Identifying Changes in Demand for Perishable Products Using Statistical Process Control and Forecasting

by

Linden Smith

Submitted in partial fulfilment of the requirements
for the degree of Master of Applied Science

at

Dalhousie University
Halifax, Nova Scotia
December 2021

Table of Contents

List of Tables

List of Figures

Abstract

Fast and accurate identification of changes in demand is crucial in the management of blood products. Canadian Blood Services (CBS) manages the collection and distribution of blood products in all parts of Canada excluding Quebec. CBS is planning a pilot project to apply pathogen reduction technology (PRT) to platelet production. The pilot will take place in the Ottawa region of Ontario prior to the eventual rollout of the technology to rest of the country. The introduction of PRT is expected to change hospital demand for platelets; however, the form of this change is unknown. Furthermore, a lag time exists between the identification of a supply-demand imbalance and the ability to address it, making it even more critical to have methods for accurately predicting demand.

The objective of this research is to determine how quickly and accurately demand changes can be detected after the introduction of PRT platelets in Ottawa, to minimize lag time and thus provide better patient health outcomes. A discrete-event simulation was used to model platelet inventory and generate data representative of possible demand shift scenarios. Process control methods were used to detect and quantify shifts in demand. To improve time to detection, forecast data points were included in the changepoint detection scheme. Several forecasting methods were tested in this role. But ultimately a Generalized Additive Model, fit using splines, was selected. It was found that statistical process control methods were effective in detecting demand shifts of any form and that forecasting decreased the time to detection, with only slight increases in the false alarm rate. Experiments showed that when the magnitude of the demand shift increased, the detection rate increased and the time to detection decreased. However, the consequences of a hidden demand shift are substantially less for shifts of smaller magnitude, mitigating the risk due to increased detection time. The results of this thesis will be useful in minimizing the patient impact of PRT.

List of Abbreviations

1. CBS – Canadian Blood Services
2. PRT – Pathogen Reduction Technology
3. ARIMA – Autoregressive Integrated Moving Average
4. ANN – Artificial Neural Network
5. LSTM – Long Short-Term Memory
6. SPC – Statistical Process Control
7. NEON – Northeastern Ontario
8. ZIP – Zero Inflated Poisson
9. OCPD – Online Changepoint Detection
10. GAM – Generalized Additive Model
11. REML – Restricted Maximum Likelihood
12. MAPE – Mean Absolute Percentage Error
13. RMSE – Root Mean Square Error

Glossary

1. **ABO Group:** Blood type determined by the presence or lack of A and B genes. The possible blood types are A, B, AB, and O.
2. **Rh Antigen:** Compound, potentially present on the surface of blood cells. Blood is classified as 'Rh positive' if it is present, or 'Rh negative' if it is not.
3. **Pooled platelets:** A blood product derived from the whole blood donations of 4 donors with the same phenotype.
4. **Apheresis:** Procedure for extracting a specific component from a donor's blood and returning the remaining components. Apheresis platelets come from a single donor.
5. **BAC-T:** Automated microbial detection system. Currently used by Canadian Blood Services to test for the presence of pathogens.
6. **PRT:** Pathogen Reduction Technology
7. **Online Changepoint Detection:** The accurate detection of sudden changes in a time series as new data is acquired.
8. **Statistical Process Control Charts**: An application of OCPD used for monitoring in quality control. Often referred to by its primary tool the Shewhart Control Chart.
9. **ARIMA/ARMA:** Autoregressive Integrated Moving Average/Autoregressive Moving Average. Forecasting method based on the idea that current value is a function of some number of recent observed values. The number and importance of each point in the forecast is model specific, but generally decreases further into the past.
10. **Local Regression:** A term used to describe regression on sequential groups of points in a larger data set. Describes linear and non-linear methods, but generally refers to smooth non-linear curves.
11. **GAM:** Generalized Additive Model, Regression technique based on the concept of adding models together to predict outcomes.
12. **Prophet:** Prophet is a time-series model which includes long-term data patterns, seasonality, and the non-periodic effect of holidays. These features are combined using a Generalized Additive Model (GAM).
13. **Artificial Neural Network:** Collection of statistical models connected in a network. Individual models, or nodes, are re-weighted in the overall model based on iterative feedback.
14. **Lasso Regression:** Lasso regression is a multiple linear regression technique, which selects a subset of variables to base the model on

Acknowledgements

I would first like to thank my supervisor Dr. John Blake for his expertise, patience, and guidance during a challenging year, as well as for giving me the opportunity to pursue research.

I would like to acknowledge Canadian Blood Services, for their financial support of this work, as well as all the CBS employees who gave me their time and insights over the last two years. Special thanks, Ken McTaggart and Chantal Couture who provided invaluable feedback and assistance.

Finally, I would like to thank my friends and family for their encouragement and understanding during this process, including Daniel Campbell who convinced me to study Industrial Engineering and only teased me a little when I did.

Chapter 1 Introduction

In all Canadian provinces, except Quebec, Canadian Blood Services (CBS) is responsible for the collection, processing, storage, and transport of blood products [1]. CBS uses a network of regional facilities in its supply chain; blood product inventory is held in a central location for each region and exclusively distributed to hospitals within the area. Each day, orders come in from hospitals for a quantity of blood product units, typically specifying the blood type or other parameters, they wish to receive. Not all hospitals order daily, with some smaller hospitals only ordering every few days. In addition, pauses to ordering on the weekend are commonplace. In emergency situations additional units may be dispatched, but most products are shipped on a regular daily cycle. Hospitals maintain their own local stores of blood products for distribution to patients. In general, inventory and usage data in hospitals is not visible to CBS or other interested parties. Thus, once a unit has been shipped from CBS its final use and ultimate disposition is not generally known or only known much after the fact.

Before being released to hospitals, blood products are tested for the presence of transmissible diseases and/or bacterial contamination. The probability of bacterial contamination in platelets is estimated to be less than 1 in 1,000 [2]. While transfusion related infections are rare, they still present a risk, and any reduction in the transmission of pathogens improves both health outcomes and patient confidence in the transfusion process.

In this report we consider two types of blood products: pooled platelets and apheresis platelets. A unit of pooled platelets is, in Canada, a combination of buffy coat platelets derived from four different donors, all of whom have the same ABO blood type. Apheresis platelets are collected from a single donor, typically in one session, using an apheresis collection/production process. Apheresis is a procedure where whole blood is collected from a donor and separated into its components. A single component, in this case platelets, can be removed while the remaining components are returned to the

donor's blood stream. Both types of blood products currently have a maximum shelf-life of 7 days, after which they are outdated and must be disposed.

Pooled platelets are typically allocated for general-use transfusion, such as in trauma cases. Apheresis platelets are generally reserved for patients with weakened immune systems, or patients who are otherwise more likely to have an adverse reaction. Pooled platelets have an estimated cost of approximately $70/unit while, due to more difficult collection and processing, apheresis platelets have an internally estimated cost of $400/unit. Platelet donors and recipients do not need to have the same blood type, within the guidelines for blood-plasma ABO substitution, for a transfusion to be considered safe and effective. This allows for flexibility in filling orders and contributes to lowering the quantity of outdates. Units approaching expiration can thus be used to fill demand for a different, compatible type.

Platelet ABO compatibility is different from that of red blood cells and uses the same rules as blood plasma, as seen below in Table 1 [3]. Rh compatibility, positive or negative, is more complicated. When pooled platelets are used, Rh-positive patients can receive either positive or negative components, while Rh-negative patients should preferably receive negative components. The Rh factor is less important for apheresis transfusion although donor and patient having the same type is preferred. In exceptional circumstances when inventory is very low, Rh-negative patients may be given Rh-positive components.

Table 1 Comparison of blood group compatibility for red blood cells and platelets. Platelets use the same compatibility rules as blood plasma

Patient ABO Group	Compatible Red Blood Cells	Compatible Platelets
O	O	O, A, B, AB
A	A, O	A, AB
B	B, O	B, AB
AB	O, A, B, AB	AB

In addition to blood group substitution, apheresis units can be substituted for pooled platelets; however pooled platelets are not generally substituted for apheresis platelets. Given the large difference in cost between the two types of products, the expiration of a unit of apheresis platelets represents a significant loss, compared to pooled platelets. To reduce the cost of waste, apheresis units that are about to expire may be preferentially substituted for orders of pooled platelets.

As noted above, CBS retains shipping information for the blood type, product type, and quantity of units sent to hospitals, but does not retain actual demand information and only limited order information. As substitution of both blood and product type is common, the distinction between quantity shipped and quantity demanded is unknown. While substitution is a valuable tool to reduce costs and waste, not retaining the original demand makes it difficult to estimate true demand and thus has the potential to inflate the perceived usage of apheresis units. In the presence of product substitution, inventory data can be used as a stand-in for hospital usage information.

The short shelf life of blood products means the level of inventory held by CBS must closely match demand. This makes platelet inventory sensitive to small changes in both demand and supply. A small, sustained, positive difference in inventory quickly leads to high wastage, while a small, sustained, negative difference quickly reduces the service level to patients.

At present, platelet units are tested for bacterial contamination, using the BAC-T Alert system, a non-destructive testing system [4]. Using BAC-T, the risk of transfusing a bacterial contaminated unit is estimated at less than 1 in 1,000. However, CBS has committed to introducing Pathogen Reduction Technology (PRT) to platelet processing with the intent of reducing the chance of transfusion related infections. The proposed system functions by combining blood product with amotosalen and exposing the mixture to ultraviolet light. The process causes mis-links in the DNA of pathogens in the blood product, preventing the organism from reproducing, thus effectively sterilizing the product [5].

While there are significant benefits to PRT treatment, it presents short term problems for Canada's platelet supply chain. It has been hypothesized that PRT treated units may have a lower platelet count than untreated units [4], potentially requiring an increase in the number of units used in transfusion. Physicians could be reticent to use pooled platelets as a new product, thus shifting demand *away* from pooled and into apheresis platelets. While the exact form of the response is unknown, the introduction of Pathogen Reduction Technology is expected to influence demand for platelet products from hospitals. Fewer platelets per unit might lead to an increase in demand for pooled platelets, a shift in demand from pooled platelets to apheresis, a combination of both or no change at all. However, the sensitivity of inventory to demand changes increases the likelihood of a service gap if there is an undetected shift in demand for one product or another.

Chapter 1.1 Objectives

It is known that platelet demand will change because of the introduction of PRT, what is not known precisely is the amount and form of the change. This leads to the question: How well can unexpected changes in demand be detected? The objective of this thesis is to provide an answer to this question. To this end, this report describes the development and evaluation of tools to detect and estimate platelet demand shifts, in the Northeastern Ontario (NEON) region, following the implementation of PR technology. As information is only available on shipments, and not hospital demand, inventory will be used as a proxy for demand. While this study focusses on the impact of PRT, the results are broadly applicable and can be adapted to demand shifts in all types of products.

Detection techniques are evaluated, in this study, on two metrics: Probability of Correct Identification, and Time to Detection, shown in Table 2.

Table 2 Project objectives

Objective Name	Objective Type
Probability of Correct Identification	Maximization
Detection Time	Minimization

Correctly identifying signals is key to detection however, variance in demand and collections for 16 different identifiable platelet sub-types, with partial substitution between product types and blood groups, means that the correct identification of a demand shift signal is not guaranteed with any method. A detection signal can be received when there has been no shift in demand: A false positive. Conversely no detection signal may be received when there has been a shift in demand: A false negative.

The time needed to detect a shift in demand contributes to the response time. The longer a change in demand goes unaddressed, the more likely a lack of supply is to either cause shortages, for patients, or to increase outdates. Time to detection is then an important aspect of performance for this work. The timescale of a response to an unexpected change in demand is not hours or days, but in weeks or months, depending on the severity of the change. Fortunately for patients, inventory acts as buffer for changes in demand. Thus, if a change is small, patients may not be affected by it before there is a response in place. As for large changes, a quick detection and response is critical, but inventory is expected to provide a safety buffer until a resolution can be found.

The two objectives above are in competition, as the sensitivity of a monitoring technique is inversely related to detection time and positively correlated to the number of false positive signals [6]. A detection technique will be selected based on how well it fulfills both objectives of the project.

Chapter 2 Literature Review

This section reviews published material related to blood product inventory management. A brief overview of the development timeline of research in this area is given. Next, important characteristics of the research are summarized by geographic region and scope. The merits of research contributions to different aspects of blood product inventory management are then discussed.

A search was conducted for papers pertaining to blood supply chain management, and platelets specifically. Once several landmark papers had been identified, the citations of each were used to locate other relevant research.

The short shelf life of blood products, and platelets specifically, has made inventory control and optimization an important topic of study. The primary goals of research on platelet inventory management are to reduce waste and shortage. Most research focusses on the selection of platelet ordering policies. While inventory and demand monitoring are a key component of selecting order quantities, there is a gap in knowledge regarding transient changes in demand. There is very little published research exploring the detection of changes in demand for platelet products in real time. There is some tangentially related research on modelling the effect of natural disasters on blood supply chains, however disasters do not require detection. Situations such as the introduction of PRT platelets are not common, and thus represent a gap in research.

Research in this area can be categorized by the method used. These methods include Forecasting, Simulation, and integrated Operations Research methods. Work in the 'Simulation' category purely use simulation, whereas work in 'Operations Research' may use simulation in conjunction with optimization models.

Chapter 2.1 Research Timeline

Modern research on blood product inventory management began in the 1950's and 1960's [7][8]. Over the next two decades a small volume of both practical and theoretical work was published. There has been a slow, but steady increase in research in this area since. A literature review on platelet management by Flint [9], published in 2019, claimed that 70 percent of citations on the subject had been published since 2010.

The popularity of different types of research approaches for blood product inventory management has risen and fallen over the last several decades. Work incorporating forecasting was first performed in the late 1970's and 1980's, before disappearing for many years. In the mid 2010's the mainstream adoption of machine learning methods, especially Neural Networks, spurred renewed interest in the topic.

Research based solely on simulation has risen in popularity since the 1980's. A key work by Hesse [10] in the late 1990's became widely cited not only in purely simulation based work, but also in studies using Operations Research. In the 2010's simulation was, for the most part, used on problems of wider scope than was previously possible.

Work using Operations Research techniques has become significantly more developed since 2000. Clearly, increased computing power and information availability has made mathematical optimization of platelet inventory more feasible. While the popularity of research on platelet inventory management, and the different methods used therein, have waxed and waned over time, the goals have remained consistent: reducing wastage and shortage over the long term.

Chapter 2.2 Research Geography and Scope

Blood product management in different areas of the world tends to share important characteristics. For instance, most countries where research has taken place have a single centralized blood authority composed of regional blood centres. However, blood supply chains in different countries operate differently because of their population demographics, geography, and the structure of their institutions. The United States, for instance, represents a significant departure from the general trend of a centralized blood management agency, though other distributed examples exist [11].

The scope of research on blood inventory product management can be divided into three categories: Hospital-level, Regional-level, and Supply Chain-level. Differences between countries should be considered when comparing the scope and results of each work. For instance, the small geographic size of the Netherlands may result in different infrastructure than that in more sparsely populated Canada.

Table 3 below describes the characteristics of each of the citations used, including method and country of study. 65% of the relevant research reviewed was performed in the United States and Canada, with 20% being specifically done in Canada. 10% of research is focused on Europe, and an additional 10% of research coming from East Asia. The remaining 15% of research studies platelet inventory management in South America, Africa, or Western Asia. These numbers may not reflect the distribution of global research accurately, as only work available in English was considered.

Table 3 Characteristics of reviewed literature

First Author (Year)	Method	Citation type	Country of Study
McCullough (1978) [7]	Forecasting	Journal Article	USA
Critchfield (1985) [12]	Forecasting	Journal Article	USA
Filho (2012) [13]	Forecasting	Conference Paper	Brazil
Lestari (2017) [14]	Forecasting	Conference Paper	Indonesia
Khaldi (2017) [15]	Forecasting	Conference Paper	Morocco
Shih (2019) [16]	Forecasting	Journal Article	Taiwan
Motamedi (2021) [17]	Forecasting	Journal Article	Canada
Katz (1983) [18]	Simulation	Journal Article	USA
Hesse (1997) [10]	Simulation	Conference Paper	USA
Atkinson (2012) [19]	Simulation	Journal Article	USA
Asllani (2014) [20]	Simulation	Journal Article	USA
Blake (2017) [21]	Simulation	Journal Article	Canada
Blake (2004) [22]	Operations Research	Conference Paper	Canada
Haijema (2007)[23]	Operations Research	Journal Article	The Netherlands
Van Dijk (2009)[24]	Operations Research	Journal Article	The Netherlands
Zahiri (2013) [25]	Operations Research	Journal Article	Iran
Abdulwahab (2014) [26]	Operations Research	Journal Article	Canada
Gunpinar (2015) [27]	Operations Research	Journal Article	USA
Civelek (2015) [28]	Operations Research	Journal Article	USA
Guan (2017) [29]	Operations Research	Conference Paper	USA

Chapter 2.3 Forecasting

Research in this category attempted to improve waste and shortage metrics by using forecasted demand to inform production decisions. Forecasting methods have been used throughout the history of research on blood product inventory management. Recently there has been an increase in interest in forecasting as a potential application for Machine Learning methods.

McCullough [7] used a two-day moving average approach to determine the number of platelets to order daily at the hospital level. Critchfield [12] used several forecasting methods at the hospital level. These included the method of McCullough, several other types of moving average model, and Winter's prediction model. It was found that Winter's model performed the best and reduced total costs and labour requirements once implemented.

Filho [13] used an Autoregressive Integrated Moving Average (ARIMA) model to forecast demand across regional blood centres and the whole supply chain. The goal of this research was the creation of an automated tool that could be used by managers in different regional blood centers, but few results were provided on the effectiveness of the ARIMA model. Similarly, Lestari [14] used several autoregressive methods to predict demand for several different blood products, finding that a simple moving average model best forecast apheresis demand. Lestari does not conclude if forecasting is effective in improving supply chain efficiency.

Khaldi [15] was among the first to apply machine learning to forecasting for platelet demand. Artificial Neural Networks (ANN) are used to forecast demand for several products at the regional blood centre level. A variety of exogenous variables were tested along with previous inventory values. ARIMA models were used as a benchmark for performance. The performance of the ANN models was found to far exceed that of the ARIMA models. However, Khaldi notes that the latter is far more interpretable. An

additional limitation is that demand data was aggregated to the monthly level, perhaps leading to a loss of fidelity.

Shih [16] compared time-series methods to machine learning methods across the entire Taiwanese blood supply. ARIMA, Exponential Smoothing Models, and Holt-Winters were compared with ANNs and Multiple Regression. They ultimately found that several of the time-series methods performed similarly, however Multiple Regression outperformed the Artificial Neural Network. When the time-series methods were compared to machine learning the results were inconclusive, with different time series models and regression performing better on different data sets. Shih suggests averaging the predicted values from each method.

Motamedi [17] uses a number of methods including ARIMA, ANN, Prophet, and Lasso Regression, to forecast demand at the hospital level, on a daily basis. Prophet is a time-series model which includes long-term data patterns, seasonality, and the non-periodic effect of holidays. These features are combined using a Generalized Additive Model (GAM). Lasso regression is a multiple linear regression technique, which selects a subset of variables to include in the model. At the day-by-day level the additional parameters of the Prophet model give it a 10% increase in performance above ARIMA. It was found that both Lasso Regression and the Neural Network model provided better results than either Prophet or ARIMA, with the Neural Network having the lowest forecast error. The author notes that the ANN has a much longer time, and greater memory complexity, compared to other methods, indicating that it may not scale well to regional or national levels.

The transient portion of a change in demand is not explicitly discussed in any of the works based on forecasting. The more complicated ANN models could be trained to create forecasts in the presence of possible changepoints, however this is not discussed in the research above.

Chapter 2.4 Simulation

While related to Operations Research, methods in this category of the review used purely simulation techniques to model platelet inventory management.

Katz [18] created a discrete event simulation of inventory for a regional blood centre. It was found that as the production quantity selected became more heavily influenced by risk aversion (i.e. increased) the number of shortages decreased, but the number of wasted units also increased. This research was conducted using a 3-day shelf life for platelets and predicted that if the shelf life were to be increased to 5 or 7 days, shortages would be minimized and outdates would be zero. While this result has been shown to be false in more recent contexts, Katz laid a foundation for the simulation of blood product inventory.

Hesse [10] used a discrete event simulation to evaluate different ordering policies at the regional level. Ordering policy improvements were found using an (s, S, t) replenishment model. It was found that the greatest improvement in performance happened at large hospitals and when small rural hospitals were clustered together.

Atkinson's [19] work is focused on the benefits of fresher red blood cell transfusion, and how it may affect blood supply chains. While not directly related to platelet inventory management, Atkinson uses an inventory and hospital simulation to determine how changes in the demand affect outdates and transfusion efficacy. It was found that there is a significant trade off between cost and transfusion efficacy when demand is close to, or greater than, supply. Small imbalances between supply and demand were found to quickly build into significant problems, even at the national level. Platelets have a much shorter shelf life than red blood cells exacerbating the issue.

Asllani [20] designed a decision support system which simulated the collections and demands for apheresis platelets in a large regional blood centre. This includes both ABO priority and shelf life. Several scenarios were tested, with the goal of reducing wastage and shortage. The scenarios tested involved a change in collections, as demand was considered an independent variable. It was found that collecting fewer A+ apheresis platelets, and not collecting on weekends reduced waste by 7%.

Blake [21] examined the inventory impact of increasing the shelf-life of platelets, and how different ordering policies could be used to improve the supply chain under these new conditions. Different ordering polices were required to best reduce waste for each of the values of shelf life, but significant reductions were possible for all increases in shelf life. It was also found that individual hospital characteristics greatly influenced how they were affected by a change in shelf life.

The collected research on platelet inventory management does not consider changepoints in the data. While different demand levels are considered in some works, the transient period surrounding changes in demand has not been studied.

Chapter 2.5 Operations Research

Work in this category finds optimal solutions to platelet inventory problems using techniques such as Stochastic Dynamic Programming (SDP), Approximate Dynamic Programming (ADP), and Integer Stochastic Programming (ISP), often in conjunction with simulation methods. Where an exact solution to optimization was not possible, simulation could be used to find near-optimal solutions.

Blake [22] developed a Dynamic Programming model to identify locally optimal ordering policies for both the blood authority and consumers. Blake reduced the dimensionality of the problem by combining orders into groups of units. It was found that the accuracy of the model was a function of both the planning horizon, and the group size. The longer the planning horizon and smaller the group size, the more accurate the model was. This approach was able to reduce both waste and shortages.

Haijema [23] created a Markov Decision Process formulation for platelet ordering policies at the regional level. The state space and demand are downsized, to make the approach feasible. A simulation approach was then used to search for the single best ordering policy. It was found that the simulation provided near optimal results in both the downsized and full-scale problems. Van Dijk [24] used the same model to evaluate group preferences and an extended shelf life.

Abdulwahab [26] used both Linear Programming and Approximate Dynamic Programming to develop a model of a single hospital and blood bank that used eight blood types. Their approach was able to find an optimal solution without downsizing, reducing waste and shortage at steady state. Similarly, Gunpinar [27] used a Stochastic Integer Programming model to model hospital level inventory, finding an optimal solution.

Civelek [28] follows much the same structure as Haijema [23] and Van Dijk [24] with the addition of a critical level protection policy. This policy protects fresh units against excessive substitution. It was found that doing so could help smooth out inventory and reduce costs. Civelek notes that while demand was assumed to be stationary, their heuristic search should be extended to non-stationary demand.

Guan [29] analyzed factors in platelet usage at the hospital level to determine what drove demand. These included units transfused in the previous days/weeks, census data, and complete blood count for inpatients. An optimization model was created, based on the relevant demand factors to collect platelets based on this information. It was found that this strategy reduced wastage when tested on historical data. Interestingly, this model could easily continue to be used even when a change in demand occurs.

Work using Operations Research is largely focussed on providing ordering policies for hospitals which minimize cost and waste under the assumption of steady state demand. While changepoints are not explicitly considered, the methods of each paper could be used to find new ordering policies for different demand levels.

Chapter 2.6 Analysis of Literature

Previous work on blood product inventory management has focussed on decision support with the goal of reducing waste and shortage. Inventory monitoring is an important component of these models; however, it is performed with the assumption that the properties of demand do not have changepoints. Previous research has dealt with improvement over the long term, by reducing the average amount of wastage and/or shortage. There is then, a significant gap in research on monitoring itself, and the transient component of changes in the blood supply chain.

Chapter 3 Methods

Chapter 3.1 Methods Overview

To achieve the objectives of this study two key functions are required. First, a detection method suitable for changes in inventory is needed; Statistical Process Control (SPC), a quality control technique will be used to monitor inventory. Second, evaluation data is required to test the detection method; however, because data representative of possible demand changes does not exist, a custom discrete event simulation was used to simulate demand change.

A change detection method and evaluation data (i.e., looking at data as the process evolves) can be used to fulfill the primary objectives of this project. However, any opportunity to increase performance is valuable and thus the SPC methods can be supplemented with short-term forecasting to reduce detection time for true positive signals, as will be demonstrated here. To increase decision support capability, once a demand change is detected, the magnitude of the change, size and direction, will also be estimated using a technique that compares the mean value of recently observed points against established process properties. For ease of integration, all components of this project were developed in R using R Studio.

Figure 1 below describes how the project components are used in testing. First, the user inputs the characteristics of a demand change experiment. Those characteristics are passed to the inventory simulation, which outputs replications of data representative of that change. Next, the data is passed to a forecasting model which creates predictions for all points in the time series. The expected values from the forecast, and the actual data, are used by the changepoint detection algorithm to determine if, and when, a change is detected. If a change is detected, its properties are estimated, and the results are output to the user. The flow of project components while deployed, Figure 2, is much the same as in testing, but data comes from the CBS inventory database and the system runs in perpetuity.

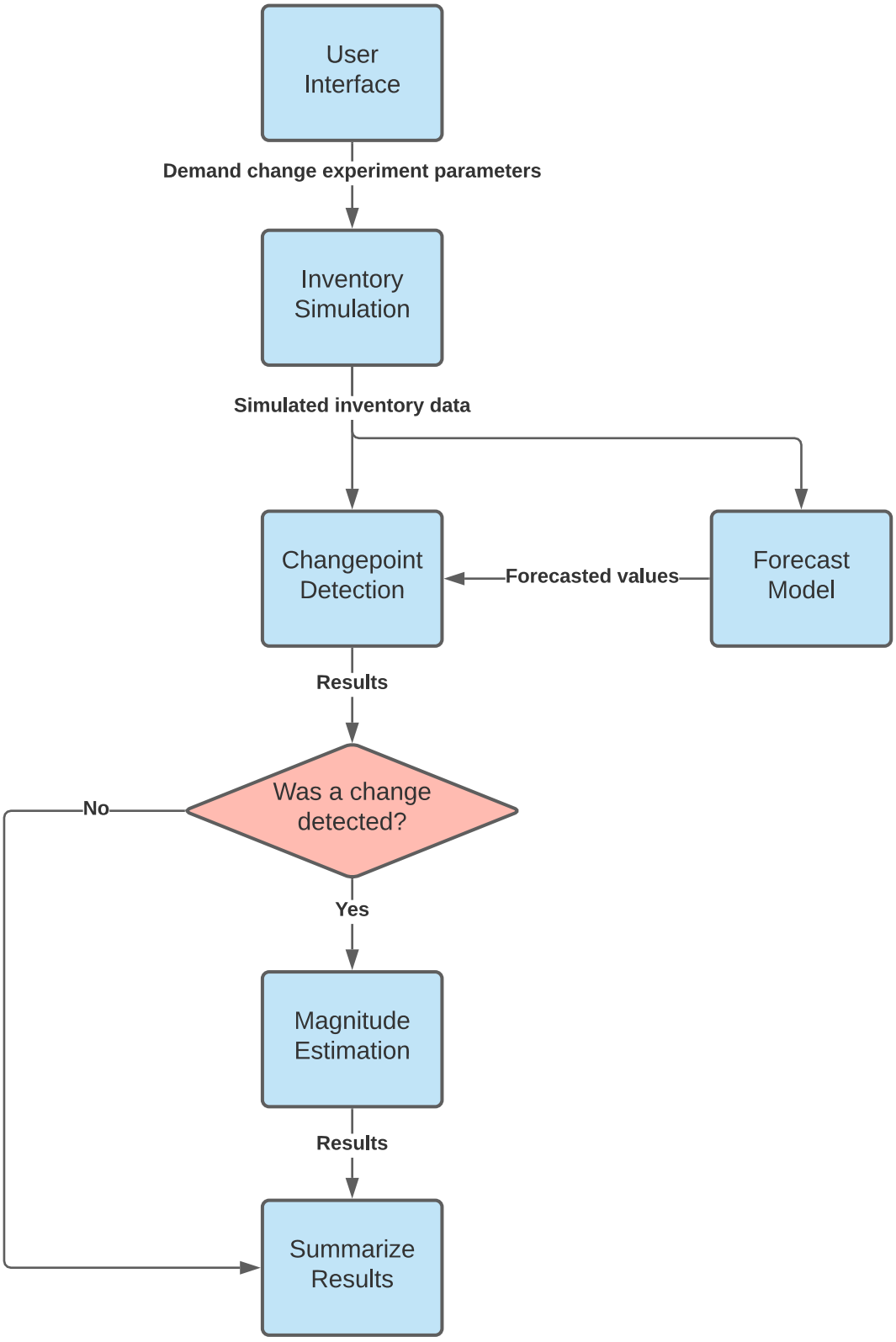

Figure 1 Block diagram describing the use of project components during evaluation

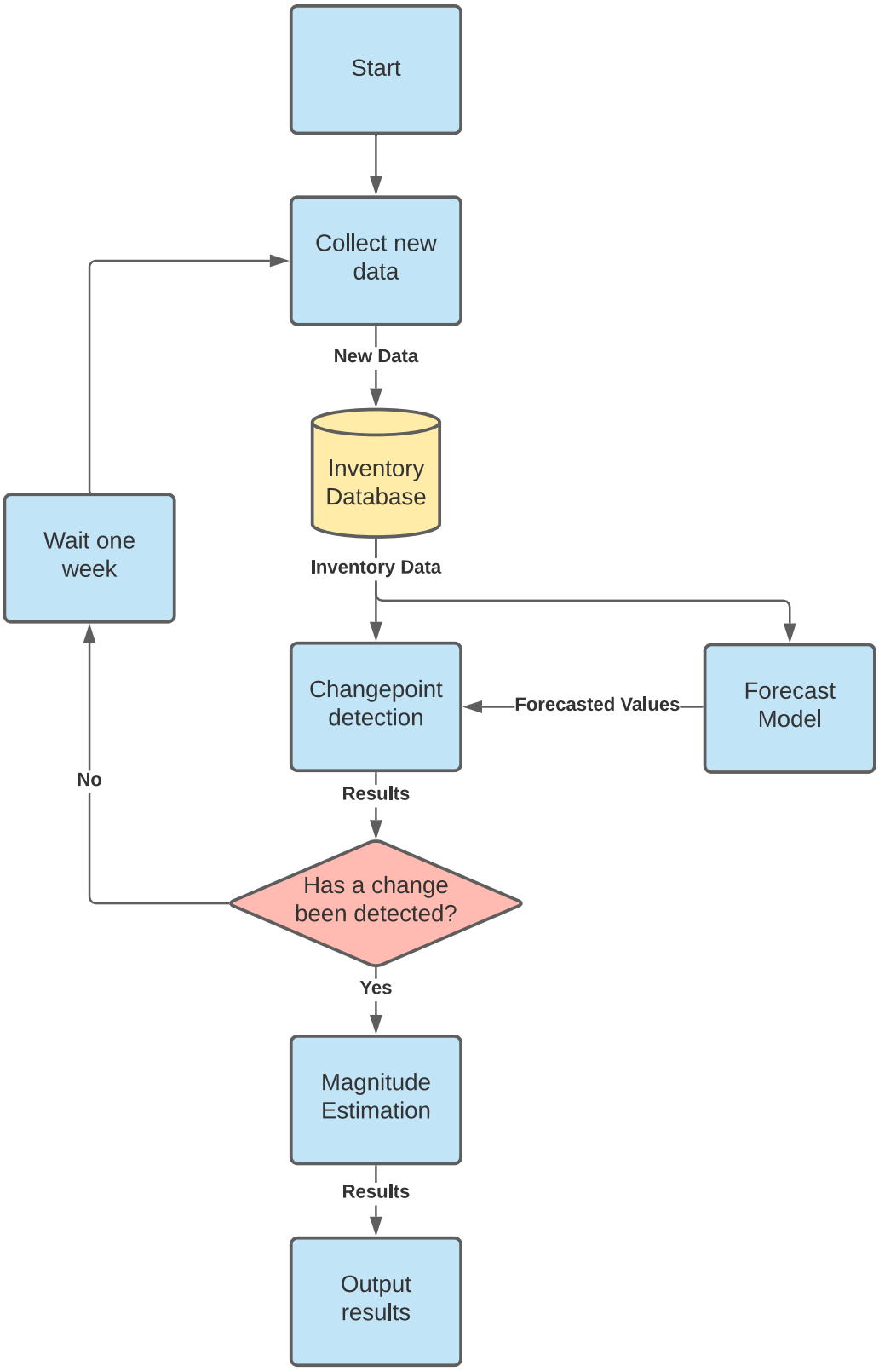

Figure 2 Block diagram describing how project components will be deployed by CBS

Chapter 3.2 Inventory Simulation

An inventory simulation was purpose built for testing change point detection methods; the simulation creates representative demand data. This approach allows the use of several assumptions that reduced model complexity. The simulation proceeds one day at a time, ignoring any time-of-day dependence of demand, collections, and shipping, instead treating them as a set of events occurring in a fixed sequence each day. All new collections arrive at the beginning of the day, as does the daily demand.

Further, the simulation model treats collections and demand in aggregate instead of simulating individual points of use. Thus, all collections come from a single source representing the combined output of all mobile and permanent clinics. Similarly, the demand of all hospitals is represented by a single item. As inventory is the focus of this study, these simplifications greatly reduce computation time, while preserving the ability to evaluate system wide changes.

Chapter 3.2.1 Distribution Fitting

The data used to populate the simulation comes from CBS' Ottawa Centre (also known as NEON or Northeastern Ontario) and covers the 2019 calendar year. The calendar year was selected, as opposed to the 2019 fiscal year, to avoid anomalies due to the COVID-19 pandemic.

The assumption was made that the daily quantity demanded of each blood type is independent of the previous day's quantity demanded, when considered at the regional level. While rare platelet types may be required on consecutive days at specific hospitals, when viewed in aggregate, demand data may be considered independent. This assumption was tested by evaluating the 1-day lag autocorrelation for the collections and demand of each product, included in Appendix A. These results show that the autocorrelation level is insignificant for all products. The Poisson distribution is typically used to model independent arrivals from large populations, however the frequency of days with zero

orders was much higher than in a typical Poisson process. This affected all blood, and product types, but was especially acute for rarer blood types. To this end, a zero-inflated Poisson distribution was fit to both the daily quantity collected, and demand, for each blood and product type segment[30]. The zero-inflated Poisson model includes an additional parameter defining the probability of the number of arrivals being zero. The model assumes that the increased number of zeros are generated by a separate process than the one that generates count values. In this case, the two processes dictate whether an order or collection happens at all, and the quantity order/collected if an order occurs. The parameters of each distribution were fit in R, with the 'fitdistrplus' library.

As an example, theoretical and empirical distributions for orders of A+ pooled platelets are compared below in Figure 3. The zero-inflated Poisson distribution appears to fit the data well, both graphically and when compared using a Chi-square test. The distribution fitting results for orders and collections of all 16 products are included in Appendix A and show that the ZIP model appears to fit all blood/ product types well.

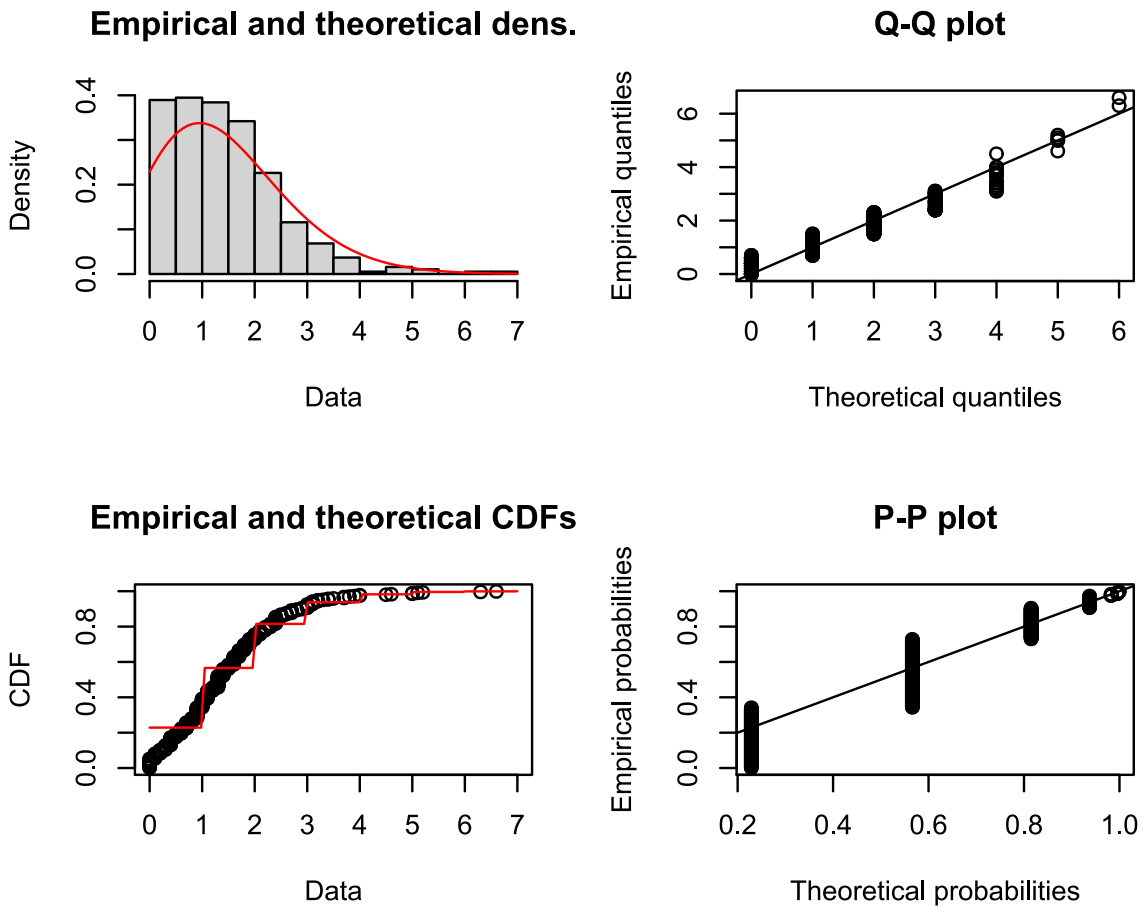

Figure 3 Comparisons of the empirical and theoretical zero-inflated Poisson distribution for orders of A+ pooled platelets

Chapter 3.2.2 Simulation Structure

An overview of the simulation structure is shown in Figure 4. The simulation is composed of three objects: collections, inventory, and demand. An inventory database keeps track of daily throughput, information that is then used to determine how orders should be filled.

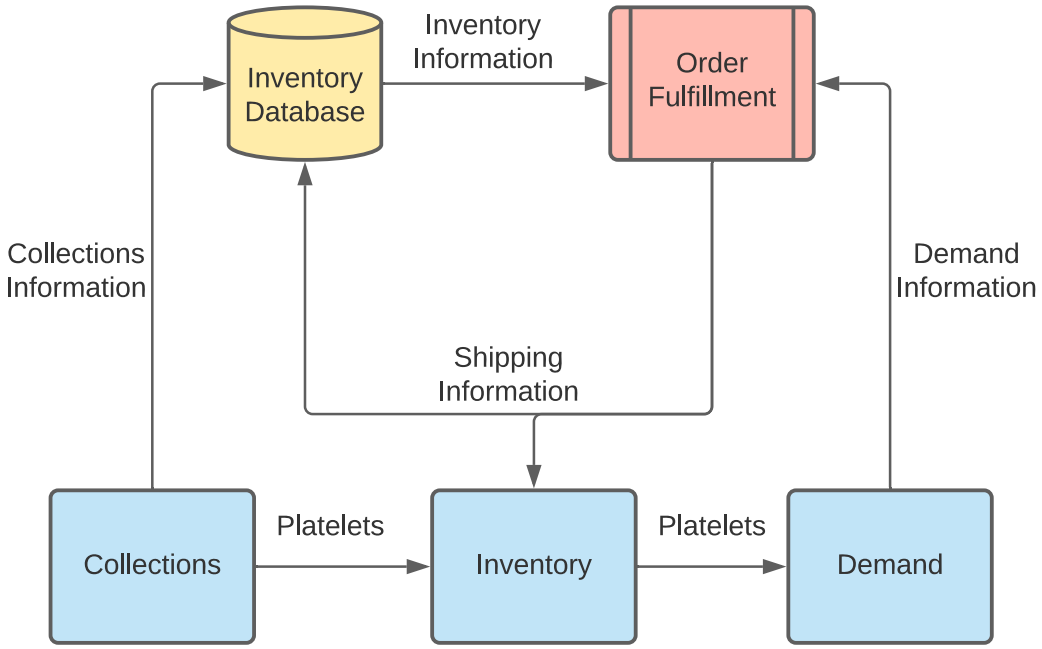

Figure 4 Simulation Block Diagram

The form of each of the three top-level objects is shown below in Equation 1. Arrows indicate the direction of the flow of units. 'Allocation' refers to the order fulfillment process described in the next section.

Equation 1 Descriptions of top-level simulation objects

$$\begin{bmatrix} A^{+}_{pooled} \\ \vdots \\ O^{-}_{apheresis} \end{bmatrix} \rightarrow \begin{bmatrix} shelf\ life\ (5) & \cdots & shelf\ life\ (1) \\ \vdots & \ddots & \vdots \\ shelf\ life\ (5) & \cdots & shelf\ life\ (1) \end{bmatrix} \xrightarrow{allocation} \begin{bmatrix} A^{+}_{pooled} \\ \vdots \\ O^{-}_{apheresis} \end{bmatrix}$$

Collections, on the left, are created daily and represented by a vector showing the amount collected of each ABO type, for each product. Each collected unit is placed in inventory, in the middle of the diagram, with a remaining shelf life of 5 (or 7), days. The diagram shows the units of each product that have just been placed in inventory on the left, depicted as 'shelf life (5)'. Units that are about to expire are shown on the right as 'shelf life (1)'. At the end of each day, if they are not allocated to fill orders, units have their shelf life decreased, moving from left to right in the same row of the inventory

matrix. If they have not been used when they have a remaining shelf life of 0 days, units are marked as outdated and exit the simulation. Units that have been collected and placed in inventory, must complete testing before they can be released to hospitals. This is modelled in the simulation by making those units inaccessible to the allocation process for the duration of testing. Each day, demand for platelets, shown on the right of Equation 1, is created using the same method as collections. Units in inventory are used to meet demand, modelled in the simulation by subtracting a unit from its position in inventory and from the demand vector. When all demand has been satisfied and all elements in the demand vector are zero, then the simulation proceeds to the next day.

Chapter 3.2.3 Inventory Allocation

The inventory allocation method for the simulation attempts to recreate the decision-making process of managers, distributing inventory to hospitals. This is accomplished using a heuristic that attempts to reduce cost, while filling all orders. The steps of the order fulfillment heuristic are as follows:

1. Exactly match apheresis inventory with apheresis demand, with priority given to units with lowest remaining shelf-life (i.e., a FIFO inventory policy)
2. If there is unsatisfied demand for apheresis units, which cannot be exactly matched, substitute a compatible apheresis unit, with priority given to units with the lowest remaining shelf-life
3. Check the shelf life of apheresis inventory. If there are any with a remaining shelf life of 0 days, use them to fill compatible orders for pooled platelets
4. Exactly match pooled inventory with pooled demand, with priority given to units with lowest remaining shelf life (FIFO)
5. If there is unsatisfied demand for pooled platelets which cannot be exactly matched substitute a compatible pooled unit. Priority given to units with least remaining shelf-life

If demand cannot be satisfied using a compatible unit, then the heuristic will fill demand with an incompatible unit while continuing to minimize cost with apheresis substitution. However, this is not a common occurrence, and it happens very rarely in the base model. In practice, if the level of inventory in the simulation is low enough that no compatible units can be found, then it is quite likely a change in

demand has already been detected. Since the purpose of the simulation is not to model the response to shifts in demand, only to create shifts for the detection process, this simplification is deemed acceptable for modelling purposes.

While the steps of the general inventory allocation heuristic are straightforward, applying it to 16 products with different compatibility, and priority adds significant complexity. The implementation of the inventory allocation heuristic is based on a 'search map' design, wherein the order in which elements of the inventory matrix are searched is pre-determined for each product. This method was selected because it condenses allocation behaviour for all products into standardized sets of rules. This streamlined the design for each of the products and allowed the use of a generic allocation algorithm structure.

Allocation is performed in two phases. The first phase deals with exactly matching apheresis orders to apheresis demand with units of the same ABO type, as seen in Figure 5. The second phase, shown in Figure 6, fills all remaining orders.

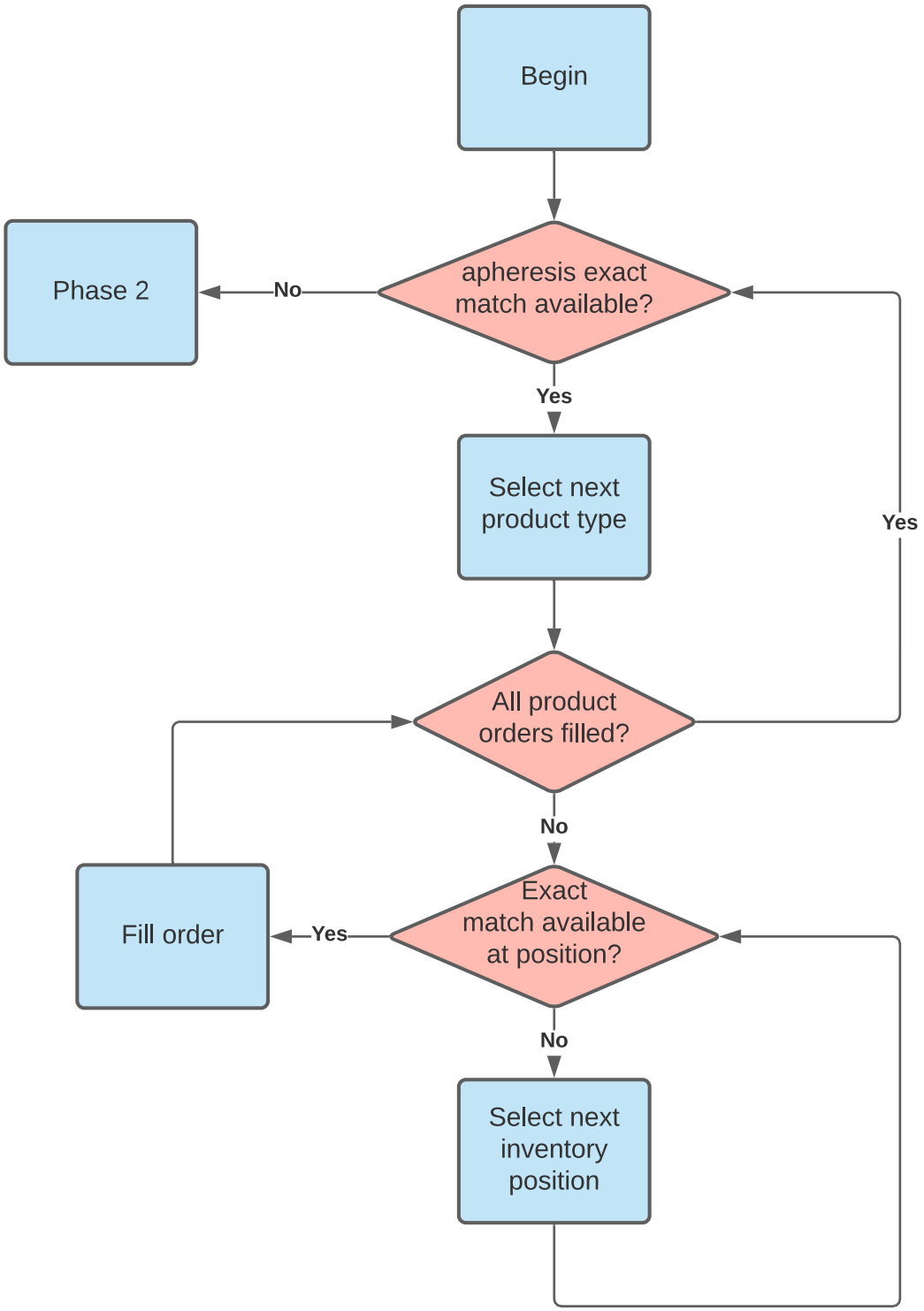

Figure 5 Inventory allocation algorithm Phase 1 flowchart

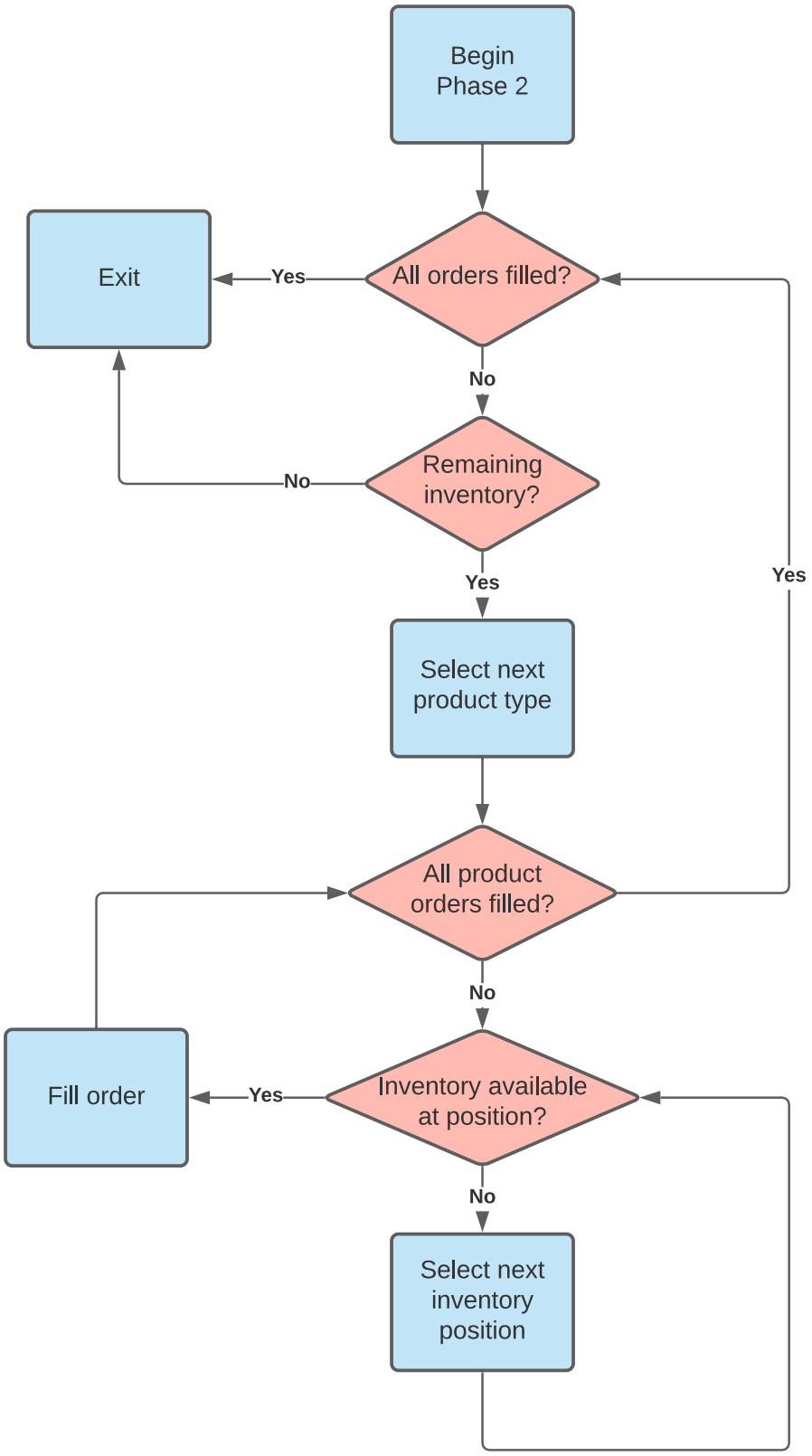

Figure 6 Inventory allocation algorithm phase 2 flowchart

A simplified example of this method with two product types, each with a shelf life of 3 days, is described below. In this example, product 1 takes the role of apheresis platelets and product 2 the role of pooled platelets. Product 1 has a higher associated cost and may be substituted for product 2 when necessary. Equation 2 shows a 2x3 inventory matrix, each element representing the held quantity of each product at a different remaining shelf life.

Equation 2 Example inventory matrix, showing remaining shelf life and product types

$$\begin{bmatrix} 3_1 & 2_1 & 1_1 \\ 3_2 & 2_2 & 1_2 \end{bmatrix}$$

The search maps for both products show the order in which the elements of the inventory matrix should be searched when allocating units. For example, Equation 3 shows the search map for product 1. The first inventory to be used is in position '1', which corresponds to the inventory matrix element for product 1 inventory on it's last day of useable shelf life. If there is no inventory in this element of the matrix position 2 is searched followed by position 3, the matrix elements containing product 1 inventory with remaining shelf life. If there is no product 1 inventory, and orders remain positions 4, 5, and then 6 will be searched for units. As previously stated, the substitution of incompatible units does not occur often in practice however handling those situations is still built into the simulation.

Equation 3 Example search map for product 1, filling the role of apheresis platelets

$$\begin{bmatrix} 3 & 2 & 1 \\ 6 & 5 & 4 \end{bmatrix}$$

Equation 4 shows the search map for product 2. Its initial search position is the same as Equation 3, first substituting remaining units of product 1 that would otherwise expire to mitigate cost. Next the search moves to product 2 inventory in positions 2, 3, 4 moving from lowest to highest remaining shelf life. In the event all product 2 inventory has been depleted, remaining product 1 inventory is searched in positions 5 and 6, although in practice this is a rarity.

Equation 4 Example search map for product 2, filling the role of pooled platelets

$$\begin{bmatrix} 6 & 5 & 1 \\ 4 & 3 & 2 \end{bmatrix}$$

These 'search maps' are the basis for how inventory is allocated in the example system below. Equation 5 shows the initial state of the system before inventory has been allocated to meet demand. The available inventory, of each product at each value of remaining shelf life, is shown in the matrix on the left, while the demand for each product is shown in the vector on the right. These two objects represent the state of the system once collections have arrived for the day.

Equation 5 Simple inventory allocation example, initial system state

$$\begin{bmatrix} 1 & 0 & 3 \\ 1 & 4 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 2 \\ 5 \end{bmatrix}$$

The 'search map' for product 1 is used to allocate 2 units to demand as seen in the inventory state of Equation 6. This reduces the inventory of product 1 on its last day of useable shelf to 1 unit and reduces the orders for product 1 to zero.

Equation 6 Simple inventory allocation example, system state after fulfilling product 1 demand

$$\begin{bmatrix} 1 & 0 & 1 \\ 1 & 4 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 0 \\ 5 \end{bmatrix}$$

Next, using the product 2 search map, expiring product 1 inventory is used to fill demand for product 2 resulting in the inventory state seen in Equation 7. This mitigates the cost of outdating a unit of product 1.

Equation 7 Simple inventory example, system state after using expiring product 2 units to fill product 1 demand

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 4 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 0 \\ 4 \end{bmatrix}$$

Lastly, the product 2 search map is then used to allocate product 2 inventory to fill the remaining demand resulting in the final system state shown in Equation 8. The only unit of product 2 with a single day of remaining shelf life is allocated to fill an order, followed by 3 units of product 2 with

2 days of remaining shelf life. At this point the simulation ages units 1 day, reducing their remaining shelf life, and moves to the collection process of the next day.

Equation 8 Simple inventory allocation example, final system state for the day

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Chapter 3.2.4 Simulation Control

The stochastic nature of both collections and demand mean that the inventory level is subject to variations. Even slight imbalances between collections and demand will cause inventory to change greatly from the desired level over an extended period. Without external control, a supply-demand imbalance will cause the inventory level to tend to zero, or grow indefinitely, depending on the direction of the imbalance. However, in practice this does not happen because people monitor the process and adjust inventory accordingly. To simulate the adjustments made by managers a feedback controller is included in the simulation, as described in Figure 7.

There is a limit to the effort the controller can exert, representing the level of adaptability of the system. For example, in practice if there is a shortage of units, surplus from other regions can be shipped into Northeastern Ontario, but there is a limit to the quantity of units that can be transferred without disadvantaging other areas.

The control effort is asymmetric for high and low inventory levels. Collections are increased more for a deficit, 12%, than they are decreased for a surplus, 5%. During development it was found that reducing collections in response to a surplus often led to a significant overshoot, reducing inventory far below the desired level. This is likely a result of naturally expiring products also contributing to a decrease in inventory.

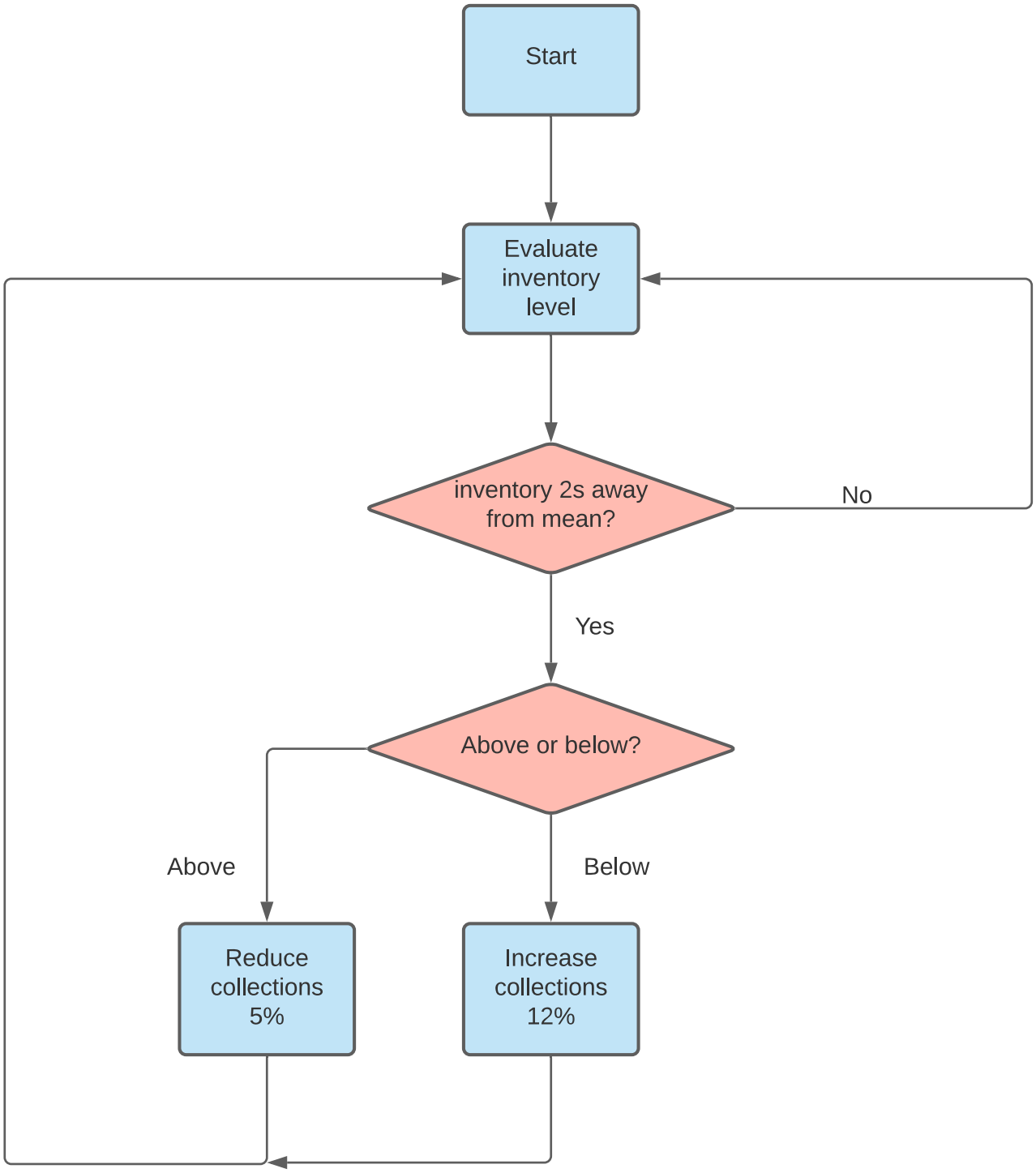

Figure 7 Simulation inventory controller flowchart

Chapter 3.2.5 Simulation Parameters

The time-dependent, autocorrelated nature of the data means that production runs must be performed in separate replications. A diagram indicating the three phases of a run is shown in Figure 8. First, the simulation is warmed-up for a period of 6 months (24 weeks). The length of the warm-up period was determined graphically using Welch's method [31]. Next, to prime the forecasting algorithms, the simulation is run in its initial state for 36 weeks. If there is a shift in demand, it is applied at the beginning of the detection period and tracked for 52 weeks.

Figure 8 Diagram indicating different phases of an inventory simulation replication

Chapter 3.3 Changepoint Detection

Platelet inventory level varies when measured at different points in time. Finding changes in the characteristics of time-series data as it becomes available is broadly referred to as Online Change Point Detection (OCPD) [32]. Several methods were tested as part of this study including a Bayesian approach to Online Changepoint detection, the results of which can be found in Appendix B. The method that best fulfilled both project objectives was a variation of the Shewhart Control Chart, an established method commonly used in Quality Control as part of Statistical Process Control (SPC) [6]. The visual representation of this technique, a Control Chart, is shown below in Figure 9.

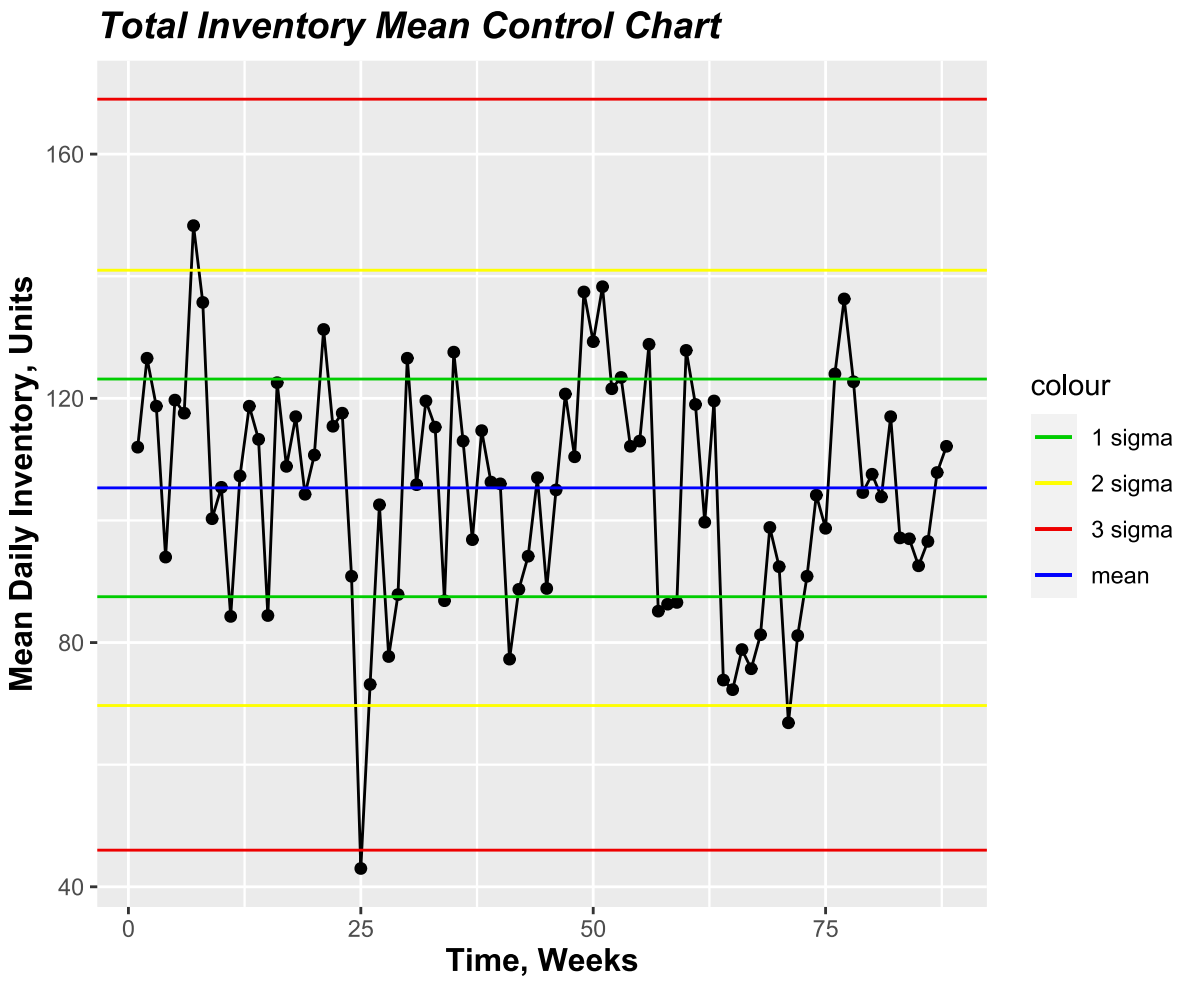

Figure 9 Shewhart Control Chart for total weekly platelet inventory

Chapter 3.3.1 Statistical Process Control

Statistical Process Control is a set of techniques used to monitor a process, with the intention of ensuring its stability. A stable process can be adjusted to reduce variability and improve performance. The process mean is often studied, as is the process range, although techniques are described in the literature to study other attributes. In this problem the mean inventory level is the quantity of interest. We note that SPC methods are commonly used in industry; Canadian Blood Services makes extensive use of SPC methods for monitoring collection, production, and inventory processes.

The first step in studying a process is to determine its statistical properties, chiefly the mean and variance. Next the sample size and sample frequency must be determined. While the choice is not always an obvious one, in this case all data from the NEON region, for each day, could be used in the analysis. SPC typically batches data into groups. This is done to 'smooth' the data and reduce the effect of outlier values. Sub-groups are selected to maximize the chance that a difference between sub-groups is the result of a real change. In this project, the subgroup of a 7-day week is used, to reduce the effect of day-of-week effects for platelet demand on the analysis.

The most widely used SPC tool is the Shewhart algorithm, represented visually by the Shewhart control chart as seen in Figure 9. The Shewhart algorithm compare the values of recent points in a series, in this case ordered by time, against established process properties. Using the Shewhart algorithm the position of points, or groups of points, relative to the mean process level, shown by a blue line in Figure 9, indicates how likely it is they would fall there by chance.

Several sets of anomaly detection rules exist, designed to sensitize the method to different forms of property changes. There is, however, a trade-off between the false alarm rate, and the mean time to detection when selecting the number of conditions included. The most basic condition evaluates whether a point falls outside of three standard deviations of the mean, indicated by red lines in Figure 9. Additional conditions look for bounded changes, such as if 2 out of 3 points in a row sit outside of two standard deviations of the mean, or if 4 out of 5 points sit outside one standard deviation of the mean. There are other conditions intended to detect other types of anomalies including prolonged bias. However, testing showed that their use greatly increasing the false positive rate. Aside from de-sensitization to changes in statistical properties by removing conditions, there are several other methods commonly used to tune SPC for use on autocorrelated data [6]. These include using the residual errors (the difference between the observation and the process means) of an autoregressive model as data points on a control chart. However, it was found, in testing, that this family of techniques had worse performance than the standard control chart as can be seen in Appendix C.

A complete list of the changepoint detection conditions included is as follows:

1. A single point more than 3 standard deviations away from the mean total inventory level
2. 2 out of 3 consecutive points more than 2 standard deviations away from the mean total inventory level
3. 4 out of 5 consecutive points more than 1 standard deviation away from the mean total inventory level
4. 2 out of 3 consecutive points more than 3 standard deviations away from the mean apheresis inventory level

The importance of additional conditions is shown below in Figure 10. The mean total daily inventory is never more than 3 standard deviations away from the mean inventory level, but a change is detected under conditions 2 and 3. The points highlighted in orange are the first to violate condition 2, while those in purple are the first to violate condition 3. The process is considered ‘out of control’ at week 39, the second orange point.

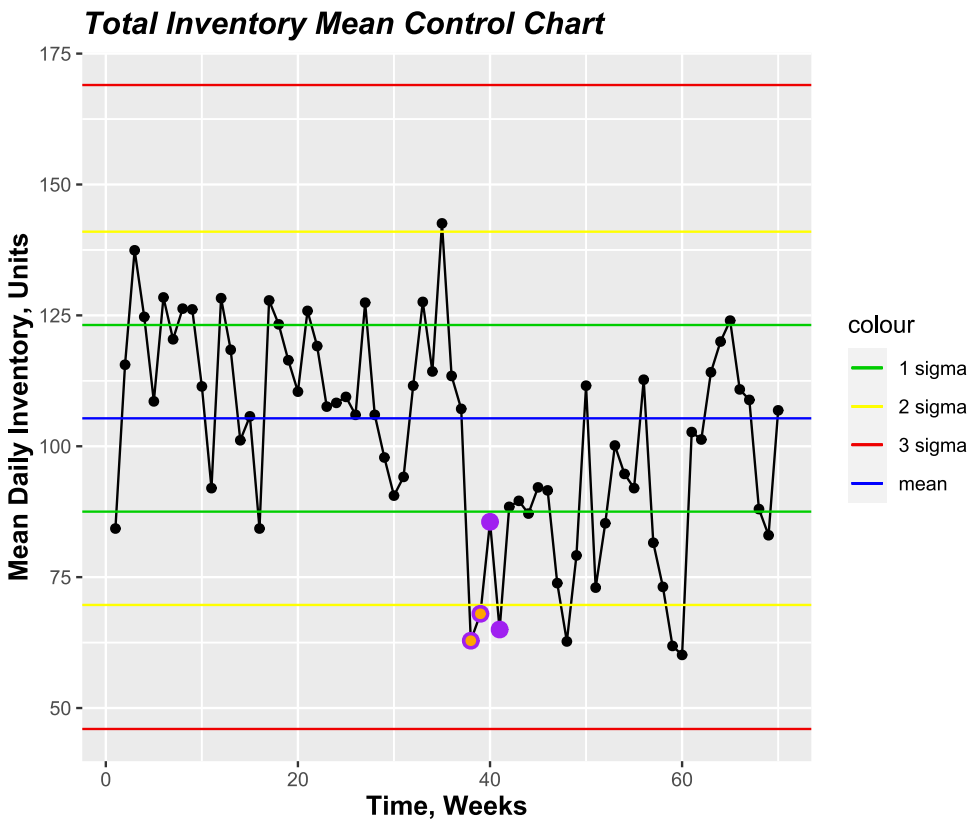

Figure 10 Total inventory mean control chart; data points are within control limits, but exhibit changes identified by rules 2 and 3. Points highlighted in orange are the first to violate condition 2, while those in purple are the first to violate condition 3

The dual inventory of pooled and apheresis platelets may be affected differently by changes in demand, as demand changes may present more in one product than the other. While pooled inventory tracks quite strongly to total inventory, demand changes which predominantly effect apheresis may go unnoticed when monitoring only total inventory. Hence, the apheresis inventory stream must be monitored, in addition to total inventory. Figure 11 demonstrates the value of tracking apheresis separately. Changes in apheresis demand patterns that do not register in total inventory, are magnified when viewing the apheresis inventory stream.

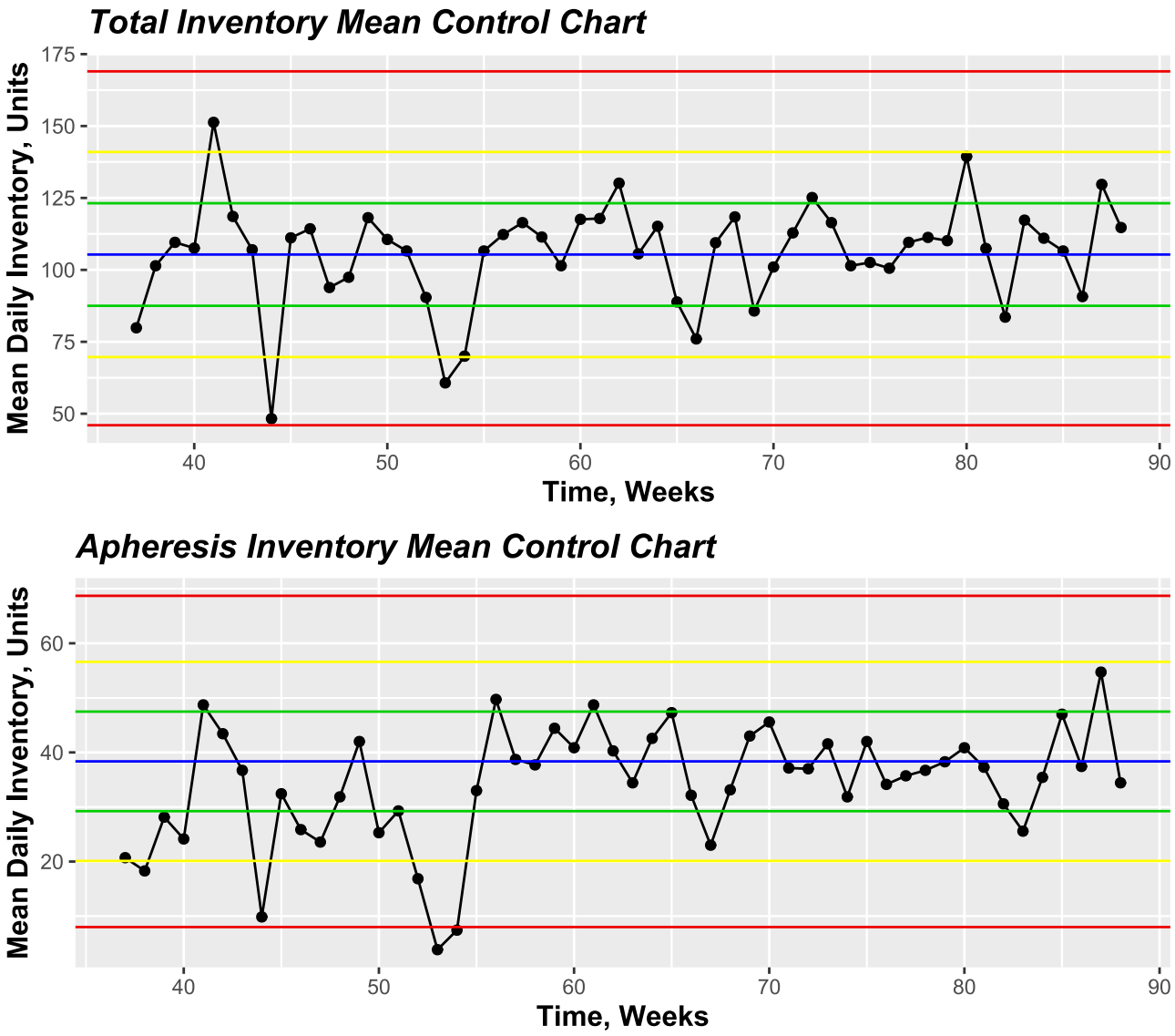

Figure 11 Total and apheresis inventory mean control charts; a change is detected in apheresis inventory but not in total

Chapter 3.4 Forecasting

The time required to find changes in a time series is limited by the rate of acquisition of new data. In this case, new data is collected daily, and analyzed weekly to reduce noise. As even a single week may matter, any opportunity to decrease the time to detection for a change in demand is important. Accordingly, in this study, we evaluate the use of local forecasting to supply the Online Changepoint Detection algorithm with additional data points, potentially reducing time to detection without having to increasing the rate of data acquisition. Several broad categories of forecasting were tried in this role and are described below. These include Linear Regression, ARIMA, and different two variations of Generalized Additive Models (GAM).

Please note, the expected value of the prediction is used in the calculations shown below. However, each predicted data point is known only with some amount of uncertainty. The range of uncertainty in a future value can be shown as a prediction interval. Accordingly, it would be reasonable to use any value within the range of the forecast point's prediction interval in the detection algorithm, not just the mean. By adjusting the value used, the data supplied to the changepoint detection algorithm could be made to provide a more (or less) conservative estimate. However, for the purposes of this analysis we will use the expected value of the forecast points.

Chapter 3.4.1 Linear Regression

Least Squares regression is a simple, powerful method included largely as a benchmark for other methods. The intention is to compare the predictive ability of more complex methods with a simple alternative to determine if the increase in computational effort is warranted. The prediction model for this method is shown below in Equation 9. Least Squares linear regression fits a straight line to data by selecting 'Beta' coefficients which minimize the sum of the squared residual errors, the difference between model values and actual values, and thus emphasizes outlying points.

Equation 9 Forecasting model for single predictor, least squares linear regression

$$\hat{y} = \beta_0 + x\beta_1$$

A graphical example is shown below in Figure 12. The orange line shows the prediction model created by fitting a linear model to a neighborhood of points. The next week's predicted value violates condition 2 of the change point detection scheme, 2 out of 3 points more than 2 standard deviations away from the mean, resulting in a positive detection signal. The actual data point for the next week also violates condition 2.

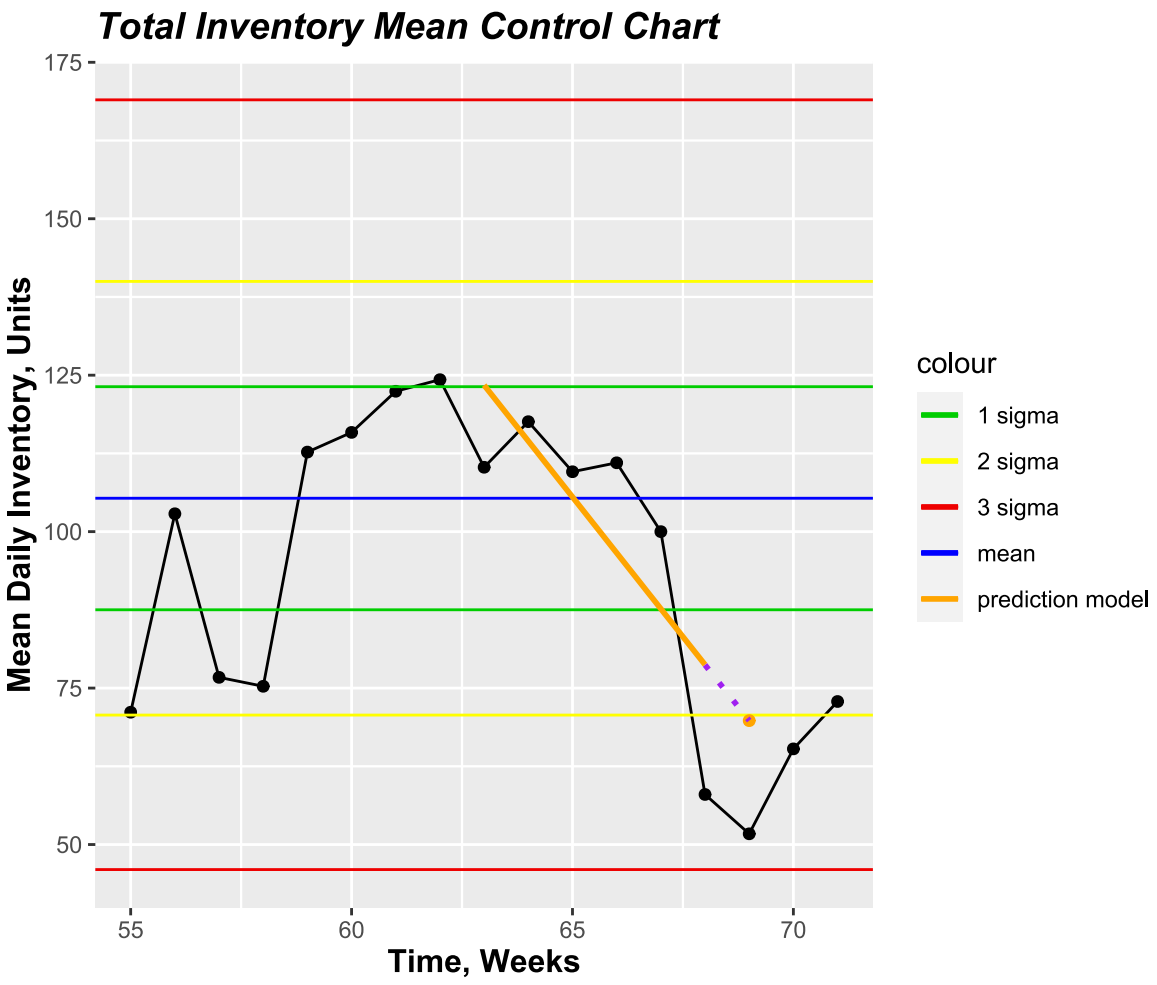

Figure 12 Total inventory mean control chart, including a linear prediction model

Chapter 3.4.2 ARIMA Family

Autoregressive Integrated Moving Average (ARIMA) models are a generalization of several forecasting techniques, allowing the convenient formulation of an autoregressive model to data that potentially shows seasonality or a trend [33]. ARIMA models are described using the notation (*p,d,q*) outlined below:

- *p* is the number of autoregressive terms, describing how many past terms are used to estimate predict future values
- *d* is the number of differences required for the series to be stationary. Indicating the order of seasonality in the data
- *q* is the number of lagged forecast errors in the prediction equation, describing the smoothness of the model

In our model, the data is already stationary making this an ARMA model since d is 0. The general ARMA forecasting model is shown below in Equation 10. If days were considered, instead of weeks, it is likely that differencing would be required.

Equation 10 Forecasting model for an ARIMA model

$$\hat{y}_t = \mu + \phi_1 y_{t-1} + \cdots + \phi_p y_{t-p} - \theta_1 e_{t-1} - \cdots - \theta_q e_{t-q}$$

An example of an ARMA model forecasting inventory data is shown in Figure 13. The orange series is the 1-step lagged forecasted value, or the predicted value at each point. However, if there is a sudden change in the values of datapoints, ARIMA models may lag. Consequently, ARMA models may not contribute greatly to a reduction in detection time, however conservative forecasting, such as this one, does not affect the false alarm rate. ARIMA models were implemented using the 'forecast' package in R.

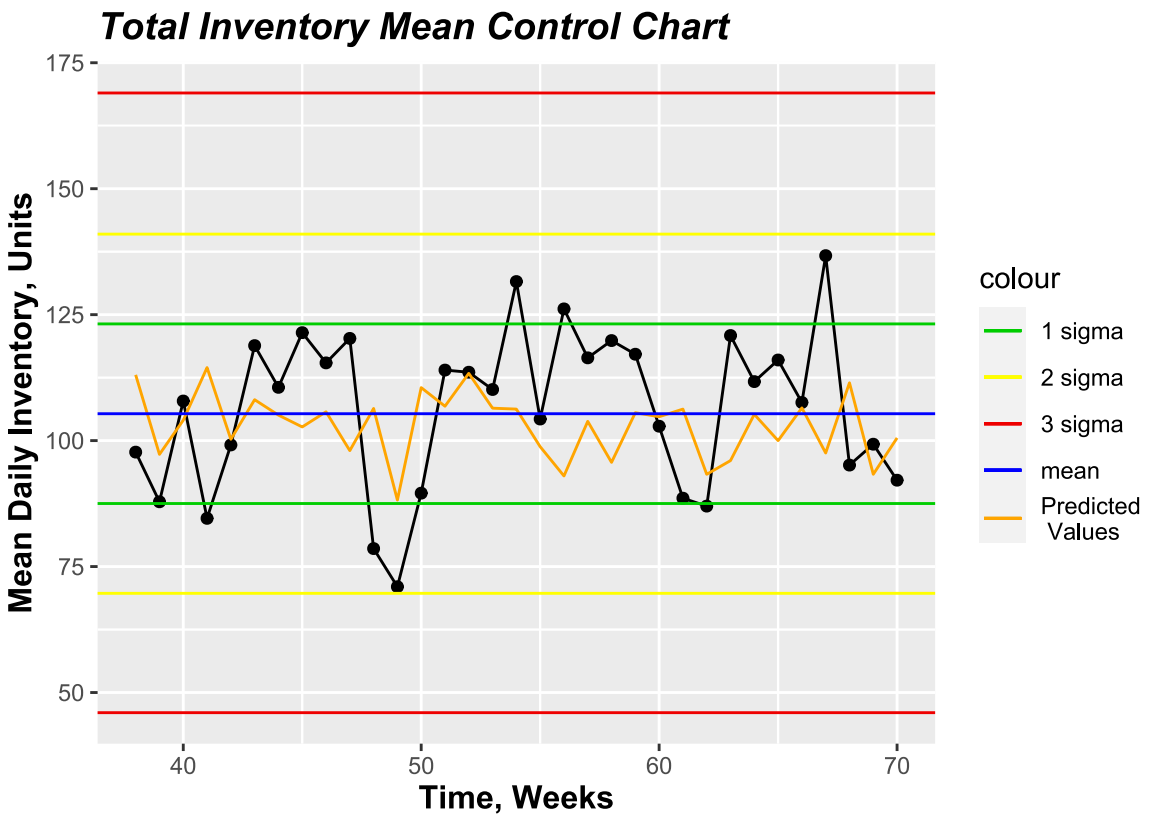

Figure 13 Total inventory mean control chart, including an ARIMA family prediction model

Chapter 3.4.3 Generalized Additive Models

GAMs are a generalized approach to forecasting that fits smooth curves to unknown models [34]. The additive component of a GAM describes how the behaviour of several predictor variables can be added together, using smooth curves, to produce a non-linear model. The forecasting model for a GAM is shown below in Equation 11, where the function, s(x), is a smooth model regressed on a single variable.

Equation 11 Generalized Additive Model prediction equation

$$g\big(E(Y)\big) = \alpha + s_1(x_1) + \cdots + s_p(x_p)$$

Individual model fitting can take place using one of several methods, usually local regression or regression splines. The smoothing parameters, which control the degree of non-linearity, for each of the individual models are selected using a back-fitting algorithm that minimizes the expected squared error of the model. A numerical approximation of a Restricted Maximum Likelihood approach, REML, is used to solve the smoothness selection problem. Backfitting is handled by the 'mgcv' package in R. While the choice of approach to the smoothness selection problem is somewhat arbitrary, there is evidence to suggest that REML performs well with smaller sample sizes [35].

In this instance there is only one predictor variable, previous inventory. This makes a GAM fit to the data equivalent to the method used to fit a single variable. However, the backfitting algorithm component of the GAM structure is still useful as a method to automatically tune the smoothing parameters of the individual model. This is useful, as the monitoring and forecasting systems are meant to operate without user input.

Both regression splines and local regression were used to fit the underlying model. Local regression fits a polynomial to a span, or neighborhood of points. The smoothness of the model is controlled by changing the degree of polynomial fit, and the span. Fitting a GAM using splines differs from polynomial regression in that data need not be fit using a single polynomial. Several piece-wise polynomials may be connected using knots instead of a single high-degree polynomial. The smoothness of the model is controlled by changing the number of knots, or degrees of freedom of the model. These two methods may yield similar results depending on the degree of smoothness; however, the piecewise functionality of splines tends to make them more susceptible to variations in data, and overfitting. The trade-off is that a spline-based model may pick up on changes in the data more quickly.

An example of a GAM forecasting model for inventory, fit using splines, is shown below in Figure 14. The orange line shows the model fit to the data, and the shaded area the 95% confidence interval. The smooth curve of the model is in between the roughness of ARIMA models, Figure 13, on one hand, and the lack of variation in linear regression, Figure 12, on the other.

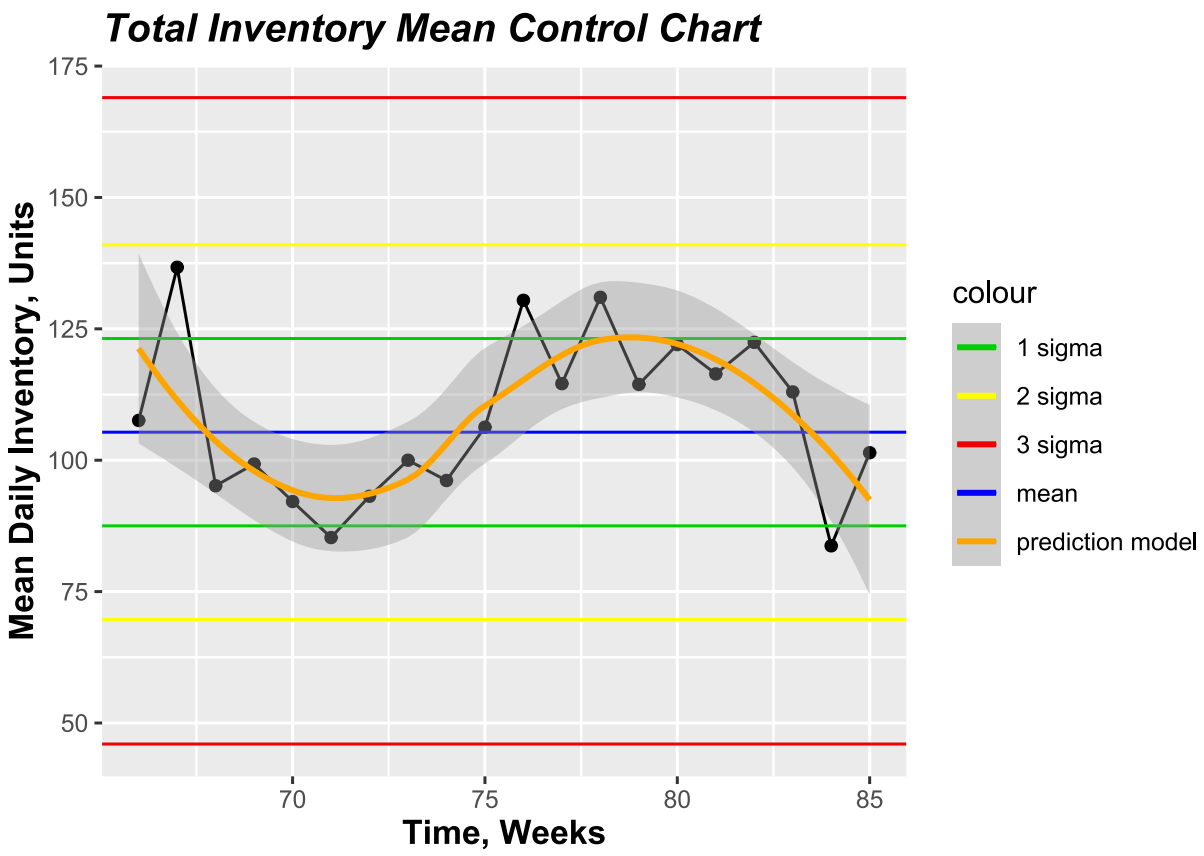

Figure 14 Total inventory mean control chart, including a spline-based GAM prediction model

Chapter 3.4.4 Evaluation of Forecasting Methods

The evaluation process for forecasting is, in this case, different than that typical of predictive modelling. Typically, models are trained and tested on different portions of the same data set to examine model fit and errors. However, in this case, prediction error is not the key performance indicator, but contribution to changepoint detection is. While demand forecasting is an important part of the blood supply chain on its own, it is not necessary in this application unless there is a significant contribution to detection metrics.

There are two categories of forecasting evaluation: base model forecasting, and demand shifted forecasting. The first category measures performance under steady state conditions, which are useful for determining how the inclusion of forecasted points effects the false positive rate of the OCPD methods. The second category of evaluation measures forecasting performance and how it relates to the improvement in the time to detection for OCPD methods.

Chapter 3.5 Magnitude Estimation

In addition to signal detection, the capability to estimate the size and direction of a detected shift in demand was built into the software package for this project. The properties of the points directly before, and directly after, a shift has been detected are compared to the established process characteristics using a difference of means test to estimate the magnitude of a shift in demand. The inventory stream where the change signal was detected is used to identify the magnitude of the change. The size of the neighbourhood used for comparison was determined by iteratively testing using replications of the simulation.

Chapter 4 Data

The data used to populate the simulation comes from CBS' Ottawa Centre covers the 2019 calendar year. To avoid intra-week variation, and autocorrelation between points, data was grouped into calendar weeks. Figure 15 below shows the autocorrelation for the daily mean inventory, while Figure *16* shows the autocorrelation for the weekly mean inventory. Note that in both figures the area between the dashed blue lines represents a level of autocorrelation where observations can be considered independent. Comparing the two, there is a significant and immediate decrease in autocorrelation when using weekly mean inventory. While the weekly inventory lag-1 autocorrelation is somewhat higher than the acceptable upper bound, the autocorrelation is modest, and the remaining data does not indicate an autocorrelation pattern.

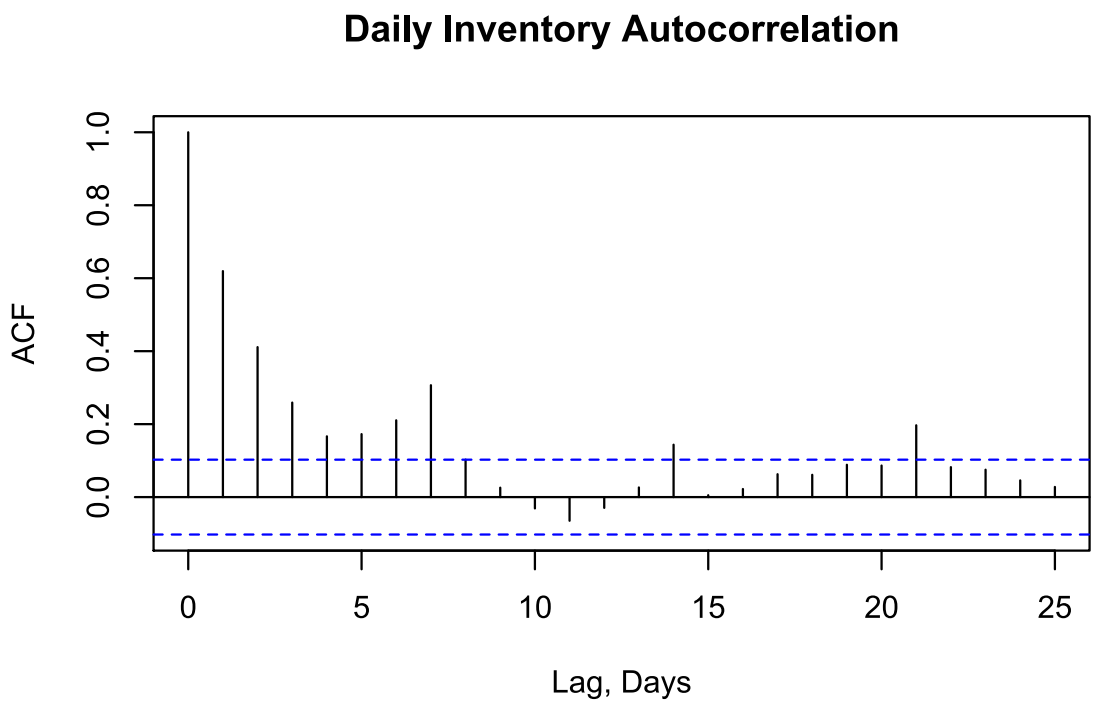

Figure 15 Autocorrelation in daily inventory

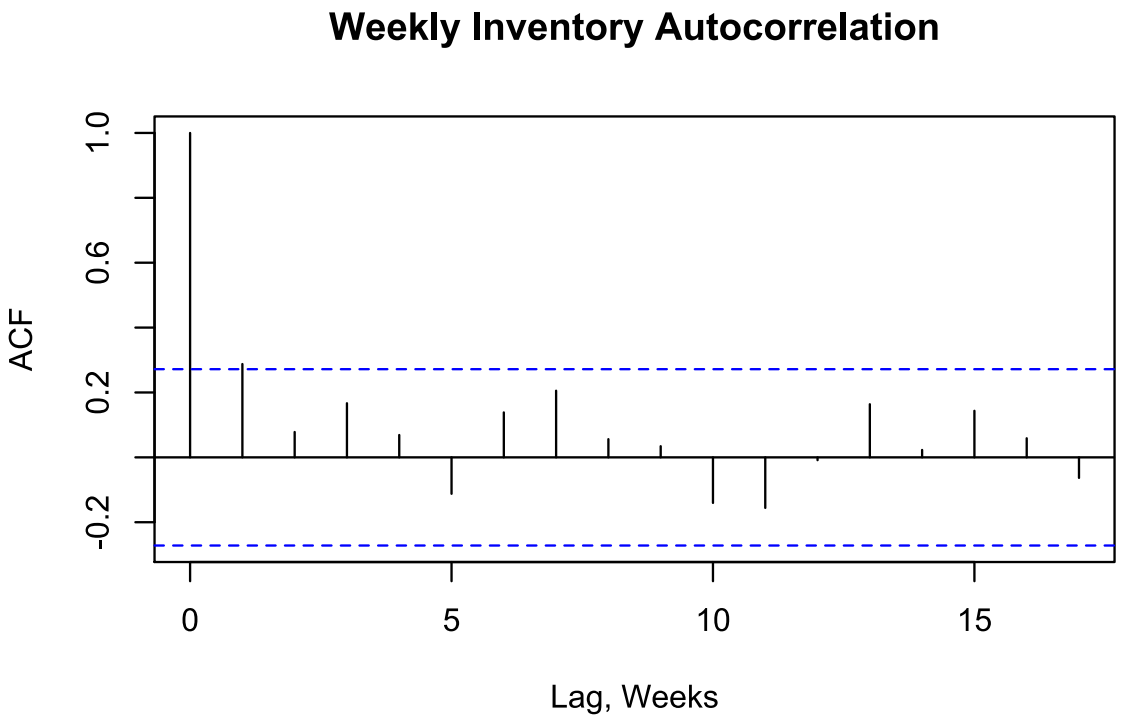

Figure 16 *Autocorrelation in weekly total inventory*

To compensate for the remaining autocorrelation in the data, the Change Point Detection methods are modified using widely accepted techniques. These include adjusting control boundaries, and the elimination of some pattern detection rules. While there is no single solution to deal with autocorrelation in OCPD problems, iterative testing with the simulation yielded an application specific method that is exploited in this project.

To determine the effect of seasonality on the data an Augmented Dickey-Fuller test was performed [33]. This test checks if a time series is stationary; if it has no trends, constant variance, and constant autocorrelation. It does this by fitting a linear model to the change in y values over a given time. The results of this test, found in Appendix D, show that when inventory data is separated into weekly samples, it can be considered stationary. As such, seasonality can be ignored in modelling.

Summary statistics for platelet inventory, by day and week, are shown below in Table 4. Some group and product type combinations are collected infrequently. For instance, pooled AB- platelets does not appear in the data set at all. While the distinction between demand and shipping makes it difficult to determine if any pooled AB- units were requested by hospitals, they are assumed to be so rare as to be not included in modelling. The 2019 NEON data set suggests that platelet inventory is ~68% pooled platelets and ~32% apheresis platelets. The relative proportion of each blood type, within each product type, approximately follows the population distribution of blood type occurrence. Information on net collections and demand for each blood and product type can be found in Appendix E. The two quantities are approximately balanced for all blood types; however, demand typically has more variance than collections.

Table 4 Daily inventory summary for the Northeastern Ontario region in 2019

Blood Type	Product Type	Mean Daily Inventory, units	Daily Inventory Stand. Dev, units	Mean Weekly Inventory, units	Weekly Inventory Stand. Dev, units
A+	Pooled	56.99	16.02	399.40	78.44
A-	Pooled	4.98	3.28	34.92	18.69
B+	Pooled	11.14	4.80	77.79	25.45
B-	Pooled	0.28	0.55	1.98	2.80
AB+	Pooled	0.64	1.37	4.40	7.87
AB-	Pooled	0.00	0.00	0.00	0.00
O+	Pooled	60.63	20.10	423.92	108.27
O-	Pooled	13.11	7.19	91.69	39.74
A+	Apheresis	13.00	5.51	91.06	25.86
A-	Apheresis	2.88	2.31	20.23	10.11
B+	Apheresis	3.64	2.57	25.50	11.69
B-	Apheresis	0.189	0.56	1.33	2.70
AB+	Apheresis	1.82	1.67	12.73	6.81
AB-	Apheresis	0.14	0.49	0.96	2.03
O+	Apheresis	15.00	6.32	105.15	28.75
O-	Apheresis	1.03	1.08	7.17	4.38

Shown in aggregate, the size difference between the two product types is obvious. Table 5 below displays the inventory for each product type as well as its coefficient of variation, the ratio of standard deviation to mean.

Table 5 Summary of platelet inventory stream properties

Property	Pooled Inventory	Apheresis Inventory
Mean	147.8	37.7
Standard Deviation	28.4	9.2
Coefficient of variation	0.19	0.24

Chapter 5 Evaluation

Chapter 5.1 Overview

Evaluation of project components proceeds as follows:

1. Validation of the base-case simulation model
2. Development of demand change experiments
3. Simulation of data representative of demand changes
4. Analysis of detection method performance in each experiment

The first three entries are described in this section, while the fourth, representing the bulk of the results, is given its own section. Validation of the base case simulation model was performed first as it underpins all other analysis. Once a valid data simulation method was in place, exploration of likely demand change scenarios was possible. Lastly, analysis of detection method performance allows validation of the final product, the detection package.

Chapter 5.2 Validation

The simulation model was verified by comparing simulation output to the parameters of the input data. While inventory was the focus, both collections and demand were also considered. The daily inventory was extracted from the input data by subtracting shipments from cumulative collections for the year. Daily inventory data was collected from the simulation by using a long-term run, a length of 10,000 days. The estimated means of these two data sets are compared in Table 6. Tested with a confidence level of 95%, the p-values indicate that there is no statistical difference between the input data and the simulation. Further details on model validation can be found in Appendix F, where it is shown that simulated collections and shipments are not statistically different than the 2019 data set.

Table 6 Simulation validation summary, the confidence interval for the difference of means between historical and simulated data is included in column 5 and 6

Inventory	Data, Daily Mean	Simulation, Daily Mean	p-value	Difference of Means 95% CI, Lower	Difference of Means 95% CI, Upper
Total	185.48	185.15	0.82	-2.59	3.24
Pooled	147.79	146.57	0.42	-1.76	4.19
Apheresis	37.70	38.59	0.07	-1.88	0.10

Chapter 5.3 Post-Validation Model Extension

To account for the change in approved shelf life of PRT platelets the released shelf life of pooled platelets was decreased from 5 to 3 days.

Chapter 5.4 Synthetic Data

There are two sets of evaluation data. The first evaluates the false positive component of the detection accuracy metric, i.e., how often a false alarm is sent. Testing on recent historical data alone does not provide a large sample, thus multiple data sets with the same statistical properties as the 2019 data set were generated. This data is referred to as Base Model data as no changes have been made to demand in this run. The second set of evaluation data assess the false negative component of the detection accuracy metric and the time to detection metric. As the form of possible demand shifts is unknown, testing of change point detection methods had to be performed on a wide variety of data to fully understand performance. Changes in demand can be described by level, type of function, and the probability they will assume a value at a given time. The latter two characteristics are described below in Table 7. Linear, or gradual, changes in demand are additionally defined by their rate of change just as stochastic changes in demand are also defined by their statistical properties.

Table 7 Descriptions of demand shift properties used in evaluation. These properties are independent of change level and inventory streams affected

F(t)	Magnitude
step	deterministic
linear	stochastic

Example changes in demand are depicted below in Figure 17. On the left a deterministic, linear, increase in demand, on the right a stochastic step increase in demand. The blue shaded region represents the variance of the process. The steady state level of the increase in demand is the same for each diagram, however the path and level of variance are different.

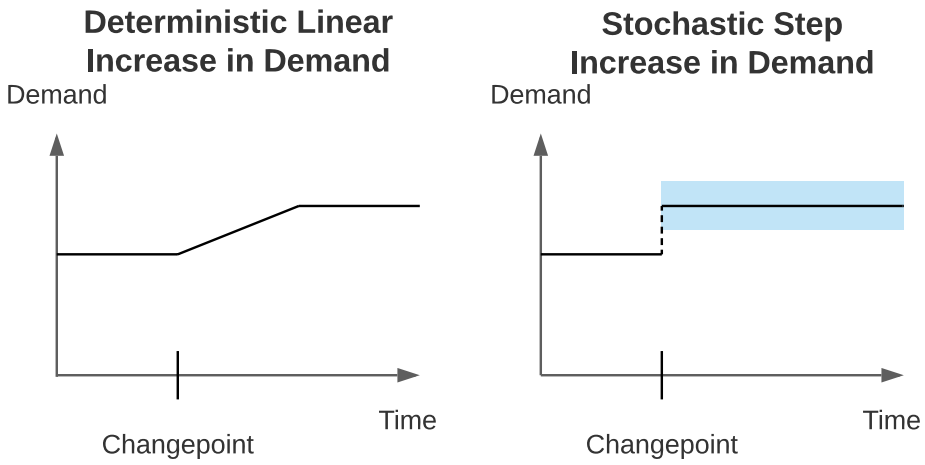

Figure 17 Examples of different possible changes in demand, when compared to a mean level

Aside from these characteristics there are several ways in which demand may be affected by the introduction of PRT platelets, and while not all of them are necessarily equally likely, access to simulation makes testing worthwhile. As a result of PRT platelets replacing BAC-T tested units there may be:

- An increase in pooled demand
- An increase in apheresis demand
- A transition of demand from pooled to apheresis

While one of these changes could occur alone, it is possible that several could occur simultaneously. For example, the effect of an increase in pooled demand may not be visible in the total inventory if there is also a transition of demand from pooled to apheresis. To account for this, data was generated to examine not only single factors but also the effect of different factors changing demand simultaneously.

Chapter 5.5 Overview of Change Experiments

Testing of demand changes can be divided into 2 categories of experiment: single factor, and multifactor. Each category evaluates detection performance under different types of conditions. The first category, single factor experiments, investigates the effect of a single type of demand change on inventory at different levels, such as an increase in pooled demand ranging from 0 to 20%. There are two subtypes of multifactor experiments: two factor, and 2^k factorial. Two factor experiments have the same structure as single factor, with an additional type of demand change present. The purpose of 2^k factorial experiments is to examine interaction among demand changes. From a computation time perspective, it is not feasible to test each combination of demand changes with each of the demand shift types and characteristics. However, we can understand each factor's effect on overall performance by testing 2 different levels of each factor in combination with each other. A full design table describing the structure of these experiments can be found in Appendix J. The last two factors examined in this experiment are the change properties described in Table 7. While these could be considered different levels of the first three factors, we elected to treat them as separate factors for ease of computations.

Chapter 6 Results

The results are separated according to their experimental categories: Base model, single factor experiments, and multifactor experiments. In each section, the performance of the detection method is evaluated, and the effect of forecasting discussed.

Chapter 6.1 Base Model

Replications of the base model were analyzed to determine the false positive rate of detection. A positive signal was detected in 3.20% of replications. Current modelling accuracy makes false alarms inevitable. However, their relative infrequency shows the reliability of these change point detection methods.

Chapter 6.1.1 Base Model Forecasting

Each of the forecasting methods was tested on the base model data, as an additional component of the changepoint detection scheme. The results of testing are compared below in Table 8.

Table 8 Performance of different forecasting methods on base case model data

Method	Base Case MAPE, %	RMSE, units of platelets	Base Case Detection Rate, %
No Forecasting	--	--	3.20
ARIMA	13.14	16.85	3.50
Linear Regression	15.45	20.00	6.50
GAM – Local Regression	16.64	21.90	31.50
GAM - Splines	14.35	18.57	4.90

Two forecasting accuracy metrics are included in the analysis, along with the detection rate. The first, Mean Absolute Percentage Error, compares the absolute difference between predicted and actual value as a percentage of the actual value. The second metric, Root Mean Squared Error, performs a similar function, but is measured in units of platelets instead. Additional results are included in Appendix G describing the time to detection of false positive signals using different forecasting techniques.

As can be seen in Table 8 the ARIMA model was found to perform best according to the forecasting accuracy metrics. In the base model, change detection signals are false positives and should be minimized. The other three forecasting methods increase the detection rate above the base amount, but not necessarily in direct proportion to forecasting accuracy. The small difference in forecasting accuracy between the Linear Regression model and GAM Local Regression model translates to a substantial increase in the false alarm rate. Conversely, the spline-based GAM model, with only slightly better forecasting accuracy than Linear Regression, has a similar false positive rate. Forecasting accuracy metrics are insufficient on their own to determine performance, thus differences in detection rates will be used to evaluate the contributions of each method.

The residual error plots for each forecast method give insight into the performance characteristics discussed above. As can be seen in Figure 18 the residual errors for the linear model are highly dispersed. This shows the inaccuracy of the model, as well as its tendency to under-estimate high values and over-estimate low values. The same general pattern can be observed in Figure 19. However, the residual values are more tightly clustered around the process mean, indicating it is more accurate than the linear model above. The residual errors for the two Generalized Additive Models, in Figure 20 and Figure 21 have similar distributions. Both of the GAM residual error plots are skewed to the lower right-hand corner, indicating that they tend to under-estimate large values.

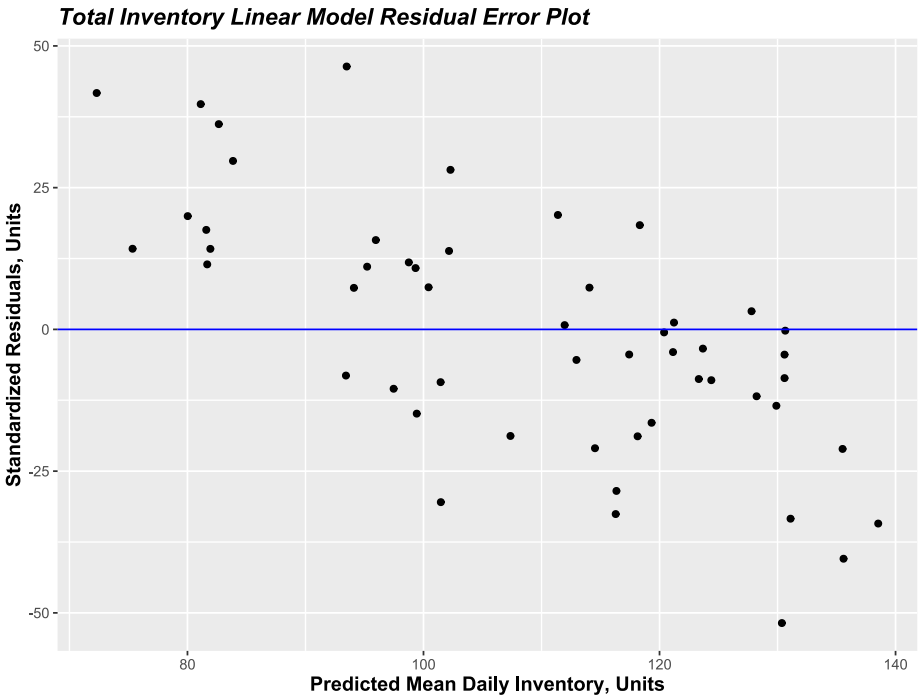

Figure 18 Example residual error plot for a linear model applied to base case total inventory data. The residual errors are highly dispersed for low and high predictions

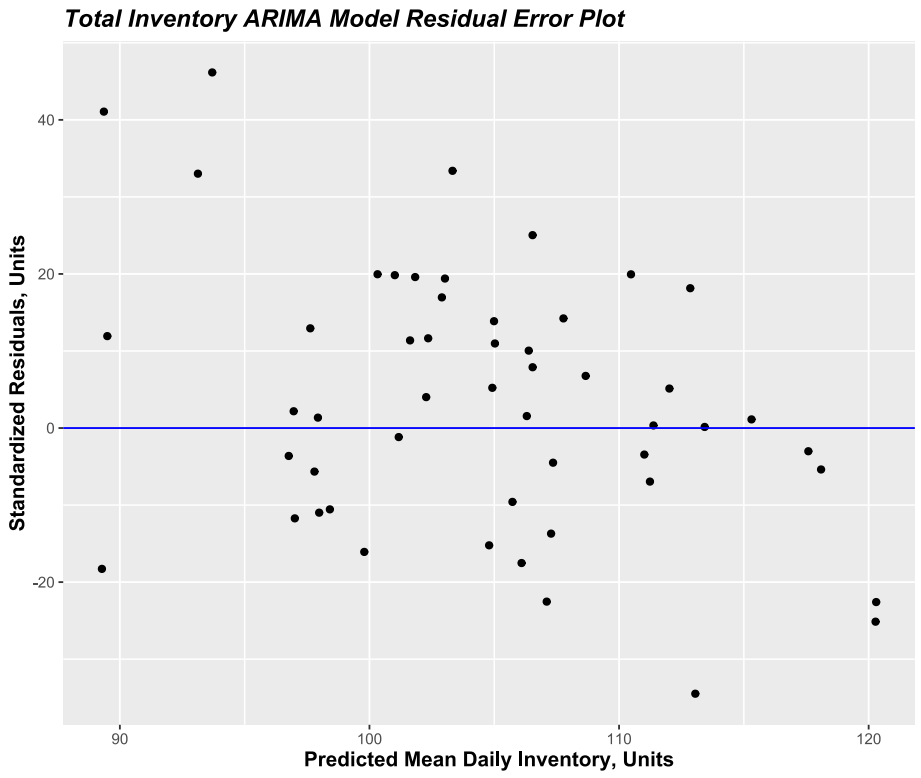

Figure 19 Example residual error plot for an ARIMA model applied to base case total inventory data. The residuals are, relatively, tightly clustered, and evenly distributed

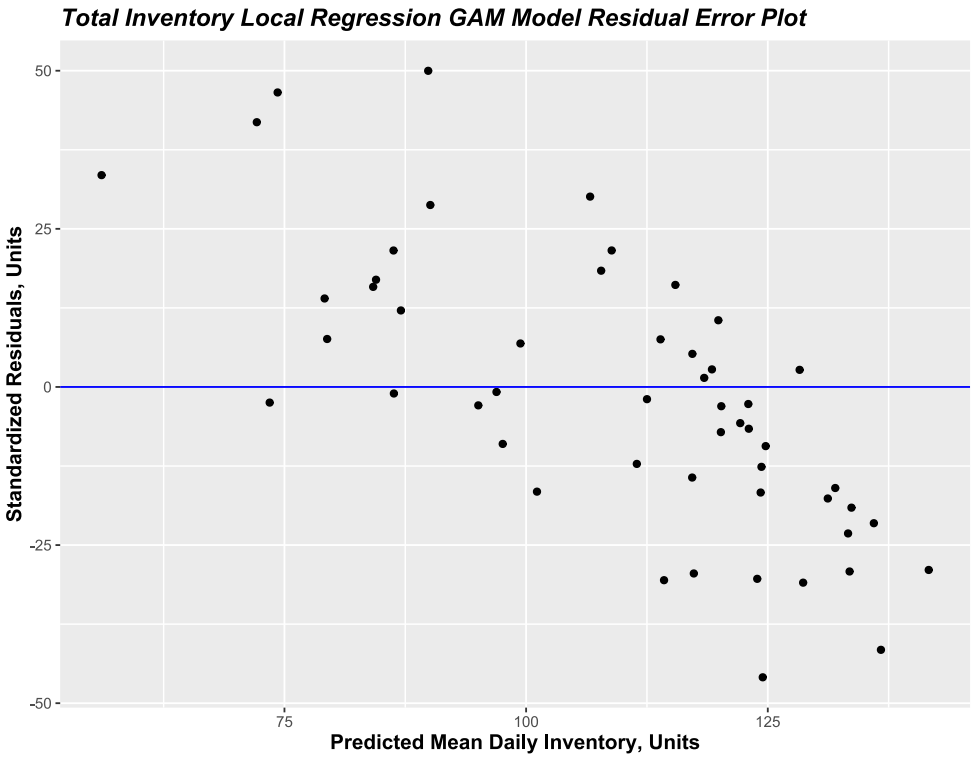

Figure 20 Example residual error plot for a Local Regression GAM Model applied to base case total inventory data. The residuals are skewed to the lower right-hand corner

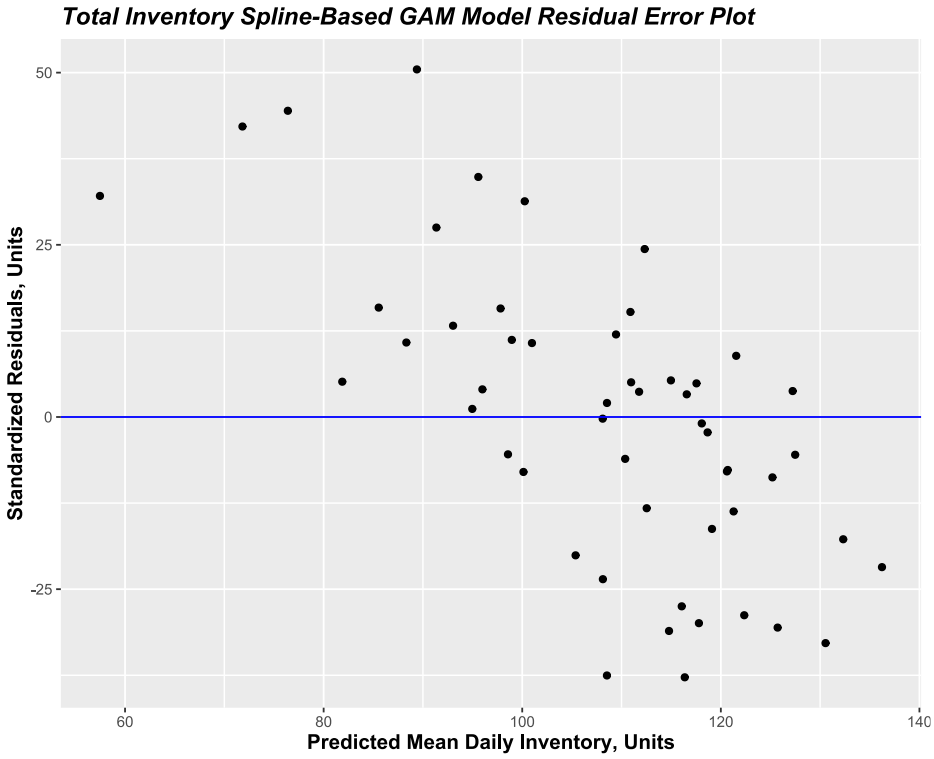

Figure 21 Example residual error plot for a spline-based GAM Model applied to base case total inventory data. The residuals are skewed to the lower right-hand corner of the plot

Chapter 6.2 Single Factor Evaluation Experiments

To evaluate the performance of detection methods, and forecasting, experiments were conducted on data with only a single factor change. The level of each factor was changed while keeping all else the same as the base model. Example results are shown below in Figure 22 and Figure 23.

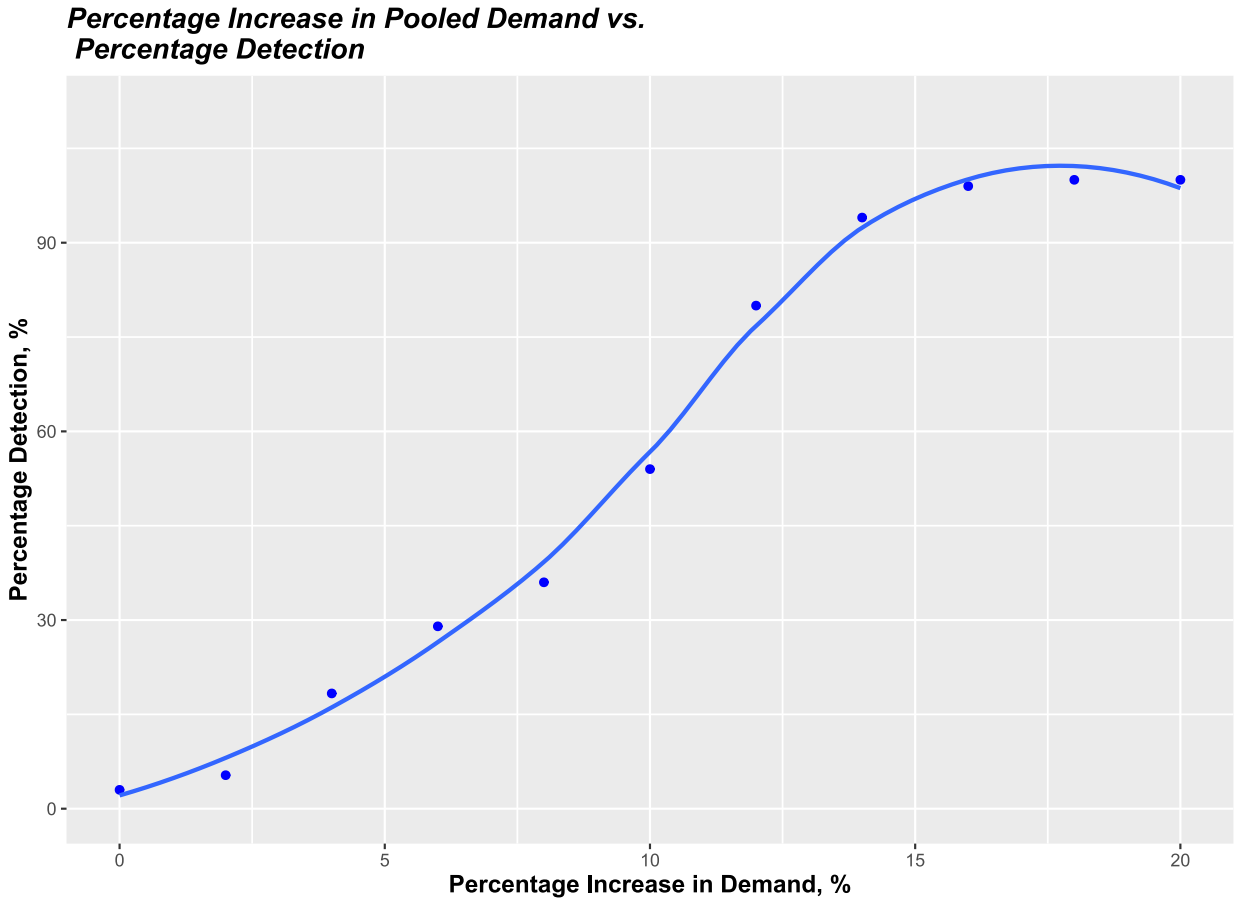

Figure 22 Percentage increase in pooled demand vs. percentage detection

Figure 22 depicts the percentage of detections when the level of increase in pooled demand is changed. These results indicate that the chance of signal detection increases rapidly, until it converges to 100% detection. The data from all single factor experiments exhibit similar trends converging to 100% detection at a level above 10% total effect on system. It should be noted that, in this case, a 10% increase in pooled demand represents a 6.8% increase in total demand.

The mean time to detection for the same single factor experiment is displayed in Figure 23. The mean time to detection decreases as the increase in pooled demand rises. The mean time to detection for the 2% level is 33.5 weeks while the mean time to detection for the 20% level is 7.33 weeks. The width of the confidence interval for each value, graphically the height of the error bars, decreases as the factor level increases. Similar trends can be found in all the single factor experiments, described in the evaluation section, and are included in Appendix H.

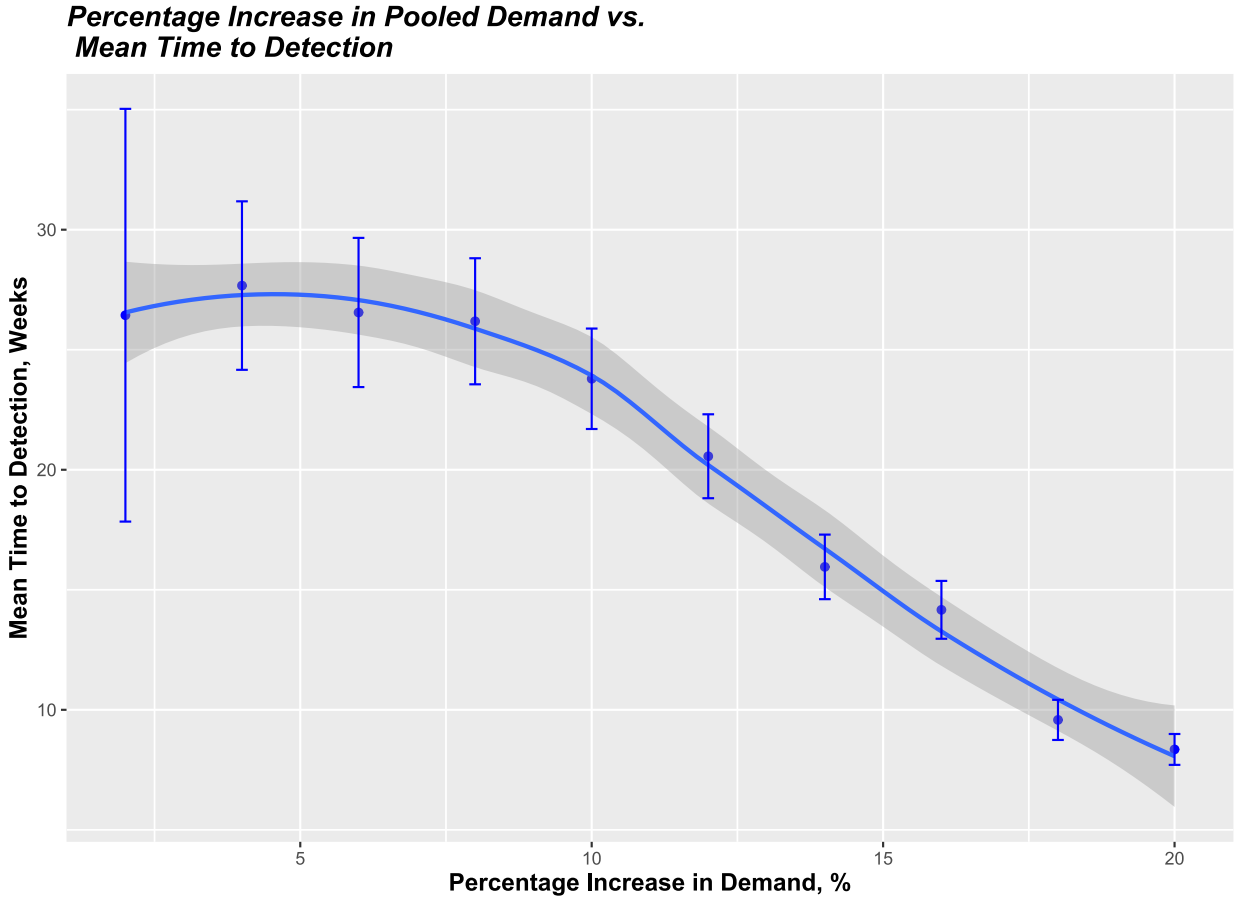

Figure 23 Percentage increase in pooled demand vs. mean time to detection. The shaded area indicates a 95% confidence interval for the trendline

The results for each of the single factor experiments are shown below. Figure 24 displays the detection rates and Table 9 through Table 11, the mean time to detection. The results show that changes to pooled demand (blue) are detected at smaller magnitudes more than other changes; detection rates converge to 100 percent asymptotically as pooled demand increases. On the other hand, an increase in apheresis demand (red) has a much lower detection rate. Finally, it may be observed that the movement of pooled demand to apheresis (green) converges to full detection more quickly than the increase in pooled demand experiments.

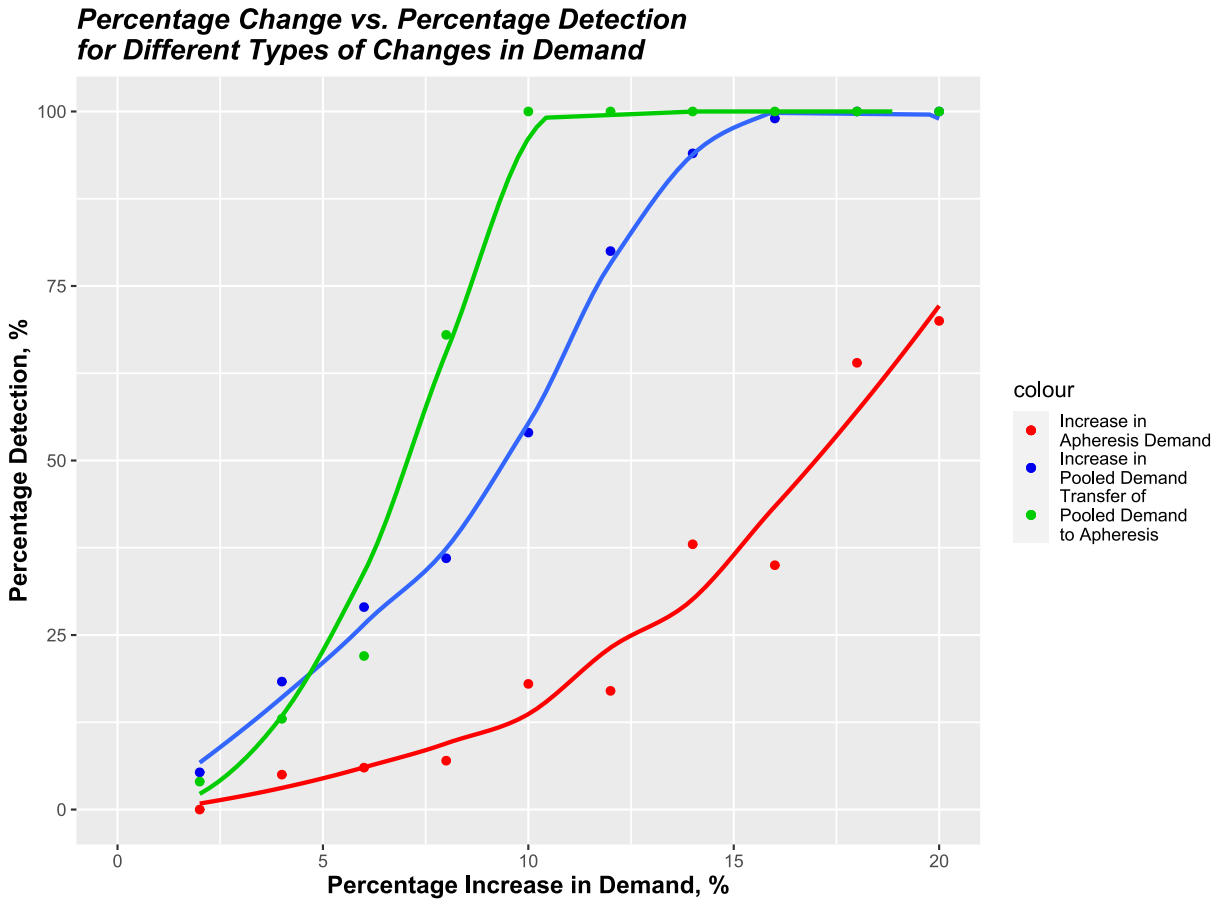

Figure 24 Percentage Change vs. Percentage Detection for different types of changes in demand. The rate of detection at each level varies by experiment

This behaviour is likely caused by two factors. The first being that changes in pooled inventory are twice as large when applied to apheresis, a function of their different inventory sizes. Second, the high variance of the apheresis inventory level prohibits the use of alternate change point detection conditions, lowering the sensitivity of apheresis monitoring to small changes. Together these conditions result in a lower detection rate for small changes and a higher detection rate for larger changes, when compared to the other experiments.

As can be seen in the results of Table 9 through Table 11, the confidence interval for all single factor experiments diminishes in size as the magnitude of the change increases. This confirms that larger changes are not only detected more quickly than small changes, but also more consistently quickly.

Table 9 Mean time to detection, increase in pooled demand single factor experiment

Increase in Pooled Demand, %	Mean Time to Detection, Weeks	95% Confidence Interval Halfwidth
2	33.50	11.14
4	19.00	6.83
6	21.18	5.78
8	24.49	4.71
10	25.04	4.31
12	20.94	2.88
14	15.86	2.45
16	10.73	1.61
18	8.99	1.48
20	7.33	1.26

Table 10 Mean time to detection, increase in apheresis demand single factor experiment

Increase in Apheresis Demand, %	Mean Time to Detection, Weeks	95% Confidence Interval Halfwidth
2	--	--
4	21.14	14.72
6	24.67	16.10
8	28.20	19.34
10	28.50	7.88
12	24.71	7.37
14	27.42	5.05
16	26.51	5.32
18	23.02	3.71
20	23.21	3.15

It may be noted that detecting small changes in apheresis demand is quite difficult, resulting in few values for detection time. There are no detections for a 2% increase in apheresis demand, hence no confidence interval is available. Despite this, the same trend in CI half width as the other experiments is present above that level.

Table 11 Mean time to detection, movement of pooled demand to apheresis demand single factor experiment

Movement of Pooled Demand to Apheresis, %	Mean Time to Detection, Weeks	95% Confidence Interval Halfwidth
2	22.00	17.47
4	25.33	7.75
6	22.47	4.79
8	20.75	3.42
10	12.43	1.63
12	8.78	1.10
14	8.72	1.29
16	6.60	0.78
18	6.09	0.57
20	3.20	0.17

Chapter 6.2.1 Single Factor Forecasting

Each of the forecasting methods discussed was tested on the single factor change experimental data. Ultimately, the GAM model fit using splines was found to produce a significant improvement in mean time to detection, while only raising the false alarm rate slightly. An example is shown below in Figure 25 where the standard changepoint detection method, blue, is compared to changepoint detection enhanced with spline-based forecasting, in red.

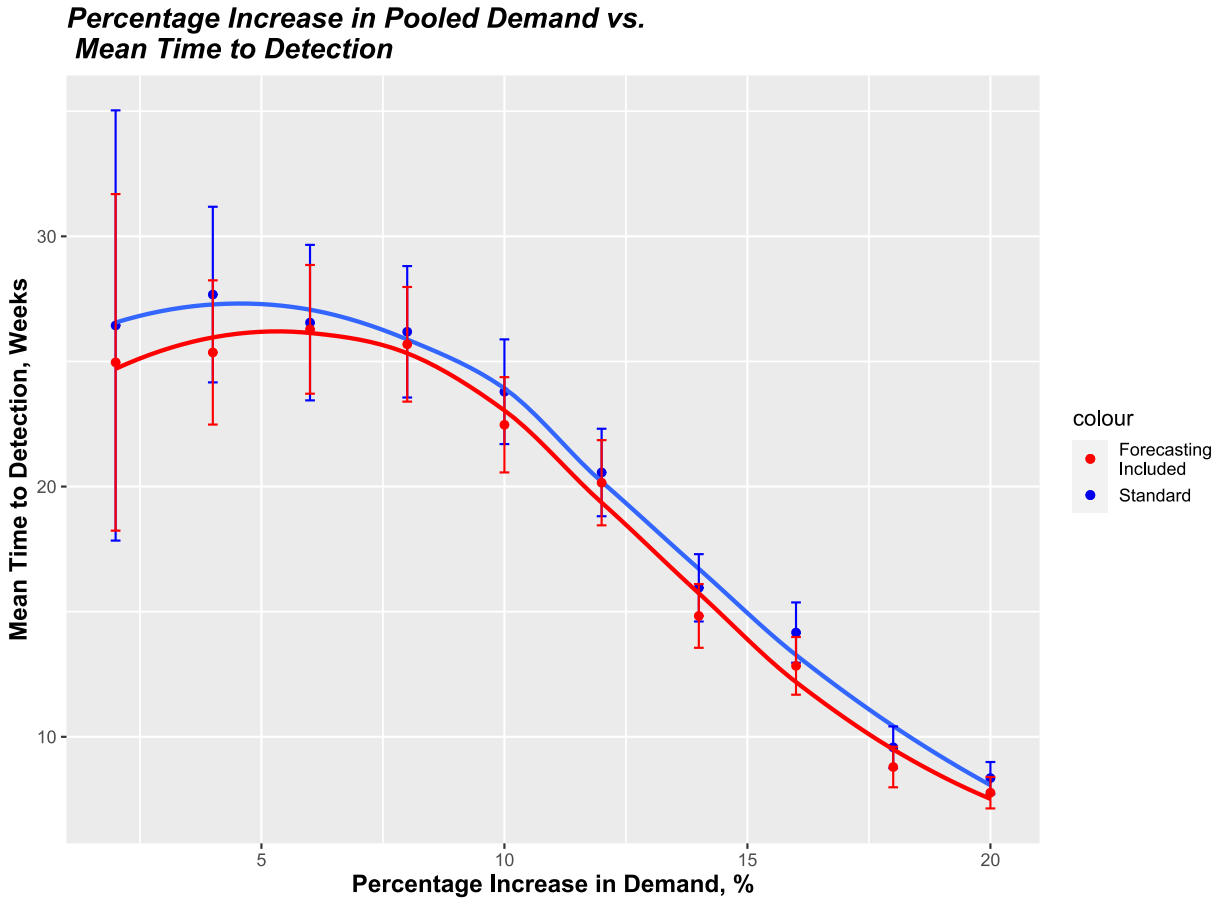

Figure 25 A comparison of the mean time to detection for statistical process control with and without forecasting

There is a significant increase in performance in the midsection of the graph where the mean time to detection is noticeably lower when forecasting is included. This result is confirmed by t-test in Appendix H, showing that there is a difference between the two sets for trials in the middle of the graph, but not on either end. This performance curve makes sense on the low end as small changes have a low probability of detection making forecasting less effective. In contrast, large changes are detected almost instantaneously, thus decreasing the value of forecasting. Overall, detection performance is consistently better when forecasting is included, even if only by a small amount. While forecasting increases the false positive rate somewhat, as seen in Table 8, it does decrease detection time. If the increase in the false positive rate can be tolerated, then forecasting improves the changepoint detection method in this case.

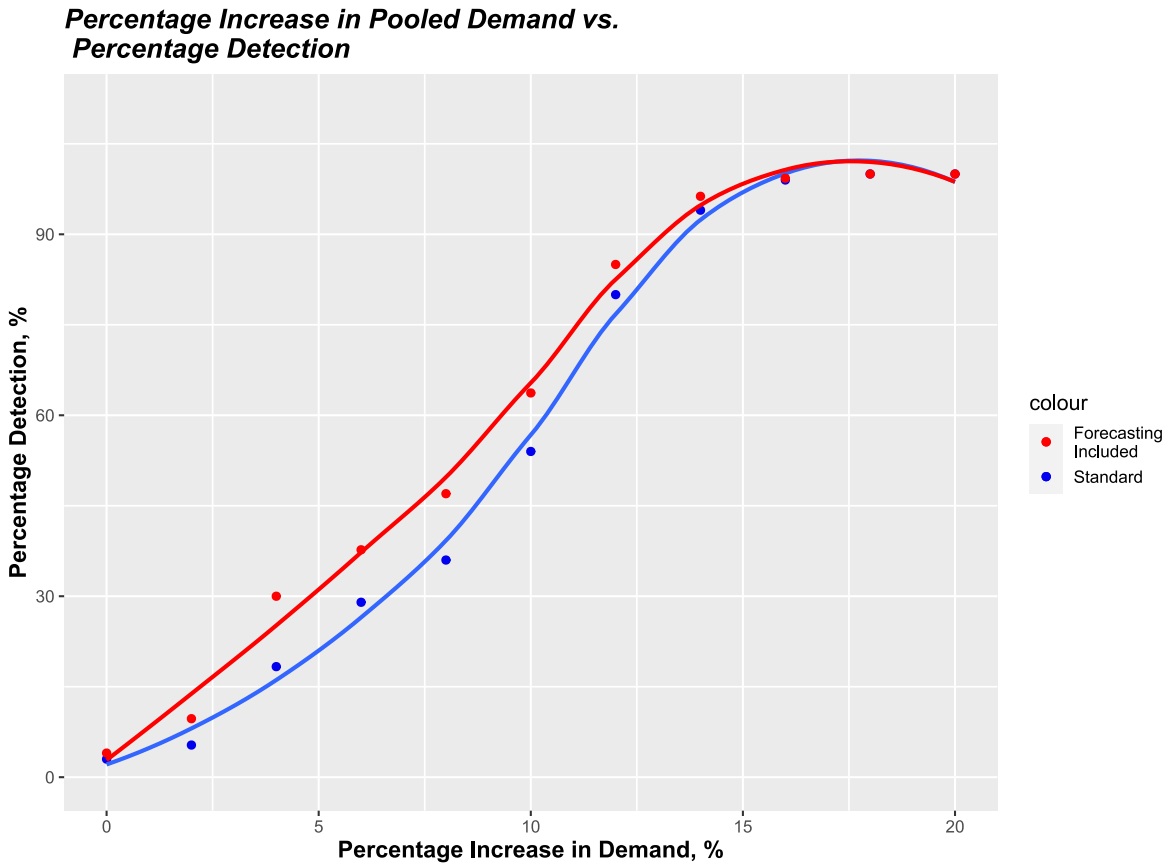

Figure 26 A comparison of the percentage detection of an increase in pooled demand for statistical process control with and without forecasting

The purpose of using forecasting is to improve the mean time to detection. While the detection rate for forecasted values in Figure 26 appears to be larger than that of the standard values, it is not statistically significant. These results are included in Appendix H for all the forecasting methods tested. It should be noted that the ability to improve detection time varies greatly depending on the method used.

Additional comparisons of all the forecasting methods, and the base detection scheme, are included below. Table 12 shows the performance change, as compared to the base model, for each of the different forecasting methods, when there is an increase in pooled demand. This is the same scenario shown graphically for a spline-based GAM in Figure 25 and Figure 26. The performance change uses the mean time to detection, measured in weeks relative to the base case. Negative numbers indicate that the Mean Time to Detection has increased. While the effect of including forecasting on detection rate is generally not statistically significant, there may still be some variation in the number of detections. For small changes in demand where the detection rate is low, each individual detection has a larger effect on detection time than for large changes where the detection rate is high. The addition of forecasting may then lead to a marginal increase in Mean Time to Detection. This does not change the significance of the results for each method; however, it does explain why the change in mean time to detection is negative for some of the experiments found in Table 12. Full results, and significance testing, for all the forecasting methods evaluated can be found in Appendix H.

The GAM model fit using Local Regression, column 4, increases performance the most. However, this effect is dampened somewhat by it's large increase in false positive rate. Between the remaining methods the GAM model fit using splines, in Column 5, clearly performs the best.

Table 12 Decrease in Mean Time to Detection for each forecasting method, when compared with the base model. Experimental results for the increase in pooled demand series of experiments

Increase in Pooled Demand, %	Linear Regression ΔMTTD, weeks	ARIMA ΔMTTD, weeks	GAM Local Regression ΔMTTD, weeks	GAM Splines ΔMTTD, weeks
2	0	0.02	8.57	9
4	-1.64	0.05	-2.04	-2.39
6	-1.5	0.07	1.95	0.23
8	0.64	0.11	2.89	0.78
10	0.02	0.09	6.54	2.64
12	0.4	0.21	6.51	2.34
14	0.31	0.07	5.75	1.46
16	0.28	0.11	2.48	0.76
18	0.62	0.04	2.62	1.26
20	0.15	0.04	1.57	0.36

Table 13 shows Mean Absolute Percentage Error for the same set of experiments. The results show that the ranking of base model forecasting accuracy does not necessarily translate to accuracy of the data representing a change in demand. The ARIMA model, which had the lowest error on base model data, performs worse than any other model on shifted demand data. The method with the best overall performance including false positives, a spline-based GAM fit, also has the lowest mean error of the methods tested. The high error associated with large changes is a product of reduced detection time. The smaller mean time to detection results in fewer points to forecast, weighting each individual forecasted point more heavily.

Table 13 Mean Absolute Percentage Error for different forecasting methods, applied to the increase in pooled demand experiments

Increase in Pooled Demand, %	Linear Regression MAPE, %	ARIMA MAPE, %	GAM Local Regression MAPE, %	GAM Splines MAPE, %
2	20.65	18.11	17.49	18.00
4	22.44	23.58	22.98	19.65
6	26.33	27.34	24.21	22.26
8	28.39	31.25	25.77	22.90
10	32.15	35.68	30.54	25.72
12	35.62	44.21	33.68	31.31
14	40.89	57.61	38.55	34.13
16	43.21	59.38	41.11	35.28
18	49.86	61.22	44.63	37.66
20	53.28	56.34	46.44	39.04

There are several methods of increasing the sensitivity of changepoint detection methods utilizing control charts; the results of this study indicate that forecasting may be an effective way to increase sensitivity, and detect changes more quickly, with only a small increase in the false alarm rate.

Chapter 6.3 Multifactor Experiments

Chapter 6.3.1 Increase in Pooled Demand and Transfer to Apheresis

To ascertain the effect of different demand shift parameters on performance, and the interaction of factors, changepoint detection was applied to data with multiple demand factors.

It is possible that the implementation of PRT platelets will change ordering patterns in more than one way. There may be an increase in the demand for pathogen-reduced pooled platelets, to make up for the lower per unit yield, accompanied by a migration of demand to apheresis platelets. Thus, a set of experiments was performed to examine the performance of detection methods to demand changes exhibiting these two patterns, with varying levels of each.

Testing several forecasting techniques on multifactor data is computationally intense, requiring significantly more resources than a single method. The results presented in this section are for Statistical Process Control enhanced with a spline-based GAM prediction model. Additional results comparing the effectiveness of methods including forecasting with standard methods are included in Appendix I and Appendix J. Here it is shown that forecasting significantly decreases time to detection but does not influence detection rate.

The detection rates for these experiments are found below in Figure 27. The contours indicate that the detection rate converges to 100% for demand shifts between 10% and 15%. It can also be observed that the effect of a migration from pooled to apheresis is detected for lower levels more often than the effect of an increase in demand for pooled platelets.

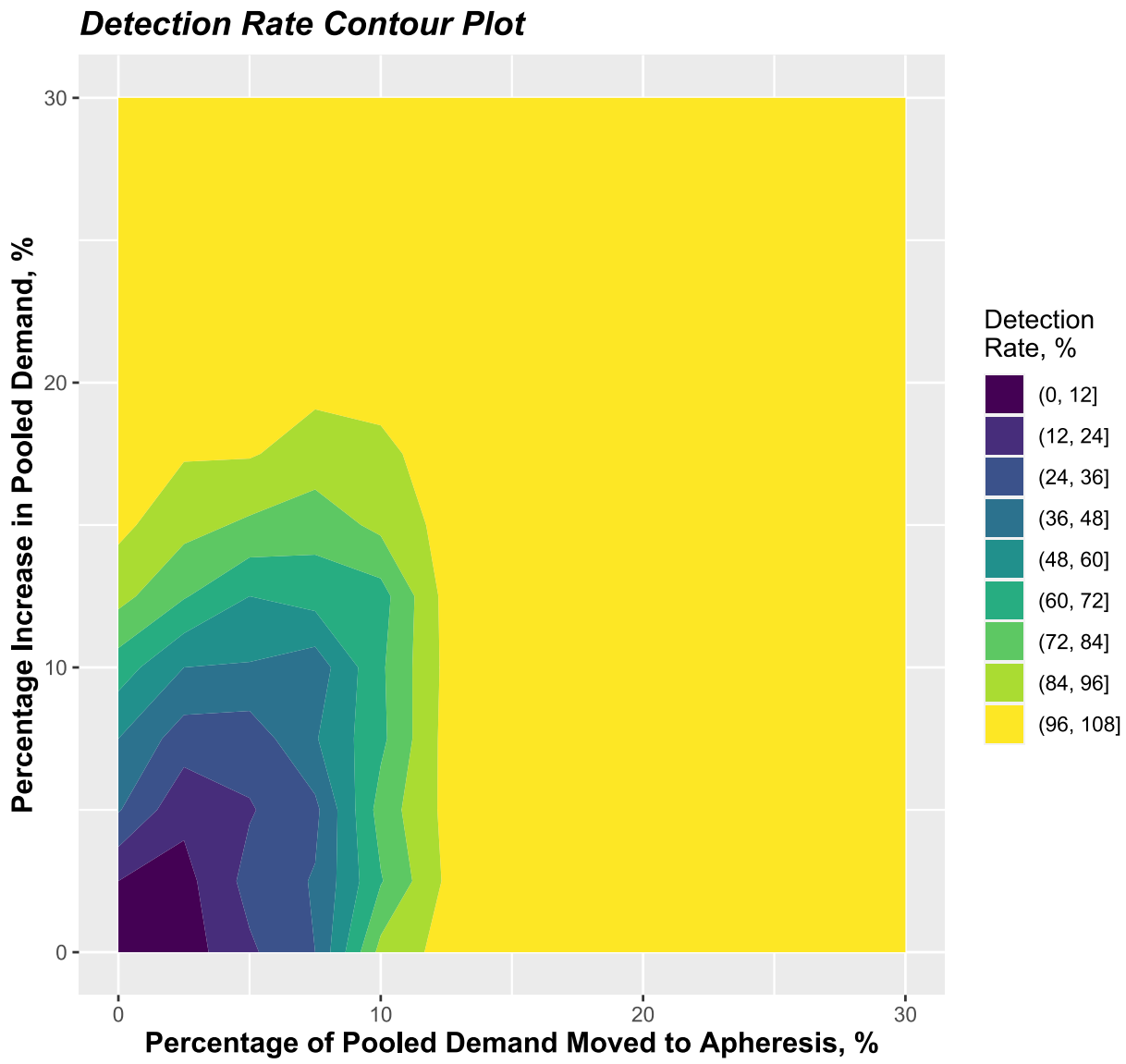

Figure 27 Contour plot of detection rate for different levels of an increase in pooled demand and pooled demand moved to apheresis. All experiments converge to full detection (yellow), however not evenly

The mean time to detection for the multiple demand shift experiment is displayed in Figure 28. While there are some local anomalies, the mean time to detection appears to decrease, with a diminishing gradient, as the magnitude of the demand shift increases. These results appear similar, but generally less favourable than, the single factor experiments.

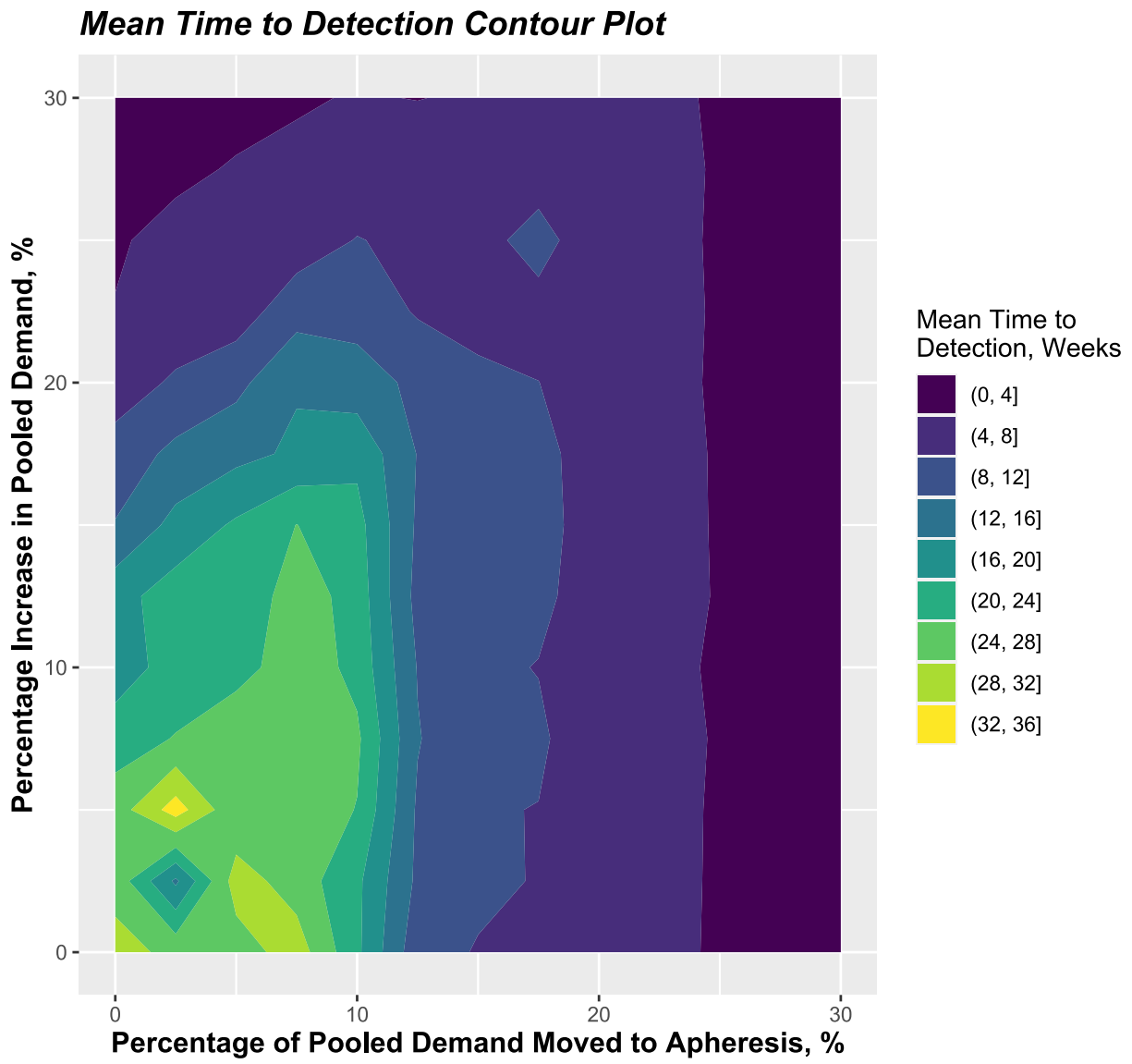

Figure 28 Contour plot of detection time for different levels of an increase in pooled demand, and pooled demand moved to apheresis. Lighter colours indicate longer detection times

Chapter 6.3.2 2^k Factorial Experiments

A set of 2^5 factorial experiments were performed to examine the interaction of demand change characteristics, as described in

Table 14. The experiments use two different levels, or types, of each factor to determine factor effect as well as interactions with other factors. A full design table describing each individual experiment is included in Appendix J.

Table 14 Descriptions of factors used in factorial level experiments. Factors 1-3 are demand shift types while factors 4 & 5 are properties

Factor				
1	2	3	4	5
An increase in pooled demand	An increase in apheresis demand	Transfer of pooled demand to apheresis	Step increase or linear increase	Deterministic or stochastic

The results of these experiments are shown graphically in Figure 29 and Figure 30. These results are organized using the two demand change properties, step/linear and deterministic/stochastic. Dividing the results using these two factors means there are 2^3 experiments in each group. Each experiment within a group has a counterpart in the other groups, the only difference being the combination of factors 4 and 5.

As can be seen in Figure 29 the experiments have a 100% detection rate, except for experiment 1 and 2 in each group. These experiments refer to the scenarios where factors 1 through 3 are all at the lower level and where only factor 1 is at the higher level, respectively. The red lines represent the mean detection rate for each group. Graphically, there is a difference in detection rates between the deterministic and stochastic changes, a result shown to be significant in the results of Appendix J. These results do not show that there is a significant difference between the detection rates in step and gradual demand change experiments.

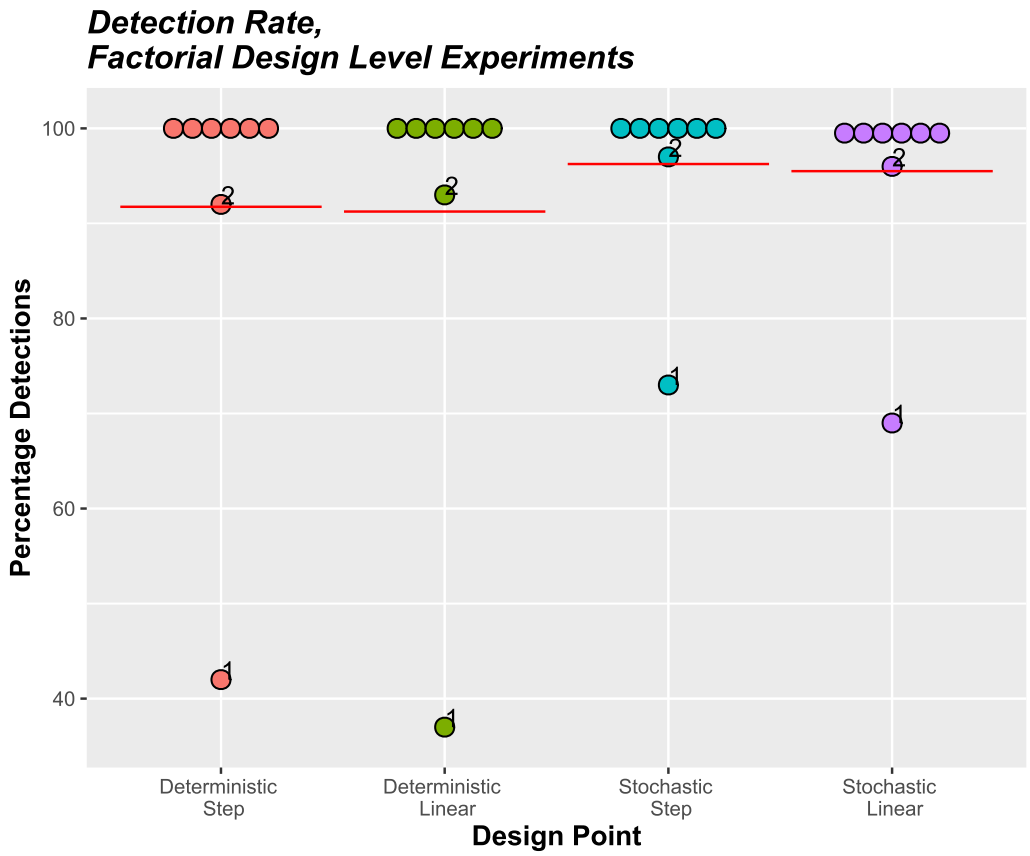

Figure 29 Detection rate for factorial level experiments grouped using factors 4 & 5. Changes in demand are all detected, except for experiments 1 and 2 in each group. The red lines indicate the mean detection rate for each group

Figure 30 shows the mean time to detection for each experiment in each group. These results are much like the results for detection rate, indicating that stochastic changes in demand have a significantly lower time to detection than deterministic changes. The difference in detection time between step and linear changes in demand is insignificant. Like detection rate, confirmation of these results can be found in Appendix J.

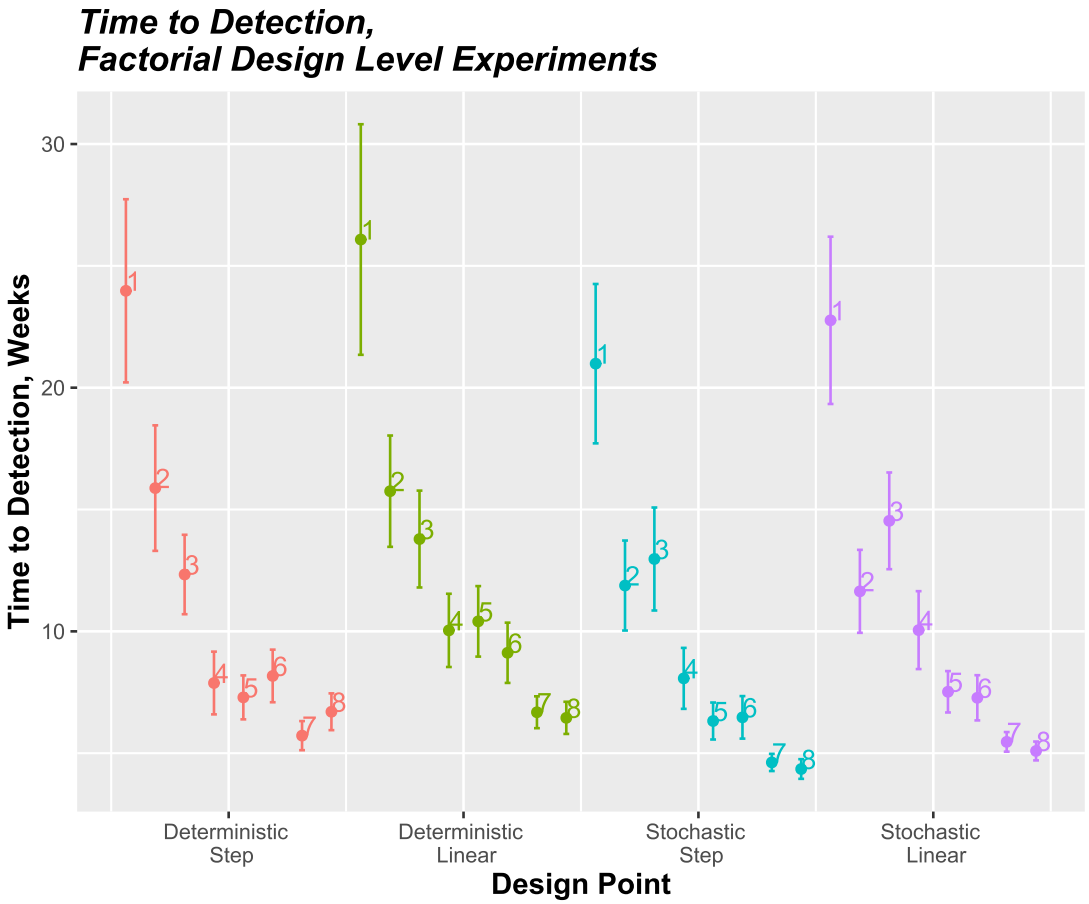

Figure 30 Mean time to detection for factorial level experiments grouped using factors 4 & 5. Each group shows the mean detection rate of each experiment, labelled by number, and repeated with different combinations of factors 4 and 5

The effect of a change in each factor on the results of the 2^5 factorial experiments are shown in Table 15. These indicate that an increase in all of factors, except a gradual change in demand (Factor 4), increase the detection rate and decrease time to detection. Factor 3, a shift of pooled demand to apheresis, has a significantly larger effect on detection time than the other factors. This result highlights the benefits of tracking inventory streams individually, as pooled platelets make up 68% of inventory and apheresis the remaining 32%. This means that a small shift of demand from pooled has a much larger effect on apheresis inventory. Tracking apheresis inventory separately makes these changes much easier to detect.

Table 15 Effect of each factor, described in Table 15, on detection rate and detection time

Factor	Effect on Detection Rate, %	Effect on Detection Time, Weeks
1	33.00	-7.69
2	38.00	-9.85
3	38.75	-14.11
4	-2.25	1.89
5	4.375	-1.64

The second level effects, how two factors can affect each other, show that more than one type of increase present at a time improves performance metrics, although the increase is bounded. For instance, detection rate cannot exceed full detection, and detection time can not be zero. Not all the second level effects are significant, for example the impact of a gradual/immediate change has a negligible effect on all types of demand change. The effects above 2nd level are not significant, meaning that if more than two factors are present at once, the result is adequately modelled by the 1st and 2nd level interactions.

Chapter 6.4 Magnitude Estimation

Estimating the magnitude of a change in demand proved to be more difficult than expected. Consider the case where there is an increase in pooled demand. The change is detected, but the response in collections is not modelled by the simulation. As inventory is a queueing network, the imbalance of collections and demand causes the inventory level to rapidly decrease, bottoming out near zero. This makes a before/after comparison difficult, as the change in inventory level is likely not representative of the change in demand.

Inventory data immediately leading up to detection can be used instead, as in Table 16, however this approach is also error prone. As can be seen in the results below the confidence interval halfwidth is larger than the mean change in all experiments, meaning that demand changes can be mistaken for a much larger change, or no change at all.

Table 16 Estimates for demand change size based on inventory for increase in pooled demand experiments

Increase in Pooled Demand, %	Mean Change Estimate, units	Estimate 95% CI Halfwidth, units
2	9.75	23.83
4	14.43	22.69
6	16.59	22.20
8	17.67	22.08
10	19.79	21.05
12	20.87	22.21
14	21.68	22.22
16	21.28	21.71
18	21.05	21.35
20	20.09	21.97

The results for other experiments are included in Appendix K, although the performance is similarly poor to the results in Table 16. If data were available directly from hospitals, estimating changes in demand would no longer be affected by the queueing behaviour of the inventory.

Chapter 7 Discussion

Chapter 7.1 Discussion of Results

The results show, as expected, that the greater the magnitude of a change is, the more quickly and consistently it is detected. While these results are encouraging from the perspective of mitigating a shortage, they are of limited use for small changes. The low detection rate for small changes means that the sample size for determining mean time to detection is smaller than the sample size for large changes. The actual distribution of detection time for small changes may differ significantly from the observed samples.

The two factor demand change experiments show that detection performance is a result of the combination of demand shifts. This is demonstrated in both section 7.3.1 and 7.3.2, the special case experiment of where there is an increase in pooled demand as well as a transfer to apheresis demand, and the 2^k factorial experiments.

Interestingly, when there is both an increase in pooled demand and a movement of pooled demand to apheresis the two changes do not cancel out and lead to worse performance. The multiplicative effect of pooled demand moving to apheresis helps bolster change detection performance and highlights the benefits of monitoring apheresis separately. The 2^k factorial experiments confirm this result showing that the movement of demand from pooled to apheresis is the most important factor in determining time to detection performance. These results also show that stochastic changes in demand are detected more consistently and quickly than deterministic changes.

The numerical values given to each demand factor, is likely less important than its relative size compared to the other factors. While a 5% movement of pooled to apheresis, is detected more quickly than a 5% increase in pooled demand, the magnitude of a demand change will not be known beforehand in the real system. The patterns in these results are useful in understanding how different changes in demand may affect detection but are not necessarily representative of how likely each change of a given magnitude is to occur.

Chapter 7.1.1 Forecasting

Including forecasted values in the set of points analyzed for change can improve the detection time. There is a trade-off with increased false positive rate, but if this can be tolerated, forecasting improves performance. This trade-off between detection sensitivity and false alarm rate is often present when adding conditions to control chart monitoring. In this case, using a change detection condition that includes forecasted values, but otherwise operates the same as an existing condition, is advantageous because it decreases detection time without directly analyzing a larger neighbourhood. Most additional conditions used to sensitize control charts consider many points in sequence., However, using these conditions substantially increases the false alarm rate in autocorrelated data, such as demand data for platelets.

Chapter 7.1.2 Validity of Results

Simulation was employed to test detection methods before use in daily operation. Assessing the validity of results obtained using simulation is difficult. On one hand, detection is performed on data from an entire health region, making the accuracy of small details in the model potentially less important. Evaluating two relatively simple metrics on aggregate data may, however, obscure modelling inadequacies caused by missing information. On the other hand, there is currently no real data available to compare to the results found using simulation. While the implementation of PRT in Ottawa will provide valuable feedback, it will not provide overwhelming evidence on the validity of results obtained using simulation. There is no evidence to show that the detection results with real data will be significantly different than those obtained by simulation. Simulation has enabled substantially more evaluation and tuning of detection methods than would otherwise be possible. Ultimately, the simulation and the detection algorithm are intended as decision support tools; a part of overall preparedness for the introduction of PRT platelets.

Chapter 7.2 Discussion of Model Limitations

Chapter 7.2.1 Inventory Data and Imprecise Substitution

The substitution of pooled and apheresis units for other compatible units is difficult to model precisely. There are currently no records of how frequently substitution occurs, or the types of units involved. Substitution is not necessarily optimal and may not even follow a set of consistent rules. Thus, any model of substitution can only be an approximation.

Chapter 7.2.2 Model Detail

The model was designed to detect changes in aggregate demand within a region. It does not account for individual hospitals being affected differently by product changes. Current modelling ignores the time of day at which collections and orders arrive and depart. Also, while detections happen on a weekly basis, there is likely still some loss of model fidelity because of this assumption.

Chapter 7.3 Future Research

Chapter 7.3.1 Comparison of Demand and Inventory Monitoring

The accuracy, and sensitivity, of change detection methods are likely limited by using inventory data instead of demand data. Inventory data has much higher autocorrelation, at both the daily and weekly level, than demand data. In addition, substitution would no longer need to be modelled, further simplifying the model, and potentially reducing noise in change signals if true demand data was available. Finally, if demand data were to be retained from hospitals, a more accurate demand change magnitude estimation procedure could be developed.

Chapter 7.3.2 Simulation Modelling of Response

Estimates of how a response to a demand change would occur could be included in future modelling. Reviews of processing capacity and flexibility of production could be used to model possible responses, and how best to avoid gaps in service.

Chapter 7.3.3 Individual Hospital Models

Modelling hospitals individually, rather than in aggregate as done here, would allow for more granular experimental changes to be made. It would also allow for more precise modelling of orders and rare blood type usage. Different hospitals have different ordering patterns, depending on the types of treatment provided. Modelling hospitals individually may increase fidelity when considering the transient component of a change in demand.

Chapter 7.3.4 Forecasting

Current forecasting models are rudimentary, predicting future inventory using only past inventory. There is significant room for improvement in forecasting, especially if demand data is used in place of inventory data. For instance, the effect of holidays and weather on demand could be included in future forecasting models.

Chapter 7.3.5 Substitution Optimization

In the future, the detection of changes in demand should be performed using demand data instead of inventory data. If this were the case, substitution would no longer need to be modelled to simulate changes in demand. Despite this, substitution between products is still an important part of inventory allocation. There are currently no consistent rules governing how products are substituted as it is performed by individual blood managers, and no records are maintained. Collecting data on platelet substitution and developing an optimization model for allocation could reduce costs and improve patient outcomes.

Chapter 8 Conclusion

Statistical Process Control is effective in detecting shifts in platelet demand. The key performance metrics, detection rate and detection time, improve as the magnitude of the shift increases. Both the rate of change, and its variability also play a role in detection performance.

A simulation model was created to generate inventory data representative of changes in hospital ordering patterns. The model, and all other project components, were built in R using R Studio. Model parameters were determined using data from the Ottawa centre for the 2019 calendar year. The simulation model was validated against the same data set.

A Statistical Process Control framework was constructed to detect change points in inventory signals. Parameters, such as control limits, were generated using 2019 inventory data. Standard SPC methods were adapted to deal with auto-correlated data more appropriately.

The SPC methods were evaluated first in the base case, where PRT platelets demand does not change, to determine the false positive rate. The false positive rate was determined to be approximately 3% at the end of the detection period.

Weekly forecast models were developed from established families of forecasting methods. The models were evaluated using historical data, base case runs of the inventory simulation, as well as data representative of demand shifts. The best performing method, a spline-based GAM model, was incorporated into the SPC analysis.

Experiments were run on data representative of changes in demand. It was found that larger shifts in demand had a higher probability of detection and a lower time to detection. Changes in demand, with an effect on the system larger than 10%, were always detected. Detection time varies greatly depending on the level of the demand shift.

The performance, and validation, of the methods developed are constrained by several issues. However, the use of simulation allows for a quick re-tooling to other datasets. The success of different methods in Statistical Process Control is often data specific, but the inclusion of a simulation component allows users to perform their own tests.

This tool was built as part of preparing for the introduction of PRT platelets but can be used to detect changes in demand regardless of the source. Once calibrated it can be run in perpetuity to help managers monitor inventory, and swiftly react to issues.

Bibliography

[1] "About us at Canadian Blood Services." https://www.blood.ca/en/about-us (accessed Oct. 27, 2021).

[2] S. Ramirez-Arcos *et al.*, "Extension of platelet shelf life with an improved bacterial testing algorithm," *Transfusion*, vol. 60, no. 12, pp. 2918–2928, 2020, doi: 10.1111/trf.16112.

[3] "Blood Component ABO Compatibility Chart | Blood Transfusion Services | UTMB Home." https://www.utmb.edu/bloodbank/component-therapy/blood-component-abo-compatibility-chart (accessed Oct. 27, 2021).

[4] J. P. Sigle *et al.*, "Comparison of transfusion efficacy of amotosalen-based pathogen-reduced platelet components and gamma-irradiated platelet components," *Transfusion*, vol. 53, no. 8, pp. 1788–1797, 2013, doi: 10.1111/j.1537-2995.2012.03959.x.

[5] L. J. Estcourt, R. Malouf, and M. F. Murphy, "Pathogen-reduced platelets for the prevention of bleeding in people of any age," *JAMA Oncol.*, vol. 4, no. 4, pp. 571–572, 2018, doi: 10.1001/jamaoncol.2017.5049.

[6] D. Montgomery, *Introduction to Statistical Quality Control*, 7th ed. Wiley-Blackwell, 2013.

[7] Steven Nahmias, "Perishable Inventory Theory : A Review Author (s): Steven Nahmias Published by : INFORMS Stable URL : http://www.jstor.org/stable/170438 REFERENCES Linked references are available on JSTOR for this article : You may need to log in to JSTOR to access the," vol. 30, no. 4, pp. 680–708, 1982.

[8] G. P. PRASTACOS, "Blood Inventory Management : An Overview of Theory and Practice Author (s): Gregory P . Prastacos Published by : INFORMS Stable URL : http://www.jstor.org/stable/2631647 REFERENCES Linked references are available on JSTOR for this article : You may need," vol. 30, no. 7, pp. 777–800, 2016.

[9] A. W. Flint, Z. K. McQuilten, G. Irwin, K. Rushford, H. E. Haysom, and E. M. Wood, "Is Platelet Expiring Out of Date? A Systematic Review," *Transfus. Med. Rev.*, vol. 34, no. 1, pp. 42–50, 2020, doi: 10.1016/j.tmrv.2019.08.006.

[10] S. M. Hesse, C. R. Coullard, M. S. Daskin, and A. P. Hurter, "Case study in platelet inventory management," *Industrial Engineering Research - Conference Proceedings*. pp. 801–806, 1997.

[11] A. F. Osorio, S. C. Brailsford, H. K. Smith, and J. T. Blake, "Designing the Blood Supply Chain: How Much, How and Where?," *Vox Sang.*, vol. 113, no. 8, pp. 760–769, 2018.

[12] G. C. Critchfield, D. P. Connelly, M. S. Ziehwein, L. S. Olesen, C. E. Nelson, and E. P. Scott, "Automatic prediction of platelet utilization by time series analysis in a large tertiary care hospital," *Am. J. Clin. Pathol.*, vol. 84, no. 5, pp. 627–631, 1985, doi: 10.1093/ajcp/84.5.627.

[13] O. S. S. Filho, W. Cezarino, and G. R. Salviano, *A decision-making tool for demand forecasting of blood components*, vol. 14, no. PART 1. IFAC, 2012.

[14] F. Lestari, U. Anwar, N. Nugraha, and B. Azwar, "Forecasting demand in blood supply chain (case study on blood transfusion unit)," *Lect. Notes Eng. Comput. Sci.*, vol. 2230, pp. 764–767, 2017.

[15] R. Khaldi, A. El Afia, R. Chiheb, and R. Faizi, "Artificial neural network based approach for blood demand forecasting: Fez transfusion blood center case study," *ACM Int. Conf. Proceeding Ser.*,

vol. Part F1294, pp. 3–8, 2017, doi: 10.1145/3090354.3090415.

[16] H. Shih and S. Rajendran, “Comparison of Time Series Methods and Machine Learning Algorithms for Forecasting Taiwan Blood Services Foundation’s Blood Supply,” *J. Healthc. Eng.*, vol. 2019, 2019, doi: 10.1155/2019/6123745.

[17] M. Motamedi, N. Li, D. G. Down, and N. M. Heddle, “Demand Forecasting for Platelet Usage: from Univariate Time Series to Multivariate Models,” 2021, [Online]. Available: http://arxiv.org/abs/2101.02305.

[18] A. J. Katz, C. W. Carter, P. Saxton, J. Blutt, and R. M. Kakaiya, “Simulation Analysis of Platelet Production and Inventory Management,” *Vox Sang.*, vol. 44, no. 1, pp. 31–36, 1983, doi: 10.1111/j.1423-0410.1983.tb04100.x.

[19] M. P. Atkinson, M. J. Fontaine, L. T. Goodnough, and L. M. Wein, “A novel allocation strategy for blood transfusions: Investigating the tradeoff between the age and availability of transfused blood,” *Transfusion*, vol. 52, no. 1, pp. 108–117, 2012, doi: 10.1111/j.1537-2995.2011.03239.x.

[20] A. Asllani, E. Culler, and L. Ettkin, “A simulation-based apheresis platelet inventory management model,” *Transfusion*, vol. 54, no. 1, pp. 2730–2735, 2014, doi: 10.1111/trf.12570.

[21] J. T. Blake, “Determining the inventory impact of extended-shelf-life platelets with a network simulation model,” *Transfusion*, vol. 57, no. 12, pp. 3001–3008, 2017, doi: 10.1111/trf.14305.

[22] J. T. Blake, S. Thompson, S. Smith, D. Anderson, R. Arellano, and D. Bernard, “Modelling Efficiency & Quality in Health Care,” *Proc. 29th Meet. EURO Work. Gr. Oper. Res. Appl. to Heal. Serv.*, pp. 47–65, 2004.

[23] R. Haijema, J. van der Wal, and N. M. van Dijk, “Blood platelet production: Optimization by dynamic programming and simulation,” *Comput. Oper. Res.*, vol. 34, no. 3, pp. 760–779, 2007, doi: 10.1016/j.cor.2005.03.023.

[24] N. Van Dijk, R. Haijema, J. Van Der Wal, and C. S. Sibinga, “Blood platelet production: A novel approach for practical optimization,” *Transfusion*, vol. 49, no. 3, pp. 411–420, 2009, doi: 10.1111/j.1537-2995.2008.01996.x.

[25] B. Zahiri, S. A. Torabi, M. Mousazadeh, and S. A. Mansouri, “Blood collection management: Methodology and application,” *Appl. Math. Model.*, vol. 39, no. 23–24, pp. 7680–7696, 2013, doi: 10.1016/j.apm.2015.04.028.

[26] U. Abdulwahab and M. I. M. Wahab, “Approximate dynamic programming modeling for a typical blood platelet bank,” *Comput. Ind. Eng.*, vol. 78, pp. 259–270, 2014, doi: 10.1016/j.cie.2014.07.017.

[27] S. Gunpinar and G. Centeno, “Stochastic integer programming models for reducing wastages and shortages of blood products at hospitals,” *Comput. Oper. Res.*, vol. 54, pp. 129–141, 2015, doi: 10.1016/j.cor.2014.08.017.

[28] I. Civelek, I. Karaesmen, and A. Scheller-Wolf, “Blood platelet inventory management with protection levels,” *Eur. J. Oper. Res.*, vol. 243, no. 3, pp. 826–838, 2015, doi: 10.1016/j.ejor.2015.01.023.

[29] L. Guan *et al.*, “Big data modeling to predict platelet usage and minimize wastage in a tertiary

care system," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 114, no. 43, pp. 11368–11373, 2017, doi: 10.1073/pnas.1714097114.

[30] T. Z. Poisson and R. Model, "Zero-Inflated Poisson Regression," pp. 1–14, 2014.

[31] M. J. Schervish, A. M. Law, and W. D. Kelton, *Simulation Modeling and Analysis.*, vol. 78, no. 383. 1983.

[32] A. Tartakovsky, I. Nikiforov, and M. Basseville, *Sequential analysis: Hypothesis testing and changepoint detection*. 2014.

[33] R. Lawton, "Time Series Analysis and its Applications," *Int. J. Forecast.*, vol. 17, no. 2, pp. 299–301, 2001, doi: 10.1016/s0169-2070(01)00083-8.

[34] T. Hastie, R. Tibshirani, and J. Friedman, *Springer Series in Statistics The Elements of Statistical Learning*, 12th ed. 2017.

[35] S. N. Wood, *Generalized additive models: An introduction with R, second edition*. 2017.

[36] R. P. Adams and D. J. C. MacKay, "Bayesian Online Changepoint Detection," 2007, [Online]. Available: http://arxiv.org/abs/0710.3742.

Appendix A: Distribution Fitting

The probability distribution of the zero inflated Poisson distribution for a random variable y_i can be described as in Equation 12. The first case describes the situation where the count is zero independent of the Poisson process, which occurs with probability π. The second case describes the Poisson process of parameter μ, including possible zero counts, which occurs with probability 1 - π.

Equation 12 Probability distribution for a Zero Inflated Poisson (ZIP) process

$$P(y_i = j) = \begin{cases} \pi_i + (1 - \pi_i)e^{-\mu_i} & if\ j = 0 \\ (1 - \pi_i)\dfrac{\mu_i^{y_i} e^{-\mu_i}}{y_i!} & if\ j > 0 \end{cases}$$

The Zero Inflated Poisson distribution parameters and fitting information for daily orders of each blood and product type are shown below in Table 17. The same information for daily collections is shown in Table 18. The differences in Pi values between collections and orders may indicate differences in the consistency of the two processes, one of the underlying difficulties associated with inventory management.

Table 17 Zero Inflated Poisson distribution parameters for daily orders, separated by blood and product type

Blood Type	Product Type	μ, Mu	π, Pi	Loglikelihood	AIC
A+	Pooled Platelets	14.76	2.44e-10	-544.43	1092.87
A-	Pooled Platelets	1.52	0.15	-627.77	1259.54
B+	Pooled Platelets	3.82	0.20	-939.77	1883.53
B-	Pooled Platelets	1.04	0.0093	-227.40	458.80
AB+	Pooled Platelets	1.13	0.85	-174.35	352.70
AB-	Pooled Platelets	--	--	--	--
O+	Pooled Platelets	15.38	2.20e-08	-571.35	1146.69
O-	Pooled Platelets	4.01	0.23	-1025.22	2054.44
A+	Apheresis Platelets	4.36	0.17	-1011.03	2026.06
A-	Apheresis Platelets	1.76	0.46	-513.60	1031.21
B+	Apheresis Platelets	1.78	0.46	-525.23	1054.47
B-	Apheresis Platelets	1.35	0.0096	-262.82	529.64
AB+	Apheresis Platelets	1.54	0.64	-367.10	738.20
AB-	Apheresis Platelets	1.58	0.97	-191.66	387.31
O+	Apheresis Platelets	4.78	0.16	-1100.41	2204.82
O-	Apheresis Platelets	0.96	0.68	-257.28	518.57

Table 18 Zero Inflated Poisson distribution parameters for daily collections, separated by blood and product type

Blood Type	Product Type	Mu	Sigma	Loglikelihood	AIC
A+	Pooled Platelets	14.60	8.79e-08	-550.55	1105.09
A-	Pooled Platelets	1.61	0.20	-603.15	1210.30
B+	Pooled Platelets	3.62	0.16	-862.50	1728.96
B-	Pooled Platelets	1.36	0.0093	-227.40	458.80
AB+	Pooled Platelets	1.51	0.0087	-1275.02	2554.05
AB-	Pooled Platelets	--	--	--	--
O+	Pooled Platelets	15.22	5.85e-08	-593.73	1191.47
O-	Pooled Platelets	3.80	0.19	-949.38	1902.75
A+	Apheresis Platelets	4.50	0.20	-901.68	1807.36
A-	Apheresis Platelets	1.97	0.53	-487.80	979.60
B+	Apheresis Platelets	2.39	0.60	-481.58	967.16
B-	Apheresis Platelets	1.91	0.0097	-124.35	252.70
AB+	Apheresis Platelets	1.73	0.677	-348.19	700.39
AB-	Apheresis Platelets	1.90	0.0097	-109.72	223.45
O+	Apheresis Platelets	4.78	0.17	-938.86	1881.72
O-	Apheresis Platelets	1.80	0.0083	-356.01	716.02

As can be seen below in Table 19, autocorrelation of demand and collections with values from the previous day are low, typically less than 0.15. This indicates that the assumption of independence for both demand and collections is sound, at least at the aggregate scale.

Table 19 Autocorrelation of demand and collections with the respective quantity from the previous day. The low values seen in the table below indicate that the assumption of independent observations holds up at the regional level

Blood Type	Product Type	Demand Autocorrelation with Previous Day	Collections Autocorrelation with Previous Day
A+	Pooled platelets	-0.09	0.10
A-	Pooled platelets	-0.04	0.16
B+	Pooled platelets	-0.13	0.14
B-	Pooled platelets	0.07	-0.01
AB+	Pooled platelets	0.08	0.14
AB-	Pooled platelets	--	--
O+	Pooled platelets	-0.01	0.17
O-	Pooled platelets	0.07	0.16
A+	Apheresis platelets	-0.11	-0.13
A-	Apheresis platelets	-0.12	0.11
B+	Apheresis platelets	-0.10	-0.13
B-	Apheresis platelets	0.04	-0.03
AB+	Apheresis platelets	-0.10	-0.04
AB-	Apheresis platelets	0.03	-0.03
O+	Apheresis platelets	-0.09	-0.14
O-	Apheresis platelets	-0.11	-0.08

Appendix B: Bayesian Online Changepoint Detection

In addition to Statistical Process Control Charts, a Bayesian method was also tested. An online Bayesian approach to changepoint detection was developed by Adams [36] and can be accessed using the R library 'ocp'. This method calculates a predictive distribution for the next time steps, based on the prior probability distribution of recent data in the current production run. The difference between the predictive and prior distributions determines if a change has occurred.

This approach was first tested on the base model and its performance shown below in Table 20. The false positive rate for BOCPD is significantly higher than that for the statistical process control chart.

Table 20 Evaluation of Bayesian Online Changepoint Detection on base model data

Method	Detection Rate, %	Mean Time to Detection, weeks	Mean Time to Detection, 95% CI Halfwidth, weeks
SPC	3.0	32.25	10.77
BOCPD	20.5	27.82	4.88

The performance of the online Bayesian approach was also tested on sets of single factor demand shift data as can be seen in Table 21 through Table 23. Equivalent data for Statistical Process Control Charts can be found in Appendix H.

Comparing the two methods it becomes clear there are significant differences in performance between Bayesian Online Changepoint Detection and control charts. Table 24 shows the results of a Chi-square test comparing the detection times in each of the single factor experiments. In each case the results show that there is a statistical difference between the methods. While this can be partially explained by the higher false positive rate of BOCPD the detection rate for large changes is lower than that of statistical process control charts, indicating less consistent performance.

Similar results can be found when comparing the time to detection for each method. The t-test results shown in Table 25 indicate that statistical process control charts have significantly better performance in each of the scenarios, except for an increase in apheresis demand, where the results of the two methods are indistinguishable from each other.

Table 21 Performance of Bayesian Online Changepoint Detection on increase in pooled demand experiments

Increase in Pooled Demand	Detection Rate, %	Mean Time to Detection, Weeks	Mean Time to Detection 95% CI Halfwidth, Weeks
2%	18	28.56	5.69
4%	19	23.10	5.06
6%	29	21.07	4.35
8%	28	25.32	5.34
10%	25	27.64	6.17
12%	36	24.17	3.65
14%	44	25.25	3.65
16%	47	23.70	3.87
18%	60	22.20	3.14
20%	63	20.86	3.11

Table 22 Performance of Bayesian Online Changepoint Detection on increase in apheresis demand experiments

Increase in Apheresis Demand	Detection Rate, %	Mean Time to Detection, Weeks	Mean Time to Detection 95% CI Halfwidth, Weeks
2%	21	27.00	5.57
4%	19	30.53	6.31
6%	18	29.28	6.50
8%	24	26.58	5.30
10%	31	23.39	4.77
12%	33	25.00	3.90
14%	36	26.39	4.01
16%	35	23.91	5.56
18%	53	18.94	3.12
20%	51	20.01	3.11

Table 23 Performance of Bayesian Online Changepoint Detection on movement of pooled demand to apheresis experiments

Pooled Demand Moved to Apheresis	Detection Rate, %	Mean Time to Detection, Weeks	Mean Time to Detection 95% CI Halfwidth, Weeks
2%	22	26.77	4.28
4%	35	25.14	4.41
6%	36	19.89	4.21
8%	43	16.81	3.27
10%	73	14.67	2.82
12%	79	10.77	2.19
14%	77	11.83	2.73
16%	90	9.41	2.12
18%	82	10.45	2.54
20%	99	3.31	0.45

Table 24 Chi-square test comparison of detection rates for Bayesian Online Changepoint Detection and Statistical Process Control Charts across 3 sets of experiments

Experiment	X^2	Degrees of Freedom	p-value
Increase in Pooled Demand	26.91	9	1.44e-3
Increase in Apheresis Demand	51.29	9	6.15e-8
Pooled Demand Moved to Apheresis	39.89	9	7.95e-6

Table 25 T-test comparison of detection times for Bayesian Online Changepoint Detection and Statistical Process Control Charts across 3 sets of experiments

Experiment	t-value	Degrees of Freedom	p-value	Mean of the Differences, units	Mean Difference 95% CI Halfwidth, units
Increase in Pooled Demand	-10.58	972.00	2.2e-16	8.71	1.62
Increase in Apheresis Demand	0.74	519.00	0.46	0.85	2.27
Pooled Demand Moved to Apheresis	-3.55	1223.00	4.1e-4	2.10	1.16

Appendix C: Residual Error Control Charts

A common technique used in statistical process control when data is significantly autocorrelated is to directly model the correlative structure of the data with a time series model, as described in Montgomery [6]. With autocorrelation removed from the data, an individual control chart can then be applied to the model residuals. For this approach to work the time series model chosen must accurately represent the autocorrelation of the process data. An ARIMA (2, 0, 2) model was used as the underlying model for this approach.

This approach was first tested on the base model and its performance shown below in Table 26. The false positive rate for the residual error control chart is close to that of the standard control chart, albeit slightly smaller.

Table 26 Performance of time-series model residual error control charts on base model data

Method	Detection Rate, %	Mean Time to Detection, weeks	Mean Time to Detection, 95% CI Halfwidth, weeks
SPC	3.2	32.25	10.77
BOCPD	3.0	25.71	14.62

The performance of the residual error control chart approach was also tested on sets of single factor demand shift data as can be seen in Table 27 through Table 29. Equivalent data for standard control charts can be found in Appendix H.

Comparing the two methods it becomes clear there are significant differences in performance between the standard and residual error control charts. Table 30 shows the results of a Chi-square test comparing the detection times in each of the single factor experiments. In each case the results show that there is a statistical difference between the methods. The detection rate for the residual error control chart is much lower, showing that while the low false positive rate is attractive, it does not maximize overall detection accuracy.

The time to detection for each of the two control charts was also compared. The t-test results shown in Table 31 indicate that statistical process control charts have significantly better performance in each of the scenarios.

Table 27 Performance of time-series residual change monitoring on increase in pooled demand experiments

Increase in Pooled Demand	Detection Rate, %	Mean Time to Detection, Weeks	Mean Time to Detection 95% CI Halfwidth, Weeks
2%	3	34.33	16.54
4%	5	19.60	12.56
6%	6	30.33	11.20
8%	9	30.00	12.75
10%	12	24.91	7.27
12%	17	31.35	6.97
14%	19	20.68	8.30
16%	26	21.61	6.74
18%	28	19.11	5.91
20%	31	17.87	6.27

Table 28 Performance of time-series residual change monitoring on increase in apheresis demand experiments

Increase in Apheresis Demand	Detection Rate, %	Mean Time to Detection, Weeks	Mean Time to Detection 95% CI Halfwidth, Weeks
2%	3	25.00	43.10
4%	0	--	--
6%	3	13.33	24.89
8%	0	--	--
10%	4	25.75	30.65
12%	3	10.67	8.72
14%	1	29.00	--
16%	4	31.50	20.15
18%	6	20.33	18.98
20%	3	28.00	30.12

Table 29 Performance of time-series residual change monitoring on movement of pooled demand to apheresis experiments

Pooled Demand Moved to Apheresis	Detection Rate, %	Mean Time to Detection, Weeks	Mean Time to Detection 95% CI Halfwidth, Weeks
2%	2	7.50	--
4%	2	16.50	--
6%	3	39.33	18.81
8%	5	17.60	25.26
10%	9	14.44	9.04
12%	7	20.71	14.70
14%	10	34.00	12.01
16%	19	22.11	9.07
18%	13	22.15	10.68
20%	16	14.75	8.85

Table 30 Chi-square test comparison of detection rates for time-series residual control charts with standard control charts across 3 sets of experiments

Experiment	X^2	Degrees of Freedom	p-value
Increase in Pooled Demand	31.22	9	1.44e-5
Increase in Apheresis Demand	61.29	9	6.59e-9
Pooled Demand Moved to Apheresis	49.72	9	7.65e-8

Table 31 T-test comparisons of time to detection for time-series residual control charts with standard control charts across 3 sets of experiments

Experiment	t-value	Degrees of Freedom	p-value	Mean of the Differences, weeks	Mean Difference 95% CI Halfwidth, weeks
Increase in Pooled Demand	-5.60	209.00	6.54e-08	7.63	2.68
Increase in Apheresis Demand	0.72	285.00	0.47	2.13	5.81
Pooled Demand Moved to Apheresis	-5.92	91.00	5.53e-08	10.97	3.68

Appendix D: Stationarity Testing

Table 32 Results of the Augmented Dickey-Fuller test with two different values of K. Both cases show that weekly inventory data is stationary

K value	Dickey-Fuller Value	Lag Order	p-value	Alternative Hypothesis
--	-1.7321	9	0.6917	Stationary
0	1.7321	0	0.99	Stationary

The autocorrelation of inventory for each day of the first three weeks is shown below in Table 33. The autocorrelation of the historical data has a much smaller sample size, 1 year, compared to the simulation validation set, 50 years. Simulation validation showed that there were no significant differences in the daily inventory levels of the two series. It may be noted that the simulation does not model holidays, or weather, two factors which may affect daily autocorrelation.

Table 33 Autocorrelation of daily inventory with the inventory of previous days. Lag 1, indicates the correlation of inventory with the day immediately before

Lag, Days	Autocorrelation, Historical Data	Autocorrelation, Simulation
0	1.00	1.00
1	0.62	0.85
2	0.41	0.67
3	0.26	0.47
4	0.17	0.27
5	0.17	0.07
6	0.21	-0.07
7	0.30	-0.16
8	0.10	-0.19
9	0.03	-0.17
10	-0.03	-0.11
11	-0.07	-0.03
12	-0.03	0.06
13	0.03	0.14
14	0.14	0.18
15	0.01	0.20
16	0.02	0.19
17	0.06	0.16
18	0.06	0.10
19	0.06	0.04
20	0.09	-0.01

The autocorrelation of weekly inventory is shown below in Table 34. The historical data and simulated data exhibit similar values for the first several weeks. The lag 1 autocorrelation is equal to, or greater than, 0.25. A general guideline in statistical process control, is that autocorrelation of 0.25 or greater can significantly affect the performance of change detection[6].

Table 34 Autocorrelation of weekly inventory with the inventory of previous weeks. Lag 1 indicates the correlation of inventory with the previous week

Lag, Weeks	Autocorrelation, Historical Data	Autocorrelation, Simulation
0	1.00	1.00
1	0.29	0.25
2	0.08	0.06
3	0.17	0.11
4	0.07	-0.02
5	-0.11	-0.02
6	0.14	0.00
7	0.21	0.03
8	0.06	0.06
9	0.03	-0.05
10	-0.14	-0.08
11	-0.16	0.03
12	-0.01	0.02
13	0.16	0.03
14	0.02	0.03
15	0.14	0.06
16	0.06	-0.03
17	-0.06	-0.02

Appendix E: 2019 Northeastern Ontario Inventory Data Summary

Table 35 Northeastern Ontario daily inventory summary, 2019 calendar year

Blood Type	Product Type	Mean Daily Inventory, units	Daily Inventory Stand. Dev, units
A+	**Pooled**	56.99	16.02
A-	**Pooled**	4.98	3.28
B+	**Pooled**	11.14	4.80
B-	**Pooled**	0.28	0.55
AB+	**Pooled**	0.64	1.37
AB-	**Pooled**	0.00	0.00
O+	**Pooled**	60.63	20.10
O-	**Pooled**	13.11	7.19
A+	**Apheresis**	13.00	5.51
A-	**Apheresis**	2.88	2.31
B+	**Apheresis**	3.64	2.57
B-	**Apheresis**	0.189	0.56
AB+	**Apheresis**	1.82	1.67
AB-	**Apheresis**	0.14	0.49
O+	**Apheresis**	15.00	6.32
O-	**Apheresis**	1.03	1.08

Table 36 Northeastern Ontario daily inventory properties, 2019 calendar year

Property	Pooled Inventory	Apheresis Inventory
Mean	147.8	37.7
Standard Deviation	28.4	9.2
coefficient of variation	0.19	0.24

Table 37 Northeastern Ontario daily collections and demand summary, 2019 calendar year

Blood Type	Product Type	Mean Orders	Stand. Dev Orders	Mean Collections	Stand. Dev Collections
A+	Pooled	14.76	10.67	14.60	10.25
A-	Pooled	1.29	1.71	1.29	1.48
B+	Pooled	3.06	3.09	3.04	2.58
B-	Pooled	0.08	0.28	0.08	0.28
AB+	Pooled	0.17	0.65	0.17	0.58
AB-	Pooled	0.00	0.00	0.00	0.00
O+	Pooled	15.38	11.72	15.22	12.26
O-	Pooled	3.08	3.73	3.07	3.15
A+	Apheresis	3.64	3.46	3.60	3.01
A-	Apheresis	0.94	1.43	0.93	1.38
B+	Apheresis	0.96	1.47	0.96	1.57
B-	Apheresis	0.06	0.30	0.06	0.34
AB+	Apheresis	0.56	1.06	0.56	1.02
AB-	Apheresis	0.05	0.29	0.05	0.31
O+	Apheresis	4.01	4.01	3.97	3.16
O-	Apheresis	0.30	0.67	0.30	0.69

Appendix F: Simulation Validation

The simulation was validated using three quantities: daily collections, daily shipments, and daily inventory. These three quantities comprise the input, output, and mean daily size of the system, respectively. Having successfully validated all three, it may be assumed that the simulated system is indistinguishable from the actual system. Validation is performed on total, pooled, and apheresis quantities to ensure that product behaviour is correct. Collection's validation can be seen below in Table 38, Shipment validation in Table 39, and daily inventory in Table 40. For comparisons of historical and simulated autocorrelation, see Appendix D.

Table 38 Simulation collections validation t-test comparison by product type

Inventory Type	Data Daily Mean Collections	Simulation Daily Mean	95%, Upper CI	95%, Lower CI	t-value	Degrees of Freedom	p-value
Total	52.36	51.25	4.01	-0.86	1.31	374.00	0.25
Pooled	40.11	39.74	3.16	-1.37	0.77	366.00	0.38
Apheresis	14.24	13.83	1.14	-0.21	1.96	375.00	0.08

Table 39 Simulation shipments validation t-test comparison by product type

Inventory Type	Data Daily Mean Shipments	Simulation Daily Mean	95%, Upper CI	95%, Lower CI	t-value	Degrees of Freedom	p-value
Total	48.40	46.76	4.19	-0.93	1.26	369.00	0.21
Pooled	37.94	37.03	3.09	-1.27	0.82	370.00	0.41
Apheresis	10.46	9.73	1.48	-0.03	1.90	373.00	0.06

Table 40 Simulation inventory validation t-test comparison by product type

Inventory Type	Data Daily Mean Inventory	Simulation Daily Mean Inventory	95%, Upper CI	95%, Lower CI	t-value	Degrees of Freedom	p-value
Total	185.48	185.21	3.20	-2.64	0.1858	401.00	0.85
Pooled	147.79	146.85	3.92	-2.04	0.6211	391.00	0.54
Apheresis	37.70	38.36	0.32	-1.65	-1.324	422.00	0.19

Appendix G: Base Model Results

Table 41 Comparison of performance for Statistical Process Control Charts on base model data, with and without forecasting

Method	Detection Rate, %	Mean Time to Detection, Weeks	Time to Detection 95% CI Halfwidth, Weeks
Standard	3.20	32.25	10.77
Linear Regression	6.5	28.61	10.23
ARIMA	3.20	32.00	10.76
GAM, Local Regression	31.50	22.79	3.69
GAM, Splines	4.90	27.59	6.85

Table 42 Comparison of forecasting accuracy on base model data

Method	Base Case MAPE, %	RMSE, units of platelets	Base Case Detection Rate, %
No Forecasting	--	--	3.20
ARIMA	13.14	16.85	3.20
Linear Regression	15.45	20.00	6.50
GAM, Local Regression	16.64	21.90	31.50
GAM, Splines	14.35	18.57	4.90

Appendix H: Single Factor Experiment Results

Standard Results

The results for the standard statistical process control chart on single factor demand experiments are shown below in Table 43 through Table 47. The results for a gradual increase in demand, as opposed to a step input, are shown in Table 46. The results for an increase in the variability of orders are shown in Table 47. These results show that a gradual change in demand is harder to detect, while one with higher variability is easier to detect.

Table 43 Increase in pooled demand experiment results using Statistical Process Control Charts

Increase in Pooled Demand, %	Detection Rate, %	Mean Time to Detection, Weeks	95% Confidence Interval Halfwidth
2	4	33.50	11.13
4	13	19.00	6.83
6	22	21.18	5.78
8	39	24.49	4.71
10	53	25.04	4.31
12	79	20.94	2.88
14	97	15.86	2.45
16	98	10.73	1.61
18	100	8.99	1.48
20	100	7.33	1.26

Table 44 Increase in apheresis demand experiment results using Statistical Process Control Charts

Increase in Apheresis Demand, %	Detection Rate, %	Mean Time to Detection, Weeks	95% Confidence Interval Halfwidth
2	0	--	--
4	5	21.14	--
6	6	24.67	--
8	7	28.20	--
10	18	28.50	7.88
12	17	24.71	7.37
14	38	27.42	5.05
16	35	26.51	5.32
18	64	23.02	3.71
20	70	23.21	3.15

Table 45 Movement of pooled demand to apheresis experiment results using Statistical Process Control Charts

Movement of Pooled Demand to Apheresis, %	Detection Rate, %	Mean Time to Detection, Weeks	95% Confidence Interval Halfwidth
2	4	22.00	17.47
4	13	25.33	7.75
6	22	22.47	4.79
8	68	20.75	3.42
10	100	12.43	1.63
12	100	8.78	1.10
14	100	8.72	1.29
16	100	6.60	0.78
18	100	6.09	0.57
20	100	3.20	0.17

Table 46 Linear increase in pooled demand experiment results using Statistical Process Control Charts. Slope = 1/3

Gradual Increase in Pooled Demand, %	Detection Rate, %	Mean Time to Detection, Weeks	95% Confidence Interval Halfwidth
2	6	28.00	15.61
4	18	28.83	7.74
6	27	26.33	6.37
8	37	28.62	4.53
10	61	22.54	3.29
12	77	18.69	2.91
14	91	16.42	2.57
16	98	12.73	1.81
18	100	8.51	1.09
20	100	7.68	0.96

Table 47 Increase in demand variance experiment results using Statistical Process Control Charts

Increase in Standard Deviation	Detection Rate, %	Mean Time to Detection, Weeks	95% Confidence Interval Halfwidth
0.02	5	23.40	21.01
0.04	3	25.67	48.57
0.06	8	26.75	14.72
0.08	5	26.40	22.36
0.10	9	18.00	7.59
0.12	12	30.33	11.78
0.14	13	25.00	8.33
0.16	14	28.50	9.49
0.18	30	25.10	5.55
0.20	29	23.69	5.57

Linear Regression Results

The results for statistical process control charts including a linear forecasting model on single factor demand experiments are shown below in Table 48 through Table 50. A Chi-square test comparison of detection rates to the standard values is included in Table 51. A t-test comparison to standard values for time to detection is presented in Table 52.

The results show that there is no significant difference in detection rate for any of the experiments. When comparing detection times, there are meaningful differences for both the increase in pooled demand, and movement of pooled demand experiments.

Table 48 Increase in pooled demand experiment results using Statistical Process Control Charts including linear forecasting

Increase in Pooled Demand, %	Detection Rate, %	Mean Time to Detection, Weeks	95% Confidence Interval Halfwidth
2	6	33.00	13.25
4	15	18.40	5.89
6	27	20.48	4.86
8	46	24.30	4.27
10	58	22.60	4.01
12	82	19.65	2.82
14	97	15.07	2.44
16	98	10.17	1.59
18	100	8.10	1.39
20	100	6.76	1.21

Table 49 Increase in apheresis demand experiment results using Statistical Process Control Charts including linear forecasting

Increase in Apheresis Demand, %	Detection Rate, %	Mean Time to Detection, Weeks	95% Confidence Interval Halfwidth
2	1	15.00	--
4	6	24.67	16.10
6	8	18.63	13.69
8	8	32.25	11.36
10	20	26.40	7.51
12	18	24.56	6.92
14	39	26.46	4.88
16	36	26.41	5.16
18	64	22.79	3.70
20	72	21.65	2.91

Table 50 Movement of pooled demand to apheresis experiment results using Statistical Process Control Charts including linear forecasting

Movement of Pooled Demand to Apheresis, %	Detection Rate, %	Mean Time to Detection, Weeks	95% Confidence Interval Halfwidth
2	6	21.83	17.28
4	19	25.84	7.38
6	34	22.47	4.79
8	68	20.75	3.42
10	100	12.43	1.63
12	100	8.78	1.10
14	100	8.70	1.29
16	100	6.60	0.78
18	100	6.09	0.57
20	100	3.20	0.17

Table 51 Chi-square test comparison of detection rates for linear model forecasting enhanced control charts with standard control charts across 3 sets of experiments

Experiment	X^2	Degrees of Freedom	p-value
Increase in Pooled Demand	2.16	9	0.98
Increase in Apheresis Demand	12.33	9	0.20
Pooled Demand Moved to Apheresis	3.82	9	0.92

Table 52 T-test comparisons of time to detection for linear model forecasting enhanced control charts with standard control charts across 3 sets of experiments

Experiment	t-value	Degrees of Freedom	p-value	Mean of the Differences, units	Mean Difference 95% CI Halfwidth, units
Increase in Pooled Demand	5.66	137.00	8.56e-8	9.57	3.34
Increase in Apheresis Demand	0.57	527.00	0.56	0.72	2.47
Pooled Demand Moved to Apheresis	7.40	114.00	2.5e-11	15.91	4.26

ARIMA results

The results for statistical process control charts including a linear forecasting model on single factor demand experiments are shown below in Table 53 through Table 55. A Chi-square test comparison of detection rates to the standard values is included in Table 56. A t-test comparison to standard values for time to detection is presented in Table 57.

The results show that there is no significant difference between the detection rate including ARIMA and the standard method. The results also show that there is no significant difference in the time to detection.

Table 53 Increase in pooled demand experiment results using Statistical Process Control Charts including ARIMA forecasting

Increase in Pooled Demand, %	Detection Rate, %	Mean Time to Detection, Weeks	95% Confidence Interval Halfwidth
2	4	33.50	11.18
4	13	19.00	6.83
6	22	21.18	5.78
8	40	23.95	4.71
10	53	25.04	4.31
12	79	20.94	2.88
14	97	15.86	2.45
16	98	10.73	1.61
18	100	8.99	1.48
20	100	7.33	1.26

Table 54 Increase in apheresis demand experiment results using Statistical Process Control Charts including ARIMA forecasting

Increase in Apheresis Demand, %	Detection Rate, %	Mean Time to Detection, Weeks	95% Confidence Interval Halfwidth
2	0	--	--
4	6	25.33	16.37
6	6	24.67	16.10
8	7	21.14	14.72
10	18	28.50	7.88
12	17	24.71	7.37
14	38	27.42	5.05
16	35	26.51	5.32
18	64	23.02	3.71
20	70	23.21	3.15

Table 55 Movement of pooled demand to apheresis experiment results using Statistical Process Control Charts including ARIMA forecasting

Movement of Pooled Demand to Apheresis, %	Detection Rate, %	Mean Time to Detection, Weeks	95% Confidence Interval Halfwidth
2	6	22.00	17.47
4	19	25.33	7.75
6	34	22.47	4.79
8	70	20.75	3.42
10	100	12.43	1.63
12	100	8.78	1.10
14	100	8.72	1.29
16	100	6.60	0.78
18	100	6.09	0.57
20	100	3.20	0.17

Table 56 Chi-square test comparison of detection rates for ARIMA model forecasting enhanced control charts with standard control charts across 3 sets of experiments

Experiment	X^2	Degrees of Freedom	p-value
Increase in Pooled Demand	0.01	9	1
Increase in Apheresis Demand	0.02	9	1
Pooled Demand Moved to Apheresis	3.79	9	0.92

Table 57 T-test comparisons of time to detection for ARIMA model forecasting enhanced control charts with standard control charts across 3 sets of experiments

Experiment	t-value	Degrees of Freedom	p-value	Mean of the Differences, units	Mean Difference 95% CI Halfwidth, units
Increase in Pooled Demand	0.15	92.00	0.99	0.37	0.54
Increase in Apheresis Demand	0.57	258.00	0.98	0.81	0.96
Pooled Demand Moved to Apheresis	-0.24	121.00	0.95	0.92	1.01

GAM, Local Regression Results

The results for statistical process control charts including a linear forecasting model on single factor demand experiments are shown below in Table 58 through Table 60. A Chi-square test comparison of detection rates to the standard values is included in Table 61. A t-test comparison to standard values for time to detection is presented in Table 62.

The results show that there is a significant difference in detection rate for all the experiments, likely a result of the extremely high detection sensitivity of the method. When comparing detection times, there are meaningful differences for both the increase in pooled demand, and movement of pooled demand experiments.

Table 58 Increase in pooled demand experiment results using Statistical Process Control Charts including local regression-based GAM forecasting

Increase in Pooled Demand, %	Detection Rate, %	Mean Time to Detection, Weeks	95% Confidence Interval Halfwidth
2	29	24.93	6.01
4	56	21.04	3.41
6	63	19.23	3.62
8	73	21.60	3.39
10	88	18.50	3.13
12	94	14.43	2.50
14	99	10.11	1.69
16	100	8.25	1.30
18	100	6.37	1.07
20	100	5.76	0.96

Table 59 Increase in apheresis demand experiment results using Statistical Process Control Charts including local regression-based GAM forecasting

Increase in Apheresis Demand, %	Detection Rate, %	Mean Time to Detection, Weeks	95% Confidence Interval Halfwidth
2	30	26.70	6.57
4	36	20.69	6.20
6	33	25.81	4.83
8	29	23.59	6.24
10	40	24.15	4.87
12	45	26.27	4.57
14	63	21.95	4.04
16	56	22.93	4.11
18	77	20.57	2.95
20	84	19.55	2.90

Table 60 Movement of pooled demand to apheresis experiment results using Statistical Process Control Charts including local regression-based GAM forecasting

Movement of Pooled Demand to Apheresis, %	Detection Rate, %	Mean Time to Detection, Weeks	95% Confidence Interval Halfwidth
2	36	21.42	4.93
4	41	21.80	4.72
6	66	19.82	3.64
8	87	18.78	3.08
10	100	11.22	1.50
12	100	7.85	1.11
14	100	7.75	1.25
16	100	6.06	0.79
18	100	5.48	0.57
20	100	3.02	0.20

Table 61 Chi-square test comparison of detection rates for local regression-based GAM forecasting enhanced control charts with standard control charts across 3 sets of experiments

Experiment	X^2	Degrees of Freedom	p-value
Increase in Pooled Demand	42.09	9	3.17e-6
Increase in Apheresis Demand	53.05	9	2.86e-8
Pooled Demand Moved to Apheresis	54.96	9	1.24e-8

Table 62 T-test comparisons of time to detection for local regression-based GAM forecasting enhanced control charts with standard control charts across 3 sets of experiments

Experiment	t-value	Degrees of Freedom	p-value	Mean of the Differences, units	Mean Difference 95% CI Halfwidth, units
Increase in Pooled Demand	11.09	428.00	2.2e-16	10.57	1.87
Increase in Apheresis Demand	1.93	537.00	0.06	2.18	2.22
Pooled Demand Moved to Apheresis	10.14	344.00	2.2e-16	13.72	2.66

GAM, Splines Results

The results for statistical process control charts including a linear forecasting model on single factor demand experiments are shown below in Table 63 through Table 65. A Chi-square test comparison of detection rates to the standard values is included in Table 66. A t-test comparison to standard values for time to detection is presented in Table 67.

The results show that there is no significant difference in detection rate for any of the sets of experiments. When comparing detection times, there are meaningful differences for both the increase in pooled demand, and movement of pooled demand experiments.

Table 63 Increase in pooled demand experiment results using Statistical Process Control Charts including spline-based GAM forecasting

Increase in Pooled Demand, %	Detection Rate, %	Mean Time to Detection, Weeks	95% Confidence Interval Halfwidth
2	6	24.50	15.73
4	23	21.39	5.82
6	37	21.41	4.57
8	53	23.71	4.27
10	60	22.40	3.99
12	85	18.60	2.78
14	99	14.40	2.31
16	100	9.97	1.52
18	100	7.73	1.39
20	100	6.97	1.28

Table 64 Increase in apheresis demand experiment results using Statistical Process Control Charts including spline-based GAM forecasting

Increase in Apheresis Demand, %	Detection Rate, %	Mean Time to Detection, Weeks	95% Confidence Interval Halfwidth
2	11	24.80	11.15
4	2	9.00	5.52
6	8	23.63	13.21
8	10	25.00	9.68
10	23	28.04	6.57
12	20	26.35	6.08
14	43	25.72	4.79
16	40	25.60	5.10
18	65	21.64	3.49
20	73	21.07	3.01

Table 65 Movement of pooled demand to apheresis experiment results using Statistical Process Control Charts including spline-based GAM forecasting

Movement of Pooled Demand to Apheresis, %	Detection Rate, %	Mean Time to Detection, Weeks	95% Confidence Interval Halfwidth
2	14	21.50	8.81
4	22	25.43	6.81
6	38	23.58	4.61
8	74	20.03	3.30
10	100	12.10	1.57
12	100	8.71	1.11
14	100	8.67	1.29
16	100	6.60	0.78
18	100	5.85	0.55
20	100	3.16	0.182

Table 66 Chi-square test comparison of detection rates for spline-based GAM forecasting enhanced control charts with standard control charts across 3 sets of experiments

Experiment	X^2	Degrees of Freedom	p-value
Increase in Pooled Demand	2.38	9	0.98
Increase in Apheresis Demand	12.51	9	0.19
Pooled Demand Moved to Apheresis	11.24	9	0.26

Table 67 T-test comparisons of time to detection for spline-based GAM forecasting enhanced control charts with standard control charts across 3 sets of experiments

Experiment	t-value	Degrees of Freedom	p-value	Mean of the Differences, units	Mean Difference 95% CI Halfwidth, units
Increase in Pooled Demand	5.45	169.00	1.72e-7	9.10	3.29
Increase in Apheresis Demand	0.89	541.00	0.38	1.10	2.43
Pooled Demand Moved to Apheresis	6.17	120.00	9.47e-9	14.23	4.57

Appendix I: Two Factor Experiments

Table 68 Increase in pooled demand and movement of pooled demand to apheresis two-factor experiment results

Experiment	Detections	Mean Time to Detection, Weeks	95% CI, Lower	95% CI, Upper
1	7.00	32.57	18.48	46.67
2	5.00	27.60	0.00	1.00
3	24.00	25.79	19.77	31.81
4	31.00	22.87	17.44	28.30
5	60.00	20.73	17.39	24.08
6	82.00	21.37	18.58	24.15
7	95.00	13.14	11.09	15.18
8	100.00	11.06	9.31	12.81
9	100.00	7.31	6.17	8.45
10	100.00	4.67	3.98	5.36
11	100.00	3.93	3.42	4.44
12	100.00	3.43	3.02	3.84
13	100.00	3.12	2.82	3.42
14	3.00	24.00	0.00	1.00
15	3.00	7.67	0.00	1.00
16	10.00	35.90	27.21	44.59
17	20.00	27.15	20.38	33.92
18	35.00	24.43	19.70	29.16
19	61.00	23.64	19.73	27.55
20	81.00	20.54	17.55	23.54
21	96.00	14.94	12.68	17.20
22	100.00	10.82	9.14	12.50
23	100.00	6.32	5.27	7.37
24	100.00	5.56	4.62	6.50
25	100.00	3.91	3.38	4.44
26	100.00	3.05	2.77	3.33
27	20.00	25.10	18.59	31.61
28	24.00	29.17	23.28	35.06
29	18.00	23.94	16.88	31.01
30	26.00	28.88	22.62	35.15
31	38.00	22.63	17.59	27.67
32	46.00	21.15	17.28	25.02
33	72.00	22.85	19.57	26.12
34	93.00	17.62	14.64	20.61
35	100.00	14.37	12.01	16.73
36	100.00	6.92	5.95	7.89

37	100.00	6.08	5.24	6.92
38	100.00	4.60	3.92	5.28
39	100.00	3.66	3.22	4.10
40	34.00	30.26	25.29	35.24
41	32.00	25.63	20.72	30.53
42	28.00	25.61	20.20	31.01
43	41.00	27.44	22.61	32.26
44	32.00	28.47	23.26	33.68
45	56.00	27.02	22.92	31.12
46	64.00	23.84	20.49	27.20
47	88.00	18.20	15.45	20.96
48	96.00	17.03	14.44	19.62
49	100.00	11.90	9.89	13.91
50	100.00	6.76	5.51	8.01
51	100.00	4.54	3.99	5.09
52	100.00	4.21	3.67	4.75
53	85.00	21.05	17.94	24.15
54	71.00	20.85	17.41	24.28
55	76.00	24.29	21.36	27.22
56	65.00	25.65	21.96	29.33
57	64.00	21.91	18.38	25.43
58	65.00	22.08	18.78	25.37
59	83.00	20.98	17.96	23.99
60	90.00	21.19	18.27	24.11
61	98.00	16.21	13.72	18.71
62	99.00	12.59	10.70	14.48
63	100.00	8.72	7.49	9.95
64	100.00	5.69	4.87	6.51
65	100.00	4.37	3.84	4.90
66	100.00	9.36	8.14	10.58
67	98.00	11.19	9.55	12.84
68	99.00	11.52	9.97	13.06
69	100.00	12.38	10.70	14.06
70	98.00	11.98	10.36	13.60
71	100.00	11.01	9.56	12.46
72	100.00	11.91	10.18	13.64
73	100.00	12.26	10.27	14.25
74	100.00	12.41	10.64	14.18
75	99.00	8.80	7.36	10.24
76	100.00	7.85	6.85	8.85
77	100.00	5.03	4.24	5.82
78	100.00	4.63	3.93	5.33
79	100.00	7.77	6.87	8.67

80	100.00	8.71	7.57	9.85
81	100.00	8.45	7.38	9.52
82	100.00	8.52	7.49	9.55
83	100.00	8.74	7.64	9.84
84	100.00	8.58	7.37	9.79
85	100.00	8.69	7.54	9.84
86	100.00	9.34	7.97	10.71
87	100.00	8.95	7.68	10.22
88	100.00	7.39	6.50	8.28
89	100.00	7.37	6.40	8.34
90	100.00	6.14	5.35	6.93
91	100.00	4.44	3.84	5.04
92	100.00	7.62	6.70	8.54
93	100.00	7.80	6.76	8.84
94	100.00	7.95	7.01	8.89
95	100.00	8.58	7.44	9.72
96	100.00	7.89	7.00	8.78
97	100.00	8.84	7.65	10.03
98	100.00	9.91	8.63	11.19
99	100.00	9.26	8.28	10.24
100	100.00	8.53	7.34	9.72
101	100.00	7.24	6.32	8.16
102	100.00	9.66	8.35	10.97
103	100.00	6.81	5.73	7.89
104	100.00	5.79	5.06	6.52
105	100.00	5.52	4.97	6.07
106	100.00	6.05	5.40	6.70
107	100.00	5.79	5.18	6.40
108	100.00	5.56	5.05	6.07
109	100.00	6.53	5.79	7.27
110	100.00	6.37	5.52	7.22
111	100.00	5.72	5.19	6.25
112	100.00	6.44	5.71	7.17
113	100.00	6.54	5.85	7.23
114	100.00	6.42	5.69	7.15
115	100.00	6.18	5.44	6.92
116	100.00	5.83	5.02	6.64
117	100.00	5.67	4.90	6.44
118	100.00	5.33	4.88	5.78
119	100.00	5.63	5.07	6.19
120	100.00	5.43	4.94	5.92
121	100.00	5.78	5.25	6.31
122	100.00	5.38	4.88	5.88

123	100.00	6.23	5.61	6.85
124	100.00	6.29	5.57	7.01
125	100.00	6.18	5.47	6.89
126	100.00	6.07	5.38	6.76
127	100.00	6.45	5.68	7.22
128	100.00	5.80	5.20	6.40
129	100.00	6.08	5.41	6.75
130	100.00	6.00	5.25	6.75
131	100.00	3.37	3.23	3.51
132	100.00	3.33	3.22	3.44
133	100.00	3.46	3.32	3.60
134	100.00	3.55	3.38	3.72
135	100.00	3.35	3.19	3.51
136	100.00	3.60	3.41	3.79
137	100.00	3.50	3.34	3.66
138	100.00	3.45	3.30	3.60
139	100.00	3.21	3.07	3.35
140	100.00	3.35	3.15	3.55
141	100.00	3.44	3.25	3.63
142	100.00	3.56	3.34	3.78
143	100.00	3.33	3.01	3.65
144	100.00	3.52	3.37	3.67
145	100.00	3.41	3.25	3.57
146	100.00	3.29	3.17	3.41
147	100.00	3.37	3.25	3.49
148	100.00	3.39	3.24	3.54
149	100.00	3.51	3.32	3.70
150	100.00	3.37	3.19	3.55
151	100.00	3.46	3.31	3.61
152	100.00	3.40	3.24	3.56
153	100.00	3.45	3.29	3.61
154	100.00	3.26	3.09	3.43
155	100.00	3.13	2.95	3.31
156	100.00	3.17	2.95	3.39
157	100.00	3.24	3.15	3.33
158	100.00	3.21	3.12	3.30
159	100.00	3.32	3.20	3.44
160	100.00	3.16	3.07	3.25
161	100.00	3.26	3.14	3.38
162	100.00	3.21	3.12	3.30
163	100.00	3.19	3.07	3.31
164	100.00	3.27	3.12	3.42
165	100.00	3.30	3.13	3.47

166	100.00	3.34	3.15	3.53
167	100.00	3.11	2.95	3.27
168	100.00	3.18	3.02	3.34
169	100.00	3.06	2.89	3.23

Table 69 Increase in pooled demand and movement of pooled demand to apheresis two-factor experiment results, with the inclusion of spline-based GAM forecasting

Experiment	Detections	Mean Time to Detection, Weeks	95% CI, Lower	95% CI, Upper
1	10	29.40	17.10	41.70
2	12	26.58	17.15	36.02
3	37	25.92	21.12	30.72
4	48	22.23	17.81	26.65
5	66	17.79	14.67	20.91
6	88	18.54	15.81	21.27
7	99	12.21	10.28	14.15
8	100	9.63	8.12	11.14
9	100	5.96	5.08	6.84
10	100	4.10	3.56	4.64
11	100	3.73	3.22	4.23
12	100	3.08	2.72	3.44
13	100	2.83	2.58	3.08
14	6	27.00	11.61	42.39
15	8	15.25	7.33	23.17
16	15	33.86	27.05	40.66
17	30	24.20	18.99	29.41
18	48	21.83	18.03	25.64
19	73	21.95	18.39	25.50
20	88	17.24	14.53	19.95
21	97	13.01	10.95	15.07
22	100	8.59	7.22	9.96
23	100	5.46	4.58	6.34
24	100	4.74	3.85	5.63
25	100	3.50	2.99	4.01
26	100	2.83	2.56	3.10
27	22	25.95	19.64	32.27
28	28	29.89	24.45	35.33
29	23	24.78	18.36	31.21
30	29	27.79	21.84	33.75
31	47	22.11	17.85	26.36
32	60	20.45	17.08	23.82
33	82	20.60	17.52	23.68

34	97	14.90	12.31	17.48
35	100	10.89	9.04	12.74
36	100	5.95	5.05	6.85
37	100	5.30	4.50	6.10
38	100	4.20	3.55	4.85
39	100	3.18	2.79	3.57
40	36	30.03	25.13	34.93
41	37	26.14	21.51	30.76
42	33	26.82	21.76	31.88
43	47	25.94	21.33	30.54
44	41	26.71	22.14	31.27
45	65	26.34	22.59	30.09
46	77	24.06	21.07	27.06
47	91	16.64	13.90	19.39
48	99	15.63	13.07	18.18
49	100	10.50	8.64	12.36
50	100	5.84	4.84	6.84
51	100	4.18	3.64	4.72
52	100	3.92	3.38	4.46
53	88	20.81	17.76	23.85
54	71	20.85	17.41	24.28
55	77	23.83	20.86	26.80
56	69	24.75	21.15	28.36
57	70	22.79	19.31	26.27
58	67	22.21	18.77	25.65
59	87	21.38	18.30	24.46
60	94	19.00	16.22	21.78
61	99	13.70	11.50	15.89
62	99	10.56	8.79	12.34
63	100	8.18	6.95	9.41
64	100	5.19	4.37	6.01
65	100	4.05	3.52	4.58
66	100	9.36	8.14	10.58
67	98	11.19	9.55	12.84
68	99	11.41	9.87	12.96
69	100	12.26	10.57	13.95
70	99	11.79	10.18	13.40
71	100	10.72	9.27	12.17
72	100	11.38	9.68	13.08
73	100	11.83	9.89	13.77
74	100	11.14	9.42	12.86
75	100	7.62	6.35	8.89
76	100	6.98	6.06	7.90

77	100	4.46	3.80	5.13
78	100	3.98	3.39	4.57
79	100	7.77	6.87	8.67
80	100	8.71	7.57	9.85
81	100	8.25	7.25	9.26
82	100	8.43	7.39	9.47
83	100	8.65	7.55	9.75
84	100	8.44	7.28	9.60
85	100	8.49	7.33	9.65
86	100	9.31	7.94	10.68
87	100	8.71	7.45	9.97
88	100	6.88	5.98	7.77
89	100	6.96	5.97	7.95
90	100	5.60	4.84	6.36
91	100	4.10	3.50	4.70
92	100	7.49	6.60	8.39
93	100	7.80	6.76	8.84
94	100	7.92	6.98	8.86
95	100	8.58	7.44	9.72
96	100	7.89	7.00	8.78
97	100	8.76	7.56	9.96
98	100	9.64	8.36	10.92
99	100	9.05	8.08	10.02
100	100	8.03	6.97	9.09
101	100	6.95	6.02	7.88
102	100	9.12	7.83	10.41
103	100	6.58	5.50	7.66
104	100	5.45	4.72	6.18
105	100	5.52	4.97	6.07
106	100	6.05	5.40	6.70
107	100	5.79	5.18	6.40
108	100	5.56	5.05	6.07
109	100	6.48	5.73	7.23
110	100	6.32	5.49	7.15
111	100	5.72	5.19	6.25
112	100	6.21	5.48	6.94
113	100	6.49	5.79	7.19
114	100	6.08	5.36	6.80
115	100	5.89	5.14	6.64
116	100	5.29	4.51	6.07
117	100	5.25	4.51	5.99
118	100	5.33	4.88	5.78
119	100	5.63	5.07	6.19

120	100	5.43	4.94	5.92
121	100	5.75	5.21	6.29
122	100	5.32	4.81	5.83
123	100	6.16	5.52	6.80
124	100	6.14	5.41	6.87
125	100	6.17	5.45	6.89
126	100	5.96	5.25	6.67
127	100	6.20	5.45	6.95
128	100	5.60	5.00	6.20
129	100	5.78	5.11	6.45
130	100	5.68	4.94	6.42
131	100	3.37	3.23	3.51
132	100	3.33	3.22	3.44
133	100	3.46	3.32	3.60
134	100	3.54	3.36	3.72
135	100	3.35	3.19	3.51
136	100	3.58	3.39	3.77
137	100	3.49	3.33	3.65
138	100	3.43	3.28	3.58
139	100	3.19	3.04	3.34
140	100	3.28	3.07	3.49
141	100	3.34	3.13	3.55
142	100	3.43	3.20	3.66
143	100	3.06	2.80	3.32
144	100	3.52	3.37	3.67
145	100	3.41	3.25	3.57
146	100	3.29	3.17	3.41
147	100	3.36	3.24	3.47
148	100	3.39	3.23	3.55
149	100	3.50	3.31	3.69
150	100	3.37	3.19	3.55
151	100	3.40	3.24	3.56
152	100	3.39	3.22	3.56
153	100	3.37	3.19	3.55
154	100	3.21	3.04	3.38
155	100	3.11	2.93	3.29
156	100	3.08	2.86	3.31
157	100	3.24	3.15	3.33
158	100	3.21	3.12	3.30
159	100	3.32	3.20	3.44
160	100	3.16	3.07	3.25
161	100	3.25	3.14	3.36
162	100	3.19	3.09	3.29

163	100	3.19	3.07	3.31
164	100	3.24	3.09	3.39
165	100	3.27	3.09	3.45
166	100	3.31	3.11	3.51
167	100	3.05	2.88	3.22
168	100	3.07	2.91	3.24
169	100	2.98	2.80	3.16

Table 70 t-test comparison of two-factor experiment detection times with and without forecasting

Experiment	t-value	Degrees of Freedom	p-value	Mean of the Differences, weeks	Mean Difference 95% CI Halfwidth, weeks
Two-Factor Increase	3.45	81.00	6.62e-8	4.50	2.31

Table 71 Chi-square test comparison of two-factor experiment detection rates with and without forecasting

Experiment	X^2	Degrees of Freedom	p-value
Two-Factor Increase	27.80	168	0.99

Appendix J: 2^k Factorial Experiment Design Tables

2^k Factorial Experiment Design Tables

1. Factor 1: Increase in pooled demand
2. Factor 2: Increase in apheresis demand
3. Factor 3: Pooled demand moves to apheresis
4. Factor 4: Rate of change
5. Factor 5: Change distribution

Table 72 Factorial level factor explanations

Level	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
-	5%	5%	5%	Step	Deterministic
+	15%	15%	15%	Linear	Stochastic

Table 73 Factorial level experiment design table sigs correspond to the level in the previous table

Design Point	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
1	-	-	-	-	-
2	+	-	-	-	-
3	-	+	-	-	-
4	+	+	-	-	-
5	-	-	+	-	-
6	+	-	+	-	-
7	-	+	+	-	-
8	+	+	+	-	-
9	-	-	-	+	-
10	+	-	-	+	-
11	-	+	-	+	-
12	+	+	-	+	-
13	-	-	+	+	-
14	+	-	+	+	-
15	-	+	+	+	-
16	+	+	+	+	-
17	-	-	-	-	+
18	+	-	-	-	+
19	-	+	-	-	+
20	+	+	-	-	+
21	-	-	+	-	+
22	+	-	+	-	+
23	-	+	+	-	+
24	+	+	+	-	+
25	-	-	-	+	+
26	+	-	-	+	+
27	-	+	-	+	+
28	+	+	-	+	+
29	-	-	+	+	+
30	+	-	+	+	+
31	-	+	+	+	+
32	+	+	+	+	+

Table 74 2k factorial experiment results, including forecasting

Design Point	Percentage Detections	Mean Time to Detection, Weeks	95% CI, Lower	95% CI, Upper
1	42	23.98	20.22	27.73
2	92	15.88	13.30	18.45
3	100	12.33	10.70	13.96
4	100	7.88	6.59	9.17
5	100	7.29	6.39	8.19
6	100	8.17	7.09	9.25
7	100	5.72	5.12	6.32
8	100	6.70	5.95	7.45
9	37	26.08	21.35	30.81
10	93	15.75	13.47	18.04
11	100	13.79	11.80	15.78
12	100	10.04	8.54	11.54
13	100	10.41	8.96	11.86
14	100	9.12	7.88	10.36
15	100	6.68	6.03	7.33
16	100	6.45	5.79	7.11
17	73	20.99	17.72	24.26
18	100	11.88	10.03	13.73
19	97	12.97	10.86	15.08
20	100	8.07	6.82	9.32
21	100	6.32	5.56	7.08
22	100	6.47	5.60	7.34
23	100	4.62	4.27	4.97
24	100	4.35	3.95	4.75
25	69	22.76	19.33	26.20
26	96	11.64	9.94	13.34
27	99	14.54	12.55	16.52
28	100	10.05	8.45	11.65
29	100	7.52	6.67	8.37
30	100	7.27	6.35	8.20
31	100	5.46	5.06	5.87
32	100	5.09	4.70	5.48

Table 75 Primary effect of each demand factor on the rate of detection, forecasting included

Factor	Effect on Detection Rate, %	t-value	p-value	95% Confidence Interval Halfwidth	Significance
E_1	33.00	4.75	2.18e-4	14.74	Yes
E_2	38.00	5.47	5.16e-5	14.74	Yes
E_3	38.75	5.58	4.18e-5	14.74	Yes
E_4	-2.25	-0.325	0.75	14.74	No
E_5	16.75	2.41	0.03	14.74	Yes

Table 76 Secondary effects of each demand factor on the rate of detection, forecasting included. If two factors are present at the same time, the primary effect of each is summed with the secondary effect

Factor	Effect on Detection Rate, %	t-value	p-value	95% Confidence Interval Halfwidth	Significance
E_{12}	-19.50	-3.137	6.36e-3	13.33	Yes
E_{13}	-20.50	-3.298	4.54e-3	13.33	Yes
E_{14}	0.50	0.08	0.94	13.33	No
E_{15}	-6.00	-0.965	0.35	13.33	No
E_{23}	-24.25	-3.90	1.27e-3	13.33	Yes
E_{24}	1.750	0.28	0.78	13.33	No
E_{25}	-9.75	-1.57	0.14	13.33	No
E_{34}	1.25	0.20	0.84	13.33	No
E_{35}	-8.75	-1.41	0.18	13.33	No
E_{45}	-0.25	-0.04	0.97	13.33	No

The intercept values for the detection rate, and time to detection, models are 52.94, and 23.08, respectively. Interaction terms are summed when each effect is present. For instance, if Factor 1 and Factor 2 were present the mean detection time would be the sum of the intercept value, the primary effect of each factor, and the effect the two factors have on each other.

Table 77 Primary effect of each demand factor on time to detection, forecasting included

Factor	Effect on Detection Time, Weeks	t-value	p-value	95% Confidence Interval Halfwidth	Significance
E_1	-7.69	-7.81	7.55e-7	2.09	Yes
E_2	-9.85	-10.00	2.76e-8	2.09	Yes
E_3	-14.10	-14.32	1.53e-10	2.09	Yes
E_4	1.89	1.91	0.07	2.09	No
E_5	-2.33	-2.37	0.03	2.09	Yes

Table 78 Secondary effects of each demand factor on time to detection, forecasting included. If two factors are present at the same time, the primary effect of each is summed with the secondary effect

Factor	Effect on Detection Time, Weeks	t-value	p-value	95% Confidence Interval Halfwidth	Significance
E_{12}	2.71	3.08	7.22e-3	1.87	Yes
E_{13}	6.98	7.92	6.28e-7	1.87	Yes
E_{14}	-0.87	-1.00	0.33	1.87	No
E_{15}	-0.51	-0.58	0.57	1.87	No
E_{23}	5.22	5.93	2.12e-5	1.87	Yes
E_{24}	-0.01	-0.02	0.99	1.87	No
E_{25}	2.17	2.47	0.03	1.87	No
E_{34}	-0.29	-0.33	0.75	1.87	No
E_{35}	-0.08	-0.09	0.93	1.87	No
E_{45}	-0.21	-0.24	0.81	1.87	No

Table 79 Standard 2k factorial experiment results

Design Point	Percentage Detections	Mean Time to Detection, Weeks	95% CI, Lower	95% CI, Upper
1	40	23.63	19.80	27.46
2	84	18.25	15.46	21.04
3	100	12.84	11.05	14.63
4	100	8.50	7.24	9.76
5	100	7.29	6.39	8.19
6	100	8.32	7.25	9.39
7	100	5.78	5.19	6.37
8	100	6.72	5.97	7.47
9	34	25.35	20.40	30.31
10	87	18.06	15.49	20.63
11	100	14.07	12.13	16.01
12	100	10.51	9.03	12.00
13	100	10.41	8.96	11.86
14	100	9.73	8.47	10.99
15	100	6.68	6.03	7.33
16	100	6.70	6.03	7.37
17	66	21.88	18.43	25.32
18	100	12.69	10.80	14.58
19	96	13.00	10.95	15.05
20	100	8.86	7.41	10.31
21	100	6.35	5.60	7.10
22	100	6.77	5.79	7.76
23	100	4.65	4.30	5.00
24	100	4.44	4.04	4.84
25	64	24.30	20.80	27.80
26	94	12.84	11.05	14.63
27	99	15.75	13.55	17.95
28	100	10.81	9.21	12.41
29	100	7.62	6.77	8.47
30	100	7.47	6.56	8.38
31	100	5.48	5.08	5.88
32	100	5.20	4.82	5.58

Table 80 Primary effect of each demand factor on the rate of detection

Factor	Effect on Detection Rate, %	t-value	p-value	95% Confidence Interval Halfwidth	Significance
E_1	32.13	4.83	1.85e-4	14.10	Yes
E_2	42.38	6.37	9.27e-6	14.10	Yes
E_3	43.38	6.52	7.04e-6	14.10	Yes
E_4	-1.88	-0.28	0.78	14.10	No
E_5	16.38	2.46	0.03	14.10	Yes

Table 81 Secondary effects of each demand factor on the rate of detection. If two factors are present at the same time, the primary effect of each is summed with the secondary effect

Factor	Effect on Detection Rate, %	t-value	p-value	95% Confidence Interval Halfwidth	Significance
E_{12}	-19.50	-3.28	4.73e-3	12.61	Yes
E_{13}	-20.75	-3.49	3.03e-3	12.61	Yes
E_{14}	0.25	0.04	0.97	12.61	No
E_{15}	-3.50	-0.59	0.56	12.61	No
E_{23}	-28.25	-4.75	2.18e-4	12.61	Yes
E_{24}	1.75	0.29	0.77	12.61	No
E_{25}	-10.50	-1.77	0.10	12.61	No
E_{34}	1.00	0.17	0.87	12.61	No
E_{35}	-9.25	-1.56	0.14	12.61	No
E_{45}	-0.25	-0.04	0.97	12.61	No

Table 82 Primary effect of each demand factor on time to detection

Factor	Effect on Detection Time, Weeks	t-value	p-value	95% Confidence Interval Halfwidth	Significance
E_1	-6.16	-6.23	1.19e-5	1.82	Yes
E_2	-9.93	-10.04	2.60e-8	1.82	Yes
E_3	-14.49	-14.65	1.08e-10	1.82	Yes
E_4	1.94	1.97	0.07	1.82	No
E_5	-1.85	-1.87	0.04	1.82	Yes

Table 83 Secondary effects of each demand factor on time to detection. If two factors are present at the same time, the primary effect of each is summed with the secondary effect

Factor	Effect on Detection Time, Weeks	t-value	p-value	95% Confidence Interval Halfwidth	Significance
E_{12}	2.02	2.29	0.03	1.88	Yes
E_{13}	6.42	7.26	1.90e-6	1.88	Yes
E_{14}	-0.93	-1.06	0.31	1.88	No
E_{15}	-1.34	-1.51	0.15	1.88	No
E_{23}	5.54	6.27	1.12e-5	1.88	Yes
E_{24}	-0.02	-0.03	0.98	1.88	No
E_{25}	2.19	2.48	0.03	1.88	Yes
E_{34}	-0.38	-0.43	0.67	1.88	No
E_{35}	-0.32	-0.36	0.72	1.88	No
E_{45}	0.08	0.09	0.93	1.88	No

The intercept values for the detection rate, and time to detection, models are 48.69, and 23.21, respectively.

Table 84 t-test comparison of mean time to detections for 2k factorial experiments, with and without forecasting

Experiment	t-value	Degrees of Freedom	p-value	Mean of the Differences, weeks	Mean Difference 95% CI Halfwidth, weeks
2^k Factorial Experiments	11.13	497	2.20e-16	11.84	2.09

Table 85 Chi square comparison of detection rate for 2k factorial experiments, with and without forecasting

Experiment	X^2	Degrees of Freedom	p-value
2^k Factorial Experiments	1.11	31	0.99

Appendix K: Magnitude Estimation

Table 86 Estimate of demand change size based on inventory monitoring for increase in pooled demand experiments

Increase in Pooled Demand, %	Mean Change Estimate, Abs. units	Estimate 95% CI Halfwidth, Abs. units	Mean Change Estimate, Relative %	Estimate 95% CI Halfwidth, Relative %
2	9.75	23.83	9.26	22.62
4	14.43	22.69	13.69	21.54
6	16.59	22.20	15.75	21.08
8	17.67	22.08	16.77	20.96
10	19.79	21.05	18.79	19.98
12	22.21	20.87	21.08	19.81
14	22.22	21.68	21.10	20.59
16	21.28	21.71	20.21	20.61
18	21.05	21.35	19.99	20.27
20	20.09	21.97	19.08	20.86

Table 87 Estimates of demand change size based on inventory monitoring for increase in apheresis demand experiments

Increase in Apheresis Demand, %	Mean Change Estimate, Abs. units	Estimate 95% CI Halfwidth, Abs. units	Mean Change Estimate, Relative %	Estimate 95% CI Halfwidth, Relative %
2	5.70	10.32	14.86	26.91
4	4.82	12.93	12.58	33.72
6	6.81	8.75	17.75	22.82
8	11.10	10.37	28.93	27.04
10	13.36	9.83	34.84	25.64
12	12.71	10.61	33.12	27.66
14	14.39	10.15	37.52	26.48
16	14.65	10.36	38.18	26.99
18	17.06	9.69	44.47	25.27
20	17.06	10.19	44.46	26.55

Table 88 Estimates of demand change size based on inventory monitoring for movement of pooled demand to apheresis experiments

Movement of Pooled Demand to Apheresis, %	Mean Change Estimate, Abs. units	Estimate 95% CI Halfwidth, Abs. units	Mean Change Estimate, Relative %	Estimate 95% CI Halfwidth, Relative %
2	10.43	23.73	9.90	22.52
4	10.70	19.21	10.15	18.23
6	10.59	15.86	10.05	15.05
8	12.97	16.05	12.32	15.23
10	14.52	15.48	13.79	14.70
12	13.88	15.30	13.18	14.52
14	13.83	16.03	13.13	15.22
16	13.44	17.28	12.76	16.41
18	13.64	16.98	12.95	16.12
20	9.96	20.08	9.45	19.06