UCL for the T^2 chart with the spline method. The UCL for the T^2 chart is the same as the other nonparametric methods because we are only using the $\hat{y}$ values for the semiparametric approach. The PoS for the EWMA and CUSUM charts using the kernel and spline methods for the linear data are the same from section 2.4.4 except the CUSUM PoS values decreases for the NIST data and the wavelet method for the NASA data yields a much higher PoS. Under the semiparametric method EWMA and the CUSUM charts using the DuPont data return the same PoS for each nonparametric method, but show an overall increase for the CUSUM chart and an overall decrease for the EWMA chart. The T^2 chart returns the same results for the PoS as the nonparametric example for all techniques for the linear data and a decrease for the nonlinear data, except for the wavelet method.

	NASA		NIST		DuPont	
Method	SP MSE	NP MSE	SP MSE	NP MSE	SP MSE	NP MSE
Parametric	n/a	104.12	n/a	0.0049	n/a	0.0044
Gaussian Kernel	263.82	4296.72	0.3859	0.8663	0.0417	0.0457
P-Splines	104.12	104.12	0.0049	0.0049	0.0098	0.0104
Haar Wavelets	6521.47	67363.67	n/a	n/a	0.0016	0.0016

Table 2.4: Mean Squared Error comparison for semiparametric and nonparametric methods

	PoS								
	NASA			**NIST**			**DuPont**		
Method	EWMA	CUSUM	T^2	EWMA	CUSUM	T^2	EWMA	CUSUM	T^2
Gaussian Kernel	0.9091	0.8182	0.5455	0.6667	0.6667	0.5	0.5455	0.9545	0.7727
P-Splines	0.9091	0.8182	0.5455	0.6667	0.6667	0.5	0.5455	0.9545	0.7727
Haar Wavelets	0.8182	0.8182	0.0909	n/a	n/a	n/a	0.5455	0.9545	0.7727

Table 2.5: Probability of Signal Results for control charts using semiparametric methods

When using the semiparametric approach on linear data, we recommend the use of the P-spline to calculate the nonparametric $\hat{y}$ values and the Haar wavelet on nonlinear data. In table 2.4 the P-spline gave the lowest MSE for the linear data compared to its other nonparametric counterparts and the Haar wavelet gave the lowest for the nonlinear data. Here we still recommend the use of the EWMA control chart for our linear data since it resulted in the larger *PoS* for the NASA data and remained consistent for the NIST data for drift shift detection but recommend the CUSUM control chart for the nonlinear data for the same reasons (table 2.5).

2.6. Conclusion and Recommendations

This paper detailed the foundation of statistical process control and the use of control charts in profile monitoring using a linear and nonlinear model. We focused on the commonly used control charts for univariate and multivariate data: CUSUM, EWMA, mCUSUM, mEWMA, and T^2. It is common that the distribution of parameters of a dataset is unknown, so we introduced three nonparametric techniques along with a semiparametric technique for when there is some known information about the data. We conclude on the performance of each chart and provide recommendations for their use under parametric, nonparametric, and semiparametric conditions. Note, all recommendations are based on results obtained from the data sets explored in this text which are considered small in relation to other applications. They should not be generalized onto data sets with a non-similar structure.

We recommend the P-spline method with a linear model on univariate data; it gave the best estimates for the nonparametric and semiparametric approaches. This may be in part to the lack of curvature in the data used, but it still gave the best results for the MSE. The P-splines were easy to implement, and the only notable drawback is the high bias or large variance potential for bad knot selection. As for nonlinear univariate data, we recommend the Haar wavelet; it

gave the best estimates for the nonparametric and semiparametric approaches. They were also easy to implement and its only drawbacks are the repercussions of choosing an inappropriate resolution.

In sections 2.4.4 and 2.5.1, the CUSUM and EWMA charts gave comparable results for PoS. Both are memory-based charts which track the small/gradual changes that occur between profiles. Each is simple to implement, and we recommend that either chart be used when tracking drift shifts in data. We observed that the CUSUM chart performed best for our nonlinear data while the EWMA chart performed best for our linear data. For the memoryless charts, we conducted our comparison using the Hotelling's T^2 charts seen in section 2.3.2 along with the basic T^2 in sections 2.4.4 and 2.5.1. The T^2_{MVE} and T^2_{MCD} proved to be robust, but more computationally extensive to generate than the traditional T^2 charts. Since we did not have a clear understanding on the in-control/out-of-control profiles of the data used, it is difficult to conclude which of the five methods performed best.

All data, their descriptions, and additional materials used can be found via the following link https://github.com/cjones808/Practitioners-Guide-Datasets.git.

Bibliography

Woodall, W. H., & Montgomery, D. C. (2014). Some current directions in the theory and application of statistical process monitoring. *Journal of Quality Technology*, *46*(1), 78–94.

Woodall, W. H. (2007). Current research on profile monitoring. *Producao*, *17*(3), 420–425.

Abdel-Salam, A.-S. G., Birch, J. B., & Jensen, W. A. (2013). A semiparametric mixed model approach to phase i profile monitoring. *Quality and Reliability Engineering International*, *29*(4), 555–569.

Montgomery, D. C. (1990). *Introduction to statistical quality control* (Second). John Wiley & Sons.

Montgomery, D. (2013). *Introduction to statistical quality control* (Seventh). John Wiley & Sons.

Shewhart, W. A. (1926). Quality control charts. *The Bell System Technical Journal*, *5*(4), 593–603.

Page, E. S. (1954). Continuous inspection schemes. *Biometrika*, *41*(1/2), 100–115.

Page, E. (1961). Cumulative sum charts. *Technometrics*, *3*(1), 1–9.

Roberts, S. (1959). Control chart tests based on geometric moving averages. *Technometrics*, *1*(3), 239–250.

Hotelling, H. et al. A generalized t test and measure of multivariate dispersion. In: *Proceedings of the second berkeley symposium on mathematical statistics and probability*. The Regents of the University of California. 1951.

Woodall, W. H., & Ncube, M. M. (1985). Multivariate cusum quality-control procedures. *Technometrics*, *27*(3), 285–292.

Lowry, C. A., Woodall, W. H., Champ, C. W., & Rigdon, S. E. (1992). A multivariate exponentially weighted moving average control chart. *Technometrics*, *34*(1), 46–53.

Pignatiello Jr, J. J., & Runger, G. C. (1990). Comparisons of multivariate cusum charts. *Journal of quality technology*, *22*(3), 173–186.

Rousseeuw, P. J. (1984). Least median of squares regression. *Journal of the American statistical association*, *79*(388), 871–880.

Vargas, N. J. A. (2003). Robust estimation in multivariate control charts for individual observations. *Journal of Quality Technology*, *35*(4), 367–376.

Rousseeuw, P. J., & Van Zomeren, B. C. (1990). Unmasking multivariate outliers and leverage points. *Journal of the American Statistical association*, *85*(411), 633–639.

Van Aelst, S., & Rousseeuw, P. (2009). Minimum volume ellipsoid. *Wiley Interdisciplinary Reviews: Computational Statistics*, *1*(1), 71–82.

Woodruff, D. L., & Rocke, D. M. (1993). Heuristic search algorithms for the minimum volume ellipsoid. *Journal of Computational and Graphical Statistics*, *2*(1), 69–95.

Butler, R., Davies, P., Jhun, M et al. (1993). Asymptotics for the minimum covariance determinant estimator. *The Annals of Statistics*, *21*(3), 1385–1400.

Davies, L. et al. (1992). The asymptotics of rousseeuw's minimum volume ellipsoid estimator. *The Annals of Statistics*, *20*(4), 1828–1843.

Rousseeuw, P. J., & Driessen, K. V. (1999). A fast algorithm for the minimum covariance determinant estimator. *Technometrics*, *41*(3), 212–223.

Hardin, J., & Rocke, D. M. (2004). Outlier detection in the multiple cluster setting using the minimum covariance determinant estimator. *Computational Statistics & Data Analysis*, *44*(4), 625–638.

Mahmoud, M. A., Parker, P. A., Woodall, W. H., & Hawkins, D. M. (2007). A change point method for linear profile data. *Quality and Reliability Engineering International*, *23*(2), 247–268.

Parker, P, Morton, M, Draper, N, & Line, W. A single-vector force calibration method featuring the modern design of experiments. In: *39th aerospace sciences meeting and exhibit*. 2001, 170.

Croarkin, C., Tobias, P., Filliben, J., Hembree, B., Guthrie, W. et al. (2006). Nist/sematech e-handbook of statistical methods. *NIST/SEMATECH, July. Available online: http://www. itl. nist. gov/div898/handbook*.

Williams, J. D. (2011). Parametric nonlinear profiles. *Statistical Analysis of Profile Monitoring*, 129–156.

Gomaa, A.-S., & Birch, J. B. (2019). A semiparametric nonlinear mixed model approach to phase i profile monitoring. *Communications in Statistics-Simulation and Computation*, *48*(6), 1677–1693.

Williams, J., Birch, J. B., Woodall, W. H., & Ferry, N. M. (2007). Statistical monitoring of heteroscedastic dose-response profiles from high-throughput screening. *Journal of Agricultural, Biological, and Environmental Statistics*, *12*(2), 216.

Jensen, W. A., & Birch, J. B. (2009). Profile monitoring via nonlinear mixed models. *Journal of Quality Technology*, *41*(1), 18–34.

Nadaraya, E. A. (1964). On estimating regression. *Theory of Probability & Its Applications*, *9*(1), 141–142.

Watson, G. S. (1964). Smooth regression analysis. *Sankhyā: The Indian Journal of Statistics, Series A*, 359–372.

Bierens, H. J. (1988). The nadaraya-watson kernel regression function estimator.

Mays, J. E., Birch, J. B., & Alden Starnes, B. (2001). Model robust regression: Combining parametric, nonparametric, and semiparametric methods. *Journal of Nonparametric Statistics*, *13*(2), 245–277.

Burden, R. L., & Faires, J. D. (2010). *Numerical analysis (9th)*. Brooks/Cole.

Kauermann, G., Claeskens, G., & Opsomer, J. D. (2009). Bootstrapping for penalized spline regression. *Journal of Computational and Graphical Statistics*, *18*(1), 126–146.

Ruppert, D., Wand, M. P., & Carroll, R. J. (2009). Semiparametric regression during 2003–2007. *Electronic journal of statistics*, *3*, 1193.

Mosteller, F., & Wallace, D. L. (1963). Inference in an authorship problem: A comparative study of discrimination methods applied to the authorship of the disputed federalist papers. *Journal of the American Statistical Association*, *58*(302), 275–309.

Stone, M. (1974). Cross-validatory choice and assessment of statistical predictions. *Journal of the royal statistical society. Series B (Methodological)*, 111–147.

Geisser, S. (2017). *Predictive inference*. Routledge.

Friedman, J. H., & Silverman, B. W. (1989). Flexible parsimonious smoothing and additive modeling. *Technometrics*, *31*(1), 3–21.

Breiman, L. (1993). Fitting additive models to regression data. *Computational statistics and data analysis*, *15*, 13–46.

Denison, D., Mallick, B., & Smith, A. (1998). Automatic bayesian curve fitting. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *60*(2), 333–350.

Molinari, N., Durand, J.-F., & Sabatier, R. (2004). Bounded optimal knots for regression splines. *Computational statistics & data analysis*, *45*(2), 159–178.

Lolive, D., Barbot, N., & Boeffard, O. Melodic contour estimation with b-spline models using a mdl criterion. In: *Proceedings of the 11th international conference on speech and computer (specom)*. 2006, 333–338.

Piri, S., Abdel-Salam, A.-S. G., & Boone, E. L. (2019). A wavelet approach for profile monitoring of poisson distribution with application. *Communications in Statisitics-Simiulation and Computation*, 1–12.

O'Sullivan, F. (1986). A statistical perspective on ill-posed inverse problems. *Statistical science*, 502–518.

Abdel-Salam, A.-S. G. (2009). *Profile monitoring with fixed and random effects using nonparametric and semiparametric methods* (Doctoral dissertation). Virginia Tech.

De Boor, C., De Boor, C., Mathématicien, E.-U., De Boor, C., & De Boor, C. (1978). *A practical guide to splines* (Vol. 27). Springer-Verlag New York.

Eilers, P. H., & Marx, B. D. (1996). Flexible smoothing with b-splines and penalties. *Statistical science*, 89–102.

Eilers, P. H., Marx, B. D., & Durbán, M. (2015). Twenty years of p-splines. *SORT: statistics and operations research transactions*, *39*(2), 0149–186.

Reis, M. S., & Saraiva, P. M. (2006). Multiscale statistical process control of paper surface profiles. *Quality Technology & Quantitative Management*, *3*(3), 263–281.

Jeong, M. K., Lu, J.-C., & Wang, N. (2006). Wavelet-based spc procedure for complicated functional data. *International Journal of Production Research*, *44*(4), 729–744.

Zou, C., Zhou, C., Wang, Z., & Tsung, F. (2007). A self-starting control chart for linear profiles. *Journal of Quality Technology*, *39*(4), 364–375.

Chicken, E., Pignatiello Jr, J. J., & Simpson, J. R. (2009). Statistical process monitoring of nonlinear profiles using wavelets. *Journal of Quality Technology*, *41*(2), 198–212.

Chang, S. I., & Yadama, S. (2010). Statistical process control for monitoring non-linear profiles using wavelet filtering and b-spline approximation. *International Journal of Production Research*, *48*(4), 1049–1068.

Nikoo, M., & Noorossana, R. (2013). Phase ii monitoring of nonlinear profile variance using wavelet. *Quality and Reliability Engineering International*, *29*(7), 1081–1089.

Daubechies, I. (1992). *Ten lectures on wavelets* (Vol. 61). Siam.

Ogden, T. (2012). *Essential wavelets for statistical applications and data analysis*. Springer Science & Business Media.

Burrus, C. S., Gopinath, R, & Guo, H. (1998). *A primer introduction to wavelets and wavelet transforms* (Vol. 3).

Olkin, I., & Spiegelman, C. H. (1987). A semiparametric approach to density estimation. *Journal of the American Statistical Association, 82*(399), 858–865.

Ke, C., & Wang, Y. (2001). Semiparametric nonlinear mixed-effects models and their applications. *Journal of the American Statistical Association, 96*(456), 1272–1298.

Einsporn, R. L. (1987). *Hatlink: A link between least squares regression and nonparametric curve estimation* (Doctoral dissertation). Virginia Polytechnic Institute and State University.

Einsporn, R. L., & Birch, J. B. (1993). *Improving predictions by combining least squares and kernel regression prediction weights* [93-11:21]. Department of Statistics, Virginia Polytechnic Institute; State University.

Burman, P., & Chaudhuri, P. On a hybrid approach to parametric and nonparametric regression. In: *Nonparametric statistical methods and related topics: A festschrift in honor of professor pk bhattacharya on the occasion of his 80th birthday*. World Scientific, 2012, pp. 233–256.

Mays, J. E., Birch, J. B., & Einsporn, R. L. (2000). An overview of model-robust regression. *Journal of statistical computation and simulation, 66*(1), 79–100.

Nottingham, Q. J., & Birch, J. B. (2000). A semiparametric approach to analysing dose–response data. *Statistics in medicine, 19*(3), 389–404.

Waterman, M. J., Birch, J. B., & Abdel-Salam, A.-S. G. (2015). Several nonparametric and semiparametric approaches to linear mixed model regression. *Journal of Statistical Computation and Simulation, 85*(5), 956–977.

Siddiqui, Z., & Abdel-Salam, A.-S. G. (2019). A semiparametric profile monitoring via residuals. *Quality and Reliability Engineering International.*

Chapter 3

Generalized Bayesian CUSUM Charts using Different Loss Functions

Abstract

We propose a Bayesian approach to the cumulative sum (CUSUM) control chart to perform under different distributions without the need of a data transformation. Different loss functions are considered to inform the framework of the charts for the Normal and Poisson conjugate cases. Using a comprehensive simulation study, we assess the method via a sensitivity analysis for the control chart decision parameters, shift sizes, and distribution hyper-parameters; where performance measurements are average run length (ARL), standard deviation of the run length (SDRL), average time to signal (ATS), and standard deviation of time to signal (SDTS). For CUSUM literature, a data transformation is often required when using non-Guassian data, so we also consider a comparative study with the classical CUSUM chart. To showcase their efficacy on over-dispersed count data, we model a count series of respiratory disease related hospitalizations in São Paulo, Brazil and implement our Bayesian charts.

Keywords: Profile Monitoring; Bayesian; Loss Functions, CUSUM; Sensitivity Analysis; Quality Control.

3.1. Introduction

In statistical process monitoring (SPM), control charts are a vital visual tool used to monitor a process and alert of any discrepancies. Many control charts have been developed to encompass the different processes to monitor, mainly falling into two categories: memory-less and memory-based. A memory-less chart does not take into consideration previous observations and are best used to detect large/sudden changes. Memory-based charts compute the current statistic using the previous' and are ideal for detecting small/gradual changes. The Shewhart control chart, first outlined in Shewhart, 1926, is a frequently used memory-less chart while the cumulative sum (CUSUM) (Page, 1954) and exponentially weighted moving average (EWMA) (Roberts, 1959) charts are well-known memory-based charts.

In Riaz et al., 2017, the EWMA control chart is considered under a Bayesian approach. They conduct a comparison analysis on control limits found using the classical approach and the Bayesian approach under a symmetric loss function and 2 asymmetric loss functions: squared error loss function (SELF), precautionary loss function (PLF), and Linex loss function (LLF). In their article they use a normal conjugate prior on data that is normally distributed and for each of the loss functions, a sensitivity analysis is done for the choice of hyper-parameter and its effect on the performance of the Bayesian EWMA chart. The measurement tools used for their comparative purposes are the average run length (ARL) and standard deviation run length (SDRL).

In this paper, we conduct analysis of a Bayesian CUSUM chart via ARL, SDRL, average time to signal (ATS), and standard deviation of time to signal (SDTS) using a Normal likelihood and

prior and a Poisson likelihood and Gamma prior distribution under the same loss functions. We first conduct analysis of the Bayesian CUSUM chart under the Normal conjugate case. Then, we complete an analysis of the control chart under a Poisson likelihood and Gamma prior. Lastly, we conduct a real data analysis of count data to show application of our method. Our objective is to assess the choice of loss function in relation to distribution and deduce generality of the Bayesian CUSUM control chart.

3.2. Loss Functions

A loss function is represented as $L(\theta,\hat{\theta})$ and is a function of the parameter of interest, θ, and an estimate of the parameter, $\hat{\theta}$. Loss functions measure how bad the current parameter estimate is, and typically the larger the loss the worse the parameter estimate. In decision theory it takes on a slightly different role, which is to attain the best estimator of the parameter. In this study we consider three different loss functions: Squared Error, Precautionary, and Linex. We construct a Bayesian CUSUM chart under the different loss functions considering the best estimators which is found using equation 4.3.

$$\hat{\theta}^* = \min_{\hat{\theta}} E_\theta[L(\hat{\theta},\theta] \tag{3.1}$$

where, $\hat{\theta}^*$ is the estimator which minimizes the expected loss

The Squared Error Loss Function (SELF) is a commonly used loss function because its simplicity. It is considered symmetric because it assigns equal weight to both positive and negative error and defined as $L(\theta,\hat{\theta}) = (\theta-\hat{\theta})^2$ with a Bayes estimator of $\hat{\theta^*} = E[\theta|x]$. The Precautionary Loss Function (PLF) is an asymmetric loss function as it weights the error in the positive direction differently than in the negative direction. It was introduced in Norstrom, 1996 in order to

prevent overestimation of the parameter and typically produces better estimators for low failure rate problems. Its form is $L(\theta,\hat{\theta}) = \frac{(\theta-\hat{\theta})^2}{\hat{\theta}}$ with its Bayes estimator given as $\hat{\theta^*} = \sqrt{E[\theta^2|x]}$. The last loss function that we consider here is another asymmetric function introduced in Varian, 1975 as the Linex Loss Function (LLF). It is asymmetric as it signifies overestimation rather than underestimation, and is often used with estimating the location parameter. It is defined as $L(\theta,\hat{\theta}) = (e^{c(\hat{\theta}-\theta)} - c(\hat{\theta}-\theta) - 1)$ with a Bayes estimator of $\hat{\theta^*} = -\frac{1}{c} lnE[e^{-c\theta}]$ (Zellner, 1986), where c is a constant such that if $c > 0$ overestimation has more significance than underestimation and if $c \to 0$ it approaches symmetry.

3.3. Bayesian Framework

We define the necessary pieces for our Bayesian approach in equations 4.1 and 4.2.

$$p(\theta|\mathbf{X}) = \frac{\mathbf{p}(\mathbf{X}|\theta)\mathbf{p}(\theta)}{\mathbf{p}(\mathbf{X})} \tag{3.2}$$

	Definition
θ	Parameter(s) of interest
$\mathbf{X}$	Data
$P(\theta)$	Prior distribution
$P(\mathbf{X})$	Marginal distribution data
$P(\mathbf{X}\|\theta)$	Likelihood function

Table 3.1: Bayesian Inference Definitions

$$p(y|\mathbf{X}) = \int p(\mathbf{y}|\theta)\mathbf{p}(\theta|\mathbf{X})\mathbf{d}\theta \tag{3.3}$$

where, $p(\mathbf{y}|\theta)$ is the likelihood function for the future data

3.3.1. Normal Conjugate

In table 3.2, we provide the Bayes estimators for each of the loss functions. Within table 3.2 and 3.3, σ_0^2 represents the prior variance and σ^2 represents the known variance of the likelihood distribution. The calculation of the variances for the posterior, posterior predictive, and Bayes estimator of the posterior predictive distributions remains the same under each loss function as what is seen in Riaz et al., 2017. For both tables, n is the sample size, $\bar{x}$ is the sample mean, μ_0 is the prior mean, σ_0^2 is the prior variance, σ^2 is the known variance, and c is the Linex loss function symmetry constant.

Loss Function	Bayes Estimator
Squared Error	$\mu_{SELF,NN} = \frac{n\bar{x}\sigma_0^2+\sigma^2\mu_0}{\sigma^2+n\sigma_0^2}$
Precautionary	$\mu_{PLF,NN} = \sqrt{\frac{(\sigma^4+\sigma^2\sigma_0^2(n+1))(\sigma^2+n\sigma_0^2)+(n\bar{x}\sigma_0^2+\sigma^2\mu_0)^2}{(\sigma^2+n\sigma_0^2)^2}}$
Linex	$\mu_{LLF,NN} = \frac{n\bar{x}\sigma_0^2+\sigma^2\mu_0}{\sigma^2+n\sigma_0^2} - \frac{c}{2}\left(\sigma^2 + \frac{\sigma^2\sigma_0^2}{\sigma^2+n\sigma_0^2}\right)$

Table 3.2: Bayes Estimators for Loss Functions

Distribution	Variance
Posterior	$\sigma_n^2 = \frac{\sigma^2\sigma_0^2}{\sigma^2+n\sigma_0^2}$
Posterior Predictive	$\sigma_{ppd}^2 = \sigma^2 + \sigma_n^2$
Bayes Estimator of Posterior Predictive	$\sigma_{\tilde{Y}}^2 = \frac{\sigma^2}{n} + \sigma_n^2$

Table 3.3: Variances Based on Distribution

3.3.2. Poisson Conjugate

The Gamma distribution is the known conjugate for the Poisson likelihood function. The posterior distribution under each of the loss functions is Gamma with shape parameter $n\bar{x} + \alpha$ and inverse scale parameter $n + \beta$ ($Gamma(n\bar{x} + \alpha, n + \beta)$). The posterior predictive and Bayes estimator of the posterior predictive distributions are derived to be Negative Binomial (see 3.10) also with shape parameter $n\bar{x} + \alpha$ and inverse scale parameter $n + \beta$ ($NegBin(n\bar{x} + \alpha, n + \beta)$). Similarly as stated in section 3.3.1 the variances for each of the distributions in table 4.3 remains the same under each loss function. In the table, α and β are the shape and inverse scale parameters respectively, with sample size n, and sample mean $\bar{x}$. Prior information of the mean and variance are used to solve for the α and β values. The mean for each of the distributions is what differs based on the Bayes' estimator equations for the given loss functions and are shown in table 4.2. For both tables n is the sample size, $\bar{x}$ is the sample mean, α is the shape parameter, β is the inverse scale parameter, and c is the Linex loss function symmetry constant.

Loss Function	Bayes Estimator
Squared Error	$\mu_{SELF,PG} = \frac{n\bar{x}+\alpha}{n+\beta}$
Precautionary	$\mu_{PLF,PG} = \frac{\sqrt{(n\bar{x}+\alpha)(n\bar{x}+\alpha)^2}}{n+\beta}$
Linex	$\mu_{LLF,PG} = -\frac{1}{c} ln\left[\frac{\beta^\alpha}{(c+\beta)^\alpha}\right]$

Table 3.4: Bayes Estimators for Loss Functions

Distribution	Variance
Posterior	$\sigma^2_{PD} = \frac{n\bar{x}+\alpha}{(n+\beta)^2}$
Posterior Predictive	$\sigma^2_{PPD} = \frac{n\bar{x}+\alpha}{(n+\beta)^2}(n+\beta+1)$
Bayes Estimator of Posterior Predictive	$\sigma^2_{\bar{Y}} = \frac{n\bar{x}+\alpha}{n(n+\beta)^2}(n+\beta+1)$

Table 3.5: Variances Based on Distribution

3.4. Control Chart Schematics

The control limits for the Bayesian CUSUM control charts follow the same form as the standard CUSUM chart. The mean and standard deviation used for the chart limits are based on the mean and standard deviation obtained using the Bayes estimator posterior predictive distribution under each loss function. Note that the plotting statistics in equation 3.4 also follow the standard form for the CUSUM control chart.

CUSUM Posterior Predictive Control Limits:

$$\begin{aligned} UCL/LCL &= \pm h * \sigma_{\bar{Y}} \\ CL &= \mu_{LF} \\ c_i &= \max[0, c_{i-1} + (\bar{y}|x) - \mu_{LF}] \end{aligned} \tag{3.4}$$

where, μ_{LF} and $\sigma_{\bar{Y}}$ are the mean and standard deviation from the Bayes estimator posterior predictive distribution respective to the loss function used, $(\bar{y}|x)$ is the Bayes estimator for the set of predicted data, and h determines the control limits. For the Poisson conjugate case, we only consider upper limits for our control charts. This is because we are concerned with an unusual rise in counts rather than a drop in counts.

3.5. Simulations and Results

In this section, sensitivity analyses of the hyper-parameters and sample size are conducted for the Bayesian CUSUM control chart for both the Normal and Poisson conjugate cases. The charts for the Normal and Poisson cases are designed to obtain $ARL_0 = 370$ and $ARL_0 = 500$ respectively and are assessed for different shift sizes ($\delta = 0$ to 2.5 by 0.25). For the simulations under both analysis, $m = 10,000$ iterations were completed to calculate the performance measurements.

3.5.1. Normal Conjugate

The Guassian conjugate-based CUSUM chart uses different values of h for each loss function to obtain the desired $ARL_0 = 370$. When generating the initial data we use a standard normal distribution ($N(0,1)$), where the out-of-control mean is $\mu_1 = \mu + \delta$ and choices for the hyper-parameters are $\mu_0 = \{5,10,15\}$ and $\sigma = \{2,4,6\}$. For the sample size analysis we choose $n = \{5,10,20,30\}$.

	$\mu_0 = 5, \sigma_0 = 2$				$\mu_0 = 10, \sigma_0 = 4$				$\mu_0 = 15, \sigma_0 = 6$			
Shifts	ARL	SDRL	ATS	SDTS	ARL	SDRL	ATS	SDTS	ARL	SDRL	ATS	SDTS
						SELF						
0	382.56	314.006	2.10E-07	1.85E-06	377.585	310.835	2.06E-07	1.84E-06	381.314	311.04	1.96E-07	1.80E-06
0.25	25.1489	6.73527	2.18E-07	1.89E-06	25.112	6.68524	2.29E-07	1.98E-06	25.226	6.69327	1.67E-07	1.66E-06
0.5	12.2902	2.39349	1.98E-07	1.80E-06	12.3517	2.42792	2.09E-07	1.85E-06	12.2861	2.36657	1.97E-07	1.80E-06
0.75	7.9762	1.32674	1.79E-07	1.71E-06	7.9855	1.33375	2.10E-07	1.85E-06	7.9705	1.31796	2.13E-07	1.87E-06
1	5.8455	0.883872	2.29E-07	1.94E-06	5.8419	0.878012	2.12E-07	1.86E-06	5.8669	0.897321	2.18E-07	1.89E-06
1.25	4.5644	0.663365	1.90E-07	1.76E-06	4.5644	0.673092	1.94E-07	1.78E-06	4.5722	0.668421	2.18E-07	1.89E-06
1.5	3.7126	0.549	1.73E-07	1.68E-06	3.7255	0.555653	2.29E-07	3.63E-06	3.7143	0.549432	2.24E-07	1.92E-06
1.75	3.1044	0.411462	2.01E-07	1.81E-06	3.1065	0.412017	2.34E-07	1.96E-06	3.1091	0.405706	1.93E-07	1.78E-06
2	2.6992	0.468529	2.04E-07	1.83E-06	2.6991	0.469637	1.93E-07	1.78E-06	2.6859	0.473963	1.91E-07	1.77E-06
2.25	2.2077	0.405661	2.03E-07	1.82E-06	2.2124	0.409251	2.16E-07	1.97E-06	2.2192	0.413946	2.27E-07	1.93E-06
2.5	2.0136	0.143579	2.00E-07	1.81E-06	2.013	0.143635	2.15E-07	1.87E-06	2.0134	0.145672	1.96E-07	1.79E-06
						PLF						
0	386.091	317.271	1.73E-07	1.68E-06	383.807	308.288	2.36E-07	1.96E-06	382.062	309.744	2.19E-07	1.90E-06
0.25	25.0807	6.82654	1.90E-07	1.76E-06	25.2625	6.75559	2.30E-07	1.94E-06	25.3573	6.90383	2.06E-07	1.84E-06
0.5	12.293	2.43589	2.17E-07	1.88E-06	12.2928	2.40106	1.98E-07	1.80E-06	12.3003	2.41058	1.66E-07	1.66E-06
0.75	7.9805	1.33301	1.65E-07	1.65E-06	7.9887	1.33902	2.06E-07	1.88E-06	7.9786	1.31436	1.93E-07	1.78E-06
1	5.8491	0.897513	1.98E-07	1.80E-06	5.8279	0.899267	1.96E-07	1.80E-06	5.8588	0.884343	1.83E-07	1.73E-06
1.25	4.5609	0.6682	2.03E-07	1.83E-06	4.5705	0.667705	2.23E-07	1.91E-06	4.583	0.666417	1.65E-07	1.65E-06
1.5	3.7241	0.550617	2.23E-07	1.91E-06	3.7171	0.55106	1.90E-07	1.77E-06	3.7208	0.557716	1.81E-07	1.73E-06
1.75	3.1132	0.415194	1.95E-07	1.79E-06	3.1046	0.402068	1.89E-07	1.76E-06	3.1096	0.411081	2.07E-07	1.84E-06
2	2.6909	0.471123	2.23E-07	1.91E-06	2.7013	0.468058	2.26E-07	1.92E-06	2.6941	0.469601	1.80E-07	1.72E-06
2.25	2.2127	0.409706	1.83E-07	1.73E-06	2.2143	0.410336	1.96E-07	1.79E-06	2.2166	0.411928	1.92E-07	1.77E-06
2.5	2.0132	0.146375	2.00E-07	1.81E-06	2.0151	0.149238	1.99E-07	1.81E-06	2.016	0.144028	1.81E-07	1.72E-06
						LLF						
0	382.633	310.757	2.16E-07	1.88E-06	379.757	312.008	1.98E-07	1.80E-06	381.621	311.286	2.05E-07	1.84E-06
0.25	25.2814	6.8093	1.90E-07	1.77E-06	25.1894	6.71061	1.55E-07	1.60E-06	25.222	6.70407	1.66E-07	1.65E-06
0.5	12.2979	2.42684	2.14E-07	1.87E-06	12.2589	2.39922	1.81E-07	1.72E-06	12.2623	2.35586	1.86E-07	1.75E-06
0.75	7.9646	1.31482	2.00E-07	1.81E-06	7.9809	1.32058	1.82E-07	1.73E-06	7.9879	1.32399	2.06E-07	1.84E-06
1	5.8316	0.88828	1.88E-07	1.76E-06	5.8337	0.889294	1.83E-07	1.73E-06	5.8465	0.890471	2.02E-07	1.82E-06
1.25	4.5581	0.669645	2.14E-07	1.87E-06	4.5811	0.660926	1.46E-07	1.55E-06	4.5675	0.671896	1.85E-07	1.74E-06
1.5	3.723	0.555041	2.04E-07	1.83E-06	3.7261	0.549253	1.74E-07	1.69E-06	3.7283	0.551615	1.88E-07	1.76E-06
1.75	3.1041	0.404306	2.15E-07	1.88E-06	3.108	0.406615	1.96E-07	1.80E-06	3.1026	0.41361	1.81E-07	1.72E-06
2	2.6947	0.470204	2.07E-07	1.84E-06	2.6936	0.472355	1.81E-07	1.72E-06	2.6927	0.472934	2.34E-07	1.96E-06
2.25	2.2067	0.405679	2.04E-07	1.83E-06	2.2029	0.402159	1.41E-07	1.53E-06	2.2195	0.414391	1.86E-07	1.75E-06
2.5	2.0152	0.144114	2.43E-07	2.01E-06	2.0132	0.136476	1.64E-07	1.64E-06	2.016	0.144028	2.08E-07	1.85E-06

Table 3.6: ARL, SDRL, ATS, and SDTS Values for Bayesian CUSUM Chart Hyper-Parameter Sensitivity Analysis with Normal Conjugate

Table 3.6 shows simulation results from the Bayesian CUSUM control chart under a normal conjugate prior. The prior mean and variance were adjusted for sensitivity analysis of the hyper-parameters while applying shifts to the in-control mean (μ_0). For all of the loss functions, the initial shift in the mean returns a drastic drop in the ARL and SDRL. After the first shift, as the shifts increase the ARL and SDRL gradually decrease. This shows that a relatively small shift is detectable when using the Bayesian CUSUM control chart and detection is consistent over

several loss functions. Results also show that as you adjust the input for the hyper-parameters, this control chart performs consistently. ATS and SDTS values, in seconds, are not significantly different within each loss function, but overall the Linex loss function has values that are noticeably smaller than the squared error and precautionary loss functions.

		Shifts											
n	h	0	0.25	0.5	0.75	1	1.25	1.5	1.75	2	2.25	2.5	
							SELF						
5	14	375.399	37.1613	18.255	11.9493	8.8191	6.9268	5.6775	4.784	4.1443	3.6073	3.1807	ARL
		(308.228)	(11.9545)	(4.21535)	(2.3375)	(1.54802)	(1.11177)	(0.867118)	(0.717038)	(0.599731)	(0.551984)	(0.44525)	SDRL
		2.07E-07	2.18E-07	1.76E-07	2.06E-07	2.01E-07	2.46E-07	1.69E-07	2.16E-07	2.11E-07	2.11E-07	1.80E-07	ATS
		(1.86E-06)	(1.89E-06)	(1.70E-06)	(1.84E-06)	(1.81E-06)	(2.01E-06)	(1.67E-06)	(1.88E-06)	(1.86E-06)	(1.87E-06)	(1.72E-06)	SDTS
10	13.6	381.085	25.2921	12.3169	8.002	5.8504	4.583	3.7143	3.1071	2.6879	2.2137	2.015	ARL
		(313.336)	(6.77449)	(2.4378)	(1.34781)	(0.885223)	(0.662202)	(0.547244)	(0.410889)	(0.47444)	(0.410161)	(0.140624)	SDRL
		2.60E-07	1.67E-07	2.19E-07	1.97E-07	1.75E-07	1.89E-07	1.82E-07	1.97E-07	1.78E-07	1.78E-07	1.72E-07	ATS
		(2.06E-06)	(1.66E-06)	(1.89E-06)	(1.80E-06)	(1.69E-06)	(1.76E-06)	(1.73E-06)	(1.80E-06)	(1.71E-06)	(1.71E-06)	(1.68E-06)	SDTS
20	13	369.377	16.9652	8.1781	5.2669	3.809	2.9699	2.2954	2.0001	1.8121	1.2501	1.0119	ARL
		(299.4)	(3.82957)	(1.40384)	(0.79364)	(0.562245)	(0.379465)	(0.456441)	(0.126095)	(0.390632)	(0.43307)	(0.108436)	SDRL
		1.90E-07	1.57E-07	2.11E-07	2.14E-07	1.92E-07	2.07E-07	1.93E-07	2.03E-07	1.50E-07	1.91E-07	2.09E-07	ATS
		(1.76E-06)	(1.60E-06)	(1.86E-06)	(1.88E-06)	(1.77E-06)	(1.84E-06)	(1.78E-06)	(1.83E-06)	(1.57E-06)	(1.77E-06)	(1.85E-06)	SDTS
30	13	373.031	13.5429	6.4694	4.1425	2.9955	2.1861	1.9625	1.4297	1.0215	1	1	ARL
		(304.37)	(2.84098)	(1.02921)	(0.60116)	(0.386885)	(0.389187)	(0.193633)	(0.495033)	(0.145044)	(0)	(0)	SDRL
		2.33E-07	2.25E-07	2.08E-07	2.14E-07	2.09E-07	2.31E-07	2.59E-07	1.97E-07	2.14E-07	2.70E-07	2.24E-07	ATS
		(1.95E-06)	(1.92E-06)	(1.84E-06)	(1.87E-06)	(1.85E-06)	(1.95E-06)	(2.06E-06)	(1.79E-06)	(1.87E-06)	(2.10E-06)	(1.91E-06)	SDTS
							PLF						
5	8	375.604	37.207	18.2401	11.9825	8.7833	6.9291	5.6663	4.7933	4.1237	3.6066	3.1778	ARL
		(310.609)	(12.1227)	(4.24745)	(2.34252)	(1.53191)	(1.10873)	(0.860781)	(0.703829)	(0.595649)	(0.554289)	(0.440213)	SDRL
		2.09E-07	1.94E-07	2.10E-07	2.01E-07	1.70E-07	1.64E-07	1.56E-07	1.96E-07	1.87E-07	1.83E-07	2.08E-07	ATS
		(1.85E-06)	(1.79E-06)	(1.85E-06)	(1.82E-06)	(1.67E-06)	(1.64E-06)	(1.60E-06)	(1.79E-06)	(1.75E-06)	(1.73E-06)	(1.85E-06)	SDTS
10	5.8	377.371	25.1675	12.2698	7.9768	5.8525	4.5655	3.7289	3.1078	2.6988	2.2135	2.0124	ARL
		(303.993)	(6.62654)	(2.41851)	(1.33726)	(0.892381)	(0.662955)	(0.557499)	(0.41398)	(0.469338)	(0.410022)	(0.138008)	SDRL
		1.88E-07	2.17E-07	2.01E-07	2.05E-07	1.87E-07	2.10E-07	1.84E-07	2.08E-07	1.65E-07	1.84E-07	1.93E-07	ATS
		(1.75E-06)	(1.90E-06)	(1.82E-06)	(1.83E-06)	(1.75E-06)	(1.85E-06)	(1.74E-06)	(1.85E-06)	(1.65E-06)	(1.74E-06)	(1.78E-06)	SDTS
20	4	368.62	16.9351	8.1502	5.2623	3.8054	2.9668	2.2932	1.9982	1.8066	1.2553	1.0122	ARL
		(297.546)	(3.80162)	(1.39422)	(0.797307)	(0.566684)	(0.388713)	(0.455449)	(0.131136)	(0.394964)	(0.43603)	(0.109778)	SDRL
		1.88E-07	1.96E-07	2.04E-07	1.98E-07	3.03E-07	1.76E-07	2.16E-07	2.10E-07	1.97E-07	2.08E-07	1.71E-07	ATS
		(1.76E-06)	(1.79E-06)	(1.83E-06)	(1.80E-06)	(2.22E-06)	(1.70E-06)	(1.88E-06)	(1.86E-06)	(1.80E-06)	(1.85E-06)	(1.68E-06)	SDTS
30	3.41	367.25	13.5895	6.5236	4.1415	3.0024	2.185	1.9616	1.447	1.0208	1.0001	1	ARL
		(295.354)	(2.83235)	(1.0291)	(0.603389)	(0.3914)	(0.389326)	(0.19423)	(0.497183)	(0.142714)	(0.0099995)	(0)	SDRL
		1.80E-07	2.03E-07	1.80E-07	2.09E-07	2.17E-07	2.38E-07	1.98E-07	1.91E-07	2.22E-07	1.98E-07	1.88E-07	ATS
		(1.72E-06)	(1.82E-06)	(1.72E-06)	(1.85E-06)	(1.89E-06)	(1.97E-06)	(1.80E-06)	(1.77E-06)	(1.90E-06)	(1.80E-06)	(1.76E-06)	SDTS
							LLF						
5	9	376.357	37.1803	18.2026	11.9679	8.8241	6.9209	5.6815	4.8087	4.1361	3.6015	3.1848	ARL
		(314.896)	(12.0124)	(4.27094)	(2.32152)	(1.53217)	(1.10256)	(0.868595)	(0.721183)	(0.592095)	(0.547264)	(0.444352)	SDRL
		1.52E-07	1.56E-07	2.10E-07	2.15E-07	1.97E-07	1.75E-07	1.75E-07	1.93E-07	1.75E-07	1.71E-07	1.53E-07	ATS
		(1.58E-06)	(1.61E-06)	(1.86E-06)	(1.88E-06)	(1.80E-06)	(1.69E-06)	(1.69E-06)	(1.78E-06)	(1.69E-06)	(1.68E-06)	(1.59E-06)	SDTS
10	7.7	378.459	25.3499	12.2567	7.9641	5.8521	4.5691	3.7212	3.1041	2.6955	2.2024	2.0124	ARL
		(312.249)	(6.70681)	(2.37525)	(1.33799)	(0.884435)	(0.670392)	(0.551245)	(0.409955)	(0.469446)	(0.402038)	(0.129793)	SDRL
		1.76E-07	1.60E-07	2.19E-07	1.86E-07	1.81E-07	1.81E-07	1.89E-07	1.84E-07	2.19E-07	1.90E-07	1.75E-07	ATS
		(1.70E-06)	(1.62E-06)	(1.90E-06)	(1.75E-06)	(1.73E-06)	(1.72E-06)	(1.76E-06)	(1.74E-06)	(1.89E-06)	(1.77E-06)	(1.70E-06)	SDTS
20	6.8	371.933	16.9739	8.1679	5.2633	3.8142	2.9723	2.2925	2.0014	1.8113	1.2497	1.014	ARL
		(310.142)	(3.86676)	(1.3851)	(0.791564)	(0.560605)	(0.382011)	(0.45513)	(0.116611)	(0.39127)	(0.432839)	(0.11749)	SDRL
		1.80E-07	2.58E-07	1.53E-07	1.83E-07	1.98E-07	1.94E-07	2.01E-07	1.69E-07	2.04E-07	1.85E-07	1.90E-07	ATS
		(1.72E-06)	(2.05E-06)	(1.58E-06)	(1.73E-06)	(1.80E-06)	(1.79E-06)	(1.82E-06)	(1.67E-06)	(1.83E-06)	(1.74E-06)	(1.76E-06)	SDTS
30	6.7	368.33	13.5945	6.5128	4.1572	3.0067	2.1906	1.967	1.4585	1.0224	1.0001	0.9999	ARL
		(304.737)	(2.84954)	(1.02442)	(0.604722)	(0.382172)	(0.392774)	(0.182513)	(0.498275)	(0.147981)	(0.0099995)	(0.0099995)	SDRL
		1.88E-07	1.68E-07	2.27E-07	1.63E-07	1.71E-07	1.58E-07	1.92E-07	1.70E-07	1.65E-07	1.69E-07	2.31E-07	ATS
		(1.76E-06)	(1.67E-06)	(1.92E-06)	(1.64E-06)	(1.68E-06)	(1.61E-06)	(1.77E-06)	(1.67E-06)	(1.65E-06)	(1.67E-06)	(1.95E-06)	SDTS

Table 3.7: ARL, SDRL, ATS, and SDTS Values for Bayesian CUSUM Chart Sample Size Sensitivity Analysis with Normal Conjugate

Similar trends about the Bayesian CUSUM chart that were noticed in table 3.6 can also be seen in table 3.7. By decreasing the h value as the sample size increases, we obtain ARL_0 values around 370. This is intuitive since increasing sample size typically delays detection, thus shrinking the bounds for the control limits allows for timely detection. As before, we notice that there is an immediate drop in ARL when the initial mean shift is applied and a gradual decrease after for all loss functions. We notice that as we increase n while applying the mean shifts, the drop in the ARL is larger in the larger samples. This shows that detection ability grows with sample size. ATS and SDTS are also effected by increasing sample size; detection time decreases as sample size increases.

3.5.2. Poisson Conjugate

Simulation results from tables 3.8 and 3.9 show performance measurements for the Bayesian CUSUM control chart under a Poisson conjugate prior. We designed a Bayesian CUSUM chart to obtain an in-control ARL of 500 and assess their performances with imposed shifts defined as $\mu_1 = \mu_0 * (1+\delta)$. We use values $\mu_0 = [10,15,20]$ and $\sigma_0 = [4,6,8]$ to solve for our values of α_0 and β_0 to obtain our prior information. These values are $\alpha = [16,36,64]$ and $\beta = \left[\frac{5}{8}, \frac{5}{12}, \frac{5}{16}\right]$ as calculated in section 3.8.1.2. Our initial data is generated from a Poisson distribution with parameter $\lambda = 25$.

	$\alpha_0 = 16, \beta_0 = \frac{5}{8}$				$\alpha_0 = 36, \beta_0 = \frac{5}{12}$				$\alpha_0 = 64, \beta_0 = \frac{5}{16}$			
Shifts	ARL	SDRL	ATS	SDTS	ARL	SDRL	ATS	SDTS	ARL	SDRL	ATS	SDTS
						SELF						
0	520.043	520.771	2.28E-07	5.00E-06	523.698	523.309	2.26E-07	6.75E-06	502.157	500.218	2.40E-07	1.98E-06
0.25	26.1833	25.8757	2.17E-07	7.46E-06	28.8763	28.5943	2.31E-07	6.72E-06	28.7685	28.5028	2.06E-07	1.84E-06
0.5	6.33065	5.91093	1.66E-07	6.53E-06	5.68287	5.46734	2.17E-07	6.93E-06	6.17329	5.91276	2.16E-07	1.88E-06
0.75	2.64614	2.48548	1.86E-07	6.94E-06	2.56805	2.46631	2.47E-07	5.60E-06	2.58994	2.51808	2.07E-07	1.85E-06
1	1.46543	1.48035	2.01E-07	7.18E-06	1.38482	1.45086	2.28E-07	6.51E-06	1.42274	1.48283	2.11E-07	1.86E-06
1.25	0.86422	1.02001	2.01E-07	7.22E-06	0.84758	1.01721	2.18E-07	4.05E-06	0.83408	1.02107	1.99E-07	1.80E-06
1.5	0.59422	0.796293	2.00E-07	7.18E-06	0.57116	0.784395	2.23E-07	6.93E-06	0.57079	0.78851	1.94E-07	1.78E-06
1.75	0.4083	0.64004	2.07E-07	7.27E-06	0.41768	0.64444	2.17E-07	5.96E-06	0.38137	0.625961	2.19E-07	1.89E-06
2	0.29026	0.536423	1.96E-07	7.13E-06	0.28799	0.534052	1.83E-07	6.60E-06	0.27276	0.522496	2.12E-07	1.86E-06
2.25	0.19789	0.438052	2.09E-07	7.36E-06	0.20633	0.447435	1.79E-07	6.79E-06	0.19119	0.433701	2.02E-07	1.82E-06
2.5	0.15763	0.389721	2.07E-07	7.27E-06	0.14413	0.373439	2.19E-07	5.75E-06	0.1492	0.382465	2.36E-07	1.98E-06
						PLF						
0	497.945	498.277	2.49E-07	7.98E-06	498.043	497.059	2.40E-07	7.85E-06	503.656	504.027	2.28E-07	5.35E-06
0.25	28.1911	27.7613	2.00E-07	7.18E-06	25.8255	25.549	2.20E-07	7.54E-06	28.5426	28.2013	2.20E-07	5.39E-06
0.5	6.13445	5.87661	2.58E-07	8.15E-06	5.76852	5.52295	2.10E-07	7.36E-06	6.40804	6.22733	2.28E-07	5.39E-06
0.75	2.63193	2.5283	2.10E-07	7.36E-06	2.54016	2.46418	1.98E-07	7.13E-06	2.49894	2.50251	2.07E-07	5.15E-06
1	1.3573	1.45241	2.39E-07	7.81E-06	1.48763	1.50498	2.19E-07	7.50E-06	1.43639	1.51323	2.22E-07	5.31E-06
1.25	0.89725	1.04922	2.19E-07	7.54E-06	0.87801	1.03212	2.15E-07	7.45E-06	0.87516	1.05679	2.16E-07	5.92E-06
1.5	0.58849	0.79213	2.34E-07	7.73E-06	0.57087	0.780319	2.56E-07	8.11E-06	0.57353	0.796626	2.13E-07	5.16E-06
1.75	0.40317	0.640347	2.24E-07	7.63E-06	0.40193	0.638187	1.86E-07	6.94E-06	0.41875	0.655758	2.09E-07	5.19E-06
2	0.29456	0.536968	2.31E-07	7.72E-06	0.27371	0.522564	2.39E-07	7.78E-06	0.27141	0.522959	2.10E-07	5.16E-06
2.25	0.2008	0.444071	2.37E-07	7.77E-06	0.19407	0.43537	2.49E-07	7.99E-06	0.1925	0.435435	2.25E-07	5.32E-06
2.5	0.15608	0.38838	2.61E-07	8.19E-06	0.16315	0.398788	1.95E-07	7.13E-06	0.14069	0.373117	2.24E-07	5.33E-06
						LLF						
0	511.163	514.059	1.94E-07	6.27E-06	511.677	511.93	1.77E-07	6.28E-06	506.904	507.725	2.19E-07	7.14E-06
0.25	29.4728	29.1113	1.96E-07	6.17E-06	30.1145	29.7191	1.92E-07	6.57E-06	23.7271	22.7815	2.22E-07	7.44E-06
0.5	6.84833	6.502	1.35E-07	4.94E-06	6.72874	6.34084	2.05E-07	6.78E-06	5.18563	4.67402	2.21E-07	7.41E-06
0.75	2.76337	2.63444	2.14E-07	6.45E-06	2.78938	2.62285	2.24E-07	6.98E-06	2.45799	2.18901	2.37E-07	7.74E-06
1	1.53064	1.54693	1.83E-07	4.69E-06	1.54514	1.54216	2.08E-07	6.97E-06	1.40388	1.3565	2.34E-07	7.64E-06
1.25	0.90341	1.05374	2.09E-07	6.44E-06	0.98446	1.09237	2.47E-07	7.56E-06	0.90978	0.977599	2.03E-07	7.19E-06
1.5	0.6385	0.829095	1.84E-07	5.68E-06	0.62572	0.815889	2.20E-07	7.08E-06	0.66242	0.798536	2.36E-07	7.69E-06
1.75	0.39512	0.634461	2.02E-07	5.92E-06	0.41974	0.655102	2.67E-07	7.80E-06	0.4376	0.642407	2.09E-07	7.24E-06
2	0.29178	0.536455	1.80E-07	6.84E-06	0.28497	0.531509	2.54E-07	7.45E-06	0.33495	0.555732	2.06E-07	7.23E-06
2.25	0.19793	0.441717	2.09E-07	7.32E-06	0.21774	0.46001	2.05E-07	7.02E-06	0.21106	0.447318	2.07E-07	7.21E-06
2.5	0.15885	0.393162	1.94E-07	5.52E-06	0.1721	0.406548	2.23E-07	7.16E-06	0.16768	0.39563	2.61E-07	7.37E-06
						Classical						
0	480.483	68.2775	2.71E-07	2.20E-06	511.303	74.7793	4.37E-07	6.31E-06	512.415	84.0906	1.99E-07	6.46E-06
0.25	231.78	17.6915	2.12E-07	1.86E-06	137.648	13.9298	4.57E-07	6.30E-06	76.9748	10.7891	2.11E-07	6.44E-06
0.5	89.2072	7.8145	2.21E-07	1.90E-06	64.247	6.6106	1.16E-06	1.09E-05	46.7017	5.78231	2.79E-07	8.24E-06
0.75	69.9729	5.7303	2.27E-07	1.92E-06	55.72	5.13864	9.90E-07	1.05E-05	50.7438	4.90322	1.91E-07	5.34E-06
1	20.2967	2.7477	2.40E-07	1.99E-06	23.912	3.079	4.02E-07	6.18E-06	31.5221	3.4653	1.78E-07	5.89E-06
1.25	18.9389	2.43786	2.38E-07	1.97E-06	35.325	3.28106	5.45E-07	6.97E-06	27.7714	2.97509	1.60E-07	4.65E-06
1.5	13.0537	1.91218	2.47E-07	2.01E-06	21.301	2.44998	4.21E-07	6.10E-06	35.4437	3.15728	2.63E-07	7.42E-06
1.75	18.6242	2.16448	2.42E-07	2.00E-06	29.634	2.65067	5.90E-07	7.48E-06	19.3307	2.21868	2.41E-07	6.49E-06
2	41.7386	3.09142	2.24E-07	1.92E-06	27.005	2.58978	4.39E-07	6.97E-06	7.1397	1.35329	3.60E-07	9.27E-06
2.25	26.841	2.39189	2.30E-07	1.94E-06	54.275	3.31382	6.72E-07	7.92E-06	7.2748	1.31083	2.20E-07	6.93E-06
2.5	8.0692	1.3175	2.23E-07	1.91E-06	8.927	1.39631	7.78E-07	7.76E-06	8.2121	1.35378	1.71E-07	5.31E-06

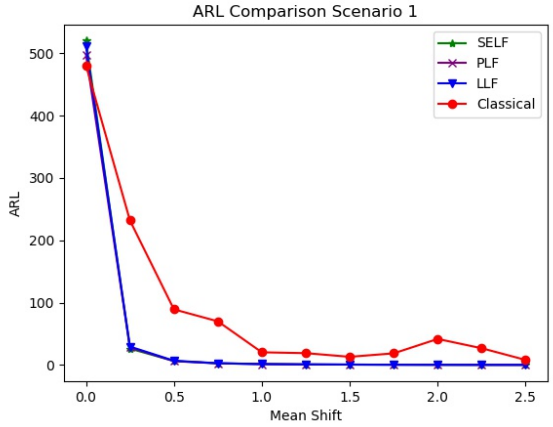
Table 3.8: ARL, SDRL, ATS, and SDTS Values for Bayesian CUSUM Chart Hyper-Parameter Sensitivity Analysis with Poisson Conjugate

We modified values for h to obtain an in-control ARL of around 500. Our prior α and β were adjusted for a sensitivity analysis of the hyper-parameters while applying shifts to the in-control sample mean. Similar to what is observed in section 3.5.1, there is a substantial drop in both the ARL and SDRL after the initial shift in the sample mean. This exhibits the chart's ability to detect small changes, regardless of the loss function that is used. We also note the consistent performance of the chart despite the choice of hyper-parameter. For the ATS we notice that results are not significantly different between each of the loss functions or the hyper-parameters. However, as you increase the α_0 while decreasing the β_0 we notice that the SDTS under the squared error and precautionary loss functions decreases. We also see that the SDTS under the squared error loss function is consistently lower than of its counterparts. Unlike what we see with the normal conjugate case, our ARL and SDRL are closer in value; this is because the use of the difference in distributions.

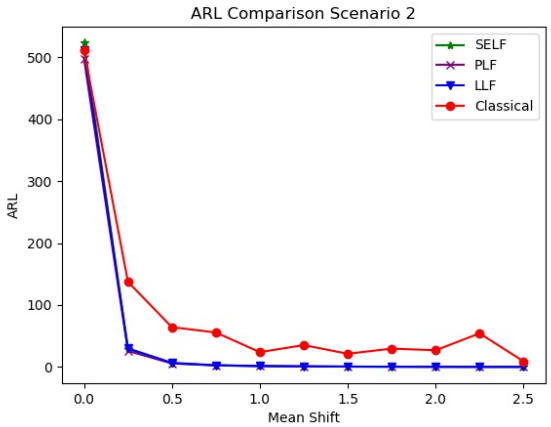
		Shifts											
n	h	0	0.25	0.5	0.75	1	1.25	1.5	1.75	2	2.25	2.5	
							SELF						
5	15	496.298	32.2171	15.6029	9.9113	7.5339	5.6834	4.8584	4.0162	3.5874	3.2056	2.5967	ARL
		(456.631)	(14.6619)	(5.95053)	(3.7745)	(2.80757)	(2.2271)	(1.92649)	(1.63295)	(1.53856)	(1.38807)	(1.24388)	SDRL
		1.71E-07	2.00E-07	1.32E-07	2.30E-07	2.14E-07	2.37E-07	2.87E-07	2.46E-07	2.67E-07	1.56E-07	2.49E-07	ATS
		(6.02E-06)	(6.91E-06)	(5.83E-06)	(7.44E-06)	(7.38E-06)	(7.35E-06)	(8.37E-06)	(7.84E-06)	(8.25E-06)	(5.89E-06)	(7.83E-06)	SDTS
10	5.6	473.362	26.3559	5.38841	2.46537	1.3864	0.85102	0.56017	0.42677	0.29204	0.2053	0.14904	ARL
		(474.171)	(25.6893)	(5.00599)	(2.27279)	(1.38406)	(0.985893)	(0.763662)	(0.638543)	(0.526149)	(0.443703)	(0.376891)	SDRL
		4.34E-07	5.14E-07	5.86E-07	5.65E-07	5.50E-07	5.90E-07	5.92E-07	5.90E-07	5.39E-07	5.13E-07	6.04E-07	ATS
		(8.21E-06)	(7.60E-06)	(8.87E-06)	(9.01E-06)	(1.13E-05)	(8.10E-06)	(9.35E-06)	(9.19E-06)	(1.01E-05)	(1.13E-05)	(8.50E-06)	SDTS
20	25.6	507.147	8.1758	3.1675	1.9069	1.176	0.8272	0.581	0.4307	0.2965	0.1923	0.1012	ARL
		(490.265)	(4.43099)	(1.54947)	(0.97736)	(0.715279)	(0.603937)	(0.556272)	(0.515167)	(0.462588)	(0.395627)	(0.302256)	SDRL
		2.69E-07	2.34E-07	2.47E-07	2.49E-07	2.44E-07	3.04E-07	2.45E-07	3.22E-07	3.00E-07	3.20E-07	2.89E-07	ATS
		(2.09E-06)	(1.95E-06)	(2.01E-06)	(2.02E-06)	(2.00E-06)	(2.22E-06)	(2.00E-06)	(2.27E-06)	(2.20E-06)	(2.26E-06)	(2.16E-06)	SDTS
30	20.05	507.798	5.60247	1.52877	0.68668	0.30691	0.12068	0.04604	0.01606	0.0062	0.002	0.00075	ARL
		(504.214)	(4.3489)	(1.1703)	(0.688368)	(0.49185)	(0.329175)	(0.210096)	(0.125706)	(0.0784956)	(0.0446766)	(0.0273759)	SDRL
		2.61E-07	2.50E-07	2.62E-07	2.70E-07	2.77E-07	2.92E-07	2.71E-07	2.77E-07	2.59E-07	2.55E-07	2.55E-07	ATS
		(2.06E-06)	(2.02E-06)	(2.06E-06)	(2.09E-06)	(2.12E-06)	(2.17E-06)	(2.10E-06)	(2.12E-06)	(2.06E-06)	(2.04E-06)	(2.04E-06)	SDTS
							PLF						
5	15	485.852	32.0134	15.7883	10.1827	7.62142	5.87656	4.9891	4.03255	3.38085	3.10012	2.78037	ARL
		(441.163)	(14.464)	(6.05076)	(3.77094)	(2.82046)	(2.25137)	(1.92472)	(1.64968)	(1.46048)	(1.37412)	(1.24799)	SDRL
		3.65E-07	4.62E-07	7.03E-07	4.33E-07	3.43E-07	3.41E-07	3.98E-07	3.43E-07	3.37E-07	3.31E-07	3.51E-07	ATS
		(4.35E-06)	(4.46E-06)	(6.62E-06)	(4.91E-06)	(3.89E-06)	(4.01E-06)	(4.34E-06)	(4.62E-06)	(4.34E-06)	(4.44E-06)	(4.45E-06)	SDTS
10	5.65	503.703	25.2912	5.77501	2.57142	1.34871	0.90072	0.60849	0.4113	0.3099	0.228	0.16009	ARL
		(505.085)	(24.6767)	(5.24267)	(2.32112)	(1.36733)	(1.00502)	(0.780647)	(0.630248)	(0.541297)	(0.465571)	(0.392481)	SDRL
		1.94E-07	2.15E-07	2.55E-07	2.35E-07	2.19E-07	2.63E-07	2.06E-07	2.47E-07	2.29E-07	2.22E-07	2.15E-07	ATS
		(7.02E-06)	(7.25E-06)	(7.89E-06)	(7.62E-06)	(7.36E-06)	(8.01E-06)	(7.02E-06)	(7.80E-06)	(7.55E-06)	(7.36E-06)	(7.24E-06)	SDTS
20	26	525.639	8.40693	3.34964	1.88663	1.30903	0.89353	0.58188	0.43516	0.28612	0.18427	0.10412	ARL
		(516.751)	(4.54815)	(1.59026)	(0.963669)	(0.724341)	(0.606856)	(0.55519)	(0.517664)	(0.458143)	(0.389351)	(0.305743)	SDRL
		2.46E-07	2.42E-07	2.27E-07	2.11E-07	2.23E-07	2.00E-07	2.19E-07	2.25E-07	2.18E-07	1.93E-07	2.11E-07	ATS
		(7.94E-06)	(7.90E-06)	(7.68E-06)	(7.41E-06)	(7.59E-06)	(7.18E-06)	(7.50E-06)	(7.63E-06)	(7.46E-06)	(7.03E-06)	(7.41E-06)	SDTS
30	19.68	525.041	5.96356	1.55198	0.66585	0.2828	0.11563	0.04548	0.01557	0.00519	0.0021	0.00067	ARL
		(523.562)	(4.71551)	(1.19914)	(0.693104)	(0.477958)	(0.323481)	(0.208882)	(0.123885)	(0.0718545)	(0.0457776)	(0.0258757)	SDRL
		2.41E-07	2.62E-07	2.70E-07	2.56E-07	2.44E-07	2.55E-07	2.35E-07	2.41E-07	2.45E-07	2.47E-07	2.46E-07	ATS
		(7.73E-06)	(8.23E-06)	(8.36E-06)	(8.07E-06)	(7.94E-06)	(8.07E-06)	(7.81E-06)	(7.90E-06)	(7.98E-06)	(7.94E-06)	(7.98E-06)	SDTS
							LLF						
5	12	538.844	28.2244	13.2126	8.57527	6.00356	4.79641	4.21672	3.16248	2.7339	2.39553	1.98138	ARL
		(506.553)	(14.7659)	(5.7545)	(3.5996)	(2.56895)	(2.0908)	(1.80989)	(1.50578)	(1.35128)	(1.23558)	(1.1248)	SDRL
		1.95E-07	2.15E-07	2.18E-07	2.24E-07	2.21E-07	1.82E-07	2.08E-07	2.09E-07	1.78E-07	2.03E-07	2.14E-07	ATS
		(6.85E-06)	(7.26E-06)	(7.31E-06)	(7.16E-06)	(7.25E-06)	(6.65E-06)	(7.11E-06)	(7.11E-06)	(6.52E-06)	(7.01E-06)	(7.14E-06)	SDTS
10	5.33	519.013	30.9964	6.4453	2.71976	1.50492	0.92406	0.62997	0.43218	0.29822	0.21673	0.14393	ARL
		(518.483)	(30.6673)	(6.08091)	(2.53317)	(1.49928	(1.04987)	(0.811891)	(0.65547)	(0.541041)	(0.458146)	(0.375199)	SDRL
		2.09E-07	2.87E-07	2.24E-07	1.94E-07	1.82E-07	1.80E-07	1.82E-07	2.13E-07	1.91E-07	2.19E-07	2.26E-07	ATS
		(5.37E-06)	(7.58E-06)	(6.27E-06)	(6.86E-06)	(5.67E-06)	(5.44E-06)	(4.83E-06)	(5.43E-06)	(7.00E-06)	(5.61E-06)	(7.08E-06)	SDTS
20	20.1	509.815	8.08598	2.77656	1.4704	0.99445	0.60287	0.37462	0.25715	0.13602	0.0824	0.0489	ARL
		(506.393)	(5.18941)	(1.59694)	(0.931431)	(0.712263)	(0.594927)	(0.510881)	(0.446211)	(0.345136)	(0.275482)	(0.215983)	SDRL
		2.40E-07	2.55E-07	2.26E-07	2.49E-07	2.37E-07	2.57E-07	2.32E-07	2.25E-07	2.19E-07	2.20E-07	2.21E-07	ATS
		(6.95E-06)	(7.53E-06)	(6.75E-06)	(7.69E-06)	(7.56E-06)	(7.20E-06)	(6.90E-06)	(7.63E-06)	(7.54E-06)	(3.62E-06)	(7.55E-06)	SDTS
30	20	476.63	5.55564	1.49389	0.64454	0.31084	0.12206	0.04281	0.01663	0.00516	0.00173	0.0008	ARL
		(474.056)	(4.30695)	(1.15642)	(0.675831)	(0.493314)	(0.331544)	(0.202873)	(0.128037)	(0.0716476)	(0.0415573)	(0.028273)	SDRL
		4.41E-07	3.27E-07	2.78E-07	2.84E-07	3.01E-07	2.99E-07	3.96E-07	2.82E-07	2.82E-07	3.00E-07	2.58E-07	ATS
		(5.30E-06)	(4.61E-06)	(4.08E-06)	(4.22E-06)	(4.26E-06)	(4.73E-06)	(5.21E-06)	(4.24E-06)	(4.04E-06)	(4.81E-06)	(5.11E-06)	SDTS

Table 3.9: ARL, SDRL, ATS, and SDTS Values for Bayesian CUSUM Chart Sample Size Sensitivity Analysis with Poisson Conjugate

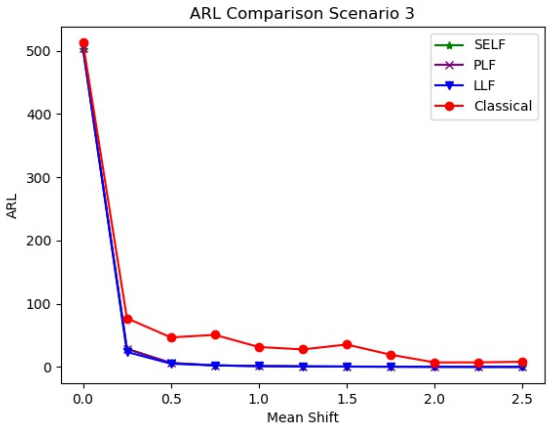
The sample size analysis uses the same values for n as in section 3.5.1 and the values for h are determined based on our desired $ARL_0 = 500$, are seen in table (reference table). Results shown in table 3.9 continue similar patterns as table 3.8. There is again a noticeable drop in the ARL as we impose our first mean shift, while the SDRL remains close in value to the ARL. Our ATS is relatively small regardless of the sample size. The SDTS for the squared error loss function appears to decrease as our sample size increases, but remains consistent for the other loss functions.



(a) $\alpha = 16$, $\beta = \frac{5}{8}$

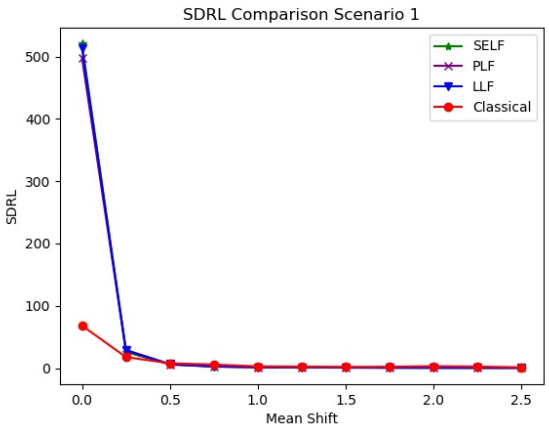


(b) $\alpha = 36$, $\beta = \frac{5}{12}$

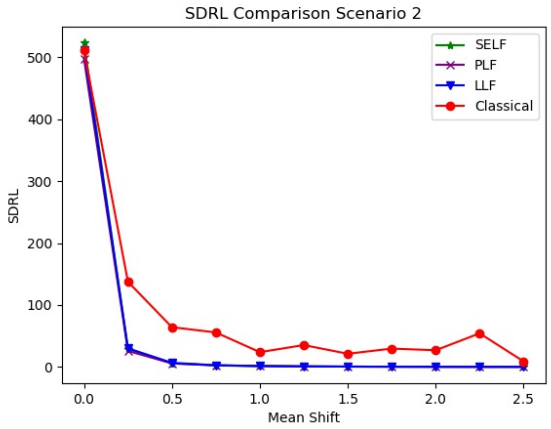


(c) $\alpha = 64$, $\beta = \frac{5}{16}$

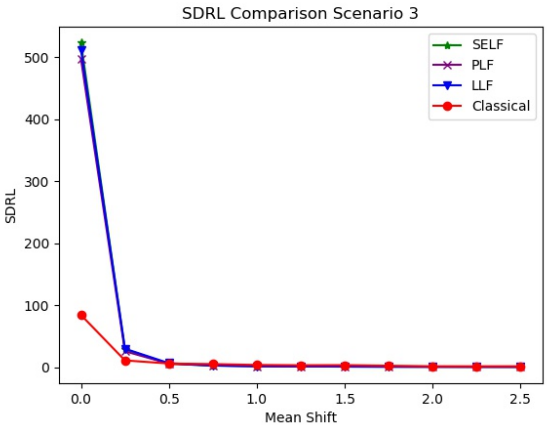
Figure 3.1: ARL Comparisons



(a) $\alpha = 16$, $\beta = \frac{5}{8}$



(b) $\alpha = 36$, $\beta = \frac{5}{12}$



(c) $\alpha = 64$, $\beta = \frac{5}{16}$

Figure 3.2: SDRL Comparisons

We show a visual comparison of the performance measurements between each of the Bayesian charts under the loss functions along with the classical CUSUM chart method in figures 4.1 and 4.2. Our plots show that regardless of the loss function used, there is a drastic decline in the ARL after a small shift is imposed and the trend continues to decline until it approaches zero. The classical chart exhibits a similar initial decline, but the difference in the in-control ARL and first out-of-control ARL occurrence is not as large. In figures 4.1a, 4.1b, and 4.1c the classical method ARL results fluctuate as the applied mean shifts are added. This is also seen for the

SDRL in scenario 2 (figure 4.2b), however, in figures 4.2a and 4.2c the classical results for the SDRL show consistency.

3.6. Real Data Analysis

We consider our method on a dataset found at www2.datasus.gov.br seen in Alencar et al., 2017 and Urbieta et al., 2017. The data is a count series of respiratory disease related hospitalizations for people over 65 years old in São Paulo, Brazil. Originally, the data was collected to represent daily counts from January 2006-December 2011, but we chose to look at the weekly number of hospitalizations, totaling 313 weeks. Shown in figure 4.3b is the week plotted against the average weekly count, observing that the weekly number of hospitalizations has a positive linear trend and a seasonal pattern over time. That is, there are a greater number of hospitalizations in the São Paulo winter months (June-August) than in their summer months (December-February).

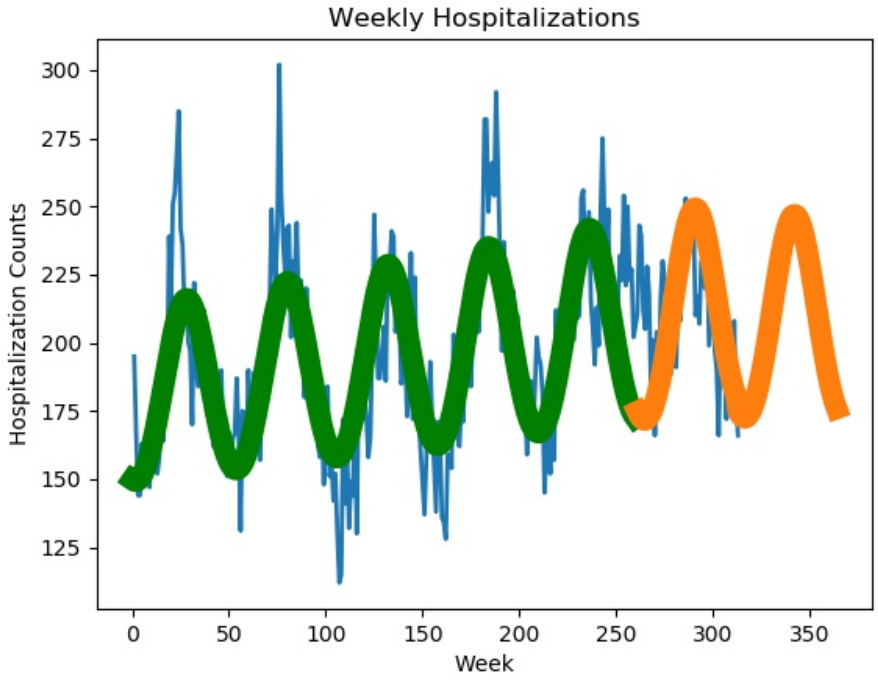


Figure 3.3: Weekly Hospitalizations with predicted values

Since we have count data with dispersion present, we consider the negative binomial distribution with our generalized linear model. We account for the seasonal pattern using the sine and cosine functions. Population is used as an offset variable along with the log link function to properly model hospitalization rate per 100,000 inhabitants. We assume that counts for January 2006-December 2010 are non-epidemic and use this to construct the model. We validate our model using the weekly data for 2011 and predict weekly counts for 2012.

$$ln\left[100,000 * \frac{\mu_{0,t}}{P_t}\right] = \beta_0 + \beta_1 \sin\left(\frac{2\pi t}{52}\right) + \beta_2 \cos\left(\frac{2\pi t}{52}\right) \tag{3.5}$$

where, $\mu_{0,t}$ is the non-epidemic average number of hospitalizations and P_t is the population size for week t, and 52 represents the seasonal period. This model follows the same make up as what is seen in Urbieta et al., 2017, however, since we do not consider daily counts we do

not use day-of-week dummy variables.

We use this data to assess the performance of our control chart and track the weekly number of hospitalizations for January 2006 to December 2011. We partition the data into two sections: training data and data to validate. Our training data is the weekly counts for January 2006-December 2010. The upper control limit is estimated using training data and we apply shifts to the weekly data for 2011 to assess the charts detection capability.

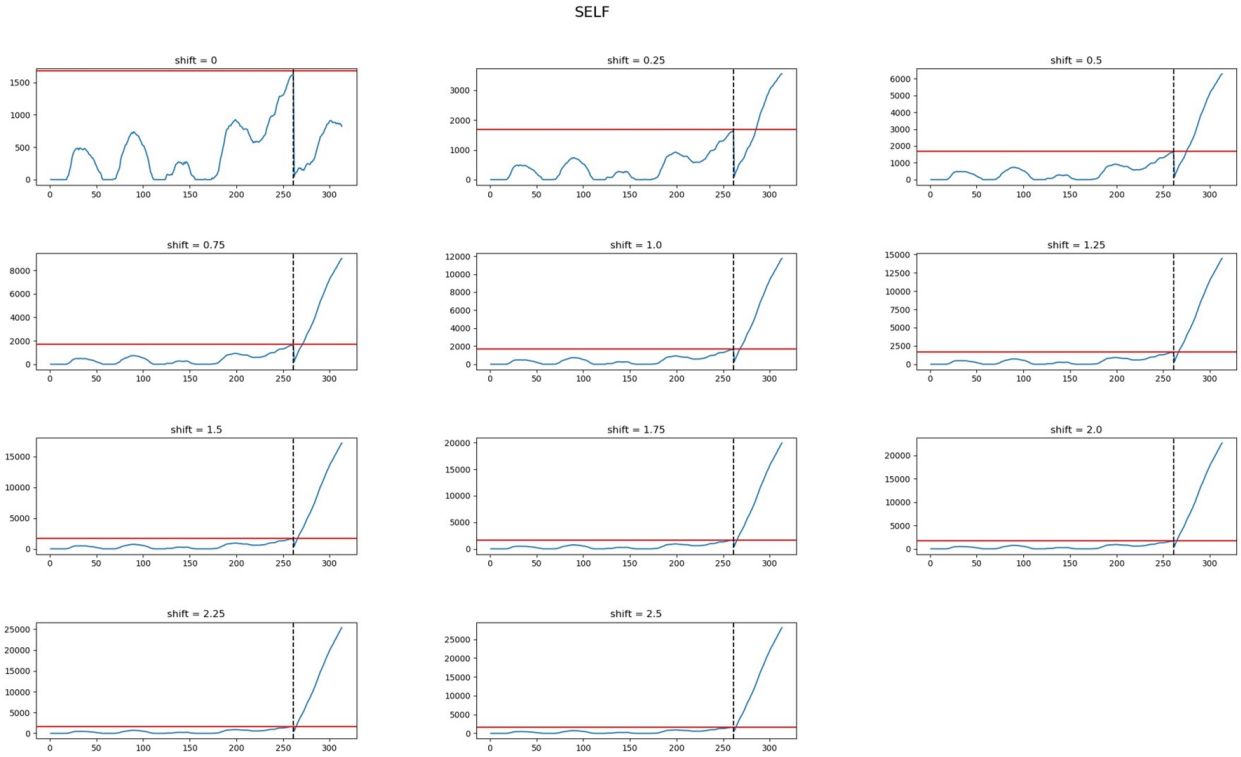


Figure 3.4: Squared Error Loss Function CUSUM Chart Results

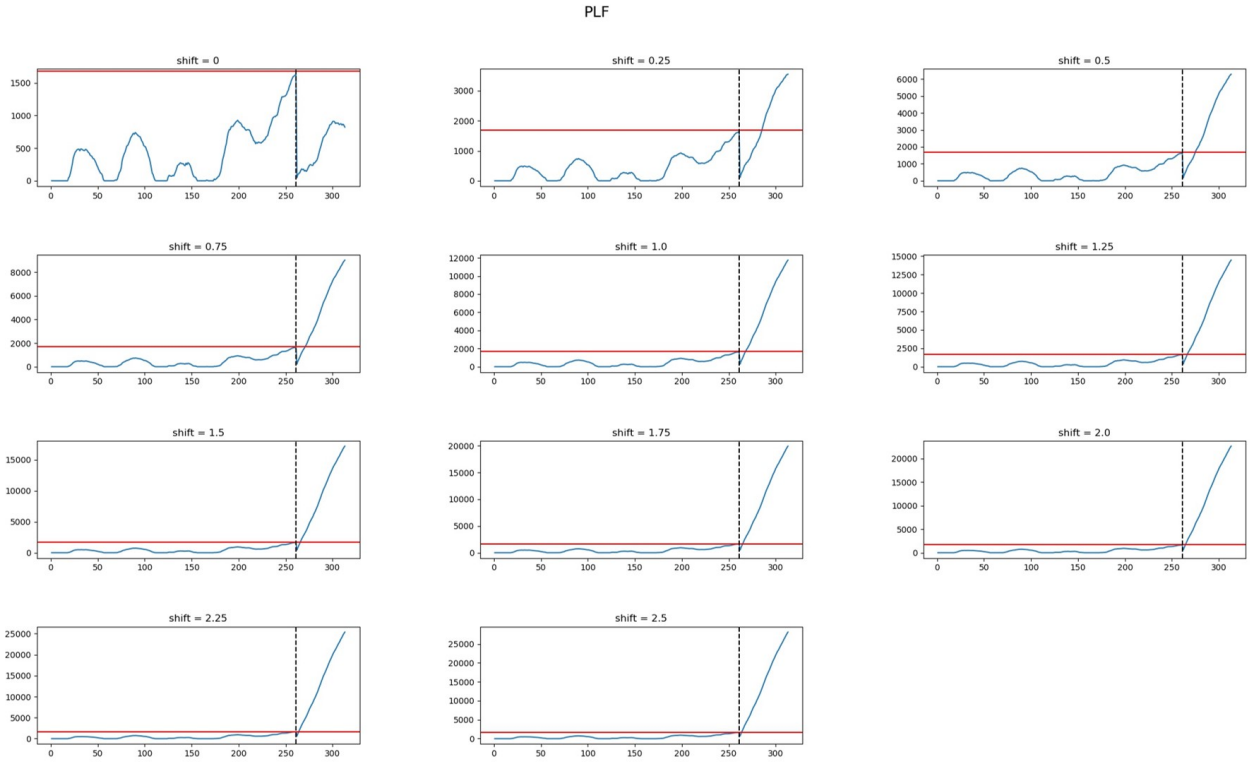


Figure 3.5: Precautionary Loss Function CUSUM Chart Results

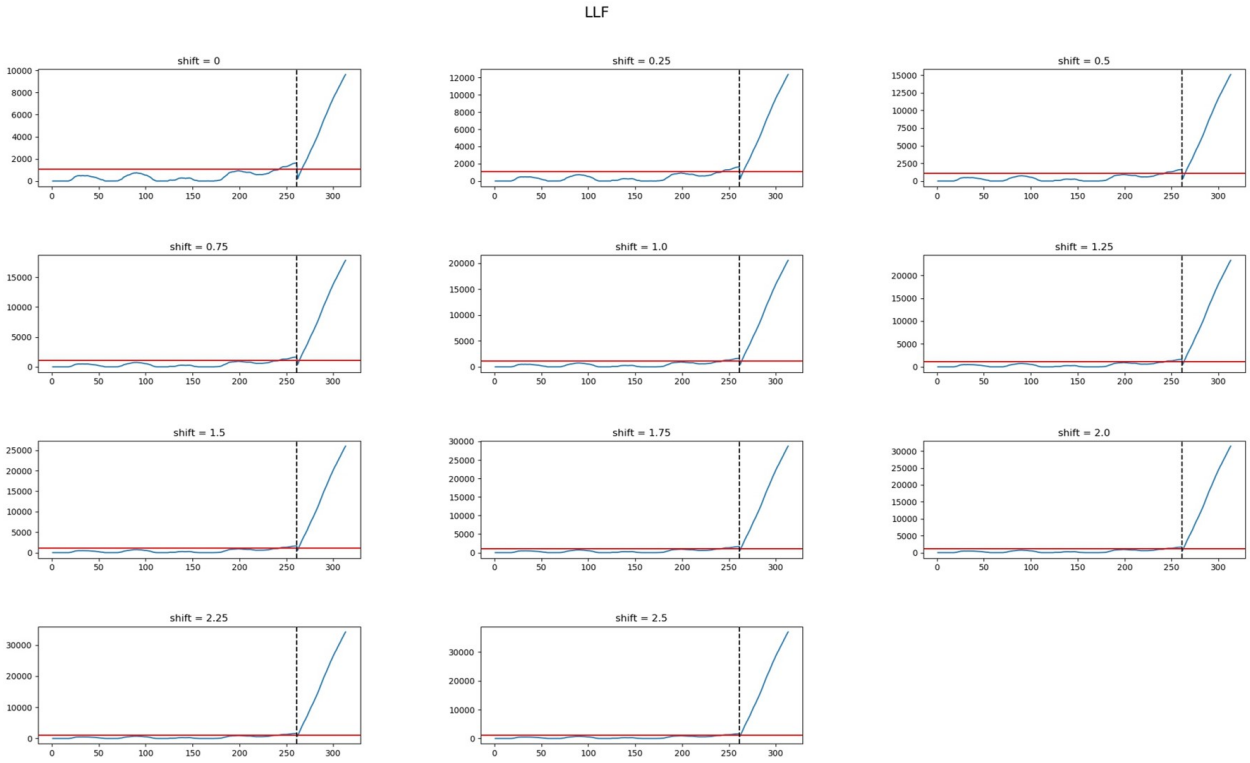


Figure 3.6: Linex Loss Function CUSUM Chart Results

Figures 3.4, 3.5, and 3.6 track the weekly number of hospitalizations for January 2006 to December 2011. Without an induced shift the weekly data for 2011 falls below the upper control limit under the SELF and PLF chart, but once a shift occurs both charts signal out-of-control. However, under the LLF chart a breech of the control limit occurs regardless if a shift was imposed, and this is likely a result of the chosen value for c. For each variation of the chart, as we increase the shift size the out-of-control detection of the charts occurs earlier. We observe a delay in detection at the beginning of 2011 because of the seasonal affects. Although we increase the counts via our shifts, in application it is not alarming that hospitalizations would increase by 25% or 50% per 100,000 inhabitants.

3.7. Discussion

We constructed Bayesian CUSUM control charts with the use of the self error, precautionary, and linex loss functions. Our objective was to assess the detection capabilities of the charts under the Gaussian conjugate and Poisson conjugate cases. The performance measurements for our study were the ARL, SDRL, ATS, and SDTS. Based on these measurements we recommend the use of any of the three CUSUM control charts. Each performed well in obtaining the desired ARL_0 and detecting a shift in the mean, while maintaining a relatively low ATS. We applied the charts to a count series dataset and analyzed their real-world capabilities. All the charts quickly detected the out-of-control occurrences, and the squared error and precautionary loss functions performed similarly in tracking in-control and out-of-control occurrences. The linex loss function signaled early in the in-control and out-of-control scenarios. This is because it is exponential in nature due to the chosen value of its constant, c and signifies overestimation over underestimation. Overall, we recommend the use of any of the control charts under different distributions and regardless of the sample size and hyper-parameters.

Bibliography

Shewhart, W. A. (1926). Quality control charts. *The Bell System Technical Journal*, *5*(4), 593–603.

Page, E. S. (1954). Continuous inspection schemes. *Biometrika*, *41*(1/2), 100–115.

Roberts, S. (1959). Control chart tests based on geometric moving averages. *Technometrics*, *1*(3), 239–250.

Riaz, S., Riaz, M., Hussain, Z., & Abbas, T. (2017). Monitoring the performance of bayesian ewma control chart using loss functions. *Computers & Industrial Engineering*, *112*, 426–436.

Norstrom, J. G. (1996). The use of precautionary loss functions in risk analysis. *IEEE Transactions on reliability*, *45*(3), 400–403.

Varian, H. R. (1975). A bayesian approach to real estate assessment. *Studies in Bayesian econometric and statistics in Honor of Leonard J. Savage*, 195–208.

Zellner, A. (1986). Bayesian estimation and prediction using asymmetric loss functions. *Journal of the American Statistical Association*, *81*(394), 446–451.

Alencar, A. P., Lee Ho, L., & Albarracin, O. Y. E. (2017). Cusum control charts to monitor series of negative binomial count data. *Statistical methods in medical research*, *26*(4), 1925–1935.

Urbieta, P., Lee HO, L., & Alencar, A. (2017). Cusum and ewma control charts for negative binomial distribution. *Quality and Reliability Engineering International*, *33*(4), 793–801.

3.8. Appendix

3.8.1. Poisson-Gamma Derivations

Poisson Likelihood:

$$\begin{aligned} x|\lambda &\sim Poisson(\lambda) \\ f(x|\lambda) &= \frac{e^{-n\lambda}\ \lambda^{\sum_{i=1}^{n} x_i}}{\prod_{i=1}^{n} x_i!} \end{aligned} \tag{3.6}$$

Gamma Prior:

$$\begin{aligned} \lambda &\sim Gamma(\alpha, \beta) \\ f(\lambda) &= \frac{\beta^{\alpha}}{\Gamma(\alpha)} \lambda^{\alpha-1} e^{-\beta\lambda} \end{aligned} \tag{3.7}$$

Gamma Posterior:

$$\begin{aligned} \lambda|x &\sim Gamma(n\bar{x}+\alpha, n+\beta) \\ f(\lambda|x) &= \frac{\beta^{\alpha}}{\Gamma(\alpha)} \lambda^{(n\bar{x}+\alpha)-1} e^{-(n+\beta)\lambda} \end{aligned} \tag{3.8}$$

Posterior Predictive for Poisson Conjugate:

$$\begin{aligned} y|x &\sim Poisson(\lambda) \\ f(y|x) &= \frac{e^{-\lambda}\ \lambda^{y}}{y!} \end{aligned} \tag{3.9}$$

$$
\begin{aligned}
f(y|x) &= \int f(\lambda|x) f(y|\lambda) d\lambda \\
&= \int \frac{(\beta+n)^{n\bar{x}+\alpha}}{\Gamma(n\bar{x}+\alpha)} e^{-(n+\beta)\lambda} \lambda^{(n\bar{x}+\alpha)-1} \frac{e^{-\lambda}\ \lambda^{y}}{y!} d\lambda \\
&= \frac{(\beta+n)^{n\bar{x}+\alpha}}{\Gamma(n\bar{x}+\alpha)\, y!} \int e^{-(n+\beta)} e^{-\lambda} \lambda^{(n\bar{x}+\alpha)-1} \lambda^{y} d\lambda \\
\text{let c} = \frac{(\beta+n)^{n\bar{x}+\alpha}}{\Gamma(n\bar{x}+\alpha)\, y!},\ & c \int e^{-(n+\beta)\lambda-\lambda} \lambda^{(n\bar{x}+\alpha)-1+y} d\lambda \\
&= c \int e^{-(n+\beta+1)\lambda} \lambda^{(n\bar{x}+y+\alpha)-1} d\lambda \\
&= \frac{(\beta+n)^{n\bar{x}+\alpha}}{\Gamma(n\bar{x}+\alpha)\, y!} \cdot \frac{\Gamma(n\bar{x}+\alpha)+y}{(\beta+n+1)^{n\bar{x}+\alpha+y}} \\
&= \frac{\Gamma(n\bar{x}+\alpha+y)}{\Gamma(n\bar{x}+\alpha)\, y!} \cdot \frac{(\beta+n)^{n\bar{x}+\alpha}}{(\beta+n+1)^{n\bar{x}+\alpha+y}} \\
&= \frac{(n\bar{x}+\alpha+y)!}{(n\bar{x}+\alpha)!\, y!} \cdot \left(\frac{\beta+n}{\beta+n+1}\right)^{n\bar{x}+\alpha} \left(\frac{1}{\beta+n+1}\right)^{y} \\
&\binom{n\bar{x}+\alpha+y-1}{y-1} \cdot \left(1-\frac{1}{\beta+n+1}\right)^{n\bar{x}+\alpha} \left(\frac{1}{\beta+n+1}\right)^{y}
\end{aligned}
\tag{3.10}
$$

$$
\begin{aligned}
f(y|x) &= \binom{n\bar{x}+\alpha+y-1}{y-1} \left(1-\frac{1}{\beta+n+1}\right)^{n\bar{x}+\alpha} \left(\frac{1}{\beta+n+1}\right)^{y} \\
y|x &\sim NegBin(n\bar{x}+\alpha, n+\beta)
\end{aligned}
\tag{3.11}
$$

3.8.1.1. Loss Functions Best Estimators

SELF:

$$
\begin{aligned}
\hat{\lambda}^{*} &= E[\lambda|x] = \frac{n\bar{x}+\alpha}{n+\beta}, \\
\left[\mu_{SELF,PG} = \hat{\lambda}^{*} = \frac{n\bar{x}+\alpha}{n+\beta}\right]
\end{aligned}
\tag{3.12}
$$

PLF:

$$\begin{aligned}
\hat{\lambda}^* &= \sqrt{E[\lambda^2|x]} \\
VAR[\lambda|x] &= E[\lambda^2|x] - E[\lambda|x]^2 \\
\frac{n\bar{x}+\alpha}{(n+\beta)^2} &= E[\lambda^2|x] - \left(\frac{n\bar{x}+\alpha}{n+\beta}\right)^2 \\
\frac{n\bar{x}+\alpha}{(n+\beta)^2} &= E[\lambda^2|x] - \frac{(n\bar{x}+\alpha)^2}{(n+\beta)^2} \\
E[\lambda^2|x] &= \frac{(n\bar{x}+\alpha)+(n\bar{x}+\alpha)^2}{(n+\beta)^2} \\
\sqrt{E[\lambda^2|x]} &= \sqrt{\frac{(n\bar{x}+\alpha)+(n\bar{x}+\alpha)^2}{(n+\beta)^2}} \\
\left[\mu_{PLF,PG} = \hat{\lambda}^* = \frac{\sqrt{(n\bar{x}+\alpha)+(n\bar{x}+\alpha)^2}}{(n+\beta)}\right]
\end{aligned} \tag{3.13}$$

LLF:

$$\begin{aligned}
\hat{\lambda}^* &= -\frac{1}{c} ln\left[E[e^{-c\lambda}]\right] \\
E[e^{-c\lambda}] &= \int e^{-c\lambda} \cdot \frac{\beta^\alpha}{\Gamma(\alpha)} \lambda^{\alpha-1} e^{-\beta\lambda} d\lambda \\
&= \frac{\beta^\alpha}{\Gamma(\alpha)} \int e^{-c\lambda} \lambda^{\alpha-1} e^{-\beta\lambda} d\lambda \\
&= \frac{\beta^\alpha}{\Gamma(\alpha)} \int e^{-(c+\beta)\lambda} \lambda^{\alpha-1} d\lambda \\
&= \frac{\beta^\alpha}{\cancel{\Gamma(\alpha)}} \cdot \frac{\cancel{\Gamma(\alpha)}}{(c+\beta)^\alpha} \\
\left[\mu_{LLF,PG} = \hat{\lambda}^* = -\frac{1}{c} ln\left[\frac{\beta^\alpha}{(c+\beta)^\alpha}\right]\right]
\end{aligned} \tag{3.14}$$

3.8.1.2. Solving for α and β

Given $\mu_0 = [10, 15, 20]$ and $\sigma_0^2 = [16, 36, 64]$:

$\underline{\mu_0 = 10,\ \sigma_0^2 = 16}$

$$\begin{aligned} 10 &= \frac{\alpha}{\beta} \Rightarrow \alpha = 10\beta \\ 16 = \frac{\alpha}{\beta^2} \Rightarrow 16 &= \frac{10\beta}{\beta^2} \Rightarrow \left[\beta = \frac{5}{8}\right] \\ 10 &= \frac{\alpha}{\frac{5}{8}} \Rightarrow [\alpha = 16] \end{aligned} \tag{3.15}$$

$\underline{\mu_0 = 15,\ \sigma_0^2 = 36}$

$$\begin{aligned} 15 &= \frac{\alpha}{\beta} \Rightarrow \alpha = 15\beta \\ 36 = \frac{\alpha}{\beta^2} \Rightarrow 36 &= \frac{15\beta}{\beta^2} \Rightarrow \left[\beta = \frac{5}{12}\right] \\ 15 &= \frac{\alpha}{\frac{5}{12}} \Rightarrow [\alpha = 36] \end{aligned} \tag{3.16}$$

$\underline{\mu_0 = 20,\ \sigma_0^2 = 64}$

$$\begin{aligned} 20 &= \frac{\alpha}{\beta} \Rightarrow \alpha = 20\beta \\ 64 = \frac{\alpha}{\beta^2} \Rightarrow 64 &= \frac{15\beta}{\beta^2} \Rightarrow \left[\beta = \frac{5}{16}\right] \\ 20 &= \frac{\alpha}{\frac{5}{16}} \Rightarrow [\alpha = 64] \end{aligned} \tag{3.17}$$

Chapter 4

Generalized Bayesian EWMA Charts using Different Loss Functions

Abstract

In profile monitoring, control charts are used to visually observe process behaviors. Often there are control charts that are best suited for varying processes. A Bayesian approach to the popular exponentially weighted moving average (EWMA) control chart is proposed to perform under different distributions. Different loss functions are considered to inform the framework of the control chart for the Poisson conjugate case. The proposed method is assessed under a simulation study where the performance measurements are the average run length (ARL), standard deviation of the run length (SDRL), average time to signal (ATS), and standard deviation of time to signal (SDTS). Further assessment of the general use of the method is done via a sensitivity analysis for the control chart decision parameters, out-of-control shift sizes, and distribution hyper-parameters. Once the performance and generalization of the chart is considered, we model a count series of respiratory disease related hospitalizations for people over 65 years old

in São Paulo, Brazil. The control chart is implemented with an analysis of the hospitalization data to showcase the efficacy of our method on over-dispersed count data.

Keywords: Profile Monitoring; Bayesian; Loss Functions, EWMA; Sensitivity Analysis; Quality Control.

4.1. Introduction

Profile monitoring is an area within statistical process monitoring (SPM) that is characterized by the use of control charts to visually monitor a process. The exponentially weighted moving average (EWMA) (Roberts, 1959) and cumulative sum (CUSUM) (Page, 1954) control charts are frequently used charts because of their memory-based properties. Both charts have been studied in conjunction with a variety of Bayesian techniques such like, Abbas et al., 2019, Ali, 2020, Tsiamyrtzis and Hawkins, 2008, Aslam and Anwar, 2020, and Noor-ul Amin and Noor, 2021.

Our objective is to apply Bayesian techniques proposed in Riaz et al., 2017 and Jones et al., 2021 to the EWMA chart for monitoring non-Gaussian data. This paper will proceed as follows: presentation of the Bayesian methods and control chart framework under these methods, simulation analysis of the proposed charting method under the Poisson likelihood-Gamma prior case, application of methods on real count data, and discussion of all results.

4.2. Bayesian Inference

Bayes theorem combines the likelihood function, which is determined by the data, with the expert-chosen prior distribution to determine what the posterior distribution will be (equation

4.1).

$$p(\theta|\mathbf{X}) = \frac{\mathbf{p}(\mathbf{X}|\theta)\mathbf{p}(\theta)}{\mathbf{p}(\mathbf{X})} \tag{4.1}$$

where, θ is the parameter of interest, $\mathbf{X}$ is the data, $P(\theta)$ is the prior distribution, $P(\mathbf{X})$ is the marginal distribution of the data, and $P(\mathbf{X}|\theta)$ is the likelihood function. Once the posterior distribution is found, it is be combined with the likelihood of future y data and integrated with respect to θ to obtain the posterior predictive distribution (equation 4.2).

$$p(y|\mathbf{X}) = \int p(\mathbf{y}|\theta)\mathbf{p}(\theta|\mathbf{X})\mathbf{d}\theta \tag{4.2}$$

where, $p(\mathbf{y}|\theta)$ is the likelihood function for the future data

4.2.1. Loss Functions

In Bayesian statistics, a loss function can be used to obtain the best estimator for the parameter of interest. The idea is to minimize the expected loss presented using a specified loss function, with the general form seen in equation 4.3.

$$\hat{\theta}^* = \min_{\hat{\theta}} E_\theta[L(\hat{\theta},\theta] \tag{4.3}$$

where, $\hat{\theta}^*$ is the estimator which minimizes the expected loss

Loss Function	$L(\theta,\hat{\theta})$ =	$\hat{\theta}^*$ =
Squared Error	$(\theta-\hat{\theta})^2$	$E[\theta\|x]$
Precautionary	$\frac{(\theta-\hat{\theta})^2}{\hat{\theta}}$	$\sqrt{E[\theta^2\|x]}$
Linex	$(e^{c(\hat{\theta}-\theta)}-c(\hat{\theta}-\theta)-1)$	$-\frac{1}{c}lnE[e^{-c\theta}]$

Table 4.1: Bayes Estimators for Loss Functions

We study the three loss functions that are shown in table 4.1 for our method. The first is the Squared Error loss function (SELF) which is a symmetric loss function known for its simplicity and easy application. Next we study the Precautionary loss function (PLF), first outlined in Norstrom, 1996 to prevent parameter overestimation. Unlike the SELF, this loss function weights positive and negative errors differently and is known to produce better estimators for low failure rate problems. The Linex loss function (LLF), our final choice, is similar to the PLF in that it is also asymmetric aiming to identify parameter overestimation. The constant, c, in the LLF is user defined such that if $c > 0$ overestimation takes presidence, but if $c \to 0$ the loss function becomes symmetric.

4.2.2. Poisson Conjugate

The Poisson distribution is best for modeling count data, and when used under Bayesian inference its conjugate prior is the Gamma distribution. When combined under equation 4.1, we obtain the following posterior distribution: $Gamma(n\bar{x}+\alpha, n+\beta)$, with shape = $n\bar{x}+\alpha$ and inverse scale = $n+\beta$. Once the posterior distribution is obtained, we are able to derive the posterior predictive distribution to be $NegBin(n\bar{x}+\alpha, n+\beta)$ under equation 4.2. The means for both the posterior and posterior predictive distributions are derived under each of the loss functions and are seen in table 4.2. Since our method tracks the Bayes estimator of a profile, we

need the variance of the Bayes Estimator of the posterior predictive distribution, $\sigma^2_{\tilde{Y}}$. Note that because we track profile averages as our Bayes Estimator, our μ under each loss function will be different, but our variances will remain the same regardless of the loss function used.

Loss Function	Bayes Estimator
Squared Error	$\mu_{SELF,PG} = \frac{n\bar{x}+\alpha}{n+\beta}$
Precautionary	$\mu_{PLF,PG} = \frac{\sqrt{(n\bar{x}+\alpha)(n\bar{x}+\alpha)^2}}{n+\beta}$
Linex	$\mu_{LLF,PG} = -\frac{1}{c} ln\left[\frac{\beta^\alpha}{(c+\beta)^\alpha}\right]$

Table 4.2: Bayes Estimators for Loss Functions

Distribution	Variance
Posterior	$\sigma^2_{PD} = \frac{n\bar{x}+\alpha}{(n+\beta)^2}$
Posterior Predictive	$\sigma^2_{PPD} = \frac{n\bar{x}+\alpha}{(n+\beta)^2}(n+\beta+1)$
Bayes Estimator of Posterior Predictive	$\sigma^2_{\tilde{Y}} = \frac{n\bar{x}+\alpha}{n(n+\beta)^2}(n+\beta+1)$

Table 4.3: Variances Based on Distribution

where, n is the sample size, $\bar{x}$ is the sample mean, α is the shape parameter, β is the inverse scale parameter, and c is the Linex loss function symmetry constant.

4.2.3. Chart Framework

In equation 4.4, μ_{LF} and $\sigma_{\tilde{Y}}$ are the mean and standard deviation from the Bayes estimator posterior predictive distribution respective to the loss function and τ is a user-defined constant

determining the length of memory. For the charting statistic, z_i, $(\bar{y}|x)$ is the Bayes estimator for the set of predicted data.

EWMA Posterior Predictive Control Limits:

$$\begin{aligned} UCL/LCL &= \mu_{LF} \pm L\sigma_{\bar{Y}}\sqrt{\frac{\tau}{2-\tau}} \\ z_i &= \tau(\bar{y}|x) + (1-\tau)z_{i-1} \end{aligned} \tag{4.4}$$

In the Poisson conjugate case, we consider only the upper control limit since we are interested in an increase in counts.

4.3. Sensitivity Analyses

We conduct a sensitivity analysis for the hyper-parameters and the sample size for the Bayesian EWMA control chart under the Poisson conjugate to test its performance. For both analyses, our performance measurements are the average run length (ARL), standard deviation of the run length (SDRL), average time to signal (ATS), and the standard deviation of the time to signal (SDTS). We assess the performance by first obtaining the desired in-control ARL, then determine which method has the greatest decrease once the initial shift is imposed.

We impose 11 relatively small shift sizes denoted as $\delta = 0$ to 2.5 by 0.25, where the shift size is defined as $\mu_1 = \mu_0 * (1+\delta)$. Note that $\delta = 0$ indicates our desired in-control ARL of around 500 ($ARL_0 = 500$) and we use $m = 10{,}000$ iterations to calculate our results. The initial data is generated using Poisson($\lambda = 25$).

4.3.1. Hyper-Parameter Analysis

We selected 3 choices for our hyper-parameter simulations derived from $\mu_0 = [10, 15, 20]$ and $\sigma_0 = [4, 6, 8]$. Using these values for μ_0 and σ_0, we solved for the prior values of α_0 and β_0 as: $\alpha_0 = [16, 36, 64]$ and $\beta_0 = \left[\frac{5}{8}, \frac{5}{12}, \frac{5}{16}\right]$. The classical EWMA chart was also studied as a comparison of its capabilities on count data versus our Bayesian methods.

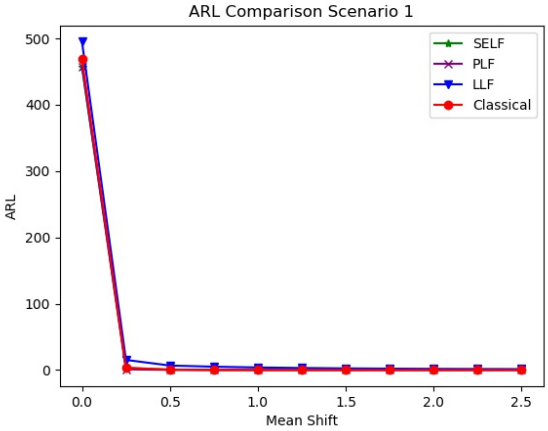
	$\alpha_0 = 16, \beta_0 = \frac{5}{8}$				$\alpha_0 = 36, \beta_0 = \frac{5}{12}$				$\alpha_0 = 64, \beta_0 = \frac{5}{16}$			
Shifts	ARL	SDRL	ATS	SDTS	ARL	SDRL	ATS	SDTS	ARL	SDRL	ATS	SDTS
						SELF						
0	457.812	584.427	2.83E-07	5.90E-06	503.518	987.898	1.93E-07	1.78E-06	536.307	2725.14	2.64E-07	2.40E-06
0.25	3.1181	5.08246	2.39E-07	5.51E-06	0.4183	1.57915	2.18E-07	1.86E-06	0.0245	0.430697	2.33E-07	1.95E-06
0.5	0.375	0.809058	2.36E-07	5.31E-06	0.0363	0.244504	2.07E-07	1.84E-06	0.0022	0.0616049	1.96E-07	1.79E-06
0.75	0.2198	0.554877	2.75E-07	5.97E-06	0.0469	0.278029	2.19E-07	1.90E-06	0.0006	0.0282779	2.71E-07	2.10E-06
1	0.0389	0.202452	3.54E-07	6.60E-06	0.0065	0.081595	2.01E-07	1.82E-06	0.0004	0.019996	2.36E-07	2.01E-06
1.25	0.0238	0.156312	2.23E-07	5.35E-06	0.0006	0.0244875	2.17E-07	1.89E-06	0.0001	0.0099995	2.19E-07	1.90E-06
1.5	0.0077	0.0874112	1.96E-07	4.97E-06	0.0007	0.0264483	2.06E-07	1.84E-06	0	0	1.96E-07	1.80E-06
1.75	0.0031	0.0555913	2.31E-07	5.38E-06	0.0001	0.0099995	2.71E-07	2.10E-06	0	0	2.04E-07	1.83E-06
2	0.0007	0.0264483	2.35E-07	5.51E-06	0.0001	0.0099995	2.13E-07	1.87E-06	0	0	2.33E-07	1.95E-06
2.25	0.0006	0.0244875	1.69E-07	4.63E-06	0	0	1.97E-07	1.80E-06	0	0	2.38E-07	1.97E-06
2.5	0.0002	0.0141407	2.78E-07	5.98E-06	0	0	1.98E-07	1.80E-06	0	0	2.45E-07	2.00E-06
						PLF						
0	458.006	721.484	2.25E-07	1.92E-06	564.132	1231.69	2.68E-07	5.45E-06	520.414	3476.57	8.96E-07	6.42E-06
0.25	0.9618	2.42271	2.47E-07	2.01E-06	0.6846	2.78538	2.44E-07	2.00E-06	0.0187	0.337861	8.65E-07	6.29E-06
0.5	0.2256	0.73014	2.29E-07	1.94E-06	0.0895	0.476329	2.06E-07	1.84E-06	0.0019	0.0607979	8.90E-07	7.16E-06
0.75	0.0299	0.190804	2.35E-07	1.96E-06	0.0087	0.111464	2.21E-07	1.90E-06	0.0003	0.0173179	9.73E-07	6.47E-06
1	0.0122	0.114242	2.31E-07	1.95E-06	0.0017	0.041196	2.21E-07	5.39E-06	0	0	8.61E-07	6.31E-06
1.25	0.0112	0.109884	2.19E-07	1.90E-06	0.0003	0.0173179	1.82E-07	6.88E-06	0	0	7.87E-07	5.85E-06
1.5	0.0006	0.0244875	2.32E-07	1.96E-06	0.0001	0.0099995	2.60E-07	8.23E-06	0	0	8.58E-07	6.13E-06
1.75	0.0001	0.0099995	2.26E-07	1.93E-06	0	0	2.08E-07	7.36E-06	0	0	7.49E-07	5.54E-06
2	0.0001	0.0099995	2.15E-07	1.88E-06	0.0001	0.0099995	2.52E-07	4.10E-06	0	0	8.54E-07	5.61E-06
2.25	0.0001	0.0099995	2.39E-07	1.98E-06	0	0	2.19E-07	1.90E-06	0	0	6.09E-07	5.55E-06
2.5	0	0	2.46E-07	2.00E-06	0	0	2.01E-07	1.81E-06	0	0	3.75E-07	4.36E-06
						LLF						
0	495.076	473.145	4.16E-07	4.90E-06	522.696	527.33	3.24E-07	2.42E-06	0	0	2.60E-07	8.23E-06
0.25	15.2008	5.3951	3.24E-07	2.30E-06	6.6035	4.99663	3.14E-07	2.26E-06	0	0	1.82E-07	6.89E-06
0.5	6.7224	1.65636	4.47E-07	3.65E-06	2.163	1.50307	3.48E-07	2.39E-06	0	0	1.57E-07	6.37E-06
0.75	4.9824	1.16829	4.05E-07	2.58E-06	1.123	0.878562	3.10E-07	2.25E-06	0	0	2.23E-07	6.94E-06
1	3.813	0.912705	3.38E-07	2.35E-06	0.6609	0.660539	3.32E-07	2.32E-06	0	0	2.31E-07	6.98E-06
1.25	3.0444	0.770343	2.83E-07	2.15E-06	0.3802	0.529385	3.43E-07	2.36E-06	0	0	2.22E-07	5.93E-06
1.5	2.4436	0.671133	3.48E-07	2.38E-06	0.2101	0.419473	2.97E-07	2.20E-06	0	0	2.12E-07	7.37E-06
1.75	2.0438	0.600568	3.35E-07	2.34E-06	0.1009	0.302191	3.93E-07	2.54E-06	0	0	1.61E-07	6.38E-06
2	1.7136	0.570942	3.12E-07	2.25E-06	0.0654	0.248038	3.19E-07	2.29E-06	0	0	3.39E-07	9.38E-06
2.25	1.5086	0.546558	3.40E-07	2.38E-06	0.0335	0.179938	2.92E-07	2.18E-06	0	0	2.08E-07	7.36E-06
2.5	1.3524	0.503998	3.35E-07	2.33E-06	0.0158	0.1255	3.09E-07	2.25E-06	0	0	2.86E-07	8.63E-06
						Classical						
0	469.116	574.412	2.93E-07	2.18E-06	445.19	547.614	4.40E-07	3.84E-06	447.05	542.769	2.41E-07	1.99E-06
0.25	3.3321	5.24559	3.44E-07	2.37E-06	6.8104	9.65327	3.47E-07	2.38E-06	4.282	6.43761	2.41E-07	1.99E-06
0.5	0.6292	1.59019	3.63E-07	2.43E-06	0.4254	1.22883	3.36E-07	2.39E-06	0.6291	1.56567	2.61E-07	2.07E-06
0.75	0.0837	0.435998	3.05E-07	2.23E-06	0.3659	1.08463	3.61E-07	2.42E-06	0.128	0.566053	2.73E-07	2.11E-06
1	0.034	0.25699	2.88E-07	2.17E-06	0.0771	0.407376	3.63E-07	2.44E-06	0.0518	0.322671	2.51E-07	2.03E-06
1.25	0.0138	0.14356	2.79E-07	2.14E-06	0.0091	0.114966	3.22E-07	2.29E-06	0.0138	0.144947	2.54E-07	2.04E-06
1.5	0.0034	0.0691986	2.83E-07	2.16E-06	0.0032	0.0692081	3.19E-07	2.28E-06	0.0025	0.0607762	2.47E-07	2.01E-06
1.75	0.002	0.050951	2.59E-07	2.06E-06	0.0022	0.0528693	2.88E-07	2.17E-06	0.0014	0.0446994	2.56E-07	2.05E-06
2	0.0003	0.0173179	2.66E-07	2.09E-06	0.0007	0.0299918	3.26E-07	2.30E-06	0.0018	0.0468696	2.21E-07	1.90E-06
2.25	0	0	3.16E-07	2.27E-06	0	0	2.97E-07	2.20E-06	0.0002	0.0141407	2.18E-07	1.89E-06
2.5	0.0001	0.0099995	2.78E-07	2.13E-06	0	0	3.35E-07	2.33E-06	0	0	2.25E-07	1.92E-06

Table 4.4: ARL, SDRL, ATS, and SDTS Values for Bayesian EWMA Chart Hyper-Parameter Sensitivity Analysis with Poisson Conjugate

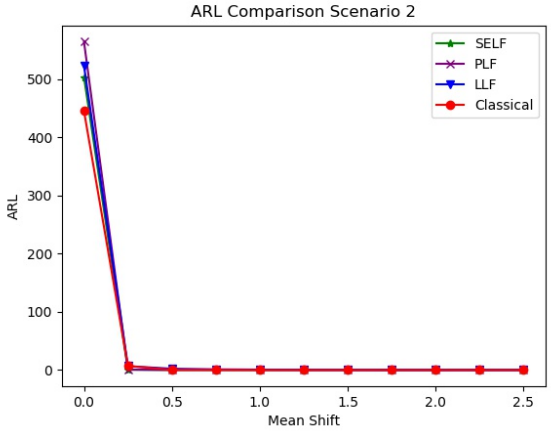
Loss function specific results for the hyper-parameter sensitivity analysis are given in table

4.4. Results for the squared error method showed a steep decrease in the ARL after the initial shift change of 0.25 was introduced and as we increased the hyper-parameter, this decrease became larger. We also noticed that as we increased the hyper-parameter, the SDRL increases and quicker detection also happens based on the ATS. These findings also translate to the precautionary case, except as the hyper-parameter increases under the PLF, the ATS increases. For the linex case, getting results for the third condition under the analysis with $n = 20$ and $\tau = 0.15$ was difficult. This is similar to the classical EWMA under each of the hyper-parameters. Results were received, but the number of simulation runs were significantly more than when using the SELF or PLF methods.

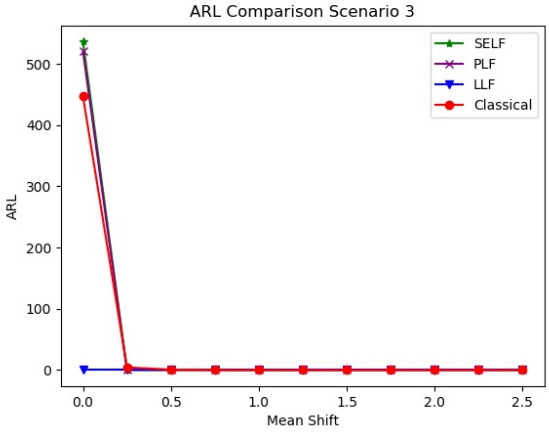
Besides the classical method, we conclude that the LLF method performed the worst under the hyper-parameter analysis. The ARL results suggest that the PLF was the best performer, but based on all four performance measurements, the SELF was the top performer.



(a) $\alpha = 16$, $\beta = \frac{5}{8}$

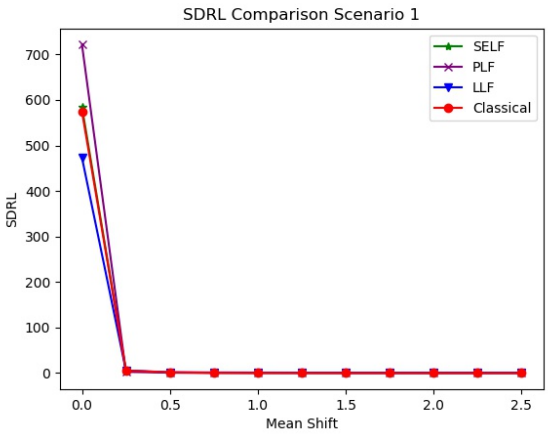


(b) $\alpha = 36$, $\beta = \frac{5}{12}$

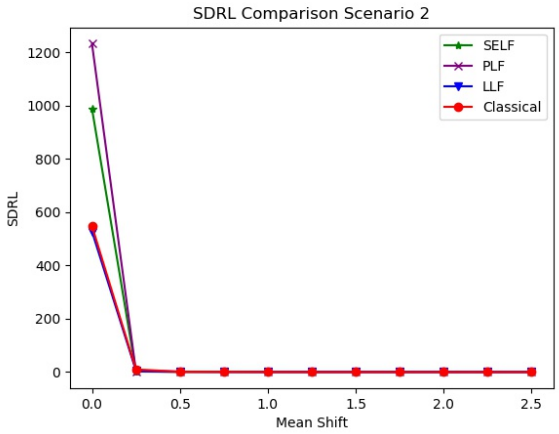


(c) $\alpha = 64$, $\beta = \frac{5}{16}$

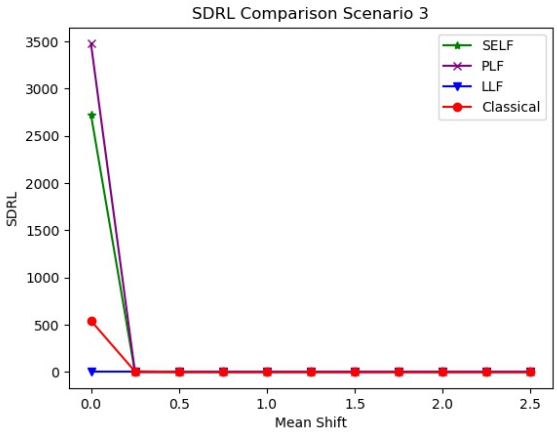
Figure 4.1: ARL Comparisons



(a) $\alpha = 16$, $\beta = \frac{5}{8}$



(b) $\alpha = 36$, $\beta = \frac{5}{12}$



(c) $\alpha = 64$, $\beta = \frac{5}{16}$

Figure 4.2: SDRL Comparisons

Figures 4.1 and 4.2 show the ARL and SDRL results from table 4.4 for each of the loss functions and the classical method.

4.3.2. Sample Size Analysis

The sample size sensitivity analysis was conducted using $n = \{20, 30, 40, 50\}$, setting $\tau = 0.15$ and adjusting the chart parameter, L, to attain $ARL_0 \approx 500$. Results from the analysis are in table

4.5.

		Shifts											
n	L	0	0.25	0.5	0.75	1	1.25	1.5	1.75	2	2.25	2.5	
						SELF							
20	-3.7	495.851	0.8169	0.0132	0.0032	0.0006	0.0006	0.0001	0.0001	0.0001	0	0	ARL
		(534.0910)	(1.9407)	(0.9694)	(0.7192)	(0.5475)	(0.5286)	(0.5093)	(0.2843)	(0.2984)	(0.3869)	(0.4422)	(SDRL)
		3.83E-07	3.37E-07	3.45E-07	3.04E-07	3.69E-07	3.81E-07	3.73E-07	3.69E-07	2.52E-07	3.74E-07	4.92E-07	ATS
		(6.70E-06)	(5.75E-06)	(5.33E-06)	(4.99E-06)	(5.07E-06)	(6.01E-06)	(5.29E-06)	(5.23E-06)	(5.01E-06)	(5.07E-06)	(5.23E-06)	(SDTS)
30	78	551.883	7.4582	4.1796	2.9312	2.1155	1.6517	1.4455	1.0171	0.9431	0.826	0.7363	ARL
		(313.336)	(6.77449)	(2.4378)	(1.34781)	(0.885223)	(0.662202)	(0.547244)	(0.410889)	(0.47444)	(0.410161)	(0.140624)	(SDRL)
		2.60E-07	1.67E-07	2.19E-07	1.97E-07	1.75E-07	1.89E-07	1.82E-07	1.97E-07	1.78E-07	1.78E-07	1.72E-07	ATS
		(7.14E-06)	(6.51E-06)	(8.30E-06)	(6.37E-06)	(5.10E-06)	(4.31E-06)	(8.79E-06)	(9.19E-06)	(7.40E-06)	(7.43E-06)	(1.00E-05)	(SDTS)
40	91	571.68	10.735	5.8342	3.8241	2.3965	2.0413	1.5571	1.1684	1.0935	0.9508	0.8813	ARL
		(551.8080)	(3.0739)	(1.3374)	(0.8558)	(0.5859)	(0.5093)	(0.5201)	(0.3858)	(0.3202)	(0.2675)	(0.3377)	(SDRL)
		3.70E-07	2.91E-07	2.74E-07	2.74E-07	2.79E-07	3.96E-07	3.18E-07	2.97E-07	3.12E-07	2.81E-07	2.91E-07	ATS
		(2.45E-06)	(2.18E-06)	(2.12E-06)	(2.12E-06)	(2.14E-06)	(2.53E-06)	(2.29E-06)	(2.20E-06)	(2.25E-06)	(2.14E-06)	(2.18E-06)	(SDTS)
50	150.5	519.9830	10.1315	5.8238	3.6264	2.8496	2.2708	1.8720	1.4312	1.1828	1.0606	0.9870	ARL
		(507.6630)	(2.2237)	(1.0494)	(0.6668)	(0.5562)	(0.4817)	(0.4150)	(0.4965)	(0.3873)	(0.2501)	(0.1779)	(SDRL)
		3.79E-07	2.97E-07	3.65E-07	4.22E-07	2.37E-07	2.68E-07	2.96E-07	2.37E-07	2.21E-07	2.70E-07	2.26E-07	ATS
		(5.69E-06)	(4.19E-06)	(4.94E-06)	(7.88E-06)	(5.97E-06)	(2.09E-06)	(6.55E-06)	(7.39E-06)	(6.93E-06)	(4.15E-06)	(7.38E-06)	(SDTS)
						PLF							
20	-1.3	533.02	0.9092	0.23	0.0582	0.0191	0.0115	0.0005	0.0012	0.0002	0.0002	0	ARL
		(843.3900)	(2.3949)	(0.7468)	(0.2814)	(0.1495)	(0.1121)	(0.0224)	(0.0374)	(0.0141)	(0)	(0)	(SDRL)
		3.43E-07	1.90E-07	2.86E-07	2.94E-07	2.12E-07	2.58E-07	2.35E-07	2.36E-07	2.60E-07	3.12E-07	2.60E-07	ATS
		(8.68E-06)	(6.90E-06)	(8.63E-06)	(8.64E-06)	(7.37E-06)	(3.25E-06)	(1.96E-06)	(7.81E-06)	(8.23E-06)	(9.02E-06)	(8.23E-06)	(SDTS)
30	35	571.1220	9.3364	3.8706	2.5543	1.5837	1.1444	0.8933	0.6677	0.4379	0.3461	0.1830	ARL
		(560.1300)	(4.0645)	(1.3741)	(0.9362)	(0.6414)	(0.5039)	(0.4511)	(0.4957)	(0.4999)	(0.4772)	(0.3867)	(SDRL)
		7.47E-07	9.50E-07	1.14E-06	8.62E-07	9.14E-07	9.58E-07	6.96E-07	8.69E-07	9.00E-07	1.14E-06	9.97E-07	ATS
		(5.19E-06)	(5.49E-06)	(1.02E-05)	(6.24E-06)	(7.47E-06)	(5.55E-06)	(4.68E-06)	(5.20E-06)	(5.61E-06)	(7.61E-06)	(6.87E-06)	(SDTS)
40	86	453.1480	10.3685	4.5154	2.8399	2.3914	1.7211	1.3709	1.1544	1.0223	0.9054	0.7773	ARL
		(435.9740)	(3.0808)	(1.0119)	(0.6610)	(0.5891)	(0.5139)	(0.4908)	(0.3794)	(0.2720)	(0.3182)	(0.4175)	(SDRL)
		6.89E-07	9.27E-07	8.86E-07	9.99E-07	9.41E-07	9.89E-07	9.55E-07	8.66E-07	9.93E-07	9.63E-07	9.23E-07	ATS
		(3.23E-06)	(3.72E-06)	(3.62E-06)	(4.46E-06)	(3.78E-06)	(4.64E-06)	(3.85E-06)	(3.61E-06)	(6.19E-06)	(3.80E-06)	(3.68E-06)	(SDTS)
50	157	499.929	10.4253	6.3638	4.0256	2.9648	2.3144	1.9576	1.5789	1.2491	1.0912	0.9895	ARL
		(478.347)	(2.28837)	(1.13236)	(0.703523)	(0.564235)	(0.493105)	(0.390643)	(0.498773)	(0.433416)	(0.294419)	(0.166102)	(SDRL)
		5.58E-07	4.17E-07	4.60E-07	3.51E-07	3.62E-07	5.35E-07	2.73E-07	3.43E-07	1.93E-07	3.33E-07	4.24E-07	ATS
		(6.25E-06)	(6.07E-06)	(5.73E-06)	(5.08E-06)	(4.97E-06)	(5.50E-06)	(7.00E-06)	(9.04E-06)	(6.79E-06)	(8.96E-06)	(3.63E-06)	(SDTS)
						LLF							
20	9	569.539	6.9163	2.1641	1.1105	0.6095	0.371	0.2089	0.11	0.056	0.0334	0.0194	ARL
		(564.402)	(5.01854)	(1.49183)	(0.881073)	(0.64731)	(0.522072)	(0.420786)	(0.315119)	(0.23079)	(0.179679)	(0.137926)	(SDRL)
		3.15E-07	7.62E-07	7.25E-07	6.69E-07	5.58E-07	5.42E-07	9.25E-07	5.88E-07	5.54E-07	5.68E-07	5.29E-07	ATS
		(2.34E-06)	(7.55E-06)	(3.41E-06)	(3.60E-06)	(3.03E-06)	(2.99E-06)	(9.43E-06)	(3.10E-06)	(2.99E-06)	(3.03E-06)	(2.94E-06)	(SDTS)
30	51	490.549	9.766	4.7005	3.0601	2.1493	1.6644	1.309	1.0308	0.9438	0.8303	0.6172	ARL
		(472.874)	(3.371)	(1.301)	(0.861677)	(0.644678)	(0.57059)	(0.493071)	(0.372091)	(0.3642)	(0.406327)	(0.490371)	(SDRL)
		3.44E-07	3.28E-07	3.27E-07	3.40E-07	3.53E-07	3.36E-07	5.33E-07	4.09E-07	4.01E-07	3.78E-07	3.57E-07	ATS
		(2.37E-06)	(2.31E-06)	(2.31E-06)	(2.35E-06)	(2.39E-06)	(2.34E-06)	(7.28E-06)	(2.57E-06)	(2.56E-06)	(2.48E-06)	(2.44E-06)	(SDTS)
40	104.8	550.602	10.1719	5.6892	3.5607	2.7699	2.1549	1.6974	1.3898	1.1228	1.0275	0.9842	ARL
		(542.454)	(2.58313)	(1.13649)	(0.717715)	(0.599962)	(0.48673)	(0.499633)	(0.491992)	(0.341057)	(0.249286)	(0.230544)	(SDRL)
		3.07E-07	3.84E-07	6.50E-07	3.58E-07	3.71E-07	3.67E-07	5.29E-07	5.07E-07	3.58E-07	4.29E-07	3.62E-07	ATS
		(2.24E-06)	(2.50E-06)	(1.16E-05)	(2.42E-06)	(2.45E-06)	(2.44E-06)	(4.22E-06)	(2.92E-06)	(2.41E-06)	(2.79E-06)	(2.43E-06)	(SDTS)
50	169.5	454.27	9.9823	5.8387	4.0167	2.9774	2.4399	2.0352	1.7157	1.3985	1.1213	1.0251	ARL
		(432.075)	(2.00589)	(0.960564)	(0.680603)	(0.53767)	(0.515546)	(0.3717)	(0.464407)	(0.489998)	(0.329221)	(0.182948)	(SDRL)
		3.24E-07	2.97E-07	3.41E-07	3.50E-07	3.28E-07	2.86E-07	4.39E-07	3.03E-07	3.26E-07	3.31E-07	3.08E-07	ATS
		(2.30E-06)	(2.21E-06)	(2.36E-06)	(2.39E-06)	(2.31E-06)	(2.16E-06)	(6.30E-06)	(2.22E-06)	(2.30E-06)	(2.32E-06)	(2.24E-06)	(SDTS)

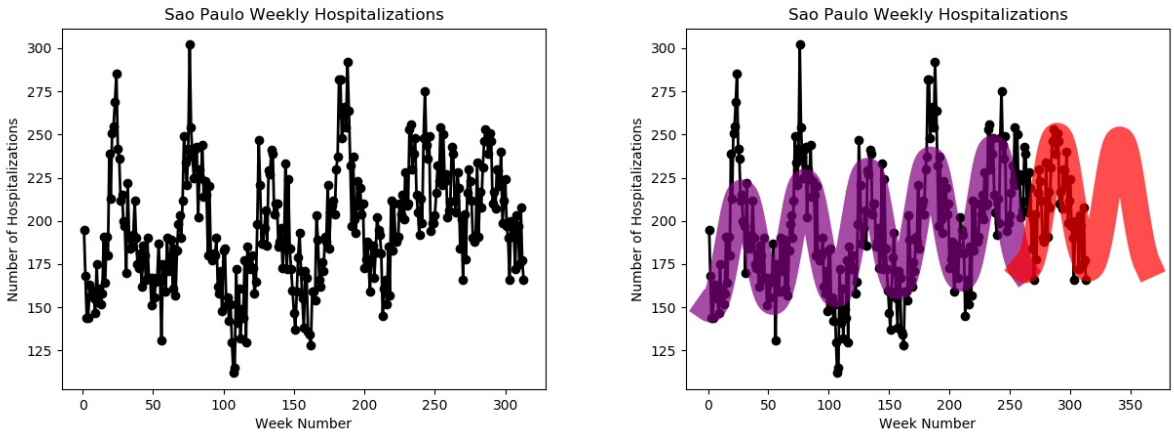
Table 4.5: ARL, SDRL, ATS, and SDTS Values for Bayesian EWMA Chart Sample Size Sensitivity Analysis with Poisson Conjugate

We see from table 4.5 that for the SELF, PLF, and LLF methods, as the sample size increased, the L tuning parameter had to be increased to reach the desirable ARL. Also, we note that there wasn't a significant change in ATS, but we did take note that it does take slightly less time for detection with the sample size increase. The most interesting result for all of the loss functions was as we increase the sample size, as we implemented the initial mean shift, the distance between the ARL_0 and out-of-control ARL grew smaller. That is, the larger the sample size, the likelihood of the chart to signal decreases.

Observation of the individual performance of the each loss function shown that under the SELF we recorded the lowest values for L. For the PLF, the SDRL values decreased as we increased the sample size and the LLF performed sub-par compared to the other loss functions based on ARL values.

4.4. Data Application Study

The chosen data is a count series that has been seen in Alencar et al., 2017, Urbieta et al., 2017, and Jones et al., 2021. It features daily count records of respiratory disease related hospitalizations for senior citizens in São Paulo, Brazil and is available to the public at www2.datasus.gov.br. Daily data was collected from January 2006 to December 2011, but since our method tracks Bayesian estimators, we chose to observe the average of the weekly counts (amounting to 313 weeks).



(a) Raw data (b) Modeled data

Figure 4.3: Weekly Hospitalizations in São Paulo

The data displays a positive linear trend with a seasonal pattern (figure 4.3). The seasonal trend reflects an increase of hospitalizations in June - August compared to December - February, which are the winter and summer months in São Paulo, respectively.

We use a negative binomial distribution under a generalized linear model and account for the seasonal pattern using sine and cosine functions. We model hospitalization rate per 100,000 inhabitants using an offset variable for the population size and the log link function. To validate our model, we assume a non-epidemic range from January 2006 to December 2010 as our in-control data and validate with data from January to December 2011. Once validation is complete, we predict values for January to December 2012 and impose shift changes on January 2011 - December 2012 data to represent an epidemic. Our objective is to emulate and provide quick detection of an epidemic outbreak. Our model is given in equation 4.5.

$$ln\left[\frac{\mu_{0,t}}{P_t}100,000\right] = \beta_0 + \beta_1 \sin\left(\frac{2\pi t}{52}\right) + \beta_2 \cos\left(\frac{2\pi t}{52}\right) \tag{4.5}$$

where $t = 1,\ldots,313$ weeks, $\mu_{0,t}$ is the average of non-epidemic hospitalizations for week t,

and P_t is the population size at week t.

In figures 4.4, 4.5, and 4.6 below are the results for the hospitalization data under each of the loss functions with shift increments added. Under the SELF and PLF cases, we notice both charts detect an epidemic after the initial shift of 0.25 is implemented. For the LLF chart, it detects an epidemic without any shift, which indicates the user-defined c parameter for the LLF may need to be adjusted.

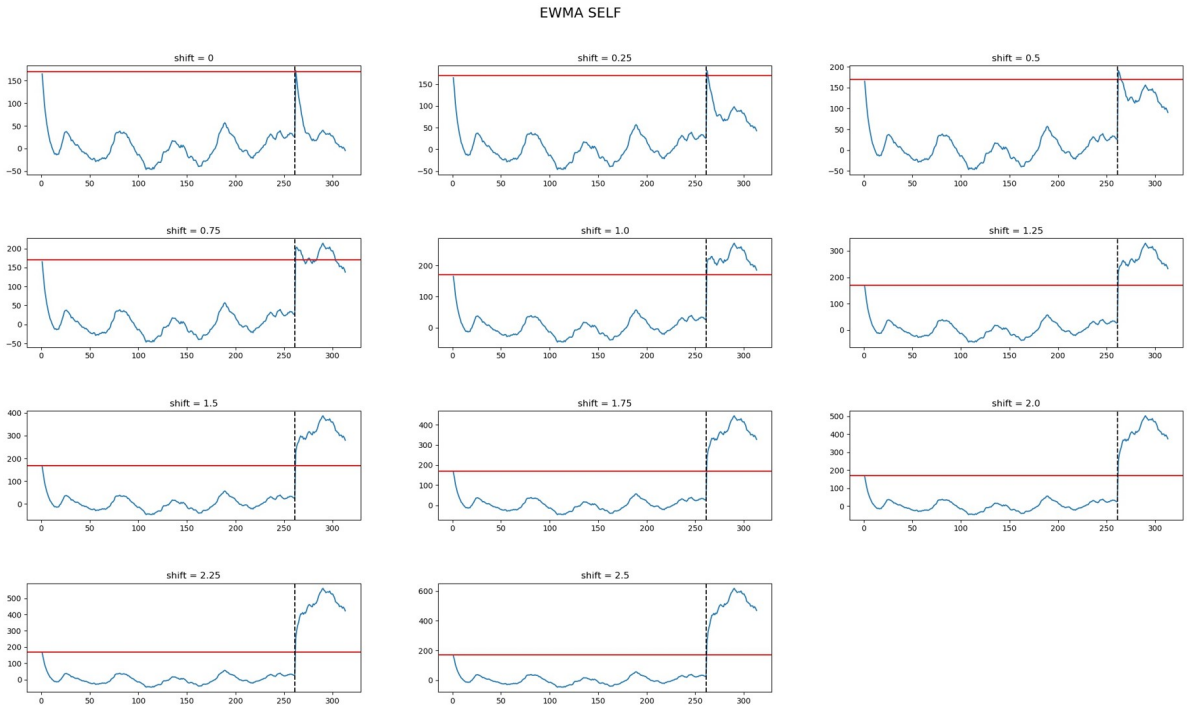


Figure 4.4: Squared Error Loss Function EWMA Chart Results

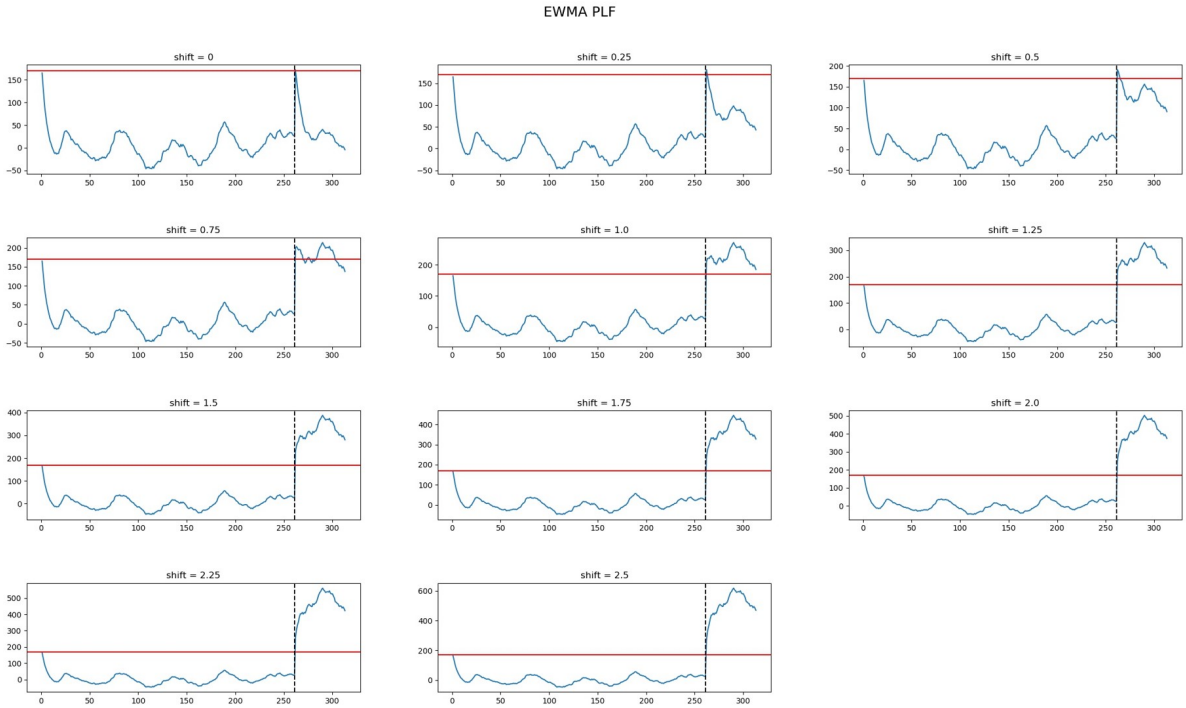


Figure 4.5: Precautionary Loss Function EWMA Chart Results

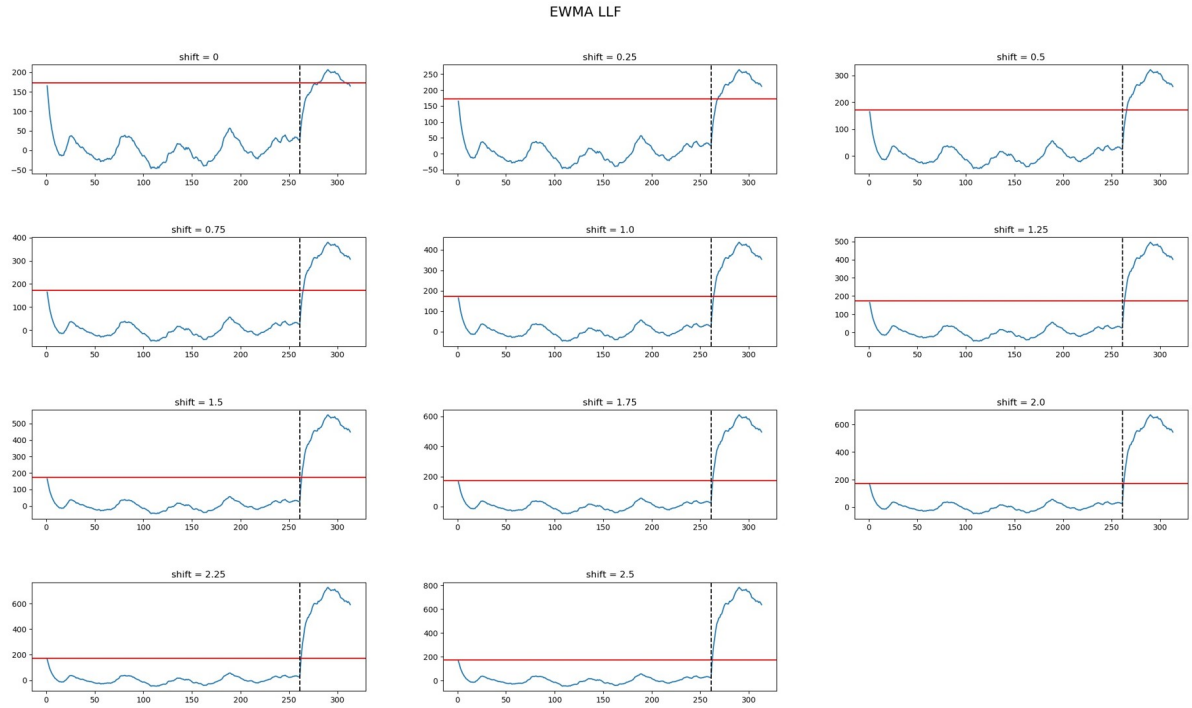


Figure 4.6: Linex Loss Function EWMA Chart Results

4.5. Conclusions and Recommendations

In our work, we examined the detection capabilities of the EWMA chart under different loss functions for the Poisson-Gamma conjugate case. We conducted simulations, implementing shifts on the in-control mean, while adjusting the hyper-parameter values and the sample size. The performance measurements used were the average run length (ARL), standard deviation of the run length (SDRL), average time to signal (ATS), and the standard deviation of the time to signal (SDTS), which were recorded for the in-control case and each of the out-of-control shifts. After simulation results were analyzed, we tested the charts' ability to detect epidemic-like instances on respiratory disease related hospitalizations for seniors in São Paulo, Brazil.

From the simulation study and real data analysis, the squared error and precautionary

loss functions performed similarly, with the PLF having better ARL results and the SELF outperforming overall. The linex loss function had sub-optimal results in both simulation studies and the real data analysis compared to the other options. In the hyper-parameter study, calculations for the ARL and SDRL for $\alpha_0 = 64, \beta_0 = \frac{5}{16}$ resulted in either zeros or unending simulations. Compared to the classical EWMA chart, the SELF and PLF are ideal substitutes when considering count series data, while the LLF performed roughly the same. It's also important to note that simulations for the classical EWMA required significantly more runs than the Bayesian charts. We recommend the use of the squared error loss function Bayesian control chart as it performed the best in most of the measurement criteria. The precautionary loss function would be the best option for use of an asymmetric loss function, and we only recommend the linex loss function when a valid study is done to attain the c value.

Bibliography

Roberts, S. (1959). Control chart tests based on geometric moving averages. *Technometrics*, *1*(3), 239–250.

Page, E. S. (1954). Continuous inspection schemes. *Biometrika*, *41*(1/2), 100–115.

Abbas, T., Ahmad, S., Riaz, M., & Qian, Z. (2019). A bayesian way of monitoring the linear profiles using cusum control charts. *Communications in Statistics-Simulation and Computation*, *48*(1), 126–149.

Ali, S. (2020). A predictive bayesian approach to ewma and cusum charts for time-between-events monitoring. *Journal of Statistical Computation and Simulation*, *90*(16), 3025–3050.

Tsiamyrtzis, P., & Hawkins, D. M. (2008). A bayesian ewma method to detect jumps at the start-up phase of a process. *Quality and Reliability Engineering International*, *24*(6), 721–735.

Aslam, M., & Anwar, S. M. (2020). An improved bayesian modified-ewma location chart and its applications in mechanical and sport industry. *PloS one*, *15*(2), e0229422.

Noor-ul Amin, M., & Noor, S. (2021). An adaptive ewma control chart for monitoring the process mean in bayesian theory under different loss functions. *Quality and Reliability Engineering International*, *37*(2), 804–819.

Riaz, S., Riaz, M., Hussain, Z., & Abbas, T. (2017). Monitoring the performance of bayesian ewma control chart using loss functions. *Computers & Industrial Engineering*, *112*, 426–436.

Jones, C. L., Abdel-Salam, A.-S. G., & Mays, D. (2021). *Generalized bayesian cusum charts using different loss functions* [Under Review].

Norstrom, J. G. (1996). The use of precautionary loss functions in risk analysis. *IEEE Transactions on reliability*, *45*(3), 400–403.

Alencar, A. P., Lee Ho, L., & Albarracin, O. Y. E. (2017). Cusum control charts to monitor series of negative binomial count data. *Statistical methods in medical research*, *26*(4), 1925–1935.

Urbieta, P., Lee HO, L., & Alencar, A. (2017). Cusum and ewma control charts for negative binomial distribution. *Quality and Reliability Engineering International*, *33*(4), 793–801.

Chapter 5

Bayesian mEWMA and mCUSUM Control Charts on Nonparametric and Semiparametric Linear Regression Models

Abstract

Bayesian multivariate cumulative sum (mCUSUM) and multivariate exponentially weighted moving average (mEWMA) charts are proposed in this work to monitor count series data. The charts are informed by the squared error loss function (SELF) and are tested under a simulation study while tuning the choice of hyper-parameters and sample sizes respectively. The charting methods are compared using the average and standard deviation of the run length (ARL and SDRL) and the average and standard deviation of the time to signal (ATS and SDTS). Both charts are assessed when used on data that is modeled using parametric, nonparametric, and semiparametric regressioin methods. The regression techniques considered are the penalized splines (p-splines) for the nonparametric and the model robust regression 1 (MRR1) for the

semiparametric. Model fits are compared using the mean squared error (MSE), Akaike information criterion (AIC), and Bayesian information criterion (BIC). The charts' ability to monitor live data is examined on suicide counts data modeled using each of the regression techniques.

Keywords: Profile Monitoring; Bayesian; Squared Error Loss Function, mCUSUM; mEWMA; Sensitivity Analysis; Quality Control; P-spline; MRR1

5.1. Introduction

Profile monitoring is a area in statistical quality control (SQC) that allows for the use of graphical tools to observe changes in a process. These graphical tools, known as control charts, come in many variations to best suit the data application and distribution, where many have been made to accommodate Gaussian data. Classical charts like the Shewhart, cumulative sum (CUSUM), and exponentially weighted moving average (EWMA) have been frequented as charts that monitor univariate data. Along with the classical approach, recent research has broadened on the Bayesian approach to monitoring univariate Guassian data. Riaz et al., 2017 introduces an approach using loss functions to inform a Bayesian EWMA chart, while Noor-ul Amin and Noor, 2021 and Jones et al., 2021 extended this idea to an adaptive EWMA (AEWMA) chart and the CUSUM chart respectively.

The Hotelling's T^2, multivariate CUSUM and EWMA (mCUSUM and mEWMA) have been advantageous alternatives to monitor data adorning multivariate characteristics. Our work expands the Bayesian charting methods from the aforementioned articles to the mCUSUM and mEWMA charts after estimating model parameters using techniques mentioned in Jones et al., 2020. In section 5.2 we introduce the loss function we use to inform our Bayesian charts, section 5.3 details the parametric, nonparametric, and semiparametric methods implemented to

estimate a model. We then define our Bayesian multivariate charts in section 5.4 and conduct a simulation analysis study followed by application of our methods on suicide count data to inform suicide prevention strategies in sections 5.5 and 5.6 respectively. Recommendations and discussion on the performance of the proposed methods are featured in section 5.7.

5.2. Squared Error Loss Function

The use of loss functions in Bayesian statistics has proven useful in obtaining the Bayes estimator, or best estimate, of a given parameter. To obtain the best estimator for the parameter, the loss function minimizes the expected loss from using parameter estimate. The first step in this process is for the user to choose which loss function best suites their needs. In Jones et al., 2021, the Guassian and Poisson conjugate cases are considered to recommend generalized Bayesian CUSUM and EWMA control charts. Of the loss functions used, the authors determined the precautionary loss function (PLF) and squared error loss function (SELF) performed similarly under their Bayesian CUSUM charts and *EWMA article* recommended the SELF as it performed best overall. Based on the results and recommendations from these articles, our work will focus on using the squared error loss function.

We define the Bayes estimator as a value that minimizes the expected loss as: $\hat{\theta}^* = \min_{\hat{\theta}} E_\theta[L(\hat{\theta},\theta]$. The squared error loss function is defined as $L(\hat{\theta},\theta) = (\theta-\hat{\theta})^2$ and its expected loss is defined as $E[\theta|x]$. The form of this loss function makes it symmetric, where both the negative and positive errors are weighted equally.

5.3. Regression Methods

For our methods, we consider a Poisson likelihood, $Poisson(\lambda)$, with a Negative Binomial prior, $Neg-Bin(\alpha,\beta)$, resulting in a Gamma posterior distribution where the shape and inverse scale parameters are $(n\bar{x}+\alpha)$ and $(n+\beta)$ respectively ($Gamma(n\bar{x}+\alpha, n+\beta)$). Combining the posterior distribution with the likelihood function of future data, we derive the posterior predictive distribution to be $Neg-Bin(n\bar{x}+\alpha, n+\beta)$.

A log link function is used on the posterior predictive data to obtain a simple linear model of the form $y_i = \beta_0 + \beta_1 X_i + \epsilon_i$ for $i = 1,2,\ldots,n$ = number of observations. The subsequent subsections detail how we use nonparametric and semiparametric methods to estimate values for β_0 and β_1, which we track in our Bayesian control charts. These coefficient estimates are defined as a vector of estimates: $\hat{\boldsymbol{\beta}} = \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{bmatrix}$. Since we are interested in more than one estimate, we will use multivariate control charts for monitoring. Section 5.4 details the charts we use and their Bayesian framework.

We assess the estimation capability of the parametric, nonparametric, and semiparametric methods using three performance measurement tools: mean squared error (MSE), Akaike's information criterion (AIC), and Bayesian information criterion (BIC). Each of the criteria is minimized, meaning that the model with the lowest value under the respective criteria should be chosen.

5.3.1. Nonparametric

Nonparametric (NP) techniques are frequented when a parameters are unknown for a specified distribution or when the underlying distribution is unknown. There are many different NP methods, such as kernels and splines, which have a multitude of variants based on the users'

needs. Penalized splines (p-splines) are a variant of B-splines that obtain a smoother fit by imposing a penalty factor. The conception of p-splines were initially detailed in O'Sullivan, 1986 and its popularity grew with discussion papers such as Eilers and Marx, 1996. Eilers et al., 2015 provides an in-depth review of the use of p-splines from 1995 to 2015 in their discussion paper. In their practitioners guide, Jones et al., 2020 reviews several NP regression methods alongside control charting techniques. A recommendation for the p-spline method with a linear model on univariate data was made based on optimal results and simple implementation. We will use the p-spline method as defined in equation 5.1 as our NP method.

$$\hat{y} = \boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{X} + \Lambda^{2p}D)^{-1}\boldsymbol{X}^T y \tag{5.1}$$

where Λ is the penalizing parameter, p is the total number of columns in the $\boldsymbol{X}$ matrix, and D is a lower diagonal matrix based on the number of knots.

5.3.2. Semiparametric

Under semiparametric (SP) regression, aspects from a parametric approach are combined with results of a NP approach using a convex combination to form a clearer picture of the data behavior. Works such as Ruppert et al., 2003, Keele and Keele, 2008, Ruppert et al., 2009 give detailed explanations on the uses of SP methods. We use the model robust regression 1 (MRR1) method introduced in Einsporn, 1987 and Einsporn and Birch, 1993 for modeling mean response, conceptualized in Burman and Chaudhuri, 2012 and Mays et al., 2000, with its term coined in the latter.

$$\hat{y}^{MRR1} = (1-\eta)\hat{y}^{P} + \eta\hat{y}^{NP} \quad \text{for } \eta \in [0,1] \tag{5.2}$$

where η is the mixing parameter such that if $\eta = 0$, the model is soley parametric and if $\eta = 1$, the model is NP. The MRR1 method seen in equation 5.2 uses estimates of y determined from both the parametric ($\hat{y}^{P}$) and NP ($\hat{y}^{NP}$) models, joining them using the user-defined mixing parameter to obtain the SP y estimates. When applying the MRR1 method to real data it's acceptable to choose η based on the parametric and nonparametric fits, however, Mays et al., 2001 derived a data-driven equation (eq. 5.3) to optimize the mixing parameter.

$$\eta = \frac{(\hat{y}^{NP}_{-i,i} - \hat{y}^{P}_{-i,i})^{T}(y - \hat{y}^{P})}{(\hat{y}^{NP} - \hat{y}^{P})^{T}(\hat{y}^{NP} - \hat{y}^{P})} \tag{5.3}$$

where $\hat{y}^{P}_{-i,i}$ and $\hat{y}^{NP}_{-i,i}$ are calculated by leaving out the i^{th} observation when estimating the $\hat{y}$ value at x_i for the parametric and nonparametric methods respectively.

5.4. Multivariate Charts

As stated in section 5.3, we are monitoring the changes that occur in the vector $\hat{\boldsymbol{\beta}}$, containing two coefficient estimates. The CUSUM (Page, 1954, Page, 1961) and EWMA (Roberts, 1959) charts are widely known and used charts for tracking univariate observations. Their use, however, can be extended to the multivariate case as seen in the following subsections.

5.4.1. mCUSUM

As the multivariate extension of the CUSUM, the mCUSUM chart (Woodall and Ncube, 1985) has the same structure. The chart sums the previous statistic calculations together and obtains the current location's statistic. Our Bayesian mCUSUM chart retains the same mechanics as the classical version, but we replace the in-control mean with our Bayes estimator, μ_{SELF}.

$$
\begin{aligned}
&mC_i = max\{0, (\boldsymbol{C_i^T \Sigma_0^{-1} C_i})^{\frac{1}{2}} - k\}, \\
&\text{for } i = 1, 2, \ldots, m = \text{number of profiles} \\
&\text{where,} \\
&\boldsymbol{C_i} = (\underline{\hat{\boldsymbol{\beta}}}_i - \boldsymbol{\mu}_{SELF}) \\
&k = \frac{1}{2}\sqrt{(\hat{\boldsymbol{\beta}}_i - \boldsymbol{\mu}_{SELF})^T \boldsymbol{\Sigma}^{-1} (\hat{\boldsymbol{\beta}}_i - \boldsymbol{\mu}_{SELF})}
\end{aligned}
\tag{5.4}
$$

5.4.2. mEWMA

The multivariate EWMA (mEWMA) (Lowry et al., 1992) weights the previous statistic with a user defined variable, r, where $r \in [0, 1]$. Similar to the EWMA chart if the weight parameter is 1, then a memory-less chart is produced, in this case the Hotelling's T^2 chart. A r value closer to 0 places more weight on the previous statistics, giving the chart a longer memory. Again, we begin with using the same structure of the classical chart and substitute our Bayes estimator for the in-control mean.

$$\boldsymbol{z_i} = r \cdot \hat{\boldsymbol{\beta}}_{\boldsymbol{i}} + (1-r)\boldsymbol{z_{i-1}},$$

$$\text{for } i = 1, 2, \ldots, m = \text{number of profiles,}$$

$$\text{where,} \tag{5.5}$$

$$T_i^2 = (\boldsymbol{z_i} - \boldsymbol{\mu}_{SELF})^T \boldsymbol{\Sigma}_{\boldsymbol{z_i}}^{-1} (\boldsymbol{z_i} - \boldsymbol{\mu}_{SELF})$$

$$\boldsymbol{\Sigma}_{\boldsymbol{z_i}} = cov(z_i) = \frac{r}{2-r}\boldsymbol{\Sigma_0}$$

5.5. Analysis

We assess the capability of our Bayesian mCUSUM and mEWMA control charts using simulation studies. This section will provide results and comparisons from a sensitivity analysis done by tuning hyper-parameters (section 5.5.1). Our charts are designed to reach an in-control ARL of 500 without a shift, and are assessed under different shift sizes ranging from 0 to 2.5 with 0.25 step sizes ($\delta = 0$ to 2.5 by 0.25). As a standard in control charting, we calculate $m = 10,000$ simulations to shrink the standard error. The initial data of size $n = 20$ is drawn from $Poisson(\lambda = 25)$ with shift sizes calculated as $\hat{\boldsymbol{\beta}}_{ooc} = \hat{\boldsymbol{\beta}}_{ic} * (1+\delta)$, where $\hat{\boldsymbol{\beta}}_{ic}$ and $\hat{\boldsymbol{\beta}}_{ooc}$ are the in-control and out-of-control $\hat{\boldsymbol{\beta}}$ values, respectively.

5.5.1. Hyper-Parameter Sensitivity Analysis

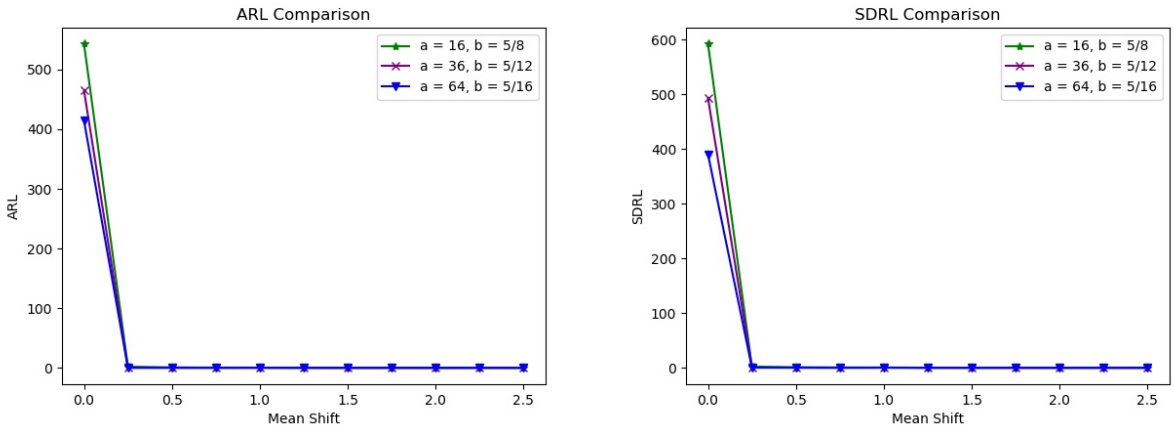
Under our hyper-parameter study, we select three choices for our mean and standard deviation as $\mu_0 = [10, 15, 20]$ and $\sigma_0 = [4, 6, 8]$, respectively. We solve for our priors, α_0 and β_0, using the mean and standard deviation and obtain $\alpha_0 = [16, 36, 64]$ and $\beta_0 = \left[\frac{5}{8}, \frac{5}{12}, \frac{5}{16}\right]$. Each sample drawn is of size $n = 20$ from the posterior predictive Negative-Binomial distribution.

5.5.1.1. mEWMA

	$\alpha_0 = 16, \beta_0 = \frac{5}{8}$							$\alpha_0 = 36, \beta_0 = \frac{5}{12}$							$\alpha_0 = 64, \beta_0 = \frac{5}{16}$						
Shifts	ARL	SDRL	ATS	SDTS	MSE	AIC	BIC	ARL	SDRL	ATS	SDTS	MSE	AIC	BIC	ARL	SDRL	ATS	SDTS	MSE	AIC	BIC
											Parametric										
0	543.15	592.852	8.87E-05	0.00012365	0.351716	0.771192	-21.6351	465.25	493.571	7.74E-05	2.06E-05	0.349271	2.54348	-39.3579	414.392	390.733	0.000106142	4.83E-05	0.356772	1.95147	-33.4379
0.25	2.09	2.53809	6.77E-05	0.000114236	0.940765	-0.189589	-12.0273	0.3	0.685565	8.40E-05	3.85E-05	0.971847	-1.2443	-1.48021	0.154	0.407779	8.03E-05	2.68E-05	1.00736	-1.39201	-0.00311628
0.5	0.82	1.30675	6.77E-05	0.000120049	2.82648	-3.35169	19.5938	0.17	0.401373	9.01E-05	5.48E-05	2.7724	-3.78423	23.9191	0.189	0.497272	8.79E-05	4.93E-05	2.73005	-3.88011	24.878
0.75	0.46	0.887919	7.03E-05	0.00011563	5.48978	-4.84705	34.5474	0.11	0.343366	7.75E-05	1.96E-05	5.78616	-5.28818	38.9587	0.05	0.217945	7.52E-05	2.27E-05	5.73704	-5.16691	37.7459
1	0.44	0.765768	7.81E-05	0.000119358	9.27219	-6.24413	48.5181	0.13	0.39128	7.83E-05	2.86E-05	9.82264	-6.54198	51.4966	0.176	0.461545	8.09E-05	3.59E-05	9.61263	-6.54123	51.4892
1.25	0.18	0.40939	7.29E-05	0.000116943	14.0486	-7.06412	56.718	0.07	0.255147	8.03E-05	2.50E-05	13.4814	-7.1086	57.1628	0.059	0.243965	8.02E-05	2.52E-05	14.9658	-7.28178	58.8946
1.5	0.21	0.453762	6.25E-05	0.000111213	20.2595	-7.76993	63.7761	0.02	0.14	7.75E-05	2.09E-05	20.5145	-7.83463	64.4231	0.026	0.171243	8.08E-05	3.83E-05	20.7503	-7.87306	64.8074
1.75	0.21	0.515655	8.33E-05	0.000126944	29.6012	-8.5088	71.1648	0.03	0.170587	8.12E-05	2.13E-05	26.4954	-8.31993	69.2761	0.013	0.113274	8.78E-05	4.77E-05	27.7965	-8.42742	70.351
2	0.1	0.331662	8.07E-05	0.000120455	34.8635	-8.8541	74.6178	0.02	0.14	7.68E-05	2.00E-05	38.3375	-9.11179	77.1947	0.015	0.121552	7.77E-05	2.38E-05	37.6731	-9.06734	76.7502
2.25	0.29	0.604897	8.07E-05	0.000120451	52.2758	-9.37342	79.8111	0.02	0.14	8.58E-05	3.08E-05	49.5177	-9.56514	81.7282	0.008	0.0890842	8.13E-05	3.20E-05	49.5922	-9.53729	81.4497
2.5	0.17	0.510979	8.33E-05	0.000121488	59.0063	-9.85301	84.6069	0.01	0.0994987	9.46E-05	7.40E-05	57.9132	-9.92115	85.2884	0.012	0.108885	9.27E-05	5.40E-05	58.5002	-9.93969	85.4737

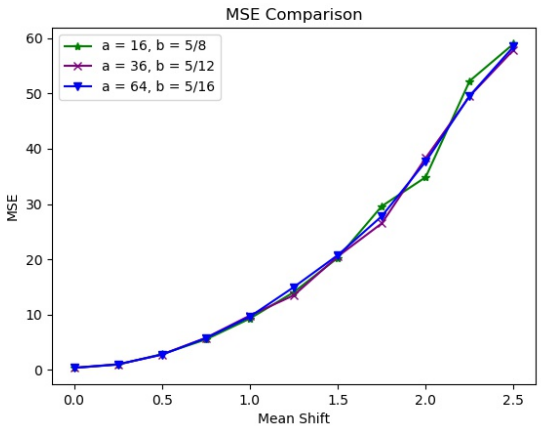
Table 5.1: mEWMA Hyper-Parameter Sensitivity Analysis

Table 5.1 above, along with figures 5.1 and 5.2 below, show results for the Bayesian mEWMA chart proposed in section 5.5. Values within the table and figure 5.1a suggest that regardless of the values chosen for the hyper-parameters, the Bayesian mEWMA chart will detect when an out-of-control situation occurs. This is seen from the sharp decrease in the ARL_0 after a shift is applied. The MSE, AIC, and BIC comparisons are useful when determining if a change in hyper-parameter influences the fit of the model. In this study, we notice that in the scenario which $\alpha = 16$, $\beta = \frac{5}{8}$ there is a slight inconsistency for each fit comparison method. Both figures 5.2b and 5.2c show an irregularity once a shift is imposed on the data, and figure 5.2a sees a similar irregularity with larger shifts. The MSE and BIC values suggest that the as the out-of-control shifts increase the model that is fit becomes worse, but the AIC suggest the opposite.

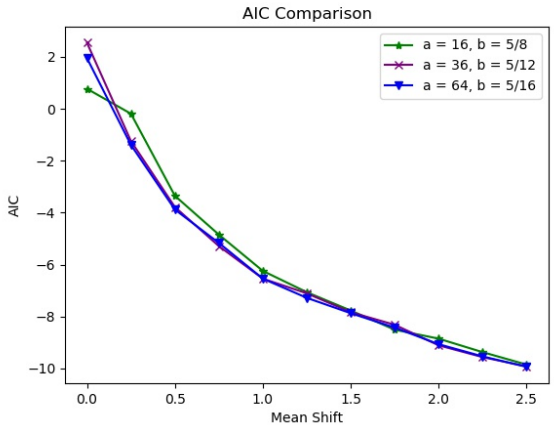


(a) ARL Comparisons (b) SDRL Comparisons

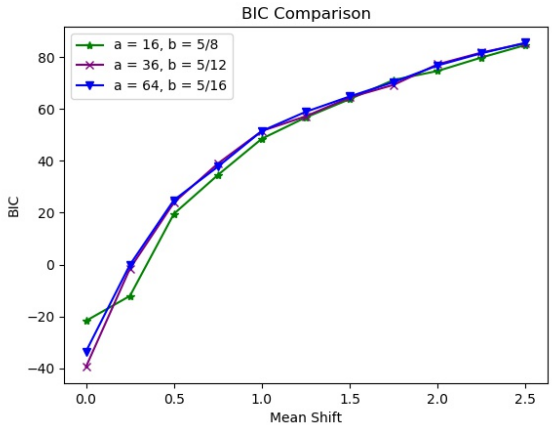
Figure 5.1: ARL and SDRL Comparisons



(a) MSE Comparisons



(b) AIC Comparisons



(c) BIC Comparisons

Figure 5.2: Model Fit Comparisons

5.6. Real Data Application

There has been worldwide stigma surrounding mental health throughout history, making it difficult or impossible to obtain the resources and information necessary for treatment. Effects of mental health disorders, such as depression, has shown in some cases to lead to death by suicide. In September 2014, the World Health Organization (WHO) recognized that suicide was a global problem, releasing their first suicide prevention plan with their objective being "to

prioritize suicide prevention on the global public health and public policy agendas and to raise awareness of suicide as a public health issue" (WHO et al., 2014). Studies done by Zortea et al., 2020, Leaune et al., 2020, Gunnell et al., 2020, and Pirkis et al., 2021 have looked at how the Covid-19 public health emergency has impacted mental health and suicide rates. Cénat et al., 2020 concluded that regardless of the gender, group or region, the mental health of populations affected by the coronavirus are being adversely impacted. We aim to use our Bayesian charting techniques to track suicide rate trends from previous years to prepare for a potential increase due to the current public health emergency. Implications of our method could aid in targeting areas/countries for immediate suicide prevention resources (i.e increase governmental funding/grants for organizations to provide increased mental health treatments).

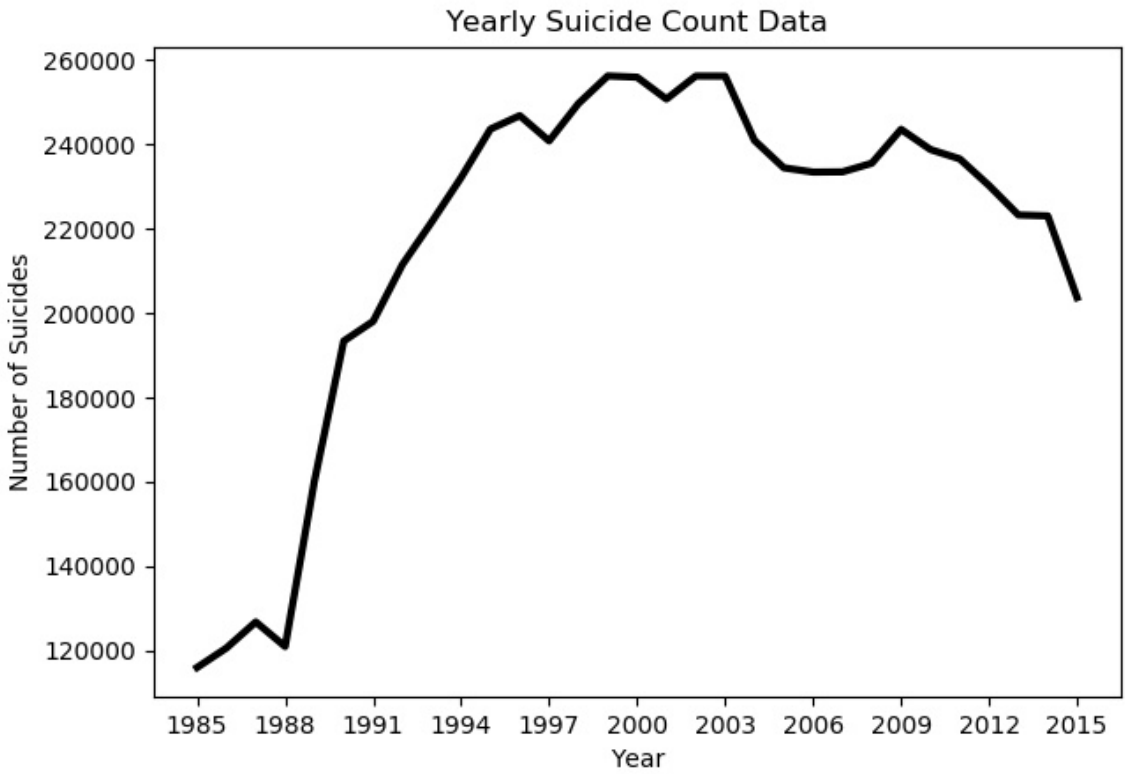
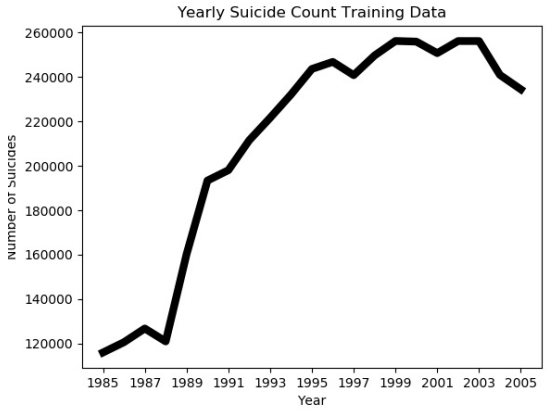


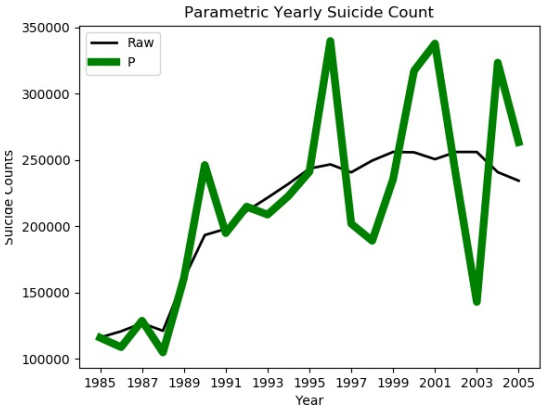
Figure 5.3: Raw Data

$$counts_i = \beta_0 + \beta_1 year_i + \beta_2 GDP_i \tag{5.6}$$

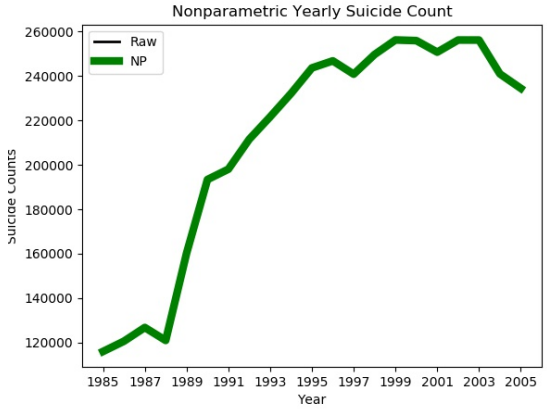
Figure 5.3 shows the global raw suicide counts per year. The suicide data we use can be found at Suicide Data Source. The data retains suicide counts from 1985 to 2016 for 101 countries of both males and females ranging from ages 15 to 75+ along with gross domestic product (GDP) information. We use suicide counts and GDP amounts from 1985-2005 for our in-control training data, and information from 2006-2015 as validation data, excluding 2016 because it was a zero column. We estimated a linear model using parametric, nonparametric, and semiparametric techniques for each year. Once we complete each regression method, we demonstrate the Bayesian mCUSUM and mEWMA charts' ability to detect out-of-control occurrences for each method by implementing various shift changes. Our goal is to aid in suicide prevention through detection and identification of irregularly high suicide counts. In the validation data, we notice suicide counts much lower than that in the training data, to capture the ability of our method to detect out-of-control occurrences, we apply a range of larger shift changes to emulate an epidemic-like situations.



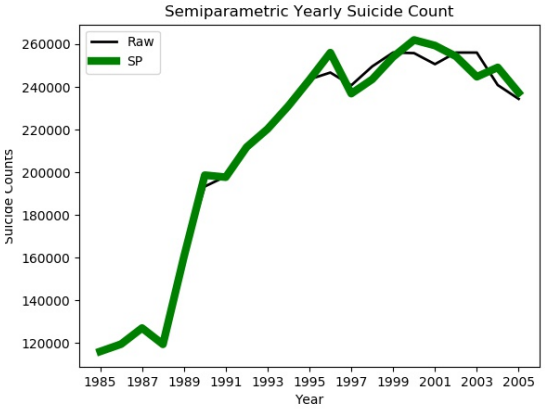
(a) Raw Training Data



(b) Parametric



(c) Nonparametric



(d) Semiparametric

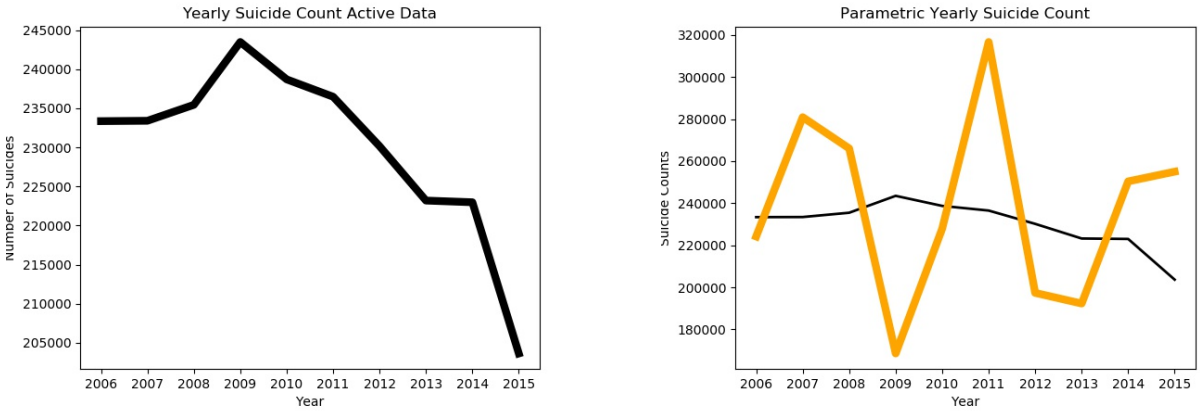
Figure 5.4: Training Data Regression Methods

Method	MSE	AIC	BIC
Parametric	2.38156e+09	-45.0233	456.899
P-spline	5.52196e-16	-32.6859	327.356
MRR1	2.38156e+07	-38.4174	387.537

Table 5.2: Training Data MSE Values

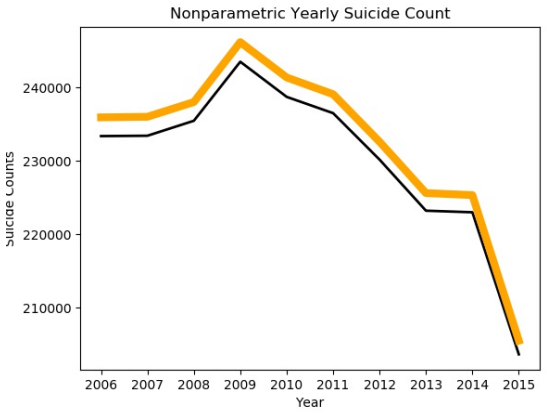
For the p-spline, since we are considering linear models, we identified two knots to be the

endpoints for each profile. Based on plots shown in figure 5.4, along with the MSE and BIC values from table 5.2, the parametric method performed the worst while the nonparametric p-spline method gave the lowest MSE/BIC and closest fit to the true data. We chose $\eta = 0.9$ as our mixing parameter for the MRR1 semiparametric technique because the nonparametric method had a better fit than the parametric. We use the the variance-covariance matrix and $\bar{\boldsymbol{\beta}}$, the average of our in-control $\hat{\boldsymbol{\beta}}$ values, to inform our control charts.

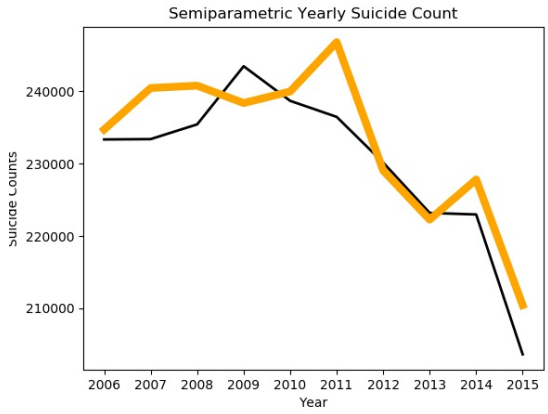


(a) Raw Training Data

(b) Parametric



(c) Nonparametric



(d) Semiparametric

Figure 5.5: Active Data Regression Methods

Method	MSE	AIC	BIC
Parametric	2.08504e+09	-43.5213	441.127
P-spline	6.10164e+06	-31.8533	318.614
MRR1	2.86851e+07	-37.2895	375.694

Table 5.3: Active Data MSE Values

Figure 5.5 and table 5.3 gave similar results to what we saw with the training data. That is, the p-spline regression method provided a better model than the MRR1 and parametric methods.

5.7. Discussion and Future Work

We explored the Bayesian mEWMA chart to monitor data with multiple variables. The squared error loss function was used to inform our Bayesian control charts, testing the tracking capabilities of these charts under a hyper-parameter sensitivity study. In the hyper-parameter analysis we noted that the choice of hyper-parameter had little effect on the ability of the Bayesian mEWMA chart, as well as the fit of the parametric model.

A natural next step in this work would be to apply this Bayesian mEWMA chart to models fit with the p-spline and MRR1 regression techniques and asses its tracking ability with a hyper-parameter sensitivity analysis. An analysis of how sample size affects the charts ability to detect out-of-control happenings would also beneficial for practitioners to determine if this method would be appropriate for their specific data. The classical mEWMA and mCUSUM charts are commonly compared memory-based charts, so a comparison of the Bayesian mEWMA and Bayesian mCUSUM charts proposed here would be an innate direction to take. Once these comparisons are executed via simulations, applying both Bayesian charts to the suicide count

data is advised to test their real-world capabilities.

Bibliography

Riaz, S., Riaz, M., Hussain, Z., & Abbas, T. (2017). Monitoring the performance of bayesian ewma control chart using loss functions. *Computers & Industrial Engineering, 112*, 426–436.

Noor-ul Amin, M., & Noor, S. (2021). An adaptive ewma control chart for monitoring the process mean in bayesian theory under different loss functions. *Quality and Reliability Engineering International, 37*(2), 804–819.

Jones, C. L., Abdel-Salam, A.-S. G., & Mays, D. (2021). *Generalized bayesian cusum charts using different loss functions* [Under Review].

Jones, C. L., Abdel-Salam, A.-S. G., & Mays, D. (2020). Practitioners guide on parametric, nonparametric, and semiparametric profile monitoring. *Quality and Reliability Engineering International*.

O'Sullivan, F. (1986). A statistical perspective on ill-posed inverse problems. *Statistical science*, 502–518.

Eilers, P. H., & Marx, B. D. (1996). Flexible smoothing with b-splines and penalties. *Statistical science*, 89–102.

Eilers, P. H., Marx, B. D., & Durbán, M. (2015). Twenty years of p-splines. *SORT: statistics and operations research transactions, 39*(2), 0149–186.

Ruppert, D., Wand, M. P., & Carroll, R. J. (2003). *Semiparametric regression*. Cambridge university press.

Keele, L., & Keele, L. (2008). *Semiparametric regression for the social sciences* (Vol. 230). Wiley Online Library.

Ruppert, D., Wand, M. P., & Carroll, R. J. (2009). Semiparametric regression during 2003–2007. *Electronic journal of statistics*, *3*, 1193.

Einsporn, R. L. (1987). *Hatlink: A link between least squares regression and nonparametric curve estimation* (Doctoral dissertation). Virginia Polytechnic Institute and State University.

Einsporn, R. L., & Birch, J. B. (1993). *Improving predictions by combining least squares and kernel regression prediction weights* [93-11:21]. Department of Statistics, Virginia Polytechnic Institute; State University.

Burman, P., & Chaudhuri, P. On a hybrid approach to parametric and nonparametric regression. In: *Nonparametric statistical methods and related topics: A festschrift in honor of professor pk bhattacharya on the occasion of his 80th birthday*. World Scientific, 2012, pp. 233–256.

Mays, J. E., Birch, J. B., & Einsporn, R. L. (2000). An overview of model-robust regression. *Journal of statistical computation and simulation*, *66*(1), 79–100.

Mays, J. E., Birch, J. B., & Alden Starnes, B. (2001). Model robust regression: Combining parametric, nonparametric, and semiparametric methods. *Journal of Nonparametric Statistics*, *13*(2), 245–277.

Page, E. S. (1954). Continuous inspection schemes. *Biometrika*, *41*(1/2), 100–115.

Page, E. (1961). Cumulative sum charts. *Technometrics*, *3*(1), 1–9.

Roberts, S. (1959). Control chart tests based on geometric moving averages. *Technometrics*, *1*(3), 239–250.

Woodall, W. H., & Ncube, M. M. (1985). Multivariate cusum quality-control procedures. *Technometrics*, *27*(3), 285–292.

Lowry, C. A., Woodall, W. H., Champ, C. W., & Rigdon, S. E. (1992). A multivariate exponentially weighted moving average control chart. *Technometrics*, *34*(1), 46–53.

WHO et al. (2014). *Preventing suicide: A global imperative*. World Health Organization.

Zortea, T. C., Brenna, C. T., Joyce, M., McClelland, H., Tippett, M., Tran, M. M., Arensman, E., Corcoran, P., Hatcher, S., Heise, M. J. et al. (2020). The impact of infectious disease-related public health emergencies on suicide, suicidal behavior, and suicidal thoughts. *Crisis*.

Leaune, E., Samuel, M., Oh, H., Poulet, E., & Brunelin, J. (2020). Suicidal behaviors and ideation during emerging viral disease outbreaks before the covid-19 pandemic: A systematic rapid review. *Preventive medicine*, 106264.

Gunnell, D., Appleby, L., Arensman, E., Hawton, K., John, A., Kapur, N., Khan, M., O'Connor, R. C., Pirkis, J., Caine, E. D. et al. (2020). Suicide risk and prevention during the covid-19 pandemic. *The Lancet Psychiatry*, *7*(6), 468–471.

Pirkis, J., John, A., Shin, S., DelPozo-Banos, M., Arya, V., Analuisa-Aguilar, P., Appleby, L., Arensman, E., Bantjes, J., Baran, A. et al. (2021). Suicide trends in the early months of the covid-19 pandemic: An interrupted time-series analysis of preliminary data from 21 countries. *The Lancet Psychiatry*.

Cénat, J. M., Blais-Rochette, C., Kokou-Kpolou, C. K., Noorishad, P.-G., Mukunzi, J. N., McIntee, S.-E., Dalexis, R. D., Goulet, M.-A., & Labelle, P. (2020). Prevalence of symptoms of depression, anxiety, insomnia, posttraumatic stress disorder, and psychological distress among populations affected by the covid-19 pandemic: A systematic review and meta-analysis. *Psychiatry research*, 113599.

Chapter 6

Discussion

6.1. Conclusion

In this dissertation we explored profile monitoring techniques using both a classical and Bayesian approach under various distributional assumptions. In chapter 2 we reviewed published works on statistical process control, supplying step-by-step applications of key monitoring techniques. This comprehensive review and guide provides practitioners a tool in learning SPC methods under parametric, nonparametric, and semiparametric regression. From this project, we conclude that the p-spline method resulted in better fits for linear data, while the Haar wavelet method fit the nonlinear data best. The memory-based charts, CUSUM and EWMA, had the highest PoS values under each data example, signifying their capability in detecting more out-of-control profiles. With these conclusions, we recommend the use of p-splines on estimating linear data and the CUSUM and EWMA charts for detecting small changes.

Chapters 3 and 4 considered the squared error (SELF), precautionary (PLF), and linex (LLF) loss functions in constructing Bayesian CUSUM and EWMA control charts, respectively. We de-

rived best estimators of the sample means for a normal conjugate and a Poisson conjugate using the various loss functions and used these, along with their variances, to inform our Bayesian charts. A demonstration of these charts was done via two simulation studies; the first was a sensitivity analysis of how the hyper-parameter choice influenced charting capabilities, and the second of the sample size's influence. The simulation results proved that our methods were adequate in detecting small increases in sample means. Particularly, the SELF CUSUM chart had the quickest detection. When all charts were administered on real hospitalization count data, both the SELF and PLF CUSUM charts had preferable outcomes over every other chart variation. We give recommendation for the squared error loss function for either chart, but emphasize it's use with the CUSUM.

In our final project, chapter 5, we compiled recommendations for methods utilized in chapters 2, 3, and 4 for monitoring linearly regressed data with Bayesian multivariate control charts. We estimated linear models to obtain model coefficients using parametric regression, penalized splines (p-splines), and model robust regression 1 (MRR1) methods. Since we exclusively considered count data in this chapter, we implemented the log link function to create a linear pattern from the data before estimating the model. Once the link function was applied and coefficients were estimated, we compared each regression method's estimation capability using the mean squared error (MSE), Akaike information criterion (AIC), and Bayesian information criterion (BIC). The p-spline regression method returned the smallest MSE and BIC values when estimating $\hat{y}$'s, while the parametric estimates resulted in greater MSE's and BIC's. Because we obtained poor estimates from the parametric method, we adjusted our MRR1 mixing parameter to rely more on the nonparametric $\hat{y}$'s in the data application, but the semiparametric values still gave a higher MSE and BIC values.

While in chapter 2 we focused on Phase I analysis, in chapters 3, 4, and 5 we emphasized our work under Phase II. We chose to direct our research to provide methods for the practitioner

of a process.

6.2. Future Work

Chapter 5 provides plausible next steps in its final section that include a comparison analysis between the parametric, p-spline, and MRR1 regression techniques. This analysis should be done while tuning the hyper-parameters and sample sizes to test the Bayesian mEWMA charts' tracking ability. To further allow for recommendation of the methods proposed, a study of the Bayesian mCUSUM chart should also be conducted and results compared to that from the Bayesian mEWMA. After both charts are tested using simulation studies, they should be applied to the suicide data that was introduced in chapter 5. This provides real-world application results to fully assess these charts.

Our work considered tracking parametric, nonparametric, and semiparametric regression method results in profile monitoring. Although our work has proven to provide a viable option in quick detection of count data, we acknowledge the application and methodological shortcomings. We primarily focused our research on linear models, but many applications do not follow a linear pattern. A future research path would be to test our methods on nonlinear data without the use of a link function. This would benefit practitioners by allowing direct implementation of a control chart and the ability to interpret results without reversing the link function. In addition to monitoring nonlinear data, we suggest the combination of our Bayesian methods with other control charts. That is, the use of loss functions to inform other control charts. We believe this would give practitioners the option to use a chart that would work best for their data, but also providing benefits from using a Bayesian technique.

Vita

Chelsea Jones, born in Norfolk, VA on December 27, 1993, graduated from Norview High School in 2012 with her advanced diploma. She received a Bachelor of Science degree in Actuarial Science with a minor in Computer Science from The Virginia State University in 2016. Chelsea began her journey in the Systems Modeling and Analysis PhD program at Virginia Commonwealth University in 2017, after being awarded the Southern Regional Education Board (SREB) Institutional Doctoral Scholar fellowship, being 1 of 9 scholars awarded campus-wide. Dr. Jones was selected for the Fulbright Scholar Award to research Monitoring of Epidemic Incidences in the Philippines for the 2021-2022 academic year. Chelsea has written a successful grant that will allow her to intern at Lawrence Berkeley National Laboratory (LBNL) researching Sampling Uncertainty Quantification of Network Traffic Predictions made by Deep Learning Models upon graduation.

Bibliography

Mahmoud, M. A., Parker, P. A., Woodall, W. H., & Hawkins, D. M. (2007). A change point method for linear profile data. *Quality and Reliability Engineering International*, *23*(2), 247–268.

Croarkin, C., Tobias, P., Filliben, J., Hembree, B., Guthrie, W. et al. (2006). Nist/sematech e-handbook of statistical methods. *NIST/SEMATECH, July. Available online: http://www. itl. nist. gov/div898/handbook*.

Abdel-Salam, A.-S. G. (2009). *Profile monitoring with fixed and random effects using nonparametric and semiparametric methods* (Doctoral dissertation). Virginia Tech.

Abbas, T., Abbasi, S. A., Riaz, M., & Qian, Z. (2019a). Phase ii monitoring of linear profiles with random explanatory variable under bayesian framework. *Computers & Industrial Engineering*, *127*, 1115–1129.

Adibi, A., Montgomery, D. C., & Borror, C. M. (2014). Phase ii monitoring of linear profiles using a p-value approach. *Int. J. Quality Engineering and Technology*, *4*(2), 97–106.

Amiri, A., Eyvazian, M., Zou, C., & Noorossana, R. (2012). A parameters reduction method for monitoring multiple linear regression profiles. *International Journal of Advanced Manufacturing Technology*.

Amiria, A., Kooshab, M., Azhdaric, A., & Wang, G. (2014). Phase i monitoring of generalized linear model-based regression profiles. *Journal of Statistical Computation and Simulation*.

Bierens, H. J. (1988). The nadaraya-watson kernel regression function estimator.

Breiman, L. (1993). Fitting additive models to regression data. *Computational statistics and data analysis*, *15*, 13–46.

Burden, R. L., & Faires, J. D. (2010). *Numerical analysis (9th)*. Brooks/Cole.

Burman, P., & Chaudhuri, P. On a hybrid approach to parametric and nonparametric regression. In: *Nonparametric statistical methods and related topics: A festschrift in honor of professor pk bhattacharya on the occasion of his 80th birthday*. World Scientific, 2012, pp. 233–256.

Burrus, C. S., Gopinath, R, & Guo, H. (1998). *A primer introduction to wavelets and wavelet transforms* (Vol. 3).

Butler, R., Davies, P., Jhun, M et al. (1993). Asymptotics for the minimum covariance determinant estimator. *The Annals of Statistics*, *21*(3), 1385–1400.

Chang, S. I., & Yadama, S. (2010). Statistical process control for monitoring non-linear profiles using wavelet filtering and b-spline approximation. *International Journal of Production Research*, *48*(4), 1049–1068.

Chicken, E., Pignatiello Jr, J. J., & Simpson, J. R. (2009). Statistical process monitoring of nonlinear profiles using wavelets. *Journal of Quality Technology*, *41*(2), 198–212.

Cuevas, A., Febrero, M., & Fraiman, R. (2006). On the use of the bootstrap for estimating functions with functional data. *Computational statistics & data analysis*, *51*(2), 1063–1074.

Cuevas, A., Febrero, M., & Fraiman, R. (2007). Robust estimation and classification for functional data via projection-based depth notions. *Computational Statistics*, *22*(3), 481–496.

Daubechies, I. (1992). *Ten lectures on wavelets* (Vol. 61). Siam.

Davies, L. et al. (1992). The asymptotics of rousseeuw's minimum volume ellipsoid estimator. *The Annals of Statistics*, *20*(4), 1828–1843.

De Boor, C., De Boor, C., Mathématicien, E.-U., De Boor, C., & De Boor, C. (1978). *A practical guide to splines* (Vol. 27). Springer-Verlag New York.

Denison, D., Mallick, B., & Smith, A. (1998). Automatic bayesian curve fitting. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *60*(2), 333–350.

Eilers, P. H., & Marx, B. D. (1996). Flexible smoothing with b-splines and penalties. *Statistical science*, 89–102.

Eilers, P. H., Marx, B. D., & Durbán, M. (2015). Twenty years of p-splines. *SORT: statistics and operations research transactions*, *39*(2), 0149–186.

Einsporn, R. L. (1987). *Hatlink: A link between least squares regression and nonparametric curve estimation* (Doctoral dissertation). Virginia Polytechnic Institute and State University.

Einsporn, R. L., & Birch, J. B. (1993). *Improving predictions by combining least squares and kernel regression prediction weights* [93-11:21]. Department of Statistics, Virginia Polytechnic Institute; State University.

Febrero, M., Galeano, P., & González-Manteiga, W. (2007). A functional analysis of nox levels: Location and scale estimation and outlier detection. *Computational Statistics*, *22*(3), 411–427.

Febrero, M., Galeano, P., & González-Manteiga, W. (2008). Outlier detection in functional data by depth measures, with application to identify abnormal nox levels. *Environmetrics: The official journal of the International Environmetrics Society*, *19*(4), 331–345.

Ferraty, F., & Vieu, P. (2006). *Nonparametric functional data analysis: Theory and practice*. Springer Science & Business Media.

Fitzenberger, B., Koenker, R., & Machado, J. A. (2013). *Economic applications of quantile regression*. Springer Science & Business Media.

Fraiman, R., Meloche, J., García-Escudero, L. A., Gordaliza, A., He, X., Maronna, R., Yohai, V. J., Sheather, S. J., McKean, J. W., Small, C. G. et al. (1999). Multivariate l-estimation. *Test*, *8*(2), 255–317.

Fraiman, R., & Muniz, G. (2001). Trimmed means for functional data. *Test*, *10*(2), 419–440.

Friedman, J. H., & Silverman, B. W. (1989). Flexible parsimonious smoothing and additive modeling. *Technometrics*, *31*(1), 3–21.

Geisser, S. (2017). *Predictive inference*. Routledge.

Gomaa, A.-S., & Birch, J. B. (2019). A semiparametric nonlinear mixed model approach to phase i profile monitoring. *Communications in Statistics-Simulation and Computation*, *48*(6), 1677–1693.

Gupta, S, Montgomery, D., & Woodall, W. (2006). Performance evaluation of two methods for online monitoring of linear calibration profiles. *International journal of production research*, *44*(10), 1927–1942.

Hardin, J., & Rocke, D. M. (2004). Outlier detection in the multiple cluster setting using the minimum covariance determinant estimator. *Computational Statistics & Data Analysis*, *44*(4), 625–638.

Jeong, M. K., Lu, J.-C., & Wang, N. (2006). Wavelet-based spc procedure for complicated functional data. *International Journal of Production Research*, *44*(4), 729–744.

Kauermann, G., Claeskens, G., & Opsomer, J. D. (2009). Bootstrapping for penalized spline regression. *Journal of Computational and Graphical Statistics*, *18*(1), 126–146.

Ke, C., & Wang, Y. (2001). Semiparametric nonlinear mixed-effects models and their applications. *Journal of the American Statistical Association*, *96*(456), 1272–1298.

Kim, K., Mahmoud, M. A., & Woodall, W. H. (2003). On the monitoring of linear profiles. *Journal of Quality Technology*, *35*(3), 317–328.

Koenker, R., & Bassett Jr, G. (1978). Regression quantiles. *Econometrica: journal of the Econometric Society*, 33–50.

Koosha, M., & Amiri, A. (2013). Generalized linear mixed model for monitoring autocorrelated logistic regression profiles. *International Journal of Advanced Manufacturing Technology*.

Li, Z., & Wang, Z. (2010). An exponentially weighted moving average scheme with variable sampling intervals for monitoring linear profiles. *Computers & Industrial Engineering*, *59*(4), 630–637.

Liu, R. Y. (1990). On a notion of data depth based on random simplices. *The Annals of Statistics*, 405–414.

Lolive, D., Barbot, N., & Boeffard, O. Melodic contour estimation with b-spline models using a mdl criterion. In: *Proceedings of the 11th international conference on speech and computer (specom)*. 2006, 333–338.

López-Pintado, S., & Romo, J. (2009). On the concept of depth for functional data. *Journal of the American Statistical Association*, *104*(486), 718–734.

Lowry, C. A., Woodall, W. H., Champ, C. W., & Rigdon, S. E. (1992). A multivariate exponentially weighted moving average control chart. *Technometrics*, *34*(1), 46–53.

Mahalanobis, P. C. On the generalized distance in statistics. In: National Institute of Science of India. 1936.

Maleki, M. R., Amiri, A., & Castagliola, P. (2018). An overview on recent profile monitoring papers (2008–2018) based on conceptual classification scheme. *Computers & Industrial Engineering*, *126*, 705–728.

Mays, J. E., Birch, J. B., & Einsporn, R. L. (2000). An overview of model-robust regression. *Journal of statistical computation and simulation*, *66*(1), 79–100.

Mays, J. E., Birch, J. B., & Alden Starnes, B. (2001). Model robust regression: Combining parametric, nonparametric, and semiparametric methods. *Journal of Nonparametric Statistics*, *13*(2), 245–277.

Meyer, K. (2005). Random regression analyses using b-splines to model growth of australian angus cattle. *Genetics Selection Evolution*, *37*(6), 473.

Molinari, N., Durand, J.-F., & Sabatier, R. (2004). Bounded optimal knots for regression splines. *Computational statistics & data analysis*, *45*(2), 159–178.

Montgomery, D. C. (1990). *Introduction to statistical quality control* (Second). John Wiley & Sons.

Montgomery, D. (2013). *Introduction to statistical quality control* (Seventh). John Wiley & Sons.

Mosteller, F., & Wallace, D. L. (1963). Inference in an authorship problem: A comparative study of discrimination methods applied to the authorship of the disputed federalist papers. *Journal of the American Statistical Association*, *58*(302), 275–309.

Nadaraya, E. A. (1964). On estimating regression. *Theory of Probability & Its Applications*, *9*(1), 141–142.

Nikoo, M., & Noorossana, R. (2013). Phase ii monitoring of nonlinear profile variance using wavelet. *Quality and Reliability Engineering International*, *29*(7), 1081–1089.

Nottingham, Q. J., & Birch, J. B. (2000). A semiparametric approach to analysing dose–response data. *Statistics in medicine*, *19*(3), 389–404.

Ogden, T. (2012). *Essential wavelets for statistical applications and data analysis*. Springer Science & Business Media.

Oja, H. (1983). Descriptive statistics for multivariate distributions. *Statistics & Probability Letters*, *1*(6), 327–332.

Olkin, I., & Spiegelman, C. H. (1987). A semiparametric approach to density estimation. *Journal of the American Statistical Association*, *82*(399), 858–865.

O'Sullivan, F. (1986). A statistical perspective on ill-posed inverse problems. *Statistical science*, 502–518.

Page, E. S. (1954). Continuous inspection schemes. *Biometrika*, *41*(1/2), 100–115.

Page, E. (1961). Cumulative sum charts. *Technometrics*, *3*(1), 1–9.

Parker, P, Morton, M, Draper, N, & Line, W. A single-vector force calibration method featuring the modern design of experiments. In: *39th aerospace sciences meeting and exhibit*. 2001, 170.

Pignatiello Jr, J. J., & Runger, G. C. (1990). Comparisons of multivariate cusum charts. *Journal of quality technology*, *22*(3), 173–186.

Piri, S. (2017). Parametric, nonparametric and semiparametric approaches in profile monitoring of poisson data.

Piri, S., Abdel-Salam, A.-S. G., & Boone, E. L. (2019). A wavelet approach for profile monitoring of poisson distribution with application. *Communications in Statisitics-Simiulation and Computation*, 1–12.

Ramsay, J. O., & Silverman, B. W. (2007). *Applied functional data analysis: Methods and case studies*. Springer.

Reis, M. S., & Saraiva, P. M. (2006). Multiscale statistical process control of paper surface profiles. *Quality Technology & Quantitative Management*, *3*(3), 263–281.

Riaz, M., Mahmood, T., Abbasi, S. A., Abbas, N., & Ahmad, S. (2017). Linear profile monitoring using ewma structure under ranked set schemes. *International Journal of Advanced Manufacturing Technology*.

Roberts, S. (1959). Control chart tests based on geometric moving averages. *Technometrics*, *1*(3), 239–250.

Rousseeuw, P. J. (1984). Least median of squares regression. *Journal of the American statistical association*, *79*(388), 871–880.

Rousseeuw, P. J., & Van Zomeren, B. C. (1990). Unmasking multivariate outliers and leverage points. *Journal of the American Statistical association*, *85*(411), 633–639.

Rousseeuw, P. J., & Driessen, K. V. (1999). A fast algorithm for the minimum covariance determinant estimator. *Technometrics*, *41*(3), 212–223.

Ruppert, D. (2002). Selecting the number of knots for penalized splines. *Journal of computational and graphical statistics*, *11*(4), 735–757.

Ruppert, D., Wand, M. P., & Carroll, R. J. (2009). Semiparametric regression during 2003–2007. *Electronic journal of statistics*, *3*, 1193.

Saghaei, A., Mehrjoo, M., & Amiri, A. (2009). A cusum-based method for monitoring simple linear profiles. *The International Journal of Advanced Manufacturing Technology, 45*(11-12), 1252.

Shewhart, W. A. (1926). Quality control charts. *The Bell System Technical Journal, 5*(4), 593–603.

Siddiqui, Z., & Abdel-Salam, A.-S. G. (2019). A semiparametric profile monitoring via residuals. *Quality and Reliability Engineering International.*

Singh, K. (1991). A notion of majority depth. *Unpublished document.*

Soleimani, P., Noorossana, R., & Amiri, A. (2009). Simple linear profiles monitoring in the presence of within profile autocorrelation. *Computers & Industrial Engineering, 57*(3), 1015–1021.

Soleymanian, M., Khedmati, M, & Mahlooji, H. (2013). Phase ii monitoring of binary response profiles. *Scientia Iranica. Transaction E, Industrial Engineering, 20*(6), 2238.

Stone, M. (1974). Cross-validatory choice and assessment of statistical predictions. *Journal of the royal statistical society. Series B (Methodological)*, 111–147.

Sun, Y., & Genton, M. G. (2011). Functional boxplots. *Journal of Computational and Graphical Statistics, 20*(2), 316–334.

R. (2010). R development core team 2010: R: A language and environment for statistical computing, version r 3.3. 2. *Available at www. r-project. org. Accessed September, 9*, 2017.

Tjahjowidodo, T. et al. (2017). A direct method to solve optimal knots of b-spline curves: An application for non-uniform b-spline curves fitting. *PloS one, 12*(3), e0173857.

Tukey, J. W. (1970). Exploratory data analysis.(limited preliminary edition), vol. 1, ch. 5.

Tukey, J. W. Mathematics and the picturing of data. In: *Proceedings of the international congress of mathematicians, vancouver, 1975. 2*. 1975, 523–531.

Vaghefi, A, Tajbakhsh, S. D., & Noorossana, R. (2009). Phase ii monitoring of nonlinear profiles. *Communications in Statistics—Theory and Methods, 38*(11), 1834–1851.

Van Aelst, S., & Rousseeuw, P. (2009). Minimum volume ellipsoid. *Wiley Interdisciplinary Reviews: Computational Statistics*, *1*(1), 71–82.

Vardi, Y., & Zhang, C.-H. (2000). The multivariate l1-median and associated data depth. *Proceedings of the National Academy of Sciences*, *97*(4), 1423–1426.

Vargas, N. J. A. (2003). Robust estimation in multivariate control charts for individual observations. *Journal of Quality Technology*, *35*(4), 367–376.

Waterman, M. J., Birch, J. B., & Abdel-Salam, A.-S. G. (2015). Several nonparametric and semiparametric approaches to linear mixed model regression. *Journal of Statistical Computation and Simulation*, *85*(5), 956–977.

Watson, G. S. (1964). Smooth regression analysis. *Sankhyā: The Indian Journal of Statistics, Series A*, 359–372.

Wells, L. J., Megahed, F. M., Niziolek, C. B., Camelio, J. A., & Woodall, W. H. (2013). Statistical process monitoring approach for high-density point clouds. *Journal of Intelligent Manufacturing*, *24*(6), 1267–1279.

Woodall, W. H., Spitzner, D. J., Montegomery, D. C., & Gupta, S. (2004). Using control charts to monitor process and product quality profiles. *Journal of Quality Technology*, *36*(3), 309.

Woodall, W. H. (2007). Current research on profile monitoring. *Producao*, *17*(3), 420–425.

Woodall, W. H., & Montgomery, D. C. (2014). Some current directions in the theory and application of statistical process monitoring. *Journal of Quality Technology*, *46*(1), 78–94.

Woodall, W. H., & Faltin, F. W. (2019). Rethinking control chart design and evaluation. *Quality Engineering*, *31*(4), 596–605.

Woodruff, D. L., & Rocke, D. M. (1993). Heuristic search algorithms for the minimum volume ellipsoid. *Journal of Computational and Graphical Statistics*, *2*(1), 69–95.

Yeh, A. B., Huwang, L., & Li, Y.-M. (2009). Profile monitoring for a binary response. *IIE Transactions*, *41*(11), 931–941.

Zand, A., Yazdanshenas, N., & Amiri, A. (2013). Change point estimation in phase i monitoring of logistic regression profile. *The International Journal of Advanced Manufacturing Technology*, *67*(9-12), 2301–2311.

Zou, C., Zhou, C., Wang, Z., & Tsung, F. (2007). A self-starting control chart for linear profiles. *Journal of Quality Technology*, *39*(4), 364–375.

Zou, C., Tsung, F., & Wang, Z. (2008). Monitoring profiles based on nonparametric regression methods. *Technometrics*, *50*(4), 512–526.

Zou, C., Qiu, P., & Hawkins, D. (2009). Nonparametric control chart for monitoring profiles using change point formulation and adaptive smoothing. *Statistica Sinica*, 1337–1357.

Norstrom, J. G. (1996). The use of precautionary loss functions in risk analysis. *IEEE Transactions on reliability*, *45*(3), 400–403.

Prabhu, S. S., & Runger, G. C. (1997). Designing a multivariate ewma control chart. *Journal of Quality Technology*, *29*(1), 8–15.

Saghir, A., Lin, Z., & Chen, C.-W. (2015). The properties of the geometric-poisson exponentially weighted moving control chart with estimated parameters. *Cogent Mathematics*, *2*(1), 992381.

Varian, H. R. (1975). A bayesian approach to real estate assessment. *Studies in Bayesian econometric and statistics in Honor of Leonard J. Savage*, 195–208.

Zellner, A. (1986). Bayesian estimation and prediction using asymmetric loss functions. *Journal of the American Statistical Association*, *81*(394), 446–451.

Hotelling, H. et al. A generalized t test and measure of multivariate dispersion. In: *Proceedings of the second berkeley symposium on mathematical statistics and probability*. The Regents of the University of California. 1951.

Jones, C. L., Abdel-Salam, A.-S. G., & Mays, D. (2020). Practitioners guide on parametric, nonparametric, and semiparametric profile monitoring. *Quality and Reliability Engineering International*.

Keele, L., & Keele, L. (2008). *Semiparametric regression for the social sciences* (Vol. 230). Wiley Online Library.

Ruppert, D., Wand, M. P., & Carroll, R. J. (2003). *Semiparametric regression*. Cambridge university press.

Cénat, J. M., Blais-Rochette, C., Kokou-Kpolou, C. K., Noorishad, P.-G., Mukunzi, J. N., McIntee, S.-E., Dalexis, R. D., Goulet, M.-A., & Labelle, P. (2020). Prevalence of symptoms of depression, anxiety, insomnia, posttraumatic stress disorder, and psychological distress among populations affected by the covid-19 pandemic: A systematic review and meta-analysis. *Psychiatry research*, 113599.

Gunnell, D., Appleby, L., Arensman, E., Hawton, K., John, A., Kapur, N., Khan, M., O'Connor, R. C., Pirkis, J., Caine, E. D. et al. (2020). Suicide risk and prevention during the covid-19 pandemic. *The Lancet Psychiatry*, *7*(6), 468–471.

Jones, C. L., Abdel-Salam, A.-S. G., & Mays, D. (2021). *Generalized bayesian cusum charts using different loss functions* [Under Review].

Leaune, E., Samuel, M., Oh, H., Poulet, E., & Brunelin, J. (2020). Suicidal behaviors and ideation during emerging viral disease outbreaks before the covid-19 pandemic: A systematic rapid review. *Preventive medicine*, 106264.

Noor-ul Amin, M., & Noor, S. (2021). An adaptive ewma control chart for monitoring the process mean in bayesian theory under different loss functions. *Quality and Reliability Engineering International*, *37*(2), 804–819.

Pirkis, J., John, A., Shin, S., DelPozo-Banos, M., Arya, V., Analuisa-Aguilar, P., Appleby, L., Arensman, E., Bantjes, J., Baran, A. et al. (2021). Suicide trends in the early months of the covid-19 pandemic: An interrupted time-series analysis of preliminary data from 21 countries. *The Lancet Psychiatry*.

WHO et al. (2014). *Preventing suicide: A global imperative*. World Health Organization.

Woodall, W. H., & Ncube, M. M. (1985). Multivariate cusum quality-control procedures. *Technometrics*, *27*(3), 285–292.

Zortea, T. C., Brenna, C. T., Joyce, M., McClelland, H., Tippett, M., Tran, M. M., Arensman, E., Corcoran, P., Hatcher, S., Heise, M. J. et al. (2020). The impact of infectious disease-related public health emergencies on suicide, suicidal behavior, and suicidal thoughts. *Crisis*.

Abbas, T., Ahmad, S., Riaz, M., & Qian, Z. (2019b). A bayesian way of monitoring the linear profiles using cusum control charts. *Communications in Statistics-Simulation and Computation*, *48*(1), 126–149.

Alencar, A. P., Lee Ho, L., & Albarracin, O. Y. E. (2017). Cusum control charts to monitor series of negative binomial count data. *Statistical methods in medical research*, *26*(4), 1925–1935.

Ali, S. (2020). A predictive bayesian approach to ewma and cusum charts for time-between-events monitoring. *Journal of Statistical Computation and Simulation*, *90*(16), 3025–3050.

Aslam, M., & Anwar, S. M. (2020). An improved bayesian modified-ewma location chart and its applications in mechanical and sport industry. *PloS one*, *15*(2), e0229422.

Das, N. (2003). Study on implementing control chart assuming negative binomial distribution with varying sample size in a software industry. *ASQ Software Quality Professional*, *6*, 38–39.

Hoffman, D. (2003). Negative binomial control limits for count data with extra-poisson variation. *Pharmaceutical Statistics: The Journal of Applied Statistics in the Pharmaceutical Industry*, *2*(2), 127–132.

Rossi, G, Lampugnani, L, & Marchi, M. (1999). An approximate cusum procedure for surveillance of health events. *Statistics in medicine*, *18*(16), 2111–2122.

Schwertman, N. C. (2005). Designing accurate control charts based on the geometric and negative binomial distributions. *Quality and Reliability Engineering International*, *21*(8), 743–756.

Tsiamyrtzis, P., & Hawkins, D. M. (2008). A bayesian ewma method to detect jumps at the start-up phase of a process. *Quality and Reliability Engineering International*, *24*(6), 721–735.

Urbieta, P., Lee HO, L., & Alencar, A. (2017). Cusum and ewma control charts for negative binomial distribution. *Quality and Reliability Engineering International*, *33*(4), 793–801.

Wise, M. (1946). The use of the negative binomial distribution in an industrial sampling problem. *Supplement to the Journal of the Royal Statistical Society*, *8*(2), 202–211.

Yu, F.-J., Yang, Y.-Y., Wang, M.-J., & Wu, Z. (2011). Using ewma control schemes for monitoring wafer quality in negative binomial process. *Microelectronics Reliability*, *51*(2), 400–405.

Yun, M., & Youlin, Z. (1996). Q control charts for negative binomial distribution. *Computers & industrial engineering*, *31*(3-4), 813–816.