FORECASTING DEMAND OF MAGAZINES AND MODELLING SEASONALITY

by

Abdullah Aktel

B.S., Industrial Engineering, Yıldız Technical University, 2005

Submitted to the Institute for Graduate Studies in
Science and Engineering in partial fulfillment of
the requirements for the degree of
Master of Science

Graduate Program in Industrial Engineering
Boğaziçi University

2008

Dedicated to
my family

ACKNOWLEDGEMENTS

I would like to gratefully acknowledge the enthusiastic supervision of Assoc. Prof. Wolfgang Hörmann. His orientation, guidance, insightful criticisms and patient encouragement contributed very much for the realization of this thesis.

I am deeply indebted to Mehper Cihangir Palavuzlar from Dergi Pazarlama Planlama A.Ş. for providing me all kinds of data eagerly and for their valuable contributions.

Special thanks go to Hakan Bekdaş and Baykar Silahlı for their support. I would like to thank Tubitak for providing scholarship during my master programme.

Finally, I wish to extend my warmest thanks to my family for their support.

ABSTRACT

FORECASTING DEMAND OF MAGAZINES AND MODELLING SEASONALITY

The main aim of this thesis is to find a sensible way to model the seasonality and forecast the demand of magazines automatically. Demand forecasting in magazine industry is very complex and historic delivery and sale data are often short, unstable and particularly perturbed by numerous factors. Generating forecasts from these large numbers of time series requires some degree of automation and simple forecasting models.

The first part of the thesis explains basic forecasting notions. Especially, the need for an automatic forecasting system is emphasized and the steps of automatic forecasting study are explained. Also, a statistical analysis is done to decide the suitable smoothing model alternatives. Finally, initialization and parameter optimization procedures are discussed. In the second part, demand estimation and handling of the censored demand in case of sellout is analyzed. In addition, the two main strategies used for planning are mentioned: top-down and bottom-up. The third part presents new forecasting methods based on combining forecasts and grouping similar characteristic endpoints by using real data. The last part explains the data organization and calculation of MAD and lost sales by using R and describes the some important algorithms that are used in magazine forecasting.

.

ÖZET

DERGİLERİN TALEP TAHMİNİ VE MEVSİMSELLİĞİN MODELLENMESİ

Bu tezin ana amacı, mevsimselliği modelleyebilmek için mantıklı bir yol bulmak ve dergilerin talebini otomatik olarak tahmin etmektir. Dergi endüstrisinde talep tahmini çok kompleks ve geçmiş satış ve sevk verileri genellikle kısa, kararsız ve özellikle de bi çok faktör tarafından bozulmuştur. Bu kadar geniş bir zaman serisinden tahminler üretmek otomasyon ve basit tahmin modelleri gerektirir.

Bu tezin ilk kısmı, temel tahminleme nosyonlarını açıklamaktadır. Özellikle, otomatik tahminleme sistemine duyulan ihtiyaç üzerinde durulmuş ve otomatik tahminleme çalışmasının basamakları açıklanmıştır. Ayrıca, uygun düzeltme modeli alternatiflerine karar verebilmek için bir istatistiksel analiz yapılmıştır. Son olarak, başlangıç ve parametre optimizasyonu tartışılmıştır. İkinci kısımda, talep tahmini ve yok satma durumunda görülemeyen gizli talebin kestirimi incelenmiştir. İlaveten planlamada kullanılan iki ana strateji olan yukarıdan aşağıya ve aşağıdan yukarı planlama açıklanmıştır. Üçüncü kısım tahminlerin birleştirilmesi ve benzer karakteristikteki son satıcı noktalarının gruplanmasına dayanan yeni tahminleme metotlarını gerçek veriler kullanılarak sunmaktadır. Son kısım ise veri organizasyonunu, en küçük mutlak sapmanın ve yok satmanın R programlama dili kullanılarak hesaplanmasını ve dergi tahminlemesinde kullanılan bazı önemli algoritmaları anlatmaktadır.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS /ABBREVATIONS

c	Printing cost per copy
c^{μ}	Change in mean
d	Demand
e_t	Forecast error
h	Cost of disposal of per unsold copy
I_t	Smoothed seasonal index or factor at the endpoint of period t.
m	Number of periods in the forecast lead-time
i	Index for the issue
j	Index for the endpoint
l	Lost sales
p	Sales price per copy
r	Return
S_t	Smoothed level of the series, computed after Y_t is observed
T_t	Smoothed trend at the end of period t
u	Price at which unsold copies are returned
y	Delivery
$\hat{y}$	Draft delivery
Y_t	Observed value of the time series in period t.
$\hat{Y}_t(m)$	Forecast for m periods ahead from origin t.
α	Smoothing parameter for the level of the series
β	Smoothing parameter for trend
γ	Smoothing parameter for seasonal factors
μ	Mean of demand
σ	Standard deviation of demand
τ	Parameter that determines service level provided
τ^*	Optimum value of τ
β_0	The level of the series
p	Number of periods in the seasonal cycle

AIC	The information criterion of Akaike
AR	Autoregressive model
ARMA	Autoregressive moving average model
ARIMA	Autoregressive integrated moving average model
MAD	Mean absolute deviation
MAPE	Mean absolute percentage error
MSE	Mean squared error
NORM	Normal distribution
SARIMA	Seasonal ARIMA model
SSE	Sum of squared error

1. INTRODUCTION

Many business and economic time series exhibit seasonal and trend variations. Seasonality is a periodic and recurrent pattern caused by factors such as weather, holidays, repeating promotions, as well as the behavior of economic agents. Although seasonal variations are perhaps the most significant component in a seasonal time series, a stochastic trend is often accompanied with the seasonal variations and can have a significant impact on various forecasting methods. In practice, standard forecasting methods that are designed to cope with seasonal demand often are not applicable.

Magazines have some characteristic features which are different from other perishable products. Each magazine issue can be thought as a different product because the magazines' cover, contents and promotions change from issue to issue. Due to the effects of promotions, special issues and seasonality the magazine demand is not stationary. It is widely accepted that magazine demand has mostly impulsive characteristic, because the demand is created by the available copies at the endpoint. The supply-demand balance is difficult to obtain. The accuracy of magazine forecasting is very important to prevent oversupply or undersupply. In Turkey, the magazine demand show seasonal characteristics and this makes it difficult to forecast the demand accurately.

"Dergi Pazarlama Planlama" (DPP) is the biggest magazine distribution planning company in Turkey and controls ¾ of the whole circulation. The company tries to find the optimum quantity of the circulation of approximately 150 magazines and distributes them to the almost 12.000 endpoints. Forecasting the demand of the endpoints requires a practical and effective forecasting system. Dergi Pazarlama Planlama A.Ş. wants to improve their forecasting system by modeling seasonality in the magazine demand. They want to decrease lost sales by using a better forecasting system. Although existing literature proposes valuable solutions for modeling perishable products' demands, there is no detailed work about modeling the seasonality in magazine demand.

The magazine planning process starts with the determination of the 'circulation' by the publisher. Past sales, number of endpoints to which the magazine is distributed, the special properties of the issue like cover, content, price and promotions are considered while determining the circulation. The magazine copies are distributed to the sales endpoints, after the planning is done. If the demand is less than the delivery amount, then the remaining copies are sent back and are called ‘returns’. If the actual demand exceeds the available number of copies sent, then a ‘sellout’ situation occurs and the difference between the demand and delivery constitutes the ‘lost sales’.

In the magazine planning literature, the main concern is the distribution planning of magazines. In the article of Bell (1978), a distribution procedure for periodicals is investigated. In the study of Artto and Pylkkänen (1999), an effective and practically applicable supply decision procedure for magazine single copy distribution is developed. In the work of Özgün (2006), magazine distribution planning system is explained by using Turkish magazine sale data. Özgün (2006) uses simple exponential smoothing as a forecasting method of magazine distribution planning system. However, forecasting the magazine demand accurately is an important part of the magazine planning procedure. So it is natural to ask if more sophisticated forecasting methods can improve the planning process.

A first obvious question is how to model seasonality. This has long been a major research topic that has significant practical implications. The choice of an appropriate method depends on the context, in particular on the objectives, the analyst’s skill, data characteristics, and number of series to be forecasted. There is a distinction between automatic and non-automatic forecasting systems. A simple automatic forecasting method, such as exponential smoothing, can be used when there are large numbers of series to be forecasted. However, if there are not so many time series traditional approaches such as classic decomposition methods, and seasonal autoregressive integrated moving average (SARIMA) models can be used to reach higher accuracy in forecasting seasonal time series.

In the article of Hyndman *et al.* (2002), a new approach to automatic forecasting based on an extended range of exponential smoothing methods is developed. Dekker *et al.* (2004), presents alternative forecasting methods that are based on using demand information from a higher aggregation level and on combining forecasts.

A good automatic forecasting method is also required in using best initialization methods and optimal smoothing parameters. Taylor (1981), examines initialization procedures proposed in the literature, analyzes their weighting of initial predictions and available historical data, and proposes a new initialization procedure. Gardner (1985), discusses class of exponential smoothing methods and explains the initialization methods of the each model.

The aim of this thesis is to find a sensible way to model the seasonality and forecast the demand of magazines automatically. Our aim is not to use the very best and most complicated model for forecasting each time series. Our goal is to provide a list of simple candidate models that will forecast the large majority of the time series well. We tried different approaches to model seasonality and forecast the demand accurately. We developed a new combined approach that compares different forecasting models' errors and selects the model for forecasting that gives minimum error. For organizing and forecasting large data effectively, we use the free programming language and statistical package, R. Specifically, the thesis provides suggestions about initialization and parameter selection for exponential smoothing and Holt-Winters.

In this thesis, special attention is given to real data. However, although we used real data for arriving at conclusions, we avoided developing company-specific solutions. Using real data always brings some inherent difficulties. In practice, there always exist problems with missing or wrong data, opened and closed endpoints and delays in obtaining information about sales. We give special attention to solve these difficulties.

This thesis is organized as follows. Chapter 2 explains the basic forecasting concepts and summarizes the smoothing techniques from the magazine forecasting perspective. Here smoothing techniques are stated in more detail where model formulations, properties, recommended parameters and starting values are discussed. Chapter 3 starts with the problem of censored demand. Next, a discussion on the demand distribution is included. In the last section, the distribution strategies are explained. Chapter 4 presents new forecasting methods based on combining forecasts and grouping similar characteristic endpoints by using real data. Chapter 5 explains the data organization by using R and describes some important algorithms that are used in magazine forecasting. Finally; conclusions, Appendix and references are set at the end of the thesis, respectively.

2. FORECASTING

It is clear that forecasting is an important part of the magazine distribution planning. Forecasting must be used to find the expected magazine demand. Forecasting may be defined as the prediction of future based on the past. Estimation of future events and conditions are called 'forecasts', and the act of making such estimations is called 'forecasting' (Bowerman e*t al.,* 2005). Since all organizations operate in an atmosphere of uncertainty, forecasts have always been necessary. Almost every organization needs forecasts. Forecasting is used in finance, marketing, personnel, and production areas, in government and profit-seeking organizations, in small social clubs, and in national political parties to determine goals and targets, to understand environment and uncertainties in the future and to make decisions about the level of risk (Hanke and Wichern, 2005). For a better understanding of the forecasting procedure, the basic forecasting notions are described.

2.1. Forecasting Steps

To understand a general forecasting study, the identification of forecasting steps is very important. The forecasting steps can be summarized as:

- Problem formulation and data collection
- Data manipulation and cleaning
- Model building and evaluation
- Model implementation (the actual forecast)
- Forecast evaluation

2.2. Forecasting Methods

We need quantitative forecasting methods in our thesis. Quantitative methods can be applied when sufficient information about the past is available in the form of numerical data. Qualitative forecasting methods are not useful for our problem.

2.2.1. Quantitative Forecasting Methods

Quantitative forecasting methods involve the analysis of historical data in an attempt to predict future values of a variable. In this thesis study, past sale and delivery amounts of the magazines are used to predict thousands of endpoints' future magazine demand. Quantitative forecasting models can be grouped into two categories: namely causal models and time series models. Forecasting many endpoints' demands forces us to use time series models as a forecasting model. So, causal models will not be explained.

2.2.1.1. Time Series Models. The aim of these forecasting models is to determine the pattern in the historical data series and extrapolate that pattern into the future in order to produce forecasts.

2.2.2. Automatic Forecasting

Demand forecasting in magazine industry is very complex and historic sale data are often short, unstable and particularly perturbed by numerous factors. Generating forecasts from these large numbers of time series requires some degree of automation. Although a skilled analyst can forecast a single time series by applying good judgment based on his or her knowledge and experience, the detailed modeling of individual series would not be cost-effective. Common problems are faced when forecasting large numbers of time series;

- No skilled analyst is available
- Many forecasts must be generated
- Frequent forecast updates are required

- Time-stamped data must be converted to time series data

- The forecasting model is not known for each time series.

Since we want to develop a fully automatic forecasting system for the magazine distribution planning process, we need an automatic forecasting system. Automatic forecasting describes a forecasting system, which does not require some initial specifications and a skilled user intervention, and includes only the input of an observed time series to generate a set of forecasts. Therefore, to forecast the large numbers of time series effectively an automatic forecasting system is required. Automatic forecasting is usually performed on each time series independently. For each time series for each candidate model, the parameter estimates (weights) should be optimized for best results. An automatic forecasting study is conceptually the same as a normal forecasting study. However, there are some special differences. Automatic forecasting study steps are the following:

<u>2.2.2.1. Accumulation Step</u>. For automatic forecasting, accumulation is the most important decision because the software makes most of the remaining decisions. Accumulation is namely, converting the time-stamped data into time series data is based on a particular frequency. Time-stamped data can be accumulated to form hourly, daily, weekly, monthly, or yearly time series. Additionally, the sum, mean, median, minimum, maximum, standard deviation, and other statistics can be used to accumulate the transactions within a particular time period.

<u>2.2.2.2. Diagnostic Step</u>. In this step, the potential list of candidate models are judged whether appropriate to a particular time series or not. For example the time series that have trends should be forecasted with models that have a trend component, or time series that are non-linear should be transformed for use with linear models, or time series that are intermittent should be forecasted with intermittent models etc…If it is known, a priori, that a time series has a particular characteristic, then the diagnostics should be overridden and the appropriate model should be used.

We will use the diagnostic step to decide the seasonality type and components of the smoothing model.

<u>2.2.2.3. Model Selection Step</u>. For automatic forecasting of large numbers of time series, only the most robust and simple models should be used. The goal is not to find the very best model for forecasting each time series. The goal is to provide a list of candidate models that will forecast the large majority of the time series well. After the candidate models have been identified by the diagnostics, each model is fit to the data with the holdout sample excluded by using the training set of the data. After model fitting, the multi-step-ahead forecasts are made in the test region for performance evaluation. One or any of the model selection criterions (*MSE, MAD, MAPE, AIC* etc…) can be used to select the best performing model from the candidate models.

We will use MAD as a model selection criterion.

<u>2.2.2.4. Forecasting Step</u>. After deciding the best forecasting model from candidate models, the selected model is fit to the full range of the data. Forecasts are made to determine the optimal model parameters and the best starting values. To do decision-making based on the forecasts, the analyst must decide whether to base the decision on the

predictions, lower confidence limits, upper confidence limits or the distribution (predictions and prediction standard errors).

2.2.2.5. Evaluation Step. After forecasting, the forecasts are compared with actual data. If the forecasts do not predict the actual data well, more detailed investigation should be done by the analyst.

2.2.2.6. Performance Step. After forecasting future periods, the actual data becomes available as time passes. After obtaining real data, forecasts can be judged according to some performance criteria. Some useful measures of forecast performance are for example, *MSE, MAD, MAPE and AIC* etc…Another useful measure of forecast performance is determining whether the newly available data falls within the previous forecasts' confidence limits.

We will use MAD and lost sales as a performance measure.

2.2.2.7. Implementation. There are many forecasting programs that can be used in a forecasting study. However, to do an automatic forecasting study, more advanced programs such as SAS, R should be used. Since R is a free and open source language, R is used in our automatic forecasting study.

2.3. Measuring Forecast Errors

Since all forecasting situations involve some degree of uncertainty, some error in forecasting must be expected. Denoting the actual value of the variable of interest at time t as y_t , the predicted value as $\hat{y}_t$ and subtracting the predicted value of $\hat{y}_t$ from the actual value y_t we can obtain the forecast error e_t. That is,

the forecast error for a particular forecast $\hat{y}_t$ is :

$$y_t - \hat{y}_t \tag{2.1}$$

There are two important methods to measure forecasting errors. These methods involve averaging some function of the difference between an actual value and its forecast value.

The mean squared error (*MSE*) is the most popular method for evaluating a forecasting technique. Each error is squared; then they are summed and divided by the number of observations.

$$MSE = \frac{1}{n} \sum_{t=1}^{n} \left(y_t - \hat{y}_t \right)^2 \tag{2.2}$$

Another method for evaluating a forecasting technique uses the average of the absolute errors. The mean absolute deviation (*MAD*):

$$MAD = \frac{1}{n} \sum_{t=1}^{n} \left| y_t - \hat{y}_t \right| \tag{2.3}$$

The *MSE* approach penalizes large forecasting errors because the errors are squared. Thus a technique that produces many moderate errors may well be preferable to one that usually has small errors but few very large errors. However, *MAD* is most useful when the analyst wants to measure forecast error in the same units as the original series. Deciding which error measure to use depends on the data structure. The magazine data cost structure is linear. When there is a sell out or return in the magazine forecasting system, the penalty is in the same unit for the distributor company. There is no need to punish the system with the square of the forecast error. Therefore, the *MAD* method is used in our thesis as a forecast measure because of the linear cost structure of the magazine demand.

2.4. Selection of the Forecasting Method

The selection of the forecasting method for a given situation is a decision problem. There are several criteria that are used to assess the competing methods. These are;

- The pattern of the data

- The time frame
- The cost of forecasting
- The accuracy desired
- The availability of data
- The ease of operation and understanding

Identification and understanding of the patterns in the data is an important factor influencing the selection of a forecasting technique. There are four types of data patterns: stationary, trend, seasonal, and cyclical. In a general forecasting study, many methods can be used to forecast these types of data. However, only automatic methods can be used to forecast magazine demand because we have too many time series to be forecasted. In such situations, naive methods, simple averaging methods, moving average models, and autoregressive moving average (ARMA) models can be used to forecast stationary series. Moving averages, Holt's linear exponential smoothing, and autoregressive integrated moving average (ARIMA) models can be used to forecast series with trend. Winter's exponential smoothing, and SARIMA models can be used to forecast seasonal series.

Since magazine demand patterns show changeable characteristic, the standard method that were explained above were not always suitable for our time series. It has been observed that the magazine demand characteristics change according to the endpoints. Although some endpoints show stationary characteristics, some of them show seasonal characteristics. As our thesis' requirements we need to select the most appropriate choice from a set of possibilities which are SARIMA and smoothing methods. It is widely accepted that SARIMA is the best choice being a linear model and capable of representing both stationary and non-stationary time series. However, our magazine data are neither long nor stable enough to apply SARIMA in our case. Therefore smoothing methods were used in our thesis to forecast magazines' demands.

2.5. Smoothing Methods

As smoothing methods have several practical advantages in short-range forecasting, these methods are widely used in industry. Model formulations are relatively simple and practical, model components and parameters have some intuitive meaning to the user and also only limited data storage is required. Smoothing methods can be classified according to non-seasonal or seasonal. And seasonal models can be grouped in to additive and multiplicative models. Although our magazine data show changeable characteristic according to the different endpoints, we will focus on the non-seasonal (simple exponential smoothing) and multiplicative seasonal models (i.e. multiplicative Holt-Winters model).

2.5.1. Simple Exponential Smoothing Method

If the mean (or the level) of a time series remains constant, then no trend model

$$y_t = \beta_0 + \varepsilon_t \text{ (or equivalently, } y_t = \mu + \varepsilon_t\text{)}$$

may be used to describe the data. Where ε_t is a random component with mean zero and variance σ^2_t. The level (β_0) is assumed to be constant in any local segment of the series but may change slowly over time. In simple exponential smoothing we use:

$$S_t = \alpha Y_t + (1-\alpha) S_{t-1} \tag{2.4}$$

$$\hat{Y}_t(m) = S_t \tag{2.5}$$

S_t is an unbiased estimator of the level as well as the forecast for any period ahead. The simple exponential smoothing method is used for forecasting a time series when there is no trend or seasonal pattern, but the mean (or level) of the time series y_t is slowly changing over time. In simple exponential smoothing, which is one special case of ARIMA models, namely ARIMA(0,1,1), the equal weighting scheme may not be appropriate when the mean of the time series is changing slowly over time. The simple exponential smoothing method will give the most recent observation the largest weight, instead of giving equal weights to each observation. This method is based on smoothing past values of a series in an exponentially decreasing manner. We can describe simple exponential smoothing by equation (2.4)

$$S_t = \alpha Y_t + (1-\alpha) S_{t-1} \tag{2.4}$$

Here α is a smoothing constant between zero and one. The value assigned to α is an important factor that effects forecasting. If it is desired that predictions are stable and random variations smoothed, a small value of α is required. If a rapid response to a real change in the pattern of observations is desired, a larger value of α is appropriate. There are two special α values which are α = 0, and α = 1. α = 0 means that the forecasted value equals to the first smoothed value. So, α = 1 means that the forecasted value equals to the last observation. Estimation of α is an iterative procedure that minimizes the mean squared error (*MSE*). Also an initial forecast (S_0) is required to start up an exponential smoothing forecasting system. The first observation or the mean of the observations can be used as an initial forecast (S_0). We will discus the selection of S_0 and α below.

2.5.2. Multiplicative Holt-Winters Method

In simple exponential smoothing the level of the time series is assumed to be changing occasionally. In some situations, the observed data have some other components, namely trend or seasonal pattern. If a time series has a linear trend with a fixed growth rate, β_1, and a fixed seasonal pattern, SN_t, with increasing (multiplicative) variation, the time series may be described by the multiplicative model

$$y_t = (\beta_0 + \beta_1 t) \times SN_t \times IR_t$$

Here IR_t is an irregular component, the level at time T-1 is $\beta_0 + \beta_1$ (T-1), and the level at time T is $\beta_0 + \beta_1 T$ showing that the growth rate for the level is β_1.

There are two types of Holt-Winters methods. The additive Holt-Winters method is used for time series with constant (additive) seasonal variation, whereas the multiplicative Holt-Winters method is used for time series with increasing (multiplicative) seasonal variation. Since the magazines' demands show increasing or decreasing seasonal variation according to their level, the multiplicative model will be used for forecasting in this study. The multiplicative Holt-Winters method is appropriate when a time series has a linear trend with a multiplicative seasonal pattern for which the level, growth rate, and the seasonal pattern may be changing rather then being fixed. To implement the multiplicative Holt-Winters method, we let S_{t-1} denote the estimate of the level in time t-1, and we let T_{t-1} denote the estimate of the growth rate in time t-1. Then, suppose that we observe a new time series value Y_t in time period t, and let I_{t-p} denote the "most recent" estimate of the seasonal factor for the season corresponding to time period t. Here p denotes the number of seasons in a year (p = 12 for monthly data), and thus t-p denotes the time period occurring one year prior to time period t. Furthermore, the subscript t-p of I_{t-p} denotes the fact that the time series value in time period t-p was the most recent time series value observed in the season being analyzed and thus the most recent time series value used to help find I_{t-p}. Then the multiplicative formulations are;

$$S_t = \alpha(Y_t / I_{t-p}) + (1-\alpha)(S_{t-1} + T_{t-1}) \tag{2.6}$$

$$T_t = \beta(S_t - S_{t-1}) + (1-\beta)T_{t-1} \tag{2.7}$$

$$I_t = \gamma(Y_t / S_t) + (1-\gamma)I_{t-p} \tag{2.8}$$

$$\hat{Y}_t(m) = (S_t + mT_t)I_{t-p+m} \tag{2.9}$$

For applying multiplicative Holt Winters method to our magazine data we had to add a fixed value, ten, which can be a different positive number, to all endpoints' demands because if an endpoint demand is zero then that endpoint future demand can not be calculated by using multiplicative Holt-Winters method. Therefore, a fixed value was added before the forecasting, and this added value was subtracted after the forecasting phase.

Up to now two smoothing methods were explained, simple exponential smoothing model and multiplicative Holt-Winters model, and these models can be used to forecast our magazine data. However some statistical analysis must be done using real magazine data to see which smoothing methods are more suitable to our data. After some analysis it was seen that the magazine data mostly fit to Simple Exponential Smoothing Method and No Trend Multiplicative Holt-Winters Method. You can find the codes of the following R implementation about the model selection in the Appendix.

2.5.3. Deciding The Suitable Smoothing Model Alternatives

This smoothing method selection analysis includes two steps, deciding about the seasonality type and deciding about the components of the models. Firstly, the seasonality type of the model has to be decided if it is additive or multiplicative. It is known that, a time series with a constant variation fits the additive model, time series with increasing or decreasing variation fits the multiplicative model. Knowing this information, we developed a function, *decide_multiplicative_additive*, which calculates the standard deviations and averages of each endpoint magazine sales and draws a graphic which gives us information about the model if it is multiplicative or additive. If the standard deviations of each endpoint magazine sales changes according to the averages then it can be concluded that these magazine data fit the multiplicative model; if not it can be concluded that these magazine data fit the additive model. For this purpose, data of the Hey Girl Magazine were used. By using the *decide_multiplicative_additive* function a graphic for Hey Girl Magazine data is drawn;

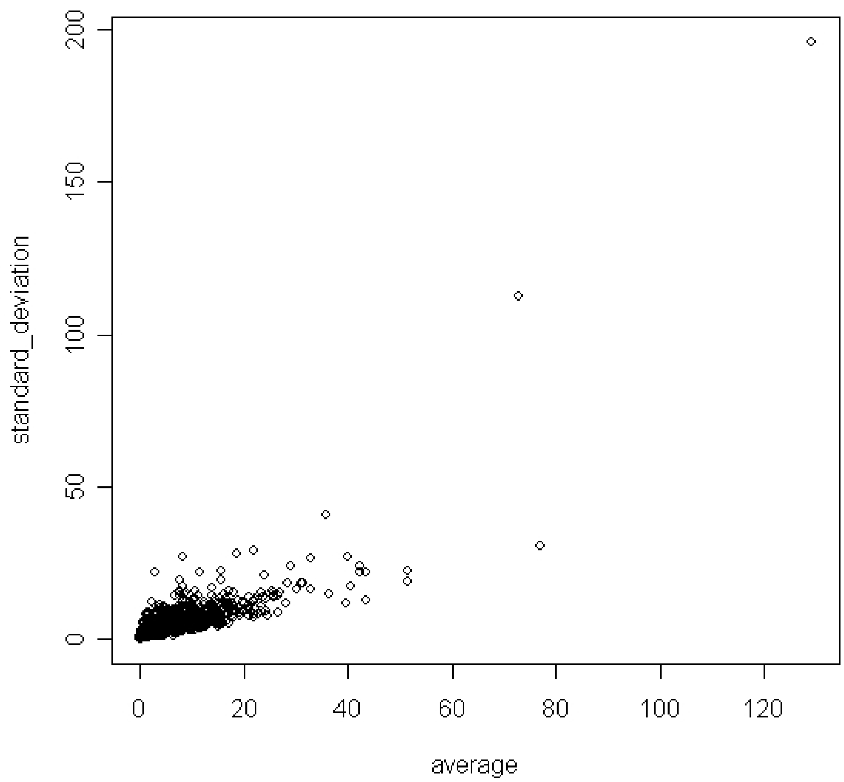


Figure 2.1. Hey Girl magazine seasonality type

And the ratio of standard deviation divided by average is, 1.05. As you can clearly understand from the graphic and the ratio, the standard deviations of the endpoints increase as the endpoints' averages increasing. We plotted this graph for different magazines, and a similar structure was seen for all magazines.

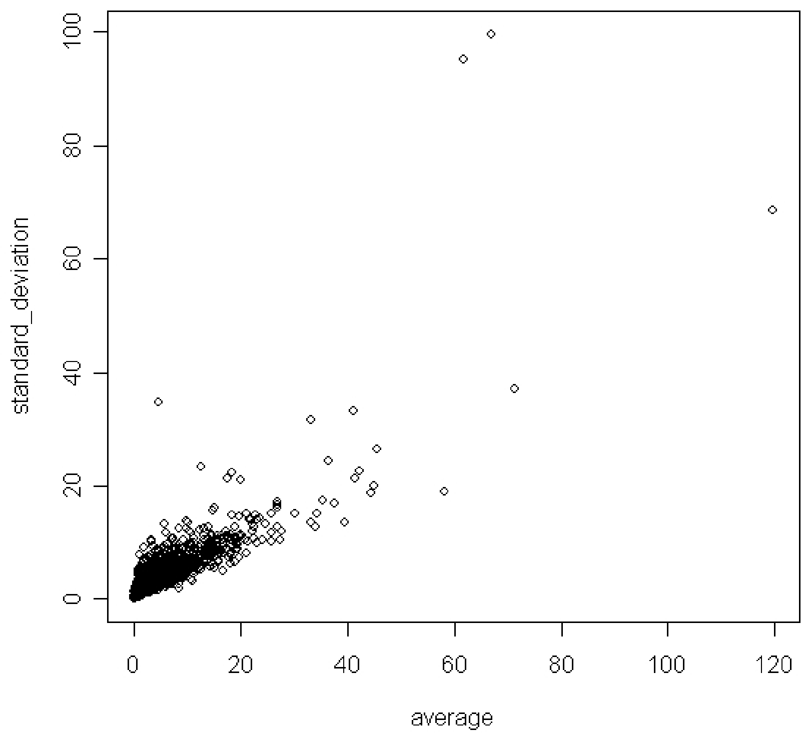


Figure 2.2. Blue Jean magazine seasonality type

And the ratio of standard deviation divided by average for Blue Jean magazine is, 1.02. Therefore, the multiplicative model is used in this forecasting study.

The second step is deciding the components of the forecasting model. To decide the model components the magazine time series are fitted to the multiplicative Holt-Winters model. If a magazine has any level, trend or seasonal component, then its level (α), trend (β) or seasonal (γ) smoothing constants have significant values. In the No trend model, β=0. By using *decide_smoothing_model* function, which draws the histograms of the smoothing coefficients, the time series are fitted to the multiplicative Holt-Winters model.

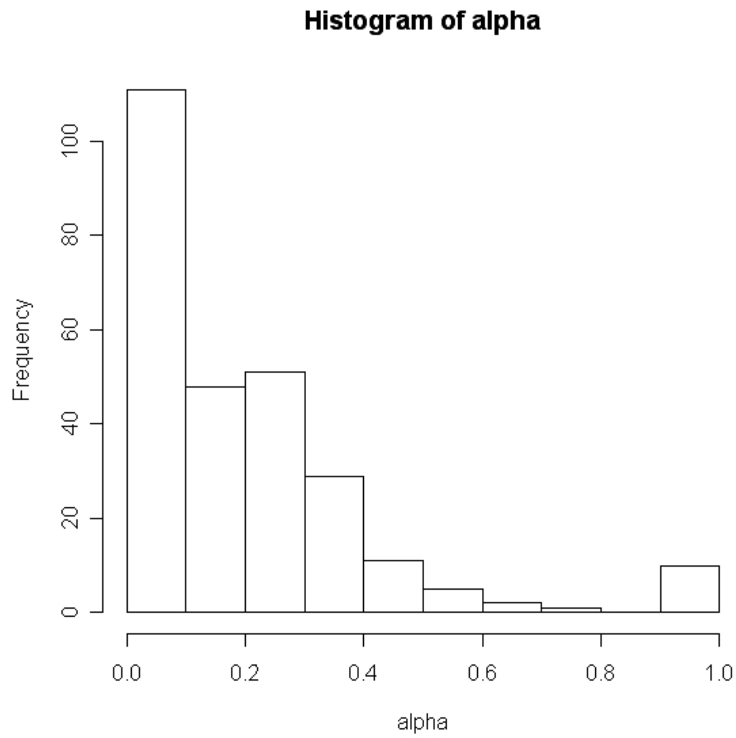


Figure 2.3. α smoothing coefficient of Hey Girl magazine

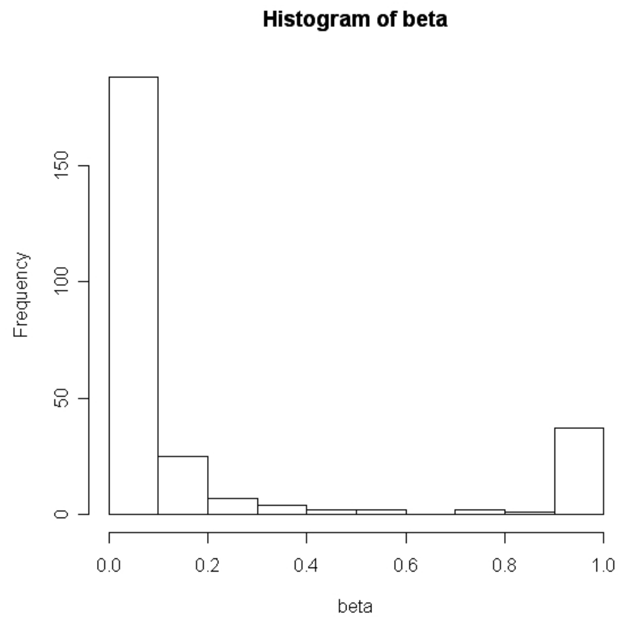


Figure 2.4. β smoothing coefficient of Hey Girl magazine

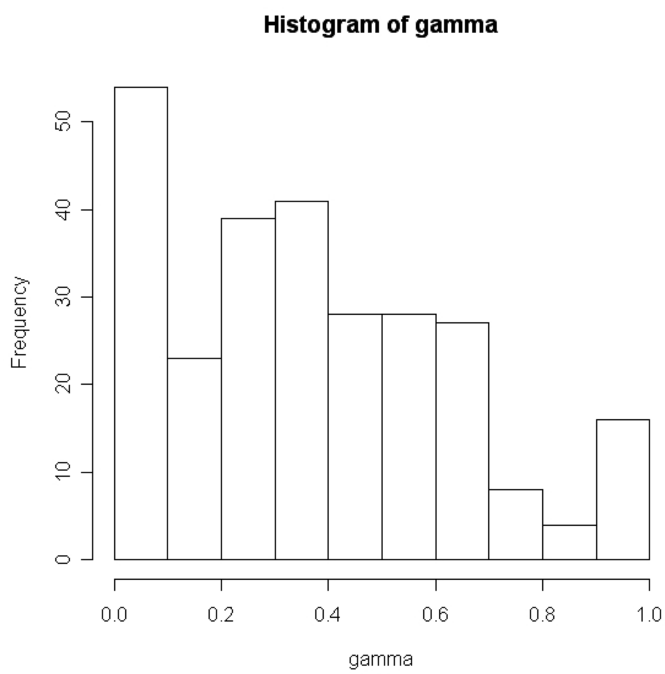


Figure 2.5. γ smoothing coefficient of Hey Girl magazine

As you can see in the graphics the magazine data have significant alpha (level) and gamma (seasonal) smoothing coefficient values, but the trend component (beta) is mainly close to zero. This experiment was done for different magazines and the results were similar. Therefore, Simple Exponential Smoothing Method and No Trend Multiplicative Holt-Winters Method are used as the two smoothing models in this forecasting study.

2.5.3.1. No Trend Multiplicative Holt-Winters Method. The only difference between the Multiplicative Holt-Winters Model and No Trend Multiplicative Holt-Winters Method is that $\beta=0$.

$$S_t = \alpha(Y_t / I_{t-p}) + (1-\alpha)S_{t-1} \tag{2.10}$$

$$I_t = \gamma(Y_t / S_t) + (1-\gamma)I_{t-p} \tag{2.11}$$

$$\hat{Y}_t(m) = S_t I_{t-p+m} \tag{2.12}$$

2.6. Initialization and Parameter Optimization Procedures

An initial forecast (S_0) is required to start up an exponential smoothing forecasting system. Procedures for initializing exponential smoothing forecasts have been developed by a number of researchers. There appears to be no empirical evidence favoring any particular method. Makridakis and Wheelwright (1978) suggests S_0 (initial forecast) $=Y_1$ (first observation). Brown (1962) suggests two approaches for initialization. The first approach applies when some historical data are available. For this case, simply using the mean of data for starting value is suggested. In situations, where no historical data are available, Brown recommends using a prediction as initial data point. Bowerman and O'Connell (1979) prefer the average of the first six observations. Ledolter and Abraham (1984) proposed backcasting to obtain the starting value, since this leads to the same forecasts as the ARIMA(0,1,1) model estimated by unconditional least squares. Backcasting is done by reversing the time order of the data and using the most recent data point to start the smoothing process.

$$S_j^* = \alpha Y_j + (1-\alpha) S_{j+1}^* \quad \text{where initially} \quad S_{n+1}^* = Y_n .$$

The last available smoothed value S_1^*, which is 'backforecast' of Y_0 from $Y_1,\ldots\ldots,Y_n$ can then be used as the initial S_0 to start the recursion.

It is apparent that the quality of forecasts derived from simple exponential smoothing will depend upon both the value S_0 used to initialize the procedure and upon the smoothing constant, α. The approach for selecting the smoothing constant usually proceeds by minimizing the one-step-ahead squared forecast errors:

$$SSE(\alpha) = \sum_{i=1}^{n} (Y_i - S_{i-1})^2$$

for the historic data. In the forecasting literature, SSE(α) is also referred to as the *ex-post* sum of squares function.

The choice of the particular initialization will influence the value of α which is obtained from the minimization of SSE(α), especially if the series is of short length. For example, a bad choice of S_0 will require a larger smoothing constant in order to discount bad initialization.

Cohen (1963) presented a Bayesian method of initializing time series forecasts. This method results in a forecast which is a blend of the initial prediction (prior estimate of the mean) and an average of the available time series data. The smoothing constant is dynamically adjusted as more data become available. The smoothing constant becomes a smoothing factor, A_t. For time period t, A_t is determined by

$$A_t = \frac{1}{(u+t)}$$

where u is a constant equal to σ^2_x/σ^2_e , in which σ^2_x is the variance of the time series data and σ^2_e is the variance of the prior estimate of the mean. If the Cohen's method is applied to a constant time series model, the resulting equation for the Bayesian estimate of the mean has the same form as an exponential smoothing model. (Taylor, 1981).

To obtain starting values for the additive seasonal factors, a linear regression on dummy variables can be used. Heuristic algorithms can be used for estimating starting values for both additive and multiplicative seasonality. Classical time series decomposition methods are more objective than the heuristics and require about the same computational effort. The seasonal factors from decomposition correspond directly to the I_t values in the seasonal models. One of the methods for estimating S_0 can then be applied to the deseasonalized data (Gardner Jr, 1985).

Although there are many initialization and parameter optimization procedures, we need to select the most appropriate choice from the set of possibilities. In our thesis, we mainly focused on two basic initialization methods which are using the first observation as an initial forecast or using the averages of the observations as an initial forecast to start forecasting. And the MAD is used in the parameter optimization phase as an error measure due to the linear cost structure of the magazine data. To find the best initialization, we did some statistical analysis by using simulated data. We compared the different initialization methods according to their forecast error as you can see below.

2.6.1. A Simulation Study About The Initialization and Parameter Optimization

In this simulation study it is aimed to find the best initialization method and the optimum smoothing parameters. In order to do this, different initialization methods were compared according to their forecast errors. When calculating forecast errors, the MAD (Mean Absolute Deviation) is used as an error measure. Before we discuss the results of the study, we explain a number of basic issues.

Firstly, the simulated data are created similar to our magazine data characteristic. Since magazine data show changeable characteristics according to the different endpoints, different models were used to create simulated data. The following models were used in this simulation study; Norm(), AR(1), SARIMA(0,1,1)(0,1,1)$_{12}$, SARIMA(1,0,0)(1,0,0)$_{12}$.

After creating the simulated data, two alternative initialization methods were compared to find the best method that will be used in the magazine forecasting. We mainly focused on two classic simple initialization approaches: using the first observation as an initial forecast or using the averages of the observations as an initial forecast to start forecasting. And also we developed an alternative methodology that combines the two approaches.

When finding forecast errors the two forecasting models were used, Simple Exponential Smoothing and No Trend Multiplicative Holt-Winters and we used MAD for the parameter optimization procedure by optimizing the smoothing parameters between zero and one.

In this simulation, data of length 58 were simulated and this simulation was repeated 500 times. The following table shows the results of this simulation study.

Table 2.1. Results of the simulation study for finding the best initialization method

Simulated Data Type	CMB Error	HW Error	EX Error	HW1 Error	HW2 Error	EX1 Error	EX2 Error
iid normal $\mu=5$	0.771	0.793	0.780	0.796	1.026	0.807	0.792
AR(1) $\rho=0.3$	0.800	0.815	0.812	0.815	1.054	0.827	0.818
AR(1) $\rho=0.9$	0.800	0.803	0.805	0.805	1.235	0.806	0.805
ARIMA(0,1,1)(0,1,1)$_{12}$ $w_1=0.5$, $w_2=0.5$	1.113	1.114	1.479	1.249	1.149	1.479	1.480
ARIMA(0,1,1)(0,1,1)$_{12}$ $w_1=0.4$, $w_2=0.6$	1.137	1.137	1.530	1.274	1.173	1.531	1.531
ARIMA(0,1,1)(0,1,1)$_{12}$ $w_1=0$, $w_2=0$	0.911	0.911	1.361	1.045	0.928	1.364	1.361
ARIMA(0,1,1)(0,1,1)$_{12}$ $w_1=0.1$, $w_2=0.8$	1.172	1.172	1.674	1.318	1.196	1.675	1.675
ARIMA(1,0,0)(1,0,0)$_{12}$ $\rho_1=0.3$, $\rho_2=0.7$	0.816	0.817	1.056	0.848	0.858	1.068	1.064
ARIMA(1,0,0)(1,0,0)$_{12}$ $\rho_1=0$, $\rho_2=0.8$	1.182	1.182	1.646	1.259	1.229	1.650	1.646
ARIMA(1,0,0)(1,0,0)$_{12}$ $\rho_1=0.5$, $\rho_2=0.5$	0.863	0.868	0.961	0.890	0.979	0.966	0.966

Some of the expression from the table should be explained to understand the results of this study better. HW1 Error means that the forecasting method of this alternative is No Trend Multiplicative Holt-Winters and it uses first value as an initial forecast whereas HW2 uses average of the observations as an initial forecast, EX1 Error means that the forecasting method of this alternative is Simple Exponential Smoothing and it uses first value as an initial forecast whereas EX2 uses average value of the observations. And also three other columns were added; HW Error, EX Error and CMB error. HW Error selects the minimum result from the HW1 Error and HW2 Error, and EX Error selects the minimum result from the EX1 and EX2. Finally, CMB (combined method) selects the minimum result from the HW Error and EX Error**.** The combined method works by comparing the forecast errors of the initialization methods according to their MAD and then selects the one which has minimum forecast error. For instance, if the smallest forecasting error of a simulated time series is EX1 Error then CMB method selects the initial value of this time series is first observation and forecasting model is simple exponential smoothing. In other words, CMB method is the combination of the initialization alternatives with minimum forecast error.

As you can see in the table, the results show us the combination of the initialization methods gives us the minimum error values for each simulated time series. The simulated time series almost reflect the real situation and the combined initialization approach better for all situations. The combined method is not only the best method for normal distribution but also the best method for trending or seasonal time series. Therefore, in the forecasting phase the combined initialization method will be used as an initialization method.

3. MAGAZINE DISTRIBUTION

3.1. Demand Estimation

To forecast the future magazine demand using past delivery and sales amount as a source of data is not enough. If the past sales data are directly used to forecast the demand of a future issue, the forecast will be biased low because we only know a lower bound for the actual demand. The real demand is censored, when all the copies sent to an endpoint are sold. This bias will be larger for endpoints having frequent stockouts. Therefore, the sales data should be adjusted that they reflect the expected demand without bias.

Bell (2000) adjusts the demand by replacing the sales figure with the conditional expectation if a sellout occurred. Denoting the endpoint *j*'s demand for issue *i* by d_{ij} and the number of copies delivered by y_{ij}, and if the demand distribution is discrete then, the expected demand in case of sellout is calculated as:

$$E[d_{ij} \mid d_{ij} \geq y_{ij}] = \frac{\sum_{y_{ij}}^{\infty} xf(x)dx}{\sum_{y_{ij}}^{\infty} f(x)dx} \tag{3.1}$$

where $f(x)$ is the density of the demand distribution of the endpoints for issue i.

3.1.1. A Discussion About The Magazine Demand Distribution

As customers usually decide to buy a magazine after seeing the cover, having a look at the articles inside, considering the promotions of the magazines, the demand has mostly impulsive characteristics. There will always be an extra demand due to the advertisements, news, recommendations or as a habitual behavior, which may be higher than the demand created by pure existence of magazine copies in the endpoint. Thus, demand is defined to be the number of copies that would be sold if enough number of copies were available. Because of the impulsive demand assumption, each customer can be thought to be exposed to a Bernoulli trial. Therefore, the demand distribution will be binomial, which can be approximated by the Poisson distribution if the probability to buy is small. Lots of studies claim that the demand observations follow approximately the Poisson distribution (Bell, 1978). However, for endpoints having larger amounts of sales and rare stockouts without trend, normal distribution can be used to approximate the demand distribution (Artto and Pylkkänen, 1998).

Due to the effects of promotions, special issues, price and seasonality, magazine sales are not stationary and sellouts are more frequent than we would expect from a stationary time series. The sales show variation from magazine to magazine, from city to city and from month to month. Because of the fact that demand is not stationary and many stockouts are observed, each issue should be thought as coming from a new demand distribution and only one observation per distribution is available. Therefore, it is impossible to fit a distribution to the sales data and test the assumptions about the demand distribution.

The Poisson assumption as endpoint demand distribution seems to be sensible. It is also advantageous that the Poisson distribution requires only a single parameter, namely the mean.

3.1.2. Estimating Future Demand

After estimating the past demands from the past sales data, these data can be used to forecast the parameters of a future issue's demand. Bell (1978) forecasts the mean ($\hat{\mu}_{ij}$) and standard deviation ($\hat{\sigma}_{ij}$) of the future issue's demand by exponential smoothing separately as he assumes normal distribution for the endpoint demand.

$$\hat{\mu}_{ij} = \alpha d_{i-n,j} + (1-\alpha)\hat{\mu}_{i-n,j} \tag{3.3}$$

$$M\hat{A}D_{ij} = \beta \left| d_{i-n,j} - \hat{\mu}_{i-n,j} \right| + (1-\beta) M\hat{A}D_{i-n,j} \tag{3.4}$$

where $\hat{\mu}_{ij}$ is the mean demand forecast for issue i, $d_{i\text{-}n,j}$ is the latest available (n periods before) actual (or adjusted, in case of sellout) demand figure, $\hat{\mu}_{i-n,j}$ is the mean demand forecast for the issue for which demand data are available, α and β are the smoothing constants of exponential smoothing and *MAD* is the mean absolute deviation of demand.

If, instead of normal distribution, Poisson distribution is assumed for endpoint demand, only the mean should be forecasted. In that case,

$$\hat{\mu}_{ij} = \alpha d_{i-n,j} + (1-\alpha)\hat{\mu}_{i-n,j} \tag{3.5}$$

$$\hat{\sigma}_{ij} = \sqrt{\hat{\mu}_{ij}} \tag{3.6}$$

Although Bell prefers exponential smoothing as the forecasting method due to its simplicity, we have also tried other forecasting methods such as Holt-Winters and ARIMA models to obtain better forecasting system.

3.2. Distribution Strategies

To plan the circulation of a magazine either bottom-up or top-down approach can be used. In the bottom-up strategy, after determining the number of copies to be delivered to each endpoint, the circulation is determined by adding up these values. In the top-down strategy, first the total demand is estimated from historical sales figures. After determining the circulation level by using demand and cost data, the total number of copies printed must be distributed to the endpoints (Özgün, 2006).

3.2.1. The Bottom-Up Approach

The Newsvendor problem is to choose the amount of copies sent to the outlet so that expected contribution is maximized (Artto and Pylkkänen, 1999). The magazine distribution planning problem is an application of the Newsvendor Problem. There the newsvendor buys a certain number of copies in the morning at a fixed price and returns the remaining copies at the end of the day at a lower price. It is assumed that the demand distribution for the newspaper is known. Then the problem is, to determine the optimum number of copies to buy in the morning.

$$\text{Number of Copies}= F^{-1}\left(\frac{p-c}{p-u}\right) \tag{3.7}$$

where p is the sales price of a copy, c is the buying cost and u is the price at which unsold copies are returned to distributor.

However, the classical Newsvendor Problem sees the costs from the newsvendor perspective. From the publisher's and distributor's perspective, the returned copies are not refunded. In addition, returning the copies causes extra handling and disposal cost. Instead of a discount for the returned copies ($-u$), we have an extra cost for the returns. Then, the optimum delivery amount becomes: $F^{-1}((p-c)/(p+h))$ where h is the variable cost of disposal of an unsold copy (including cost of handling, returning and storage minus their salvage value).

<u>3.2.1.1. The Optimum Circulation</u>. Bell (1978) showed that the optimum circulation is obtained by summing up the optimal amounts for all single endpoints. In a single endpoint the optimal amount leads to a probability of selling out equal to $(c + h) / (p + h)$.

Bell assumes that the demand in endpoint j follows a normal distribution with mean $\hat{\mu}_{ij}$ and standard deviation $\hat{\sigma}_{ij}$ which are estimated by exponential smoothing. The optimum number of copies to be sent to endpoint j is then determined as:

$$y_{ij} = \hat{\mu}_{ij} + \tau^{*}\hat{\sigma}_{ij} \tag{3.8}$$

where τ^{*} denotes the $(1- (c + h) / (p + h))$ quantile of the standard normal distribution.

Then, the optimal circulation is calculated as:

$$y_i = \sum_{j} y_{ij} \tag{3.9}$$

The solution of the Newsvendor problem can be applied to other demand distributions as well.

<u>3.2.1.2. The returns and the lost sales</u>. Although knowing optimum circulation is an important factor in the magazine planning, it is not enough. The two important measures namely return and lost sales concepts are frequently used in magazine industry. To be able to judge the sensitivity of the forecasting, knowing the expected amount of copies that will return and expected number of potential customers that will not be able to find a copy is critical for the decision of circulation.

If the delivered amount is higher than the demand, then this difference becomes 'return'. However, if the demand is higher than the delivery, then this creates a 'lost sale'

$$r_{ij}(\tau) = \max\left\{0, y_{ij}(\tau) - d_{ij}\right\} \tag{3.10}$$

$$l_{ij}(\tau) = \max\left\{0, d_{ij} - y_{ij}(\tau)\right\} \tag{3.11}$$

If $d_{ij} \neq y_{ij}$ return or lost sale occurs at an endpoint. This relation can be expressed as follows.

$$y_{ij}(\tau) - d_{ij} = r_{ij}(\tau) - l_{ij}(\tau) \tag{3.12}$$

3.2.2. The Top-Down Approach

In the top-down approach, determining the delivery amounts of endpoints in the given circulation is the basic problem. It should be known which endpoint will receive what percentage of the total number of printed copies to do this allocation. Using past sales is the most sensible way of determining these weights. Although it is the bottom-up approach that directly utilizes the demand forecasts of endpoints, also the top-down approach must use the demand forecasts. This means that, an automatic forecasting method should be used to have estimates of endpoints demands. After, this forecasted demand can be used to calculate the delivery. The steps can be summarized as;

- Forecasts the demands $(\hat{\mu}_j)$ of endpoints from past demands.

- Add "restitution margin" $(\tau\hat{\sigma}_j)$ to the forecasted demand to cover unexpected demand and calculate the draft deliveries $(\tilde{y}_j = \hat{\mu}_j + \tau\hat{\sigma}_j)$.

- Adjust the draft deliveries in order to equate their sum to the circulation amount.

- Calculate the actual deliveries (y_j).

4. FORECASTING MAGAZINE DEMAND AND CALCULATING LOST SALES

Up to now we explained magazine forecasting steps theoretically. This chapter proposes a new forecasting method based on the theoretical methods by using real data. While developing the method, it was tried to be as generic as possible.

The complete magazine planning process includes demand estimation, forecasting and the determination of the delivery amounts. The data set used here belongs to monthly popular Turkish magazines that have average circulation of around 25,000. They are distributed to more than 8,000 endpoints nationwide. To do forecasting experiments for each endpoint we need enough data for each endpoint. Therefore we calculated the sum of the first twelve and last twelve magazine sales. If these sums are greater than ten then the endpoint is selected for our data. After this filtering we have about 2000 remaining endpoints with 58 months of data. By using these sales, the planning of a future issue will be presented. We present the results using the Poisson distribution as demand distribution. By the help of the electronic versions of the files used in this chapter, which are available on the compact disc, the reader can experiment with different data parameters.

4.1. Forecasting Magazine Demand

Our aim is to develop a fully automatic forecasting system that can be used in forecasting magazine demand. The first step is the estimation of past demands.

4.1.1. Estimation Of Past Demands

The endpoint demands in case of sellout should be adjusted by using the formula given in Equation 3.1 for the Poisson distribution assumption. Since the Poisson assumption seems to be sensible as endpoint demand distribution, Equation 3.1 will be used to calculate past demands.

To estimate the past demands two time series are needed, the delivery and the sale data. Note that the parameters of the assumed distribution should be estimated for every endpoint. Moreover, these parameters should be updated when new data are available. For that reason, the demand estimation and parameter forecasting should be done simultaneously. Therefore, a function was developed to do these calculations namely *estimate_past_demands* which calls the function *condexp*.

The *estimate_past_demands* function starts by determining the initial values of the algorithm. The first demand and the first smoothed values are taken as the first sale values. Since the aim of these calculations is estimating past demands, we ignore where real sales' values are not available at the beginning of the time series. If there are some points for which sales data are not available (either because no copies are sent to that endpoint or due to data errors), the demand is taken as zero. If the sales were less than the delivery, then the sales amount are equal to the demand. But, if they are the same, the sales should be replaced by the conditional mean of the demand, given that the demand is greater than or equal to the number of copies of the sold issue. If the demand of endpoint *j* is assumed to follow the Poisson distribution, then the calculation of conditional expectation is done by the *condexp* function according to the Equation 4.1.

$$E\left[d_{ij} \mid d_{ij} \geq y_{ij}\right] = \frac{\mu_{ij} - \sum_{k=1}^{y_{ij}-1} k \cdot g(k, \mu_{ij})}{1 - G(k, \mu_{ij})} \tag{4.1}$$

where y_{ij} is the delivery, $g(k, \mu_{ij})$ is the probability density function of the Poisson distribution with the estimated mean μ_{ij} and $G(k, \mu_{ij})$ is the cumulative density function. You can find this function codes in the next chapters and you can find the txt file containing all results of the delivery, sales and demand data on the compact disc.

4.1.2. Forecasting Future Demand

Bell (1978) prefers simple exponential smoothing to forecast magazine demand for its minimum computation and data storage requirement. However, simple exponential smoothing is not useful to forecast time series which have seasonality and trend component. In order to make the forecast error minimal the smoothing coefficient should be optimized. Moreover, if the smoothing coefficients are optimized, and the best initialization methods are used to start forecasting the forecast errors may be reduced significantly.

For forecasting time series that have seasonality component with minimum forecasting error, seasonal forecasting methods should be used. Indeed, SARIMA is the best choice to forecast seasonal time series and is capable of representing both seasonality and non seasonality. In order to use SARIMA for this magazine data, our time series should be both long (at least 100 months) and stable. However, our data are neither long enough (58 months) nor stable enough to use SARIMA for forecasting. Hence, we decided to use the No Trend Holt-Winters method which is useful to get good forecasts for unstable and short series.

Our aim is to forecast the magazine demand with minimum error. There are two alternative ways to forecast the seasonal magazine data with minimum error. The first one is assuming that all endpoints are seasonal. Then No Trend Holt-Winters smoothing method should be used to forecast all endpoints. In addition, the smoothing coefficients of No Trend Holt-Winters method are optimized between zero and one to increase forecasting accuracy. The following table shows us the comparison of the No Trend Holt-Winters forecast results and Simple Exponential Smoothing forecast results.

Table 4.1. MAD tables of Simple Exponential Smoothing Method and No Trend Holt-Winters Method for all endpoints and for months 49-58

MAGAZINES	SIMPLE EXPONENTIAL SMOOTHING MAD	NO TREND HOLT-WINTERS MAD
HEY GIRL	1.554	1.801
BLUE JEAN	1.816	2.233
PC NET	2.064	2.172
FORM SANTE	1.663	1.849
ATLAS	2.030	2.420
ELLE	2.336	2.449
MAISON FRANCHAISE	1.968	2.129
BARBIE	1.918	2.170
BURDA	2.147	2.261
GO-GIRL	1.716	1.904
LEZZET	2.289	2.634
GOAL	1.612	1.768
İSTANBUL LIFE	2.297	2.606
MİLLİYET SANAT	1.880	2.056

As it is seen on the table, the MAD results of No Trend Holt-Winters method are larger than those of Simple Exponential Smoothing. The results indicate that we have to use another method as primitive Simple Exponential Smoothing performs better than No Trend Holt-Winters.

A second alternative is that all endpoints are classified depending on their sale trend which often increases especially in summer months or decreases in winter months. Since lots of seasonal endpoints show summer seasonal characteristics we decided to focus on summer seasonal endpoints. These endpoints are determined and a summer seasonality index between zero and one is assigned to each endpoint. If an endpoint summer seasonality index is close to one then this endpoint is defined to be “a summer endpoint” and No Trend Holt-Winters method is used to forecast these endpoints. However, this method has some subjective characteristics. The seasonality coefficients are user defined and although some endpoints have seasonal pattern, these endpoints may not be forecasted by seasonal models perfectly. In addition, the seasonality definitions may change from magazine to magazine and this prevents making a good forecasting system that works well all the time for all magazines. Instead of using this system as a main system it will be used as an assistant to our forecasting system and tried to develop a new combined method to be used in magazine forecasting.

4.1.2.1. Combined Approach. In this method, we combined the Simple Exponential Smoothing and No Trend Holt-Winters Methods to minimize the error. Before we explain the method, there are a number of basic issues that must be explained.

In this forecasting study, the magazine demand data (58 months) was separated into two sets namely, training and test sets, to check the developed method more effectively. If the data are not separated, the developed method has to be selected by using the all historic forecasts and this may mislead our future forecasts. Therefore, forecasting models are fitted to the data by using the training set (first 48 months of the data) and the forecast errors are produced by using the rest of the data as test set. Another important issue is the parameter optimization of the smoothing coefficients. In simple exponential smoothing, the smoothing coefficient is chosen a fixed value usually 0.2 and used for all the time series. However, if a parameter optimization is done in the training set of each time series, the optimal parameters can be used in the test set to forecast the demand of endpoints. Normally, the parameter optimization should allow all values between zero and one but we had to use smaller bounds for this optimization because some extreme parameter values which were found in the training set did not fit to test set and the forecasting error values were too high. Therefore, we had to restrict optimization to the interval (0.2, 0.5) to decrease the forecast error to reasonable levels. In addition, we had to use a coefficient between one and two to leverage the No Trend Holt-Winters Method forecasts in the method selection phase to increase our forecasting accuracy.

The combined approach works by comparing the forecast errors of the smoothing methods, according to their MAD and then selects the one which has minimum forecast error. These smoothing methods are: Simple Exponential Smoothing I, Simple Exponential Smoothing II and No Trend Holt-Winters Methods. The difference between Simple Exponential Smoothing I and II is their initialization method. In Simple Exponential Smoothing I, the initial value is the first observation whereas in Simple Exponential Smoothing II the initial value is the average value of observations. As we look at the results of the function (*cmprMAD*), the combined method gives almost for all magazines less error than Simple Exponential Smoothing.

Table 4.2. MAD tables of Simple Exponential Smoothing method and Combined method for all endpoints and for months 49-58

MAGAZINES	NUMBER OF ENDPOINTS	SIMPLE EXPONENTIAL SMOOTHING MAD	COMBINED METHOD MAD
HEY GIRL	1703	1.554	1.532
BLUE JEAN	1390	1.816	1.788
PC NET	1907	2.064	1.961
FORM SANTE	673	1.663	1.647
ATLAS	1743	2.030	2.032
ELLE	1262	2.336	2.269
MAISON FRANCHAISE	633	1.968	1.956
BARBIE	986	1.918	1.919
BURDA	1448	2.147	2.129
GO-GIRL	1334	1.716	1.689
LEZZET	726	2.289	2.218
GOAL	827	1.612	1.607
İSTANBUL LIFE	379	2.297	2.282
MİLLİYET SANAT	545	1.880	1.859

Although we reduced forecast error by using this combined method, still simple exponential smoothing gives better results for some magazines (ATLAS and BARBIE). There is neither parameter optimization nor modeling seasonality in Exponential Smoothing but it gives less or almost equal forecasting error for some magazines. The unstability in some endpoints seems to cause this situation. In addition, mainly endpoints with low average sale contribute to this situation. The combined approach learns the pattern in the time series and forecasts depending on this pattern. Since the unstable endpoints present a random structure, Bell's (1978) Simple Exponential Smoothing works better on these time series. To solve this problem we added another idea.

4.1.2.2. Grouping Similar Characteristic Endpoints. The magazine data are separated into three groups according to their average sales. The first group includes less than five average sales (low sales group), the second group includes between five and 10 average sales (medium sales group), the last group includes more than ten average sales (high sales group). Moreover, summer seasonality indexes are used in this combined method to increase the forecasting quality. After this grouping process, the combined approach is applied to the high sale group. However, before applying the combined approach on the medium and the low sales groups, the endpoints with high summer seasonality index are selected and the combined approach is applied to the selected group and simple exponential smoothing is applied to the rest of the endpoints which are the medium and the low sale with low summer seasonality index. The MAD results of the magazines for each group are given in the Table 4.3.

We used four years of the data and forecasted 10 months for 14 magazines to produce MAD values. All calculations for one magazine take approximately four minutes.

Table 4.3. Comparison of the MAD results of simple exponential smoothing method and combined method for three endpoint groups and for months 49-58

Hey Girl	**Number of Endpoints**	**Simple Exponential Smoothing MAD**	**Combined Method MAD**
High	239	3.472	3.254
Summer Medium	24	3.594	3.439
Summer Low	128	1.325	1.291
Average Error		2.777	2.623
Blue Jean	**Number of Endpoints**	**Simple Exponential Smoothing MAD**	**Combined Method MAD**
High	231	3.812	3.638
Summer Medium	16	3.563	3.777
Summer Low	65	1.530	1.540
Average Error		3.324	3.208
Pc Net	**Number of Endpoints**	**Simple Exponential Smoothing MAD**	**Combined Method MAD**
High	442	4.151	3.780
Summer Medium	3	2.251	3.376
Summer Low	16	1.241	1.240
Average Error		4.037	3.690
Form Sante	**Number of Endpoints**	**Simple Exponential Smoothing MAD**	**Combined Method MAD**
High	51	4.868	4.799
Summer Medium	5	5.993	3.463
Summer Low	53	1.696	1.609
Average Error		3.377	3.187

Table 4.3. Comparison of the MAD results of simple exponential smoothing method and combined method for three endpoint groups and for months 49-58 continue

Atlas	**Number of Endpoints**	**Simple Exponential Smoothing MAD**	**Combined Method MAD**
High	288	4.841	4.810
Summer Medium	435	0.107	0.097
Summer Low	7	12.632	12.646
Average Error		2.095	2.076
Elle	**Number of Endpoints**	**Simple Exponential Smoothing MAD**	**Combined Method MAD**
High	182	6.689	6.322
Summer Medium	19	5.271	4.047
Summer Low	82	1.509	1.385
Average Error		5.093	4.738
Maison Franchaise	**Number of Endpoints**	**Simple Exponential Smoothing MAD**	**Combined Method MAD**
High	92	4.865	4.807
Summer Medium	8	5.672	5.264
Summer Low	31	1.756	1.647
Average Error		4.179	4.087
Barbie	**Number of Endpoints**	**Simple Exponential Smoothing MAD**	**Combined Method MAD**
High	144	4.375	4.300
Summer Medium	254	0.115	0.110
Summer Low	57	1.310	1.360
Average Error		1.613	1.593
Burda	**Number of Endpoints**	**Simple Exponential Smoothing MAD**	**Combined Method MAD**
High	257	5.002	4.934
Summer Medium	2	2.140	1.925
Summer Low	35	1.282	1.264
Average Error		4.540	4.477
Go-Girl	**Number of Endpoints**	**Simple Exponential Smoothing MAD**	**Combined Method MAD**
High	131	4.542	4.205
Summer Medium	21	3.876	3.807
Summer Low	85	1.216	1.177
Average Error		3.290	3.084
Lezzet	**Number of Endpoints**	**Simple Exponential Smoothing MAD**	**Combined Method MAD**
High	89	5.974	5.495
Summer Medium	3	2.866	2.708
Summer Low	19	1.661	1.581
Average Error		5.152	4.750
Goal	**Number of Endpoints**	**Simple Exponential Smoothing MAD**	**Combined Method MAD**

Table 4.3. Comparison of the MAD results of simple exponential smoothing method and combined method for three endpoint groups and for months 49-58 continue

High	77	3.747	3.625
Summer Medium	7	3.483	3.232
Summer Low	39	1.360	1.309
Average Error		2.975	2.869
İstanbul Life	**Number of Endpoints**	**Simple Exponential Smoothing MAD**	**Combined Method MAD**
High	77	5.272	5.160
Summer Medium	5	4.341	4.284
Summer Low	27	1.405	1.344
Average Error		4.272	4.175
Milliyet Sanat	**Number of Endpoints**	**Simple Exponential Smoothing MAD**	**Combined Method MAD**
High	63	5.489924	5.323
Summer Medium	4	3.663	2.914
Summer Low	15	1.701	1.607
Average Error		4.707	4.526

As it is seen on the MAD comparison tables, the Combined Method average error values are less than the Simple Exponential Smoothing average error values for all magazines. The following table summarizes percentage reductions of MAD for all selected endpoints in the MAD values;

Table 4.4. Percentage reductions of combined method in the MAD values for all endpoints and for moths 49-58

MAGAZINE	NUMBER OF ENDPOINTS	IMPROVEMENT
HEY GIRL	1703	%2,2
BLUE JEAN	1390	%1,4
PC NET	1907	%4,0
FORM SANTE	673	%1,8
ATLAS	1743	%0,3
ELLE	1262	%3,4
MAISON FRANCHAISE	633	%0,9
BARBIE	986	%0,4
BURDA	1448	%0,5
GO-GIRL	1334	%2,1
LEZZET	726	%2,6
GOAL	827	%0,9
İSTANBUL LIFE	379	%1,2
MİLLİYET SANAT	545	%1,4

The forecasting error reduction changes from magazine to magazine and we get the maximum improvement in the Pc Net Magazine. If we look at the combined approach theoretically, this system is better than simple exponential smoothing. Although the improvement is not large, companies usually forecast many time series and even a small error reduction in their forecasting system increases their sales. Thus for big companies, sometimes even one per cent improvement can be important for their forecasting system. Before advising the combined method to the company we have to look at the influence on the expected lost sales.

4.2. Calculation Of Lost Sales

As we explained in chapter three, the bottom-up or the top-down approach can be used to find the deliveries. The top-down approach can be used easily to find the deliveries and also it follows a logical order. In practice, the circulation decision is taken by the publisher so most practitioners apply the top-down procedure to find deliveries. Therefore, the top-down approach was used to calculate the deliveries in this forecasting study.

After forecasting the future demand by using combined approach, the delivery amounts of endpoints for the future issue can be determined by using top-down approach. After forecasting the demand the restitution margin ($\tau\hat{\sigma}_j$) should be added that works like a buffer against the unexpected demand. Denoting the forecasted mean of endpoint *j* by $\hat{\mu}_j$, the standard deviation by $\sqrt{\hat{\mu}_j}$ and the parameter that shows the service level provided by τ, the draft delivery amount is calculated as: $\tilde{y}_j = \hat{\mu}_j + \tau\sqrt{\hat{\mu}_j}$

Due to the promotion, price or seasonality effects on the publisher's expectations, the sum of draft deliveries $(\sum \tilde{y}_j)$ may differ from the circulation that is determined independently by the publisher. So, the draft deliveries should be adjusted in order to equate their sum to the circulation amount.

4.2.1. Equating The Draft Delivery To The Circulation

To equate the draft delivery $\tilde{y}_j = \hat{\mu}_j + \tau\sqrt{\hat{\mu}_j}$ to the circulation, a coefficient (c^{μ}) is defined. This coefficient works as leverage to increase or decrease the total delivery. Thus the actual delivery is obtained by multiplying draft deliveries with the (c^{μ}) coefficient, using:

$$y_j = c^{\mu}\hat{\mu}_j + \tau\sqrt{c^{\mu}\hat{\mu}_j} \tag{4.2}$$

To find (c^{μ}), the R function, *uniroot* is used which searches for a root of a given intervals.

To see the performance of the combined method on the expected lost sales, we used simple exponential smoothing method as a benchmark. When we are comparing the lost sales of the simple exponential smoothing and combined method, we need to equate the delivery amounts of the two methods. Firstly, we calculated the magazine deliveries by using bottom-up approach for exponential smooth method. And sum of the deliveries constitutes the circulation amount. After finding the circulation amount, we calculated the draft deliveries by using top-down approach for combined method. After that, we adjust the draft deliveries in order to equate their sum to the circulation amount. Finally, we obtained the actual deliveries from equation 4.2.

After obtaining the actual deliveries (y_j), expected returns and lost sales can be calculated by using equation 3.10 and 3.11. The following table shows the delivery, return and lost sale values for magazines.

We used four years of the data and forecasted 10 months for 14 magazines to produce lost sales. All calculations for one magazine take approximately four minutes.

Table 4.5. Comparison of the simple exponential smoothing and combined method delivery, return and lost sale values for the three endpoint groups and for months 49-58.

Hey Girl	**Simple Exponential Smoothing Method**			**Combined Method**		
	Delivery	**Return**	**Lost Sales**	**Delivery**	**Return**	**Lost Sales**
High	33389	2431	870	33414	2468	765
Summer Medium	2241	204	169	2242	218	158
Summer Low	7159	608	264	7141	638	280
Total	42789	3243	1303	42797	3324	1203
Blue Jean	**Simple Exponential Smoothing Method**			**Combined Method**		
	Delivery	**Return**	**Lost Sales**	**Delivery**	**Return**	**Lost Sales**
High	36826	2693	551	36827	2918	548
Summer Medium	1427	143	63	1433	142	60
Summer Low	3698	358	113	3692	370	112
Total	41951	3194	727	41952	3430	720
Pc Net	**Simple Exponential Smoothing Method**			**Combined Method**		
	Delivery	**Return**	**Lost Sales**	**Delivery**	**Return**	**Lost Sales**
High	78232	5617	641	78226	5995	602
Summer Medium	182	20	8	184	15	3
Summer Low	760	68	24	768	76	24
Total	79174	5705	673	79178	6086	629
Form Sante	**Simple Exponential Smoothing Method**			**Combined Method**		

Table 4.5. Comparison of the simple exponential smoothing and combined method delivery, return and lost sale values for the three endpoint groups and for months 49- 58 continue

High	10499	671	373	10506	683	373
Summer Medium	432	56	80	431	39	56
Summer Low	2488	242	166	2505	243	176
Total	13419	969	619	13442	965	605
Atlas	**Simple Exponential Smoothing Method**			**Combined Method**		
	Delivery	**Return**	**Lost Sales**	**Delivery**	**Return**	**Lost Sales**
High	73153	4330	1549	73154	4419	1550
Summer Medium	44622	2986	691	44623	3051	689
Summer Low	1426	100	51	1431	111	42
Total	119201	7416	2291	119208	7581	2281
Elle	**Simple Exponential Smoothing Method**			**Combined Method**		
	Delivery	**Return**	**Lost Sales**	**Delivery**	**Return**	**Lost Sales**
High	49231	3338	1315	49247	3339	1194
Summer Medium	2437	258	213	2437	201	158
Summer Low	5388	461	207	5406	441	226
Total	57056	4057	1735	57090	3981	1578
Maison Franchaise	**Simple Exponential Smoothing Method**			**Combined Method**		
	Delivery	**Return**	**Lost Sales**	**Delivery**	**Return**	**Lost Sales**
High	20006	1354	335	19998	1364	334
Summer Medium	953	91	110	954	86	90
Summer Low	1887	167	103	1890	164	104
Total	22846	1612	548	22842	1614	528
Barbie	**Simple Exponential Smoothing Method**			**Combined Method**		
	Delivery	**Return**	**Lost Sales**	**Delivery**	**Return**	**Lost Sales**
High	24173	1827	643	24177	1924	635
Summer Medium	559	69	55	564	79	56
Summer Low	3263	275	87	3267	293	89
Total	27995	2171	785	28008	2296	780

Table 4.5. Comparison of the simple exponential smoothing and combined method delivery, return and lost sale values for the three endpoint groups and for months 49-58 continue

Burda	**Simple Exponential Smoothing Method**			**Combined Method**		
	Delivery	**Return**	**Lost Sales**	**Delivery**	**Return**	**Lost Sales**
High	130204	4127	2384	130200	4330	2370
Summer Medium	306	12	5	307	12	6
Summer Low	3319	196	46	3322	198	46
Total	133829	4335	2435	133829	4540	2422
Go-Girl	**Simple Exponential Smoothing Method**			**Combined Method**		
	Delivery	**Return**	**Lost Sales**	**Delivery**	**Return**	**Lost Sales**
High	39336	1832	948	39329	1848	839
Summer Medium	3285	222	223	3277	237	201
Summer Low	8833	480	240	8829	499	225
Total	51454	2534	1411	51435	2584	1265
Lezzet	**Simple Exponential Smoothing Method**			**Combined Method**		
	Delivery	**Return**	**Lost Sales**	**Delivery**	**Return**	**Lost Sales**
High	35382	1777	594	35383	1938	580
Summer Medium	511	26	24	512	30	19
Summer Low	2250	139	44	2246	144	55
Total	38143	1942	662	38141	2112	654
Goal	**Simple Exponential Smoothing Method**			**Combined Method**		
	Delivery	**Return**	**Lost Sales**	**Delivery**	**Return**	**Lost Sales**
High	22463	999	383	22460	1016	379
Summer Medium	1164	69	70	1165	71	66
Summer Low	3328	217	115	3340	214	117
Total	26955	1285	568	26965	1301	562
İstanbul Life	**Simple Exponential Smoothing Method**			**Combined Method**		
	Delivery	**Return**	**Lost Sales**	**Delivery**	**Return**	**Lost Sales**
High	33907	1320	699	33900	1386	681
Summer Medium	1009	59	66	1012	65	58

Table 4.5. Comparison of the simple exponential smoothing and combined method delivery, return and lost sale values for the three endpoint groups and for months 49-58 continue

Summer Low	2032	130	71	2026	137	75
Total	36948	1509	836	36938	1588	814
Milliyet Sanat	**Simple Exponential Smoothing Method**			**Combined Method**		
	Delivery	**Return**	**Lost Sales**	**Delivery**	**Return**	**Lost Sales**
High	29345	1097	488	29342	1117	490
Summer Medium	653	40	34	648	34	16
Summer Low	1056	75	71	1053	71	76
Total	31054	1212	593	31043	1222	582

As it is seen from these tables, the combined method lost sale values are less than primitive simple exponential smoothing lost sale values for all magazines. The combined approach not only reduced MAD values but also reduced lost sales. As it is understood from these results we may say that the combined approach is a practical and automatic way to forecast magazine demand reducing MAD and lost sales. The following table summarizes percentage reductions of the lost sales.

Table 4.6. Percentage reductions of combined method in the lost sale values for all endpoints and for months 49-58

MAGAZINE	NUMBER OF ENDPOINTS	IMPROVEMENT
HEY GIRL	1703	%5,0
BLUE JEAN	1390	%0,9
PC NET	1907	%3,8
FORM SANTE	673	%1,0
ATLAS	1743	%0,3
ELLE	1262	%5,8
MAISON FRANCHAISE	633	%2,1
BARBIE	986	%0,3
BURDA	1448	%0,3
GO-GIRL	1334	%4,9
LEZZET	726	%0,5
GOAL	827	%0,4
İSTANBUL LIFE	379	%1,5
MİLLİYET SANAT	545	%0,9

As it can be seen from the table we get 0,3-5,8 per cent reduction in the lost sales for the magazines. We can advise our combined method as a forecasting system for the distributor company

5. TECHNICAL REPORT ABOUT DATA ORGANIZATION AND R ALGORITHMS

This chapter gives information about magazine data organization with R and explains the main functions that were written to decide about the best forecasting techniques for magazine demand.

5.1. Data Organization With R

The real magazine data are so large. One aim of this thesis was to find a sensible data organization for magazine data. Since R is a statistical software and useful tool to handle data effectively, R program was used to do this data organization and to forecast the magazine demand. Hey Girl magazine is used as demonstration example.

Firstly, the magazine data must be imported into R. Magazine data include past delivery and sale information in a big matrix. Its rows show endpoint codes and its columns show months. To manipulate data, the Hey Girl magazine data were separated into three parts: Hey Girl Sent, Hey Girl Sold and Hey Girl Endpoint Vector. By using “read.table” command the data are imported into R. Hey Girl Sent matrix is stored in “sent”, Hey Girl Sold matrix is stored in “sold” and the Hey Girl magazine endpoint codes are stored in “HG_ENDPVECT”. We used:

```
sent<-read.table(file=file.choose(),header=TRUE)
sold<-read.table(file=file.choose(),header=TRUE)
HG_ENDPVECT<- read.table(file=file.choose(),header=TRUE)
```

To increase the calculations speed the transposes of both sent and sold matrixes are taken by using “t ( )” command and “sent” and “sold” matrixes are stored in “HG_SENT” and “HG_SOLD”. Using:

```
HG_SENT<-t(sent)
HG_SOLD<-t(sold)
```

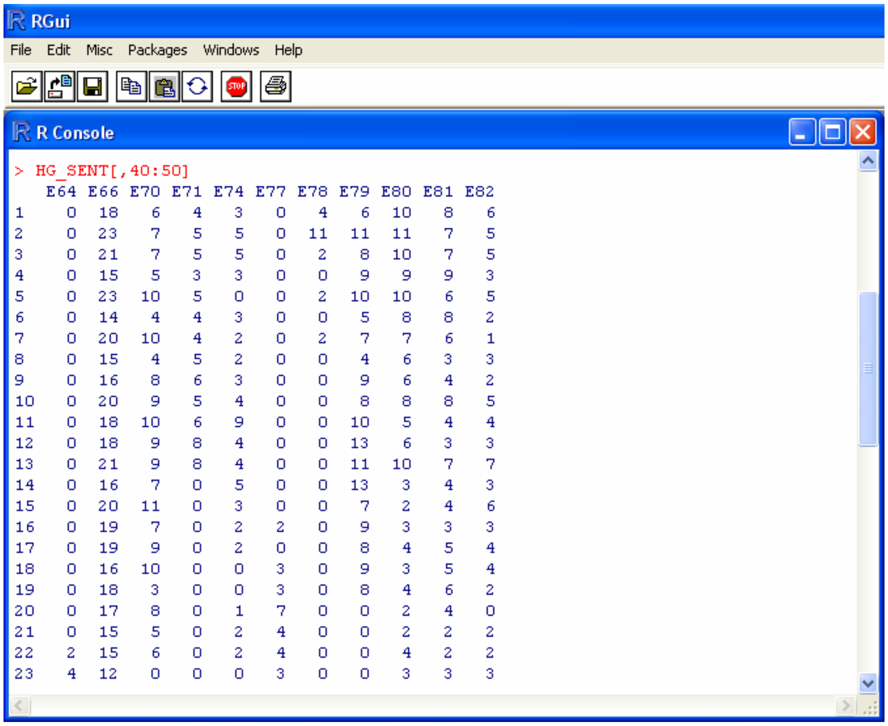
"HG_SENT" and "HG_SOLD" matrixes are changed into data frames to increase the effectiveness of our data base. Therefore, we used:

```
HG_SOLD<-data.frame(HG_SOLD)
HG_SENT<-data.frame(HG_SENT)
```

The "dimnames" command is used to assign endpoint codes to "HG_SOLD" and "HG_SENT" data frames' columns.

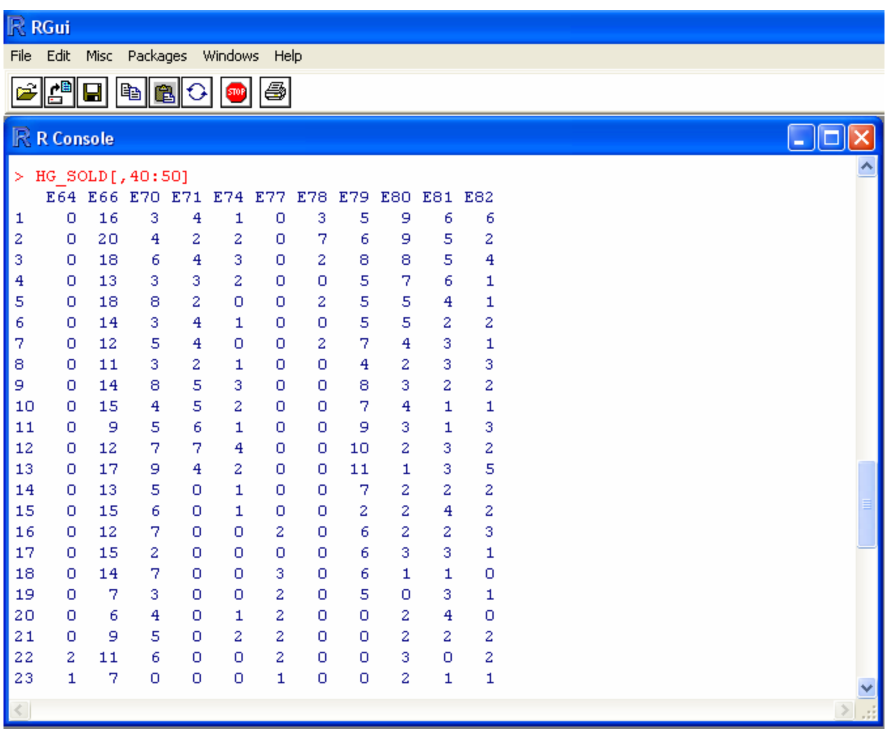
```
attributes(HG_SOLD)$dimnames<-
list(1:length(HG_SOLD[,1]),paste("E",HG_ENDPVECT[[1]][1:length(HG_ENDPVECT[[
1]])],sep=""))
attributes(HG_SENT)$dimnames<-
list(1:length(HG_SENT[,1]),paste("E",HG_ENDPVECT[[1]][1:length(HG_ENDPVECT[[
1]])],sep=""))
```

The following R windows show some endpoints delivery and sale values for Hey Girl Magazine.



	E64	E66	E70	E71	E74	E77	E78	E79	E80	E81	E82
1	0	18	6	4	3	0	4	6	10	8	6
2	0	23	7	5	5	0	11	11	11	7	5
3	0	21	7	5	5	0	2	8	10	7	5
4	0	15	5	3	3	0	0	9	9	9	3
5	0	23	10	5	0	0	2	10	10	6	5
6	0	14	4	4	3	0	0	5	8	8	2
7	0	20	10	4	2	0	2	7	7	6	1
8	0	15	4	5	2	0	0	4	6	3	3
9	0	16	8	6	3	0	0	9	6	4	2
10	0	20	9	5	4	0	0	8	8	8	5
11	0	18	10	6	9	0	0	10	5	4	4
12	0	18	9	8	4	0	0	13	6	3	3
13	0	21	9	8	4	0	0	11	10	7	7
14	0	16	7	0	5	0	0	13	3	4	3
15	0	20	11	0	3	0	0	7	2	4	6
16	0	19	7	0	2	2	0	9	3	3	3
17	0	19	9	0	2	0	0	8	4	5	4
18	0	16	10	0	0	3	0	9	3	5	4
19	0	18	3	0	0	3	0	8	4	6	2
20	0	17	8	0	1	7	0	0	2	4	0
21	0	15	5	0	2	4	0	0	2	2	2
22	2	15	6	0	2	4	0	0	4	2	2
23	4	12	0	0	0	3	0	0	3	3	3

Figure 4.1. Hey Girl magazine delivery values for the selected endpoints



	E64	E66	E70	E71	E74	E77	E78	E79	E80	E81	E82
1	0	16	3	4	1	0	3	5	9	6	6
2	0	20	4	2	2	0	7	6	9	5	2
3	0	18	6	4	3	0	2	8	8	5	4
4	0	13	3	3	2	0	0	5	7	6	1
5	0	18	8	2	0	0	2	5	5	4	1
6	0	14	3	4	1	0	0	5	5	2	2
7	0	12	5	4	0	0	2	7	4	3	1
8	0	11	3	2	1	0	0	4	2	3	3
9	0	14	8	5	3	0	0	8	3	2	2
10	0	15	4	5	2	0	0	7	4	1	1
11	0	9	5	6	1	0	0	9	3	1	3
12	0	12	7	7	4	0	0	10	2	3	2
13	0	17	9	4	2	0	0	11	1	3	5
14	0	13	5	0	1	0	0	7	2	2	2
15	0	15	6	0	1	0	0	2	2	4	2
16	0	12	7	0	0	2	0	6	2	2	3
17	0	15	2	0	0	0	0	6	3	3	1
18	0	14	7	0	0	3	0	6	1	1	0
19	0	7	3	0	0	2	0	5	0	3	1
20	0	6	4	0	1	2	0	0	2	4	0
21	0	9	5	0	2	2	0	0	2	2	2
22	2	11	6	0	0	2	0	0	3	0	2
23	1	7	0	0	0	1	0	0	2	1	1

Figure 4.2. Hey Girl magazine sale values for the selected endpoints

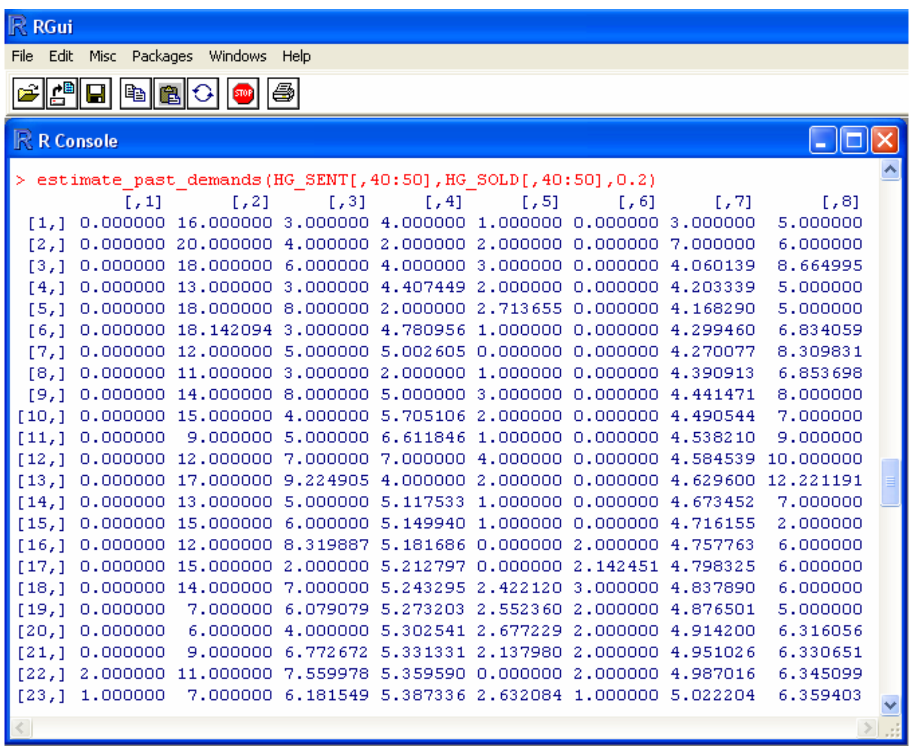
The magazine data must be filtered to start the forecasting. Therefore, HG_SENT and HG_SOLD data frames are filtered by using "*filter( )*" function.

```
filtre<-filter(HG_SOLD,12,47,58)
HG_SOLD<-HG_SOLD[,filtre]
HG_SENT<-HG_SENT[,filtre]
```

After filtering, the past demand is produced by using "*estimate_past_demands*" function and stored into HG_DEMAND.

```
HG_DEMAND<-estimate_past_demands(HG_SENT,HG_SOLD,0.2)
```

The following R window shows Hey Girl Magazine demand values for the same endpoints.




```
> estimate_past_demands(HG_SENT[,40:50],HG_SOLD[,40:50],0.2)
          [,1]      [,2]     [,3]     [,4]     [,5]     [,6]     [,7]      [,8]
 [1,] 0.000000 16.000000 3.000000 4.000000 1.000000 0.000000 3.000000  5.000000
 [2,] 0.000000 20.000000 4.000000 2.000000 2.000000 0.000000 7.000000  6.000000
 [3,] 0.000000 18.000000 6.000000 4.000000 3.000000 0.000000 4.060139  8.664995
 [4,] 0.000000 13.000000 3.000000 4.407449 2.000000 0.000000 4.203339  5.000000
 [5,] 0.000000 18.000000 8.000000 2.000000 2.713655 0.000000 4.168290  5.000000
 [6,] 0.000000 18.142094 3.000000 4.780956 1.000000 0.000000 4.299460  6.834059
 [7,] 0.000000 12.000000 5.000000 5.002605 0.000000 0.000000 4.270077  8.309831
 [8,] 0.000000 11.000000 3.000000 2.000000 1.000000 0.000000 4.390913  6.853698
 [9,] 0.000000 14.000000 8.000000 5.000000 3.000000 0.000000 4.441471  8.000000
[10,] 0.000000 15.000000 4.000000 5.705106 2.000000 0.000000 4.490544  7.000000
[11,] 0.000000  9.000000 5.000000 6.611846 1.000000 0.000000 4.538210  9.000000
[12,] 0.000000 12.000000 7.000000 7.000000 4.000000 0.000000 4.584539 10.000000
[13,] 0.000000 17.000000 9.224905 4.000000 2.000000 0.000000 4.629600 12.221191
[14,] 0.000000 13.000000 5.000000 5.117533 1.000000 0.000000 4.673452  7.000000
[15,] 0.000000 15.000000 6.000000 5.149940 1.000000 0.000000 4.716155  2.000000
[16,] 0.000000 12.000000 8.319887 5.181686 0.000000 2.000000 4.757763  6.000000
[17,] 0.000000 15.000000 2.000000 5.212797 0.000000 2.142451 4.798325  6.000000
[18,] 0.000000 14.000000 7.000000 5.243295 2.422120 3.000000 4.837890  6.000000
[19,] 0.000000  7.000000 6.079079 5.273203 2.552360 2.000000 4.876501  5.000000
[20,] 0.000000  6.000000 4.000000 5.302541 2.677229 2.000000 4.914200  6.316056
[21,] 0.000000  9.000000 6.772672 5.331331 2.137980 2.000000 4.951026  6.330651
[22,] 2.000000 11.000000 7.559978 5.359590 0.000000 2.000000 4.987016  6.345099
[23,] 1.000000  7.000000 6.181549 5.387336 2.632084 1.000000 5.022204  6.359403
```

Figure 4.3. Hey Girl magazine demand values for the selected endpoints

Now the past demand values can be used to decide best forecasting technique.

5.1.1. Calculation of MAD and Lost Sale Values

For applying our forecasting method it is necessary to decompose the endpoints into the groups "high", "summer medium", "summer low" and rest. By using the "*find_groups*" function the HG_DEMAND data are separated into three groups and stored in HG_DEMAND_10, HG_DEMAND_5_10 and HG_DEMAND_5. After that, to find the summer seasonal endpoints with medium and low sale values "*Summer_Ep*" function is used. After this grouping process, the MAD values and the lost sale values are calculated by using "*cmprMAD*", "*Lost_Sale_adj_Comparison*" functions. The following figure shows the Hey Girl magazine MAD, delivery, return and lost sale values for each endpoint groups.

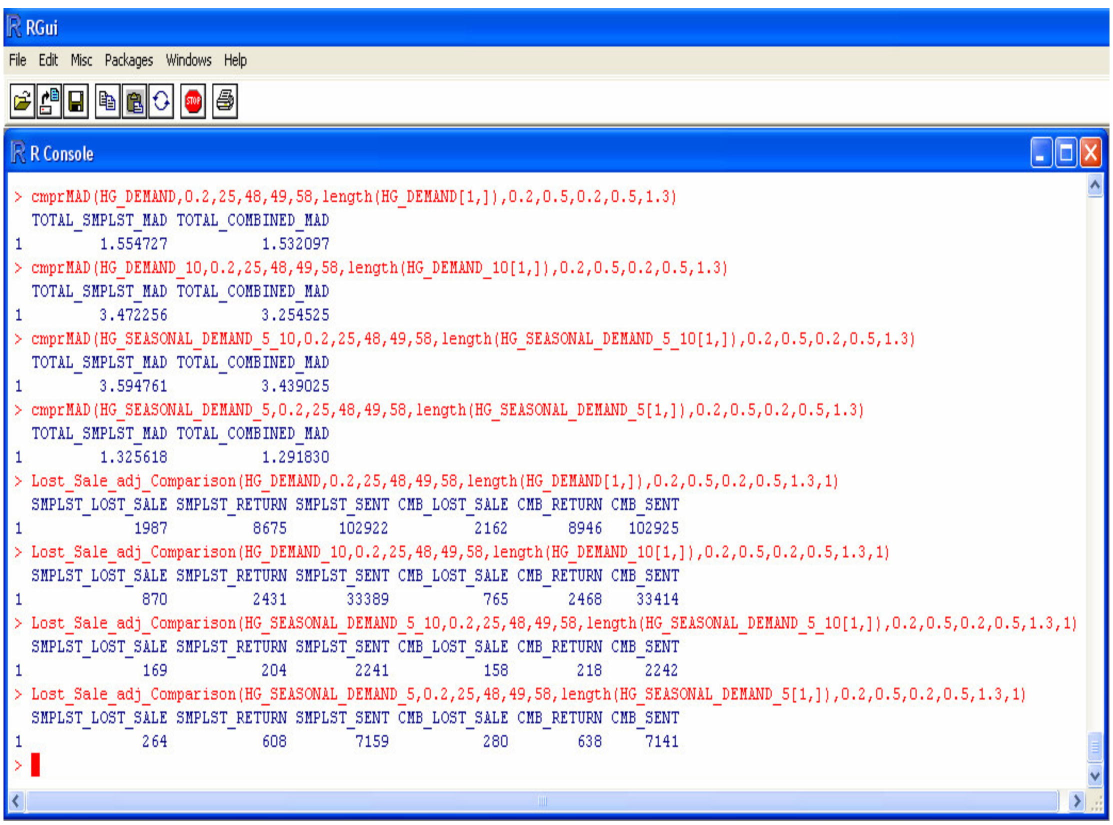


Figure 4.4. Calculation of MAD, delivery, return and lost sale values

After calculating Hey Girl magazine endpoint groups' MAD and lost sale values, the percentage reductions of the combined method for the Hey Girl magazine can be found. The improvement in MAD is calculated %2,2 for Hey Girl magazine;

(1.554727*1703-(3.472256*239+3.594761*24+ 1.325618*128))
=1561.878
(1.554727*1703-
(3.254525*239+3.439025*24+1.291830*128+1561.878))/(1.554727*1703)
=0.022

The improvement in lost sale values for Hey Girl magazine is calculated to be %0,5;

(1987-(870+169+264)) =684
(1987-(765+158+280+684))/1987=0.050

5.2. Some Important Functions

The algorithms will be explained according to the magazine forecasting order. The first algorithm, *estimate_past_demands*,

5.2.1. Estimate Past Demands Function

estimate_past_demands is used to produce past demand estimates. *estimate_past_demands* function needs the delivery, sale values and alfa smoothing parameter to estimate the past demands. If the sales are less than the delivery, then the sales amount are equal to the demand. But, if they are the same, the sales should be replaced by the conditional mean of the demand and *condexp* function is used to calculate the expected demand in case of sell out. The inner for loop of the *estimate_past_demands* function starts calculations by using simple exponential smoothing formula from the second data point because the initial values of the algorithm are assigned at the beginning. And the outer for loop repeats these calculations for other selected endpoints

```
estimate_past_demands <-function(sent,sold,alfa){
#sent… past delivery values
#sold… past sale values
```

```
#alfa… alpha smoothing parameter of the exponential smoothing
forecast<-matrix(nrow=length(sold[,1]),ncol=length(sold[1,]))
demand<-matrix(nrow=length(sold[,1]),ncol=length(sold[1,]))
        for(j in 1:length(sold[1,])){
        forecast[1,j]=sold[1,j]
        demand[1,j]=sold[1,j];
                for(i in 2:length(sold[,1])){
                forecast[i,j]=alfa*demand[i-1,j]+(1-alfa)*forecast[i-1,j]
                        if(sold[i,j]==0){
                        demand[i,j]=0
                        }
                        if(sent[i,j]>sold[i,j]){
                        demand[i,j]=sold[i,j]
                        }
                        if(sent[i,j]==sold[i,j]){
                        demand[i,j]=condexp(sent[i,j],forecast[i,j])
                        }
                demand[i,j][demand[i,j]==Inf]<-sold[i,j]
                demand[i,j][is.na(demand[i,j])]<-0;
                demand[i,j]=max(demand[i,j],sold[i,j])
                }
        }
print(demand)
}
```

Figure 4.5. estimate_past_demands function

<u>5.2.1.1. Condexp Function.</u> *condexp* function calculates conditional mean of demand based on Equation 4.1. When estimating conditional mean of demand, *condexp* function uses simple exponential smoothing formula to produce forecasted past demand. After the past demand is estimated, these demands can be used to forecast the future demand. Before forecasting future demand we have to decide which forecasting method is used in this study.

```
condexp<-function(sent,forecast){
```

```
#forecast…forecasted past demand
(forecast-sum(dpois((1:(sent-1)),forecast)*(1:(sent-1))))/(1-sum(dpois((1:(sent-
1)),forecast)))
}
```

Figure 4.6. condexp function

The second algorithm selects forecasting method by using *cmprMAD* function.

5.2.2. CmprMAD Function

The *cmprMAD* function compares Simple Exponential Smoothing I, Simple Exponential Smoothing II and No Trend Holt-Winters and selects the best method with minimum MAD. The method decision is done by using training set (first 48 months) of the magazine demand data. The for loop of the *cmprMAD* function calculates the *traning_mad* and *forecasted_mad* values of all the forecasting methods. The method selection is done by comparing *traning_mad* values of the forecasting methods. After method selection, the *forecasted_mad* values of the selected methods are assigned to the *selected_mad* values.

```
cmprMAD<-
function(ts,alf,nmin1,nmax1,nmin2,nmax2,length,rmin1,rmax1,rmin2,rmax2,ratio){
# selects the best method by using combined approach
# ts... estimated past demand.
# alf ... exponential smoothing parameter
# nmin1... nmax1   first, last observation for which MAD is evaluated in the training set
# nmin2...nmax2    first,last observation for which MAD is evaluated in the test set
#length...the number of endpoints for which MAD is evaluated
#rmin1... rmax1 optimization interval of alfa parameter
#rmin2... rmax2 optimization interval of gama parameter
#ratio... the coefficient of HoltWinters Training MAD in the selection phase
simplest_forecasted_MAD<-NULL
opt_ex1_training_mad<-NULL
opt_ex1_forecasted_MAD<-NULL
p<-NULL
```

```
opt_ex1<-as.list(p)
opt_ex2_training_mad<-NULL
opt_ex2_forecasted_MAD<-NULL
p<-NULL
opt_ex2<-as.list(p)
opt_ex2_training_mad<-NULL
q<-NULL;
opt_hw<-as.list(p)
opt_hw_training_mad<-NULL
opt_hw_forecasted_MAD<-NULL
selected_mad<-NULL
ts_hw<-(10+ts)
        for(i in 1:length) {
        simplest_forecasted_MAD[i]=expsmMAD(ts[,i],alf,ts[1,i],nmin2,nmax2)
        opt_ex1[[i]]=optimise(expsmMAD,c(rmin1,rmax1),ts=ts[,i],initialval=ts[1,i],nmin=
nmin1,nmax=nmax1)
        opt_ex1_training_mad[i]<-opt_ex1[[i]]$objective
        opt_ex1_forecasted_MAD[i]=expsmMAD(ts[,i],opt_ex1[[i]]$minimum,ts[1,i],nmi
n2,nmax2)
        opt_ex2[[i]]=optimise(expsmMAD,c(rmin1,rmax1),ts=ts[,i],initialval=mean(ts[1:n
max1,i]),nmin=nmin1,nmax=nmax1)
        opt_ex2_training_mad[i]<-opt_ex2[[i]]$objective
        opt_ex2_forecasted_MAD[i]=expsmMAD(ts[,i],opt_ex2[[i]]$minimum,mean(ts[1:
nmax1,i]),nmin2,nmax2)
        opt_hw[[i]]=optim(c(0.2,0.2),holtwintersMAD,ts=ts_hw[,i],nmin=nmin1,nmax=nm
ax1,method=c("L-BFGS-B"),lower=c(rmin1,rmin2),upper=c(rmax1,rmax2))
        opt_hw_training_mad[i]<-opt_hw[[i]]$value
        opt_hw_forecasted_MAD[i]=holtwintersMAD(ts_hw[,i],opt_hw[[i]]$par,nmin2,n
max2)

        if(min(opt_ex1_training_mad[i],opt_ex2_training_mad[i],ratio*opt_hw_training_m
ad[i])==opt_ex1_training_mad[i]) {
                selected_mad[i]=opt_ex1_forecasted_MAD[i]
```

```
        }
        else
if(min(opt_ex1_training_mad[i],opt_ex2_training_mad[i],ratio*opt_hw_training_mad[i])=
=opt_ex2_training_mad[i]) {
                selected_mad[i]=opt_ex2_forecasted_MAD[i]
        }
        else {
                selected_mad[i]=opt_hw_forecasted_MAD[i]
        }
}
TOTAL_SMPLST_MAD<-mean(simplest_forecasted_MAD )
TOTAL_COMBINED_MAD<-mean(selected_mad)
print(data.frame(TOTAL_SMPLST_MAD,TOTAL_COMBINED_MAD))
}
```

Figure 4.7. cmprMAD function

While selecting the best forecasting method, optimized smoothing coefficients are used that are produced from the training set. After this selection the forecasted MAD is calculated by using test set (last 10 months) of the data. To do this method selection *cmprMAD* function uses *expsmMAD* and *holtwintersMAD* functions.

5.2.2.1. ExpsmMAD Function. *expsmMAD* function calculates the MAD by using simple exponential smoothing

```
expsmMAD<-function(ts,alf,initialval,nmin,nmax){
# calculates MAD for simple exponential smoothing forecasts
# ts... time series as vector
# alf ... exponential smoothing parameter
# initialval.. single starting values
# nmin... nmax   first, last observation for which MAD is evaluated
y<-initialval
for(i in 1:(nmax-1)) y[i+1]=alf*ts[i]+(1-alf)*y[i]
mean(abs(y[nmin:nmax]-ts[nmin:nmax]))
```

```
}
```

Figure 4.8. expsmMAD function

<u>5.2.2.2. HoltwintersMAD Function</u>. *holtwintersMAD* function calculates MAD by using No Trend Holt-Winters method.

```
holtwintersMAD<-function(ts,parvec,nmin,nmax){
# calculates MAD for holtwinters forecasts
# ts... time series as vector
# parvec ... parameter vector of holtwinters
# nmin... nmax   first, last observation for which MAD is evaluated
y<-NULL
l<-NULL
sn<-NULL
l[1]<-mean(ts[1:12])
sn[1:12]<-ts[1:12]/l[1]
y[1]=(l[1])*sn[1];
        for(i in 2:(nmax-1)){
        l[i]=parvec[1]*(ts[i]/sn[i])+(1-parvec[1])*(l[i-1])
        sn[i+11]=parvec[2]*(ts[i-1]/l[i-1])+(1-parvec[2])*sn[i-1]
        y[i+1]=(l[i])*sn[i+1]
        }
mean(abs(y[nmin:nmax]-ts[nmin:nmax]))
}
```

Figure 4.9. holtwintersMAD function

6. CONCLUSIONS

The objective of this thesis was to develop an automatic forecasting system to increase the effectiveness of magazine distribution systems. Classical smoothing methods and several alternative forecasting methods were analyzed. Simple exponential smoothing was used as a benchmark.

As we wanted to model the seasonality in the magazine demand we first tried Holt-Winters for all endpoints. However, Holt-Winters performed very poorly. It was clearly worse than simple exponential smoothing. Therefore, we developed a "combined" method to model seasonality. The combined method compares the historic errors of exponential smoothing and Holt-winters to select the best method for each endpoint. We also used the best initialization methods and optimal smoothing parameters to get better forecasts. However, the performance was still disappointing. The reason seems to be that the demand time series are unstable.

In a next step, we restricted parameter optimization sub-intervals to exclude extreme cases close to zero and one. In addition, we grouped the data according to their average sales and seasonality characteristics. We applied the combined method to the high, summer medium and summer low groups and we used simple exponential smoothing for the rest of the data. Thanks to the combined method, we obtained an average reduction of 3.4 per cent for the MAD. When we applied this new method for to the distribution of 14 magazines in 10 consecutive months using the combined bottom-up and top-down approach, the lost sales were reduced by 1.2 per cent compared to a distribution procedure using simple exponential smoothing. This improvement may look small but we should not forget that it is reached without extra costs.

The second aim of this thesis was to find a sensible data organization for magazine data by using "R", a programming language and statistical package. We saw that using R is not only fast but also simple for forecasting large data. The R system has an extensive library of packages. Many calculations and analyses can be handled by using R's built-in functions without coding. Moreover, the R system can handle very large data sets without problems. So the second main finding of this thesis is that R is very useful for organizing and forecasting magazine data. Forecasting one magazine can be done in less than a minute by using our R statistical package. With Excel, a similar analysis would have been much more cumbersome as our basis data consisted of a 58×11000 matrix for each magazine.

APPENDIX A: ADDITIONAL R CODES

The following R code is used for filtering the data:

```
filter<-function(data,n1=12,n2=47,n3=58){
#data… time series as vector.
#n1, n2, n3… the months of the endpoints that filtering will be done in these inervals.
res1<-NULL;
	for(i in 1:length(data[1,])){
		if( sum(data[1:n1,i])>10 & sum(data[n2:n3,i])>10 ){
		res1[i]=names(data[i])
		}
	}
son=res1[!is.na(res1)];
print(son);
}
```

Figure A.1. filter function

The following R code is used for grouping endpoints:

```
find_groups<-function(data,val1,val2){
#nmin1... nmax1 first, last value for which endpoitns are grouped between in this interval.
res1<-NULL
	if(val1==0&val2>0){
		for(i in 1:length(data[1,])){
			if(mean(data[,i])>=val2){
			res1[i]=names(data[i])
			}
		}
	}
	else if(val1>0&val2>0){
		for(i in 1:length(data[1,])){
			if(mean(data[,i])>=val1&mean(data[,i])<val2){
			res1[i]=names(data[i])
			}
		}
```

```
          }
          else if(val1>0&val2==0){
                    for(i in 1:length(data[1,])){
                              if( mean(data[,i])<val1){
                              res1[i]=names(data[i])
                              }
                    }
          }
son=res1[!is.na(res1)];
print(son);
        }
```

Figure A.2. find_groups function

The following R code is used for finding summer seasonal endpoints:

```
Summer_Ep<-function(vect,months,pervec){
total=apply(vect,2,sum);
upto<-length(vect[,1])/12
res1<-NULL;
z=1;
          for(i in 1:length(months)){
                    for(k in 0:(upto-1)){


                    res1[z]=months[i]+k*12;
                    z=z+1;
                    }
           }
sum1=apply(vect[res1,],2,sum);Summer_Coef=(sum1/total);
summer_ep<-Summer_Coef[Summer_Coef>=pervec]
print(names(summer_ep))
}
```

Figure A.3. Summer_Ep function

The following R code is used for producing future demand by using Simple Exponential Emoothing:

```
expsmFOR<-function(ts,alf,initialval,nmin,nmax){
# calculates MAD for simple exponential smoothing forecasts
# ts... time series as vector
# alf ... exponential smoothing parameter
# initialval.. single starting values
# nmin... nmax   first, last observation for which forecasts are evaluated
y<-initialval
for(i in 1:(nmax-1)) y[i+1]=alf*ts[i]+(1-alf)*y[i]
y[nmin:nmax]
}
```

Figure A.4. expsmFOR function

The following R code is used for producing future demand by using No Trend Holt Winters:

```
holtwintersFOR<-function(ts,parvec,nmin,nmax){
# calculates MAD for holtwinters forecasts
# ts... time series as vector
# parvec ... parameter vector of holtwinters
# nmin... nmax   first, last observation for which forecasts are evaluated
y<-NULL
l<-NULL
sn<-NULL
l[1]<-mean(ts[1:12])
sn[1:12]<-ts[1:12]/l[1]
y[1]=(l[1])*sn[1];
        for(i in 2:(nmax-1)){
        l[i]=parvec[1]*(ts[i]/sn[i])+(1-parvec[1])*(l[i-1])
        sn[i+11]=parvec[2]*(ts[i-1]/l[i-1])+(1-parvec[2])*sn[i-1]
        y[i+1]=(l[i])*sn[i+1]
        }
y[nmin:nmax]
}
```

Figure A.5. holtwintersFOR function

The following R code is used for producing delivery, return and lost sales:

```
Lost_Sale_adj_Comparison<-
function(ts,alf,nmin1,nmax1,nmin2,nmax2,length,rmin1,rmax1,rmin2,rmax2,ratio,to){
# selects the best method by using combined approach
# ts... time series as vector
# alf ... exponential smoothing parameter
# nmin1... nmax1   first, last observation for which MAD is evaluated in the training set
# nmin2...nmax2    first, last observation for which forecasts are evaluated in the test set
#length...the number of endpoints for which MAD is evaluated
#rmin1... rmax1 optimization interval of alfa parameter
#rmin2... rmax2 optimization interval of gama parameter
#ratio... the coefficient of HoltWinters Training MAD in the selection phase
ts_hw<-(10+ts)
simplest_training_MAD<-NULL
simplest_forecasted_FOR<-matrix(nrow=(nmax2-nmin2)+1,ncol=length)
simplest_forecasted_SENT<-matrix(nrow=(nmax2-nmin2)+1,ncol=length)
simplest_lost_sale<-NULL
simplest_return<-NULL
opt_ex1_training_mad<-NULL
opt_ex1_forecasted_FOR<-matrix(nrow=(nmax2-nmin2)+1,ncol=length)
p<-NULL
opt_ex1<-as.list(p)
opt_ex2_training_mad<-NULL
opt_ex1_forecasted_FOR<-matrix(nrow=(nmax2-nmin2)+1,ncol=length)
p<-NULL
opt_ex2<-as.list(p)
opt_ex2_training_mad<-NULL
opt_ex2_forecasted_FOR<-matrix(nrow=(nmax2-nmin2)+1,ncol=length)
```

```
q<-NULL;
opt_hw<-as.list(p)
opt_hw_training_mad<-NULL
opt_hw_forecasted_FOR<-matrix(nrow=(nmax2-nmin2)+1,ncol=length)
selected_for<-matrix(nrow=(nmax2-nmin2)+1,ncol=length)
sent<-matrix(nrow=(nmax2-nmin2)+1,ncol=length)
lost_sale<-NULL;
return<-NULL;
	for(i in 1:length) {
	simplest_training_MAD[i]=expsmMAD(ts[,i],alf,ts[1,i],nmin1,nmax1)
	simplest_forecasted_FOR[,i]=expsmFOR(ts[,i],alf,ts[1,i],nmin2,nmax2)
	simplest_forecasted_SENT[,i]=simplest_forecasted_FOR[,i]+to*sqrt(simplest_fore
casted_FOR[,i])
	simplest_lost_sale[i]=sum(max(0,(ts[nmin2:nmax2,i]-
simplest_forecasted_SENT[,i])))
	simplest_return[i]=sum(max(0,(simplest_forecasted_SENT[,i]-
ts[nmin2:nmax2,i])))
	opt_ex1[[i]]=optimise(expsmMAD,c(rmin1,rmax1),ts=ts[,i],initialval=ts[1,i],nmin=
nmin1,nmax=nmax1)
	opt_ex1_training_mad[i]<-opt_ex1[[i]]$objective
	opt_ex1_forecasted_FOR[,i]=expsmFOR(ts[,i],opt_ex1[[i]]$minimum,ts[1,i],nmin
2,nmax2)
	opt_ex2[[i]]=optimise(expsmMAD,c(rmin1,rmax1),ts=ts[,i],initialval=mean(ts[1:n
max1,i]),nmin=nmin1,nmax=nmax1)
	opt_ex2_training_mad[i]<-opt_ex2[[i]]$objective
	opt_ex2_forecasted_FOR[,i]=expsmFOR(ts[,i],opt_ex2[[i]]$minimum,ts[1,i],nmin
2,nmax2)

opt_hw[[i]]=optim(c(0.2,0.2),holtwintersMAD,ts=ts_hw[,i],nmin=nmin1,nmax=nmax1,me
thod=c("L-BFGS-B"),lower=c(rmin1,rmin2),upper=c(rmax1,rmax2))
	opt_hw_training_mad[i]<-opt_hw[[i]]$value
	opt_hw_forecasted_FOR[,i]=holtwintersFOR(ts_hw[,i],opt_hw[[i]]$par,nmin2,nma
x2)
```

```
	if(min(opt_ex1_training_mad[i],opt_ex2_training_mad[i],ratio*opt_hw_training_m
ad[i])==opt_ex1_training_mad[i]) {
		selected_for[,i]=opt_ex1_forecasted_FOR[,i]
		sent[,i]=selected_for[,i]+to*sqrt(selected_for[,i])
	}
	else
if(min(opt_ex1_training_mad[i],opt_ex2_training_mad[i],ratio*opt_hw_training_mad[i])=
=opt_ex2_training_mad[i]) {
		selected_for[,i]=opt_ex2_forecasted_FOR[,i]
		sent[,i]=selected_for[,i]+to*sqrt(selected_for[,i])
	}
	else {
		selected_for[,i]=opt_hw_forecasted_FOR[,i]
		sent[,i]=selected_for[,i]+to*sqrt(selected_for[,i])
	}
}
adj<-function(del,cir,to,coef){
sum(coef*del+to*sqrt(coef*del))-cir
}
cm=uniroot(adj,c(0,3),tol=0.0001,del=selected_for,cir=sum(simplest_forecasted_SENT),to
=1)$root
sent=selected_for*cm + to*sqrt(selected_for*cm)
rsent=round(sent)
min_val=min(sent-rsent)
max_val=max(sent-rsent)
for(i in 1:length){
	for(j in 1:length(sent[,1])){
		if(sent[j,i]-rsent[j,i]==min_val){
		rsent[j,i]=rsent[j,i]-1
		}
		if(sent[j,i]-rsent[j,i]==max_val){
		rsent[j,i]=rsent[j,i]+1
```

```
                    }
          }
}
for(i in 1:length) {
                    lost_sale[i]=sum(max(0,(ts[nmin2:nmax2,i]-rsent[,i])))
                    return[i]=sum(max(0,(rsent[,i]-ts[nmin2:nmax2,i])))
}
print(data.frame(SMPLST_LOST_SALE=round(sum(simplest_lost_sale)),SMPLST_RET
URN=round(sum(simplest_return)),SMPLST_SENT=round(sum(simplest_forecasted_SE
NT)),CMB_LOST_SALE=round(sum(lost_sale)),CMB_RETURN=round(sum(return)),C
MB_SENT=sum(rsent)))
}
```

Figure A.6. Lost_Sale_adj_Comparison function

The following R code is used for deciding seasonality type:

```
decide_multiplicative_additive<-function(ts){
#ts … time series as vector
standard_deviation<-apply(ts,2,sd);
average<-apply(ts,2,mean);
plot(average,standard_deviation);
}
```

Figure A.7. decide_multiplicative_additive function

The following R code is used for deciding smoothing model:

```
decide_smoothing_model<-function(ts){
res1<-NULL;
res2<-NULL;
res3<-NULL;
       for(i in 1:length(ts[1,])){
       res1[i]=HoltWinters(ts[,i],seasonal=c("multiplicative"))$alpha;
       res2[i]=HoltWinters(ts[,i],seasonal=c("multiplicative"))$beta;
```

```
        res3[i]=HoltWinters(ts[,i],seasonal=c("multiplicative"))$gamma;
        }
alpha=res1;
beta=res2
gamma=res3;
hist(alpha);
windows();
hist(beta);
windows();
hist(gamma);
}
```

Figure A.8. decide_smoothing_model function

The following R codes are used in initialization and parameter optimization simulation study:

```
#model1;  iid ts
ts500<-matrix(nrow=500,ncol=58)
for(i in 1:500){
ts500[i,]<-round(rnorm(58,5));
}
#model2 AR(1), ro=0.3
ts_ar1_0.3<-matrix(nrow=500,ncol=58)
for(i in 1:500) {
ts_ar1_0.3[i,]=round(arima.sim(list(order = c(1,0,0), ar = 0.3), n = 58)+15)
}
#model3  AR(1), ro=0.9
ts_ar1_0.9<-matrix(nrow=500,ncol=58)
for(i in 1:500) {
ts_ar1_0.9[i,]=round(arima.sim(list(order = c(1,0,0), ar = 0.9), n = 58)+15)
}
#model4 ARIMA(0,1,1)(0,1,1)12  versiyon1;   w1=0.5, w2=0.5; versiyon2; w1=0.4,
w2=0.6; versiyon3; w1=0, w2=0, versiyon4; w1=0.1, w2=0.8
sim_sairma_model4<-function(n,w1,w2,m,rt){
```

```
        Y<-matrix(nrow=rt,ncol=n);
        for(j in 1:rt){
        Y[j,1:13]= round(arima.sim( n = 13, list( ar = c(0,0), ma = c(w1, w2) ) )+m);
                for(i in 14:n){
                        Y[j,i]=round(Y[j,i-1]+Y[j,i-12]-Y[j,i-13]-w1*rnorm(1)-
w2*rnorm(1)+rnorm(1));
                }
        }
print(Y)
}
model4_versiyon1<-sim_sairma_model4(58,0.5,0.5,100,500)
model4_versiyon2<-sim_sairma_model4(58,0.4,0.6,100,500)
model4_versiyon3<-sim_sairma_model4(58,0,0,100,500)
model4_versiyon4<-sim_sairma_model4(58,0.1,0.8,100,500)
#model5 ARIMA(1,0,0)(1,0,0)12 versiyon1;  versiyon1; ro1=0.3, ro2=0.7  versiyon2;
ro1=0, ro2=0 , versiyon3; ro1=0, ro2=0.8, versiyon4; ro1=0.5, ro2=0.5
sim_sairma_model5<-function(n,a,b,m,rt){
        Y<-matrix(nrow=rt,ncol=n);
        for(j in 1:rt){
        Y[j,1:12]= round(arima.sim( n = 12, list( ar = c(0.2, 0.7), ma = c(0, 0) ) )+m);
                for(i in 13:n){
                        Y[j,i]=round(a*Y[j,i-1]+b*Y[j,i-12]+rnorm(1));
                }
        }
print(Y)
}
model5_versiyon1<-sim_sairma_model5(58,0.3,0.7,12,500)
model5_versiyon2<-sim_sairma_model5(58,0,0,50,500)
sim_sairma_model5<-function(n,ro1,ro2,m,rt){
        Y<-matrix(nrow=rt,ncol=n);
        for(j in 1:rt){
        Y[j,1:12]= round(arima.sim( n = 12, list( ar = c(ro1, ro2), ma = c(0, 0) ) )+m);
                for(i in 13:n){
```

```
                    Y[j,i]=round(ro1*Y[j,i-1]+ro2*Y[j,i-12]+rnorm(1));
            }
        }
print(Y)
}
model5_versiyon3<-sim_sairma_model5(58,0,0.8,20,500)
model5_versiyon4<-sim_sairma_model5(58,0.5,0.5,12,500)
#EXPONENTIAL SMOOTHING FORMULATIONS
#Versiyon1-First Value is used as an initial value.
ex_smooth1_1<-function(vect,alfa){
l<-NULL;
l[1]=vect[1];
        for(i in 1:36){
                l[i+1]=alfa*vect[i]+(1-alfa)*l[i]
        }
        AD=abs(vect[13:36]-l[13:36])
print(mean(AD))
}
ex_smooth1_2<-function(vect,alfa){
l<-NULL;
l[1]=vect[1];
        for(i in 1:36){
                l[i+1]=alfa*vect[i]+(1-alfa)*l[i]
        }
        AD=abs(vect[13:36]-l[13:36])
print(mean(AD))
}
#Versiyon2-Average of the all data is used as an initial value.
ex_smooth2_1<-function(vect,alfa){
l<-NULL;
l[1]=mean(vect);
        for(i in 1:36){
                l[i+1]=alfa*vect[i]+(1-alfa)*l[i]
```

```
    }
    AD=abs(vect[13:36]-l[13:36])
print(mean(AD))
}
ex_smooth2_2<-function(vect,alfa){
l<-NULL;
l[1]=mean(vect);
    for(i in 1:36){
        l[i+1]=alfa*vect[i]+(1-alfa)*l[i]
    }
    AD=abs(vect[13:36]-l[13:36])

print(mean(AD))
}
#HOLT WINTERS FORMULATIONS
#Versiyon1-First Value is used as an initial value and seasonality indices are set to 1.
holt_winters1_1<-function(vect,a){
l<-NULL;
sn<-NULL;
l[1]=(a[1]);
sn[1:12]=1;
y<-NULL;
y[1]=(l[1])*sn[1];
    for(i in 2:36){
        l[i]=vect[1]*(a[i]/sn[i])+(1-vect[1])*(l[i-1]);
        sn[i+12-1]=vect[2]*(a[i-1]/l[i-1])+(1-vect[2])*sn[i-1];
        y[i+1]=(l[i])*sn[i+1];
    }
    AE=abs(a[13:36]-y[13:36])
print(mean(AE))
}
holt_winters1_2<-function(vect,a){
l<-NULL;
```

```
sn<-NULL;
l[1]=(a[1]);
sn[1:12]=1;
y<-NULL;
y[1]=(l[1])*sn[1];
	for(i in 2:36){
		l[i]=vect[1]*(a[i]/sn[i])+(1-vect[1])*(l[i-1]);
		sn[i+12-1]=vect[2]*(a[i-1]/l[i-1])+(1-vect[2])*sn[i-1];
		y[i+1]=(l[i])*sn[i+1];
	}
	AE=abs(a[13:36]-y[13:36])
print(mean(AE))
}

#Versiyon2-Average of first 12 Value is used as an initial value and seasonality indices are found by dividing first 12 values to the initial value.
holt_winters2_1<-function(vect,a){
l<-NULL;
sn<-NULL;
l[1]=mean(a[1:12]);
sn[1:12]=a[1:12]/l[1];
y<-NULL;
y[1]=(l[1])*sn[1];
	for(i in 2:36){
		l[i]=vect[1]*(a[i]/sn[i])+(1-vect[1])*(l[i-1]);
		sn[i+12-1]=vect[2]*(a[i-1]/l[i-1])+(1-vect[2])*sn[i-1];
		y[i+1]=(l[i])*sn[i+1];
	}
	AE=abs(a[13:36]-y[13:36])
print(mean(AE))
}
holt_winters2_2<-function(vect,a){
l<-NULL;
```

```
sn<-NULL;
l[1]=mean(a[1:12]);
sn[1:12]=a[1:12]/l[1];
y<-NULL;
y[1]=(l[1])*sn[1];
	for(i in 2:36){
		l[i]=vect[1]*(a[i]/sn[i])+(1-vect[1])*(l[i-1]);
		sn[i+12-1]=vect[2]*(a[i-1]/l[i-1])+(1-vect[2])*sn[i-1];
		y[i+1]=(l[i])*sn[i+1];
	}
	AE=abs(a[13:36]-y[13:36])
print(mean(AE))
}
cmpr_mthds<-function(time_series){
MAD_holt_winters1_1<-NULL;
v<-NULL;
MAD_holt_winters1_2<-as.list(v)
MAD_holt_winters2_1<-NULL;
q<-NULL;
MAD_holt_winters2_2<-as.list(q)
MAD_ex_smooth1_1<-NULL;
MAD_ex_smooth1_2<-NULL;
MAD_ex_smooth2_1<-NULL;
MAD_ex_smooth2_2<-NULL;
holt_winters1<-NULL;
holt_winters2<-NULL;
exponential_smoothing1<-NULL;
exponential_smoothing2<-NULL;
holt_winters<-NULL;
exponential_smoothing<-NULL;
final_forecast<-NULL;
fd<-NULL;
hw1=0;
```

```
hw2=0;
ex1=0;
ex2=0;
hw=0;
ex=0;
for(i in 1:500) {
	MAD_holt_winters1_1[i]=optim(c(0.3,0.25),holt_winters1_1,a=time_series[i,],met
hod=c("L-BFGS-B"),lower=c(0,0),upper=c(1,1))$value
	MAD_holt_winters1_2[[i]]=optim(c(0.3,0.25),holt_winters1_1,a=time_series[i,],m
ethod=c("L-BFGS-B"),lower=c(0,0),upper=c(1,1))$par
	MAD_holt_winters2_1[i]=optim(c(0.3,0.25),holt_winters2_1,a=time_series[i,],met
hod=c("L-BFGS-B"),lower=c(0,0),upper=c(1,1))$value
	MAD_holt_winters2_2[[i]]=optim(c(0.3,0.25),holt_winters2_1,a=time_series[i,],m
ethod=c("L-BFGS-B"),lower=c(0,0),upper=c(1,1))$par
	MAD_ex_smooth1_1[i]=optimise(ex_smooth1_1,c(0,1),vect=time_series[i,])$obje
ctive

MAD_ex_smooth1_2[i]=optimise(ex_smooth1_1,c(0,1),vect=time_series[i,])$minimum
	MAD_ex_smooth2_1[i]=optimise(ex_smooth2_1,c(0,1),vect=time_series[i,])$obje
ctive

MAD_ex_smooth2_2[i]=optimise(ex_smooth2_1,c(0,1),vect=time_series[i,])$minimum

	holt_winters1[i]=holt_winters1_2(MAD_holt_winters1_2[[i]],time_series[i,])

	holt_winters2[i]=holt_winters2_2(MAD_holt_winters2_2[[i]],time_series[i,])

	exponential_smoothing1[i]=ex_smooth1_2(time_series[i,],MAD_ex_smooth1_2[i])

	exponential_smoothing2[i]=ex_smooth2_2(time_series[i,],MAD_ex_smooth2_2[i])
		if(MAD_holt_winters1_1[i]<MAD_holt_winters2_1[i]){
		holt_winters[i]=holt_winters1_2(MAD_holt_winters1_2[[i]],time_series[i,])
		hw1=hw1+1;
```

```
        }
        else {
        holt_winters[i]=holt_winters2_2(MAD_holt_winters2_2[[i]],time_series[i,])
        hw2=hw2+1;
        }
        if(MAD_ex_smooth1_1[i]<MAD_ex_smooth2_1[i]){

    exponential_smoothing[i]=ex_smooth1_2(time_series[i,],MAD_ex_smooth1_2[i])
        ex1=ex1+1;
        }
        else {

    exponential_smoothing[i]=ex_smooth2_2(time_series[i,],MAD_ex_smooth2_2[i])
        ex2=ex2+1;
        }
        if(holt_winters[i]<exponential_smoothing[i]){
        final_forecast[i]=holt_winters[i];
        hw=hw+1;
        }
        else {
        final_forecast[i]=exponential_smoothing[i];
        ex=ex+1;
        }
}
print( data.frame(
Cmb=mean(final_forecast),HW=mean(holt_winters),EX=mean(exponential_smoothing),H
W_EP=hw,EX_EP=ex,HW1=mean(holt_winters1),HW2=mean(holt_winters2),EX1=mea
n(exponential_smoothing1),EX2=mean(exponential_smoothing2),HW_EP1=hw1,HW_EP
2=hw2,EX_EP1=ex1,EX_EP2=ex2 ) )
}
```

Figure A.9. Initialization and parameter optimization functions

REFERENCES

Artto, K. A. and E. Pylkkänen, "An effective procedure for distribution of magazines", *International Transactions in Operation Research,* Vol. 6, pp. 289-320, 1999.

Bell, P. C., "A new procedure for distribution of periodicals", *Journal of the Operational Research Society*, Vol. 29, No. 5, pp. 427-434, 1978.

Bell, P. C., "Forecasting demand variation when there are stockouts", *Journal of the Operational Research Society*, Vol. 51, No. 5, pp. 358-363, 2000.

Bowerman, B. L., R. T. O'Connell and A. B. Koehler, 1979, *Time Series and Forecasting*: *An Applied Approach*, Duxbury Press, North Scituate.

Bowerman, B. L., R. T. O'Connell and A. B. Koehler, 2005, *Forecasting, Time Series and Regression,* Thomson Brooks/Cole, Belmont.

Brown, R. G., 1962, *Smoothing, Forecasting and Prediction of Discrete Time Series*, Prentice-Hall, Englewood Cliffs, NJ.

Broze, L. and G. Mélard, "Exponential Smoothing: Estimation by Maximum Likelihood", *Journal of Forecasting*, Vol. 9, pp. 445-455, 1990.

Cohen, G. D., "A note on exponential smoothing and autocorrelated inputs", *Operation Research*, Vol. 11, pp. 361-367, 1963.

Dekker, M., K. V. Donselaar and P. Ouwehand, "How to use aggregation and combined forecasting to improve seasonal demand forecasts", *International Journal of Production Economics*, Vol. 90, pp. 151-167, 2004.

Farnum, N. R., "Exponential Smoothing: Behavior of the ex-post Sum of Squares near zero and one", *Journal of Forecasting*, Vol. 11, pp. 47-56, 1992.

Gardner, E. S. Jr., “Exponential Smoothing: The State of the Art”, *Journal of Forecasting*, Vol. 4, pp. 1-28, 1985.

Hanke, J. E. and D. W. Wichern, 2005, *Business Forecasting,* Pearson Prentice Hall, New Jersey

Hyndman, R. J., A. B. Koehler, R. D. Snyder and S. Grose, “A state space framework for automatic forecasting using exponential smoothing methods”, *International Journal of Forecasting*, Vol. 18, pp. 439-454, 2002.

Leonard, M., *Large Scale Automatic Forecasting*: *Millions of Forecasts* http://support.sas.com/rnd/app/papers/largescale.pdf

Ledolter, J. and B. Abraham, “Some Comments on the Initialization of Exponential Smoothing”, *Journal of Forecasting*, Vol. 3, pp. 79-84, 1984.

Makridakis, S. and Wheelwright, S. C., 1978, *Forecasting*: *Methods and Applications*, Wiley, New York.

Ord, K. and S. Lowe, “Review: Automatic Forecasting”, *The American Statistician*, Vol. 50, pp. 88-94, 1996.

Özgün, O., *Distribution Planning of Magazines*, Master Thesis, Boğaziçi University, 2006.

Taylor, S. G., “Initialization of Exponential Smoothing Forecasts”, Vol. 13, pp. 199-205, 1981.