OPTIMIZING SMOOTHING PARAMETERS FOR THE TRIPLE EXPONENTIAL FORECASTING MODEL

BY

HARISH B NARASINGARAJ

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in Industrial Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2016

Urbana, Illinois

Adviser:

Professor Rakesh Nagi

ABSTRACT

Exponential smoothing has always been a popular topic of research in forecasting. The triple exponential smoothing in particular involves modeling a function that is a combination of level, trend and seasonal factors. While simulating the model, each of the factors is associated with a parameter whose value has a significant impact on the accuracy of the forecast, yet optimizing these parameters for a time series has received relatively little attention in literature. In this thesis we will explore the results of multi-step forecasting by using parameters optimized through an algorithm centered around h-step ahead errors. An empirical study conducted on forecasting the monthly time series from the M3-Competition across a range of horizons gave us promising results. We show that this method proves to be better than the standard Holt-Winters procedure for the entire forecasting horizon in five out the six categories of data considered . We also show that this method significantly improves the accuracy over the short term forecasting horizon when compared to the automated Holt-Winters procedure used by experts in the M3 competition. Encouraged by these results, we recommend replicating this methodology to other models of the triple exponential smoothing in the future.

Key words: Holt Winters; Triple Exponential Smoothing parameters; M3 Competition; Parameter optimization.

To my mother, for her unconditional love, support and sacrifices.
To my father, for showing me the meaning of hard work and sincerity.
To my teachers, for encouraging and guiding me on my journey.
To my friends, for molding me into who I am.

ACKNOWLEDGMENTS

Immeasurable gratitude and appreciation is extended to the following people without whom this thesis would not have been possible.

Professor Rakesh Nagi, Head and Donald Biggar Willett Professor of Industrial and Enterprise Systems Engineering, for his support, advise, guidance and valuable academic insight that was instrumental in the completion and success of this study.

Dr. Stas Grishin, Executive Vice President of Research and Product Management, Terra Technology, for the opportunity and guidance to tackle a real time industrial problem in forecasting, which provided the basis for the idea expressed in this paper.

Mr. Melih Temel, Senior Product Manager, Terra Technology, for his patience and guidance that helped me overcome numerous difficulties while working on the idea for this paper.

I want to take this opportunity to thank all the faculty members on the department of Industrial Engineering for their help and encouragement.

This acknowledgment would not be complete without thanking my friends, especially **Banu and Vats** who have stood by me as a pillar of support during all my endeavors.

Finally I have to thank my parents **Narasingaraj and Meena** who have made innumerable sacrifices so that I could pursue my dreams.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

Exponential smoothing forecasting methods originated from the work of Brown (1959, 1963)[1, 4] and Holt *et al.* (1960)[15] who were creating forecasting models for inventory control systems. From then, there has been an ocean of papers published in the field. Forecasting has become a cornerstone of optimization in today's world. Estimating future values has become key to estimating future expense of resources, thereby offering opportunities to optimize related resource utilization. This is especially seen in the area of inventory control. Despite its simplicity, for the last few decades Holt's method with modifications has been the most popular approach for forecasting a trend based series. However in cases where there is a seasonal complexity present, the Holt-Winters triple exponential smoothing method is predominantly employed. These methods are still popularly employed because of the simplicity in which they can be applied in an automated manner to large data sets with varying trend. The method follows an additive or multiplicative trend and seasonality estimation based on modeling the level, growth rate and seasonal components. Each of these components is associated with a parameter that needs to be estimated based on minimization of errors. There has been a lot of work on different models with exponential forecasting. Pegel (1969)[27] laid down the basics of time series models which has been added to by Gardner (1985)[13] and reviewed again by Hyndman *et al.* (2002)[16]. There is comparatively less literature on optimizing these parameters for forecasting time series models. Values of these parameters largely influence the accuracy of the forecasts. Research has been done on making the parameters adaptive by Williams (1988)[30]. Gardner and McKenzie (1985)[13] introduce an extra parameter to dampen the projected trend which often helps improve the accuracy of the forecasts and introduced more variation to the forecast trend. Recently there has been more research on steady state models proposed by Hyndman *et al.* (2002)[16]. Modifications to these models based on the minimization of multi-step ahead error and auto-regressive estimates of forecasts such as work done by Tucker *et al.* (2013)[23] are receiving more attention.

In most of the popular methods associated with triple exponential smoothing, a single set of parameters are used to predict an intercept and slope (adaptive in some cases) which in turn is used to predict the forecast values in the required horizon. There have been some modifications to this method to introduce more variability. Motivated by such modifications, this thesis is aimed at exploring the parameter optimization process. In this thesis, we optimize the parameters by minimizing multi-step forecast errors to forecast values at different steps in the forecasting horizon. By applying this process to the additive trend and multiplicative seasonality model we will obtain multiple sets of intercepts, slopes and seasonal factors, each of which will lead to a different set of forecasts. We will apply this algorithm to the monthly time series from the M3-Competition where we can compare its performance to the traditional method of Holt-Winters along with the results obtained by, Makridakis *et al.* (2000)[22]. Estimating multiple sets of parameters to forecast values when seasonality is present should lead to a comparable if not more robust forecasting performance, especially in the short term. Since the Holt-Winters model is one of the most easy to implement model, this study could provide a basis for improving its multi-step forecasting efficiency based on parameter selection.

In Chapter 2 of the thesis, we will review the literature on exponential smoothing method employed by Holt-Winters, its different models and modifications to its methodology. Chapter 3 introduces the algorithms and methodology we use to select the parameters while implementing the triple exponential smoothing method. In Chapter 4 we do an empirical analysis using a large data set from the M3-Competition to compare our results with the those obtained from the traditional method. Finally in Chapter 5, we provide a summary of our findings and conclusions with suggestions for possible future research areas.

CHAPTER 2

LITERATURE REVIEW

Time series forecasting started with the need to optimize inventory systems. From its introduction by Brown (1959, 1963)[1, 4] and Holt *et al.*(1957)[15] the methodology of forecasting took a radical turn when Winters (1960)[32] modified it and introduced the exponential form of the equation. He summarized the combined effects of level, trend and seasonality to introduce this basic form.

$$S_t = \alpha(X_t/I_{t-L}) + (1-\alpha)(S_{t-1} + T_{t-1}) \tag{2.1}$$

$$T_t = \beta(S_t - S_{t-1}) + (1-\beta)T_{t-1} \tag{2.2}$$

$$I_t = \gamma(X_t/S_t) + (1-\gamma)I_{t-L} \tag{2.3}$$

$$X_t(m) = (S_t + mT_t)I_{t-L+m} \tag{2.4}$$

Here X_t is the actual observations and $X_t(m)$ is the m-step ahead forecast. α, β and γ are the smoothing parameters determined by minimizing forecast errors during the iterative process. This method estimates the local trend T_t and seasonality I_t by smoothing out the successive ratios and difference of the local level S_t. L represents the length of the period associated with seasonality. The forecast function is the sum of level and projected growth multiplied with seasonality. This set of equations introduced more than 50 years ago has become one of the most popular methods used even in today's world because of its simplicity, effectiveness and ease of implementation. With gaining popularity the methodology of Holt-Winters forecasting has constantly been a popular topic for research and this eventually lead to different models of the exponential smoothing to be formulated. Pagel (1969)[27] talks about different combinations of trend and cyclic effects in additive and multiplicative form. According to Pagel, trend and seasonality, if present, consist of either an additive or multiplicative effect. This results in 9 combinations of the model. The model mentioned in equations (2.1 to

2.4) can be classified into the additive trend and multiplicative seasonality model. Following more research on models Gardner and Mckenzie (1985)[13] proposed an additional damping factor ϕ that could be combined with multiplicative or additive trend. The basic form of this damped Holt equation is as follows.

$$S_t = \alpha(X_t) + (1-\alpha)(S_{t-1} + \phi T_{t-1}) \tag{2.5}$$

$$T_t = \beta(S_t - S_{t-1}) + (1-\beta)\phi T_{t-1} \tag{2.6}$$

$$X_t(m) = (S_t + \sum_{i=1}^{m} \phi^i T_t) \tag{2.7}$$

Table 2.1: Represents the different models of the exponential smoothing method, adapted from Hyndman *et al.* (2002)[16]

Formulae for recursive calculations and point forecasts

Trend component	Seasonal component N (none)	 A (additive)	 M (multiplicative)
N (none)	$P_t = Y_t$ $Q_t = l_{t-1}$ $\phi = 1$ $F_{t+h} = l_t$	$P_t = Y_t - s_{t-m}$ $Q_t = l_{t-1}$ $T_t = Y_t - Q_t$ $\phi = 1$ $F_{t+h} = l_t + s_{t+h-m}$	$P_t = Y_t / s_{t-m}$ $Q_t = l_{t-1}$ $T_t = Y_t / Q_t$ $\phi = 1$ $F_{t+h} = l_t s_{t+h-m}$
A (additive)	$P_t = Y_t$ $Q_t = l_{t-1} + b_{t-1}$ $R_t = l_t - l_{t-1}$ $\phi = 1$ $F_{t+h} = l_t + hb_t$	$P_t = Y_t - s_{t-m}$ $Q_t = l_{t-1} + b_{t-1}$ $R_t = l_t - l_{t-1}$ $T_t = Y_t - Q_t$ $\phi = 1$ $F_{t+h} = l_t + hb_t + s_{t+h-m}$	$P_t = Y_t / s_{t-m}$ $Q_t = l_{t-1} + b_{t-1}$ $R_t = l_t - l_{t-1}$ $T_t = Y_t / Q_t$ $\phi = 1$ $F_{t+h} = (l_t + hb_t)s_{t+h-m}$
M (multiplicative)	$P_t = Y_t$ $Q_t = l_{t-1}b_{t-1}$ $R_t = l_t / l_{t-1}$ $\phi = 1$ $F_{t+h} = l_t b_t^h$	$P_t = Y_t - s_{t-m}$ $Q_t = l_{t-1}b_{t-1}$ $R_t = l_t / l_{t-1}$ $T_t = Y_t - Q_t$ $\phi = 1$ $F_{t+h} = l_t b_t^h + s_{t+h-m}$	$P_t = Y_t / s_{t-m}$ $Q_t = l_{t-1}b_{t-1}$ $R_t = l_t / l_{t-1}$ $T_t = Y_t / Q_t$ $\phi = 1$ $F_{t+h} = l_t b_t^h s_{t+h-m}$
D (damped)	$P_t = Y_t$ $Q_t = l_{t-1} + b_{t-1}$ $R_t = l_t - l_{t-1}$ $\beta < \phi < 1$ $F_{t+h} = l_t +$ $(1 + \phi + \cdots + \phi^{h-1})b_t$	$P_t = Y_t - s_{t-m}$ $Q_t = l_{t-1} + b_{t-1}$ $R_t = l_t - l_{t-1}$ $T_t = Y_t - Q_t$ $\beta < \phi < 1$ $F_{t+h} = l_t +$ $(1 + \phi + \cdots + \phi^{h-1})b_t + s_{t+h-m}$	$P_t = Y_t / s_{t-m}$ $Q_t = l_{t-1} + b_{t-1}$ $R_t = l_t - l_{t-1}$ $T_t = Y_t / Q_t$ $\beta < \phi < 1$ $F_{t+h} = [l_t +$ $(1 + \phi + \cdots + \phi^{h-1})b_t]s_{t+h-m}$

With more clarity on multiple models of forecasts as given in Table 2.1, Bates and Granger (1969)[2] introduced a new methodology which combines different sets of forecasts each of which could be calculated by using a different model. Their research also explored a method to determine weights that could be assigned to forecasts from various models to create a resultant forecast that yielded a lower mean square error than any of the individual ones. The method was simple yet empirically proven to be effective. At this point research into modifications of the existing methodologies for forecasting on the time series models received an enormous boost. Newbold and Granger (1974)[25] proposed an autoregressive model for forecasting using Holt-Winters exponential smoothing. They also pursued more research into combining various forecasts which proved to be effective. Additional and more robust research was done on combining forecasts from various models by Winkler and Makridakis (1983)[31]. They proposed five procedures to determine weights of different combinations of forecasts of which two were empirically proven to be successful.

The Box-Jenkins model was introduced by Box and Jenkins (1970)[3]. This model was based on a class of models called autoregressive integrated moving average (ARIMA) processes. Box and Jenkins popularized the ARIMA process by applying them to non-stationary and seasonal data through a set of techniques that helped identify the model and estimate the necessary parameters. Chatfield (1977)[5] wrote a paper to compare the Box-Jenkins model with the Holt-Winters model and concluded that the Holt-Winters had a superior performance 27 to 42 percent of the time. Some of the important reasons he stated for his conclusions include the following: (1) The ARIMA process did not have an equivalent for any multiplicative form of Holt-Winters; (2) The Box-Jenkins method does not necessarily identify the correct model that should be used; (3) The accuracy of the forecasts largely depend on the computational methods used to identify the parameters. Further it was noted that the additive model of Holt-Winters forecasting could be represented by the ARIMA model. Chatfield (1978)[6] continued exploring the Holt-Winters forecasting procedure in detail. He pointed out that in usual empirical studies the automated version of Holt-Winters is applied and a non-automated model which uses subjective judgment, when employed can prove to provide better results. This subjective judgment can be used to determine the nature of seasonality or trend. While going into the details of this procedure Chatfield analyzes auto-correlation coefficients for one step ahead forecast errors to indicate the effectiveness of the forecasts. He goes on to add an autoregressive parameter to fit a first-order autoregressive model to the Holt-Winters one step ahead forecast errors. This method was found to be effective in reducing the overall mean square error in four out of the five series the model was tested on. Chatfield analyzed initial starting values for mean, trend and seasonal factors that could be used in a non-automated Holt-Winters model and their effects on improving the forecasting accuracy. It was found that initial values in the iterative forecasting equations had a significant impact on the accuracy of the forecasts produced. Ledolter and Abraham (1984)[18] continued exploring the initial parameters and its effect on exponential smoothing. They proposed an initialization which uses back-forecasts and explored its relation with reducing the mean square error while forecasting. Backforecasts have their time order reveresed and the forecasting process is written in its reverse form. The proposed equation in their paper can be used to calculate back forecasts at any previous time.

Makridakis *et al.* (1982)[20] reported the results of a forecasting competition popularly known as the M-Compeition where seven experts in each of 24 methods forecasted 1001 series of data for six to eighteen steps in the forecasting horizon. The forecasting competition encouraged judgmental, explanatory and extrapolative methods as well a combination of all of them. The methods were compared across different categories and types of trend which included yearly, quarterly and monthly data. Based on forecasting horizon, it was found that Holt, Brown and Holt-Winters performed well for shorter forecasting horizons. In the deseasonalized Holt method, h-step ahead errors were calculated and for predicting multiple forecasts h-steps into the future, however there was not much difference observed in the accuracy of forecasts while following this methodology. Also an adaptive response rate exponential smoothing methodology (ARRES), as described in the same paper, did not produce any results that were significantly more accurate. The paper also explored deseasonalizing methods using the CENSUS II methodology described by McKenzie (1984)[24], which failed to do better than the traditional moving average method. The main conclusions made by the competition were as follows: (1) Statistically sophisticated and more complex methods do not always outperform simpler methodologies; (2) Relative rankings of various measures varies with the accuracy measure used;

(3) A combination of various methods on an average outperforms the individual methods; (4) Length of the forecasting horizon largely influences the accuracy of the method employed.

The Makridakis competition was critiqued as epiricism without theory. This was largely due to the empirical methodology employed by some of the forecasting methods that did not have a proper theoretical framework. Based on the results, some models seemed to outperform others and more research was needed to explore and understand why they did so. This was especially true with the models proposed by Parszen (1979)[26] and Lewandowski (1982)[19]. Gardner and McKenzie (1985)[13] studied the success factors behind Parzen and Lewandowski's models and proposed a damping factor to be added into the exponential smoothing procedure. The purpose of the damping factor was to eliminate erratic trends that might lead to inaccurate forecasts especially over long lead times. This method was shown to be superior to other forecasting procedure using the M-competitions data. The reason for this model's success was that most time series have an underlying assumption that the trend value will continue unabated regardless of the lead time. Gardner *et al.*[13] challenged that assumption. Following a similar thought process, Williams (1987)[30] proposed an adaptive forecasting model where the smoothing parameters are determined through an adaptive algorithm instead of using a constant set to test the data. Based on his tests Williams found his proposed model was more likely to perform better than the non adaptive version of the Holt-Winters model. Cipra (1992)[8] modified the Holt-Winters methodology to a weighted regression problem that still retains its recursive nature. Another similar technique proposed by Williams and Miller (1999)[29] was to improve the accuracy of long term forecasts by including judgmental adjustments within the exponential smoothing forecasts.

With the introduction of multiple models on exponential smoothing in forecasting, Gardner and McKenzie (1988)[9] analyzed through experiments the effect of variances of the differences, that time series could have on model selection. They discuss an objective model identification procedure to help deduce the relevant models based on time series data. Gardner (1990)[10] continued this research on model selection and gauged its impact on inventory control systems. Models were evaluated based on inventory costs and service levels. The study bolstered the importance of model selection required to determine the amount of inventory investment needed to support a particular service level. Gardner (2006)[12] does a more detailed analysis on all the proposed methodologies in exponential smoothing, summarizing over 50 years of research in his paper.

Following the success of the first M competition, the M2 and M3 competitions were conducted to explore new methodologies and validate existing ones. The main purpose of the M2 competition as described by Makridakis *et al.* (1993) [21] was to prove that the first M competition was not just an empirical study. The M2 competition consisted of 29 actual series, six of which were macro economic in nature and the rest were from four companies. Forecasters were allowed to ask for more qualitative information from the participating companies. The results were no different from that observed in the first M competition. The M3 competition however consisted of 3003 series of data which served as a repository for empirical studies on forecasting. The M3 competition had four different categories of data which included annual, monthly, weekly and other. 24 well known methods were employed to test the data and results were recorded. The main conclusions of the M3 competition as summarized by Makridakis *et al.* (2000)[22] were: (1) The observation that statistically complex methods were not necessarily more accurate than simpler ones was reiterated by the results of the competition; (2) The competition validates the use of some of the new methods like ForecastPro which is empirically based and eclectic in nature; (3) Some methodologies were found to outperform all others when a particular trend is involved and more detailed studies are required to find the reason for such behavior. The database of the M3 competition served as an empirical test to numerous papers proposing different methodologies for forecasting.

Encouraged by the availability of a robust database for testing, Hyndman *et al.* (2002)[16] provided a different approach to exponential forecasting by using equivalent state space models. This helped with increasing the calculation of likelihood and estimating the type of model that could be used. Estimating prediction intervals and random simulations were also easier when the state space model was used. Following the advent of the state space models, a lot of research began focusing around its framework. Kirkendall (2006)[17] proposed a method to analyze and classify the observations in exponential smoothing into steady, outlier, and level shift models, using statespace sequential data processing. Gelper *et al.* (2010)[14] applied a version of the

Kalman filter to the state space model associated with exponential and Holt-Winters smoothing. The method yields a recursive updating scheme that results in more robust forecasts. The paper also proposed a more robust method of selection of parameters. Another relevant study was done by Mcelroy and Wildi (2013)[23] which focused on the minimization of a multi-step ahead error criterion that is based on the asymptotic average of squared forecast errors. The paper derives a non-linear function of the parameter using a generalized KullbackLeibler measure on state space models.

While there has been a large amount of literature on sophisticated variations to time series forecasting, the variation in parameters while determining multi-step ahead forecasts has been minimal. The method of determining various steps ahead by smoothing out parameters based on errors in the corresponding seasons of the Holt-Winters triple exponential smoothing has not received enough attention. This thesis makes a contribution by leveraging the M3 competitions data for an empirical study of the aforementioned research gap.

CHAPTER 3

METHODOLOGY

As mentioned previously in Chapter 2, while there exists various modifications to the triple exponential smoothing methodology, where the smoothing parameters are made adaptive or are derived from a combination of various user defined errors there has been no robust study on a method that constantly changes the smoothing parameters for forecasting at every step in the horizon. We wish to follow up on this through experimenting on the above mentioned process to the Holt-Winters triple exponential smoothing in an automated manner. The process derived by Holt (3.1 to 3.3) is a simple iterative equation where X_t is the actual observations and $X_t(m)$ is the m-step ahead forecast derived by smoothing the level S_t and trend T_t. α represents the level smoothing parameter and β represents the trend smoothing parameter.

$$S_t = \alpha(X_t) + (1 - \alpha)(S_{t-1} + T_{t-1}) \tag{3.1}$$

$$T_t = \beta(S_t - S_{t-1}) + (1 - \beta)T_{t-1} \tag{3.2}$$

$$X_t(m) = (S_t + mT_t)I_{t+m} \tag{3.3}$$

When a seasonal element is involved there can be two ways to modify Holt's procedure. The entire data set can be deseasonalized based on derived seasonal factors after which Holt's methodology can be applied and the resulting forecast can be reseasonalized. A more common method is to use the Holt-Winters forecasting procedure (3.4 to 3.7) which in turn is an extension of the the original Holt's equations to include seasonal smoothing I_t through a smoothing paramater γ, where L is the length of the period. These three parameters largely impact the accuracy of the forecasts. This is the model we will use to experiment with because of the ease with which it can be automated, especially when modifications are made to it.

$$S_t = \alpha(X_t/I_{t-L}) + (1-\alpha)(S_{t-1} + T_{t-1}) \tag{3.4}$$

$$T_t = \beta(S_t - S_{t-1}) + (1-\beta)T_{t-1} \tag{3.5}$$

$$I_t = \gamma(X_t/S_t) + (1-\gamma)I_{t-L} \tag{3.6}$$

$$X_t(m) = (S_t + mT_t)I_{t-L+m} \tag{3.7}$$

While there has been considerable literature published on studies that derive the optimal parameters to be used based on specific industries, Chatfield (1978)[6] overrides that statement through studies on different time series with varying trends and seasonality, He mentions that the parameter values are not based on industry but are based on the degree with which seasonality and trend vary. Further literature has always concentrated on combining various methods and establishing criteria for user defined errors which can be minimized. In this paper we will follow two methods as described below, and simulate the procedure on the M3 competition's database for monthly forecasts and compare the results with other methods published for the same database. We simulated these forecasts in an automated manner by running scripts codded on Matlab.

Method 1 - **HW**: Traditional Holt-Winters that uses an additive trend and multiplicative seasonality. The smoothing parameters used to derive the forecast for the entire horizon are fixed and lie between 0 and 1. The parameters are smoothed through a grid search with 0.05 increments. the minimized error is the one step ahead forecast's Mean Absolute Percentage Error (MAPE). The error equations can be represented as below.

$$F_t = (S_{t-1} + T_{t-1})I_{t-L} \tag{3.8}$$

$$M = \frac{1}{N}\sum_{t=1}^{N} |\frac{F_t - S_t}{S_t}| \tag{3.9}$$

Here F_t represents the one step ahead forecasts taken for any time T and M represents the error that is required to be minimized. N represents the total data points or time stamps associated with the data available for testing. Minimizing this MAPE error over for the testing data for the smoothing parameters using the three-dimensional grid will result in the optimal values that can be used to forecast values for the entire horizon. One of the reasons we are using the MAPE error over the more frequently used sum of squares error is because the MAPE error will consider the size of the data point that is being forecasted while deriving the error. We found this an important criteria to consider after a few discussions with companies that use forecasting. This method would emphasize better accuracy in cases when data points are relatively smaller in the series. The actual deviation in the error would be seen relative to the size of the data point being forecasted.

Method 2 - **HW-New**: Here we use the a method similar to the one mentioned in method 1. The additive trend with multiplicative seasonality version of Holt-Winters is used to calculate the forecasts and the smoothign parameters are optimized through a grid search from 0 to 1 using 0.05 increments. However the error that is minimized varies as given below.

$$F_t(m) = (S_{t-1}(m) + T_{t-1}(m))I_{t-L}(m) \tag{3.10}$$

$$M(m) = \frac{1}{N}\sum_{t=1}^{N}\left|\frac{F_t(m) - S_t(m)}{S_t(m)}\right| \tag{3.11}$$

Where $F_t(m)$ represents the one step ahead forecasts taken for any time t while deriving the actual forecast $X_t(m)$, which is the m-step ahead forecast given by

$$X_t(m) = (S_t + mT_t)I_{t-L+m} \tag{3.12}$$

$M(m)$ represents the error that will be minimized to compute the forecast m steps ahead and t here represents the time. N similar to before represents the total data points or time stamps associated with the data available for testing which is independent of the forecast step. We must also mention that this methodology follows the reasoning that data points taken m steps in separation will follow different smoothing patterns which will lead to different smoothing parameters. For this reason m has to be less than the seasonal period L, and if the value is m^0 such that m^0 is greater than L we replace the value of m^0 with m such that $m = m^0 - L$.

$$\begin{aligned} F_t(m^0) &= (S_{t-1}(m^0) + T_{t-1}(m^0))I_{t-L}(m^0) \\ &= (S_{t-1}(m) + T_{t-1}(m))I_{t-L}(m) \end{aligned} \tag{3.13}$$

By applying this reduction the parameters based on the error reduction will be equal to that derived for m which is less than L. Hence it is only required to find the smoothing parameters for the forecasts up to L steps ahead and repeat the values.

When we are to consider the initial values for these methods, it must be understood that initial values themselves influence the parameter selection as much as the errors that are minimized. Different initial values lead to different parameters which will result in substantially different forecasts as described by Chatfield and Yar (1988)[7]. It thus important to chose our values for initial smoothed level, trend and seasonal components carefully. As mentioned in our literature review Ledolter and Abraham (1984)[18] gave considerable theoretical arguments to use backcasting, however Makridakis *et al.*, (2000)[22] empirically showed that backcasting gave poor values based on the results of the M3 competition. Gardener and Mckenzie (1985)[13] also proposed a simple linear regression on time that produced both initial level and trend. However if the series consisted of more than a simple linear trend then this method would be not make much sense and in our case, all the data series are monthly values and assumed to have a seasonal component. So for all purposes of our experiments we represent our initial trend value by calculating simple averages of the first few months which is similar to the approach used by Gardner (1999)[11]. We adopted the initial growth by calculating 1/12 of the difference between the same months in the first and second year. The average of the difference was then taken to the value of T_0. The initial level determined the starting point of our forecast and instead of deriving S_0 and we set S_1 to the same value as the first observation X_1. For deriving the seasonal components we used an approach of taking the average of the ratios of each of the actual observation in the month to the average of all data points in the period represented by the length of a year. However when calculating the seasonal indices we did not restrict ourselves to the first few years instead we calculated it for the entire data set of the testing data for which whole period data points were available. This was done based on testing the data and accuracy of forecasts by using different values of seasonal indices initialized based on selecting different time ranges.

Some of the methods that are also mentioned for comparative purposes are those used by Makridakis *et al.* (2000) and (1982)[20, 22]. The Holt-Winters method used here is a two or three parametric model based on the presence of seasonality determined by an auto-correlation factor. If seasonality was not observed Holt's 2 parameter model was used, otherwise the additive trend and multiplicative seasonality model of Holt-Winters was used. The parameters were optimized based on a grid search by minimizing the mean square error MSE (3.14). We shall refer to this method as - **HW M3**

$$MSE = \frac{1}{N}\sum_{t=1}^{N}(F_t - S_t)^2 \qquad (3.14)$$

We will also use some of the more sophisticated methods used in the M3 competition as described in Makridakis *et al.* (2000)[22] to compare the short term forecasts, which are more relevant to all practical applications in the industry.

CHAPTER 4

EMPIRICAL EXPERIMENTS AND RESULTS

After formulating the algorithms as described in Chapter 3, an empirical analysis was conducted to investigate their forecasting accuracy. 1,428 monthly time series from the most recent forecasting competition, namely the M3 competition was used as the testing data for our empirical experimentation. The data points varied in each of the series from 48 to 126, the median of which was 115 and we forecasted 18 months of data for each of the series. It must be mentioned that none of the data points contained zeroes. The data from the competition belonged to 6 different categories namely industrial, finance, demographic, micro-economic, macro-economic and others. This allowed for a more detailed industry specific analysis of our methods. We must also mention that both the yearly and quarterly data from the M3 competition were not used because for practical purposes automated procedures are rarely applied to such series of low frequencies. Also, we wanted to focus on the seasonal variation of exponential smoothing methodology which is most applicable to the monthly data points by considering years to represent the cyclic seasonal periods. While comparing our results we used the Symmetric Mean Absolute Percentage Error (SMAPE) which was the one error measure which was reported in detail by Makridakis *et al.* (2000)[22] in their presentation on the results of the M3-Competition and is given by

$$M = \frac{1}{N}\sum_{t=1}^{N} \left|\frac{F_t - X_t}{\frac{S_t + X_t}{2}}\right| \tag{4.1}$$

This measure has an advantage over the MAPE as it does not give large errors when the forecast is significantly high while the actual observation is close to zero. Also, due to its symmetric nature the large percentage differences that occurs when the forecast value F_t is more that the actual value X_t and the actual value is more that the forecast, are avoided. To represent the data in a simple manner we segmented the forecasts into three buckets, short-term (0 to 6 months), medium-term (7 to 12 months) and long-term (13 to 18 months) forecasts. Since the short term is more important to industries and has direct implications on their short term resource utilization we give this more emphasis and represent the forecasts in this bucket in a more detailed manner. We will also explore a more detailed analysis of the different categories of data at the end of this Chapter.

Table 4.1: Summary of SMAPE results from all methods with the 1,428 data series from the M3 competition

	Forecasting Horizon			
Method	1 to 6	7 to 12	13 to 18	Overall
WH M3	13.44	15.06	19.33	15.94
WH	13.13	16.38	21.55	17.02
WH New	12.90	16.13	20.59	16.54

Table 4.2: Detailed SMAPE results for the first 6 months of forecasts

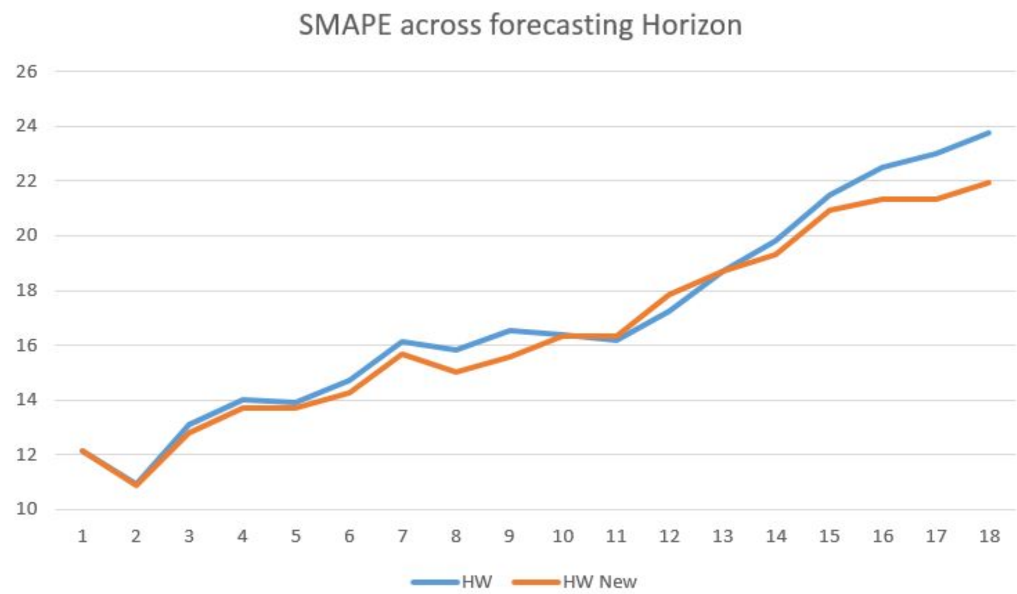
	Forecasting Horizon					
Method	1	2	3	4	5	6
WH M3	12.50	11.79	13.80	14.96	13.91	**13.71**
WH	12.13	10.94	13.09	13.99	13.90	14.73
WH-New	**12.13**	**10.88**	**12.80**	**13.70**	**13.68**	14.23

From Table 4.1 we can see that the HW-New method which we propose in this research performs considerably better in the short term. For the long and medium term the method HW M3 defined by Makridakis *et al.* (2000)[22] performs better. This could be because of the fact that HW M3 method involves both 2 and 3 parameter based prediction which could result in better accuracy in the long term. An important inference that needs to be made is that when compared to a traditional implementation of the automated model for Holt-Winters method, namely the HW method described in this paper, the method with modifications the HW-New performs significantly better. Within the short-term forecasting, when these values are inserted into the Table 4.3 consisting of results for 24 methods summarized by Makridakis *et al* (2000)[22] the rankings based on SMAPE for HW M3, HW and HW-New are 18, 12 and 9 respectively.

Table 4.3: Details of short term forecasts as derived in Makridakas *et al.* (2000)[22]

Method	Forecasting Horizon 1 to 6	Rank
Nave2	15.13	25
Single	13.44	18
Holt	13.11	10
Dampen	12.67	6
Comb SHD	12.79	8
BJ Automatic	12.74	7
Autobox1	13.42	17
Autobox2	13.52	22
Autobox3	13.47	21
Robust-Trend	15.42	26
ARARMA	13.59	23
Automat ANN	12.55	4
Flores/Pearce1	13.93	24
Flores/Pearce2	13.21	16
PP-autocast	13.11	10
ForecastPro	11.82	2
SmartFcs	12.58	5
Theta-sm	13.2	14
Theta	11.8	1
RBF	13.18	13
ForecastX	12.31	3
AAM1	13.2	14
AAM2	13.45	20
HW	**13.13**	**12**
HW M3	**13.44**	**18**
HW-New	**12.9**	**9**

Figure 4.1: Average SMAPE of forecasts for 1428 monthly data series taken from the M3 competition



A more detailed analysis of how the HW-New method performs when compared to the traditional version of the HW can be inferred from a graph representing all 18 months of forecast. From this graph we can see that the HW-New, which is a derived by optimizing different smoothing parameter for every step of the forecast, clearly outperforms the standard method.

We will continue this comparative analysis in a more detailed manner to observe how different segments of the M3 competitions data perform when the standard and the modified methods are applied to them. We have represented the results of average SMAPE taken across the respective data and represented in tables and charts. The tables contain the data summarized in buckets as described before and the charts contain the average for every step forecasted in the forecasting horizon.

From the results of the tables and graphs it can be seen that the new method proposed clearly outperforms the traditional one in nearly every segment except the others category. This is especially clear across industrial and micro-economic data. A possible reasoning could be because of the strong presence of seasonality with such data that offers a different smoothing of level and trend when applied to a constant number of steps ahead.

Table 4.4: Average SMAPE of forecasts for 111 demographical monthly data series taken from the M3 competition

Method	Forecasting Horizon			
	1 to 6	7 to 12	13 to 18	Overall
HW	5.86	10.61	14.09	10.18
HW-New	6.08	9.69	14.65	10.14

Figure 4.2: Average SMAPE of forecasts for 111 demographical monthly data series taken from the M3 competition

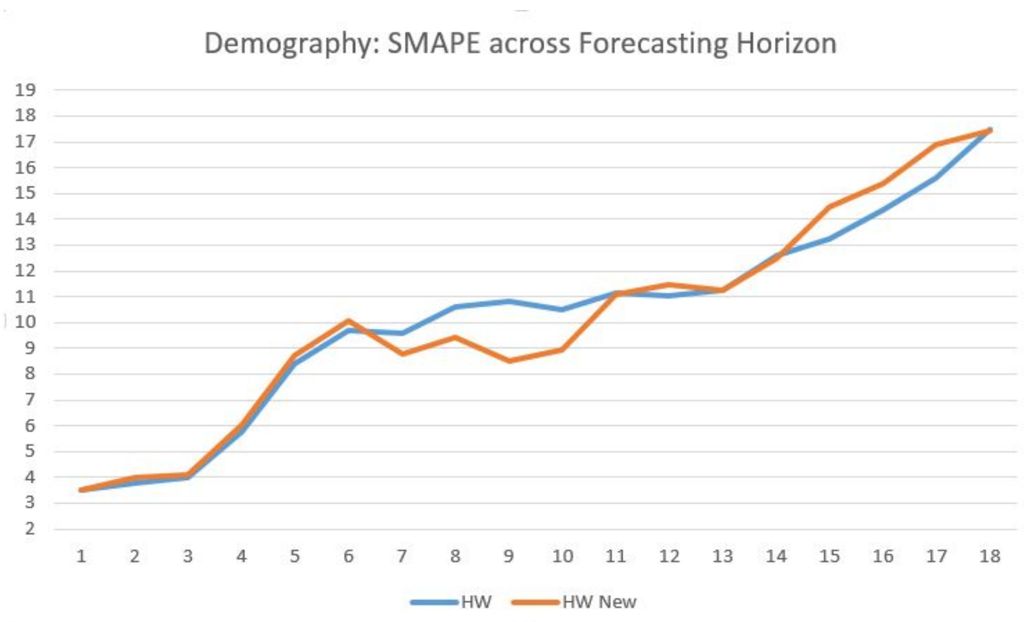


Table 4.5: Average SMAPE of forecasts for 145 financial monthly data series taken from the M3 competition

	Forecasting Horizon			
Method	1 to 6	7 to 12	13 to 18	Overall
HW	10.52	14.86	19.55	14.98
HW-New	10.15	15.15	19.28	14.86

Figure 4.3: Average SMAPE of forecasts for 145 financial monthly data series taken from the M3 competition

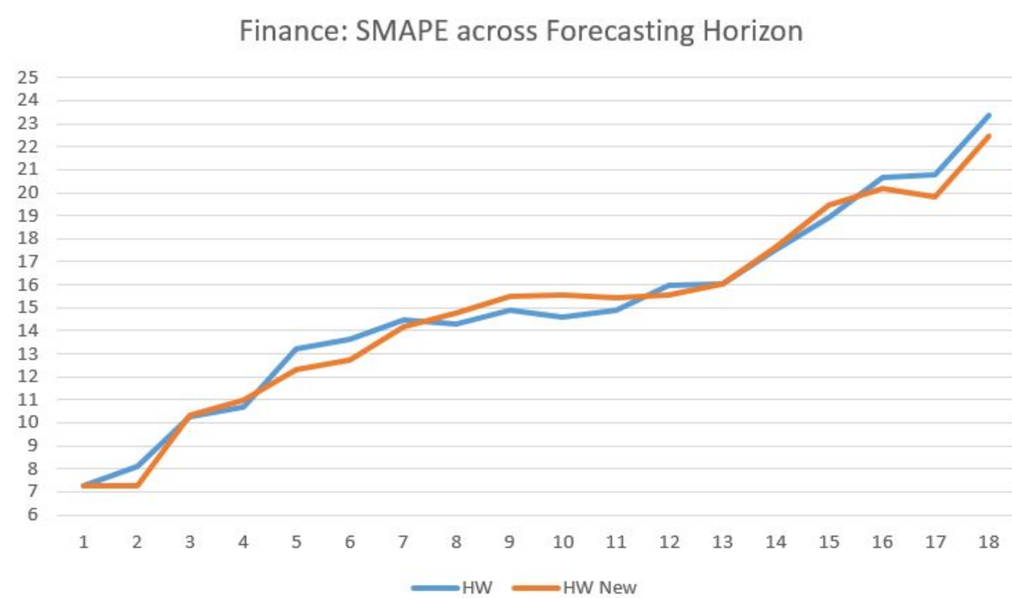


Table 4.6: Average SMAPE of forecasts for 334 industrial monthly data series taken from the M3 competition

	Forecasting Horizon			
Method	1 to 6	7 to 12	13 to 18	Overall
HW	9.97	14.83	18.68	14.49
HW-New	9.62	14.36	17.10	13.69

Figure 4.4: Average SMAPE of forecasts for 334 industrial monthly data series taken from the M3 competition

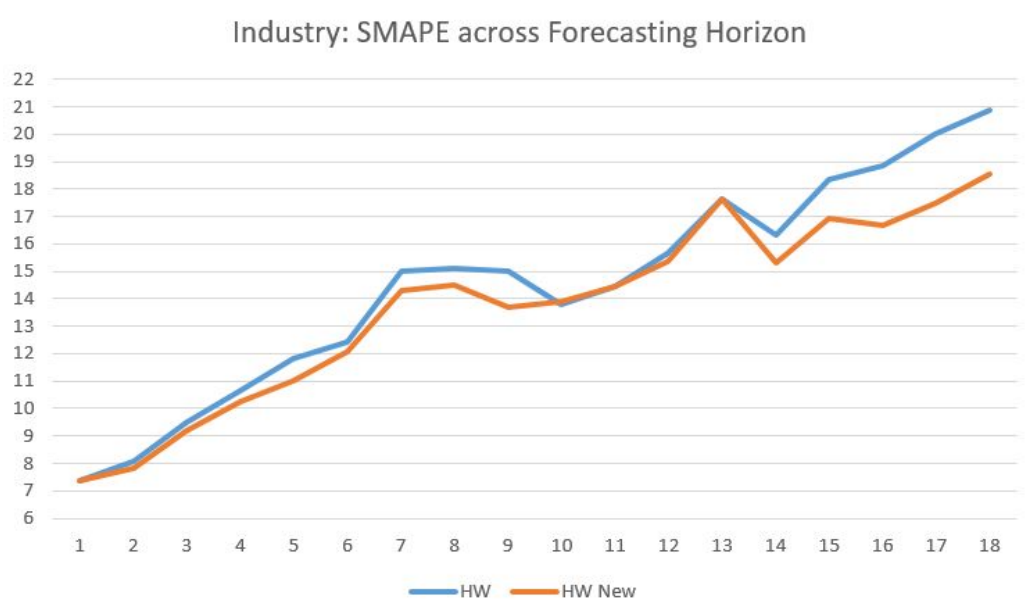


Table 4.7: Average SMAPE of forecasts for 312 macro-economic monthly data series taken from the M3 competition

	Forecasting Horizon			
Method	1 to 6	7 to 12	13 to 18	Overall
HW	23.58	25.66	33.94	27.73
HW-New	23.13	24.98	31.95	26.69

Figure 4.5: Average SMAPE of forecasts for 312 macro-economic monthly data series taken from the M3 competition

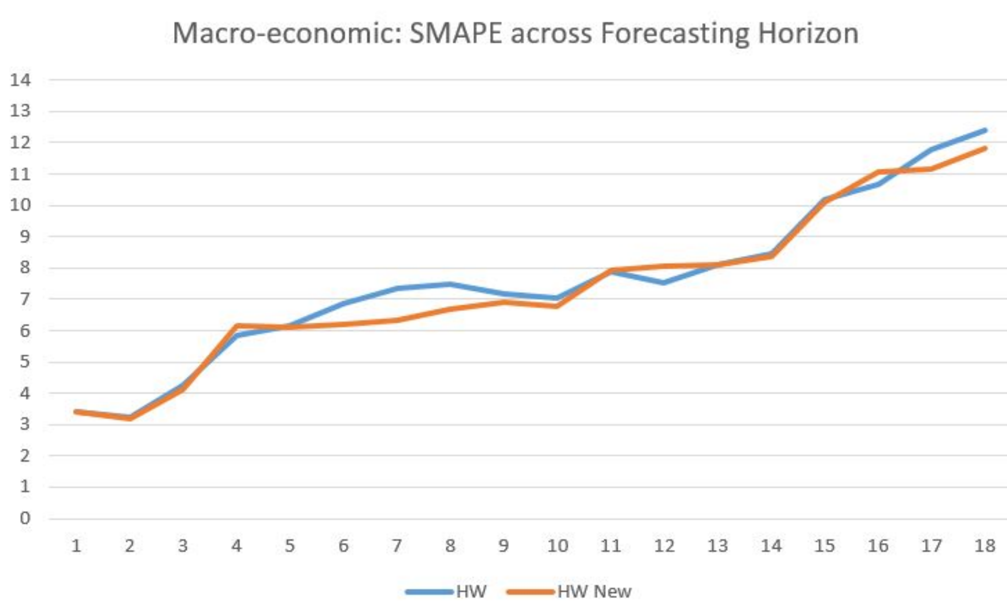


Table 4.8: Average SMAPE of forecasts for 474 micro-economic monthly data series taken from the M3 competition

	Forecasting Horizon			
Method	1 to 6	7 to 12	13 to 18	Overall
HW	4.95	7.41	10.26	7.54
HW-New	4.86	7.11	10.11	7.36

Figure 4.6: Average SMAPE of forecasts for 474 micro-economic monthly data series taken from the M3 competition

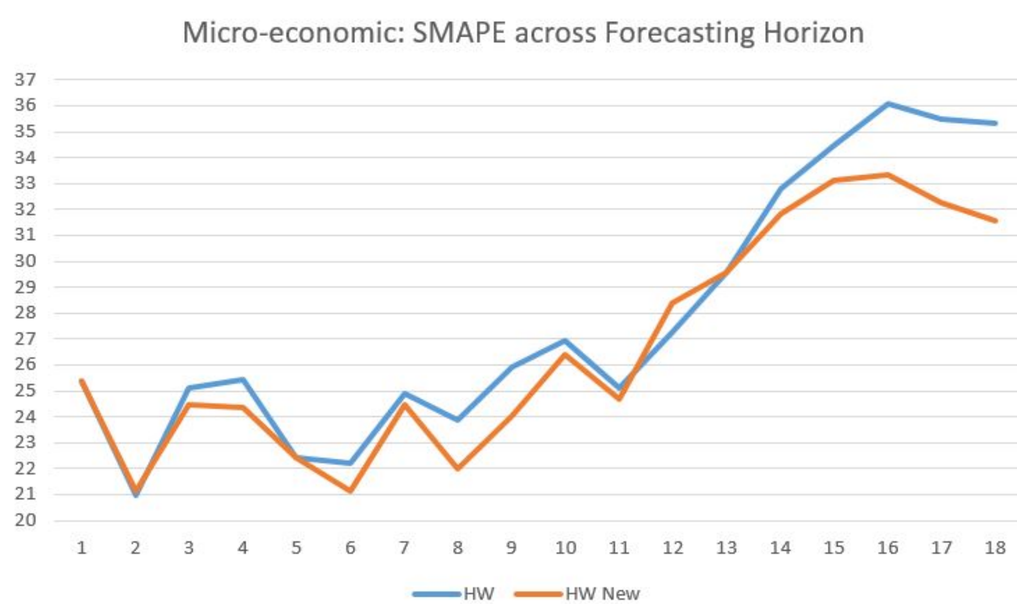
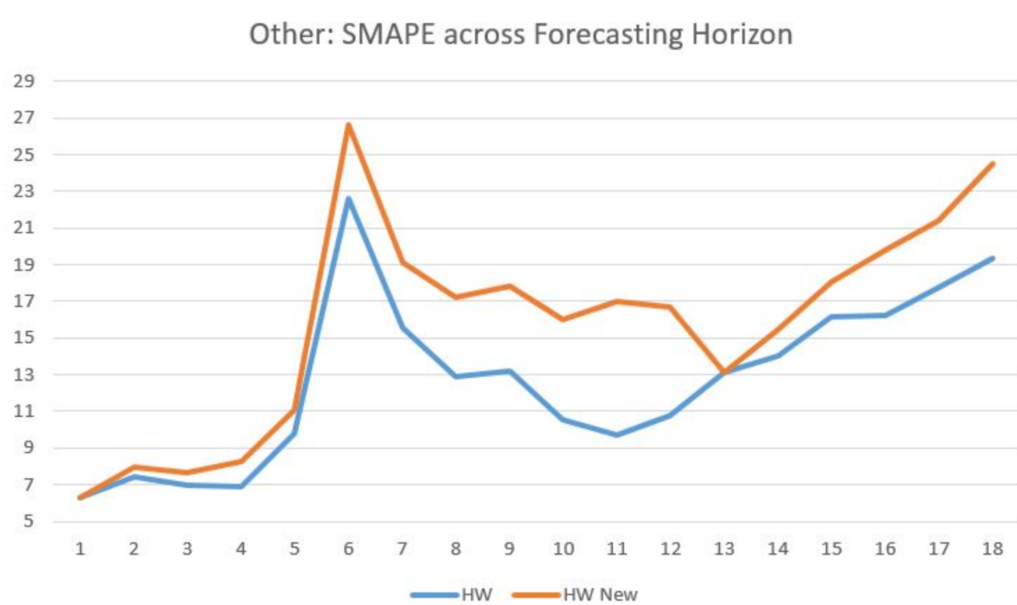


Table 4.9: Average SMAPE of forecasts for 52 other category monthly data series taken from the M3 competition

	Forecasting Horizon			
Method	1 to 6	7 to 12	13 to 18	Overall
HW	10.00	12.11	16.11	12.74
HW-New	11.30	17.30	18.74	15.78

Figure 4.7: Average SMAPE of forecasts for 52 other category monthly data series taken from the M3 competition



CHAPTER 5

SUMMARY AND CONCLUSIONS

Triple exponential forecasting and, in particular, Holt-Winters method continues to be a corner stone of forecasting even though it has been more than 50 years since its inception. This popularity can be attributed to its simple methodology that can be easily implemented in an automated manner to a large data set. While there has been a significant amount of research that alters the core of the triple exponential method, there has been relatively lesser focus on the smoothing parameters that need to be optimized. The majority of research that has focused on optimizing the smoothing parameters propose a single set of parameters to be used for the entire forecasting horizon. We wish to challenge this idea and explore a method that estimates a different set of smoothing parameters for different steps in the forecasting horizon.

In this thesis, we have implemented the traditional Holt-Winters method for forecasting and proposed modifications to the parameter optimization process. We identify different smoothing parameters based on minimizing the error associated with the step being forecasted in the forecast horizon. We have used 1,428 monthly time series data from the M3 competition to compare with both the Holt-Winters from the results in the competition and the standard method in which Holt-Winters procedure is applied. We found that the new method clearly outperformed the standard method across the horizon. While analyzing the average SMAPE of the forecasts across the different categories of data, we found that the improvement of the new method was most significant while forecasting for industrial and micro-economic data. The new method was in fact inferior to the standard method when it came to data in the "other" category of the M3 database. We concluded that this could be caused by strong effects of seasonality and more research is needed in this area to analyze when the modified method will have a high probability of yielding better results. This could be based on the correlation factor, as used by Makridakis *et al.* (1982)[20], to gauge seasonality. When considering the short term forecast values, the new method once again significantly outperformed the other methods which can be seen by the change in ranks projected. This is extremely important as most industries find short term projections as important factors that influence their resource optimization.

When considering long term forecasting the new method has higher SMAPE error than the method employed in Makridakis *et al.* (2000)[22] but still remains lesser than the standard Holt-Winters forecasting procedure. As explained before, a possible reason for this could be the fact that Makridakis uses a combination of a 2- and 3-parameter model while forecasting. In the traditional method, the trend is constant and increases in an additive manner while forecasting a constant number of steps ahead. Forecasts in the later part of long forecasting horizons are affected heavily by the trend, and when multiplied with the seasonal factor, these values could result in large errors, which can be avoided by using the 2-parameter model. A possible solution to circumvent this problem and reduce long term forecasting errors could be using the modified method with damped exponential smoothing as implemented by Gardner *et al.*(1985)[13] and Taylor (2003)[28]. By applying a damping factor the additive part of the trend would reduce with each forecasting step which should minimize the overall SMAPE. It would be interesting to explore how the damping factor itself would vary, while forecasting multiple steps ahead in the forecasting horizon and if the method explored in this paper would prove to be better than one that uses constant smoothing and damping parameters to forecast over the entire horizon.

In conclusion, we feel that the results for the 1,428 series indicate that our modifications, which encourage a different set of parameters for each of the steps in the forecasting horizon, when compared with the traditional application of Holt-Winters forecasting, yields a relatively higher accuracy rate. This is especially seen when applied to an automated procedure where short term forecasting is important.

CHAPTER 6

REFERENCES

[1] *Statistical Forecasting for Inventory Control*, New York, 1959. McGraw-Hill.

[2] J. M. Bates and C. W. J. Granger. The combination of forecasts. *Operational Research Society*, 20(4):451–468, Dec 1969.

[3] G. E. P. Box and G. M. Jenkins. *Time Series Analysis: Forecasting and Control*. Holden-Day, San Francisco, 1970.

[4] R. Brown. *Smoothing, Forecasting and Prediction of Discreet Time Series*. Prentice Hall, Englewood Cliffs, N.J., 1963.

[5] C. Chatfield. Some recent developments in time-series analysis. *Journal of the Royal Statistical Society. Series A (General)*, 140(4):492–510, 1977.

[6] C. Chatfield. The holt-winters forecasting procedure. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 27(3):264–279, 1978.

[7] C. Chatfield and M. Yar. Holt-winters forecasting: Some practical issues. *Journal of the Royal Statistical Society. Series D (The Statistician)*, 37(2):129–140, 1988.

[8] T. Cipra. Robust exponential smoothing. *Journal of Forecasting*, 11:5769, 1992.

[9] J. Everette S. Gardner. A simple method of computing prediction intervals for time series forecasts. *Management Science*, 34(4):541–546, Apr 1988.

[10] J. Everette S. Gardner. Evaluating forecast performance in an inventory control system. *Management Science*, 36(4):490–499, Apr 1990.

[11] J. Everette S. Gardner. Rule-based forecasting vs. damped trend exponential smoothing. *Management Science*, 45:1169–1176, 1999.

[12] J. Everette S. Gardner. Exponential smoothing: The state of the art part ii. *International Journal of Forecasting*, 22(4):637–666, Sept 2006.

[13] J. Everette S. Gardner and E. McKenzie. Forecasting trends in time series. *Management Science*, 31(10):1237–1246, Oct 1985.

[14] S. Gelper, R. Fried, and C. Croux. Robust forecasting with exponential and holtwinters smoothing. *Journal of Forecasting*, 29:285300, Sept 2010.

[15] C. C. Holt, F. Modigliani, J. F. Muth, and H. A. Simon. *Planning Production, Inventories, and Work Force*. Prentice Hall, Englewood Cliffs, N.J., 1960.

[16] R. J. Hyndman, A. B. Koehler, R. D. Snyder, and S. Grose. A state space framework for automatic forecasting using exponential smoothing methods. *International Journal of Forecasting*, 8:439–454, 2002.

[17] N. Kirkendall. Monitoring for outliers and level shifts in kalman lter implementations of exponential smoothing. *Journal of Forecasting*, 11:543560, 2006.

[18] J. Ledolter and B. Abraham. Some comments on the initialization of exponential smoothing. *Journal of Forecasting*, 3:79–84, 1984.

[19] R. Lewandowski. Sales forecasting by forsys. *Journal of Forecasting*, 1:205214, 1982.

[20] S. Makridakis, A. Andersen, R. Carbone, R. Fildes, M. Hibon, R. Lewandowski, J. Newton, E. Parzen, and R.Winkler. The accuracy of extrapolation (time series) methods: Results of a forecasting competition. *Journal of Forecasting*, 1:111–153, 1982.

[21] S. Makridakis, C. Chatfield, M. Hibon, M. Lawrence, T. Mills, K. Ord, and L. F. Simmons. The m2-competition: A real-time judgmentally based forecasting study. *International Journal of Forecasting*, 9:5–22, 1993.

[22] S. Makridakis and M. Hibon. The m3-competition: results, conclusions and implications. *International Journal of Forecasting*, 16:451–476, 2000.

[23] T. McElroy and M. Wildi. Some comments on the initialization of exponential smoothing. *International Journal of Forecasting*, 29:378–394, 2013.

[24] S. K. McKenzie. Concurrent seasonal adjustment with census x-11. *Journal of Business and Economic Statistics*, 2(3), Jul 1984.

[25] P. Newbold and C. W. J. Granger. Experience with forecasting univariate time series and the combination of forecasts. *Journal of the Royal Statistical Society. Series A (General)*, 137(2):131–165, 1974.

[26] E. Parzen. Time series and whitening filter estimation. *TIMS Studies in Management Science*, 12:149165, 1979.

[27] C. C. Pegels. Exponential forecasting: Some new variations. *Management Science*, 15(5):311–315, Jan 1969.

[28] J. W. Taylor. Exponential smoothing with a damped multiplicative trend. *International Journal of Forecasting*, 19:715–725, 2003.

[29] D. W. Williams and D. Miller. Level-adjusted exponential smoothing for modeling planned discontinuities. *International Journal of Forecasting*, 15:273–289, 1999.

[30] T. M. Williams. Adaptive holt-winters forecasting. *The Journal of the Operational Research Society*, 38(6):553–560, Jun 1987.

[31] R. L. Winkler and S. Makridakis. The combination of forecasts. *Journal of the Royal Statistical Society. Series A (General)*, 146(2):150–157, 1983.

[32] P. R. Winters. Forecasting sales by exponentially weighted moving averages. *Management Science*, 6(3):324–342, Apr 1960.

APPENDIX A

MATLAB SCRIPT FOR ALGORTITHMS USED

```
% Code to run traditional holt-winters
%triple exponential smoothing
e_flag=0;
num_cols =150;
w_s_cols=[1 4];
w_f_cols=[2 3 5 6];
format=[];
for I=1:num_cols
    if any(I==w_s_cols)
        format=[format '%s'];
    elseif any(I==w_f_cols)
        format=[format '%f'];
    else
        format=[format '%*s'];
    end
end
f1 = fopen('M3Comp.csv');
% data read from csv - from the international journal of
%forecasting data base
%https://forecasters.org/resources/time-series-data
hist=textscan(f1,format,'delimiter',',','HeaderLines',1);
fclose(f1);

[p,~] = size(hist{1});
display(p);

data = csvread('M3Comp.csv',1,6);
```

```
sp = 1;
per = 12/sp;


for z=1:1


    tm = hist{2}(z);
    fm = hist{3}(z);
    fms = fm/sp;
    mon = tm-fms;
    spt = mon/sp;
    ad = data(z,1:tm);
    d = zeros(spt,1);
    ac = zeros(fms,1);
    f = zeros(fms,1);
    fd = zeros(spt,1);
    yrs = floor(mon/12);
    tyrs = 2;

    for k = 1:mon
        j = ceil(k/sp);
        d(j) = d(j)+ ad(k);
    end
    for k = 1:fms
        j = ceil(k/sp);
        ac(j) = ac(j)+ ad(k+mon);
    end



    t = zeros(spt,1);
    si = zeros(spt+per,1);
    sa = zeros(per,yrs);
    yra = zeros(yrs,1);
    yrt = zeros(yrs,1);
```

```
fd(1) = d(1);
tr =0;

for i = 1:per
    tr = tr + ((d(i+per) - d(i))/per);
end

t(1) = tr/per;

for i =1:12*yrs
    y = ceil(i/12);
    yrt(y) = yrt(y) + ad(i);
    m = rem(i,12);
    if m==0
        m=12;
    end
    j = ceil(m/sp);
    sa(j,y) = sa(j,y) + ad(i);

end

for i = 1:yrs
    yra(i) = yrt(i)/per;
    if(yra(i)==0)
        e_flag =1;
        yra(i)=1;
    end
end

for i =1:per;
    for j = 1:yrs
        si(i) =  si(i) +(sa(i,j)/yra(j));
    end
    if(si(i)==0)
        e_flag=1;
        si(i) = 1;
```

```
    else
        si(i) = si(i)/yrs;
    end
end

[a, b, c, mape] = calcabc(spt,per,d,fd,t,si);



if (fd(1)==0)
    si(1+per) = 1;
    e_flag = 1;
else
    si(1+per) = c * d(1)/fd(1) + (1-c) * si(1);
end

for i = 2:spt
 fd(i) = a * d(i)/si(i) + (1-a)*(fd(i-1)+t(i-1));
 t(i) = b * (fd(i)- fd(i-1)) + (1-b)* t(i-1);
 if (fd(i)==0 || d(i)==0)
    e_flag = 1;
    si(i+per) = 1;
 else
    si(i+per) = c * d(i)/fd(i) + (1-c) * si(i);
 end
end

if ( fd(spt)<0)
    fd(spt) =10;
    t(spt) =0;
    e_flag=1;
elseif((fd(spt) + fms*t(spt))<0)
    t(spt) = (-fd(spt)+10)/lag;
    e_flag=2;
end

for i = 1:fms
```

```
            k = rem(i,per);
            if(k==0)
                k=12;
            end
        f(i) = (fd(spt) + i*t(spt))* si(spt+k);
        if f(i)<0
              e_flag=1;
              f(i) = 0;
        end
    end

    name = char(hist{1}(z));
    type = char(hist{4}(z));

    fid = fopen('Full_M3_Data.csv','w');
    % use repmat to construct repeating formats
    % ( numColumns-1 because no comma on last one)
    fprintf(fid,('%s,%s,'),name,type);
    numFmt = repmat('%f,',1,spt-1);
    fprintf(fid,[numFmt,'%f'],d);
    numFmt = repmat('%f,',1,fms-1);
    fprintf(fid,[numFmt,'%f\n'],f);
    fclose(fid);

    fid = fopen('M3_WH_Forecast.csv','w');
    % use repmat to construct repeating formats
    % ( numColumns-1 because no comma on last one)
    fprintf(fid,('%s,'),name);
    fprintf(fid,('%f,'),hist{3}(z));
    numFmt = repmat('%f,',1,fms-1);
    fprintf(fid,[numFmt,'%f\n'],f);
    fclose(fid);

end
```

```
% Code to smooth level trend and seasonality parameters
function [a,b,c,min] = calcabc(spt,per,d,fd,t,si)

    min = 1000000000;
    a =0;
    b =0;
    c =0;
    fs = zeros(spt);

p =0.05;
q =0.05;
r =0.05;

x = 0;


while x<1.01
y=0;
while y<1.01
z=0;
while z<1.01

if (fd(1)<=0)
si(1+per) = 1;
else
si(1+per) = z * d(1)/fd(1) + (1-z) * si(1);
end
fs(1) = fd(1)*si(1);

for i = 2:spt
fd(i) = x * d(i)/si(i) + (1-x)*(fd(i-1)+t(i-1));
t(i) = y * (fd(i)- fd(i-1)) + (1-y)* t(i-1);
if (fd(i)<=0 || d(i)<=0)
si(i+per) = 1;
else
si(i+per) = z * d(i)/fd(i) + (1-z) * si(i);
```

```
end
fs(i) = (fd(i-1)+t(i-1))*si(i);
end

mape = calmap(d,fs,spt);

if(mape < min + 0.00001)
min = mape;
a = x;
b = y;
c = z;
end

z = z + r;
end
y = y + q;
end
x = x + p;
end

end



% Code to run modified holt-winters triple exponential
% smoothing data from Hist

e_flag=0;
num_cols = 150;
w_s_cols=[1 4];
w_f_cols=[2 3 5 6];
format=[];
for I=1:num_cols
    if any(I==w_s_cols)
        format=[format '%s'];
    elseif any(I==w_f_cols)
        format=[format '%f'];
```

```
        else
            format=[format '%*s'];
        end
    end
    f1 = fopen('M3Comp.csv');
    hist=textscan(f1,format,'delimiter',',','HeaderLines',1);
    fclose(f1);

    [p,~] = size(hist{1});
    display(p);

    data = csvread('M3Comp.csv',1,6);

    sp = 1;
    per = 12/sp;
    lag =6;



    for z=1:p


        tm = hist{2}(z);
        fm = hist{3}(z);
        fms = fm/sp;
        mon = tm-fms;
        spt = mon/sp;
        ad = data(z,1:tm);
        d = zeros(spt,1);
        ac = zeros(fms,1);
        f = zeros(fms,1);
        fd = zeros(spt,1);
        yrs = floor(mon/12);
        tyrs = 2;

        for k = 1:mon
```

```
    j = ceil(k/sp);
    d(j) = d(j)+ ad(k);
end
for k = 1:fms
    j = ceil(k/sp);
    ac(j) = ac(j)+ ad(k+mon);
end


for l=1:lag

    t = zeros(spt,1);
    si = zeros(spt+per,1);
    sa = zeros(per,yrs);
    yra = zeros(yrs,1);
    yrt = zeros(yrs,1);

    fd(1) = d(1);
    tr =0;

    for i = 1:per
        tr = tr + ((d(i+per) - d(i))/per);
    end

    t(1) = tr/per;

    for i =1:12*yrs
        y = ceil(i/12);
        yrt(y) = yrt(y) + ad(i);
        m = rem(i,12);
        if m==0
            m=12;
        end
        j = ceil(m/sp);
        sa(j,y) = sa(j,y) + ad(i);
```

```
    end

    for i = 1:yrs
        yra(i) = yrt(i)/per;
        if(yra(i)==0)
            e_flag =1;
            yra(i)=1;
        end
    end

    for i =1:per;
        for j = 1:yrs
            si(i) =  si(i) +(sa(i,j)/yra(j));
        end
        if(si(i)==0)
            e_flag=1;
            si(i) = 1;
        else
            si(i) = si(i)/yrs;
        end
    end
[a, b, c, mape_lag] = calcabc_whl(spt,per,d,fd,t,si,l);


    if (fd(1)==0)
      si(1+per) = 1;
      e_flag = 1;
    else
      si(1+per) = c * d(1)/fd(1) + (1-c) * si(1);
    end



  for i = 2:spt
      fd(i) = a * d(i)/si(i) + (1-a)*(fd(i-1)+t(i-1));
      t(i) = b * (fd(i)- fd(i-1)) + (1-b)* t(i-1);
      if (fd(i)<=0 || d(i)<=0)
```

```
            e_flag = 1;
            si(i+per) = 1;
        else
            si(i+per) = c * d(i)/fd(i) + (1-c) * si(i);
        end
end


    if ( fd(spt)<0)
        fd(spt) =10;
        t(spt) =0;
        e_flag=1;
    elseif((fd(spt) + lag*t(spt))<0)
        t(spt) = (-fd(spt)+10)/lag;
        e_flag=2;
    end

    for i = 1:fms
            k = rem(i,per);
            if(k==0)
                k=12;
            end
        f(i) = (fd(spt) + i*t(spt))* si(spt+k);
            if f(i)<0
                e_flag=1;
                f(i) = 0;
            end
    end

    name = char(hist{1}(z));
    type = char(hist{4}(z));
    if l==1

        fid = fopen('L1_WHL.csv','a');
        fprintf(fid,('%s,'),name);
        fprintf(fid,('%f,'),hist{3}(z));
```

```
    numFmt = repmat('%f,',1,fms-1);
    fprintf(fid,[numFmt,'%f\n'],f);
    fclose(fid);

elseif l==2

    fid = fopen('L2_WHL.csv','a');
    fprintf(fid,('%s,'),name);
    fprintf(fid,('%f,'),hist{3}(z));
    numFmt = repmat('%f,',1,fms-1);
    fprintf(fid,[numFmt,'%f\n'],f);
    fclose(fid);

elseif l==3

    fid = fopen('L3_WHL.csv','a');
    fprintf(fid,('%s,'),name);
    fprintf(fid,('%f,'),hist{3}(z));
    numFmt = repmat('%f,',1,fms-1);
    fprintf(fid,[numFmt,'%f\n'],f);
    fclose(fid);

elseif l==4

    fid = fopen('L4_WHL.csv','a');
    fprintf(fid,('%s,'),name);
    fprintf(fid,('%f,'),hist{3}(z));
    numFmt = repmat('%f,',1,fms-1);
    fprintf(fid,[numFmt,'%f\n'],f);
    fclose(fid);

elseif l==5

    fid = fopen('L5_WHL.csv','a');
    fprintf(fid,('%s,'),name);
    fprintf(fid,('%f,'),hist{3}(z));
```

```
    numFmt = repmat('%f,',1,fms-1);
    fprintf(fid,[numFmt,'%f\n'],f);
    fclose(fid);

elseif l==6

    fid = fopen('L6_WHL.csv','a');
    fprintf(fid,('%s,'),name);
    fprintf(fid,('%f,'),hist{3}(z));
    numFmt = repmat('%f,',1,fms-1);
    fprintf(fid,[numFmt,'%f\n'],f);
    fclose(fid);

elseif l==7

    fid = fopen('L7_WHL.csv','a');
    fprintf(fid,('%s,'),name);
    fprintf(fid,('%f,'),hist{3}(z));
    numFmt = repmat('%f,',1,fms-1);
    fprintf(fid,[numFmt,'%f\n'],f);
    fclose(fid);

elseif l==8

    fid = fopen('L8_WHL.csv','a');
    fprintf(fid,('%s,'),name);
    fprintf(fid,('%f,'),hist{3}(z));
    numFmt = repmat('%f,',1,fms-1);
    fprintf(fid,[numFmt,'%f\n'],f);
    fclose(fid);

elseif l==9

    fid = fopen('L9_WHL.csv','a');
    fprintf(fid,('%s,'),name);
    fprintf(fid,('%f,'),hist{3}(z));
```

```
            numFmt = repmat('%f,',1,fms-1);
            fprintf(fid,[numFmt,'%f\n'],f);
            fclose(fid);

        elseif l==10

            fid = fopen('L10_WHL.csv','a');
            fprintf(fid,('%s,'),name);
            fprintf(fid,('%f,'),hist{3}(z));
            numFmt = repmat('%f,',1,fms-1);
            fprintf(fid,[numFmt,'%f\n'],f);
            fclose(fid);

        elseif l==11

            fid = fopen('L11_WHL.csv','a');
            fprintf(fid,('%s,'),name);
            fprintf(fid,('%f,'),hist{3}(z));
            numFmt = repmat('%f,',1,fms-1);
            fprintf(fid,[numFmt,'%f\n'],f);
            fclose(fid);

        elseif l==12

            fid = fopen('L12_WHL.csv','a');
            fprintf(fid,('%s,'),name);
            fprintf(fid,('%f,'),hist{3}(z));
            numFmt = repmat('%f,',1,fms-1);
            fprintf(fid,[numFmt,'%f\n'],f);
            fclose(fid);
        end
    end
end
```

```
% Code for smoothing parameters in the modified
% version of Holt-Witners algorithm
function [a,b,c,min] = calcabc_whl(spt,per,d,fd,t,si,lag)

    min = 1000000000;
    a =0;
    b =0;
    c =0;
    fs = zeros(spt);



    p =0.05;
    q =0.05;
    r =0.05;

    x = 0;



while x<1.01
y=0;
while y<1.01
z=0;
while z<1.01

if (fd(1)<=0)
si(1+per) = 1;
else
si(1+per) = z * d(1)/fd(1) + (1-z) * si(1);
end

fs(1) = fd(1)*si(1);

for i = 2:spt-lag
fd(i) = x * d(i)/si(i) + (1-x)*(fd(i-1)+t(i-1));
t(i) = y * (fd(i)- fd(i-1)) + (1-y)* t(i-1);
if (fd(i)<=0 || d(i)<=0)
```

```
si(i+per) = 1;
else
si(i+per) = z * d(i)/fd(i) + (1-z) * si(i);
end
fs(i+lag-1) = (fd(i-1)+lag*t(i-1))*si(i+lag-1);
end

fs(spt) = (fd(spt-lag)+lag*t(spt-lag))*si(spt);

mape = calmap(d(lag+1:spt),fs(lag+1:spt),spt-lag);

if(mape < min + 0.00001)
min = mape;
a = x;
b = y;
c = z;
end

z = z + r;
end
y = y + q;
end
x = x + p;
end
end

% calculating the MAPE error
function [mape] = calmap(d,fd,spt)

    wm = 0;
    for i = 1:spt
        m = 100 * abs((d(i) - fd(i))/d(i));
        wm = wm + m;
    end

    mape = wm/spt;
```

```
end

%Calculating accuracy between created file and results
%taken from database

e_flag=0;
num_cols = 150;
w_s_cols=[1 4];
w_f_cols=[2 3 5 6];
format=[];
for I=1:num_cols
    if any(I==w_s_cols)
        format=[format '%s'];
    elseif any(I==w_f_cols)
        format=[format '%f'];
    else
        format=[format '%*s'];
    end
end
f1 = fopen('M3Comp.csv');
hist=textscan(f1,format,'delimiter',',','HeaderLines',1);
fclose(f1);

[p,~] = size(hist{1});
display(p);

data1 = csvread('M3A.csv',0,2);
data2 = csvread('L12_WHL_MS.csv',0,2);
sp = 1;
per = 12/sp;
for z=1:p

    tm = hist{2}(z);
    tm = tm/sp;
    fm = hist{3}(z);
```

```
    fms = fm/sp;
    mon = tm-fms;
    spt = mon/sp;
    ad1 = data1(z,1:fms);
    ad2 = data2(z,1:fms);
    ac = zeros(fms,1);
    yrs = floor(mon/12);
    tyrs = 2;

    for i=1:fms
        ac = calc_smape(ad1,ad2,fms);
    end

    name = char(hist{1}(z));
    type = char(hist{4}(z));

    fid = fopen('Acc_L12_WHLMS.csv','a');
    fprintf(fid,('%s,%s,'),name,type);
    numFmt = repmat('%f,',1,fms-1);
    fprintf(fid,[numFmt,'%f\n'],ac);
    fclose(fid);

end

% calculating SMAPE
function [ac] = calc_smape(d,fd,fms)

ac = zeros(fms,1);
for i = 1:fms
ac(i) = 100 * (abs(d(i) - fd(i)))/(abs(d(i) + fd(i))/2);
end

end
```