FORECASTING NATURAL GAS CONSUMPTION BY TIME SERIES METHODS

by

Mustafa AKKURT

July 2009

FORECASTING NATURAL GAS CONSUMPTION BY TIME SERIES METHODS

by

Mustafa AKKURT

A thesis submitted to

the Graduate Institute of Sciences and Engineering

of

Fatih University

in partial fulfillment of the requirements for the degree of

Master of Science

in

Industrial Engineering

July 2009

Istanbul, Turkey

APPROVAL PAGE

I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Science.

____
Assist. Prof. Dr. Mehmet ŞEVKLİ
Head of Department

This is to certify that I have read this thesis and that in my opinion it is fully adequate, in scope and quality, as a thesis for the degree of Master of Science.

____
Assist. Prof. Dr. Ömer Fahrettin DEMİREL
Supervisor

Examining Committee Members

Assist. Prof. Dr. Ömer Fahrettin DEMİREL	____________________
Assist. Prof. Dr. Fahrettin ELDEMİR	____________________
Prof. Selim ZAİM	____________________

It is approved that this thesis has been written in compliance with the formatting rules laid down by the Graduate Institute of Sciences and Engineering.

____
Assoc. Prof. Dr. Nurullah ARSLAN
Director

July 2009

FORECASTING NATURAL GAS CONSUMPTION BY TIME SERIES METHODS

Mustafa AKKURT

M.S. Thesis – Industrial Engineering
July 2009

Supervisor: Assist. Prof. Dr. Ömer Fahrettin DEMİREL

ABSTRACT

Forecasting is the way of making predictions about a variable. Time series methods are utilized to make forecasting. The bootstrap method is a method of computational inference that simulates the creation of new data by resampling from a single data set. In this study natural gas consumptions of Istanbul and Turkey are forecasted by using time series methods. The natural gas consumption data are provided by Istanbul Gaz Dagitim A.S. (IGDAS), Boru Hatlari ile Petrol Tasima A.S. (BOTAS), and International Energy Agency (IEA). Time series methods such as exponential smoothing, Winters' forecasting and Box-Jenkins methods are used to predict future natural gas consumption values of Istanbul and Turkey on different time periods. These methods are compared in order to determine the best fitted one to related time series data. Different natural gas consumption scenarios about Istanbul are generated by using bootstrap method.

Keywords: Natural Gas Consumption, Forecasting, Time Series, Winters' method, Box-Jenkins methods, Bootstrap, Scenario Generation.

ZAMAN SERİLERİ METOTLARIYLA DOĞALGAZ TÜKETİMİNİN TAHMİNİ

Mustafa AKKURT

Yüksek Lisans Tezi – Endüstri Mühendisliği
Temmuz 2009

Tez Yöneticisi: Yrd. Doç. Dr. Ömer Fahrettin DEMİREL

ÖZ

Tahminleme bir değişkene ait öngörüler yapmanın yoludur. Zaman serileri metotları tahminler yapmak için zaman serileri verilerinin analizidir. Bootstrap, bir veri setinden tekrar örnekleme ile yeni bir seri üretimini benzeten sayısal bir metottur. Bu çalışmada İstanbul ve Türkiye'nin doğalgaz tüketimleri zaman serileri metotları kullanılarak tahmin edilmiştir. Doğalgaz tüketim verileri İstanbul Gaz Dağıtım A.Ş. (İGDAŞ), Boru Hatları ile Petrol Taşıma A.Ş. (BOTAŞ) ve Uluslararası Enerji Ajansı (IEA) 'dan temin edilmiştir. Üssel yumuşatma, Winters' metodu ve Box-Jenkins metotları gibi zaman serileri metotları kullanılarak gelecek İstanbul ve Türkiye doğalgaz tüketim değerleri değişik zaman aralıkları üzerinden tahmin edilmiştir. Bu metotlar ilgili zaman serisi verisine uyan en iyi metodu belirlemek için karşılaştırılmıştır. İstanbul için değişik doğalgaz tüketim senaryoları bootstrap metodu kullanılarak üretilmiştir.

Anahtar Kelimeler: Doğalgaz Tüketimi, Tahminleme, Zaman Serileri, Winters metodu, Box-Jenkins Metotları, Bootstrap, Senaryo Üretimi.

Dedicated to my parents

ACKNOWLEDGEMENT

I would like to express my gratitude to my supervisor Assist. Prof. Dr. Ömer Fahrettin DEMİREL whose help, stimulating suggestions and encouragement helped me during the course of research and writing of this thesis.

I also want to thank to Assist. Prof. Dr. Fahrettin ELDEMİR and Prof. Selim ZAİM for their significant contributions to my research endeavors.

I would also like to thank to The Graduate Institute of Sciences and Engineering of Fatih University and TÜBİTAK for funding my thesis. Moreover, I would like to thank IGDAS stuff who meets my data demands.

I express my thanks and appreciation to my parents for their understanding, motivation and patience. Lastly, but in no sense the least, I am thankful to all colleagues and friends who support me through all these years.

TABLE OF CONTENTS

ABSTRACT iii
ÖZ iv
DEDICATION v
ACKNOWLEDGEMENT vi
TABLE OF CONTENTS vii
LIST OF TABLES x
LIST OF FIGURES xi
LIST OF SYMBOLS AND ABBREVIATIONS xiii
CHAPTER 1 INTRODUCTION 1
1.1 NATURAL GAS CONSUMPTION FORECASTING 3
1.1.1 Importance of Natural Gas Usage 3
1.1.2 Review of Studies about Natural Gas Consumption 5
1.1.3 Characteristics of Natural Gas Consumption 6
CHAPTER 2 FORECASTING TIME SERIES DATA 8
2.1 WHAT IS FORECASTING? 8
2.2 STATIONARY AND NON-STATIONARY DATA 9
2.3 ANALYSIS OF TIME SERIES 9
2.3.1 Classical Decomposition of Time Series 10
2.3.2 Box-Jenkins Methods 10
2.3.2.1 Model Identification 11
2.3.2.2 Model Parameters Estimation 11
2.3.2.3 Diagnostic Checking 11
2.3.2.4 Autoregressive Model 12
2.3.2.5 Moving Average Model 12
2.3.2.6 Mixed Autoregressive Moving Average (ARMA) Model 12
2.3.2.7 Autoregressive Integrated Moving Average (ARIMA) Model 13
2.3.2.8 The Box-Jenkins Seasonal (SARIMA) Model 14

CHAPTER 3 TIME SERIES DATA BOOTSTRAP 15
3.1 BOOTSTRAP OF TIME SERIES DATA AND BOOTSTRAP METHODS 15
3.1.1 Parametric Bootstrap Method 15
3.1.2 Nonparametric Bootstrap Methods 16
3.1.2.1 Moving Block Bootstrap 16
3.1.2.2 Linked Blockwise Bootstrap 16
3.1.2.3 Matched-Block Bootstrap 17
3.1.2.4 Stationary Bootstrap 17
3.1.2.5 Threshold Bootstrap 18
3.1.2.6 Nearest Neighbor Bootstrap 18
3.1.2.7 Subseries Method 18
3.2 SCENARIO GENERATION BY BOOTSTRAP 19
3.3 FORECASTING BY BOOTSTRAP SCENARIOS 20
CHAPTER 4 FORECASTING NATURAL GAS CONSUMPTION DATA OF ISTANBUL 21
4.1 DAILY NATURAL GAS CONSUMPTION DATA OF ISTANBUL 21
4.1.1 The Plot of Daily Natural Gas Consumption of Istanbul 21
4.1.2 Parameter Search 24
4.2 MONTHLY NATURAL GAS CONSUMPTION DATA OF ISTANBUL 27
4.2.1 The Plot of Monthly Natural Gas Consumption Data of Istanbul 27
4.2.2 Forecasting Monthly Natural Gas Consumption of Istanbul 27
4.2.3 The Box-Jenkins Seasonal (SARIMA) Model 33
CHAPTER 5 FORECASTING NATURAL GAS CONSUMPTION DATA OF TURKEY 37
5.1 MONTHLY NATURAL GAS CONSUMPTION DATA OF TURKEY 37
5.1.1 The Plot of Natural Gas Consumption Data of Turkey 37
5.1.2 Forecasting Monthly Natural Gas Consumption 39
5.1.3 Autoregressive Integrated Moving Average Model for Forecasting Monthly Natural Gas Consumption of Turkey 40
5.2 ANNUAL NATURAL GAS CONSUMPTION DATA OF TURKEY 43

5.2.1 The Plot of Natural Gas Consumption Data of Turkey 43
5.2.2 Forecasting the Annual Natural Gas Consumption of Turkey with Trend Analysis 43
5.2.3 Forecasting the Annual Natural Gas Consumption of Turkey with Autoregressive Integrated Moving Average Model 50
5.3 Forecasting the Last Four Years of Annual Natural Gas Consumption of Turkey and Comparing with Literature Study 60
CHAPTER 6 GENERATING CONSUMPTION SCENARIOS BY BOOTSTRAP ... 62
6.1 ESTIMATING THE BEST BLOCK SIZE 62
6.2 HOW TO DEAL WITH NON-STATIONARITY 63
6.2.1 Removal of Non-Stationary Components of Monthly Natural Gas Consumption of Istanbul.. 64
6.2.2 Estimation of Best Block Size 72
6.3 GENERATING BOOTSTRAPS AND SIBLING SERIES OF MONTHLY NATURAL GAS CONSUMPTION DATA OF ISTANBUL 74
6.4 PRODUCTION OF MONTHLY NATURAL GAS CONSUMPTION SCENARIOS OF ISTANBUL 76
CHAPTER 7 CONCLUSIONS 79
APPENDIX A SAMPLES OF BOOTSTRAPS, SIBLING SERIES AND MONTHLY CONSUMPTION SCENARIOS OF ISTANBUL 82
REFERENCES 82

LIST OF TABLES

Table 2.1 Several Time Series Methods 9
Table 4.1 Parameters’ test results 24
Table 4.2 Forecasting methods’ results 29
Table 4.3 Winters’ Multiplicative Method results 29
Table 4.4 Parameter search 32
Table 4.5 Comparisons of forecasts and real value of 2008 and 2009 33
Table 4.6 Comparison of forecasts by using Winters’ method and SARIMA 34
Table 5.1 SmartForecasts results 39
Table 5.2 Winters’ multiplicative method forecasts 39
Table 5.3 Comparison of two forecasting types 42
Table 5.4 Error values of different models 45
Table 5.5 Forecasts about 2007 and 2008 45
Table 5.6 Parameter search on double exponential smoothing 46
Table 5.7 Comparison of double exponential smoothing models 46
Table 5.8 Coefficients of ARIMA models 57
Table 5.9 Results of ARIMA models 57
Table 5.10 Comparison between ARIMA model and double exponential smoothing ... 58
Table 5.11 Comparison table of methods 61
Table 6.1 Golden deltas’ normalized zero-crossing rates 72
Table 6.2 Kolmogrov-Smirnov (K-S) 2-sample test results 73
Table 6.3 Bootstrap forecasts of 2008 and 2009 76
Table 6.4 Comparison of Winters’ method and Bootstrap based forecast 77

LIST OF FIGURES

Figure 1.1 The total consumption of natural gas at residents for Turkey (MENR).......... 4
Figure 4.1 Daily natural gas consumption of Istanbul (IGDAS)................................... 23
Figure 4.2 Winters' method plot with 0,01; 0,02; 0,30 smoothing constants................. 25
Figure 4.3 Forecast and real values of 2007 with 0,01;0,02;0,30 smoothing constants. 26
Figure 4.4 Monthly natural gas consumption of Istanbul from 1993 to 2009................ 28
Figure 4.5 Winters' method plot with 0,15; 0,15; 0,15 parameters............................... 31
Figure 4.6 ARIMA(0,1,1)(0,1,1) residuals' ACF and PACF plots 36
Figure 5.1 Monthly natural gas consumption of Turkey from 1999 to 2008 (IEA)....... 38
Figure 5.2 Residual ACF and PACF plots of ARIMA(0,0,2)(1,1,0)............................. 41
Figure 5.3 Annual natural gas consumption of Turkey from 1987 to 2008 (BOTAS)... 44
Figure 5.4 Trend analysis plot of annual natural gas consumption of Turkey by applying regression model... 47
Figure 5.5 Trend analysis plot of annual natural gas consumption of Turkey by applying quadratic trend model 48
Figure 5.6 Plot of annual natural gas consumption in Turkey by applying double exponential smoothing model... 49
Figure 5.7 Time series plot after first difference of annual natural gas consumption of Turkey... 51
Figure 5.8 Autocorrelation function plot of first difference of annual natural gas consumption of Turkey.. 52
Figure 5.9 Time series plot after second difference of annual natural gas consumption of Turkey... 54
Figure 5.10 Autocorrelation function plot of second difference of annual natural gas consumption of Turkey.. 55
Figure 5.11 Partial autocorrelation function plot of second difference of annual natural gas consumption of Turkey.. 56
Figure 5.12 ARIMA(1,2,1) residual ACF and PACF plots... 59

Figure 6.1 Natural logarithms of monthly natural gas consumption data of Istanbul 65
Figure 6.2 The plot of seasonally differenced logged natural gas consumption data of Istanbul.. 67
Figure 6.3 First difference of logged and seasonally differenced natural gas consumption data of Istanbul 69
Figure 6.4 Autocorrelation function plot of logged, seasonally differenced and first differenced natural gas consumption data of Istanbul 70
Figure 6.5 Partial autocorrelation function plot of logged, seasonally differenced and first differenced natural gas consumption data of Istanbul... 71
Figure 6.6 K-S Statistics results related with block sizes 73
Figure 6.7 Autocorrelation function plot of a bootstrap 74
Figure 6.8 Partial autocorrelation function plot of a bootstrap....................................... 75
Figure 6.9 Average of consumption scenarios as forecasts of 2008 and 2009 months .. 78

LIST OF SYMBOLS AND ABBREVIATIONS

SYMBOL/ABBREVIATION,

A	:	Actual visitor number in the time series
Ψ	:	The trend factor
C	:	The cyclical factor
ξ	:	The seasonal factor
Θ	:	The irregular factor
AR(*p*)	:	Autoregressive process of order *p*
MA(*q*)	:	Moving average process of order *q*
ARMA (*p, q*)	:	Mixed autoregressive moving average model with orders of *p* and *q*
ARIMA(p, d, q)	:	Autoregressive integrated moving average model with orders of *p*, *d*, and *q*
SARIMA $(p,d,q)\times(P,D,Q)_s$	:	Seasonal autoregressive integrated moving average model with non-seasonal orders of *p, d, q* and seasonal orders of *P, D, Q*
BOTAS	:	Boru Hatlari ile Petrol Tasima A.S.
IGDAS	:	Istanbul Gaz Dagitim A.S.
IEA	:	International Energy Agency
MENR	:	Ministry of Energy and Natural Sciences

CHAPTER 1

INTRODUCTION

Forecasting is very important in many types of organizations since predictions of future events must be incorporated into the decision making process. The government of a country must be able to forecast such things as air quality, water quality, unemployment rate, inflation rate, and welfare payments in order to formulate its policies. Business firms, in particular, require forecasts of many events and conditions in all phases of their operations. Forecasters use past data and must base the forecast on the results of this analysis. Forecasters use past data in the following way. First, the forecaster analyzes the data in order to identify a pattern that can be used to describe it. Then this pattern is extrapolated, or extended, into the future in order to prepare a forecast. The basic strategy is employed in most forecasting techniques and rests on the assumption that the pattern that has been identified will continue in the future (Bowerman, O'Connell, Koehler, 2004).

The choice of forecasting method depends on a variety of considerations, including:

- How the forecast is to be used.
- The type of time series and its properties, such as presence/absence of trend and/or seasonality. Some series are very regular and hence 'very predictable', but others are not. As always, a time plot of the data is very helpful.
- How many past observations are available.
- The length of the forecasting horizon. For example, short-term forecasting.

- The skill and experience of the analyst and the computer programs available. The analyst should select a method he feels ‘happy’ with, and also consider the possibility of trying more than one method (Chatfield, 1996).

Monitoring and controlling forecasts is very important to make accurate predictions. Different types of error measurements such as MAD can be employed to see forecasting errors.

Regardless of the system, each company faces that forecasts are seldom perfect. This means that outside factors that we cannot predict or control often impact the forecast. Companies need to allow for this reality (Heizer and Render, 2001).

Demand forecasting is very important to enable efficient and economic operation of the relevant system. Energy consumption forecasting is a type of demand forecasting. Natural gas production and transmission businesses require accurate consumption forecasts to satisfy maximum amount of natural gas needs.

Usually, weather conditions and consumer behaviors affect the natural gas consumption values strongly. These factors are taken into considerations by forecasters and different types of forecasting methods are applied to find out future values.

In this study different kinds of univariate forecast techniques are employed in order to investigate which technique is appropriate to estimate future natural gas consumption values. In addition to these techniques we used bootstrap method to generate consumption scenarios. This approach is realized to draw better forecast range compared to other forecasting methods’ prediction limits.

We have daily and monthly Istanbul natural gas consumption data provided by Istanbul Gaz Dagitim A.S.(IGDAS) and Boru Hatlari ile Petrol Tasima A.S. (BOTAS). Monthly and annual natural gas consumption data of Turkey are provided by Boru Hatlari ile Petrol Tasima A.S. (BOTAS) and International Energy Agency (IEA) (BOTAS and IGDAS Reports). Box-Jenkins methods such as ARIMA and SARIMA are utilized to make forecasts. Trend analysis and Winters’ method are also applied to fitting time series data of natural gas consumption. These methods are compared in order to suggest a forecast method. Using bootstrap method we generated natural gas consumption scenarios about Istanbul.

1.1 NATURAL GAS CONSUMPTION FORECASTING

Turkey is an important candidate to be the ‘‘energy corridor’’ in the transmission of the abundant oil and natural gas resources of the Middle East and Middle Asia countries to the Western market. Furthermore, Turkey is planning to increase its oil and gas pipeline infrastructure to accommodate its increased energy consumption. Naturally, Turkish natural gas usage is projected to increase remarkably in coming years, with the prime consumers, expected to be industry and power plants (Kilic, 2006).

Estimation of natural gas demand is an important part of gas production and transmission business. The challenges of this forecasting are the volatility of consumer profile, the strong dependency on weather conditions and the lack of historical data (Viet and Mandziuk, 2000).

1.1.1 Importance of Natural Gas Usage

Energy is one of the most important inputs required to maintain social and economical improvement in a country. It is necessary that energy demand should be performed at the right time economically, and be of good quality and respectful of increasing environmental consciousness in order to preserve national development and a high standard of living. Natural gas is an alternative energy source that has cleanliness, burning easiness, high thermal value and resource availability (Aras and Aras, 2003).

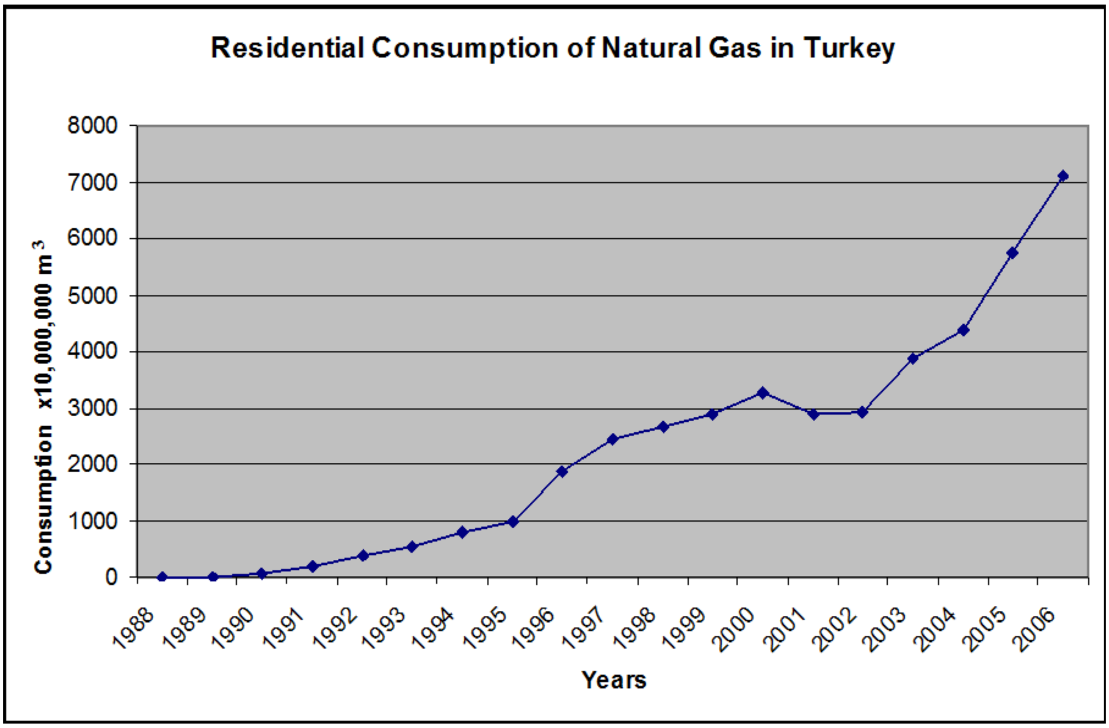

Figure 1.1 The total consumption of natural gas at residents for Turkey (MENR)

All over the world, the use of natural gas is projected to nearly double between 1999 and 2020, providing a relatively clean fuel for efficient new gas turbine power plants. The largest increases in gas use are expected in Central and South America and in developing Asia, and the developing countries as a whole are expected to add a larger increment to gas use by 2020 than are the industrialized countries. Turkish energy consumption has risen dramatically over the past 20 years due to the combined demands of industrialization and urbanization. For instance, residential use of natural gas increases through years (Figure 1.1). Turkey is located at a strategic place between the Middle and Near East, where rich oil and natural gas reserves prevail, and the Western world, where the main energy consumption takes place. Turkey is also situated near the Caspian Sea, where natural gas and oil production are expected to increase substantially. Turkey has made a remarkable contribution to the stability of the region and still continues to maintain this policy. It is accepted that creating a balanced international cooperation setting is an important factor for acquiring more reliable energy supply (Ozturk and Hepbasli, 2003).

1.1.2 Review of Studies about Natural Gas Consumption

There is wide range of studies about forecasting natural gas consumption in the literature. Liu and Lin (1991) estimated the residential consumption of natural gas in Taiwan by using linear transfer function method. Brown and Matin (1995) made a study about development of feed-forward artificial neural network models to forecast daily gas consumption in Wisconsin. Durmayaz et al. (2000) estimated the residential heating energy requirement and fuel consumption in Istanbul based on degree-hours method. Khotanzad et al. (2000) has used the artificial neural network (ANN) forecasters with application the prediction of daily natural gas consumption needed by gas utilities. Gumrah et al. (2001) analyzed the factors and their relationships that influencing the gas consumption in Ankara, and they suggested a model based on degree-day concept including annual number of customers, average degree days, and the usage per customer. Sarak and Satman (2003) forecast the residential heating natural gas consumption in Turkey by using degree-day method. Aras and Aras (2003) have described an approach to obtain appropriate models for forecasting residential monthly natural gas consumption in terms of time series analyses and degree-day method. Viet and Mandziuk (2003) analyzed and tested the several approaches to prediction of natural gas consumption with neural and fuzzy neural systems for natural gas load in two different regions of Poland. Siemek et al. (2003) implemented the Hubbert model based upon Starzman modification to describe the possible scenario of the development of the Poland gas sector. Liu et al.(2004) used the support vector regression (SVM) technique for natural gas load forecasting of Xi’an city, and they compared the result with the 7-lead day forecasting of neural network based model. Gil and Deferrari (2004) presented a generalized model which predicts mainly the residential and commercial natural gas consumption in urban areas of Argentina, for the short and intermediate ranges of time. Brown et al. (2005) presented the mathematical models for gas forecasting in their study. Gutiérrez et al. (2005) used Gompertz-type innovation diffusion process as a stochastic growth model of natural gas consumption in Spain and compared stochastic logistic innovation modeling and stochastic lognormal growth modeling of a non-innovation diffusion process. Al-Fattah (2006) presents a methodology for developing forecasting models for predicting U.S. natural gas production, proved reserves, and annual depletion to year 2025 using time series modeling approach. Kenisarin and Kenisarina (2006) investigated the energy saving potential in the residential sector of Uzbekistan. Ivezić (2006) showed the results of investigation of an artificial neural network (ANN) model for short term natural gas consumption forecasting. This methodology uses multilayer artificial neural networks to incorporate historical weather and consumption data. Wong-Parodi et al. (2006) compared the accuracy of the forecasts for the natural gas prices of Energy Information Administration's short term energy outlook and the futures market for the period from 1999 to 2004. Potocnik et al. (2007) proposed a strategy to estimate forecasting risk of natural gas consumption in Slovenia. This strategy combines an energy demand forecasting model, an economic incentive model and a risk model. Sanchez-Ubeda and Berzosa (2007) presented a model based on decomposition approach to capture demand patterns in a very large number of different historical profiles. Ediger and Akar (2007) used ARIMA and SARIMA methods to estimate the future primary energy demand of Turkey from 2005 to 2020. Kızılaslan and Karlık (2009) used seven neural networks algorithms as forecasting models they tried to find the best solution on forecasting of monthly natural gas consumption.

1.1.3 Characteristics of Natural Gas Consumption

The usage of the natural gas can be classified into 3 groups; the residential users, the industrial users and the commercial users. The demand characteristics of these three categories differ significantly. The residential customer demands are typically temperature sensitive, increasing on weekends. The commercial customers are also typically temperature sensitive, but decreasing on weekends. Industrial customer demand is much less temperature sensitive, decreasing significantly on weekends. Historically, many methods have been used to predict daily demand. Gas controllers have used methods such as looking at use patterns on similar historical days and scatter plots of use versus temperature. Often these methods are only successfully applied by experts with years of experience at a Local Distribution Company (LDC). LDC firms are taken into consideration as distributors for cities. Importation and distribution of natural gas to cities is undertaken presently by BOTAS. The city gas distribution companies such as IGDAS in Istanbul, EGO in Ankara, BURSAGAZ in Bursa, ESGAZ in Eskisehir and IZGAZ in Izmit are the distributors of natural gas (Aras and Aras, 2003).

LDC faces many challenges in the business of supplying gas to their customers. The gas supply system of an LDC consists of gate stations, compressors, gas storage, and customers. The LDC must operate these systems to assure delivery of gas in adequate volumes at required pressures under all circumstances. For efficient, economical, and safe operation, the daily gas demanded by the customers must be known in advance with some degree of accuracy. The customer base of an LDC consists of many individual customers, each with unique demand characteristics. Customers use gas for space heating, known as heating load, for heating water, drying, cooking and baking, and other processes, known as base load, and for electric power generation (Brown et al., 2005).

CHAPTER 2

FORECASTING TIME SERIES DATA

2.1 WHAT IS FORECASTING?

Forecasting is a methodology that helps us for the estimation of the future. Forecasting is an important activity in economics, commerce, marketing and various branches of science (Chatfield, 2000). A forecast may be short-range, medium-range or long-range forecast. Short-range is usually 3 months, medium-range is up to 3 years and long-range is over 3 years. There are three main types of forecasts; economic forecasts, technological forecasts and demand forecasts (Heizer and Render, 2001).

Generally, there are two different forecasting approaches; one is qualitative methods and the other is quantitative methods.

Qualitative methods are subjective and numerical or statistical calculations are not needed for these types of methods. These methods are such as executive opinion, market research and Delphi method.

Quantitative methods are based on mathematical calculations. They grouped into two classes; casual models and time series models. Casual models investigate the relationship between independent and dependent data. Linear regression is the common form of this model (Stevenson, 2007).

Time series models examine the patterns through the past data and estimate the future. This type of data is based on a sequence of evenly spaced (weekly, monthly, quarterly, and so on) data points. Time series forecasting requires time series analyses. That means breaking down past data into components and then projecting them forward. Several time series methods are presented in Table 2.1.

Table 2.1 Several Time Series Methods

Method	Explanation
Simple Moving Average	Simply takes average of past periods
Single Exponential Smoothing	Forecast is the weighted average of the last forecast and the current value of data
Holt’s Linear Exponential Smoothing Technique	This technique allows smoothing trend and slope by using different smoothing constants
Winter’s Method	This is a kind of exponential smoothing adjusted for trend and seasonal variation.

2.2 STATIONARY AND NON-STATIONARY DATA

Stationarity of a data set guarantees spatially invariant statistical properties, such as mean and variance. Stationarity has always played a major role in the time series analysis. Broadly speaking a time series is said to be stationary if there is no systematic change in mean (no trend), if there is no systematic change in variance, and if strictly periodic variations have been removed. In contrast, non-stationary data may have a trend or seasonality. Most of the probability theory of time series is concerned with stationary time series, and for this reason time-series analysis often requires one to turn non-stationary series into a stationary one so as to use this theory (Chatfield, 1996).

2.3 ANALYSIS OF TIME SERIES

The analysis of time series helps to detect regularities in the observations of a variable and derive ‘laws’ from them, and/or exploit all information included in this variable to better predict future developments (Kirchgässner and Wolters, 2007). There are many methods to model and forecast time series. These models are univariate and multivariate models. Univariate models refer to a time series that consists of single (scalar) observations recorded sequentially over equal time increments. Common approaches about univariate models are decomposition of time series, autoregressive (AR) time series model, moving average (MA) model and Box-Jenkins method. Extensions of these models to deal with vector-valued data available under the heading of multivariate time series models.

2.3.1 Classical Decomposition of Time Series

The decomposition into unobserved components that depend on different causal factors, as it is usually employed in the classical time series analysis, was developed by Persons (1919). He distinguished four different components: a long-run development, the trend, a cyclical component with periods of more than one year, the business cycle, a component that contains the ups and downs within a year, the seasonal cycle, and a component that contains all movements which neither belong to the trend nor to the business cycle nor to the seasonal component, the residual.

A multiplicative relationship among the components is assumed by this method as follows:

$$A_t = \Psi_t \cdot C_t \cdot \xi_t \cdot \Theta_t \tag{2.1}$$

where
A = actual visitor number in the time series
Ψ = the trend factor
C = the cyclical factor
ξ = the seasonal factor
Θ = the irregular factor

The first three components are deterministic which are called “signal”, while the last component is a random variable, which is called “noise”. To be able to make a proper forecast, one must know to what extent each component is present in the data (Kirchgässner and Wolters, 2007).

2.3.2 Box-Jenkins Methods

Box and Jenkins (1970) described a new approach to time series analyses. This approach identifies a specific model on the basis of certain statistical figures. They assumed that there was a common stochastic model for the whole generation process of time series. Stages of Box-Jenkins methods are model identification, model parameters estimation, and diagnostic checking.

2.3.2.1 Model Identification

Data should have constant mean, variance, autocorrelation in order to be examined which member of the class ARIMA processes appears to be most appropriate. First of all, take the difference of the data until they are stationary. This is achieved by examining the correlograms of various differenced series until one is found which comes down to zero ‘fairly quickly’ and from which any seasonal cyclic effect has been largely removed. If the data are non-seasonal, an ARMA model can now be fitted. If the data are seasonal, then the SARIMA model may be fitted.

2.3.2.2 Model Parameters Estimation

Secondly, the parameters of this model are estimated. After selection of the model, different orders of the model can be tested to estimate the parameters.

2.3.2.3 Diagnostic Checking

Thirdly, the specification of the model is checked by statistical tests. This essentially consists of examining the residuals from the fitted model to see if there is any evidence of non-randomness. If specification errors become obvious, the specification has to be changed and the parameters have to be re-estimated. The correlograms of the residuals is calculated then can be seen how many coefficients are significantly different from zero and whether any further terms are indicated for the ARIMA model. This procedure is re-iterated until it generates a model that satisfies the given criteria. If the fitted model appears to be inadequate, then alternative ARIMA models may be tried until a satisfactory one is found.

This model can finally be used for forecasts. This univariate forecasting method is to find an appropriate formula so that the residuals are as small as possible and exhibit no pattern. The model-building process involves a few steps, repeated as necessary, to end up with a specific formula that replicates the patterns in the series as closely as possible and also produces accurate forecasts (Chatfield, 1996; Kirchgässner and Wolters, 2007).

2.3.2.4 Autoregressive Model

A time series is said to follow an autoregressive (AR) model of order *p* if the current value of the series can be expressed as a linear function of the previous values of the series plus a random term (Al-Fattah, 2006). Suppose that $\{Z_t\}$ is a purely random process with mean zero and variance σ_Z^2. Then a process $\{X_t\}$ is said to be an autoregressive process of order *p* if

$$X_t = \alpha_1 X_{t-1} + \ldots + \alpha_p X_{t-p} + Z_t \tag{2.2}$$

This is rather like a multiple regression model, but X_t is regressed not on independent variables but on past values of X_t; hence the prefix "auto". An autoregressive process of order *p* is generally abbreviated to an AR(*p*) process (Chatfield, 1996).

2.3.2.5 Moving Average Model

A moving average model is conceptually a linear regression of the current value of the series against the white noise or random shocks of one or more prior values of the series (NIST/SEMATECH e-Handbook of Statistical Methods). Suppose that $\{Z_t\}$ is a purely random process with mean zero and variance σ_Z^2. Then a process $\{X_t\}$ is said to be a moving average process of order *q* (abbreviated to an MA(*q*) process) if

$$X_t = \beta_0 Z_t + \beta_1 Z_{t-1} + \ldots + \beta_q Z_{t-q} \tag{2.3}$$

where $\{\beta i\}$ are constants. The *Z*s are usually scaled so that β_0=1 (Chatfield, 1996).

2.3.2.6 Mixed Autoregressive Moving Average (ARMA) Model

This mixed process contains *p* AR terms and *q* MA terms. An ARMA (*p, q*) process is given by

$$X_t = \alpha_1 X_{t-1} + \ldots + \alpha_p X_{t-p} + Z_t + \beta_1 Z_{t-1} + \ldots + \beta_q Z_{t-q} \tag{2.4}$$

2.3.2.7 Autoregressive Integrated Moving Average (ARIMA) Model

Most of the time series are non-stationary. As mentioned above, stationary time series have constant mean and variance. In order to fit a stationary model, it is necessary to remove non-stationary sources of variation (Chatfield, 1996). Removing trend, seasonality and other variations, then transforming back to original series result autoregressive integrated moving average model (ARIMA). ARIMA(*p, d, q*) is written as:

$$w_t = \mu + \frac{\theta(B)}{\phi(B)}\varepsilon_t \tag{2.5}$$

where

p is the order of the autoregressive component, d the order of the differencing, and q the order of the moving-average process,

w_t = is the response series or a difference of the response series,

μ = a constant or intercept,

B = the backshift operator (i.e. $Bx_t = x_{t-1}$),

$\theta(B)$= the moving average operator, represented as a polynomial in the backshift operator: $\theta(B)$= 1-$\theta_1 B - ... - \theta_q B^q$,

$\phi(B)$= the autoregressive operator, represented as a polynomial in the backshift operator: $\phi(B)$= 1-$\phi_1 B - ... - \phi_p B^q$ and

ε_t = the random error or shock (Al-Fattah, 2006).

2.3.2.8 The Box-Jenkins Seasonal (SARIMA) Model

In practice, many time series contain a seasonal periodic component which repeats every s observations. For example, with monthly observations, where s = 12, we may typically expect X_t to depend on terms such as X_{t-12}, and perhaps X_{t-24}, as well as terms such as $X_{t-1}, X_{t-2}, \ldots$ Box and Jenkins have generalized the ARIMA model to deal with seasonality, and define a general multiplicative seasonal ARIMA model (abbreviated SARIMA model) as

$$\phi_p(B)\Phi_P(B^s)W_t = \theta_q(B)\Theta_Q(B^s)Z_t \tag{2.6}$$

where B denotes the backward shift operator, $\phi_p, \Phi_P, \theta_q, \Theta_Q$ are polynomials of order p, P, q, Q respectively, Z_t denotes a purely random process, and

$$W_t = \nabla^d \nabla_s^D X_t \tag{2.7}$$

The variables $\{W_t\}$ are formed from the original series $\{X_t\}$ not only by simple differencing (to remove trend) but also by seasonal differencing, ∇_s, to remove seasonality. The model in equations above is said to be a SARIMA model of order $(p,d,q)\times(P,D,Q)_s$ (Chatfield, 1996).

CHAPTER 3

TIME SERIES DATA BOOTSTRAP

In this chapter, time series data bootstrap and its methods are explained deeply. Studies about scenario generation by bootstrap and forecasting by bootstrap scenarios are other topics of this section.

3.1 BOOTSTRAP OF TIME SERIES DATA AND BOOTSTRAP METHODS

The bootstrap is a method of computational inference that simulates the creation of new data by resampling from a single data set. Although developed first for cross-sectional data, in recent years several bootstrap methods for dependent data have developed. Like the conventional bootstrap, these methods focus on estimates derived from bootstrap replicates.

There are two approaches to bootstrapping dependent data. One is the model based approach (parametric method), and the other is model free approach (nonparametric methods).

3.1.1 Parametric Bootstrap Method

The parametric bootstrap method first fits a model to the dependent data, and then resamples residuals instead of the original data. A bootstrap sample is generated by using the resampled residuals and the parameters of the model (Romano and Thombs, 1996). Efron and Tibshirani (1986) gave an example of parametric bootstrapping of time series data in which they fit a first order autoregressive model to a data set and calculated the residuals. They applied the parametric bootstrap method by resampling the residuals in order to estimate the standard error of ϕ, which is the estimated parameter of AR (1) process. Thombs and Schucany (1990) applied the parametric bootstrap to autoregressive processes for finding the prediction intervals. Unlike the Box-Jenkins method, which is a classical way of obtaining prediction intervals, the method they proposed did not require normality of the residuals. Extensions of the parametric bootstrap to the mixed autoregressive moving average (ARMA) processes were considered by Thombs (1987) and Souza and Neto (1996). The main difficulty in parametric method is fitting a reasonable model to the data. It requires the modeler to know the correct underlying dependence structure of the series (Carlstein, 1993). ARMA models are not able to model basic features of many real time series, and fitting models beyond ARMA is a very difficult task (Kunsch, 1989).

3.1.2 Nonparametric Bootstrap Methods

The most appealing point of the bootstrap technique is that it is a nonparametric method and, therefore using model free resampling techniques for a dependent data will be more appropriate (Carlstein, 1993)

3.1.2.1 Moving Block Bootstrap

Technique was introduced by Kunch (1989) and Liu and Singh (1992) independently. Let $X_1, X_2, \ldots, X_n$ be the sample of a stationary time series and B_i be the block of b consecutive observations starting from i^{th} observation. $B_i = (X_i, \ldots, X_{i+b-1})$, where $i = 1,2, \ldots, n\text{-}b\text{+}1$, forms $n\text{-}b\text{+}1$ overlapping blocks from the original sample. Resampling n/b blocks with replacement from the set $\{B_1, B_2, \ldots, B_{n-b+1}\}$ produces the bootstrap sample (Romano and Thombs, 1996). In the moving blocks bootstrap (MBB) method, the reason for blocking the series is to preserve the dependence structure within blocks and pass it to the bootstrap sample. The assumption made in MBB is that, for block size b, observations more than b time units apart are independent. MBB becomes the conventional bootstrap when the block size is 1.

3.1.2.2 Linked Blockwise Bootstrap

In moving blocks bootstrap, the dependence structure of the data is disregarded near the block end points. To correct this problem, Kunch and Carlstein (1990) proposed a modification to the moving blocks bootstrap. In the linked blockwise bootstrap, the p nearest neighbors of the final observation of the first block are identified after randomly drawing the first block from the data. Then one of these p nearest neighbors is selected randomly as the starting observation of the second block.

3.1.2.3 Matched-Block Bootstrap

Another modification to the MBB was proposed by Carlstein et al. (1998). They were trying to match the blocks by resampling them according to a Markov chain whose transition probabilities depended on the series. They proposed two types of matching algorithms: the first algorithm is based on a kernel estimate of the conditional lag one distribution, and the second is a fitted autoregression of small order. They made a simulation study comparing the performance of matched-block bootstrap versus moving blocks bootstrap. They generated series from two AR (1) processes, and tried to estimate the variance of the sample mean by bootstrapping. They used the mean squared error (MSE) of the algorithm of the variance as a measure of accuracy. The simulation experiment showed that the matched-block bootstrap gave lower MSE than the MBB. The main conclusion of the study was that matching the blocks reduces the bias significantly but has little effect on variance.

3.1.2.4 Stationary Bootstrap

Stationary bootstrap (SB) was proposed by Politis and Romano (1994) for resampling weakly dependent stationary time series. Just like moving block bootstrap, the stationary bootstrap also uses blocks for resampling. However, in the SB the number of observations in each block is not fixed number b, rather it is a random number from geometric distribution (Leger et al., 1992). SB wraps the series around in a circle before resampling; therefore, unlike the MBB, a block can begin with the last observation of the series. Then it resamples the blocks whose starting points have a discrete uniform distribution $\{1, 2,\ldots, n)$ where n is the number of observations. The average block size will be $1/p$, as the block size has a geometric distribution with design parameter p.

3.1.2.5 Threshold Bootstrap

This method was introduced by Kim et al.(1993b,c) as a generalization of binary bootstrap. This method was modified and theoretically by Park (1997). Park and Willemain (1999) proposed the threshold bootstrap and the threshold jackknife for weakly dependent time series. Willemain et al. (2001) presented the simulation output analysis using the threshold bootstrap. The threshold can either be the sample mean or the sample median of the series. Data sequences above and below the threshold form high and low runs, respectively. A cycle consists of one successive high and low run.

3.1.2.6 Nearest Neighbor Bootstrap

The nearest neighbor bootstrap was developed by Lall and Sharma (1996) for resampling hydrologic time series. It can be applied to both univariate and multivariate processes. This method assumes a known dependence structure, i.e., the number of lags upon which the future flow will depend. The number of lags will determine the dimension of a “feature vector”. For instance if the serial dependence is limited to three previous lags, then the feature vector will be (X_{t-1}, X_{t-2}, X_{t-3}). Then the k nearest neighbors of the feature vectors are determined from the k closest vectors in terms of weighted Euclidean distance. The simulated value is resampled from the successors of these k vectors. Instead of resampling uniformly, a discrete resampling kernel, which decreased monotonically with the increase in distance from feature vector, was used.

There are two choices required in applying nearest neighbor bootstrap. The first is the determining the number of lags to consider, which is a difficult task in real series. The second is choosing the number of neighbors of k.

3.1.2.7 Subseries Method

The subseries method was proposed by Carlstein (1986) to estimate the variance of the statistic of interest. Although this method does not rely on resampling, one could resample the subseries and compute bootstrap replicates (Hall et al., 1995). Let X_1, $X_2,\ldots, X_n$ be a stationary series. The subseries are formed by dividing this series into k non-overlapping subseries of size l, such that $k * l = n$. The subseries S_i, i=1, 2,…,k consists of l consecutive observations starting from $X_{(i-1)l+1}$:S_i={$X_{(i-1)l+1}$, $X_{(i-1)l+2}$,…,$X_{(i-}$

$_{1)l+1}\}$. We want to estimate the Var*(t)*, given that $t = t(X_1, X_2, \ldots, X_n)$ is the statistic of interest. The subseries replicate of t is $t_i = t(S_i)$, i = 1,2,…,k. (Demirel, 2000)

Then the estimator for Var*(t)* is

$$Var(t) = l \cdot \sum_{i=1}^{k} (t_i - \bar{t})^2 / k \tag{3.1}$$

where $$\bar{t} = \sum_{i=1}^{k} t_i / k \tag{3.2}$$

3.2 SCENARIO GENERATION BY BOOTSTRAP

Scenarios are the possible events which will occur in the future. A simulation is mainly based on scenario(s). Model simulation was used to investigate whether scenario generation would show such all options for the best preference. A good scenario generation method should be able to generate large numbers of realistic scenarios. If one is designing a new system or product, inputs are scenarios representing the conditions with which the new system must cope. When the scenarios take the form of univariate stationary time series, the moving blocks bootstrap has the potential to be a good automatic scenario generator. Demirel and Willemain (2002a) determined the proper bootstrap block length. They have developed a method of setting the block length based on the distribution of a statistic computed from zero crossing counts. Higher Order Crossing (HOC) counts of the zero crossings made by the mean-centered series and mean-centered series first, second, third and so on differences (Kedem,1993). To test whether this way of setting the block length results in realistic scenarios, they performed two Turing tests. These visualization experiments confirmed that, when a bootstrap is optimally tuned, it is difficult for sophisticated subjects to identify a bootstrap sample plotted among several real samples.

Simulation modelers frequently face a choice between fidelity and variety in their input scenarios. Using an historical trace provides only one realistic scenario. Using the input modeling facilities in commercial simulation software may provide any number of unrealistic scenarios. Demirel and Willemain (2002b) eased this dilemma by developing a way to use the moving blocks bootstrap to convert a single trace into an unlimited number of realistic input scenarios. They did this by setting the bootstrap block size to make the bootstrap samples mimic independent realizations in terms of the distribution of distance between pairs of inputs. They measured the distance using a new statistic computed from zero crossings. They estimated the best block size by scaling up an estimate computed by analyzing subseries of the trace.

The need for large numbers of realistic scenarios applies to every domain in which simulation is used as part of the system design process. To be useful, scenarios should mimic the underlying data generating process by reflecting the auto- and cross-correlations of the historical data. Huang and Willemain (2006) describe a new scenario generation procedure based on the nearest-neighbor bootstrap. They also propose a new performance evaluation criterion for multivariate time-series scenario generators based on the distribution of a composite correlation discrepancy measure. They illustrate the new method and measure by generating simulated scenarios for the US Treasury yield curve.

3.3 FORECASTING BY BOOTSTRAP SCENARIOS

In this research, we developed on idea of using the bootstrap scenarios to forecast natural gas consumption data. Moving block bootstrap is used to produce bootstrap series of original data.

In order to bootstrap the data, we transformed non-stationary natural gas consumption data into stationary one. After transformation, we produced 100 bootstrap series. These series transformed back into non-stationary series by several calculation steps. Subsequently, we have 100 sibling series. Those series are used to generate 100 different forecasted series by the Winters’ method. That 100 different forecasted series are averaged in our study to produce predictions of 2008 and 2009 months.

CHAPTER 4

FORECASTING NATURAL GAS CONSUMPTION DATA OF ISTANBUL

In this chapter we will examine the daily and monthly natural gas consumption data of Istanbul. Several time series methods are applied and the best of them will be recommended for each data.

4.1 DAILY NATURAL GAS CONSUMPTION DATA OF ISTANBUL

Daily natural gas consumption records exhibits the total daily usage of natural gas in Istanbul. This information includes all consumption data of natural gas for residential and industry usage at Istanbul.

4.1.1 The Plot of Daily Natural Gas Consumption of Istanbul

In order to understand the characteristics of daily natural gas consumption of Istanbul, the daily data is plotted through the period 01.01.2004-31.12.2006. As clearly seen in the Figure 4.1, there is seasonality and trend throughout the days. The seasonal component seems strongly; it is like a top of mountain that starts with November and ends with April months and it is like a bottom of U shape that starts with May and ends with October. That means consumption reaches maximum at winter days. On the other hand it has minimum values through the summer days. Also the plot exhibits an increasing trend. That occurs as a result of new consumers added to natural gas usage system in Istanbul.

For the purpose of handling seasonality and trend, Winters’ forecasting method is selected in this part of study. This method smoothes the data by using Holt-Winters exponential smoothing and calculates the estimates for level, trend and seasonal components of time series (MINITAB).

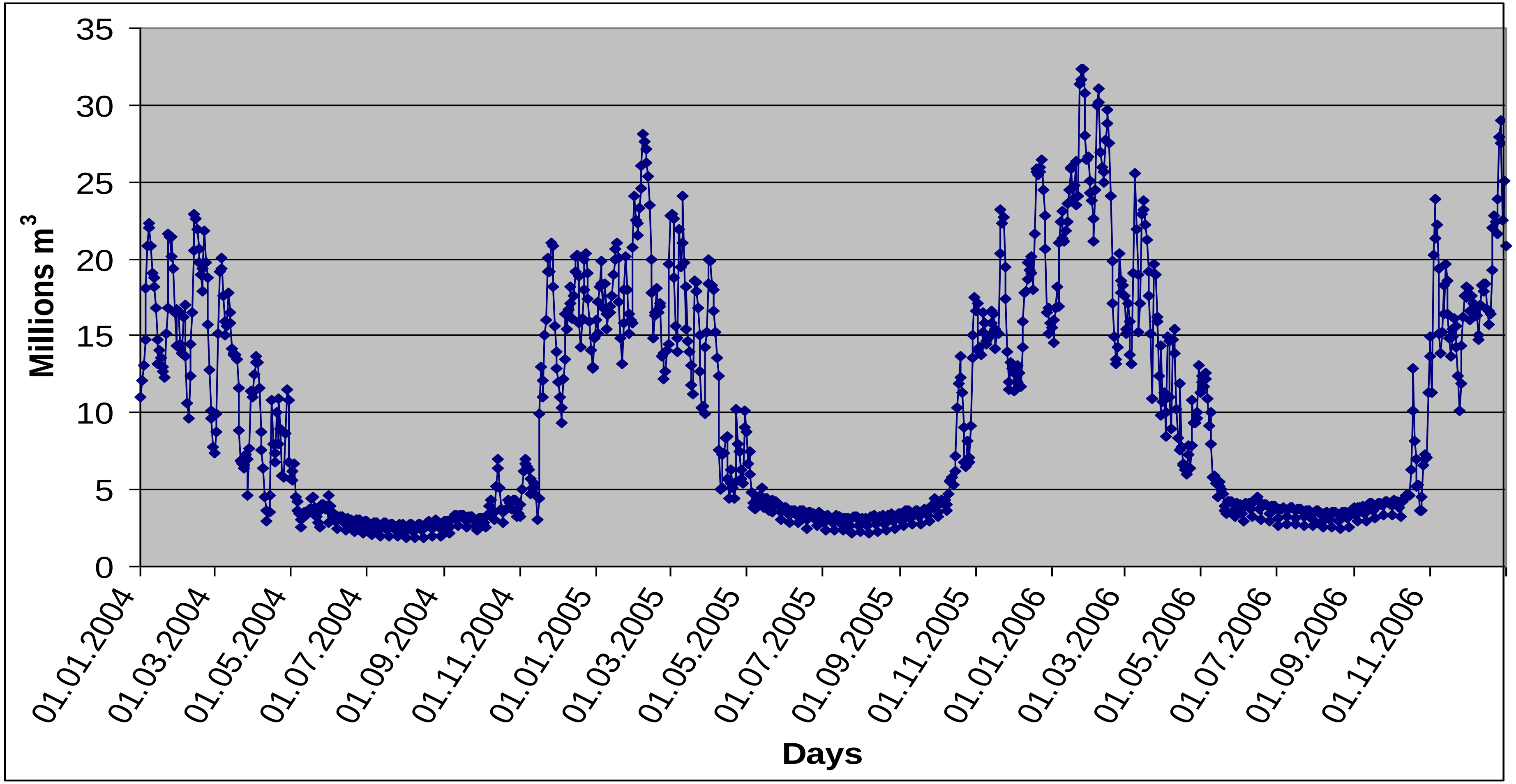

Figure 4.1 Daily natural gas consumption of Istanbul (IGDAS)

4.1.2 Parameter Search

As seen clearly seen above in Figure 4.1, seasonality has great impact on the natural gas consumption data but trend and level terms have not as much as impact like seasonal factor. For stable forecasts the parameter values usually considered between 0,1 and 0,3 in literature. We have decided to fix the seasonal factor to 0,3 and the level parameter to search with 0,01; 0,05; 0,06 and trend parameter to search with 0,01; 0,02; 0,03. The parameter search has shown in Table 3.1.

Table 4.1 Parameters' test results

Parameters	MAD
(0,01;0,01;0,30)	2.345.313
(0,01;0,02;0,30)	*2.035.112*
(0,01;0,03;0,30)	3.001.006
(0,05;0,01;0,30)	2.097.866
(0,05;0,02;0,30)	2.091.689
(0,05;0,03;0,30)	2.069.182
(0,06;0,01;0,30)	2.126.273
(0,06;0,02;0,30)	2.174.865
(0,06;0,03;0,30)	2.210.376

The second line of Table 4.1 with 0,01; 0,02; 0,30 parameters produces better MAD than others. Figure 4.2 and Figure 4.3 are the related plots of 0,01; 0,02; 0,30 smoothing constants that show how these constants fit well to real values.

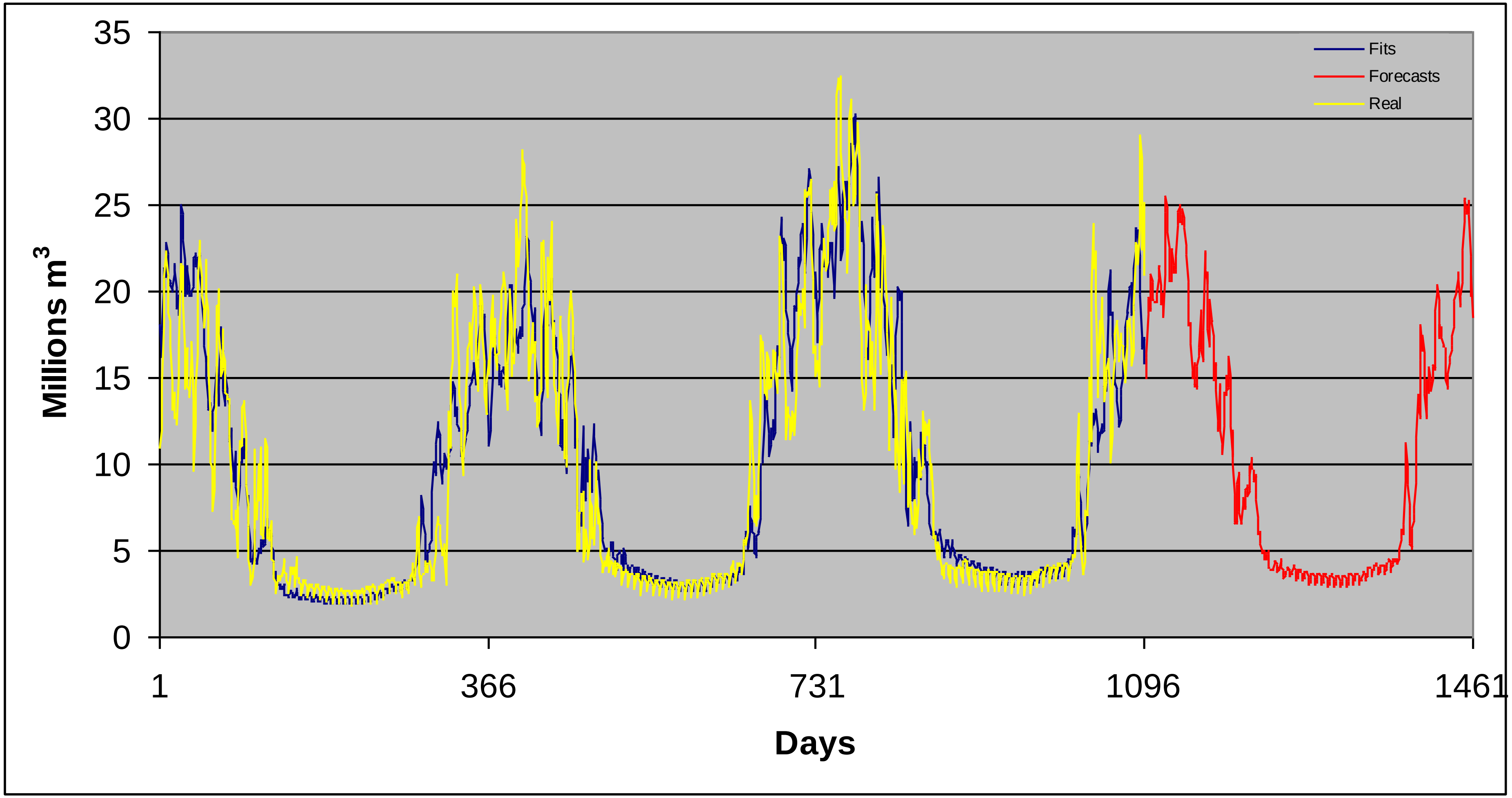

Figure 4.2 Winters' method plot with 0,01; 0,02; 0,30 smoothing constants

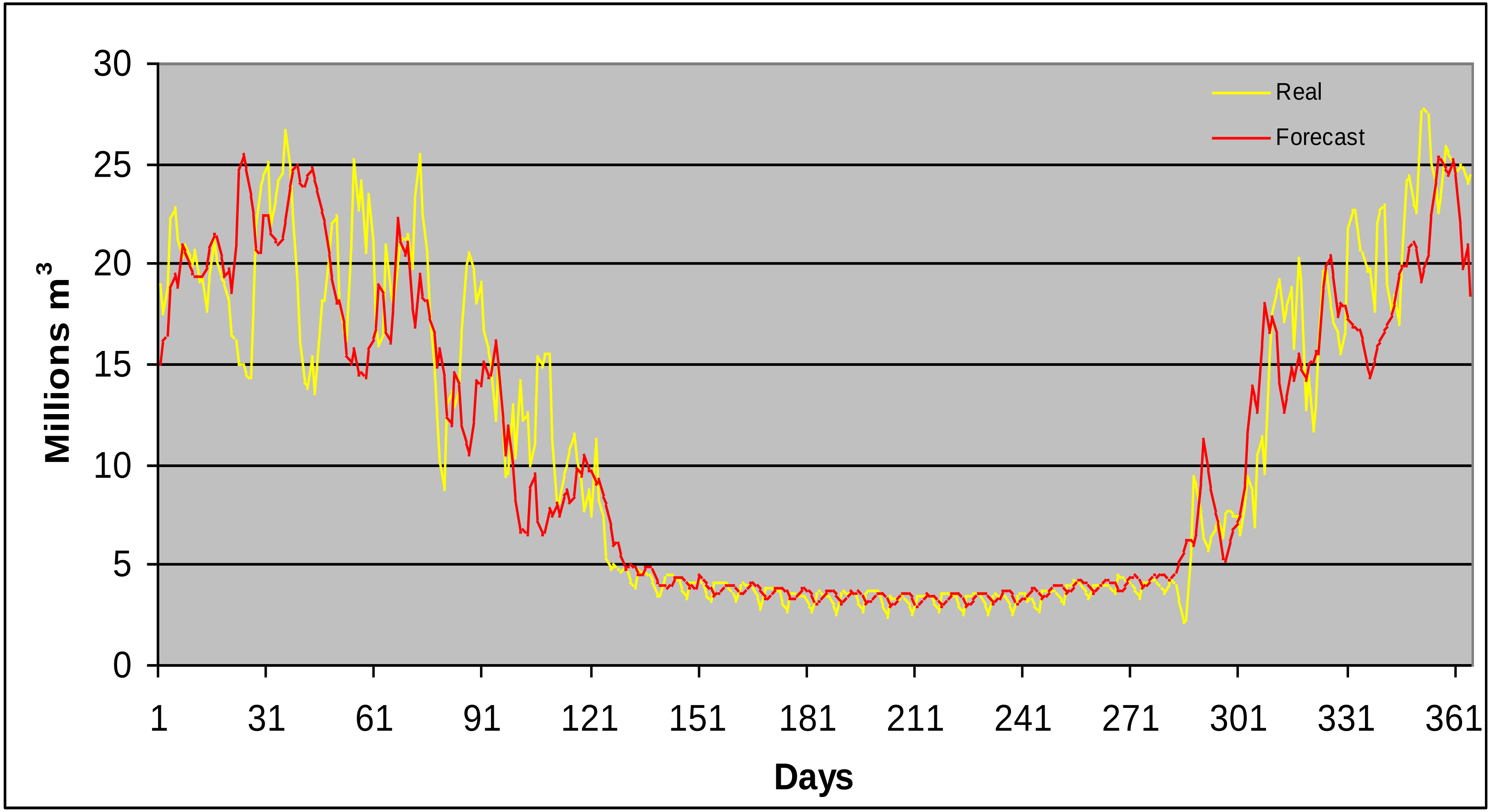

Figure 4.3 Forecast and real values of 2007 with 0,01;0,02;0,30 smoothing constants

When we examine the plot of the data it is clear that the daily natural gas consumption data of Istanbul depends on temperature. The summer consumption values are very low and winter consumption values are high. That is the result of the majority of consumption data is from households. For a better forecasting some should use temperature values in the forecasting model. We do think the results obtained with using temperature values will outperform than our time series method.

4.2 MONTHLY NATURAL GAS CONSUMPTION DATA OF ISTANBUL

Monthly natural gas consumption of Istanbul consists twelve data points for a year. The data is from 10 IGDAS natural gas station consumption and natural gas taken directly from TPAO (Turkish Petroleum Corporation).

4.2.1 The Plot of Monthly Natural Gas Consumption Data of Istanbul

The plot of monthly data illustrates an increasing trend and seasonality in Figure 4.4. Seasonal component of this series is very strong. It appears that every year consumption increases with the rising demand.

4.2.2 Forecasting Monthly Natural Gas Consumption of Istanbul

At this point of study firstly, predictions are made by using SmartForecasts software. SmartForecasts is a forecasting software specifically designed to forecast data for businesses.

The maximum number of data points SmartForecasts can handle is 108 for forecasting 12 points. We used 108 natural gas consumption data points starting from 1999 to 2007 in order to use SmartForecasts. SmartForecasts makes automatic search for which forecasting method is suitable for monthly consumption data. Program compares the methods and reports as a list in Table 4.2.

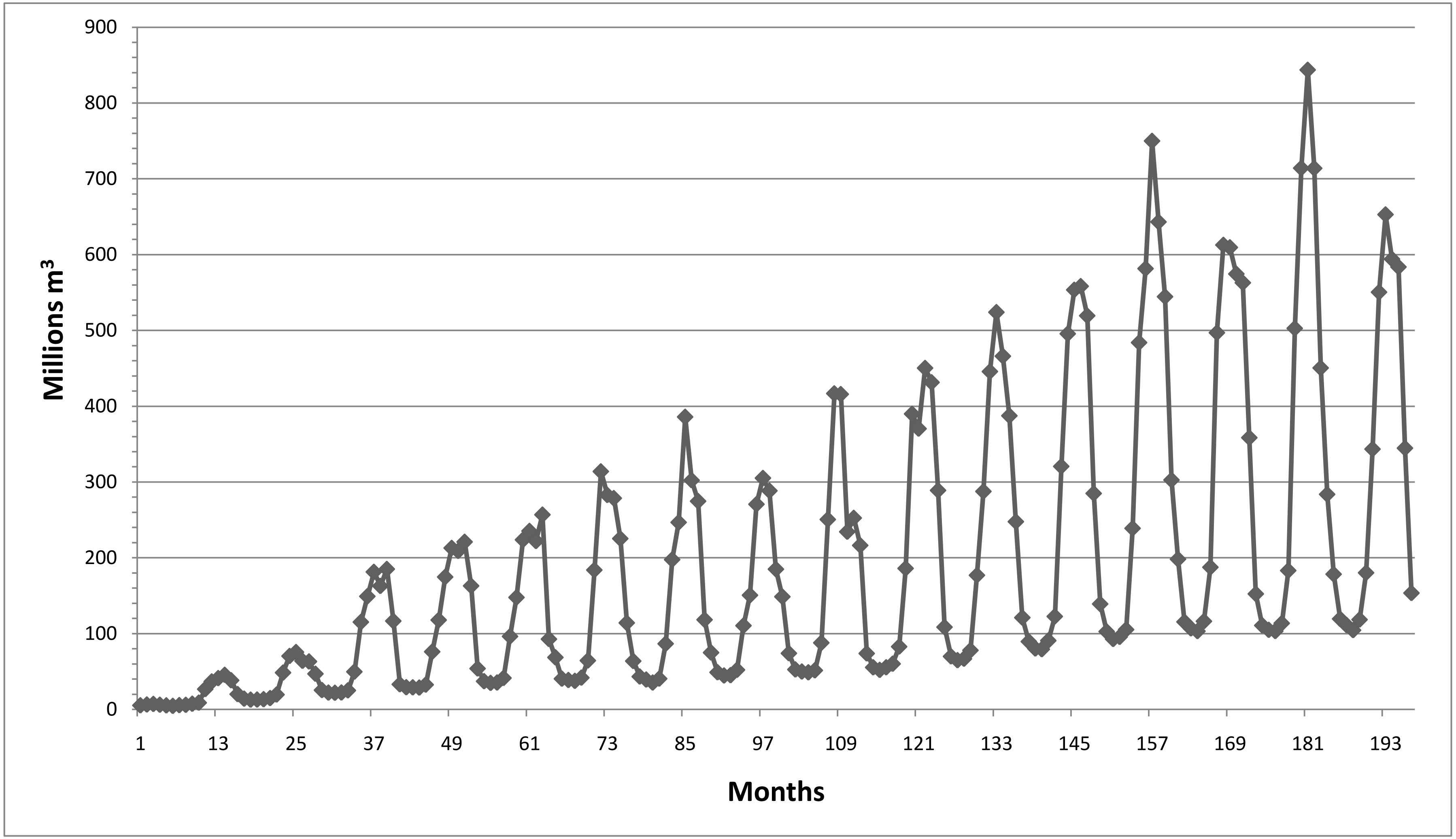

Figure 4.4 Monthly natural gas consumption of Istanbul from 1993 to 2009

Table 4.2 Forecasting methods' results

Rank	% Worse	Avg Error	Forecasting Method
1	(winner)	38773028.00	Winters' Multiplicative, weights = 15% 15% 15%
2	48.0%	57374576.00	Winters' Additive, weights = 9% 9% 9%
3	280.0%	147354096.00	Simple Moving Average of 12 periods
4	296.1%	153581968.00	Double Exponential Smoothing, weight = 3%
5	300.8%	155396880.00	Single Exponential Smoothing, weight = 13%
6	309.8%	158897072.00	Linear Moving Average of 12 periods

According to above results, Winters' multiplicative method presents the minimum average error with 0,15;0,15;0,15 smoothing constants. Forecasts about the year 2008 and 5 months of 2009 are shown below in Table 4.3 using Winters' multiplicative method.

Table 4.3 Winters' Multiplicative Method results

Forecasts of V1 using Winters' Multiplicative Method			
with weights = 15% 15% 15% and seasonal cycle length = 12 periods.			
Based on 108 cases: C1 to C108 .			
Time Period	**Lower Limit**	**Forecasts**	**Upper Limit**
C109	675,062,464	741,058,304	807,054,144
C110	577,301,696	648,262,976	719,224,256
C111	493,846,624	566,773,632	639,700,672
C112	273,310,112	348,020,320	422,730,528
C113	89,844,608	166,103,424	242,362,240
C114	36,632,200	113,786,528	190,940,848
C115	24,535,288	104,504,560	184,473,824
C116	21,498,724	103,196,392	184,894,064
C117	27,264,040	114,185,544	201,107,040
C118	114,357,648	203,004,400	291,651,136
C119	361,471,904	451,654,368	541,836,800
C120	582,065,472	676,583,296	771,101,120
C121	713,710,401	787,053,347	860,396,294
C122	593,261,168	667,663,106	742,065,044
C123	526,155,075	601,648,259	677,141,443
C124	286,329,419	362,944,725	439,560,030
C125	94,996,521	172,763,488	250,530,455

With the 0,15; 0,15; 0,15 smoothing constants, forecasts are plotted in Figure 4.5. Blue line is the real, red line is the lower limit, purple line is the upper limit and green line is the forecasts.

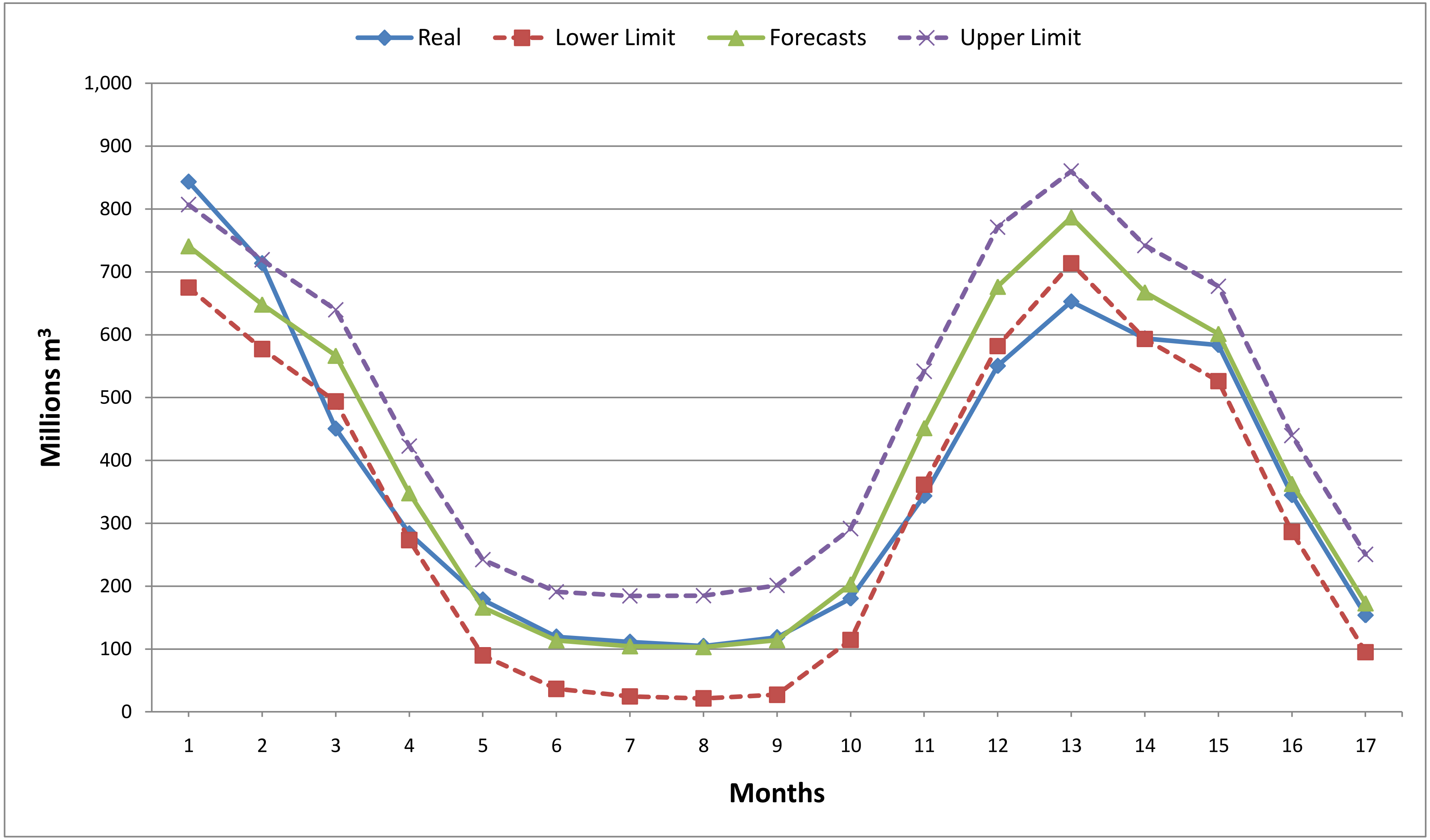

Figure 4.5 Winters' method plot with 0,15; 0,15; 0,15 parameters

SmartForecasts calculations are accepted as starting point, and then we investigated the possible better parameter combination. January, February and March, which also known as the peak consumption months, are selected for the determination of better parameters with lower MAD. Using smoothing constants with levels respectively 0,05; 0,15; 0,30; MINITAB Winters' multiplicative method gives forecasts about 3 months of 2008 and MAD results are calculated in Table 4.4 with the assist of real values of these 3 months.

Table 4.4 Parameter search

Parameters	MAD
[0,05;0,05;0,05]	124.241.060
[0,05;0,05;0,15]	94.381.252
[0,05;0,05;0,30]	86.995.994
[0,05;0,15;0,05]	114.872.310
[0,05;0,15;0,15]	85.848.622
[0,05;0,15;0,30]	94.072.682
[0,05;0,30;0,05]	98.046.461
[0,05;0,30;0,15]	*67.472.348*
[0,05;0,30;0,30]	109.465.478
[0,15;0,05;0,05]	79.053.843
[0,15;0,05;0,15]	78.991.745
[0,15;0,05;0,30]	101.072.490
[0,15;0,15;0,05]	90.485.810
[0,15;0,15;0,15]	90.385.094
[0,15;0,15;0,30]	111.384.939
[0,15;0,30;0,05]	93.498.502
[0,15;0,30;0,15]	92.975.757
[0,15;0,30;0,30]	113.077.317
[0,30;0,05;0,05]	95.389.196
[0,30;0,05;0,15]	83.242.962
[0,30;0,05;0,30]	92.554.031
[0,30;0,15;0,05]	97.146.907
[0,30;0,15;0,15]	84.741.688
[0,30;0,15;0,30]	93.578.490
[0,30;0,30;0,05]	92.614.970
[0,30;0,30;0,15]	79.782.189
[0,30;0,30;0,30]	100.316.148

0,05; 0,30; 0,15 constants give the minimum MAD. Comparisons of real and forecast values are listed in Table 4.5. Winters' method MAD with 0,05 level, 0,30 trend and 0,15 seasonality constant is lower than SmartForecasts Winters' Multiplicative MAD. Note that 0,05; 0,30; 0,15 Winters' method's limits are in Table

4.6 and Winters' method with 0,15; 0,15; 0,15 limits are in Table 4.3. It is clear that Winters' method with 0,05; 0,30; 0,15 parameters produce better forecasts of 2008 and 5 months of 2009.

Table 4.5 Comparisons of forecasts and real value of 2008 and 2009

Year	Month	Real	Winters' Method with 0,05; 0,30; 0,15	Error	Winters' Method with 0,15; 0,15; 0,15	Error
2008	**January**	843,618,178	849,079,233	5,461,056	741,058,304	-102,559,874
2008	**February**	713,833,255	639,215,416	-74,617,839	648,262,976	-65,570,279
2008	**March**	450,535,818	572,873,968	122,338,150	566,773,632	116,237,814
2008	**April**	283,844,118	344,767,357	60,923,239	348,020,320	64,176,202
2008	**May**	178,506,683	162,962,311	-15,544,372	166,103,424	-12,403,259
2008	**June**	119,554,899	111,505,827	-8,049,072	113,786,528	-5,768,371
2008	**July**	111,280,773	102,472,124	-8,808,649	104,504,560	-6,776,213
2008	**August**	104,610,386	100,599,219	-4,011,167	103,196,392	-1,413,994
2008	**September**	118,546,882	110,956,076	-7,590,806	114,185,544	-4,361,338
2008	**October**	180,355,252	196,323,181	15,967,929	203,004,400	22,649,148
2008	**November**	343,461,253	434,005,451	90,544,198	451,654,368	108,193,115
2008	**December**	550,464,497	623,020,610	72,556,113	676,583,296	126,118,799
2009	**January**	652,857,140	809,169,997	156,312,857	787,053,347	134,196,207
2009	**February**	594,104,138	609,052,253	14,948,115	667,663,106	73,558,968
2009	**March**	583,689,727	545,734,590	-37,955,137	601,648,259	17,958,532
2009	**April**	344,758,520	328,369,585	-16,388,935	362,944,725	18,186,205
2009	**May**	153,624,534	155,180,680	1,556,146	172,763,488	19,138,954
			MAD	40,534,383	**MAD**	117,716,283

4.2.3 The Box-Jenkins Seasonal (SARIMA) Model

The time plot of monthly natural gas consumption of Istanbul has strong seasonal component over the periods. Observations with s = 12, one can expect X_t to depend on terms of X_{t-12} and perhaps X_{t-24}, as well as terms such as $X_{t-1}, X_{t-2}, \dots$ (Chatfield,1996).

SPSS has the capability of estimating the models for autoregressive moving average and produces the forecasts. In this part of study, SPSS is used to explore best-fitting ARIMA model.

Monthly natural gas consumption observations of Istanbul from 1993 to 2007 are entered into SPSS spreadsheet. Time series analyses function is selected under the related menu of SPSS. Time series modeler makes a deep search to find the suitable forecasting method.

Then SPSS prints the output onto screen. ARIMA(0,1,1)(0,1,1) is the recommended model for monthly natural gas consumption of Istanbul. SPSS takes the first difference of original data set. MA(1) parameter estimate is 0,302. Also SPSS takes the first seasonal difference of the original data set. Seasonal MA(1)'s parameter estimate is 0,810.

So the ARIMA(0,1,1)(0,1,1) model turns out to be

$$W_t = (1 - 0{,}302B)(1 - 0{,}810B^{12})Z_t \tag{4.1}$$

Model's residual ACF and PACF plots are shown in Figure 4.6. Winters' method with 0,05; 0,30; 0,15 smoothing constants is preferred in the previous section of this chapter. Here comparisons of these two models are listed in Table 4.6.

Table 4.6 Comparison of forecasts by using Winters' method and SARIMA against real value of 2008 and 2009

Year	Month	Real	Winters' Method with 0,05; 0,30; 0,15	Lower Limit	Upper Limit	Error
2008	**January**	843,618,178	849,079,233	762,841,024	935,317,443	5,461,056
2008	**February**	713,833,255	639,215,416	549,886,076	728,544,755	-74,617,839
2008	**March**	***450,535,818***	572,873,968	479,970,935	665,777,000	122,338,150
2008	**April**	283,844,118	344,767,357	247,861,442	441,673,272	60,923,239
2008	**May**	178,506,683	162,962,311	61,675,195	264,249,426	-15,544,372
2008	**June**	119,554,899	111,505,827	5,506,093	217,505,560	-8,049,072
2008	**July**	111,280,773	102,472,124	-8,529,442	213,473,690	-8,808,649
2008	**August**	104,610,386	100,599,219	-15,656,070	216,854,508	-4,011,167
2008	**September**	118,546,882	110,956,076	-10,772,217	232,684,369	-7,590,806
2008	**October**	180,355,252	196,323,181	68,930,863	323,715,499	15,967,929
2008	**November**	343,461,253	434,005,451	300,782,447	567,228,454	90,544,198
2008	**December**	550,464,497	623,020,610	483,821,203	762,220,017	72,556,113
2009	**January**	***652,857,140***	809,169,997	663,866,446	954,473,547	156,312,857
2009	**February**	594,104,138	609,052,253	457,532,258	760,572,248	14,948,115
2009	**March**	583,689,727	545,734,590	387,899,118	703,570,063	-37,955,137
2009	**April**	344,758,520	328,369,585	164,131,025	492,608,144	-16,388,935
2009	**May**	153,624,534	155,180,680	-15,538,718	325,900,077	1,556,146
					MAD	*40,534,383*
Year	**Month**	**Real**	**SARIMA**	**Lower Limit**	**Upper Limit**	**Error**
2008	**January**	843,618,178	798,647,973	535,528,306	1,147,501,684	-44,970,205

Year	Month	Real	SARIMA	Lower Limit	Upper Limit	Error
2008	**February**	713,833,255	739,313,322	451,848,264	1,144,483,456	25,480,067
2008	**March**	450,535,818	684,978,877	386,531,404	1,127,806,900	234,443,059
2008	**April**	283,844,118	423,930,695	222,717,789	736,241,159	140,086,577
2008	**May**	178,506,683	209,141,816	102,902,894	380,854,887	30,635,133
2008	**June**	119,554,899	147,538,166	68,292,610	280,453,278	27,983,267
2008	**July**	111,280,773	139,178,714	60,824,182	275,178,774	27,897,941
2008	**August**	104,610,386	140,143,408	57,993,596	287,363,082	35,533,022
2008	**September**	118,546,882	158,518,529	62,266,656	336,271,678	39,971,647
2008	**October**	180,355,252	279,276,671	104,348,712	611,629,028	98,921,419
2008	**November**	343,461,253	650,057,690	231,456,363	1,467,103,124	306,596,437
2008	**December**	550,464,497	950,541,121	323,031,230	2,207,208,563	400,076,624
2009	**January**	652,857,140	1,066,779,663	337,116,429	2,586,608,677	413,922,523
2009	**February**	594,104,138	991,273,003	294,736,631	2,489,676,429	397,168,865
2009	**March**	583,689,727	921,900,003	258,450,426	2,393,410,351	338,210,276
2009	**April**	344,758,520	572,723,932	151,661,573	1,534,172,978	227,965,412
2009	**May**	153,624,534	283,619,517	71,056,651	782,631,404	129,994,983
					MAD	*117,716,283*

ARIMA(0,1,1)(0,1,1) is referred as SARIMA in Table 4.6. SARIMA model has worse MAD value than Winters' multiplicative method. Also there are bigger gaps between forecast limits in SARIMA compared to Winters' method ones. Only two real values of 2008 March and 2009 February are out of Winters' method's related forecast limits. As the forecasting period lengthens or increases, forecasting accuracy decreases in prediction of monthly natural gas consumption of Istanbul.

It is reasonable that Winters' multiplicative method should be employed to forecast the next monthly gas consumptions of Istanbul with the level, trend and seasonal constants that are 0,05; 0,30 and 0,15 correspondingly.

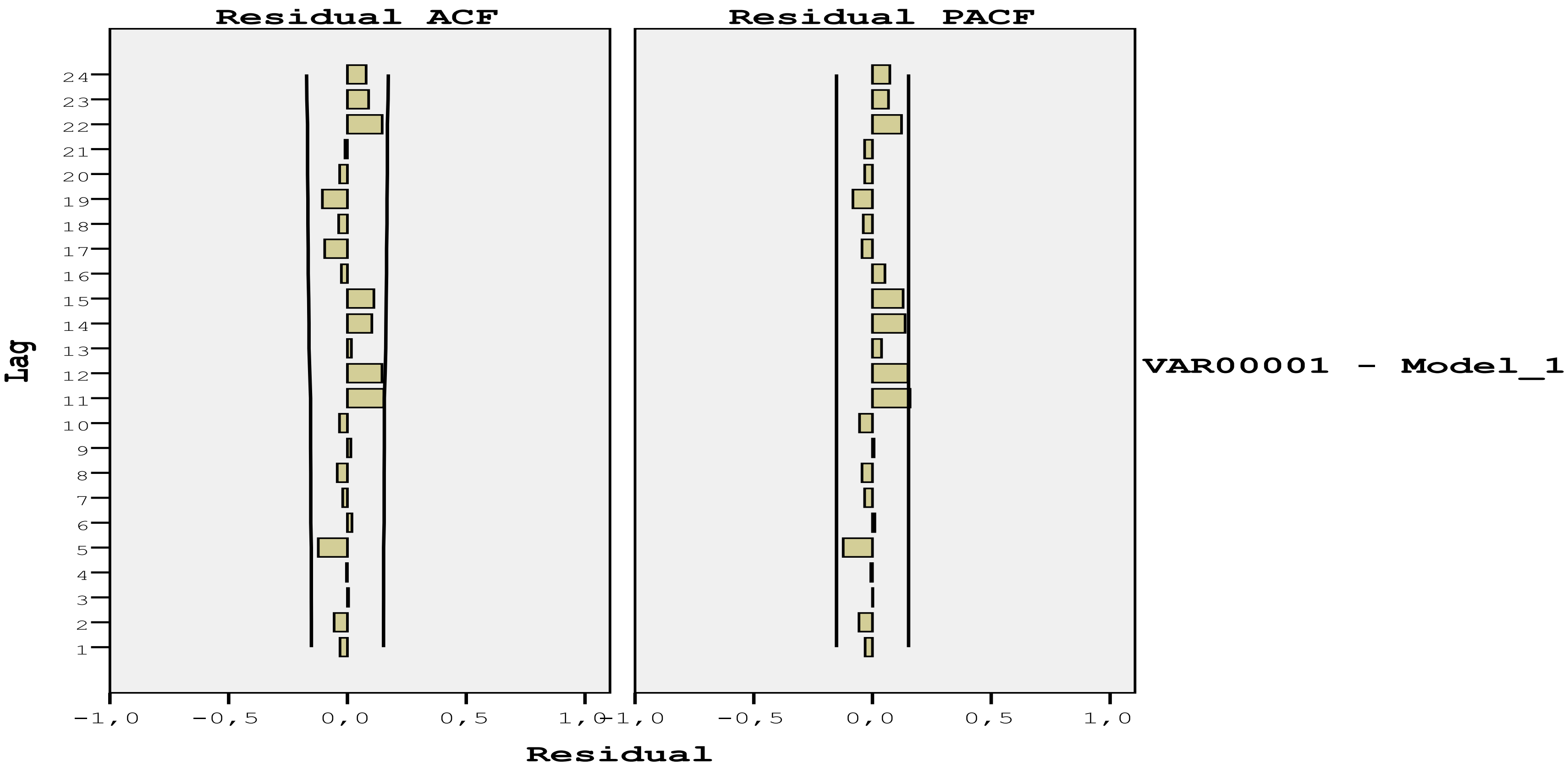

Figure 4.6 ARIMA(0,1,1)(0,1,1) residuals' ACF and PACF plots

CHAPTER 5

FORECASTING NATURAL GAS CONSUMPTION DATA OF TURKEY

In this chapter, natural gas consumption data of Turkey will be examined on monthly and annual time basis. BOTAS and IEA are the providers of the data.

5.1 MONTHLY NATURAL GAS CONSUMPTION DATA OF TURKEY

At this part of study monthly natural gas consumption data of Turkey are taken into account to make predictions about the future. The data covers from 1999 to 2008 years and these 120 points are observed and provided through that period by International Energy Agency (IEA).

5.1.1 The Plot of Natural Gas Consumption Data of Turkey

Natural gas consumption based on months of Turkey exhibits seasonality and increasing trend among nine years as clearly seen in Figure 5.1. As other natural gas consumption data, this one has same pattern; increasing consumption in winters and decreasing consumption in summer periods.

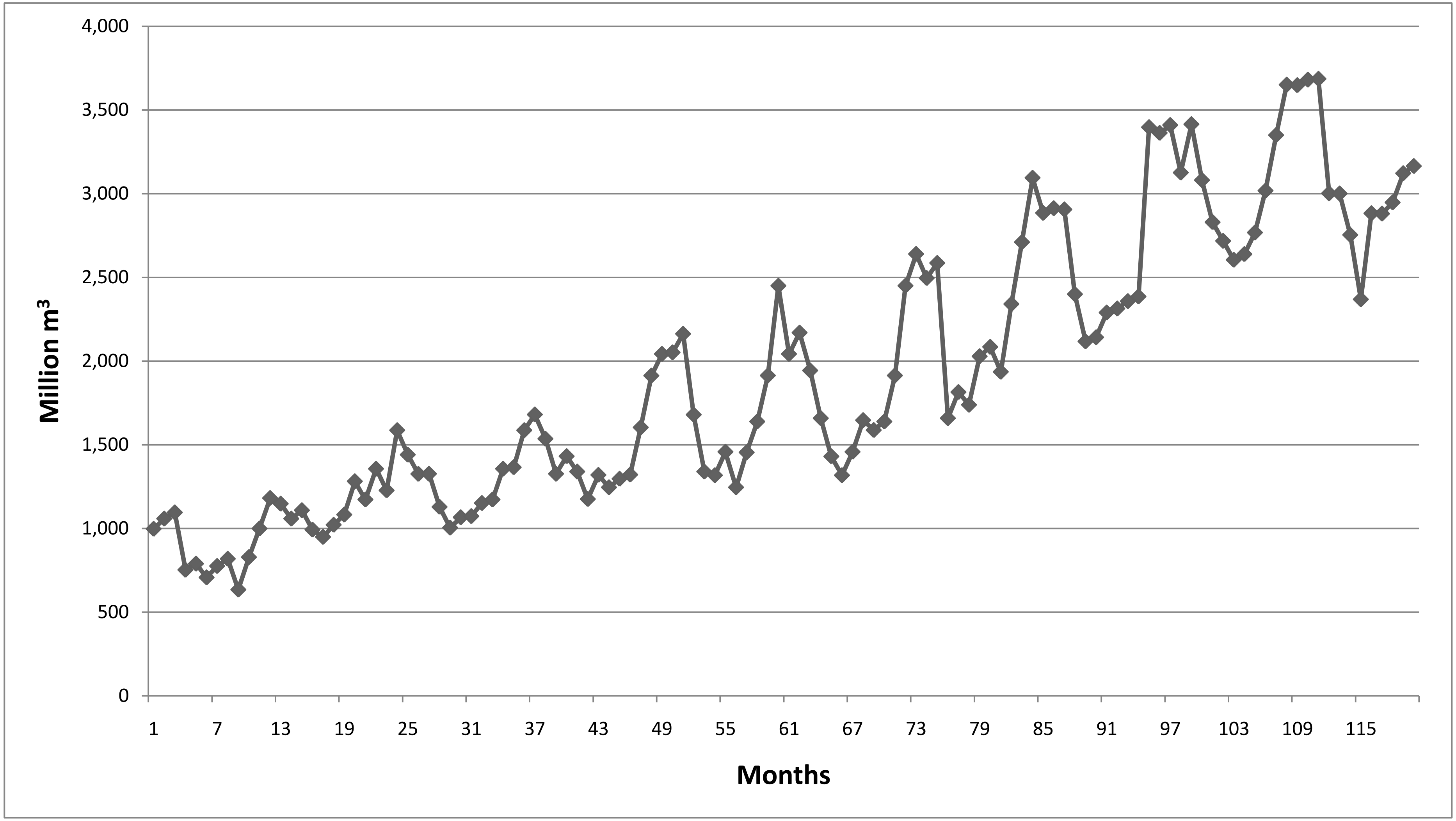

Figure 5.1 Monthly natural gas consumption of Turkey from 1999 to 2008 (IEA)

5.1.2 Forecasting Monthly Natural Gas Consumption

SmartForecasts is run for automatic search on forecasting methods. The data entered into SmartForecasts is from 1999 to 2007. We have natural gas consumption data of 2008, and it would be used for validation. SmartForecasts gives the below results in Table 5.1.

Table 5.1 SmartForecasts results

Rank	% Worse	Avg Error	Forecasting Method
1	(winner)	160.10	Winters' Multiplicative, weights = 26% 26% 26%
2	33.8%	214.21	Winters' Additive, weights = 54% 54% 54%
3	37.1%	219.49	Simple Moving Average of 1 periods
4	37.4%	219.93	Single Exponential Smoothing, weight = 97%
5	76.2%	282.05	Double Exponential Smoothing, weight = 69%
6	82.4%	292.01	Linear Moving Average of 12 periods

Winters' multiplicative method with weights 0,26; 0,26; 0,26 is suggested by SmartForecasts. That method's forecasts about 2008 are presented in Table 5.2.

Table 5.2 Winters' multiplicative method forecasts

2008	Real	Winters' Method with 0,26; 0,26; 0,26	Lower Limit	Upper Limit	Error
January	3647	3,831	3,480	4,182	184
February	3680	3,622	3,261	3,982	-58
March	3685	3,574	3,202	3,945	-111
April	3001	2,874	2,491	3,258	-127
May	***3000***	2,579	2,182	2,975	-421
June	2753	2,442	2,031	2,852	-311
July	2368	2,552	2,127	2,977	184
August	2882	2,576	2,135	3,017	-306
September	2880	2,557	2,100	3,015	-323
October	2947	2,778	2,303	3,252	-169
November	3121	3,241	2,749	3,733	120
December	3164	3,657	3,146	4,167	493
				MAD	234

5.1.3 Autoregressive Integrated Moving Average Model for Forecasting Monthly Natural Gas Consumption of Turkey

At this part of study, monthly natural gas consumption data of Istanbul from 1999 to 2007 are entered into SPSS sheet for determining the Box-Jenkins model.

SPSS Time Series function recommends ARIMA(0,0,2)(1,1,0) model. Model constant is estimated as 263,224. MA(1) parameter estimate is -0,473; MA(2) is also -0,471. Seasonal AR(1) parameter estimate is -0,425. Residual ACF and PACF plots are given in Figure 5.2.

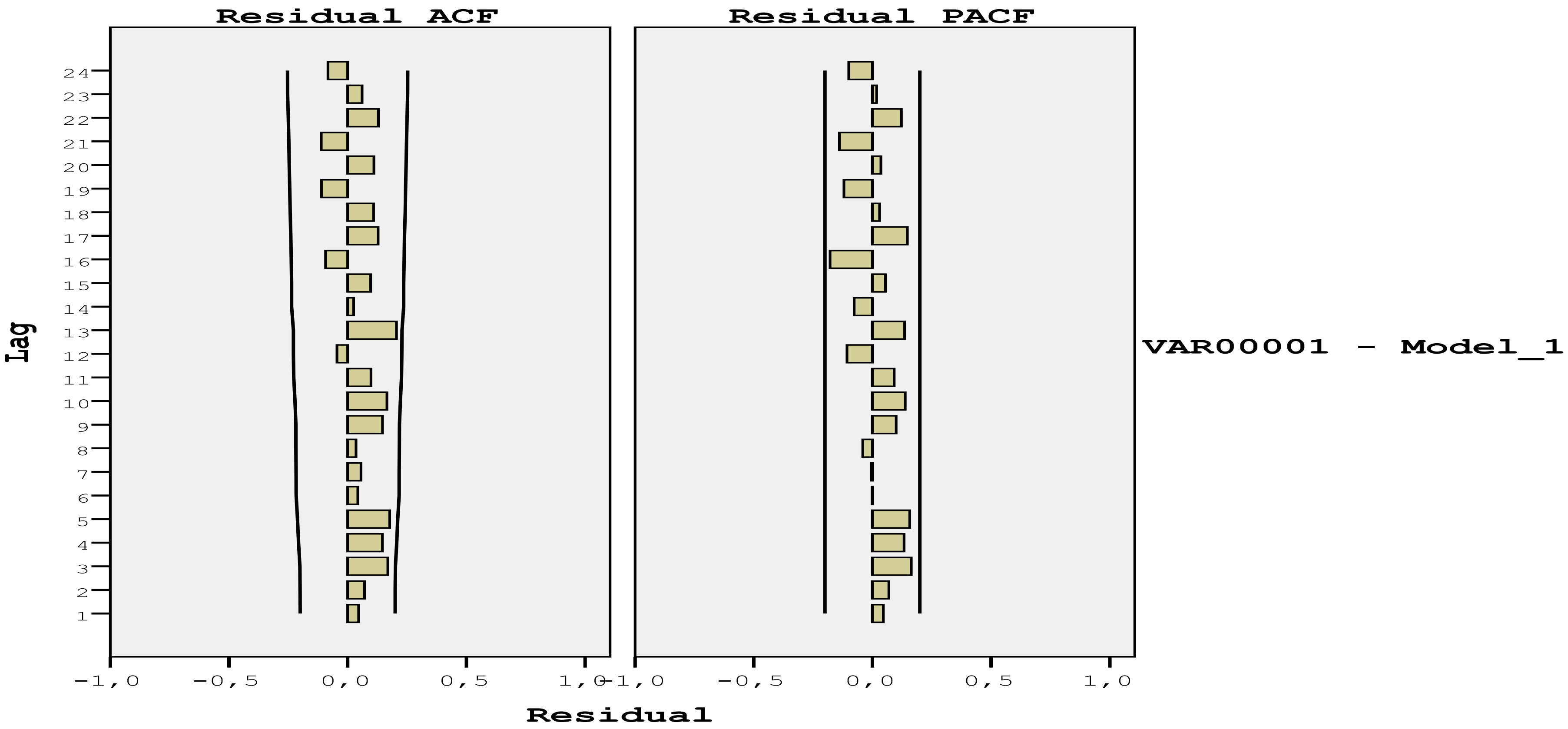

Figure 5.2 Residual ACF and PACF plots of ARIMA(0,0,2)(1,1,0)

Winters multiplicative method with weights 0,26; 0,26; 0,26 and ARIMA(0,0,2)(1,1,0) with their 2008 forecasts are compared below in Table 5.3. (ARIMA(0,0,2)(1,1,0) is abbreviated as SARIMA.)

Table 5.3 Comparison of two forecasting types

2008	Real	Winters' Method with 0,26; 0,26; 0,26	Lower Limit	Upper Limit	Error
January	3647	3,831	3,480	4,182	184
February	3680	3,622	3,261	3,982	-58
March	3685	3,574	3,202	3,945	-111
April	3001	2,874	2,491	3,258	-127
May	***3000***	2,579	2,182	2,975	-421
June	2753	2,442	2,031	2,852	-311
July	2368	2,552	2,127	2,977	184
August	2882	2,576	2,135	3,017	-306
September	2880	2,557	2,100	3,015	-323
October	2947	2,778	2,303	3,252	-169
November	3121	3,241	2,749	3,733	120
December	3164	3,657	3,146	4,167	493
				MAD	233.95

2008	Real	SARIMA	Lower Limit	Upper Limit	Error
January	3647	3,364	2,889	3,839	-283
February	3680	3,323	2,746	3,900	-357
March	3685	3,332	2,683	3,982	-353
April	3001	2,971	2,321	3,621	-30
May	3000	2,715	2,065	3,365	-285
June	2753	2,625	1,975	3,274	-128
July	2368	2,553	1,904	3,203	185
August	2882	2,586	1,936	3,236	-296
September	2880	2,701	2,051	3,351	-179
October	2947	2,916	2,266	3,565	-31
November	3121	3,357	2,707	4,006	236
December	3164	3,604	2,954	4,254	440
				MAD	233.64

Winters' multiplicative method and SARIMA have nearly same MAD values. Winters' method has one real value out of its limits at 2008 May. SARIMA has slightly better MAD value. SARIMA should be selected for forecasting method in monthly natural gas consumption of Turkey estimation.

5.2 ANNUAL NATURAL GAS CONSUMPTION DATA OF TURKEY

Turkey has a grown need for energy sources. Among the increment of natural gas demand in the electricity, fertilizer, residential, industry usages; the overall natural gas consumption rises rapidly.

Annual natural gas consumption data of Turkey is provided by BOTAS. It includes yearly natural gas consumption values from 1987 to 2008 of Turkey.

5.2.1 The Plot of Natural Gas Consumption Data of Turkey

The diagram about annul data of natural gas consumption shows an obvious increasing trend and no seasonality in Figure 5.3. Trend analysis would be meaningful to explore future expectations about annual natural gas consumption of Turkey.

5.2.2 Forecasting the Annual Natural Gas Consumption of Turkey with Trend Analysis

Trend analysis can be done with different ways of time series models. Commonly used models are linear trend model and quadratic trend model. Linear trend model accounts for linearity in the trend. Quadratic trend model accounts for curvature in the trend.

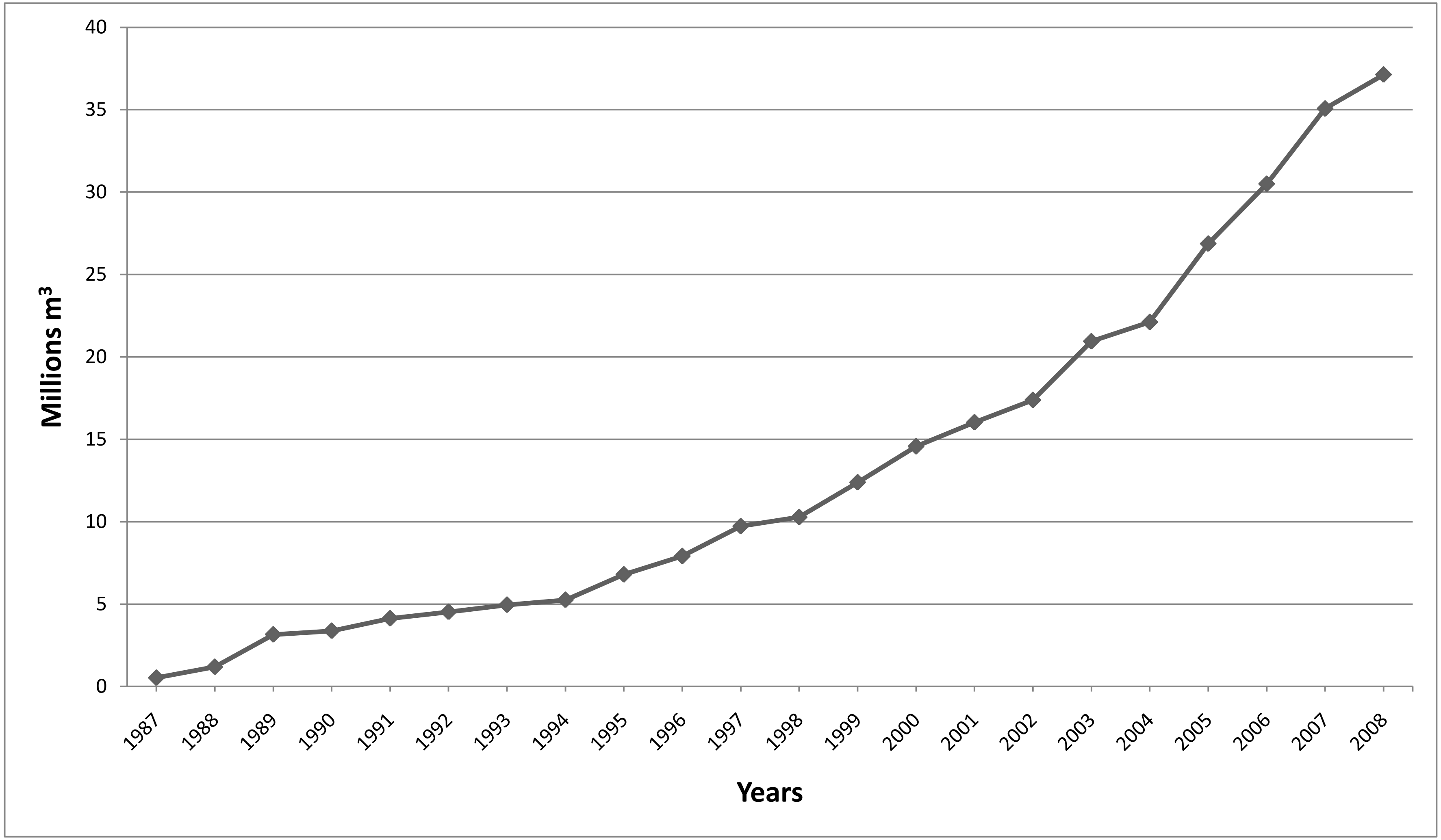

Figure 5.3 Annual natural gas consumption of Turkey from 1987 to 2008 (BOTAS)

According to above models, observations between 1987 and 2006 are used to calculate the forecast about 2007. Also double exponential smoothing is investigated whether or not suitable for annual natural gas consumption of Turkey. Note that double exponential smoothing's level and trend weights are set to optimal by automatically in MINITAB. The plots of all three methods are drawn as Figure 5.4, Figure 5.5 and Figure 5.6 on the next pages respectively. Error values of these models are listed in the Table 5.4.

Table 5.4 Error values of different models

Methods	MAPE	MAD	MSE
Regression	57	2.042	5.750.200
Quadratic Trend	25	676	686.926
Double Exponential Smoothing with 0,23 and 3,43 weights	13	676	741.555

Because the mean absolute percentage error (MAPE) and the mean absolute deviation (MAD) values of quadratic trend and double exponential smoothing are lower than regression in Table 5.4 they are appropriate methods for predictions. Also SmartForecasts is used to guess 2007 consumption. SmartForecasts runs automatic search to select forecasting method with minimum error. Program decided to use double exponential smoothing method with both 0,71 weights. We also investigated the long-term forecasting by making the prediction of 2008 using same parameters with above methods. Forecasts about 2007 and 2008 are shown below in Table 5.5. Note that DES means Double Exponential Smoothing with related parameters. LL is lower limit and UL is upper limit of forecasts.

Table 5.5 Forecasts about 2007 and 2008

Years	Real	DES (0,23; 3;43)	UL	LL	Error	Quadratic Trend	Error	DES (0,71; 0,71)	UL	LL	Error
2007	35,064	34,394	36,049	32,738	-670	31,842	-3,222	34,031	36,440	31,623	-1,033
2008	37,128	39,251	54,445	24,058	2,123	34,928	-2,200	37,864	40,922	34,805	736
				MAD	1,397	**MAD**	2,711			**MAD**	884

Both plots and forecast values demonstrate that double exponential smoothing model fits very well to annual natural gas consumption data. Moreover double exponential smoothing method has minimum error and acceptable lower and upper limits. Tracking the pattern is the concerning criteria which gives an idea about the future. Quadratic trend method and SmartForecasts do not fit as well as double exponential smoothing. So the double exponential smoothing model is suggested to make forecasts in the long term by using 0,71 and 0,71 weights.

For further improvement, we made a parameter search on double exponential smoothing. For level and trend constants we tried 0,2; 0,4 and 0,6 values. Table 5.6 results are shown and also compared with SmartForecasts recommended parameters.

Table 5.6 Parameter search on double exponential smoothing

DES		
Level	**Trend**	**MAD**
0.2	0.2	7811
0.2	0.4	5826
0.2	0.6	3635
0.4	0.2	4482
0.4	0.4	2754
0.4	0.6	2001
0.6	0.2	3019
0.6	0.4	1671
0.6	*0.6*	*866*

(0,6; 0,6) combination has a lower MAD value than others. Its limits and MAD value in Table 5.7 are compared with double exponential smoothing with (0,71; 0,71) parameters.

Table 5.7 Comparison of double exponential smoothing models

DES (0.6; 0.6)						DES (0.71; 0.71)			
Year	**Real**	**Forecast**	**LLr**	**UL**	**Error**	**Forecast**	**LLr**	**UL**	**Error**
2007	35,064	33,437	30,870	36,004	(1,627)	34,031	36,440	31,623	(1,033)
2008	37,128	37,023	33,961	40,086	(105)	37,864	40,922	34,805	736
				MAD	866			**MAD**	884

Double exponential smoothing with 0.6 and 0.6 parameters has better MAD value. So we can take that model as our forecasting model.

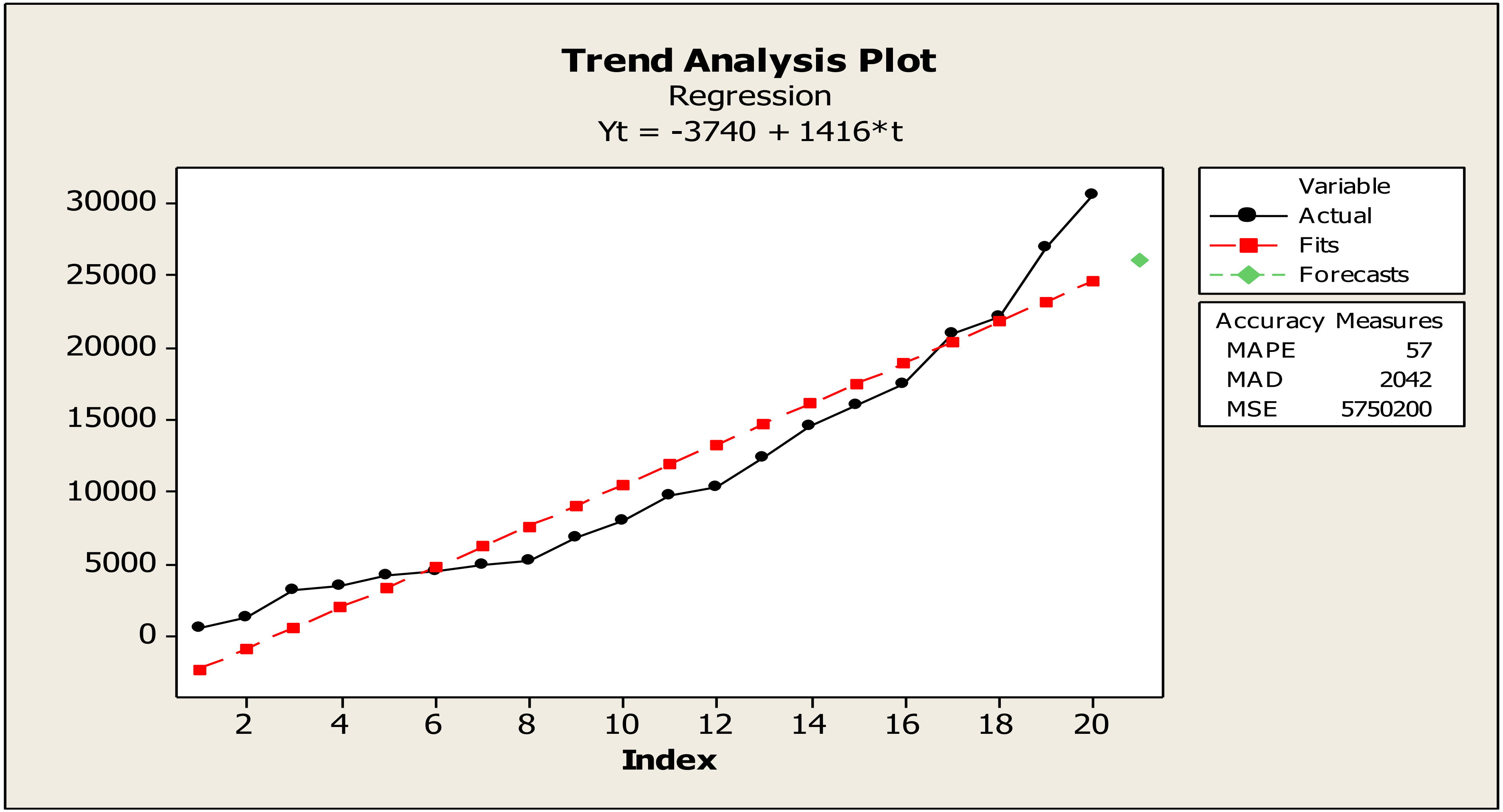

Figure 5.4 Trend analysis plot of annual natural gas consumption of Turkey by applying regression model

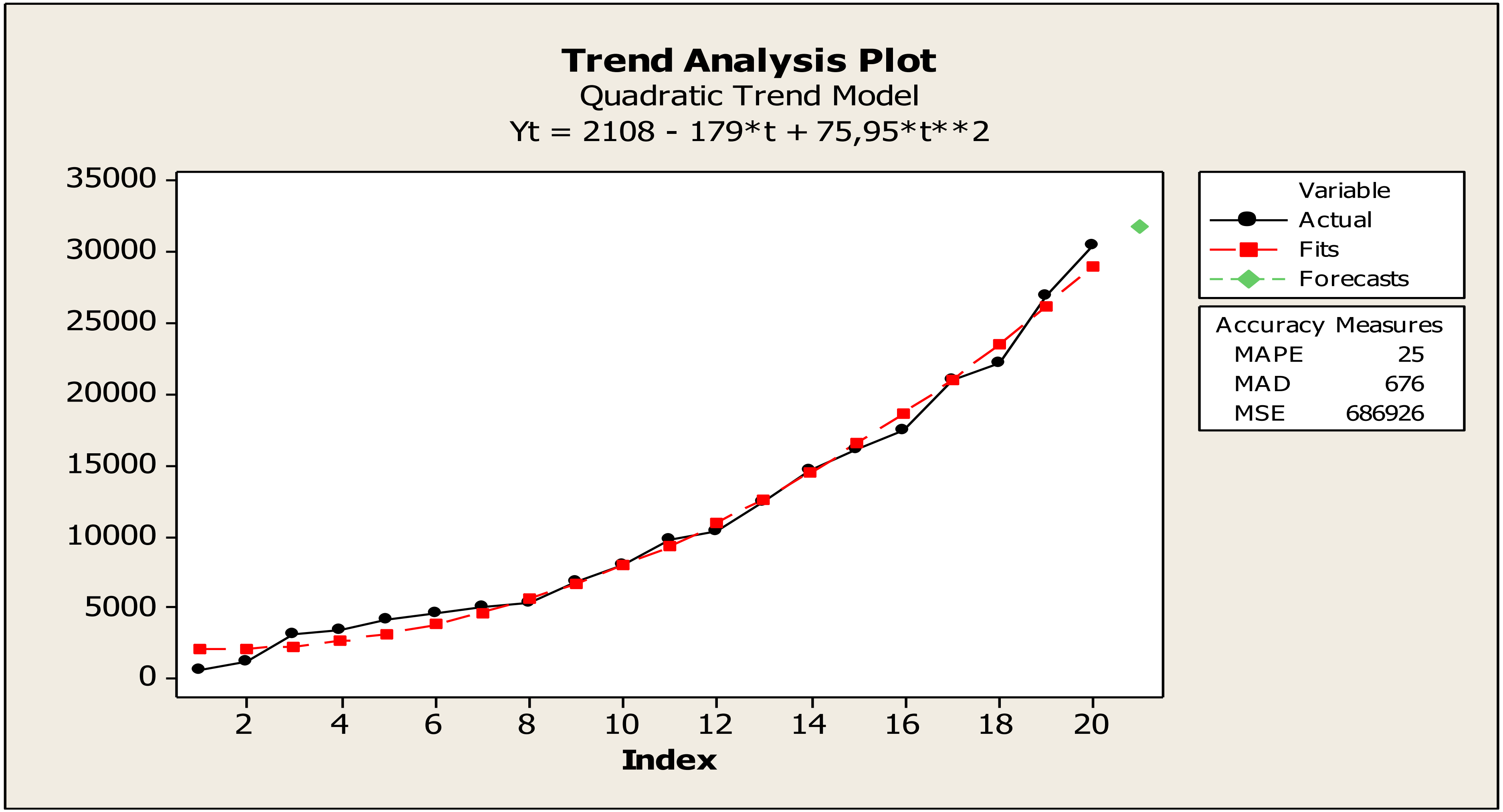

Figure 5.5 Trend analysis plot of annual natural gas consumption of Turkey by applying quadratic trend model

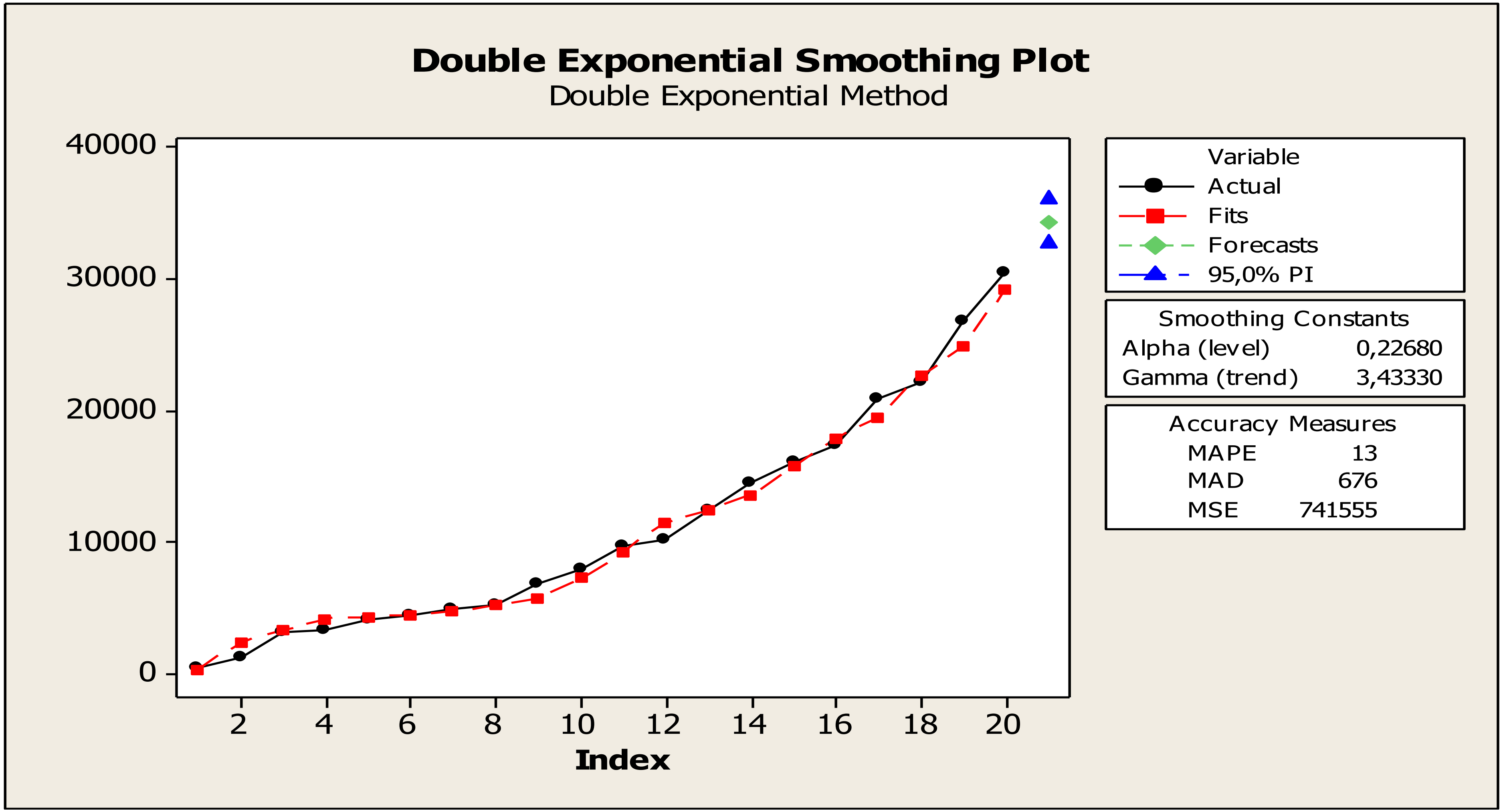

Figure 5.6 Plot of annual natural gas consumption in Turkey by applying double exponential smoothing model

5.2.3 Forecasting the Annual Natural Gas Consumption of Turkey with Autoregressive Integrated Moving Average Model

As clearly seen in Figure 5.3, annual natural gas consumption of Turkey has an increasing trend. That makes the data non-stationary one. This kind of series must be transform into stationary form in order to make time series analysis according to Box-Jenkins procedure.

Differencing is one of the transformation methods to make the data stationary. It is particularly useful for removing a trend. For non-seasonal data, first-order differencing is usually sufficient to attain apparent stationarity, so that the new series $\{y_1, \ldots, y_{N-1}\}$ is formed from the original series $\{x_1, \ldots, x_N\}$ by

$$y_t = x_{t+1} - x_t = \nabla x_{t+1} \tag{5.1}$$

First-order differencing is widely used. Occasionally second-order differencing is required using the operator ∇^2, where (Chatfield, 1996)

$$\nabla^2 x_{t+2} = \nabla x_{t+2} - \nabla x_{t+1} = x_{t+2} - 2x_{t+1} + x_t \tag{5.2}$$

Firstly, we take the first difference of the data and then we plot the new data in Figure 5.7. Time series plot of first difference of the data has a trend. Autocorrelation function (ACF) is reducing slowly in Figure 5.8. The autocorrelation function shows high correlation coefficients with lags. The slow decline of the ACF and time plot of first difference of data suggests that second differencing is required.

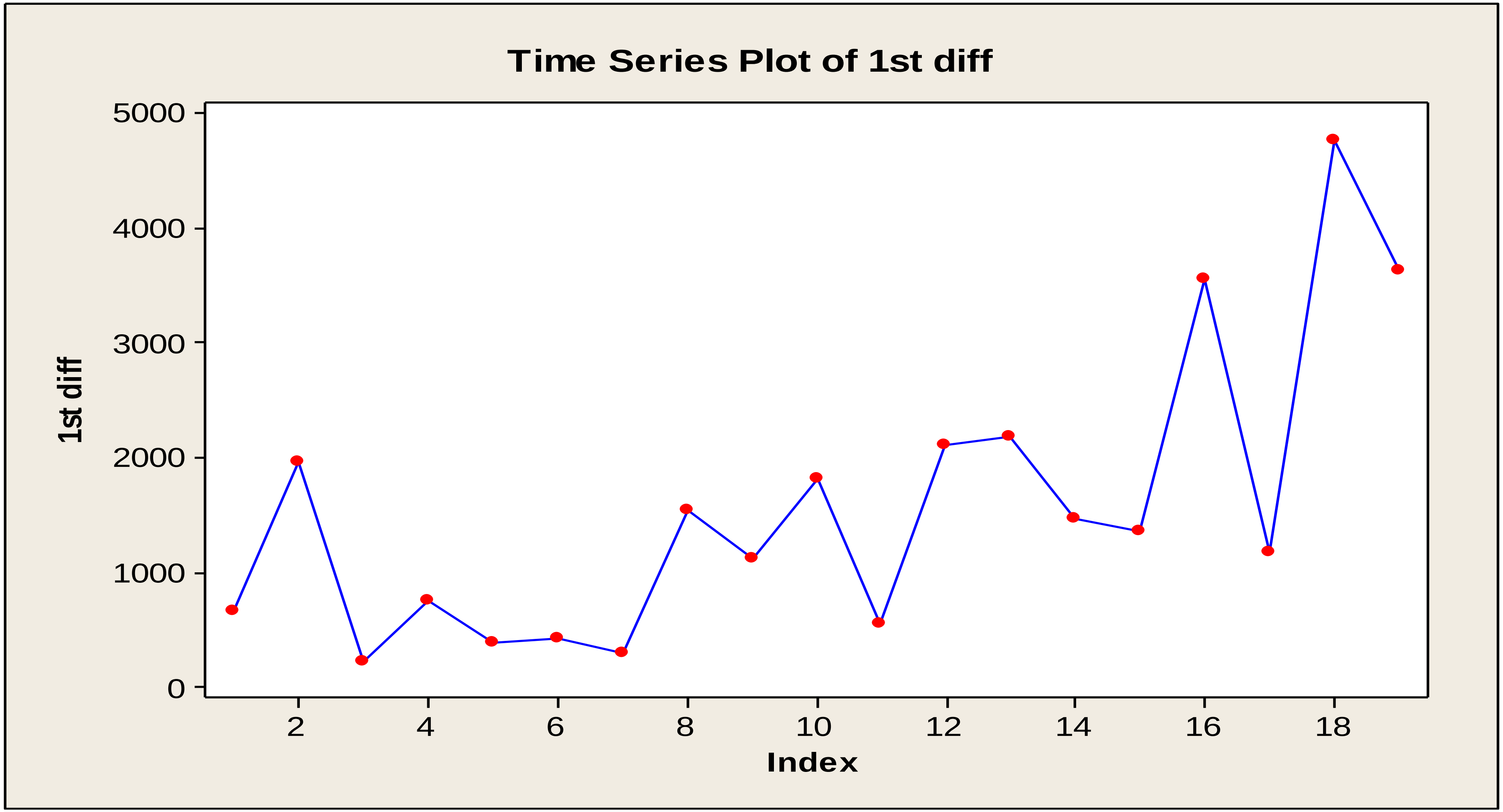

Figure 5.7 Time series plot after first difference of annual natural gas consumption of Turkey

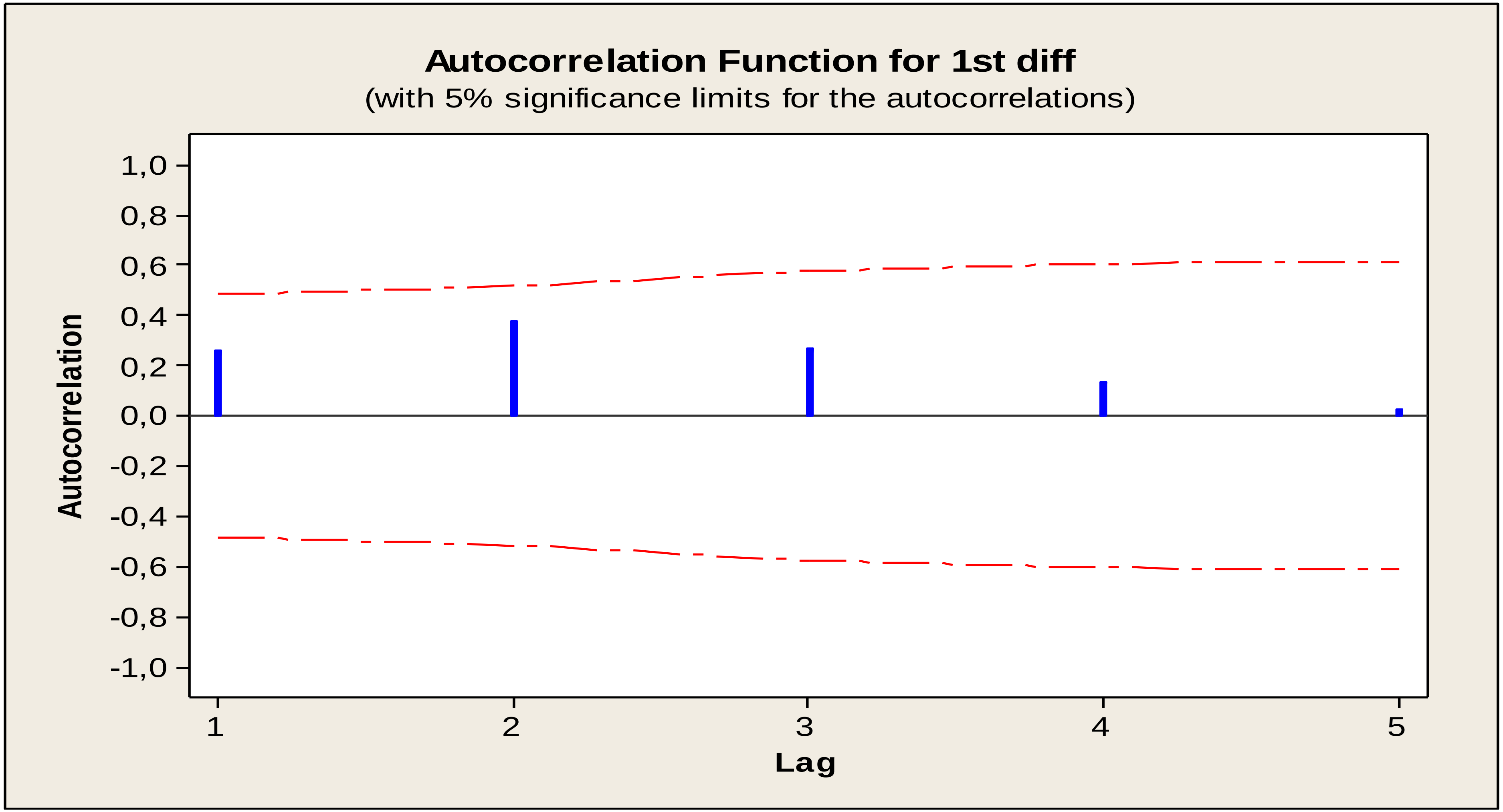

Figure 5.8 Autocorrelation function plot of first difference of annual natural gas consumption of Turkey

Second difference of the data is plotted in Figure 5.9. Now it has no trend, stationarity is obviously seen in Figure 5.9. Its autocorrelation and partial autocorrelation function plots are shown in Figure 5.10 and Figure 5.11 respectively. ACF decreases rapidly and changes its sign at each lag level.

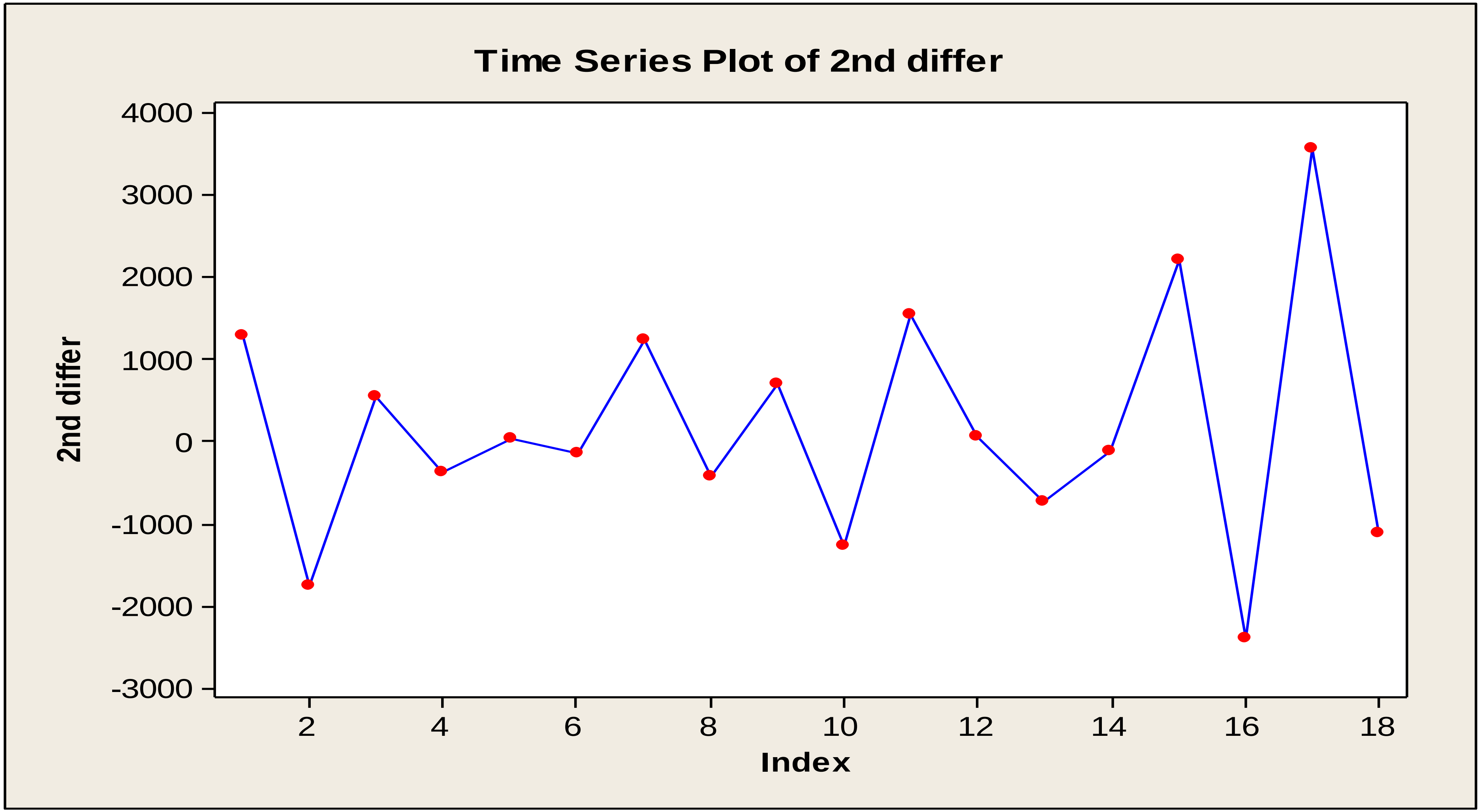

Figure 5.9 Time series plot after second difference of annual natural gas consumption of Turkey

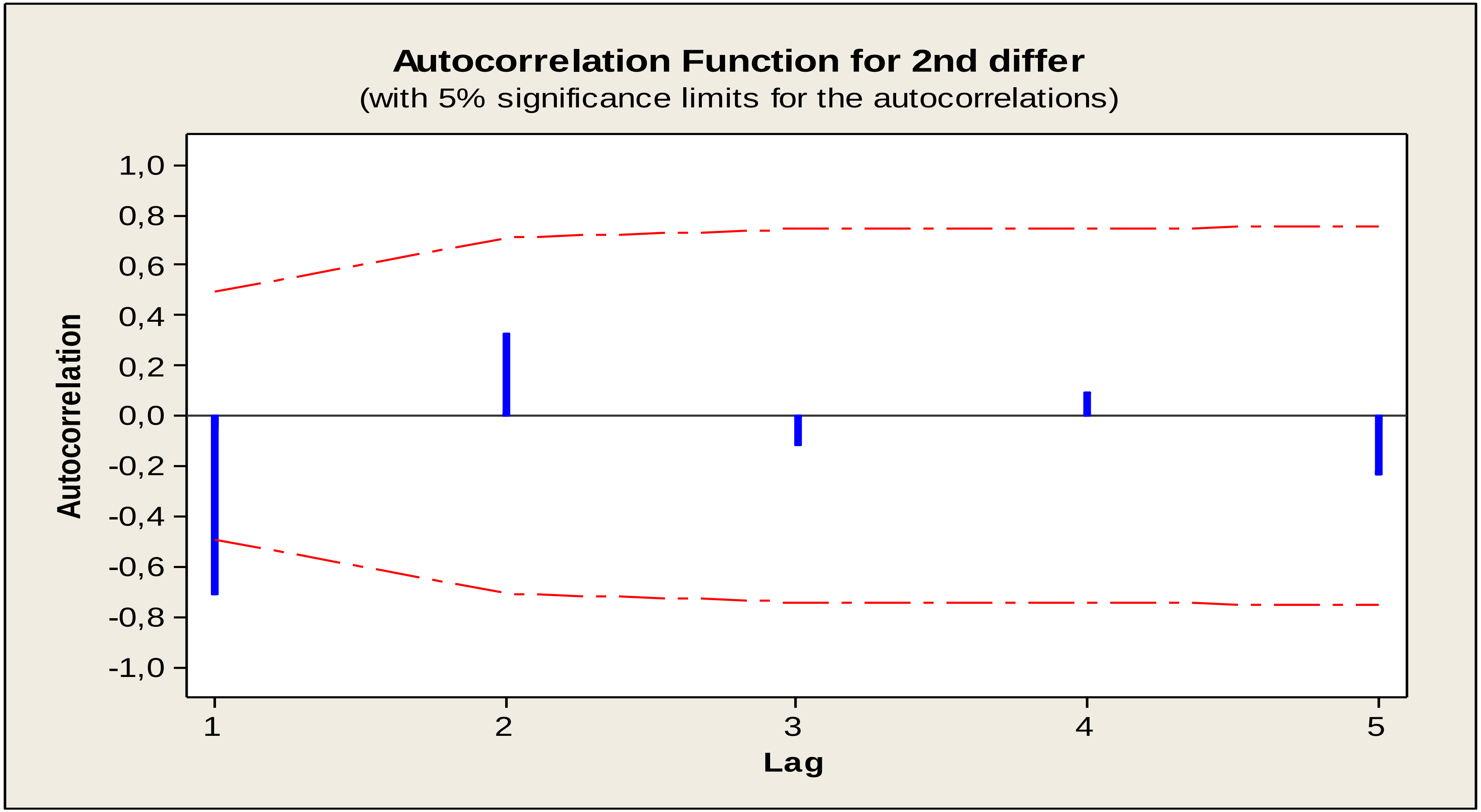

Figure 5.10 Autocorrelation function plot of second difference of annual natural gas consumption of Turkey

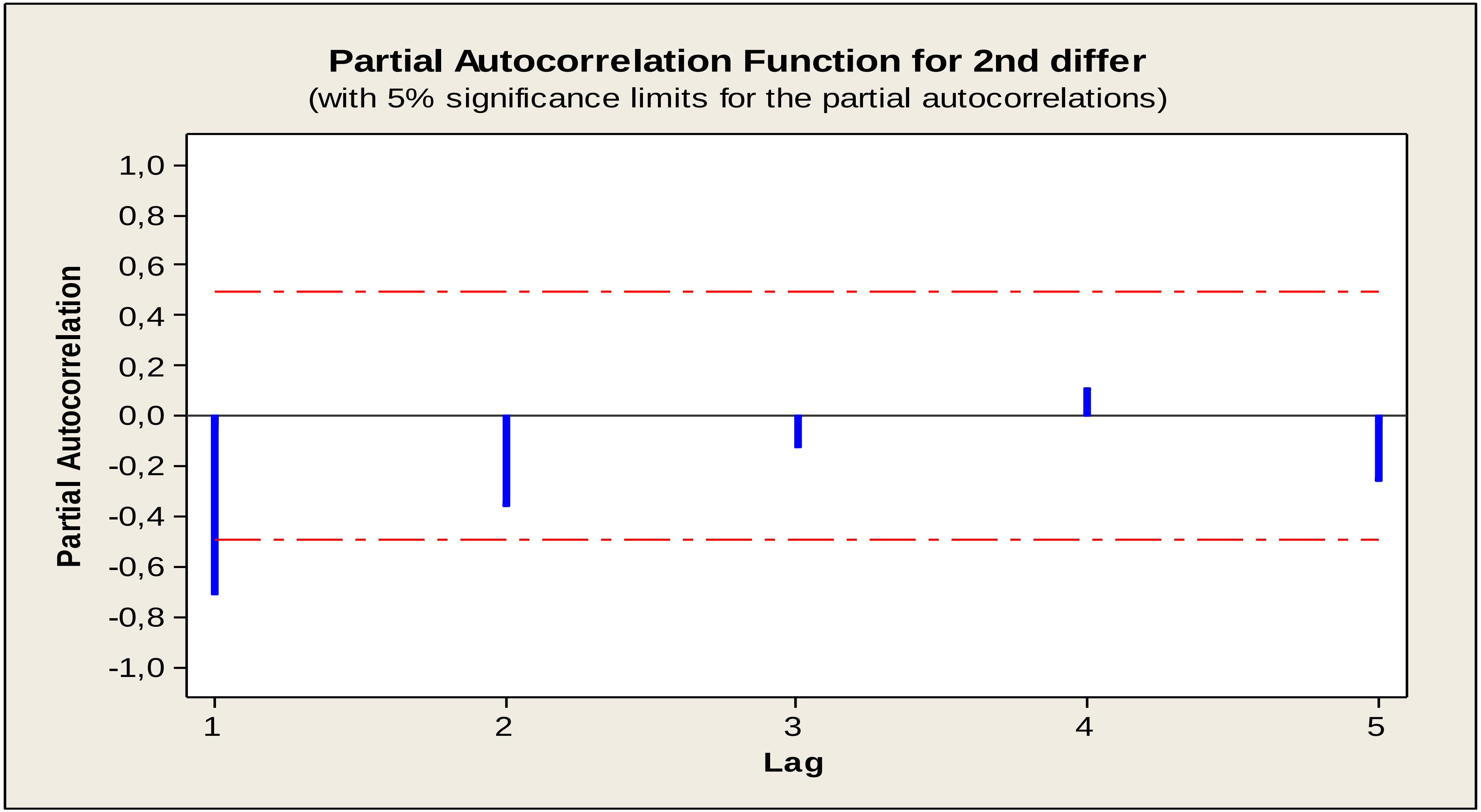

Figure 5.11 Partial autocorrelation function plot of second difference of annual natural gas consumption of Turkey

Second order differencing is enough for the transformation of data to stationary one. Order of autoregressive process of the ARIMA model is determined by the number of partial autocorrelation function coefficients. Similarly order of moving average of the ARIMA model is decided by the number of autocorrelation function coefficients. The peak points of these functions show the order of models (Goktas, 2005).

The criterion mentioned in the previous paragraph is employed to find out the orders of annual natural gas consumption ARIMA model. Autocorrelation function plot of the second difference in Figure 5.10 has the peak point at the first lag. Moving average order should be first order. Partial correlation function plot of second difference in Figure 5.11 has the peak at the first lag. So autoregressive process order should be first order for the model. As a result of that analysis, ARIMA(1,2,1) is decided to make annual natural gas consumption forecasts of Turkey.

In order to check our findings, we run SPSS which is used for the seeking the ARIMA model of monthly gas consumption of Istanbul before. SPSS made an automatic search on orders of ARIMA and it suggested the ARIMA(1,2,0) model for annual gas consumption data of Turkey. Its residual ACF and PACF plots are shown as Figure 5.12. Coefficients of ARIMA models are listed in Table 5.6. Both models are shown with their statistics in Table 5.7. ARIMA(1,2,1) model is fitting to 2007 and 2008 with lower MAD than ARIMA(1,2,0) model.

Table 5.8 Coefficients of ARIMA models

	Coefficients	
Model	**AR(1)**	**MA(1)**
ARIMA(1,2,0)	-0,708	
ARIMA(1,2,1)	-0,642	0,194

Table 5.9 Results of ARIMA models

Years	Real	ARIMA(1,2,0)	UL	LL	Error	ARIMA(1,2,1)	UL	LL	Error
2007	35,064	34,920	37,080	32,760	-144	34,872	37,103	32,642	-192
2008	37,128	38,782	42,311	35,253	1,654	38,725	42,331	35,119	1,597
				MAD	899			**MAD**	894

As a conclusion, ARIMA(1,2,1) model compared with double exponential smoothing in terms of forecast and limit values in Table 5.8. Double exponential smoothing with 0,60 and 0,60 parameters error value is lower than ARIMA(1,2,1). As a result, double exponential smoothing produces better long term forecasts than ARIMA for annual natural gas consumption of Turkey.

Table 5.10 Comparison between ARIMA model and double exponential smoothing

Years	Real	DES (0.6; 0.6)	UL	LL	Error	ARIMA(1,2,1)	UL	LL	Error
2007	35,064	33,437	36,004	30,870	-1,627	34,872	37,103	32,642	-192
2008	37,128	37,023	40,086	33,961	-105	38,725	42,331	35,119	1,597
				MAD	866			**MAD**	894

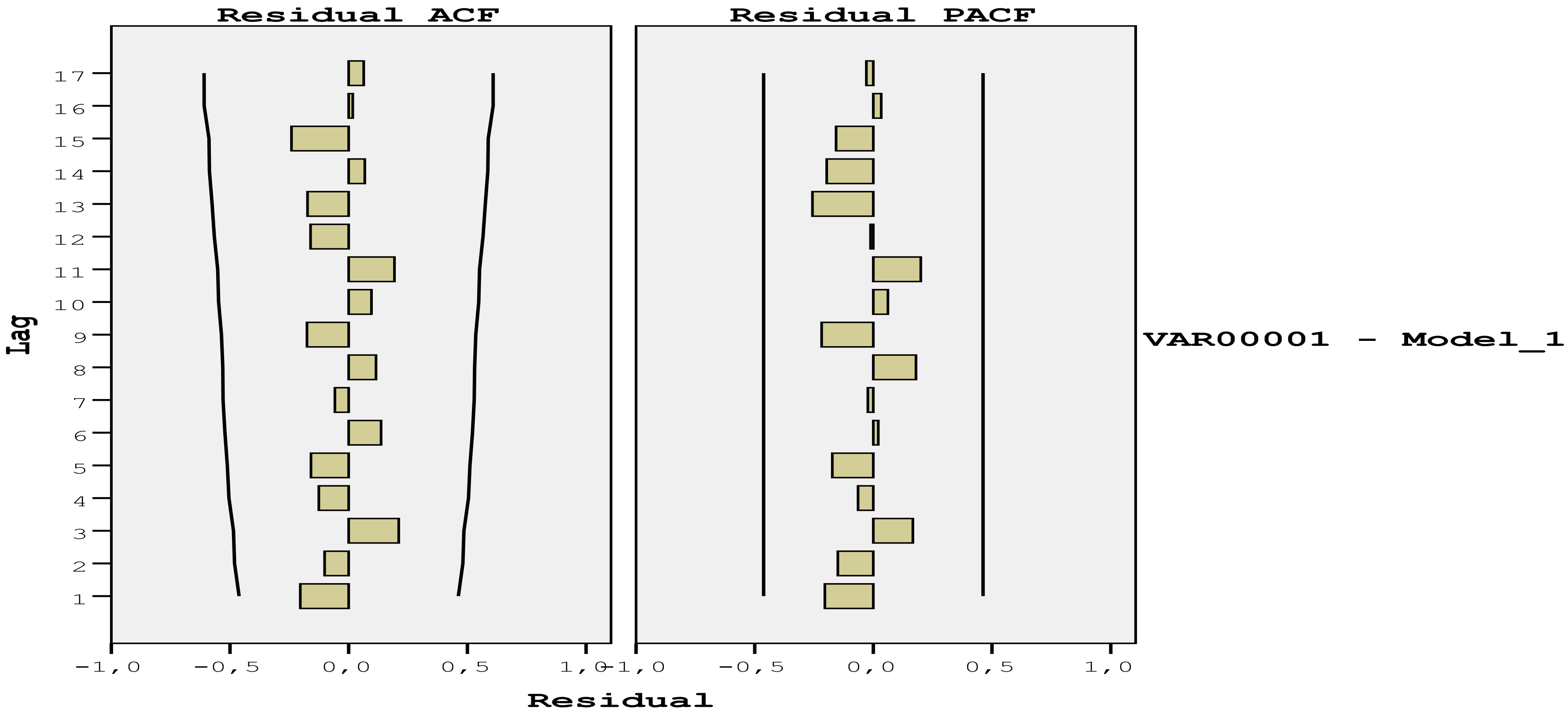

Figure 5.12 ARIMA(1,2,1) residual ACF and PACF plots

5.3 Forecasting the Last Four Years of Annual Natural Gas Consumption of Turkey and Comparing with Literature Study

Ediger and Akar (2007) made a study about forecasting of primary energy demand by fuel in Turkey. They used autoregressive integrated moving average (ARIMA) and seasonal ARIMA (SARIMA) methods to predict future primary energy demand of Turkey from 2005 to 2020.

The results of their have shown that the average annual growth rates of individual energy sources and total primary energy will decrease in all cases except wood and animal–plant remains which will have negative growth rates. The decrease in the rate of energy demand may be interpreted that the energy intensity peak will be achieved in the coming decades. Another interpretation is that any decrease in energy demand will slow down the economic growth during the forecasted period. At the end of their study, they proposed some policy recommendations.

One of the findings of Ediger and Akar is that natural gas will continue to be a key element of the Turkish energy system in the future. In order to understand the future changes, they forecasted the consumption values of natural gas period between 2005 and 2020.

This section of our study provides the comparison of forecast values between our study and the Ediger and Akar's study. Annual natural gas consumption data of Turkey from 1987 to 2004 is used and SPSS software is run to determine the autoregressive integrated moving average model. Also same data are employed to make forecasts using double exponential smoothing. Note that same data and time period is used by Ediger and Akar. Table 5.9 represents the comparison table of forecasts related to applied methods.

Table 5.11 Comparison table of methods

Methods	2005	2006	2007	2008	MAD
Ediger and Akar's Study Forecast	22,319	24,155	26,569	28,378	7,032
ARIMA(0,1,0) Forecast	23,378	24,648	25,917	27,187	7,105
Upper Control Limit of ARIMA(0,1,0)	25,244	27,286	29,149	30,919	
Lower Control Limit of ARIMA (0,1,0)	21,512	22,009	22,686	23,456	
Double Exponential Smoothing Forecast	24,623	26,919	29,216	31,512	4,320
Upper Control Limit of Double Exp. Smoothing	26,096	29,830	33,589	37,353	
Lower Control Limit of Double Exp. Smoothing	23,150	24,008	24,843	25,671	
Real Values	26,865	30,493	35,064	37,128	

SPSS suggested the ARIMA(0,1,0) model to make forecasts. Predicted values of Ediger and Akar’s study are slightly better than ARIMA(0,1,0). On the other hand, double exponential smoothing MAD and forecasts are far better than all other methods. Double exponential smoothing with 0,41 and 1,37 parameters should be considered as the forecasting method of 2005, 2006, 2007 and 2008.

CHAPTER 6

GENERATING CONSUMPTION SCENARIOS BY BOOTSTRAP

Scenarios are projected sequences of events, detailed plans or possibilities of related subject. Generating scenarios is the important part of strategic planning process. Bootstrap is a resampling method that generates lots of sibling series from a series, so one could use them to produce scenarios.

The original sample represents the population from which it was drawn. So resamples from this sample represent what we would get if we took many samples from the population.

Bootstrapping requires stationary data. Non-stationary data can be transformed into stationary data by different methods in order to apply bootstrap.

We apply Demirel and Willemain (2002)'s scenario generation technique to generate bootstrap scenarios. This technique requires a certain time series length. Therefore we used monthly natural gas consumption data of Istanbul which has enough observations to apply the method.

6.1 ESTIMATING THE BEST BLOCK SIZE

First we need to determine best block size to generate new data series. Subseries technique is a solution for this problem. The steps of subseries technique are explained briefly below:

1. Divide the original series into $s \geq 5$ subseries with m observations each. (Choosing $s = 5$ produces 10 reference values of Δ. There is a tradeoff here: more subseries provide a smoother estimate of the reference

distribution, but subseries that contain too few observations yield unstable estimates of Δ.)

2. Compute $\binom{s}{2}$ values of Δ using all pairs of subseries.

3. Create r = 100 or more bootstrap samples of each subseries, beginning with a block size of one and incrementing by one. More bootstrap samples give better results.

4. For each block size, compute the r values of Δ* between each of the subseries and its bootstrap samples.

5. Compare the sample distributions of Δ and Δ* using the K-S two sample test.

6. The block size which gives the minimum K-S statistic value D is the best block size for subseries, $\hat{b}_m$.

7. Estimate the best block size for the whole series by using the square root rule, $\hat{b}_n = s^{1/2}\,\hat{b}_m$. (Demirel and Willemain, 2002b)

Δ is the index of difference. The Δ statistic is a measure of the distance between two sample time series. It is based on the characterization of series by higher order crossing (HOC) counts (Kedem, 1993). Higher order crossings (HOC) are counts of the zero crossings made by the mean-centered series and its first, second, and higher differences. Any stationary process can be characterized by its HOC (Demirel and Willemain, 2002b).

6.2 HOW TO DEAL WITH NON-STATIONARITY

Monthly natural gas consumption of Istanbul is a non-stationary data which is shown in Figure 4.4. It should be transformed into stationary one and then get back to non-stationary data to make analyses. Non-stationary natural gas consumption data of Istanbul has three components which are seasonality, trend and increasing variance. Removal of these components results the stationary data. Then above steps of subseries technique are used for prediction of best block size.

6.2.1 Removal of Non-Stationary Components of Monthly Natural Gas Consumption of Istanbul

In this data as shown in Figure 4.4, we have three components of non-stationarity. 1st component is seasonality. Data is monthly seasonal. 2nd component is trend. Data has increasing trend. 3rd component is non-constant variance. Data has increasing variability. We start with non-constant variance. Box and Jenkins (1970) took logarithms to make the seasonal effect additive. Increasing variance would be removed by taking logarithm of the natural gas consumption data of Istanbul.

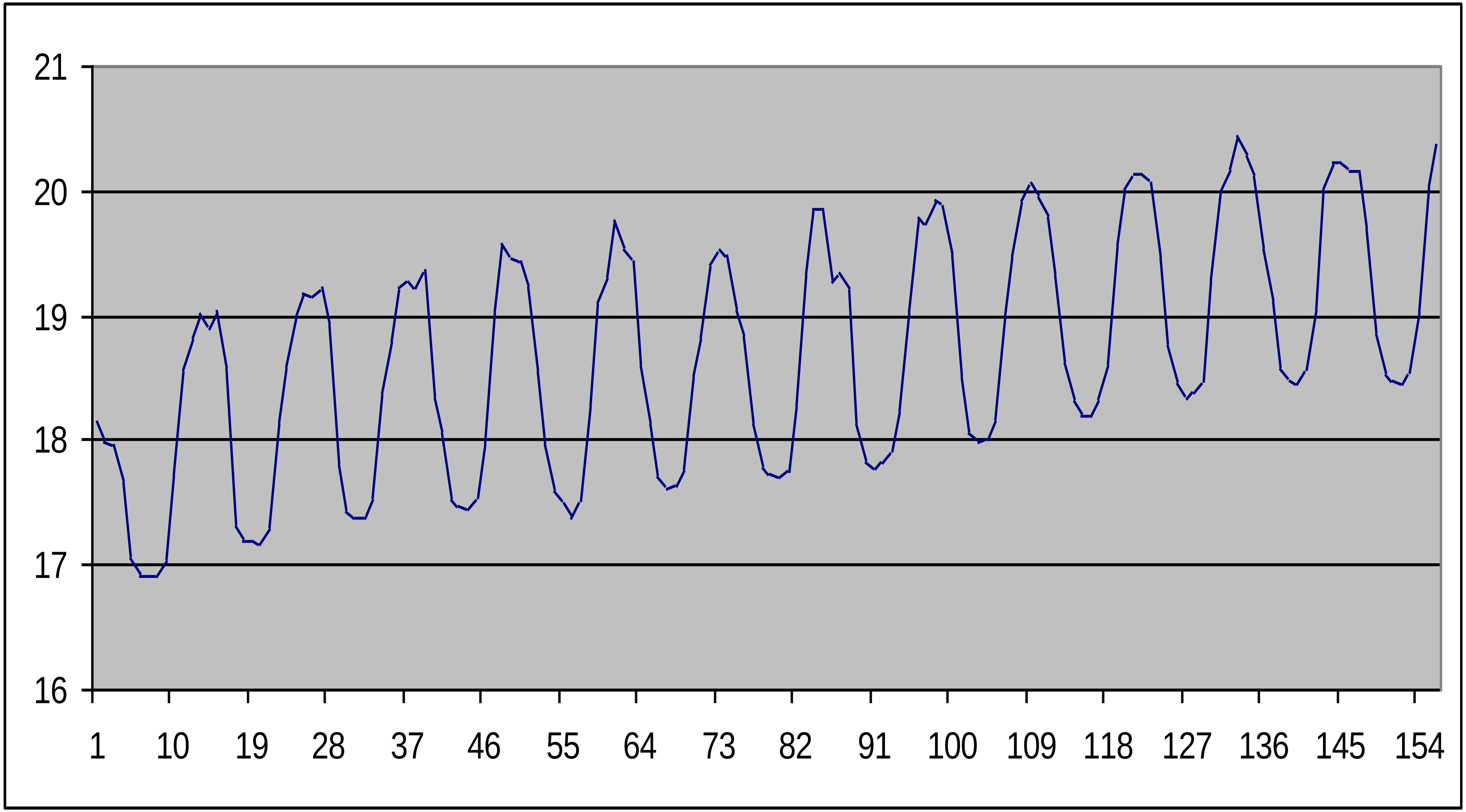

Figure 6.1 Natural logarithms of monthly natural gas consumption data of Istanbul

Figure 6.1 shows the plot of logged data of natural gas consumption of Istanbul. The data has still seasonal and trend components. Firstly it is guaranteed that we should take the seasonal differencing in order to eliminate the seasonality. Seasonal differencing is calculated by the taking difference of each month’s logged value from seasonal factor. Seasonal factor is computed i.e. as the summation of first months of years and then dividing this sum to the number of first months. Seasonal differencing is applied to natural logarithms of the data. Results are shown on the next page as Figure 6.2.

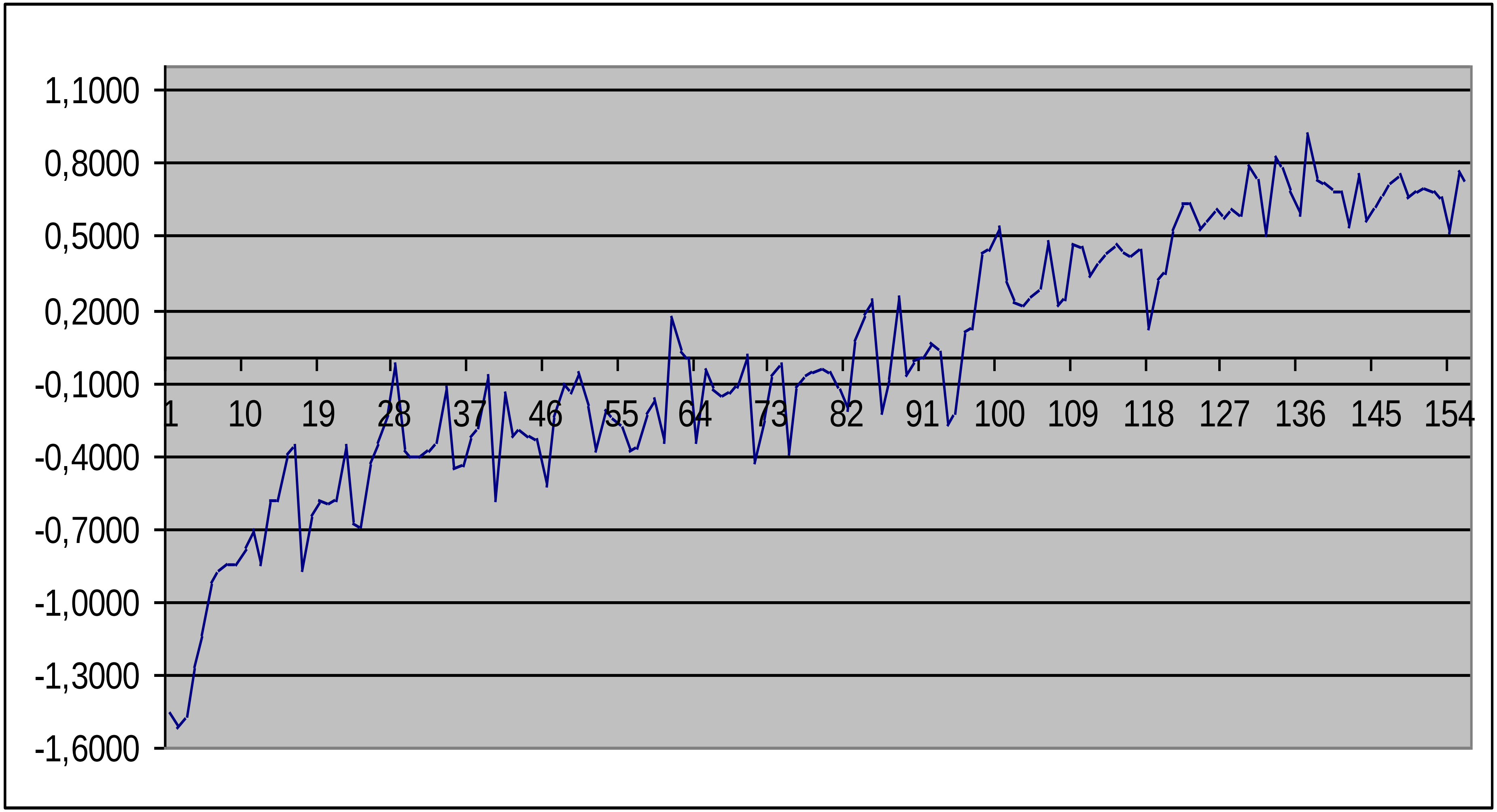

Figure 6.2 The plot of seasonally differenced logged natural gas consumption data of Istanbul

Figure 6.2 shows the plot of seasonally differenced logged natural gas consumption data of Istanbul. Increasing variance and seasonality are removed as a result of previous computations. The last component that should be removed is the trend of data. First difference of logged and seasonally differenced data would eliminate the trend. It is shown on the next page as Figure 6.3.

Also the ACF and PACF plots of final stationary data are shown in Figure 6.4 and 6.5 respectively. Spikes of the lags are in the limits of autocorrelation function as shown in Figure 6.4. Moreover ACF dies down fairly quickly and cuts off rapidly. In addition, PACF plot in Figure 5.5 supports the idea that we reach a stationary data.

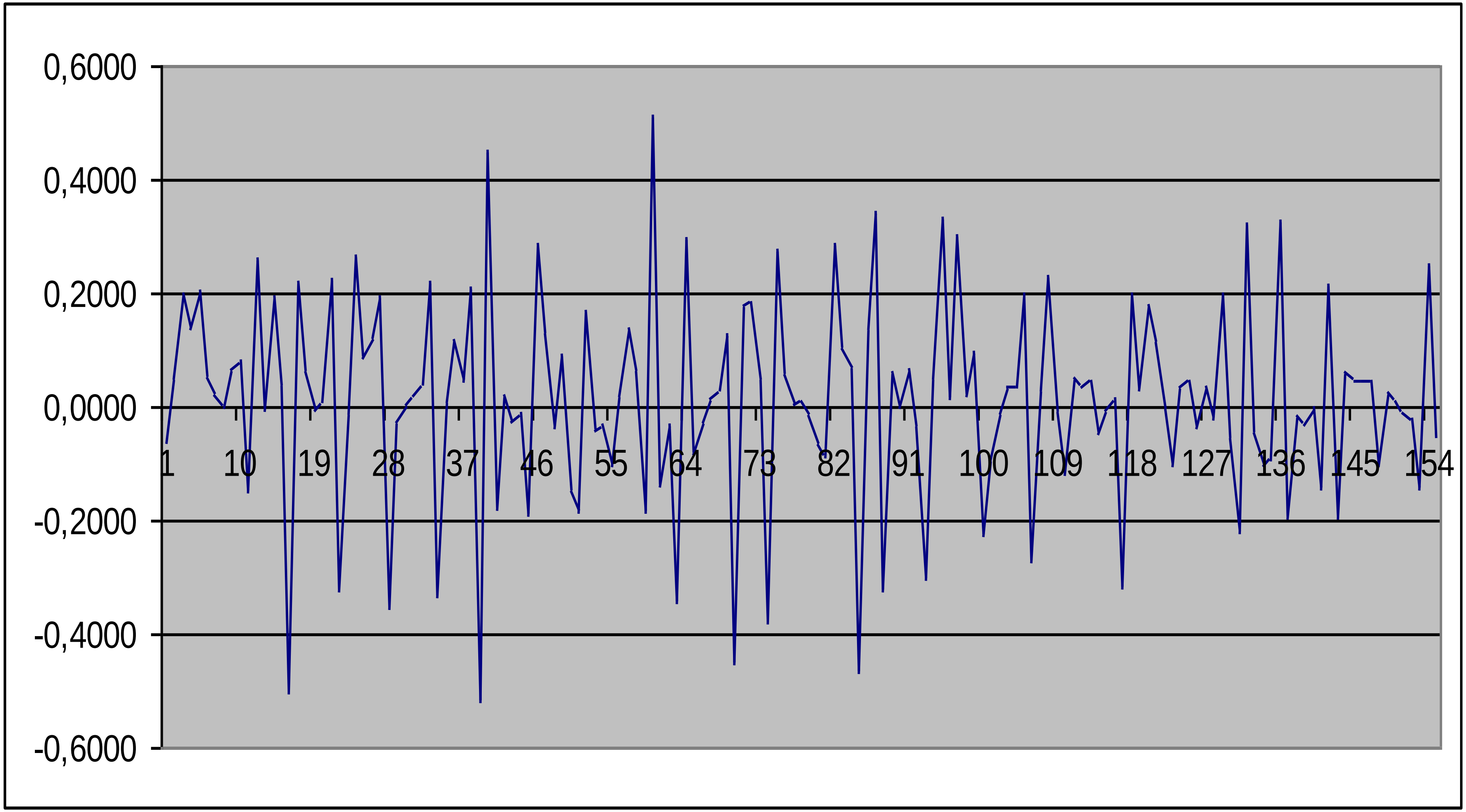

Figure 6.3 First difference of logged and seasonally differenced natural gas consumption data of Istanbul

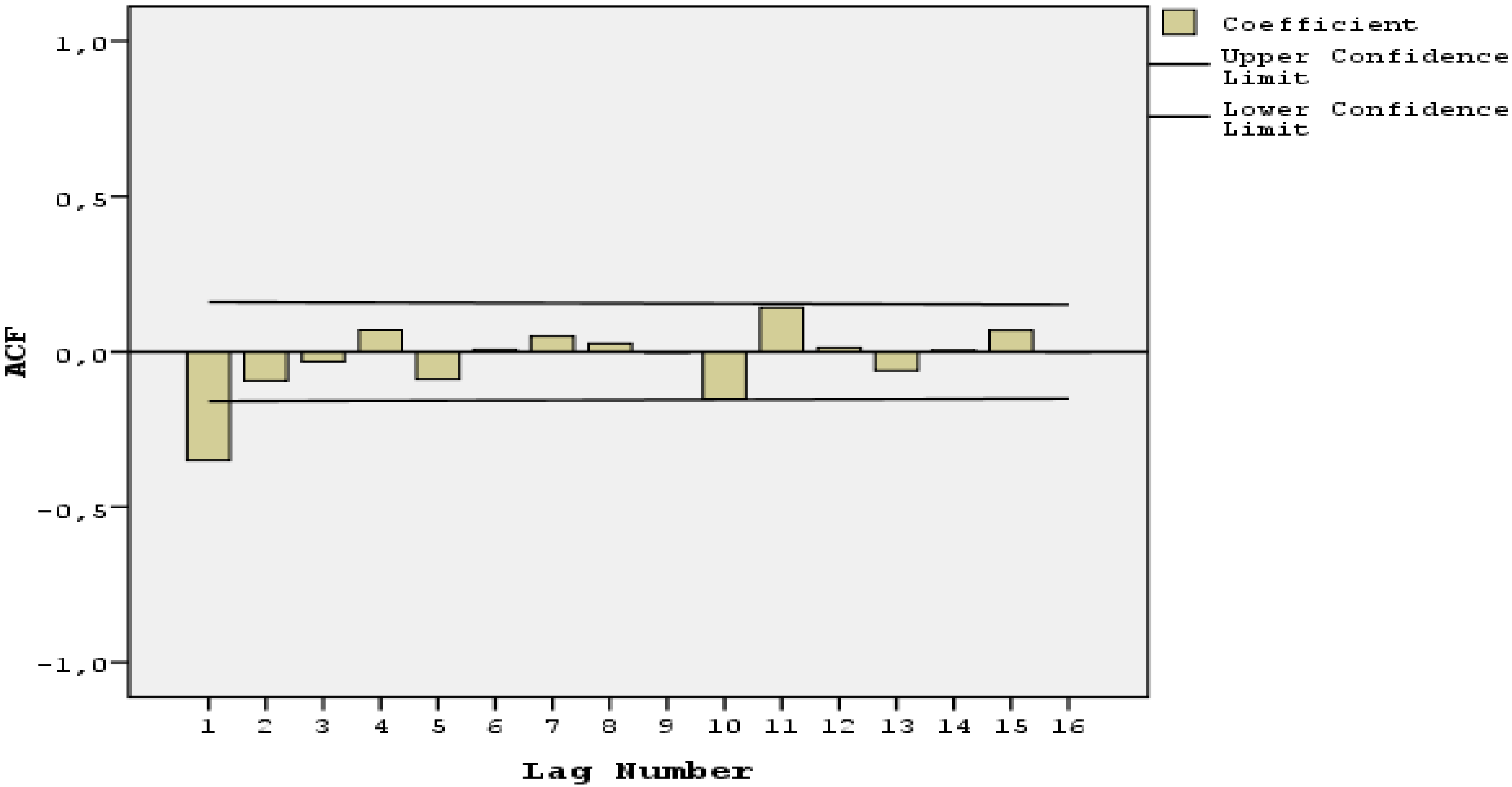

Figure 6.4 Autocorrelation function plot of logged, seasonally differenced and first differenced natural gas consumption data of Istanbul

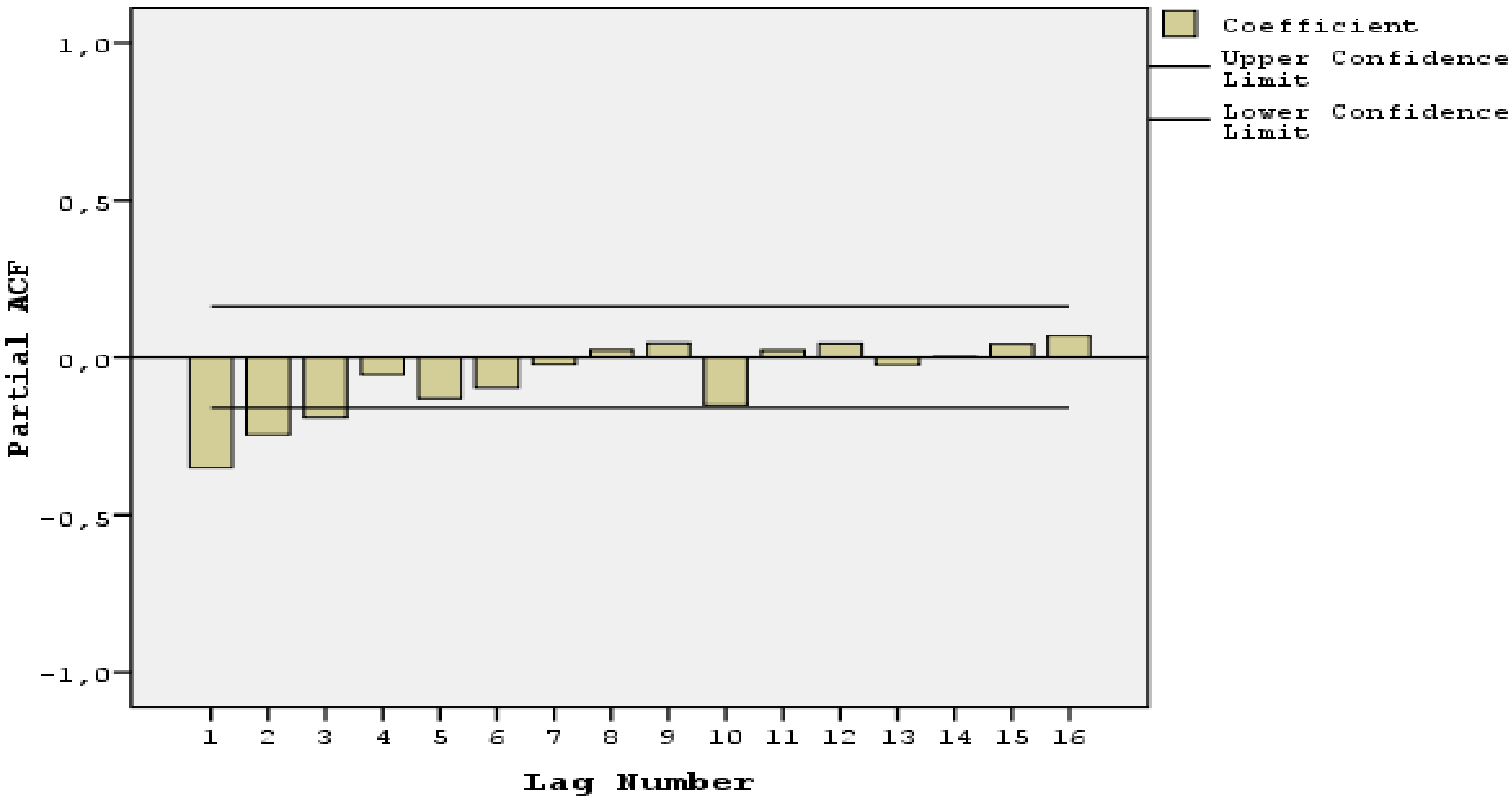

Figure 6.5 Partial autocorrelation function plot of logged, seasonally differenced and first differenced natural gas consumption data of Istanbul

6.2.2 Estimation of Best Block Size

After getting the stationary data, best block size is investigated in order to produce different monthly natural gas consumption scenarios of Istanbul.

The stationary data is divided into 5 subseries firstly. From each pair of subseries which is $\binom{5}{2} = 10$, delta values are calculated. These are called “golden deltas” and shown in Table 6.1. An Excel file with formulas is created to calculate all these steps to produce golden deltas.

Table 6.1 Golden deltas’ normalized zero-crossing rates

Delta 1	4.81
Delta 2	0.60
Delta 3	0.65
Delta 4	2.61
Delta 5	6.70
Delta 6	3.23
Delta 7	1.33
Delta 8	0.80
Delta 9	3.93
Delta 10	1.68

The second issue is about generating bootstraps of each subseries. We set up Excel files to generate r = 1000 bootstraps of subseries and compute r values of Δ^* between each of the subseries and its bootstraps. We start with block size = 1 and increment by one up to 5. So we got 1000 different Δ^* for each subseries with one block size. Totally we have 5000 Δ^* for each block size.

The third step is the comparison between golden deltas and Δ^*’s. This evaluation of deltas is made with Kolmogrov-Smirnov (K-S) 2-sample test. SPSS can handle two sample K-S test. The minimum K-S statistic value D is chosen for the determination of best block size. This block size is symbolized as $\hat{b}_m$. The best block size of monthly natural gas consumption of Istanbul is 3 as shown in Table 6.2 and Figure 6.6.

Table 6.2 Kolmogrov-Smirnov (K-S) 2-sample test results

Block Size	K-S Statistics
1	0.1978
2	0.1668
3	*0.1600*
4	0.1748
5	0.2410

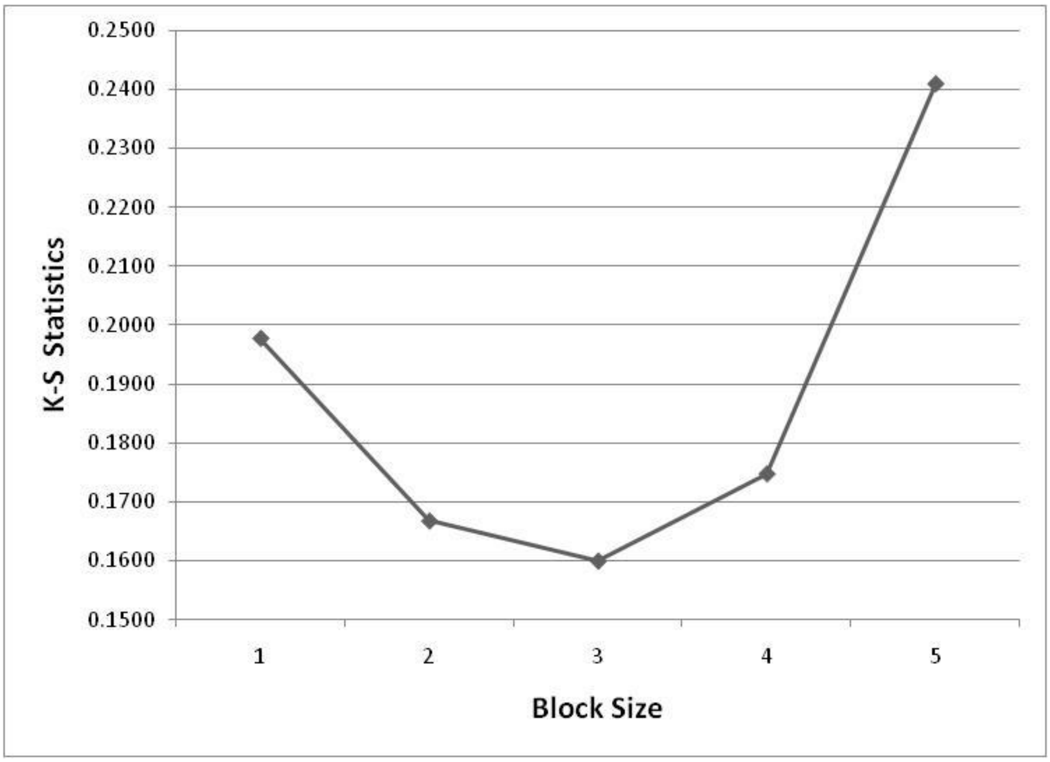

Figure 6.6 K-S Statistics results related with block sizes

$\hat{b}_m$ is used in square root rule, $\hat{b}_n = s^{1/2}\,\hat{b}_m$. $\hat{b}_m = 3$ and $\hat{b}_n = 6.7082$ for our data. If we rounded up the $\hat{b}_n$, the result is 7. In order to generate scenarios of monthly natural gas consumption of Istanbul, block size is equal to 7.

Finally we have concluded the best block size estimation, and we could go further with generating sibling series of original monthly natural gas consumption of Istanbul.

6.3 GENERATING BOOTSTRAPS AND SIBLING SERIES OF MONTHLY NATURAL GAS CONSUMPTION DATA OF ISTANBUL

In the previous section, we have calculated the best block size for the scenarios of monthly natural gas consumption of Istanbul. We set up an Excel file with macros based on Visual Basic codes. That file generates 100 bootstraps of our original consumption data, and then takes them back to original like (sibling) series by using moving block bootstrap. ACF and PACF plots of a bootstrap show us that, in this study, we preserve the dependence structure within blocks and pass it to the bootstrap sample as understand through Figure 6.7 and Figure 6.8. Those plots are nearly same with Figure 6.4 and Figure 6.5

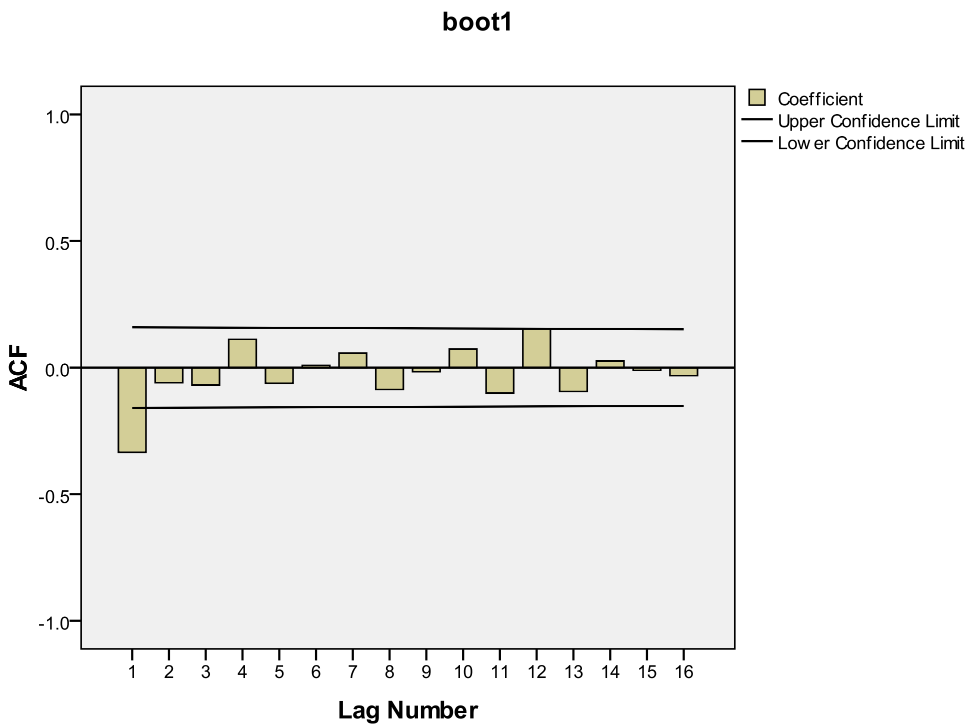

Figure 6.7 Autocorrelation function plot of a bootstrap

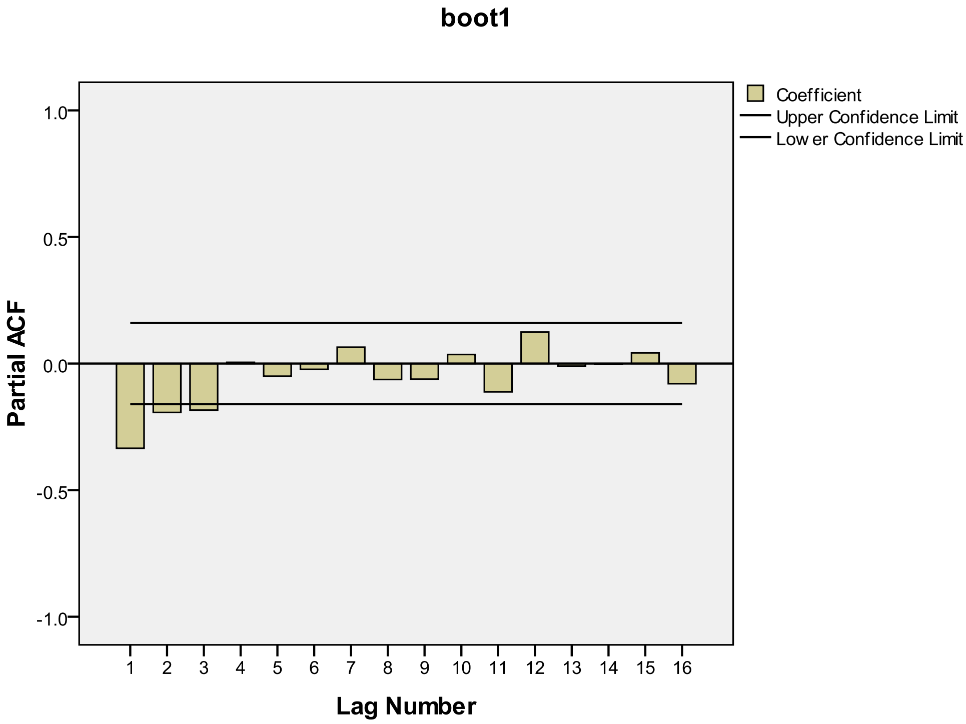

Figure 6.8 Partial autocorrelation function plot of a bootstrap

Sibling series are generated from bootstraps. We take a bootstrap data set and make reverse calculations of what we have done in the removal of non-stationary components section. Reverse calculation starts with the production of the first differentiated data set. At that step, we add up previous data point to current data point in order to get 1^{st} differentiated data set or “detrended data set”. Then, seasonal index is added on detrended data set. So we get seasonality added data set. As the final step, we applied the exponential function that returns e raised to the power of a number which is the sibling series’ data point.

Eventually, we have generated 100 bootstraps and 100 sibling series of monthly natural gas consumption data of Istanbul.

6.4 PRODUCTION OF MONTHLY NATURAL GAS CONSUMPTION SCENARIOS OF ISTANBUL

After determining the best block size and generating bootstraps and sibling series, we could produce different monthly scenarios about the natural gas consumption of Istanbul.

In the 4th chapter, Winters' method with 0.05, 0.30 and 0.15 parameters is the recommended estimation technique for monthly natural gas consumption of Istanbul. These parameters are employed for the production of 100 scenarios.

A MINITAB macro that uses above Winters' method is written and run for our generation of 100 scenarios. Samples are shown in Appendix A. Average of the 100 scenarios could be recommended as the forecast of 2008 and 2009 months which is listed in Table 6.3.

Table 6.3 Bootstrap forecasts of 2008 and 2009

Year	Months	Bootstrap based Forecast	Real	Error
2008	January	848,226,260	843,618,178	4,608,082
2008	February	768,579,496	713,833,255	54,746,241
2008	March	728,386,441	450,535,818	277,850,623
2008	April	452,581,345	283,844,118	168,737,227
2008	May	218,704,507	178,506,683	40,197,824
2008	June	153,703,880	119,554,899	34,148,981
2008	July	147,288,895	111,280,773	36,008,122
2008	August	148,533,172	104,610,386	43,922,786
2008	September	170,882,829	118,546,882	52,335,947
2008	October	321,548,798	180,355,252	141,193,546
2008	November	689,343,330	343,461,253	345,882,077
2008	December	1,062,275,441	550,464,497	511,810,944
2009	January	987,801,205	652,857,140	334,944,065
2009	February	892,393,276	594,104,138	298,289,138
2009	March	841,755,027	583,689,727	258,065,300
2009	April	524,505,659	344,758,520	179,747,139
2009	May	253,490,831	153,624,534	99,866,297
			MAD	169,550,255

In order to compare the forecast performance of bootstrap based Winters' forecast with the recommended method for monthly natural gas consumption in Chapter 4, results are shown in Table 6.4.

Bootstrap based forecast has good performance from May 2008 to September 2008 as seen in Figure 6.9. After October 2008 with the effect of economical crisis natural gas consumption decreases and the forecasts are overestimated from the real.

Table 6.4 Comparison of Winters' method and Bootstrap based forecast

Year	Month	Real	Winters' Method with 0,05; 0,30; 0,15	Error	Bootstrap based Forecast	Error
2008	January	843,618,178	849,079,233	5,461,056	848,226,260	4,608,082
2008	February	713,833,255	639,215,416	-74,617,839	768,579,496	54,746,241
2008	March	450,535,818	572,873,968	122,338,150	728,386,441	277,850,623
2008	April	283,844,118	344,767,357	60,923,239	452,581,345	168,737,227
2008	May	178,506,683	162,962,311	-15,544,372	218,704,507	40,197,824
2008	June	119,554,899	111,505,827	-8,049,072	153,703,880	34,148,981
2008	July	111,280,773	102,472,124	-8,808,649	147,288,895	36,008,122
2008	August	104,610,386	100,599,219	-4,011,167	148,533,172	43,922,786
2008	September	118,546,882	110,956,076	-7,590,806	170,882,829	52,335,947
2008	October	180,355,252	196,323,181	15,967,929	321,548,798	141,193,546
2008	November	343,461,253	434,005,451	90,544,198	689,343,330	345,882,077
2008	December	550,464,497	623,020,610	72,556,113	1,062,275,441	511,810,944
2009	January	652,857,140	809,169,997	156,312,857	987,801,205	334,944,065
2009	February	594,104,138	609,052,253	14,948,115	892,393,276	298,289,138
2009	March	583,689,727	545,734,590	-37,955,137	841,755,027	258,065,300
2009	April	344,758,520	328,369,585	-16,388,935	524,505,659	179,747,139
2009	May	153,624,534	155,180,680	1,556,146	253,490,831	99,866,297
			MAD	40,534,383	**MAD**	169,550,255

In Table 5.4, MAD value of Winters' method with 0,05; 0,30; 0,15 is lower than bootstrap based forecast. This is showing that bootstrap based forecasting is outperformed by Winters' method.

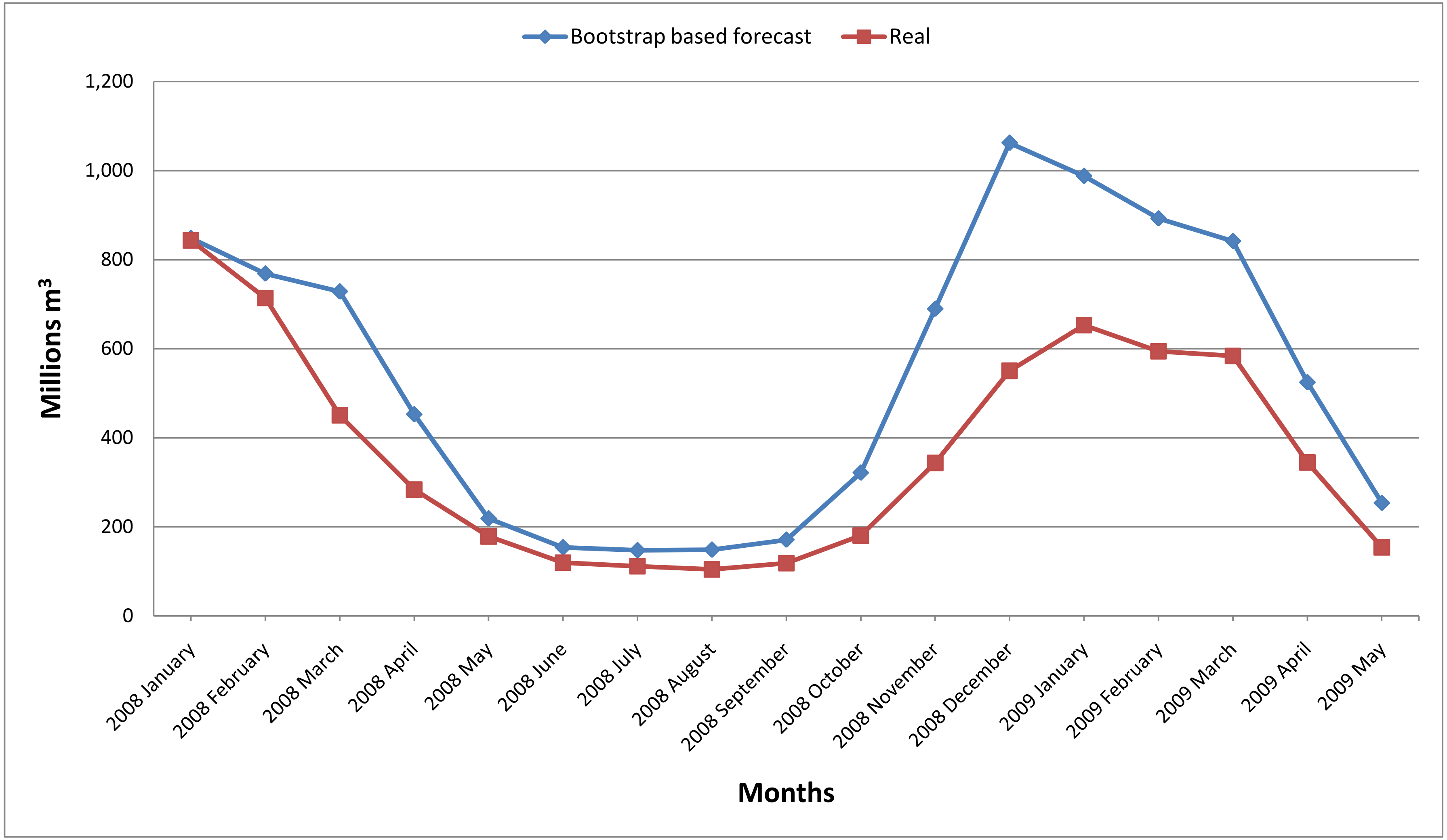

Figure 6.9 Average of consumption scenarios as forecasts of 2008 and 2009 months

CHAPTER 7

CONCLUSIONS

In this study, natural gas consumption data of Istanbul and Turkey are examined in terms of time series analysis on daily, monthly and monthly, annual basis respectively. Istanbul consumption data between 1992 and 2009 years is provided from IGDAS. Turkey consumption data between 1987 and 2008 years is provided from BOTAS and IEA.

Demand forecasting is an important issue for service and manufacturing industries. Time series analysis is one of the most popular forecasting methods. Recent studies about the natural gas consumption are the applications of time series analysis methods.

In Chapter 4, several time series methods are investigated by using SPSS, Minitab, SmartForecasts or on hand manual calculations. Winters' method is overall recommendation of software packages for daily and monthly consumption data of Istanbul. In application of Winters' method, it is hard to determine the smoothing constants. In order to handle that situation, we made parameter searches on data, and then we use these parameters to forecast coming time periods. On daily basis, consumption of natural gas depends on temperature. Also major consumers of natural gas are households. So, one should use temperature effect on his/her forecasting technique for further studies. On monthly basis, data has seasonality and trend, so Winters' method is applied via SmartForecasts and Minitab. Parameters search is done to reach better smoothing constants with lower mean absolute deviation (MAD). In addition to Winters' method, seasonal autoregressive integrated moving average model (SARIMA) is applied on that data. SARIMA is one the Box-Jenkins methods which is explained in Chapter 2. Finally, Winters' method and SARIMA are compared, we find out forecasts of Winters' method with 0,05; 0,30; 0,15 have lower MAD and better limits than SARIMA. As the forecasting period lengthens or increases, forecasting accuracy decreases in prediction of monthly natural gas consumption of Istanbul.

Natural gas consumption of Turkey is examined on monthly and annual basis in Chapter 5. Monthly consumption of Turkey has seasonality and trend as shown in Figure 5.1. We run SmartForecasts to investigate which forecasting method is more appropriate for that type of data, and then it turns Winters' method as the best one. Besides, SPSS is used to establish an autoregressive integrated moving average model (ARIMA). It recommended SARIMA, and then it is compared with Winters' method. They have same MAD value and also nearly same limits. Winters' method has a real value out of its limits. SARIMA(ARIMA(0,0,2)(1,1,0)) can be used for forecasting of monthly natural gas consumption of Turkey. Annual natural gas consumption has an increasing trend only. Trend analysis is applied thorough annual consumption data of Turkey. Double exponential smoothing fits well that data. Also ARIMA model is applied by manual calculations and SPSS ARIMA tool. ARIMA results are compared with double exponential smoothing, ARIMA has higher MAD value. At the end of Chapter 5, Ediger and Akar's study (2007) is compared with ours in terms of forecasting 4 years of annual natural consumption of Turkey. Double exponential smoothing forecasts in our study are closer to real values than Ediger and Akar's (2007) proposed ARIMA method.

Bootstrap of time series data have developed in recent years. Bootstrapping is the generation of sibling series of an original data. In our study at Chapter 6, we make a research on bootstrapping natural gas consumption data. We used the monthly natural gas consumption data of Istanbul because it has enough data points for bootstrap. Original data is non-stationary one. In order to bootstrap the data, we transformed that data into stationary one. After transformation, we produced 100 bootstrap series. These series transformed back into non-stationary series by several calculation steps. Subsequently, we have 100 sibling series. Those series are used to generate 100 different forecasted series by the Winters' method which is chosen as the best estimation technique for the monthly natural gas consumption in Chapter 4. That 100 different forecasted series are averaged in our study to produce predictions of 2008 and 2009 months. Those predictions are not close as SARIMA estimations to monthly natural gas consumption data of Istanbul. Winters' method performs well compared to all other methods for that consumption data.

Time series methods depend on time and past values of data. Those methods take into consideration past pattern of the data, and then make predictions based on that pattern. Natural gas consumption data is affected from various factors such as season, temperature, weekend-weekday, etc. In addition, economical crisis affects the consumption rate. Peak natural gas consumption of Istanbul occurs generally at 1^{st}, 2^{nd}, 3^{rd}, 11^{th}, 12^{th} months' of a year. The 1^{st}, 2^{nd}, 3^{rd} months' of 2008, consumptions are higher than 2007's same period. By the beginning of economical crisis, 11^{th}, 12^{th} months' of 2008, consumptions remained below of same period of 2007. At the 1^{st}, 2^{nd}, 3^{rd} months' of 2009, consumptions are very close to 2007's same months. As a result of those changes of data, time series forecasting methods overestimate the consumption values. This yields worse MAD values.

For further assignments, one should seek for the methods which consider all natural gas consumption factors. This will produce better forecasts but never same as real values.

APPENDIX A

SAMPLES OF BOOTSTRAPS, SIBLING SERIES AND MONTHLY CONSUMPTION SCENARIOS OF ISTANBUL

No.	Bootstrap 1	Bootstrap 2	Sibling Series 1	Sibling Series 2	Forecasts 1	Forecasts 2
1	0.20059	0.03632	75,687,921	75,687,921	715,681,905	811,252,139
2	0.02983	0.04506	83,281,408	70,665,300	662,635,993	650,215,511
3	0.17739	-0.04360	80,397,177	69,265,047	546,715,580	585,825,770
4	0.11365	-0.00746	58,408,569	40,344,002	363,500,334	322,944,577
5	-0.00713	0.01789	30,867,253	18,888,744	172,514,915	154,261,004
6	-0.10320	-0.31938	21,666,954	13,594,704	127,299,777	113,814,241
7	0.03506	0.20059	18,311,277	9,255,563	117,146,524	104,423,300
8	0.22665	0.00908	18,871,853	11,256,129	122,280,638	100,906,109
9	-0.32240	0.11747	26,619,096	12,772,632	144,724,300	105,066,061
10	-0.01278	0.04696	35,836,149	26,695,777	257,683,483	196,954,551
11	0.26469	0.21054	75,616,447	59,797,359	547,085,732	415,412,293
12	0.08640	-0.52043	147,815,748	110,731,386	810,139,438	567,990,336
13	0.12084	0.44997	150,708,724	61,538,838	810,518,263	557,196,089
14	0.19256	-0.18035	153,117,066	86,891,209	749,484,104	441,134,131
15	0.01197	0.13029	173,935,886	67,981,060	617,596,465	392,262,564
16	-0.01630	-0.45006	107,099,154	47,116,221	410,123,904	213,218,965
17	-0.06546	0.17987	49,701,665	14,170,228	194,408,153	100,321,102
18	-0.08513	0.18707	32,910,723	11,991,932		
19	0.28737	0.04855	28,320,768	13,547,893		
20	0.10292	-0.37873	37,564,834	14,152,371		
21	0.06472	0.27847	46,819,429	10,896,699		
22	-0.09310	0.06403	92,828,035	26,753,138		
23	-0.00996	-0.18358	180,755,773	60,957,754		
24	0.03460	0.51389	268,483,672	76,111,863		
25	0.03669	-0.13909	259,919,832	118,994,947		
26	0.19938	-0.03053	242,761,981	93,224,051		
27	-0.27432	-0.34468	277,657,730	84,724,345		
28	0.03252	0.29565	128,402,158	36,518,406		

No.	Bootstrap 1	Bootstrap 2	Sibling Series 1	Sibling Series 2
29	0.05261	0.03252	62,569,531	23,151,333
30	0.02234	0.22865	46,623,722	16,908,247
31	0.00097	-0.01224	44,673,435	19,913,135
32	0.06421	-0.11953	44,498,190	19,574,635
33	0.07987	0.04988	53,355,157	19,531,256
34	-0.14785	0.03632	107,401,071	38,153,600
35	0.26221	0.04506	197,990,561	84,557,796
36	-0.04860	0.03428	386,076,949	132,701,205
37	-0.10138	-0.02006	343,923,571	128,427,271
38	-0.09013	0.19851	279,794,868	113,331,395
39	0.32836	-0.05387	239,571,811	129,509,479
40	-0.19270	-0.22195	202,412,375	74,662,596
41	-0.01331	0.32126	78,743,223	28,208,237
42	-0.02905	-0.04860	54,932,790	27,497,522
43	0.13979	-0.52043	49,998,255	24,542,776
44	0.34245	0.44997	57,218,679	14,513,670
45	-0.32297	-0.18035	90,617,316	25,594,416
46	0.06094	0.02038	121,924,809	39,715,734
47	0.00220	-0.02360	276,950,804	86,627,850
48	0.06849	-0.00993	416,397,988	126,928,400
49	-0.03624	-0.18942	417,014,649	117,528,280
50	0.26221	-0.34468	362,093,123	87,555,370
51	-0.00303	0.29565	440,997,750	58,120,088
52	0.19377	-0.08009	267,496,646	47,525,164
53	0.03499	-0.02527	153,157,144	20,692,113
54	-0.50505	0.01374	112,132,385	14,263,597
55	0.22056	0.02856	63,405,824	13,549,800
56	0.05983	0.13029	78,665,955	13,874,231

No.	Bootstrap 1	Bootstrap 2	Sibling Series 1	Sibling Series 2
57	-0.00746	0.01789	93,911,598	17,772,197
58	0.01789	-0.31938	173,229,425	33,624,366
59	-0.31938	0.20059	376,908,749	52,214,835
60	0.20059	0.02983	410,849,229	95,732,707
61	0.02983	0.17739	469,562,797	92,238,269
62	0.17739	0.11365	435,566,797	99,164,252
63	0.11365	-0.00713	487,338,158	104,099,730
64	0.13834	0.20036	332,188,806	62,885,519
65	0.06403	0.14075	179,940,296	36,243,627
66	-0.18358	0.20743	135,623,445	29,495,379
67	0.51389	0.05261	105,765,737	34,007,704
68	-0.13909	0.02234	175,952,801	35,669,374
69	-0.03053	0.00097	172,161,907	41,015,455
70	-0.34468	0.06421	310,329,165	76,297,927
71	-0.19626	0.22865	469,864,798	173,877,690
72	0.06175	-0.01224	579,283,719	327,867,274
73	0.04585	-0.11953	576,244,887	302,884,542
74	0.04461	0.04988	543,157,531	241,975,926
75	0.04592	0.03632	532,154,242	238,326,449
76	-0.10341	0.04506	338,982,445	150,363,921
77	0.02802	-0.04360	144,189,459	74,195,663
78	-0.09013	-0.00529	104,833,668	50,216,011
79	0.32836	0.00845	89,762,900	46,804,106
80	-0.19270	0.22665	124,042,693	46,970,575
81	-0.01331	-0.32240	115,035,036	66,252,860
82	-0.02905	-0.01278	210,957,480	89,193,388
83	0.00177	0.26469	437,950,307	188,203,457
84	-0.14230	0.08640	658,179,028	367,901,905

No.	Bootstrap 1	Bootstrap 2	Sibling Series 1	Sibling Series 2
85	-0.13909	-0.13909	533,877,621	375,102,297
86	-0.03053	-0.03053	418,255,026	293,865,888
87	-0.34468	-0.34468	380,120,608	267,072,654
88	0.29565	0.29565	163,841,911	115,115,290
89	-0.08009	-0.08009	103,869,777	72,978,882
90	-0.02527	-0.02527	67,780,135	47,622,308
91	0.01374	0.01374	61,925,254	43,508,670
92	-0.04360	0.14075	62,474,758	43,894,751
93	-0.00746	0.20743	67,254,150	56,817,972
94	0.01789	0.05261	124,057,070	129,931,434
95	-0.31938	0.02234	269,920,627	292,688,837
96	0.20059	0.00097	294,226,871	449,015,907
97	0.02983	0.06421	336,274,191	420,319,669
98	0.17739	0.07987	311,928,188	403,526,761
99	-0.00993	-0.33244	349,003,895	409,542,778
100	-0.18942	0.00908	210,240,266	178,698,762
101	0.28784	0.11747	82,056,863	85,060,913
102	0.12903	0.04696	77,360,457	67,631,013
103	-0.03737	0.21054	82,469,938	66,416,903
104	0.09015	-0.52043	79,056,842	81,580,646
105	-0.14691	0.44997	97,283,744	54,515,174
106	0.03428	0.03460	156,093,047	158,884,266
107	-0.02006	0.03669	345,234,922	351,520,944
108	0.19851	0.19938	507,636,273	547,063,080
109	-0.05387	-0.27432	578,977,959	624,488,675
110	-0.22195	0.03252	493,937,264	427,361,457
111	0.32126	0.22865	370,695,836	413,674,716
112	-0.04860	-0.01224	310,982,869	316,344,297

No.	Bootstrap 1	Bootstrap 2	Sibling Series 1	Sibling Series 2
113	-0.08009	-0.10138	139,731,210	147,403,547
114	-0.02527	-0.09013	91,181,483	94,162,147
115	0.01374	0.32836	83,305,182	80,625,504
116	0.02856	-0.19270	84,044,404	111,415,793
117	0.13029	-0.01331	97,243,716	103,325,069
118	-0.45006	-0.02905	205,868,184	189,483,108
119	0.17987	0.00177	280,527,538	393,369,248
120	0.04461	0.19256	503,781,452	591,179,833
121	0.04592	-0.35280	492,619,911	670,259,952
122	-0.10341	-0.02618	464,365,422	424,064,949
123	0.02802	0.00270	392,364,066	387,081,855
124	0.01024	0.01948	245,503,479	236,142,005
125	-0.01028	0.04015	116,996,166	113,578,654
126	-0.02169	0.22268	81,866,443	83,585,295
127	0.05261	0.26221	75,063,110	97,853,922
128	0.02234	-0.00303	78,730,810	126,569,018
129	0.00097	0.19377	90,530,886	141,892,369
130	0.06421	0.03499	168,407,711	320,078,342
131	0.07987	-0.50505	383,789,508	708,431,670
132	-0.14785	0.22056	623,639,073	641,370,870
133	0.26221	0.05983	503,060,730	747,811,059
134	-0.33244	-0.06546	588,714,669	714,797,237
135	0.00908	-0.08513	395,609,777	627,321,816
136	0.11747	0.28737	242,889,464	350,524,185
137	0.04696	0.10292	128,852,402	220,388,181
138	0.21054	0.06472	95,473,773	172,695,938
139	-0.52043	-0.46821	110,423,027	172,635,461
140	0.44997	0.13979	65,300,005	107,562,055

No.	Bootstrap 1	Bootstrap 2	Sibling Series 1	Sibling Series 2
141	-0.14785	-0.09013	115,154,575	139,097,322
142	0.26221	0.32836	184,593,121	236,221,825
143	-0.00303	-0.19270	512,788,703	701,080,829
144	0.19377	-0.01331	766,961,548	867,422,247
145	0.03499	-0.02905	870,609,950	800,475,668
146	-0.50505	0.00177	811,758,435	700,066,903
147	0.22056	-0.14230	459,012,787	657,121,175
148	-0.02905	-0.09013	348,184,540	346,772,929
149	0.00177	0.32836	159,536,295	149,474,843
150	-0.14230	-0.19270	112,985,933	146,746,685
151	0.21458	-0.01331	91,825,671	113,401,186
152	-0.19626	-0.02905	113,246,468	111,355,119
153	0.06175	0.00177	104,650,080	121,630,650
154	0.04585	-0.14230	206,872,395	226,440,562
155	-0.00529	0.04506	462,870,773	419,759,638
156			690,740,111	658,751,916

REFERENCES

Al-Fattah, S.M., “Time Series Modeling for U.S. Natural Gas Forecasting”, *E-Journal of Petroleum Management Economics*, Vol.1, 2006.

Aras, H. and Aras, N., “Forecasting Residential Natural Gas Demand”, *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects*, Vol. 26:5, pp.463-472, 2004.

BOTAS Annual Report 2005, Boru Hatları ile Petrol Tasima A.S (BOTAS), http://www.botas.gov.tr/icerik/docs/faalrapor/2005/fr2005_full.pdf, 2005.

BOTAS Annual Report 2006, Boru Hatları ile Petrol Tasima A.S (BOTAS), http://www.botas.gov.tr/icerik/docs/faalrapor/2006/fr2006_full.pdf, 2006.

BOTAS Annual Report 2007, Boru Hatları ile Petrol Tasima A.S (BOTAS), http://www.botas.gov.tr/icerik/docs/faalrapor/2007/fr2007_full.pdf, 2007.

Box, G.E.P. and Jenkins, G.M. *Time Series Analysis, Forecasting and Control*, San Francisco: Holden-Day (revised edn published 1976), 1970.

Bowerman, B.L., O’Connell, R. and Koehler, A., *Forecasting, Time Series and Regression, 4th Edition*, South-Western College Pub., 2004.

Brown, R.H. and Matin, I., “Development of Artificial Neural Network Models to Predict Daily Gas Consumption”, *Proceedings of the 1995 IEEE IECON 21st International Conference*, Orlando, Vol.2, pp.1389-1394, 1995.

Brown, R.H., Marx, B.M and Corliss, G.F., *Mathematical Models for Gas Forecasting*, GasDay Report No.129, 2005.

Carlstein, E., “The Use of Subseries Values for Estimating the Variance of General Statistic from A Stationary Sequence”, *The Annals of Statistics*, Vol.14, pp.1171-1179,1986.

Carlstein, E., Do, K.-A., Hall, P., Hesterberg, T., and Kunsch, R., "Matched-Block Bootstrap for Dependent Data", *Bernoulli*, Vol.4(3), pp.305-328, 1998.

Carlstein, E., Resampling *Techniques for Stationary Time Series: Some Recent Developments*, in D. Brillinger (Eds.), pp.75-85, Springer-Verlag, New York, 1993.

Chatfield, C. *The Analyses of Time Series: An Introduction, Fifth Edition*, Chapman & Hall, 1996.

Chatfield, C., *Time-Series Forecasting*, Chapman & Hall/CRC, 2000.

Demirel, O.F., *Applications of the Bootstrap in Scenario Generation*, Ph.D. Thesis, Rensselaer Polytechnic Institute, 2000.

Demirel., O.F., and Willemain, T.R., "A Turing Test of Bootstrap Scenarios", *Journal of Computational and Graphical Statistics*, Vol.11(4), pp.896-909, 2002a

Demirel., O.F., and Willemain, T.R., "Generation of Simulation Input Scenarios Using Bootstrap Methods", *Journal of the Operational Research Society*, Vol.53, pp.69-78, 2002b.

Durmayaz, A., Kadioglu, M. and Sen, Z., "An Application of the Degree-Hours Method to Estimate the Residential Heating Energy Requirement and Fuel Consumption in Istanbul", *Energy*, Vol.25, pp.1245-1256, 2000.

Ediger, V.S. and Akar, S. "ARIMA Forecasting of Primary Energy Demand by Fuel in Turkey", *Energy Policy*, Vol.35, pp.1701-1708, 2007.

Efron, B. and Tibshirani, R., "Bootstrap Methods for Standard Errors, Confidence Intervals, and Other Measures of Statistical Accuracy", *Statistical Science*, Vol.1, pp.64-77, 1986.

Gil, S. and Deferrari, J., "Generalized Model of Prediction of Natural Gas Consumption", *Journal of Energy Resources Technology*, Vol.126, pp.90-98, 2004.

Gumrah, F., Katircioglu, D., Aykan, Y., Okumus, S. and Kilincer, N. "Modeling of Gas Demand Using Degree-Day Concept: Case Study for Ankara", *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects*,Vol.23:2, pp.101-114, 2001.

Gutierrez, R., Nafidi, A., and Sanchez G. R., “Forecasting Total Natural-Gas Consumption in Spain by Using the Stochastic Gompertz Innovation Diffusion Model”, *Applied Energy*, Vol.80, pp.115-124, 2005.

Hall, P., Horowitz, J.L., and Jing, B., “On Blocking Rules for the Bootstrap with Dependent Data”, *Biometrika*, Vol.82, pp.561-74, 1995.

Heizer, J. and Render, B., *Principles of Operations Management*, *6th Edition*, Pearson / Prentice Hall, 2006.

Huang, H., and Willemain, T.R., “Input Modeling for Financial Simulations Using the Bootstrap”, *Journal of Simulation*, Vol.1, pp.39-52, 2006.

Ivezic, D., “Short-Term Natural Gas Consumption Forecast”, *FME Transactions*, Vol.34, pp.165-169, 2006.

Kedem, B., *Time Series Analysis by Higher Order Crossings*, IEEE Press: Piscataway, NJ., 1993.

Kenisarian, M. and Kenisarina, K., “Energy Saving Potential in the Residential Sector of Uzbekistan”, *Energy*, Vol.32, pp.1319-1325, 2007.

Khotanzad, A., Elragal, H. and Lu, T.L., “Combination of Artificial Neural-Network Forecasters for Prediction of Natural Gas Consumption”, *IEEE Transactions on Neural Networks*, Vol.11, pp.464-473, 2000.

Kilic, A.M., “Turkey’s Natural Gas Necessity, Consumption and Future Perspectives”, *Energy Policy*, Vol.34, pp.1928-1934, 2006.

Kim, Y., Willemain, T.R., Haddock, J., and Runger, G., “The Threshold Bootstrap: A New Approach to Simulation Output Analysis”, *Proceedings of the 1993 Winter Simulation Conference*, G. Evans, M. Mollaghasemi, E. Russell and W.Biles (Eds.), pp. 498-502, IEEE Press: Piscataway, 1993c.

Kim, Y., Willemain, T.R., Haddock, J., and Runger, G., *Simulation Output Analysis Using the Threshold Bootstrap*, Technical Report No.37-93-378, Department of Decision Sciences and Engineering Systems, Rensselaer Polytechnic Institute, Troy, New York, 1993b.

Kirchgassner, G. and Wolters, J., *Introduction to Modern Time Series Analysis*, Springer, 2008.

Kizilaslan R. and Karlik B., "Combination of Neural Networks Forecasters For Monthly Natural Gas Consumption Prediction", *Neural Network World*, Vol. 19, pp. 191-199, 2009.

Kunsch, H. and Carlstein, E., *The Linked Blockwise Bootstrap for Serially Dependent Observations*, Tech. Rep., Dept. of Statist., Univ. of North Carolina, Chapel Hill, 1990.

Kunsch, H., "The Jackknife and the Bootstrap for General Stationary Observations", *The Annals of Statistics*, Vol.17, pp. 1217-1241, 1989.

Lall, U. and Sharma, A., "A Nearest Neighbor Bootstrap for Resampling Hydrologic Time Series", *Water Resources Research*, Vol.32, pp.679-693, 1996.

Leger, C., Politis, D.N., and Romano, J.P., "Bootstrap Technology and Applications", *Technometrics*, Vol.16, pp.378-398, 1992.

Liu, H., Liu, H., Zheng, G., Liang, Y. and Ni, Y., "Research on Natural Gas Load Forecasting Based on Support Vector Regression", *Fifth World Congress on Intelligent Control and Automation*, 15-19 June 2004, Vol.4, 3591A-3595, 2004.

Liu, L.M. and Lin, M.W., "Forecasting Residential Consumption of Natural Gas Using Monthly and Quarterly Time Series" *International Journal of Forecasting*, Vol.7, pp.3-16, 1991.

Liu, R. and Singh, K., "Moving Blocks Jackknife and Bootstrap Capture Weak Dependence", *Exploring the Limits of Bootstrap* , R. LePage and L. Billard (Eds.), Wiley, New York, 1992.

Ministry of Energy and Natural Resources, *MENR Web Site*, http://www.enerji.gov.tr/index.php, 2008.

Moving Average Model for Box-Jenkins Approach, *NIST/SEMATECH e-Handbook of StatisticalMethods*, http://www.itl.nist.gov/div898/handbook/pmc/section4/pmc444.htm#MA, 2006.

Natural Gas Consumption Data Archieve, *International Energy Agency*, http://www.iea.org/Textbase/stats/surveys/archives.asp, 2009.

Ozturk, H.K. and Hepbasli, A., "The Place of Natural Gas in Turkey's Energy Sources and Future Perspectives", *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects*, Vol.25, pp.293-307, 2003.

Park, D. and Willemain, T.R., “The Threshold Bootstrap and Threshold Jackknife”, *Computational Statistics and Data Analysis*, 1999.

Park, D., *The Threshold Bootstrap For Time Series Analysis*, Unpublished Dissertation, Dep. of Decision Sciences and Engineering Systems, Rensselaer Polytechnic Institute, 1997.

Politis, D.N., and Romano, J.P., “The Stationary Bootstrap”, *Journal of the American Statistical Association*, Vol.89, pp. 1303-1313, 1994.

Potocnik, P., Thaler, M., Govekar, E., Grabec, I., Poredos, A., “Forecasting Risks of Natural Gas Consumption in Slovenia”, *Energy Policy*, Vol.35., pp.4271-4282, 2007.

Romano, J.P. and Thombs, L.A., “Inference for Autocorrelations Under Weak Assumptions”, *Journal of the American Statistical Association*, Vol.51, pp. 590-600, 1996.

Sanchez-Ubeda, E. and Berzosa, A., “Modeling and Forecasting Industrial End-Use Natural Gas Consumption”, *Energy Economics*, Vol.29, pp.710-742, 2007.

Sarak, H. and Satman, A., “The Degree-Day Method to Estimate the Residential Heating Natural Gas Consumption in Turkey: A Case Study”, *Energy*, Vol.28, pp.929-939, 2003.

Siemek, J., Nagy, S., and Rychlicki, S., “Estimation of Natural-Gas Consumption in Poland Based on the Logistic-Curve Interpretation”, *Applied Energy*, Vol.75, pp.1-7, 2003.

Souza, R.C. and NETO, A.C., “A Bootstrap Simulation Study in ARMA(p,q) Structures”, *Journal of Forecasting*, Vol.15, pp. 343-353, 1996.

Stevenson, W., *Operations Management, 10th Edition*, McGraw-Hill/Irwin, 2008.

Goktas, O., *Teorik ve Uygulamalı Zaman Serileri Analizi*, Besir Kitabevi,2005.

Thombs, L.A. and Schucany, W.R., “Bootstrap Prediction Intervals for Autoregression”, *Journal of the American Statistical Association*, Vol.85, pp.486-492, 1990.

Thombs, L.A., *Prediction for ARMA(p,q) Series Using the Bootstrap*, Instit. Math. Statist. Bulletin, pp.225, 1987.

Turkiye ve Istanbul'da Gaz Tuketimi Analizi Raporu, Istanbul Gaz Dagitim A.S. (IGDAS), December 2006.

Viet, N.H., and Mandziuk, J., "Neural and Fuzzy Neural Networks for Natural Gas Consumption Prediction", *IEEE 13th Workshop on Neural Networks for Signal Processing*, 17-19 September 2003, pp.759-768, 2003.

Willemain, T. R., Park, D. S., Kim,Y.B., Shin, K.I., "Simulation Output Analysis Using the Threshold Bootstrap", European Journal of Operational Research, Vol.34, pp.17-28, 2001.

Wong-Parodi, G., Dale, L. and Lekov, A., "Comparing Price Forecast Accuracy of Natural Gas Models and Futures Markets", *Energy Policy*, Vol.34, pp.4115-4122, 2006.