MODELING MONTHLY ELECTRICITY DEMAND IN TURKEY FOR 1990-2006

A THESIS SUBMITTED TO
GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

DUYGU KÜÇÜKBAHAR

IN PARTIAL FULFILLMENT OF THE REQIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
INDUSTRIAL ENGINEERING

FEBRUARY 2008

Approval of the thesis:

MODELING MONTHLY ELECTRICITY DEMAND IN TURKEY FOR 1990 - 2006

submitted by **DUYGU KÜÇÜKBAHAR** in partial fulfillment of the requirements for the degree of **Master of Science in Industrial Engineering Department, Middle East Technical University** by,

Prof. Dr. Canan Özgen
Dean, Graduate School of **Natural and Applied Sciences** _______________

Prof. Dr. Nur Evin Özdemirel
Head of Department, **Industrial Engineering** _______________

Prof. Dr. Gülser Köksal
Supervisor, **Industrial Engineering Dept., METU** _______________

Examining Committee Members:

Assoc. Prof. Dr. Esra Karasakal
Industrial Engineering Dept., METU _______________

Prof. Dr. Gülser Köksal
Industrial Engineering Dept., METU _______________

Dr. Özlem Türker
Industrial Engineering Dept., Cankaya University _______________

M.S. Nuh Nadi Bakır
Director of Board, ERE Mühendislik Construction and Trade Co. _______________

Dr. Merih Aydınalp Köksal
Environmental Engineering, Hacettepe University _______________

Date: February 6 , 2007

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last name: DUYGU KÜÇÜKBAHAR

Signature:

ABSTRACT

MODELING MONTHLY ELECTRICITY DEMAND IN TURKEY FOR 1990-2006

KÜÇÜKBAHAR, DUYGU

M.S., Department of Industrial Engineering

Supervisor: Prof. Dr. Gülser Köksal

February 2008, 165 pages

Factors such as economical development, rapid increase in population and climate change increased electricity demand in Turkey as well as in other countries. Thus, using the correct methods to estimate short, medium and long term electricity demand forms a basis for the countries to develop their energy strategy. In this study, monthly electricity demand of Turkey is estimated. First, the effect of natural gas price and consumption to electricity demand and elasticities are searched with a simple regression model. Although, natural gas is known as a substitute of electricity, natural gas consumption and natural gas over electricity price ratio are found to be nearly inelastic. Second part includes two models and cointegration relation is investigated in nonstationary industry production index, electricity consumption per capita and electricity prices series in the first one. An error correction model is then formed with an additional average temperature variable and 12 months electricity demand is forecasted. In the second one, heating degree-days and cooling degree-days are used instead of the average temperature variable and a new error correction model is formed.

The first model performs better than the second one, indicating the seasonality of electricity consumption during a year. The results of both models are also compared with previous studies to investigate the effect of different weather variables.

Key Words: Unit Root Test, Electricity Demand, Cointegration, Error Correction Model, Natural Gas

ÖZ

1990-2006 YILLARI İÇİN TÜRKİYE'DEKİ AYLIK ELEKTRİK TALEBİNİN MODELLENMESİ

KÜÇÜKBAHAR, DUYGU

Yüksek Lisans, Endüstri Mühendisliği Bölümü

Tez Yöneticisi: Prof. Dr. Gülser Köksal

Şubat 2008, 165 sayfa

Ekonomik gelişmeler, hızlı nüfus artışı ve iklim değişikliği gibi birçok etken, elektriğe olan talebin Türkiye'de hızla artmasına sebep olmuştur. Kısa, orta ve uzun dönemli elektrik tüketiminin doğru metotlarla tahmin edilmesi, ülkelerin enerji stratejilerini belirlenmesinde temel arz edecektir. Bu çalışmada Türkiye'deki ortalama aylık elektrik tüketimi tahmin edilmiştir. İlk aşamada elektriğin ikamesi olan doğalgaz tüketimi ve fiyatının, elektrik tüketimine olan etkisi basit regresyon modeli ile araştırılmış ve esneklikler incelenmiştir. Her ne kadar doğal gaz, elektriğin ikamesi olarak görülse de doğalgaz tüketimi ve doğal gaz fiyatının elektrik fiyatına oranının elektrik tüketimine karşı esnek olmadığı görülmüştür. İkinci aşama, iki değişik model içermektedir. Bunlardan ilkinde sanayi üretim endeksi, ortalama sıcaklık ve elektrik fiyatı bağımsız değişken olarak ele alınmış, durağan olmayan serilerde koentegrasyon ilişkisi incelenmiş ve son olarak hata düzeltme modeli kullanılarak 12 aylık elektrik talep tahmininde bulunulmuştur. İkincisinde ise ortalama sıcaklık değişkeni yerine ısıtma gün derece ve soğutma gün derece değişkenleri kullanılarak bir hata düzeltme modeli oluşturulmuştur. İki modelin sonuçları birbirleriyle

karşılaştırıldığında ilk modelin daha uygun sonuçlar verdiği görülmüştür. Buna ek olarak, modellerin sonuçları, farklı sıcaklık değişkenlerinin elektrik tüketimine olan etkisini araştırmak amacıyla geçmişte yapılan çalışmalarla karşılaştırmıştır.

Anahtar Kelimeler: Birim Kök Testi, Elektrik Tüketimi, Koentegrasyon, Hata Düzeltme Modelleri, Doğal Gaz

To My Family

ACKNOWLEDGMENTS

I would like to express my deepest gratitude to my supervisor Prof. Dr. Gülser Köksal for her guidance, encouragement and endless support.

I particularly would like to thank to Özlem Türker for all her valuable advices, comments and suggestions throughout the study and for sharing all the statistical knowledge with me. I owe much to her patience and insight.

I wish to thank to the members of my examining committee; Esra Karasakal, Nuh Nadi Bakır and Merih Aydınalp Köksal for their contributions to this study.

I would also like to express my sincere thanks to my bosses and colleagues at ERE for their patience and endless support. I am grateful to my homemate Meryem Tatlıer and my friends Demet Çetiner and Deniz Yanık for their encouragement and patience throughout thesis.

Finally, I would like to express my appreciation for my family for the endless love and understanding they gave me throughout my life. I appreciate also Hasan Onur Beygo for his invaluable support, kindness and for being in my life.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVATIONS

ACF	:	Autocorrelation Function
ADF	:	Augmented Dickey Fuller
ADL	:	Autoregressive Distribution Lag Modeling
AIC	:	Akaike Information Criteron
ANN	:	Artificial Neural Networks
AR	:	Autoregressive
ARCH	:	Autoregressive Conditional Heteroskedasticity
ARCH-M	:	Autoregressive Conditional Heteroskedasticity in Mean
ARIMA	:	Autoregressive Integrated Moving Average Model
ARMA	:	Autoregressive Moving Average
BOT	:	Build-Operate-Transfer
BO	:	Build-Operate
BOTAS	:	Petroleum Pipeline Corporation
BTC	:	Baku-Tbilisi-Ceyhan
CDD	:	Cooling Degree Days
CONS	:	Electricity consumption per capita series between the years 1990 and 2006
CONS00	:	Electricity consumption per capita series between the years 2000 and 2006
CRDW	:	Cointegrating Regression Durbin Watson
CUSUM	:	Cumulative Sum of Squares
DPT	:	State Planning Organization
DSI	:	State Hydraulic Works
DW	:	Durbin Watson
ECM	:	Error Correction Models
EDAS	:	Electricity Distribution Company
EGARCH	:	Exponential Generalized Autoregressive Conditional Heteroskedasticity
EIA	:	Energy Information Agency
EIE	:	General Directorate of Electrical Power Resources Survey and Development Administration
EMM	:	Electricity Market Module
EMO	:	Chamber of Electrical Engineers
ENPEP	:	The Energy and Power Evaluation Program
EPDK	:	Additionally Energy Market Regulatory Authority
ETKB	:	Ministry of Energy and Natural Resources
EUAS	:	Electricity Generation Company

FIGARCH	:	Fractionally Integrated Generalized Autoregressive Conditional Heteroskedasticity
GARCH	:	Generalized Autoregressive Conditional Heteroskedasticity
GASCONS00	:	Natural gas consumption between the years 2000 and 2006
GDP	:	Gross Domestic Product
GHG	:	Green House Gases
HDD	:	Heating Degree Days
HEGY	:	Hylleberg, Engle, Granger and Yoo
IEA	:	International Energy Agency
IGARCH	:	Integrated Generalized Autoregressive Conditional Heteroskedasticity
INDEX	:	Industry production index
KPSS	:	Kwiatkowski, Phillips, Schmidt, Shin
LM	:	Lagrange Multiplier
MA	:	Moving Average
MAED	:	Model for Assessment of Energy Demand
MAED	:	Mean Absolute Error
MAPE	:	Mean Absolute Percentage Error
MLE	:	Maximum Likelihood Estimation
MMST	:	Million Short Tons
MTA	:	General Directorate of Mineral Research and Exploration
MTOE	:	Million Tons of Oil Equivalent
MYTM	:	National Load Dispatch Center
NEMS	:	National Energy Modeling System
OECD	:	Organization of Economic Coordination and Development
OLS	:	Ordinary Least Squares
OPEC	:	Organization of the Petroleum Exporting Countries
PACF	:	Partial Autocorrelation Function
PRICE	:	Real electricity price
PRICERATIO00	:	Natural gas price / electricity price between the years 2000 and 2006
RBFN	:	Radial Basis Function Networks
RMSE	:	Root Mean Squared Error
SIC	:	Schwarz (Bayesian) Information Criterion
TARCH	:	Threshold Autoregressive Conditional Heteroskedasticity
TEAS	:	Turkish Electricity Company
TEDAS	:	Turkish Electricity Distribution Company
TEIAS	:	Turkish Electricity Transmission Company
TEK	:	Turkish Electricity Corporation
TEMP	:	Average Temperature
TETAS	:	Turkish Electricity Trading Company

TIC	:	Theil Inequality Coefficient
TKI	:	Turkish Hard Coal Enterprises
TMMOB	:	Union of Chambers of Turkish Engineers and Architects
TOBB	:	The Union of Chambers and Commodity Exchanges of Turkey
TOR	:	Transfer of Operating Rights
TSO	:	Transmission System Operator
TUFE	:	Consumption Price Index
TUIK	:	Turkish Statistical Institute
UCTE	:	Union for the Coordination of Transmission of Electricity
UNCCC	:	United Nations Climate Change Charter
UNFCCC	:	United Nations Framework Convention on Climate Change
VAR	:	Vector Autoregression
VECM	:	Vector Error Correction Models
WASP	:	Wien Automatic System Planning Package

1 TWh = 10^3 GWh = 10^6 MWh = 10^9 kWh
1 m^3 natural gas = 39 mega joules = 10.8

CHAPTER I

INTRODUCTION

One of the most discussed issues in the world currently is energy supply security and running out primary energy resources emanating from climate change, environmental pollution and technological growth. Turkey, being a developing country, is also facing with these problems since, parallel with the economical growth, usage of energy resources increased in Turkey, too. Recent works performed by Governmental Bodies state that electricity supplied, which is a secondary energy resource that does not have any substitute in many of its usage areas, will not be able to meet the demand in year 2013 with an optimistic scenario (Turkish, Electricity Transmission Company [TEIAS], 2007, a). In addition to its difference from other products and services for being considered as both an input and an output, electricity is an important resource for not being stored, that is, it should be used whenever produced. All of these indicators accelerated the reform and privatization movements in Turkey, especially after 2001, when the Electricity Market Law (Law number: 4628) was enacted. The Law aimed to supply electricity with low cost and high quality to the customers. Following this Law, electricity demand forecast studies gained more importance, as the problem with the electricity supply and demand balance was mainly because of the wrong policies due to wrong demand forecasts performed previously. The models used by Government in the past did not perform well and always forecasted demand more than the actual consumption because of technical problems and bad assumptions, resulting with excess capacity, wrong investments like Build-Operate-Transfer (BOT), Build-Operate (BO) and Transfer of Operating Rights (TOR) projects, external dependence, higher electricity prices and an uncompetitive environment (Keleş, 2005). This example and many other examples in the world showed that high forecast may cause constructing too many plants, high level of reserves and high costs, whereas low forecasts may result in shortage of electricity, increase in prices, poor service quality and electricity cut-offs.

Electricity demand forecasting studies begun with long term forecasting in the world. Long term forecasts are typically performed for 5 years to 25 years and they are used in generation, transmission and tariff planning and scheduling, feasibility studies and developing strategies such as expansion of utilities or making new investments. Economic variables such as population, income and price play an important role in consumption in the long term. Medium term forecasts are for a few weeks to few years and they consider both socio-economic and temperature variables. Medium term forecasting can be used for fuel procurement, scheduling unit maintenance and diversity interchanges. In addition to long and medium term forecasts, particularly free market conditions brought the requirement of good electricity demand estimates in the short term because electricity prices are set hourly by the electricity generators in the new pool system in many countries such as UK, Sweden and at last in Turkey [1]. Short-term electricity demand forecasts are from a few minutes to few weeks forecasts. They are useful in controlling and scheduling power plants to generate electricity. These short term demand forecasts help to decide which equipments should be operated to meet the demand at that period, how to optimize generation thus minimize cost with supply security. Temperature, day of the week, football matches, sunlight, holidays and religious days all have important impacts on electricity consumption in the short term.

Various methods have been developed for energy and electricity demand forecasting studies during the last decade. First studies included statistical models, however artificial neural networks (ANN) and neural fuzzy logics outspreaded in recent years. End-use method and disaggregation, econometric methods such as error correction models (ECM) and regression, time series methods such as smoothing methods, decomposition, and hybrid approaches are the most well-known approaches used in electricity demand forecasting.

The number of electricity demand forecast models is limited in Turkey. Additionally, these are long term forecasts making yearly estimates. However, the importance of medium term forecasts should be considered as monthly data measures seasonality and fluctuations in a year and helps to provide supply security. Moreover, work done by public sector, using the Model for Assessment of Energy Demand (MAED) always gave inaccurate results with high percentage of errors, which leads to wrong planning and wrong investments (Kumbaroğlu, 2006). Most of the other studies cover only electricity consumption data from the past and do not use economical indicators, which have great effect on electricity consumption. This study aims to overcome these problems by making monthly forecasts using econometric and recent data. Furthermore, it intends to realize the seasonal fluctuations in the electricity consumption to assist generators to plan and schedule power plants in medium term, develop policies for supply and demand balance, cope with flexible requirements, help the generators to offer prices and make bids in the newly defined electricity trading system. Public sector can also benefit from this study for scheduling power plants, scheduling maintenance and planning fuel procurement.

Chapter II gives general information about the energy and electricity sector in Turkey. Firstly, energy resources and their usage are explained and then, developments in electricity sector since 1900's, when it was first started to be used in Turkey, is defined regarding the structure of electricity sector.

In Chapter III, forecasting methodologies and previous works done about electricity demand in Turkey and in other countries are described briefly. Additionally, the model used currently by the Government is explained. This chapter also covers the details of the cointegration and ECM, which is the methodology used in this study for its good performance. First of all, the series are tested for stationarity using seasonal unit root tests. Then, if the series have cointegrating relation, they can be modelled with ECMs. A few, yearly electricity demand forecasting studies are performed using ECM in the past in Turkey.

Next in Chapter IV, effect of natural gas consumption and natural gas price to electricity consumption is searched for being a substitute of electricity with 2000-2006 data and natural gas variables are found to be significant. Other substitutes of electricity such as oil and coal are disregarded in this study because of lack of data. Natural gas variables are not used while modeling electricity demand because of the short data period. In the first step of the modeling, Hylleberg, Engle, Granger and Yoo (HEGY) test, being different from former studies, is used to investigate seasonal unit roots in the monthly series. Secondly, two cointegration relation and error correction models are formed for electricity demand using different indicators for weather variable [2] and both short and long run elasticities are investigated. Monthly data between 1990 and 2005 is used in the study and 2006 data is used to compare forecasted and actual values. Apart from other electricity demand forecast studies in Turkey, conditional variance is also investigated with Autoregressive Conditional Heteroskedasticity test. The models and forecasts are evaluated using statistical indicators such as percentage error and mean absolute percentage error. The modeling and analysis are performed using E-Views 4.1 package program.

Finally, in Chapter VI, evaluation of the study and conclusion are given as well as future directions.

CHAPTER 2

ENERGY AND ELECTRICITY IN TURKEY

2.1 ENERGY

Energy is one of the key stones of development and globalization and that is why countries that are rich in fuels get stronger. Fossil fuels have great importance in the world's energy production. In 18^{th} century the countries that used their coal reserves in industry developed economically more than others. However, in 20^{th} century petroleum and natural gas gained importance and Organization of the Petroleum Exporting Countries (OPEC) is founded by 13 countries to organize and coordinate the oil supply and stabilize the petroleum market, which now controls the prices. A cartel for natural gas is discussed for years and is said to be founded in the near future by natural gas exporting countries like Russia, Iran and Qatar.

Total energy supply in the world in 2004 was 11059 Mtoe, whereas it was 6035 Mtoe in 1973 according to the Key World Energy Statistics published by International Energy Agency (IEA) in 2006, which shows energy utilization increased nearly twice in the world (Figure 2.1). In 2004, nearly %85 of world's energy demand was met from fossil fuels like coal, natural gas and petroleum and nearly % 25 of total supply was from coal. Oil has the biggest share both in 1974 and 2004. Usage of both natural gas and nuclear energy increased in 30 years time. Hydro and other renewable energy resources still have small shares. However it is a fact that the interest in renewable energy will be increasing around the world in the near future since researches performed all over the world found out that primary energy resources will be depleted.

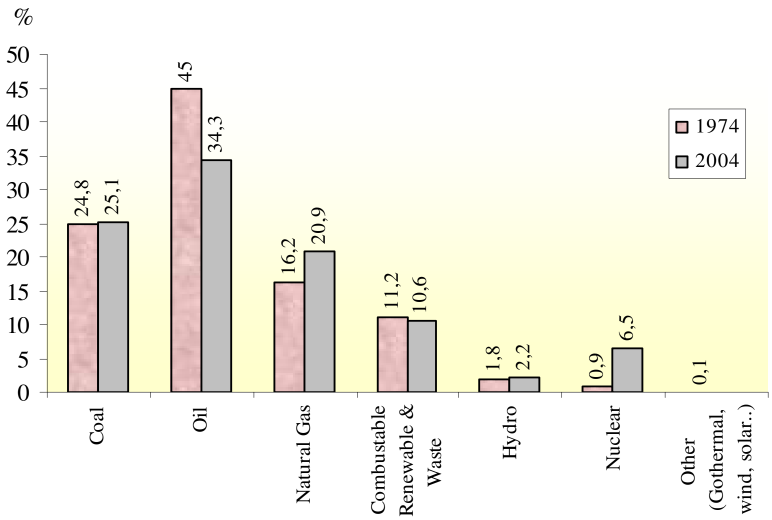

Figure 2.1 – Total Primary Energy Supply in the World (IEA, 2006a)

Excess demand for energy, scarcity of resources and environmental effects of fossil fuels caused alternative fuels to be considered today. For instance, petroleum, natural gas and coal reserves are said to run out in 40-50, 60-67 and 240-250 years time respectively. Additionally, fossil fuels have the disadvantage of green house gases (GHG) emission at higher rates. To illustrate, 1 kWh electricity production from a coal fired power plant is equivalent to 900 – 1000 grams of CO_2 emission. CO_2 emission values in gCO_2/kWh from other resources are given in the next page (Ministry of Energy and Natural Resources [ETKB], 2005).

Type of Power Plant	CO2 Emission
Coal	975
Fuel Oil	742
Natural Gas	608
Combined Heat and Power	518
Solar	53
Wind	29
Nuclear	22
Geothermal	15
Hydro	11

Energy resources can be divided into two as primary and secondary energy resources; where primary energy resources are the resources that produce energy and secondary energy resources are produced by using primary energy resources. Energy resources according to their types are given in Figure 2.2 (Union of Chambers of Turkish Engineers and Architects [TMMOB], 2006).

Primary energy resources are divided into three as new and renewable resources, nuclear energy and traditional resources. With developments in technology and increase in demand for sustainable and clean energy, renewable energy resources started to be preferred in power plants. For instance 5000 kW wind turbines can be manufactured now, while in 1980's their power was just around 100 kW. Wave energy will be used in near future to meet the energy demand in many countries as a clean and efficient resource. In Japan, which is a country dependent to petroleum imports, the target is to built a 5000 MW photovoltaic (PV) farm in 2016 and the energy from that farm will be mostly used in households (Kızak, 2006). Secondary energy resources can be named as hydrogen and electricity synthetic oil.

Among the primary and traditional energy resources, Saudi Arabia has the biggest share in producing crude oil with 519 Million Tons (Mt) forming the 13.2% of the world total in 2005. Natural gas is produced and exported at most in Russia and United States. The amount produced in these countries in 2005 accounted for %21.8 and %18 of the total production respectively. On the other hand, United States was the leading importer of natural gas in 2004. Total amount of hard coal produced in the world was 4973 Mt in 2005; of which China and United States has the largest shares, with 2226 Mt and 951 Mt. %46 of the nuclear energy was produced in United States and France in 2004. China, Canada and Brazil were the major producers of hydroelectricity in the world with 543 TWh, 341 TWh and 321 TWh and shares 12.6%, 12.1% and 11.1%, respectively (IEA, 2006b).

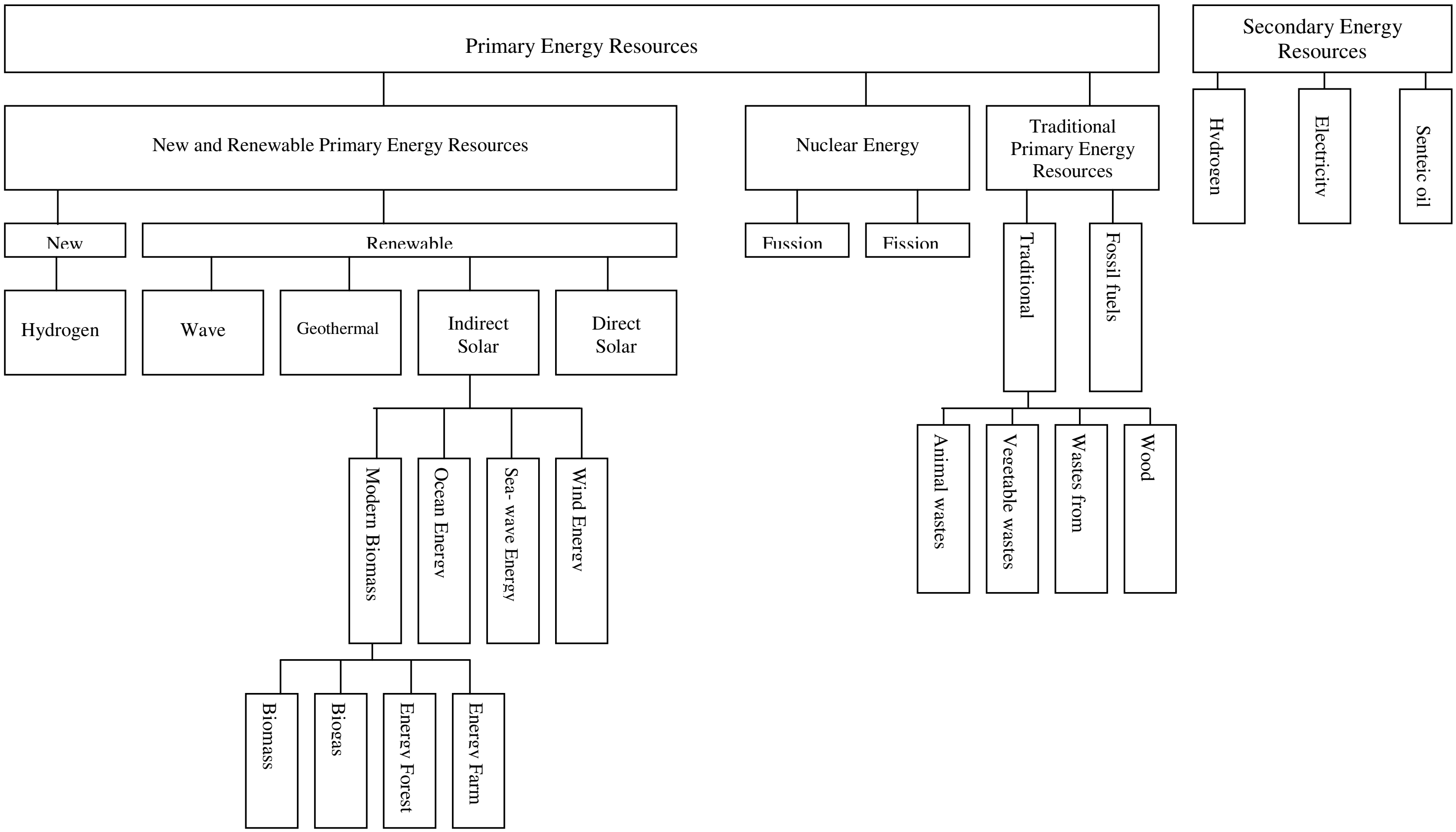

Figure 2.2 : Primary and Secondary Energy Resources (TMMOB, 2006)

On the other hand, demand side is growing rapidly too with the increase in fuel prices and uncertainties in economy. The demand in developing countries like Brazil, China and India is expected to have 70% of the total demand in 25 years time. On the other hand, 70% of the EU's energy demand will be met by imported products in 20-30 years time (The Union of Chambers and Commodity Exchanges of Turkey [TOBB], 2007). It is estimated that, energy demand in 2020 will increase 65% compared to today's demand and it will increase %250 in 2050 (Kaştan, 2006). In Figure 2.3, which is drawn with the data from 2006 Key World Energy Statistics, the total energy consumption in the world in 1974 and 2004 is compared according to energy resources. Although the percentage of oil consumption is the highest, it decreased in 2004 because of new energy resources. Another important point is that consumption of secondary energy resources increased more than consumption of primary energy resources.

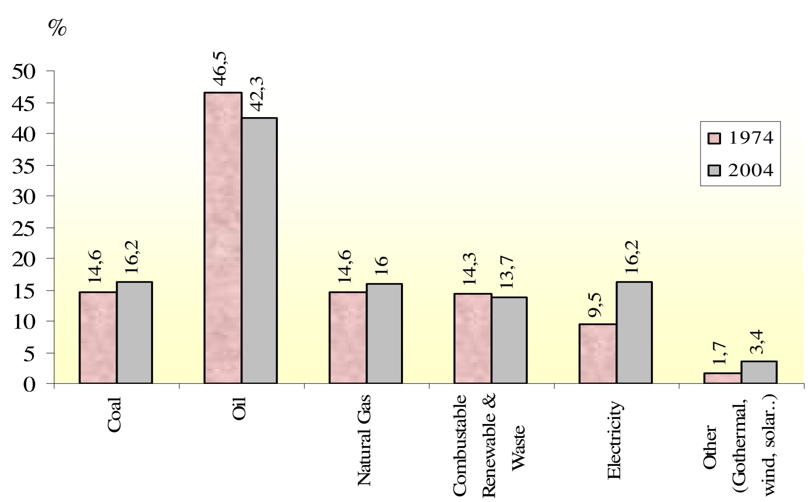

Figure 2.3 – Total Energy Consumption in the World (IEA, 2006a)

2.2 ENERGY RESOURCES IN TURKEY

Basic energy policy in Turkey is supplying high quality, economic and reliable energy to the consumers. Initiatives of the energy policy are determined with the 5 years progress plans and include supplying economic, reliable and sufficient energy, encouraging new investments to meet the increasing demand, energy supply security and to reach social and economic development targets.

Primary energy resources of Turkey are hard coal, lignite, asphaltite, oil, natural gas, hydroelectric, wind and geothermal energy. Although energy resources in Turkey are limited, being one of the largest countries with its area of 779452 km^2 and its strategic location made the country a natural energy bridge between Middle East and Europe (Energy Information Agency [EIA], 2006a). Ceyhan, located in the south part of Turkey became an energy terminal, where power plants are constructed, Baku-Tbilisi-Ceyhan (BTC) Pipeline, Kirkuk-Ceyhan Pipeline ends and Iraq oil is transferred. Moreover, Bosphorus is the main transfer point of the fuels within Turkey. BTC Pipeline is the first direct line where crude oil is transferred from Caspian Sea to Mediterranean without passing Russia. The pipeline cost $4 billion to build and it is 1100 mile long. Another pipeline Kirkuk-Ceyhan is the line where northern Iraq oil is exported. Moreover, with its growing population and economy, energy activities gained more importance in Turkey in recent years (EIA, 2006b). In January 2006, the oil reserves in Turkey were 300 million barrel, of which 43000 barrels per day is produced in the first nine months of 2006. Oil reserves are located in southeastern part of Turkey, especially in Hakkari. In the last years, energy generated from oil decreased in Turkey according to the high production costs, high inflation rates and economical crisis.

Turkey is especially new and significant user of natural gas. Although its reserves are limited with 300 billion cubic feet, the consumption was 793 billion cubic feet in 2004 and the Government signed contracts for additional natural gas for the year 2010. Natural gas market in Turkey is managed by state owned company, Petroleum Pipeline Corporation (BOTAS). Existing pipelines that are used to transfer natural gas to Turkey are Blue Stream and Iraq Turkey pipelines. Both of the lines are 750 miles long and their maximum capacity is 565 and 495 billion cubic feet per year. Blue Stream transport the natural gas from Russia to Turkey underneath the Black Sea and Iran-Turkey pipeline starts from Tabriz and ends in Ankara. There are some other proposed or under construction projects for natural gas such as South Caucasus Pipeline (Baku-Tbilisi-Erzurum) from Shah Deniz in Azerbaijan and Turkey-Greece Interconnector, where Turkey will be the transit way from Azerbaijan to Greece (EIA, 2006c). Natural gas in Turkey is used for electricity generation, in industry as fertilizer and mostly in households for heating in the winter.

Total coal reserves in Turkey is 4614 million short tons (Mmst) in 2004 and most of this is lignite and subbituminuos coal reserves. Nearly, %40 of the lignite is obtained from Afsin-Elbistan basin and hard coal is only available in Zonguldak. Although both private and public sector companies can take part in production, processing and distribution activities of lignite, hard coal production is done only by Turkish Hard Coal Enterprises (TKI) (EIA, 2006d).

One of the most important resources in Turkey is hydraulic with its 125 billion kWh economic potential and 430 billion kWh gross potential. Asphaltit reserves exist in Şırnak and Silopi and 82 million tons reserve is detected in this area. Considerable geothermal energy potential was detected in Turkey since 1962, especially in West (Denizli, Aydın, Manisa, Afyon, Balıkesir), Middle (Ankara, Kırşehir) and East (Sivas, Van) Anatolia. Turkey is also in the seventh rank in the world with its geothermal potential (Yiğitgüden, 1999). On the other hand, wind energy became popular after 2001, when Energy Market Law is enacted. Currently 58 generation licenses are issued for wind energy (EPDK, 2008). More detailed information about supply and consumption of primary energy resources in Turkey is given in Appendices A and B (Teknik Yayıncılık).

2.3 ELECTRICITY

Electricity consumption of a country is one of the most important signs of its economic, industrial and social development. Thus, electricity consumption per capita is higher in developed and industrialized countries compared to developing countries. Table 2.1 shows electricity demand per capita in Organization of Economic Coordination and Development (OECD) Countries. Demand for electricity increases promptly in the world because this secondary energy resource is clean and easy to use. On the other hand, electricity consumption is increasing faster in developing countries as a result of rapid increase in population, high rates of urbanization and economic growth, which are most used indicators of energy consumption. However, up to date data indicate that the rate of electricity consumption is much more than rate of population growth. Percentage increase in electricity consumption, gross domestic product (GDP) and population in the world is given in Table 2.2, whereas increase in electricity consumption and population in Turkey is given in Table 2.3. The difference between the rates is mainly as a result of economical growth in Turkey (ETKB, 2004).

Table 2.1 Electricity Consumption Per Capita in OECD Countries (ETKB, 2004)

Electricity Consumption Per Capita (kWh/capita)			
World	2516	**France**	7689
OECD	8204	**Germany**	7030
Iceland	28126	**Netherlands**	6923
Norway	24650	**Slovenia**	6835
Canada	17179	**Denmark**	6629
Finland	16784	**Checz Republic**	6224
Luxemburg	16509	**Spain**	5924
Sweden	15420	**Italy**	5644
Australia	11126	**Cyprus**	5415
Belgium	8579	**Greece**	5150
Austria	7850	**Turkey**	1766

Table 2.2. Average Growth Rates of Population, Electricity Consumption and GDP between 1971 and 2001 (ETKB, 2004)

	Austria	Spain	Portugal	Italy	Germany	England	Turkey	Korea
Population	0,27	0,55	0,52	0,23	0,16	0,17	2,12	1,22
Electricity Consumption	3,07	4,75	5,95	3,31	1,64	1,43	8,92	11,35
GDP	2,63	2,97	3,41	2,47	2,12	2,31	3,83	7,21

Table 2.3 - Increase in Electricity Consumption and Population (ETKB, 2004)

Year	Electricity Consumption (%)	Population (%)
1970	42,02	11,13
1975	45,83	12,46
1980	34,61	9,81
1985	30,55	11,70
1990	36,55	10,29
1995	30,53	8,22
2000	31,44	9,25
2005	24,54	5,91

Since this energy resource cannot be stored, it should be consumed whenever it is produced. This requires performing the electricity supply, planning, scheduling, transmission and distribution activities optimally while generating it with low cost and high quality. Electricity generation cost is composed of investment, operating and fuel costs in general. Among these, fuel cost takes an important place since fossil fuels have higher costs and fuel cost of renewable such as wind, solar and hydraulic is zero. When marginal costs are compared, electricity generated from hydraulic resources has the lowest cost. Lignite, natural gas, coal, wind, petroleum and nuclear power plants follow it (Akbank, 2006). Electricity generation investment costs from various fuels in USA are indicated in Table 2.4 for the years 2000 and 2050. A key point about investment costs is that, the costs will be decreasing instead of increasing with technological development in the following years (Üzmen, 2007).

Table 2.4 - Investment Costs of Fuels in USA (Üzmen, 2007)

Fuel Type	Investment Costs ($/kW)	
	2000	**2050**
Coal	1000-1650	1000-1650
Petroleum	600-800	440-730
Natural Gas	710-1150	640-910
Nuclear	1600-2800	1200-1640
Biomass	1570-1760	1240-1300
Solar (PV)	2900-5100	1150-1780
Wind	1400	750

Additionally Energy Market Regulatory Authority (EPDK) declared the unit investment costs for the year 2007 which are given in Table 2.5. The investment costs show that renewable energy resources are the most expensive, being the solar and geothermal the biggest.

Table 2.5 - Total Unit Investment Costs in Turkey for Electricity (EPDK, 2007)

Fuel	Total Unit Investment Cost (YTL/MW)
Coal	1.250.000
Natural gas/LPG	1.000.000
Fuel Oil / Nafta	1.000.000
Hydraulic	1.600.000
Wind	2.000.000
Geothermal	2.100.000
Biomass	1.700.000
Biogas	1.900.000
Solar	4.200.000
Other (except nuclear)	1.400.000

2.4 ELECTRICITY IN TURKEY

2.4.1 History of Electricity

In Turkey, electricity was firstly generated in Tarsus with a 2 kW dynamo in September 1902 and it was distributed to the village. However, the first organized electricity generation was in Istanbul Silahtarağa Power Plant which was founded by Hungarian Ganz Corporation, Banque Generale de Credit and Banque de Brexellese. The consortium was named as Ottoman Electricity Corporation and electricity generated started to be distributed to Istanbul in 11 February 1914. Ottoman Electricity Corporation was bought by Government in 1 July 1938. The first province to use electricity during Turkish Republic was Adapazarı, in 1923. In the following years, in 1935s, Etibank, General Directorate of Mineral Research and Exploration (MTA) and General Directorate of Electrical Power Resources Survey and Development Administration (EIE) were founded to take part in electricity generation. In 1948, the first regional plant, Zonguldak Çatalağzı Thermal Power Plant was founded. Sarıyar Power Plant is integrated to the system and Northwest Anatolia Interconnected System was formed in 1956. Additionally, Northwest Anatolia Electricity Turkish Company, Çukurova Electricity Company, Kepez Electrical Power Plant Trading Company were founded between 1952 and 1956. In 1960's, regional companies which are responsible for electricity generation, transmission, trading and distribution were named as Etibank Electricity Enterprises Institution. ETKB was founded in 1962 with the aim of national energy policy. Hydroelectrical power plants, operated by State Hydraulic Works (DSI) were transferred to Etibank in 1967. In 1970, Turkish Electricity Corporation (TEK) that can be described as a vertically integrated monopoly was founded with the Law numbered 1312 and took over the operation of the power plants.

In the following years, Government could not make new investments since it could not find appropriate foreign credits. Moreover, those were the days when liberalization and privatization movements begun in the World. In 1980's, privatization works started with the Law numbered 3096 and monopoly ended. Private companies that get necessary permissions from the Ministry and had right to sign contracts would be able to generate, transmit and distribute electricity. BO, BOT, TOR and autoproducer model came into force during this period. With some of these models; private sector were given privileges to construct and/or operate power plants and to sell the electricity produced to the Government with a tariff decided by both sides and purchase guarantee. At the end, those companies would transfer the power plants to the Government with all of its rights. On the other hand, autoproducers construct and operate power plants and they use the electricity they produced on their own or for their affiliates. They can also trade nearly %50 of the electricity that they generate according to the Commission Ruling of EPDK (26744), published on January 2, 2008. In 12 August 1993, TEK was rebuilt again as Turkey Electricity Generation and Transmission Company (TEAS) and Turkey Electricity Distribution Company (TEDAS) (Kulalı, 1997).

As mentioned before, Turkish electricity sector has been dominated by state-run monopolies for years. However, recent reform programs brought important changes to the sector, which allowed private investments. Government introduced the Electricity Market Law (4628) in 2001 with the purpose of:

"*ensuring the development of a financially sound and transparent electricity market operating in a competitive environment under provisions of civil law and the delivery of sufficient, good quality, low cost and environment-friendly electricity to consumers and to ensure the autonomous regulation and supervision of this market*"

and founded EPDK as a regulator body. TEAS was separated into three as TEIAS, Turkish Electricity Trading Company (TETAS) and Electricity Generation Company (EUAS). Afterwards, Government introduced a new Law, numbered 4646, on Natural Gas Market that concerns the liberalization of the natural gas market and formation of a financially sound, stable and transparent market.

In May, 2005 Law on utilization of renewable energy resources for the purpose of generating electrical energy (5346) was published:

"to expand the utilization of renewable energy resources for generating electrical energy, to benefit from these resources in secure, economic and qualified manner, to increase the diversification of energy resources, to reduce greenhouse gas emissions, to assess waste products, to protect the environment and to develop the related manufacturing sector for realizing these objectives."

Finally, Electricity Market Balancing and Settlement Regulation, which was published in 2004, aimed to stabilize supply and demand of electricity. This law also intended to form a new free electricity market where the players would take their place in the spot market, generators would make bids and the prices would be set hourly by a merit system like the one in England. With those laws and reforms, the confidence in energy sector and the amount of private sector investments is expected to increase. Today, EUAS controls nearly 50% of the electricity generation, TEIAS owns the entire transmission network, and 21 distribution companies, which were formed with separation of TEDAS controls the distribution.

2.4.2 Electricity Market Structure

The electricity market in Turkey forms 3% of the GDP and this account for 18 billion YTL when calculated with the wholesale energy prices. To meet the 6%-9% increase in yearly demand, increase the investments by private sector and strengthen the competitiveness in industry, electricity prices should be decreased and a competitive market structure should be formed (Deloitte, 2006). The electricity market consists of generation, transmission, distribution and trading activities in common. The structure of the electricity market, which is given in Figure 2.4 (Starodubtsev, 2006) is the free market structure that is planned to come into force.

Generation: Electricity generation activities in Turkey are done by both of the public and private sectors. The main electricity generators are state owned EUAS and its partners, private sector that runs BO, BOT and TOR projects, autoproducers and private generation companies with generation licenses. Some of the electricity demand is still supplied by imports.

EUAS: is the electricity generation company established as a successor company of TEAS. Generation is mainly based on natural gas, coal and hydroelectric power plants. The generated amount by EUAS in 2006 was 84530 GWh of which 68526.3 GWh was from domestic resources that are coal, hydro and geothermal. The remaining production was from natural gas and liquid fuels. Today EUAS and its affiliates own 21 thermal power plants and 106 hydroelectric power plants (EUAS, 2007).

Mobile Power Plants: Mobile Power Plants are thermal power plants owned by the Government in Turkey. They are generally operated under emergency conditions such as natural disasters or when there are blackouts. They can float in the sea or move in the land and by this way, they can be transferred to the place of urgency. The biggest disadvantage of these power plants is that it is too expensive to generate electricity with them and their damage to environment (Chamber of Electrical Engineers [EMO], 2006). There are 11 mobile power plants constructed in Turkey, however most of them are not operating now.

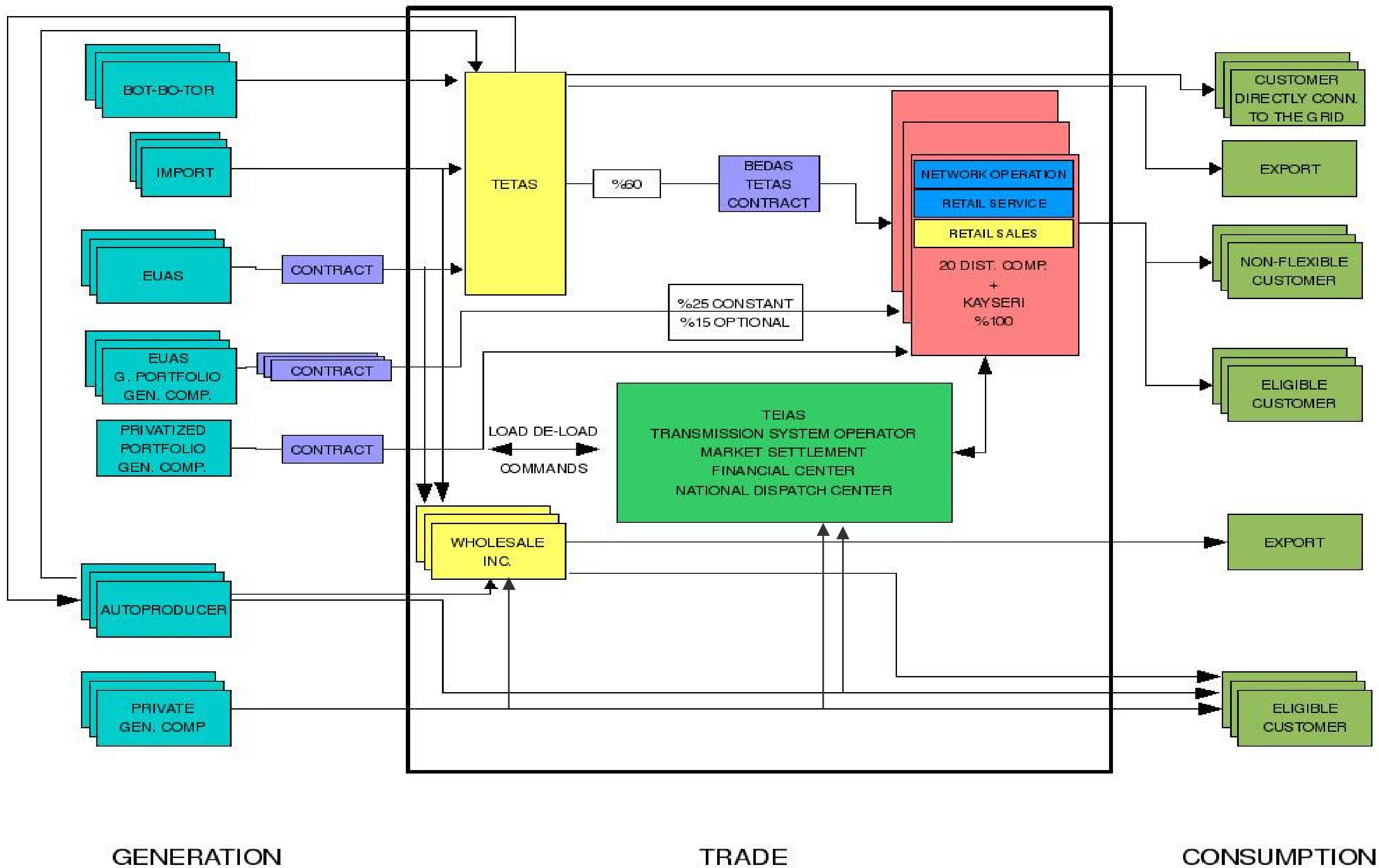

Figure 2.4 – Electricity Market Structure (Starodubtsev, 2006)

BOT: Build-Operate-Transfer model is first started to be discussed in 1984. The model predicts that investments are done by the private sector; subsequently they operate the power plant and sell the electricity to the public companies with a tariff decided with a privilege agreement. At the end of the agreement period, private sector transfers the ownership of the operating power plant to the Government with all the maintenance completed. This model provides a new financial choice to the Government and prevents budget deficit of the Government. Secondly, it prompts privatization and allows making investments in Turkey with foreign funds.

Third, it is admitted that private sector can construct power plants more efficiently, with low cost and in a short period of time to maximize their revenue. However, the model has disadvantages as well as its advantages. At the end of the agreement period, the power plants are transferred to the Government and this requires forming new bodies in public sector to operate these power plants. Additionally, the model is sensitive to political and economical stability because agreements are signed for longer periods of time such as 15 years and it does not predict any changes during this period (Imre, 2001). In April 2007, the number of operating BOT hydroelectric power plants were 18, with 3909 GWh yearly generation some of which are Sucatı, Yamula and Birecik Power Plants.

BO: The law on building and operating electrical power plants and regulating electricity sales with Build – Operate model (4283) is enacted in 1997 aiming to allow private companies to own, generate and operate power plants. However, the power plants operating with hydraulic, nuclear, geothermal and other renewable resources are out of this scope. The generating companies may get treasury guarantee for the electricity they generated according to their contracts with TEAS.

TOR*:* Transfer of Operating Rights is a kind of privatization where the power plants owned by Government are transferred to the private sector for a certain period. The owner of the plants is still the Government and the private sector makes profit by operating the plant and selling the electricity generated there.

Autoproducer and Autoproducer Groups: They are "*any legal entity engaged in electricity generation primarily for its own needs (or affiliates). Autoproducers can sell the excess electricity they produced under liberal market conditions"* (Electricity Market Law, 2001). In general these facilities are integrated with an industrial facility and they use hot gases, vapor or other wastes to generate electricity. The number of autoproducers increased after 1990's since generating the electricity by them was more profitable. In 2006, electricity generated by autoproducers was about 10% of the total production in Turkey (TEIAS, 2006). By December 2007, the number of autoproducer and autoproducer group licensed firms is 201.

Private Generator Companies: They are "*Any legal entity, except for autoproducers and autoproducer groups, engaged in generation of electricity and the sale of the electricity it has generated"* (Electricity Market Law, 2001). Private Generator Companies took part in electricity market after the Electricity Market Law was enacted in 2001. With this law, they are entitled to construct and operate power plants under liberalized market conditions with generation license. As of December 2007, the number of generation licenses that are issued by EPDK is 516, which means the private companies took their place in the liberalized electricity market. Electricity produced in 2006 with respect to generators is given in Figure 2.5.

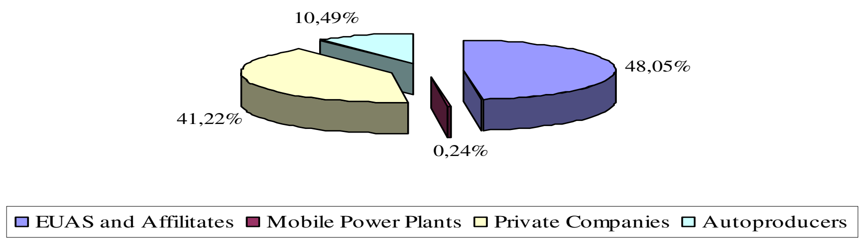

Figure 2.5 Electricity Generation Companies and Their Share (TUIK, 2007a)

Transmission: *Transmission is the transport of electricity through lines higher than 36 kV* (Electricity Market Law, 2001). TEIAS, Turkish Electricity Transmission Company Inc., established as a successor company of TEAS is in charge of transmission of electricity since 2001. TEIAS is also responsible for planning the investments of new transmission lines, controlling the operation, maintenance and rehabilitation of the existing systems, obtaining the necessary electromechanical equipment for networks, dealing with educational and R&D works for the construction and operating the systems, cooperating with other companies to conduct these works, making agreements with all companies that are connected to the grid, preparing transmission and system utilization tariffs, etc. (TEIAS, 2007a).

Distribution: "*Distribution is the transport of electricity through 36 kV or lower lines. Distribution Companies are any legal entities engaged in electricity distribution in a certain geographical region"* (Electricity Market Law, 2001). These companies are licensed to distribute electricity in their region and they should also act as a retailer where consumers cannot purchase electricity from another supplier. The distribution companies are responsible for providing electricity distribution and connection services to all system users, replacing or expanding the capacity of the facilities if necessary and preparing electricity demand forecasts in their region. In the framework of liberalization, distribution system, %90 of which is driven by TEDAS, was decided to be privatized and Turkey was separated into 21 distribution regions in 2006. Government declared that privatization of the distribution companies is in the pipeline. 8 distribution companies before the new structure were: TEDAS, Trakya EDAS, Boğaziçi EDAS, Körfez EDAS, Meram EDAS, Sakarya EDAS, Başkent EDAS and Kayseri and it's around. Figure 2.6 shows the new distribution scheme of electricity in Turkey (TEDAS, 2007).

1. Dicle Elektrik Dağıtım A.Ş.
2. Vangölü Elektrik Dağıtım A.Ş.
3. Aras Elektrik Dağıtım A.Ş.
4. Çoruh Elektrik Dağıtım A.Ş.
5. Fırat Elektrik Dağıtım A.Ş.
6. Çamlıbel Elektrik Dağıtım A.Ş.
7. Toroslar Elektrik Dağıtım A.Ş.
8. Meram Elektrik Dağıtım A.Ş.
9. Başkent Elektrik Dağıtım A.Ş.
10. Akdeniz Elektrik Dağıtım A.Ş.
11. Gediz Elektrik Dağıtım A.Ş.
12. Uludağ Elektrik Dağıtım A.Ş.
13. Trakya Elektrik Dağıtım A.Ş.
14. İstanbul Elektrik Dağıtım A.Ş.
15. Sakarya Elektrik Dağıtım A.Ş.
16. Osmangazi Elektrik Dağıtım A.Ş.
17. Boğaziçi Elektrik Dağıtım A.Ş.
18. Kayseri ve Civarı Elektrik Dağıtım A.Ş.
19. Menderes Elektrik Dağıtım A.Ş.
20. Göksu Elektrik Dağıtım A.Ş.
21. Yeşilırmak Elektrik Dağıtım A.Ş.

Figure 2.6. – Distribution Regions (TEDAS, 2007)

<u>Trading:</u> Market Players for electricity in Turkey include generators, traders and consumers. Generator companies that can sell the electricity they produce are:

Private Generator Companies: They can sell the electricity they produced

- to wholesaler companies by bilateral agreements which are the commercial agreements between real persons and legal entities for the purchase and/or sale of electricity under the provisions of civil law without requiring Board approval. Wholesale companies are any legal entity engaged in the wholesale, import, export, trade of electricity energy and/or capacity and the sale of the same to the eligible consumers. Eligible consumers has the right to choose their supplier
- to retailer companies that are any legal entities engaged in import of electricity and/or capacity and retail sale to consumers, excluding those directly connected to the transmission system and in providing retail sale services to consumers with bilateral agreements. Distribution companies are retail companies and have retail sales licenses.

- to Electricity Market Settlement Mechanism. Electricity that is generated by private sector companies can be sold to the market balance and settlement mechanism. This free market mechanism is first introduced in November 3rd 2004 with the Electricity Market Balance and Settlement Regulation (2004). After 1.5 years trial period, in August 2006, the mechanism described in this Regulation started to be applied in the electricity market. The Regulation aimed real time balancing of electricity that is supplying enough and continuous electricity with low cost and at high quality. In this system, electricity producers that have the properties indicated in the Regulation, give hourly bids for the electricity they will produce during the next month. They also inform the Market Settlement Center about the amount of electricity they will produce, its technical properties etc. These bids are sorted from the lowest to the highest price and price is set at the point where the supply meets the demands and all the generators sell the electricity at that price. The Settlement Center works under TEIAS.

- to eligible consumers, who have the right to choose their supplier.

EUAS: It sells the electricity to TEDAS and TETAS with bilateral agreements. If it produces more than the amount sold by bilateral agreements, it can sell it to the Market Settlement Center.

Autoproducers: The biggest part of the electricity produced by autoproducers is used by themselves or by their partners. However, they can sell the excess amount under free market conditions like private generation companies.

BOT – BO – TOR: They sell the electricity to TETAS by bilateral agreements.

TETAS, Wholesaler and Retailer Companies: They act as the trading companies by buying and selling electricity. TETAS can sell the electricity to TEDAS, to the customers directly connected to the grid or it is interested in export activities. Wholesalers are private companies who have the right and license to buy and sell electricity. Retailer companies are the same as the distribution companies and they have the obligation to sell electricity to eligible and non-eligible consumers such as households and industry. Market Settlement Center does not make any profit from buying and selling activities and it is a pool system.

Eligible and Non-eligible Consumers: *"Non-eligible consumers are any real person or legal entity that can purchase electricity energy and/or capacity only from retail sale companies or from a distribution company holding a retail sale license in its region*" (Electricity Market Law, 2001). Households are an example of non-eligible consumers. On the other hand, eligible consumers have the right to choose their supplier. There are some conditions to be an eligible customer such as being directly connected to the grid or exceeding the eligible consumer limit in previous year or in the current year. Eligible consumer limit is published each year in January. In 2007, this limit is declared as 3 million kWh, which means if the industrial facility uses more than this amount, it has the right to choose its electricity supplier. This limit is further reduced to 1.2 million kWh in January 2008.

Import and Export: TETAS is in charge of import and export activities in Turkey. Net import, increased from 175.5 GWh to 549 GWh between the years 1990 and 2004. On the other hand, export increased from 906.8 GWh to 1798.1 GWh which can be seen in Tables 2.6 and 2.7 (TEIAS, 2006).

Table 2.6 - Distribution of Imported Electrical Energy (GWh) (TEIAS, 2006)

COUNTRY	1990	1995	2000	2001	2002	2003	2004	2005
Bulgaria	0,2		3054,4	3475,3	3107,7	927,9		
Former-USSR	53,5							
Georgia	121,8		177	463,2	92,7			101,1
Azerbaijan								
Iran			244,1	236,7	50,1			
Turkmenistan			26,8	23,3		23,5	389,4	448
TOTAL	175,5	0	3502,3	4198,5	3250,5	951,4	389,4	549,1

Table 2.7 - Distribution of Exported Electrical Energy (GWh) (TEIAS, 2006)

COUNTRY	1990	1995	2000	2001	2002	2003	2004	2005
Bulgaria	506,1	0	0	0	0	0	0	0
Albania	83,9	0	0	0	0	0	0	0
Romania	195	0	0	0	0	0	0	0
Georgia	121,8	178,3	0	0	0	0	0	9,3
Azerbaijan	0	494,7	437,3	432,8	435,1	401,6	378,7	384,1
Iraq	0	22,9	0	0	0	186	765,6	1404,7
TOTAL	906,8	695,9	437,3	432,8	435,1	587,6	1144,3	1798,1

Another issue on the agenda of the Government is Turkey's membership to Union for the Coordination of Transmission of Electricity (UCTE). *UCTE is the association of transmission system operators in continental Europe, providing a reliable market base by efficient and secure electric power highways* (UCTE, 2007). When this membership is approved, Turkish Grid (Transmission System) will be connected and be part of European Grid and thus Turkish Private Energy Producers will be able to export their green energy to European markets as written in the Kyoto Protocol and United Nations Climate Change Charter (UNCCC) because European countries have commitments and they need renewable energy for CO_2 emission reduction. Turkey, being rich with its renewable resources and waiting for new investment for these reserves, will be a good location for foreign investors to generate and trade electricity to their country.

To give more detailed information, in 1992, United Nations Framework Convention on Climate Change (UNFCCC) was signed in Rio de Janeiro, to prevent the climate change resulting from human being and to minimize its effects. All OECD countries except Turkey signed the agreement because Turkey seemed as a developed country and should meet some technical and financial liabilities, including CO_2 emissions according to the protocol. Therefore, Turkey asked for being in developing country status. The agreement aims to decrease the emission level in the year 2000 to 1990s level. In December 1997, a new protocol was prepared in Kyoto, Japan. In this Kyoto Protocol, target year was renewed as 2008-2010 instead of 2000. An important aspect of this protocol was a new emission market, where all of the member countries will have emission quotas and countries can trade emission. The protocol mentions cautions that can be applied such as using hydraulic resources, renewable resources and nuclear energy (Yiğitgüden, 1999).

2.4.3 Electricity Prices

Electricity prices in Turkey are settled according to bilateral agreements, retail tariffs declared by the Government and market players in the spot market. In addition to bilateral agreements, with the enactment of Electricity Market Balancing and Settlement Regulation mentioned previously, market players took their place in the spot market. Today, spot market provides the possibility of placing purchase and sales bids for single hours and block bids. The equilibrium price that is determined on this system is the market price which is set by bilateral auction of suppliers as well as consumers (ERE Hydroelectricity, 2006). According to the Commission Ruling of EPDK (1428/38), the average wholesale electricity price of Turkey in 2007 is 9.67 Ykr/kWh. Spot prices are declared for day, peak and night each month. Day hours are between 06:00 – 17:00, peak hours are between 17:00-22:00 and electricity consumption is the largest during these hours. Night hours are between 22:00 - 06:00. For instance spot market prices in November 2007 were 14.48 Ykr/kWh for day, 15.79 Ykr/kWh for peak, 10.70 Ykr/kWh for night hours (TEIAS, 2007c). The values are high since in winter electricity consumption increases due to heating purposes and the demand exceeds the supply. So the bidders give higher prices. Spot market prices are around 12-13 Ykr/kWh on the average.

On the other hand distribution companies sell the electricity to non-eligible end users at the prices given below:

Household electricity prices ~ 13 Ykr/kWh

Industrial electricity prices ~ 11Ykr/kWh

According to the Commission Ruling published by EPDK (1425), wholesale price of TETAS to distribution companies is 9.53 Ykr/kWh. In Figures 2.7 and 2.8 household and industrial electricity prices in EU countries are given.

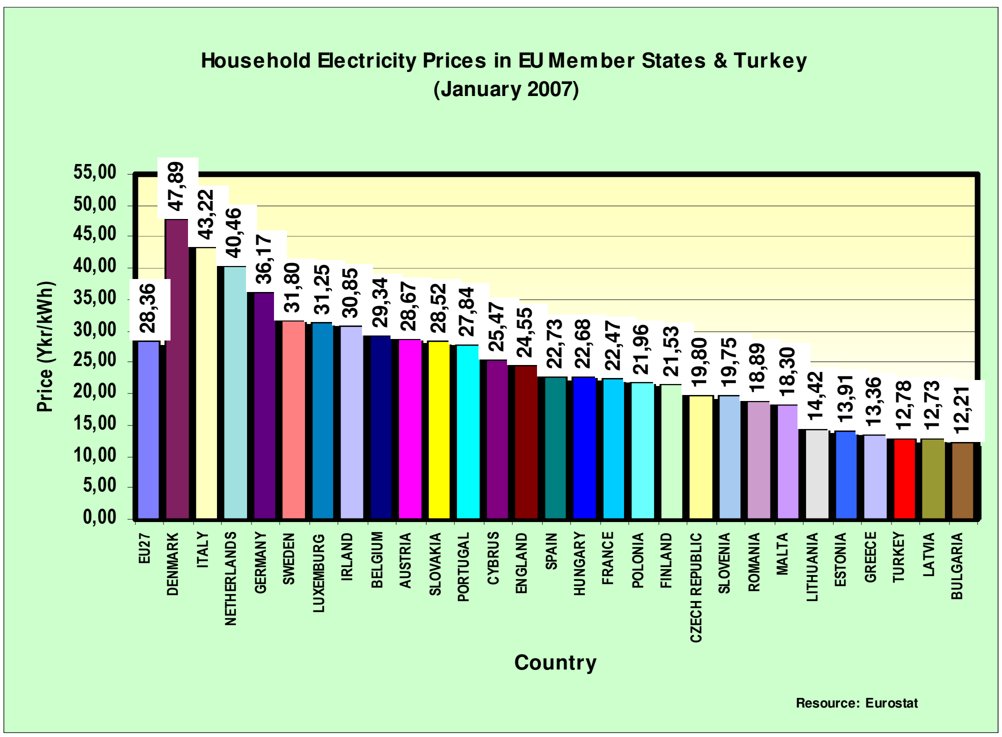

Figure 2.7 – Household Electricity Pricesin EU Member States & Turkey (Eurostat, 2008)

When electricity prices in EU countries and in Turkey are compared, Turkey's electricity prices are lower. However, when the ratios between household and industrial prices are calculated, it seems that industrial users in Turkey are buying the electricity at higher prices.

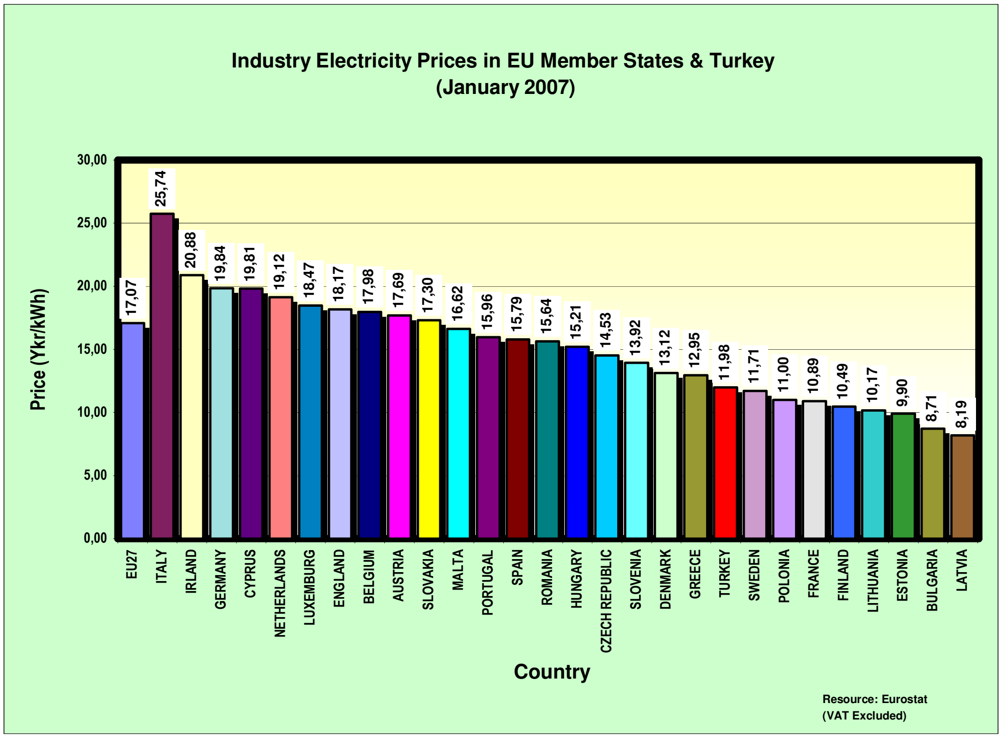

Figure 2.8– Industry Electricity Prices in EU Member States & Turkey (Eurostat, 2008)

2.4.4 Privatization and Reform in Electricity Sector

Liberalization and reform movements gained more importance with the enactment of Electricity Market Law in 2001, and then in 17 March 2004, Electricity Sector Reform and Privatization Strategy Document (High Planning Council, 2004) is published aiming to decrease cost of electricity, maintain supply security, prevent theft and losses and decrease the level of losses to OECD level, make the public sector make investments and reflect these benefits to the consumers (Privatization Strategy Document, 2004). The main reason of reform was rapid increase in electricity consumption and supply and demand balance. Secondly, Government will not be able to make new investments in the future to meet the demand and the vertical integrated structure was inefficient. Additionally, this was a necessary step for the EU membership (Erdoğdu, 2006a). In the first step, distribution side would be privatized and as described before, distribution is separated to 21 regions. The second step was to privatize Istanbul Anatolia Side, Baskent EDAS and Sakarya EDAS regions. However, bidding for these regions is delayed two times. The Government stated that privatization will cause increase in electricity prices. Secondly, Government indicated that electricity infrastructure was a significant issue and should be built by them. But this delay caused lack of confidence and doubts in the market. Moreover, supply and demand balance is one of the most important issues discussed nowadays. Delay in distribution privatization will prevent investments in manufacturing side because first of all, the demand side should be modified and then new investments can be done in the manufacturing side. Today, privatization of both the supply and distribution side at the same time is being discussed by the Government. Some of the EUAS's power plants are waiting to be privatized and for this reason portfolio generation companies are founded.

2.4.5 Theft and Losses in Turkey

One of the targets of reform process is decreasing the theft and losses in electricity market. This percentage is high in Turkey compared to the other OECD Countries. The theft and losses ratio was 21.58 % in 2000, 21.43% in 2001, 20.86 % in 2002, 19.93 % in 2003, 18.58% in 2004 and 17.80 % in 2005. Theft and losses are mostly in the south-eastern part of Turkey. The biggest losses and theft were in Mardin and Şırnak with %71 in 2004. On the other side, Bursa and Birecik has the smallest losses in Turkey with 5% (Bakır, 2007).

High theft and loss ratio in electricity causes variation in prices in different regions. These losses are hidden by cross subsidies and people in cities with low theft-losses pay some part of the bills of people who live in cities with high theft-losses (Gümüşdere, 2004). To overcome this problem, uniform retail tariff system is applied by TEDAS. Table 2.8 consists of information about losses and theft in percentages in different distribution regions.

Table 2.8 – Losses in Different Distribution Regions (Atiyas and Nuez, 2007)

Distribution Company	(%)	Distribution Company	(%)
Vangölü	63,8	Yeşilırmak	9,5
Dicle	57,8	Trakya	9,3
Aras	29,4	Göksu	9,3
Çoruh	12,3	Akdeniz	8,9
Boğaziçi	12,3	Uludağ	8,8
Fırat	11,7	Çamlıbel	8,5
Toroslar	10,9	Meram	7,8
İstanbul	10,2	Osmangazi	7,2
Sakarya	10,1	Menderes	7,1
Başkent	9,6	Gediz	6,5

2.4.6 Electricity Demand in Turkey

In 17 October 1973, Government started programmed power outages. According to Arabul and Selçuk (2000) the reasons of this bottleneck were in the implementation of plans and programs, maintenance and restoring activities, timing of the investments, problems in import of fuels and insufficient demand forecasts and problems in the supply and demand balance. Today there is a similar bottleneck in electricity supply resulting nearly from similar reasons. Turkey's electricity demand in 2006 was 168 billion kWh and increasing nearly 6-9% per annum according to TEIAS data. In Table 2.9 electricity demand and percentage increase for the past ten years is given. Electricity demand is affected with economical crisis in 2001 so there is a decline in this period (TEIAS, 2007b).

According to the electricity demand projection published by TEIAS on 20th July 2007, gross electricity consumption in Turkey was 160.8 GWh in 2005 and 174.2 GWh in 2006 with an increase of 7.2 % and 8.3 % respectively with respect to previous year and is still showing an increase.

Table 2.9 – Peak Power and Energy Demand of Turkey (TEIAS, 2007b)

	Peak Power Demand (Mw)	Increase (%)	Energy Demand (Gwh)	Increase (%)
1996	15231	7.5	94789	10.8
1997	16926	11.1	105517	11.3
1998	17799	5.2	114023	8.1
1999	18938	6.4	118485	3.9
2000	19390	2.4	128276	8.3
2001	19612	1.1	126871	-1.1
2002	21006	7.1	132553	4.5
2003	21729	3.4	141151	6.5
2004	23485	8.1	150018	6.3
2005	25174	7.2	160806	7.2

Due to the forecasts with low and high scenarios, electricity consumption will increase 6% and 8% per year between 2007 and 2016 (TEIAS, 2007b). These kinds of studies prove that electricity supply and demand balance and supply security became a serious problem today, because currently installed power plants will not be able to meet the whole demand in 3 or 4 years time. Therefore, good demand forecasts are required to schedule and plan the electricity production and investments and to follow the exact route. There is also a requirement to encourage the investors to enter the market and increase electricity generation to balance the supply and demand. Figure 2.9

shows electricity supply and demand projections for the next years according to the Production Capacity Projection.

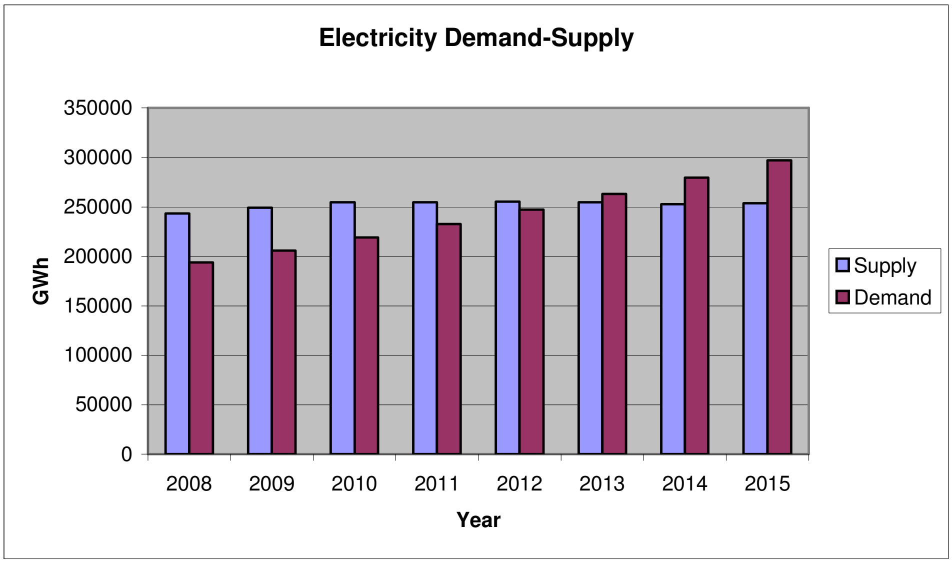

Figure 2.9 – Demand and Supply Projections (TEIAS, 2007b)

CHAPTER 3

LITERATURE REVIEW AND BACKGROUND

Electricity demand forecasting is an important issue today, where most of the countries suffer from electricity supply security due to the climate change, technological growth and scarce energy resources. Electricity prices are also affected from unbalanced demand and supply and they may go up, as there will be lack of resource shortly. Considering this, underestimating electricity demand can cause undercapacity and poor quality of service including black outs. On the other hand, overestimating the electricity demand can cause authorization of a plant that may not be needed and additional unnecessary investments in the long term. That is, supply and demand balance is also important while setting tariffs and for demand side management.

Electricity demand is affected by various factors like temperature, income, GDP, population and price. Forecasts can be done for short, medium and long terms. In general economic indicators do not effect electricity consumption as much as temperature in the short term. Additionally, day of the week, football matches, sunlight, holidays, religious days characterize electricity consumption. Short term forecasts are from a few hours to few weeks forecasts. They are useful in controlling and scheduling power plants to generate electricity. These short term demand forecasts help to decide which equipments should be operated to meet the demand at that period and how to minimize cost with supply security.

Medium term forecasts should consider both socio-economic and temperature variables. Medium term forecasting can be used for fuel procurement, scheduling unit maintenance and diversity interchanges. Medium term forecasts are from a few weeks to few months forecasts. Mostly used type of forecasts is the long term forecasts which are useful for generation and transmission planning, tariff planning, feasibility studies, and expansion of utilities (Syed and Saleh, 1997). Long term forecasts are from 5 to 25 years.

There are several methods to estimate demand. In general, not a single method is used to make forecasts; instead the methods are combined to get better results. In section 3.1 forecasting methods are described in brief. In section 3.2, cointegration and ECM, being the methodology used in this study is discussed. Sections 3.3 and 3.4 cover previous studies about electricity demand.

3.1 METHODS FOR ESTIMATING ELECTRICITY DEMAND

In the first part of this section, a brief explanation is given for the time series and then methods for estimating electricity demand are explained.

3.1.1 Basic Definitions

Times series, *is a sequence of data points that are measured typically at successive times and usually spaced at uniform time intervals* (Wikipedia, 2007a). Time series are mostly used in statistical analysis and econometrics and they are useful to make policy decision. Time series can be separated into two according to their recording period; first being continuous time series, which can be recorded continuously such as electrical signals, voltage and vibration and second type being as discontinuous time series, which can be recorded periodically such as interest rate and sales volume (Kadılar, 2005, p.1).

Time series analysis can be performed in short, medium and long time horizons. Developing a time series using the data points is a **stochastic process**, that is defined as an ordered sequence of random variables $\{x(s,t), s\in S, t\in T\}$, such that for each $t\in T$, $x(\bullet,t)$ is a random variable on the sample space S and for each $s\in S$, $x(s,\bullet)$ is a realization of stochastic process on index set T. Generally, a stochastic process or realization of the stochastic process is shown with the notation $\{x(t), t\in T\}$. (Banerjee et al., 1993, p.10-11).

Time series are separated into two as stationary and nonstationary series according to the deviation from the mean. Stationarity can also be divided into two as weak and strict stationarity. A stochastic process is **strictly stationary** if all existing moments of the process are constant in time, that is the joint and conditional probability distributions of the process' do not change in time. Strictly stationary process is indicated as:

Let any subset of time be $(t_1,t_2\ldots t_n)$ of T,

h is a real number where $t_i+h \in T$, $i=1,2,\ldots.n$, then

$F(x(t_1), x(t_2),\ldots\ldots\ldots x(t_n)) = F(x(t_1+h), x(t_2+h)\ldots.. x(t_n+h))$

Where $F(\bullet)$ is the joint distribution function of n values.

The process $\{X_t\}$ is **weakly (covariance / second order) stationary** when:

$E[x(t_i)] = E[x(t_i+h)] = \mu = \text{constant}$

$E[(x(t_i))^2] = E[(x(t_i+h))^2] = \mu_2 = \text{constant}$

$E[x(t_i)\, x(t_j)] = E[x(t_i+h)\, x(t_j+h)] = \mu_{ij} = \text{constant}.$

Thus, means and variance of series are constant over time and the covariance between two periods depends only the gap between the periods, and not the actual time at which covariance is considered (Charemza and Deadman, 1997, p.84-98; Kadılar,2005, p.20-21).

If these conditions are not satisfied, the series are nonstationary and when a series is nonstationary, spurious and a meaningless pattern may occur. For instance, output of the regression shows that R-squared and t statistics are high enough indicating the variables are significant and the data fits the model well. On the other hand, Durbin Watson (DW) statistics, which is used to test for the existence of autocorrelation in the series, can be low. So, to get accurate results, nonstationarity should be eliminated. Usually, weak stationary is investigated in time series because it is difficult to obtain strict stationary series where the distribution of the series is unchanged in time.

Nonstationarity may be because of the tendency of a series to move in one direction, which is defined as trend. When nonstationarity of the series is as a result of stochastic effect or random shock, trend is stochastic. When the mean of the process is a specific function of time which may be expressed as:

$y_t = \mu_t + \varepsilon_t$

$\mu_t = \alpha + \beta.t$

the process has a deterministic trend.

By moving the time series data by one or more periods, lagged time series are obtained. One period lagged time series of X_t is X_{t-1}. Lagged series have the same structure with the original series that is if there is a seasonal variation in time series, there still exist a seasonal variation in the lagged series, in condition that the lag period is not too long.

B is the lag operator where $By_t = y_{t-1}$

Some of the properties of the lag operators are (Kutlar, 2000,p. 8-9):

- Lag of a constant is the constant itself

  $Bc=c$

- Distributive property

  $(B^i+B^j)y_t=B^iy_t+B^jy_t=y_{t-i}+y_{t-j}$

- Combination property

$B^iB^jy_t=B^i(B^jy_t)=y_{t-i-j}$

- Lag operators can take negative values

  $B^{-i}\ y_t = y_{t+i}$ or if j=-1, $B^j\ y_t = y_{t-j} = y_{t+j}$

If a process can be made stationary by differencing, it is called an **integrated process.** A process differenced d times to achieve stationarity is integrated of order d and is shown as $\Delta^d x_t$, where Δ^d is the differencing operator, $(1-B)^d$.
If there is still trend or seasonal variation in the series, a second differencing can be applied which is shown as:

$\Delta^2 x_t = (1-B)^2\ x_t = \Delta\ x_t - \Delta\ x_{t-1}$

In most cases second differencing is done when the series are exponential.

If the data has seasonal fluctuations, it may be seasonally integrated. Birchenhall et al (1988) defined seasonal integration as "*A non-deterministic series X is said to be integrated of order (d,D), denoted $X_t \sim I(d,D)$, if the series has a stationary, invertible ARMA representation after one-period differencing d times and seasonally differencing D times.*"

Random walk process is defined as $x_t=x_{t-1} + \varepsilon_t$ where ε_t are identical and independent random variables. Random walk is a stochastic and nonstationary process. Another nonstationary process is random walk with drift where μ is different from zero.

$x_t=x_{t-1} + \varepsilon_{t} + \mu$

The process without drift can be also shown as $(1-B)x_t= \varepsilon_t$ which is a process integrated of order 1.

On the other hand, **white noise process** is a stationary process, which has 0 mean and uncorrelated over time. The process can be shown as:

$\{x(t), t\in T\}$, where $E\ [x(t)]=0$, $E\ [(x(t))^2]=\sigma^2 < \infty$ and

$E[x(t), x(t+h)]=0$ where $h \neq 0$.

An important tool for time series data is **autocorrelation function (ACF)** that gives the relationship between the time series and its lagged values. The values of ACF are calculated as:

$$r_k = \frac{\sum_{t=1}^{T-k} (x_t - \bar{x})(x_{t+k} - \bar{x})}{\sum_{t=1}^{T} (x_t - \bar{x})^2}$$

where

x_t : original time series

x_{t+k} : k period lagged time series

$\bar{x}$: average of the time series

r_k : autocorrelation of k^{th} lag

ACF takes values between 1 and -1.

ACF versus lag number gives the ACF graph or correlogram (Kadılar, 2005, p.13).

On the other hand, **partial autocorrelation (PACF)** gives the relationship between two lags of the series X_t and X_{t-k}, when the intermediate values X_{t-1}, $X_{t-2} \ldots X_{t-k+1}$ are constant. All of the partial autocorrelation coefficients form PACF (Kadılar, 2005, p.16).

$$r_{kk} = \frac{r_k - \sum_{j=1}^{k-1} (r_{k-1,j})(r_{k-j})}{1 - \sum_{j=1}^{k-1} (r_{k-1,j})(r_j)}$$

where r_k is the autocorrelation coefficient and

$r_{kj} = r_{k-1,j} - (r_{kk})(r_{k-1,k-j})$

PACF versus lag number gives the PACF graph and PACF takes values between -1 and 1.

The equations will be estimated using **Ordinary Least Squares (OLS)** in this study since it is one of the mostly used estimators in statistics and OLS is suitable for the linear models (Akdi, 2003, p.116). Additionally, according to the Gauss–Markov theorem*, "in a linear model where the errors have zero expectation, are uncorrelated and have equal variances, the best linear unbiased estimators of the coefficients are the least squares estimators*" (Wikipedia, 2007b).

Assume a regression equation:

$Y_t = \beta_0 + \beta_1 X_t + e_t$, t=1,2….n and e_t white noise

OLS estimators are obtained by minimizing the $\sum \varepsilon_t^2$ that yields:

$$\hat{\beta}_1 = \frac{\sum_{t=1}^{n}(Y_t - \overline{Y})(X_t - \overline{X})}{\sum_{t=1}^{n}(X_t - \overline{X})^2}, \ \hat{\beta}_0 = \overline{Y} - \hat{\beta}_1 \overline{X}$$

where X_t and Y_t are the time series

.

Maximum Likelihood Estimation (MLE) is another technique used in statistics and econometrics, to obtain the estimators. It chooses the parameters considering the maximization of the likelihood of the sample.

In some cases more than one model can be formed for a data set. There are some model selection criterions to decide which model is better among various models. **Akaike Information Criteron (AIC)** and **Schwarz (Bayesian) Information Criterion (SIC)** are the ones that are frequently used. AIC is based on the one step forecast mean square error and is expressed as:

$$AIC = \ln|\Sigma| + (2N/T)$$

where $|\Sigma|$ is the determinant of residual variance-covariance matrix, N is the total number of parameters estimated, T is the number of observations

SIC can be shown as:

$$SIC = \ln|\Sigma| + \{(N \ln T)/T\}$$

where $|\Sigma|$, N and T are the same as AIC. The lowest the value of AIC and SIC, the better the model.

The forecasts from the models are evaluated and compared using some statistical indicators which are Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE) and Theil Inequality Coefficient (TIC). The calculation of these statistics is shown below, where the forecast sample is j = T+1, T+2….T+h, $\hat{y}_t$ is the forecasted value, *y* is the actual value and t is the period.

$$RMSE = \sqrt{\sum_{t=T+1}^{T+h} (\hat{y}_t - y_t)^2 / h}$$

$$MAE = \sum_{t=T+1}^{T+h} |\hat{y}_t - y_t| / h$$

RMSE and MAE are used to compare forecast for different models with the same series. It is better to have small errors

$$MAPE = \sum_{t=T+1}^{T+h} \left| \frac{\hat{y}_t - y_t}{y_t} \right| / h$$

$$TIC = \frac{\sqrt{\sum_{t=T+1}^{T+h} (\hat{y}_t - y_t)^2 / h}}{\sqrt{\sum_{t=T+1}^{T+h} (\hat{y}_t)^2 / h} + \sqrt{\sum_{t=T+1}^{T+h} (y_t)^2 / h}}$$

TIC is always between 0 and 1, when it is closer to 0, the forecasts are more accurate. Other measures to show the accuracy of the forecast are bias, variance and covariance proportions obtained from the decomposition of MSE as

$$\sum_{t=T+1}^{T+h}(\hat{y}_t - y_t)^2 / h = ((\sum \hat{y}_t / h) - \overline{y})^2 + (s_{\hat{y}} - s_y)^2 + 2(1-r)s_{\hat{y}}s_y$$

where $\sum \hat{y}_t / h, \overline{y}, s_{\hat{y}}, s_y$ are the means and standard deviations of $\hat{y}_t$ and y, r is the correlation between $\hat{y}$ and y.

Bias Proportion: $\dfrac{((\sum \hat{y}_t / h) - \overline{y})^2}{\sum (\hat{y}_t - y_t)^2 / h}$

Bias proportion shows the difference between the actual and forecasted mean.

Variance Proportion: $\dfrac{(s_{\hat{y}} - s_y)^2}{\sum (\hat{y}_t - y_t)^2 / h}$

Variance proportion shows the difference between forecast variation and actual variation.

Covariance Proportion: $\dfrac{2(1-r)s_{\hat{y}}s_y}{\sum (\hat{y}_t - y_t)^2 / h}$

Covariance proportion shows the remaining unsystematic forecast error. These three sum up to 1. Small values of bias and variance proportion are preferred for good forecasts.

3.1.2 Methods

There are various methods used by the researchers to estimate electricity demand. The most well known methods can be sequenced as average percentage increase method, trend models, disaggregating approach, econometric methods, time series methods, artificial neural networks and integrated models, which are combination of different methods and many submodules.

Average percentage increase method is one of the oldest and easiest methods among them. Average percentage increase in actual consumption for previous years is found and forecast for future periods are calculated according to this percentage increase. However, this method assumes that the electricity consumption increases linearly. Moreover, it does not consider other economic variables that have influence on electricity consumption.

Trend models are also one of the firstly used methods to estimate electricity demand. Consumption is estimated just as a function of time and no other economic variables (income, price etc.) exist in the model. The past values are projected to estimate the future values. This method is simple and easy to use so it is preferred as groundwork. Sometimes these kinds of models are called exponential growth models (Rhys, 1984). The simple model can be written as:

$E_t=E_0*(1+e)^t$

where

E_t: electricity demand in year t

E_0: electricity demand in base year

e: average demand increase in previous years

t: time

Disaggregating (bottom up) approach is used for over 30 years and each year with extension of data its usage increased. Large and detailed number of judgments is required for this approach. The name **end-use method** is also used for this approach as it takes into account the various uses of electricity in different sectors. To illustrate, in residential sector electricity is used for lighting, heating, refrigerators and cooking. At the end, the separate demands for each home appliance are added to get total electricity demand. This model is appropriate where new technologies are introduced but it requires collecting a large number of data and disaggregating the demand carefully. To summarize, end use method is a bottom up approach which is mostly appropriate when detailed data for consumption exist. The most important limitation of this method is data availability about household appliances. Additionally, usage of household appliances takes place when people are at home, so all these limitations should be taken into account during estimation. Household electricity consumption can be written as:

$$D_t = \sum_{i=1}^{n} K_{it} W_{it}$$

D_t: electricity demand of households in year t

K_{it}: number of household appliances of type i, in year t

W_{it}: electricity consumption of household appliances of type i, in year t (Çakır, 2002)

In general the level of income determines the number of household appliances, however, this number is generally found by surveys . Parti and Parti (1980) state the biggest disadvantage of this method as:

"*…..they are based upon theoretical considerations, rather than observed consumer behavior and cannot be adjusted in any systematic way for regional differences or changes in price, income, or household size as can the current econometric estimates. The primary disadvantage of the use of direct appliance metering is its great cost.*"

Econometric methods are started to be used during the early 1960's as an extension of statistical methods (Meetamehra, 2002). Relation between demand for electricity and some economical indicators such as population, income, price of power and alternative fuels are investigated with econometric methods. Therefore, the method combines statistical methods and economic theory. This method requires identifying the explanatory variables correctly and enough length of the data period. GDP and consumption approach, Error Correction Models (ECM), Autoregressive Distributed Lag Modeling (ADL), Vector Autoregression (VAR) and Vector Error Correction Models (VECM) are various types of econometric models.

GDP and consumption approach can be considered as an econometric model**,** which is a simple method that considers the relation between GDP and electricity consumption. The estimated ratio (i.e. K_e) which is calculated as given below is used for future forecasts. The formulation can be written as:

$K_e=E_e/G_e$

$D_t=D_0*(1+(K_e*G_t))^t$

Where

G_e : average GDP rate in the base period

E_e : average consumption rate in the base period

K_e : ratio of GDP rate to consumption rate in the base period

D_t : electricity consumption in year t.

D_0 : electricity consumption in the base year

G_t : expected GDP rate in year t (Çakır,2002)

The disadvantages of this method is the direct relation between GDP and consumption, because a rapid increase in GDP will cause a rapid increase in consumption and this method disregards the technological changes, habits, demographical changes in time and only investigates the relationship between electricity consumption and GDP (Keleş, 2005).

Second type of econometric approach to estimate electricity demand is ECM method. The first step to model electricity demand by **ECM** is searching whether the time series are stationary or not. If they are nonstationary and integrated at the same order, cointegration relation among the variables is investigated. **Cointegration** is one of the latest and most popular methods that is started to be used in 1980's as an important tool in econometric modeling. Long run relationship between demand and the economic factors such as consumption, GDP and price can be modeled with cointegration and once the cointegrating relationship between the variables is approved, ECM can be used as the appropriate model for the variables. One of the disadvantages of this technique is that unit root and cointegration tests for the series might give misleading results. When the explanatory variables used in the model have mutual relation with the dependent variable such as consumption and price may both affect each other. In such cases another model called VAR model is more appropriate. Third, technological developments are not considered in the model and cointegration approach assumes constancy in the long run. Cointegration and ECM are explained in more detail in section 3.2.

Davidson et al (1978) modeled consumption as a general and large model in the beginning. Then, they reduced its size testing linear and nonlinear restrictions. This general modeling technique is named as **ADL** modeling, where a dependent variable is regressed on its own and explanatory variables' lagged values in addition to the explanatory variables' current values. The model can be expressed as:

$$y_t = a_1y_{t-1} + a_2y_{t-2} + a_3y_{t-3} + \ldots . aky_{t-k} + \beta_0x_t + \beta_1x_{t-1} + \ldots . B_nx_{t-n} + \varepsilon_t$$

or

$$a(B)\, y_t = b(B)\, x_t + \varepsilon_t$$

where y_t is the dependent variable and x_t is the explanatory variable and B is the lag operator.

Another econometric model is **VAR,** which is different from others for containing multiple equations. Some of the variables can be explained by other variables whereas some variables can act only as explanatory variables. The variables that are explained by others are called endogenous variables in a system of equations, whereas the explanatory variables are called exogenous. This relationship between variables can be shown with more than one equation. For instance, income and consumption may have mutual relationship and affect each other. One of the latest developments in this area was in 1980's by Sims (Charemza and Deadman, 1997, p.167). In the general unrestricted VAR model, each variable is regressed on lagged values of itself and other endogenous variables. Lag length that will be used in the equations is important and it is chosen considering the autocorrelation restriction. An example VAR model containing two variables and two lags is indicated below:

$$\begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} a_1 & b_1 \\ c_1 & d_1 \end{bmatrix} \begin{bmatrix} x_{t-1} \\ y_{t-1} \end{bmatrix} + \begin{bmatrix} a_2 & b_2 \\ c_2 & d_2 \end{bmatrix} \begin{bmatrix} x_{t-2} \\ y_{t-2} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}$$

where $E(\varepsilon_{1t}) = E(\varepsilon_{2t}) = 0$, $E(\varepsilon^2_{1t_{1t}}) = \sigma_{11}$, $E(\varepsilon^2_{2t_{1t}}) = \sigma_{22}$, $E(\varepsilon_{1t_{1t}} . \varepsilon_{2t}) = \sigma_{12}$

and error terms are uncorrelated.

Lag length in VAR models may be large, which increases the number of estimated parameters at the same time. To overcome this difficulty and eliminate some of the coefficients, the hypothesis that the coefficients are jointly equal to zero is tested. In VAR models, Granger causality is used to test zero restrictions. Granger causality can be described as given below:

"x is a Granger cause of y (denoted as $x \to y$), if present y can be predicted with better accuracy by using past values of x rather than by not doing so, other information being identical."

VAR estimates can be modeled as a **VECM** if the cointegrating relationship among the variables is verified. VECM is a restricted VAR which is used with nonstationary and cointegrated series (Quantitative Microsoftware, 2000, p.547). The cointegration term (i.e error correction term) is integrated to the VECM which restricts long run behaviour of the endogenous variable. While working with VAR and VECMs, Johanson cointegration test procedure that is explained in section 3.2 is preferred.

Let a simple cointegration equation be: $x_t = \beta y_t$

The vector error correction model is:

$$\Delta y_t = \alpha_1 (x_{t-1} - \beta y_{t-1}) + \varepsilon_{1,t}$$
$$\Delta x_t = \alpha_1 (x_{t-1} - \beta y_{t-1}) + \varepsilon_{2,t}$$

The right hand side variable is the error correction term. In the long run, this term is zero. α_i is the speed of adjustment.

In addition to econometric methods, **Time Series Methods** are also preferred while performing electricity demand forecasting studies for their simplicity and reliability. This method can be defined as an econometric model where the only explanatory variable used is lagged values of the variable that will be estimated. At least 20 years data is required to get good results. Time series methods do not describe cause-effect relationship. Time series methods can be divided into three as exponential smoothing, Box - Jenkins and decomposition methods.

Time series is composed of trend, seasonality, cycles and irregularity components in **decomposition method**. Trend is the change in mean by time. The series has an increasing or decreasing pattern if it has a trend. Seasonal variation can be defined as the intra year movements. There is a seasonal variation in a series if there is an increase or decrease in that season of each year. The length of this fluctuation is called as period. Cyclical variations are the between year movements. Some series have increasing decreasing pattern in cycles, such as 2 years time periods. Series that do not fit with any of the above definitions and have irregular pattern may have irregularity. Irregularity is the random fluctuation of the series. Irregular series cannot be used for forecasting because no one can imagine what will be in the future. Example equations for decomposition method are given in (1) and (2). (Kutlar, 2000,p.8 Kadılar, 2005,p.60-63)

$Zt= T_t+S_t+I_t+\varepsilon_t \quad t = 1,2,\ldots T \quad (1)$

$Zt= T_t.S_t.I_t.\varepsilon_t \quad t = 1,2,\ldots T \quad (2)$

The first equation is called the additive model whereas the second equation is called the multiplicative model. When the magnitude of the seasonal component is rather constant regardless of changes in trend, additive model is more suitable. On the other hand, if seasonal component varies with changes in trend, a multiplicative model is preferred.

Exponential smoothing is another time series method where weights are assigned to time series data giving biggest weight to the recent data. Smoothing techniques can be separated as; simple, Holt's and Winter's smoothing techniques. In this method, current and preceding observations are given higher weights and older observations are given exponentially smaller weights. The formula of the exponential smoothing is:

$$S_t = \alpha.X_t + (1-\alpha).S_{t-1}$$

where S is the smoothed value that is forecasted, X is the observations and α is the weight. To sum up, *each smoothed value is the weighted average of previous observations and weights decrease exponentially depending on the value of* α (Statsoft, 2006).

Finally, **Box-Jenkins** developed Autoregressive Integrated Moving Average (ARIMA) methodology in 1976 and the method became very popular for its power and flexibility. The steps of the Box-Jenkins methodology in general are:

- Editing the data and making transformation if necessary.

- Deciding whether the series is stationary. The stationary of the series can be searched by investigating the autocorrelation function (ACF) and partial autocorrelation function (PACF) graphs of the series searching whether the lag values are between boundaries. Recently, some stationary tests such as Phillips Peron, Dickey Fuller and Augmented Dickey Fuller tests are used to get more accurate results, which will be explained later in more detail.
- Differencing the series if they are non-stationary
- Deciding the levels of Autoregressive Moving Average (ARMA) by investigating correlograms, correlation and autocorrelation functions.
- Estimating the parameters
- Validation of the model (Kadılar, 2005,p.185-233)

<u>Autoregressive Models, AR (p)</u>: The AR(p) is a stochastic stationary process, where the time series is dependent to the past p values (Kutlar,2000). AR(p) process can be defined by the following equation:

$y_t= m+ \alpha_1 y_{t-1} + \alpha_2 y_{t-2} +\ldots\ldots \alpha_p y_{t-p} + u_t$

where, m is a constant, p is the degree of the function and u_t is the error term which is iid $(0, \sigma^2)$. As explained before, if the series is stationary expected value of the series remains constant independent of the time which is expressed as:

$E(y_t) = E(y_{t-1}) = E(y_{t-2})\ldots\ldots= \mu$ and $\mu= \alpha_1\mu + \alpha_2\mu$ ----- $\alpha_p\mu + m$

So, for the series to be stationary, expected value should be finite and

$1- \alpha_1 - \alpha_2-\ldots\ldots\alpha_p<1$

Mostly used AR(p) models are AR(1) and AR(2). When the model is AR(1), then the equation becomes $y_t= m+ \alpha_1 y_{t-1}+ u_t$

The autocorrelation function, for k=0 becomes 1 and converges to 0.

If $\mid \alpha_1$ 1>I , then the series is stationary.

For AR(1) process, if α=1, the series is nonstationary and unit root exists. In the case of nonstationarity, integration may occur. In Figure 3.1, ACF and

PACF for the AR(1) process is given. It can be seen that PACF becomes 0 after a certain period (one peak value) and ACF exponentially decays to 0 for an AR(1) process.

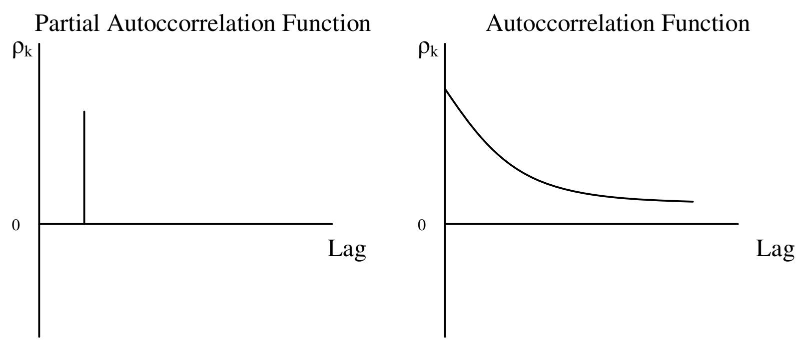

Figure 3.1 – Autocorrelation and Partial Autocorrelation Function of AR(1) Process

AR(p) model can be also written with lag operators as indicated below:

$$(1-\phi_1 B-\phi_2 B^2 \ldots -\phi_p B^p) z_t = \varepsilon_t$$

<u>Moving Average Models, MA(q)</u>: The dependent variable is found by adding the lagged values of the error term in a MA(q) model.

The MA(q) equation can be written as:

$$y_t = \mu + u_t - \theta_1 u_{t-1} - \theta_2 u_{t-2} \ldots\ldots\ldots\ldots - \theta_q u_{t-q}$$

where q is the degree of the function. The parameters of the equation can be positive or negative. The error terms in the equation had normal distribution with covariance equal to 0.

$E(u_t)=0$, $E(u_t^2) = \sigma^2$

When the model is MA(1), then the equation becomes;

$$y_t = \mu + u_t - \theta_1 u_{t-1}$$

If $| \theta_1 1 > |$, then the series is stationary.

For MA(1), ACF becomes 0 after a certain period and partial PACF approaches to 0.

In figure 3.2, a drawing for ACF and PACF is given for a MA(1) model.

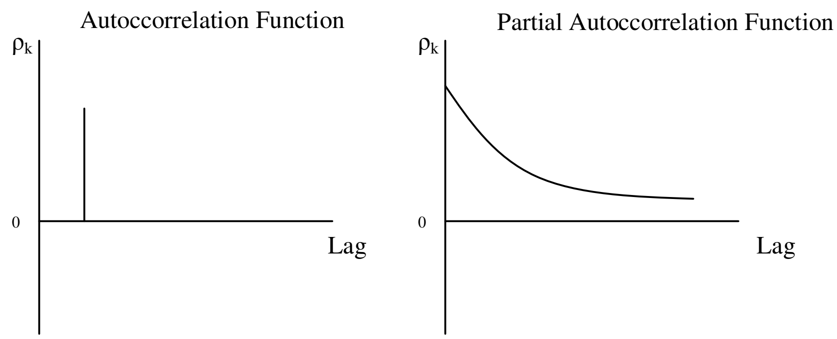

Figure 3.2 – Autocorrelation and Partial Autocorrelation Function of MA(1) Process

The representation of MA(q) model with lag operators is given below.

$z_t = (1-\theta_1 B - \theta_2 B^2 ....\theta_q B^q)\varepsilon_t$

AR and MA equations can be inverted and MA(1) equation can be written as AR(∞) model as stated following

$y_t = u_t (1 - \theta B)$ (MA(1) equation)

$u_t = y_t (1 - \theta B)^{-1} = y_t (1 - \theta B - \theta^2 B^2 ....)$ which is an infinite order AR process. θ should be between [-1,1] for the MA process to be invertible.

ARMA(p,q):

The ARMA(p,q) equation can be written as

$y_t = m + \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + ... \alpha_p y_{t-p} + u_t - \theta_1 u_{t-1} - \theta_2 u_{t-2} .... - \theta_q u_{t-q}$

For series to be stationary, mean, variance and covariance should be independent of time as before.

The model can be diagnosed using autocorrelation and partial autocorrelation functions. If the autocorrelation function has q values outside the confidence interval, the model is said to have degree of q. If the partial autocorrelation function has p values outside the confidence interval and goes to 0, the AR model is said to have degree of p. If both of the functions makes peak and then decrease, model is ARMA (p,q)

Autoregressive Integrated Moving Average Model, ARIMA (p,d,q): Time series is ARIMA(p,d,q) if a series is differenced d times to make it stationary and than ARMA(p,q) model is fitted to this differenced data. After the values for p,d,q and the model parameters are estimated, checking the model adequacy begins. This is done by investigating the autocorrelation and partial autocorrelation of residuals obtained from the estimated model. If they are not statistically significant, the data is truly modelled. Finally, assumptions of normality, independency (i.e. by ACF and PACF) and constant variance are checked.

The representation of ARIMA(p,q) model with lag operators is:

$$(1-\phi_1 B-\phi_2 B^2 \ldots-\phi_p B^p)(1-B)^d z_t=(1-\theta_1 B-\theta_2 B^2 \ldots\ldots..\theta_q B^q)\varepsilon_t$$

Seasonal Autoregressive Integrated Moving Average Model, SARIMA (p,d,q)(P,D,Q): Seasonal Box – Jenkins models are similar to nonseasonal models, however, the procedure for SARIMA is more complex. Both seasonal and periodic differencing may be required for the seasonal models. The general SARIMA(p,d,q)(P,D,Q) is written as

$$(1-\phi_1 B-\phi_2 B^2 \ldots-\phi_p B^p)(1-\Phi_1 B^s-\Phi_2 B^{2s} \ldots..-\Phi_p B^{Ps})(1-B)^d (1-B^s)^D z_t=$$
$$(1-\theta_1 B-\theta_2 B^2 \ldots\ldots..\theta_q B^q)(1-\Theta_1 B^s-\Theta_2 B^{2s} \ldots\ldots.\Theta_Q B^{Qs})\varepsilon_t$$

Another electricity demand forecasting model that is begun to be used recently is **artificial neural networks (ANN)**. The disadvantage of neural networks is that their performance decreases when the series are nonstationary, which is a situation faced in many econometric series. However, they are good when the data set is small and they are also advantageous over statistical methods in that they are externally parameterized and complex relations can be modeled by this approach (Thang, 2004). ANN approach consists of five stages that are formulation, identification, construction, validation and deployment.

Finally, most of the models for energy demand consist of many submodules and these models are called **integrated models**. The most well-known and utilized models in the world are PRIMES and ENPEP.

<u>PRIMES</u>**:** is an energy model about energy supply and demand and it is used in member states of European Union. The model is good in that it allows technological changes to be integrated to the model. It is a general decision and policy making tool in energy sector and also helps to make forecast, analyze policies and develop scenarios. PRIMES was firstly used in 1997. Among its various submodules, the supply module simulates the operation and capacity expansion. Demand module determines the dispatching activities and capacity expansion in a national level; however national works can be integrated to EU level. Cost evaluation modules and price setting mechanisms are related with the supply model and total revenue and pricing is studied here. Demand is analyzed in residential, industry, commercial and transport sectors in all fuel types. Statistical indicators published by Eurostat are prepared with this model (Capros, 2005).

<u>The Energy and Power Evaluation Program (ENPEP):</u> is a set of analytical tools to plan energy and environmental issues, which is formed by Argonne National Laboratories and International Atomic Energy Agency, supported by the World Bank, US Energy Ministry and Hungarian Electricity Authority. The submodules of ENPEP are used in more than 90 countries now and also used in Turkey. ENPEP analyses the entire energy system in an integrated framework and investigates sectoral effects between residential, electricity, oil

and coal sector. Figure 3.3 gives the relationship between the modules of ENPEP (Sakaryalı et al, 2000).

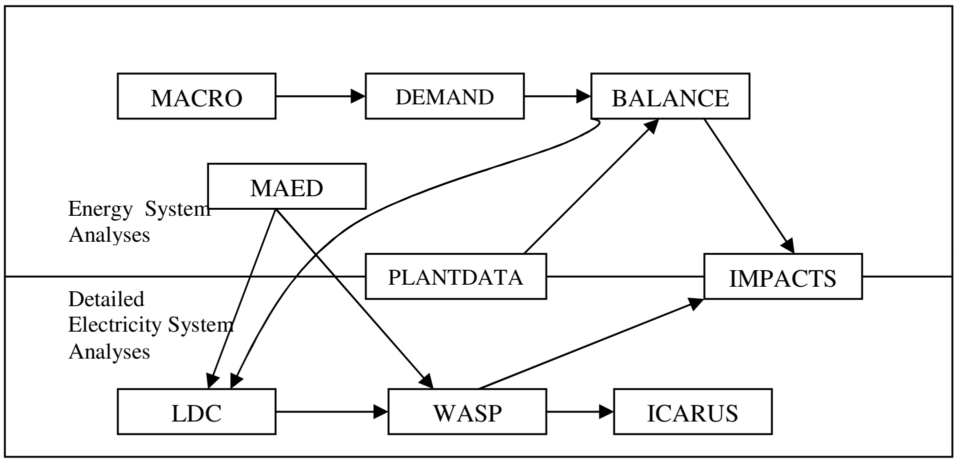

Figure 3.3 – Modules of ENPEP (Sakaryalı et al, 2000)

ENPEP includes 9 modules which are explained as given below by Quintanilla et al in 2002.

LDC: characterizes electrical load over time. Information generated by LDC is used in Wien Automatic System Planning Package (WASP) module for electricity generation.

ELECTRIC (WASP): determines the minimum cost expansion plan for electric generating system.

ICARUS: carries out detailed production cost and reliability calculations for a specified electricity generating system.

IMPACTS: estimates environmental and resource requirements for the BALANCE and/or WASP modules.

MAED : describes the electrical demand.

PLANTDATA: provides a library of technical data on electrical generating plants to be used in other modules.

MACRO: allows user to specify macroeconomic growth.
DEMAND: projects energy demand with the information provided in MACRO.
BALANCE: computes equilibrium energy supply and demand balances over the study period.

National Energy Modeling System (NEMS): EIA in USA uses a long term (20 to 25 years) forecasting tool to make projections and develop strategies; called NEMS which covers various simulation models to study major energy supply, demand, general domestic macroeconomic conditions and oil markets (Boedecker et al, 2000). NEMS Electricity Market Module (EMM) includes capacity planning, generation, transmission and pricing activities for different fuels. The decisions are optimized by choosing the minimum fuel, operation and maintenance and environmental costs meeting the electricity demand and environmental constraints. EMM uses the information received from fuel supply module, demand module, macroeconomic module and NEMS system module. On the other hand, other modules use the pricing outputs formed by the EMM module (EIA, 2001).

3.2 METHOD USED IN THE THESIS

Mentioning several types of demand forecasting models and comparing them, cointegration and ECM method is decided to be used in this study. This technique is superior to other methods in that it takes out economic, demographic or social variables that affect electricity demand. Also the method is straightforward to use and easy to interpret. Moreover, time series are generally nonstationary and they are usually cointegrated with one or more of the determinants (Fouquet et al , 1997).

3.2.1 Stationarity

To form a good model and prevent spurious regression, stationarity of the series should be developed. Stationarity of the series can be tested using ACF and PACF graphs in general. However, it is not appropriate to test stationary just by looking at graphs, so unit root tests have been developed aiming to decide the level of integration in the series. Following, some of these unit root tests are described.

Dickey & Fuller (1979) developed a method to test the stationarity assuming that the time series are generated by a pure autoregressive process of order 1, AR(1).

They estimated the regression equation:

$\Delta y_t = \delta\, y_{t-1} + \varepsilon_t$

or

$y_t = (1+\delta)\, y_{t-1} + \varepsilon_t$

where $\rho = (1+\delta)$

Dickey & Fuller (1979) used the test hypothesis:

Ho: $\delta = 0$

Ha: $\delta < 0$ which implies $\rho < 1$ and y_t is integrated of order 0 (i.e. the series is nonstationary if $\rho = 1$).

Dickey & Fuller (1979) also extended their study using deterministic terms. The stochastic process with a drift and an intercept are expressed as:

$\Delta y_t = \mu + \delta\, y_{t-1} + \varepsilon_t$

$\Delta y_t = \mu + \delta\, y_{t-1} + \gamma t + \varepsilon_t$

where μ is constant or intercept and t is a drift. The procedure is same with the first regression equation but critical values change as there are additional variables. For this reason, Dickey & Fuller calculated critical values for different sample sizes by Monte-Carlo simulation, which is used to evaluate quantities by experimenting as it's difficult to evaluate them analytically.

Augmented Dickey Fuller (ADF) test is one of the mostly used tests for unit roots. Dickey Fuller test is calculated under the assumption of AR(1) process, which does not representative in all cases. If AR(1) test is used when Y_t follows an AR(p) process, it causes the error term, ε_t, to be autocorrelated so that Dickey Fuller test will no more be valid. ADF test is calculated under AR(p) process which is represented as:

$$Y_t = a_0 + a_1 Y_{t-1} + a_2 Y_{t-2} + a_3 Y_{t-3} \ldots\ldots\ldots\ldots + a_{p-1} Y_{t-p+1} + a_p Y_{t-p} + \varepsilon_t$$

The procedure is the same with DF process, however the equation:

$$\Delta Y_t = a_0 + \gamma Y_{t-1} + \sum_{i=2}^{p} \beta_i \Delta Y_{t-i+1} + \varepsilon_t$$

is estimated where $\gamma = -(1 - \sum_{i=1}^{p} a_i), \beta_i = \sum_{j=1}^{p} a_j$ and $\gamma=0$ is tested like DF test.

Former unit root testing studies covered the ARIMA(p,d,q) models, where d=1 and p and q are known. **Said and Dickey** (1984) discussed the possibility to test for unit roots when p and q are unknown with null hypothesis that Ho: $\rho = 1$.

The model described below is estimated by Said and Dickey (1984)

$Y_t = \rho Y_{t-1} + Z_t$ (t=1,2…..)

$Z_t = \alpha Z_{t-1} + e_t + \beta e_{t-1}$ (t=……..-2,-1,0,1,2…..)

where $|\alpha| < 1$, $|\beta| < 1$, $Y_0=0$ and e_t are iid random variables.

Null hypothesis of $|\rho| = 1$ is tested.

Hasza and Fuller (1982) tested unit roots for monthly data. They used the model $(1-B)(1-B^{12})y_t = \varepsilon_t$

and estimated the equation that is given below

$y_t = \alpha_1 y_{t-1} + \alpha_2 y_{t-12} + \alpha_3 y_{t-13} + \varepsilon_t$

$y_t = \emptyset_1 y_{t-1} + \emptyset_2 (y_{t-12} - \emptyset_1 y_{t-13}) + \varepsilon_t$

and tested restrictions $[\alpha_1, \alpha_2, \alpha_3] = [1,1,-1]$

$[\varnothing_1, \varnothing_2] = [1,1]$ with F statistics.

The disadvantage of this model is that under the null hypothesis, the test imposes two unit roots at 0 frequency and the performance of the test may change when there are unit roots at only some of the seasonal frequencies. To conclude, the test suffers from low power and residual autocorrelation in the errors.

Dickey et al (1984) developed a unit root test for seasonal roots named Dickey, Hasza and Fuller test. The simplest seasonal time series model can be expressed as:

$y_t = \alpha_s y_{t-s} + e_t$ where t=1,2… and e_t is iid $(0,\sigma^2)$. If the data is monthly, s=12 and when the data is quarterly, s=4. The test hypothesis is $\alpha=1$, that is there is a seasonal unit root against the alternative that $\alpha<1$.

The test was later extended for higher order models.

The procedure for higher order series is described below:

- y_t is regressed on its lagged values and the equation is estimated.

$$\Delta_s y_t = \sum_{i=1}^{p} \lambda_i \Delta_s y_{t-i} + \varepsilon_t$$

- z_t is created from $y_t, y_{t-1}, \ldots, y_{t-h}$ as

$$z_t = y_t - \sum_{i=1}^{p} \lambda_i y_{t-i}$$

where $\hat{\lambda}_i$'s are OLS estimates of λ_i

- z_t is substituted into the below equation and δ and δ_i 's are estimated.

$$\Delta_s z_t = \delta . z_{t-s} + \sum_{i=1}^{p} \delta_i \Delta_s y_{t-i} + \varepsilon$$

The null hypothesis states that there is a seasonally integrated process. If the estimate of δ is below the lower critical value (i.e. H_o is rejected), than there is not any seasonal unit root or stochastic seasonality that can be removed by taking seasonal differences. If H_o is not rejected; seasonal differencing is done to reach stationarity instead of higher orders of periodic (regular) differencing (Charemza and Deadman, 1997, p. 105-106). A major disadvantage of this test is it does not allow for seasonal unit roots at all of the frequencies.

As most of the series contains seasonal components**, Hylleberg et al** (1990) improved a new method to test cyclical movements at different frequencies. Seasonality can be modelled with three types of equations that are purely deterministic seasonal process, stationary seasonal process and integrated seasonal process. HEGY test aims to search for the seasonal unit roots in a univariate series.

The procedure proposed by HEGY test for quarterly data is described below:

- The polynomial of the backshift operators can be expressed as:

$(1-B^4) = (1-B)(1+B)(1-iB)(1+iB) = (1-B)(1+B)(1-B^2)$ where the unit roots are 1,-1, i and –i and frequencies are 0, 2 cycle per year, 4 cycle per year and annual cycle respectively. The first root indicates nonseasonal root, whereas others indicate seasonal roots and cycles.

- The data is generated by equation $\varphi(B)x_t = \varepsilon_t$ and the null hypothesis states that all roots of $\varphi(B) = 0$, lies outside the unit circle.
- The equation formed by restructuring $\varphi(B)$ is expressed as

  $$\varphi^*(B) y_{4t} = \pi_1 y_{1t-1} + \pi_2 y_{2t-1} + \pi_3 y_{3t-2} + \pi_4 y_{3t-1} + \varepsilon_t$$

  where

  $$y_{1t} = (1 + B + B^2 + B^3)x_t = S(B)x_t$$
  $$y_{2t} = -(1 - B + B^2 - B^3)x_t$$
  $$y_{3t} = -(1 - B^2)x_t$$
  $$y_{4t} = -(1 - B^4)x_t$$

- The equation is estimated and π values are tested for periodic and seasonal unit roots.

For the first root, π_1=0 is tested whereas for the second root, π_2=0 is tested against the alternative of no unit roots. π_3 and π_4 are tested jointly with an F test for annual cycle. There is not a seasonal unit root if π_2 and either π_3 or π_4 are different from zero. To summarize, all π_i values have to be negative and if they are not, the series have certainly unit root and they are nonstationary. If the π_i values are negative, they should be compared with the critical values to conclude if there is a unit root or not. The critical values for sample sizes 48, 100, 136 and 200 can be obtained from Hylleberg et al (1990).

HEGY test can be extended by using deterministic variables such as a constant. However, the critical values change when additional variables are used in the equation which can be shown as:

$$\varphi^*(B) y_{4t} = \pi_1 y_{1t-1} + \pi_2 y_{2t-1} + \pi_3 y_{3t-2} + \pi_4 y_{3t-1} + \mu_t + \varepsilon_t$$

Bealieu and Miron (1993) extended the quarterly HEGY procedure and derived the test for monthly data. They used Monte Carlo simulations to compute finite sample critical values, which are also used in this study. HEGY test is superior compared to other unit root test in that it takes all the periodic and seasonal frequencies into account and it allows finding unit roots at any seasonal frequency without being dependent to other unit roots. In brief, the test procedure is as indicated below (Beaulieu and Miron, 1993). First, new series are generated, where x_t is the series which will be investigated.

$$y_{1t} = (1 + B + B^2 + B^3 + B^4 + B^5 + B^6 + B^7 + B^8 + B^9 + B^{10} + B^{11})x_t$$
$$y_{2t} = -(1 - B + B^2 - B^3 + B^4 - B^5 + B^6 - B^7 + B^8 - B^9 + B^{10} - B_{11})x_t$$
$$y_{3t} = -(B - B^3 + B^5 - B^7 + B^9 - B^{11})x_t$$
$$y_{4t} = -(1 - B^2 + B^4 - B^6 + B^8 - B^{10} + B^{11})x_t$$
$$y_{5t} = -1/2\,(1 + B - 2B^2 + B^3 + B^4 - 2B^5 + B^6 + B^7 - 2B^8 + B^9 + B^{10} - 2B^{11})x_t$$
$$y_{6t} = \sqrt{3}/2\,(1 - B + B^3 - B^4 + B^6 - B^7 + B^9 - B^{10} + B^{11})x_t$$
$$y_{7t} = -1/2\,(1 - B - 2B^2 - B^3 + B^4 + 2B^5 + B^6 - B^7 - 2B^8 - B^9 + B^{10} + 2B^{11})x_t$$
$$y_{8t} = -\sqrt{3}/2\,(1 + B - B^3 - B^4 + B^6 + B^7 - B^9 - B^{10})x_t$$
$$y_{9t} = -1/2\,(\sqrt{3} - B + B^3 - \sqrt{3}B^4 + 2B^5 - \sqrt{3}B^6 + B^7 - B^9 + \sqrt{3}B^{10} - 2B^{11})x_t$$
$$y_{10t} = 1/2\,(1 - \sqrt{3}B + 2B^2 - \sqrt{3}B^3 + B^4 - B^6 + \sqrt{3}B^7 - 2B^8 + \sqrt{3}B^9 - B^{10})x_t$$
$$y_{11t} = 1/2\,(\sqrt{3} + B - B^3 - \sqrt{3}B^4 - 2B^5 - \sqrt{3}B^6 - B^7 + B^9 + \sqrt{3}B^{10} + 2B^{11})x_t$$
$$y_{12t} = -1/2\,(1 + \sqrt{3}B + 2B^2 + \sqrt{3}B^3 + B^4 - B^6 - \sqrt{3}B^7 - 2B^8 - \sqrt{3}B^9 - B^{10})x_t$$
$$y_{13t} = (1 - B^{12})x_t$$

Then the equation:

$\varphi(B)y_{13t} = \sum_{k=1}^{12} \pi_k y_{k,t-1} + \varepsilon_t$ is estimated where

$\varphi(B)$ is the polynomial of backshift operator

ε_t is the white noise process

π_i values are calculated from the equation by OLS and the results are compared with critical values found by Monte Carlo Simulation to test π_i values.

One sided test hypothesis for 0 and π frequencies is:

H_o: $\pi_k = 0$

H_a: $\pi_k < 0$

If π_i is 0, the test hypothesis is accepted indicating that there is a unit root at zero frequency. The procedure is also same for π frequency.

Unit roots at other frequencies are tested in pairs. First of all even k values are tested with a two sided hypothesis.

H_o: $\pi_k = 0$ (there is unit root at that frequency)

H_a: $\pi_k \neq 0$

If one fails to reject H_o , there may be a unit root at that frequency and odd values of k are tested with a one sided hypothesis:

H_o: $\pi_{k-1} = 0$ (if there is unit root at that frequency)

H_a: $\pi_{k-1} < 0$

The seasonal unit roots are:

-1 , ±i , -1/2 (1±√3i) , 1/2 (1±√3i), -1/2 (√3±i) , 1/2 (√3±i)

whose frequencies are π , ±π/2 , ±2π/3 , ±π/3 , ±5π/6 , ±π/6 and which shows 6,3,9,8,4,2,10,7,5,1 and 11 cycles per year respectively.

Another method to test the stationarity at seasonal frequencies is applying F statistics by testing $\pi_{k-1} = \pi_k = 0$. If π_k is different from 0 for k=2 and for at least one member of each pairs {3,4}, {5,6},{7,8},{9,10},{11,12}, there exists no unit root in the series.

The test can be extended by adding seasonal dummy variables, a trend and a constant to the equation and can be written as:

$$\varphi(B)y_{13t} = \sum_{k=1}^{12} \pi_k y_{k,t-1} + m_0 t + m_1 + \sum_{k=1}^{12} m_k S_{kt} + \varepsilon_t$$

Philips and Peron test is similar to Dickey Fuller test except it allows weakly dependent and heterogeneously distributed error terms. The procedure developed by Philips and Peron (1998) is that firstly a first order autoregression with a constant and a time trend (if necessary) is estimated. Then the appropriate Z statistics is calculated. The test regression is the AR(1) process indicated below:

$$\Delta y_t = \alpha + \beta y_{t-1} + \varepsilon_t$$

The asymptotic distribution of the test is the same with ADF t statistics.

The standard unit root tests proposes the null hypothesis of unit root against the alternative and if there is not a strong evidence for the alternative, the test is accepted indicating unit root in the series. On the other hand, **Kwiatkowski, Phillips, Schmidt, Shin (KPSS) Test** developed in 1992 uses the null hypothesis that a series is trend stationary instead of having a unit root. So, it is very useful to support the results of other unit root test with KPSS (Aggarwala and Kyaw, 2005). KPSS tests the null hypothesis that the random walk equation has zero variance. The series can be expressed as:

$y_t = \xi t + r_t + \varepsilon_t$ which includes a deterministic trend; random walk r_t and a stationary error.

$r_t = r_{t-1} + u_t$ (u_t is iid$(0,\sigma^2_u)$). The stationary hypothesis is $\sigma^2_u=0$.

HEGY test for monthly data is used in this study its ability to consider the seasonality. Additionally, ADF test is used as an alternative to HEGY test and for being reliable and a commonly used method.

3.2.2 Cointegration

Granger and Newbold (1974) indicated that spurious regression occurs while working with nonstationary time series. Cointegration is important in econometrics as it is the complementary idea of meaningful versus spurious regression and it allows describing the existence of equilibrium among time series. A set of cointegrating variables can also be represented as an ECM where a term representing the deviation of observed values from the long-run equilibrium enters the model.

Cointegration techniques have begun to be used in many studies in 1990's after Engle and Granger developed cointegration procedure in 1987. Cointegration is then used in various areas like purchasing power parity, transmission of macroeconomic shocks across countries, currency substitution and market efficiency (Craig and Rush, 1991).

"If Y is a series such that d^{th} differences $(1 - B)^d Y$ are stationary, it is called integrated and denoted I(d)" that is when the series is I(0) it is stationary. When a vector of series is at the same order whereas their linear combination is I(0), the series are said to be cointegrated and the series can be modeled as ECM. Engle et al (1989) extended their research in following years with seasonal integration and cointegration at zero and seasonal frequencies and estimated electricity demand both in the short run and in the long with regression and an error correction model.

In another way, Engle and Granger (1987) defined the cointegration as:

Time series x_t and y_t are cointegrated of order d,b where $d \geq b \geq 0$

$x_t, y_t \sim CI(d,b)$

if

1. both series are integrated of order d
2. there exists a linear combination of these variables, say $\alpha_1.x_t + \alpha_2.y_t$ which is integrated of order d-b

and $[\alpha_1,\alpha_2]$ is called cointegrating vector (Charemza and Deadman, 1997, 127)

A simple example for cointegration defined by Engle and Granger (1987) is given below:

Let $\{x_t\}$ and $\{y_t\}$ be two series both of which are I(1).

$x_t + \alpha y_t = e_t$ (1)

$x_t + \beta y_t = u_t$ where $\alpha \neq \beta$

$u_t = u_{t-1} + \varepsilon_{1t}$

$e_t = \rho e_{t-1} + \varepsilon_{2t}$

$|\rho| < 1$

$(\varepsilon_{1t}, \varepsilon_{2t})$ is identically and independently distributed as a bivariate normal with E(ε_{1t})=E(ε_{2t})=0,

Var(ε_{1t})=σ_{11}, Var(ε_{2t})=σ_{22},

Cov($\varepsilon_{1t}, \varepsilon_{2t}$)=$\sigma_{12}$

Equations indicated above are solved for x_t and y_t

$$x_t = \alpha(\alpha-\beta)^{-1} u_t - \beta(\alpha-\beta)^{-1} e_t$$
$$y_t = (\alpha-\beta)^{-1} e_t - (\alpha-\beta)^{-1} u_t$$

u_t is random walk and x_t and y_t depend on u_t and those two series are I(1). However, $x_t + \alpha y_t$ is I(0) because e_t is stationary. The vector [1,α] is the cointegrating vector and $x_t + \alpha y_t$ is the equilibrium relationship (Banerjee et al., 1993, p.137).

To sum up, Engle and Granger tests the estimated residuals of equation (1) for stationarity considering the test hypothesis of no cointegration against the alternative. Mostly used unit root tests for residuals include ADF, Cointegrating Regression Durbin Watson test (CRDW) and Dickey Fuller tests where ADF and Dickey Fuller tests are discussed before.

CRDW test includes computing DW statistics for the estimated deviations from the long run path.

$CRDW = \sum_{t=2}^{T} (\hat{u}_t - \hat{u}_{t-1})^2 / \sum_{t=1}^{T} \hat{u}_t^{\,2}$ where $\hat{u}_t$ is the OLS residuals of cointegrating equation. When the value of test statistics is smaller than the critical value, the possibility of rejecting the cointegration relation increases.

In addition to Engle and Granger's method, Johanson and Juselius (1990) cointegration approach is one of the most popular methods that is generally used with multiple equations and in Vector Autoregression models. This procedure includes maximum likelihood estimates of unconstrained cointegrating vectors and makes distinction between one or more cointegrating vectors (Siklos and Wohar, 1996). Johansen test gives two test statistics that are trace test statistics which tests the hypothesis that there are at most r cointegrating vectors and maximum eigenvalue test statistics which tests the hypothesis of r cointegrating vectors against the alternative of (r+1) cointegrating vectors. If there is a conflict between the results of both statistics, maximum eigenvalue test statistics is preferred (Banerjee et al, 1993).

Johansen trace test procedure is as follows:

- In the first step, ΔZ_t is regressed on its lagged values, $\Delta Z_{t-1}, \Delta Z_{t-2}, \ldots \Delta Z_{t-k+1}$

  Residuals from these equations at time t are represented with $R_{0t.}$

  ΔZ_{t-k} is regressed on $\Delta Z_{t-1}, \Delta Z_{t-2}, \ldots \Delta Z_{t-k+1}$

  Residuals from these equations at time t are represented with $R_{kt.}$

- In the second step, four n x n matrices is computed from the second moments and cross products of residuals R_{0t} and R_{kt} and they are indicated as S_{00},S_{0k},S_{k0},S_{kk} where

  $S_{ij} = T^{-1}\sum_{t=1}^{T} R_{it} R'_{jt}$ i,j=0,k (T:sample size)

- In the third step equation given below is solved to find the roots and eigenvalues of the polynomial equation in μ, where $\hat{\mu}_i$ are eigenvalues

$$\left| \mu S_{kk} - S_{k0} S_{oo}^{-1} S_{0k} \right| = 0$$

Finally, for each $\hat{\mu}_i$,

$$LR_{tr}(r/k) = -T \sum_{i=r+1}^{k} \ln(1-\hat{\mu}_i)$$

LR statistics, which tests the null hypothesis that there are at most r cointegrating vectors, is calculated. The test starts from r=0 and if it is rejected goes on with r≤1, r≤2 (Charemza and Deadman, 1997, pp176-178).

On the other hand maximum eigenvalue test statistics is calculated as

$$LR_{\max}(r/r+1) = -T\log(1-\hat{\mu}_{r+1})$$

$$LR_{\max}(r/r+1) = LR_{tr}(r/k) - LR_{tr}(r+1/k)$$

where the notations are the same with trace statistics.

3.2.3 Error Correction Models

If the variables in an equation are cointegrated, then there is an adjustment process that prevents errors to get larger in the long run. That is, cointegration is required to obtain an ECM. Engle and Granger (1987) proposed a model which uses OLS in estimating regression. The variables should be all at the same order, for instance I(1) or the dependent variable should be I(1) whereas explanatory variables are CI(d+1,d) to test for cointegration.

When first situation is considered and both y_t and x_t are I(1).

$y_t = \beta . x_t + u_t$

Cointegrating relation is investigated by Dickey Fuller or ADF test and estimates of u_t is found to be stationary and cointegrating vector is found to be $\left(1, -\hat{\beta}\right)$

Then the short run model with an error correction mechanism can be written as:

$$\Delta y_t = \alpha_1 \Delta x_t + \alpha_2 (y_{t-1} - \beta x_{t-1}) + \varepsilon_t$$

where ε_t : error term and $\alpha_2 < 0$

A disadvantage of Engle-Granger technique is said to be that whether it is suitable to estimate the equation by OLS or not as both x_t and y_t are nonstationary. However, it is shown that properties of OLS are good enough (Charemza and Deadmand, 1997, p. 131-134).

3.2.4 Autoregressive Conditional Heteroskedasticity (ARCH)

There are several studies in literature which investigates the variability such as standard deviation, deviation from trend and coefficient of variation. ARCH is one among them that is firstly introduced by Engle in 1982 as a model to measure the conditional variance of a series as a function of past errors when the variance of the model depends on its past values. ARCH procedure allows the conditional variance to change over time while the unconditional variance remained constant. Conditional variance does not depend on the past and one-period forecast variance is assumed to be constant in traditional methods which ignored that both the conditional mean and conditional variance may change over time.

The ARCH model of Engle is given by:

$$y_t = \varepsilon_t h_t^{1/2}$$

$$h_t = \alpha_0 + \alpha_1 y^2{}_{t-1}$$

with $V(\varepsilon_t)=1$ and h_t is the conditional variance. The unconditional variance is either zero or infinity.

Engle propose using maximum likelihood estimation instead of OLS as it is less efficient while investigating ARCH effect.

In many studies only unconditional variance of the random variable is considered, however, there are some cases where the volatility should be modeled and forecasted. By this way, more accurate intervals and more efficient estimators can be obtained. There are several reasons to estimate volatility; first being the requirement to analyze the risk of holding an asset, second to obtain more accurate intervals by modeling the variance of the errors, third to obtain more efficient estimators by handling the heteroskedasticity properly.

ARCH LM test, which is a Lagrange Multiplier (LM) test, is used to check the autoregressive conditional heteroskedasticity in the residuals. The test statistics is computed from the auxiliary test regression indicated below:

$$e^2{}_t = \beta_0 + \left(\sum_{s=1}^{q} \beta_s e^2{}_{t-s} \right) + v_t$$

where e_t is the residuals. Above equation is a regression of squared residuals on a constant and lagged squared residuals up to order q. The null hypothesis is the unavailability of ARCH up to order q in the residuals. LM statistic is asymptotically distributed as $\chi^2(q)$ (Quantitative Microsoftware, 2000, p.361)

The success of ARCH models according to Bera and Higgins (1993) are their simplicity and ease. The models also take care of the clustered errors, nonlinearity and changes in ability to forecast.

Bollerslev (1986) extended ARCH model, introducing a more general procedure, which is called **Generalized ARCH (GARCH)**. ε_t denotes a real-valued discrete time stochastic process and Ψ_t denotes the information set of all information through time t. Then, the GARCH(p,q) process is given as:

$\varepsilon_t \mid \Psi_{t-1} \sim N(0,h_t)$

$$h_t = \alpha_0 + \sum_{i=1}^{q} \alpha_t \varepsilon^2{}_{t-i} + \sum_{i=1}^{p} \beta_i h_{t-i}$$

$$= \alpha_0 + A(L)\varepsilon_t{}^2 + B(L)h_t$$

where $p \geq 0$, $q > 0$,

$\alpha_0 > 0$, $\alpha_i \geq 0$ $i=1,2,\ldots q$

$\beta_i \geq 0$ $i=1,2,\ldots p$

The equation consist a constant (α_0), news about volatility from previous periods (ε_{t-i}) and last periods' forecast variance (h_{t-i}).

If p=0, the process becomes an ARCH(q) process and if p=q=0, ε_t is white noise. Therefore, GARCH process allows more flexible lagged structure then ARCH, which can be considered as a learning mechanism.

Engle et al (1987) defined another model **ARCH in Mean (ARCH-M)**, where the conditional variance is introduced into the mean equation as:

$$y_t = x'_t \gamma + \sigma_t^2 + \varepsilon_t$$

In finance, the expected return of an asset depends on the expected risk and ARCH-M model is appropriate in this situation.

Glosten et al (1993) defined **Threshold ARCH (TARCH)** model as:

$$\sigma^2{}_t = \omega + \alpha\varepsilon^2{}_{t-1} + \gamma\varepsilon^2{}_{t-1} d_{t-1} + \beta\sigma^2{}_{t-1}$$

where dummy variable d_t=1 if ε_t>0, and 0 otherwise because in the model good and bad news had different effects on conditional variance. If there are good news, the effect is α, however the effect of bad news is $\gamma + \alpha$.

Nelson developed (1991) **Exponential GARCH (EGARCH)** model which considers the sign of the volatility. As it can be seen from the equation the variance responds differently to negative and positive residuals and the effect of positive or negative shocks differ. The model is:

$$\log(\sigma^2{}_t) = \omega + \beta \log \sigma^2{}_{t-1} + \alpha \left| \frac{\varepsilon_{t-1}}{\sigma_{t-1}} \right| + \gamma \frac{\varepsilon_{t-1}}{\sigma_{t-1}}$$

The variance of the EGARCH model is conditioned on its lagged values and the standardized error term(ε_t/σ_t).

TARCH and EGARCH models are adequate to model the volatility in case of asymmetry. These kinds of models were extended later and some other models such as IGARCH, FIGARCH models are introduced, where Integrated GARCH (IGARCH) models which have most of the features of the unit root processes for the mean and Fractionally Integrated GARCH (FIGARCH) models, which can explain the observed temporal dependencies in financial market uncertainty. The IGARCH model is given by

$$\Phi(B)(1-B)\varepsilon^2{}_t = \omega + \left[1-\beta(B)\right]v_t$$

where $\Phi(B) \equiv \left[1-\alpha(B)-\beta(B)\right](1-B)^{-1}$ is of order m-1 and FIGARCH is obtained by replacing the first differencing operator with fractional differencing operator (Baillie et al , 1996).

ARCH models became very popular recently for their strength in improving the forecasts. In general, they are used in research areas such as risk analysis, portfolio selection, monetary growth, stock return and pricing. The most popular ARCH model that is used in financial market volatility is the GARCH(1,1) model (Hamilton and Susmelb, 1994). Caporale and Doroodian (1994) applied a GARCH(1,1) model for exchange rate in US. The results showed that exchange rate volatility has a negative effect on trade flows. Alper and Yilmaz (2004) worked on the impact of financial globalization on stock market volatility in Turkey with a GARCH model and found strong ARCH effects in residuals for İstanbul stock exchange. GARCH(1,1) model again gave best results among other GARCH models. Zhang (2007) also used

GARCH model to estimate the demand structure of electronic products and used multivariate GARCH(1,1) model as it allows time varying covariance. The simulation performed during research indicated that sample size affects the accuracy of the model and mean squared error decreases with increasing sample size.

Moreover, Apergis (1997) studied the relation between inflation uncertainty and demand for money in Greece. He used ARCH model for its advantage of estimating conditional mean and variance jointly. The outcome supported the previous outcomes that inflation volatility affects money demand.

Worthington et al (2005) used daily spot prices from 13 December 1998 to 30 June 2001 in half-hourly basis to search for the volatility of daily spot prices in 5 regions of Australian electricity market utilizing a multivariate GARCH process. The results indicated that there is a national market for electricity in Australia; however the regional electricity spot markets are not integrated. The results also showed that there are strong ARCH and GARCH effects because own volatility and cross-volatility spillovers are significant for nearly all markets.

Duran and Şahin (2006) studied the volatility and its direction between Istanbul stock exchange services, financial, industrial and technology indexes using EGARCH model. Conditional variances obtained from the EGARCH are used as volatility indicators in a VAR model.

Çifter et al (2007) modeled the risk for the interest rates using GARCH and some other methods such as Pareto Analysis. The results indicated that the GARCH model is effective in short run whereas other methods were better in the long run because GARCH models are said to ignore structural breaks and changes in regime.

Ferreira et al (2006) studied the asymmetric effect of shocks in stock market to stock returns by EGARCH and TARCH models in various countries because negative shocks are said to generate more volatility compared to positive shocks. Both the model captured asymmetry. Some market had high α value indicating that the effect of shocks stays long.

The previous studies indicate that ARCH was ignored in electricity demand studies. However, in this study ARCH effect is also considered as for the modeling is in the medium term.

3.2.5 Validity of the Model

The models are formed under some assumptions such as normality, heteroskedasticity or autocorrelation, so the residuals have to be tested to approve that the correct model is formed.

Jarque-Bera test statistics is used to test **normality** of the residuals. This test depends on the skewness and kurtosis and explained with the null hypothesis;

H_o: The series is normally distributed

H_a: The series is not normally distributed

Hypothesis is tested with

$$\chi^2 = \frac{N-k}{6}\left(S^2 + \frac{(K-3)^2}{4}\right)$$

where S is skewness, which measures of asymmetry of the distribution of the series around its mean and should be around 0. K is kurtosis, which measures flatness of distribution of the series and k is the number of estimated coefficients used to create the series and is around 3. As can be described form the equation, the test measures the divergence of skewness and kurtosis from the normal distribution. Test statistics is distributed as χ^2 with 2 degrees of freedom (Quantitative Microsoftware, User Guide, 2000, p.153).

In general Durbin Watson test is applied to test for the existence of **serial correlation** in the residuals. DW test statistics is calculated as:

$$d = \frac{\sum_{t=2}^{T}(e_t - e_{t-1})^2}{\sum_{t=1}^{T} e_t^{\,2}}$$

where e_t is the error term.

It is expected to have DW values between 1.5 and 2.5. Bigger values than 2.5 show a negative autocorrelation in the series, whereas lower values than 1.5 indicates a positive autocorrelation in the series. However, in this study, Breusch Godfrey serial correlation LM test is used to test for serial correlation, as Durbin Watson test does not give accurate results when lagged values of the dependent variable is included in the model (Quantitative Microsoftware, 2000, p.360-361). The null hypothesis is that there is no serial correlation in the residuals up to lag order k.

A well-known test for **heteroskedasticity** is White's Heteroskedasticity Test, where the residuals from a regression are estimated by least squares and then the auxiliary regression, where the squared residuals on all possible cross products of the regressors are estimated (Quantitative Microsoftware, User Guide, 2000, p.362).

The parameters in the model are also tested for **stability** within some subsamples of data. The test that is used in this study for stability of the parameters is CUSUM test, developed by Brown et al (1975). The test is based on cumulative sum of the recursive residuals and if the cumulative sum is outside the %5 critical value range, it is concluded that there is some instability. The values are calculated by the statistics:

$$W_t = \sum_{r=k+1}^{t} w_r / s$$

where t=k+1,…,T , w is the recursive residual, s is the standard error of regression. Another alternative to CUSUM test is CUSUM of square test, which is based on below formula:

$$W_t = \left(\sum_{r=k+1}^{t} {w_r}^2 \right) / \left(\sum_{r=k+1}^{T} {w_r}^2 \right)$$

The test gives a graph of s_t against t and compares the results within %5 critical lines.

3.3 ELECTRICITY DEMAND FORECASTING STUDIES IN TURKEY

Although forecasting models in the world are used since 1950's, electricity demand forecast studies in Turkey were first performed in 1960's. First studies were mathematical models that consist of curve-fitting models and they were studied in universities. Petrol crisis in 1973 was a turn-over point for the modeling because simple models were transferred into complex models that consider relation between energy and economy, energy density and investments (Çağlar, 2006). Since then, electricity demand forecast studies were done by ETKB in coordination with State Planning Organization (DPT).

With the incitement of World Bank in 1984, **MAED** and **WASP III** that were presented by the IEA for developing countries and that are sub modules of ENPEP, begun to be used for demand forecasting and production planning activities. In the Long Term Electricity Demand Report, which is published by ETKB (2004), MAED is defined as a simulation model that evaluates the medium and long term energy and electricity demand of a country.

Both medium (1993-1997) and long-term (1996-2010) production and consumption analysis were carried out in 1990's by using WASP and MAED by ETKB in cooperation with DPT. The aim of the study was minimizing the cost of the energy mix that will meet electrical energy demand (TEK, 1994). However, those were programs used in developed countries and were not efficient for Turkey. The error percentage of the forecast was about 17% for 3 years, 25% for 10 year, 44% for 15 years in 2003 (Kumbaroğlu, 2006). The reason for such a variation was not only the shocks such as economic crisis, which Turkey faced in previous years but also integrating the energy-economy relation, supply-demand relation and effects of the prices to the model later instead of using these key factors inside the model. Another reason was the unreliable demographic and economical parameters used in the model. Additionally, because of the import and export in Turkey, policies and prices of other countries should be taken in account. Derman (2006) indicated that wrong demand forecasts may be results of

- Not realizing the electricity generation, transmission, distribution investments indicated in demand projections,
- Breakdown of the networks and big power plants,
- Long period drought,
- Economical crisis,
- Natural disasters in industrialized and populated cities (especially earthquakes).

Electricity demand forecasting gained more importance after the enactment of Electricity Market Law in 2001. The law states that distribution companies in each region will be responsible for demand forecasts in that region, and the transmission company TEIAS will collect those forecasts to estimate the total electricity demand in Turkey and prepare demand projections. In addition to Electricity Market Law Electrical Energy Sector Reform and Privatization Strategy Document was admitted in 17^{th} March 2004, aiming to remove the short and medium term uncertainties, being one of the targets. From that day, electricity demand studies accelerated in collaboration with EPDK, ETKB, DPT and Undersecretariat of Treasury. MAED that is already used by ETKB is decided to be used with some improvements as the appropriate tool for electricity demand forecasting and a demand forecast is done in 2004. The main determinants of electricity consumption are taken as GDP, population, demographical changes and market growth and the forecasts are done for household, industry, agriculture and transportation sectors basically. Model can consist of other variables such as weather, climate conditions, seasonal changes, income per capita, employment, geographical properties of the country and technological developments. The model uses a large amount of economic, social and technical data and makes projections regarding the energy policy and targets of the country. A base year, 1990, and two types of parameters, fixed and technical parameters, are used in the model where fixed parameters are required for the restructuring of the base year and technical parameters are necessary for projections. The accuracy of the model is tested with a control year (ETKB, 2004). Electricity demand projections in each sector including agriculture, transportation, household and industry can be done for 5 or 10 years time intervals. In the study, 2020 is chosen as the last year. Demands of the in between years are determined by interpolation and/or according to GDP rates. 3 scenarios, first regarding the contribution of the agriculture, industry, transportation, household and services to GDP; second showing the sensitivity of the electricity demand to changes in manufacturing sub sectors and third (being the former scenario) where GDP values are obtained from DPT, are used to determine the electricity demand. The projections performed by using three different scenarios can be seen in Table 3.1.

Table 3.1 – Demand Projections using MAED (ETKB, 2004)

Demand Projections (GWh)				
	2005	**2010**	**2015**	**2020**
Former Scenario	168.262	269.842	409.531	570.521
Scenario I	163.191	242.021	356.202	499.489
Scenario II	159.399	216.750	294.563	406.530

When the results are compared, former scenario has the highest and scenario 2 has the lowest demand forecasts. Extended knowledge about MAED is explained subsequently.

MAED is a simulation model based on scenario approach which is initially introduced in DOS system and distributed to 40 countries and then converted to EXCEL which is more user friendly (International Atomic Energy Agency, 2006). It is for medium and long term forecasts and not applicable on short term. The inputs used by MAED are energy sector data including socioeconomic and technological scenario assumptions, substitutional energy uses, process efficiencies and hourly load characteristics whereas the outputs are useful energy demand, final energy demand, electricity demand, hourly electric load and load duration curves. Some examples for inputs can be GDP as an economic input, population, growth rate, urbanization, and labor force as demographic inputs, household size, mobility, car ownership, transport modes as life style and energy intensity, efficiency and penetrations as technologic inputs. The MAED methodology includes reconstructing the energy consumption by sectors and fuels in the base year where final energy consumption of the country is broken down into major consuming sectors such as agriculture, residential, transportation and then energy consumption in each sector is disaggregated such as motor fuels for transportation. Secondly, different scenarios are established and energy demand forecasts from these scenarios are analyzed. In this stage, the social, economic and technical factors that affect the electricity demand in each subsector are identified to develop scenarios. Each scenario is used to evaluate energy demand. Base year selected for MAED study is important and it should include recent economic and energy data which should be consistent and reliable. In the base year, conditions and energy consumption should be stable and there should not be any unusual events on that year. Some examples for scenario inputs can be population, GDP, technological improvements etc. (ADICA, 2005a). Price of energy is not considered in the model explicitly, but it can be used by the planner while developing scenarios. Additionally, fossil fuel consumption is not divided into its type such as oil, gas and coal.

The model includes 4 modules that are energy demand calculations, hourly electric power demand, electric load duration curve and load modulation coefficients. The first module makes energy demand projections by using the different social, economic and technological scenarios. Hourly electric power demand module finds out the hourly electric load on the grid system from the electricity demand of each sector. Electric load duration curve module uses the hourly loads and creates the load duration curve of the power system which is necessary for the WASP module. The last module is an auxiliary module. It uses the load curves to analyze the past evolution of the coefficients (ADICA, 2005b).

To sum up, the previous works completed by ETKB on energy demand includes the works done in 1989, 1990, 1993, 1997, 1999, 2003 and 2004 using MAED and WASP. All of the forecasts showed that, both gross and net electricity demand forecasts differed significantly from the actual consumption, with an average of %20 in the long run as illustrated in Figures 3.4 and 3.5.

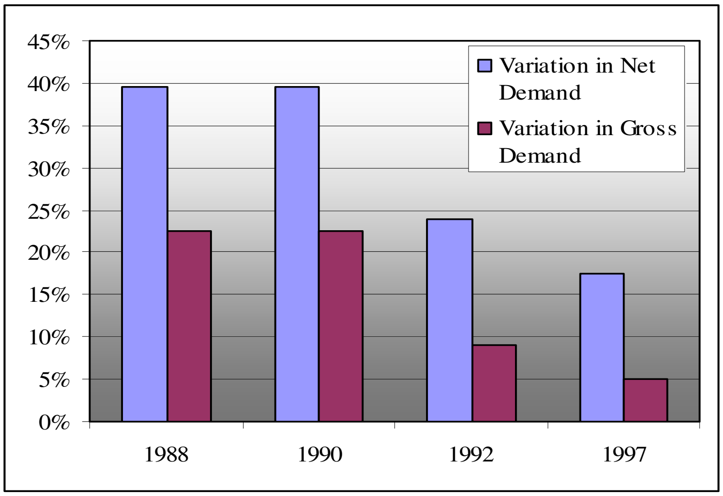

Figure 3.4 – Deviation of forecasts from the actual demand (Forecast up to 2000)

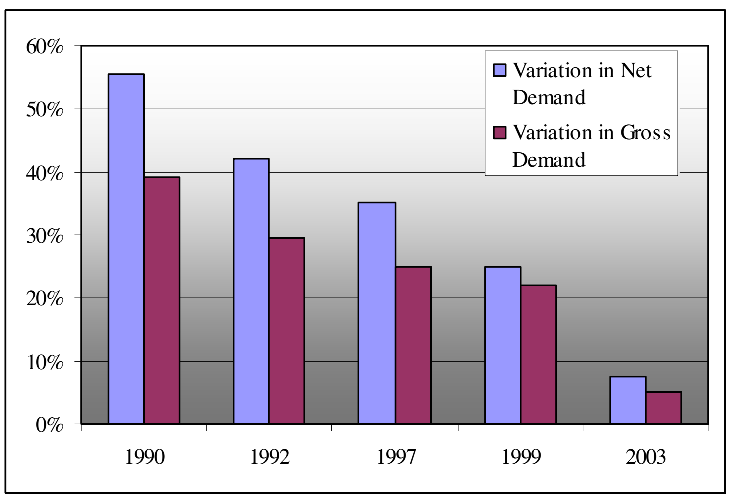

Figure 3.5 – Deviation of forecasts from the actual demand (Forecast up to 2005)

Variation of gross electricity demand was less than the variation of net electricity demand forecasts because of increasing theft and losses in recent years. Additionally, the economical crisis in years 2000 and 2001, lack of detailed data, wrong estimates of fuel prices and GDP by DPT were other reasons of this variation. All of these reasons caused high demand estimates for electricity, which then caused making unnecessary investments and excess capacity in 2000's or idle power plants (Keleş, 2005).

The last improvements in public sector about electricity demand forecasts included Regulation Concerning Electrical Energy Demand Forecast (26129), which was published on 04/04/2006 by EPDK. The regulation consists of the method and period of forecasting, data set that will be used, validation and verification of the forecast and results. According to the regulation, distribution companies perform forecasting studies by using data collected from industrial zones. Forecasts are done for 10 year periods each year for pessimistic, base and optimistic scenarios. Data used includes former values of the demand and demographical, technological, scientific, economical, social and environmental data. The method should well reflect the relation between variables, design should be understandable and easy, consistent with economic theories. Furthermore, the model and variables should be verified statistically and should also be stable. Tests (at most %10 confidence interval) to control the validity should be done also. Forecasts prepared by distribution companies are sent to TEIAS and collected there. Finally, TEIAS present the forecasts to EPDK for approval before they are published in the TEIAS web site. However, it is a fact that the procedure does not operate as written in the Regulation because of various reasons such as lack of data in distribution regions.

On the other side, electricity demand forecasting studies are not limited with Governmental bodies. There are several studies done in universities or by private sector companies in short, medium and mostly in long term basis which will be mentioned in the following paragraphs.

Taking benefit of statistical analysis packages such as SYSTAT, STATGRAPHICS, MICRO TSP, Gökçe (1991) forecasted yearly electricity demand in Turkey for the period 1989-2000. She used yearly electricity consumption data between 1923 and 1988 in logarithm form. She divided the modeling techniques into two, first being as single time series models like classical regression, regression model with ARMA(p,q) error structure, exponential smoothing methods and ARIMA(p,d,q) modeling with Box-Jenkins technique. The outcome of the models showed that ARMA(1,10) regression model, indicated below has the least minimum square errors.

$Y_t = 3{,}991 + 0{,}102\ T + U_t$

$U_t = 0{,}959\ U_{t-1} - 0{,}786\ a_{t-10} + a_t \quad t = 1924 \ldots 1988$

Y_t= electricity consumption in year t

T= trend variable

U_t = stationary process (ARMA (1,10))

a_t= error term.

Second modelling technique is described as causal model, having additional variables that are GDP, real electricity prices, wholesale price index and population. This time the data covered the period between 1970 and 1988. The research indicated that there is strong relationship between electricity consumption and economic variables. Trying various models, ARMA (1,0) model gave the best results.

The model:

$Y_t = -18.49 + 1.403\ N_t + 1.09\ GDP_t + U_t$

$U_t = 0.710\ U_{t-1} + \varepsilon_t \quad t=1971 \ldots 1988$

Where

N_t =electricity price in year t

GDP_t =gross domestic product in year t

Regression results indicated %1 change in GDP results in %1.04 change in electricity consumption. Thus; there is a GDP elasticity of price. On the other hand, each %1 change in consumption is a result of %1.4 decrease in price.

Because there is not so much substitute of electricity, changes in price do not effect the electricity consumption visibly.

Toptaş (1992) used Box-Jenkins time series model to estimate monthly peak load demand in Turkey. The data covered 144 months electricity consumption from January 1979 to December 1990. Seasonal variation occurred as monthly data is used in the model. The resulting model was seasonal ARIMA, SARIMA $(0,1,0)_1$ x $(1,1,0)_{12}$ which is expressed in the full form as $(1\text{-}ØB^{12})\Delta\Delta^{12}X_t = e_t$. Toptaş also developed simple linear regression, curve linear regression, exponential regression, moving averages, simple exponential smoothing and Winter's seasonal exponential smoothing methods and the results indicated that curve linear regression and exponential regression is better than simple linear regression whereas Winter's exponential smoothing method is superior to simple exponential smoothing and Holt's exponential smoothing. Also, moving averages is good at estimating only short term forecasts because of being beneficial for 1 period ahead. Mean absolute percentage error and root mean square error are used to evaluate the forecasting performance of the models and it is found out that the best results are obtained by Box-Jenkins technique.

Kırkgül (1993) investigated the relationship between electricity consumption, GDP, price and population. The data collected proves that increase in GDP causes a parallel increase in electricity consumption between years 1973 and 1984. When income increased, usage of electrical appliances increased too. Also this is supported by technological development. Electricity consumption is also affected by electricity price because increase in electricity price may decrease usage of electrical household appliances. In fact, substitutes of electricity play an important role when affect of electricity price on consumption is searched. For instance, there is no substitute of electricity for lighting and cooling, so price becomes insensitive here. The findings indicate that correlation between electricity consumption and population is high.

Kırkgül estimated electricity consumption between 1992 and 2000 using average percentage increase, macro approach according to GDP and compared the results with the MAED model, which is used by ETKB. The first method is not appropriate for long term forecasting as variation increases each year. The second method predicts a coefficient for the relationship between GDP and consumption and estimates the consumption using this coefficient. Both method ignores the changes in technology, changes in habits and sets a certain ratio for demand forecasting which gives misleading results in the long run.

Similar to Kırkgül (1993), Şengün (1994) examined yearly electricity demand in two subgroups; household and industry using economic and social data such as GDP, population, electricity prices and production of home appliances for years 1950-1991. Additional variables like marriage rate, price of substitutes (petroleum, coal, natural gas) and price and storage of house appliances are thought to be used in the model but they are ignored due to the lack of data. After comparing both linear and exponential equation systems and regressions, exponential demand equations are found to be superior and used as the appropriate tool in her study. The results of the forecast are analyzed using t, F and Durbin Watson test statistics. The projection period consisted 1992-2000 being 40879.95 MWh for households 86896.80 MWh for industry in 2000. The outcomes also suggested both industry and household electricity consumption has high GDP elasticity whereas price elasticity is very low.

Özdoğan and Arıkol (1994) presented a more detailed study about the fuel and electricity consumption in sectoral basis (textile, food and cement sectors) considering process, space and water heating, illumination and power requirements in each sector using a bottom-up approach. After calculating the specific electricity consumption (e.g. for illumination), the results are used to estimate total electricity demand in each sector. However, they focused on analyzing the utilization of fuel in each sector, not on energy demand projection. It was found out that fuel-based energy usage decreases respectively at food, cement and textile sectors and space heating requirements decrease from textile to food and cement. Water heating for domestic purposes is lower. Motive power usage increases from food to cement and textile sectors. Illumination is much more in textile because of long labor-intensive hours.

Bakırtaş et al (2000) investigated the possible causality between demand and income and found out that income, electricity consumption, GNP, population, consumer price index, electricity price and income are important when elasticities are considered. However, electricity consumption also depends on urbanization, life styles and quality of life. Additionally, they estimated electricity consumption for the years 1997-2010 using Box-Jenkins methods.

Apart from other studies in Turkey, Tak (2002) forecasted both the total yearly electricity demand in Turkey and the demand in subgroups including industry, household, commercial sites, government offices and general lighting using error correction model technique. The yearly data used consisted of electricity consumption, electricity prices, wholesale price index, GNP, population, total production amount of household appliances and number of buildings between the years 1970 and 2000. In general, the estimating procedure included searching for the stationarity of the time series data using ADF unit root test, setting long run relationship by using least squares method, cointegration analysis using Engle Granger procedure and error correction models. The variables used to estimate total electricity demand were consumption per capita (C), GDP per capita (G), average real electricity price (P), trend (T) and dummy (D) for the years 1994 and 1999.
The cointegration equation for the variables was:

$$\ln(C) = -13.363 - 0.221 \ln(P) - 1.810 \ln(G) + 0.273 \ln(T) + 0.123 D + U_t$$

and the error correction model for the total energy demand was:

$$dln(C) = 0.036 + 0.632 \, dln(G) + 0.237 \, dln(C(-1)) - 0.189 \, U(-1) + \varepsilon_t$$

where U is the cointegration variable and ε_t are errors.

The results indicated that long run price elasticity of electricity demand was -0.22 which is as expected and long run income elasticity was 1.81. Price is not elastic in the short term but GDP elasticity is 0.63. In addition to that, each year, %18.9 of the long term deviations are corrected by error correction model. Electricity demand in the subgroups is estimated similarly. As for the industrial electricity demand model, electricity consumption in industry is taken as the dependent variable, and industrial electricity price and GDP for industry are taken as explanatory variables. To estimate electricity consumption in commercial sites, electricity consumption, real electricity price, GDP for commercial sites and trend are used. For households, electricity consumption per capita, GDP per capita, average electricity price in households, manufacture of household appliances per capita are used as variables. For Government offices, electricity consumption in Government offices and GDP are used and finally to estimate electricity consumption for lighting, demand for lighting and number of buildings are considered. After modeling stage, projections for years 2001 – 2005 were done by using three development scenarios being low, medium and high development and according to the results, total electricity consumption in Turkey is estimated to increase from 98295 GWh in 2000 to 127032 GWh in 2005 with low development scenario; to 138169 GWh in 2005 with mid-term development scenario and to 150.10 GWh in 2005 with high development scenario. The insufficient side of the model was the ignorance of cross effects of the variables in the model.

As most of the previous studies on electricity demand, Çakır (2002) utilized ARIMA (p,d,q) and exponential smoothing methods to forecast total electricity demand in Turkey. In the study, electricity consumption data between 1970 and 1998 and E-Views 3.1 software are used to test stationarity of the series. All of the models are tested for break point efficiency. According to break point efficiency, the most successful model was Brown's exponential smoothing method.

One of the improvements in forecasting was ANN in the last ten years. Hamzaçebi and Kutay (2004) used ANN to estimate yearly electricity demand in Turkey. Neural networks gave better results compared to regression and Box-Jenkins methods. The advantage of ANN is that they allow analyzing the forecast with smaller number of data.

In 2006, Erdoğdu utilized cointegration technique and partial adjustment model which is the dynamic version of reduced form model to investigate the relationship between real electricity prices, real GDP per capita and electricity consumption (2006b). The data was quarterly time series data for years 1984 - 2004. The appropriate model is found to be:

$\ln E_t = \text{Ø}_0 + \text{Ø}_1 \ln P_t + \text{Ø}_2 \ln Y_t + \text{Ø}_3 \ln P_{t-2} + \text{Ø}_4 t + \text{Ø}_5 \ln E_{t-2} + \varepsilon_t$

where E_t is the electricity consumption per capita at time t

P_t is the price of electricity at time t

Y_t is the GDP per capita at time t

t is the time variable

The short and run long elasticities obtained from the model are given below:

	Short run	Long run
Price elasticity :	-0.041	-0.297
Income elasticity :	0.057	0.414

In the second stage of his study, Erdoğdu forecasted yearly electricity demand in Turkey for 2005-2014 using yearly electricity consumption data covering the period 1923-2004. The model is developed using ARIMA methodology. The results showed forecasted net electricity consumption will be 129311

MWh in 2005, 155667 MWH in 2010 and 160090 in 2015. Net electricity consumption value in 2005 was 130262.9 which show the percentage error is %0.7.

Ediger and Akar (2006) estimated energy demand in Turkey, using ARIMA and SARIMA methods for the yearly time series data between 1950 and 2003. Total energy demands including the fuel types; hard coal, lignite, petrocoke, wood, animal and plant remains, oil, asphaltite, natural gas, hydropower, electricity and solar is modeled. Seasonal components are added to some of the equations, forming a seasonal ARIMA model. The outcomes illustrated that demand of each type of fuel would increase in the next 15 years except wood. Share of fossil fuels in total energy demand will increase also, however the highest rate of increase will be in natural gas as expected.

Recently Sanlı (2007), a member of World Energy Council Turkish National Committee, is working on a web-based electricity demand forecasting analyze program. The structure of the program is given in Figure 3.6. The first module includes graphical representations of 6 time series: GDP, economic growth, net electricity consumption, gross electricity consumption, population and installed capacity from 1970 to 2005.

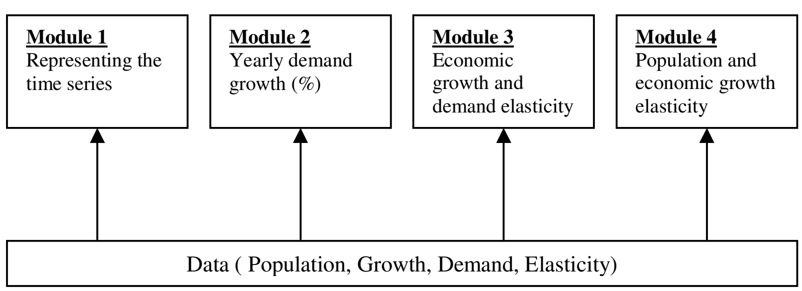

Figure 3.6 – Structure of the Program (Sanlı, 2007)

In the second module, demand is analyzed by entering the yearly demand growth rate according to the formula:

This year's demand = Previous year's demand * (1 +Growth Rate % / 100)

Third module involves economic growth rate and demand growth/economic growth ratio instead of demand growth rate. The fourth module has 2 parts; templates and options. The demand is a function of GDP, population and demand for previous period and is defined as:

$e^y=s.e^a.e^b.e^c$

Halıcıoğlu (2007) used real income per capita, real residential electricity price, electricity consumption and urbanization rate as variables in her study. She used bounds testing and ADL approach because of having small sample size, endogenity problems and for its advantage of being irrespective whether the regressors are I(0) or I(1). Annual data covering the period 1968 and 2005 is used and all the variables are found to be I(1). In the first step of model formation, an ADL model is formed and lag number is chosen as 2. Next, bounds F test is applied and the results showed there exist long run relationship among the variables, then ADL model with error correction term is estimated. Income, urbanization and price elasticity are found to be 0.70, 1.34 and -0.52, respectively, in the long run.

There are several other researches that do not consider the electricity demand forecast directly but that are interested in the causality relationship between electricity consumption and GDP, temperature and price. The direction of the relationship between GDP and consumption, defined as causality relationship has always been a research issue. Effect of temperature changes to electricity load in Turkey is investigated in 2005 by Gölbaşı MYTM. The method followed was finding monthly weighted average temperature for 65 provinces for every hour using electricity consumption and temperature, and then weighted average temperature for Turkey. Finally, for each month of 2005, the effect of weighted average temperature on consumption is calculated. The results showed that load changes are really high in February and March. Moreover, temperature increase causes increase in load from May till October. In other months increase in temperature causes decrease in load (MYTM, 2005).

3.4 ELECTRICITY DEMAND FORECASTING STUDIES IN THE WORLD

Energy forecasts form a basis for financial planning, scheduling and investing. Thus, in many countries in the world and especially in USA, electricity demand forecasts are important for defining the policies of the countries.

Houthakker (1951) used double logarithmic model in his residential demand forecast study for UK. The model covered average income per household, marginal price of electricity, marginal price of gas, average holding of heavy domestic equipment and annual consumption per customer. The model showed that income elasticity is higher compared to price elasticity and cross price elasticity with values 1.17, -0.89 and 0.21, respectively.

Fisher and Keisen (1962) were the first to distinguish between short and long run residential electricity demand using adjustment model. In the short run, the income elasticity in rural areas was low or negative; on the other hand, a rise in income caused increase in electricity consumption in urban areas. In the long run, the variables that have effects on consumption were changes in population, number of households, changes in income, changes in stock appliances, but not the price of electricity and household appliances.

Baxter and Rees (1968) derived three types of models for industrial electricity demand; first being the general form, where the inputs are labor, capital, oil, gas, coal and electricity. The second linear model considered the changes in fuel technology and used trend as an exogenous variable. Finally, the third model is formed with the assumption that electricity consumption is directly related to the output. The results are investigated and it is found that, electricity demand is responsive to the changes in output and the fuel technology, however it is not responsive to price as expected.

Wilson (1971) studied residential demand of electricity and used a regression model consisting of average electricity consumption per household, price of electricity, average price of natural gas, median family income, number of degree days and number of rooms per household. Elasticities are estimated in the long run and the results differed from Fisher and Kaysen's and Baxter and Rees' founding in that price significantly affected consumption.

Different from other studies, Anderson (1973) included various substitutes of electricity, such as coal, oil and natural gas in his double logarithmic model. Anderson also estimated industrial energy demand in USA with an adjustment model using price of coal, coke, oil, electricity and average wage rate of production workers.

Mount et al (1973) modeled residential, commercial and industrial electricity demand both for long and short run by an adjustment model which includes lagged dependent variable, price, population, temperature and income. Long run results show that demand is price elastic for all three type customers. Income is inelastic for residential and industrial demand and elastic for commercial demand. Population shows evidence of unit elasticity. Gas and appliance prices are found to be inelastic.

Another research about residential, industrial and commercial demand was performed by Murray et al (1978) for different regions of Virginia. The regression data included monthly consumption data by customer type, yearly income data, which was interpolated in the same way as employment data, and yearly population data that was interpolated linearly, monthly temperature, electricity and alternative fuel prices between January 1958 and December 1973. The outcomes indicated that alternative fuel prices have effect on electricity consumption and there is a price and income elasticity of electricity demand and rise in real income and electricity price will lead to deterioration of load factors of electric utilities.

Analogous to previous studies, disaggregation and regression methods are used by Skinner (1984), who collected electricity consumption data by metering system, bills, questionnaire and by placing data recorders to different kinds of customers. Domestic sector sales are disaggregated as refrigerators, cooking, space heating and water heating. Industrial sector is broken down by industry type such as iron, steel, food and drink. Commercial sector is broken down as shops, public buildings and offices. Load curves are expresses as demand estimation coefficients. Then these demand estimate coefficients are multiplied by appropriate population consumption and estimates of total consumption are obtained.

In Republic of Ireland, the model is based on forecasting the extra electricity that will be sold next year. Simple regression models that get total, extra domestic and extra non-domestic electricity sales as variables are used. The increment in domestic sector is found by personal consumption of goods and services, whereas extra non-domestic electricity sales are found by GDP. Forecasts are done by Transmission System Operators (TSO). The advantage of the model is its simplicity and compatibility with small data set (TSO of Ireland, 2002).

Barakat et al (1990) applied three time series techniques that are decomposition, SARIMA and Winter's Exponential Smoothing methods to calculate peak load demand in utilities. In the decomposition methods, time series is decomposed as cyclical variation, seasonality and trend.

SARIMA model for monthly data is expressed as:

$$(1-\phi_1 B)(1-\phi_1 B^{12})(1-B)(1-B^{12})LOG(Z_t) = (1-\varepsilon_1 B - \varepsilon_2 B^2)(1-\phi_1 B^{12})e_t$$

where:

B: backward shift operator

Z_t : time series of monthly peak demand

e_t: random error term

The remaining parameters are constant coefficients.

Seasonal time series can be analyzed by Winter's method and this method is based on smoothing equations for trend, seasonality and irregularity. The equations are shown below:

$$(SAV)_t = \alpha \frac{(Z_t)}{I_{t-s}} + (1-\alpha)\left[(SAV)_{t-1} + T_{t-1}\right]$$

$$Tt = \gamma\left[(SAV)_t - (SAV)_{t-1}\right] + (1-\gamma)T_{t-1}$$

$$I_t = \beta \frac{(Z_t)}{(SAV)_t} + (1-\beta)I_{t-s}$$

where

Zt : Data values which contain seasonality

$(SAV)_t$: Smoothed average value of series that do not contain seasonality

I_t: seasonal adjustment factor

T_t: trend component

S: length of seasonality

α, β, γ : Smoothing parameters

Monthly consumption data between 1974 and 1987 is used to estimate peak demand between 1981 and 1987 with two years periods (1974-1980 data to estimate 1981 and 1982, 1974-1981 data to estimate 1982-1983) and the results are compared with the actual values. Although decomposition method performed better than other two methods in predicting the annual peak demand in the first years, Winter's method performance was better between 1984-1987, which shows none of the methodologies were better and stable.

Silk and Joutz (1997) included electricity appliance stock in addition to the mostly used temperature, price and income variables in their residential electricity demand study for US. Additionally, fuel oil price is included in the model. They used cointegration and error correction method for being advantageous in evaluating short and long run relations and elasticity. The data included yearly data from 1949 to 1993. The cointegration equation showed the price and income elasticity of demand was nearly 0.50, which are small compared to other studies and which can be a result of using stock variable in the equation. On the other hand, short run elasticities were about 0.25. This difference is a result of the changes in the long run and short run behaviors of prices. The estimates were higher compared to other estimates done by US energy agencies.

Wong and Rad (1998) forecasted electricity demand of Hong Kong using Box-Jenkins methodology and SPSS statistical package. Yearly electricity consumption series is differenced two times to make the series stationary. By analyzing the ACF and PACF graphs, the model is found to be ARIMA(2,2,0).

Since 2000's, neural networks based methods took their place in electricity demand forecasting studies. Charytoniuk and Chen (2000) used a neural network based model instead of using traditional methods because of the dynamic environment in electricity sector and the difficulty to obtain past data. The total demand is calculated by summing up the demand of several custom groups. The inputs used in the model included demand affecting factors like time of the day and weather variables. Secondly the neural network architecture is decided and multilayer feed forward neural network are chosen because of their good performance. Accuracy of the forecasts in neural networks depends on forecasted customer characteristics, group size, customer classification system and the degree of the survey.

Christodoulakis et al (2000) separated the whole energy demand both as sectoral and energy type. They used data between 1974-1994 to estimate liquid fuels, electricity and solid fuels consumption in households, transportation, industry and commercial sectors. Finally, they used the outputs to search CO_2 emission in Greece. The total electricity demand was a function of electricity price and sectoral outputs. Two stage error correction models is used for electricity demand forecast, first stage being an equation in levels for the long run relationship and second stage being a short run equation in differences. The own price elasticity in the short run was 0.14. However, it was insignificant in the long run equation.

Larsen and Nesbakken (2002) studied end-use consumption in Norwegian Households by an engineering model named ERAD and an econometric approach. ERAD has technical knowledge on building structure of the houses, thus the requirement of electricity for heating, illumination, hot water and appliances. For instance, the space-heating requirement is calculated as a function of insulation standard for windows, walls, roofs and floors, indoor and outdoor temperature etc. After calculating the entire requirement, total electricity demand is calculated by summing these values. Secondly, the econometric model included heating degree days, household characteristics and electricity prices as variables. This model is better than ERAD in that end user parameters are estimated directly instead of making assumptions.

Engineering model estimates gave higher results in space heating, water heating and cooking compared to the econometric model. On the other hand econometric model gave higher results for lighting, washing and drying. Therefore, the results from both models were considerably different and it is not possible to say that any of the models is advantageous than the other one.

Kamerschen and Porter (2004) estimated residential, industrial and total electricity demand by partial adjustment model, which is adopted by Houthakker and Taylor (1970) and simultaneous equation model, which is adopted by Halvorsen (1975). The variables for both residential and industrial electricity demand included annual residential electricity sales per customer, real expected marginal price of residential electricity; real annual GDP, real price of residential natural gas, Heating Degree Day (HDD) and Cooling Degree Day (CDD). They estimated 3 different models in each sector, first considering HDD, then CDD and finally using both HDD and CDD. The results show that temperature plays a more significant role in residential electricity consumption and price is more elastic in residential sector as well. Residential price elasticity changes between -0.85 and -0.94. On the other hand, industrial demand elasticity is between -0.34 and -0.55. Additionally, simultaneous equation model gave more appropriate results compared to flow adjustment method.

Fortum, a Swedish electricity generation company, made a short term demand forecast study in Norway, Sweden, Finland and Denmark for 10 days in 2004 (Laukkanen, 2004). Different from other studies with Box-Jenkins model they used weather as external variables and considered the special days such as bank holidays, Christmas and football matches. The most important variable that affects electricity demand in Nordic countries is temperature. Especially in the winter, demand for electricity is very high as these countries are located near to the North Pole. Additionally, industrial production is important in those countries. To illustrate, paper industry consumes a big share and when they are in holiday in winter electricity consumption decrease significantly. Special days such as Independence Day or weekends are important and electricity consumption decrease in these days.

Mamun and Nagasaka (2004) used an ANN method in their long term demand forecast study for Japan, which is called Radial Basis Function Networks (RBFN)[3]. GDP, population, summer and winter days, index of industrial production, oil price, electricity price and maximum electric power are used as inputs in the neural network. Also contribution factor, which is the sum of the absolute values of the weights of an input, is used to determine the influence level of each input. RBFN included three layers which are input layer where there is source nodes to connect the network to its environment, hidden layer where there a hidden units that provide a set of basis function and output layer which is a linear combination of hidden function. Economic factors were forecasted yearly in addition to the electricity demand forecast with the model for years 2001 and 2010. RBFN gave more accurate results compared to other studies (96.5 % accuracy). Loads in Japan are estimated to increase %1.39 per annum till 2015.

Thang (2004) used neural fuzzy approaches to estimate short term load forecast of Australia. The forecasting is done hourly for the next day. The feedforward backpropagation neural-network is used as the suitable type of model since it is good at modelling complex and non-linear relationship between inputs and outputs. Separate models are derived for weekend and weekdays as electricity demand is low in weekends. Additionally, holidays are removed from the model to get more precise results. The outcomes show the model that uses just the electricity consumption data could meet the demands and can produce good forecasts.

One of the few studies that used monthly data to estimate electricity demand was Hondroyiannis (2004). The approach is said to differ from other studies for using weighted monthly average weather variable. The income, price and weather variables are found to be integrated of order 1. Cointegration is tested using Johansen maximum likelihood approach and VAR is used to decide the appropriate lag length. Two dummy variables are included in the model, one for the reduction of VAT from %18 to %8 starting from January 1999 and second for Christmas Holidays. The resulting cointegration equation showed that income, price and temperature elasticity were 1.56, -0.41 and -0.19, respectively. Finally Vector Error Correction Model is estimated and variables of this equation are established as the 4^{th} and 12^{th} lag of the electricity consumption and income variables, where the sign of 4^{th} lag is negative and others are positive. The main finding of this study was that temperature is significant in the long run, but not in the short run.

Away from other studies, which make forecasts for the long term, McSharry et al (2005) predicted peak demand and used daily electricity consumption data in Netherlands between 1991 and 2000. The consumption is assumed to be affected by time, special days and weather conditions as in other studies. Temperature, wind speed, luminosity and cooling power are used to define weather variables. The demand is thought to change quadratically with time. The model has various advantages such as it is capable of taking into account various variables like day of week, seasonality and weather variables and simulated probabilistic weather data can be used when there is not enough data available.

Narayan and Smyth (2005) used bounds testing approach to cointegration and Autoregressive Distribution Lag Model to estimate yearly electricity demand in Australia. Two models are formed with income, temperature, electricity price and natural gas price data. In the first model price data is used separately and in the second model the ratio of electricity price to natural gas is used. A long run relationship is found and short and long run coefficients are estimated since the variables are found to be cointegrated. The adjustment coefficients are significant in both of the model with values 0.03 and 0.10, approving long run relation. Income elasticity of the models is 0.323 and 0.408 respectively. The electricity price elasticity is -0.541 in the long run whereas it is -0.263 in the short run in the first model. Natural gas price is insignificant in the first model, but in the second model price ratio is significant. Temperature has positive sign and significant only in the first model.

As discussed, cointegration techniques became popular recently. Mozumder and Marathe (2005) used electricity consumption per capita and GDP per capita for 1971-1999 in their research. Testing the stationarity of the series by ADF test and finding cointegrating relationship between the two variables by Johansen trace test statistics, VECM is written as:

$\Delta PCGDP_t = \alpha + \Sigma\beta_i \, \Delta \, PCGDP_{\,t-i} + \Sigma\gamma_i \, \Delta \, PCEC_{\,t-i} + \delta\varepsilon_{\,t-1} + \mu_t$

$\Delta PCEC_t = \alpha + \Sigma\beta_i \, \Delta \, PCEC_{\,t-i} + \Sigma\gamma_i \, \Delta \, PCGDP_{\,t-i} + \delta\varepsilon_{\,t-1} + \mu_t$

PCGDP: per capita GDP

PCEC: per capita electricity consumption

After estimating VEC model, Granger causality test results indicated that PCGDP granger causes PCEC but not vice versa.

Hor (2005) questioned the impact of weather on electricity demand. Electricity demand increases both in winter and summer due to increase in use of heaters and lighting or coolers and air conditioning. A simple linear regression model is formed as;

$E = \beta_0 + \beta_1 T_{mm}$

Where E is the predicted electricity demand, β_0 and β_1 are constants; T_{mm} is mean monthly temperature in the model. The results showed that sensitivity to temperature is more significant in the spring and autumn compared to the winter and summer. This unexpected result can be because of the fact that lighting load does not change even at lower temperatures and in the winter there is a base level of comfort and increase in heating is not a linear function any more.

Taylor et al (2006) compared six univariate methods which are: multiplicative seasonal ARIMA, exponential smoothing, ANN, principal component analysis (PCA) approach, which resembles decomposition and regression, a seasonal version of the standard naive random walk benchmark model and a benchmark. 30 weeks of hourly data is used in Rio and 30 weeks of hourly data is used in England and Wales. To get smoother series, extreme values in special days, like holidays or religious days were removed. In the first model, autocorrelation is eliminated utilizing the ACF and PACF and the best model is chosen using Schwarz Information Criterion. Second, exponential smoothing method is applied, which is known for its robustness and simplicity. Third, a type of neural networks, single hidden layer feedforward network, that are usually preferred in estimating electricity demand, is used. This network has k inputs that are connected to each of m units in a single hidden layer, which is connected to an output. The fourth model, PCA aims to decrease the dimension of variables, which are highly correlated, to a smaller set of data. The simplistic benchmark model is poor in that it cannot perform well when long periods will be estimated, so this method is used to form a base for the benchmark model. The outputs showed that naive benchmark method is the best among others. Unlike other studies about load demand, the neural network showed a poor performance. The exponential smoothing method is superior to ARIMA and PCA methods

Song et al (2006) used integrated approach to forecast electricity demand combining fuzzy approach for weekends where the demand is low, with exponential smoothing for weekdays. This hybrid algorithm estimates 24 hours load demand according to type of day. Additionally, weather impact is considered and integrated to the model. For instance, exponential smoothing method is applied in spring, winter and fall, where electricity consumption is highly dependent to temperature. In the summer months again fuzzy approach is used to estimate demand. The outcomes showed that both the forecast errors and mean absolute percentage errors are low showing the model is accurate enough. This method is new in that it used separate load forecasting methods during a week.

As it can be seen various methods are used to estimate electricity demand in different countries. One of the most significant finding is that most of the studies consider long time horizons and make long-term forecasts. However, short and medium term forecasts are as important as long term forecasts because electricity market is being privatized and liberalized in many countries. Prices are set hourly and planning and scheduling activities gained more weight as of supply security and excess demand issues. So, enough consideration should be given to medium and short term forecasts.

CHAPTER 4

MODELLING MONTHLY ELECTRICITY DEMAND IN TURKEY FOR 1990-2006

In this study, cointegration and error correction model is used to estimate monthly electricity demand in Turkey for being an econometric method that considers social, demographic and economic variables instead of just using the dependent variable and its lagged values as time series methods, which ignores the causal relationship between the estimated variable and other econometric variables. With this method, both long and short run relationship between the variables are investigated. Additionally, time series are generally nonstationary and cointegration eliminates the loss of generality which is caused by differencing because if all or a subset of the variables are I(1), there may exist a linear combination of the variables which is stationary, I(0). Another method, bottom up approach is not preferred also because of the requirement of a detailed data set. On the other hand, cointegration and ECM is easy to use and interpret. The only disadvantage of ECM is that it ignores mutual relationship between the variables. Although it is believed that VAR and VECM perform better in this kind of situations, the performance of VAR methodology is still discussed.

The model is constructed with monthly data because forecast periods are usually yearly and there are only a few studies in Turkey for medium term electricity demand forecast. However, the importance of mid-term forecast cannot be disregarded. Private sector companies need these forecasts to plan energy transactions, determine the amount of production and the price to offer for electricity in the following months. If the predictions state that the electricity demand of the following month will be high, the electricity generators may offer higher prices for the electricity they sell, as it is certain that they will sell the electricity they produced. Good forecasts are also necessary for revenue management and financial planning. Public sector companies need good forecast results to stabilize the supply and demand of electricity and avoid inevitable electricity cut-offs or excess electricity. Because the Government still owns %50 of the electricity generation and state owned electricity generation company, EUAS produces the biggest part, it can regulate the production amount of the following months. Furthermore, network operators benefit from these forecasts to support investment decisions, allow smaller margins in the network and reduce risk. In brief, good forecasts are necessary for all of the market players since high forecast errors may result in constructing too many and unnecessary units and high operating costs.

Former studies mostly include income and price variables in their demand estimating equations. Although electricity is a resource that cannot be replaced with another energy resource, substitutes of electricity should not be ignored while performing electricity demand forecasts. In USA natural gas has started to be used as a clean and cheap resource in 1980's, which later caused an increase in natural gas prices. Usage of natural gas in households increased recently in Turkey after 1990's for cooking and hot water during the year and heating in winter. In a liberalized market, an increase in natural gas prices should lead to increase in demand for electricity. Thus, well-known substitutes of electricity such as natural gas, petroleum and coal should be considered as much as other explanatory variables like electricity price, income or population in a model.

The data used in this study covered monthly electricity consumption, natural gas consumption, natural gas price, electricity price, population, consumer price index, industry production index, temperature, heating degree days and cooling degree days. Any substitute of electricity such as coal or oil is not included in this study because monthly data is not available for these substitutes. Additionally, just 7 years data from 2000 to 2006 is available for natural gas consumption and price which is not a long period for seasonal data. Thus, instead of modeling demand with a small data set (of sample size 72) for the years 2000-2006, effect of natural gas price over electricity price ratio and natural gas consumption to electricity consumption is investigated with a simple linear regression model in the first step to see the relationship between these two substitutes and search whether these variables are significant, between 2000 and 2006.

In the second step of this study electricity demand is modelled with the monthly data between January 1990 and December 2005 without natural gas data. 2006 data of the variables is also available and is used to compare the actual and forecasted values. Electricity consumption per capita, real electricity price, industry production index, average temperature, HDD and CDD are used as dependent and independent variables. The variables are all in logarithm form to investigate the elasticities directly.

Monthly natural gas consumption data (kWh) between years 2000 and 2006 was obtained from BOTAS in m^3. Then, it is converted to kWh as given below: 1 m^3 natural gas = 10.8 kWh

As monthly data before the year 2000 cannot be obtained, the data is restricted for 7 years. Consumption is divided by population to obtain natural gas consumption per capita.

Monthly natural gas prices (YTL/kWh) for years 2000-2006 were obtained from Turkish Statistical Institute (TUIK) as well as electricity price data.

Population (person): Yearly population data is obtained from TUIK. Data consisting years between 1983 and 2004 is obtained from the book "Statistical Indicators 1923-2004" published by TUIK (2005a), whereas years 2005 and 2006 are obtained from the web site of TUIK (2007c). If a population has a constant birth rate through time and is never limited by food or disease, it has exponential growth (Wikipedia, 2007c). Therefore it is assumed that population increases exponentially so yearly data was converted to monthly data by the given formula.

$P = P_o * \exp^{KT}$

where P: population estimated
P_o: initial size of the population
K : constant of growth rate
T : time

Monthly Electricity Consumption (kWh/capita): Monthly electricity consumption data between 1990 and 2006 is obtained from Gölbaşı National Load Dispatch Center (MYTM), a branch of Turkish Electricity Transmission Company, where operation of the power plants and electricity network in Turkey are watched and organized continuously, information consisting the region and the amount of electricity consumption, breakdowns in the network are recorded and hourly electricity demand forecast studies are performed. By dividing the monthly electricity consumption to monthly population, monthly electricity consumption per capita is obtained.

Consumption Price Index (TUFE): Monthly consumption price index is obtained from the website of TUIK (2007b). 1987 is taken as base year (1987=100). This index is used to convert the current electricity prices to real electricity prices and to remove the effect of inflation.

<u>Electricity Price (YTL/kWh):</u> Monthly electricity prices were obtained from Turkish Statistical Institute in Ankara. Electricity prices were converted to real prices using TUFE.

<u>Industry Production Index:</u> Monthly industry production index values between 1990 and 2006 are taken from the web pages of Turkish Central Bank (2007).

<u>Temperature (^{0}C):</u> Monthly temperature for 82 cities in Turkey is obtained from Turkish State Meteorology Institute in Ankara (2007). A weighted average temperature is calculated with the area of the cities in km^2 obtained from Turkey's Statistical Yearbook, 2004 published by TUIK (2005b).

<u>Heating Degree Variable :</u> Another variable that may effect electricity demand and used as an exogenous variable to estimate electricity demand in several studies is HDD, which reflects the energy needed to heat a building (house or business) and is calculated in daily basis (Wikipedia, 2007d). Calculation of HDD differs in different countries; however, Eurostat offered a common methodology to be used in European countries that is also used in this study (Turkish Meteorological Institute, 2006).

$$HDD = \left\{ \begin{array}{ll} (18^0\mathrm{C} - \mathrm{T_m})\ \mathrm{x\ d} & \mathrm{if\ T_m} \leq 15^0\mathrm{C} \\ 0, & \mathrm{else} \end{array} \right\}$$

where 15^{0}C is the heating threshold, d is the number of days. Monthly and yearly values are found by summing the daily values.

However, in this study, only monthly temperatures were available. Therefore, instead of using daily data and then adding daily values to obtain monthly data, monthly data is calculated directly by the below formula, which is similar to the one for HDD.

$$HDV = \begin{cases} (18^0\mathrm{C} - \mathrm{T_m}) & \text{if } \mathrm{T_m} \leq 15^0\mathrm{C} \\ 0, & \text{else} \end{cases}$$

where T_m is the monthly temperature data.

<u>Cooling Degree Variable</u>: CDD reflects the energy required to cool a building and calculated daily like HDD. CDD is generally used in civil sector to calculate the energy demand during construction and for design issues. There is not a formal method to calculate the CDD, but mostly used method is as follows:

$$CDD = \begin{cases} (\mathrm{T_m} - 22^0\mathrm{C}) \text{ x d} & \text{if } \mathrm{T_m} \geq 22^0\mathrm{C} \\ 0, & \text{else} \end{cases}$$

where 22^0C is the cooling threshold, d is the number of days. Monthly and yearly values are found by adding the daily values.

However, in this study, only monthly temperatures were available. Therefore, instead of using daily data and then adding daily values to obtain monthly data, monthly data is calculated directly by the below formula, which is similar to the one for CDD.

$$CDV = \begin{cases} (\mathrm{T_m} - 22^0\mathrm{C}) & \text{if } \mathrm{T_m} \geq 22^0\mathrm{C} \\ 0, & \text{else} \end{cases}$$

where T_m is the monthly temperature data.

Graphs of the series are given in Figures 4.1 – 4.8.

Natural gas price over electricity price ratio does not seem to have a trend or seasonal pattern. Recently, gas prices increased rapidly whereas electricity prices remained the same due to political reasons. This explains the increase in this ratio during the previous years. On the other hand, petrol prices increased from 20$ to 100$ per barrel recently, which also cause increase in natural gas prices.

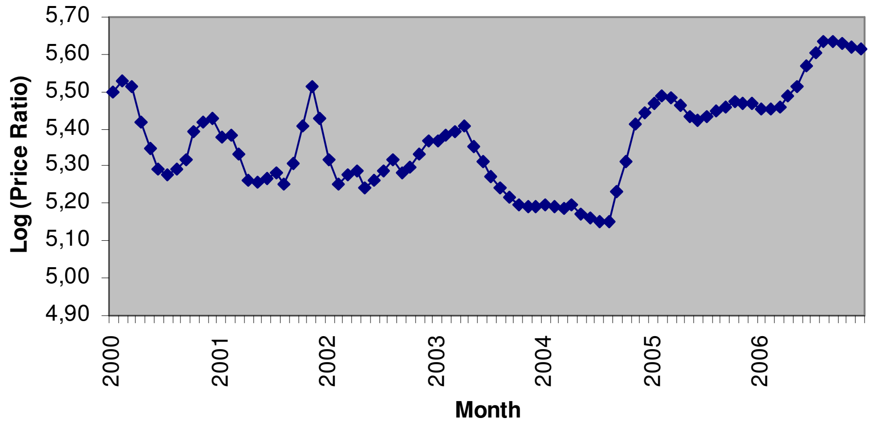

Figure 4.1 – Log (Natural Gas Price/Electricity Price)

It seems that there exists a slow upward trend and seasonal pattern in natural gas consumption data. Natural gas is especially used in winter for heating purposes since 1990's. In the summer electricity is preferred for cooling purposes and for air conditioners. The use of natural gas is increasing year by year as natural gas becomes available in more cities in Turkey.

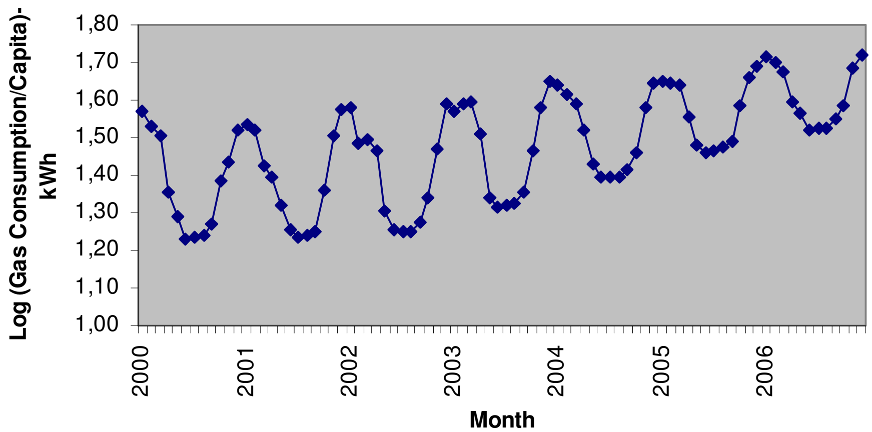

Figure 4.2 – Log (Natural Gas Consumption/Capita)

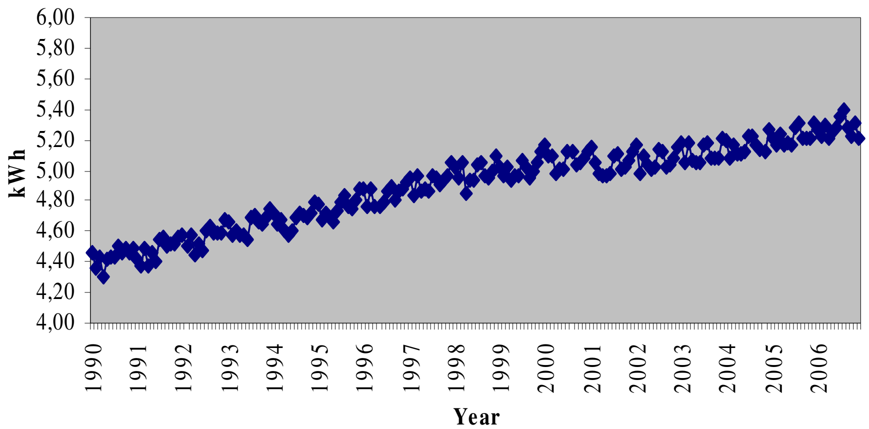

Figure 4.3 – Log (Electricity Consumption / Capita)

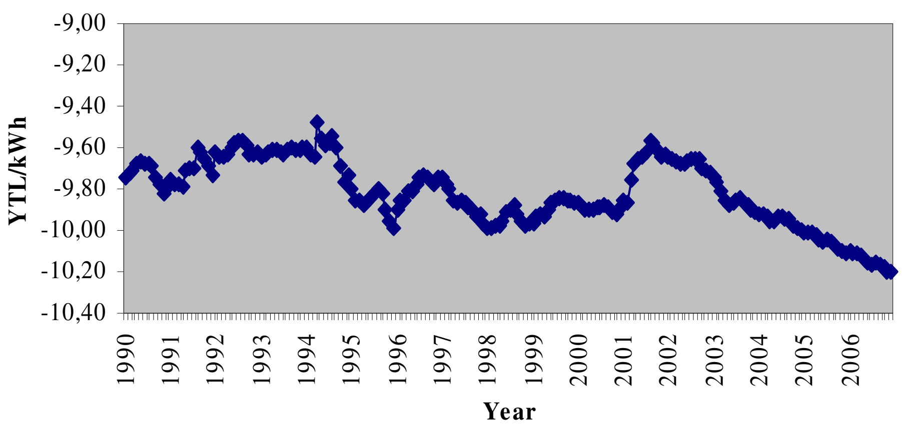

Figure 4.4 – Log (Real Electricity Price)

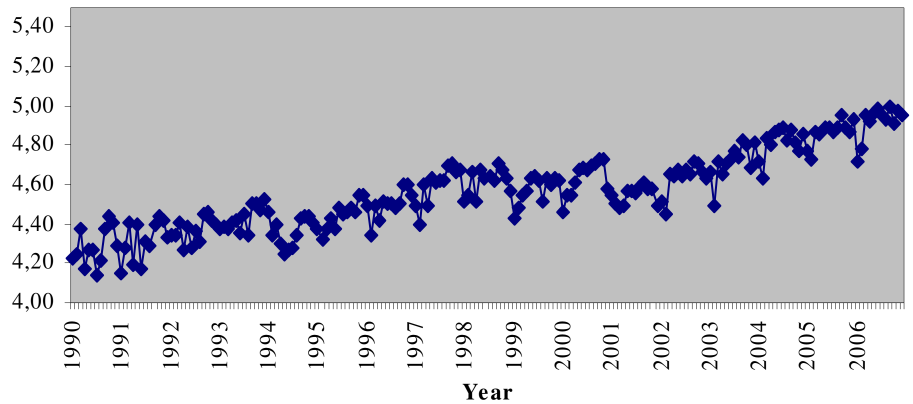

Figure 4.5 – Log(Industry Production Index)

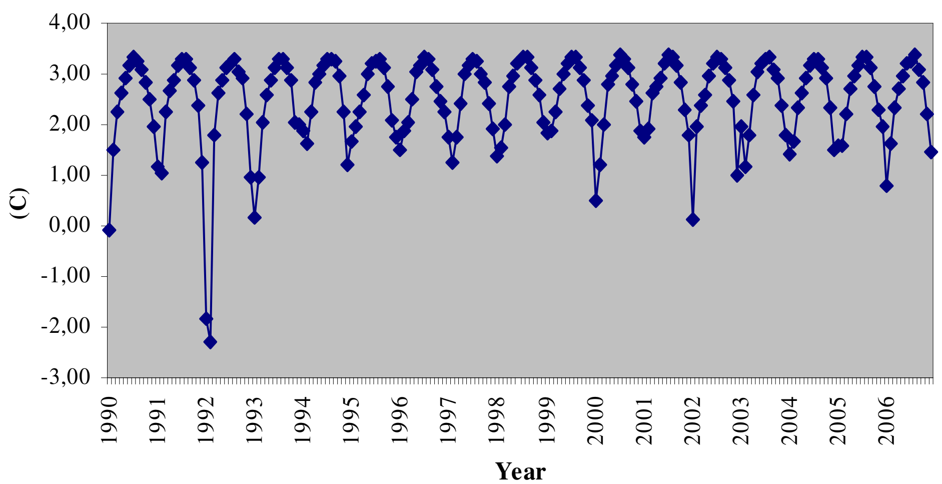

Figure 4.6 – Log (Temperature)

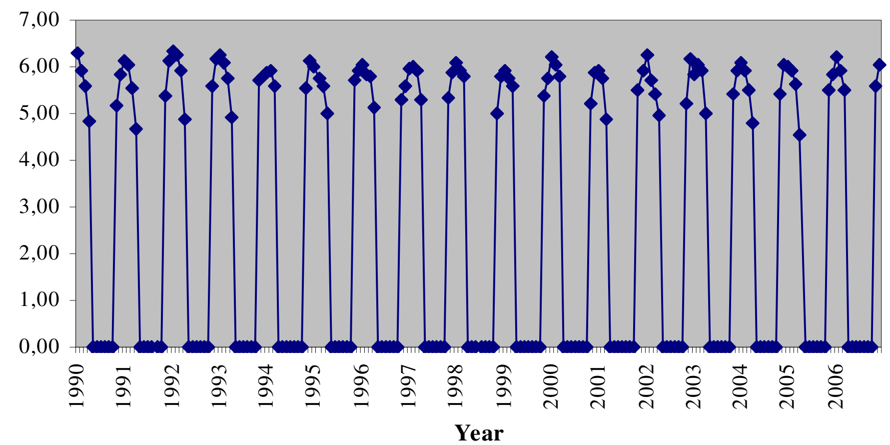

Figure 4.7 – Log (HDV)

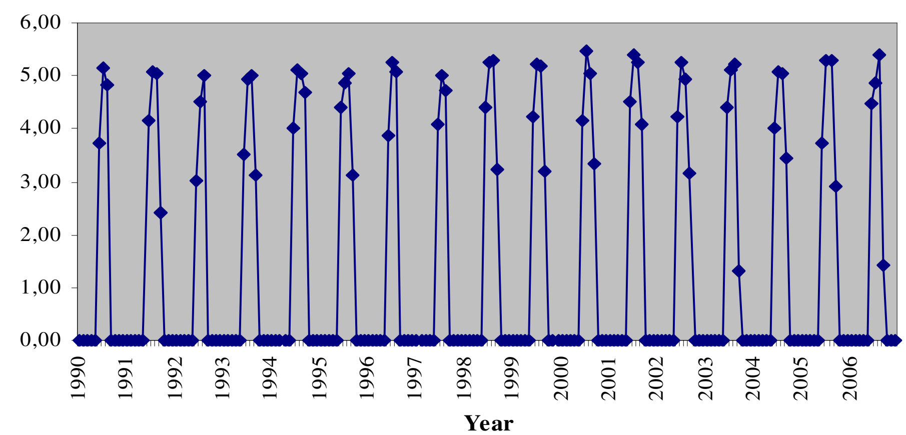

Figure 4.8 – Log (CDV)

Logarithms of the data are taken to investigate elasticities. However, it is known that ln(x) is defined only if x>0 and both heating and cooling degree days may have 0 values in a year. The number of heating degree days in summer is zero and

number of cooling degree days in the winter is zero. So to overcome this problem, 1 is added to HDV and CDV data for each month and then logarithm is calculated.

There is an upward trend in electricity consumption and industry production index series. On the other hand electricity price graph indicates a downward trend in last years that is a result of Governmental policy since electricity price did not change for 5 years. Temperature, HDV and CDV has seasonal patterns, which are expected.

4.1 –NATURAL GAS AND ELECTRICITY AS SUBSTITUTES

The number of electricity demand forecasting studies that take natural gas consumption or natural gas price as an explanatory variable is notably low compared to the ones that take income and price as variables. Walasek, (1981) used natural gas price as an explanatory variable in his flow adjustment model of demand in US. The results indicated that small changes in electricity and natural gas price may effect electricity consumption in an unexpected way in some small regions. This can be a result of availability of natural gas in some regions and a misspecification problem. The results also indicate that both short and long run elasticities of natural gas price is much lower compared to elasticities for electricity price and income. Barnett et al (2005) regressed electricity price and natural gas price, income, population, time and manufacturing output on electricity demand in their multi market model for United States to estimate electricity demand, supposing equal signs for electricity consumption and natural gas prices as they are substitutes. However, the results showed gas price does not really affect electricity demand, so it is not an effective substitute. On the other hand, Dunlap (2006) defined natural gas as an effective substitute for electricity in winter for heating purposes. However, use of natural gas in household decrease in warmer months in a year. In fact, most of the work about price elasticity of natural gas and electricity is done with yearly data to avoid seasonal effects. Dunlap estimated a log-log model with two stages least squares regression using monthly data from January 1990 to May 2005. The variables were price of electricity and gas and dummy variables. The model analyzed the seasonal effect for natural gas by using two variables for natural gas price; being nonwinter and winter price. The results demonstrated that substitution effect in the winter was considerable according to other months.

4.1.1 Stationarity

While developing a time series model, it should be known that whether the stochastic process obtained changes depending on time or not. If the series is not stationary, it will not be suitable to write a simple model expressing the past and future structure of the series because autocorrelations deviate from 0 and spurious patterns occur (Kutlar, 2000, p.12, 153-157). So, it is required to test the stationarity of the series before forming a forecast model.

As data is available just for 7 years and amount of data is limited, ADF test is used to search for unit roots of monthly time series data. The null hypothesis that the series has unit root is tested against the alternative that there is no unit root in the series, which is indicated as:

H_o: series has a unit root

H_a: series has no unit root

ADF test is applied including some explanatory variables as trend and intercept (stochastic and deterministic terms) and nonstationary series is differenced once to obtain stationary series. It can be concluded that gas consumption and gas price/electricity price ratio is I(1) and electricity consumption does not have unit root, as given in Table 4.1.

Table 4.1 – ADF Test Statistics

	Log (Electricity Consumption)	Log (Gas Consumption)	Log (Gas Price/Electricity Price)
t statistics	-4.754	-2.182	-1.371
Probability	0.001	0.492	0.058
t stat after differencing	-	-9.893	-3.620
Prob, after differencing	-	0.000	0.035
Order of Integration	I(0)	I(1)	I(1)

Critical values at 1%, 5% and 10% level are -4.089, -3.473 and -3.164 respectively,

4.1.2 Regression Model

After this procedure, regression equation is calculated using E-Views 4.1 package program. A simple regression equation is calculated in the beginning, which can be shown as:

$$CONS00 = 5.134 + 0.265*\Delta(PRICERATIO00) + 0.070*\Delta(GASCONS00)$$

where CONS00: electricity consumption per capita

Δ (PRICERATIO00): differenced series of natural gas price / electricity price

Δ (GASCONS00): differenced series of natural gas consumption

The findings of this model point out that differenced price ratio series is insignificant at 5% level with a t value of 0.894. Additionally R^2 and adjusted R-squared values are both 0.069 and 0.046 respectively which are low for a good model.

Durbin Watson statistics is 0.55 which is also low and indicates the positive correlation in the residuals. P value of the F statistics is 0.55 so the null hypothesis that all the slope coefficients in the equation is zero is not rejected. Finally the sign of the gas consumption variables is not correct. All these findings show us that the model is misspecified.

Next, one period lagged values of independent variables and 12th lag of consumption are used in the model considering that effects of these variables is seen one period later.

CONS00 = -0.472 + 0.266 * Δ(PRICERATIO00(-1)) – 0.052 * Δ(GASCONS00(-1)) + 1.099 * CONS00(-12)

The results are better than the first one, with a high R2 of 0.816 and adjusted R-squared of 0.809. Durbin Watson statistics is 1.18, that is still smaller than 2, however as there is a lagged value of the dependent variables, using another serial correlation test will give better results. Although PRICERATIO00 variable still seems insignificant at 5% level with t values 1.828 it will not be dropped from the model as its effect to electricity consumption is searched in this study. Additionally, the variable is significant at 10% level. Signs of the variables are correct and the model seems appropriate; however the distribution of residuals and autocorrelation in residuals should be well analyzed. Jarque-Bera test statistics is applied to the residuals to see whether they are normally distributed or not. The test statistics is 19.296 with a p value near to 0, which strongly rejects the null hypothesis of normal distribution. Secondly Breusch-Goldfrey Serial Correlation LM test indicates there is a serial correlation in the residuals with F statistics 2.437 and Observed R-squared with value of 24.698. This shows that standard errors are estimated in the wrong way and residuals are not efficient as well as they are misleading. These also indicate that the model is misspecified. The residuals from the model are seen in Figure 4.9.

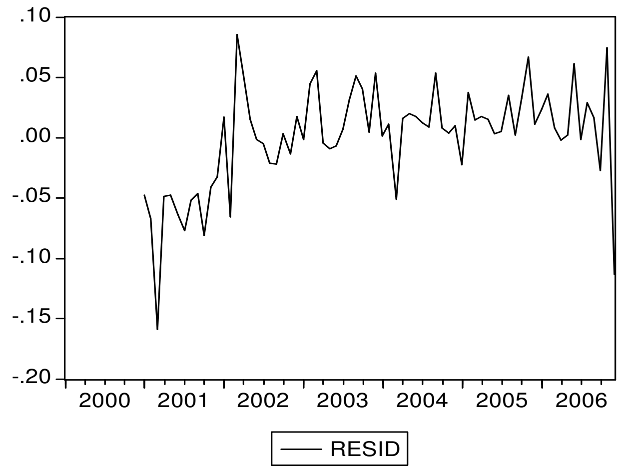

Figure 4.9 - Residual Graph

The graph of the residuals show that there are peak points in the residual series which may be a result of any structural changes or crisis in fuels, economy, etc and effects the electricity consumption, so a dummy variable is added to the model for year 2001. The effects of financial crisis in February 2001 effected income, GDP, value of money, industries and finally electricity consumption. On the other hand, constant is eliminated from the model for being insignificant compared to the critical values at 5% level. After experimenting with various models, the final model becomes:

$$\text{CONS00} = \underset{\substack{(0.131)\\(2.289)}}{0.300} * \Delta(\text{PRICERATIO00}(-1)) - \underset{\substack{(0.018)\\(-2.509)}}{0.046} * \Delta(\text{GASCONS00}(-1))$$

$$+\underset{\substack{(0.001)\\(1070.503)}}{1.007} * \text{CONS00}(-12) - \underset{\substack{(0.041)\\(-3.978)}}{0.162} * D_1$$

R^2 shows success of regression in predicting values of dependent variable within the sample whereas adjusted R-squared penalizes for addition of irrelevant regressors, both should be around 0.80. The results suggest the model fit the data well because R^2 is 0.847 whereas adjusted R-squared is 0.840 which are high enough. AIC is -3.527 and SIC is -3.400 that is better than the previous models. p value of F statistics is near zero and the null hypothesis that all the slope coefficient is 0 is rejected.

Serial Correlation in Residuals: DW statistics in this equation is 1.24 and low. However, in this study, Breusch Godfrey serial correlation LM test is used to test for serial correlation, as there are lagged values of the dependent variable in the model. Breusch Godfrey statistics strongly indicate that there is no serial correlation in the residuals with F-statistics equal to 1.286 and probability 0.237, and Observed R-squared equal to 42.568 with probability 0.209.

Normality: Jarque Bera test statistics is used as normality test in this study. According to the results and the graph given in Figure 4.10, null hypothesis of normality is accepted.

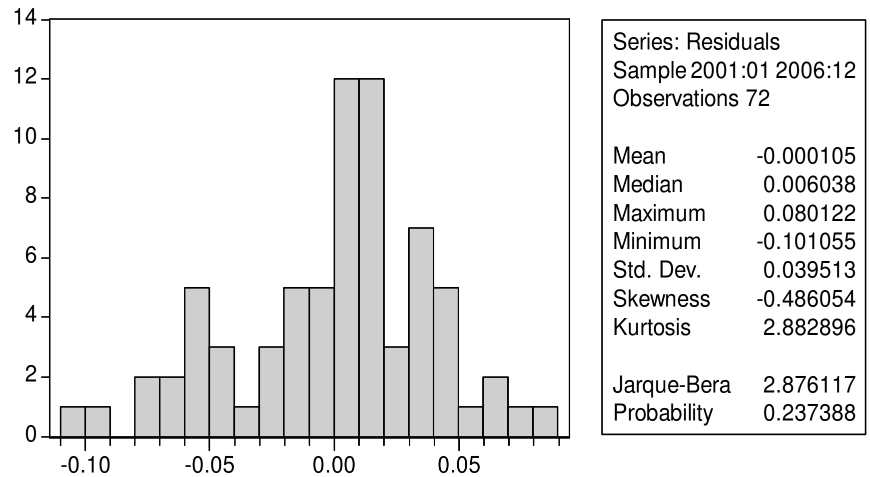

Figure 4.10 – Jarque- Bera Test Statistics for Regression Equation

The assumption of the normality of the residuals holds, with Jarque-Bera test statistics value of 2.876.

White's Heteroskedasticity Test: The results show null hypothesis of no heteroskedasticity is accepted with the F-statistics equal to 1.139 and probability 0.349 and Obs. R-squared equal to 12.438 with probability 0.332.

ARCH effect: ARCH is mentioned in detail in Chapter 3. The volatility cannot be modeled as test statistics show no ARCH effect with probabilities 0.164 and 0.169 for F statistics observed R-squared respectively.

The resulting model suggests that both of the variables are significant at 5% level, which indicates that natural gas is a substitute of electricity and this variable should be considered while forming a model. On the other hand, natural gas consumption and natural gas price / electricity price ratio have correct signs. There is a negative relationship between electricity and natural gas consumption as they are substitutes and when natural gas consumption increases 1 unit, electricity consumption decreases 0.046. Price ratio and consumption have positive relationship because increase in price ratio may be thought as increase in natural gas price / decrease in electricity price which cause increase in electricity consumption. It seems that changes in prices effect electricity consumption more than the natural gas consumption.

However, as mentioned in the beginning of this section, the data set is too small to compete with seasonal fluctuations in an error correction model. So, being conscious about the impact of natural gas on electricity consumption, the cointegration technique is applied without natural gas data.

4.2 MODELLING ELECTRICTY DEMAND IN TURKEY FOR 1990-2006

In this section, the stationary of the variables will be searched by using unit root tests. Secondly, if there is a cointegration relation between the nonstationary variables, error correction models will be used.

As most of the electricity demand forecasting studies utilizes GDP as an explanatory variable, monthly electricity demand in Turkey is firstly modeled using GDP, electricity price and temperature data. Only quarterly GDP data is available for Turkey in Turkish Central Bank database and quarterly values are converted to monthly data by interpolation. Finally, these values are divided by the population to get real GDP/capita values and when the graph for monthly real GDP/capita is investigated, a seasonal pattern where production increase in winter and decrease in summer months occurred, which may not be really the case. The model formed with GDP series had several problems such as wrong signs and unexpected values and the main reason for this inefficient model and misleading results is thought to be converting quarterly GDP data to monthly data and the loss of properties of the data during this process in addition to the seasonal pattern of GDP data, which is not really correct. So, instead of converting quarterly GDP data to monthly data, monthly industry production index data, available in Turkish Central Bank is used. Although GDP cannot be presented only by industry production, the correlation between those two variables, which is equal to 0.693, indicates that industry production data mostly represent GDP data. Moreover, Figure 4.11 shows that electricity consumption in industry has the biggest share in Turkey in the 4th quarter of 2006 (TUIK, 2007a). This also indicates that using industry production index is appropriate for this study.

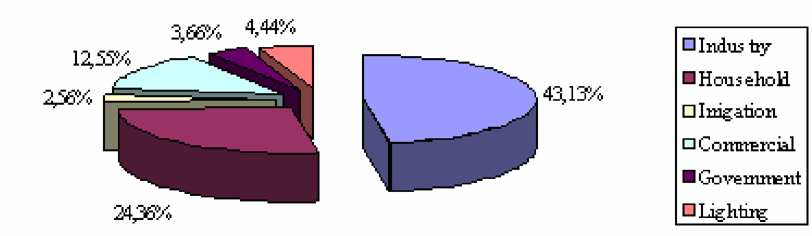

Figure 4.11 – Electricity Consumption w.r.t Consumer Type (TUIK, 2007a)

Electricity demand model is formed in two ways; firstly average temperature is used as an explanatory variable and secondly HDV and CDV are used instead of average temperature. Then, the results of the two models are compared. In fact, it is discussed that HDD and CDD is not a good way to estimate energy demand as heating or cooling requirements may change due to the structure of the buildings (whether insulated or not), personal choices like when they feel comfortable and number of sunny days. This is because heat requirement does not change linearly with temperature and electricity used in households or buildings does not only limited with heating. There are other electricity requirements such as cooking, household appliances etc.

4.2.1 Stationarity

ADF test and HEGY test, developed by Hylleberg et al in 1990 and extended by Bealieu and Miron (1993) for monthly data are used to search for unit roots in time series data.

In this study, firstly y_{1t}, y_{2t}..y_{12t} series were generated using the formula indicated by Beaulieu and Miron (1993). Using these new generated series, 3 years (36 months) lagged dependent variable, a constant, trend and seasonal dummies for 12 months where appropriate, series were tested against the alternative that there exists no unit root in time series data. Lagged dependent variables that are insignificant (smaller than %15) were deleted from the model by using Schwarz Information Criteria as in Bealieu and Miron's study (1993).

HEGY test results and π values for the test are indicated in Table 4.2, shaded areas indicate that there exists a unit root at that frequency. Critical values for the test statistics at %5 level are given in Table 4.3.

Test results indicate that in most of the series, π_1 is equal to 0 and there exists unit root in logarithmic values of consumption per capita, real electricity prices and industry production index. Average temperature series do not have any unit at any frequencies. Moreover, the results show that none of the series have unit roots at any seasonal frequencies. HDV and CDV series do not have any unit root at any periodic or seasonal frequencies.

Table 4.2– HEGY Test Results for Variables

DEP VAR	CONS	TEMP	INDEX	PRICE	HDV	CDV
EXP VAR	constant +trend + seas. dum	constant + seas. dum.	constant +trend+ seas dum.	constant +trend+ seas dum.	constant + seas. dum	constant + seas. dum
Π1	3.081	-4.974	0.529	-1.531	-4.593	-3.605
Π2	-8.373	-5.861	-7.820	-5.576	-4.996	-4.208
Π3	-2.850	-5.863	-4.650	-4.818	-6.476	-5.517
Π4	-3.906	-0.662	-5.658	-8.173	-0.608	-0.502
Π5	-6.733	-8.388	-6.746	-4.616	-7.834	-6.121
Π6	-3.496	-1.872	-0.958	3.081	-0.734	-0.916
Π7	-7.648	-6.693	-9.063	-8.390	-6.701	-5.962
Π8	-6.849	-2.056	-7.054	-8.104	-1.262	-1.439
Π9	-8.597	-7.428	-9.156	-4.917	-6.554	-5.378
Π10	-0.379	0.236	0.152	4.367	1.162	0.898
Π11	-5.477	-7.848	-6.007	-6.656	-6.849	-5.667
Π12	1.610	0.896	3.941	0.166	1.626	0.868

Table 4.3 – Critical Values at 5% level

Explanatory Variable	constant+trend	constant + seasonal dummy	constant +trend+seas dummy
Π1	-3.320	-2.760	-3.280
Π2	-1.880	-2.750	-2.760
Π3	-1.880	-3.250	-3.240
Π4	-1.610	-1.850	-1.850
Π5	-1.880	-3.250	-3.240
Π6	-1.610	-1.850	-1.850
Π7	-1.880	-3.250	-3.240
Π8	-1.610	-1.850	-1.850
Π9	-1.880	-3.250	-3.240
Π10	-1.610	-1.850	-1.850
Π11	-1.880	-3.250	-3.240
Π12	-1.610	-1.850	-1.850

The results are checked with ADF test statistics.

ADF test results are given Table 4.4.

Table 4.4 – ADF Test Results for Series

	t statistics	Prob	Mc Kinnon Critical values		
			%1 level	%5 level	%10 level
CONS.	-1.566	0.803	-4.007	-3.434	-3.141
TEMP.	-2.868	0.051	-3.464	-2.876	-2.575
INDEX	-2.018	0.588	-4.007	-3.434	-3.141
PRICE	-2.037	0.577	-4.004	-3.432	-3.140
HDD	-2.051	0.265	-3.464	-2.876	-2.575
CDD	-2.467	0.125	-3.464	-2.876	-2.575

t statistics indicate that there is unit root in all of the series at 5% level . To control level of integration, the ADF test statistics is applied to first difference of the series. The results rejected the null hypothesis of unit root for all of the series and they are found to be integrated of order 1.

Table 4.5 – ADF Test Results for Differenced Series

	t statistics	Probability	Critical values		
			%1 level	%5 level	%10 level
Δ(CONS.)	-5.201	0.000	-4.007	-3.434	-3.141
Δ(TEMP.)	-15.912	0.000	-3.464	-2.876	-2.575
Δ(INDEX)	-5.012	0.000	-4.007	-3.434	-3.141
Δ(PRICE)	-11.724	0.000	-4.004	-3.432	-3.140
Δ(HDD)	-29.513	0.000	-3.464	-2.876	-2.575
Δ(CDD)	-33.551	0.000	-3.464	-2.876	-2.575

However, the outcome found by ADF test did not justified the results of HEGY test indicating roots in all of the series except average temperature, HDV and CDV at %5 level. The reason of this difference is disregarding the seasonal variation in weather series while performing ADF test. So, Philips-Perron test is also applied to these series. The results shows that t values for average temperature, HDV and CDV are -5.643, -6.533 and -4.766 and nearly with 0 probabilities whereas the test critical value at 5% level is -2.876. So, the results show that the unit root hypothesis is rejected.

Therefore, results of HEGY test is used in this study since it considers both seasonal variation and deterministic trend.

4.2.2 Cointegration

As most of the series in econometrics, consumption, production index and price series are nonstationary too. Differencing is one of the methods to remove stationarity as mentioned before. However, differencing the series till stationarity is achieved may cause loss of long run properties in series and is not the best solution at all. So, cointegration is used in this study.

Engle and Granger Test procedure is applied to the series; consumption, production index and price for both of the models. Engle and Granger test procedure includes two steps commonly. First of all the order of integration of variables is tested since both the dependent and independent variables should be at the same order. Secondly, cointegrating vector and residuals are estimated and residuals are tested for unit root. That is, when all variables are integrated of the same order, a cointegrating equation can be found and residuals are tested for unit root with ADF test.

The series average temperature, HDV and CDV do not have any unit roots so cointegration equations for both of the model are the same including consumption, industry production index and price series. If the test results indicate a cointegrating relation among these variables, first model will be formed with these variables and additionally with the average temperature variable that is added to the model later. Second model will be formed again with these cointegrating variables in addition to HDV and CDV.

The cointegration equation for both of the models is indicated below with standard errors and t values in parentheses, respectively.

$$CONS = \underset{(-1.376)}{\underset{(0.580)}{-0.798}} + \underset{(18.117)}{\underset{(0.063)}{1.145}}*INDEX - \underset{(-0.611)}{\underset{(0.078)}{0.047}}*PRICE$$

where CONS : log (electricity consumption per capita)

INDEX: log (industry production index)

PRICE: log (real electricity price)

Signs of the variables are as expected, that is consumption is positively related with Industry Production Index whereas electricity price is negatively correlated with consumption. The result shows that the coefficient of index is above one, which is the situation in most of the countries and price elasticity is 0.047, which is a low value compared to industry production index. As it is the same in many countries, electricity prices are regulated by the Government in Turkey and whatever the price is, electricity demand increases. Other variables are significant at 5% level and all of the variables have correct signs. According to the procedure described, the next level is to search whether the residuals are I(0) or not.

Residuals from the cointegrating equation are tested for unit roots, ADF test procedure with lag length 4 is used which is chosen as the best lag length by SIC. Additionally, Philips Perron unit root test is applied to residuals. The test statistics rejects the null hypothesis that there is a unit root in the residuals with values indicated in Table 4.6 and Table 4.7.

Table 4.6 – ADF Test Results for Residuals

ADF Test Statistics		t statistics	prob
		-3.014	0.035
Critical Values	**%1 level**	-3.463	
	%5 level	-2.876	
	%10 level	-2.575	

Table 4.7 – Philips-Perron Test Results for Residuals

Philips-Perron Test Statistics		t statistics	prob
		-7.289	0.000
Critical Values	**%1 level**	-3.463	
	%5 level	-2.877	
	%10 level	-2.574	

Therefore, although the series have unit root at zero level and they are integrated of order 1, the linear combination of them is I(0) and there exist a cointegration relation between the series.

4.2.3 ECM with Average Temperature Variable (ECM-1)

If there is a long run relation between the series, they can be modelled as ECM. This kind of models includes constant, deterministic components, lagged values of variables and the cointegrating coefficient, called speed of adjustment. This error correction term indicates necessary period of time for system to come into equilibrium when there is a shock, which makes system deviate from equilibrium. (Türker, 1999)

ECM for the first type of variables is given below with standard errors and t values in parentheses, respectively.

Δ(CONS) = -0,044*RESIDUAL(-1) – 0,126*Δ(PRICE(-3)) + 0,259*Δ(INDEX) +

(0,016)	(0,046)	(0,028)
(-2,758)	(-2,746)	(9,313)

0,095*Δ(INDEX(-10))-0,300*Δ(CONS(-1))-0,205*Δ(CONS(-2))-0,292*Δ(CONS(-3))

(0,025)	(0,041)	(0,049)	(0,043)
(3,797)	(-7,398)	(-4,114)	(-6,731)

-0,177*Δ(CONS(-4))-0,215*Δ(CONS(-10)+0,260*Δ(CONS(-12))+0,013*TEMP(-12)

(0,036)	(0,044)	(0,039)	(0,001)
(-4,791)	(0,039)	(6,575)	(13,711)

– 0,067*S_1 – 0,060*S_2 – 0,046*S_3 -0,075*S_4 - 0,069*S_5 - 0,086*D_1 - 0,129*D_2 -

(0,009)	(0,008)	(0,008)	(0,009)	(0,010)	(0,020)	(0,020)
(-7,419)	(-7,032)	(-5,496)	(-8,649)	(-6,597)	(-4,241)	(6,416)

0,069*D_3 - 0,084*D_4

(0,020)	(0,020)
(-3,422)	(-4,156)

where CONS : electricity consumption per capita

PRICE: real electricity price

INDEX: industry production index

TEMP: average temperature

Δ: difference operator

S_1, S_2, S_3, S_4, S_5,: Seasonal dummy variables for February, April, June, September, May

D_1, D_2, D_3 and D_4: Dummy variables for the 4th month of 1998, 1st and 4th month of 2000 and 3rd month of 2001 respectively.

R^2 is equal to 0.933 and adjusted R-squared value is 0.925 which indicates the model fits the data well. Seasonal dummies and 4 dummy variables for the 4h month of 1998, 1st and 4th months of 2000 and 3rd month of 2001 are used in the equation. Turkey is a country that faced many problems and fluctuations in its history. In 1998, inflation decreased and economic activities slowed down in Turkey, but there was not a crisis in the country. However, 1998 was a crisis year for the world. Asia, Russia and Latin America crisis showed their effects in Turkey also especially in some of the sectors such as clothing, steel-iron or suitcase production because export to other countries and capacities of the plants decreased (9 Eylül University, 1998). 2000 and 2001 were the years of financial and foreign currency crisis, which caused economic recession, variation in labor force and production and service sectors in Turkey. In 1999, the income level decreased 11.6 % and the country lived one of the worst years in its history. Thus, in the beginning of 2000, Government started a new program about the economic policy and dominated the financial market, which later caused variation in daily interest rates. This new program had an effect on income and consumption in the first months of 2000. However, the result was an economic crisis in November 2000. However, the economic program implemented caused changes in the market and this caused to add dummy variables in this model. After 2000 crisis, in February 2001, Turkey faced with a new currency crisis where the value of 1$ increased nearly %40 only in 10 days time.

While modeling, the estimators and test statistics are derived under some assumptions such as normality, autocorrelation or heteroskedasticity. So, these assumptions are tested for the models to be valid.

Normality of the Residuals: The output of the Jarque-Bera test statistics can be seen at the graph given in Figure 4.12.

The probability of Jarque-Bera test is 0.541, indicating the null hypothesis of normal distribution is not rejected, Also, looking at the χ^2 table, $\chi^2_{0,025;2}$ is equal to 7.378 which is a bigger number than 0.798 and an indicator of the acceptation of null hypothesis at %5 level.

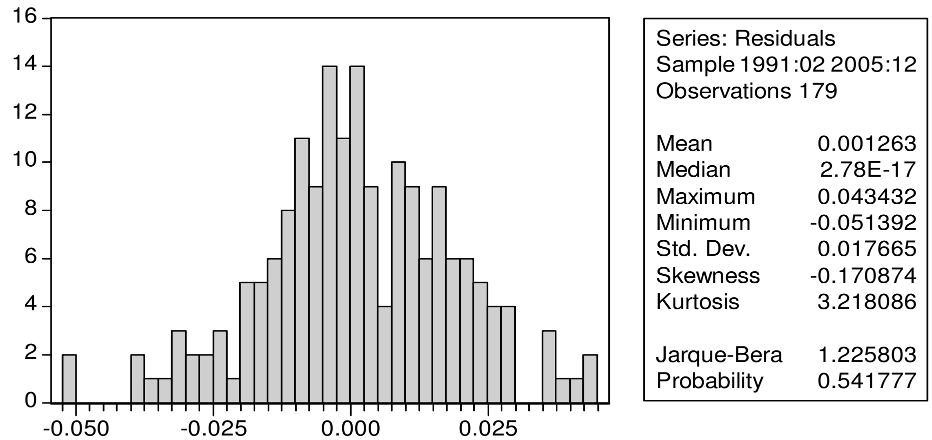

Figure 4.12- Jarque-Bera Test Statistics for ECM-1

Autocorrelation*:* Another condition that should be investigated is the existence of autocorrelation in the residuals. The results of the study, using a lag number 12 as the data is monthly, are shown below indicating no serial correlation in residuals.

Breusch-Godfrey Serial Correlation LM Test:

F-statistic:	1.579	Probability: 0.104
Obs*R-squared:	19.749	Probability: 0.072

Heteroskedasticity: White's heteroskedasticity test is used in this study.

White Heteroskedasticity Test:

F-statistic:	1.022	Probability 0.445
Obs*R-squared:	33.784	Probability 0.429

Two test statistics are reported; F-statistics for comparison and Obs*R-squared statistics as Whites test statistics (number of observation * Centered R-squared from the test regression). According to the results, the null hypothesis that there is no heteroskedasticity in residuals against the alternative that there is heteroskedasticity in some of the residuals is accepted.

ARCH Test: ARCH effect is investigated in the residuals. If there is ARCH effect in the residuals the series will be modeled using ARCH models.

<u>ARCH LM Test:</u>

F-statistic: 0.943 Probability : 0.505

Obs*R-squared: 11.435 Probability : 0.492

The results do not reject the null hypothesis indicating that there is no ARCH effect in residuals up to order q.

Stability of the Parameters*:* The parameters in the model are also tested for stability within some subsamples of data by CUSUM and CUSUM of square tests. The test results in Figure 4.13 indicated stability of the parameters and there is not a structural change.

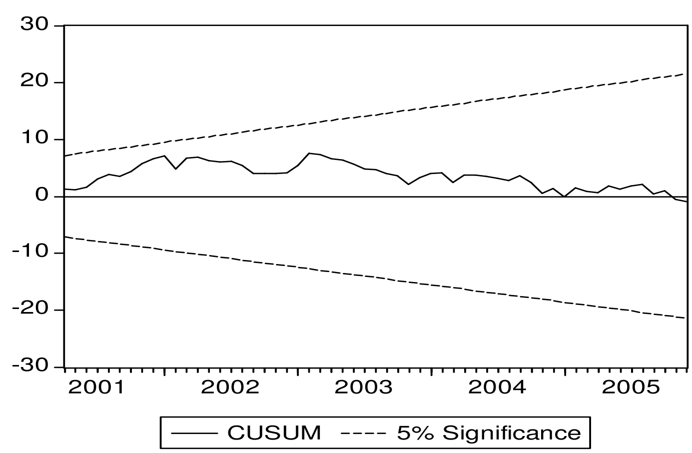

Figure 4.13 – CUSUM Test for ECM-1

The CUSUM of square test gives a graph of S_t against t and compares the results within %5 critical lines. The result again indicates the stability of the parameters.

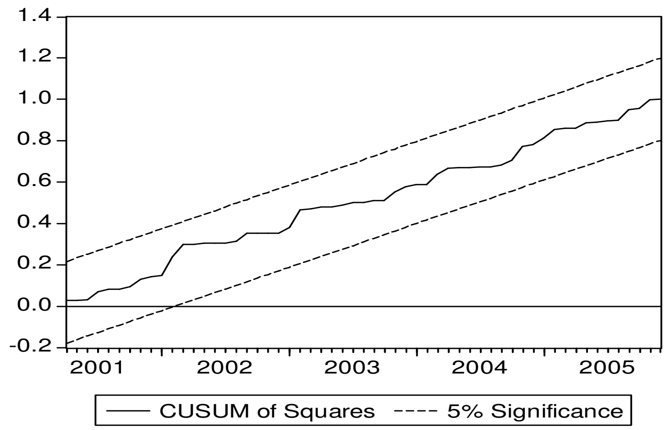

Figure 4.14 – CUSUM of Squares Test for ECM-1

Secondly, the electricity demand model is formed using HDD and CDD instead of average temperature.

4.2.4 ECM with HDV and CDV variables (ECM-2)

ECM is formed by using electricity consumption, industry production index, electricity price, HDV and CDV as variables. The resulting model can be seen below with standard errors and t-values in parenthesis.

$$\Delta(CONS) = \underset{\substack{(0,016) \\ (-2,504)}}{-0,039*RESIDUAL(-1)} + \underset{\substack{(0,028) \\ (8,333)}}{0,237*\Delta(INDEX)} + \underset{\substack{(0,028) \\ (11,717)}}{0,070*\Delta(INDEX(-10))}$$

$$-\underset{\substack{(0,046) \\ (-1,468)}}{0,068*\Delta(PRICE(-3))} - \underset{\substack{(0,048) \\ (-9,306)}}{0,449*\Delta(CONS(-1))} - \underset{\substack{(0,039) \\ (-3,874)}}{0,154*\Delta(CONS(-2))} - 0,181*$$

$$\underset{\substack{(0,042) \\ (-4,282)}}{\Delta(CONS(-3))} - \underset{\substack{(0,030) \\ (-4,777)}}{0,143*\Delta(CONS(-4))} - \underset{\substack{(0,044) \\ (-3,614)}}{0,159*\Delta(CONS(-10)} + \underset{\substack{(0,040) \\ (5,706)}}{0,231*\Delta(CONS(-12))} +$$

$$\underset{\substack{(0,000) \\ (11,717)}}{0,007*HDD(-12)} - \underset{\substack{(0,009) \\ (-10,136)}}{0,099*S1} - \underset{\substack{(0,009) \\ (-3,629)}}{0,031*S2} - \underset{\substack{(0,008) \\ (-6,127)}}{0,048*S3} - \underset{\substack{(0,009) \\ (-3,944)}}{0,037*S4} + \underset{\substack{(0,007) \\ (7,974)}}{0,059*S5} + \underset{\substack{(0,007) \\ (7,963)}}{0,055*S6}$$

$$-\underset{\substack{(0,004) \\ (-5,355)}}{0,041*S7} - \underset{\substack{(0,020) \\ (-4,251)}}{0,086*D1} + \underset{\substack{(0,020) \\ (5,339)}}{0,107*D2} - \underset{\substack{(0,020) \\ (-2,537)}}{0,051*D3}$$

where CONS : electricity consumption per capita

PRICE: real electricity price

INDEX: industry production index

HDV: heating degree variable

CDV: cooling degree variable

Δ: differencing operator

$S_1, S_2, S_3, S_4, S_5, S_6, S_7$: seasonal dummy variables February, March, April, May, July, August and September

D_1, D_2, D_3: dummy variables for the 4th month of 1998, 1st month of 2000 and 3rd month of 2001 respectively. The dummy variables added are the same as the ones in the first model and they are included in the model with similar reasons.

R^2 value is 0.937 and adjusted R-squared is 0.929 indicating the data fit the model well. The model is discussed in more detail in the "comparison of the models" section of this Chapter.

The next step is checking whether the model is valid or not.

Normality of the Residuals: The results show 0.576 probability for Jarque-Bera test statistics, indicating the null hypothesis of normal distribution is not rejected. Also, X^2 table shows that $X^2_{0,025;2}$ is equal to 7.378 which is bigger than the Jarque-Bera value. This shows that the null hypothesis is not rejected at %5 level.

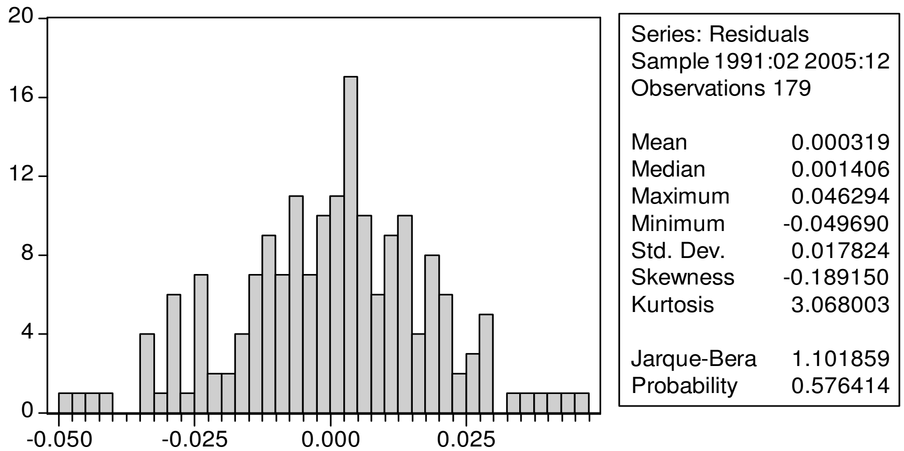

Figure 4.15 - Jarque-Bera Test Results for ECM-2

Autocorrelation: The null hypothesis, there is no serial correlation in the residuals up to lag order 12 is tested with Breusch-Godfrey Serial Correlation LM Test, whose results are given next.

Breusch-Godfrey Serial Correlation LM Test:

F-statistic:	1.512	Probability	0.126
Obs*R-squared:	19.747	Probability	0.072

White's Heteroskedasticity Test: White's Heteroskedasticity Test is used to test for heteroskedasticity. The results indicate the null hypothesis that there is no heteroskedasticity in residuals against the alternative that there is heteroskedasticity in some of the residuals is not rejected.

White Heteroskedasticity Test:

F-statistic:	0.857	Probability 0.688
Obs*R-squared:	28.301	Probability 0.654

ARCH Test: ARCH LM test results do not reject the null hypothesis indicating that there is no ARCH effect in residuals up to order 12.

ARCH LM Test:

F-statistic:	1.159	Probability: 0.317
Obs*R-squared:	13.838	Probability: 0.311

The results approved the null hypothesis that there is no ARCH effect in the residuals.

Stability of the Parameters: CUSUM and Squares of CUSUM test is used as stability tests. The test is based on cumulative sum of the recursive residuals and if the cumulative sum is outside the %5 critical value range, it is concluded that there is some instability. The graphs 4.16 and 4.17 indicate that the parameters are stable.

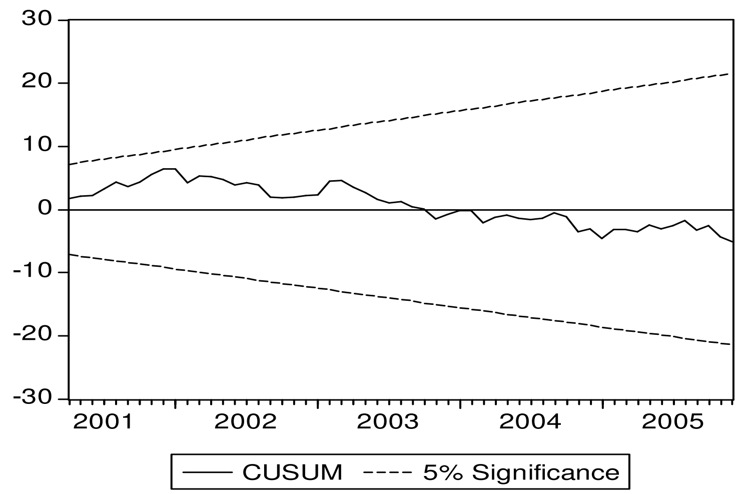

Figure 4.16 – CUSUM Test for ECM-2

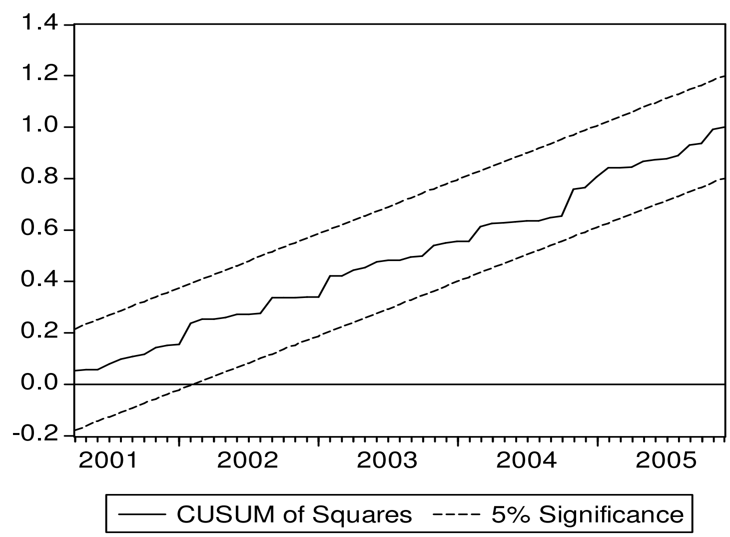

Figure 4.17 – CUSUM of Squares Test for ECM-2

4.2.5 Comparison of the Models

The cointegrating relationship for the electricity demand forecasting model indicates that industry production index is positively correlated with consumption and price is negatively correlated with consumption. Production index elasticity of consumption is 1.145 in the long run. In many countries GDP elasticity of electricity consumption is more than one as it is the case in this study too. According to Table 4.8, elasticity is over 1 in developing countries, whereas in developed countries it is equal to or smaller than 1 (ETKB, 2004). However, price elasticity is low compared to industry production index elasticity and it is 0.047 in the long run. This means that electricity consumption per capita does not respond to changes in price as much as in industrial production. This is expected, since electricity does not have substitutes in most of its usage areas such as lighting and for household appliances. Electricity is an interesting input and output as it cannot be stored and should be consumed when it is supplied. Thus, whatever the price is people use electricity for their daily needs or to produce something.

Table 4.8 – GDP Elasticity of Electricity (ETKB, 2004)

	ELASTICITIES							
	Austria	**Spain**	**Portugal**	**Italy**	**Germany**	**England**	**Turkey**	**Korea**
1980	1.36	2.08	2.02	0.61	0.82	2.94	-1.82	-1.67
1985	1.85	1.87	2.18	1.05	1.32	1.47	-1.82	-1.67
1990	0.93	0.99	2.93	0.72	-0.54	1.74	0.86	1.55
1991	1.92	1.05	3.60	0.50	-0.76	-1.65	7.01	1.10
1992	-0.36	3.85	6.46	1.71	0.57	-1.05	1.86	1.91
1993	-3.68	0.66	0.17	-0.13	0.87	0.25	1.15	1.87
1994	0.39	1.59	1.28	3.10	0.32	0.13	-1.09	3.27
1995	1.95	1.76	2.02	0.82	1.19	0.86	1.39	1.24
1996	1.54	0.97	3.89	0.18	1.98	1.49	1.54	1.66
1997	0.33	1.67	2.04	1.42	-0.11	-0.17	1.50	1.78
1998	0.60	1.39	3.20	0.77	0.61	1.16	2.61	0.36
1999	0.75	1.94	5.01	0.70	0.05	0.82	-0.83	1.00
2000	1.34	1.70	1.84	1.28	-1.18	0.74	1.12	1.16

Additionally, the electricity market is not fully liberalized in Turkey yet and there was a vertical integrated monopoly for years where electricity generation, transmission and distribution activities were owned by the same public company, TEK. The consumers do not have the right to choose their electricity suppliers or the price they will buy electricity. Additionally, governmental policies set electricity prices and the prices are regulated. The most significant example of this situation is that electricity prices are the same for four years. All of these reasons, affect the relation between electricity consumption and price and conflicts with the general economic theory where demand and price acts oppositely. On the other hand, Turkey is a developing country and parallel to income, electricity used for industrial purposes and electricity consumption increase too.

The first Error Correction Model includes error correction term, lagged values of real price, industry production index, consumption, temperature, seasonal dummies and 4 additional dummy variables. It is expected to have (–) and (+) signs for price and income data respectively, and the signs are correct. Third lag of price variable is found significant with coefficient 0.126. On the other hand, monthly electricity consumption depends on the industry production index variable and its 10^{th} lag. Both of these variables are positive, indicating that an increase in production will increase electricity consumption. However, in the short run, income elasticity is lower with value 0.259. Lagged values of dependent variable are also used, as electricity consumption mostly depends on its lagged values. 12^{th} lag of average temperature series is also found significant in the model, but with a low coefficient, 0.013. Error correction term is significant at 5% level and coefficient of adjustment is 0.044, which indicates that if there is a shock for the system, it will take about 23 months for the system to reach to equilibrium, which is nearly two years time.

5 seasonal dummies are used in the model for February, April, June, September and May to capture the deterministic seasonal behavior and to get more precise result as the data is seasonal data and the series contains seasonal patterns. Also four other dummy variables for 4^{th} month of 1998, 1^{st} and 4^{th} months of 2000 and 3^{rd} month of 2001 are used since electricity consumption change owing to some external shocks such as fuel crisis, economical crisis inside or outside the country.

The variables used in the second model are similar to the ones in the first model except the average temperature variable. Instead of this variable HDV and CDV series are used in the model. However, CDV is omitted from the equation since it was insignificant regarding the t test. As in the first model, all of the variables have correct signs. Industry production index have positive values, but the short term elasticity is smaller compared to the long run with value 0.237. Lagged dependent variables are also used in the model. The 12^{th} lagged of HDV is used and it has positive sign as expected, which means when number of heating degree days increase, consumption increase, but with a low coefficient 0.007. This time seven seasonal dummy variables are used for the months February, March, April, May, July, August and September and also three dummy variables are used for April 1998, January 2000 and March 2001. The increase in number of seasonal dummy variables may depend on to the HDV and CDV series, which show seasonal patterns. Adjustment of coefficient is 0.039 in the second model which is smaller than the first model. This means adjustment after a shock will be slower in the second model.

To sum up, the data to forecast the model is monthly data between 1990 and 2005. 2006 electricity consumption data is also available and is used to compare the forecast results with the actual ones. Two error correction models are formed, the first one including average temperature as an explanatory variable and the second one including HDV and CDV. R^2 and adjusted R-squared values are closer to each other in both of the models, being smaller in the first model. The values of model selecting criteria AIC and SIC are stated as:

	The First Model	The Second Model
AIC:	-4.956	-5.000
SIC:	-4.599	-4.627

They are close to each other. One year demand forecast and actual electricity demands per capita for 2006 is calculated for both of the models and outputs can be seen in Tables 4.9 and 4.10. Additionally percentage of deviation from the actual values is also given in these tables.

Table 4.9 – Actual and Forecasted Electricity Demand/Capita in 2006 (Model 1 - with Avg, Temp)

Month (2006) (TEMP)	Electricity Demand/Capita		% Error
	Actual	Forecast	
January	193.951	190.255	-0.019
February	185.816	185.013	-0.004
March	198.163	199.213	0.005
April	182.154	185.897	0.021
May	189.799	184.587	-0.027
June	195.976	184.400	-0.059
July	210.927	205.299	-0.027
August	221.764	211.845	-0.045
September	195.494	195.983	0.003
October	185.027	194.686	0.052
November	203.107	202.256	-0.004
December	183.926	216.235	0.176

Table 4.10 – Actual and Forecasted Electricity Demand/Capita in 2006 (Model 2 - with HDV and CDV)

Month (2006) (HDD-CDD)	Electricity Demand/Capita		% Error
	Actual	Forecast	
January	193.951	192.251	-0.009
February	185.816	184.273	-0.008
March	198.163	197.867	-0.001
April	182.154	179.273	-0.016
May	189.799	181.011	-0.046
June	195.976	182.235	-0.070
July	210.927	202.764	-0.039
August	221.764	206.357	-0.069
September	195.494	190.402	-0.026
October	185.027	188.347	0.018
November	203.107	197.782	-0.026
December	183.926	211.948	0.152

Looking at the outputs, it is not certain which model performs better, because in some of the months second model has less percentage error while the first one is better in other months. As the data is monthly and electricity consumption changes seasonally, the errors do not increase with time in a year. In December, the forecast results give nearly 18% error in the first model, and 15% in the second one, which indicate that the forecasted consumption value considerably deviates from the actual consumption.

Electricity demand especially increases in the summer months, in July and August and in the winter, in November and December according to the forecasts, which is predictable. The use of electricity, as well as the electricity prices, increases in summer months because of air conditioners. Oppositely, in the winter use of electrical heaters increase which cause an increase in electricity consumption. Moreover, economic activities rise during these months. That's why demand and supply balance is mostly a problem in summer and winter and electricity prices in the new pool system in Turkey are high during these months.

The important indicators of the accuracy of the models are shown in Table 4.11.

Table 4.11 – Comparison of the Errors

	Model with Avg. Temp	**Model with HDV and CDV**
RMSE	0.055	0.055
MAE	0.036	0.039
MAPE	0.681	0.756
TIC	0.005	0.005
Bias Prop	0.006	0.058
Var. Prop	0.004	0.006
Covar. Prop	0.989	0.935

All of the statistical indicators are smaller in the electricity demand model with average temperature. Particularly, RMSE and MAE can be used to evaluate and compare the efficiency of the models. RMSE values are the same and 0.055 for both of the models, whereas MAE is 0.0036 and 0.039 for the first and second model respectively. To conclude, the results show that both of the models are good enough to make forecasts with low statistical values. However, the first model is better in predicting the monthly demand. Since the only model that estimated monthly electricity demand in Turkey was in 1992, it is not possible to compare its results with these models.

CHAPTER 5

CONCLUSION AND FUTURE WORK

Alternative energy resources gained importance since current energy resources will not be able to meet the demand in the near future. Additionally, increasing energy demand, population growth and rapid economical growth made this issue more significant, especially in developing countries. The world economy grew %3.3 per annum between the years 1971 and 2002, whereas electricity demand increased %3.7, which shows the parallel relationship between GDP and electricity demand in the world. Electricity differs from other products and services because it cannot be stored and replaced by another energy resource. That is why supply and demand balance in electricity sector should be well defined. Most of the demand forecasting studies performed by the Government and private sector include long-term demand forecasts for their simplicity and to avoid seasonality problems. However, market liberalization, security of supply and environmental problems displayed the importance of medium or short term demand projections as the market and prices are more flexible today.

In this study, natural gas is examined as a significant substitute of electricity, especially for heating purposes in households. It is important because its usage increased quickly in Turkey during the last twenty years in big cities. Electricity demand is regressed on natural gas price over electricity price ratio and natural gas consumption using the data between 2000 and 2006. The results showed that cross price elasticity of electricity consumption is 0.300 and gas consumption elasticity is 0.046 in the long run, which means although they are significant, they do not affect electricity consumption reasonably. These small coefficients may be a result of non-spread use of natural gas in

Turkey or other substitutes of electricity that are not included in the model. Additionally, both of the natural gas and electricity prices are pressurized by the Government. Natural gas price in Turkey is also influenced by the outer shocks because of dependency to import. For instance, increase in petroleum price or the export policies of Russia shapes the price. To sum up, the electricity and natural gas markets are not operating under a fully liberalized and free market conditions yet. However, an important finding was that the natural gas variables are significant and whatever their coefficients are, they should be included in the electricity demand forecasting model to obtain better results. Natural gas variables cannot be added to this electricity demand modeling study because of the short period of the data. Therefore, as the time passes and monthly data is collected, natural gas series should be integrated to the error correction model. Additionally, other fuels should also be considered in the model if data can be obtained for them.

In the second part of the thesis, two medium term electricity demand forecasting models are formed by using the GDP, price and weather variables with recent data, for the years between 1990 and 2006. The first model covered average temperature as an explanatory variable, but apart from other works in Turkey, the second model covered HDV and CDV, other variables remaining the same. Seasonal and deterministic dummy variables are used to cope with seasonality. Monthly consumption, estimated for the year 2006 is compared with the actual data and the models are compared in between. The results showed that the first model is somewhat better. Short and long run elasticity of price and industry production index are different being small in the short run. GDP elasticity is bigger than 1 in the long run, which shows demand is highly elastic to GDP. On the other hand, price is nearly inelastic because electricity demand does not have too many substitutes and people cannot respond to changes in the electricity price. Furthermore, electricity prices are regulated by the Government in Turkey, which removes the known relationship between demand and price.

Cointegration and ECM method is preferred among various demand forecasting methods since it distinguishes between the long and short run effects, avoids filtering problem of series and simple to use. Furthermore, this study is one of the few medium term electricity demand forecasting studies that use monthly data and ECM method. However, ECM is disadvantageous in that it ignores the mutual relationship between the variables. For instance, changes in production index level may affect electricity consumption as well as changes in consumption might affect production index. The Granger causality test also approved the relation between industry production index and consumption. So, the most important problem faced during this study includes the causality relation between the variables. Further research can be done using VAR and VECM, which takes into account this situation. Another difficulty faced during this research was about the availability of the monthly data for the variables because electricity demand forecasts require sophisticated technical background and detailed data, however the current conditions of the market do not allow gathering all the data necessary. Also, estimating electricity demand requires the projections for social and economic variables such as GDP and population and deviation in these projections causes the model to deviate much more. The model is good in that the results obtained are reliable compared to the actual consumption values. However, it is advised to update the model each year as soon as new and reliable data is obtained.

To conclude, better forecasts bring better policies and strategies and they are vital for supply security for both of the public sector, which still owns nearly %50 of the production and private sector, whose role in the electricity market is becoming more significant.

REFERENCES

- ADICA (2005a) *MAED Methodology and Approach for Model Application*, accessible at August 11, 2007, http://www.adica.com/media/downloads/MAED%20Methodology.pdf
- ADICA (2005b) *MAED: Model for Analysis of the Energy Demand*, accessible at August 11, 2007, http://www.adica.com/media/downloads/MAED_Summary.pdf
- Aggarwala, R. and Kyaw, N. A. (2005) Equity market integration in the NAFTA region: Evidence from unit root and cointegration tests, *International Review of Financial Analysis,* 14, 393–406
- AKBANK (2006), *Sektörler: Elektrik Enerjisi Sektörü*, accessible at December 11,2007, http://www.exi26.com/Article.asp?PageID=832.
- Akdi, Y. (2003) *Zaman serileri analizi: birim kökler ve kointegrasyon*, Ankara Bıçaklar Kitabevi
- Alper C. E. and Yılmaz, K. (2004) Volatility and contagion: evidence from the Istanbul stock exchange, *Economic Systems,* 28, 353–367
- Anderson, K.P., (1971). Toward Econometric Estimation of Industrial Energy Demand: An Experimental Application to the Primary Metals Industry, The Rand Corporation, R-719-NSF, USA
- Anderson, K.P., (1973). Residential Energy Use: An Econometric Analysis, The Rand Corporation, R-1297-NSF, USA
- Apergis, N. (1997) Inflation uncertainty, money demand and monetary deregulation: Evidence from a univariate ARCH model and cointegration tests, *Journal of Policy Modeling,* 19, 279-293
- Arabul, H. and Selçuk, N. (2000), ASO, ISO, 2000, DPT, Elektrik enerjisinde ulusal politika: 2000 yılında Türkiye'de elektrik enerjisinde mevcut durum, sorunlar ve çözüm önerileri , Ankara, Laga Basım Yayın Reklamcılık

- Atiyas, I. and Nuez, J. (2007). Second generation structural reforms: de-regulation and competition in infrastructure industries; The evolution of the Turkish energy, telecommunication and transport sectors in light of EU harmonisation, İstanbul, Jamanak Matbaası
- Baillie R. T., Boilerslev, T. , Mikkelsen, H. O. (1996) Fractionally integrated generalized autoregressive conditional heteroskedasticity, *Journal of Econometrics,* 74, 3-30
- Bakır (2007), Türkiye'de Özel Sektör, Hidroelektrik Yatırımlarının Serüveni, 11 Nisan 2007'de Zonguldak Karaelmas Üniversitesi'nde verdiği konferans sunumu
- Bakırtaş, T., Karbuz, S. and Bildirici M. (2000) *An econometric analyses of electricity demand in Turkey*, Middle East Technical University Studies in Development, Ankara
- Banerjee, A., Dolado, J. J., Galbbraith, J. W. and Hendry, D.F. (1993). *Cointegration, error correction and the econometric analysis of nonstationary data, USA,* Oxford University Press
- Barakat, E.H., Qayyum, M.A., Hamed, M.N., Al Rashed, S.A. (1990). Short-term peak demand forecasting in fast developing utility with inherit dynamic load characteristics, *IEEE Transactions on Power Systems,* 5.
- Barnett, A.H. , Reutter, K. A. and Thompson, H. (2005). The first step in restructuring the US electric industry, *Energy Economics,* 27, 225– 235
- Baxter, R. E. and Rees, R. (1968). Analysis of the industrial demand for electricity, *The* Economic Journal, 78, 277-298.
- Beaulieu J. J. and Miron A. J. (1993). Seasonal Unit roots at aggregate US data, *Journal of Econometrics,* 55 , 305-328
- Bera, A. K., and Higgins, M. L. (1993). ARCH models: Properties, estimation and testing, *Journal of Economic Surveys,* 7, 307-366.
- Birchenhall, C. R. and Osborn, D.R. Chui, A.P.L, Smith, J.P. (1988) Seasonality and the order of integration for consumption, *Oxford Bulletin of Economics and Statistics*, 50,4

- Boedecker, E., Cymbalsky, J. and Wade, S. (2002) *Modeling distributed electricity generation in the NEMS buildings models*, accessible at July 26, 2007, http://www.eia.doe.gov/oiaf/analysispaper/electricity_generation.html
- Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroskedasticity, *Journal of Econometrics,* 31, 307-327.
- Brown, R.L., Durbin, J. and Evans, J.M. (1975). Techniques for testing the constancy of regression relationships over time, *Journal of the Royal Statisical Society,* 37, 149-163.
- Caporale, T. and Doroodian, K. (1994) Exchange rate variability and the flow of international trade, *Economics Letters,* 46, 49-54,
- Capros, P. (2005) *The PRIMES energy system model summary description*, National Technical University of Athens, European Commission Joule-III Program, accessible at August 11,2007 http://www.e3mlab.ntua.gr/manuals/ PRIMsd.pdf
- Charemza, W.W. and Deadman D.D. (1997) *New directions in econometric practice*, (2nd edition), USA, Edward Elgar Publishing Limited
- Charytoniuk, W. and , Chen M.S. (2000). Neural-Network-Based Demand forecasting in a Deregulated Environment, *IEEE Transactions on Industry Applications*, 36, 3
- Christodoulakis, N. M., Kalyvitis, U. S. C., Lalas, D. P., Pesmajoglou, S. (2000) Forecasting energy consumption and energy related CO_2 emissions in Greece: An evaluation of the consequences of the Community Support Framework II and natural gas penetration, *Energy Economics,* 22, 395-422
- Commission Ruling numbered: 1424/38, dated:17/12/2007, EPDK, accessible at December 26, 2007, http://www.epdk.org.tr/mevzuat/kurul/elektrik/1424_38/1424_38.doc

- Commission Ruling numbered: 1425, dated:17/12/2007, EPDK, accessible at December 26, 2007, http://www.epdk.org.tr/tarife/elektrik/ toptansatis/ tetas/ 1425.doc
- Commission Ruling numbered: 26744, dated 31/12/2007 EPDK, accessible at January 2, 2008, http://www.epdk.org.tr/tarife/elektrik/ gecisdonemi/ 1455/1455.doc
- Craig S. H. and Rush, M. (1991) Cointegration: how short is the long run?, *Journal of International Money and Finance,* 10, 571-581
- Çağlar, A.O. (2006) Gözlük Mü Teleskop Mu?, *Global Enerji*, 22,40-45
- Çakır, Ö. (2002) Türkiye elektrik enerjisi talebinin zaman serisi teknikleri ile tahmini, Master's thesis, Marmara University, İstanbul
- Çifter, A., Özün, A. and Yılmazer, S. (2007) Beklenen kuyruk kaybı ve genelleştirilmiş pareto dağılımı ile riske maruz değer öngörüsü: faiz oranları üzerine bir uygulama, *Bankacılar Dergisi,* 60
- Davidson, J. E. H. et al (1978). Econometric modelling of the aggregate time-series relationship between consumers' expenditure and income in the United Kingdom, Economic Journal, 88(352), 661-92
- Deloitte (2006). *Türkiye elektrik enerjisi piyasası gelişmeler beklentiler Rapor*, İstanbul, accessible at on September 27, 2007, http://www.deloitte.com/dtt/cda/doc/content/turkey%28tr%29_er_Turkiye _%20elektrik_enerjisi_piyasasi_041206.pdf
- Derman, T. (2006) Ekonomik Krizleri Tahmin Modeliniz var mı, *Kaynak Elektrik Dergisi*, 209, 79
- Dickey, D.A. and Fuller, W. A. (1979) Distribution of the estimators for autoregressive time series with a unit root, Journal of the American Statistical Association, 74, 427-431
- Dickey, D.A., Hasza, D.P. and Fuller, W.A. (1984). Testing for unit roots in seasonal time series, *Journal of the American Statistical Association*, 79, 355-367

- Dunlap, N. (2006) *Seasonal variation in the cross-price effects of natural gas and electricity,* accessible at March 21, 2007, http://dnichols.wustl.edu/ honors2006/NDunlap%20Thesis.pdf
- Duran, S. and Şahin, A. (2006) İMKB Hizmetler, Mali, Sınai ve Teknoloji Endeksleri Arasındaki İlişkinin Belirlenmesi, *Sosyal Bilimler Araştırmaları Dergisi*, 1, 57-70
- Ediger V. and Akar S. (2007) ARIMA forecasting of primary energy demand by fuel in Turkey, *Energy Policy,* 35, 1701–1708
- EIA (2001) *Electricity market module of the NEMS, Model documentation report*, US Department of Energy Washington, DOE/EIA-M068(2001)
- EIA (2006a), *Country analysis briefs, Turkey, background* , Accessible at December 10, 2007
  http://www.eia.doe.gov/cabs/Turkey/Background.html
- EIA (2006b), *Country analysis briefs, Turkey, oil* , accessible at December 10, 2007, http://www.eia.doe.gov/cabs/Turkey/ Oil.html
- EIA (2006c), *Country analysis briefs, Turkey, coal* , Naturalgas accessible at December 10, 2007, http://www.eia.doe.gov/ cabs/Turkey/Coal.html
- EIA (2006d), *Country analysis briefs, Turkey, naturalgas* , Naturalgas Accessible at December 10, 2007, http://www.eia.doe.gov/cabs/ Turkey/Naturalgas.html
- Electricity Market Balance and Settlement Regulation , Published in the Official Gazette No. 25632 dated 03/11/ 2004
- Electricity Market Eligible Consumer Regulation, Published in the Official Gazette No. 24866 dated 04/09/ 2002
- Electricity Market Law, Law No: 4628, Ratification Date: 20/02/2001, Enactment Date: 03/03/2001
- EMO (2006) *Mobil Santral Nedir? Ne Değildir?,* accessible at December 17, 2007, http://www.emo.org.tr/resimler/ekler/ a172e964907a97d_ ek.pdf?dergi=1, 2006

- Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation, *Econometrica,* 50, 987-1007.
- Engle, R.F. and Granger, C.W.J (1987). Cointegration and error correction: Representation, Estimation and Testing, *Econometrica,* 55, 251-276
- Engle, R.F., Lilien D., Robins R. (1987), Estimating time varying risk premia in the term structure: The ARCH-M model, *Econometrica,* 50,987-1007
- Engle, R.F., Granger, C.W.J and Hallman, J.J. (1989). Merging short and long-run forecasts; an application of seasonal cointegration to monthly electricity sales forecasting, *Journal of Econometrics,* 40, 45-62.
- Erdoğdu, E. (2006a). Energy market reforms in Turkey: An economic analysis, Master Thesis, *University of Surrey, Energy Economics and Policy*, United Kingdom
- Erdoğdu, E. (2006b) Electricity demand analysis using cointegration and ARIMA modeling: A case study of Turkey, *Energy Policy,* 35 1129-1146
- EPDK (2008), Tesis, Yakıt Tipine Göre Lisanslar, Accessible at on February 11, 2008, http://www.epdk.org.tr/lisans/elektrik/lisansdatabase/verilentesistipisorgula.asp
- ERE Hydroelectricity Generation and Trade Corp. (2006) *Info-Memorandum*, Printed by ERE.
- ETKB (2004), *Türkiye uzun dönem elektrik enerjisi talep çalışması raporu,* Ankara, Government Publishing Office
- ETKB General Directorate of Energy Works (2005) , *Enerji Sektöründe Sera Gazı Azaltımı Çalışma Grubu Raporu*, Ankara, accessible at November 10, 2007, http://www.iklim.cevreorman.gov.tr/raporlar/Enerji.pdf

- EUAS (2007). EUAS üretim tesisleri 2006 yılı haziran ayı kümülatif enerji üretim tablosu (kWh), accessible at December 15, 2007, http://www.euas.gov.tr/
- Ferreira N. B., Menezes, R. and Mendes D. A. (2006) *Asymmetric conditional volatility in international stock markets*, Department of Quantitative Methods, IBS-ISCTE Business School, ISCTE, Av. Forcas Armadas, Portugal
- Fisher, F.M. and Kaysen, C.A. (1962). A study in econometrics: The demand for electricity in US, Amsterdam, Nord Holland Publishing Co
- Fouquet, R., Pearson, P., Hawdon D., Robinson, C., Stevens, P. (1997) The future of UK final user energy demand, *Energy Policy,* 25, 231-240.
- TEK (1994) *WASP modeli ile Türkiye uzun dönem üretim-tüketim incelemesi (1996-2010)*, Ankara, Government Publishing Office
- TSO of Ireland (2002), A new methodology for forecasting long term electricity demand for the Republic of Ireland, Ireland, Public Document
- Glosten, L. R., Jaganathan, R., Runkle D. (1993) On the relation between the expected value and the volatility of the normal excess return on stocks, *Journal of Finance,* 48,1779-1801
- Goerten, J. and Clement, E. (2007), Electricity prices for EU households and industrial consumers on 1 January 2007, Statistics in Focus, Environment and Energy, accessible at August 23, 2007, http://epp.eurostat.ec.europa.eu/cache/ITY_OFFPUB/KS-SF-07-080/EN/ KS-SF-07-080-EN.PDF, 80
- Gökçe, S. (1991). *Türkiye'de elektrik enerjisi talep tahmin modelleri*, Master's thesis, Hacettepe University , Ankara
- Granger C. W. J. and Newbold P. (1974). Spurious regression in econometrics, *Journal of Econometrics,* 2, 111-120
- Gümüşdere, E. (2004) Theft and losses in Turkish electricity sector; Empirical analysis and implications for tariff design, Sabancı University, İstanbul

- Halıcıoğlu, F. (2007). Residential electricity demand dynamics in Turkey, Energy Economics, 29, 199–210
- Halvorsen, R. (1975). Residential demand for electricity. Rev. Econ. Stat. 57, 12–18.
- Hamilton J. D. and Susmelb, R., (1994) Autoregressive conditional heteroskedasticity and changes in regime, *Journal of Econometrics,* 64, 307-333
- Hamzaçebi, C. and Kutay F. (2004) Yapay sinir ağları ile türkiye elektrik enerjisi tüketiminin 2010 yılına kadar tahmini, *Gazi Üniversitesi Mühendislik Mimarlık Fakültesi Dergisi,* 19, 227-233
- Hasza, D.P., Fuller, W.A., (1982). Testing for nonstationary parameter specifications in seasonal time series models, *The Annals of Statistics,* 10, 1209-1216.
- High Planning Council (17/03/2004), Electrical Energy Market Reform and Privatization Strategy Document
- Hondroyiannis, G. (2004). Estimating residentisl demand for electricity in Greece, *Energy Economics,* 26, 319-334
- Hor, C. L., (2005). Analyzing the impact of weather variables on monthly electricity demand, *IEEE Transactions on Power Systems,* 20 , 2078 - 2085
- Houthakker, H.S., (1951). Some calculations of electricity consumption in Great Britain, *Journal of the Royal Statistical Society*, 114, 249-270
- Houthakker, H.S., Taylor, L.D., 1970. Consumer demand in the United States: analyses and projections., Harvard University Press, Cambridge, MA.
- Hylleberg, S., Engle, R.F., Granger, C.W.J. and Yoo, B.S. (1990). Seasonal integration and cointegration, *Journal of Econometrics,* 44, 215-238
- Imre, E. (2001) Türkiye'de Yap-İşlet-Devret modeli; yasal çatısı, uygulaması, Yüksek Denetleme Kurulu, accessible at January 7 , 2007,

http://www.ydk.gov.tr/seminerler/turkiyede_yid_modeli.htm#t2

- International Atomic Energy Agency, 2006, *Model for Analysis of Energy, Demand* (MAED-2), *User Manual*, Computer Manual Series No. 18, VIENNA, Printed by the IAEA
- IEA (2006a), *Key World Energy Statistics*, accessible at June,18 2007 http://www.iea.org/Textbase/nppdf/free/2006/ Key2006.pdf
- IEA (2006b), *2006 World Energy Outlook,* ISBN: 92-64-10989-7, France, IEA Publishing
- Johansen, S. and Juselius, K (1990) Maximum likelihood estimation and inference on cointegration- with applications to the demand for money, *Oxford Bulletin of Economics and Statistics,* 52, 169-210
- Kadılar, C. (2005), *SPSS uygulamalı zaman serilerine giriş*, Ankara: Hacettepe Üniversitesi Fen Fakültesi İstatistik Bölümü
- Kamerschen D. R., Porter D. V. (2004). The demand for residential, industrial and total electricity: 1973–1998, *Energy Economics*, 26 , 87–100
- Kaştan, Y. (2006) *Enerji Kaynaklarının Türkiye Siyasi Yapısına Etkisi, Bildiriler Kitapçığı* , Turkey 10th Energy Congress, Istanbul
- Keleş, M. S. (2005). *Elektrik enerjisi talep tahminleri ve Türkiye ekonomisine olan etkileri,* Thesis of specialization, Undersecretariat of Treasury, Ankara
- Kırkgül, S. (1993). *Türkiye elektrik enerjisi talep tahmini,* Master's thesis, Yıldız Technical University, İstanbul
- Kızak, V. (2006) Güneş enerjisinde japon rüzgarı, *Ekoloji Magazin Dergisi*, 11
- Kulalı, İ (1997). *Türkiye'de elektrik enerjisi sektörü ve özelleştirme çalışmaları*, Thesis of specialization, State Planning Organization, Ankara
- Kumbaroğlu, G. (2006) Enerji Modellerimiz Geliştirilmelidir, *Global Enerji,* 22, 38-39

- Kutlar, A. (2000), *Ekonometrik zaman serileri teori ve uygulama,* Ankara, Baran Ofset
- Kwiatkowski, D., Phillips, P.C.B., Schmidt, P. and Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root, *Journal of Econometrics,* 54, 159-178.
- Larsen, B. M. and Nesbakken, R. How to quantify household electricity end-use consumption, *Research Department Discussion Papers* No. 346, *Statistics*, Norway
- Laukkanen, A.(2004) The use of special day information in a demand forecasting model for Nordic power market, Mat-2.108, *Independent research projects in applied mathematics*, Systems Analysis Laboratory, Helsinki University of Technology.
- Law on building and operating electrical power plants and regulating electricity sales with Build-Operate Model, Law No:4283, Ratification date: 16/7/1997, Enactment Date: 29/08/1997
- Law on utilization of renewable energy resources for the purpose of generating electrical energy, Law No. 5346, Ratification Date: 10.05.2005, Enactment Date: 18.05.2005
- Mamun M.A., Nagasaka K. (2004). Artificial neural networks applied to long-term electricity demand forecasting, *IEEE Computer Society,* Proceedings of the Fourth International Conference on Hybrid Intelligent Systems
- McSharry, P. E., Bouwman, S. and Bloemhof, G., (2005). Probabilistic forecasts of the magnitude and timing of peak electricity demand*, IEEE Transactions on Power Systems*, 20, 1166-1172
- Meetamehra, A. (2002) *Demand Forecasting for Electricity,* accessible at December 18, 2006, http://www.teriin.org/pub/papers/ft30.pdf
- Mount, T. D., Chapman, L. D. and Tyrrell, T. J. (1973). *Electricity Demand in the United States: An Econometric Analysis*, Oak Ridge Natural Laboratory, Oak Ridge

- Mozumder, P. and Marathe, A., (2005). Causality relationship between electricity consumption and GDP in Bangladesh, *Energy Policy,* 35, 395-402
- Murray, P. M., Spann, R., Pulley, L., Beauvais, E. (1978). The demand for electricity in Virginia, *The Review of Economics and Statistics*, 60, 585-600
- Narayan, P.K. and Smyth , R. (2005), The residential demand for electricity in Australia: an application of the bounds testing approach to cointegration, *Energy Policy*, 33 , 467–474,
- Nelson, D. B. (1991) Conditional heteroskedasticity in asset returns: A new approach, *Econometrica,* 59, 347-370
- Özdoğan, S. and Arıkol, M. (1995). Energy and exergy analyses of selected Turkish industries, *Energy*, 20, 73-80
- Parti, M. and Parti, C. (1980). The total and appliance-specific conditional demand for electricity in the household sector. *Bell Journal of Economics,* 11, 309–321.
- Phillips, P. C. B. and Perron, P. (1988). Testing for a unit root in time series regression, *Biometrika,* 75, 335-46
- Regulation Concerning Electricity Demand Forecast, Published in the Official Gazette No. 26129 dated 04/04/2006
- Quantitative Microsoftware, LLC (2000) Eviews 4 User's Guide, USA
- Quintanilla, J., Aguilar, V., Serrato, G., Ortega, R., Martín del Campo, C., Mar, E., Conde, L.A. , Gutiérrez, S. (2002) , *ENPEP model and* CO_2 *emissions in Mexico*, Proceedings of the US Economic, Environmental and Energy Modelling Workshop, Mexico
- Rhys, J. M. W. (1984). Techniques for forecasting electricity demand. *The Statistician*, 33, 23-33
- Said, S. E. and Dickey, D. (1984). Testing for unit roots in autoregressive-moving average models of unknown order, *Biometrika,* 71, 599-607

- Sakaryalı, S., Sevgör, S., Erdoğ, F., and Yıldıran, M. (2000), *Impacts modeli ile uzun döneml enerji kaynakları arz-talep planlama çalışmalarında çevre faktörünün incelenmesi,* Türkiye 8. Enerji Kongresi Tebliğleri Vol. 1, p. 40, Dünya Enerji Türk Milli Komitesi, Ankara
- Sanlı, B. (October 23, 2007), *İnternet tarayıcısı üzerinde çalışabilen elektrik talep tahmini analiz programı,* TMMOB Turkey 6th Energy Symposium, Ankara
- Siklos, P. L. and. Wohar, M.E (1996) Cointegration and the term structure: a multicountry comparison, *International Review of Economics and Finance*, 5, 2 l-34
- Silk, J. I. and Joutz, F. L. (1997). Short and long-run elasticities in US residential electricity demand: a co-integration approach*, Energy Economics,* 19, 493-513
- Skinner, N. H. (1984). Load research and ıts application to electricity demand forecasting, *The Statistician,* 33, 65-73
- Song K. B. , Ha S. K., Park J. W. , Kweon D. J. And Kim K. H. (2006) Hybrid load forecasting method with analysis of temperature sensitivity, *IEEE Transactions on Power Systems,* 21
- Starodubtsev, I. (2006) Reform of the turkish electrical energy sector: basic principles and interim results, accessible at November 10,2007, http://www.dundee.ac.uk/cepmlp/journal/html/Vol17/article17_13.php
- Statsoft (n.d) *Time series analysis*, accessible at October 23, 2006, http://statsoft.com/textbook/sttimser.html#arima
- Şengün, M. (1994). *Türkiye Elektrik Enerjisi Ekonometrik Talep Tahmini,* Master's thesis, İstanbul University, İstanbul
- Syed H. I. and Saleh M. A. (1997) Principles of electricity demand forecasting. II. Applications, *Power Engineering Journal*, 11, 91-95
- Tak, S. (2002) *Elektrik enerjisi talep tahmin metotları ve Türkiye için ekonometrik bir uygulama,* Master's thesis, Atatürk University, Erzurum

- Taylor, J. W., Menezes, L.M. and McSharry, P. E. (2006). A comparison of univariate methods for forecasting electricity demand up to a day ahead, *International Journal of Forecasting,* 22, 1-16.
- TEDAS, (2007) ,Map for Electricity Distribution, Accessible at November 24, 2007 http://www.tedas.gov.tr/158,Elektrik_Dagitim_Harita.html
- TEIAS (2006), *Electricity production and transmission statistics of Turkey*, accessible at July 15, 2007, http://www.teias.gov.tr/ istatistik2005/index.htm
- TEIAS (2007a) Purpose and Activity Field of the Establishment, accessible at November 24, 2007, www.teias.gov.tr
- TEIAS (2007b), a, *Türkiye Elektrik Enerjisi 10 Yıllık Üretim Kapasite Projeksiyonu (2007-2016)*, Ankara, Government Printing Office
- TEIAS (2007c) *Saatlik system marjinal fiyatları*, accessible at December 2, 2007, https://pmum.teias.gov.tr/UzlasmaWeb/
- Thang, K. F. (2004) MATLAB@ implementation of neural and neuro-fuzzy approaches for short-term electricity demand forecasting, 2004 International Conference on Power System Technology – POWERCON, Singapore
- TMMOB Chamber of Mechanical Engineers (2006)*, Enerji politikaları yerli, yeni ve yenilenebilir enerji kaynakları raporu,* No: MMO/2006/417
- TOBB (2007), *Invaluable Potential of Turkey, A Smarter , cleaner and competitive way of generating electricity*, TOBB Printing Office, Ankara
- Toptaş, M. (1992). *A comparative time series study for medium term electricity demand forecasting for Turkey,* Master's thesis, Middle East Technical University, Ankara
- TUIK (2005a), *Statistical Indicators 1923-2004*, Vol.0535, Ankara, TUIK Printing Office
- TUIK (2005b) Türkiye İstatistik Yıllığı 2004, iklim ve arazi, accessible at February 10, 2007, http://www.die.gov.tr/yillik/01_Arazi.pdf

- TUIK (2007a) *Haber Bülteni*, accessible at January 23, 2008, http://www.tuik.gov.tr/PreHaberBultenleri.do?id=464.
- TUIK (2007b), *Consumer Price Index data*, accessible at January 28, 2007, http://www.tuik.gov.tr/VeriBilgi.do
- TUIK (2007c), *Population data*, accessible at January 28, 2007, http://www.tuik.gov.tr/VeriBilgi.do
- Turkish Central Bank (2007) *Industry production index data,* Retrieved June 26, 2007 www.tcmb.gov.tr
- Turkish State Meteorology Institute (2006), *Isıtma ve Soğutma Gün Dereceleri*, accessible at April 5, 2007, http://www.meteoroloji.gov.tr/2006/zirai/zirai-aylikgunderece.aspx
- Türker, Ö. (1999). *Structural time series modeling of inflation behavior in Turkey,* Master Thesis, Middle East Technical University, Ankara
- Türkiye Elektrik Kurumu Dışındaki Kuruluşların Elektrik Üretimi, İletimi, Dağıtımı ve Ticareti ile Görevlendirilmesi Hakkında Kanun, Law No: 3096 , Ratification date: 4/12/1984
- UCTE (2007) *About us*, accessible at December 10, 2007, www.ucte.org
- ÜZMEN, R.(2007) Küresel iklim değişikliğinde enerji üretiminin rolü ve nükleer enerji, accessible at February 13, 2007, http://www.enerjiajansi.com/index.php?option=com_contentandtask=viewandid=29andItemid=48
- Walasek, R.A. (1981). Regional variations in electricity demand elasticities: The situation in the central United States, *GeoJoumal Supplementary*, 3 , 37-47
- Wikipedia (2007a). *Time Series*, accessible at January 5, 2007, http://en.wikipedia.org/wiki/Time_series
- Wikipedia (2007b). *Gauss-Markov Theorem*, accessible at January 27, 2008, http://en.wikipedia.org/wiki/Gauss-Markov_theorem

- Wikipedia (2007c). *Exponential Growth,* accessible at February 8, 2007, http://en.wikipedia.org/wiki/Exponential_growth
- Wikipedia (2007d). *Heating Degree Day,* accessible at April 5, 2007, http://en.wikipedia.org/wiki/Heating_degree_day
- Wilson, J.W., (1971). Residential demand for electricity, *Quarterly Review of* Economics and Business, 11, 7-22
- Wong, Y. K. and Rad, A.B. (1998). *Gaussmarkov models for forecastıng and rısk evaluation,* IEEE Catalogue No: 98EX1370-7803-4495-2/98/$10.00 IEEE Department of Electrical Engineering Hong Kong Polytechnic University Hunghom, Hong Kong
- Teknik Yayıncılık (2006) International Energy and Environmental Technology Systems Conference Energy Statistics Book, Teknik Yayıncılık A.Ş., İstanbul
- Wothington A. C., Spratley A. K. and Higgs H. (2005) Transmission of prices and price volatility in Australian electricity spot markets: A multivariate GARCH analysis, *University of Wollongoy*, Faculty of Commerce Papers
- Yiğitgüden, Y. (1999). *Türkiye Elektrik Enerjisi Sektöründe Özelleştirme Politikaları ve Çalışmaları*, Istanbul, Midas Basım Yayın Tanıtım
- Zhang, F. (2007). An application of vector GARCH model in semiconductor demand planning, *European Journal of Operational Research*, 181, 288–297
- 9 Eylül University (1998). *1998 Dünya panoramasi ve Türkiye,* accessible at October 23, 2007, http://abegitim.org/balkir/tur/gazete/kategori1/1998dunya_eko_panorama.pdf

APPENDIX A
PRIMARY ENERGY RESOURCES SUPPLY IN TURKEY

YEAR	Coal (1000 ton)	Lignite (1000 ton)	Asphaltit (1000 ton)	Petroleum (1000 ton)	Natural gas (10 6 m3)	Hydraulic (GWh)
1970	4.573	5.782	36	3.542		3.033
1971	4.639	6.222	23	3.452		2.610
1972	4.641	7.342	168	3.388		3.204
1973	4.642	7.754	289	3.511		2.603
1974	4.965	8.354	394	3.309		3.356
1975	4.813	9.150	456	3.095		5.904
1976	4.632	11.146	443	2.595	15	8.375
1977	4.405	12.176	434	2.713	18	8.572
1978	4.295	15.122	297	2.736	22	9.335
1979	4.051	13.127	203	2.831	34	10.289
1980	3.598	14.469	558	2.330	23	11.348
1981	3.970	16.476	560	2.363	16	12.616
1982	4.008	17.804	860	2.333	45	14.167
1983	3.539	20.956	750	2.203	8	11.343
1984	3.632	26.115	225	2.087	40	13.426
1985	3.605	35.869	523	2.110	68	12.045
1986	3.526	42.284	607	2.393	457	11.873
1987	3.461	42.896	631	2.630	297	18.618
1988	3.256	35.338	624	2.564	99	28.950
1989	3.038	48.762	416	2.876	174	17.940
1990	2.745	44.407	276	3.717	212	23.148
1991	2.762	43.207	139	4.451	203	22.683
1992	2.830	48.388	213	4.281	198	26.568
1993	2.789	45.685	86	3.892	200	33.951
1994	2.839	51.533		3.687	200	30.586
1995	2.248	52.758	67	3.516	182	35.541
1996	2.441	53.888	34	3.500	206	40.475
1997	2.513	57.387	29	3.457	253	39.816
1998	2.156	65.204	23	3.224	565	42.229
1999	1.990	65.019	29	2.940	731	34.678
2000	2.392	60.854	22	2.749	639	30.879
2001	2.494	59.572	31	2.551	312	24.010
2002	2.319	51.660	5	2.442	378	33.684
2003	2.059	46.168	336	2.375	561	35.330
2004	1.946	43.709	722	2.276	708	26.084

APPENDIX A (cont'd)
PRIMARY ENERGY RESOURCES SUPPLY IN TURKEY

YEAR	Geothermal		Wind	Solar	Wood	Waste
	Elec.(GWh)	Heat(10^3TEP)	(GWh)	(1000 TEP)	(1000 ton)	(1000 ton)
1970		23			12.816	9.253
1971		38			12.189	9.316
1972		38			13.503	9.514
1973		48			13.847	9.807
1974		50			14.500	10.088
1975		56			14.562	10.495
1976		58			14.734	11.002
1977		58			14.989	11.276
1978		60			15.248	11.750
1979		60			15.506	12.258
1980		60			15.765	12.839
1981		60			16.023	12.689
1982		82			16.760	12.607
1983		100			17.086	12.748
1984	22	178			17.256	11.978
1985	6	232			17.368	11.039
1986	44	304		5	17.570	11.343
1987	58	324		10	17.693	11.059
1988	68	340		13	17.711	10.987
1989	63	342		19	17.815	10.885
1990	80	364		28	17.870	8.030
1991	81	365		41	17.971	7.918
1992	70	388		60	18.070	7.772
1993	78	400		88	18.171	7.377
1994	79	415		129	18.272	7.074
1995	86	437		143	18.374	6.765
1996	84	471		159	18.374	6.666
1997	83	531		179	18.374	6.575
1998	85	582	6	210	18.374	6.396
1999	81	618	21	236	17.642	6.184
2000	76	648	33	262	16.938	5.981
2001	90	687	62	287	16.263	5.790
2002	105	730	48	318	15.614	5.609
2003	89	784	61	350	14.991	5.439
2004	93	811	58	375	14.393	5.278

APPENDIX B
ENERGY RESOURCES CONSUMPTION IN TURKEY

YEAR	Coal (1000 ton)	Lignite (1000 ton)	Asphaltit (1000 ton)	Petroleum (1000 ton)	Natural gas (106 m3)	Hydraulic (GWh)	Wind (GWh)
1970	4727	5772	36	7579		3033	
1971	4651	6376	23	8819		2610	
1972	4638	7355	168	10215		3204	
1973	4595	7642	289	11995		2603	
1974	5031	8188	394	12132		3356	
1975	4959	8973	456	13503		5904	
1976	5005	10998	443	14992	15	8375	
1977	5057	11675	434	17230	18	8572	
1978	4696	13235	297	17010	22	9335	
1979	4898	13882	203	14796	34	10289	
1980	4630	15243	558	15309	23	11348	
1981	4522	16179	560	15090	16	12616	
1982	5044	17716	861	16127	45	14167	
1983	5336	20663	750	16705	8	11343	
1984	5678	25632	225	16990	40	13426	
1985	6189	34767	523	17270	68	12045	
1986	6545	52354	607	18688	457	11873	
1987	7220	40653	631	21239	735	18618	
1988	7525	33080	624	21302	1225	28950	
1989	6825	47557	409	21732	3162	17940	
1990	8191	45891	287	22700	3418	23148	
1991	8824	48851	139	22113	4205	22683	
1992	8841	50659	197	23660	4612	26568	
1993	8544	46086	102	27074	5088	33951	
1994	8192	51178	0	25859	5408	30586	
1995	8548	52405	66	27918	6937	35541	
1996	10892	54961	34	29604	8114	40475	
1997	12537	50474	29	29176	10072	39816	
1998	13146	64504	23	29022	10648	42229	6
1999	11362	64049	29	28862	12902	34678	21
2000	15525	64384	22	31072	15086	30879	33
2001	11176	61010	31	29661	16339	24010	62
2002	13830	52039	5	29776	17694	33684	48
2003	17535	46051	336	30669	21374	35330	61
2004	18904	44823	722	31729	22446	46084	58

APPENDIX B (cont'd)
ENERGY RESOURCES CONSUMPTION IN TURKEY

YEAR	Geothermal Electricity (GWh)	Geothermal Heat (10^3TEP)	Solar (1000 TEP)	Wood (1000 ton)	Waste (1000 ton)	Electricity Import (GWh)	Electricity Export (GWh)
1970		23		12816	9253		
1971		38		12189	9316		
1972		38		13503	9514		
1973		48		13847	9807		
1974		50		14500	10088		
1975		56		14562	10495	96	
1976		58		14734	11002	332	
1977		58		14989	11276	492	
1978		60		15248	11750	621	
1979		60		15506	12258	1044	
1980		60		15765	12839	1341	
1981		60		16023	12689	1616	
1982		82		16760	12607	1773	
1983		100		17086	12748	2221	
1984	22	178		17256	11978	2653	
1985	6	232		17368	11039	2142	
1986	44	304	5	17570	11343	777	
1987	58	324	10	17693	11059	572	
1988	68	340	13	17711	10987	381	
1989	63	342	19	17815	10885	559	
1990	80	364	28	17870	8030	176	907
1991	81	365	41	17971	7918	759	506
1992	70	388	60	18070	7772	189	314
1993	78	400	88	18171	7377	213	589
1994	79	415	129	18272	7074	31	570
1995	86	437	143	18374	6765	0	696
1996	84	471	159	18374	6666	270	343
1997	83	531	179	18374	6575	2492	271
1998	85	582	210	18374	6396	3299	298
1999	81	618	236	17642	6184	2330	285
2000	76	648	262	16938	5981	3791	437
2001	90	687	287	16263	5790	4579	433
2002	105	730	318	15614	5609	3588	435
2003	89	784	350	14991	5439	1158	588
2004	93	811	375	14393	5278	464	1144