ABSTRACT

TIPI, EVREN. Forecasting Sales of Slow-moving Items. (Under the direction of Dr. Russell E. King and Dr. Thom J. Hodgson.)

Inventory management is very difficult if the sales are slow-moving, where demand appears at random and with many time periods having no demand at all. Even though slow-moving demand is more common in the service parts business, it is also a difficult problem in the retail industry.

Exponential smoothing methods and the approach of Croston (1972) and its modifications are very popular to estimate the sales of slow-moving items. However, exponential smoothing generally leads to improper stock levels since more weight is placed on the most recent data, making the forecast highest just after a demand occurrence, and lowest just before the demand occurs again. Croston proposed an alternative method that considers not only the demand size but also the inter-arrival time between demands to overcome this situation. Croston's method and its modifications adjust the forecast only after a demand is observed, but this may not be very practical if the inter-arrival times between demands are very long. In addition, both exponential smoothing and Croston do not work well if the demand is seasonal.

In this research, we add a seasonality index to the existing exponential smoothing and Croston's modification methods. In addition, a correction factor is included that adjusts the forecast if the forecast error is more than a threshold value. Moreover, in some of our new models, we also adjust the forecast not only after a demand, like Croston, but also if the time

since the last demand exceeds the average inter-arrival time or two times the average inter-arrival time depending on the new method.

These forecasting methods are compared using sales forecast accuracy metrics and inventory control measures. Whether the goal is to have an accurate sales forecast or to optimize inventory control measures, our methods outperform the existing exponential smoothing methods, as well as the approach of Croston and its modifications.

Forecasting Sales of Slow-moving Items

by
Evren Tipi

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Industrial Engineering

Raleigh, North Carolina

2011

APPROVED BY:

Dr. Russell E. King Chair of Advisory Committee	Dr. Thom J. Hodgson Co-chair of Advisory Committee
Dr. James R. Wilson	Dr. Peter Bloomfield

Dr. Donald P. Warsing

DEDICATION

To my parents Nermin and Ersen Tipi,

who have always loved me and

encouraged me in my studies.

BIOGRAPHY

Evren Tipi was born in Kastamonu, Turkey, on February 20, 1980. For her secondary and high school education, she attended Tokat Anatolian High School and Ankara Science High School. She received her Bachelor of Science in Industrial Engineering degree from Bilkent University in 2002 with a full scholarship. Upon graduation, she began her graduate study in Industrial and Systems Engineering at North Carolina State University. She worked as a teaching assistant and research assistant in the Industrial and Systems Engineering Department. In 2004 she was selected for the "Outstanding T/A in ISE Department Award" and also received the "Mentored Teaching Assistantship Award". She received her M.I.E. in Fall 2004 and started her Ph.D. education. While continuing her education, she worked as a customer service engineer and fabrication quality engineer at a manufacturing company for two years. In addition, she also worked for one of the biggest home improvement retail companies for four years as a forecast analyst, senior project forecast analyst, network optimization planning manager, and research manager.

Her research interests include forecasting, production and inventory control, and supply chain management.

ACKNOWLEDGMENTS

I would like to extend my most sincere thanks to Dr. Russell E. King for all his help, support, guidance, trust, and advice during my Ph.D. education and also for meeting me on many weekends. I would also like to thank Dr. Thom Hodgson, for his valuable comments regarding my dissertation and encouragement through the years. I have learnt a lot from both of them. Furthermore, I also would like to thank Dr. James R. Wilson, Dr. Peter Bloomfield, and Dr. Don Warsing for thoughtful comments, guidance, and serving on the committee.

I would like to thank my parents Nermin and Ersen Tipi, and my sister Devrim Tipi Urhan for all their encouragement, support, love, and believing in me all my life. I could not complete my studies if you were not there. Furthermore, I would like to also thank my husband A. Bugrahan Akben for driving me from Charlotte to Raleigh to meet my advisors on the weekends. And lastly, I say to my little baby boy Kenan: Welcome to our life!

My thanks also go to all my friends both in Charlotte and Raleigh, in particular Nihal Gulseren Cakmakci, Ozlem Karacali Cete, and Canguzel Zulfikar for their invaluable friendship, prayers, and for bringing joy to my life.

I must express my thanks to Logistics and Distribution EVP Joseph Michael Mabry for letting me use the company’s data, and my directors Jeff Ziegler and Rob Glynn for their support. Moreover, I also thank my co-workers in forecasting department, especially Amy Symon.

Finally, I am thankful to God to let me keep my promise to him to complete my studies.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 Introduction

Demand forecasting forms the starting point for inventory planning and is one of the most critical concerns of inventory management. It is even more difficult if the demand is intermittent, where demand appears at random with many time periods having no demand at all. In addition, if a demand occurs, it can be more than one unit. Even though intermittent demand is more common in the service parts business, it is also a challenging problem in the retail industry.

In practice, single exponential smoothing (SES) is often used to forecast intermittent demand. However, this method generally leads to improper stock levels since more weight is placed on the most recent data, making the forecast highest just after a demand occurrence, and lowest just before the demand occurs again. To overcome this situation, Croston (1972) proposed an alternative method that considers not only the demand size but also the inter-arrival time between demands.

Croston's method is a benchmark in forecasting intermittent demand; thus, it has been assessed by several authors in the literature. Willemain *et al.* (1994) compared Croston's method with SES and concluded that Croston's method is robustly superior to SES. In addition, Johnston and Boylan (1996) also showed that Croston's method is always better than SES when the average inter-arrival time between demands is greater than 1.25 review intervals. However, Syntetos and Boylan (2001) showed that the original Croston method is biased, and they corrected it by multiplying the forecast for the demand per period with $1-\alpha/2$ (Syntetos and Boylan, 2005), where α is the smoothing constant. Levén and Segerstedt

(2004) use the Croston approach of only updating when there is a positive demand, but instead of updating the forecast of demand and inter-arrival time between arrivals separately, they updated the forecast for the demand per period directly using the ratio of demand size and interval. However, later in 2007 Boylan and Syntetos showed that Levén and Segerstedt's method is biased.

In addition to Croston's method and its modifications, there is another promising class of methods known as bootstrapping that is used to forecast intermittent demand. The main advantage of bootstrapping is that the mean and variance of the lead time demand distribution is forecasted directly by repeated sampling from realized demands. The main disadvantage is that this method is rather complex to implement and understand. Bootstrapping has been proposed in several papers including Snyder (2002), Smart and Willemain (2000), Smart (2002), and Willemain *et al*. (2004). In Willemain *et al*. (2004), the data from nine industrial companies were used to show that the bootstrapping method produces more accurate forecasts of the distribution of demand over a fixed lead time than do SES and Croston's method.

Even though there have been papers in the literature on modifying the Croston algorithm, very few of them investigated the "trend" in the data. One of the exceptions is in Altay *et al*. (2008). Wright's modification of Holt's method (WMH) has been adapted to forecast intermittent demand. In Altay *et al*.'s study, they have compared the performance of Syntetos and Boylan's modification to the Croston method (SB2) and WMH using a simulated environment as well as a real data set of aircraft parts demand. When testing the real data set, they observed that SB2 carries lower inventories than WHM at the expense of service level and that WMH provides superior service levels in every case with no significant total cost difference.

Another challenge is to find the appropriate accuracy measure for intermittent demand. All relative-to-the-series accuracy measures (*MAPE* (Mean Absolute Percentage Error) and *MdAPE* (Median Absolute Percentage Error)) must be excluded, since actual demand is used in the denominator of the calculation and can be zero in some periods, which will lead to division by 0. Syntetos and Boylan (2005) put accuracy measures of intermittent demand into two groups, absolute accuracy measures and accuracy measures relative to another method. In addition, Hoover (2006) analyzed other accuracy metrics for slow-moving items such as: denominator-adjusted *MAPE,* the symmetric *MAPE,* and ratio of mean absolute error to mean demand.

In this research, we add a seasonality index to the existing SES and Croston's modification methods. In addition, a correction factor (*CF*) is included that adjusts the forecast if the forecast error is more than a threshold value. Moreover, Croston's method or its modifications adjust the forecast only after a demand, but this may not be very practical if the inter-arrival times between demands are very high. In our approach, we also adjust the forecast not only after a demand, but also if the time since the last demand exceeds the average inter-arrival time or two times the average inter-arrival time, depending on the new method.

The remainder of this dissertation is organized as follows. In Chapter 2 the literature review is provided. In Chapter 3 our proposed forecasting methods for slow-moving items are introduced. Sales forecast accuracy measures used for slow-moving items are presented in Chapter 4. The experimental structure and sales forecast accuracy results are explained in Chapters 5 and 6. In addition to sales forecast accuracy metrics, inventory control financial metrics results are presented in Chapter 7. Conclusions and future work are presented in Chapter 8. Furthermore, Appendix A has the Mean Error, Mean Absolute Error, Scaled Mean Absolute Error, and Relative Geometric Root Mean Square Error paired-sample *t*-test results; and Appendix B has the Confidence Interval figures of inventory control financial metrics for *CF* parameter set combinations.

2 Literature Review

Several techniques for forecasting slow-moving items have been suggested in the literature. In this chapter we review these approaches.

2.1 Simple Exponential Smoothing

Croston (1972) shows that simple exponential smoothing is not appropriate for forecasting intermittent demand due to its bias. He uses two models to demonstrate this.

His first model assumes that demand is a deterministic μ units that occurs every p review intervals. If the first demand occurs at time t=1, then the demand at time t, Y_t, is given by:

$$Y_t = \begin{cases} \mu, & \text{if } (t-1)/p = \text{nonnegative integer,} \\ 0, & \text{otherwise.} \end{cases}$$

for $p \geq 1$.

If we forecast using the basic SES formula, then the forecast of demand for period $t+1$ that is made at the end of period t, Y_t', is calculated as:

$$Y_t' = Y_{t-1}' + \alpha e_t = \alpha Y_t + (1-\alpha) Y_{t-1}',$$

where, α is the smoothing coefficient $0 \leq \alpha \leq 1$, and $e_t = Y_t - Y_{t-1}'$ is the forecast error in period t. For simplicity, we assume that the forecast lead time or lag is zero meaning that we forecast the next period's demand.

If we update the forecast only when the demand occurs, then the expected value of the estimate is not equal to $\frac{\mu}{p}$; instead, it is:

$$E(Y_t') = \frac{\mu\alpha}{1-(1-\alpha)^p}.$$

In his second model, Croston (1972) uses a stochastic model of inter-arrival time and demand size. The demand sizes are independently distributed from a normal distribution $N(\mu,\sigma^2)$, and in every review period, the demand independently occurs with probability *1/p*. Under these assumptions, the expected demand per unit time period is:

$$E(Y_t) = \frac{\mu}{p} .$$

Croston showed that if the demand estimates are updated every period with SES, the mean and variance are as follows:

$$E\left(Y_t'\right) = \frac{\mu}{p},$$

$$V\left(Y_t'\right) = \frac{\alpha}{2-\alpha}\left(\frac{p-1}{p^2}\mu^2 + \frac{\sigma^2}{p}\right).$$

However, if this method is used for stock replenishment, most of the time the replenishment will be triggered by a demand which recently occurred in the last review interval. Thus, the estimates that are made after a demand occurs will have a biased expected value, which is:

$$E\left(Y_t'\right) = \mu(\alpha + (1-\alpha)/p).$$

The bias is equal to $\alpha(p-1)$ times the average demand.

SES is easy to calculate, but the forecasts are biased and sales promotions and the seasonality of the demand are not taken into account.

2.2 Poisson Method

In addition to SES, the Poisson distribution can be used to forecast intermittent demand (Johnston and Boylan (1996)). A random sample from a Poisson-distributed random variable with mean λ_t' can be used to estimate the demand. This method is described below.

$$\lambda_t' = \alpha Y_t + (1-\alpha)\lambda_{t-1}',$$

$$Y_t' \leftarrow \text{Poisson}(\lambda_t'),$$

where

Y_t' is the forecasted demand for period *t*+1 that is made at the end of period *t*,

Y_t is the demand size (actual demand) at the end of period *t*,

α is smoothing constant, and

λ_t' is an estimate of the expected demand for period *t*+1 that is made at the end of period *t*, and λ_0' is calculated dividing the total demand in the warm-up period by the number of periods in the warm-up period.

Similar to SES, the forecasts generated with the Poisson method are easy to calculate, but sales promotions and the seasonality of the demand are not taken into account. In addition, since the forecasts are generated using a Poisson distribution, different runs can produce different forecasts, which might not be desirable.

2.3 Croston's Method

Croston (1972) developed the benchmark for forecasting intermittent demand. He showed that SES is not appropriate for forecasting intermittent demand due to its bias, and suggested a new method. In his method, separate SES estimates of the average size of demand and the average time between demands are calculated each time a demand occurs. If there is no demand, then no changes are made to the estimates. His method is described below.

The following notation is used.

z_t' : The demand size estimate for period *t*+1 that is made at the end of period *t*,

p_t' : The average inter-arrival time period estimate between the non-zero demands for period *t*+1 that is made at the end of period *t*,

Y_t' : The forecasted demand for period *t*+1 that is made at the end of period *t*, and is calculated by dividing z_t' by p_t',

Y_t : The demand size (actual demand) at the end of period *t*,

α : The smoothing constant,

q: The time difference between period *t* and the time of the occurrence of the previous demand.

Croston's method is summarized below:

IF $Y_t = 0$ then

$z'_t = z'_{t-1}$

$p'_t = p'_{t-1}$

$q = q + 1$

ELSE

$z'_t = z'_{t-1} + \alpha(Y_t - z'_{t-1})$

$p'_t = p'_{t-1} + \alpha(q - p'_{t-1})$

$q = 1$

ENDIF

The forecast is then calculated as:

$$Y'_t = \frac{z'_t}{p'_t}.$$

The values z'_0 and p'_0 can be initialized using the historical data of the warm-up period by averaging the non-zero demand sizes and the inter-arrival times and q is set equal to zero initially.

According to Croston, using this model the expected estimate of demand per period is:

$$E(Y'_t) = E(\frac{z'_t}{p'_t}) = \frac{E(z'_t)}{E(p'_t)} = \frac{\mu}{p}.$$

Although his method is widely accepted and very well-known, Rao (1973) corrected the algebraic errors in Croston's error estimators. In addition, Syntetos and Boylan (2001) showed that Croston's method is positively biased (2001),

$$E(\frac{z'_t}{p'_t}) = E(z'_t)E(\frac{1}{p'_t}) \text{ and } E(\frac{1}{p'_t}) > \frac{1}{E(p'_t)}.$$

The expected demand per time period for $\alpha = 1$ is:

$$E(\frac{z_t'}{p_t'}) = E(z_t')E(\frac{1}{p_t'}) = \mu\left(-\frac{1}{p-1}\log(\frac{1}{p})\right) \text{ and } E(p_t') = E(p_t) = p.$$

The maximum bias is attained when $\alpha = 1$ and is equal to:

$$\mu\left(-\frac{1}{p-1}\log(\frac{1}{p})\right) - \frac{\mu}{p}.$$

If p_t' and $\frac{z_t'}{p_t'}$ were independent, then $E(Y_t')$ would be equal to $\frac{E(z_t')}{E(p_t')}$. Thus, different probabilistic assumptions will lead to different results.

Croston's method is easy to calculate; forecasts are updated only when demand occurs, which is not desirable in some cases such as the following: if the product suddenly stops selling, then the forecast should be lowered. In addition, this method is biased, and sales promotions and seasonality are not taken into account.

2.4 Syntetos and Boylan Methods 1 and 2 (SB1 and SB2)

Syntetos and Boylan (2001) showed that the Croston's original method is biased and suggested a method similar to Croston's that is unbiased (SB1) such that:

$$Y_t' = z_t' \frac{1}{p_t' c^{p_t'-1}} \text{ and } c > 100.$$

Similar to Croston's method, separate SES estimates of the average size of demand and $1/(p_t' c^{p_t'-1})$ are calculated each time a demand occurs. If there is no demand, then no changes are made to the estimates. Theoretically, if c is infinitely large ($c > 100$ is sufficient), then this estimate is unbiased. This method has very small bias since c cannot be infinite, and performs better than Croston for $\alpha > 0.15$ (Syntetos and Boylan, 2001). Syntetos and Boylan (2005) suggested another similar method to Croston's which is unbiased (SB2), and the forecast is calculated as below:

$$Y_t' = (1 - \frac{\alpha}{2}) \frac{z_t'}{p_t'} .$$

As before, separate SES estimates of the average size of demand and the average time between demands are calculated each time a demand occurs. If there is no demand, no changes are made to the estimates.

Using 3,000 observations of real intermittent demand data from the automotive industry, Syntetos and Boylan (2005) compared their new method with three other methods: a simple moving average with 13 periods, SES, and Croston's method. Their new method performed better than the three other methods on most of the error measures considered in their paper. In addition, in Syntetos and Boylan (2006) they used the same data and the same methods and concluded that SB2 performs better than the others in terms of stock control performance.

Based on the findings of Syntetos and Boylan, SB2 is unbiased and works better than Croston on most of the error measures. However, similar to Croston's method, seasonality or the impact of the sales promotions are not taken into account; and the forecasts are updated only when demand occurs, which is not desirable in some cases such as the following: if the product suddenly stops selling, then the forecast should be lowered.

2.5 Levén and Segerstedt Method (LSM)

Levén and Segerstedt (2004) suggested a new method to forecast intermittent demand. The new estimate is as follows:

$$Y'_n = Y'_{n-1} + \alpha\left(\frac{z_n}{T_n - T_{n-1}} - Y'_{n-1}\right),$$

where n is an index counting the periods in which demand occurs, Y_n is the measured demand size during the n^{th} period in which demand occurs, and T_n is the time period in which the quantity z_n is demanded. Thus, Y'_n is the forecasted mean demand rate per period calculated at the end of period T_n. Similar to Croston (1972), the forecasts are updated at the ends of periods in which demand occurs.

Levén and Segerstedt (2004) used an Erlang distribution to fit the demand data; and according to their simulation studies where these forecasts are used to manage inventory, the new method tends to yield fewer inventory shortages than a system based on exponential smoothing. However, Boylan and Syntetos (2007) showed that the LSM estimator is biased.

The method of Levén and Segerstedt (2004) works better than SES, and is computationally simpler than Croston's method from a calculation perspective. However, it is more biased than Croston's method, and promotions and seasonality are not taken into account. Similar to the Croston, SB1, and SB2 methods the forecast is updated only after a demand occurs.

2.6 Adaptation of Wright's Modification of Holt's Method

In a recent study by Altay *et al*. (2008), Wright's modification of Holt's double exponential smoothing is adapted to forecast intermittent demand by taking into consideration the trend. The results indicate that firms focusing on minimizing inventory levels as a priority should consider forecasting using the SB2 method. If their priority is high customer service, then the modified Holt's method is superior. The advantage of this method is that it has a trend component. However, similar to SES, Croston and SB2, promotions and seasonality are not taken into account. Furthermore, the forecast is updated only after a sale.

2.7 Bootstrapping and Other Methods

There are other methods such as bootstrapping or neural networks that are used to forecast intermittent demand. Bootstrapping has been proposed in Snyder (2002), Smart and Willemain (2000), Smart (2002), and Willemain *et al*. (2004). In Willemain *et al*. (2004), the data from nine industrial companies was used to show that the bootstrapping method produces more accurate forecasts of the distribution of demand over a fixed lead time than do SES and Croston's method. The main advantage of bootstrapping is that the mean and variance of the lead time demand distribution are forecasted directly by repeated sampling from realized demands. The main disadvantage of this method is that it is rather complex to understand and implement compared with Croston's method and its modifications, since bootstrapping requires development of a Markov model and simulated experimentation.

Hua *et al*. (2007) developed a new approach for forecasting the intermittent demand of spare parts. Using data sets of forty different kinds of spare parts from a petrochemical enterprise in China, they show that their method produces more accurate forecasts of lead time demand than do SES, Croston's method, and the Markov bootstrapping method. However, this method is very complex to understand and implement, and promotions and seasonality are not taken into account.

There have been some other recent papers about forecasting intermittent demand. Neural networks are used to forecast lumpy demand in Gutierrez *et al*. (2008). Syntetos *et al*. (2009) is the first published evidence that judgmental adjustments to intermittent demand can be effective. In addition, the issues related to forecasting and inventory management of a wholesale company that deals with intermittent engineering supplies is addressed in Syntetos *et al.* (2010). For a broad review of supply chain forecasting, readers should refer to Fildes *et al.* (2008).

2.8 Summary of Main Methods for Forecasting Slow-moving Items

Below is a summary of the characteristics of the existing forecasting methods used to forecast slow-moving items' sales.

Table 2.8.1 Summary of the existing slow-moving items' forecasting methods

Forecasting Method	Easy to Calculate?	Biased?	Has trend component?	Has Seasonality/ Promotions?	Forecast Update Frequency	Reference
SES	Y	Y	N	N	Every period	Croston (1972)
Croston	Y	Y	N	N	After a sale	Croston (1972)
SB1	Y	Y	N	N	After a sale	Syntetos and Boylan (2001)
SB2	Y	N	N	N	After a sale	Syntetos and Boylan (2005)
LSM	Y	Y	N	N	After a sale	Levén and Segerstedt (2004)
Bootstrapping	N	N	N	N	Every period	Willemain *et al*. (2004).
Adaptation of Wright's modification of Holt's method	N	N	Y	N	After a sale	Altay *et al*. (2008)

3 Forecasting Methods Compared

In this chapter we motivate the need for improved forecasting techniques for items with slow-moving sales and then describe the features we use to develop four new approaches.

3.1 Motivation

In order to demonstrate the problems with the existing techniques in the literature, we evaluate their performance on aggregated weekly U.S. sales across stores over a one-year period for a dryer (Item A) and a dishwasher (Item B) at Lowe’s Home Improvement Company.

Croston’s method or its modifications adjusts the forecast only after a demand is observed, but this may not be very practical if the inter-arrival times between demands are very long. Both the Croston and SB2 methods tend to over-forecast due to not adjusting the forecast often enough. In addition, neither of the Croston-based methods nor the SES method take seasonality of the demand into account. As an example, in Figures 3.1.1 and 3.1.2, the actual sales for each week is plotted along with the forecasts generated using the SES, Poisson SES, Croston, and SB2 methods.

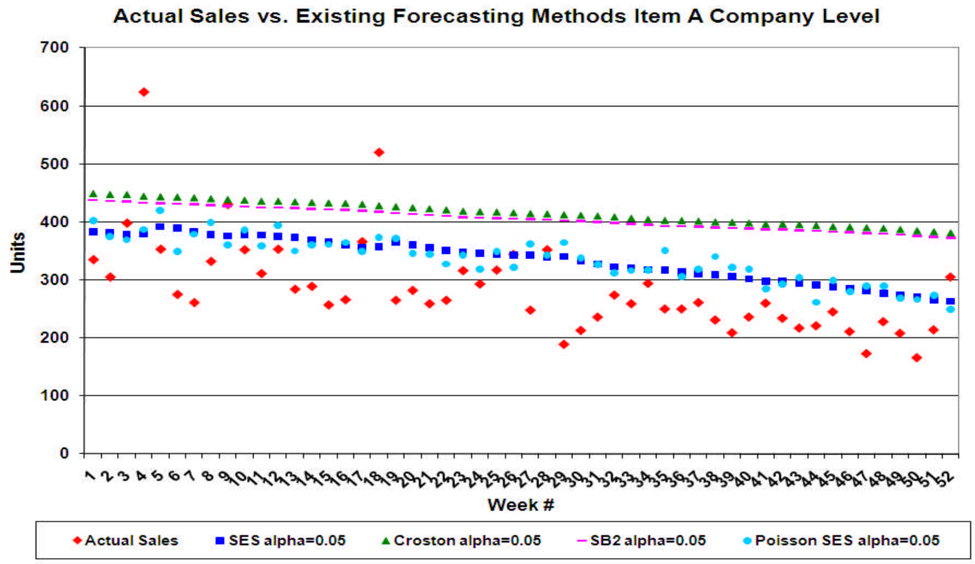

Figure 3.1.1 Company level actual sales vs. existing methods' forecast, Item A

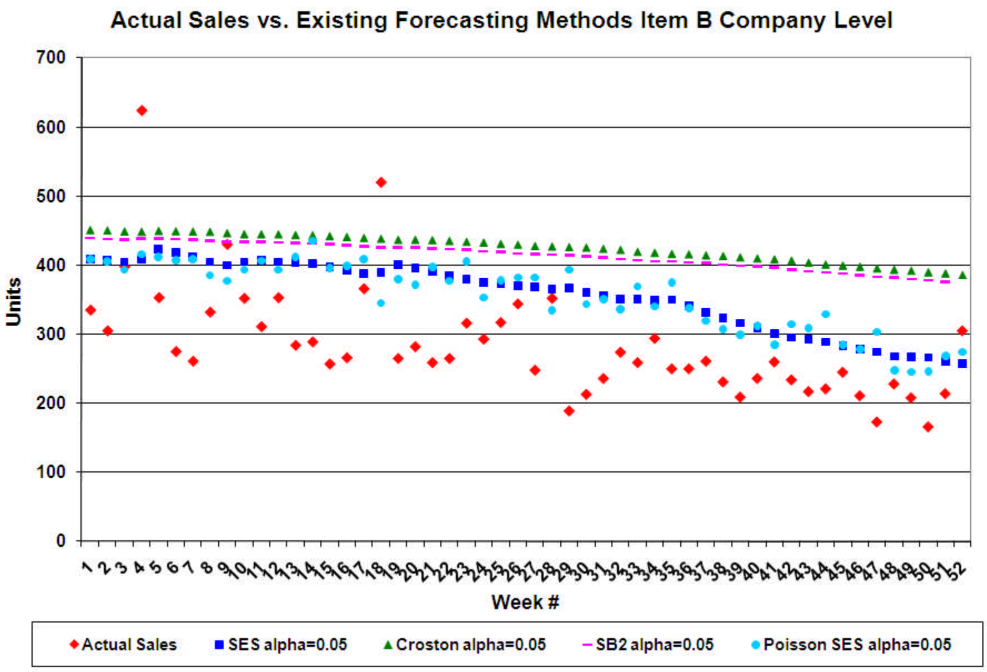

Figure 3.1.2 Company level actual sales vs. existing methods' forecast, Item B

3.2 New Features

Based upon the observations from the previous section, we develop three modifications that are used with the most promising of the existing forecasting techniques from the literature. These modifications include:

1. More frequent forecast updates,
2. Seasonal adjustments, and
3. A correction factor for poor performance.

Each is described below.

3.2.1 Forecast Update Adjustment

In some of our forecasting methods, we adjust the forecast not only after a demand, but also if the time since the last demand exceeds the average inter-arrival time or two times the average inter-arrival time, depending on the method.

3.2.2 Seasonal Adjustment

To the best of our knowledge, seasonality is not taken into account to forecast the sales of slow selling items. Thus, we also add a seasonality index (SI) to the forecast.

The SI methods calculate a forecast using seasonality indices for the promotional weeks which are, for our data, Memorial Day, July 4th, Labor Day, and Thanksgiving week. The indices for those four weeks are calculated using aggregate level country sales by item as below:

$$SI_{i,t} = \frac{Y_{i,t}}{\frac{1}{52}\sum_{t=1}^{52} Y_{i,t}},$$

where $Y_{i,t}$ is the aggregated store sales of item i in week t. The index for other non-promotional weeks is set to 1.

3.2.3 Correction Factor

The sales data consists of weekly sales of two appliance items in the stores. The first item is sold at 1,623 stores, whereas the second item is sold at 1,633 stores. If the forecast is consistently high or low compared with actual sales for the last *x* consecutive weeks for an item-store pair, then the forecasts of the SES, Croston, and SB2 methods are lowered or increased by a correction factor (*CF*). The value of *CF* is updated weekly for each item-store to be used in the forecasting methods with *CF*. Below is the notation that is used to calculate *CF* of item *i* at the k^{th} store:

$Y'_{t,i,k}$ is the forecasted demand for period t+1 at the end of period t of item i at the k^{th} store;

$Y_{t,i,k}$ is the actual demand at the end of period t of item i at the k^{th} store;

$t_{r,i,k}$ is the first week that the item i is received at the k^{th} store;

$IA_{0,i,k}$ is the time between $t_{r,i,k}$ and the first non-zero demand of item i at the k^{th} store;

$IA_{j,i,k}$ is the time between the non-zero demands j and j-1 of item i at the k^{th} store;

$CF_{t,i,k}$ is the correction factor at the end of period t for the forecasted demand for item i at the k^{th} store;

$m_{t,i,k}$ is the counter that is used in the correction factor calculation at the end of period t for item i at the k^{th} store;

m is the number used to determine when to assign CF to a value other than 1 and is equal to 2;

$t_{i,k}$ is the number of weeks used to initialize a given forecasting technique for item i at the k^{th} store;

$n_{i,k}$ is the number of weeks with non-zero demands for item i between times $t_{r,k}+1$ and $t_{r,k}+t_k$ at the k^{th} store;

t_l is the last week of data used for item i at the k^{th} store, and it is the same for all item stores;

x_i is the number of item-stores of item i (it is 1,623 for item A and 1,633 for item B.);

LL is the lower limit parameter used in CF calculation, if the aggregated sales of item i at all stores is less than or equal to LL times the aggregated forecast of item i in all stores, that week is considered a week where sales are overestimated;

UL is the upper limit parameter used in CF calculation, if the aggregated sales of item i at all stores is greater than or equal to UL times the aggregated forecast of item i in all stores, that week is considered a week where sales are underestimated; and

MCF is the maximum correction factor.

Below is the method to calculate CF for item i at the k^{th} store.

1) Initialize/set parameters.

   a) $m_{0,i,k} = 0$,

   b) Initialize the sales forecast all item-stores,

   c) $t \leftarrow 1$.

2) Update the forecast of all item-stores based on the method selected.

3) If $\sum_{k=1}^{x_i} Y_{t,i,k} \leq LL \sum_{k=1}^{x_i} Y'_{t-1,i,k}$, then $m_{t,i,k}$ =minimum ($m_{t,i,k} +$ 1, *MCF*)

   Else if $\sum_{k=1}^{x_i} Y_{t,i,k} \geq UL \sum_{k=1}^{x_i} Y'_{t-1,i,k}$, then $m_{t,i,k}$ =maximum ($m_{t,i,k} - 1$, -*MCF*)

4) $$CF_{t,i,k} = \begin{cases} (UL - 0.1)^{-\min(m, MCF)}, & if \ m_{t,i,k} \geq m \\ (LL + 0.1)^{\max(m, -MCF)}, & if \ m_{t,i,k} \leq -m \\ 1, \text{ o.w.} \end{cases}$$

5) Set $t = t + 1$

6) Return to step 2) until $t > t_l$.

The main logic used to calculate *CF* using the formula in step 4) above is making sure that the forecast is lowered via *CF* if the sales are overestimated; and forecast is increased via *CF*, if sales are underestimated. Table 3.2.3.1 displays the *CF* values for $m_{t,k}$, *MCF*, *LL* and *UL*.

Table 3.2.3.1 *CF* values based on *m*, *MCF*, *LL* and *UL*

Sales are over estimated/ Sales are under estimated	*m*	*MCF*	*LL* **=0.8**	*UL* **=1.2**
O	2	2		**0.83**
O	2	3		**0.83**
O	3	3		**0.75**
U	-2	2	**1.23**	
U	-2	3	**1.23**	
U	-3	3	**1.37**	

Table 3.2.3.2 gives the values of the parameters required in *CF* calculation used in the simulation experiments.

Table 3.2.3.2 Parameter values for *CF* calculation

Parameter	Value
UL	1.2
LL	0.8
MCF	2
M	2
$t_{i,k}$	58

These values are set based upon preliminary investigations that worked well.

3.3 Forecasting Methods Evaluated

In this dissertation we compare four of the existing techniques in the literature, SES, Croston, SB2, and SES based on Poisson forecasts with four new techniques formed by including some or all the modifications described in Section 3.2. Table 3.3.1 summarizes all methods investigated.

Table 3.3.1 Summary of the methods to be analyzed

Type	Croston and Modifications	Syntetos and Boylan and Modifications	SES
Regular	Croston (Original)	SB2 (Original)	SES (Original)
	Croston SI New	SB2 SI New	
	Croston SI New CF	SB2 SI New CF	
Poisson			Poisson SES (Original)

Each of these methods are broken into three subgroups: (i) Croston and modifications; (ii) SB2 and modifications; and (iii) SES and modifications. The methods with the label "New" have a different forecast update frequency. For Croston SI New and SB2 SI New, the forecast is also updated if the last time a sale occurred exceeds two times the average inter-arrival time between the sales. Furthermore, for Croston SI New CF and SB2 SI New CF, the forecast is also updated if the last time a sale occurred exceeds the average inter-arrival time between the sales. In addition, all of the new methods include seasonality indexes to adjust the forecast as discussed in Section 3.2.2, while the methods with the label *CF* utilize the correction factor to adjust the forecast as discussed in Section 3.2.3.

A summary of the calculations of forecasting methods is provided in Table 3.3.2.

Table 3.3.2 Summary of the new methods' calculations

Forecasting Method	**Calculations**
Croston SI New	If Y_t=0 and current interarrival time $\le$ average interarrival time, then: $z'_t = z'_{t-1}$ and $p'_t = p'_{t-1}$ and $q = q+1$ Else if Y_t=0 and current interarrival time > average interarrival time, then: $z'_t = z'_{t-1} + \alpha(\frac{Y_t}{SI_t} - z'_{t-1})$ and $p'_t = p'_{t-1} + \alpha(q - p'_{t-1})$, $q = q+1$ Else $z'_t = z'_{t-1} + \alpha(\frac{Y_t}{SI_t} - z_{t-1}')$ and $p'_t = p'_{t-1} + \alpha(q - p'_{t-1})$, $q = 1$ End if $Y'_t = \frac{z'_t}{p'_t} * SI_{t+1}$.
SB2 SI New	Same If Statement as Croston SI New. $Y'_t = (1 - \frac{\alpha}{2})\frac{z'_t}{p'_t} SI_{t+1}$
Croston SI New CF	Same If Statement as Croston SI New. $Y'_t = \frac{z'_t}{p'_t} * SI_{t+1} CF_t$,
SB2 SI New CF	Same If Statement as Croston SI New. $Y'_t = (1 - \frac{\alpha}{2})\frac{z'_t}{p'_t} SI_{t+1} CF_t$

In all methods, the minimum value of the forecast is 0.001, not 0. It is the company's policy to have a non-zero forecast if the products are not seasonal.

4 Sales Forecast Accuracy Measures

Accuracy measures can be broadly categorized in four groups: absolute, relative to a base, relative to another method, and relative to the series. The following notation is used in the next sections to explain the various accuracy measures used in this research:

Y_t': the estimate made at the end of period t of the demand in period t+1, obtained by any forecasting method;

Y_t: the actual demand in period t;

e_t: the forecast error in period t; and

N: the number of demand time periods considered for the purpose of comparison.

4.1 Absolute Accuracy Measures

Absolute accuracy measures represent the forecast error in a particular time period that is expressed in an absolute or squared form.

4.1.1 Mean Square Error

One of the absolute accuracy measures is Mean Square Forecast Error (*MSE*). It is defined as:

$$MSE = \frac{\sum_{t=1}^{N}(Y_t - Y_{t-1}')^2}{N} = \frac{\sum_{t=1}^{N}(e_t)^2}{N}. \quad (4.1)$$

MSE is the second moment (about the origin) of the error, and thus incorporates both the variance of the estimator and its bias. However, *MSE* is scale dependent and tends to heavily weight outliers. This is a result of the squaring of each term, which weights large errors more heavily than small ones. Chatfield (1992) stated that since *MSE* is scale dependent, it can be disastrous to average *MSE*s across different series.

4.1.2 Mean Signed Error

Another accuracy measure is Mean Signed Error (*ME*), defined as:

$$ME = \frac{\sum_{t=1}^{N}(Y_t - Y'_{t-1})}{N} = \frac{\sum_{t=1}^{N} e_t}{N}. \tag{4.2}$$

ME is very easy to calculate and interpret since the difference between the average *ME* given by different forecasting methods indicates how much more or less biased one method is compared with another one.

The disadvantage of this accuracy measure is that when *ME*s are averaged, the series with large numbers may dominate the final result. However, since the *ME* measure takes into account the sign of the error, scale differences will not have a great effect on the average *ME* value compared with some other absolute accuracy measures such as *MSE*. In addition, our data set does not have large numbers; and, in practice, this is not a big concern.

4.1.3 Mean Absolute Error and Scaled Mean Absolute Error

Another accuracy measure is Mean Absolute Error (*MAE*), defined as:

$$MAE = \frac{\sum_{t=1}^{N} \left| Y_t - Y'_{t-1} \right|}{N} = \frac{\sum_{t=1}^{N} \left| e_t \right|}{N} \quad . \tag{4.3}$$

The Scaled Mean Absolute Error (*SMAE*) is calculated by dividing the originally calculated *MAE* per method per series by the average demand per unit time period (for the series under concern) to eliminate scale dependencies. It is defined as:

$$SMAE = \frac{MAE}{\left(\frac{\sum_{t=1}^{N} Y_t}{N} \right)} . \tag{4.4}$$

4.2 Accuracy Measures Relative to a Base

U-Statistic, Batting Average, Geometric Mean Relative Absolute Error (GMRAE), and *MdRAE (Median Relative Absolute Error)* are some of the accuracy measures relative to a base, but they are not relevant to slow-moving items since the forecast error produced by the naïve method can be zero as well.

4.3 Accuracy Measures Relative to the Series

MAPE (Mean Absolute Percentage Error), *MdAPE* (Median Absolute Percentage Error), *MAPEFF* (Mean Absolute Percentage Error from Forecast), $MAPE_{\text{sym}}$ (Symmetric

Mean Absolute Percentage Error), *DAM* (Denominator Adjusted *MAPE*), and Ratio of *MAD/MEAN* are some of the accuracy measures relative to the series.

Even though *MAPE* is an effective way of addressing the scale dependence problem for fast selling items, it is not applicable to slow-moving items since the actual value can be zero in some periods and division by 0 will generate an error. Another option is to use $MAPE_{sym}$, which is calculated by dividing the absolute difference of the actual sales and the forecast by the average of the actual and forecast values. Details are discussed in the next sections.

4.3.1 Mean Absolute Percentage Error (*MAPE*)

MAPE is calculated as:

$$MAPE = \frac{\sum_{t=1}^{N} \left| \frac{Y_t - Y'_{t-1}}{Y_t} \right|}{N} 100\% = \frac{\sum_{t=1}^{N} \left| \frac{e_t}{Y_t} \right|}{N} 100\% . \quad (4.5)$$

MAPE is probably the most widely used scale-independent method to measure accuracy. It can be used to compare forecast performance across different data sets, since it allows averaging all *MAPE* values for a method across many series.

However, measurements based on percentage errors have the disadvantage of being infinite or undefined if there are zero values in a series, as is frequent for intermittent data. Another disadvantage of *MAPE* is it lacks statistical theory similar to *MSE* and has been rejected by Fildes (1992) since the sampling distribution for *MAPE* measured across series is often badly positively skewed and *MAPE* is also sensitive to location in that a change of origin in the data affects the *MAPE* calculations. Moreover, *MAPE* puts a heavier penalty on forecasts that are more than the actual demand rather than on those that are less than the actual demand (Madridakis, 1993). In addition, *MAPE* is highly prone to right-skewed asymmetry in actual practice, since the underlying error distributions of these measures have only positive values and no upper bound. (Smith and Sincich, 1988). This has led to the use of the "symmetric" *MAPE* ($MAPE_{sym}$) in the M3-competition (Makridakis & Hibon, 2000). $MAPE_{sym}$ is discussed in the next section.

4.3.2 Symmetric Mean Absolute Percentage Error ($MAPE_{sym}$)

$MAPE_{sym}$ is calculated as:

$$MAPE_{sym} = \frac{\sum_{t=1}^{N} \left| \frac{Y_t - Y'_{t-1}}{(Y'_{t-1} + Y_t)/2} \right|}{N} 100\% = \frac{\sum_{t=1}^{N} \left| \frac{e_t}{(Y'_{t-1} + Y_t)/2} \right|}{N} 100\% \,. \tag{4.6}$$

The symmetrical *MAPE* ($MAPE_{sym}$) was designed to deal with some of the limitations of *MAPE* (Makridakis 1993). Like *MAPE*, $MAPE_{sym}$ is an average of the absolute percent errors but these errors are computed using a denominator representing the average of the forecast and observed values. $MAPE_{sym}$ has an upper limit of 200%, offers a well-designed range to judge the level of accuracy, and should be influenced less by extreme values. It also corrects for the computational asymmetry of the forecast error. For example, a forecast equal to 150 and actual demand of 100 yields a *MAPE*-computed forecast error of 50%, while a forecast of 100 and actual demand of 150 yields a *MAPE*-computed forecast error of 33%.

However, the average of forecast and actual sales in the denominator of the $MAPE_{sym}$ yields 40% in either situation.

Moreover, if the actual value Y_t is zero, the forecast Y'_{t-1} is likely to be close to zero. In addition, if the forecast is generated by a Poisson distribution, the forecasts can also be exactly zero. Thus the measurement will still involve division by a number close to zero or zero. In addition, whenever the actual sale is 0, $MAPE_{sym}$ will have a value of 2, regardless of the forecast. Thus, the use of $MAPE_{sym}$ is not recommended.

4.3.3 Mean Absolute Percentage Error from Forecast (*MAPEFF*)

MAPEFF is calculated as (Pearson and Wallace (1999)):

$$MAPEFF = \frac{\sum_{t=1}^{N} \frac{(Y_t - Y'_{t-1})}{Y'_{t-1}}}{N} 100\% = \frac{\sum_{t=1}^{N} \frac{(e_t)}{Y'_{t-1}}}{N} 100\%.$$

Similar to $MAPE_{sym}$, if the forecast is generated by a Poisson distribution, the forecasts can also be exactly zero. Thus *MAPEFF* will still involve division by zero.

4.3.4 Denominator Adjusted Mean Absolute Percentage Error (*DAM*)

DAM is calculated as (Hoover, 2006):

$$DAM = \frac{\sum_{t=1}^{N} \frac{|Y_t - Y'_{t-1}|}{\hat{Y}_t}}{N} 100\%,$$

$$= \frac{\sum_{t=1}^{N} \frac{|e_t|}{\hat{Y}_t}}{N} 100\%,$$

where the adjusted denominator value $\hat{Y}_t$ is defined as

$$\hat{Y}_t = \begin{Bmatrix} 1, & \text{if } Y_t = 0 \\ Y_t, & \text{otherwise} \end{Bmatrix} .$$

4.3.5 Ratio of Mean Absolute Error to *MEAN* Demand *(MAE/MEAN)*

MAE/MEAN is calculated as (Hoover, 2006):

$$MAE / MEAN = \frac{\frac{\sum_{t=1}^{N} |Y_t - Y'_{t-1}|}{N}}{\frac{\sum_{t=1}^{N} (Y_t)}{N}} 100\% = \frac{\sum_{t=1}^{N} \frac{|e_t|}{N}}{\frac{\sum_{t=1}^{N} (Y_t)}{N}} 100\% \cdot$$

4.4 Accuracy Measures for Comparing Two Methods

Accuracy measures for one method relative to another method are widely used for comparisons indicating how much better or how many times one method performs better than the other method. Relative Geometric Root Mean Square Error (*RGRMSE*) is used for this purpose.

4.4.1 Relative Geometric Root Mean Square Error (*RGRMSE*)

RGRMSE for comparing methods *A* and *B* in a particular time series is defined as:

$$RGRMSE = \frac{\left(\prod_{t=1}^{N}(Y_t - Y'_{A,t-1})^2\right)^{\frac{1}{2N}}}{\left(\prod_{t=1}^{N}(Y_t - Y'_{B,t-1})^2\right)^{\frac{1}{2N}}}. \tag{4.7}$$

In addition, *RGRMSE* for comparing methods *A* and *B* across a range of series (k=1,…, κ) is defined as:

$$RGRMSE = \frac{\left(\prod_{k=1}^{\kappa}(GRMSE_{A,k})^2\right)^{\frac{1}{2\kappa}}}{\left(\prod_{k=1}^{\kappa}(GRMSE_{B,k})^2\right)^{\frac{1}{2\kappa}}} = \frac{\left(\prod_{k=1}^{\kappa}(GRMSE_{A,k})\right)^{\frac{1}{\kappa}}}{\left(\prod_{k=1}^{\kappa}(GRMSE_{B,k})\right)^{\frac{1}{\kappa}}}, \tag{4.8}$$

where $GRMSE_{i,k}$, per series for method *i* is calculated as:

$$GRMSE_{i,k} = \left(\prod_{t=1}^{N}(Y_t - Y'_{i,t-1})^2\right)^{\frac{1}{2N}}, \; i \in \{A, B\}. \tag{4.9}$$

Even though this metric is more complex, it is more robust to outliers. This measure was first suggested by Fildes (1992) and was shown to cancel out the distorting effect of large errors. However, the disadvantage of this accuracy measure is the value of *GRMSE* is zero when, in any period, the forecast is equal to the actual sales, which might occur for the forecast generated by Poisson methods, and this makes comparison not possible among series for Poisson generated forecasts.

4.5 Summary of Chapter 4

Accuracy measures can be broadly categorized in four groups: absolute, relative to a base, relative to another method, and relative to the series. In this chapter, we summarized these common sales forecast accuracy measures used in the literature. In addition, among those measures we also identified the accuracy measures that cannot be used to compare forecasting methods for slow-seller items such as *MAPE*, *GMRAE*, and *MdRAE*.

Among the methods that can be used to compare forecasting methods for slow-seller items, we will use *ME*, *MAE*, *SMAE*, and *GRMSE* /*RGRMSE* as the sales forecast accuracy measures to compare the existing and new forecasting methods for slow-seller items throughout the rest of the dissertation. These measures are selected since they are easy to calculate and interpret and due to the reasons explained in Chapter 4. In addition, these measures are commonly used in the literature for slow-seller items.

5 Experimental Structure

We compare 25 forecasting methods using sales data for two slow-selling appliance items, a dryer and a dishwasher from the home improvement retail industry. The data consists of 3,256 item-store sales. Unfortunately, 267 of the item-stores do not have enough history for comparisons; thus they are excluded from the analysis. For the remaining 2,989 item-stores, we define slow moving items as those with average weekly sales less than or equal to 0.3. Using this definition, 2,504 of the item-stores meet this definition and are used in the analysis.

Table 5.1 shows the sales series characteristics.

Table 5.1 Sales series characteristics

	Dryer (Item A), Stores with Average Weekly Sales <=0.3			Dishwasher (Item B), Stores with Average Weekly Sales <=0.3		
	Ranges		Mean Value	Ranges		Mean Value
	Min	Max		Min	Max	
Minimum non-zero demand	1	2	0.94	1	2	0.93
Maximum non-zero demand	1	5	1.43	1	6	1.44
Mean non-zero demand	1	2.66	1.04	1	2.2	1.03
Variance non-zero demand	0	4.33	0.12	0	2.5	0.11
Minimum inter-arrival time	0	52	3.6	0	52	3.06
Maximum inter-arrival time	1	52	21.92	1	52	20.19
Mean inter-arrival time	2.33	26	8.59	2.33	26	7.91
Average Demand/week	0	0.3	0.128	0	0.3	0.13

The data used is the sales of 2 appliance items at the retail stores over a 110-week period. The forecasts are generated only after the first receipt date of the item at the store. In addition, the historical data is divided into two sections. The first section is used as a warm-up period, with the remaining is used as experimental data. The sales data of the first 58 weeks after the initial receipt date of the item at the store is used for the warm-up period to initialize the forecasts. Moreover, the calculation of the accuracy and inventory metrics starts 58 weeks after the initial receipt date of the item at the store after this warm-up period.

Values of $\alpha =$ 0.05, 0.1, 0.15, and 0.2 are used as smoothing constants. In addition, similar to Syntetos and Boylan (2005), sales forecast accuracies are calculated separately for two cases: every forecast generated ("all points in time") and for the weeks right after a nonzero demand (at "issue points only").

In addition to sales forecast accuracies, inventory control metrics are compared for the different methods in Chapter 7.

6 Sales Forecast Accuracy Analysis

In this chapter, the existence of bias for each of the methods is tested using the *ME* statistic. In addition, the forecasting methods are compared using the *MAE*, *SMAE*, and *RGRMSE* accuracy metrics. The results are displayed separately for the two items, A and B (mentioned in Chapter 3), for smoothing values of $\alpha = 0.05$, 0.1, 0.15, and 0.2. The accuracy metrics are computed at the end of each week (termed "all points in time") and also for the weeks right after a nonzero demand (termed "issue points only"). In the sections that follow we rank the methods based on "all points in time" measures as well as the measures at issue points along the lines of other researchers like Syntetos and Boylan (2005).

We provide figures showing confidence intervals (*CI*s) of each method and each sales forecast accuracy measure across all smoothing parameter values.

In equations (6.1), (6.2), and (6.3) below, the value $\overline{X}_i$ is the average sales forecast accuracy measure (*ME*, *MAE*, *SMAE*, or *GRMSE*) obtained by a specific forecasting method across all series (stores) for item *i*, and is calculated as:

$$\overline{X}_i = \frac{1}{\kappa_i} \sum_{k=1}^{\kappa_i} X_{ik} \ ,$$

where X_{ik} is the sales forecast accuracy measure (*ME*, *MAE*, *SMAE*, or *GRMSE*) obtained by a specific forecasting method for item *i* at the k^{th} store, and κ_i is the number of demand series (stores) for item *i*. Note that κ_i is equal to 1,254 for item A, and 1,250 for item B.

We define s_{X_i} as the standard deviation of the sales forecast accuracy measure for item *i*, and is calculated as:

$$s_{X_i} = \sqrt{\frac{1}{\kappa_i - 1}\sum_{k=1}^{\kappa_i}(X_{ik} - \overline{X_i})^2} \ .$$

Since the number of stores are very much larger than 30 for both items, a 99% confidence interval for the true mean sales forecast accuracy measure of item *i* is given by:

$$\overline{X_i} \pm 2.58\frac{s_{X_i}}{\sqrt{\kappa_i}} \ . \tag{6.1}$$

Moreover, the lower limit of the confidence interval of the true mean sales forecast accuracy measure of item *i* is given by:

$$\overline{X_i} - 2.58\frac{s_{X_i}}{\sqrt{\kappa_i}} \ , \tag{6.2}$$

and the upper limit of confidence interval of the true mean sales forecast accuracy measure of item *i* is given by:

$$\overline{X_i} + 2.58\frac{s_{X_i}}{\sqrt{\kappa_i}} \ . \tag{6.3}$$

Since the sample sizes are quite large, the resulting confidence intervals are very tight making differences in performance between the forecasting methods quite obvious for the most part. However, in some cases, these confidence intervals overlap. Therefore, in addition to the confidence intervals, we also provide the results of two-sided *t*-tests in order to distinguish differences in two methods' performances. Specifically, we formulate the following null and alternative hypotheses.

H_0: The average sales forecast accuracy measure (*ME*, *MAE*, *SMAE*, or *GRMSE*) resulting from forecasting method "1" for item *i* is equal to the average sales forecast accuracy measure (Absolute *ME*, *MAE*, *SMAE*, or *GRMSE*) given by method "2" for item *i*.

H_1: The average sales forecast accuracy measure (*ME*, *MAE*, *SMAE*, or *GRMSE*) given by method "1" for item *i* is not equal to the average sales forecast accuracy measure (Absolute *ME*, *MAE*, *SMAE*, or *GRMSE*) given by method "2" for item *i*.

For each performance measure (*ME*, *MAE*, *SMAE*, and *GRMSE*) two-sided *t*-test statistics are calculated for each pair of forecasting methods to be compared in order to test the hypotheses. In order to define the test statistic, we first define ε_{ikm} as the sales forecast accuracy measure of item *i* at the k^{th} store using method *m*. Further, we define P_{ik} as the paired sales forecast accuracy measure difference for item *i* between methods "1" and "2" for the k^{th} store, and is calculated as:

$P_{ik} = \varepsilon_{ik1} - \varepsilon_{ik2}$, for k=1,…,κ_i .

Then, the *t*-test statistic is calculated as:

$$t = \frac{\overline{P_i} - \mu}{\frac{s_{P_i}}{\sqrt{\kappa_i}}}, \qquad (6.4)$$

where:

$$\overline{P_i} = \frac{1}{\kappa_i}\sum_{k=1}^{\kappa_i} P_{ik} \ ,$$

$$s_{P_i}^2 = \frac{1}{\kappa_i - 1}\sum_{k=1}^{\kappa_i}(P_{ik} - \overline{P_i})^2 \text{, and}$$

μ =0.

The *t*-test results of these pairwise comparisons are given in Appendix A.1-A.4, and the explanations of these results are in each sub-section.

We have such large sample sizes and thus can estimate the average percent forecast error metric with excellent accuracy; and the statistical tests for pairwise comparisons indicate that virtually any differences in the performance metrics are statistically significant, though perhaps not practically significant. This allows us to rank the methods based on the average sales forecast accuracy metric of each method. The best performance for a given method and metric is selected separately for "all points in time" and at "issue points only" across the four levels of the smoothing parameter value α. Then, these values are ranked separately for "all points in time" and at "issue points only" across the methods.

We remove the scale dependency for the *ME*, *MAE* and *SMAE* metrics by dividing by the mean demand over the testing horizon (i.e., mean *ME*/mean demand, mean *MAE*/mean demand, and mean *SMAE*/ mean demand), and we use these as the performance measures to compare the forecasting methods. We demonstrate this using the *ME* measure below.

We define ME_{ik} as the mean error obtained by a specific forecasting method of item i at the k^{th} store and $\overline{ME_i}$ as the overall average of the mean errors over all stores, i.e.

$$\overline{ME_i} = \frac{1}{\kappa_i}\sum_{k=1}^{\kappa_i} ME_{ik} \text{ ,}$$

where κ_i is the number of demand series (stores) of item *i*. Then, let Y_{ikt} be the actual demand of item *i* at the k^{th} store in period *t*, let $\overline{Y}_{ik}$ be the average actual demand over all *N* periods, i.e.

$$\overline{Y}_{ik} = \frac{1}{N}\sum_{t=1}^{N}\overline{Y}_{ikt} \ ,$$

and let $\overline{\overline{Y}}_i$ be the grand average over all stores, i.e.

$$\overline{\overline{Y}}_i = \frac{1}{\kappa_i}\sum_{k=1}^{\kappa_i}\overline{Y}_{ik}.$$

Based on these definitions we then define the mean *ME*/mean demand for item *i* as

$$\frac{\overline{ME}_i}{\overline{\overline{Y}}_i}. \tag{6.5}$$

The sales forecast accuracy and bias analysis using *ME* are reported in Section 6.1, the sales forecast accuracy analysis using *MAE* are given in Section 6.2, the sales forecast accuracy analysis using *SMAE* are given in Section 6.3, and the sales forecast accuracy analysis using *RGRMSE* are given in Section 6.4.

6.1 *ME* Results

Determination of whether each forecasting method is biased or not has been tested using the *ME* accuracy measure and confidence intervals on the *ME* values. The bias is equal to the difference between the forecast and actual sales and it is equal to the negative of *ME*. Thus, a positive *ME* value indicates that the actual sales are greater than the forecasted demand, i.e., sales are underestimated, and a negative *ME* value indicates that the sales are overestimated.

The 99% confidence intervals of the *ME* for each method calculated using equations (6.2) and (6.3) are presented separately for Items A and B as measured at “all points in time” and at “issue points only” in the Figures 6.1.1, 6.1.2, 6.1.3, and 6.1.4 below. The zero line is drawn in black to separate clearly the cases of demand underestimation (above the zero line) from the cases of demand overestimation (below the zero line). Furthermore, vertical solid lines are added to the figures to separate the results for different levels of the smoothing parameter value α and vertical dashed lines are added to the figures to separate the results for the existing (OLD) forecasting methods versus the new (NEW) methods.

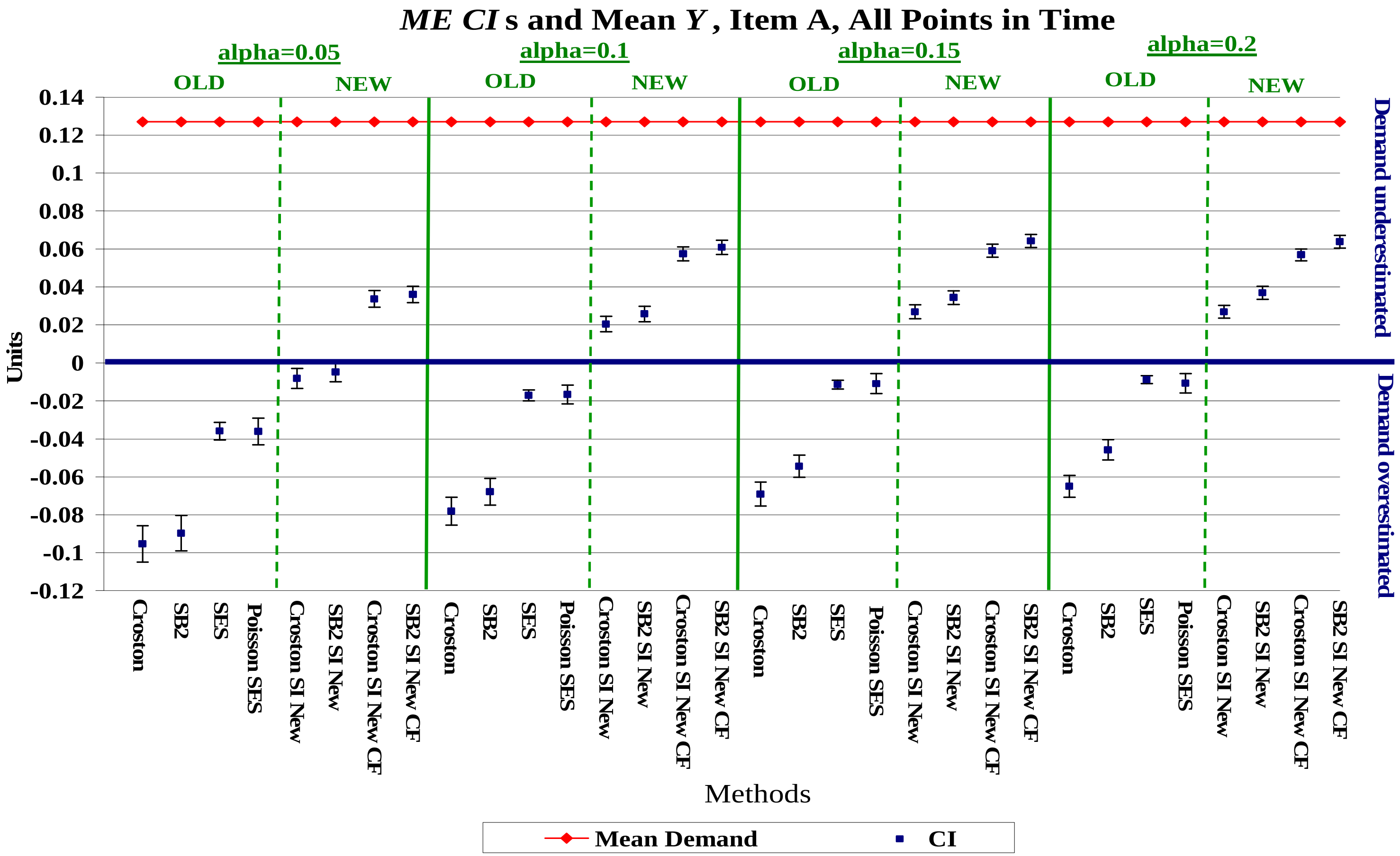

Figure 6.1.1 *ME* confidence intervals and mean demand of item A, All Points in Time

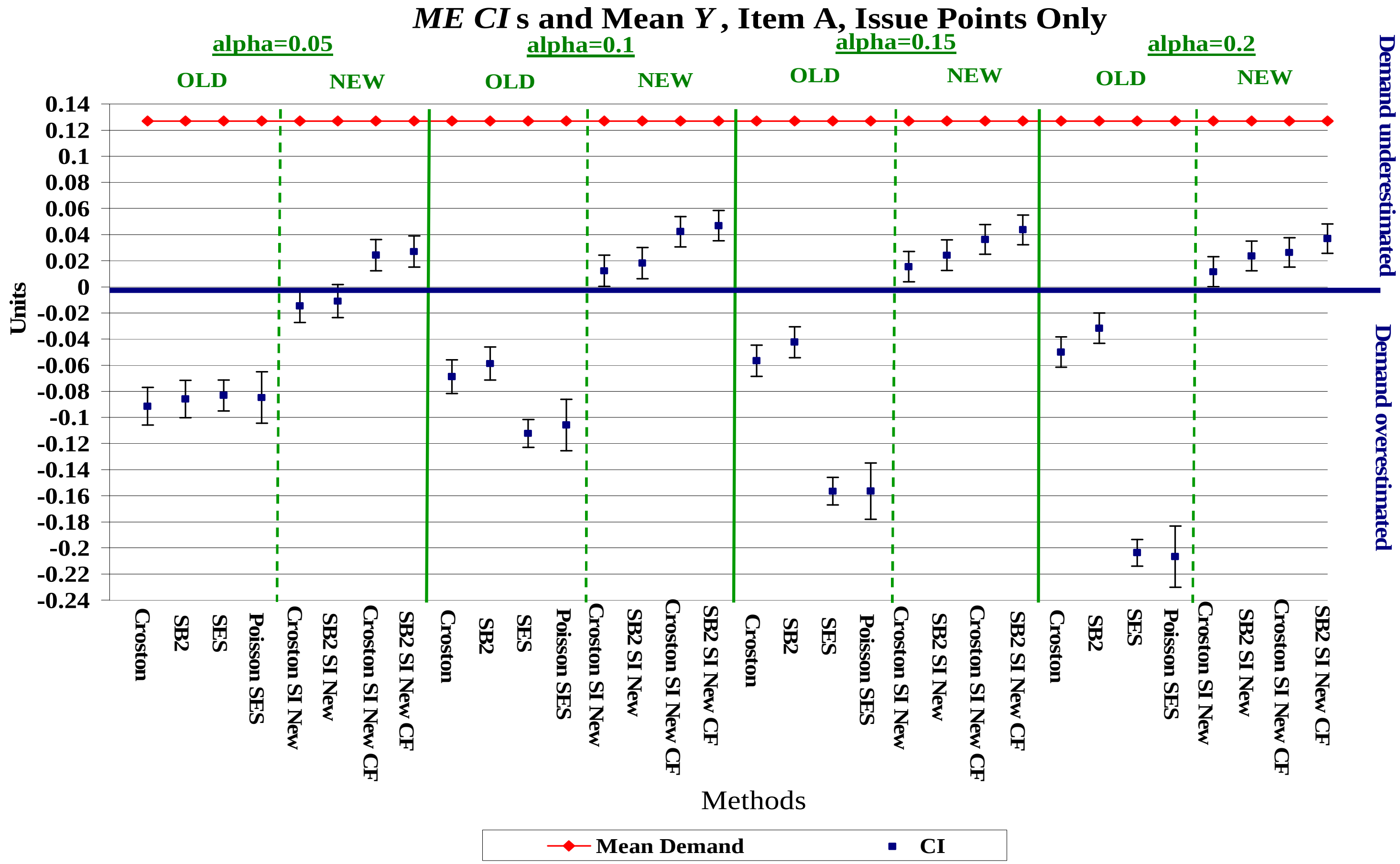

Figure 6.1.2 *ME* confidence intervals and mean demand of item A, Issue Points Only

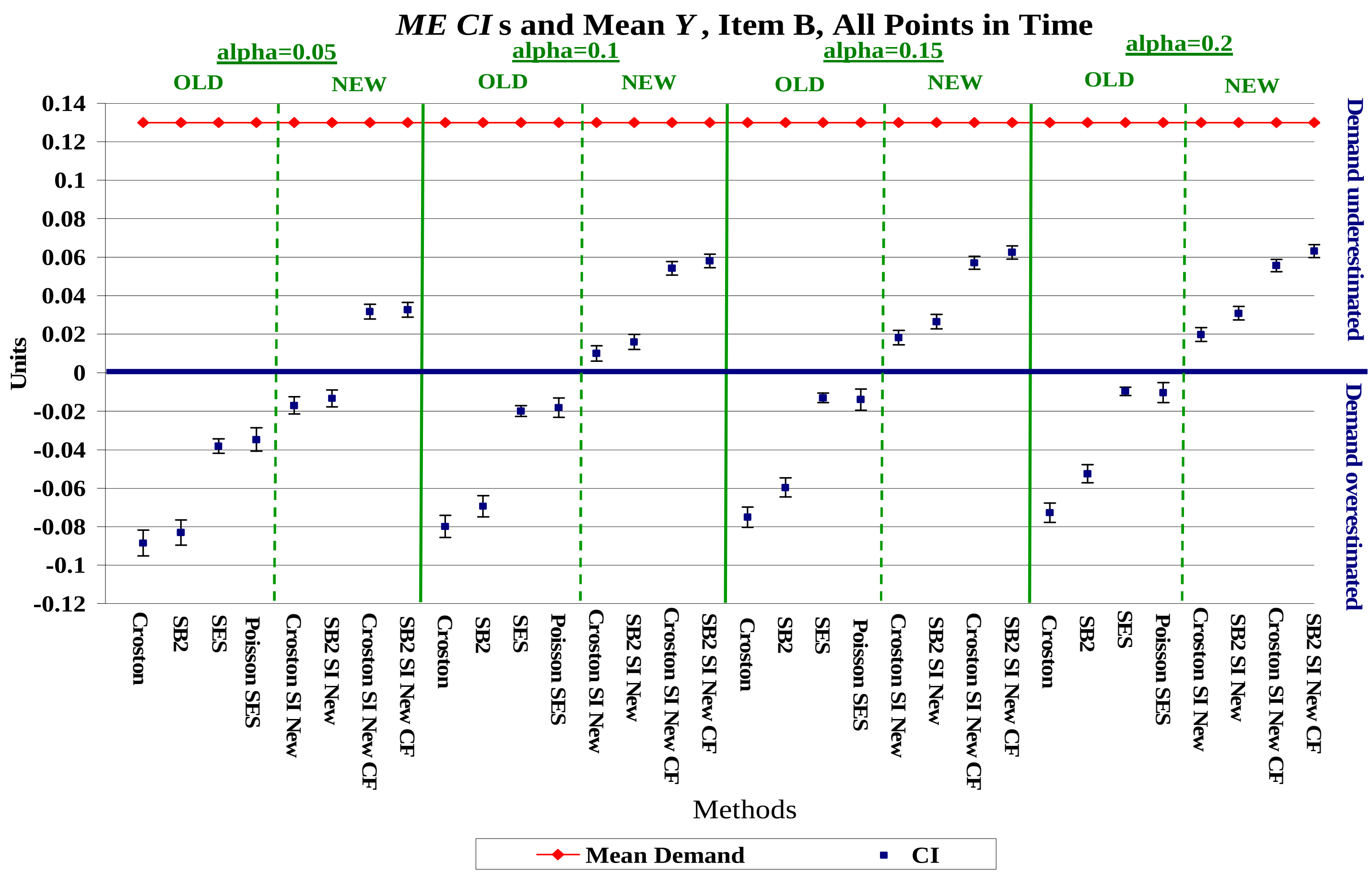

Figure 6.1.3 *ME* confidence intervals and mean demand of item B, All Points in Time

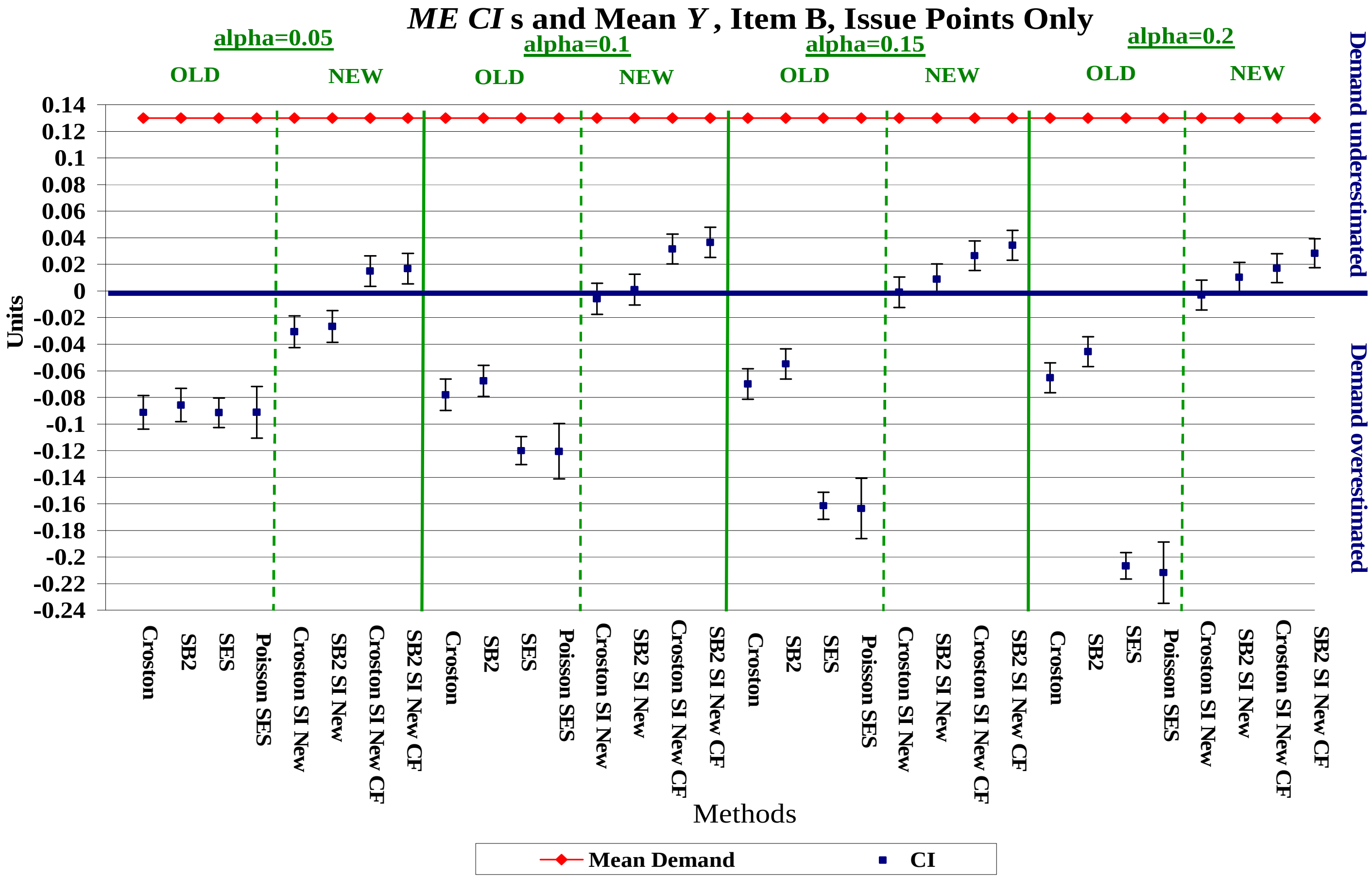

Figure 6.1.4 *ME* confidence intervals and mean demand of item B, Issue Points Only

A cursory observation of these figures shows the very tight confidence intervals we alluded to earlier with some methods consistently different than others. Below we provide a list of other observations that can be made.

First observations about the performance of the existing methods are made.

- In all cases the existing methods overestimate the demand.
- The Croston and SB2 methods have very similar performance (except for "all points in time" and α =0.15 and 0.2 values; in this case the SB2 method is better (less biased) than the Croston method), although SB2 generally yields a slightly lower forecast than Croston, but this result is not statistically significant.
- The SES and Poisson SES methods have similar performance but neither consistently results in lower forecasts than the other.
- The performance of the SES and Poisson SES methods is better (less biased) than Croston and SB2 at "all points in time"; while the opposite is true at "issue points only", except for α =0.05.

Next we make observations relative to the new methods.

- In all cases the new methods Croston SI New CF and SB2 SI New CF underestimate the demand.
- For both items, the Croston SI New and SB2 SI New methods slightly underestimate the demand at "all points in time" except for α =0.05. For Item A at "issue points only" these two methods are either unbiased (when α =0.05) or they slightly underestimate the demand for the other α values, while for Item B

they either slightly overestimate the demand (when α =0.05) or are unbiased otherwise.

- The forecasts generated by the Croston SI New and SB2 SI New methods are higher than the Croston SI New CF and SB2 SI New CF methods. This indicates that the addition of the correction factor lowers the forecast.

Additional observations include the following.

- The new methods consistently produce lower forecasts than the existing methods for both items at all levels of the smoothing coefficient (α) both at “all points in time” and at “issue points only”.
- For both items, “all points in time”, and α =0.05, the Croston SI New and SB2 SI New methods perform better than the other methods (since the confidence intervals are closer to the 0 line where the bias is zero).
- For both items, “all points in time”, and α =0.1, the SES, Poisson SES, Croston SI New and SB2 SI New methods perform better than the other methods.
- For item A, “all points in time”, and α =0.15, the SES, Poisson SES methods perform better than the other methods. For item B, “all points in time”, and α =0.15, the SES, Poisson SES, and Croston SI New methods perform better than the other methods.
- For both items, “all points in time”, and α =0.2, the SES and Poisson SES methods perform better than the other methods.

- For both items, at "issue points only", and α =0.05, the new methods perform better than the old methods.
- For both items, "issue points only", and α =0.1, the Croston SI New and SB2 SI New methods perform better than all other methods.
- For item A, at "issue points only", and α =0.15 and 0.2, no clear distinction on which methods perform the best can be made since the confidence intervals overlap. However, for item A, "issue points only", and α =0.1, 0.15, and 0.2 the SES and Poisson SES methods perform the worst.
- For item B, at "issue points only", and α =0.15 and 0.2, the Croston SI New method performs better than the other methods except for SB2 SI New method with α =0.15 and 0.2 and the Croston SI New CF method with α =0.2. For item B, "issue points only", and α =0.1, 0.15, and 0.2 the SES and Poisson SES methods perform the worst.

In addition to the confidence intervals of the *ME*, the paired-sample *t*-test results of *ME* in Appendix A.1 calculated using the equation (6.4) are used to determine whether one method is more biased than the other one. The following statements are based on the observations of these paired-sample *t*-test results, and they provide better comparison information than the confidence intervals when the confidence intervals overlap. Furthermore, these statements are based on the statistical significant differences, not based on the practical significances.

First observations about the performance for "all points in time" are made.

- For both items and "all points in time", the forecasting methods are ranked from highest forecast generating method to lowest forecast generating method as: Croston, SB2, SES / Poisson SES (the SES and Poisson SES methods are not statistically significantly different), Croston SI New, SB2 SI New, Croston SI New CF, and SB2 SI New CF methods.

Next we make observations at "issue points only."

- For both items the Croston method yields statistically significantly higher forecasts than all the new methods and the SB2 method. However, the Croston method yields statistically significantly lower forecast than the SES and Poisson methods except for α =0.05.
- For both items, the SB2 method yields statistically significantly higher forecasts than all the new methods. However, similar to Croston's method the SB2 method yields statistically significantly lower forecast than the SES and Poisson methods except for item A and α =0.05.
- The SES and Poisson SES methods are not statistically significantly different.
- The SES and Poisson SES methods yield statistically significantly higher forecasts than the new methods.
- For both items, the new methods yield statistically significantly lower forecast than the old methods.

- For both items, the new forecasting methods are ranked from highest forecast generating method to lowest forecast generating method as: the Croston SI New, SB2 SI New, Croston SI New CF, and SB2 SI New CF methods.

Since the statistical tests for pairwise comparisons indicate that any differences in the performance metrics are statistically significant and we can estimate the average *ME* with excellent accuracy due to large sample sizes, we focus on ranking the methods using the lowest absolute value of mean *ME*/mean demand among all α values for "all points in time" and at "issue points only" separately. The lowest absolute values of each method for "all points in time" and at "issue points only" are highlighted below.

Table 6.1.1 Mean *ME*/mean demand of item A

Item A, Mean *ME*/Mean Demand								
	All Points in Time				Issue Points Only			
Methods	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.15$	$\alpha = 0.2$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.15$	$\alpha = 0.2$
Existing methods								
Croston	-75.10%	-61.51%	-54.45%	-51.14%	-72.07%	-54.24%	-44.54%	-39.39%
SB2	-70.72%	-53.44%	-42.87%	-36.03%	-67.66%	-46.30%	-33.36%	-24.99%
SES	-28.26%	-13.44%	-9.02%	-6.96%	-65.45%	-88.41%	-123.25%	-160.41%
Poisson SES	-28.44%	-13.10%	-8.61%	-8.46%	-66.75%	-83.32%	-123.22%	-162.82%
New methods								
Croston SI New	-6.48%	16.09%	21.21%	21.18%	-11.56%	9.66%	12.15%	9.11%
SB2 SI New	-3.82%	20.28%	27.12%	29.06%	-8.65%	14.40%	19.08%	18.65%
Croston SI New CF	26.57%	45.26%	46.59%	44.82%	19.13%	33.32%	28.66%	20.73%
SB2 SI New CF	28.41%	47.99%	50.59%	50.33%	21.27%	36.88%	34.35%	29.11%

For item A and for “all points in time”, all existing methods overestimate the sales. On the other hand, all new methods underestimate the sales, except for the Croston SI New and SB2 SI New methods with α =0.05.

In addition, the lowest absolute value of mean *ME*/mean demand of the existing methods is achieved with α =0.2, whereas the lowest absolute value of mean *ME*/mean demand of the new methods is achieved with α =0.05. Moreover, the SB2 SI New method had the lowest absolute value of mean *ME*/mean demand among all methods.

For item A and at “issue points only,” all existing methods overestimate the sales. On the other hand, all new methods underestimate the sales, except for the Croston SI New and SB2 SI New methods with α =0.05. The lowest absolute value of mean *ME*/mean demand of the Croston and SB2 methods is achieved with α =0.2, whereas the lowest absolute value of mean *ME*/mean demand of the SES and Poisson SES methods is achieved with α =0.05. In addition, the lowest absolute value of mean *ME*/mean demand of all new methods is achieved with α =0.05, except for the Croston SI New method. Moreover, the SB2 SI New method had the lowest absolute value of mean *ME*/mean demand among all methods. All new methods perform better than the old methods at “issue points only.”

Table 6.1.2 Mean *ME*/mean demand of item B

Item B, Mean *ME*/Mean Demand								
	All Points in Time				Issue Points Only			
Methods	$\alpha=0.05$	$\alpha=0.1$	$\alpha=0.15$	$\alpha=0.2$	$\alpha=0.05$	$\alpha=0.1$	$\alpha=0.15$	$\alpha=0.2$
Existing methods								
Croston	-67.98%	-61.34%	-57.63%	-55.86%	-70.14%	-59.95%	-53.70%	-50.00%
SB2	-63.78%	-53.28%	-45.81%	-40.28%	-65.88%	-51.94%	-42.16%	-34.99%
SES	-29.31%	-15.31%	-10.03%	-7.43%	-70.30%	-92.20%	-123.95%	-158.72%
Poisson SES	-26.67%	-13.92%	-10.71%	-7.98%	-70.02%	-92.56%	-125.55%	-162.64%
New methods								
Croston SI New	-13.06%	7.69%	13.95%	15.25%	-23.51%	-4.49%	-0.78%	-2.39%
SB2 SI New	-10.24%	12.30%	20.40%	23.72%	-20.42%	0.74%	6.79%	7.86%
Croston SI New CF	24.39%	41.68%	43.80%	42.79%	11.47%	24.32%	20.40%	13.09%
SB2 SI New CF	25.13%	44.60%	48.01%	48.50%	12.96%	28.11%	26.38%	21.80%

For item B and for "all points in time," all existing methods overestimate the sales. On the other hand, all new methods underestimate the sales, except for the Croston SI New and SB2 SI New methods with α=0.05. In addition, the lowest absolute value of mean *ME*/Mean demand of the existing methods is achieved with α=0.2, whereas the lowest absolute value of mean *ME*/Mean demand of the new methods is achieved with α=0.05, except for the Croston SI New method. Moreover, the SES method had the lowest absolute value of mean *ME*/Mean demand among all methods.

For item B and at "issue points only," all existing methods overestimate the sales. On the other hand, all new methods underestimate the sales, except for the Croston SI New method all α values and the SB2 SI New method with α=0.05. The lowest absolute value of mean *ME*/mean demand of the Croston and SB2 methods is achieved with α=0.2, whereas the lowest absolute value of mean *ME*/mean demand of the SES and Poisson SES methods is achieved with α =0.05. Moreover, the SB2 SI New method had the lowest absolute value of mean *ME*/mean demand among all methods. All new methods perform better than the old methods at "issue points only."

Table 6.1.3 displays the *ME* rank of methods using the lowest absolute mean *ME*/mean demand value among all α values for "all points in time" and at "issue points only."

Table 6.1.3 Ranking of methods based on *ME*

Rank	All Points in Time		Issue Points Only	
	Item A	Item B	Item A	Item B
1	SB2 SI New	SES	SB2 SI New	SB2 SI New
2	Croston SI New	Croston SI New	Croston SI New	Croston SI New
3	SES	Poisson SES	Croston SI New CF	Croston SI New CF
4	Poisson SES	SB2 SI New	SB2 SI New CF	SB2 SI New CF
5	Croston SI New CF	Croston SI New CF	Croston	SB2
6	SB2 SI New CF	SB2 SI New CF	SB2	Croston
7	SB2	SB2	SES	Poisson SES
8	Croston	Croston	Poisson SES	SES

6.2 *MAE* Results

The 99% confidence intervals of the *MAE* for each method calculated using equations (6.2) and (6.3) are presented separately for Item A and B as measured at "all points in time" and at "issue points only" in the Figures 6.2.1, 6.2.2, 6.2.3, and 6.2.4 below.

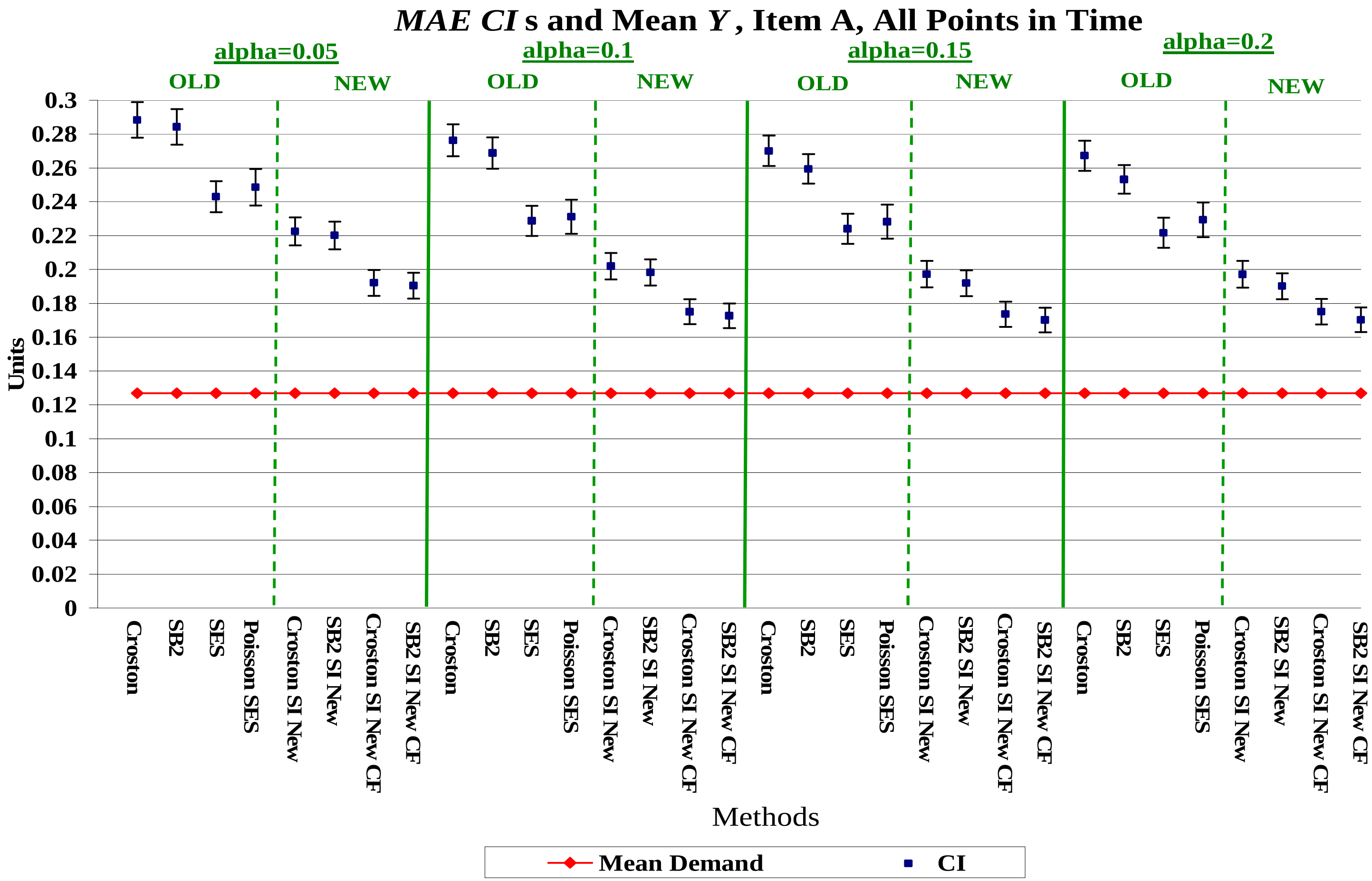

Figure 6.2.1 *MAE* confidence intervals and mean demand of Item A, All Points in Time

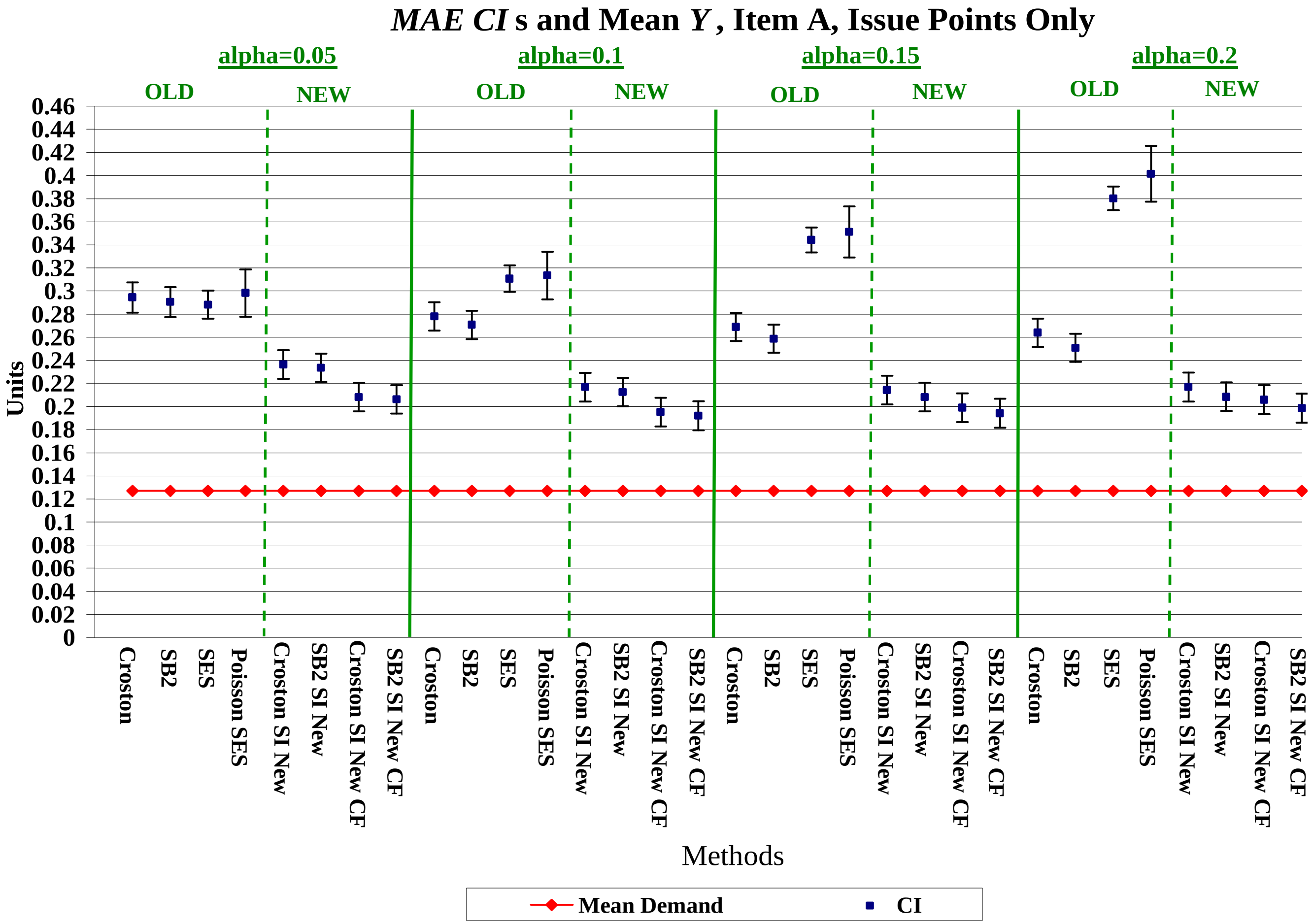

Figure 6.2.2 *MAE* confidence intervals and mean demand of Item A, Issue Points Only

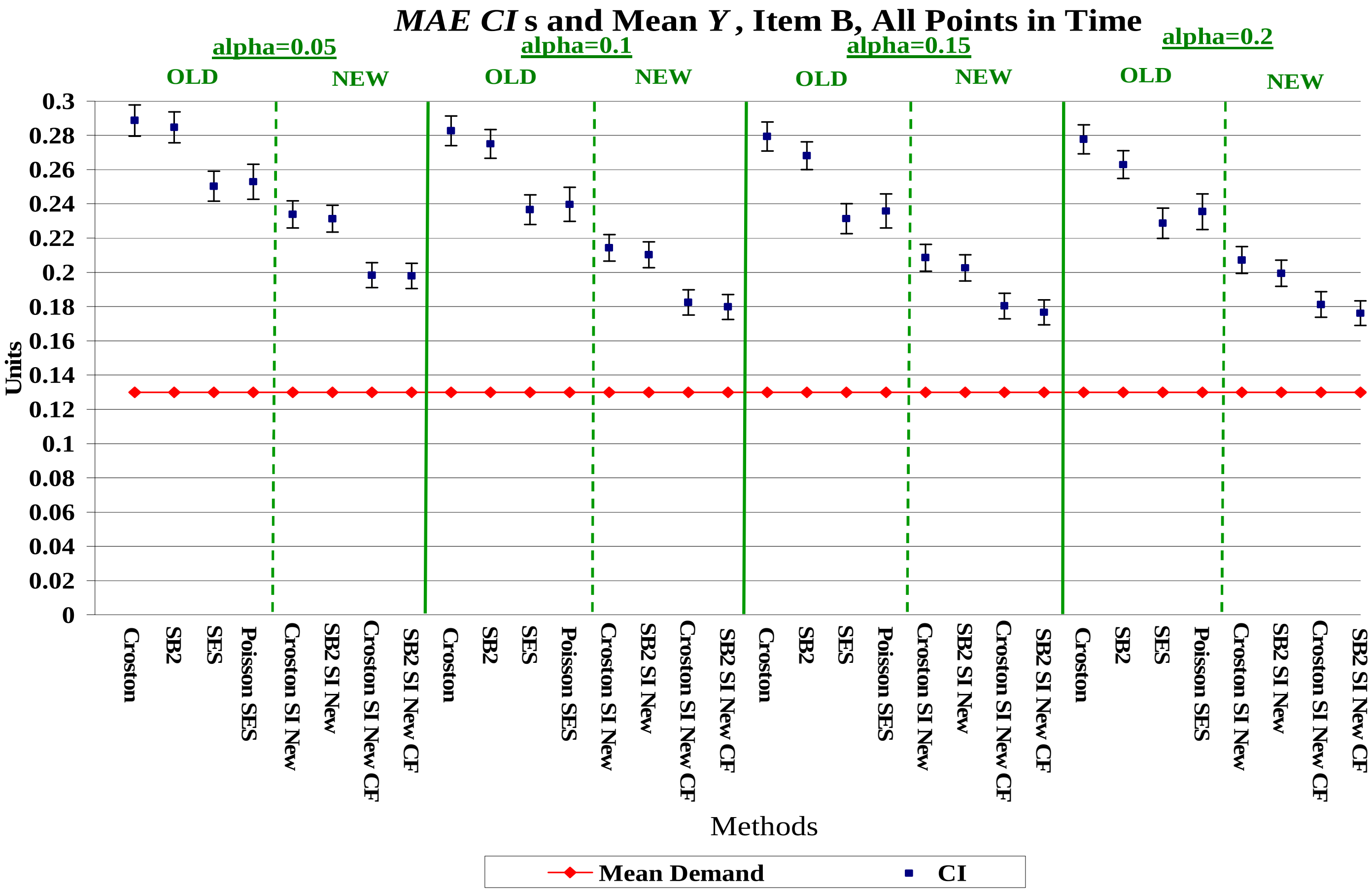

Figure 6.2.3 *MAE* confidence intervals and mean demand of Item B, All Points in Time

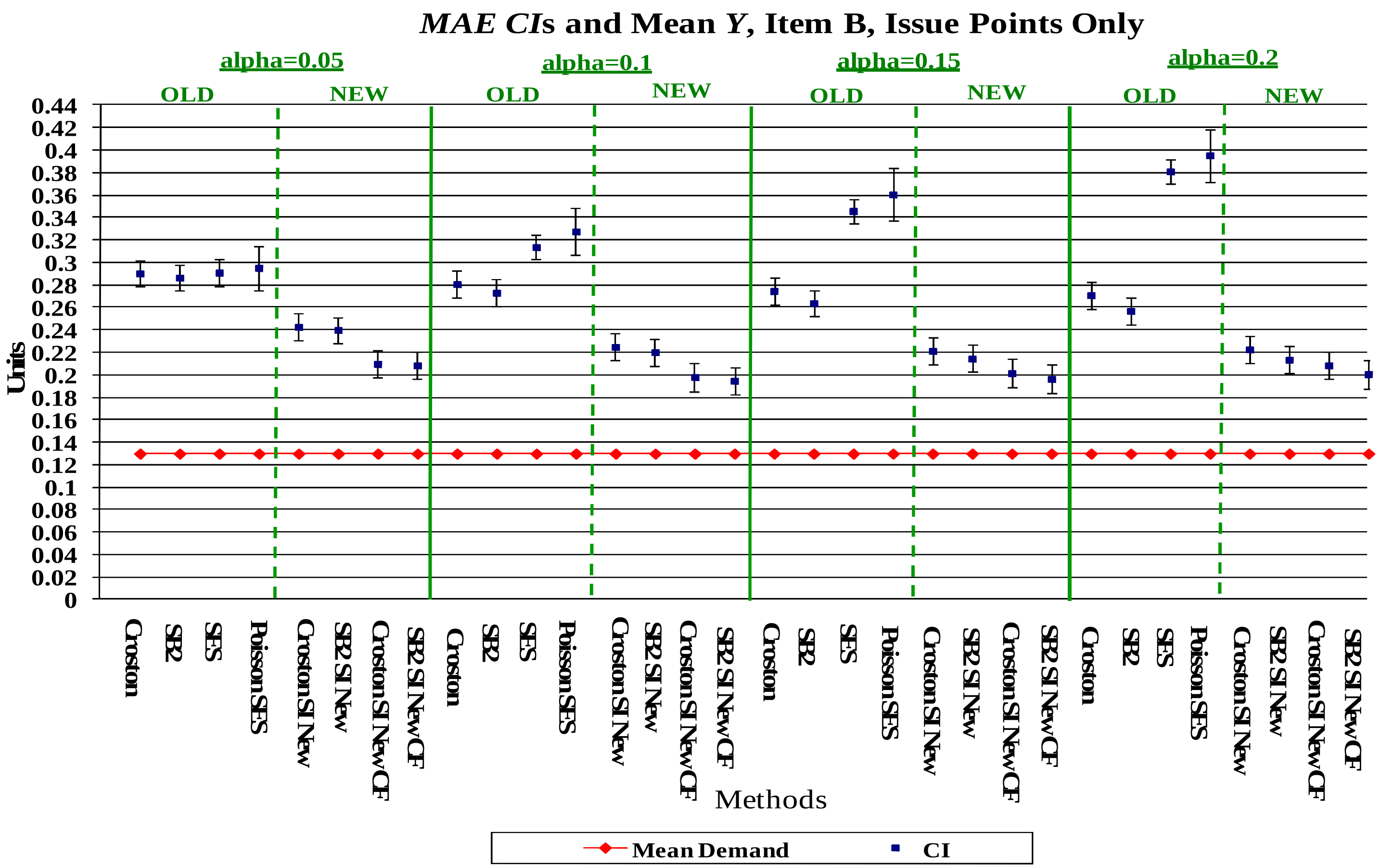

Figure 6.2.4 *MAE* confidence intervals and mean demand of Item B, Issue Points Only

Similar to the *ME* measure, the confidence intervals are very tight, with some methods consistently different than others. Below we provide a list of other observations that can be made.

First observations about the performance of the existing methods are made.

- The Croston and SB2 methods have very similar performance although SB2 generally yields a slightly lower *MAE* than Croston, but this result is not statistically significant.
- The SES and Poisson SES methods have similar performance.
- The performance of the SES and Poisson SES methods is better than Croston and SB2 at “all points in time” while the opposite is true at “issue points only”, except for α =0.05.

Next we make observations relative to the new methods.

- The Croston SI New and SB2 SI New methods have very similar performance, although the SB2 SI New method generally yields a slightly lower *MAE* than the Croston SI New method, but this result is not statistically significant.
- The Croston SI New CF and SB2 SI New CF methods have very similar performance, although the SB2 SI New CF method generally yields a slightly lower *MAE* than the Croston SI New CF method, but this result is not statistically significant.
- The Croston SI New CF and SB2 SI New CF methods perform better than the Croston SI New and SB2 SI New methods for both items and “all points in time”

and at “issue points only” with α =0.05. The new methods have similar performance for both items and at “issue points only” with α =0.1, 0.15, and 0.2, although the Croston SI New CF and SB2 SI New CF methods generally yield a slightly lower *MAE* than the Croston SI New and SB2 SI New methods, , but this result is not statistically significant.

Additional observations include the following.

- The new methods consistently perform better than the existing methods for both items at all levels of smoothing coefficient (α) both at “all points in time” and “issue points only.”
- For both items and “all points in time”, and for both items and at “issue points only” with α =0.05, the Croston SI New CF and SB2 SI New CF methods perform better than all other methods. In addition, for both items, “all points in time” the Croston and SB2 methods perform worse than all other methods.
- For both items, at “issue points only” with α =0.1, 0.15, and 0.2, the new methods have similar performance, and the SES and Poisson SES methods perform the worst.

In addition to the confidence intervals of the *MAE*, the paired-sample *t*-test results of *MAE* in Appendix A.2 calculated using the equation (6.4) are used to determine whether one method is more biased than the other one. The following statements are based on the observations of these paired-sample *t*-test results and they provide better comparison information than the confidence intervals when the confidence intervals overlap.

Furthermore, these statements are based on the statistical significant differences, not based on the practical significances.

First observations about the performance for "all points in time" are made.

- For both items and "all points in time", the forecasting methods are ranked from highest *MAE* to lowest *MAE* generating method as: Croston, SB2, SES / Poisson SES (The SES method has lower *MAE* than the Poisson SES method, but this results is not statistically significantly different for all α values and items), Croston SI New, SB2 SI New, Croston SI New CF, and SB2 SI New CF methods.

Next we make observations at "issue points only".

- For both items the Croston method yields statistically significantly higher value of *MAE* than all the new methods and the SB2 method. However, the Croston method yields statistically significantly lower values of *MAE* than the SES and Poisson SES methods for item B except for α =0.05. In addition, the Croston method yields statistically significantly lower values of *MAE* than the SES and Poisson SES methods for item A except for the Poisson SES method and α =0.05.
- For both items, the SB2 method yields statistically significantly higher values of *MAE* than all the new methods. However, similar to Croston's method the SB2 method yields statistically significantly lower values of *MAE* than the SES and Poisson SES methods for item A except for α =0.05. In addition, the SB2 method yields statistically significantly lower values of *MAE* than the SES and Poisson SES methods for item B except for α =0.05 and the Poisson SES method.

- The SES and Poisson SES methods are not statistically significantly different except for item A and α =0.2.
- The SES and Poisson SES methods yield statistically significantly higher values of *MAE* than the new methods.
- For both items, the new methods perform better than the old methods.
- For both items, the new forecasting methods are ranked from highest *MAE* generating method to lowest *MAE* generating method as: Croston SI New, SB2 SI New, Croston SI New CF, and SB2 SI New CF methods.

Since the statistical tests for pairwise comparisons indicate that any differences in the performance metrics are statistically significant and we can estimate the average *MAE* with excellent accuracy due to large sample sizes, we focus on ranking the methods using the lowest absolute value of mean *MAE*/mean demand among all α values for "all points in time" and at "issue points only" separately.

The lowest absolute values of each method for "all points in time" and at "issue points only" are highlighted below.

Table 6.2.1 Mean *MAE*/mean demand of item A

Item A, Mean *MAE*/Mean Demand								
	All Points in Time				Issue Points Only			
Methods	α =0.05	α =0.1	α =0.15	α =0.2	α =0.05	α =0.1	α =0.15	α =0.2
Existing methods								
Croston	227.16%	217.67%	212.80%	210.56%	232.02%	219.00%	211.85%	207.94%
SB2	223.98%	211.78%	204.34%	199.50%	228.83%	213.28%	203.80%	197.60%
SES	191.46%	180.22%	176.52%	174.65%	227.08%	244.78%	271.18%	299.47%
Poisson SES	195.85%	182.15%	179.82%	180.68%	235.04%	246.93%	276.68%	316.28%

Table 6.2.1 Continued

New methods								
Croston SI New	175.31%	159.10%	155.40%	155.33%	186.16%	170.77%	168.88%	170.84%
SB2 SI New	173.42%	156.14%	151.24%	149.80%	184.11%	167.46%	164.05%	164.21%
Croston SI New CF	151.33%	137.91%	136.84%	137.97%	164.01%	153.70%	156.80%	162.20%
SB2 SI New CF	150.04%	136.02%	134.08%	134.18%	162.52%	151.24%	152.89%	156.43%

For item A and for “all points in time,” all mean *MAE*/mean demand values are greater than 100%. The lowest absolute value of mean *MAE*/mean demand of all methods is achieved with α =0.2 except for the Poisson SES method. All new methods are better than existing methods. Moreover, the SB2 SI New CF method had the lowest absolute value of mean *MAE*/mean demand among all methods.

For item A and at “issue points only,” all mean *MAE*/ mean demand values are greater than 100%. All new methods are better than existing methods. Moreover, the SB2 SI New CF method had the lowest absolute value of mean *MAE*/mean demand among all methods.

Table 6.2.2 Mean *MAE*/mean demand of item B

Item B, *MAE* Confidence Interval and Mean *MAE*/Mean Demand								
	All Points in Time				Issue Points Only			
Methods	α =0.05	α =0.1	α =0.15	α =0.2	α =0.05	α =0.1	α =0.15	α =0.2
Existing methods								
Croston	221.66%	217.04%	214.47%	213.23%	223.09%	215.61%	210.93%	208.07%
SB2	218.61%	211.19%	205.89%	201.91%	220.02%	209.84%	202.63%	197.28%
SES	192.22%	181.72%	177.67%	175.64%	223.55%	240.96%	265.65%	292.63%
Poisson SES	194.24%	184.05%	181.07%	180.81%	226.73%	251.67%	276.87%	303.71%

Table 6.2.2 Continued

New methods								
Croston SI New	179.56%	164.66%	160.14%	159.13%	186.38%	172.69%	169.97%	171.03%
SB2 SI New	177.57%	161.43%	155.62%	153.21%	184.22%	169.07%	164.72%	163.93%
Croston SI New CF	152.33%	140.09%	138.54%	139.20%	161.04%	152.07%	154.89%	160.10%
SB2 SI New CF	151.98%	138.09%	135.65%	135.29%	160.14%	149.47%	150.78%	154.09%

For item B and for "all points in time," all mean *MAE*/ mean demand values are greater than 100%. The lowest absolute value of mean *MAE*/mean demand of all methods is achieved with α =0.2. All new methods are better than existing methods. Moreover, the SB2 SI New CF method had the lowest absolute value of mean *MAE*/mean demand among all methods.

For item A and at "issue points only," mean *MAE*/mean demand values are greater than 100%. All new methods are better than the existing methods. Moreover, the SB2 SI New CF method had the lowest mean *MAE*/mean demand value among all methods.

Table 6.2.3 displays the *MAE* rank of methods using the lowest mean *MAE*/mean demand value among all α values for "all points in time" and at "issue points only."

Table 6.2.3 Ranking of methods based on *MAE*

Rank	All Points in Time		Issue Points Only	
	Item A	Item B	Item A	Item B
1	SB2 SI New CF	SB2 SI New CF	SB2 SI New CF	SB2 SI New CF
2	Croston SI New CF	Croston SI New CF	Croston SI New CF	Croston SI New CF
3	SB2 SI New	SB2 SI New	SB2 SI New	SB2 SI New
4	Croston SI New	Croston SI New	Croston SI New	Croston SI New
5	SES	SES	SB2	SB2
6	Poisson SES	Poisson SES	Croston	Croston
7	SB2	SB2	SES	SES
8	Croston	Croston	Poisson SES	Poisson SES

6.3 *SMAE* Results

The 99% confidence intervals of the *SMAE* for each method calculated using equations (6.2) and (6.3) are presented separately for Item A and B as measured at "all points in time" and at "issue points only" in the Figures 6.3.1, 6.3.2, 6.3.3, and 6.3.4 below.

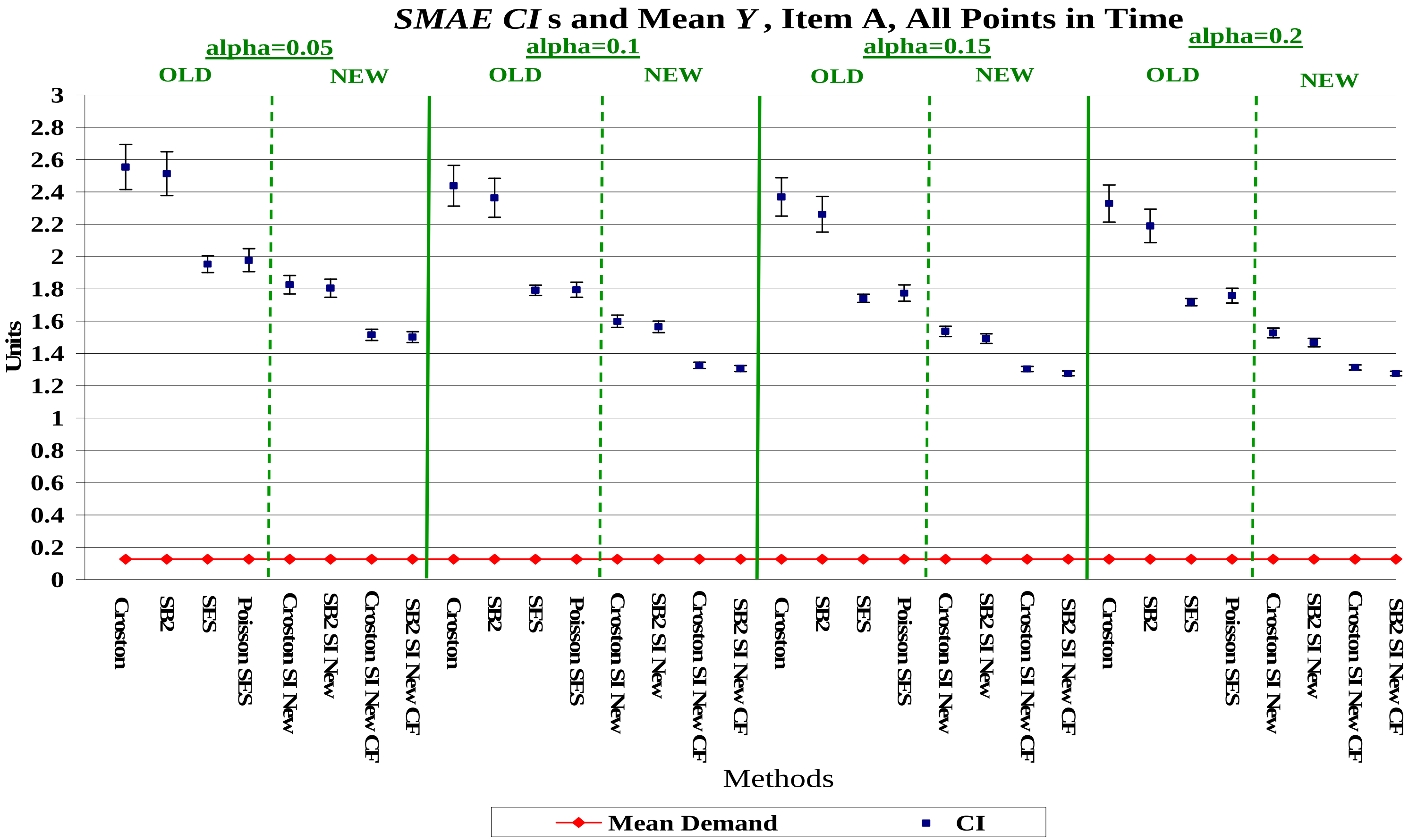

Figure 6.3.1 *SMAE* confidence intervals and mean demand of Item A, All Points in Time

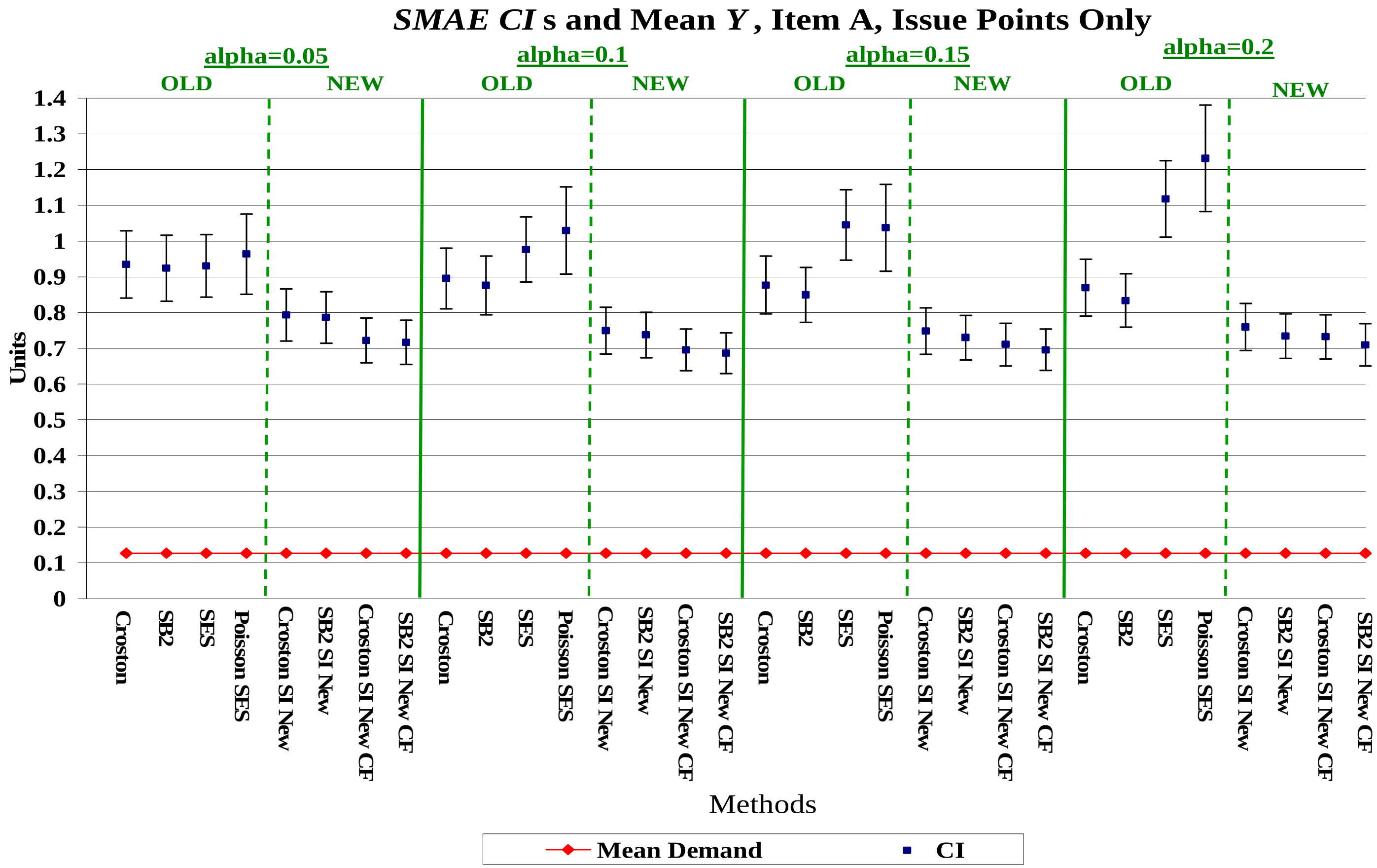

Figure 6.3.2 *SMAE* confidence intervals and mean demand of Item A, Issue Points Only

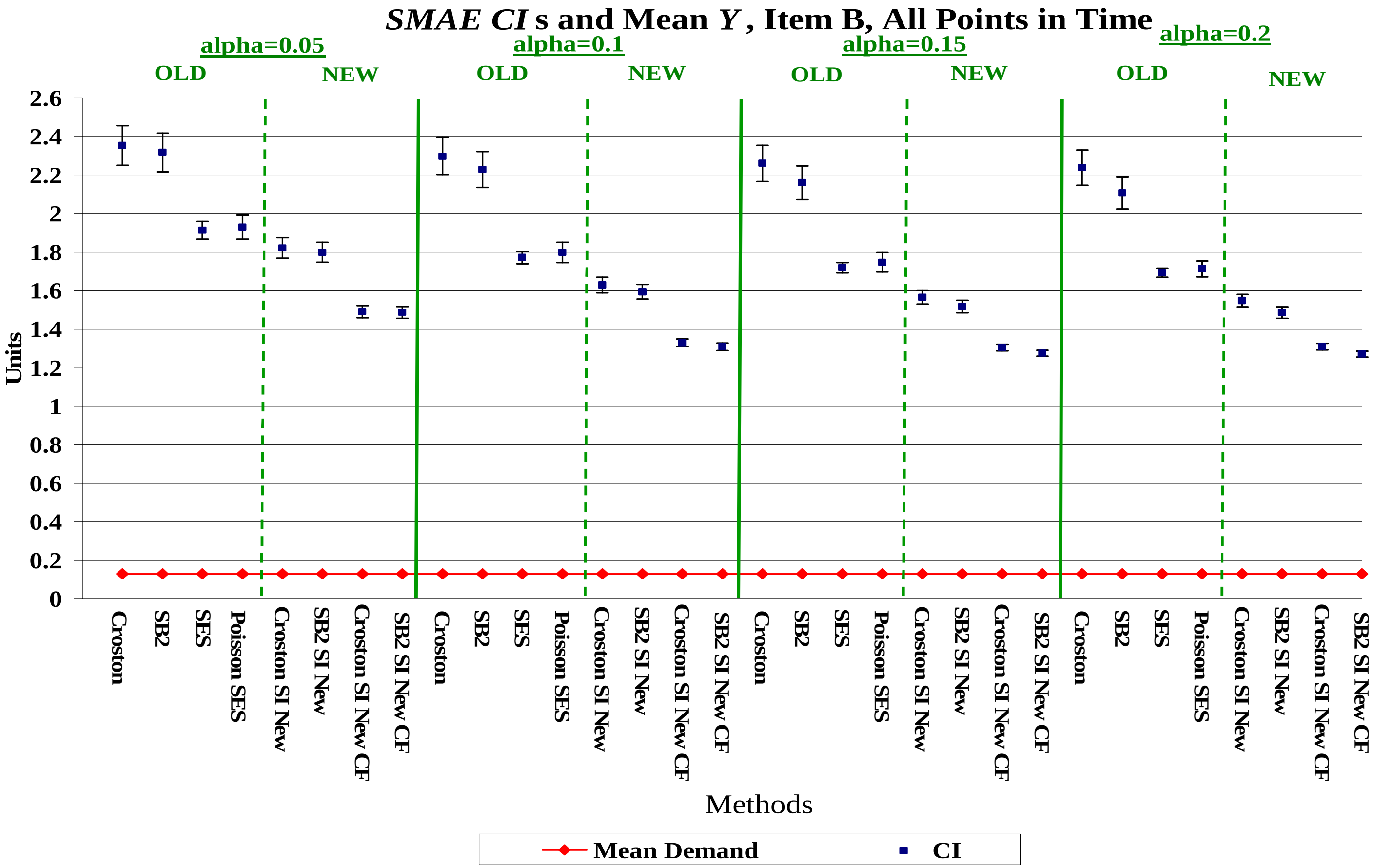

Figure 6.3.3 *SMAE* confidence intervals and mean demand of Item B, All Points in Time

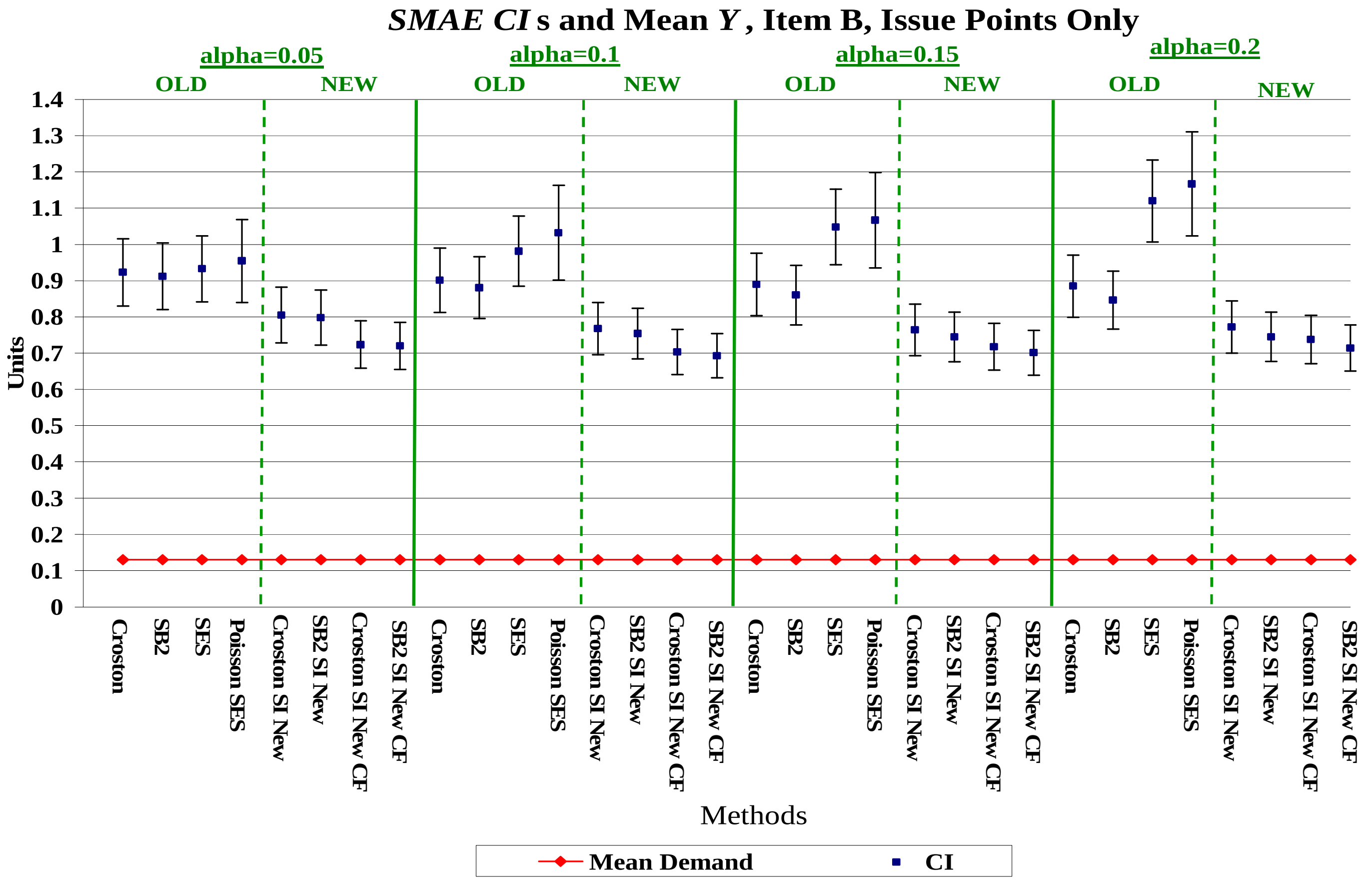

Figure 6.3.4 *SMAE* confidence intervals and mean demand of Item B, Issue Points Only

Below we provide a list of other observations that can be made using the confidence interval figures 6.3.1, 6.3.2, 6.3.3, and 6.3.4.

First observations about the performance of the existing methods are made.

- The Croston and SB2 methods have very similar performance although SB2 generally yields a slightly lower value of *SMAE* than Croston, but this result is not statistically significant.
- The SES and Poisson SES methods have similar performance.
- The performance of the SES and Poisson SES methods is better than Croston and SB2 at "all points in time," while the opposite is true at "issue points only" and α = 0.2. In addition, the Croston and SB2 methods generally yield a slightly lower value of *SMAE* than the SES and Poisson SES methods at "issue points only".

Next we make observations relative to the new methods.

- The Croston SI New and SB2 SI New methods have very similar performance although the SB2 SI New method generally yields a slightly lower value of *SMAE* than the Croston SI New method, but this result is not statistically significant.
- The Croston SI New CF and SB2 SI New CF methods have very similar performance, although the SB2 SI New CF method generally yields a slightly lower value of *SMAE* than the Croston SI New CF method, but this result is not statistically significant.
- The Croston SI New CF and SB2 SI New CF methods perform better than the Croston SI New and SB2 SI New methods for both items and "all points in time."

The new methods have similar performance for both items and at "issue points only," although the Croston SI New CF and SB2 SI New CF methods generally yield a slightly lower value of *SMAE* than the Croston SI New and SB2 SI New methods, but this result is not statistically significant.

Additional observations include the following.

- For both items and "all points in time," the forecasting methods are ranked from highest *SMAE* to lowest *SMAE* generating method as: The Croston/SB2, SES / Poisson SES, Croston SI New/SB2 SI New, and Croston SI New CF/ SB2 SI New CF methods.
- For both items and "all points in time", and for both items and at "issue points only" except for α =0.2, the Croston SI New CF and SB2 SI New CF methods perform better than the old methods. In addition, for both items "all points in time" the Croston and SB2 methods perform worse than all other methods.
- For both items, at "issue points only" with α =0.2, the SES and Poisson SES methods perform the worst.

In addition to the confidence intervals of the *SMAE*, the paired-sample *t*-test results of *SMAE* in Appendix A.2 calculated using the equation (6.4) are used to determine whether one method is more biased than the other one. The following statements are based on the observations of these paired-sample *t*-test results and they provide better comparison information than the confidence intervals when the confidence intervals overlap.

Furthermore, these statements are based on the statistical significant differences, not based on the practical significances.

First observations about the performance for “all points in time” are made.

- For both items and “all points in time,” the forecasting methods are ranked from highest *SMAE* to lowest *SMAE* generating method as: Croston, SB2, SES / Poisson SES (the SES method has lower *SMAE* than the Poisson SES method, but this result is not statistically significant), Croston SI New, SB2 SI New, Croston SI New CF, and SB2 SI New CF methods.

Next we make observations at “issue points only.”

- For both items the Croston method yields statistically significantly higher values of *SMAE* than all the new methods and the SB2 method. However, the Croston method yields statistically significantly lower values of *SMAE* than the SES and Poisson SES methods for both items except for α =0.05.
- For both items, the SB2 method yields statistically significantly higher values of *SMAE* than all the new methods. However, similar to Croston’s method the SB2 method yields statistically significantly lower values of *SMAE* than the SES and Poisson SES methods for item A except for α =0.05. In addition, the SB2 method yields statistically significantly lower values of *SMAE* than the SES and Poisson SES methods for item B except for α =0.05 and Poisson SES method.
- The SES and Poisson SES methods are not statistically significantly different except for item A and α =0.2.

- The SES and Poisson SES methods yield statistically significantly higher values of *SMAE* than the new methods.
- For both items, the new methods perform better than the old methods.
- For both items, the new forecasting methods are ranked from highest value of *SMAE* to lowest value as follows: the Croston SI New, SB2 SI New, Croston SI New CF, and SB2 SI New CF methods.

Since the statistical tests for pairwise comparisons indicate that any differences in the performance metrics are statistically significant, we focus on ranking the methods using the lowest mean *SMAE*/mean demand value among all α values for "all points in time" and at "issue points only" separately. The lowest values of each method for "all points in time" and at "issue points only" are highlighted below.

Table 6.3.1 Mean *SMAE*/mean demand of item A

Item A, Mean *SMAE*/Mean Demand								
	All Points in Time				Issue Points Only			
Methods	α =0.05	α =0.1	α =0.15	α =0.2	α =0.05	α =0.1	α =0.15	α =0.2
Existing methods								
Croston	2011.80%	1921.25%	1866.48%	1834.31%	736.50%	705.31%	690.90%	685.32%
SB2	1979.90%	1862.03%	1781.77%	1724.58%	728.08%	690.05%	669.08%	656.79%
SES	1538.84%	1411.03%	1371.48%	1353.11%	733.04%	769.22%	823.31%	880.53%
Poisson SES	1557.67%	1414.14%	1397.72%	1384.82%	759.00%	811.20%	817.28%	969.90%
New methods								
Croston SI New	1438.91%	1259.48%	1211.11%	1202.98%	625.04%	590.39%	589.49%	598.37%
SB2 SI New	1421.32%	1233.35%	1175.57%	1156.45%	619.39%	580.85%	575.24%	578.51%
Croston SI New CF	1193.94%	1044.35%	1027.38%	1034.92%	568.77%	548.13%	559.99%	576.74%
SB2 SI New CF	1182.51%	1029.00%	1005.68%	1005.27%	564.54%	540.72%	548.00%	559.07%

For item A and for “all points in time,” the lowest absolute value of mean *SMAE*/mean demand of all methods is achieved with α =0.2. All new methods are better than the existing methods. Moreover, the SB2 SI New CF method had the lowest absolute value of mean *SMAE*/mean demand among all methods.

For item A and at “issue points only,” all new methods are better than the existing methods. Moreover, the SB2 SI New CF method had the lowest absolute value of mean *SMAE*/mean demand among all methods.

Table 6.3.2 Mean *SMAE*/mean demand of item B

Item B, Mean *SMAE*/Mean Demand								
	All Points in Time				Issue Points Only			
Methods	α =0.05	α =0.1	α =0.15	α =0.2	α =0.05	α =0.1	α =0.15	α =0.2
Existing methods								
Croston	1808.26%	1765.48%	1737.41%	1719.87%	708.68%	691.81%	683.21%	679.62%
SB2	1780.81%	1712.72%	1660.38%	1618.91%	700.54%	676.39%	660.72%	649.99%
SES	1469.75%	1361.17%	1320.81%	1301.25%	716.07%	753.48%	804.85%	860.25%
Poisson SES	1483.18%	1381.73%	1342.45%	1316.06%	732.91%	792.66%	819.10%	896.14%
New methods								
Croston SI New	1399.41%	1251.76%	1202.81%	1189.60%	618.51%	589.46%	586.89%	592.94%
SB2 SI New	1382.18%	1224.68%	1165.88%	1141.70%	612.63%	579.15%	571.60%	571.94%
Croston SI New CF	1145.88%	1021.37%	1001.91%	1005.61%	555.48%	539.97%	551.17%	566.62%
SB2 SI New CF	1142.40%	1005.82%	980.07%	976.16%	552.84%	532.13%	538.56%	548.25%

For item B and for “all points in time”, the lowest absolute value of mean *SMAE*/mean demand of all methods is achieved with α =0.2 except for the Croston SI New CF method. All new methods are better than the existing methods. Moreover, SB2 SI New CF method had the lowest absolute value of mean *SMAE*/mean demand among all methods.

For item B and at “issue points only,” all new methods are better than the existing methods. Moreover, the SB2 SI New CF method had the lowest absolute value of mean *SMAE*/mean demand among all methods.

Table 6.3.3 displays the *SMAE* rank of methods using the lowest mean *SMAE*/mean demand value among all α values for “all points in time” and at “issue points only.”

Table 6.3.3 Ranking of methods based on *SMAE*

Rank	All Points in Time		Issue Points Only	
	Item A	Item B	Item A	Item B
1	SB2 SI New CF	SB2 SI New CF	SB2 SI New CF	SB2 SI New CF
2	Croston SI New CF	Croston SI New CF	Croston SI New CF	Croston SI New CF
3	SB2 SI New	SB2 SI New	SB2 SI New	SB2 SI New
4	Croston SI New	Croston SI New	Croston SI New	Croston SI New
5	SES	SES	SB2	SB2
6	Poisson SES	Poisson SES	Croston	Croston
7	SB2	SB2	SES	SES
8	Croston	Croston	Poisson SES	Poisson SES

6.4 *RGRMSE*/*GRMSE* Results

As stated before, *RGRMSE* cannot be calculated between two methods, if one of those methods’ forecasts are generated using Poisson distribution. In addition, *GRMSE* values of Poisson generated forecasts will generate the lowest *GRMSE* values by default, thus *GRMSE* values of Poisson SES will not be displayed in the figures below.

The 99% confidence intervals of the *GRMSE* for each method calculated using equations (6.2) and (6.3) are presented separately for Item A and B as measured at “all points in time” and “at issue points only” in the Figures 6.4.1, 6.4.2, 6.4.3, and 6.4.4 below.

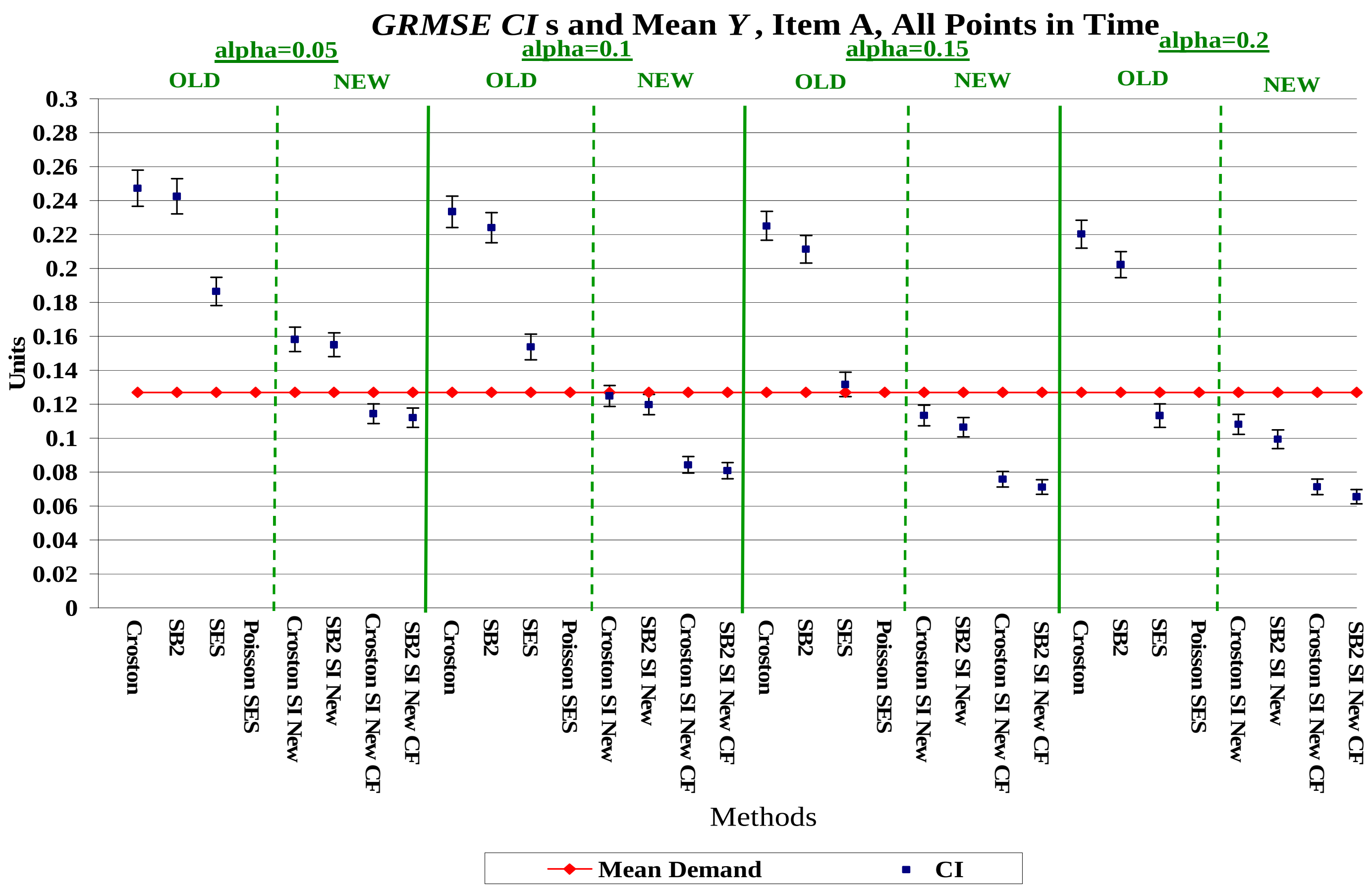

Figure 6.4.1 *GRMSE* confidence intervals and mean demand of Item A, All Points in Time

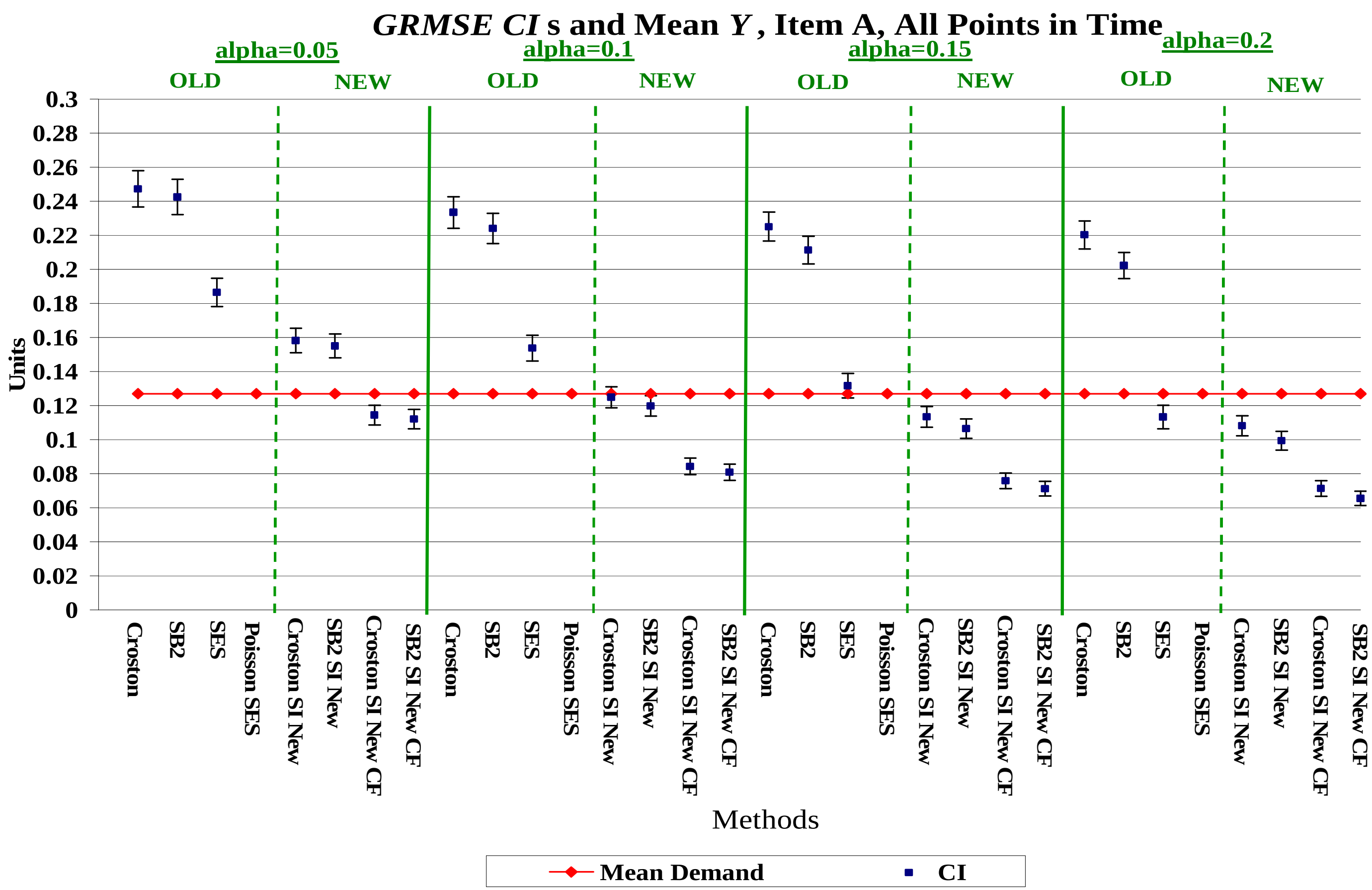

Figure 6.4.2 *GRMSE* confidence intervals and mean demand of Item A, Issue Points only

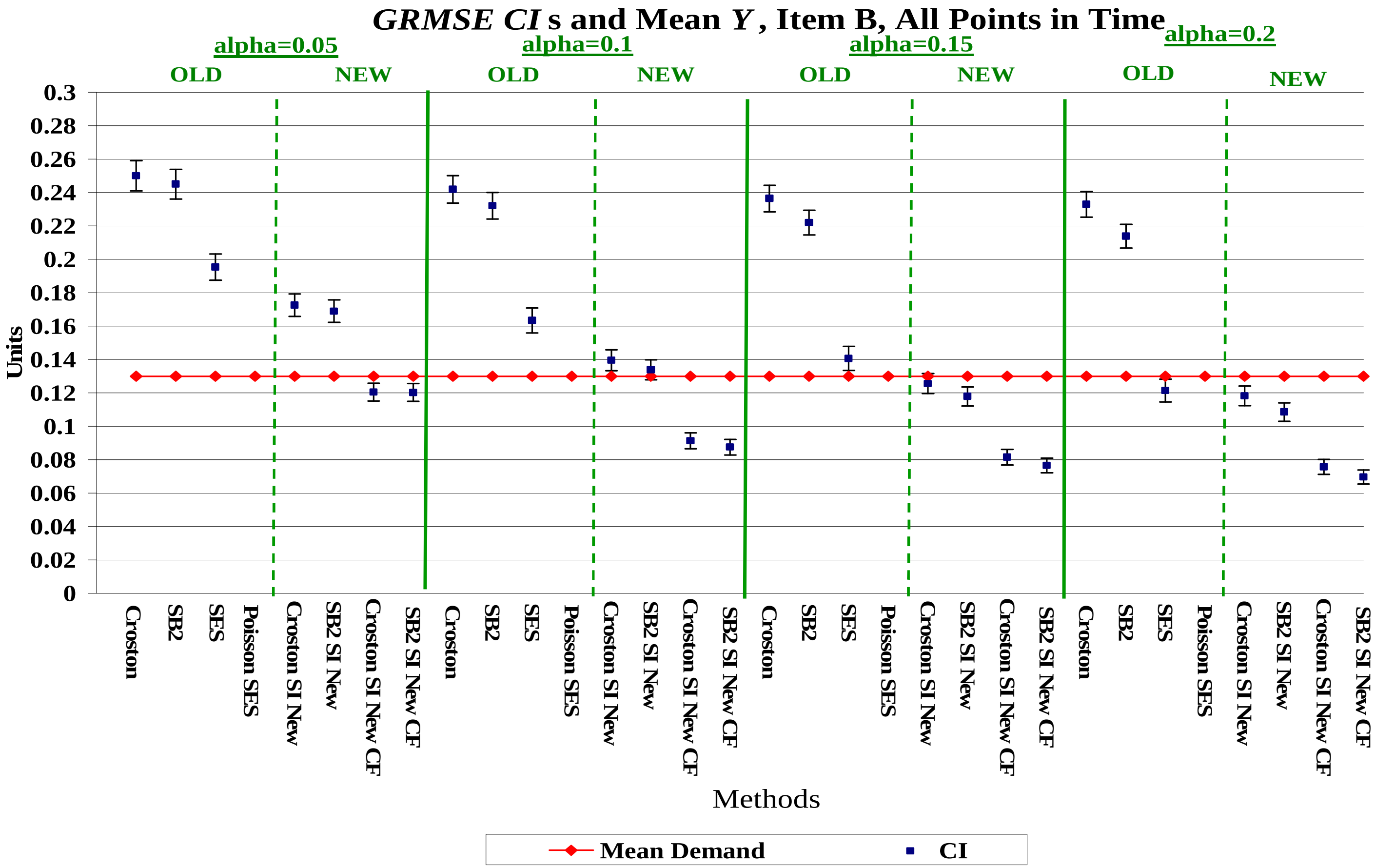

Figure 6.4.3 *GRMSE*, confidence intervals and mean demand of Item B, All Points in Time

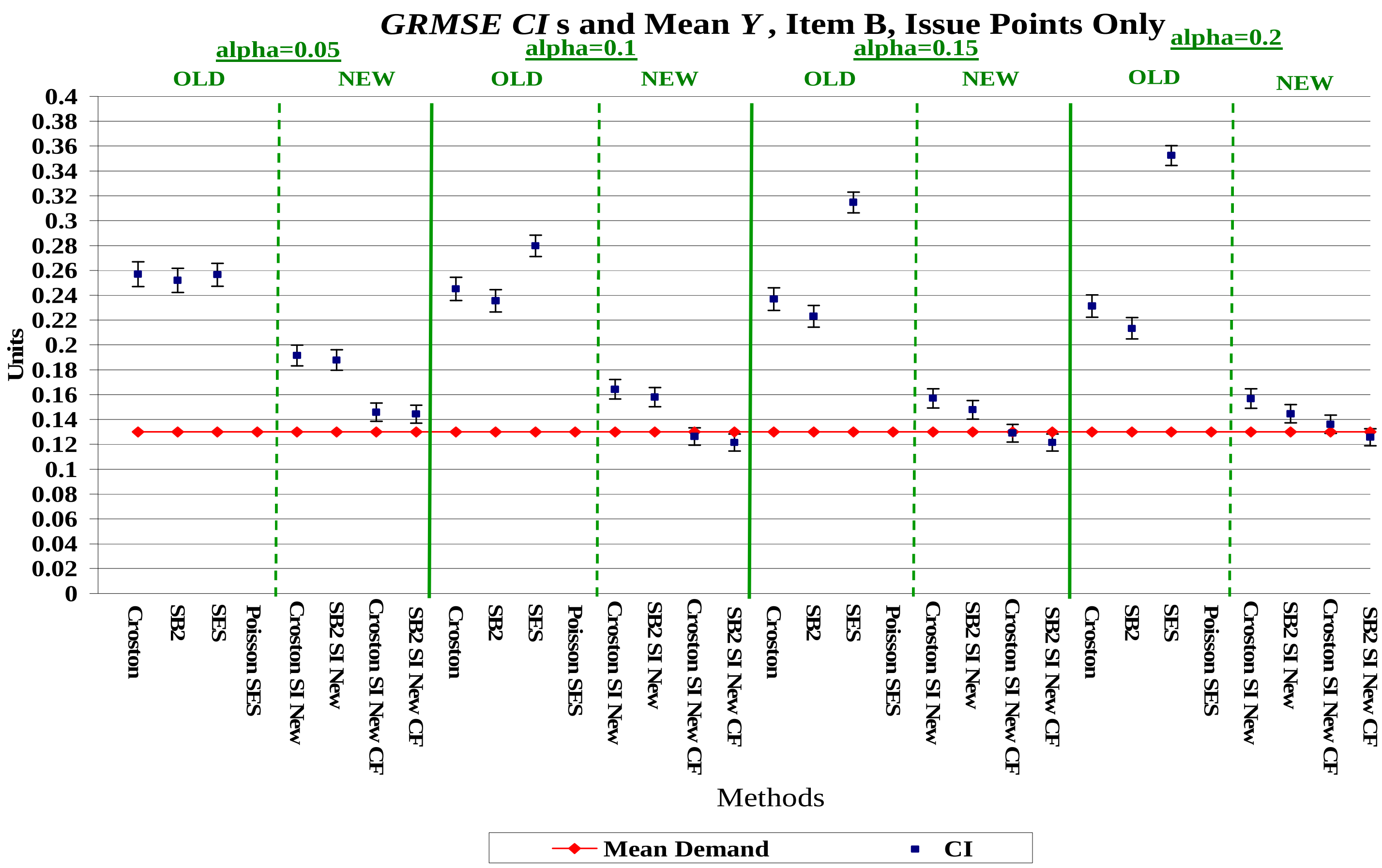

Figure 6.4.4 *GRMSE* confidence intervals and mean demand of Item B, Issue Points only

Similar to *ME* and *MAE*, the confidence intervals are very tight with some methods consistently different than others. Below we provide a list of other observations that can be made.

First observations about the performance of the existing methods are made.

- The Croston and SB2 methods have very similar performance except for α =0.02. In addition, the SB2 method performs better than the Croston method for α =0.02, and the SB2 method generally yields a slightly lower value of *GRMSE* than Croston, but this result is not statistically significant.
- The performance of the SES method is better than Croston and SB2 at “all points in time” while the opposite is true at “issue points only,” except for α = 0.05.

Next we make observations relative to the new methods.

- The Croston SI New and SB2 SI New methods have very similar performance, although the SB2 SI New method generally yields a slightly lower value of *GRMSE* than the Croston SI New method, but this result is not statistically significant.
- The Croston SI New CF and SB2 SI New CF methods have very similar performance, although the SB2 SI New CF method generally yields a slightly lower value of *GRMSE* than the Croston SI New CF method, but this result is not statistically significant.
- The Croston SI New CF and SB2 SI New CF methods perform better than the Croston SI New and SB2 SI New methods for both items and “all points in time,”

for item A at “issue points only” with α =0.05 and 0.1, and for item B at “issue points only” except for α =0.2. Furthermore, when the confidence intervals of the new methods overlap, the Croston SI New CF and SB2 SI New CF methods generally yield a slightly lower value of *GRMSE* than the Croston SI New and SB2 SI New methods, but this result is not statistically significant.

Additional observations include the following.

- The new methods consistently perform better than the existing methods for both items at all levels of smoothing coefficient (α) both at “all points in time” and “issue points only.”
- For both items and “all points in time,” and for both items and at “issue points only” with $\alpha = 0.05$, 0.1, and for item B and at “issue points only” with $\alpha = 0.15$, the Croston SI New CF and SB2 SI New CF methods perform better than all other methods. In addition, for both items and “all points in time” the Croston and SB2 methods perform worse than all other methods.
- For both items, at “issue points only” with $\alpha = 0.1$, 0.15, and 0.2, the SES method performs the worst.

In addition to the confidence intervals of the *GRMSE*, the paired-sample *t*-test results of *GRMSE* in Appendix A.2 calculated using the equation (6.4) are used to determine whether one method is more biased than the other one. The following statements are based on the observations of these paired-sample *t*-test results and they provide better comparison information than the confidence intervals when the confidence intervals overlap.

Furthermore, these statements are based on the statistical significant differences, not based on the practical significances.

First observations about the performance for “all points in time” are made.

- For both items and “all points in time”, the forecasting methods are ranked from highest value of *GRMSE* to lowest value as follows: The Croston, SB2, SES, Croston SI New, SB2 SI New, Croston SI New CF, and SB2 SI New CF methods.

Next we make observations at “issue points only.”

- For both items the Croston method yields statistically significantly higher values of *GRMSE* than all the new methods and the SB2 method. However, the Croston method yields statistically significantly lower values of *GRMSE* than the SES method except for $\alpha = 0.05$.
- For both items, the SB2 method yields statistically significantly higher values of *GRMSE* than all the new methods. However, similar to Croston’s method the SB2 method yields statistically significantly lower values of *GRMSE* than the SES method except for $\alpha = 0.05$.
- The SES method yields statistically significantly higher values of *GRMSE* than the new methods except for the Croston SI New method “all points in time” and $\alpha = 0.2$.
- For both items, the new methods perform better than the old methods.

- For both items, the new forecasting methods are ranked from highest value of *GRMSE* to lowest value as follows: the Croston SI New, SB2 SI New, Croston SI New CF, and SB2 SI New CF methods.

Since the statistical tests for pairwise comparisons indicate that any differences in the performance metrics are statistically significant, we focus on ranking the methods using the mean *GRMSE* among all α values for "all points in time" and at "issue points only" separately. The lowest values of each method for "all points in time" and at "issue points only" are highlighted below.

Table 6.4.1 Mean *GRMSE* of item A

Item A, Mean *GRMSE*								
	All Points in Time				Issue Points Only			
Methods	α =0.05	α =0.1	α =0.15	α =0.2	α =0.05	α =0.1	α =0.15	α =0.2
Existing methods								
Croston	194.78%	183.89%	177.30%	173.56%	202.10%	187.46%	178.64%	173.36%
SB2	191.02%	176.51%	166.49%	159.36%	198.29%	180.24%	168.23%	159.86%
SES	146.89%	121.13%	103.76%	89.22%	197.07%	216.73%	246.00%	276.79%
Poisson SES	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
New methods								
Croston SI New	124.65%	98.41%	89.34%	85.24%	140.37%	117.98%	114.30%	116.11%
SB2 SI New	122.12%	94.36%	83.86%	78.30%	137.67%	113.44%	107.68%	107.17%
Croston SI New CF	90.11%	66.35%	59.72%	56.16%	109.67%	93.86%	97.13%	103.60%
SB2 SI New CF	88.26%	63.64%	56.07%	51.62%	107.58%	90.27%	91.55%	95.68%

For item A and for "all points in time," the lowest absolute value of mean *GRMSE* of all methods is achieved with α =0.2. All new methods are better than the existing methods.

Moreover, the SB2 SI New CF method had the lowest absolute value of mean *GRMSE* among all methods.

For item A and at “issue points only,” all new methods are better than the existing methods. Moreover, the SB2 SI New CF method had the lowest absolute value of mean *GRMSE* among all methods.

Table 6.4.2 Mean *GRMSE* of item B

Item B, Mean *GRMSE*								
	All Points in Time				Issue Points Only			
Methods	$\alpha=0.05$	$\alpha=0.1$	$\alpha=0.15$	$\alpha=0.2$	$\alpha=0.05$	$\alpha=0.1$	$\alpha=0.15$	$\alpha=0.2$
Existing methods								
Croston	191.99%	185.75%	181.56%	178.81%	197.29%	188.19%	181.95%	177.66%
SB2	188.15%	178.23%	170.46%	164.18%	193.49%	180.87%	171.31%	163.79%
SES	150.01%	125.46%	108.03%	93.20%	196.95%	214.87%	241.65%	270.63%
Poisson SES	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
New methods								
Croston SI New	132.45%	107.14%	96.46%	90.82%	147.09%	126.21%	120.57%	120.39%
SB2 SI New	129.77%	102.75%	90.54%	83.39%	144.30%	121.35%	113.54%	111.08%
Croston SI New CF	92.46%	70.09%	62.59%	58.16%	111.95%	97.03%	99.11%	104.58%
SB2 SI New CF	92.33%	67.22%	58.75%	53.44%	110.82%	93.31%	93.38%	96.56%

For item B and for “all points in time,” the lowest absolute value of mean *GRMSE* of all methods is achieved with α =0.2. All new methods are better than the existing methods. Moreover, the SB2 SI New CF method has the lowest absolute value of mean *GRMSE* among all methods.

For item B and at "issue points only," all new methods are better than the existing methods. Moreover, the SB2 SI New CF method had the lowest absolute value of mean *GRMSE* among all methods.

Table 6.4.3 displays the *GRMSE* rank of methods using the lowest mean *GRMSE* value among all α values for "all points in time" and at "issue points only."

Table 6.4.3 Ranking of methods based on *GRMSE*

Rank	All Points in Time		Issue Points Only	
	Item A	Item B	Item A	Item B
1	SB2 SI New CF	SB2 SI New CF	SB2 SI New CF	SB2 SI New CF
2	Croston SI New CF	Croston SI New CF	Croston SI New CF	Croston SI New CF
3	SB2 SI New	SB2 SI New	SB2 SI New	SB2 SI New
4	Croston SI New	Croston SI New	Croston SI New	Croston SI New
5	SES	SES	SB2	SB2
6	SB2	SB2	Croston	Croston
7	Croston	Croston	SES	SES

6.5 Results of Sales Forecast Accuracy Metrics and Discussion

In this chapter, the existence of bias for each of the methods is tested using the *ME* statistic. In addition, the forecasting methods are compared using the *MAE*, *SMAE*, and *RGRMSE* accuracy metrics. The accuracy metrics are computed at "all points in time" and at "issue points only," for the two items, A and B separately, and for smoothing values of α =0.05, 0.1, 0.15, and 0.2.

We provide figures showing confidence intervals of each method and each sales forecast accuracy measure across all smoothing parameter values and use these confidence intervals to make comparisons between methods. Since the sample sizes are quite large, the resulting confidence intervals are very tight, making differences in performance between the forecasting methods quite obvious for the most part. However, in some cases, these confidence intervals overlap. Therefore, in addition to the confidence intervals, we also provide the results of two-sided *t*-tests in order to distinguish differences in two methods' performances. Due to the very large sample sizes available for this analysis, statistical tests indicate that any differences in the performance metrics are statistically significant. Thus, to rank methods, we compare the average performance of the methods for each accuracy metric based on "all points in time" performances as well as the performances at "issue points only." The best performance for a given method and metric is selected across the four levels of smoothing parameter value α for "all points in time" and at "issue points only" separately. We use mean *ME*/mean demand, mean *MAE*/mean demand, and mean *SMAE*/ mean demand, and mean *GRMSE* to rank methods.

One observation that needs to be mentioned again is that *GRMSE* sales forecast accuracy measure cannot be used to evaluate forecasting methods which generates integer forecasts, such as the Poisson SES. Thus, *GRMSE* could not be used to evaluate the Poisson SES method.

The new methods consistently produce lower forecasts than the existing methods for both items at all levels of the smoothing coefficient (α) both at "all points in time" and at “issue points only.” Furthermore, for both items and for “all points in time,” all existing methods overestimate the sales. On the other hand, all new methods underestimate the sales, except for the Croston SI New and SB2 SI New methods with α =0.05. Moreover, for item A the SB2 SI New method has the lowest absolute value of mean *ME*/mean demand among all methods and the SES method had the lowest absolute value of mean *ME*/mean demand among all the methods.

For both items and at “issue points only,” all existing methods overestimate the sales. On the other hand, all new methods underestimate the sales, except for the Croston SI New all alpha values and SB2 SI New methods with α =0.05. Moreover, the SB2 SI New method has the lowest absolute value of mean *ME*/mean demand among all methods. All new methods perform better than the old methods at “issue points only.”

Based on the *MAE*, *SMAE*, and *GRMSE* accuracy measures used, the new methods perform better than the old methods. In addition, the SB2 SI New CF method performs the best among all methods. Furthermore, for both items and “all points in time,” the forecasting methods are ranked from highest values of *MAE*, *SMAE*, and *GRMSE* to lowest values as follows: The Croston, SB2, SES / Poisson SES (the SES method has lower values than the Poisson SES method, but this result is not statistically significantly different and *GRMSE* cannot be measured for the Poisson SES method), Croston SI New, SB2 SI New, Croston SI New CF, and SB2 SI New CF methods.

Furthermore, for both items at "issue points only," the new forecasting methods are ranked from highest values of *MAE*, *SMAE*, and *GRMSE* to lowest values as follows: the Croston SI New, SB2 SI New, Croston SI New CF, and SB2 SI New CF methods.

In order to determine whether a method performs better than another method, one cannot only use sales forecast accuracy measures as a criterion since determination of practical significance can be difficult to determine. Instead, inventory control metrics such as adjusted margin or *GMROI* should also be taken into consideration while making comparisons, since a better sales forecast accuracy measure might not reflect better inventory metrics. This analysis is done in Chapter 7.

7 Inventory Control Analysis

In the previous chapter, the traditional forecasting accuracy measures are considered in comparing the methods. In this chapter, financial oriented measures such as fill rate, margin, inventory holding cost, adjusted margin (defined as the margin minus inventory holding cost), and gross margin return on investment (*GMROI*) are used to compare the forecasting algorithms.

7.1 A Simulation Model of the Inventory and Replenishment Process

In order to evaluate the forecasting techniques' performance in terms of the measures listed above, we simulate the orders that would be generated from each forecasting technique and the resulting inventory of an item in a store. In the next subsection, we define the ordering and replenishment policy for the simulated process. This is followed by a detailed explanation of the simulation process.

7.1.1 Inventory and Replenishment Policy

An (s, S) inventory model is used to determine the simulated order quantities in each period over the range of data for each item and store combination. We used the inventory control model of the company to find the best forecasting method based on the inventory metrics. Let i be the inventory on-hand, o be the inventory on-order, s be the reorder point, and S be order-up-to level. Then, the order quantity, q, is determined by the following (Schneider, 1978):

$$q = \begin{cases} S - i - o, & \text{if } i + o \le s, \\ 0, & \text{if } i + o > s. \end{cases}$$

Since a forecast is used to estimate future demand, the values of s and S may change over time; therefore let s_t and S_t be the reorder point and order-up-to level for period t. Given the uncertainty of demand, there are a number of ways of computing the values of the parameters (s_t and S_t) of the inventory policy. One approach is to set the value of the reorder point at period t, s_t, equal to the total forecast demand over some specified horizon plus a safety stock estimate; while the order-up-to level at period t, S_t, is equal to the total forecast demand over some larger specified horizon plus a safety stock estimate.

In general inventory models, safety stock is conventionally defined as the average level of the net stock (that is, the inventory on-hand minus backorders) just before a replenishment arrives (Silver, Pyke and Peterson (1998): p. 234). The safety stock is set to satisfy a service level requirement, which generally specifies the probability that no stockouts occur during a replenishment cycle (Silver, Pyke and Peterson (1998): p. 245). Let D be the demand over the replenishment lead time (L) plus review time (T) periods. The quantity L is the lead time between the vendor and store. Furthermore, T represents how often the inventory positions of the stores are reviewed and orders are released. Let the cycle service level (referred to as *CSL* in this dissertation) denote the required probability that no stockouts occur during the replenishment cycle of length T+L time periods. Then, the conventional method for setting ss_t, the safety stock at time t, is to solve the following equation for ss_t:

$$\text{Prob}\{D \le E[D] + ss_t\} = CSL. \tag{7.1}$$

Let $F_D(x) = \text{Prob}\{D \le x\}$ be the cumulative distribution function (c.d.f.) of *D*, with expected value $\mu_D = E[D]$ and variance $\sigma_D^2 = E[(D-\mu_D)^2]$. Then, the conventional definition of safety stock is obtained from (7.1) as

$$ss_t = F_D^{-1}(CSL) - \mu_D = k\sigma_D, \tag{7.2}$$

where the safety stock factor *k* is given by:

$$k = \frac{[F_D^{-1}(CSL) - \mu_D]}{\sigma_D}. \tag{7.3}$$

If *D* has a normal distribution, then

$$F_D^{-1}(CSL) = \mu_D + \Phi^{-1}(CSL)\sigma_D, \tag{7.4}$$

where $\Phi^{-1}(\cdot)$ is the inverse of the standard normal c.d.f. Therefore when *D* is a normal random variable, Equation (7.3) and (7.4) together yield the familiar result

$$k = \Phi^{-1}(CSL). \tag{7.5}$$

See, for example, Equation (4) of Syntetos et al. (2010).

In contrast to the conventional definition (7.2) of the safety stock ss_t, in this dissertation, we use the method used by the company providing the data which is calculated as:

$$ss_t = F_D^{-1}(CSL)\sigma_D = k'\sigma_D, \tag{7.6}$$

where σ_D is the standard deviation of demand over the replenishment lead time, and

$$k' = F_D^{-1}(CSL). \tag{7.7}$$

The differences between formulas (7.2) and (7.6) for the safety stock and between (7.3) and (7.7) for the safety factor warrant some comment. The conventional formula (7.2) for the safety stock results in a smaller value than (7.6) by the amount μ_D, the expected demand over the time interval of length *T*+*L*; and thus the safety stock defined by (7.6) is proportionally much larger than the safety stock defined by (7.2). Similarly the safety factor (7.3) is a "standardized" version of $F_D^{-1}(CSL)$, the $CSL \cdot 100\%$ percentile of the distribution of *D*. Therefore the safety factor (7.3) is dimensionless (or unitless); and the safety stock (7.2) has for its unit of measurement (scale unit) [items demanded] over the time interval of length *T*+*L* periods. By contrast, the safety factor (7.7) is expressed in terms of the measurement unit [items demanded]; and therefore the safety stock (7.6) is expressed in terms of the measurement unit [(items demanded)^2]. When *D* is normally distributed, the difference between the safety factors (7.5) and (7.7) is particularly striking – whereas (7.5) involves $\Phi^{-1}(\cdot)$, the inverse of the c.d.f for a normal random variable with mean zero and variance one, (7.7) involves $F_D^{-1}(\cdot)$, the inverse of the c.d.f of the random variable *D*, which has mean μ_D and variance σ_D^2. However, it is worth noting that since we are dealing with slow moving items μ_D is near zero, although $\sigma_D^2 \neq 1$.

In this dissertation we adopt Equations (7.6) and (7.7) as the respective definitions of the safety stock and the safety factor because these definitions are used by the retailer that supplied the data for items A and B. There is also a lead time variability factor in the retailer's safety stock calculation, however that data was not available to us, thus it could not be used.

Since the actual demand distribution $F_D(\cdot)$ is not known, we use the standard approach based on the assumption that *D* has a known distribution with mean μ_D and variance σ_D^2; and we estimate σ_D^2 as $(L+T)MSE_t$ where MSE_t is the estimated mean square error of the demand forecast for period *t* using a given forecasting technique. The safety factor (k') in equation (7.7) is calculated using a specified cycle service level, *CSL*, in our analysis. Along the lines of Syntetos *et al*. (2010), we use an exponentially smoothed value for the mean squared error, i.e.

$$MSE_t = \gamma(Y_t - Y'_{t-1})^2 + (1-\gamma)MSE_{t-1}, \quad (7.8)$$

where Y_t is actual sales in period *t*, and Y'_t is the estimate of demand computed at the end of period *t* for the demand in period *t+1* using a given forecasting technique, and γ is a smoothing coefficient. The initial value MSE_0 is set to 0.

Assuming demand over *L+T* can be reasonably assumed to follow a Poisson distribution with mean $\lambda(L+T)$, the value of the safety factor, k' is determined from Equation (7.7) such that the following is true.

$$\sum_{j=0}^{\lceil k' \rceil - 1} \frac{[\lambda(L+T)]^j}{j!} e^{-\lambda(L+T)} < CSL$$

and

$$\sum_{j=0}^{\lceil k' \rceil} \frac{[\lambda(L+T)]^j}{j!} e^{-\lambda(L+T)} \geq CSL, \quad (7.9)$$

where the ceiling function $\lceil x \rceil$ denotes the smallest integer greater than or equal to x for all real x. The demand rate λ at period t, λ_t', is estimated using the following equation,

$$\lambda_t' = \omega(Y_{t-1}) + (1-\omega)\lambda_{t-1}', \tag{7.10}$$

where ω is a smoothing coefficient. The initial value λ_0' is calculated dividing the total sales during the warm up period by the number of weeks in the warm up period.

Then, as in Syntetos (2001), the safety stock value computed in period t for the next $L+T$ periods, ss_t, is calculated as follows:

$$ss_t = \left\lceil k' \sqrt{(L+T)MSE_t} \right\rceil. \tag{7.11}$$

Let $Y_{t,P}'$ denote the forecast computed in period t-1 for the total expected demand over the time to the next P periods. Based upon the current company policy, the value of the order point at period t, s_t, is equal to the total forecast demand over the next review (T) plus lead time between vendor and store (L) periods, $Y_{t,T+L}'$, plus the safety stock estimate at period t, ss_t. In addition, the order-up-to level at period t, S_t, is equal to the total forecast demand over the lead time (L) plus additional weeks (W), $Y_{t,L+W}'$, plus the safety stock estimate at period t, ss_t. The quantity W is the additional weeks of supply that is being ordered every time an order is released and it is determined by corporate management for each item.

The values of the stocking policy in period t are computed as follows:

$$\begin{aligned} s_t &= \left\langle Y_{t,T+L}' + ss_t \right\rangle \\ S_t &= \max(\left\langle Y_{t,L+W}' + ss_t \right\rangle, s_t + 1), \end{aligned} \tag{7.12}$$

where

$$\langle x \rangle = \begin{Bmatrix} \lceil x \rceil, if \ x - \lfloor x \rfloor \geq 0.5 \\ \lfloor x \rfloor, otherwise \end{Bmatrix},$$

and the floor function $\lfloor x \rfloor = \lceil x \rceil - 1$ denotes the largest integer not exceeding *x* for all real values of *x*.

7.1.2 Simulation Process

In this section we detail the simulation process that is used to evaluate inventory-based performance measures for each of the candidate forecasting techniques. For each item *i* and store *k*, we divide the historical weekly sales data into two sets $H_{i,k}$ and $E_{i,k}$. The set $H_{i,k}$ represents the range of weeks used to initialize a given forecasting technique with historical data, while $E_{i,k}$ includes the range of weeks that are used to evaluate the performance of a given technique. For an item-store, the first week in the set $H_{i,k}$ is the week immediately after the first receipt of the item in that store. We use the first fifty-eight weeks of data after the first stocking date of the item in that store to initialize the forecasts, i.e. $|H_{i,k}| = 58$. Since the first receipt date of an item in a store can change, the number of weeks in set $E_{i,k}$ may be different for different item-store pairs.

The following is a list of parameters used in this section.

x_i is the total number of item-stores used in the simulation for item *i*;

k_i is a counter for stores for item *i*;

κ_i is the number of demand series (item-stores) used for item *i*'s inventory control metric calculations;

$Y'_{0,i,k}$ is the initial forecast for item *i* at the k^{th} store;

$\lambda'_{t,i,k}$ is the demand rate λ at period *t* for item *i* at the k^{th} store;

$t_{r,i,k}$ is the first week that the item *i* is received at the k^{th} store;

t_l is the last week of data used and it is the same for all item-stores;

$t_{i,k}$ is the number of weeks used to initialize a given forecasting technique for item *i* at the k^{th} store;

$H_{i,k}$ is the warm-up period for item *i* at the k^{th} store and contains the weeks between $t_{r,i,k}+1$ and $t_{r,i,k}+t_{i,k}$;

$E_{i,k}$ is the period that is used to calculate results for item *i* at the k^{th} store and contains the weeks between $t_{r,i,k}+t_{i,k}+1$ and t_l;

$IA'_{0,i,k}$ is the time between the first non-zero demand and $t_{r,i,k}$ for item *i* at the k^{th} store;

$IA'_{j,i,k}$ is the time between the non-zero demands *j* and *j*-1 for item *i* at the k^{th} store;

$n_{i,k}$ is the number of weeks in $E_{i,k}$ with non-zero demand for item *i* at the k^{th} store;

$TAS_{i,k}$ is the total actual sales for item *i* at the k^{th} store over the weeks in set $E_{i,k}$,

$TLS_{i,k}$ is the total lost sales for item *i* at the k^{th} store over the weeks in set $E_{i,k}$;

$w_{i,k}$ is the counter for item *i* at the k^{th} store used for weeks in set $E_{i,k}$;

$c_{i,k}$ is the cost for item *i* at the k^{th} store ;

$sp_{i,k}$ is the selling price for item i at the k^{th} store; and

$h_{i,k}$ is the inventory holding cost for item i at the k^{th} store.

The following is a list of output statistics that collected and computed:

$TM_{i,k}$ is the total margin for item i at the k^{th} store over the weeks in set $E_{i,k}$ where margin is the revenue less the cost of goods sold;

$TIHC_{i,k}$ is the total inventory holding cost for item i at the k^{th} store over the weeks in set $E_{i,k}$;

$FR_{i,k}$ is the fill rate for item i at the k^{th} store over the weeks in set $E_{i,k}$;

$YM_{i,k}$ is the yearly margin for item i at the k^{th} store over the weeks in set $E_{i,k}$, i.e. it is the total margin scaled to an annual amount;

$YIHC_{i,k}$ is the yearly inventory holding cost for item i at the k^{th} store over the weeks in set $E_{i,k}$ i.e. it is the total inventory cost scaled to an annual amount;

$YAM_{i,k}$ is the yearly adjusted margin for item i at the k^{th} store over the weeks in set $E_{i,k}$ i.e. it is the yearly margin less than yearly inventory holding cost ($YM_{i,k} - YIHC_{i,k}$);

$GMROI_{i,k}$ is the gross margin return on investment for item i at the k^{th} store over the weeks in set $E_{i,k}$, i.e. it is the yearly margin divided by the yearly inventory holding cost ($\frac{YM_{i,k}}{YIHC_{i,k}}$);

AFR$_i$ is the average of the fill rates for item *i* over all stores that are used to calculate inventory control metrics;

AYM$_i$ is the average of the yearly margins for item *i* over all stores that are used to calculate inventory control metrics;

AYIHC$_i$ is the average of the yearly inventory holding costs for item *i* over all stores that are used to calculate inventory control metrics;

AYAM$_i$ is the average of the yearly adjusted margins for item *i* over all stores that are used to calculate inventory control metrics; and

AGMROI$_i$ is the average of the GMROI values for item *i* over all stores that are used to calculate inventory control metrics.

The order of events in a week for a given item-store starts after the item is received in the store. The order of events for the simulation of a given item *i* is as follows:

1) Initialize item-store counter for item *i* to 1, $k_i \leftarrow 1$ and initialize item-store counter for inventory metrics calculations for item *i* to 0, $\kappa_i \leftarrow 0$.
2) Initialize forecast, inventory levels, and inventory metrics for item *i* at the k^{th} store.
   a) $i_{0,i,k} \leftarrow 1$
   b) $o_{0,i,k} \leftarrow 1$
   c) Initialize the sales forecast.

      If the forecasting method is SES or Poisson SES, then:

$$Y'_{0,i,k} \leftarrow \max\left(\frac{\sum_{j=t_{r,i,k}+1}^{t_{r,i,k}+t_{i,k}} Y_{j,i,k}}{t_{i,k}}, 0.001\right),$$

otherwise:

$$Y'_{0,i,k} \leftarrow \left\{\begin{array}{ll} 0.001, & if\ n_{i,k}=0, \\ \dfrac{\dfrac{\sum_{j=t_{r,i,k}+1}^{t_{r,i,k}+t_{i,k}} Y_{j,i,k}}{n_{i,k}}}{\dfrac{\sum_{j=1}^{j=n_{i,k}} IA_{j,i,k}}{n_{i,k}}}, & \text{otherwise.} \end{array}\right\}$$

d) $MSE_{0,i,k} \leftarrow 0$

e) $\lambda'_{0,i,k} \leftarrow \dfrac{\sum_{j=t_{r,i,k}}^{t_{r,i,k}+t_{i,k}} Y_{i,k,j}}{t_{i,k}}$

f) $t \leftarrow 1$

g) $w_{i,k} \leftarrow 0$

h) $S_{t,i,k} \leftarrow 1$ and $s_{t,i,k} \leftarrow 0$

i) $TAS_{0,i,k} \leftarrow 0, TLS_{0,i,k} \leftarrow 0, TM_{0,i,k} \leftarrow 0, TIHC_{0,i,k} \leftarrow 0, FR_{0,i,k} \leftarrow 0, YIHC_{0,i,k} \leftarrow 0, YMIHC_{0,i,k} \leftarrow 0,$ and $GMROI_{0,i,k} \leftarrow 0$

3) Receive the order of size $q_{t-L,i,k}$ made in period *t*-*L*, if any; update the on-hand and on-order inventory variables as follows:

$$i_{t,i,k} \leftarrow i_{t-1,i,k} + q_{t-L,i,k},$$
$$o_{t,i,k} \leftarrow o_{t-1,i,k} - q_{t-L,i,k}.$$

4) Demand for period *t*, $Y_{t,i,k}$ occurs. Demand can be zero, and if the initial on-hand is less than the demand, the difference is considered lost sales.

   The inventory on-hand is determined as follows:

$$i_{t,i,k} \leftarrow \max(i_{t,i,k} - Y_{t,i,k}, 0).$$

   Lost sales are determined as follows:

$$ls_{t,i,k} \leftarrow \max(Y_{t,i,k} - i_{t,i,k}, 0).$$

   Actual sales are determined as follows:

$$as_{t,i,k} \leftarrow Y_{t,i,k} - ls_{t,i,k}.$$

5) Update the following.

   a) The forecast $Y'_{t,i,k}$ based on the forecasting technique being evaluated.

   b) Compute $MSE_{t,i,k}$ according to Equation (7.8).

   c) Compute $k'_{t,i,k}$ according to Equation (7.9).

   d) Compute $\lambda'_{t,i,k}$ according to Equation (7.10).

   e) Compute $ss_{t,i,k}$ according to Equation (7.11).

   f) Compute $s_{t,i,k}$ and $S_{t,i,k}$ according to Equation (7.12).

6) Based on the (*s*, *S*) inventory model, place an order for this period, $q_{t,i,k}$, and update the inventory on-order.

$$q_{t,i,k} \leftarrow \begin{Bmatrix} S_{t,i,k} - i_{t,i,k} - o_{t,i,k}, & \text{if } i_{t,i,k} + o_{t,i,k} \le s_{t,i,k} \\ 0 & , \text{ otherwise.} \end{Bmatrix},$$

$$o_{t,i,k} \leftarrow o_{t,i,k} + q_{t,i,k}.$$

7) $t \leftarrow t + 1$.

8) Update the following if $t \ge t_{r,i,k} + t_{i,k} + 1$:

a) $TAS_{i,k} \leftarrow TAS_{i,k} + as_{t,i,k}$,

b) $TLS_{i,k} \leftarrow TLS_{i,k} + ls_{t,i,k}$,

c) $TM_{i,k} \leftarrow TM_{i,k} + (sp_{i,k} - c_{i,k}) * as_{t,i,k}$,

d) $TIHC_{i,k} \leftarrow TIHC_{i,k} + c_{i,k} \dfrac{(i_{t,i,k} + i_{t-1,i,k})}{2} \left(\dfrac{h_{i,k}}{52} \right)$, and

e) $w_{i,k} = w_{i,k} + 1$.

9) Repeat steps Step 2) through 8) until $t > t_l$.

10) If $w_{i,k} \ge 1$, calculate the following for item i at the k^{th} store:

a) $FR_{i,k} \leftarrow \dfrac{TAS_{i,k}}{TAS_{i,k} + TLS_{i,k}}$,

b) $YM_{i,k} \leftarrow 52 \cdot \dfrac{TM_{i,k}}{w_{i,k}}$,

c) $YIHC_{i,k} \leftarrow 52 \cdot \dfrac{TIHC_{i,k}}{w_{i,k}}$,

d) $YAM_{i,k} \leftarrow YM_{i,k} - YIHC_{i,k}$,

e) $GMROI_{i,k} \leftarrow \frac{YM_{i,k}}{YIHC_{i,k}}$, and

f) If $\frac{\sum_{t=t_{r,i,k}+t_{i,k}}^{t_l} Y_{t,i,k}}{t_l - t_{r,i,k} - t_{i,k} + 1} \leq 0.3$, then $\kappa_i \leftarrow \kappa_i + 1$.

11) $k_i \leftarrow k_i + 1$.

12) If $k_i > x_i$, go to (13),

Else, go to (1).

13) Calculate the following only for all item-stores that satisfy the following two conditions:

$$\frac{\sum_{t=t_{r,i,k}+t_{i,k}}^{t_l} Y_{t,i,k}}{t_l - t_{r,i,k} - t_{i,k} + 1} \leq 0.3 \text{ and } w_{i,k} \geq 1:$$

a) $AFR_i \leftarrow \frac{\sum_{i=1}^{\kappa_i} FR_i}{\kappa_i}$,

b) $AYM_i \leftarrow \frac{\sum_{i=1}^{\kappa_i} YM_i}{\kappa_i}$,

c) $AYIHC_i \leftarrow \frac{\sum_{i=1}^{\kappa_i} YIHC_i}{\kappa_i}$,

d) $AYAM_i \leftarrow \frac{\sum_{i=1}^{\kappa_i} YAM_i}{\kappa_i}$, and

e) $AGMROI_i \leftarrow \frac{\sum_{i=1}^{\kappa_i} GMROI_i}{\kappa_i}$,

where κ_i is the total number of stores that satisfy the two conditions above for item *i*.

The calculations are done separately for items A and B, and the results of the calculations in step 13) are used to compare the forecasting methods separately for items A and B. Table 7.1 gives the values of the parameters required in formulas used in the simulation experiments. Furthermore, the initial inventory on-hand, i_0, is set to one unit for each item-store.

Table 7.1 Parameter values for simulation and justification

Parameter	Value	Justification
T	1	Company policy
L	2	Company policy
W	1	Company policy
Target *CSL*	85%	Company policy
$h_{i,k}$	0.19	Company policy
γ	0.25	Syntetos *et al*. (2010)
ω	0.05	Croston (1972)
t_k	58	Preliminary calculations

A target *CSL* of 90% was also considered but those results are not shown since a target *CSL* of 85% gave better adjusted margin and *GMROI* results with average fill rates greater than 90%.

For each of the eight methods, the confidence intervals for the average fill rate, inventory holding cost, margin, adjusted margin, and *GMROI* are computed using the simulation data and are displayed individually in the following sections. Before defining the confidence interval calculation few definitions are required. First, let $\overline{FR_i}$ be the average fill rate obtained by a specific forecasting method and α value across all series of item *i*. It is calculated as:

$$\overline{FR_i} = \frac{1}{\kappa_i}\sum_{k=1}^{\kappa_i} FR_{i,k}$$ where κ_i is the number of demand series (stores) carrying item *i*.

Second, let s_{FR_i} be the standard deviation of the *FR* of item *i*. It is calculated as:

$$s_{FR_i} = \sqrt{\frac{1}{\kappa_i - 1}\sum_{k=1}^{\kappa_i}(FR_{i,k} - \overline{FR_i})^2}\,.$$

Then, the 99% confidence interval of fill rate of item *i* is given by:

$$\overline{FR_i} - 2.58\frac{s_{FR_i}}{\sqrt{\kappa_i}} \le FR_i \le \overline{FR_i} + 2.58\frac{s_{FR_i}}{\sqrt{\kappa_i}}\,. \qquad (7.13)$$

In addition, the results of paired sample *t*-tests are also displayed for fill rate, inventory holding cost, margin, adjusted margin, and *GMROI* in Sections 7.3-7.7. As an example, below are the null hypothesis and the alternative hypothesis for the fill rate measure:

H_0: The average fill rate given by method 1 for item *i* is equal to the average fill rate given by method 2 for item *i*.

H_1: The average fill rate given by method 1 for item *i* is not equal to the average fill rate given by method 2 for item *i*.

The Student's *t*-test is used to test the hypothesis since the number of demand series is large. In order to define the test statistic *t*, we first define the following:

$FR_{i,k,m}$: The simulated mean fill rate of item *i* at store *k* (k=1,…,κ_i) using forecasting method *m*,

$P_{i,k}$: The paired-sample mean fill rate difference of item *i* for store *k* (k=1,…,κ_i) between methods 1 and 2, i.e. $P_{i,k} = FR_{i,k,1} - FR_{i,k,2}$, and

$\overline{P_i}$: The average paired mean fill rate difference, i.e. $\overline{P_i} = \frac{1}{\kappa_i}\sum_{k=1}^{\kappa_i} P_{i,k}$, k=1,…,κ_i .

The test statistic is then calculated as follows:

$$t = \frac{\overline{P_i}}{\frac{s_{P_i}}{\sqrt{\kappa_i}}}, \tag{7.14}$$

where:

$$s_{P_i}^2 = \frac{1}{\kappa_i - 1}\sum_{k=1}^{\kappa_i}(P_{i,k} - \overline{P_i})^2 .$$

7.2 Selection of the Correction Factor and Forecast Update Frequency

Before evaluating the inventory performance measures of the methods in general, in this section the correction factor parameters used in two of the methods evaluated in this work (Croston SI New CF and SB2 SI New CF) are evaluated in order to determine the appropriate values for these parameters in the experimentation that follows in this chapter.

As stated in Section 3.2.3, the value of the correction factor (*CF*) for item *i* and at the k^{th} store at time period *t* involves the use of four input parameters: a maximum correction factor (*MCF*), an upper limit (*UL*), a lower limit (*LL*), and an update frequency. The first three of these inputs (*MCF*, *UL*, and *LL*) are used in the calculation of the correction factor as shown in the formula below:

$$CF_{t,i,k} = \begin{cases} (UL-0.1)^{-\min(m,MCF)}, & \text{if } m_{t,i,k} \geq m \\ (LL+0.1)^{\max(m,-MCF)}, & \text{if } m_{t,i,k} \leq -m \\ 1, \text{otherwise.} \end{cases}$$

The update frequency determines how often the forecast is updated.

In the experimentation of Chapter 6, the value of *CF* used with the Croston SI New CF and SB2 SI New CF methods was calculated using *UL* = 1.2, *LL* = 0.8, and *MCF* = 2. In addition, the forecast is updated not only after a demand, but also if the time after the last demand exceeds the average inter-arrival time.

In this sub-section the forecasts using the Croston SI New CF and SB2 SI New CF methods are simulated using *UL* values of 1.2 and 1.3, *LL* values of 0.7 and 0.8, *MCF* values of 2 and 3, and update frequency values of 1, 2, and 3; therefore a total of 24 parameter set combinations are compared using the inventory metrics described in the introduction of this chapter. The parameter set combinations are explicitly listed in Table B.1.1 in Appendix B.1.

The performance of a forecasting method (Croston SI New CF and SB2 SI New CF) using a given parameter set combination and a specified target cycle service level (target *CSL*=85%) is computed for each item (A and B) using the process described in Section 7.1.2. The five inventory performance metrics of this Chapter (fill rate, margin, inventory holding cost, adjusted margin, and *GMROI*) are used to evaluate the parameter set combination. Specifically, for each of the two items (A and B) and for each combination of the forecasting method and parameter set combination, simulations are run over all series (stores) for which demand data is available. Then, confidence intervals on the mean performance are computed as illustrated for the fill rate metric below. First, the mean fill rate of item *i*, using forecasting method *j*, parameter set combination *m*, and forecasting smoothing coefficient (α_n) value where $\alpha_n = 0.05n$ across all data series (stores) is calculated as:

$$\overline{FR}_{i,j,m,n} = \frac{1}{\kappa_i}\sum_{k=1}^{\kappa_i} FR_{i,j,m,n,k} ,$$

where

$FR_{i,j,m,n,k}$ is the observed fill rate from the simulation of item *i*, using forecasting method *j*, parameter set combination *m*, and forecasting smoothing coefficient (α_n) value for series (store) *k*. The α_n equals 0.05*n for n=1, 2, 3, and 4.

Then the lower limit of the confidence interval of fill rate of item *i*, using forecasting method *j*, parameter set combination *m*, and forecasting smoothing coefficient (α_n) value is given by:

$$\overline{FR}_{i,j,m,n} - 2.58\frac{s_{FR_{i,j,m,n}}}{\sqrt{\kappa_i}} ,$$

and the corresponding upper limit of the confidence interval is given by:

$$\overline{FR}_{i,j,m,n} + 2.58\frac{s_{FR_{i,j,m,n}}}{\sqrt{\kappa_i}} .$$

where

$s_{FR_{i,j,m,n}}$ is the standard deviation of the observed simulated values over all series, and is calculated as:

$$s_{FR_{i,j,m,n}} = \sqrt{\frac{1}{\kappa_i - 1}\sum_{k=1}^{\kappa_i}(FR_{i,j,m,n,k} - \overline{FR}_{i,j,m,n})^2} \ .$$

Note, κ_i is the number of stores for item *i*.

The resulting confidence intervals for the inventory metrics (fill rate, margin, inventory holding cost, adjusted margin, and *GMROI*) of these 24 parameter set combinations for Croston SI New CF and SB2 SI New CF methods for items A and B and a target *CSL* value of 85% are displayed in Appendices B.2-B.6. It can be observed that these confidence intervals are overlapping, suggesting no significance difference in performance. A similar observation can be made about the update frequency. Therefore, for the experimentation in the sections that follow, we use parameter set 3 since it is robust and often leads to the best average performance across all the adjusted margin and *GMROI* metrics. This parameter set combination has *MCF* of 2, *UL* of 1.2, *LL* of 0.8, and update frequency of 1 week.

7.3 Fill Rate Results

In this section, we evaluate the performance of the forecasting methods in terms of the mean fill rates computed in the simulations for a target *CSL* value of 85% and smoothing coefficient, α values of 0.05, 0.1, 0.15, and 0.2. The results are presented separately for items A and B. In addition, for each method, we identify the α value that generates the highest fill rate and then make pairwise comparisons using a two sided *t*-test separately for items A and B. Finally an overall ranking of the methods based on statistically significant differences are displayed in Table 7.3.3.

Confidence intervals of the mean fill rates calculated using the formulas in (7.13) are displayed in the following figures.

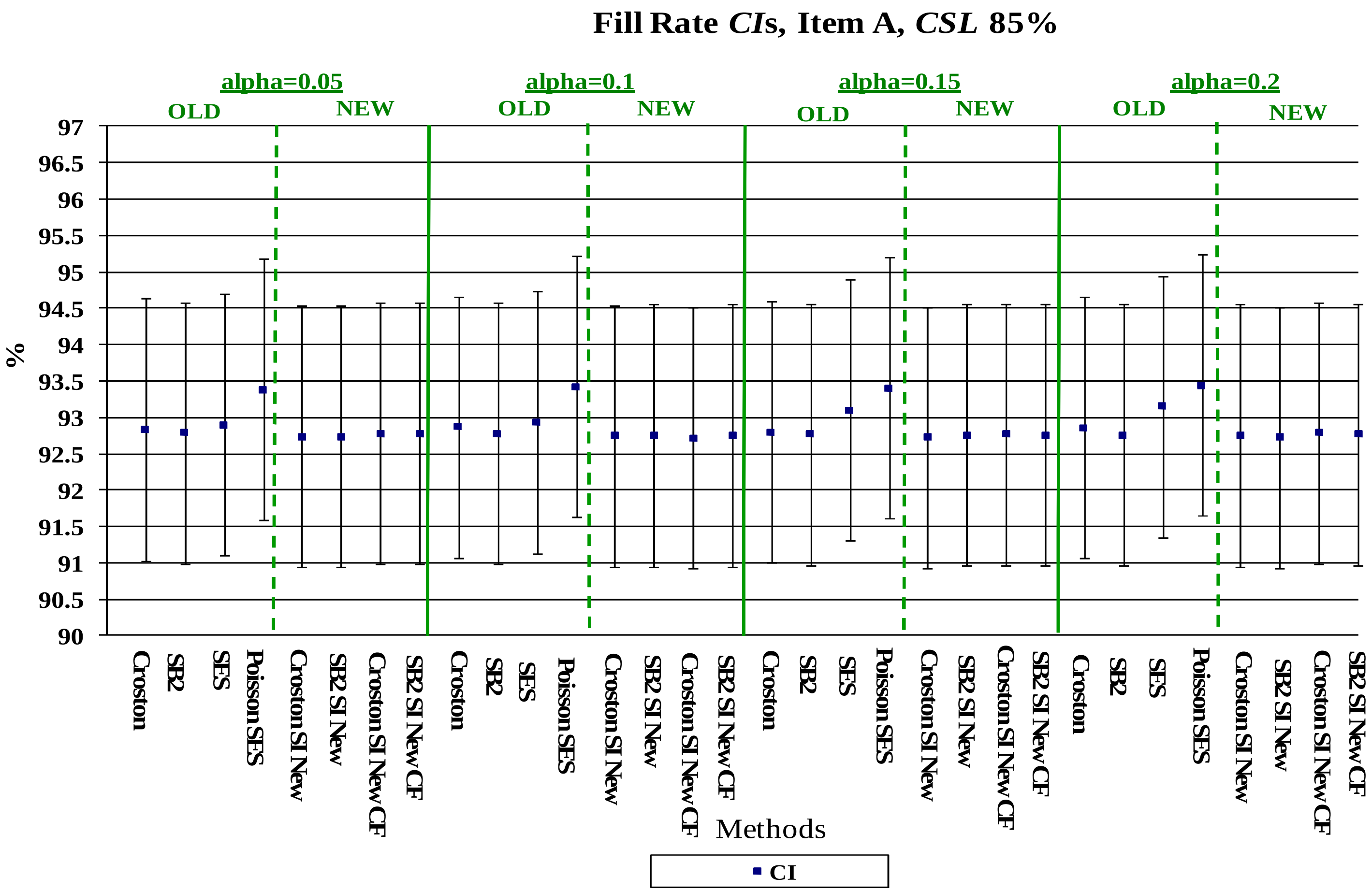

Figure 7.3.1 Item A fill rate confidence intervals, target *CSL*=85%

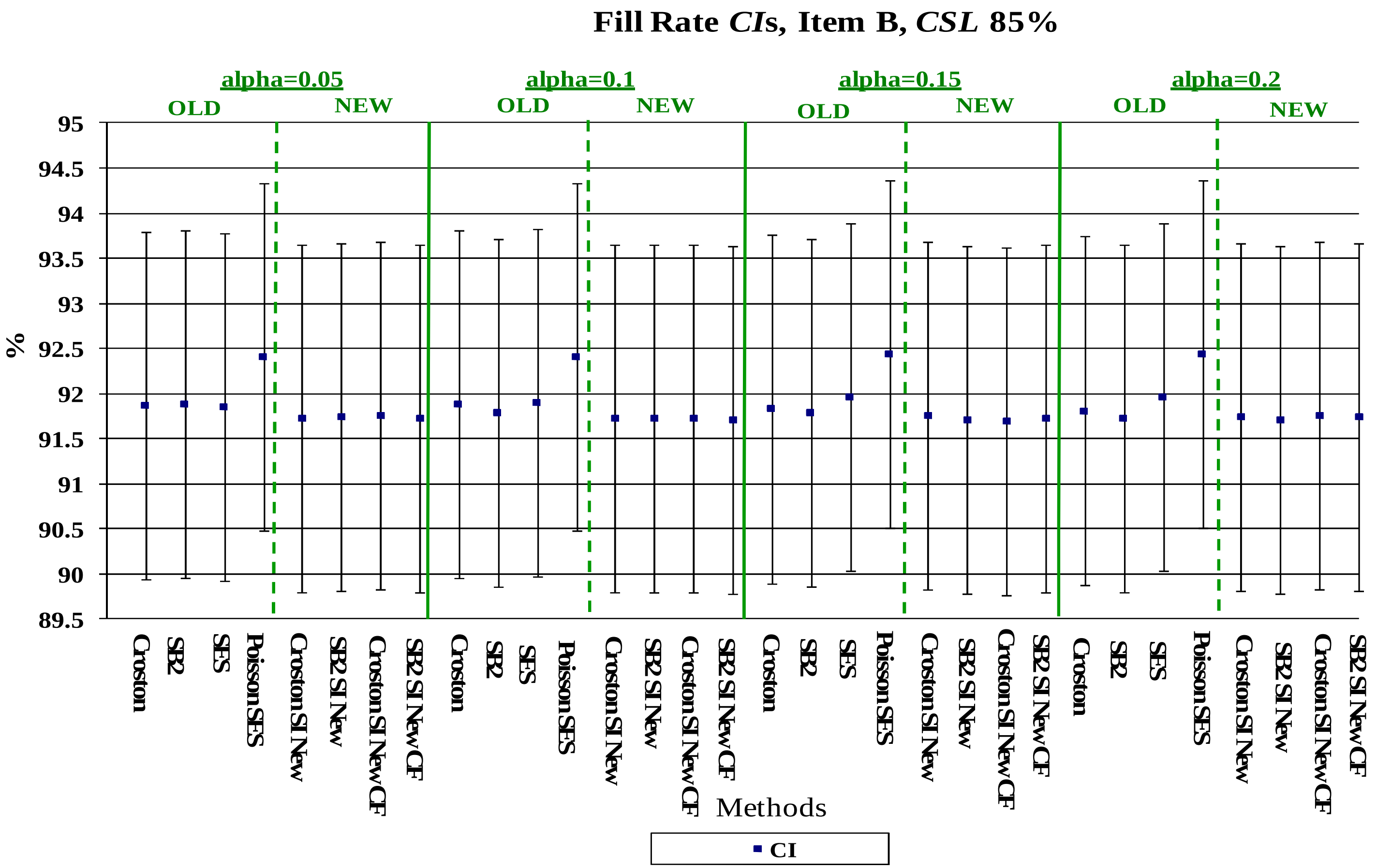

Figure 7.3.2 Item B fill rate confidence intervals, target *CSL*=85%

Since all the confidence intervals overlap, *t*-tests for paired fill rate comparisons are performed using (7.14) to determine the statistical differences between pairs of methods. For each method, the α value that generates the highest average fill rate is used for comparison. Table 7.3.1 displays the *t* values for each item sorted from smallest to largest. The values are bolded that are statistically significant (i.e. test statistic is greater than 2.58 in absolute value).

Table 7.3.1 Paired-sample *t*-test values of fill rate differences

Methods Compared Method 1- Method 2	**Item A**	**Methods Compared Method 1- Method 2**	**Item B**
SB2-Poisson SES	**-6.84**	SB2-Poisson SES	**-6.41**
Croston-Poisson SES	**-6.4**	Croston-Poisson SES	**-6.21**
SES-Poisson SES	**-4.45**	SES-Poisson SES	**-5.89**
SB2-SES	**-4.37**	SB2-SES	-1.5
Croston-SES	**-3.54**	Croston-SES	-1.43
Croston SI New-Croston SI New CF	-0.53	SB2 SI New-Croston SI New CF	-0.4
Croston SI New-SB2 SI New CF	-0.52	SB2 SI New-SB2 SI New CF	-0.06
SB2 SI New-Croston SI New CF	-0.48	Croston-SB2	-0.04
SB2 SI New-SB2 SI New CF	-0.39	Croston SI New-Croston SI New CF	0.02
Croston SI New-SB2 SI New	-0.1	Croston SI New-SB2 SI New CF	0.45
SB2-Croston SI New CF	0.03	Croston SI New CF-SB2 SI New CF	0.61
Croston SI New CF-SB2 SI New CF	0.18	Croston SI New-SB2 SI New	0.7
SB2-SB2 SI New CF	0.19	Croston-Croston SI New CF	2.55
SB2-SB2 SI New	0.52	SB2-Croston SI New	**2.77**
SB2-Croston SI New	0.66	Croston-Croston SI New	**2.82**
Croston-Croston SI New CF	1.28	SB2-Croston SI New CF	**2.94**
Croston-SB2	1.62	Croston-SB2 SI New CF	**3**
Croston-SB2 SI New CF	1.67	Croston-SB2 SI New	**3.18**
Croston-SB2 SI New	**2.64**	SB2-SB2 SI New	**3.21**
Croston-Croston SI New	**3**	SB2-SB2 SI New CF	**3.21**
SES-Croston SI New CF	**4.22**	SES-Croston SI New	**3.3**
SES-SB2 SI New CF	**4.6**	SES-Croston SI New CF	**3.35**
SES-SB2 SI New	**4.67**	SES-SB2 SI New	**3.48**
SES-Croston SI New	**4.96**	SES-SB2 SI New CF	**3.7**
Poisson SES-Croston SI New CF	**7.21**	Poisson SES-Croston SI New	**6.98**
Poisson SES-SB2 SI New CF	**7.28**	Poisson SES-Croston SI New CF	**7.11**
Poisson SES-Croston SI New	**7.33**	Poisson SES-SB2 SI New	**7.17**
Poisson SES-SB2 SI New	**7.37**	Poisson SES-SB2 SI New CF	**7.28**

For item A, the Croston and SB2 methods are worse than the SES and Poisson SES methods. In addition, the SES method is worse than the Poisson SES method. There is no statistical difference between the Croston, SB2, Croston SI New CF, and SB2 SI New CF methods. Furthermore, there is no statistical difference between any of the new methods. In addition, there is no statistical difference between the SB2 method, and the Croston SI New and SB2 SI New methods. The Croston method is better than the Croston SI New and SB2 SI New methods. The SES and Poisson SES methods are better than the new methods.

For item B, the Croston and SB2 methods are worse than the Poisson SES method. In addition, the SES method is worse than the Poisson SES method. There is no statistical difference between the Croston, SB2, and SES methods. Furthermore, there is no statistical difference between any of the new methods. The Croston and SB2 methods are better than the new methods, except there is no statistical difference between the Croston and Croston SI New CF methods. The SES and Poisson SES methods are better than new methods.

Table 7.3.2 displays the mean fill rates of the existing and new methods for items A and B based on a target *CSL* value of 85%. The best (highest) values for each method and item are highlighted.

Table 7.3.2 Mean fill rate

Mean Fill Rate								
Methods	Item A				Item B			
	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.15$	$\alpha = 0.2$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.15$	$\alpha = 0.2$
Existing Methods								
Croston	92.80%	92.90%	92.80%	92.90%	91.90%	91.90%	91.80%	91.80%
SB2	92.80%	92.80%	92.80%	92.80%	91.90%	91.80%	91.80%	91.70%
SES	92.90%	92.90%	93.10%	93.10%	91.90%	91.90%	92.00%	92.00%
Poisson SES	93.40%	93.40%	93.40%	93.40%	92.40%	92.40%	92.40%	92.40%
New Methods								
Croston SI New	92.70%	92.70%	92.70%	92.80%	91.70%	91.70%	91.80%	91.70%
SB2 SI New	92.70%	92.80%	92.80%	92.70%	91.70%	91.70%	91.70%	91.70%
Croston SI New CF	92.80%	92.70%	92.80%	92.80%	91.80%	91.70%	91.70%	91.80%
SB2 SI New CF	92.80%	92.80%	92.80%	92.80%	91.70%	91.70%	91.70%	91.70%

Table 7.3.3 displays the fill rate ranks based on the statistically significant differences; note that pairwise 1% tests were used to rank the methods, not a multiple-comparisons procedure, and hence that the chance of at least one incorrect ranking is higher than 1%.

Table 7.3.3 Summary of fill rate ranks based on statistically significant differences

Rank	**Item A**	**Item B**
1	Poisson SES	Poisson SES
2	SES	Croston, SB2, SES
3	Croston, SB2, SB2 SI New CF, Croston SI New CF	Croston, Croston SI New CF
4	SB2, SB2 SI New, Croston SI New	SB2 SI New, Croston SI New, SB2 SI New CF
5	SB2 SI New, Croston SI New, SB2 SI New CF, Croston SI New CF	SB2 SI New, Croston SI New, SB2 SI New CF, Croston SI New CF

As expected, the Poisson SES method generates forecasts that have a higher fill rate compared to other methods, since more inventory is held at the stores. Since the forecasts using the Poisson SES method are integer values (except 0.001 when the forecast is equal to 0), the *s* and *S* values are higher compared to other methods, and this increases the fill rate.

All 99% confidence intervals on the actual fill rates are above the nominal *CSL*=85%, by at least 5%, this is a clear indication that the safety stock is too large. In addition, since the mean fill rate differences between all methods are not greater than 1%, the differences are not practically significant.

7.4 Margin Results

In this section the forecasting methods are evaluated based on the margin performance metric. The resulting average margin for each technique is computed in the simulation for a target *CSL* value of 85% at α values of 0.05, 0.1, 0.15, and 0.2. The results are presented separately for items A and B. In addition, the results of pairwise margin comparisons using a two sided *t*-test are also presented separately for items A and B. Finally, a ranking of the methods based on margin is presented.

The *CI*s of the mean margin as calculated using the formulas in (7.13) are displayed in the following figures.

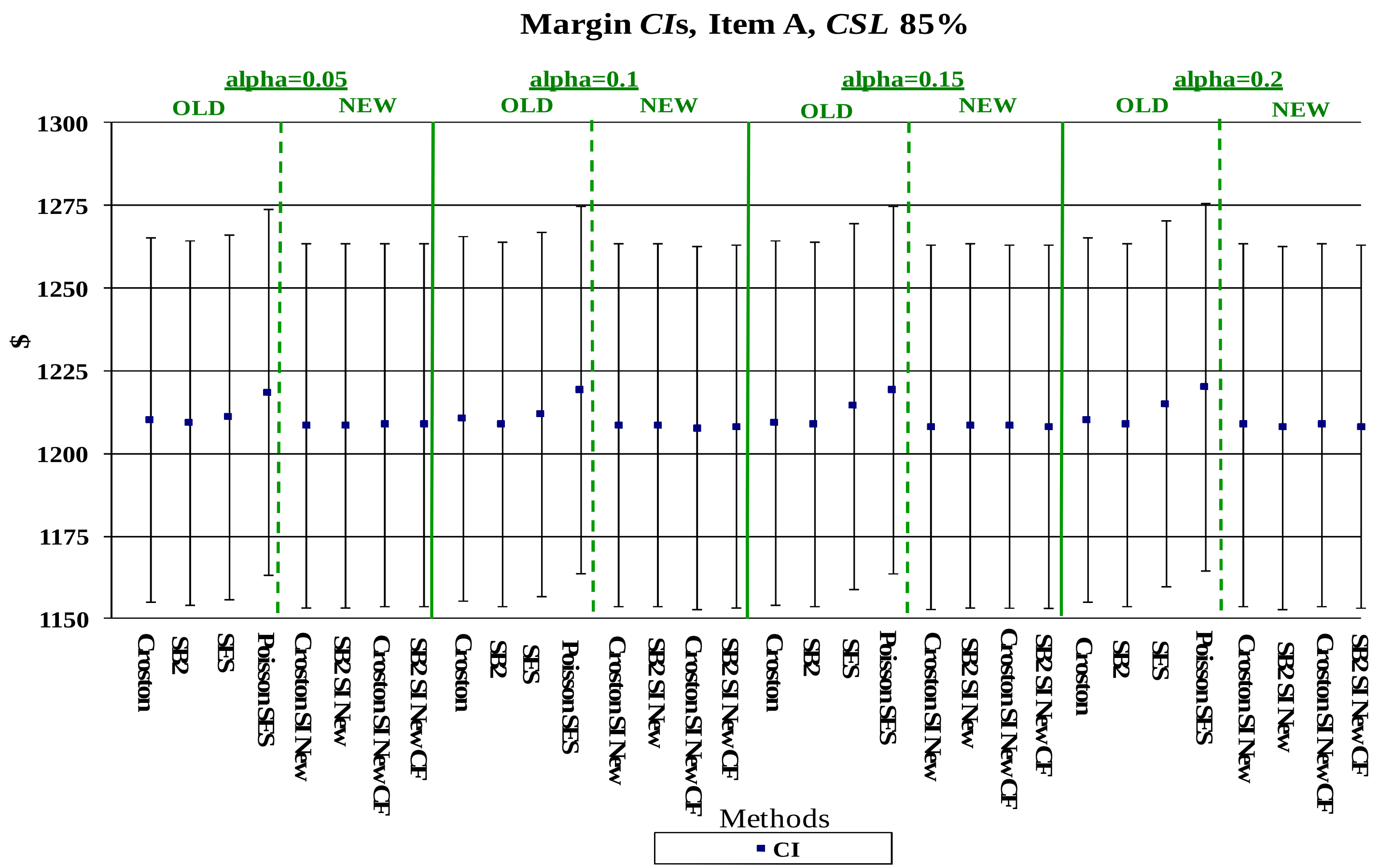

Figure 7.4.1 Item A margin confidence intervals, target *CSL*=85%

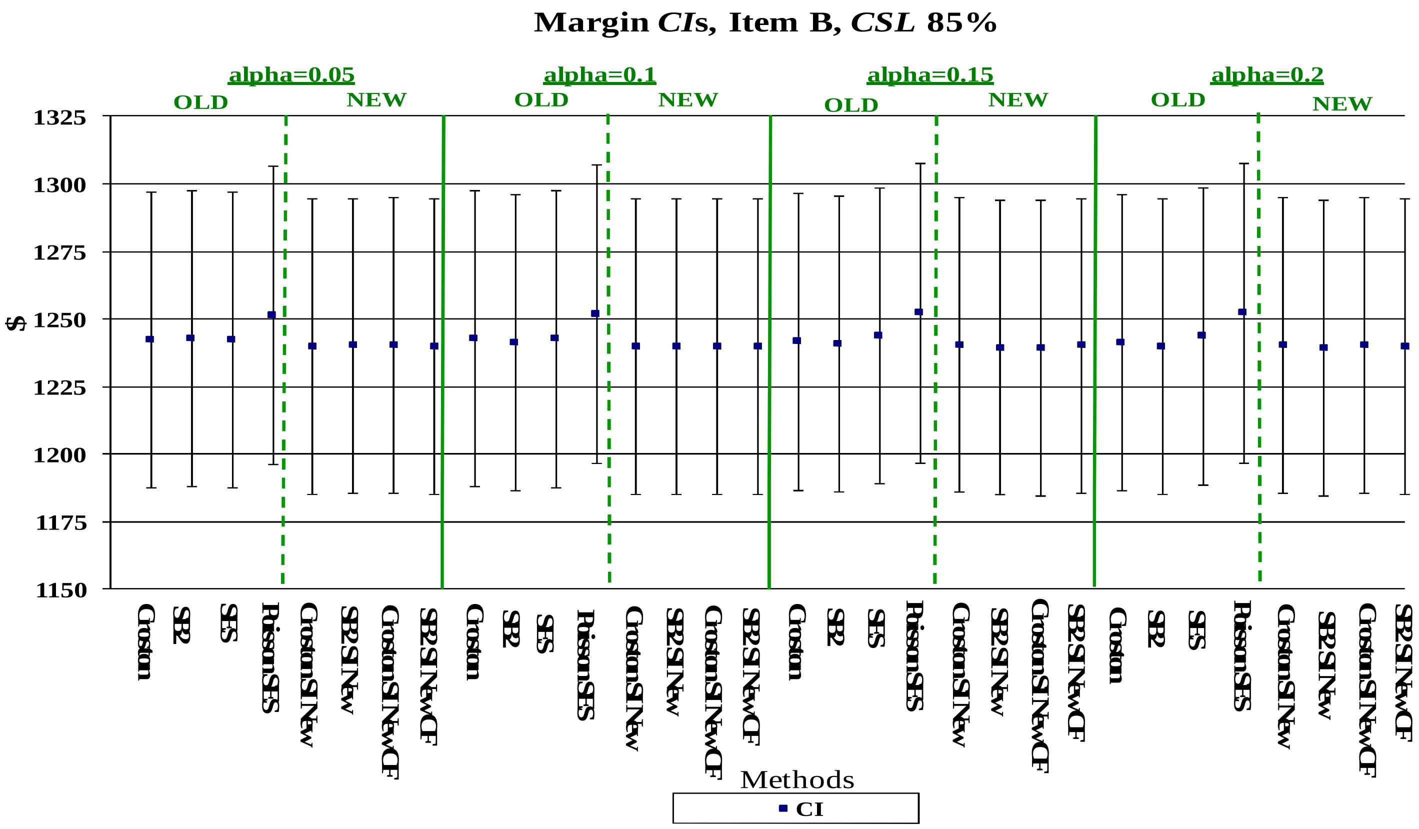

Figure 7.4.2 Item B margin confidence intervals, target *CSL*=85%

Since all the confidence intervals overlap, as before *t*-tests for paired margin comparisons are performed using (7.14) to determine the statistical differences between pairs of methods. For each method, the α value that generates the highest average fill rate is used for comparison. Table 7.4.1 displays the *t* values for each item sorted from smallest to largest. The values are bolded that are statistically significant.

Table 7.4.1 Paired-sample *t*-test values for margin differences

Methods Compared Method 1- Method 2	**Item A**	**Methods Compared Method 1- Method 2**	**Item B**
SB2-Poisson SES	**-6.96**	SB2-Poisson SES	**-6.58**
Croston-Poisson SES	**-6.49**	Croston-Poisson SES	**-6.24**
SES-Poisson SES	**-4.7**	SES-Poisson SES	**-5.93**
SB2-SES	**-4.45**	SB2-SES	-1.39
Croston-SES	**-3.77**	Croston-SES	-1.1
SB2 SI New-Croston SI New CF	-0.25	SB2 SI New-Croston SI New CF	-0.32
SB2 SI New-SB2 SI New CF	-0.25	SB2 SI New-SB2 SI New CF	0
Croston SI New-Croston SI New CF	-0.06	Croston SI New-Croston SI New CF	0.2
Croston SI New-SB2 SI New CF	-0.06	Croston SI New CF-SB2 SI New CF	0.37
Croston SI New CF-SB2 SI New CF	0	Croston-SB2	0.38
Croston SI New-SB2 SI New	0.19	Croston SI New-SB2 SI New CF	0.74
SB2-Croston SI New CF	0.57	Croston SI New-SB2 SI New	0.89
SB2-SB2 SI New CF	0.57	SB2-Croston SI New	**2.71**
SB2-Croston SI New	0.74	Croston-Croston SI New CF	**2.74**
SB2-SB2 SI New	0.82	Croston-Croston SI New	**2.85**
Croston-SB2	1.73	SB2-Croston SI New CF	**3.04**
Croston-Croston SI New CF	1.96	SB2-SB2 SI New CF	**3.04**
Croston-SB2 SI New CF	1.96	SES-Croston SI New	**3.04**
Croston-SB2 SI New	2.53	Croston-SB2 SI New CF	**3.08**
Croston-Croston SI New	**2.99**	SB2-SB2 SI New	**3.23**
SES-Croston SI New CF	**5.12**	Croston-SB2 SI New	**3.25**
SES-SB2 SI New CF	**5.12**	SES-Croston SI New CF	**3.34**
SES-Croston SI New	**5.17**	SES-SB2 SI New	**3.49**
SES-SB2 SI New	**5.18**	SES-SB2 SI New CF	**3.63**
Poisson SES-Croston SI New CF	**7.41**	Poisson SES-Croston SI New	**6.93**
Poisson SES-SB2 SI New CF	**7.41**	Poisson SES-Croston SI New CF	**7.11**
Poisson SES-Croston SI New	**7.53**	Poisson SES-SB2 SI New CF	**7.14**
Poisson SES-SB2 SI New	**7.6**	Poisson SES-SB2 SI New	**7.19**

For item A, the Croston and SB2 methods are worse than the SES and Poisson SES methods. In addition, the SES method is worse than the Poisson SES method. Furthermore, there is no statistical difference between the Croston, SB2 and the new methods, except the Croston method is better than the Croston SI New method. The SES and Poisson SES methods are better than the new methods.

For item B, the Croston and SB2 methods are worse than the Poisson SES method. In addition, the SES method is worse than the Poisson SES method. There is no statistical difference between the Croston, SB2, and SES methods. Furthermore, there is no statistical difference between the new methods. The old methods are better than the new methods.

Tables 7.4.2 displays the mean margin of the existing and new methods for items A and B based on target *CSL* value of 85%. The best (highest) values for each method and item are highlighted and are used to compare the methods with practical significance if they are statistically different.

Table 7.4.2 Mean margin

Mean Margin								
	Item A				Item B			
Methods	$\alpha=0.05$	$\alpha=0.1$	$\alpha=0.15$	$\alpha=0.2$	$\alpha=0.05$	$\alpha=0.1$	$\alpha=0.15$	$\alpha=0.2$
Existing Methods								
Croston	1210.1	1210.8	1209.4	1210.5	1242.7	1243.1	1241.8	1241.6
SB2	1209.5	1209	1209	1208.8	1243	1241.3	1241.2	1240
SES	1211.2	1212	1214.4	1215.2	1242.6	1242.8	1244.1	1243.8
Poisson SES	1218.7	1219.5	1219.3	1220.1	1251.6	1252	1252.4	1252.4

Table 7.4.2 Continued

New Methods								
Croston SI New	1208.6	1208.7	1208.1	1208.8	1240	1240.2	1240.8	1240.7
SB2 SI New	1208.6	1208.7	1208.6	1208	1240.3	1240.1	1239.7	1239.6
Croston SI New CF	1208.9	1207.7	1208.4	1208.9	1240.4	1240.1	1239.5	1240.6
SB2 SI New CF	1208.9	1208.2	1208.3	1208.3	1240.1	1240	1240.3	1240.1

Table 7.4.3 displays the margin ranks based on the statistically significant differences; note that pairwise 1% tests were used to rank the methods, not a multiple-comparisons procedure, and hence that the chance of at least one incorrect ranking is higher than 1%. This is also true for inventory holding cost, adjusted margin, and *GMROI* rankings.

Table 7.4.3 Summary of margin ranks based on statistically significant differences

Rank	**Item A**	**Item B**
1	Poisson SES	Poisson SES
2	SES	Croston, SB2, SES
3	Croston, SB2, SB2 SI New CF, Croston SI New CF, SB2 SI New	SB2 SI New, Croston SI New, Croston SI New CF, SB2 SI New CF
4	SB2, Croston SI New	
5	SB2 SI New, Croston SI New, SB2 SI New CF, Croston SI New CF	

If the difference in performance between the two methods is statistically significant and the mean margin difference between the two methods is greater than or equal to 50 cents, then the difference in performance between the methods is also practically significant.

The definition of “practically significant” is subjective and can change based on the company’s policy and 50 cents might seem low to determine practical significance. However

50 cents difference at the item-store level corresponds to $1,000 at the item country level across 2,000 stores. In addition, if the company has 1,000 such slow-seller items, 50 cents per item per store corresponds to an average difference of $1,000,000 across all stores and items. The same definition of practical significance is used for inventory holding cost and adjusted margin.

Table 7.4.4 displays the ranking of the methods for the margin metric based on the practically significant differences.

Table 7.4.4 Summary of margin ranks based on practically significant differences

Rank	**Item A**	**Item B**
1	Poisson SES	Poisson SES
2	SES	Croston, SB2, SES
3	Croston, SB2, SB2 SI New CF, Croston SI New CF, SB2 SI New	SB2 SI New, Croston SI New, Croston SI New CF, SB2 SI New CF
4	SB2, Croston SI New	
5	SB2 SI New, Croston SI New, SB2 SI New CF, Croston SI New CF	

As expected, the Poisson SES method generates forecasts have a higher margin compared to other methods, since more inventory is held at the stores. Since the Poisson SES forecasts are integer values (except 0.001 when the forecast=0), the *s* and *S* values are higher compared to other methods and this increases the margin. In addition, for item A the SES method’s margin is practically significantly better than all other methods except for the Poisson SES method.

7.5 Inventory Holding Cost Results

In the following tables and figures of Section 7.5, the confidence intervals of the mean inventory holding cost for smoothing coefficient, α values of 0.05, 0.1, 0.15, and 0.2 are presented separately for items A and B based on a target *CSL* value of 85%. In addition, the results of pairwise inventory holding cost comparisons with two sided *t*-test of the α value that generates the highest inventory holding cost are also presented separately for items A and B. The ranks of the methods based on statistically significant differences are displayed in Table 7.5.3, and inventory holding cost ranks of the methods based on practically significant differences are displayed in Table 7.5.4.

The confidence intervals of the mean inventory holding cost calculated using Equation (7.13) are displayed in the following figures.

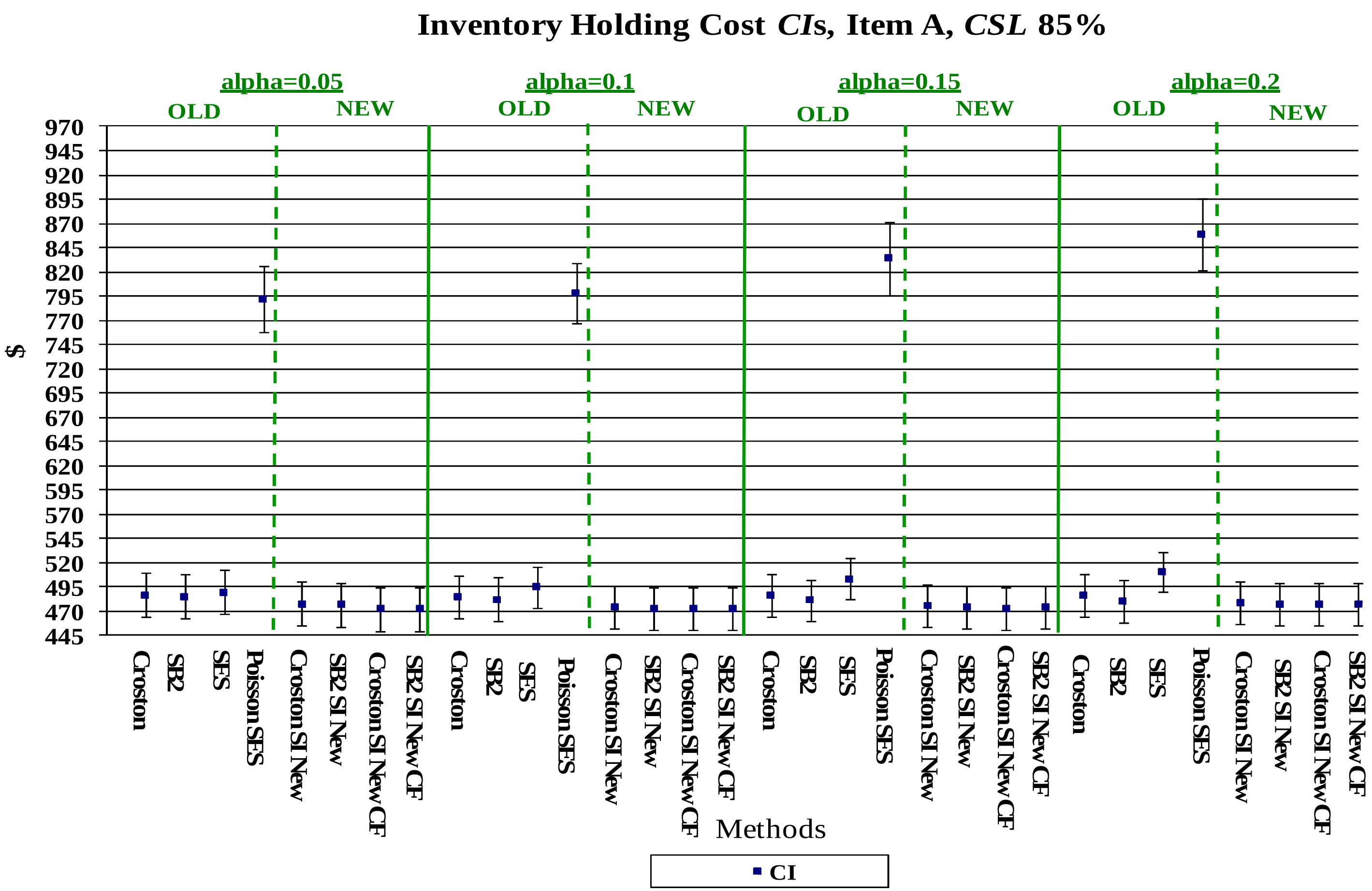

Figure 7.5.1 Item A inventory holding cost confidence intervals, target *CSL*=85%

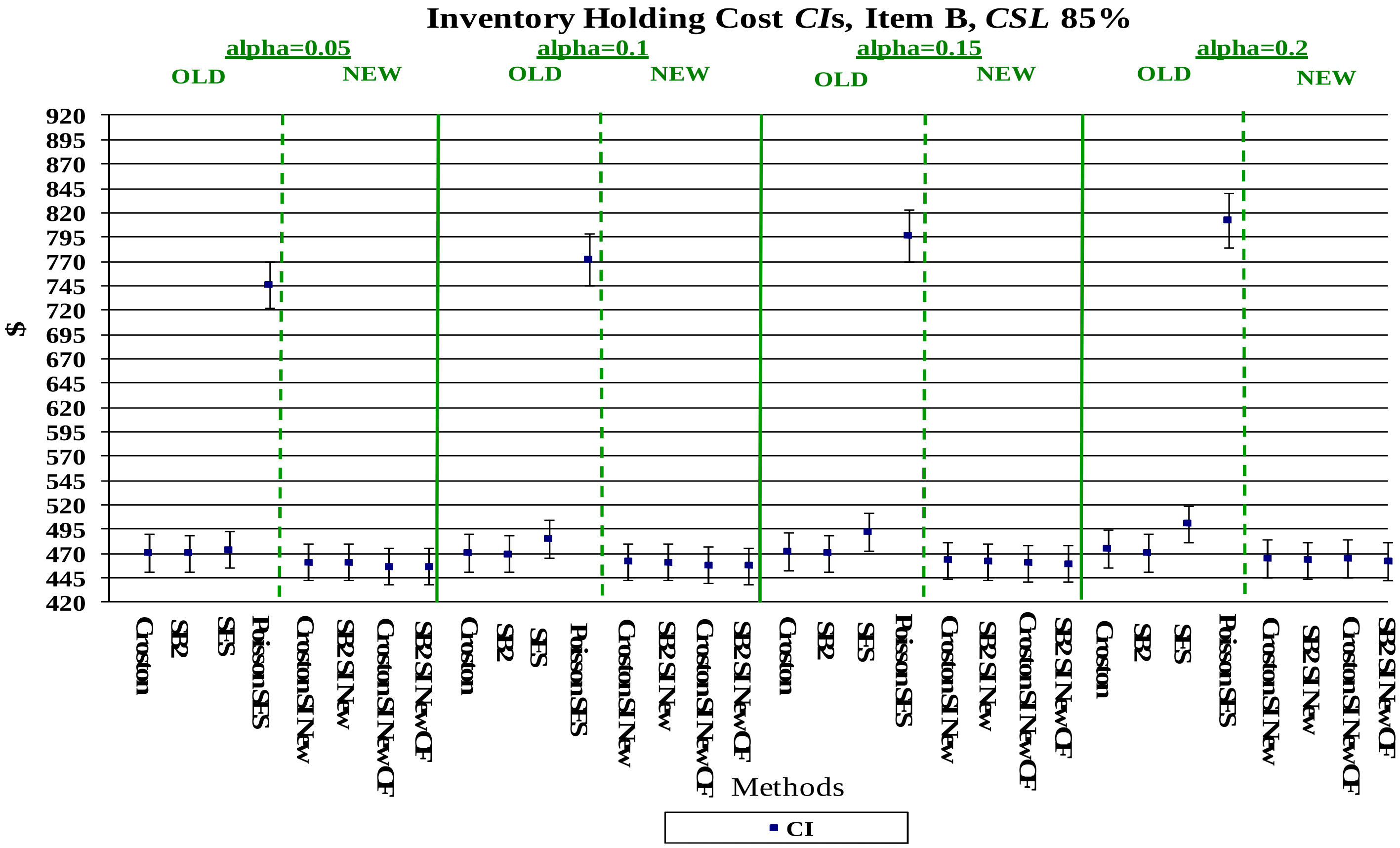

Figure 7.5.2 Item B inventory holding cost confidence intervals, 85% *CSL*

Since all the confidence intervals overlap except for those for the Poisson SES method, as before *t*-tests for paired inventory holding cost are performed using (7.14) to determine the statistical differences between pairs of methods. For each method, the α value that generates the highest average fill rate is used for comparison. Table 7.5.1 displays the *t* values for each item sorted from smallest to largest. The values are bolded that are statistically significant.

Table 7.5.1 Paired-sample *t*-test values for inventory holding cost differences

Methods Compared Method 1- Method 2	**Item A**	**Methods Compared Method 1- Method 2**	**Item B**
Croston-Poisson SES	**-32.42**	Croston-Poisson SES	**-36.57**
SB2-Poisson SES	**-32.07**	SB2-Poisson SES	**-36.34**
SES-Poisson SES	**-31.66**	SES-Poisson SES	**-35.57**
SB2-SES	**-5.83**	SB2-SES	**-3.86**
Croston-SES	**-4.14**	Croston-SES	**-2.74**
Croston SI New CF-SB2 SI New CF	-0.12	Croston SI New CF-SB2 SI New CF	-0.38
SB2 SI New-SB2 SI New CF	0.08	Croston SI New-SB2 SI New	-0.09
SB2 SI New-Croston SI New CF	0.16	Croston-SB2	1.06
Croston SI New-Croston SI New CF	1.46	Croston SI New-SB2 SI New CF	**4.21**
Croston SI New-SB2 SI New CF	1.68	Croston SI New-Croston SI New CF	**4.46**
Croston SI New-SB2 SI New	2.51	SB2 SI New-Croston SI New CF	**4.91**
Croston-SB2	**2.98**	SB2 SI New-SB2 SI New CF	**4.98**
SB2-Croston SI New CF	**4.25**	SB2-SB2 SI New	**7.24**
SB2-Croston SI New	**4.44**	Croston-SB2 SI New	**7.47**
SB2-SB2 SI New CF	**5.22**	SB2-Croston SI New	**7.64**
SB2-SB2 SI New	**6.04**	Croston-Croston SI New	**8.39**
Croston-Croston SI New	**7.26**	SES-SB2 SI New	**9.4**
Croston-SB2 SI New CF	**7.35**	SES-Croston SI New	**9.8**
Croston-Croston SI New CF	**7.55**	SB2-SB2 SI New CF	**10.08**
Croston-SB2 SI New	**8.48**	SB2-Croston SI New CF	**10.16**
SES-Croston SI New	**9.06**	Croston-SB2 SI New CF	**10.33**
SES-SB2 SI New CF	**9.23**	Croston-Croston SI New CF	**10.49**
SES-Croston SI New CF	**9.91**	SES-SB2 SI New CF	**12.2**
SES-SB2 SI New	**10.3**	SES-Croston SI New CF	**12.42**
Poisson SES-Croston SI New	**32.83**	Poisson SES-SB2 SI New	**37.28**
Poisson SES-SB2 SI New	**33.07**	Poisson SES-Croston SI New	**37.36**
Poisson SES-SB2 SI New CF	**33.12**	Poisson SES-Croston SI New CF	**37.94**
Poisson SES-Croston SI New CF	**33.22**	Poisson SES-SB2 SI New CF	**37.99**

For item A, the Croston and SB2 methods have lower inventory holding cost than the SES and Poisson SES methods. In addition, the SES method has lower inventory holding cost than the Poisson SES method. Furthermore, there is no statistical difference between the new methods. The Croston method has higher inventory holding cost than the SB2 method. The old methods have higher inventory holding cost than the new methods.

For item B, the Croston and SB2 methods have lower inventory holding cost than the SES and Poisson SES methods. In addition, the SES method has lower inventory holding cost than the Poisson SES method. Furthermore, there is no statistical difference between the Croston SI New CF and SB2 SI New CF methods. Moreover, there is no statistical difference between the Croston SI New and SB2 SI New methods. In addition, the Croston and SB2 methods are not statistically different. Both Croston SI New and SB2 SI New methods have higher inventory holding cost than Croston SI New CF and SB2 SI New CF methods. The old methods have higher inventory holding cost than the new methods.

Table 7.5.2 displays the mean inventory holding cost of the existing and new methods for items A and B based on a *CSL* of 85%. The lowest values for each method and item are highlighted below and are used to compare the methods with practical significance if they are statistically different.

Table 7.5.2 Mean inventory holding cost

Mean Inventory Holding Cost								
Methods	Item A				Item B			
	$\alpha=0.05$	$\alpha=0.1$	$\alpha=0.15$	$\alpha=0.2$	$\alpha=0.05$	$\alpha=0.1$	$\alpha=0.15$	$\alpha=0.2$
Existing Methods								
Croston	487.0	485.0	486.3	486.4	471.1	471.6	472.9	475.6
SB2	485.7	482.7	481.3	481	470.4	470.3	470.5	471.4
SES	490.2	495.0	503.5	510.9	474.5	485.4	492.7	501.3
Poisson SES	792.5	798.5	833.8	858.8	746.5	772.5	796.4	812.7
New Methods								
Croston SI New	477.6	474.5	476.3	479.5	461.4	462.2	463.3	465.3
SB2 SI New	477.0	472.6	474.8	477.6	461.5	461.5	461.8	463.4
Croston SI New CF	472.3	473.4	473.4	477.9	457.0	458.7	460.5	465.1
SB2 SI New CF	472.6	472.5	474.4	477.6	457.2	457.5	460.2	462.7

Table 7.5.3 displays the inventory holding cost ranks based on the statistically significant differences. The higher the inventory holding cost is, the higher (worse) the rank is.

Table 7.5.3 Inventory holding cost ranks based on statistically significant differences

Rank	**Item A**	**Item B**
1	Croston SI New CF, SB2 SI New CF, Croston SI New, SB2 SI New	Croston SI New CF, SB2 SI New CF
2	SB2	Croston SI New, SB2 SI New
3	Croston	Croston, SB2
4	SES	SES
5	Poisson SES	Poisson SES

An overall ranking of the methods is presented in Table 7.5.4 for the inventory holding cost metric.

Table 7.5.4 Summary of inventory holding cost ranks based on practical significance

Rank	Item A	Item B
1	Croston SI New CF, SB2 SI New CF, Croston SI New, SB2 SI New	Croston SI New CF, SB2 SI New CF
2	SB2	Croston SI New, SB2 SI New
3	Croston	Croston, SB2
4	SES	SES
5	Poisson SES	Poisson SES

The Poisson SES and SES methods have practically higher inventory holding cost than all other methods. The new methods have practically lower inventory holding cost compared to old methods.

7.6 Adjusted Margin Results

In this section the forecasting methods are evaluated based on the average adjusted margin metric. Recall adjusted margin is defined as the margin minus the inventory holding cost. Simulated performance for a target *CSL* value of 85% using smoothing coefficient, α, values of 0.05, 0.1, 0.15, and 0.2 are presented separately for items A and B. Results of pairwise adjusted margin comparisons using two sided *t*-test are also presented separately for items A and B. A ranking of the methods for this metric based on statistically significant differences are displayed in Table 7.6.3 and adjusted margin ranks of the methods based on practically significant differences are displayed in Table 7.6.4.

The confidence intervals of the mean adjusted margin calculated using the formulas in (7.13) are displayed in the following figures.

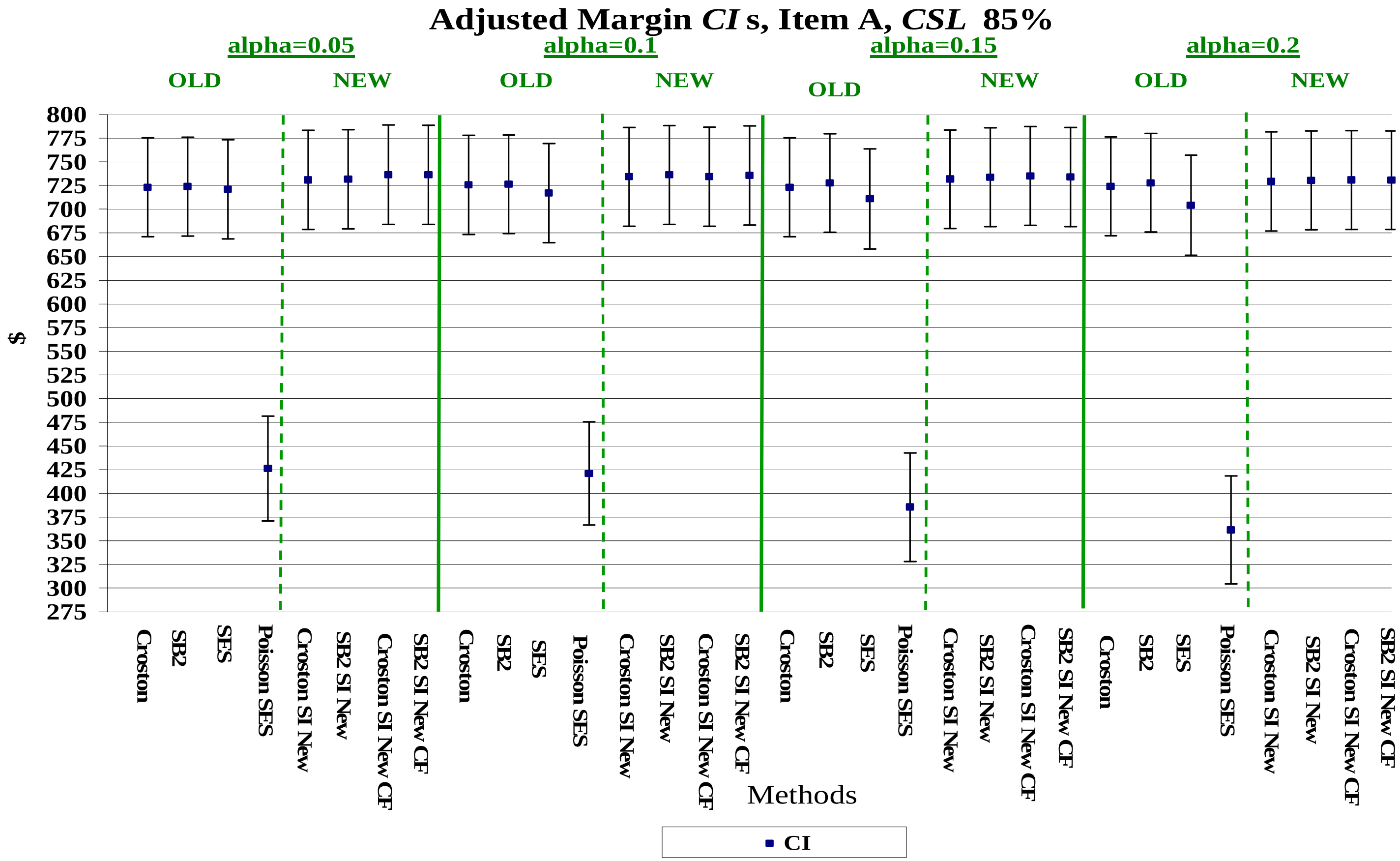

Figure 7.6.1 Item A adjusted margin confidence intervals, 85% *CSL*

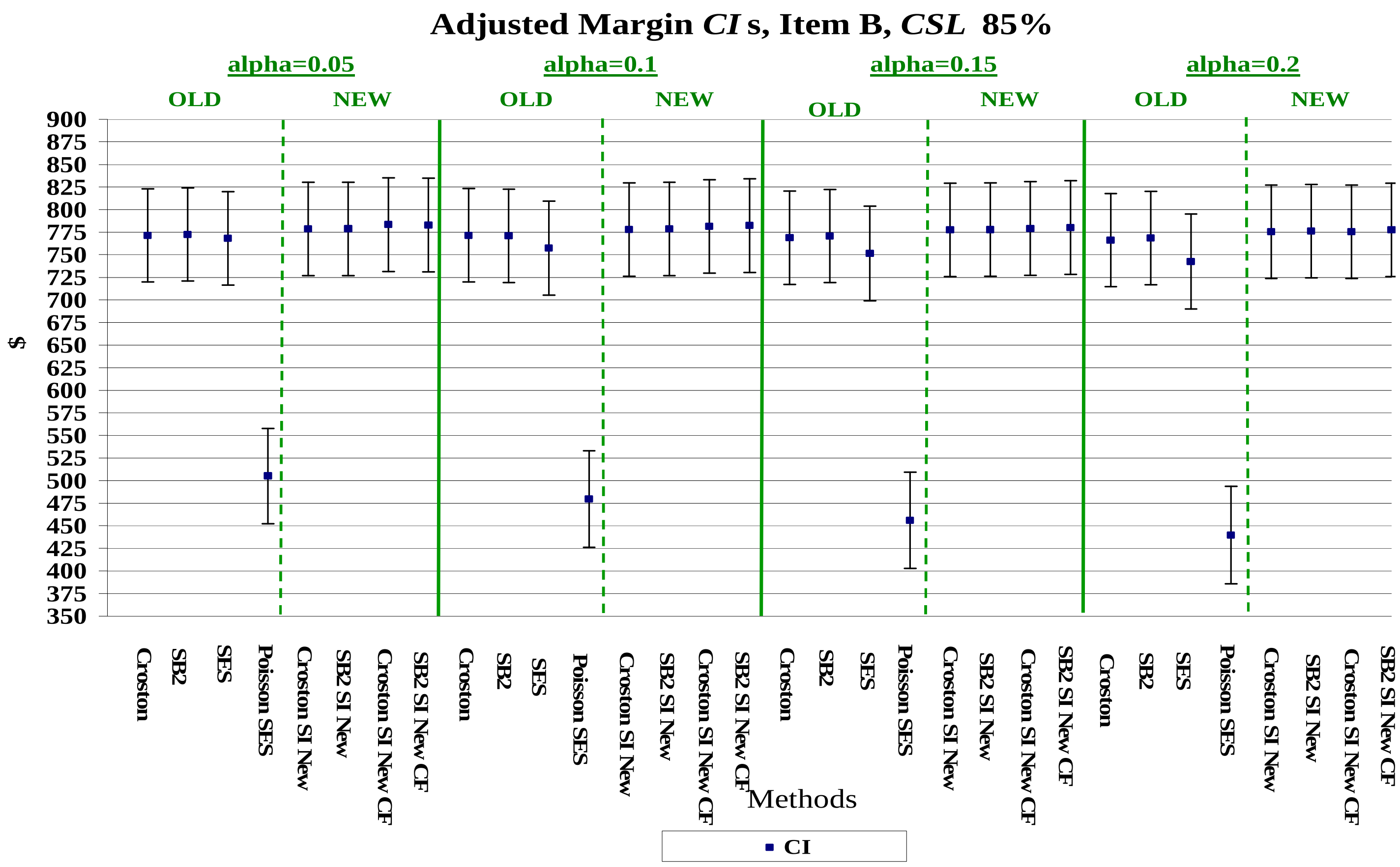

Figure 7.6.2 Item B adjusted margin confidence intervals, 85% *CSL*

Since all the confidence intervals overlap except for those for the Poisson SES method, as before *t*-tests for paired inventory holding cost are performed using (7.14) to determine the statistical differences between pairs of methods. For each method, the α value that generates the highest average fill rate is used for comparison. Table 7.6.1 displays the *t* values for each item sorted from smallest to largest. The values are bolded that are statistically significant.

Table 7.6.1 Paired-sample *t*-test values for adjusted margin differences

Methods Compared Method 1- Method 2	**Item A**	**Methods Compared Method 1- Method 2**	**Item B**
Poisson SES-SB2 SI New CF	**-31.68**	Poisson SES-Croston SI New CF	**-35.32**
Poisson SES-Croston SI New CF	**-31.62**	Poisson SES-SB2 SI New CF	**-35.3**
Poisson SES-SB2 SI New	**-31.45**	Poisson SES-SB2 SI New	**-34.75**
Poisson SES-Croston SI New	**-31.25**	Poisson SES-Croston SI New	**-34.7**
SES-SB2 SI New	**-7.75**	SES-Croston SI New CF	**-8.95**
SES-Croston SI New CF	**-7.57**	SES-SB2 SI New CF	**-8.64**
SES-SB2 SI New CF	**-7.37**	Croston-Croston SI New CF	**-7.56**
SES-Croston SI New	**-6.68**	Croston-SB2 SI New CF	**-7.12**
Croston-SB2 SI New	**-6.26**	SB2-Croston SI New CF	**-6.87**
Croston-Croston SI New CF	**-5.64**	SES-Croston SI New	**-6.56**
Croston-SB2 SI New CF	**-5.51**	SB2-SB2 SI New CF	**-6.5**
SB2-SB2 SI New	**-5.24**	SES-SB2 SI New	**-6.44**
Croston-Croston SI New	**-5.16**	Croston-Croston SI New	**-5.09**
SB2-Croston SI New	**-3.9**	Croston-SB2 SI New	**-5.02**
SB2-Croston SI New CF	**-3.87**	SB2-SB2 SI New	**-4.42**
SB2-SB2 SI New CF	**-3.87**	SB2-Croston SI New	**-4.36**
Croston SI New-SB2 SI New	-2.24	Croston SI New-Croston SI New CF	**-4.09**
Croston SI New-Croston SI New CF	-1.42	SB2 SI New-Croston SI New CF	**-3.96**
Croston-SB2	-1.41	SB2 SI New-SB2 SI New CF	**-3.6**
Croston SI New-SB2 SI New CF	-1.27	Croston SI New-SB2 SI New CF	**-3.56**
SB2 SI New-Croston SI New CF	-0.25	Croston-SB2	-2.01
SB2 SI New-SB2 SI New CF	-0.06	Croston SI New-SB2 SI New	-0.28
Croston SI New CF-SB2 SI New CF	0.51	Croston SI New CF-SB2 SI New CF	0.67
Croston-SES	**3.1**	Croston-SES	2.32
SB2-SES	**3.95**	SB2-SES	**3.04**
SES-Poisson SES	**30.47**	SES-Poisson SES	**33.5**
SB2-Poisson SES	**30.48**	SB2-Poisson SES	**34.51**
Croston-Poisson SES	**31.06**	Croston-Poisson SES	**34.53**

For item A, the old methods have lower adjusted margin than the new methods. There is no statistical difference between the new methods. In addition, there is no statistical difference between the Croston and SB2 methods. The Croston and SB2 methods have higher adjusted margin than the SES and Poisson SES methods. In addition, SES has higher adjusted margin than the Poisson SES method.

For item B, the old methods have lower adjusted margin than the new methods. In addition, the Croston SI New and SB2 SI New methods have lower adjusted margin than the Croston SI New CF and SB2 SI New CF methods. There is no statistical difference between the Croston and SB2 methods. Furthermore, there is no statistical difference between the Croston SI New and SB2 SI New methods. In addition, there is no statistical difference between the Croston SI New CF and SB2 SI New CF methods. The Croston and SES methods are statistically indifferent and the Croston method has higher adjusted margin than the Poisson SES method. The SB2 method has higher adjusted margin than the SES and Poisson SES methods. In addition, SES has higher adjusted margin than the Poisson SES method.

Table 7.6.2 displays the mean adjusted margin of the existing and new methods for items A and B for a target *CSL* value of 85%. For each method and each item, the best (highest) values across the smoothing coefficient values evaluated are highlighted and used to compare the methods.

Table 7.6.2 Mean adjusted margin

Methods	Item A				Item B			
Mean Adjusted margin								
	$\alpha=0.05$	$\alpha=0.1$	$\alpha=0.15$	$\alpha=0.2$	$\alpha=0.05$	$\alpha=0.1$	$\alpha=0.15$	$\alpha=0.2$
Existing Methods								
Croston	723.1	725.7	723.1	724.1	771.5	771.5	768.9	766.1
SB2	723.8	726.3	727.7	727.8	772.5	771	770.7	768.5
SES	721	717	710.9	704.2	768.1	757.4	751.4	742.6
Poisson SES	426.2	421	385.5	361.3	505.1	479.6	455.9	439.7
New Methods								
Croston SI New	731	734.2	731.8	729.3	778.6	778	777.5	775.4
SB2 SI New	731.6	736.1	733.8	730.4	778.7	778.7	777.9	776.2
Croston SI New CF	736.6	734.3	735.1	731	783.4	781.4	779	775.4
SB2 SI New CF	736.2	735.7	733.9	730.7	782.9	782.5	780.1	777.5

Table 7.6.3 displays the adjusted margin ranks based on the statistically significant differences.

Table 7.6.3 Adjusted margin ranks based on statistically significant differences

Rank	**Item A**	**Item B**
1	Croston SI New CF, SB2 SI New CF, Croston SI New, SB2 SI New	Croston SI New CF, SB2 SI New CF
2	SB2, Croston	Croston SI New, SB2 SI New
3	SES	Croston, SB2
4	Poisson SES	Croston , SES
5		Poisson SES

Table 7.6.4 displays the inventory holding cost ranks based on the practically significant differences.

Table 7.6.4 Summary of adjusted margin ranks based on practical significance

Rank	Item A	Item B
1	Croston SI New CF, SB2 SI New CF, Croston SI New, SB2 SI New	Croston SI New CF, SB2 SI New CF
2	SB2, Croston	Croston SI New, SB2 SI New
3	SES	Croston, SB2
4	Poisson SES	Croston , SES
5		Poisson SES

The Poisson SES method has practically lower adjusted margin than all other methods. In addition, the SES method has practically lower adjusted margin than all other methods, except the Croston method for item B.

The new methods have practically higher adjusted margin compared to old methods.

7.7 Gross Margin Return on Inventory (*GMROI*) Results

The confidence intervals of the mean *GMROI* based on a target *CSL* value of 85% for smoothing coefficient, α, values of 0.05, 0.1, 0.15, and 0.2 are presented separately for items A and B. In addition, the results of pairwise *GMROI* comparisons for the α value that generates the highest fill rate using two sided *t*-tests are also presented separately for items A and B. The ranks of the methods based on statistically significant differences are displayed in Table 7.7.3 and *GMROI* ranks of the methods based on practically significant differences are displayed in Table 7.7.4.

The confidence intervals of the mean *GMROI* calculated using the formulas in (7.13) are displayed in the following figures

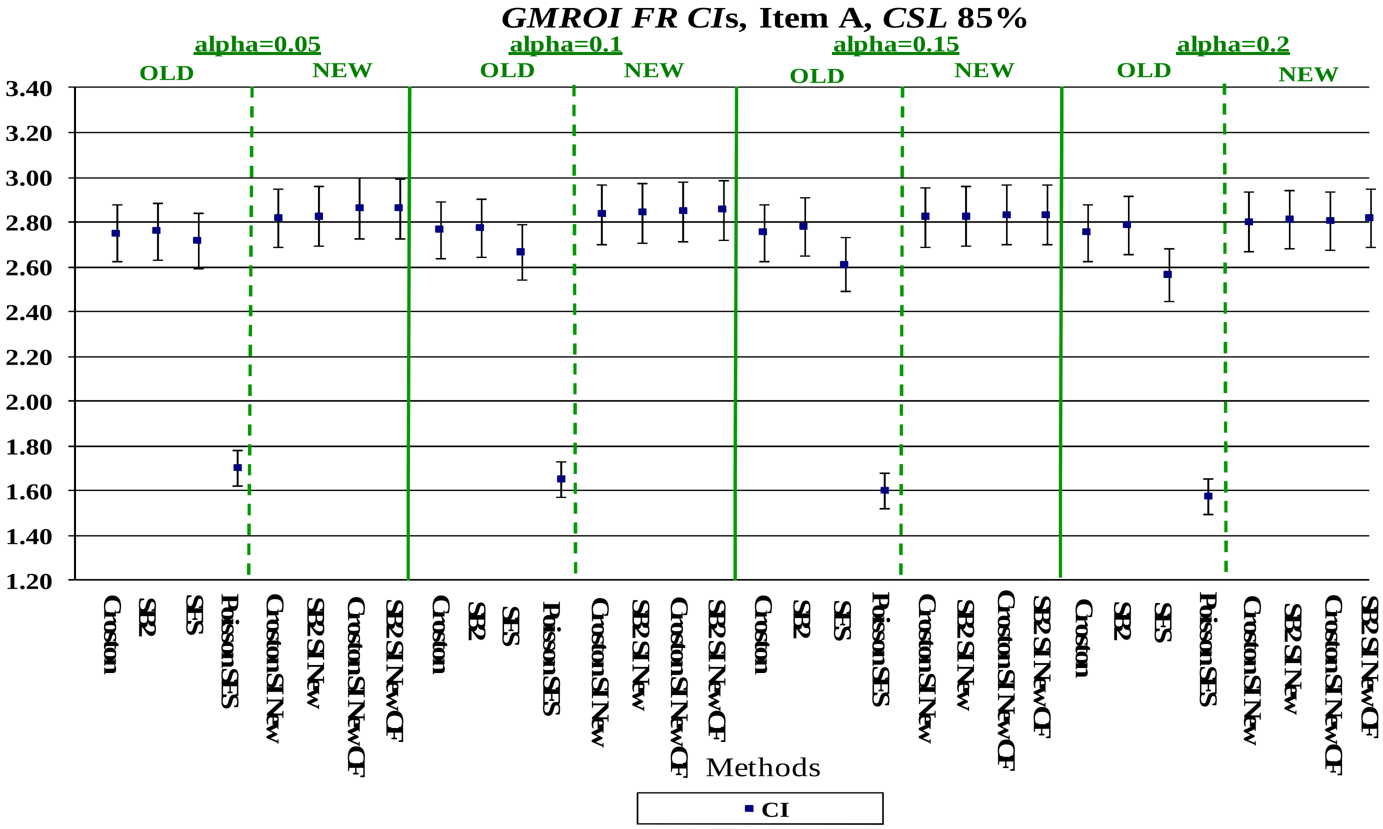

Figure 7.7.1 Item A *GMROI* confidence intervals, 85% *CSL*

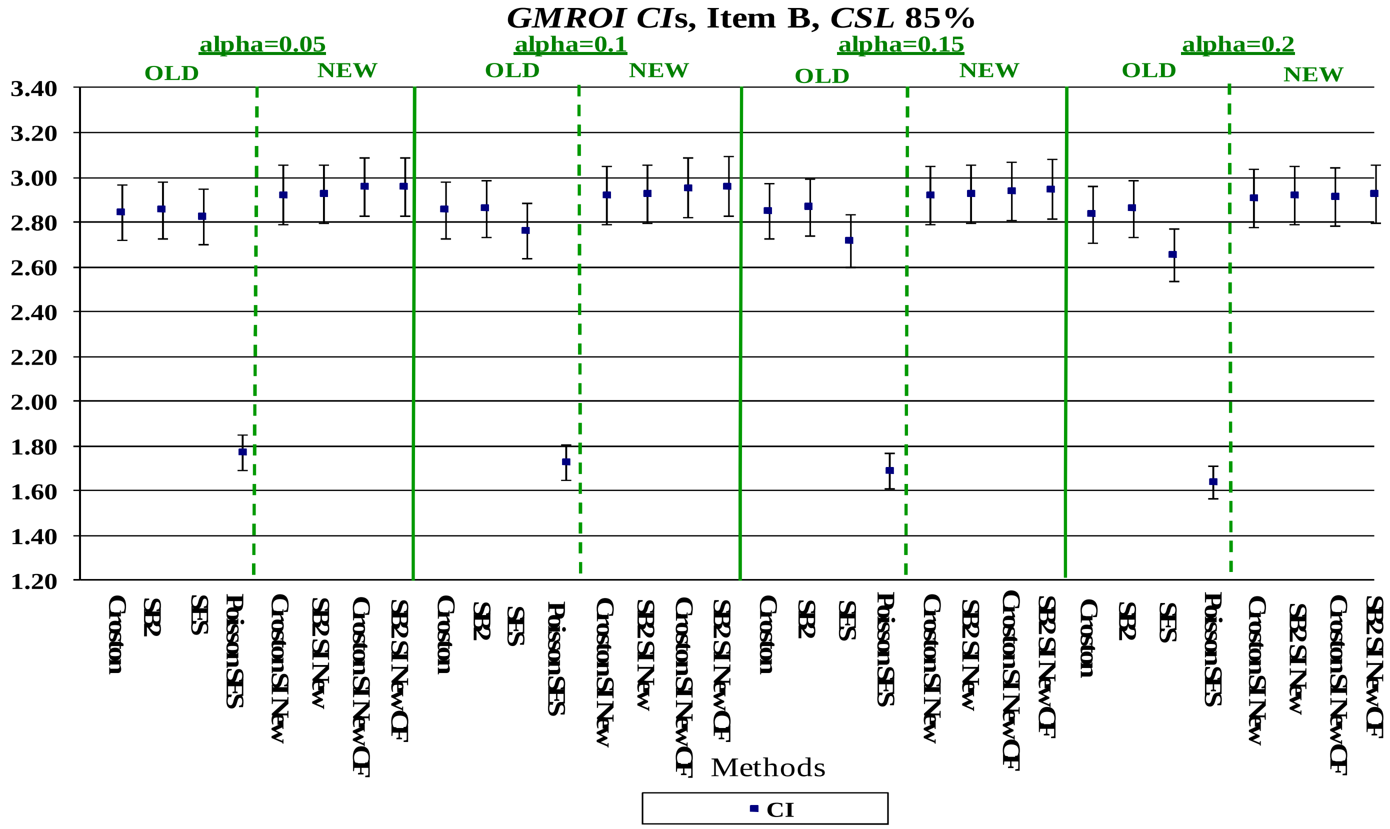

Figure 7.7.2 Item B *GMROI* confidence intervals, 85% *CSL*

Since all the confidence intervals overlap except for those for the Poisson SES method, as before *t*-tests for paired *GMROI* comparisons are performed using (7.14) to determine the statistical differences between pairs of methods. For each method, the α value that generates the highest average fill rate is used for comparison. Table 7.7.1 displays the *t* values for each item sorted from smallest to largest. The values are bolded that are statistically significant.

Table 7.7.1 Paired-sample *t*-test values for *GMROI* differences

Methods Compared Method 1- Method 2	**Item A**	**Methods Compared Method 1- Method 2**	**Item B**
Poisson SES-SB2 SI New	**-36.57**	Poisson SES-Croston SI New CF	**-37.22**
Poisson SES-SB2 SI New CF	**-36.36**	Poisson SES-SB2 SI New	**-36.06**
Poisson SES-Croston SI New	**-36.23**	Poisson SES-SB2 SI New CF	**-36.06**
Poisson SES-Croston SI New CF	**-35.43**	Poisson SES-Croston SI New	**-35.77**
SES-SB2 SI New CF	**-13.2**	SES-Croston SI New CF	**-13.53**
SES-SB2 SI New	**-12.2**	SES-SB2 SI New CF	**-12.41**
SES-Croston SI New	**-11.6**	Croston-Croston SI New CF	**-11.56**
SES-Croston SI New CF	**-11.6**	SB2-Croston SI New CF	**-10.69**
Croston-SB2 SI New CF	**-10.19**	Croston-SB2 SI New CF	**-10.49**
Croston-Croston SI New CF	**-9.53**	SB2-SB2 SI New CF	**-10**
Croston-SB2 SI New	**-9.48**	SES-SB2 SI New	**-9.91**
Croston-Croston SI New	**-8.8**	SES-Croston SI New	**-9.4**
SB2-SB2 SI New CF	**-8.47**	Croston-SB2 SI New	**-8.29**
SB2-Croston SI New CF	**-7.86**	Croston-Croston SI New	**-8.18**
SB2-SB2 SI New	**-7.8**	SB2-SB2 SI New	**-7.67**
SB2-Croston SI New	**-6.58**	SB2-Croston SI New	**-6.83**
Croston SI New-SB2 SI New CF	**-4.1**	SB2 SI New-Croston SI New CF	**-4.99**
Croston SI New-Croston SI New CF	**-3.9**	SB2 SI New-SB2 SI New CF	**-4.68**
Croston-SB2	**-3.16**	Croston SI New-Croston SI New CF	**-4.66**
SB2 SI New-SB2 SI New CF	**-2.71**	Croston SI New-SB2 SI New CF	**-4.66**
SB2 SI New-Croston SI New CF	**-2.69**	Croston-SB2	**-2.61**
Croston SI New-SB2 SI New	-2.2	Croston SI New-SB2 SI New	-0.61
Croston SI New CF-SB2 SI New CF	0.78	Croston SI New CF-SB2 SI New CF	-0.43
Croston-SES	**5.42**	Croston-SES	**3.49**
SB2-SES	**8.3**	SB2-SES	**5.13**
SES-Poisson SES	**35.47**	SB2-Poisson SES	**35.33**
SB2-Poisson SES	**35.6**	Croston-Poisson SES	**35.57**
Croston-Poisson SES	**35.65**	SES-Poisson SES	**35.69**

For both items, the old methods have lower *GMROI* than the new methods. The Croston SI New and SB2 SI New methods have lower *GMROI* than the Croston SI New CF and SB2 SI New CF methods. There is no statistical difference between the Croston and SB2 methods. In addition, there is no statistical difference between the Croston SI New and SB2 SI New methods. Furthermore, there is no statistical difference between the Croston SI New CF and SB2 SI New CF methods.

The Croston and SB2 methods have higher *GMROI* than the SES and Poisson SES methods. In addition, SES has higher *GMROI* than the Poisson SES method.

Tables 7.7.2 displays the mean *GMROI* from all methods for items A and B based on *CSL* of 85%. The best (highest) values for each method and item are highlighted and are used to compare the methods with practical significance if they are statistically different.

Table 7.7.2 Mean *GMROI*

Mean *GMROI*								
	Item A				Item B			
Methods	$\alpha=0.05$	$\alpha=0.1$	$\alpha=0.15$	$\alpha=0.2$	$\alpha=0.05$	$\alpha=0.1$	$\alpha=0.15$	$\alpha=0.2$
Existing Methods								
Croston	2.75	2.77	2.75	2.75	2.85	2.86	2.85	2.84
SB2	2.76	2.77	2.78	2.79	2.86	2.86	2.87	2.86
SES	2.72	2.67	2.61	2.57	2.83	2.76	2.72	2.65
Poisson SES	1.7	1.65	1.6	1.58	1.77	1.73	1.69	1.64
New Methods								
Croston SI New	2.82	2.84	2.82	2.8	2.92	2.92	2.92	2.91
SB2 SI New	2.83	2.84	2.83	2.81	2.93	2.93	2.93	2.92
Croston SI New CF	2.87	2.85	2.83	2.81	2.96	2.95	2.94	2.91
SB2 SI New CF	2.86	2.85	2.83	2.82	2.96	2.96	2.95	2.93

Table 7.7.3 displays the *GMROI* ranks based on the statistically significant differences. If the two methods are statistically significant and the maximum mean *GMROI* differences between the two methods are greater than or equal to 0.01, the methods are practically significant. The definition of "practically significant" is subjective and can change based on the company's policy.

Table 7.7.3 Summary of *GMROI* ranks based on statistically significant differences

Rank	**Item A**	**Item B**
1	Croston SI New CF, SB2 SI New CF	Croston SI New CF, SB2 SI New CF
2	Croston SI New, SB2 SI New	Croston SI New, SB2 SI New
3	SB2	SB2
4	Croston	Croston
5	SES	SES
6	Poisson SES	Poisson SES

Table 7.7.4 displays the *GMROI* ranks based on the practically significant differences.

Table 7.7.4 Summary of *GMROI* ranks based on practically significant differences

Rank	**Item A**	**Item B**
1	Croston SI New CF, SB2 SI New CF	Croston SI New CF, SB2 SI New CF
2	Croston SI New, SB2 SI New	Croston SI New, SB2 SI New
3	SB2	SB2
4	Croston	Croston
5	SES	SES
6	Poisson SES	Poisson SES

Poisson SES and SES have practically lower *GMROI* than all other methods. The new methods have practically higher *GMROI* compared to old methods. Croston SI New CF and SB2 SI New CF methods have practically higher *GMROI* than all other methods.

7.8 Conclusion of Inventory Control Analysis Metrics and Discussion

In this chapter fill rate, margin, inventory holding cost, adjusted margin, and *GMROI* are used as financial metrics to compare forecasting methods. Comparisons are done using the confidence intervals of the metrics, the mean metrics, and the results of pairwise metric comparisons with two sided *t*-tests for items A and B based on *CSL* of 85%. Since all the confidence intervals of metrics overlap, except for the Poisson SES method for some metrics, *t*-test values for paired metric comparisons are used to determine the statistical differences between methods.

For the inventory holding cost, margin, and adjusted margin metrics, if the difference in performance between the two methods is statistically significant and the mean difference between the two methods is greater than or equal to 50 cents, then the difference in performance between the methods is also significant. The definition of "practically significance" is subjective and can change based on the company's policy and 50 cents might seem low to determine practical significance.

As expected, Poisson SES-generated forecasts have a higher fill rate and margin compared to other methods, since more inventory is held at the stores. As expected the mean fill rates are significantly greater than the *CSL* value of 85%. Since the mean fill rate differences between all methods are not greater than 1%, the differences are not practically significant.

The old methods have practically higher inventory holding cost, and lower adjusted margin and *GMROI* than the new methods. In addition, the Poisson SES method has the highest inventory holding cost, the lowest adjusted margin, and *GMROI*. The SB2 method has higher *GMROI* than the Croston method, and the Croston method has higher *GMROI* than the SES method. Even though for item A there is no statistical difference of adjusted margin between the new methods; for item B the Croston SI New CF and SB2 SI New CF methods have better adjusted margin than the Croston SI New and SB2 SI New methods; and this result is also practically significant. Furthermore, for both items the Croston SI New CF and SB2 SI New CF methods have better *GMROI* than the Croston SI New, SB2 SI New methods; and this result is also practically significant.

To summarize, the new methods generated lower inventory holding cost and higher adjusted margin and *GMROI*. Thus, the new methods perform better than the existing methods. In addition, based on these inventory results, the SB2 SI New CF and Croston SI New CF methods are the best to forecast the sales of both items since these methods generate the highest adjusted margin (except for item A) and *GMROI*.

Before evaluating the inventory performance measures of the methods in general, in this section the correction factor parameters used in two of the methods evaluated in this work (Croston SI New CF and SB2 SI New CF) are evaluated in order to determine the appropriate values for these parameters. As stated in Section 3.2.3, the value of the correction (*CF*) for item *i* and at the k^{th} store at time period *t*, involves the use of four input parameters: a maximum correction factor (*MCF*), an upper limit (*UL*), a lower limit (*LL*), and an update frequency.

The forecasts using with the Croston SI New CF and SB2 SI New CF methods are simulated using *UL* values of 1.2 and 1.3, *LL* values of 0.7 and 0.8, *MCF* values of 2 and 3, update frequency values of 1, 2, and 3; therefore a total of 24 parameter set combinations are compared using the inventory metrics of this chapter.

In addition to the analysis above, before comparing the eight methods using the inventory control metrics, the *CF* parameter set combinations of the Croston SI New CF and SB2 SI New CF methods are determined. To achieve that, as described in Section 7.1.2, the performance of a forecasting method (Croston SI New CF and SB2 SI New CF) using a given parameter set combination and a specified target cycle service level (*CSL*=85%) is computed for each item (A and B). The five inventory performance metrics of this Chapter (fill rate, margin, inventory holding cost, adjusted margin, and *GMROI*) are used to evaluate the parameter set combinations. Specifically, for each of the two items (A and B) and for each combination of the forecasting method and parameter set combination, simulations are run over all series (stores) for which demand data is available. Then, confidence intervals on the mean performance are computed. The resulting confidence intervals overlap, and this suggests no significance difference in performance. A similar observation can be made about the update frequency. Therefore, for the experimentation, we use parameter set 3 since it often leads to the best average performance across all the adjusted margin and *GMROI* metrics. This parameter set combination has *MCF* of 2, *UL* of 1.2, *LL* of 0.8, and update frequency of 1.

8 Conclusion and Future Work

Inventory management is very difficult if the sales are slow-moving where demand appears at random and with many time periods having no demand at all. Even though slow-moving demand is more common in the service parts business, it is also a difficult problem in the retail industry.

Exponential smoothing methods and the approach of Croston (1972) and its modifications are very popular to estimate the sales of slow-moving items. However, exponential smoothing generally leads to improper stock levels since more weight is placed on the most recent data, making the forecast highest just after a demand occurrence, and lowest just before the demand occurs again. Croston proposed an alternative method that considers not only the demand size but also the inter-arrival time between demands to overcome this situation. Croston's method and its modifications adjust the forecast only after a demand is observed, but this may not be very practical if the inter-arrival times between demands are very long. In addition, both exponential smoothing and Croston do not work well if the demand is seasonal.

In this research, we add a seasonality index to the existing exponential smoothing and Croston modification methods. To the best of our knowledge, this is the only research in the literature that uses seasonality indices to adjust the forecast of slow moving items. In addition, a correction factor is included that adjusts the forecast if the forecast error is more than a threshold value. Moreover, in some of our new models, we also adjust the forecast not only after a demand like Croston, but also if the time since the last demand exceeds the average inter-arrival time or two times the average inter-arrival time depending on the new method.

The existing and new forecasting methods are compared using sales forecast accuracy metrics of *ME*, *MAE*, *SMAE*, and *GRMSE*. The accuracy metrics are computed at "all points in time" and at "issue points only," for the two items, A and B separately, and for smoothing values of $\alpha = 0.05$, 0.1, 0.15, and 0.2. Due to the very large sample sizes available for this analysis, statistical tests indicate that any differences in the performance metrics are statistically significant. In order to rank methods, we compare the average performance of the methods for each accuracy metric based on "all points in time" performances as well as the performances at "issue points only." The best performance for a given method and metric is selected across the four levels of smoothing parameter value α for "all points in time" and at issue points separately. We remove the scale dependency for the *ME*, *MAE*, and *SMAE* metrics by dividing by the mean demand over the testing horizon, i.e. mean *ME*/mean demand, mean *MAE*/mean demand, and mean *SMAE*/ mean demand. Then, these values are ranked across the methods.

The mean *GRMSE* measure is used directly to rank the methods. One observation that needs to be mentioned again is that *GRMSE* sales forecast accuracy measure cannot be used to compare forecasting methods which generates integer forecasts, such as the Poisson SES. Thus, *GRMSE* could not be used to evaluate the Poisson SES method.

For both items and for "all points in time," all existing methods overestimate the sales. On the other hand, all new methods underestimate the sales, except for the Croston SI

New and SB2 SI New methods with α =0.05. Moreover, for item A the SB2 SI New method has the lowest absolute value of mean *ME*/mean demand among all methods and the SES method had the lowest absolute value of mean *ME*/mean demand among all the methods.

For both items and at "issue points only," all existing methods overestimate the sales. On the other hand, all new methods underestimate the sales, except for the Croston SI New and SB2 SI New methods with α =0.05. Moreover, the SB2 SI New method has the lowest absolute value of mean *ME*/mean demand among all methods. All new methods perform better than the old methods at "issue points only."

Based on the *MAE*, *SMAE*, and *GMRSE* accuracy measures used, the new methods perform better than the old methods. In addition, the SB2 SI New CF method performs the best among all methods.

In order to determine whether a method performs better than another method, one cannot only use sales forecast accuracy measures as a criterion since determination of practical significance can be difficult to determine. Instead, inventory control metrics should also be taken into consideration while making comparisons, since a better sales forecast accuracy measure might not reflect better inventory metrics. Thus, in addition to sales forecast accuracy metrics, financially oriented measures such as fill rate, margin, inventory holding cost, adjusted margin, and *GMROI* are used to compare the forecasting algorithms. These comparisons are done using the confidence intervals of the metrics, the mean metrics, and the results of pairwise metric comparisons with two sided *t*-tests for items A and B based on *CSL* of 85%. Since all the confidence intervals of metrics overlap except for the Poisson

SES method, *t*-test values for paired metric comparisons are used to determine the statistical differences between methods.

As expected, Poisson SES-generated forecasts have a higher fill rate and margin compared to other methods, since more inventory is held at the stores. The mean fill rates are significantly greater than the *CSL* value of 85%. Since all fill rates are greater than the *CSL* value of 85%, and the mean fill rate differences between methods are not greater than 1%, the differences are not practically significant.

The old methods have practically higher inventory holding cost, and lower adjusted margin and *GMROI* than the new methods. In addition, the Poisson SES method has the highest inventory holding cost, the lowest adjusted margin, and *GMROI*. The SB2 method has higher *GMROI* than the Croston method, and the Croston method has higher *GMROI* than the SES method. Although for item A there is no statistical difference of adjusted margin between the new methods; for item B the Croston SI New CF and SB2 SI New CF methods have better adjusted margin than the Croston SI New and SB2 SI New methods; and this result is also practically significant. Furthermore, for both items the Croston SI New CF and SB2 SI New CF methods have better *GMROI* than the Croston SI New, SB2 SI New methods; and this result is also practically significant.

To summarize, the new methods generated lower inventory holding cost and higher adjusted margin and *GMROI*. Thus, the new methods perform better than the existing methods. In addition, based on these inventory results, the SB2 SI New CF and Croston SI

New CF methods are the best to forecast the sales of both items since these methods generate the highest adjusted margin (except for item A) and *GMROI*.

Finally, the best practice for companies with slow selling items is selecting the best sales forecast method based on the company goals using simulation techniques at least twice a year, since one method works well might not work well in the future as the company goals as well as sales behavior can change with time. In addition, similar to what was done in Chapter 7 for the Croston SI CF New and SB2 CF New methods, the input parameters *MCF*, *UL*, *LL*, and update frequency should be optimized for each slow-seller item and then these parameters should be used in comparisons.

In this research, the methods do not have any “trend” or “product life cycle” component. As part of future work, adjusting the forecast using the “trend” and also taking the “product life cycle” into consideration will be investigated.

In addition, the outliers are not removed from the data before generating the forecast. This will also be analyzed as future research. Furthermore, *CF* and seasonality index calculations can also be improved even further by using regional values instead of country level values.

We used the inventory control model of the company to find the best forecasting method based on the inventory metrics. As part of future research, the best inventory control model that gives the best adjusted margin and *GMROI* will be investigated. As an example, preliminary calculations showed that the general safety stock model explained in sub-section

7.1.1 gave better adjusted margin and *GMROI* values than the company's inventory control model.

In this research, only two appliance items are used to compare the forecasting methods. As part of future research, additional seasonal/promotional slow-moving items such as tools, rugs etc. will be used to evaluate the forecasting methods.

Moreover, the inventory model that is used in this thesis is based on a two-tier vendor to store direct network. However, in real life, most of the items are being supplied to the stores with a three-tier network (vendor to regional distribution centers (RDCs) and then to the stores). In addition to two-tier network analysis, three-tier network will be analyzed as part of the future research. To the best of our knowledge, this has never been done in comparing forecasting methods for slow selling items. In three-tier networks, the items are stocked in the RDCs and they also have a safety stock, as well as *s* and *S* levels at the RDCs. Our main assumption will be that vendors do not have any constraints to supply the orders such as production constraints, unexpected shutdowns, and holidays, etc. The challenge is to come up with the RDC shipment forecast *ss*, *s*, and *S* levels. RDC shipments can be estimated using estimated store need using a multi-echelon logic.

Finally, there are a very large number of paired-sample t-tests in Chapter 7 on which are based several conclusions about the relative performance of the various forecasting methods with respect to a variety of economic metrics. The problem with this kind of comparison is that although each paired-sample t-test is performed at the significance level of 1%, the overall "experiment wise" error rate for all these tests taken together is substantially higher than 1%. That is, the probability that there is at least one incorrect conclusion among all these pairwise comparisons is much greater than 1%. Multiple comparison procedures are specifically designed to control the overall probability of committing a type I error in making a large number of such comparisons. In the current situation, it appears that the economic performance metrics of interest would require the use of nonparametric multiple comparison procedures (that is, procedures capable of handling nonnormal responses); and this could require the use of statistical methods for which there is virtually no commercial software of the quality and widespread acceptance that is comparable. Thus, a more definitive analysis of the relative economic performance of the various forecasting methods will be part of our future research.

References

Altay, N., Rudisill, F., and Litteral, LA. (2008), Adapting Wright's modification of Holt's method to forecasting intermittent demand. International Journal of Production Economics, **111**, 389-408.

Boylan, J.E., Syntetos, A. (2007), The accuracy of a modified Croston procedure. *International Journal of Production Economics*, **107**, 511-517.

Chatfield, C. (1992), A commentary on error measures. *International Journal of Forecasting,* **8**, 99-111.

Croston JF (1972), Forecasting and stock control for intermittent demands. *Operational Research Quarterly,* **23**, 289-304.

Fildes, R. (1992), The evaluation of extrapolative forecasting methods. *International Journal of Forecasting,* **8**, 91-98.

Fildes, R., Nikolopoulos, K., Crone, S., Syntetos, A.A., (2008), Forecasting and operational research: a review. *Journal of the Operational Research Society*, **59**, 1150–1172.

Gutierrez, R.S., Solis, A.O., Mukhopadhyay, S., (2008), Lumpy demand forecasting using neural networks. International Journal of Production Economics 111, 409–420.

Hoover, J (2006), Measuring Forecast Accuracy: Omissions in Today's Forecasting Engines and Demand Planning Software. FORESIGHT: International Journal of Applied Forecasting, 4, 32-35.

Hua ZS, Zhang B, Yang J and Tan DS (2007), A new approach of forecasting intermittent demand for spare parts inventories in the process industries. *Journal of the Operational Research Society*, **58**, 52-61.

Johnston FR and Boylan JE (1996), Forecasting for items of intermittent demand. *Journal of the Operational Research Society*, **47**, 113-121.

Levén, E., Segerstedt, A. (2004), Inventory control with a modified Croston procedure and Erlang distribution. *International Journal of Production Economics*, **90**, 361-367.

Makridakis, S. (1993), Accuracy Measures: Theoretical and Practical Concerns. *International Journal of Forecasting* **9**: 527-529.

Makridakis, S., Hibon, M. (2000), The M3-competition: Results, conclusions and implications, *International Journal of Forecasting*, 16, 451–476.

Pearson, R.L., Wallace, W.A. (1999), Appropriate measures of forecast error for business decision making. Presented at 19th International Symposium on Forecasting, June 27-30 (Washington, DC, USA).

Rao, A.V. (1973), A comment on: forecasting and stock control for intermittent demands. *Operational Research Quarterly* **24**(4), 639–640.

Schneider, H. (1978), Methods for determining the reorder point of an (*s*, *S*) ordering policy. *J Opl Res Soc*, 29, 1181-1193.

Smart, C.N. (2002), Accurate intermittent demand/inventory forecasting: new technologies and dramatic results. *In: Proceedings of APICS International Conference*, p. C-08.

Smart, C.N., Willemain, T.R. (2000), A new way to forecast intermittent demand. *APICS The Performance Advantage* 64–68.

Smith, S. and T. Sincich. (1988), Stability over Time in the Distribution of Population Forecast Errors. *Demography* **25**: 461-474.

Snyder, R. (2002), Forecasting sales of slow and fast moving inventories. European *Journal of Operational Research* **140**, 684–699.

Syntetos, A.A. (2001), Forecasting of intermittent demand. Unpublished Ph.D. Thesis, Buckinghamshire Business School, Brunel University.

Syntetos, A.A., Boylan, J.E. (2001), On the Bias of intermittent demand estimates. *International Journal of Production Economics,* **71**, 457-466.

Syntetos, A.A., Boylan, J.E. (2005), The accuracy of intermittent demand estimates. *International Journal of forecasting*, **21**, 303-314.

Syntetos, A.A., Boylan, J.E. (2006), On the stock control performance of intermittent demand estimators. *International Journal of Production Economics* **103**, 36–47.

Syntetos, A.A., Boylan, J.E. (2007), The accuracy of a modified Croston procedure. *International Journal of Production Economics* **107**, 511–517.

Syntetos, A.A., Nikolopoulos, K., Boylan, J.E., Fildes, R., Goodwin, P. (2009), Integrating management judgement into intermittent demand forecasts. *International Journal of Production Economics*, 118, 72–81.

Syntetos, A.A., Babai MZ, Davies J., Stephenson D. (2010), Forecasting and stock control: A study in a wholesaling context. *International Journal of Production Economics*, 127, 103–111.

Willemain TR, Smart CN, Shockor JH and DeSautels PA (1994), Forecasting intermittent demand in manufacturing: a comparative evaluation of Croston's method. *International Journal of Forecasting*, **10**, 529-538.

Willemain, T.R., Smart, C.N., Schwarz, H.F. (2004), A new approach to forecasting intermittent demand for service parts inventories. *International Journal of Forecasting* **20**, 375-387.

APPENDICES

Appendix A The Appendix for Chapter 6

Appendix A.1 *ME* Results

Table A.1.1 Paired-sample *t*-test values for *ME* differences, item A

Item A, *ME* Comparison, *t* value								
	All Points In Time				Issue Points Only			
Methods Compared Method 1- Method 2	α values				α values			
	0.05	**0.1**	**0.15**	**0.2**	**0.05**	**0.1**	**0.15**	**0.2**
Croston-SB2	-51.80	-59.71	-64.00	-65.90	-66.14	-76.70	-82.80	-86.43
Croston-SES	-26.01	-27.68	-28.61	-29.30	-5.04	26.91	62.24	85.43
Croston-Poisson SES	-19.85	-21.86	-21.32	-19.87	-1.04	5.65	13.33	19.08
Croston-Croston SI New	-37.98	-45.65	-49.39	-50.14	-46.06	-55.47	-60.08	-61.01
Croston-SB2 SI New	-38.83	-47.29	-52.06	-54.03	-47.53	-58.45	-65.49	-70.04
Croston-Croston SI New CF	-47.07	-56.63	-60.78	-61.61	-57.99	-68.09	-70.96	-69.13
Croston-SB2 SI New CF	-47.34	-57.08	-61.60	-62.97	-58.53	-69.25	-73.65	-74.71
SB2-SES	-24.47	-24.34	-22.90	-21.06	-1.74	34.74	73.54	97.00
SB2-Poisson SES	-18.37	-18.83	-16.60	-13.43	-0.18	7.21	15.27	21.32
SB2-Croston SI New	-37.06	-43.78	-46.10	-45.00	-44.47	-52.01	-53.17	-48.41
SB2-SB2 SI New	-37.98	-45.65	-49.39	-50.14	-46.07	-55.46	-60.05	-60.99
SB2-Croston SI New CF	-46.77	-56.09	-59.71	-59.67	-57.41	-66.71	-67.42	-61.04
SB2-SB2 SI New CF	-47.07	-56.62	-60.77	-61.60	-58.00	-68.09	-70.96	-69.13
SES-Poisson SES	0.12	-0.25	-0.28	1.05	0.26	-1.01	0.00	0.38
SES-Croston SI New	-27.67	-34.38	-35.84	-34.83	-77.39	-106.7	-115.2	-116.3
SES-SB2 SI New	-30.72	-38.53	-41.43	-42.33	-81.71	-111.6	-120.6	-122.3
SES-Croston SI New CF	-59.93	-65.26	-64.54	-63.07	-112.5	-131.2	-130.9	-126.5
SES-SB2 SI New CF	-60.08	-65.14	-64.38	-63.07	-112.3	-131.7	-133.0	-130.1
Poisson SES-Croston SI New	-12.95	-18.52	-17.89	-17.34	-11.03	-18.24	-23.25	-26.70
Poisson SES-SB2 SI New	-14.47	-21.03	-21.33	-21.78	-11.62	-19.18	-24.46	-28.16
Poisson SES-Croston SI New CF	-31.17	-36.05	-33.59	-31.71	-17.24	-22.99	-26.22	-28.58
Poisson SES-SB2 SI New CF	-31.91	-37.16	-35.38	-34.16	-17.66	-23.67	-27.20	-29.84
Croston SI New-SB2 SI New	-52.99	-53.04	-53.02	-53.48	-64.14	-66.55	-68.97	-71.69
Croston SI New-Croston SI New CF	-51.46	-50.15	-46.64	-44.92	-55.45	-46.40	-35.80	-27.03
Croston SI New-SB2 SI New CF	-52.06	-51.53	-49.38	-49.07	-56.70	-50.20	-44.37	-42.04
SB2 SI New-Croston SI New CF	-50.70	-48.15	-42.18	-37.15	-53.99	-41.51	-23.98	-5.61
SB2 SI New-SB2 SI New CF	-51.46	-50.15	-46.63	-44.90	-55.49	-46.42	-35.80	-27.03
Croston SI New CF-SB2 SI New CF	-49.30	-47.45	-48.06	-49.03	-63.69	-66.87	-70.56	-73.47

Table A.1.2 Paired-sample *t*-test values for *ME* differences, item B

Item B, *ME* Comparison, *t* value								
	All Points In Time				Issue Points Only			
Methods Compared Method 1- Method 2	α values				α values			
	0.05	**0.1**	**0.15**	**0.2**	**0.05**	**0.1**	**0.15**	**0.2**
Croston-SB2	-66.29	-72.09	-74.77	-75.58	-83.33	-90.41	-94.27	-96.20
Croston-SES	-30.64	-35.12	-37.37	-38.25	0.17	32.02	65.22	86.86
Croston-Poisson SES	-20.74	-25.19	-24.69	-24.62	-0.02	6.22	11.68	18.33
Croston-Croston SI New	-43.11	-51.33	-55.50	-56.62	-44.46	-55.16	-59.84	-60.87
Croston-SB2 SI New	-44.69	-54.03	-59.51	-62.02	-47.09	-59.97	-67.78	-73.01
Croston-Croston SI New CF	-60.47	-71.36	-75.46	-75.48	-67.95	-79.11	-81.14	-78.60
Croston-SB2 SI New CF	-60.46	-71.91	-76.26	-76.75	-68.87	-80.95	-84.78	-85.37
SB2-SES	-28.20	-30.37	-29.90	-28.04	4.76	41.43	77.60	99.34
SB2-Poisson SES	-18.89	-21.43	-18.98	-17.16	0.88	7.76	13.58	20.78
SB2-Croston SI New	-41.39	-48.24	-50.58	-49.52	-41.66	-49.76	-50.34	-45.38
SB2-SB2 SI New	-43.11	-51.33	-55.50	-56.62	-44.46	-55.16	-59.83	-60.85
SB2-Croston SI New CF	-60.01	-70.68	-74.26	-73.40	-66.85	-76.92	-76.30	-68.70
SB2-SB2 SI New CF	-60.01	-71.36	-75.45	-75.48	-67.84	-79.11	-81.13	-78.59
SES-Poisson SES	-1.80	-1.05	0.47	0.40	-0.06	0.07	0.26	0.65
SES-Croston SI New	-22.16	-28.52	-29.96	-28.80	-63.57	-96.47	-109.6	-114.0
SES-SB2 SI New	-25.76	-33.61	-36.92	-38.01	-68.20	-102.5	-116.4	-121.3
SES-Croston SI New CF	-64.99	-69.01	-68.27	-65.76	-113.3	-132.1	-131.9	-127.8
SES-SB2 SI New CF	-65.06	-68.88	-67.98	-65.76	-114.5	-133.1	-134.5	-132.1
Poisson SES-Croston SI New	-8.18	-14.14	-15.18	-14.09	-10.02	-16.96	-20.39	-26.01
Poisson SES-SB2 SI New	-9.87	-17.10	-19.02	-19.05	-10.69	-17.97	-21.65	-27.68
Poisson SES-Croston SI New CF	-30.61	-36.52	-33.32	-31.17	-17.59	-22.52	-23.96	-28.70
Poisson SES-SB2 SI New CF	-30.99	-37.90	-35.24	-33.79	-17.91	-23.23	-24.94	-30.09
Croston SI New-SB2 SI New	-65.09	-60.43	-57.93	-57.05	-77.57	-74.83	-74.73	-75.60
Croston SI New-Croston SI New CF	-62.10	-58.63	-53.67	-50.25	-63.59	-52.70	-41.54	-31.85
Croston SI New-SB2 SI New CF	-62.11	-60.18	-56.33	-54.15	-64.75	-56.72	-49.73	-45.65
SB2 SI New-Croston SI New CF	-61.08	-56.38	-49.21	-42.82	-61.66	-47.63	-30.48	-12.47
SB2 SI New-SB2 SI New CF	-61.16	-58.63	-53.66	-50.24	-63.04	-52.70	-41.54	-31.85
Croston SI New CF-SB2 SI New CF	-28.80	-51.92	-50.42	-50.78	-32.53	-73.10	-74.75	-76.73

Appendix A.2 *MAE* Results

Table A.2.1 Paired-sample *t*-test values for *MAE* differences, item A

Item A, *MAE* Comparison, *t* value								
	All Points In Time				Issue Points Only			
Methods Compared Method 1- Method 2	α values				α values			
	0.05	**0.1**	**0.15**	**0.2**	**0.05**	**0.1**	**0.15**	**0.2**
Croston-SB2	54.72	65.31	71.09	73.57	47.80	55.27	59.58	62.06
Croston-SES	25.50	26.85	27.19	27.11	3.77	-20.31	-46.91	-64.61
Croston-Poisson SES	15.73	18.31	17.24	15.24	-0.59	-5.43	-10.98	-16.74
Croston-Croston SI New	37.29	43.80	46.16	45.73	34.91	41.87	45.54	46.67
Croston-SB2 SI New	38.18	45.61	49.16	50.05	35.91	43.89	49.21	52.77
Croston-Croston SI New CF	47.28	55.94	58.67	58.18	43.24	50.77	53.36	52.60
Croston-SB2 SI New CF	47.65	56.74	60.17	60.58	43.58	51.49	55.05	56.17
SB2-SES	24.04	23.82	22.30	20.39	1.39	-26.00	-55.12	-72.97
SB2-Poisson SES	14.33	15.70	13.22	9.95	-1.22	-6.56	-12.39	-18.36
SB2-Croston SI New	36.35	41.82	42.75	40.64	33.83	39.50	40.80	37.99
SB2-SB2 SI New	37.30	43.81	46.16	45.74	34.92	41.87	45.53	46.66
SB2-Croston SI New CF	46.90	55.07	56.93	55.29	42.88	49.92	51.09	47.26
SB2-SB2 SI New CF	47.29	55.96	58.67	58.18	43.25	50.77	53.36	52.60
SES-Poisson SES	-2.95	-1.50	-2.38	-4.34	-1.62	-0.42	-0.96	-2.66
SES-Croston SI New	28.66	35.86	36.69	34.69	58.76	80.55	87.05	88.29
SES-SB2 SI New	31.87	40.64	43.34	43.61	61.81	84.01	90.78	92.39
SES-Croston SI New CF	67.78	76.31	74.68	71.14	83.91	98.22	98.62	95.91
SES-SB2 SI New CF	68.23	76.65	75.26	72.36	83.64	98.36	99.85	98.26
Poisson SES-Croston SI New	12.84	16.44	16.32	16.35	9.77	14.94	18.52	22.58
Poisson SES-SB2 SI New	13.97	18.53	19.04	19.85	10.19	15.60	19.36	23.60
Poisson SES-Croston SI New CF	27.02	31.91	29.03	28.02	14.26	18.37	20.70	23.99
Poisson SES-SB2 SI New CF	27.65	33.07	30.62	30.17	14.55	18.84	21.37	24.85
Croston SI New-SB2 SI New	61.60	64.37	64.81	65.50	45.22	46.50	48.08	49.80
Croston SI New-Croston SI New CF	56.86	54.98	51.19	50.35	40.01	33.48	26.18	20.10
Croston SI New-SB2 SI New CF	57.89	57.31	55.23	56.13	40.81	36.01	31.96	30.27
SB2 SI New-Croston SI New CF	55.64	51.97	45.44	41.03	39.05	30.18	18.13	5.43
SB2 SI New-SB2 SI New CF	56.89	54.98	51.18	50.33	40.02	33.48	26.16	20.09
Croston SI New CF-SB2 SI New CF	57.80	57.87	58.57	59.60	44.35	46.12	48.56	50.50

Table A.2.2 Paired-sample *t*-test values for *MAE* differences, item B

Item B, Absolute *MAE* Comparison, *t* value								
	All Points In Time				Issue Points Only			
Methods Compared Method 1- Method 2	α values				α values			
	0.05	**0.1**	**0.15**	**0.2**	**0.05**	**0.1**	**0.15**	**0.2**
Croston-SB2	76.31	84.56	88.68	89.94	60.15	65.18	67.87	69.12
Croston-SES	29.48	32.31	32.95	32.52	-0.47	-25.17	-50.80	-67.56
Croston-Poisson SES	15.58	19.06	19.12	17.63	-0.77	-6.88	-10.71	-15.56
Croston-Croston SI New	43.04	49.86	52.40	52.10	35.01	42.68	46.33	47.36
Croston-SB2 SI New	44.71	52.82	56.91	58.18	36.81	45.99	51.78	55.69
Croston-Croston SI New CF	62.28	71.51	73.47	71.75	51.67	59.65	61.37	59.75
Croston-SB2 SI New CF	62.29	72.76	75.62	74.96	52.17	60.80	63.69	64.18
SB2-SES	27.20	28.12	26.69	24.33	-3.79	-32.02	-59.79	-76.55
SB2-Poisson SES	13.98	15.95	14.47	11.74	-1.43	-7.99	-12.09	-17.33
SB2-Croston SI New	41.28	46.63	47.36	45.09	33.08	38.95	39.73	36.54
SB2-SB2 SI New	43.04	49.86	52.40	52.10	35.01	42.68	46.33	47.36
SB2-Croston SI New CF	61.59	70.12	70.95	67.83	50.97	58.27	58.22	53.12
SB2-SB2 SI New CF	61.60	71.51	73.47	71.75	51.52	59.65	61.37	59.77
SES-Poisson SES	-1.43	-1.75	-2.38	-3.83	-0.69	-2.08	-1.85	-1.84
SES-Croston SI New	23.22	29.41	30.35	28.24	50.49	75.08	85.17	88.67
SES-SB2 SI New	26.86	34.92	37.96	38.00	53.77	79.29	89.86	93.74
SES-Croston SI New CF	74.48	80.01	77.06	71.26	86.68	100.84	101.27	98.63
SES-SB2 SI New CF	74.79	80.60	77.92	72.99	87.27	101.25	102.82	101.39
Poisson SES-Croston SI New	9.50	13.41	13.91	14.03	8.69	15.21	17.47	21.54
Poisson SES-SB2 SI New	10.79	15.65	16.87	17.76	9.16	15.91	18.34	22.69
Poisson SES-Croston SI New CF	27.45	31.09	28.32	27.43	14.18	19.19	20.02	23.46
Poisson SES-SB2 SI New CF	27.67	32.33	30.00	29.60	14.37	19.68	20.70	24.41
Croston SI New-SB2 SI New	80.36	74.24	70.55	69.49	54.07	51.86	51.77	52.34
Croston SI New-Croston SI New CF	67.81	62.48	57.08	54.28	46.06	37.73	29.57	22.48
Croston SI New-SB2 SI New CF	67.87	65.13	61.16	59.93	46.59	40.41	35.13	31.96
SB2 SI New-Croston SI New CF	66.19	59.13	51.31	45.35	44.83	34.35	22.02	9.13
SB2 SI New-SB2 SI New CF	66.29	62.47	57.07	54.26	45.48	37.74	29.57	22.48
Croston SI New CF-SB2 SI New CF	17.01	64.52	62.03	62.33	19.66	50.05	51.41	52.93

Appendix A.3 *SMAE* Results

Table A.3.1 Paired-sample *t*-test values for *SMAE* differences, item A

Item A, *SMAE* Comparison, *t* value								
	All Points In Time				Issue Points Only			
Methods Compared Method 1- Method 2	α values				α values			
	0.05	**0.1**	**0.15**	**0.2**	**0.05**	**0.1**	**0.15**	**0.2**
Croston-SB2	29.97	30.68	31.16	31.34	15.36	16.38	16.88	17.16
Croston-SES	16.06	15.74	15.43	15.14	0.49	-9.07	-15.76	-18.22
Croston-Poisson SES	13.73	14.81	13.49	12.77	-0.98	-4.35	-4.85	-9.35
Croston-Croston SI New	19.73	20.73	20.70	20.19	13.36	14.91	15.61	15.75
Croston-SB2 SI New	20.13	21.37	21.65	21.48	13.58	15.24	16.09	16.52
Croston-Croston SI New CF	24.26	24.91	24.64	24.04	14.84	15.94	16.32	16.13
Croston-SB2 SI New CF	24.45	25.24	25.18	24.83	14.89	16.03	16.52	16.60
SB2-SES	15.52	14.77	13.95	13.11	-0.73	-11.12	-16.99	-18.81
SB2-Poisson SES	13.13	13.80	11.89	10.62	-1.35	-4.96	-5.69	-10.07
SB2-Croston SI New	19.31	20.04	19.66	18.71	13.11	14.48	14.79	14.04
SB2-SB2 SI New	19.73	20.73	20.71	20.19	13.37	14.92	15.61	15.75
SB2-Croston SI New CF	24.07	24.57	24.06	23.17	14.79	15.83	16.01	15.23
SB2-SB2 SI New CF	24.26	24.91	24.65	24.05	14.84	15.94	16.32	16.13
SES-Poisson SES	-1.31	-0.26	-2.04	-2.54	-1.19	-1.84	0.24	-3.30
SES-Croston SI New	14.72	23.47	26.46	25.63	16.91	18.72	19.38	19.62
SES-SB2 SI New	17.76	28.93	34.72	36.83	17.24	18.87	19.45	19.67
SES-Croston SI New CF	54.62	71.46	76.58	76.85	18.92	19.47	19.73	19.82
SES-SB2 SI New CF	54.90	72.42	79.09	81.41	18.89	19.43	19.69	19.79
Poisson SES-Croston SI New	7.65	11.34	13.08	12.91	5.95	8.70	8.58	11.45
Poisson SES-SB2 SI New	8.81	13.45	15.72	16.54	6.18	9.00	9.05	11.85
Poisson SES-Croston SI New CF	22.79	29.75	26.56	26.86	8.19	10.01	9.62	11.89
Poisson SES-SB2 SI New CF	23.39	30.91	28.00	29.12	8.34	10.21	9.99	12.23
Croston SI New-SB2 SI New	41.23	46.00	49.72	53.45	15.00	15.59	16.03	16.33
Croston SI New-Croston SI New CF	32.39	30.24	28.98	28.90	13.75	12.51	10.74	8.58
Croston SI New-SB2 SI New CF	33.01	31.64	31.50	32.78	13.94	13.26	12.79	12.46
SB2 SI New-Croston SI New CF	31.71	28.66	26.02	24.10	13.51	11.39	6.97	0.87
SB2 SI New-SB2 SI New CF	32.39	30.24	28.98	28.90	13.75	12.51	10.73	8.57
Croston SI New CF-SB2 SI New CF	43.77	55.46	63.75	69.94	14.90	15.59	15.97	16.31

Table A.3.2 Paired-sample *t*-test values for *SMAE* differences, item B

Item B, *SMAE* Comparison, *t* value								
	All Points In Time				Issue Points Only			
Methods Compared Method 1- Method 2	α values				α values			
	0.05	**0.1**	**0.15**	**0.2**	**0.05**	**0.1**	**0.15**	**0.2**
Croston-SB2	35.92	36.51	36.99	37.38	15.80	16.27	16.47	16.58
Croston-SES	17.45	17.87	18.02	18.08	-1.47	-11.02	-16.29	-18.01
Croston-Poisson SES	13.40	15.27	15.11	14.61	-1.16	-3.95	-5.07	-7.78
Croston-Croston SI New	22.23	23.41	23.69	23.81	13.28	14.89	15.53	15.54
Croston-SB2 SI New	22.86	24.36	25.02	25.57	13.63	15.36	16.18	16.50
Croston-Croston SI New CF	28.58	29.09	29.01	28.80	15.26	16.36	16.56	16.35
Croston-SB2 SI New CF	28.58	29.53	29.69	29.77	15.33	16.44	16.72	16.69
SB2-SES	16.65	16.56	16.10	15.48	-3.17	-13.03	-17.23	-18.42
SB2-Poisson SES	12.59	13.84	13.04	12.01	-1.55	-4.54	-5.88	-8.68
SB2-Croston SI New	21.55	22.38	22.20	21.75	12.86	14.21	14.34	13.24
SB2-SB2 SI New	22.23	23.41	23.70	23.81	13.28	14.89	15.53	15.54
SB2-Croston SI New CF	28.31	28.62	28.27	27.72	15.18	16.25	16.28	15.59
SB2-SB2 SI New CF	28.31	29.09	29.02	28.81	15.26	16.36	16.57	16.36
SES-Poisson SES	-0.98	-1.69	-1.73	-1.35	-0.81	-1.58	-0.55	-1.35
SES-Croston SI New	11.18	17.31	19.39	18.27	16.51	18.21	18.74	18.95
SES-SB2 SI New	14.36	22.85	27.65	29.19	16.81	18.35	18.82	19.00
SES-Croston SI New CF	55.43	67.98	74.78	76.81	18.04	18.77	19.00	19.06
SES-SB2 SI New CF	55.29	69.07	77.29	81.54	18.06	18.72	18.95	19.02
Poisson SES-Croston SI New	5.66	9.40	10.10	9.68	5.37	7.54	8.36	10.26
Poisson SES-SB2 SI New	6.87	11.47	12.94	13.81	5.63	7.87	8.84	10.79
Poisson SES-Croston SI New CF	23.33	26.54	25.64	27.46	7.96	9.08	9.51	10.91
Poisson SES-SB2 SI New CF	23.55	27.57	27.18	30.12	8.06	9.29	9.86	11.33
Croston SI New-SB2 SI New	44.07	46.71	49.24	50.69	15.10	14.90	14.98	15.13
Croston SI New-Croston SI New CF	33.71	31.33	29.89	28.76	14.51	13.42	11.50	9.34
Croston SI New-SB2 SI New CF	33.68	32.66	32.21	32.17	14.63	13.88	12.89	12.16
SB2 SI New-Croston SI New CF	33.03	29.83	27.17	24.55	14.35	12.68	8.68	2.61
SB2 SI New-SB2 SI New CF	33.01	31.33	29.89	28.76	14.50	13.42	11.50	9.34
Croston SI New CF-SB2 SI New CF	16.46	57.57	63.62	67.38	9.64	14.67	14.96	15.33

Appendix A.4 *GRMSE* Results

Table A.4.1 Paired-sample *t*-test values for *GRMSE* differences, item A

Item A, *GRMSE* Comparison, *t* value								
	All Points In Time				Issue Points Only			
Methods Compared Method 1- Method 2	α values				α values			
	0.05	**0.1**	**0.15**	**0.2**	**0.05**	**0.1**	**0.15**	**0.2**
Croston-SB2	150.39	215.03	228.72	232.42	102.99	145.43	152.34	153.80
Croston-SES	15.70	9.17	8.74	9.68	1.90	-18.41	-44.65	-62.71
Croston-Poisson SES	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Croston-Croston SI New	21.32	8.67	7.62	7.61	21.94	24.71	27.87	30.00
Croston-SB2 SI New	21.87	8.98	7.99	8.08	22.58	25.56	29.33	32.43
Croston-Croston SI New CF	31.36	13.55	11.96	11.93	32.72	36.77	40.26	42.32
Croston-SB2 SI New CF	31.70	13.92	12.46	12.56	33.16	37.45	41.59	44.74
SB2-SES	15.25	9.03	8.63	9.57	-0.13	-23.53	-53.45	-73.74
SB2-Poisson SES	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
SB2-Croston SI New	20.79	8.53	7.51	7.50	21.30	23.84	26.30	27.22
SB2-SB2 SI New	21.36	8.85	7.89	7.97	21.97	24.75	27.93	30.06
SB2-Croston SI New CF	31.06	13.47	11.89	11.86	32.30	36.10	38.83	39.52
SB2-SB2 SI New CF	31.41	13.84	12.39	12.50	32.76	36.84	40.36	42.44
SES-Poisson SES	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
SES-Croston SI New	16.65	17.47	8.84	-0.95	29.61	27.11	26.87	27.13
SES-SB2 SI New	18.46	20.66	13.00	4.25	30.40	27.65	27.38	27.65
SES-Croston SI New CF	56.44	54.93	49.85	40.88	40.13	31.09	28.91	28.12
SES-SB2 SI New CF	58.10	56.95	52.93	45.56	40.63	31.45	29.30	28.56
Poisson SES-Croston SI New	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Poisson SES-SB2 SI New	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Poisson SES-Croston SI New CF	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Poisson SES-SB2 SI New CF	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Croston SI New-SB2 SI New	222.27	231.27	213.19	189.51	152.32	163.90	166.23	163.57
Croston SI New-Croston SI New CF	49.22	29.94	22.78	18.82	42.41	27.52	19.93	15.94
Croston SI New-SB2 SI New CF	51.13	31.60	24.53	21.00	44.53	30.44	24.64	23.44
SB2 SI New-Croston SI New CF	47.28	28.44	21.26	17.27	40.23	24.40	14.73	7.35
SB2 SI New-SB2 SI New CF	49.26	30.19	23.12	19.57	42.46	27.55	19.96	15.97
Croston SI New CF-SB2 SI New CF	254.56	231.69	178.85	153.53	157.12	178.04	174.61	166.81

Table A.4.2 Paired-sample *t*-test values for *GRMSE* differences, item B

Item B, *GRMSE* Comparison, *t* value								
	All Points In Time				Issue Points Only			
Methods Compared Method 1- Method 2	α values				α values			
	0.05	**0.1**	**0.15**	**0.2**	**0.05**	**0.1**	**0.15**	**0.2**
Croston-SB2	250.58	260.89	262.60	259.91	142.27	147.99	149.39	148.42
Croston-SES	17.38	10.82	9.78	10.24	0.23	-20.96	-45.66	-63.13
Croston-Poisson SES	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Croston-Croston SI New	26.65	12.44	8.83	8.09	23.15	24.06	27.60	30.10
Croston-SB2 SI New	27.53	12.85	9.20	8.57	24.05	25.12	29.33	32.81
Croston-Croston SI New CF	38.44	17.09	12.89	12.49	35.99	38.60	41.76	44.12
Croston-SB2 SI New CF	38.46	17.46	13.37	13.09	36.95	39.38	43.18	46.56
SB2-SES	16.73	10.60	9.65	10.12	-2.51	-27.33	-55.36	-74.87
SB2-Poisson SES	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
SB2-Croston SI New	25.76	12.16	8.67	7.95	22.23	22.95	25.72	26.98
SB2-SB2 SI New	26.67	12.58	9.04	8.44	23.15	24.08	27.65	30.15
SB2-Croston SI New CF	38.00	16.96	12.81	12.41	35.45	37.82	40.26	41.29
SB2-SB2 SI New CF	38.03	17.34	13.29	13.02	36.42	38.67	41.88	44.25
SES-Poisson SES	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
SES-Croston SI New	13.63	14.11	7.14	-2.93	27.26	26.40	26.74	26.38
SES-SB2 SI New	15.78	17.77	12.32	3.96	28.21	27.10	27.40	27.03
SES-Croston SI New CF	64.45	65.61	62.61	53.41	39.88	29.31	27.98	28.13
SES-SB2 SI New CF	64.56	68.24	66.69	59.40	41.17	29.72	28.43	28.65
Poisson SES-Croston SI New	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Poisson SES-SB2 SI New	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Poisson SES-Croston SI New CF	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Poisson SES-SB2 SI New CF	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Croston SI New-SB2 SI New	246.32	282.38	245.75	207.90	89.79	139.45	148.11	149.22
Croston SI New-Croston SI New CF	58.71	37.38	28.25	22.26	41.85	28.85	23.44	19.65
Croston SI New-SB2 SI New CF	58.90	39.24	30.11	24.41	44.25	31.34	27.71	26.66
SB2 SI New-Croston SI New CF	56.68	35.61	26.48	20.55	40.07	26.21	18.71	11.57
SB2 SI New-SB2 SI New CF	56.87	37.59	28.49	22.87	42.46	28.90	23.48	19.70
Croston SI New CF-SB2 SI New CF	6.46	282.88	211.01	168.33	12.57	162.11	168.33	154.05

Appendix B The Appendix for Chapter 7

Appendix B.1 Simulation Parameters of Croston SI New CF and SB2 SI New CF Methods

Table B.1.1 Simulation parameters of Croston SI New CF and SB2 SI New CF methods

Method #	Update Frequency (Inter-arrival time)	Week count (*MCF*)	*LL*	*UL*
1	**1**	**2**	**0.7**	**1.2**
2	**1**	**2**	**0.7**	**1.3**
3	**1**	**2**	**0.8**	**1.2**
4	**1**	**2**	**0.8**	**1.3**
5	**1**	**3**	**0.7**	**1.2**
6	**1**	**3**	**0.7**	**1.3**
7	**1**	**3**	**0.8**	**1.2**
8	**1**	**3**	**0.8**	**1.3**
9	2	2	0.7	1.2
10	2	2	0.7	1.3
11	2	2	0.8	1.2
12	2	2	0.8	1.3
13	2	3	0.7	1.2
14	2	3	0.7	1.3
15	2	3	0.8	1.2
16	2	3	0.8	1.3
17	**3**	**2**	**0.7**	**1.2**
18	**3**	**2**	**0.7**	**1.3**
19	**3**	**2**	**0.8**	**1.2**
20	**3**	**2**	**0.8**	**1.3**
21	**3**	**3**	**0.7**	**1.2**
22	**3**	**3**	**0.7**	**1.3**
23	**3**	**3**	**0.8**	**1.2**
24	**3**	**3**	**0.8**	**1.3**

Appendix B.2 The Fill Rate Confidence Interval Figures of Croston SI New CF and SB2 SI New CF Parameter Set Combinations

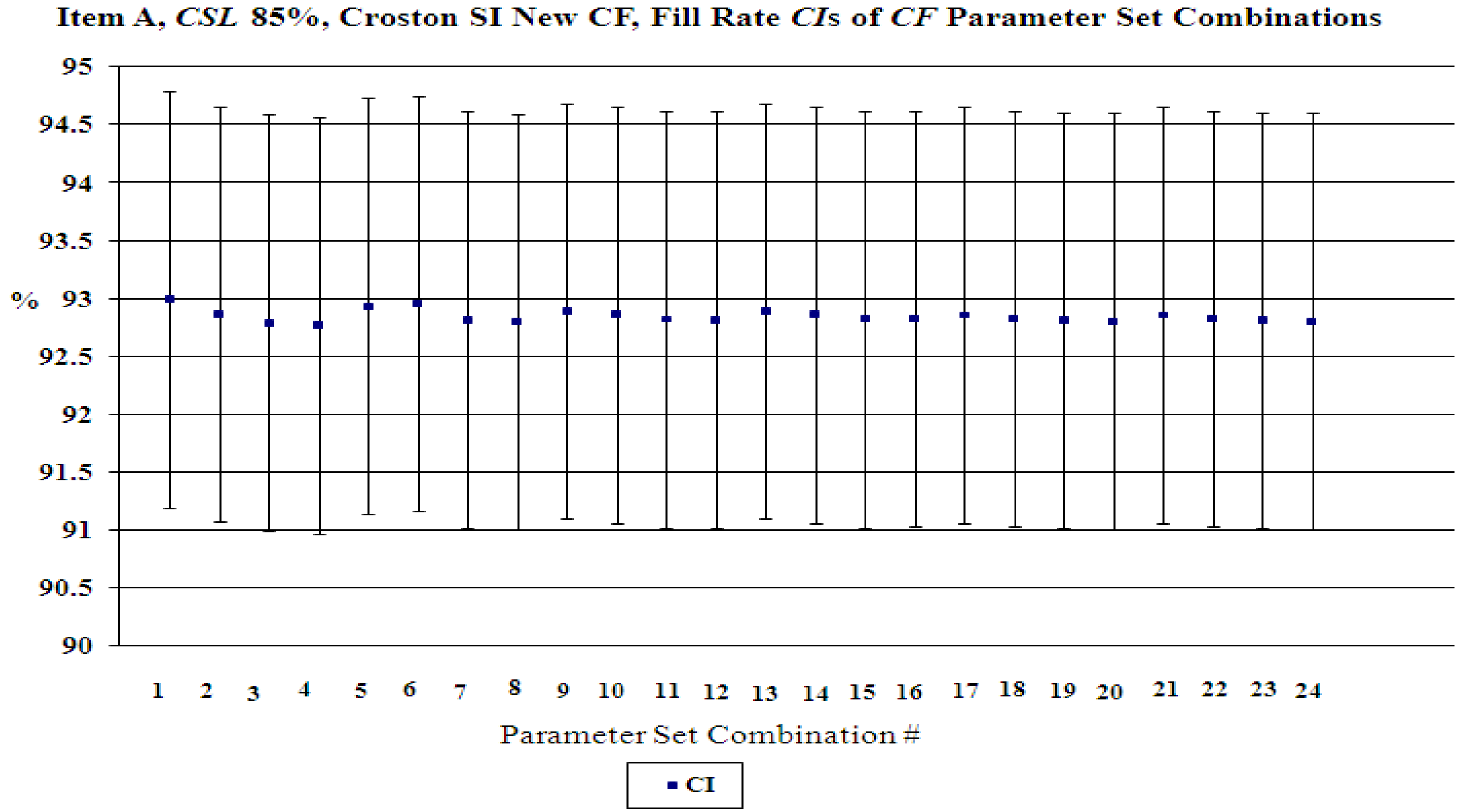

Figure B.2.1 The fill rate confidence intervals of Croston SI New CF parameter set combinations, Item A, target *CSL*=85%

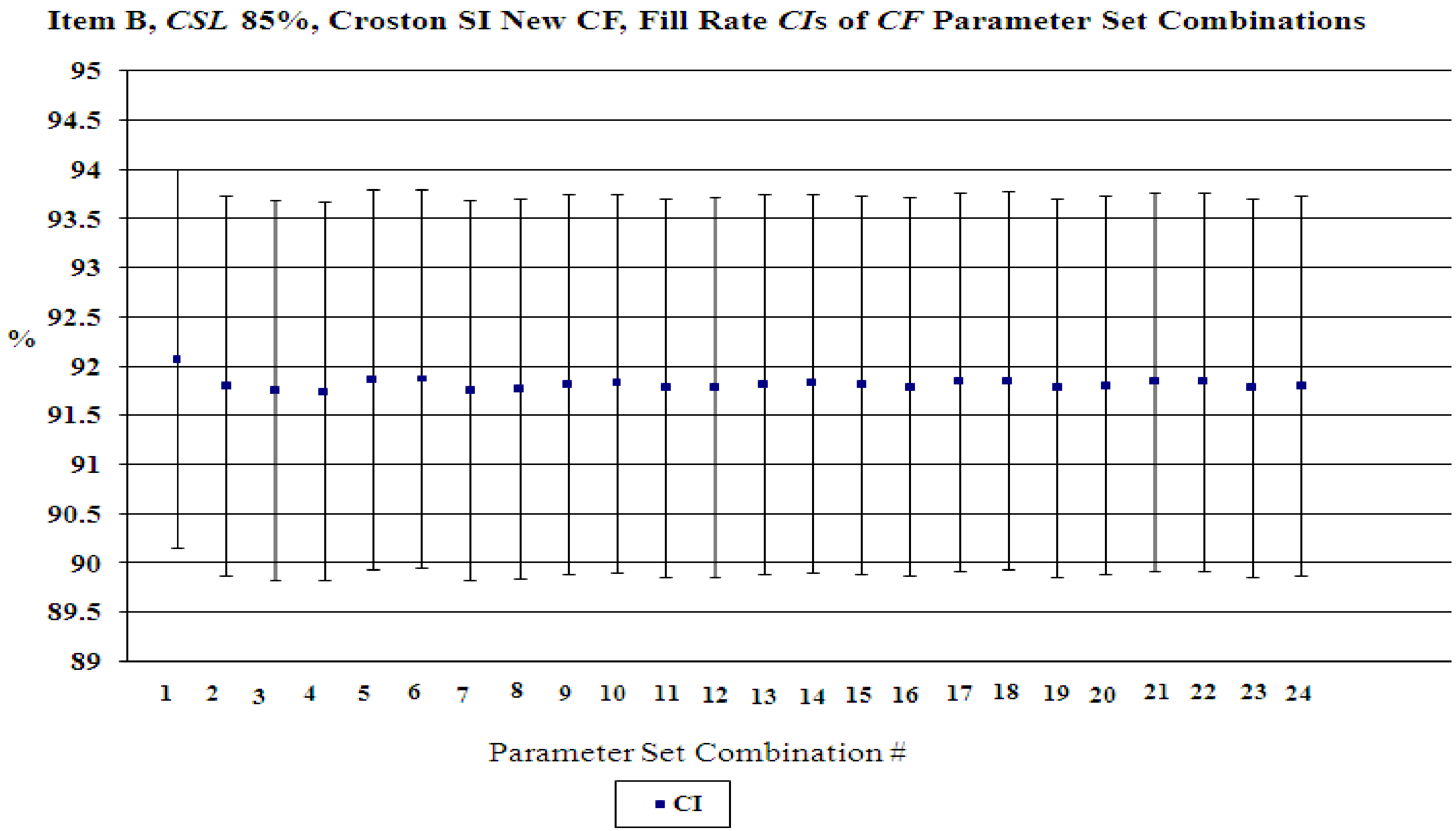

Figure B.2.2 The fill rate confidence intervals of Croston SI New CF parameter set combinations, Item B, target *CSL*=85%

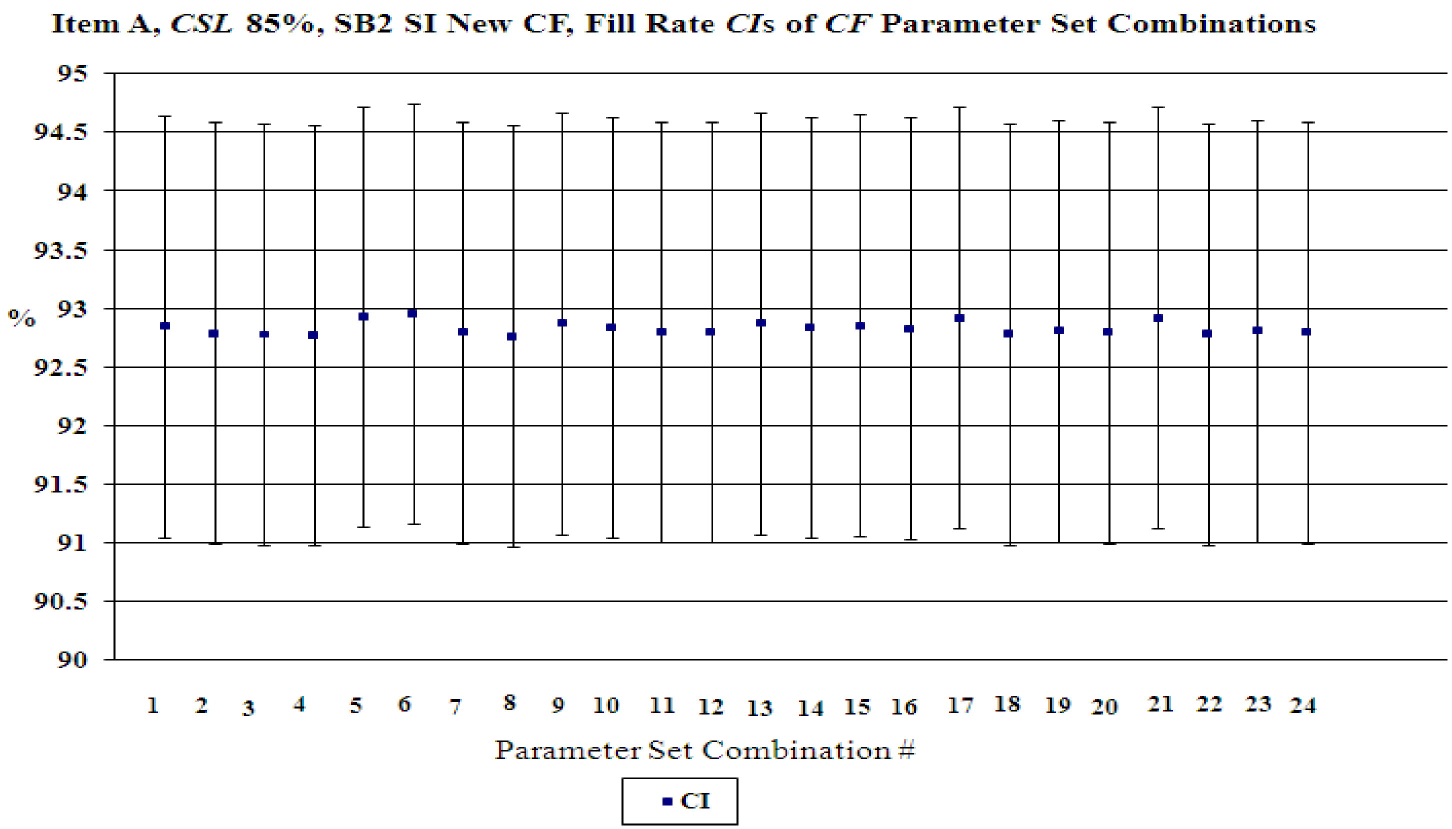

Figure B.2.3 The fill rate confidence intervals of SB2 SI New CF parameter set combinations, Item A, target *CSL*=85%

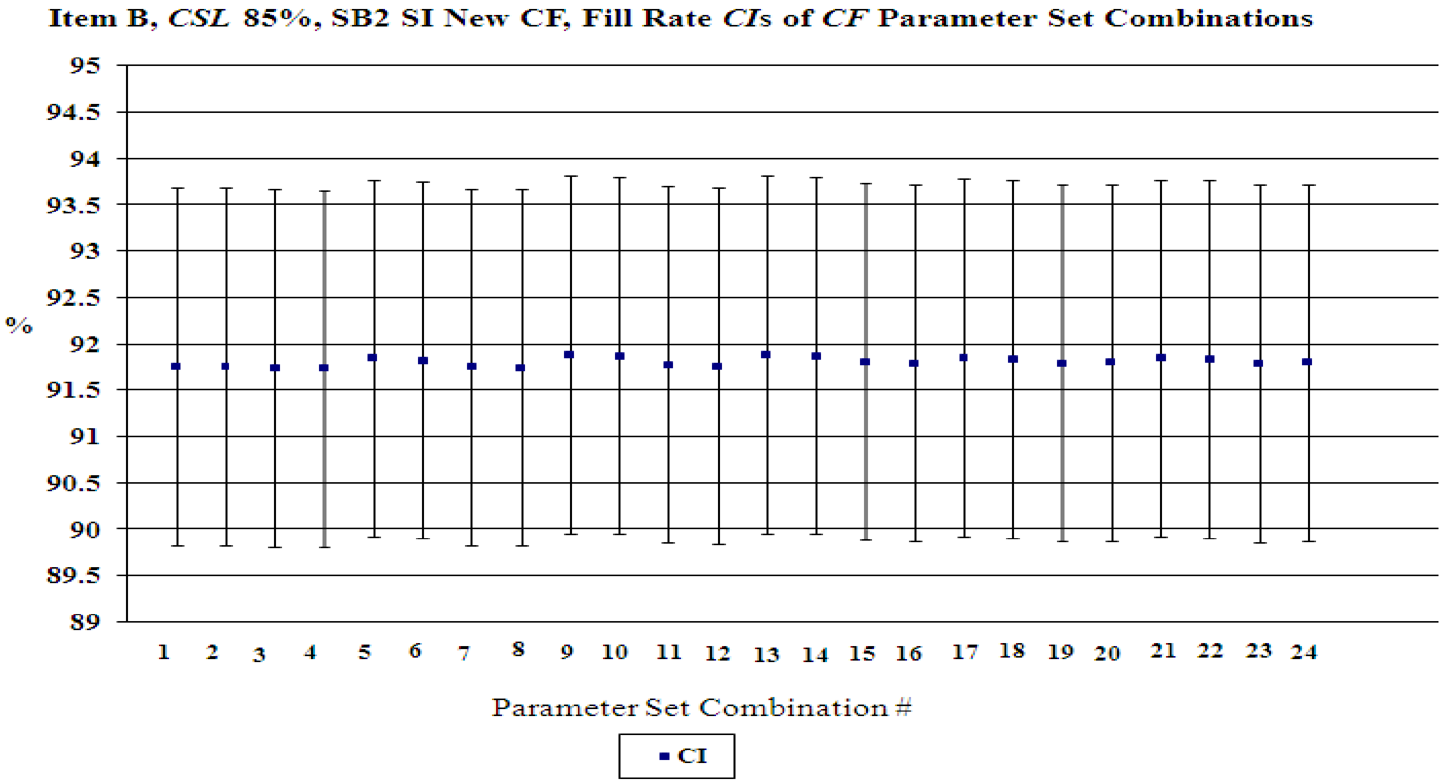

Figure B.2.4 The fill rate confidence intervals of SB2 SI New CF parameter set combinations, Item B, target *CSL*=85%

Appendix B.3 The Margin Confidence Interval Figures of Croston SI New CF and SB2 SI New CF Parameter Set Combinations

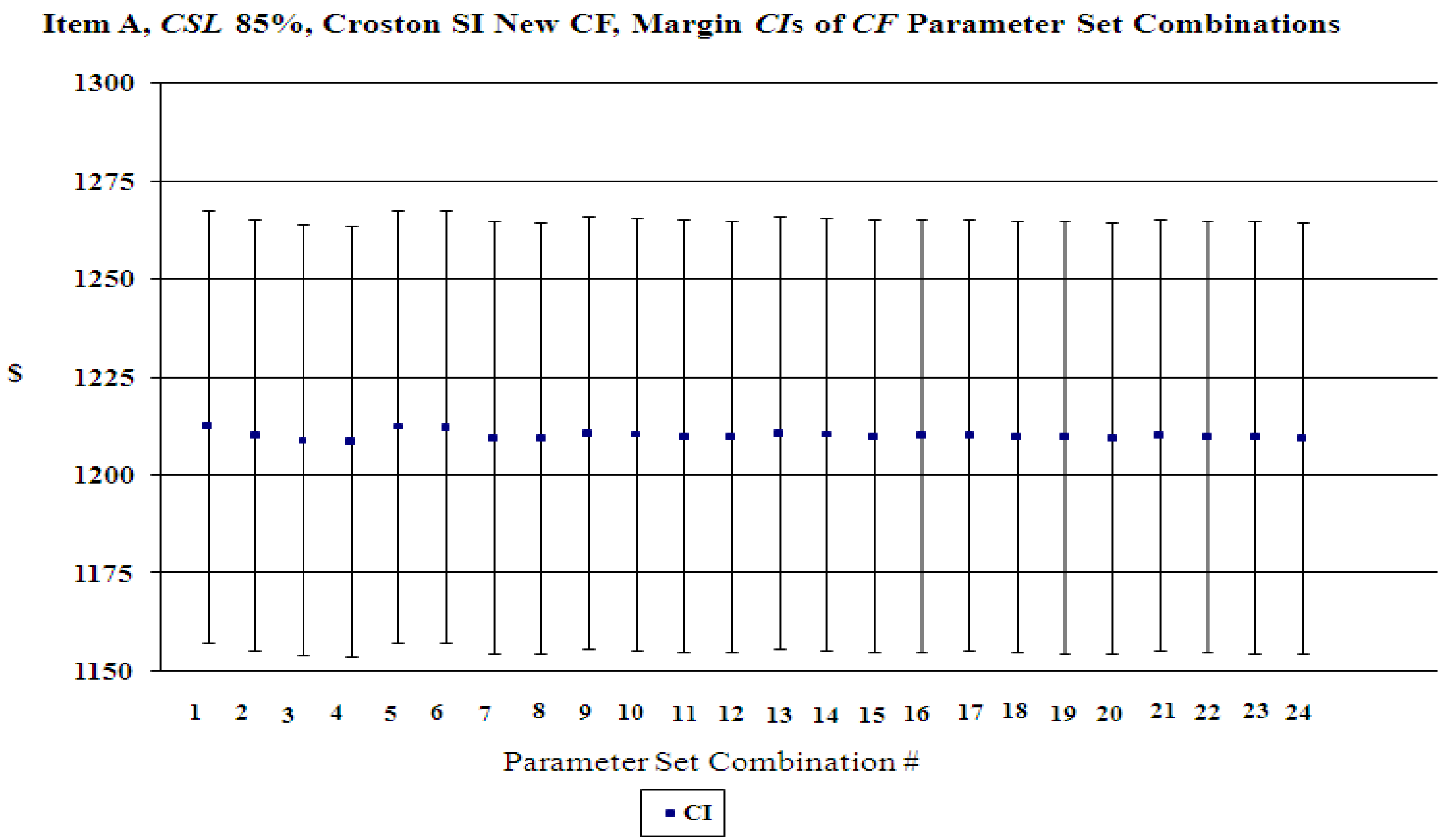

Figure B.3.1 The margin confidence intervals of Croston SI New CF parameter set combinations, Item A, target *CSL*=85%

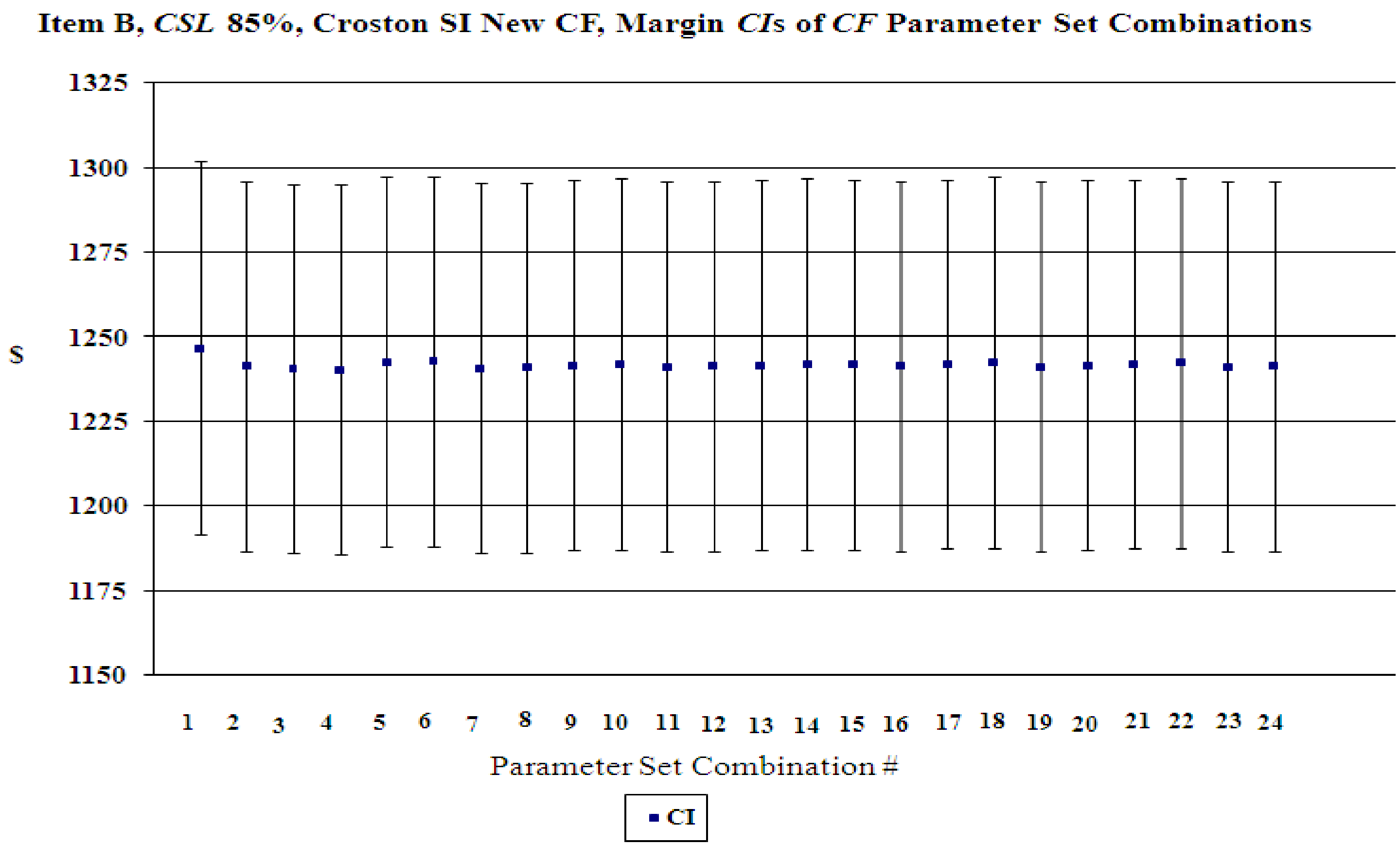

Figure B.3.2 The margin confidence intervals of Croston SI New CF parameter set combinations, Item B, target *CSL*=85%

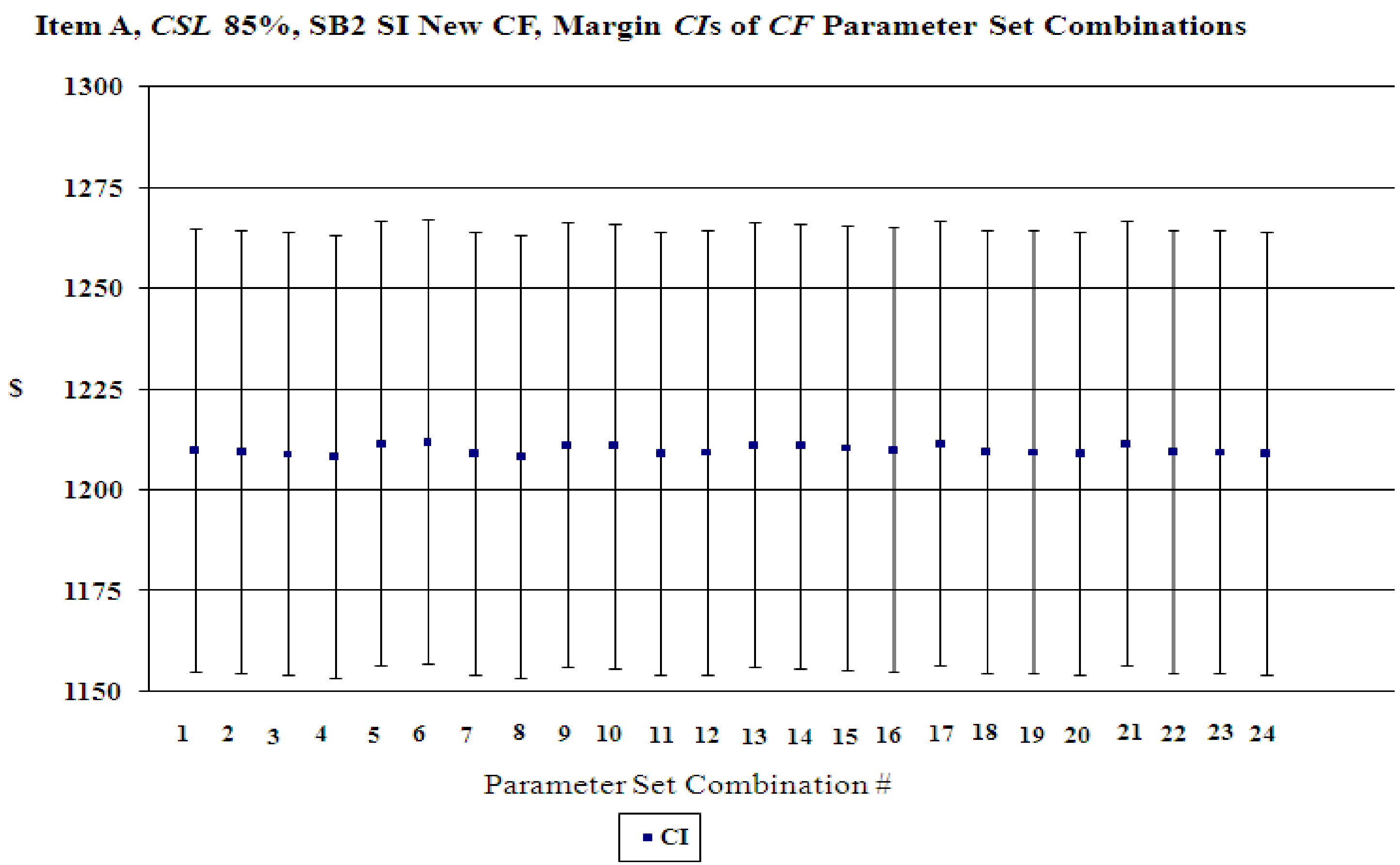

Figure B.3.3 The margin confidence intervals of SB2 SI New CF parameter set combinations, Item A, target *CSL*=85%

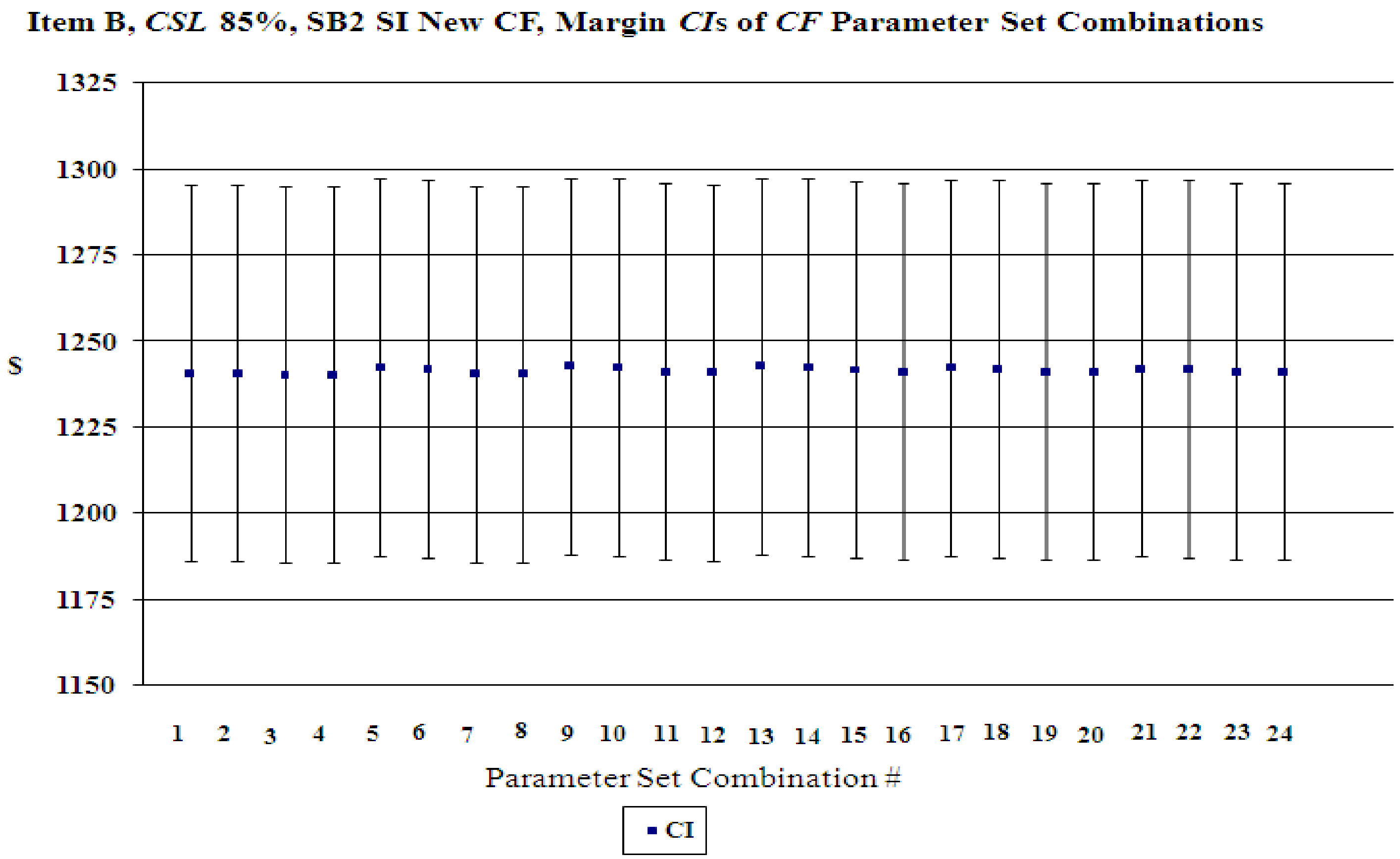

Figure B.3.4 The margin confidence intervals of SB2 SI New CF parameter set combinations, Item B, target *CSL*=85%

Appendix B.4 The Inventory Holding Cost Confidence Interval Figures of Croston SI New CF and SB2 SI New CF Parameter Set Combinations

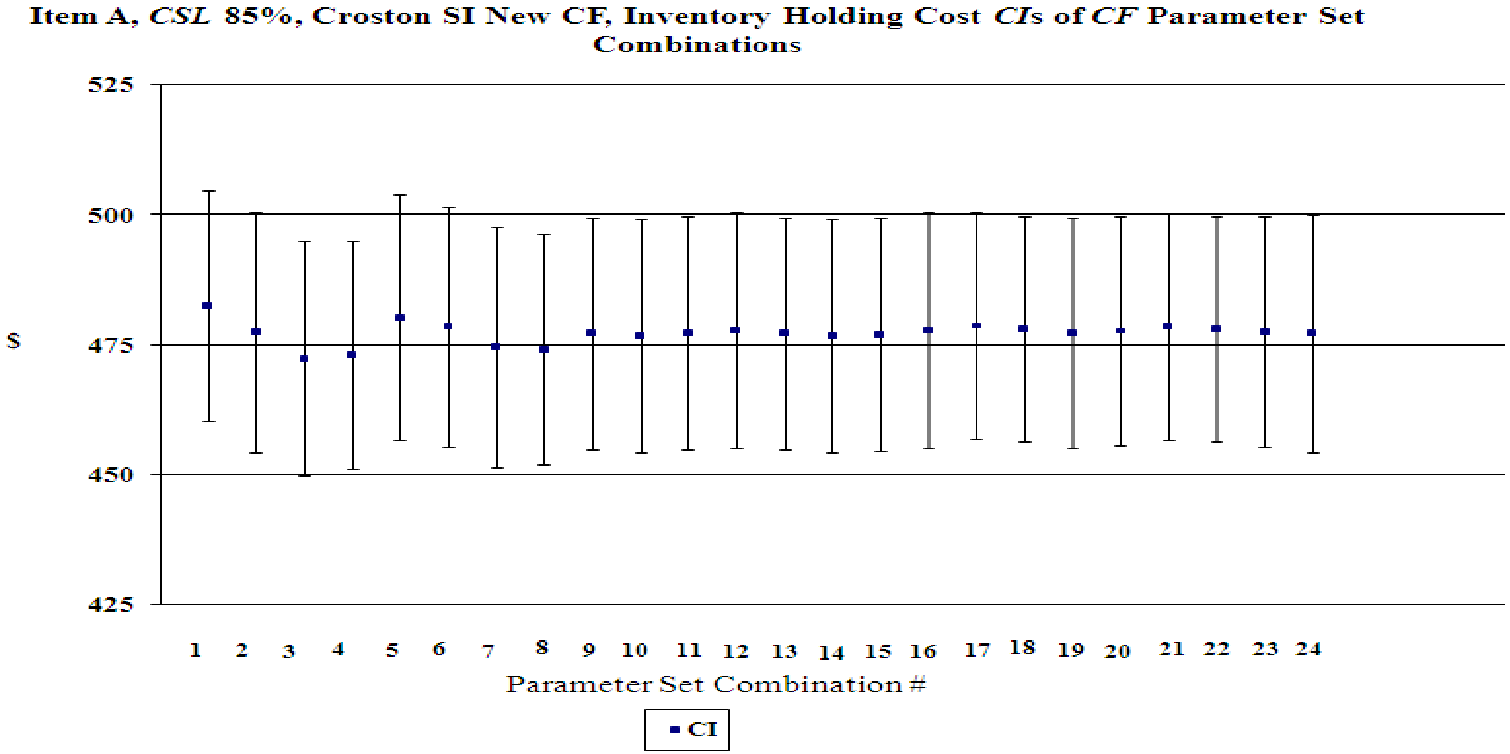

Figure B.4.1 The inventory holding cost confidence intervals of Croston SI New CF parameter set combinations, Item A, target *CSL*=85%

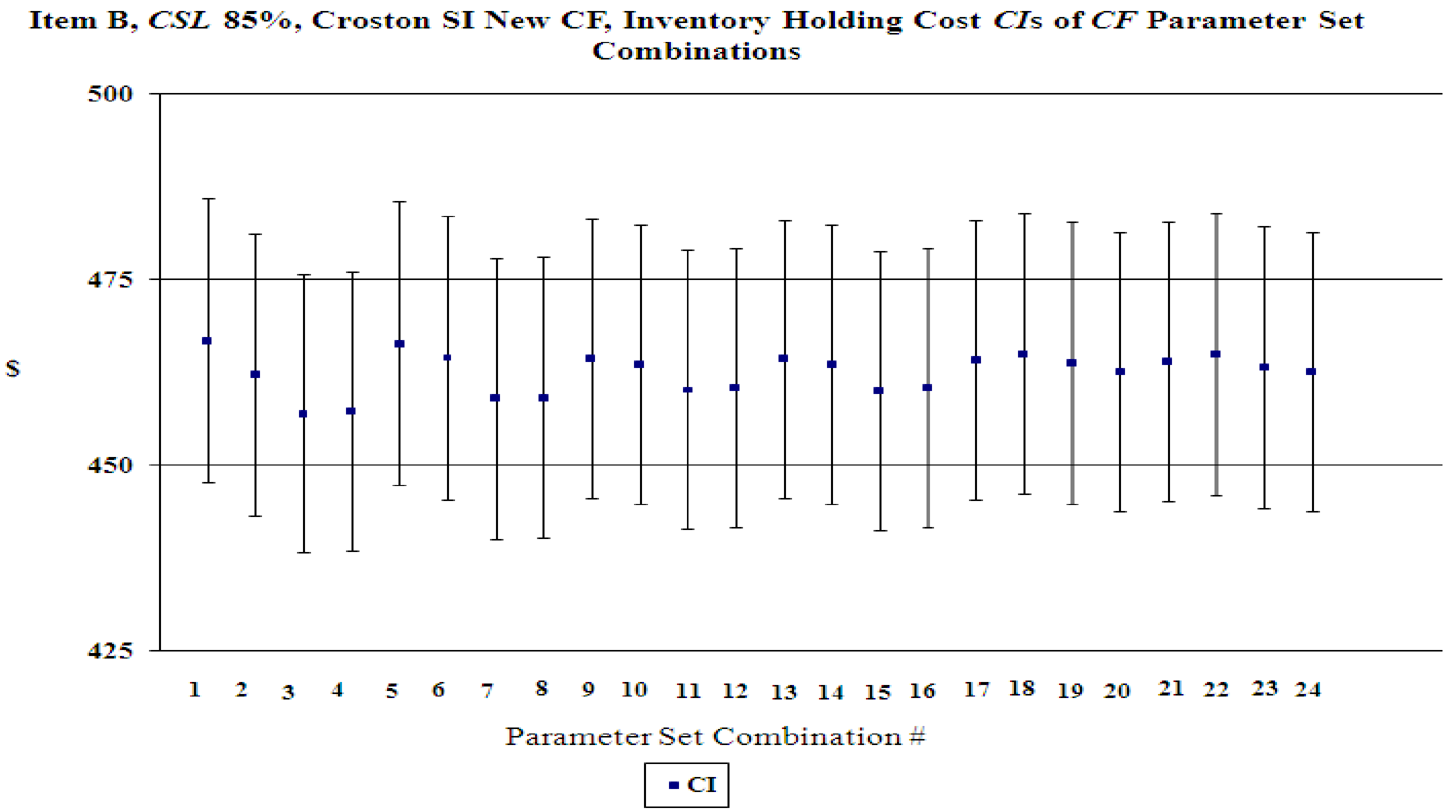

Figure B.4.2 The inventory holding cost confidence intervals of Croston SI New CF parameter set combinations, Item B, target *CSL*=85%

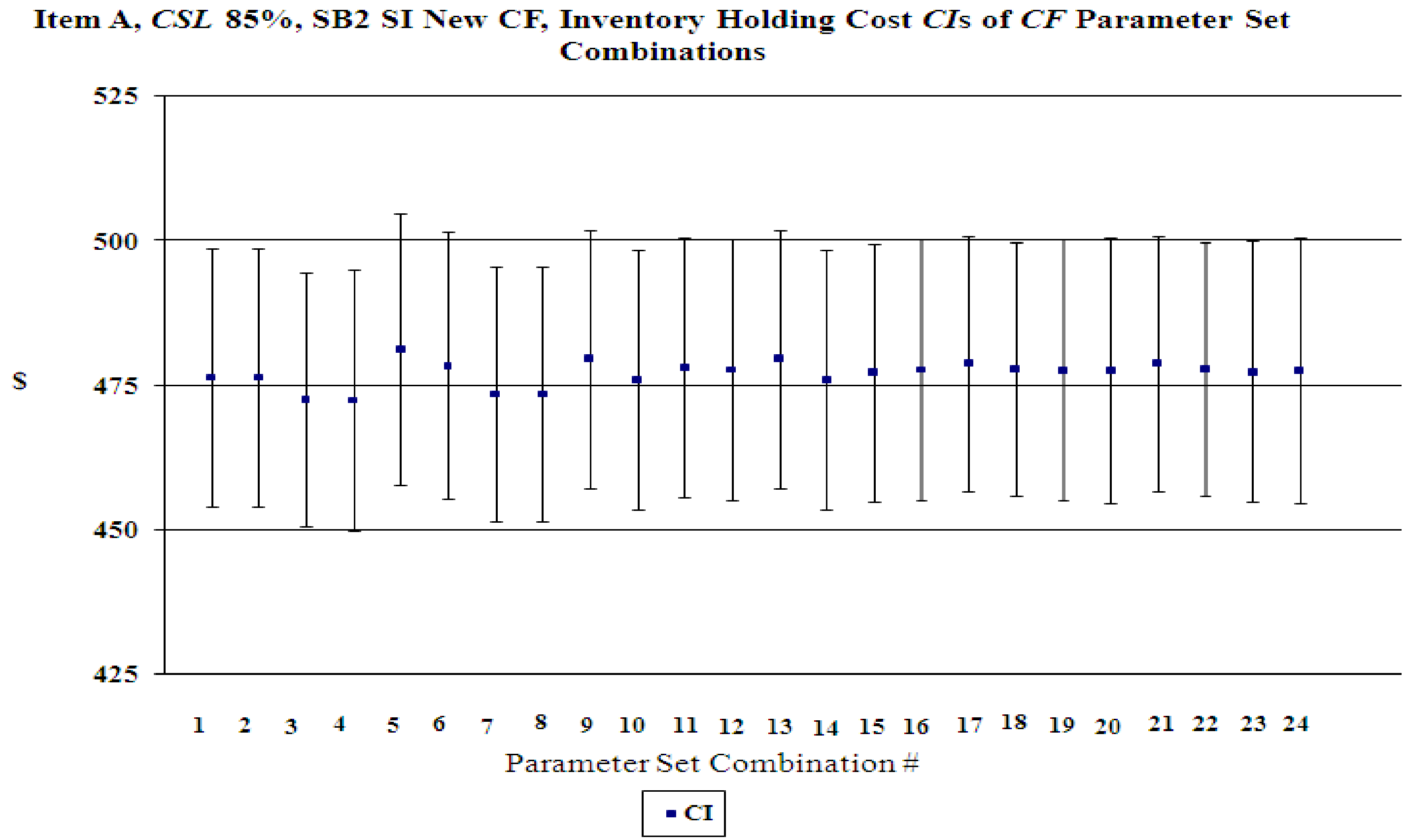

Figure B.4.3 The inventory holding cost confidence intervals of SB2 SI New CF parameter set combinations, Item A, target *CSL*=85%

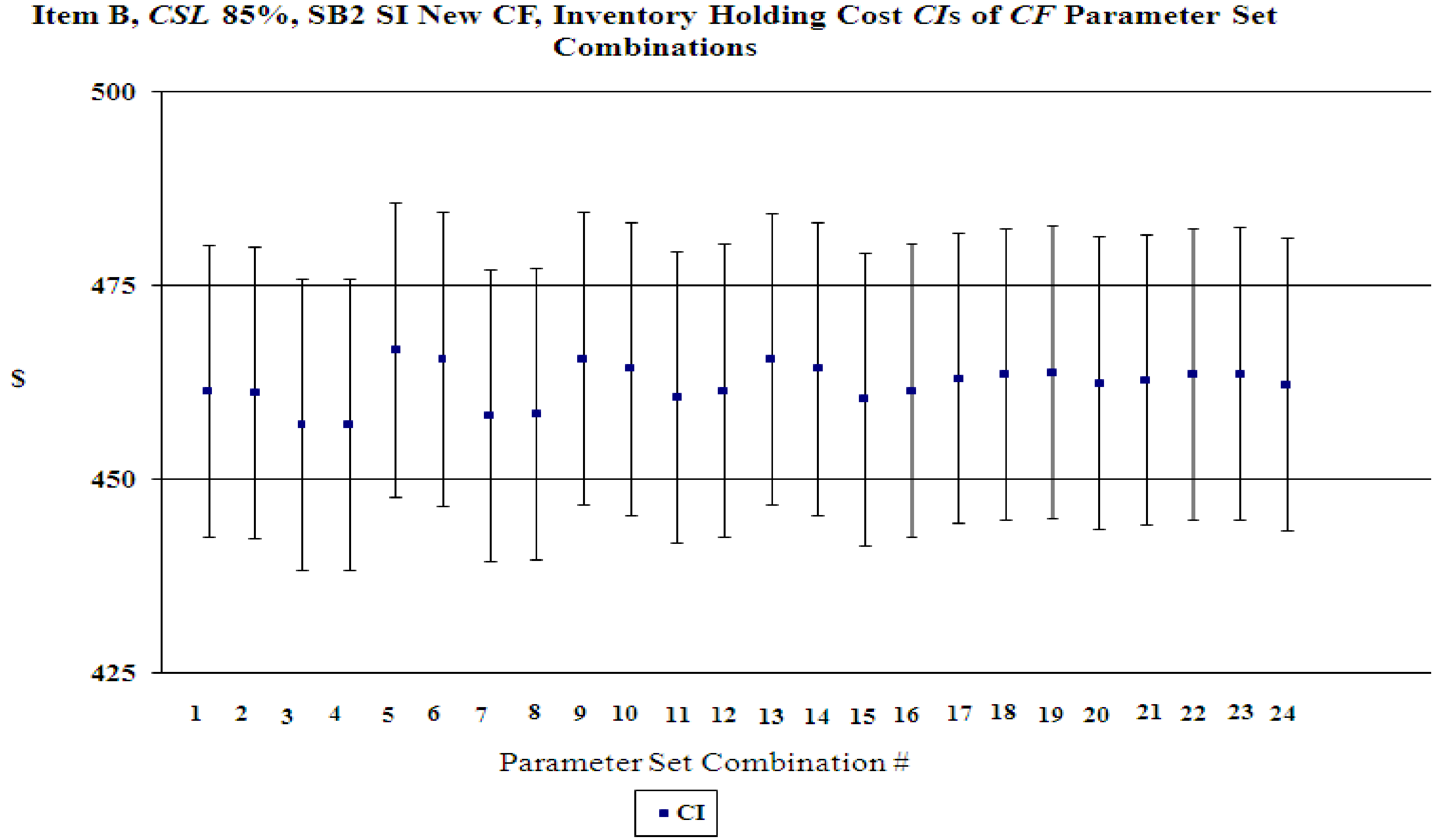

Figure B.4.4 The inventory holding cost confidence intervals of SB2 SI New CF parameter set combinations, Item B, target *CSL*=85%

Appendix B.5 The Adjusted Margin Confidence Interval Figures of Croston SI New CF and SB2 SI New CF Parameter Set Combinations

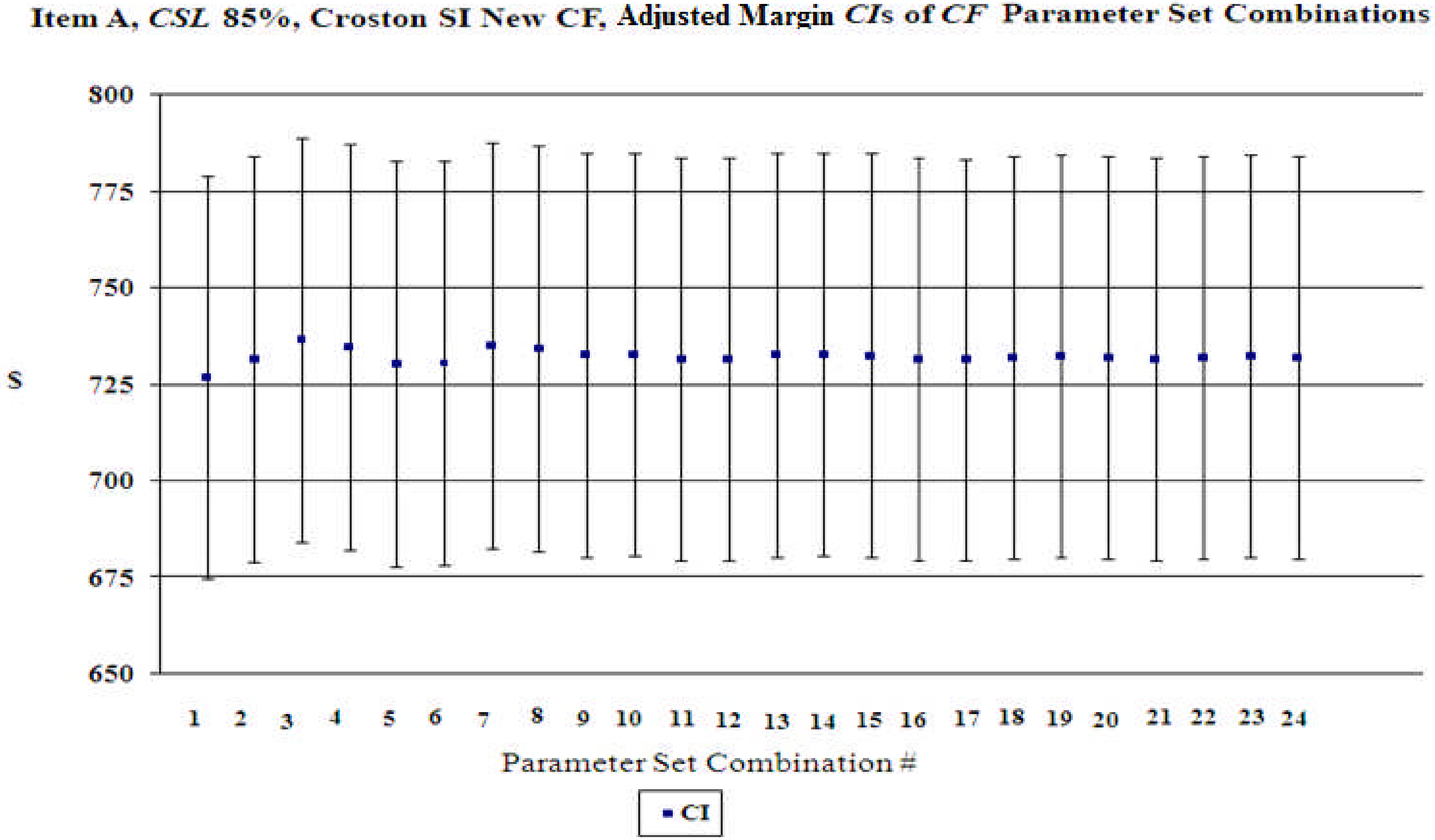

Figure B.5.1 The adjusted margin confidence intervals of Croston SI New CF parameter set combinations, Item A, target *CSL*=85%

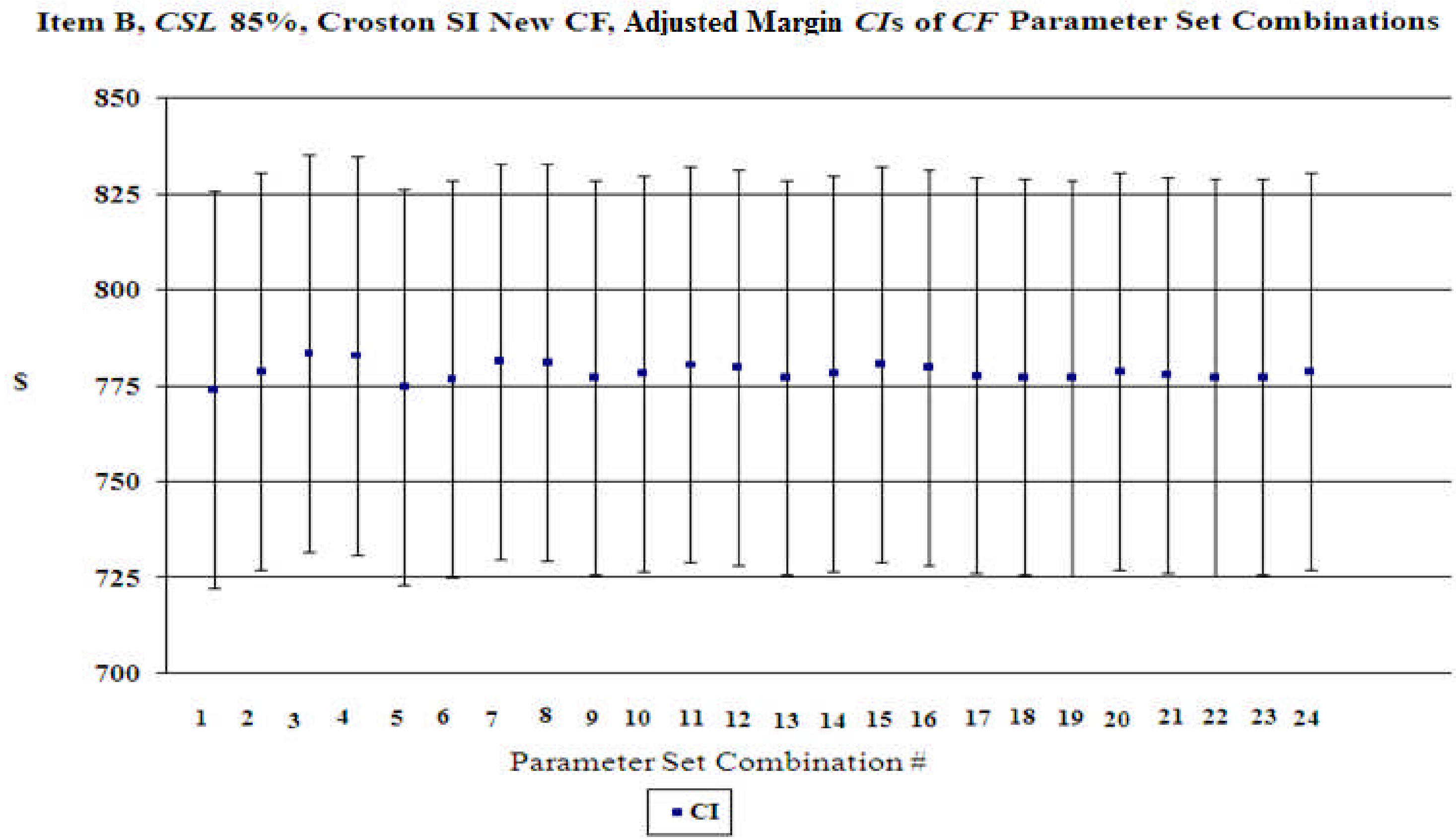

Figure B.5.2 The adjusted margin confidence intervals of Croston SI New CF parameter set combinations, Item B, target *CSL*=85%

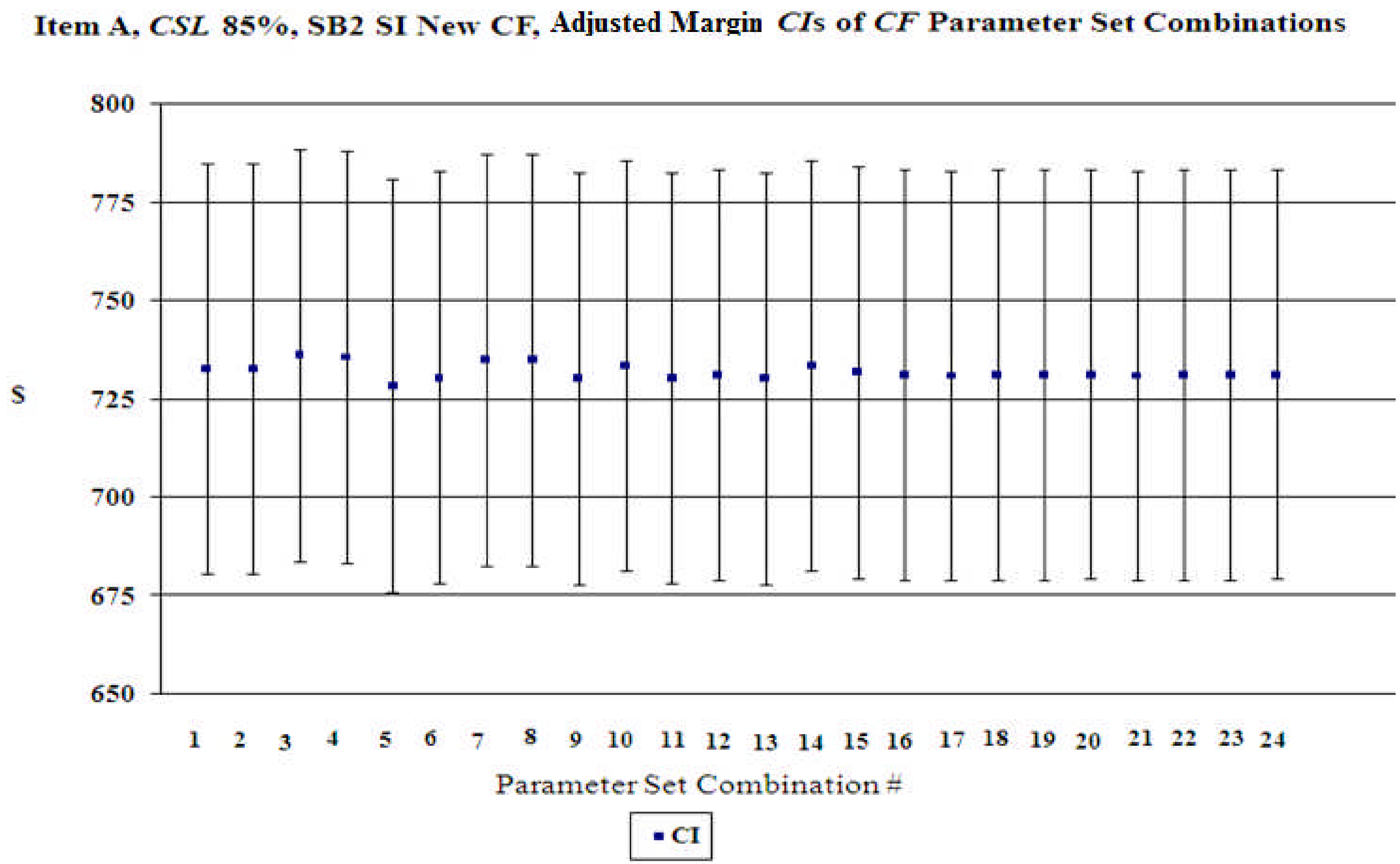

Figure B.5.3 The adjusted margin confidence intervals of SB2 SI New CF parameter set combinations, Item A, target *CSL*=85%

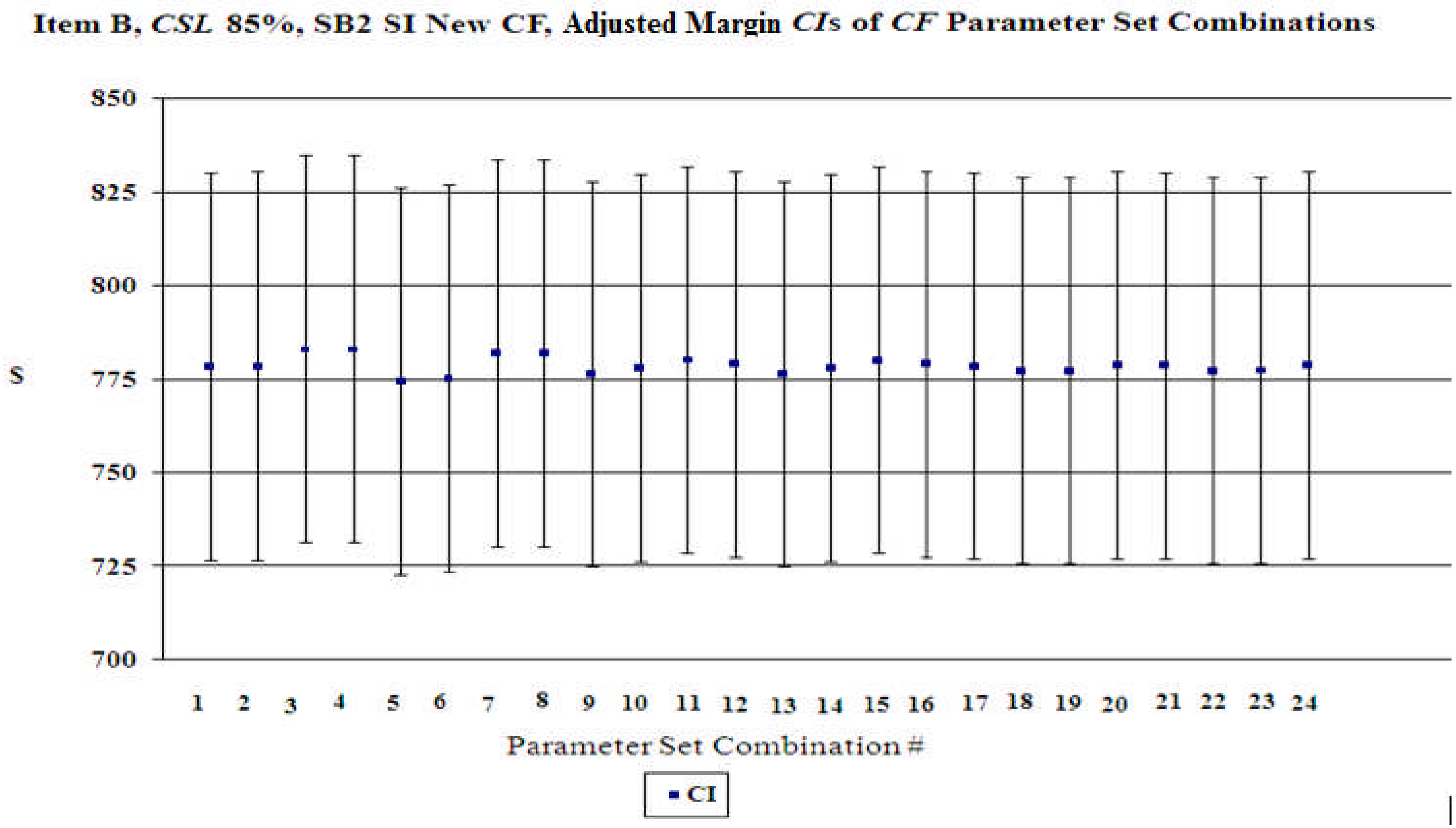

Figure B.5.4 The adjusted margin confidence intervals of SB2 SI New CF parameter set combinations, Item B, target *CSL*=85%

Appendix B.6 The *GMROI* Confidence Interval Figures of Croston SI New CF and SB2 SI New CF Parameter Set Combinations

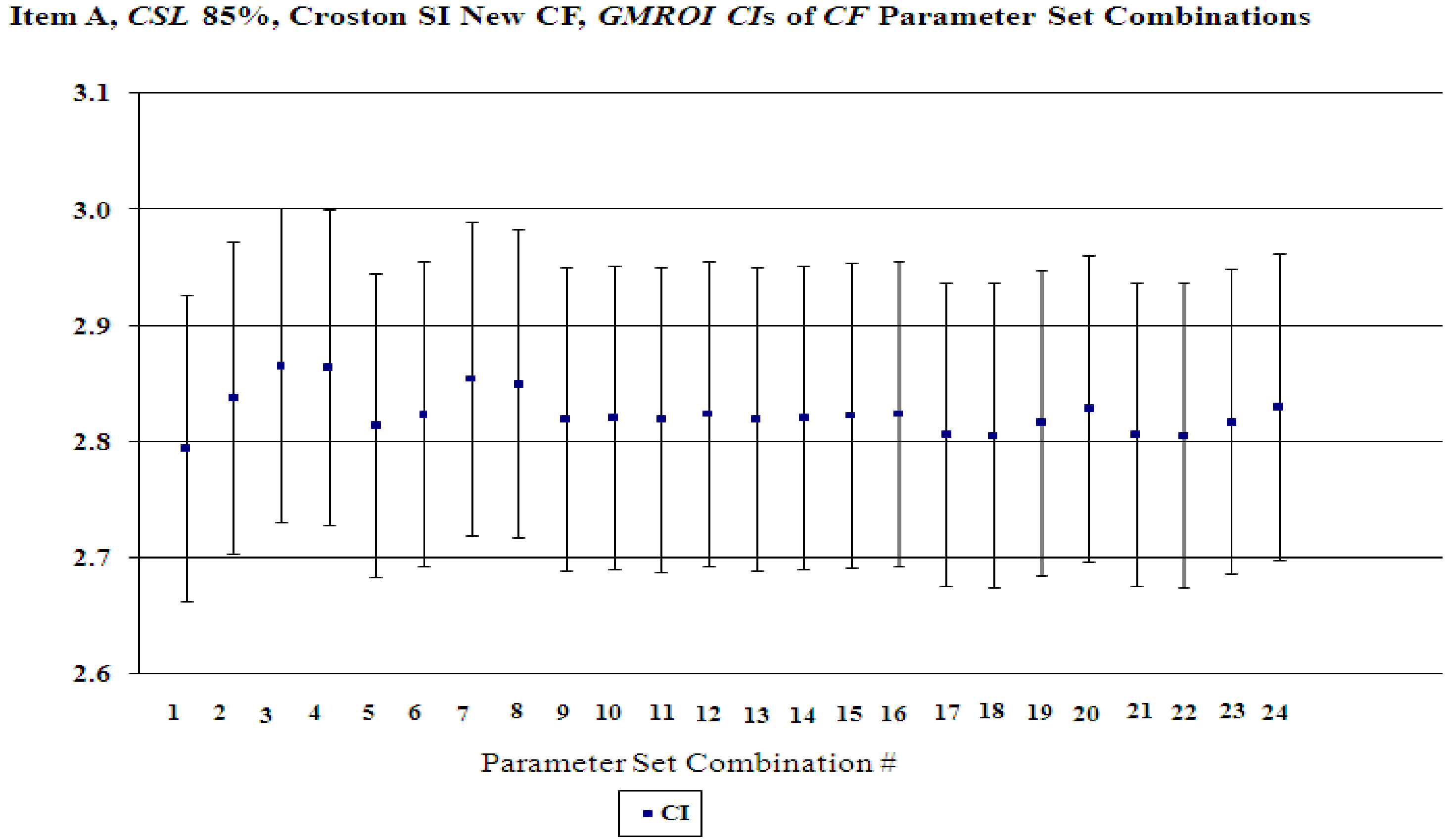

Figure B.6.1 The *GMROI* confidence intervals of Croston SI New CF parameter set combinations, Item A, target *CSL*=85%

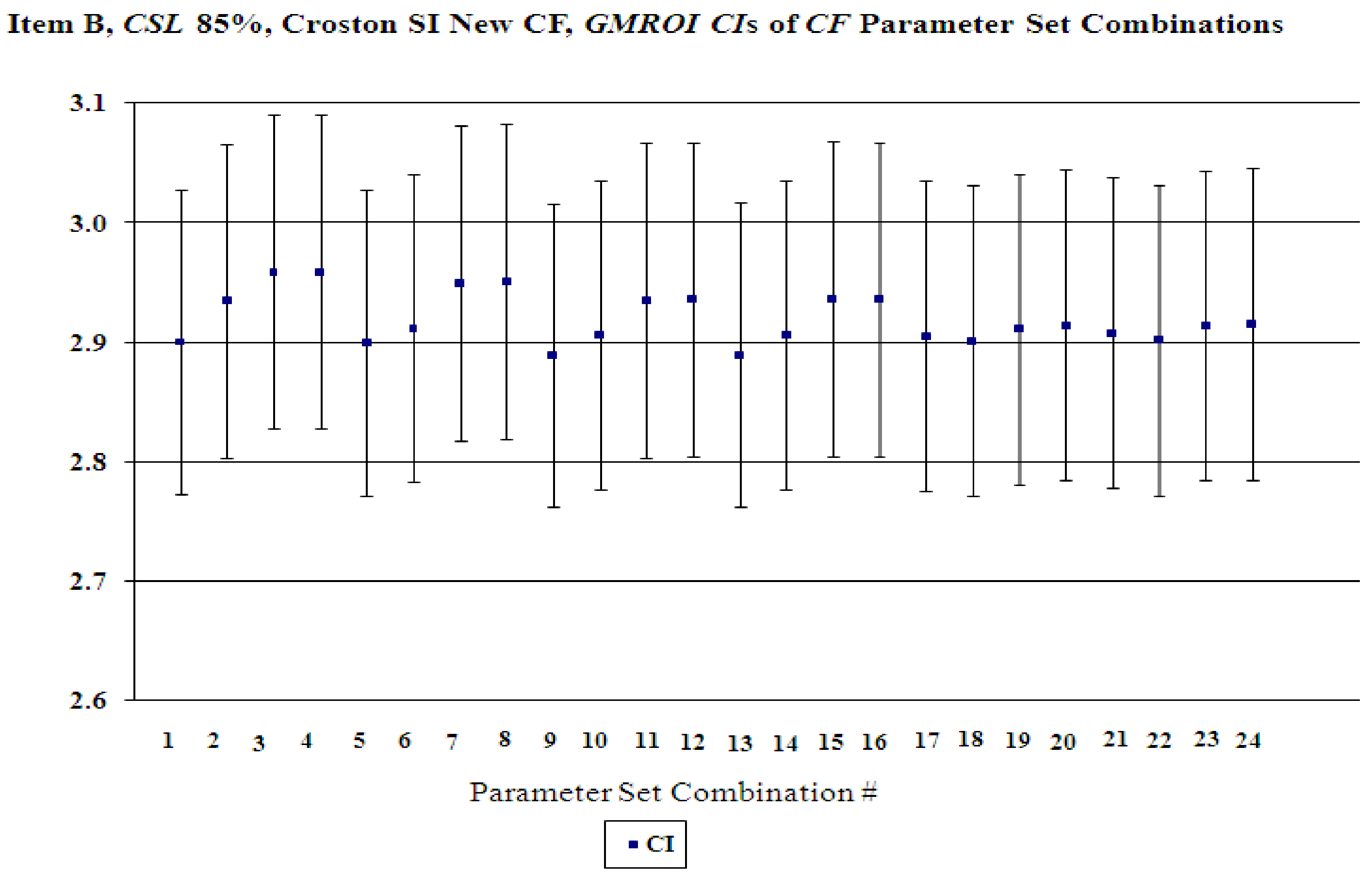

Figure B.6.2 The *GMROI* confidence intervals of Croston SI New CF parameter set combinations, Item B, target *CSL*=85%

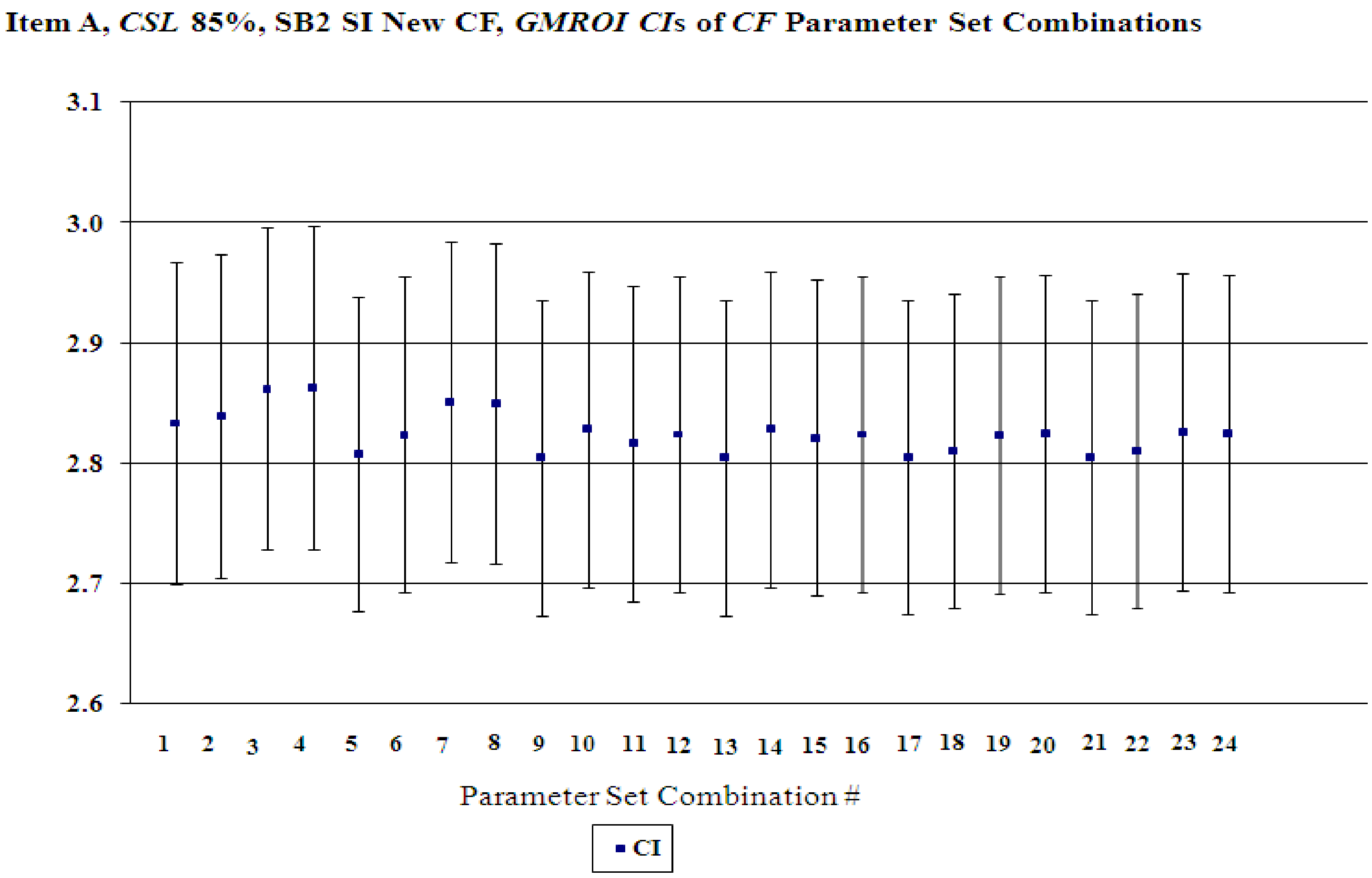

Figure B.6.3 The *GMROI* confidence intervals of SB2 SI New CF parameter set combinations, Item B, target *CSL*=85%

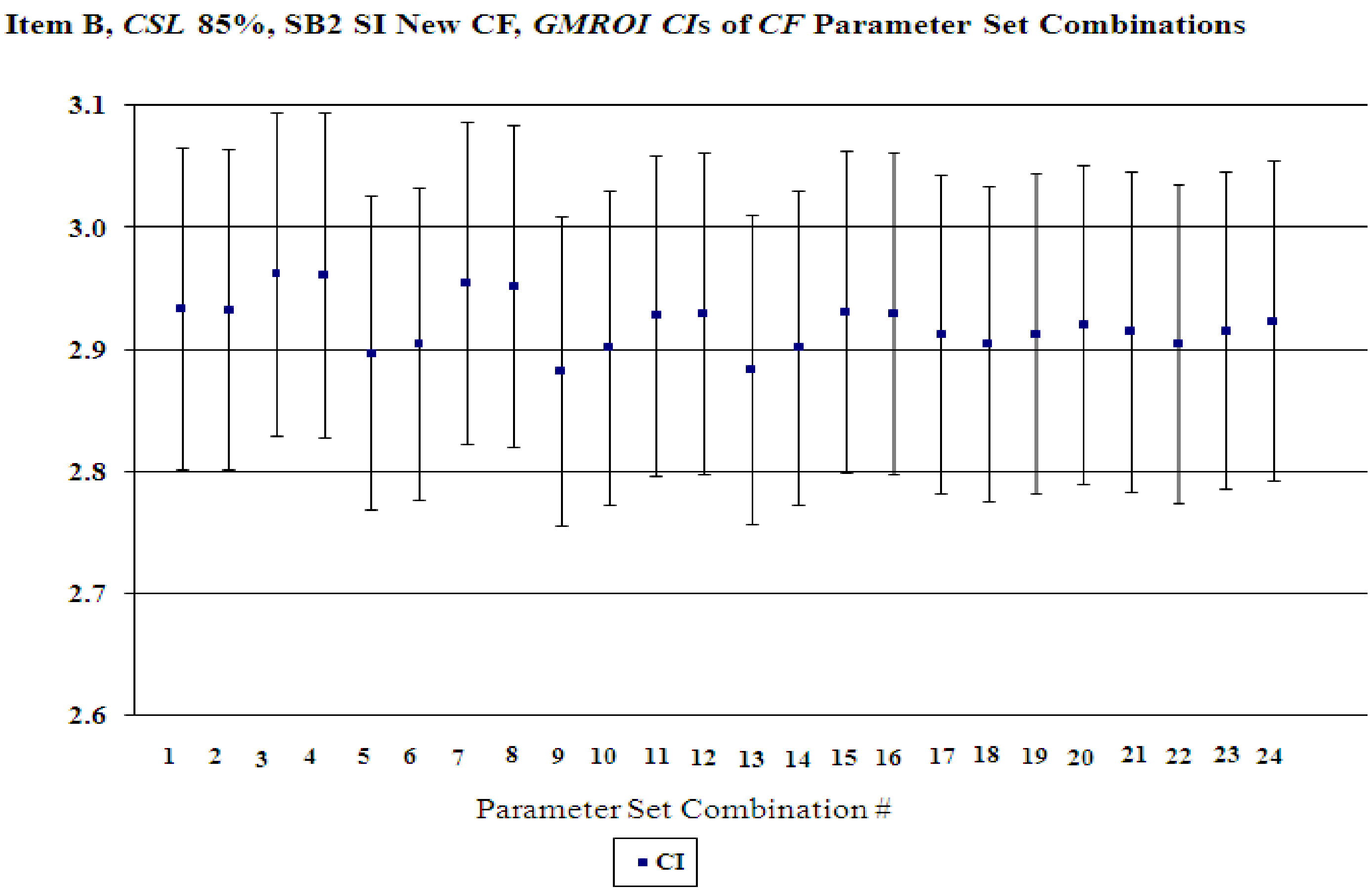

Figure B.6.4 The *GMROI* confidence intervals of SB2 SI New CF parameter set combinations, Item B, target *CSL*=85%