FORECASTING OF SHORT TERM AND MID TERM ISTANBUL NATURAL GAS COMSUMPTION VALUES BY NEURAL NETWORK ALGORITHMS

by

Recep KIZILASLAN

JUNE 2008

FORECASTING OF SHORT TERM AND MID TERM ISTANBUL NATURAL GAS COMSUMPTION VALUES BY NEURAL NETWORK ALGORITHMS

by

Recep KIZILASLAN

A thesis submitted to

the Graduate Institute of Sciences and Engineering

of

Fatih University

in partial fulfillment of the requirements for the degree of

Master of Science

in

Industrial Engineering

June 2008
Istanbul, Turkey

APPROVAL PAGE

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

Prof. Dr. Mazhar Ünsal
Head of Department

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

Prof. Dr. Bekir KARLIK
Supervisor

Examining Committee Members

Prof. Dr. Bekir KARLIK..

Prof. Dr. Mazhar ÜNSAL..

Assist. Prof. Dr. Fatih CAMCI...

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

Assist. Prof. Dr. Nurullah ARSLAN
Director

June 2008

FORECASTING OF SHORT TERM AND MID TERM ISTANBUL NATURAL GAS COMSUMPTION VALUES BY NEURAL NETWORK ALGORITHMS

Recep KIZILASLAN

M.S. Thesis – Industrial Engineering
June 2008

Supervisor: Prof. Dr. Bekir KARLIK

ABSTRACT

The aim of this study is to find a suitable natural gas energy forecasting model for hourly, daily, weekly and monthly values by using artificial neural networks(ANN). As it is known, accurate forecasting is important for both gas distributors and consumers. On the view point of distributors, with accurate forecasting the number of false alarms would be significantly decreased and tranship limits would be scheduled.By this way gas systems would be more reliable and profitable. Although accurate forecasting values are good for general consumers there will be no disconnect and breakdown etc. In this study wide factor analyzing study is done in order to find the factors that effects the gas consumptions. Founded results were applied to ANN feed forward back propogation algorithm models.The reasons behind choosing ANN are the ability of ANN to forecast future values of more than one variable at the same time and to model the nonlinear relation in the data structure. Seven different algorithm models were used and comperison of their performance were done.

Key Words: Artificial Neural Network, Forecasting, Natural Gas

KISA VE ORTA VADELİ İSTANBUL DOĞAL GAZ TÜKETİMİNİN YAPAY SİNİR AĞLARI ALGORİTMALARI İLE TAHMİN EDİLMESİ

Recep KIZILASLAN

Yüksek Lisans Tezi – Endüstri Mühendisliği
Haziran 2008

Tez Yöneticisi: Prof. Dr. Bekir KARLIK

ÖZ

Bu calışmada yapay sinir ağları algoritmaları kullanılarak saatlik, günlük, haftalık ve aylık doğal gaz enerjisi tahmini için uygun modellerin bulunması amaçlanmıştır. Bilindiği üzere doğal gaz tüketiminin doğru tahmin edilmesi hem gaz dağıtıcıları hem de tüketicileri açısından önemlidir. Doğal gaz dağıtıcısı açısından incelendiğinde, doğru tahmin değerleri sistemde olusabilecek hataları azaltır ve gaz dagıtım limitlerinin doğru şekilde programlanabilir hale getirir. Bu sayede gaz sistemleri cok daha gerçekçi ve karlı hale gelir. Ayrıca tüketici açısından bakıldığında doğru tahmin değerleri sistemde oluşabilecek hatalari azaltacağından ve bu sayede dogal gaz kesintisi olmayacağından dolayı iyidir.

Bu calışmada öncelikle gaz tüketimine etki eden faktörleri belirlemek icin kapsamlı faktor analizi calışması gerçekleştirilmistir. Bulunan sonuçlar geri beslemeli yapay sinir ağları modellerine uygulanmıştır. Yapay sinir ağlari tahmin yöntemi uygulamalarının kullanılmasının amacı algoritmaların doğrusal olmayan verilerin modellemesinde iyi sonuçlar vermesi ve birden fazla tahminin ayni anda yapılabilmesidir. Yedi değisik yapay sinir ağları algoritması uygulanıp bunların karşılastırılması yapılmıştır.

Anahtar Kelimeler: Yapay Sinir Ağları, Tahmin, Doğalgaz

DEDICATION

Dedicated to my parents for their endless support and patience during the forming phase of this thesis.

The men who moves a mountain

begins with carrying away

small stones.

Confucius

ACKNOWLEDGEMENT

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

I also want to thank to Prof. Dr. Mazhar Ünsal, Assist. Prof. Dr. Fatih CAMCI and Assist. Prof. Dr. Özgür UYSAL for their significant contributions to my research endeavors.

I would also like to thank to The Graduate Institute of Sciences and Engineering of Fatih University, IGDAS and IBB for funding my thesis.

I express my thanks and appreciation to my family for their understanding, motivation and patience. Lastly, but in no sense the least, I am thankful to all colleagues and friends who made me stay at the university a memorable and valuable experience.

TABLE OF CONTENTS

LIST OF FIGURES

FIGURE

LIST OF TABLES

TABLE

CHAPTER 1

INTRODUCTION

Istanbul is located between Europe and Asia. This geographical location makes it a natural bridge connecting Europe and Asia. In 2008, the population of Istanbul is 12.573.836, 41,86 % over the 1990 level which is 7309190 people. The average population growth rate was 2.46 % per year between 1990 and and 2008, the highest among the International Energy Agency (IAE) member countries` metropol cities. It is predicted that these population growth rates will not continue with these rates but however the population will reach almost 15000000 by 2020.

In Turkey the economy has undergone a significant shift from agriculture towards the service sector and to some extend industry, although some 30% (%43 in 1993) of the active population was still employed in agriculture in 2003 [1]. Istanbul is rapidly growing in terms of both economy and its population. At the same time its energy demand, actually for natural gas and electricity, has been increasing rapidly. Natural gas energy is very important energy for not only industry but also general consumers like residental consumers, schools, hotels, etc. So natural gas energy is vital input for technical, social and economic development of Turkey as the other countries.

In 1990s energy consumption increased about 4,5 % per year but this value is also increased to higher rates because of the economic development in Turkey actually between 2002 and 2008. Demand of natural gas energy has the most growing rate than the other energy sources like electricity, fuel oil, petroleum, coal etc. Especially residental users give natural gas energy preference over another energy sources for heating and hot water because of its ease-of-use, cheapness and being environment-friendly.

This preference to natural gas energy get some problems like accurate demand forecasting and some operating problems. The prediction of natural gas consumption is not only very important for gas distributions (IGDAS) but also important for transportation companies as well as government agencies associated to the natural gas sectors. Particularly, short range prediction like hourly, 1 to 5 days is important to ensure the normal supply of Istanbul. This type of prediction is very important for countries actually like Turkey which do not have its own natural gas energy sources and where the production sites are far from the centers of consumption. As a result of these, working on forecasting techniques that gives accurate forecast values is very important for demand planners. In this study, we are studying on forecasting of natural gas consumption in short, mid and long term periods by applying artificial neural network algorithms which required sufficient historical data.

1.1.Why Prediction of Natural Gas Consumption Is Important?

Natural gas energy plays an increasingly significant role in economy activities. *Nearly* 2/3 world energy consumption comes from crude oil and natural gas. The world’s largest traded commodity, $500 billion/year. Sharp oil price movements are likely to disturb aggregate economic activity[2].

Pipeline systems in natural gas companies are able to accumulate gas in the pipelines which allows storing enough gas in the pipelines. However accurate forecasting would allow lower pressure in the pipelines and raising production at right time[3]. During increased winter consumption there is a significant danger of a system breakdown and some of the consumers even have to be disconnect. With accurate the forecasting number of false alarms would be significantly decreased and gas delivery limitations would be scheduled. Forecasting the long-term, mid-term and short-term natural gas consumption values would make the gas system more reliable and profitable. Although accurate short-term and mid-term forecasting values are good for general consumers (especially residental consumers) because there will be no disconnect and breakdown, besides long-term forecasting values are good for natural gas companies because of low price, happy customers, low operating costs etc.

Distributors of natural gas need to predict future natural gas consumption in order to purchase a sufficient supply on contract. Distributors that offer their customers equal payment plans need to predict the consumption of each customer, twelve months in advance. None equal payment plan customers are expected to pay for their consumption as billed monthly. However, as a cost-saving measure, their meters are usually only read bimonthly. Estimates based on previous consumption patterns are used for months without readings. Often, for customers whose meters are usually inaccessible, e.g. indoor meters, meters buried under snow, etc., only two or three irregular readings per year are taken. In such cases, the distributor relies on estimates most of the time.

If accurate demand and supply forecasting values do not calculated we could not guarantee that there won't be any natural gas shortages expressly in winters because the situation is very uncertain especially in Istanbul. All of the consumers, actually residential consumers are expected to get their supplies; some industrial users with interruptible contracts may have to switch to other energy sources if demand increases further and supplies remain tight or shut down operations. As a result of these shutdowns, employees would likely be laid off. The tightness in natural gas supply could affect consumers directly through their heating and cooling bills. In the longer-term, it could also result in job losses, as manufacturers that are heavy users of gas move to countries where natural gas is readily available or cheaper.

1.2. Historical Background

Neural network techniques have been recently suggested for short-term, midterm or long term load forecasting by a large number of researchers. Asar *et al.*, Bakirtzis *et al.*, Mohammed *et al.*, A. Khotanzad *et al.*,has worked on a specification of neural network applications in the load forecasting problems [6-13].

R. H. Brown *et al.*, has developed an artificial neural-network models to predict daily gas consumption.

As the models presented in the literature are either intended for forecasting the whole daily load curve at once, or forecasting the load weekly, this division is used here in testing different models. This combination of several forecasting techniques were used and focused on minimizing the variance of the forecast errors. Alireza Khotanzad *et al.* [10-11] investigated several linear and non-linear combination approaches including ANN-based models. Nguyen and Mandziuk have worked several approaches to prediction of natural gas consumption with neural and fuzzy neural systems for a certain region of Poland [12]. Musilek *et al.* [13] has tried to solve the problem of seasonal dependency with a recurrent neural network used as a gate for a statistical mixture model.

On the application of short term forecasting like hourly or daily electric energy consumption forecasting artificial neural network techniques were used several times. Yalcinoz and Eminoglu, 2005 [17], Hsu and Chen, 2003 [18]; Beccali et al., 2004 [19]; Khotanzad et al., 1995 [20] and Khotanzad et al., 1996[21]; Chow and Leung, 1996 [26]; Hobbs et al., 1998 [27]; Lee and Park, 1992 [28]; Mohammed et al., 1995 [29] and for medium term forecasting Yalcinoz and Eminoglu, 2005 [17]; Azadeh et al., 2006a[22], Azadeh et al., 2006b[23], Azadeh et al., 2007a[24] and Azadeh et al., 2007b[25]. This is because of the ANN's ability to learn and construct a complex non-linear mapping through a set of input output examples.

CHAPTER 2

2.1. Structure of Natural Gas Consumption in Istanbul

2.1.1. Structure and Network System of Natural Gas in Istanbul

Istanbul is divided into two parts and has about 12 millions population. IGDAS distributes natural gas at city center, Silivri and Sile. Meantime two private company distributes natural gas at two small regions within city borders. There is no pipeline that distributes natural gas between the two parts of the city therefore there are two distinct networks. There is no difference between two networks in terms of the design, management or materials. Special a type of provision station is being used for Silivri and Sile. The map below shows the natural gas distribution of the Istanbul [4].

Figure 2.1. Natural gas distribution licenses in Istanbul and regions that uses natural gas

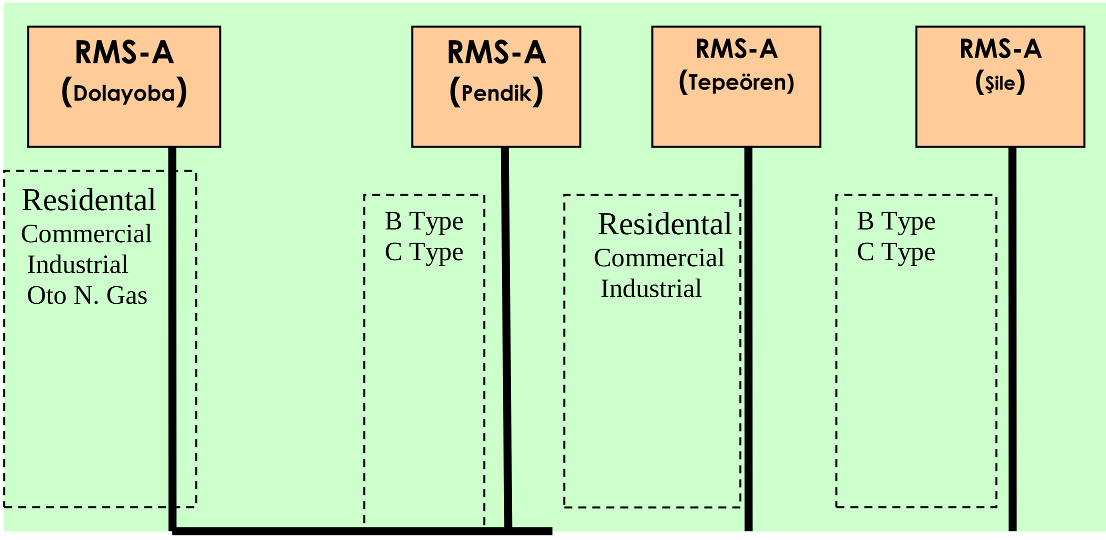


Figure.2.2. European side gas supplier stations(City Entrance-RMS)

There are eleven natural gas entrance points to the city; seven of which from European part and the rest from Anatolian part. These entrance points are being shown as dots in figure. Basic consumption sources are being shown in the schema. This is important in understanding the characteristic of consumption. More than 3 millions of people consume 4.7 billions m^3 natural gas per year (2006) in various activities. The use of natural gas is increasing and becoming widespread continuously up to now. We are expecting a decrease in this trend as we think that natural gas has been distributed to even rural areas [4].

Table 2.1. Data of March in 2007

Number of subscriber	3.378.564
Number of Gas Customer	3.039.063
Number of Gas Counter	2.288.129
Length of Steel Network	1.140.587
Length of PE Network	10.585.532
Number of BR	626
Number of Industrial customer	604
Number of Service Box	521.031

There are 10 entrance points over main distribution line where IGDAS distributes natural gas. Datas about these entrance points according to the total capacities has been summerised in the table below (BOTAS data). As seen there is 56 millions m^3 total natural gas input per day to distribution network.

Table2.2. Capacity of natural gas entrance point of Istanbul in 2007

	Maximum Capacity(Sm³/day)	Total Reserve Capacity (Sm³/day)	Unused Capacity (Sm³/day)
Igdas (Esenyurt)-1 RM/A	9.600.000	8.942.000	658.000
Igdas (Esenyurt)-2 RM/A	4.800.000	4.800.000	0
Igdas (Esenyurt)-3 RM/A	7.200.000	5.695.000	1.505.000
Igdas Ambarlı RM/A	7.200.000	0	7.200.000
Igdas RM/A	2.520.000	898.693	1.621.307
Igdas (Silivri) RM/A	2.160.000	52.838	2.107.162
European Side	33.480.000	20.388.531	13.091.469
Igdas (Pendik)-1 RM/A	9.600.000	6.019.000	3.581.000
Igdas (Pendik)-2 RM/A	7.200.000	5.205.000	1.995.000
Igdas (Tuzla) RM/A	1.080.000	144.510	935.490
Igdas -Pendik sahil- RM/A	4.800.000	2.074.000	2.726.000
Anatolian side	22.680.000	13.442.510	9.237.490
Istanbul Total	56.160.000	33.831.041	22.328.959

2.1.2. General

Since 20 years natural gas has been distributed increasingly in Istanbul. It is hard to distribute natural gas in such a complex and populous city. Under gas scarcity conditions excessive gas consumer and under excessive gas conditions need to proliferation on gas consumer are encountered continuously. The amount of gas consumption has always been important. Recently price pressure caused by annual; daily and seasional gas consumption has been important. Decreasing civil protection caused by freedom on natural gas market has become important for gas distributors and sellers. As wholesale gas sale gets free from public control these conditions are inclined to be more important. Grasping the operating supplies secondary benefits in the determination of the way the work done (companies, consumption balancing, balancing points, Network integration, season dependant management etc.). In this study the analysis of natural gas consumption of Istanbul has been done by taking the whole components of gas consumption into consideration. In this analysis, information about the function or calculation scope that can be used in determination of the ultimate value has been given.

2.1.2.1. Change in the numbers of the subscription and gas consumers

Number of gas consumer has a dynamic character that always change. This change alters according to infrastructure, campaigns and regional, cultural and social parameters. Gas usage is a important compound of preference dependent and programs that appeals people to use. Preference dependence depends on social and cultural characteristics. Appealing programs consist of infrastructure development programs, campaigns, collaboration of other institutions [4]. The whole change is not in a linear relationship. However it can be figured out by a approximate correlation in mathematics .

2.1.2.1.1. Number of Gas Customers in Regional Bases

Istanbul is divided into three parts according to active wire and consumer services management.Change in the numbers of gas consumers between 1992 and 2006 is given in the graph below. Change in the numbers of gas consumers differs from each others.

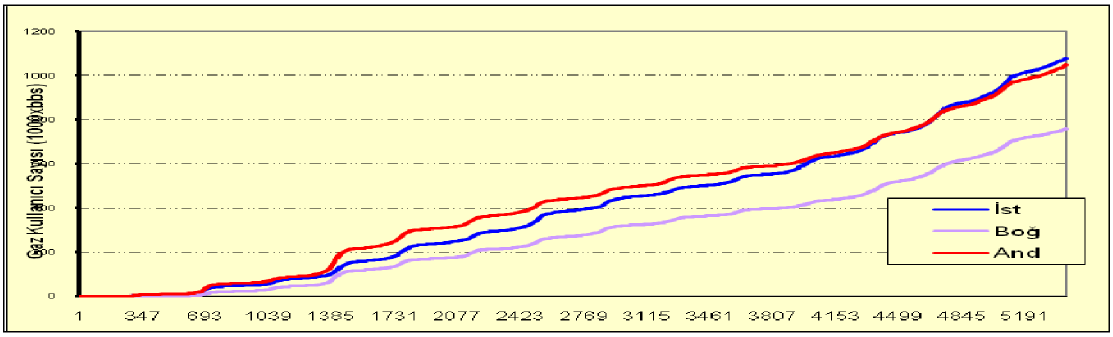


Figure 2.3. Changing of Natural Gas Customer Number Between 1992-2006

Graph shows that there is not a certain parallel change in the numbers of the gas consumers. Especially change trend in the Anatolian and Istanbul regions is more unstable. This is more related to size of the new gas catchment areas of regions. Change trend in Beyoglu is more linear.

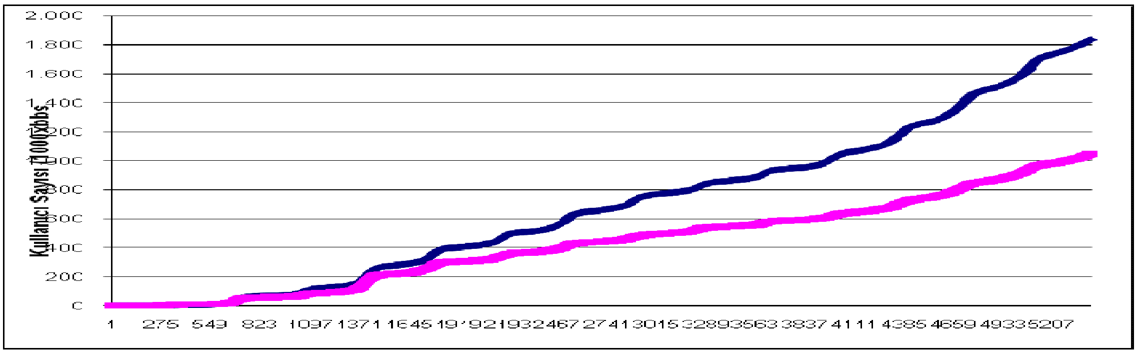


Figure 2.4. Changing of Natural Gas Customer Number of European and Anatolian Side Between 1992-2006

Number of natural gas consumers and subscribers are increasing in time gradually. Number of gas subscribers, gas consumers and change in trend curves of three regions have been indicated in Table 4.2. There is %2 deviation in the similarity of the change trends.this analysis is important for data reliability.

Change in the number of gas consumers for Anatolian and European parts between 2001-2006 has been given in table4.3. Mathematical equations that resemble these curves indicates that change in European side(**y = 496,15x + 714109**) is about twice as much as change in Anatolian side **(y = 250,35x + 470333)**
(((**496,15/250,35=1,98**))) [4].

In the later process we should examine how much these change trends can preserve themselves. The factors that affect gas consume can be enumerated in importance as temperature, number of gas consumers, rate of gas consumers that uses it for production, regional-demographic factors. Consume rates must be specified in this manner. Although there are tables according to consumer types, it is hard to recognize temperature factors and regional conditions in detailed on these tables. Therefore making some advanced investigations that enhance factor effectiveness give rise to making more accurate predictions and obtaining some precise characteristics of city's gas consume behaviors.

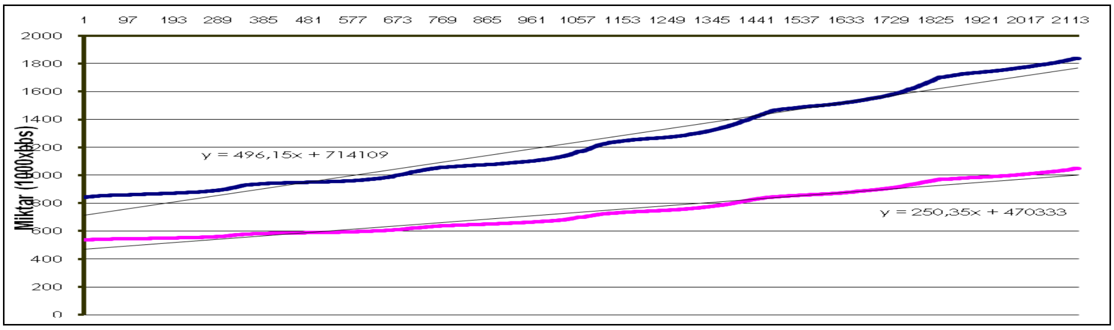


Figure 2.5. Changing of Gas Customer Number Between 2001–2006 (for Anatolian And European sides)

2.1.3. Calculations of Monthly Residental Gas Consumptions

2.1.3.1 General Calculations and Analysis

Finding out the average gas consume according to unit consumer, gives basic provides value for later analyses. Especially in a dynamic structure increase in the number of steady gas consumer affects the bigness of consume value related to gas usage. Therefore, in this study many evaluations have been done according to unit subscriber gas consume value. Tables and graphs below indicates the recent average gas consume per subscriber per month. As we expect stationary gas consume per unit subscriber in summer term graphs shows that is not. In this situation we should examine the indefinite conditions related to the compounds of the analysis [4]

Table 2.3. Monthly average consumptions for per customer between 1995-2006 (m^3/month)

	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	Average
January	281,8	321,6	259,8	240,6	240,6	285,2	195,6	223,7	192,1	241,0	214,9	259,1	**246,3**
February	222,0	287,8	259,4	221,7	229,4	212,3	182,5	132,0	242,5	212,8	215,3	213,3	**219,3**
March	208,2	315,8	266,2	255,4	175,2	193,1	109,9	139,4	225,0	168,4	195,9	175,3	**202,3**
April	125,8	180,5	188,7	68,2	78,5	68,0	83,0	115,2	141,7	96,3	95,5	88,1	**110,8**
May	39,8	17,2	34,5	44,9	28,9	35,1	29,6	24,0	24,3	35,7	35,2	50,8	**33,3**
June	16,1	12,4	15,8	11,2	9,1	13,7	16,2	13,2	16,3	21,2	20,6	21,4	**15,6**
July	14,6	9,1	6,8	13,0	8,4	8,3	12,7	10,7	13,2	12,4	17,1	19,6	**12,1**
August	13,8	10,1	10,1	9,9	4,7	8,5	10,7	12,1	13,3	15,4	16,7	18,0	**11,9**
September	15,8	11,4	13,9	11,3	9,0	13,6	13,2	13,2	19,6	19,6	17,2	20,6	**14,9**
October	77,9	78,2	82,9	34,3	44,4	56,8	35,7	25,3	48,3	31,7	71,3	71,3	**54,8**
November	214,8	134,6	139,5	147,5	135,3	84,4	142,5	88,4	130,3	118,7	164,7	164,7	**138,8**
December	268,1	209,1	217,9	261,1	168,4	175,0	254,0	209,7	206,0	189,5	194,0	194,0	**212,2**
Average	124,9	132,3	124,6	109,9	94,3	96,2	90,5	83,9	106,1	96,9	104,9	108,0	

Monthly gas consumption for 1995-2006 term has been shown in the table. To this, gas consumption ranges from minimum 4,7 m^3/month to maximum 4,7 m^3/month. Average monthly gas consumption value ranges from 83.9 to 132.3 in year. This value is important for prediction. Lowest consumption occurs in August as 11,9 m^3/month and highest consumption in January as 246,3 m^3/month according to the monthly consumption that is observed for years.

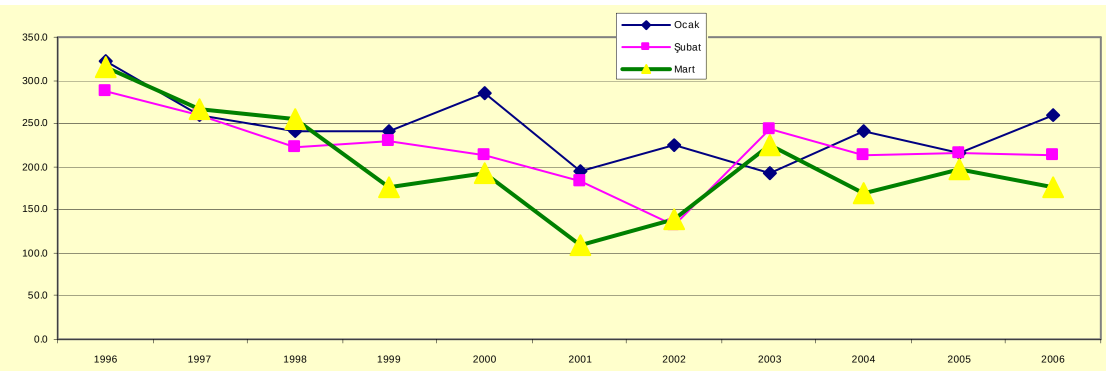


Figure 2.6. Average gas consumption for per customer in 1996-2006 January-February-March

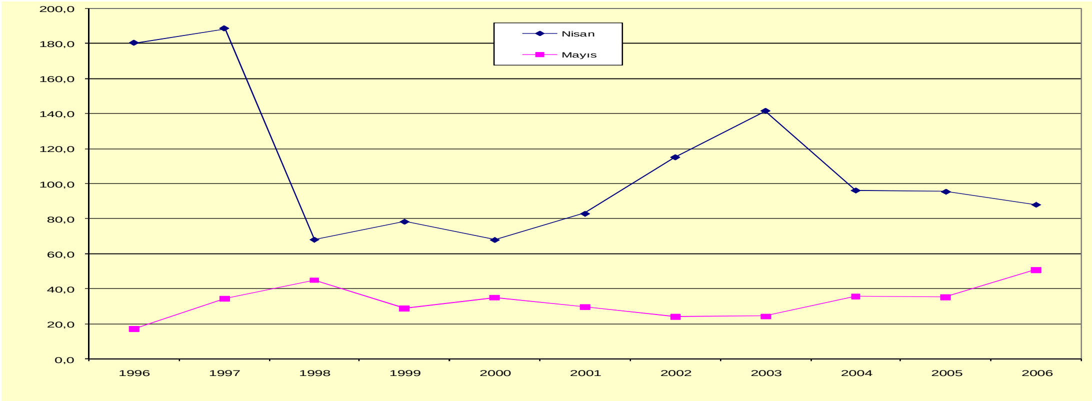


Figure 2.7. Average gas consumption for per customer in 1996-2006 April and May

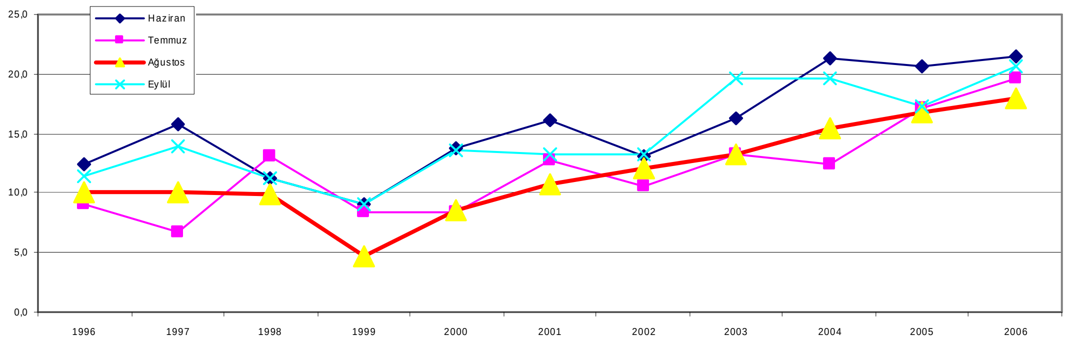


Figure 2.8. Average gas consumption for per customer in 1996-2006 June-July-August and September

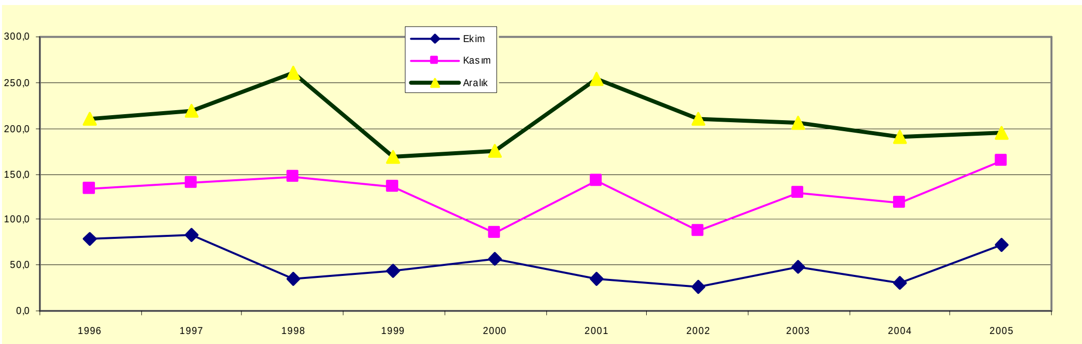


Figure 2.9. Average gas consumption for per customer in 1996-2006 October-November and December

A decreasing change in gas consumption is being recognized as three monthly changes between 1996-2006 term observed. These changes will be elaborated in heat consumption connections.

As these graphs analysed remarkable terms listed below:

1- Climatic effects are top in winters.

2- Change per average unit gas consumer alters every years (even in summers).

3-Average gas consumption amount changes even in summers. This change stems from gas usage habits

4-There is a close relationship between economical situations and usage habits. Gas usage decreased noteworthly especially in 2001 and 2002.

5-There is a wide difference even in summers in which a linear change expected. This especially makes it hard to determine the consumption amount of unit gas consumer.

6-Wide differences in summer term variations indicates that there are some factors that are independent from measurement. [4]

All these entries as forming a real correlation in determining a real value, lead important deviations determining the real correlations because they increase vagueness of the parameters. As data reliability increases forming accurate relationships will be much more possible. First calculation related to this has been made by using the correlations that have been gained by the summation of the average monthly consumption values in yearly base. Informations about this calculation are given below.

As an average method, to calculate the approximate household gas consumption dependant on the last ten years monthly based tendencies. Basic formulas (correlations) related to this is given in the table below. These correlations refer to the average changes of last ten years' mathematical data [4].

Table 2.4. Formulations of monthly consumption values. (for 1995 X=0 , 1996 x=1 etc.)

January	y = -5,9013x + 278,53	July	y = 0,9461x + 6,249
February	y = -5,7277x + 253,37	August	y = 0,9424x + 6,1156
March	y = -10,902x + 267,2	September	y = 0,9509x + 9,0887
April	y = -5,2414x + 140,88	October	y = -2,5994x + 65,199
May	y = 1,1189x + 26,021	November	y = -0,263x + 130,02
June	y = 0,9886x + 9,625	December	y = -2,4724x + 222,07

2.1.3.2. Consumption Rates –Summer Months

A comparison of consumption values for summers has been done and summarized above. As seen in table there is remarkable deviations in average monthly consumptions. Consumption amounts are very unstable in summers. Therefore it is hard to produce a common and high valid statement. In order to produce more valid statements a detailed investigation must be done over different data systems.

This graph is important because it remarks that climate independent conditions are very important on consumption. That is to say the effects that form consumption can vary in socio-economic and political conditions. Consumptions widely varies in Istanbul which have high domestic activities. Namely in mathematical expression, as a closed function, city can give different outputs.

2.1.4. Evaluation

Above graphs and tables shows that there are many undefined things taht effects the consumption of natural gas. In the last ten years these values are hardly inclined to change profoundly for unit gas consumers. However recent data show that there is a change in big oscillation. Graph shows the earthquake dependent effects and economical crises dependent consequences for last ten years clearly [4]. With the economical development of 2004 and 2005 consumption amounts are increasing. Therefore there is not a linear relation between data. It is hard to form a long term prediction as we think about the summer data. Meantime as commercial subscribers considered in general gas consumers; it must be predicted to estimate how much of the activities and changes stem from these subscribers.

2.1.4.1 Evaluation of Residental Consumption

Istanbul gas distribution has been founded and developed basicly for dwellings. This trend has not changed for years. To this number of subscribers and seasional characteristics are considered as the basic components. Other than this as explained economical crises has always been important on consumption. Yet it is better to make estimations without considering this in general physical components. Predicting these effects make it easier to determine.

Table 2.5. Natural gas consumption rates between 2000-2006 years

	2000	2001	2002	2003	2004	2005	2006
Per-1 (January-March)	867.657	682.720	758.267	1.123.980	1.231.593	1.461.737	1.749.554
Per-2 (May-August)	102.067	117.464	113.958	152.789	214.921	259.021	361.771
Per-3 (September-December)	356.355	598.722	494.888	639.693	698.199	941.373	0

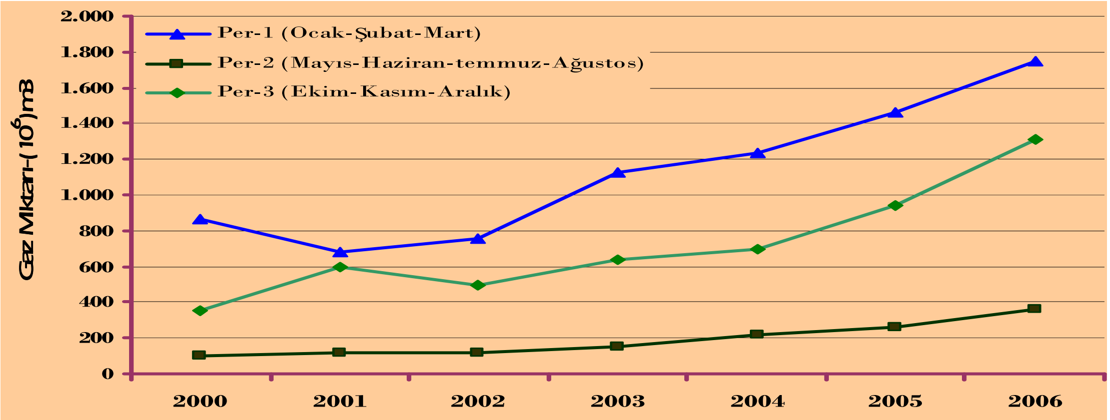


Figure 2.11. Graph of three period gas consumption changes between 2000-2006

Table 2.5. and Figure 2.10. indicate the used gas amounts between 2000-2006. According to this domestic gas consumption of Istanbul has been increasing steadily except the break in 2001and 2002. First period is January-February-March term ,second one is summer term May-June-July-August and the third one is winter term October-November –December. Trends in these periods are differ from each others and not occured in the same amont with each others. (Table-7). At this context this situation makes it hard to predict better. This change is more related to temprature changes. Besides periodical changes in gas usage transition can be efective

Table 2.6. Per-1/Per-3 ratio table

	2000	2001	2002	2003	2004	2005	2006
Per-1/Per-3	2,43	1,14	1,53	1,76	1,76	1,55	1,33

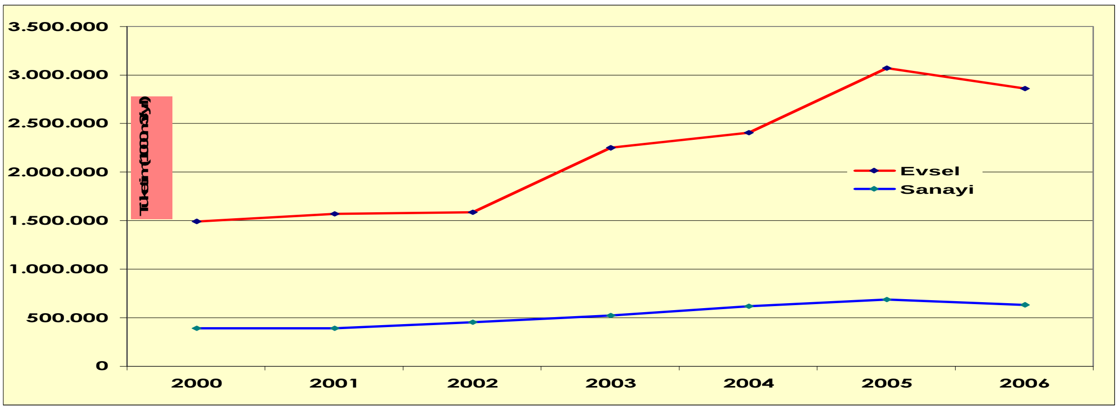


Figure 2.12. Changes of gas consumption (Residental- Commercial)

As seen in the Figure 2.12 gas selling has been linear between the crisis term 2000-2002. It is inclined to encounter with similiar results in the probable crises.

CHAPTER 3

3.1. DATA ANALYSIS

In this work several approaches are analyzed and tested to prediction of natural gas consumption of Istanbul. The data includes hourly, daily, weekly and monthly natural gas consumption and number of customer values in two different region of Istanbul which are Anatolian and European sides. Before applying the prediction methods and strategies, data must be analyzed in order to determine the factors that effects the demand of natural gas.

The most significant factors that effects the natural gas consumption of residential and commercial users are temperature, number of customer, day factor, day of the week (holiday or working day) and previous day gas consumption values etc. Hourly, weekly, daily and monthly values will be analyzed below.

3.1.1. Analysis of Hourly Values

Pipeline systems are able to accumulate gas in the pipelines which allows storing enough gas in the pipelines through the night to cover the morning consumption peeks. However, accurate forecasting would allow lower pressure in the pipelines and raising production at a right time [3]. Actually in winter seasons the consumption rate is very changeable. There may be several reasons of this like changeable behaviour of temperature in day time, number of customer and different profile of customers, different types of consumers (industry, residential, commercial etc.). These factors make it difficult to predict the natural gas consumption value. Because of this changeable behaviour of consumption in winter there would be significant danger of system breakdown and some of the consumers even have to be disconnect [3]. By determining the cause of false alarms by calculating the accurate forecasting values, number of these problems significantly decreased and delivery limitations would easly be scheduled.

Hourly data between 2004 and 2006 years are taken for the analysis. The graph below shows behaviour of the hourly consumption between 2004 and 2006.

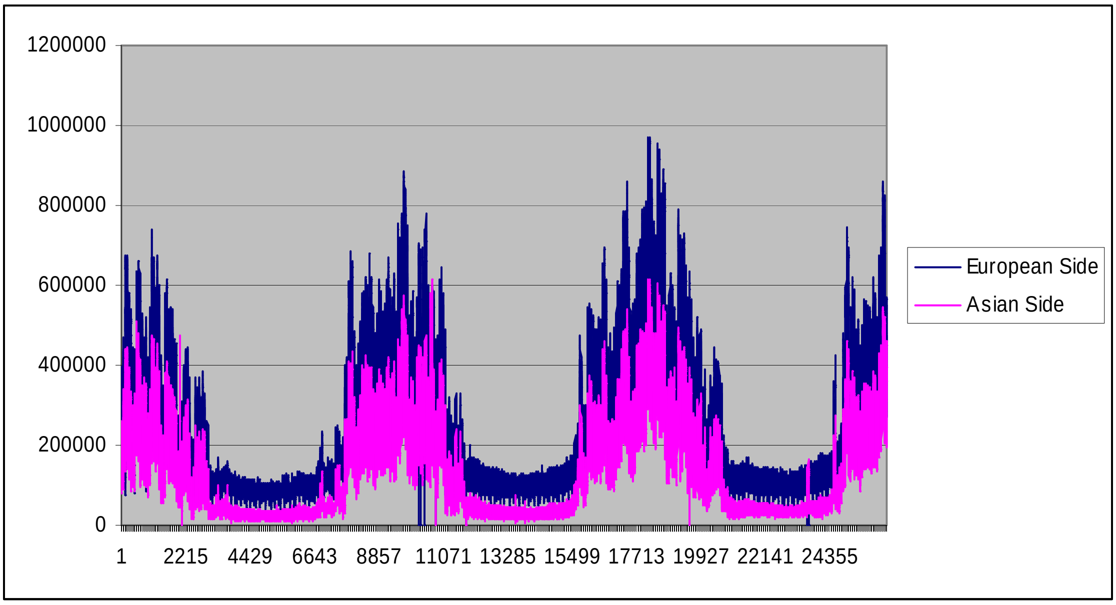


Figure 3.1. Hourly consumption values between 2004-2006

3.1.1.1 ANOVA in Hourly Values (Winter Season)

As mentioned before outside temperature is the most important factor because natural gas is mostly used for heating. There are some other several factors that effects the consumption like wind speed, customer profile, type of combi boiler etc., but all of these factors are extremely hard to predict so including them in the model would only compromise the accuracy of the model and so results. For hourly demand forecasting the factors are not only these that mentioned above but also customers in where they live in a industry region or another normal regions effects the demand of natural gas. Second selected parameters are temperature and hour factors, because a significant difference in consumption can be observed day hours and night hours.

Table 3.1. An example of hourly natural gas consumption values

January-2006	Number of Customer	Hour	Consumption (*1000 m3)	Temperature ©
1	965073	09.00	281,0	5,5
		10.00	277,0	5,7
		11.00	281,0	5,8
		12.00	278,0	5,7
		13.00	263,0	6
		14.00	251,0	6,9
		15.00	239,0	7
		16.00	241,0	6,5
		17.00	255,0	8,7
		18.00	286,0	9,5
		19.00	297,0	11,5
		20.00	305,0	12,2
		21.00	304,0	12
		22.00	297,0	12,5
		23.00	277,0	11,8
		24.00	212,0	10,2
		01.00	167,0	9
		02.00	145,0	8,1
		03.00	143,0	8
		04.00	137,0	7,6
		05.00	141,0	8
		06.00	167,0	8,2
		07.00	288,0	7,6
		08.00	335,0	7,5

An example of natural gas consumption values are shown in the table above. This type of data were taken between 2004 and 2007 years. In order to determine the hour factor for different seasons analysis of variance (ANOVA) test should be applied on the values. We have calculated ANOVA test for the month of different seasons because consumption values in different hours are not same with the month of different seasons. As a result of this we have taken a monthly hour values from a winter season and a month from a summer season. We could take the rest of the month results same with these months' results because consumption behaviours are same with them but they did not be showed here.

Temperature is measured every hour. Temperature values above are decimal so we are divided temperature values in different values in order to determine the consumption rates to different temperature limits. For example, between the temperature values of -4 ^{o}C and -2 ^{o}C are taken as -3 ^{o}C, 7 ^{o}C and 9 ^{o}C as 8 ^{o}C etc.

Analysis of variance test is applied on consumption versus hour and temperature and results are shown below.

Table 3.2. ANOVA tables for consumption versus hour and temperature

General Linear Model: Consumption versus Hour; Temperature

Factor	Type	Levels Values
Hour	**Fixed**	24 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24
Temperature	**Fixed**	11 -3 -1 1 3 5 7 9 11 13 15 17

Analysis of Variance for Consumption, using Adjusted SS for Tests

Source	DF	Seq SS	Adj SS	Adj MS	F	P	
Hour	23	6001072	6016382	261582	144,76	**0,001**	**(Significant)**
Temperature	10	3891953	3891953	389195	215,38	0,001	**(Significant)**
Error	710	1282986	1282986	1807			
Total	743	11176012					

As it is seen in the result, P-value show us that hour and temperature values are very significant.

Average natural gas consumption values according to the hour and temperature limits are as below.

Least Squares Means for Consumption

Hour	**Mean**	**SE Mean**
1	203,6	8,128
2	172,4	8,142
3	156,7	8,203

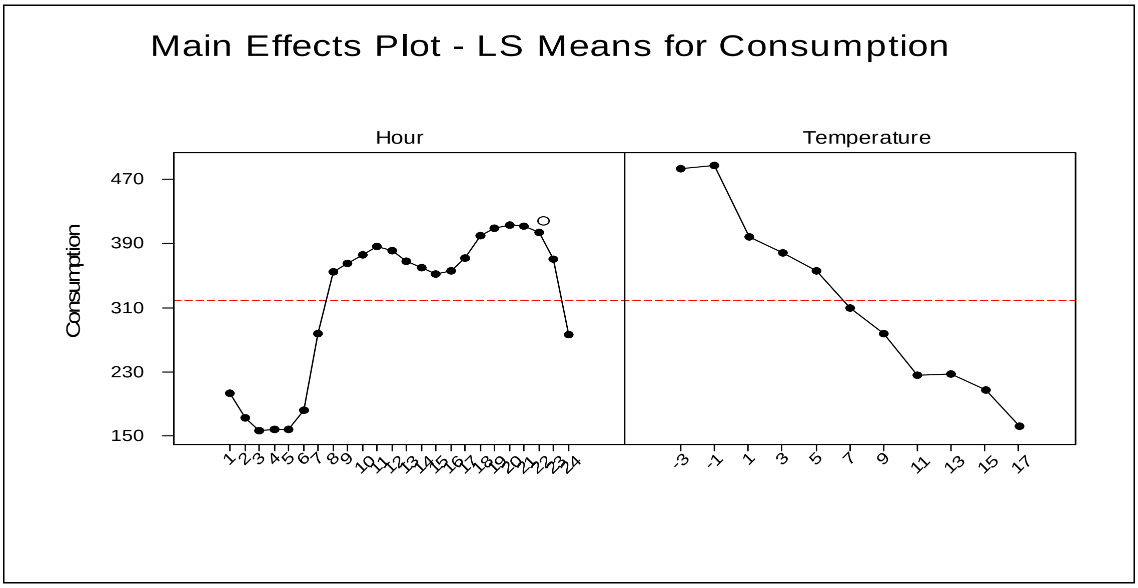
4	157,5	8,181
5	157,8	8,177
6	181,8	8,205
7	277,2	8,182
8	355,0	8,222
9	365,7	8,193
10	375,7	8,170
11	386,2	8,179
12	381,7	8,201
13	368,0	8,223
14	360,6	8,249
15	352,7	8,220
16	355,9	8,203
17	372,4	8,215
18	399,7	8,157
19	409,6	8,108
20	413,7	8,002
21	411,4	7,945
22	404,5	7,917
23	370,5	8,052
24	276,2	8,099

Temperature	***Mean***	***SE Mean***
-3	484,0	6,643
-1	487,1	4,420
1	398,6	4,501
3	378,1	2,775
5	355,8	4,564
7	310,0	5,127
9	277,3	5,687
11	225,7	7,948
13	227,3	10,159
15	207,2	15,263
17	162,6	24,914

As mentioned before, temperature is very significant factor on the natural gas consumption. As the temperature decreases demand is increases and when temperature increases it is the exact opposite of temperature decrease. But as in the figure 3.22 below this is true for only in day times. Despite temperature decrease at nights, consumption is decreases too.

On the graph, consumption values are very high between 07 and 22 hours and after 22 it is decreasing very fast. So in our model, we will attempt weights to the hours in order to define the difference.

Figure 3.2. Main effect plot of consumption versus hour and temperature



3.1.1.2. ANOVA in Hourly Values (Summer Season)

Summer season daily hours' variance analysis is shown below. Temperature values are between 1 ^{o}C and 29 ^{o}C and temperature is very significant factor in our model as it is shown in the ANOVA test results below.

Analysis of variance test is applied on consumption versus hour and temperature and results are shown below.

Table 3.3. ANOVA tables for May-2006 consumption versus hour and temperature

General Linear Model: May06 Demand versus May06 Hours; May06 Temp.

```
Factor        Type   Levels  Values
May06 Hours   fixed      24  1; 2; 3; 4; 5; 6; 7; 8; 9; 10; 11; 12; 13; 14; 15;
                             16; 17; 18; 19; 20; 21; 22; 23; 24
May06 Temp.   fixed      13  1; 7; 9; 11; 13; 15; 17; 19; 21; 23; 25; 27; 29


Analysis of Variance for May06 Demand, using Adjusted SS for Tests

Source        DF    Seq SS    Adj SS  Adj MS      F      P
May06 Hours   23  129424,1  128427,8  5583,8  16,48  0,001  Significant Factor
May06 Temp.   12   26408,8   26408,8  2200,7   6,49  0,001  Significant Factor

Error        708  239949,4  239949,4   338,9
Total        743  395782,2


S = 18,4095   R-Sq = 39,37%   R-Sq(adj) = 36,38%
```

Consumption values, according to the daily hours, summer season values do not have the same trend with winter season hourly consumption values. When demand of natural gas is very high between 07 am and 09 pm in winter season, there is a decreasing trend in summer season's daily hours between 09 am and 8 pm. The results are shown in figure 3.3 and figure 3.4 clearly.

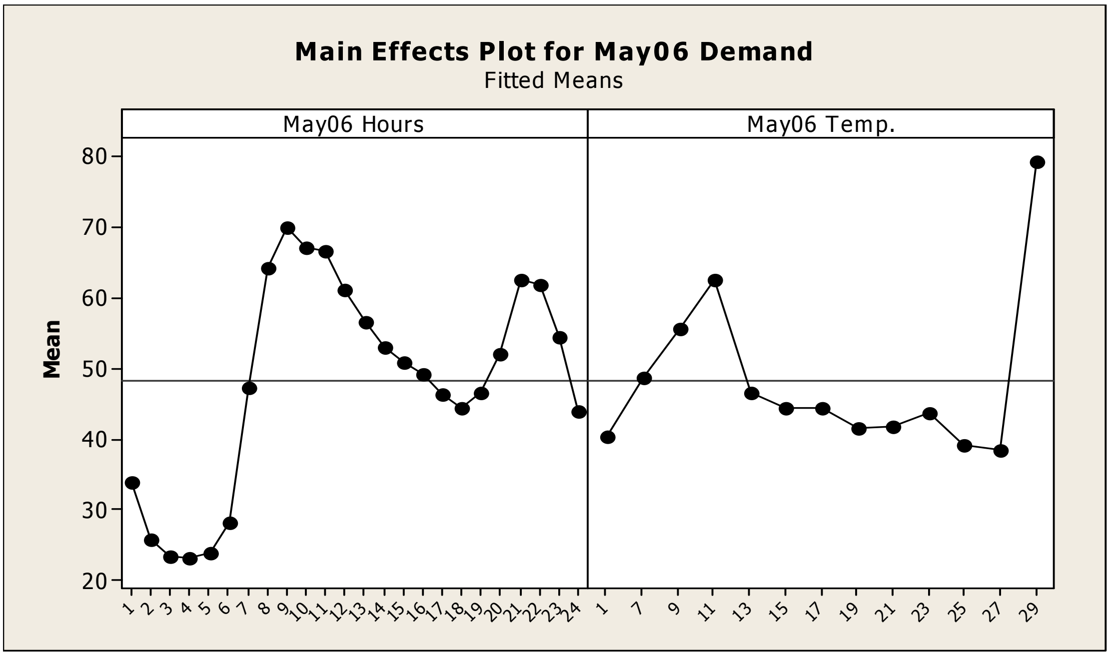


Figure 3.3. Main effect plot of consumption versus hours and temperature for May-2006 hours

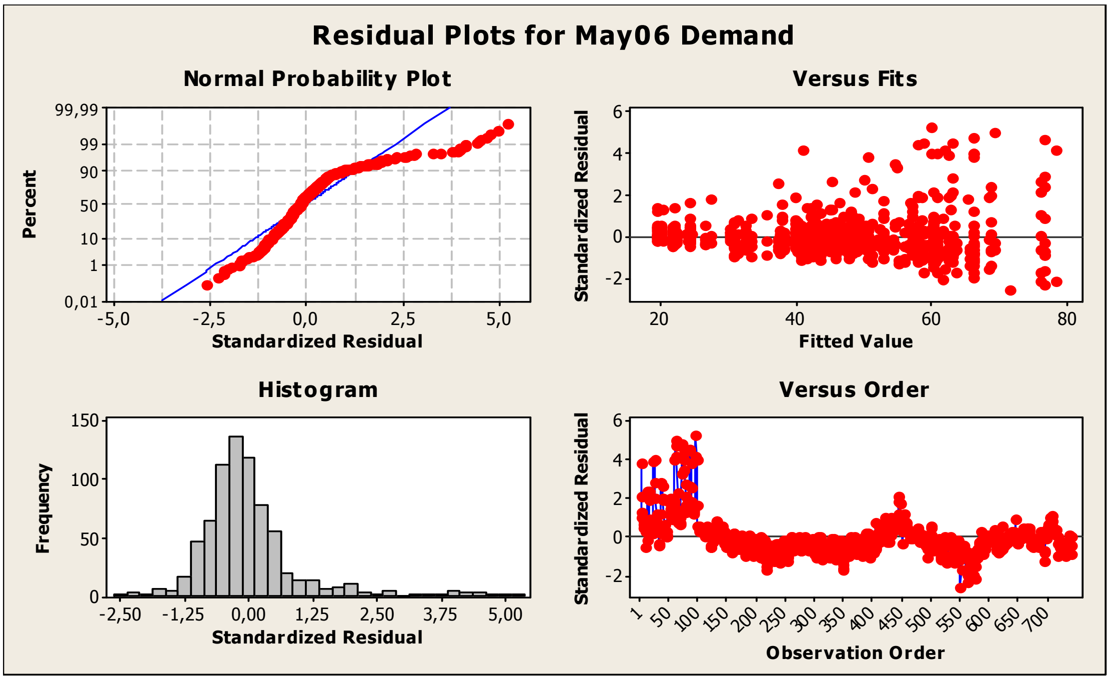


Figure 3.4. Residual effect plot for May-2006 hourly demand

3.1.2. Analysis of Daily Values

Daily natural gas consumption forecasting is another important and required thing that very useful for natural gas companies. Almost until the last years, natural gas supply was considered a public service and any demand forecasting which was undertaken tend to be over the longer term, yearly or more, concerning future demand values or natural gas prices and technical improvements. But nowadays, short-term demand and price forecasting have become increasingly very important because of the rise of the competitive natural gas suppliers markets. Short-term accurate natural gas demand forecasting studies are very important for suppliers and producers in order to maximize not only their profit but also avowing profit loses over the misjudgement of future demand movements. All of these studies are also very essential for consumers to maximize their utilities.

The data covers daily natural gas load in two different regions of Istanbul, Anatolian and European sides independently and both of them together. The graphs below shows the daily natural gas consumption values severally.

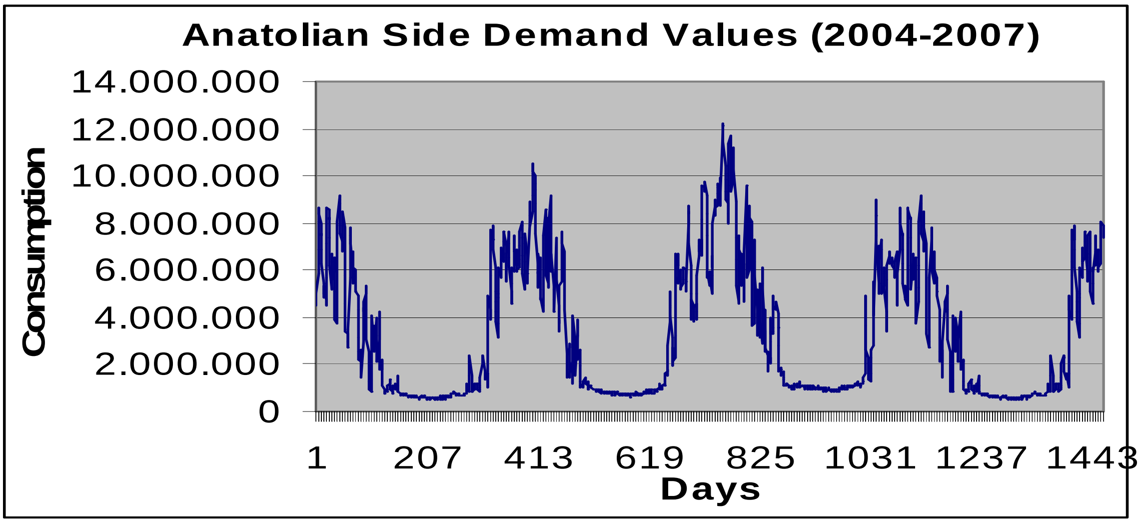


Figure 3.5. Anatolian side natural gas demand values between 2004-2007

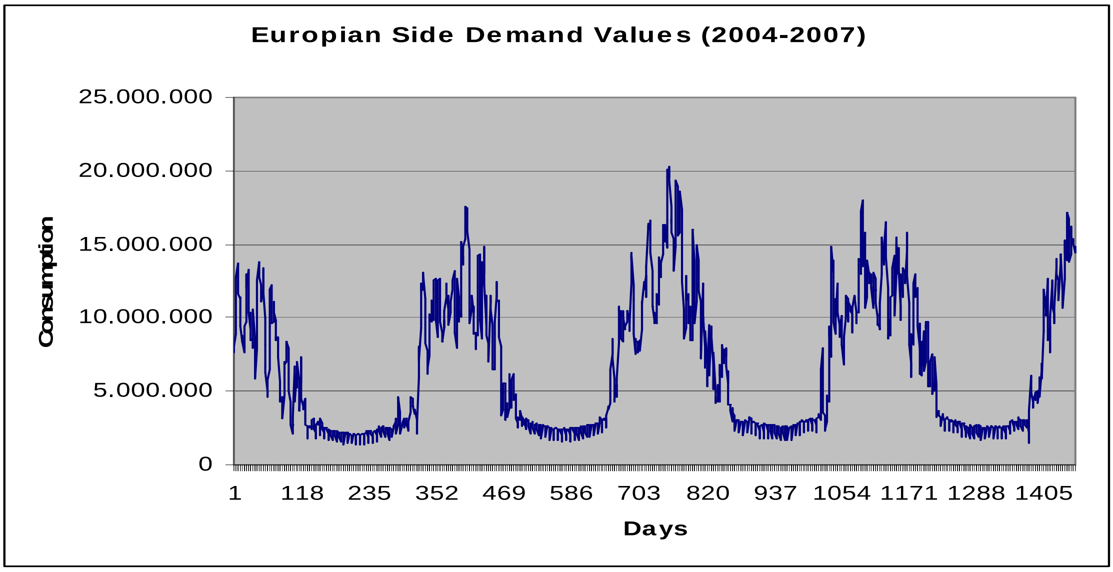


Figure 3.6. European side natural gas demand values between 2004-2007

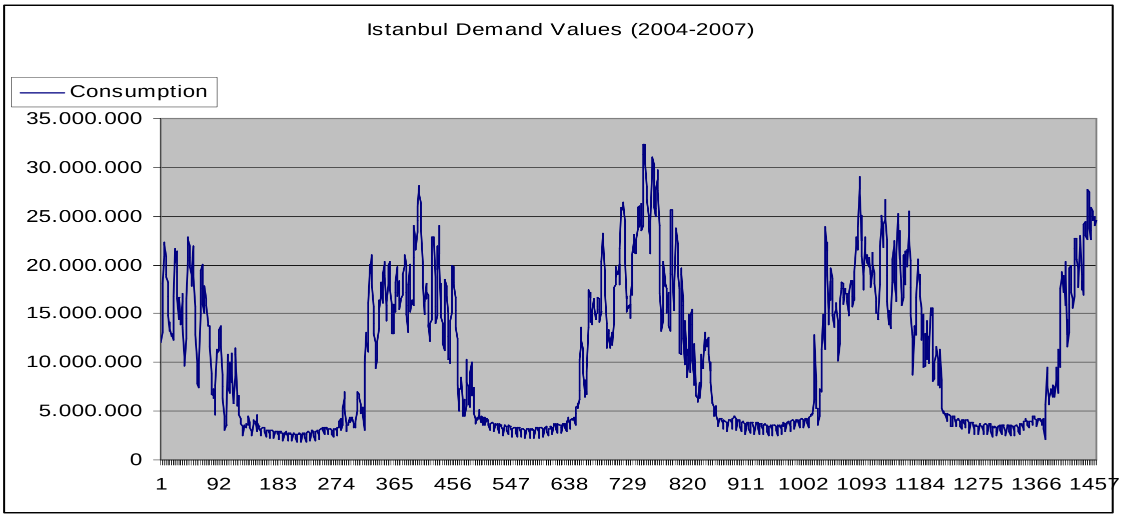


Figure 3.7. Istanbul natural gas demand values between 2004-2007

Consumer number is one of the most significant factor on the natural gas consumption. Figure 3.8 shows that consumer number always increasing day to day.

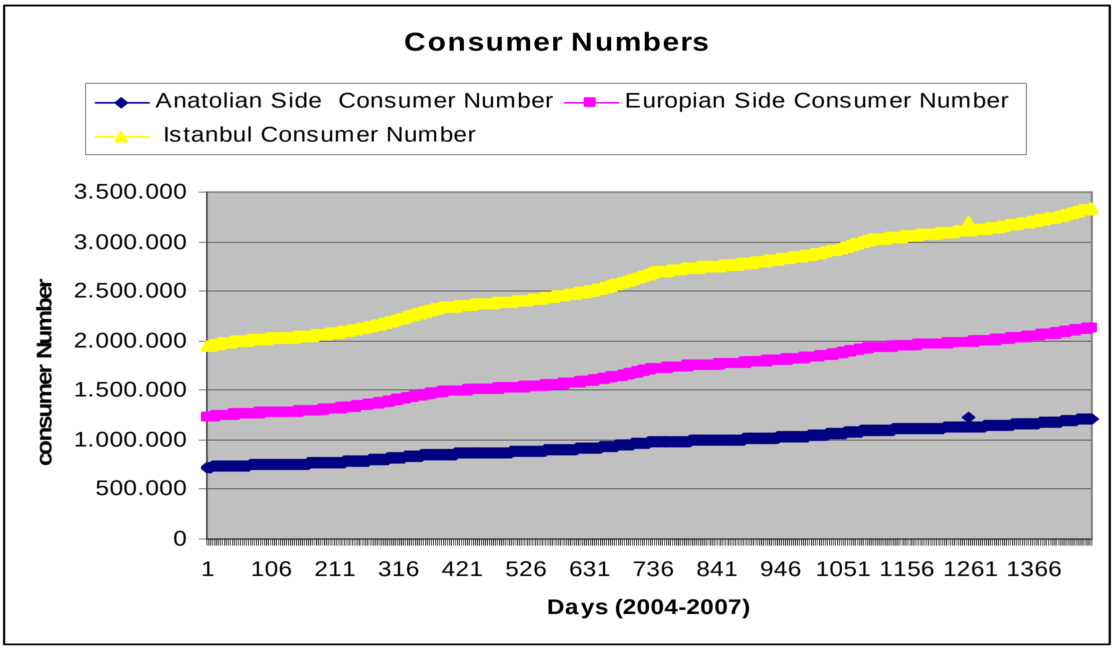


Figure 3.8. Trend of number of consumer for Istanbul and both sides

Trend analysis for number of consumers shows that it will increase to more than 4 million on year 2009 if this increasing trend continue.

Linear trend model equations are as below.

$Y(t) = 1208264 + 628*t$ (European side)

$Y(t)= 712635 + 336*t$ (Anatolian side)

The equations shows that slope of the European side is more than Anatolian side. According to these equations consumer number of European side will reach 2,6 million and Anatolian side will reach 1,5 million by 2009 years. Trend analysis plot, figure 3.9, is also shows the forecasted customer number by the year 2009.

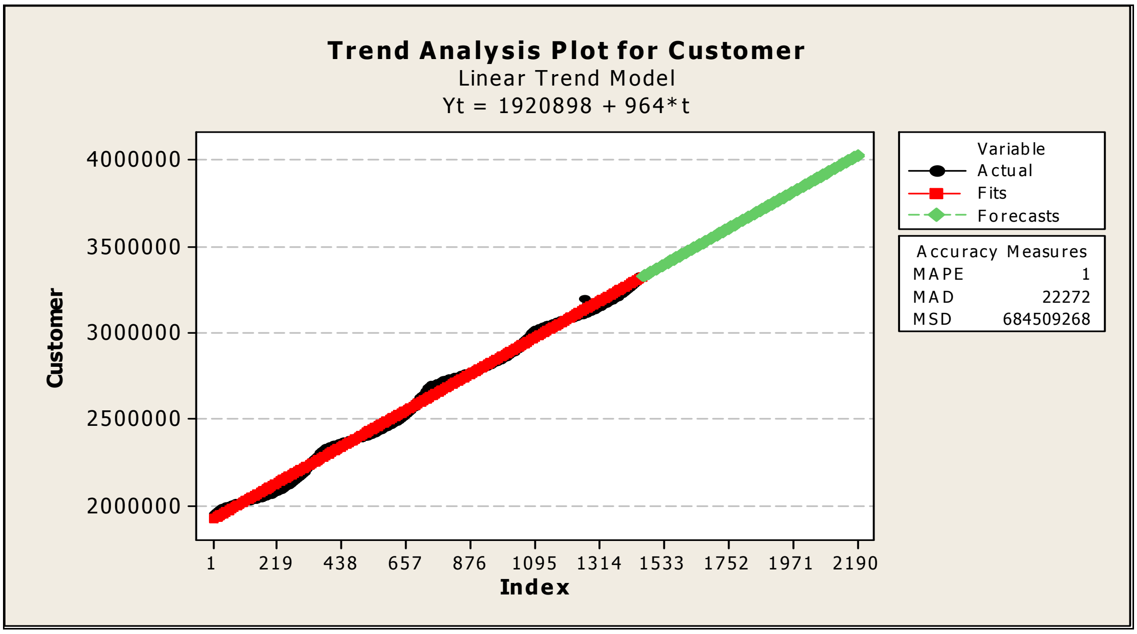


Figure 3.9. Trend analysis plot of consumer number for Istanbul

Trend analysis for number of consumers show us that it will increase to more than 4 million on year 2009 if this increasing trend continue.

As it is mentioned before, temperature is very important factor on natural gas energy consumption values. In order to define the factors that effects the consumption we have to consider the temperature values and illustrate the movements of it. The graph below shows the daily temperature values between years of 2004-2007.

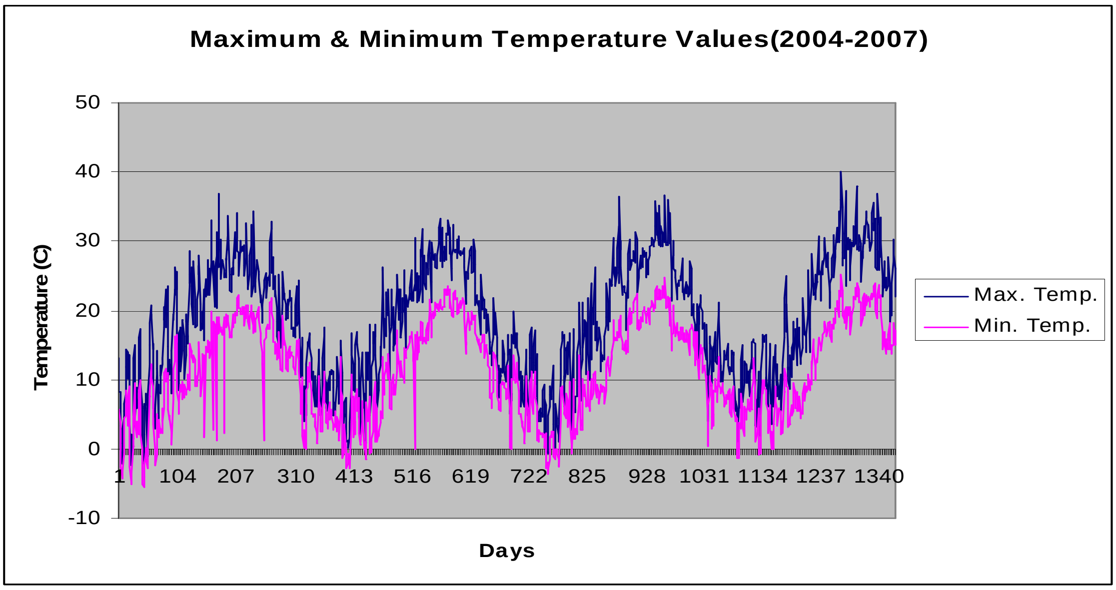


Figure 3.10. Maximum and minimum temperature values for Istanbul between years 2004-2007

Factor analysis for natural gas consumption values have been done. We can arrange the factors that effects it like:

- Maximum temperature
- Minimum temperature
- Day factor
- Day
- Month
- Year
- Previous day maximum temperature
- Previous day minimum temperature
- Number of consumers
- Previous day consumption values
- Wind Speed
- Price
- Price of other energy sources
- Types of combi boiler
- Consumers characters

etc…

All of these and some other undefined factors could effect the consumption of natural gas. Firstly we must define the factors that effects mostly. Of course there are some main effectors like wind speed, type of combi boiler, consumer characters etc. but we could not consider them because of the lack of accurate data.

Price is also another factor but price values do not change day to day so it is not important to consider it for daily forecasting values.

In our past studies, it is showed that previous day consumption values are very significant factor on tomorrow consumption values. In our model we have considered the day factor because as it is seen on the graph below consumption values on weekdays are higher than the weekend days. Figure 3.11 shows that as it is increasing on weekdays, there is a significant decrease in weekend and some holidays. In order to define this variability it must be considered in the model.

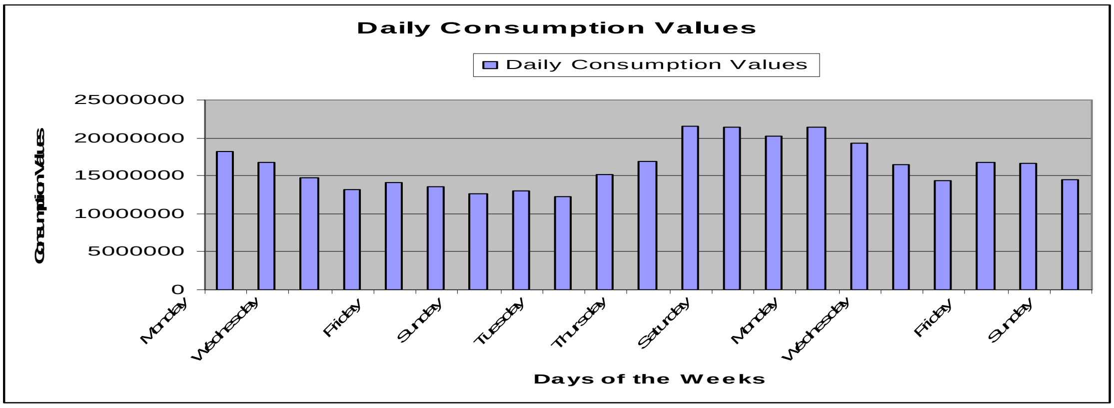


Figure 3.11. Variability on consumption according to the days of weeks.

Full regression analysis test have been done in order to determine the effects of the factors. The MINITAB statistical program was used and the results are below:

Table 3.4. ANOVA tables for consumption versus effected factors

Regression Analysis: Consumption versus Pre Days Max; Pre Days Min;

Predictor	Coef	SE Coef	T	P	
Constant	2126944	289543	7,35	0,001	*Significant*
Pre Days Max Temp	183873	14096	13,04	0,001	*Significant*
Pre Days Min Temp	84550	16380	5,16	0,001	*Significant*
Max. Temp.	-293013	12261	-23,90	0,001	*Significant*
Min. Temp.	-93024	17680	-5,26	0,001	*Significant*
Total Consumer Number	0,50508	0,09686	5,21	0,001	*Significant*
Previous day consumptiom	0,86820	0,01190	72,94	0,000	*Significant*

S = 1191286   R-Sq = 97,5%   R-Sq(adj) = 97,4%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	6	7,40102E+16	1,23350E+16	8691,77	0,000
Residual Error	1361	1,93148E+15	1,41916E+12		
Total	1367	7,59417E+16			

Source	DF	Seq SS
Pre Days Max Temp	1	5,98131E+16
Pre Days Min Temp	1	6,86225E+14
Max. Temp.	1	2,99751E+15
Min. Temp.	1	5,05269E+12
Total consumer number	1	2,95853E+15
Previous day consumption	1	7,54978E+15

As it is seen in Table 3.4 results, P-values are equal to zero. This means that, the factors are very significant for our model. If P-value is less than or equal to 0,05 we can consider it important for our model with 95% confidence interval level.

Table3.5. An example of daily natural gas consumption input values

Day	Month	Year	Day Factor	Pre Days Max Temp	Pre. Days Min. Temp	Max. Temp.	Min. Temp.	Total Consumer Number	Pre. Day Consump.	Consumption
2	1	2004	Friday	13,9	8,1	13,1	5,7	1.943.003	10.964.505	**12.079.253**
3	1	2004	Saturday	13,1	5,7	8,3	5,6	1.944.753	12.079.253	**13.078.319**
4	1	2004	Sunday	8,3	5,6	8,2	3,7	1.946.660	13.078.319	**14.747.299**
5	1	2004	Monday	8,2	3,7	3,3	0,7	1.948.223	14.747.299	**18.066.318**
11	1	2004	Sunday	4,6	-0,6	5,5	0,2	1.958.377	19.071.047	**18.772.911**
12	1	2004	Monday	5,5	0,2	6,6	4,5	1.959.697	18.772.911	**18.220.223**
13	1	2004	Tuesday	6,6	4,5	8	4,8	1.960.865	18.220.223	**16.786.761**
14	1	2004	Wednesday	8	4,8	11,8	7,5	1.962.125	16.786.761	**14.795.242**
20	1	2004	Tuesday	13,7	7,5	13,6	9	1.968.729	12.973.687	**12.291.801**
21	1	2004	Wednesday	13,6	9	11,8	2,6	1.970.076	12.291.801	**15.156.092**
22	1	2004	Thursday	11,8	2,6	3,5	-2,5	1.971.161	15.156.092	**16.831.046**
23	1	2004	Friday	3,5	-2,5	-2,4	-5	1.971.257	16.831.046	**21.599.698**
24	1	2004	Saturday	-2,4	-5	-0,3	-5,2	1.971.321	21.599.698	**21.405.378**
27	1	2004	Tuesday	1,4	0,3	4,5	0,3	1.973.715	21.433.757	**19.323.272**
28	1	2004	Wednesday	4,5	0,3	13,2	4,6	1.975.467	19.323.272	**16.514.204**
29	1	2004	Thursday	13,2	4,6	15,1	9,5	1.978.215	16.514.204	**14.390.590**
30	1	2004	Friday	15,1	9,5	8,8	3,6	1.980.497	14.390.590	**16.700.662**
31	1	2004	Saturday	8,8	3,6	4,6	3,6	1.982.656	16.700.662	**16.609.062**
1	2	2004	Sunday	4,6	3,6	8,8	1,7	1.982.706	16.609.062	**14.471.495**
2	2	2004	Monday	8,8	1,7	8,3	3,6	1.982.716	14.471.495	**14.173.108**
3	2	2004	Tuesday	8,3	3,6	6,9	4	1.982.728	14.173.108	**13.881.241**
4	2	2004	Wednesday	6,9	4	5,8	3,3	1.982.747	13.881.241	**16.188.822**
5	2	2004	Thursday	5,8	3,3	7,4	2,1	1.983.480	16.188.822	**17.039.291**

3.1.3. Analysis of Weekly Values

Weekly values were composed from the daily values. Temperature values are designed as an average of seven days average temperature values. As in the daily values average maximum and minimum temperature values were taken. Seven day average number of consumers were considered for input of consumer number.

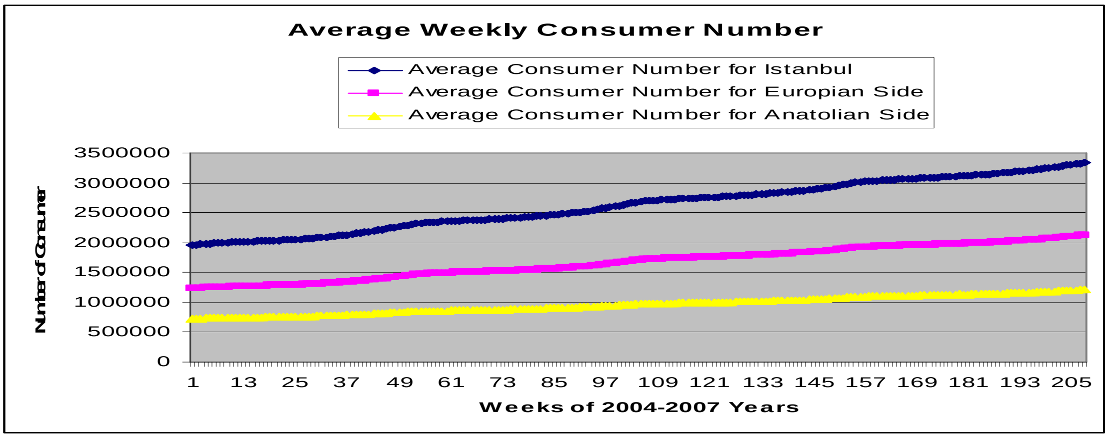


Figure 3.12. Average weekly consumer number of Istanbul between 2004-2007 years

Week factor is also important factor for natural gas consumption, winter season weekly consumption values are not same with summer season values. It is shown in the figure 3.13 that there is a significant difference between summer season and winter season weeks natural gas consumption values.

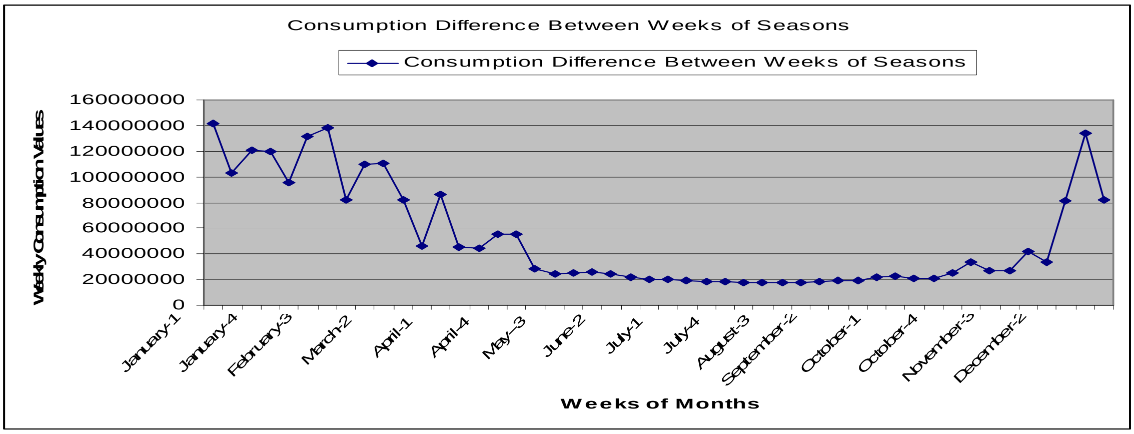


Figure 3.13. Natural gas consumption difference between seasons' weeks.

Factor analysis for weekly natural gas consumption values have been done. We can arrange the factors that effects it, like:

- Year
- Week number
- Average maximum temperature
- Average minimum temperature
- Average consumer number

etc.

As in the daily values, all of these and some other undefined factors could effect the weekly consumption of natural gas. Firstly we must define the factors that effects mostly. Of course there are some main effectors like wind speed, type of combi boiler, consumer characters etc. but we could not consider them because of the lack of accurate data so we will not consider these in weekly forecasting models.

Price is also another factor that effects it but price values do not change week to week so it is not important to consider it for weekly forecasting values.

Table 3.6 ANOVA tables for consumption versus effected factors (weekly)

Regression Analysis: Istanbul Total Consumption versus Year; Max Avg Temp...

Predictor	***Coef***	***SE Coef***	***T***	***P***
Constant	3,42549E+11	60840815990	5,63	0,001
Year	-171347367	30443271	-5,63	0,001
Istanbul average Consumer Number	507,14	85,25	5,95	0,001
Week	-3268643	563928	-5,80	0,001
Average Max. Temperature	-4755025	666850	-7,13	0,001
Average Min.Temperature	133976	813595	0,16	0,869

S = 15056737 R-Sq = 91,5% R-Sq(adj) = 91,2%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	5	4,59020E+17	9,18040E+16	404,95	0,001
Residual Error	189	4,28473E+16	2,26705E+14		
Total	194	5,01867E+17			

Source	DF	Seq SS
Year	1	4,91167E+15
Istanbul average Consumer Number	1	5,92410E+15
Week	1	3,80323E+17
Average Max. Temperature	1	6,78547E+16
Average Min.Temperature	1	6,14751E+12

Full regression analysis test have been done in order to determine the effects of the factors. The MINITAB statistical program was used and the results are above. Results shows that average minimum temperature is not significant factor to our model.

Table 3.7. An example of weekly natural gas consumption input values.

Year	Week	Weekly Average consumer Number	Average Max. Temp.	Average Min. Temp.	Weekly Total Consumption
2004	1	1953500	2,16	-1,97	141944597
2004	2	1963347	11,10	5,53	103268140
2004	3	1970222	5,99	0,41	120411441
2004	4	1977937	8,06	3,37	119443042
2004	5	1983663	10,77	5,43	95155993
2004	6	1987687	5,00	-1,30	131408447
2004	7	1991776	5,23	-4,69	138318763
2005	28	2436576	27,57	20,21	21482633
2005	29	2442037	30,03	20,76	20484797
2005	30	2447824	28,14	21,56	20990337
2005	31	2454600	31,01	22,26	20415258
2005	32	2460663	28,63	20,19	21307706
2005	33	2466588	28,63	21,19	21571447
2005	34	2472553	28,91	20,73	22016923
2006	9	2728716	13,54	5,93	113197779
2006	10	2732236	10,94	4,24	139637442
2006	11	2736181	7,83	2,57	149685168
2006	12	2739459	14,54	5,94	109392229
2006	13	2742542	16,80	7,63	82843820
2006	14	2745800	15,16	7,09	89709283
2006	15	2749240	18,61	7,73	66337264
2007	7	3047421	11,46	7,10	122919647
2007	8	3051928	10,13	3,63	142761134
2007	9	3055955	11,06	4,57	138694141
2007	10	3060249	8,67	5,13	141888610
2007	11	3064665	10,31	4,30	144236603
2007	12	3069007	18,53	9,11	84911717
2007	13	3072785	10,86	4,43	130604461
2007	14	3076214	16,09	7,01	88669998
2007	15	3079580	15,64	6,74	83556762

3.1.4. Analysis of Monthly Values

Monthly values were composed from the daily values. Temperature values are designed as an average of days temperature values. As in the daily and weekly values average maximum and minimum temperature values were taken. Monthly average number of consumers were considered for input of consumer number.

Month factor is also important factor for natural gas consumption, seasonal behaviour of monthly consumption values could be seen. It is shown in figure 3.14 that there is a significant seasonal difference between months' natural gas consumption values.

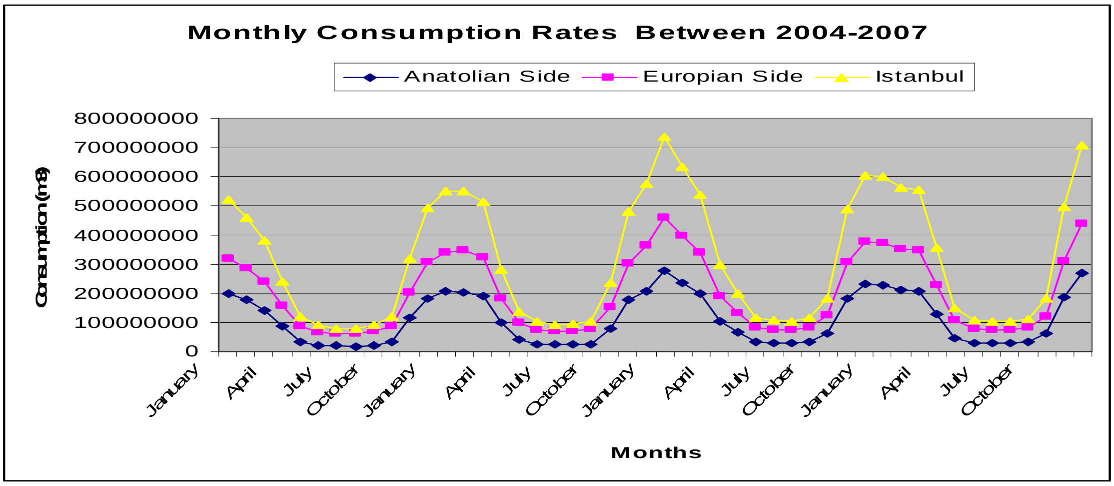


Figure 3.14. Monthly consumption rates between 2004-2007 for Istanbul.

Figure 3.14 shows that natural gas consumption values are mostly depends on seasonal factors. We could divide a year into three period. Period-1 includes January, February, March and December; Period-2 includes October, November, and April; Period-3 includes May, June, July, August, September months. Monthly consumption values in these periods are similar to each other and do not change year to year. Below ANOVA table shows the importance of periods too.

Table 3.8 ANOVAresults for consumption versus monthly periods

```
Analysis of Variance for Istanbul Natural Gas Consumption
Source     DF        SS        MS        F      P
Seasonal    2 6,531E+17 3,265E+17   130,43    0,001
Error      45 1,127E+17 2,504E+15
Total      47 7,658E+17
                                   Individual 95% CIs For Mean
                                   Based on Pooled StDev
Level       N      Mean     StDev  -----+---------+---------+---------+-
Period 1   16  3,51E+08  53889485                              (-*--)
Period 2   12  1,97E+08  76335513              (--*--)
```

```
Period 3   20  80145458  16225050   (-*-)
                                   -----+---------+---------+---------+-
Pooled StDev = 50035800          1,00E+08  2,00E+08  3,00E+08  4,00E+08
```

P-value is less than 0.05 so it means that period factor is the most significant for our model.

Below ANOVA test is also applied for month effect for natural gas consumption values. The results shows that, as mentioned before in Period analysis, month factor is also very significant for consumption and dashed lines shows that months could be arranged like Period-1 includes January, February, March and December; Period-2 includes October, November, and April; Period-3 includes May, June, July, August, September months too.

Table 3.9 ANOVA results for consumption versus months of year

```
Analysis of Variance for İstanbul Natural Gas Consumption
Source     DF        SS        MS        F        P
Month         11 7,173E+17 6,521E+16    48,42    0,001
Error      36 4,848E+16 1,347E+15
Total      47 7,658E+17
                                   Individual 95% CIs For Mean
                                   Based on Pooled StDev
Level       N      Mean     StDev  --------+---------+---------+--------
April       4  1,90E+08  29389194            (--*--)
August      4  70535517   6195346   (--*--)
December    4  3,72E+08  52861418                           (--*--)
February    4  3,46E+08  47009573                         (--*--)
January     4  3,74E+08  61511380                           (--*--)
July        4  70362311   6451544   (--*--)
June        4  76341291   6572767   (--*--)
March       4  3,13E+08  49938927                      (--*--)
May         4  1,06E+08  19194074      (--*--)
November    4  2,81E+08  52340104                   (--*--)
October     4  1,21E+08  27322605       (--*--)
Septembe    4  77931033   6169237   (--*---)
                                   --------+---------+---------+--------
Pooled StDev = 36698308             1,20E+08  2,40E+08  3,60E+08
```

Below ANOVA test is also applied for year effect for natural gas consumption values in order to understand if there is an year effect on consumption or not. P-value shows that consumption value does not change year to year. This means of course is not that consumption value is stable. It is increasing year to year but ratio of consumption values are same. So we could say there is no significant effect of year on consumption values.

Table 3.10 ANOVA results for consumption versus years

```
Analysis of Variance for İstanbul Natural Gas Consumption
Source     DF        SS        MS        F        P
Yıl         3 2,511E+16 8,371E+15     0,50    0,686
Error      44 7,406E+17 1,683E+16
Total      47 7,658E+17
                                   Individual 95% CIs For Mean
                                   Based on Pooled StDev
Level       N      Mean     StDev  ------+---------+---------+---------+
2004       12  1,62E+08  1,04E+08  (------------*------------)
2005       12  2,01E+08  1,25E+08         (------------*-----------)
2006       12  2,21E+08  1,46E+08            (------------*-----------)
2007       12  2,15E+08  1,40E+08           (------------*-----------)
                                   ------+---------+---------+---------+
Pooled StDev = 1,30E+08          1,20E+08  1,80E+08  2,40E+08  3,00E+08
```

As a result, the effected factors for monthly consumption values are period, month, average temperature, maximum consumer number, minimum consumer number.

Table 3.11. A format example of monthly natural gas consumption values

Month	Average Consumer Number	Min. Consumer Number	Max. Consumer Number	Average Temperature	Istanbul Monthly Consumption
January	1.962.843	1941375	1982656	4,4	521465100
February	1.989.775	1982706	1999121	5,6	461465218
March	2.008.051	1999720	2015421	8,0	382921070
April	2.021.793	2015984	2027615	11,9	241680214
May	2.033.715	2027624	2039158	16,3	120401074
June	2.048.218	2039852	2056784	31,7	89354950
July	2.069.546	2057722	2082235	22,3	80428831
August	2.113.143	2082428	2149429	22,6	79763902
September	2.130.512	2112805	2149429	20,4	91012641
October	2.225.547	2150843	2305382	16,6	120895431
November	2.225.088	2197828	2251401	11,3	318058294
December	2.279.450	2253488	2305382	8,1	492205178
January	2.325.817	2306431	2338429	6,4	549457812
February	2.347.923	2339551	2355047	5,5	552808810
March	2.364.278	2355614	2372067	6,8	514024050
April	2.379.015	2372725	2385684	12,1	282053702
May	2.394.072	2385689	2402334	16,2	138837378
June	2.413.125	2403076	2423213	19,7	102169315
July	2.437.702	2424000	2450502	23,5	92912471
August	2.464.884	2451391	2511922	24,4	95046390

CHAPTER 4

4.1. Artificial Neural Networks

4.1.1. What Is An Artificial Neural Network?

The terms of neural network is used to specify a network or circuits of biological neurons. It is composition of artificial neuron or nodes so in modern usage of the term it is often specified as artificial neural network. Neural networks are highly interconnected simple processing units designed in a way to model how the human brain performs a particular task [32].

Artificial neural networks (ANNs) are programs designed to simulate the way a simple biological nervous system is believed to operate. They are based on simulated nerve cells or neurons which are joined together in a variety of ways to form networks. These networks have the capacity to learn, memorize and create relationships amongst data [5].

In Marvin Minsky words, “*Artificial Intelligence is the science of making machines do things that would require intelligence if done by men.* ”

A neural network is an interconnected assambly of simple prosessing elements, *units* or *nodes*, whose functinality is loosly based on animal neuron. The processing ability of the network is stored in the inter-unit connection strengths, or *weights*, obtained by a process of adaptation to, or *learning* from, a set of training patterns. In order to see how very different this is from the processing done by conventional computers it is worth examining the underlying principles that lie at the heart of all such machines [5].

Most widely used artificial neural network type for forecasting is multi layer perceptron(MLP) model. Multilayer perceptrons are best known and most widely used kind of neural network models. Networks with interconnections that do not form any loops are called feedforward [33]. Recurrent or non-feedforward networks in which there are one or more loops of interconnections are used for some kinds of applications [34].

4.1.2 The Perceptron

The perceptron was firstly introduced by Rossenblatt that used as an enhancement off the TLU (Threshold Logic Unit). It consists of a TLU whose inputs come from a set of preprocessing *association units* (A-units). The A-units can be assigned any arbitrary Boolean functionality but are fixed - they do not learn. The rest of the node functions just like a TLU and may be therefore be trained in exactly the same way [5].

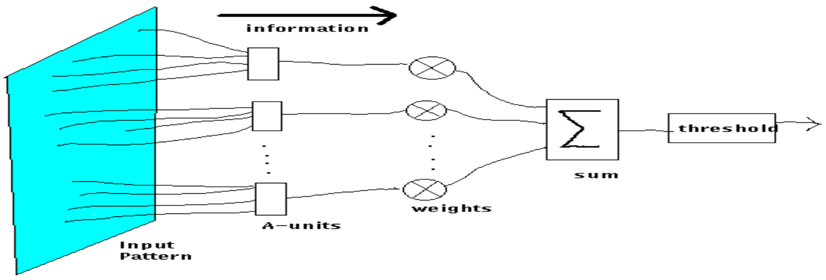


Figure 4.1. Single-layer perceptron

4.1.3. Architecture of Neural Networks

4.1.3.1. Feed Forward Networks

Feed-forward structure is a type of artificial neural networks that allows signals to pass one way only, from input to output. There is no loops (feedback) so the output of any layer does not effect that same layer. It also tend to be straight forward networks that associate inputs with outputs.

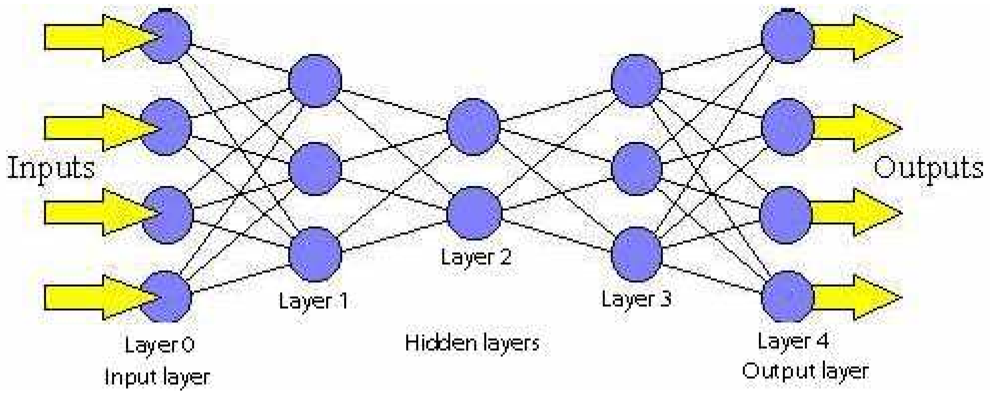


Figure 4.2. An example of a simple feedforward network

4.1.3.2. Network Properties

The most common artificial neural networks consist of three layers that called as "*input*" , "*hidden*" and "*output*". As shown in the above figure, a layer of "input" units is connected to a layer of "hidden" units, and "hidden" units is connected to a layer of "output" units.

- The activity of of the input units represent the raw information that is fed into the network.
- The activity of each hidden unit is determined by the activities of the input units and the weights on the connections between the input and the hidden units.
- The behaviour of the output units depends on the activity of the hidden units and the weights between the hidden and output units. [5]

4.1.4. The Learning Process

Every neural network posseses knowledge which is contained in the values of the connections weights. Modifying the knowledge stored in the network as a function of experience implies a learning rule for changing the values of the weights [5].

The memorisation of patterns and subsequent response of the network can be categorised into two general paradigms:

- **Associative Mapping:** The network learns to produce a particular pattern on the set of input units whenever another particular pattern is applied on the set of input units.

- **Regularity Detection:** Units learns to respond to a particular properties of input patterns. Whereas in asssociative mapping the network stores the relationships among patterns, in regularity detection the response of each unit has a particular 'meaning'. This type of learning mechanism is essential for *feature discovery* and *knowledge representation.*

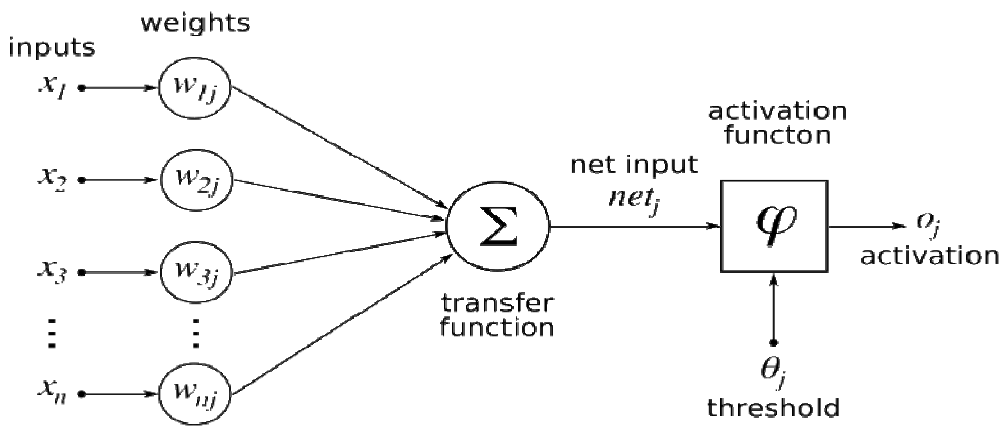


Figure 4.3. Structure of learning process

Information is stored in the weight matrix W of a neural network. Learning is the determination of the weights. Following the way learning is performed, we can distinguish two major categories of neural networks:

- **Fixed Networks** : Fixed networks in which the weights cannot be changed, ie dW/dt=0. In such networks, the weights are fixed a priori according to the problem to solve.

- **Adaptive Networks:** Adaptive networks which are able to change their weights, ie dW/dt not= 0.

4.1.4.1. Type of Learning Processes

The learning methods used for adaptive neural networks can be classified into two major categories:

4.1.4.1.1. Supervised Learning: Supervised learning which incorporates an external teacher, so that each output unit is told what its response to input signals ought to be [5].

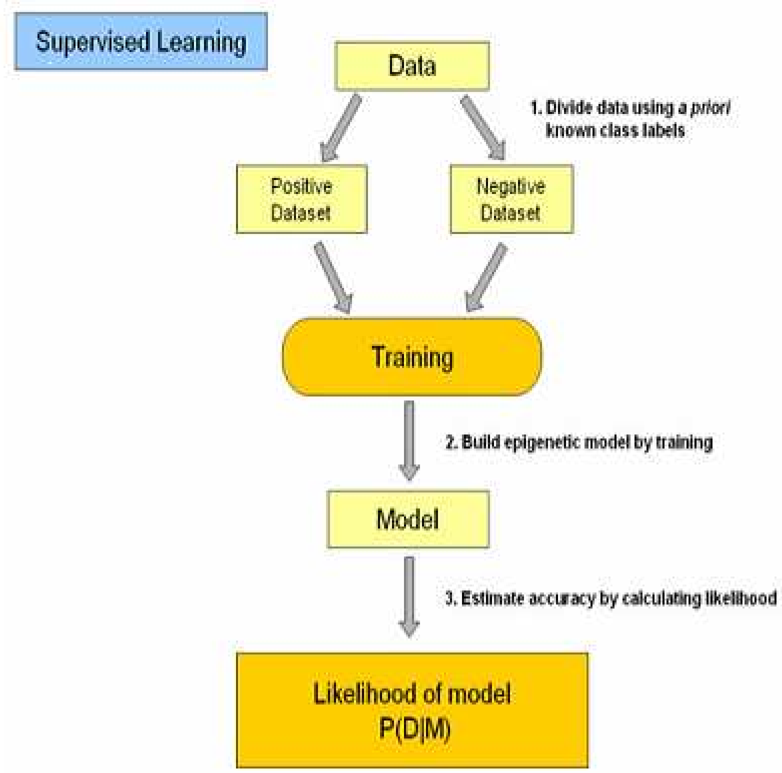


Figure 4.4. Model of supervised learning process

4.1.4.1.2. Unsupervised Learning: Unsupervised learning does not use any external teacher, means there is no input, and is based upon only local information.

4.1.5. Activation Functions

The behaviour of an ANN (Artificial Neural Network) depends on both the weights and the input-output function (Activation function) that is specified for the units. This function typically falls into one of three categories:

- **Linear Activation Function:** The output activity is proportional to the total weighted output.
- **Treshold Activation Function:** The output is set of one of two levels, depending on whether the total input is greater or less than some treshold values.
- **Sigmoid Activation Functions:** The output varies continuously but not linearly as the input changes. Sigmoid units bear a greater resemblance to real neurones than do linear or threshold units, but all three must be considered rough approximations [5].

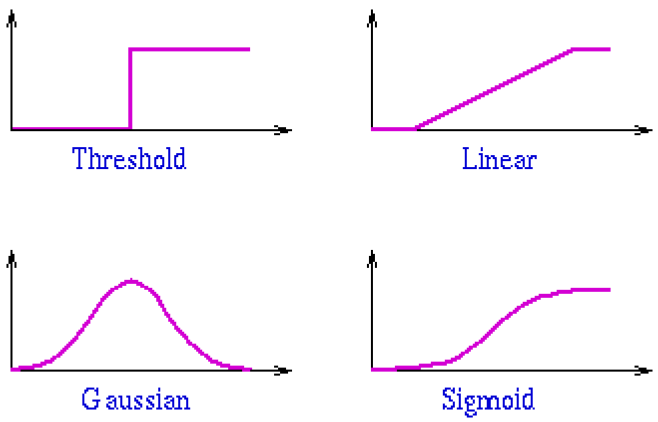


Figure 4.5. Samples of activation functions

To make a neural network that performs some specific task, we must choose how the units are connected to one another, and we must set the weights on the connections appropriately.

4.2. What Are Artificial Neural Networks Used For?

The usage of artificial intelligence are limitless but it can be classified in different sections like:

- ***Classification:*** Customer/ Market profiles, medical diagnosis, signature verification, image recognition, classification of cell types and any more.
- ***Forecasting:*** Future sales, production requirements, market performance, economic indicators, energy requirements, medical outcomes, chemical reaction products, weather, crop forecasts, environmental risk, horse races, jury panels.
- **Modeling:** Process control, systems control, chemical structures, dynamic systems, signal compression, plastics moulding, welding control, robot control, and many more.

4.3. Who needs Artificial Neural Networks?

As mentioned before, there are limitless usage area of artificial intelligence as people in finance, industry, business, education whose problems are complex, laborious, fuzzy or un-resolvable using present methods that works with or analyze data of any kind needs artificial intelligence solution methods.

4.4. Mathematical Model of ANN

There are many current learning algorithms in ANN. The most popular algorithm is back propogation algorithm. İn this project we have used feed forward back propogation algorithm. In this section we will discuss and show the mathematical model of back propogation algorithm.

4.4.1. Back-Propagation Algorithm

The multi layered feed forward network has a better ability to learn the correspondence between input patterns and teaching values from many sample data by the error back propagation algorithm. In this paper, we used three-layered feed forward neural networks and taught them by error back propagation. The output (O_{ij}) of each unit ij is defined by,

$O_{ij} = f(net_{ij})$, $net_{ij} = \sum_i w_{ij} O_i + \theta_j$ (1)

Where O_i is the output of unit i, w_{ij} is the weight of the connection from unit i to unit j, θ_j is the bias of unit j, $\sum_i$ is a summation of every unit ij whose output flows into unit j, and $f(x)$ is a monotonously increasing function. In practice, a logistic activation function $f(x) = 1/(1+\exp(-x))$ is used. When the set of m-dimensional input patterns $\{i_p = (i_{p1}, i_{p2}, \ldots, i_{pm}); p \in P\}$ where P denotes set of presented patterns, and their corresponding desired n-dimensional output patterns $\{t_p=(t_{p1}, t_{p2}, \ldots, t_{pm}); p \in P\}$ are provided, the neural network is taught to compute ideal patterns as follows. The squared error function E_p for a pattern p is defined by

$$E_p=\frac{1}{2}\sum_{j \in \text{output}} \left(t_{pj} - O_{pj}\right)^2 \qquad (2)$$

The purpose is to make $E = \sum_p E_p$ small enough by choosing appropriate w_{ji} and θ_j. To realize this purpose, a pattern $p \in P$ is chosen successively and randomly, and then w_{ji} and θ_j are changed by

$$\Delta_p w_{ji} = -\varepsilon (\partial E_p / \partial w_{ji}) \qquad (3)$$

$$\Delta_p \theta_j = -\varepsilon (\partial E_p / \partial \theta_j) \qquad (4)$$

Where, ε is a small positive constant. By calculating the right hand side of (3) and (4), it follows that

$$\Delta_p w_{ji} = \varepsilon \delta_{pj} O_{pi} \qquad (5)$$

$$\Delta_p \theta_j = \varepsilon \delta_{pj} \qquad (6)$$

Where

$$\delta_p=\begin{cases} f'(net_j)(t_{pj} - O_{pj}) \\ f'(net_j)\sum_k w_{kj}\delta_{pk} \end{cases} j \qquad (7)$$

Note that k in the above summation represents every unit k that’s output follows into unit j. In order to accelerate the computation, the momentum terms are added on (5), (6),

$$\Delta_p w_{ji} (n+1) = \varepsilon \delta_{p j} O_{pi} + \alpha \Delta_p w_{ji} (n) \qquad (8)$$

$$\Delta_p \theta_j (n+1) = \varepsilon \delta_{p j} + \alpha \Delta_p \theta_j (n) \qquad (9)$$

Where n represents the number of learning cycles, and α is a small positive value.

4.5. Different Back Propogation Learning Algorithms

Seven different training algorithm have used for comparison of ANN learning. A briefly information about seven ANN algorithms as writen below. The informations about the algorithms have been taken from Alyuda NeuroIntelligence program document file [35].

4.5.1. Quick Propagation Algorithm

Quick propagation is a heuristic modification of the back propagation algorithm invented by Scott Fahlman. This training algorithm treats the weights as if they were quasi-independent and attempts to use a simple quadratic model to approximate the error surface. In spite the fact that the algorithm hasn't theoretical foundation, it's proved to be much faster than standard back-propagation for many problems. Although sometimes the quick propagation algorithm may be instable and inclined to stuck in local minima.

4.5.2. Conjugate Gradient Descent Algorithm

Conjugate Gradient Descent is an advanced method for training multi-layer neural networks. It is based on the linear search usage in the line of an optimal network weights' change. The correction of weights is conducted once per iteration. In most cases, this method works faster than Back Propagation and provides more precise forecasting results.

4.5.3. Quasi-Newton

The network training algorithm based on Newton's method. An approximate Hessian matrix is computed during each iteration of the algorithm based on the gradients. There are two major variations, recommended and the most popular BFGS (Broyden-Fletcher-Goldfarb-Shanno) method.

4.5.4. Limited Memory Quasi-Newton

A network training algorithm, which is a variation of Quasi Newton (BFGS) that avoids the need to store Hessian matrix and thus require less memory and can be used for bigger networks.

4.5.5. Levenberg-Marquardt

Levenberg-Marquardt is an advanced non-linear optimization algorithm. It is the fastest algorithm available for multi-layer perceptrons. However, it has the following restrictions:

- It can only be used on networks with a single output unit.
- It can only be used with small networks (a few hundred weights) because its memory requirements are proportional to the square of the number of weights in the network.
- It is only defined for the sum squared error function and therefore it is only appropriate for regression problems.

4.5.6. Incremental Back Propagation

Variation of the Back Propagation where the network weights are updated after presenting each case from the training set, rather than once per iteration. This is originally invented variant of back propagation and sometimes referred to as Standard Back Propagation. It can be the most preferred algorithm for large data sets.

4.5.7. Batch Back Propagation

This is an advanced variant of Back Propagation where network weights update takes place once per iteration. Epoch means the process of passing of all training record through the network.

CHAPTER 5

5.1. FORECASTING OF NATURAL GAS CONSUPTION USING ANN

5.1.1. Preparing Data Sets

Proper preparation of our data sets is the most improtant step in working with neural networks or other forecasting models. Having a well designed or prepared data will lead us to impressive results. Unprepared data always lead to faiure in applying forecasting models actually in neural networks. There are several steps for preparing data sets.

- **Right amount of data needed.**

  If our model has not enough historical data, the neural network will not have enough information about our problem in order to train correctly. We must determine the true number of data set. It is not always true to have too much data because too much data may increase training time of neural network and even deteriote network performance. Right amount of data depends of problem and its complexity.

- **Data should not be self-contradictory.**

  If our model has self-contradictory data neural network could not train exactly. We must remove self-contradictory data or add more inputs related to our model to make data consistent. Neural neetwork would not be able to find relation between the input and outputs and can not calculate difference between data sets.

- **Inputs should have maximum influence on target.**

  First of all we must determine all parameters that effects our model and than add all parameters that determine our target significantly. By doing this better results could be calculated.

- **Data shouldn’t have missing data.**

  Missing data must need to be remove or subtitude with related values.For example in our model there are different temperature values and they are assigned as 1, 2, 7.7, 8, 16, 50, 6, 12, 5 ^{o}C etc. There is a contredictionary between temperature values, 50 ^{o}C is a wrong type value and not related with our values. So it must be removed or changed with appropriate value.

- **Data should well present our problem environmet.**

Firstly enough effected factors must be determined and included our model.the more different variants we have the more confident will be network will be performing on new data. Dataset should contain at leaset several cases for each category in each categorical column and must be well presented in different conbinations.

5.1.2. Applying Datasets to Model

The input dataset contains data that can be used to train and test neural networks. Firstly we must determine the dataset which will be in train, validation and test sections. Our dataset should have numeric, categorical, date or time format. Neural network couldnot work with categorical, date or time format values so it must be writen as a numerical values and normalized to appropriate values.

5.1.2.1.Preprosessing Dataset

As mentioned before data must be preprocessed. Preprocessing means modification of the data before it is applied to neural network.Preprocessing trnasforms data to make it suitable for neural network. Numerical values scaled and text values transformed to numerical values.

5.1.2.1.1. Encoding of Numerical Values

In our model numeric colums are scaled as:

- SF = (SRmax-SRmin)/(Xmax-Xmin)
- Xp = SRmin + (X-Xmin)* SF

Where:

- X - actual value of a numeric data in column
- Xmin - minimum actual value of the data in column

- Xmax - maximum actual value of the data in column
- SRmin - lower scaling range limit
- SRmax - upper scaling range limit
- SF - scaling factor
- Xp - preprocessed value

For input columns scaling range is [-1..1].

For the target column scaling range depends on activation function of the output layer:

Table 5.1. Formulas and scaling ranges of activation functions

Output layer activation function	**Scaling range**	**Formula**
Linear	[-1..1]	[-inf..inf].
Logistic	[0..1]	$F(x) = 1/ (1+e^{-x})$.
Hyperbolic Tangent	[-1..1]	$F(x) = (e^{x} - e^{-x})/ (e^{x} + e^{-x})$.
Softmax	[0..1]	*For classification problems*

Xmin and Xmax are usualy taken from the input dataset but ıf we know taht future data for forecasting will be lower than the minimum or higher than the maximum than presented in our inpur data file, we can change the minimum and maximum values used for scaling.

In our model number of customer and consumption values always increasing to higher values we have changed the maximum values with appropriate vales.

5.1.2.1.2. Encoding of Categorical Values

Encoding of categorical colums are encoded as binary encoding method.For example, day factors in the data were writen as Monday, Tuesday, Wednesday etc. Representing of Monday is shown as {1,0,0,0,0,0,0} and for Wednesday {0,0,1,0,0,0,0} etc.

5.1.3. Neural Network Design

In neural network applications first of all network design must be done. We need to specify the network architecture, *number of hidden layer and units in each layer,* and properties of network, *error and activation functions*.

A network with too few hidden units only roughly discovers hidden dependencies in your data whereby your network produces a significant number of errors. A network with too many hidden units will tend to memorize all your data instead of finding relations that also lead to bigger network errors. You need to find the best solution for your problem and your dataset.

In our neural network models only one hidden layer was used and for all applications number of units in each hidden layer increased until 50 and all results were saved in order to find the best model for our network. The results were showed below.

5.1.4. Training Algorithm Selection

Selection of training algorithm is the most important section in neural network applications. There are many training algorithms in literature but there is no exact single best training algorithm for neural networks. It could change according to the characteristic of the problem.

In our models, we have used several training algorithms as below and select the best algorith for our models.

- Quick propagation Algorithm
- Conjugate Gradient Descent Algorithm
- Quasi-Newton Algorithm
- Limited Memory Quasi-Newton Algorithm
- Levenberg-Marquardt Algorithm
- Incremental Back-Propagation Algorithm
- Batch Back-Propagation Algorithm

The following simple rules that is taken from www.alyuda.com proved to be quite effective for most practical purposes.

1. If you have a network with a small number of weights (usually, up to 300), Levenberg-Marquardt algorithm is efficient. Levenberg-Marquardt often performs considerably faster than other algorithms and finds better optima than other algorithms. But its memory requirements are proportional to the square of the number of weights. Another Levenberg-Marquardt limitation is that it is specifically designed to minimize the sum of squares error and cannot be used for other types of network error.
2. If you have networks with a moderate number of weights, Quasi-Newton and Limited Memory Quasi-Newton algorithms are efficient. But their memory requirements are also proportional to the square of the number of weights.
3. If your network has a large number of weights, we recommend you using Conjugate Gradient Descent. Conjugate Gradient Descent has nearly the convergence speed of second-order methods, while avoiding the need to compute and store the Hessian matrix. Its memory requirements are proportional to the number of weights.
4. Conjugate Gradient Descent and Quick Propagation are general-purpose training algorithms of choice.
5. You can use incremental and batch Back propagation for networks of any size. Back propagation algorithm is the most popular algorithm for training of multi-layer perceptrons and is often used by researchers and practitioners. The main drawbacks of back propagation are: slow convergence, need to tune up the learning rate and momentum parameters, and high probability of getting caught in local minima. Incremental back propagation can be efficient for large datasets if you properly select the learning rate and momentum. It usually performs better than batch back propagation.

5.2. Neural Network Models And Calculations

It needs to specify the network architecture, number of hidden layer and units in each layer, and properties of network, error and activation functions.

In this study the network is trained by both Mathlab Neural Network Module (nntool) and NeuroIntelligence Neural Network based program. For all models 60% of the available data is used for training, 20% for validation and 20% for testing.

In our neural network models only one hidden layer was used and for all applications number of units in each hidden layer increased until 50. Each of them replicated 10 times and all results were saved in order to find the best model for our network.

Exhausted search method is used for incremental of number of hidden layer neuron number and inverse test error is selected for fitness criteria. For output layer logistic activation function is applied.

5.2.1. ANN Forecasting Calculations for Istanbul Daily Values

Input factors for daily forecasting are taken as total customer number, day, day factor, previous day cunsomption, maximum and minimum temprature, previous day maximum and minimum temperature, month and year.

In ANN daily forecasting models there are 19 input values and one output value. Below graph shows the best 5 network models after all 50 hidden layer units were applied to the model one by one.

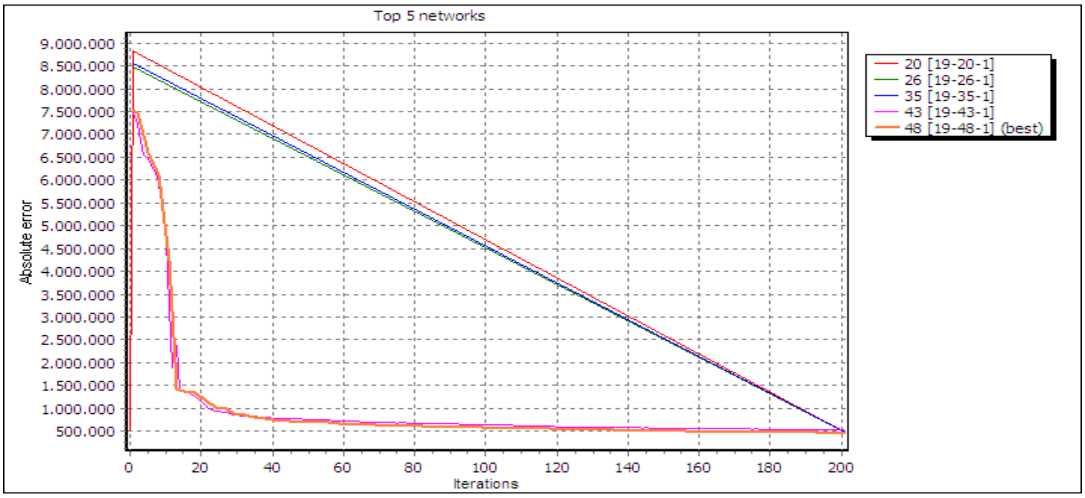


Figure 5.1. Graph of best 5 networks for Istanbul daily values

[19-48-1] architecture had the best fitness

The graph Figure 5.1. and also below results Table 5.2. shows that the model with 48 hidden neuron is best for Istanbul daily gas consumption forecasting in neural network model.

All of the training, validation, test, correlation and R-Square values for all hidden neuron nımbers are shown below Table 5.2. Most of the R-Square values which shows that input values could explain our model or not, are over 98 percent.

Table 5.1. Importance of input factors for Istanbul daily forecasting

Input column name	Importance, %
• Month	1,361611
• Day	0,168876
• Day Factor	1,211695
• Max. Temp.	26,856485
• Min. Temp.	2,269972
• Pre Days Max. Temp.	5,502207
• Pre Days Min. Temp.	0,170156
• Total Customer Number	3,073997
• Year	2,004163
• Previous Day Cunsomption	57,380838

Above Table 5.1. results shows that most important factors for daily forecasting of natural gas of Istanbul is “pravious day consumption” and “maximum temperature” values with % 57.38 and % 26.83 importance.

The results of training, validating and test values are as shown in Table 5.2.

Table 5.2. Results of ANN models for Istanbul daily forecasting.

ID	Architecture	# of Weights	Fitness	Train Error	Validation Error	Test Error	AIC	Correlation	R-Squared
1	[19-1-1]	22	0,000001	771248,25	701750,12	744147,56	6305,579	0,988697	0,977255
2	[19-2-1]	43	0,000002	624989,31	637047,81	628613,06	6151,602	0,992755	0,985535
3	[19-3-1]	64	0,000002	631466,93	609574,93	626262,68	6203,212	0,992063	0,984058
4	[19-4-1]	85	0,000002	610094,5	623204	615294	6213,121	0,992858	0,985728
5	[19-5-1]	106	0,000002	606071,93	626864,62	625607,37	6248,956	0,992823	0,985542
6	[19-6-1]	127	0,000002	591806,25	629309,37	611625,18	6268,756	0,993337	0,98661
7	[19-7-1]	148	0,000002	592129,25	619228,62	599076,87	6311,265	0,993615	0,987207
8	[19-8-1]	169	0,000002	559505,43	584179,56	589994,56	6300,447	0,993229	0,986422
9	[19-9-1]	190	0,000002	559205,43	592209,68	596662,12	6341,947	0,99436	0,988688
10	[19-10-1]	211	0,000002	522191,75	549911,25	574506,62	6320,122	0,994674	0,989223
11	[19-11-1]	232	0,000002	556925,81	587991,12	587253,62	6422,140	0,994049	0,988065
12	[19-12-1]	253	0,000002	563806,56	589626,12	590599,12	6475,584	0,994239	0,988426
13	[19-13-1]	274	0,000002	552528,68	603019,68	602973,06	6498,752	0,994174	0,988228
14	[19-14-1]	295	0,000002	529149,37	575419,81	582376,25	6500,458	0,994466	0,988941
15	[19-15-1]	316	0,000002	531953,43	596473,5	571280,75	6547,383	0,994515	0,989012
16	[19-16-1]	337	0,000002	555384,06	599686,62	590884,68	6629,556	0,993433	0,986869
17	[19-17-1]	358	0,000002	516609	608458,75	575576,81	6604,104	0,994904	0,989824

Table 5.2. Results of ANN models for Istanbul daily forecasting. (continued)

ID	Architecture	# of Weights	Fitness	Train Error	Validation Error	Test Error	AIC	Correlation	R-Squared
18	[19-18-1]	379	0,000002	516525,34	585146	584493,12	6645,953	0,994921	0,989854
19	[19-19-1]	400	0,000002	543625,93	579879,93	569286,37	6735,613	0,994266	0,988549
20	**[19-20-1]**	**421**	**0,000002**	**495111,12**	**550158,12**	**546631,62**	**6690,490**	**0,995038**	**0,990038**
21	[19-21-1]	442	0,000002	526121,5	603867,93	576761	6789,109	0,994085	0,988157
22	[19-22-1]	463	0,000002	518598,46	578124,62	583072,31	6817,686	0,994476	0,988909
23	[19-23-1]	484	0,000002	490695,09	551850,68	585190,68	6808,140	0,995304	0,990484
24	[19-24-1]	505	0,000002	524226,68	570544,5	577820,81	6911,747	0,994608	0,989221
25	[19-25-1]	526	0,000002	501009,5	575712,37	565118,62	6911,528	0,994854	0,989663
26	**[19-26-1]**	**547**	**0,000002**	**523086,12**	**579131,62**	**546027,37**	**6993,717**	**0,994182**	**0,988146**
27	[19-27-1]	568	0,000002	582550,12	600350,25	587643,5	7136,064	0,993149	0,986115
28	[19-28-1]	589	0,000002	502403,28	592618,81	556872,56	7040,117	0,994679	0,989312
29	[19-29-1]	610	0,000002	563979,43	588741,81	570806,75	7189,870	0,993823	0,987654
30	[19-30-1]	631	0,000002	525036,62	629702,62	584830,06	7165,185	0,995018	0,990054
31	[19-31-1]	652	0,000002	540026,25	587660,68	564666,75	7233,421	0,994555	0,989116
32	[19-32-1]	673	0,000002	519515,09	540357,81	563323,12	7239,332	0,994791	0,989593
33	[19-33-1]	694	0,000002	538696,06	577380,31	572217,56	7315,122	0,99436	0,988662
34	[19-34-1]	715	0,000002	531446,06	574834	578934	7344,494	0,994645	0,989253

Table 5.2. Results of ANN models for Istanbul daily forecasting. (continued)

ID	Architecture	# of Weights	Fitness	Train Error	Validation Error	Test Error	AIC	Correlation	R-Squared
35	**[19-35-1]**	**736**	**0,000002**	**496884,12**	**561382,87**	**543603,5**	**7323,822**	**0,994708**	**0,989394**
36	[19-36-1]	757	0,000002	535502,12	587535,75	581184,5	7435,580	0,994385	0,988634
37	[19-37-1]	778	0,000002	517340,06	549448,62	558102,37	7445,422	0,994594	0,989113
38	[19-38-1]	799	0,000002	544159,37	568411,43	579965,93	7534,527	0,994327	0,988602
39	[19-39-1]	820	0,000002	503319,06	565301,18	561745,81	7503,814	0,994611	0,989241
40	[19-40-1]	841	0,000002	508331,81	588527,68	573033,68	7555,051	0,994839	0,989664
41	[19-41-1]	862	0,000002	532458,5	585791,68	575046,37	7640,268	0,994417	0,988821
42	[19-42-1]	883	0,000002	570605,75	604768,37	586760,12	7746,756	0,993752	0,987514
43	**[19-43-1]**	**904**	**0,000002**	**485755,34**	**588509,06**	**545829,93**	**7638,710**	**0,99501**	**0,989982**
44	[19-44-1]	925	0,000002	532781,43	564216	580912,87	7766,833	0,994335	0,988615
45	[19-45-1]	946	0,000002	473598,68	550404,25	565170,12	7699,089	0,995259	0,990421
46	[19-46-1]	967	0,000002	542487,25	561084,5	571609,81	7867,659	0,993982	0,987916
47	[19-47-1]	988	0,000002	561273,93	594196,93	572554,93	7941,388	0,993771	0,987538
48	***[19-48-1]***	***1009***	***0,000002***	***456975,53***	***530035,68***	***528736,56***	***7791,788***	***0,995301***	***0,990617***
49	[19-49-1]	1030	0,000002	487912,06	570244,87	552619,43	7894,839	0,995082	0,990165
50	[19-50-1]	1051	0,000002	502736,81	575865	577351,68	7964,735	0,994944	0,98983

5.2.1.1. ANN Calculations of Istanbul Daily Gas Prediction Training Values

After best model for our values were defined as 19 input, 48 hidden layer neurons and 1 output number, this model is applied on all seven training algorithm that mentioned before. Best algorithm for forecasting of daily values was found as Quick Propagation Algorithm. Network is trained 200 times and figure 5.2. shows the target and output values of training data sets.

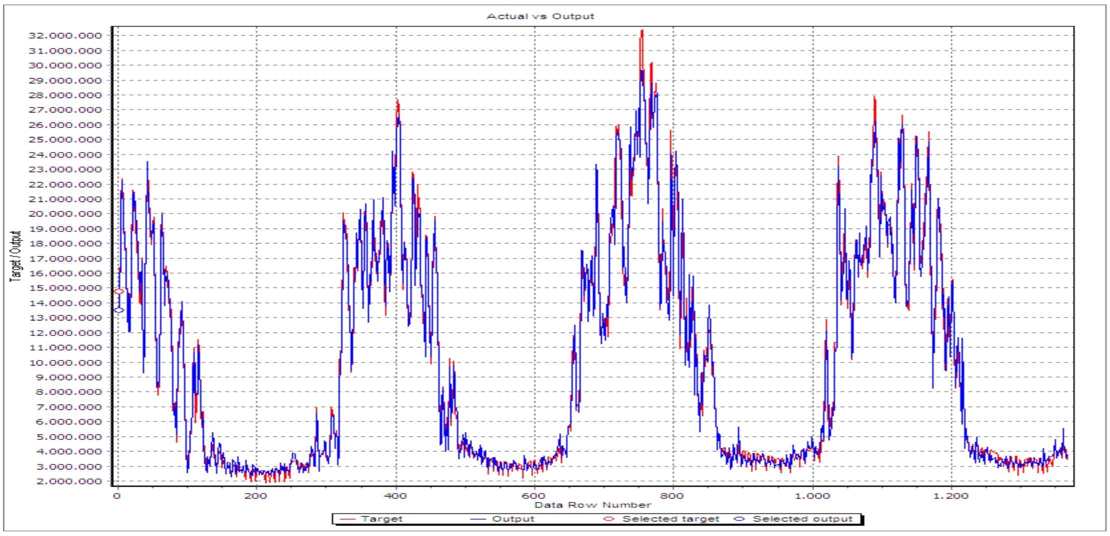


Figure 5.2. Istanbul daily values training set target- output graph

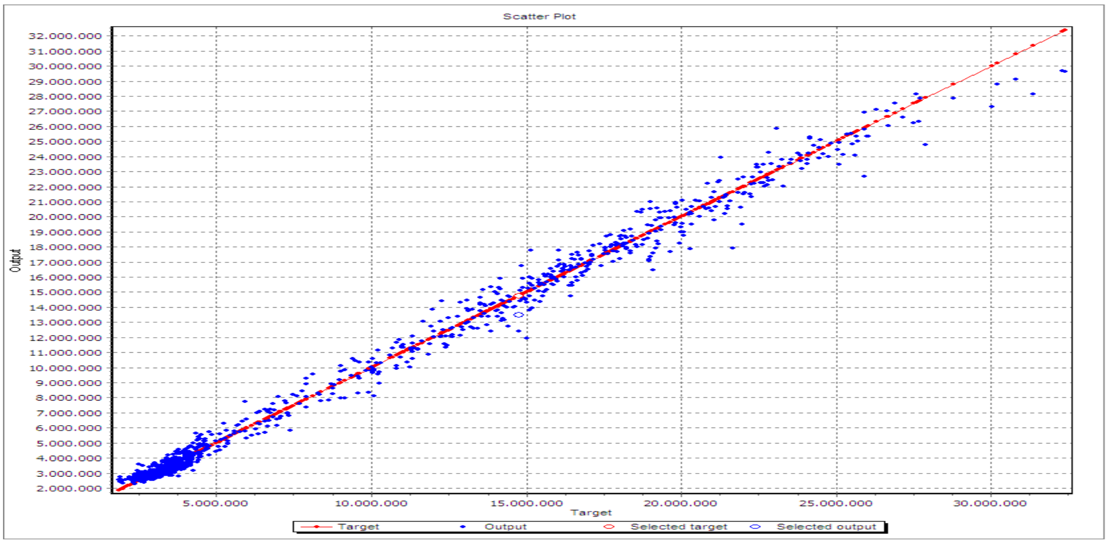


Figure 5.3. Istanbul daily values training set target- output scatter plot graph

Above graph, Figure 5.3., also shows the target and output values. Red line shows the target actual data and blue ones shows the predicted values. Deviations from the actual values are clearly seen in the table. Actual numerical values are as in Table 5.3.

Table 5.3. Istanbul daily training values results

	Target	Output	AE	ARE
Mean:	9900951,048	9872261,283	480675	0,064
Std. Dev.:	7513911,639	7451964,663	490330	0,058
Min.:	1856581	2289305,195	448,691	0,00002
Max.:	32386915	29683989,97	3759195	0,446
Correlation :	0,995832			
R-Square :	0,99151			

As it can be seen above results that for training algorithms ‘correlation’ and ‘R-square’ are all very well with the results over 99%. Another criteria for the analysis of error used for training results is Absolute Relative Error (ARE). Absolute error is the absolute value of the difference between target and output values. ARE is stands for Absolute Relative Error which is an error value that indicates the quality of neural network training. This index is calculated by dividing the difference between actual and desired values of output values by the module of desired output value.

The ARE values are changing between 0, 00002 and 0,446 with the standard deviation of 0,058. Average training error is about % 6.

Although it seems here, the actual cases are smaller than this value since there are several extreme errors in the data set greater than 15%. These are the ones as mentioned before in factor analyzing section, because of the complexity of Istanbul natural gas distribution system there are some other undefined reasons that effects the consumption of natural gas. After these extremes are excluded the ARE value is less than 5%.

Below graphs and values in the tables shows the validation and test results of Istanbul daily forecasting values.

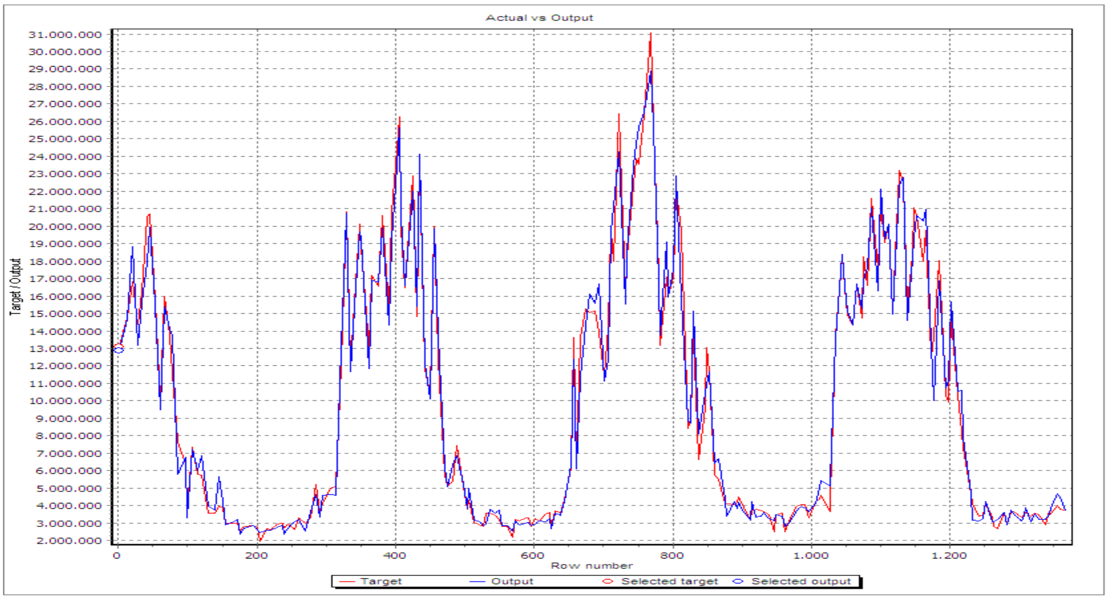


Figure 5.4. Istanbul daily values validation set target- output graph

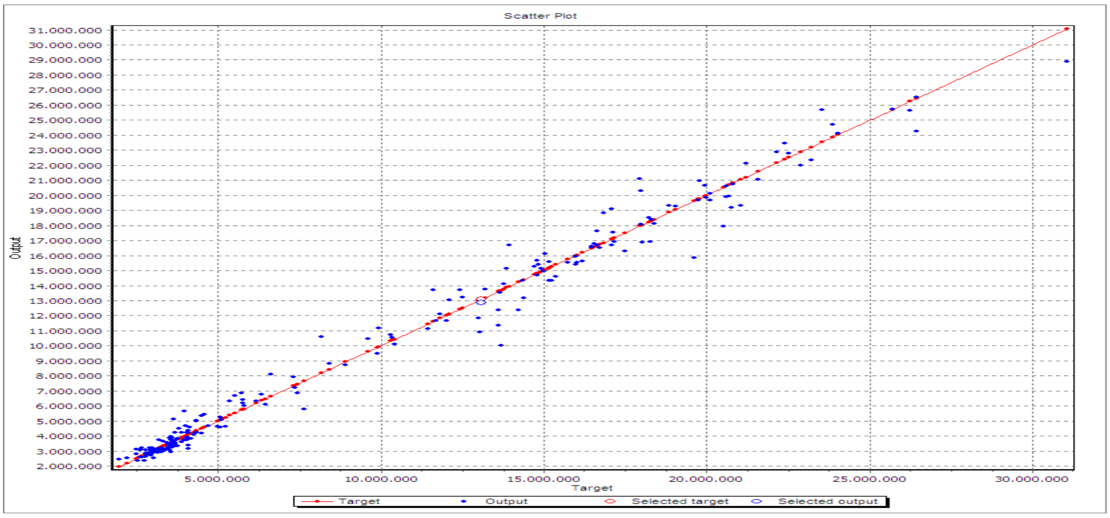


Figure 5.5. Istanbul daily values validation set target- output scatter plot graph

Table 5.4. Istanbul daily validating values results

	Target	Output	AE	ARE
Mean:	9690552,936	9703016,35	573703	0,072
Std Dev:	7181008,475	7130328,531	667737	0,069
Min:	1964488	2349123,953	21,7	0,000002
Max:	31067144	28902829,56	3762123	0,422
Correlation :	0,992459			
R-Squared :	0,984756			

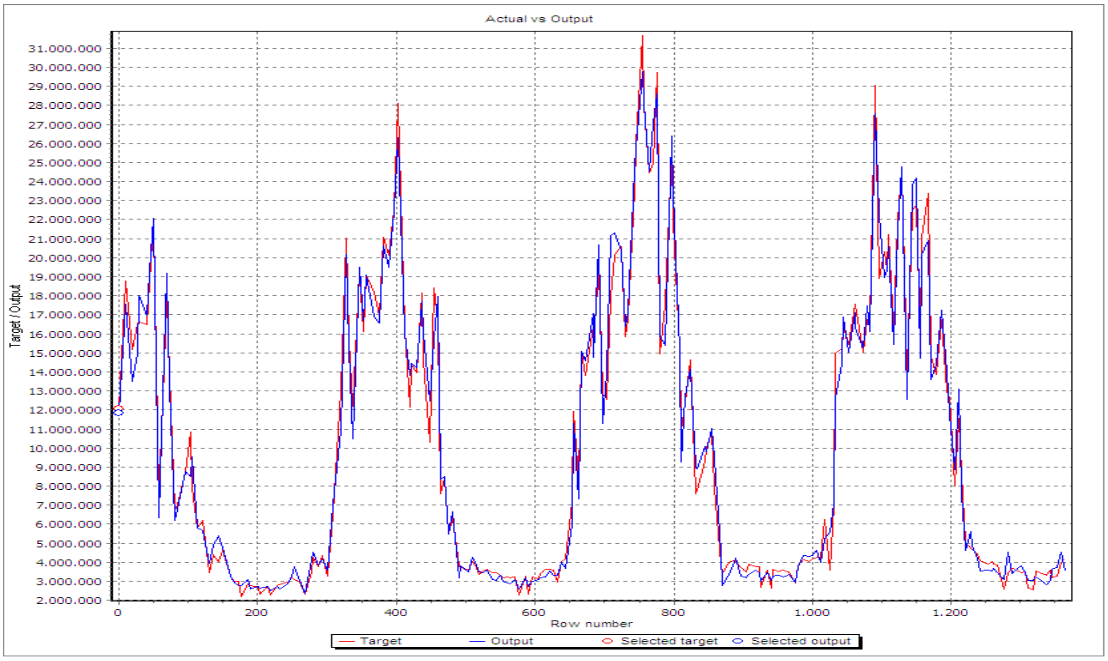


Figure 5.6. Istanbul daily values test set target- output graph

Table 5.5. Istanbul daily testing values results.

	Target	Output	AE	ARE
Mean:	9881807,945	9855060,49	630634,824	0,078
Std Dev:	7441695,131	7419953,250	622067,642	0,068
Min:	2190417	2387828,631	3174,430	0,00024
Max:	31667264	29787799,201	3318446,646	0,552
Correlation :	0,992905			
R-Squared :	0,985748			

Best criteria for the performance of network is the results of test values. The test values are randomly selected from input values. Resultant graph of target and output values are as shown in Figure 5.6. The performance of network is very well with the result of less than %7 absolute relative error value.

We have selected the test values randomly. %60 of total input values are taken for training, %20 for validating and % 20 for testing. It is not important to select the training, validation or test values manually like for example years of 2004 and 2005 for training, year 2006 for validating and year 2007 for testing. Because in our model time factors like day month and year were taken as a numerical input values so the model is trained based on this criteria. So for testing the other days consumption values like days of 2008 year by entering the time factors and other input values ANN model can easly predict the output value.

5.2.2. ANN Forecasting Calculations for Istanbul Anatolian Side Daily Dalues

In ANN daily forecasting models there are 19 input values and one output value. Below graph shows the best 5 network models after all 50 hidden layer units were applied to the model one by one.

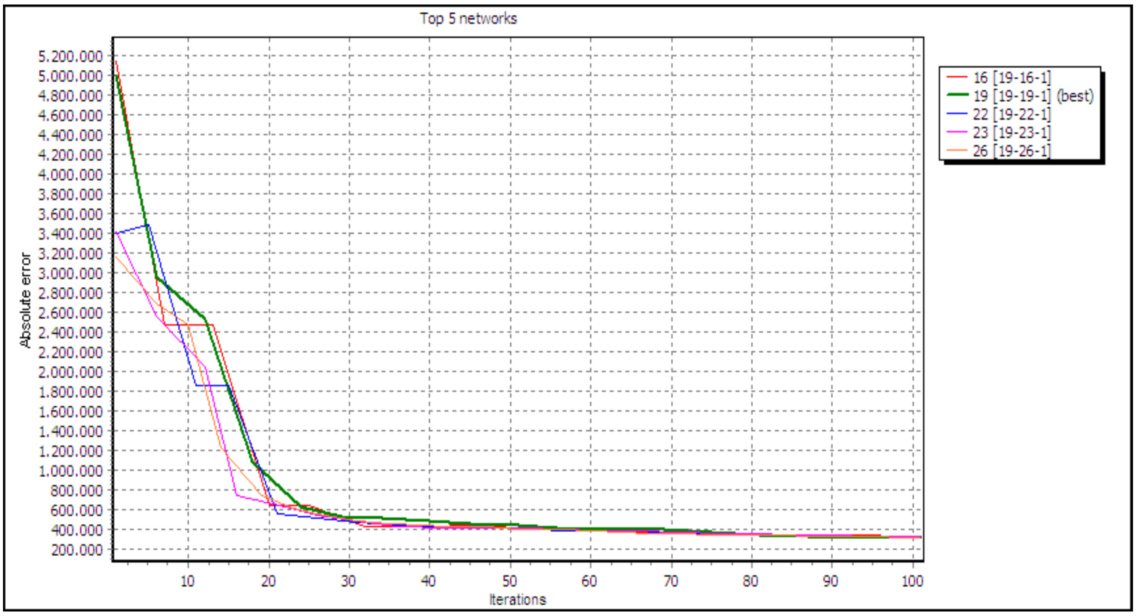


Figure 5.7. Graph of Best 5 Networks for Istanbul Anatolian Side Daily Values

[19-19-1] architecture had the best fitness

The graph Figure 5.7. shows that the model with 19 hidden neuron is best for Istanbul Anatolian side daily gas consumption forecasting in neural network model.

All of the training, validation, test, correlation and R-Square values for all hidden neuron numbers are calculated for this model again. Most of the R-Square values which shows that input values could explain our model or not, are over 96 percent.

Exhaustive search method is used for incremental of neuron number. Fitness criteria for this model is inverse test error and for verification of 50 network architectures, all of the models are retrained ten times.

Table 5.6. Importance of input factors for Istanbul Anatolian side daily forecasting

Input column name	Importance, %
• **Day**	1,112914
• **Month**	12,852396
• **Year**	10,127645
• **Day Factor**	1,089544
• **Pre. Days Max. Temp.**	6,160945
• **Pre. Days Min. Temp.**	1,554942
• **Max. Temp.**	17,501937
• **Min. Temp.**	3,15523
• **Anatolian Side Customer Number**	4,360342
• **Previous Day Consumption**	42,084105

Above Table 5.4. results shows that most important factors for daily forecasting of natural gas of Istanbul Anatolian side is "previous day consumption" and "maximum temperature" values with % 42,08 and % 17,5 importance.

The results of training, validating and test values are as shown below.

Table 5.7. Istanbul Anatolian side daily training values results

	Target	Output	AE	ARE
Mean:	3253975,49	3250877,56	259750,58	0,114528
Std Dev:	2882399,70	2840037,52	338738,81	0,15804
Min:	501000	558754,37	236,79	0,000168
Max:	12126701	11063574,63	2321515,78	2,141211
Correlation :	0,988981			
R-Squared :	0,977401			

Table 5.8. Istanbul Anatolian side daily validating values results.

	Target	Output	AE	ARE
Mean:	3179074,90	3180526,24	376653,08	0,155347
Std Dev:	2793875,88	2757799,84	523301,74	0,22299
Min:	501000	574329,78	1285,92	0,00018
Max:	12203435	10980399,06	2991469,76	1,595955
Correlation :	0,973108	**R-Squared :**	0,94534	

Table 5.9. Istanbul Anatolian side daily testing values results.

	Target	Output	AE	ARE
Mean:	3226788,22	3156187,11	389061,45	0,159205
Std Dev:	2867195,49	2756596,14	589881,92	0,252094
Min:	502000	580332,05	775,65	0,000106
Max:	11371779	10260063,83	4005180,28	2,277532
Correlation :	0,969501			
R-Squared :	0,934289			

As it can be seen above results that for training algorithms ‘correlation’ and ‘R-square’ are all very well with the results over % 97.

ARE values for training values is about %11 as shown in Table 5.7., validating and testing are about %15 as in the tables 5.8 and 5.9. In some values maximum ARE is very extreme. These are the ones as mentioned before there may be some false in data values or the conditions for these days may be different than the other normal days, like the weather is more cold than normal, holiday, crisis in city etc.

Although it seems here again, the actual cases are smaller than this value since there are several extreme errors in the data set greater than 30%. These are the ones as mentioned before in factor analyzing section, because of the complexity of Istanbul natural gas distribution system there are some other undefined reasons that effects the consumption of natural gas. And also some of the days’ input values may be false. After these extremes are excluded the ARE value would be less than 10%.

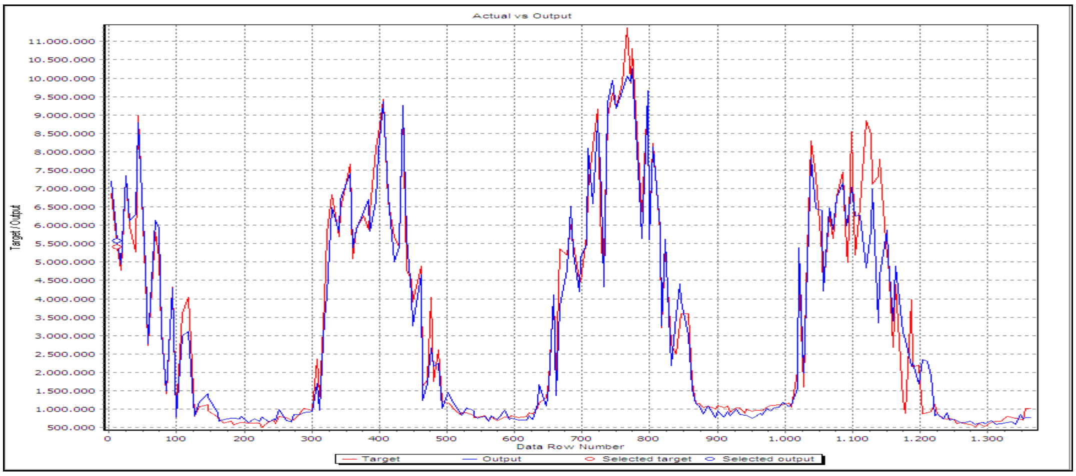


Figure 5.8. Graph of test results for Istanbul Anatolian side daily values

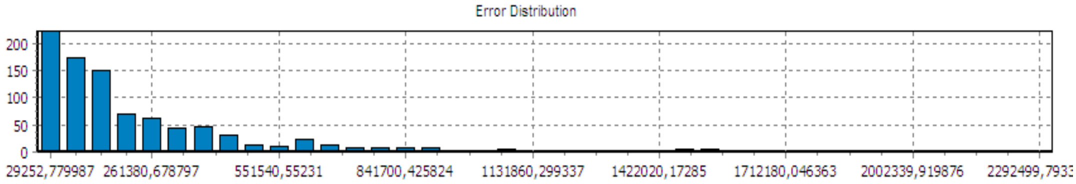


Figure 5.9. Histogram of error distribution for prediction of Anatolian side values

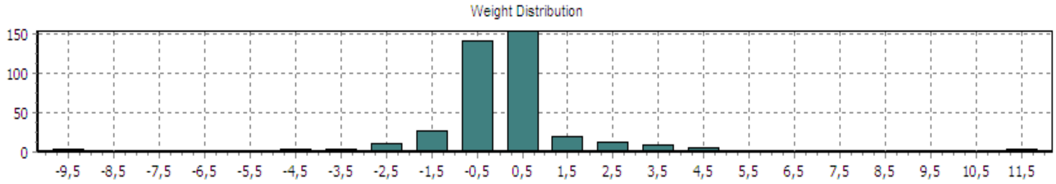


Figure 5.10. Histogram of weigth distributions for prediction of Anatolian side values

Above graphs shows test results of istanbul Anatiolian side. As seen in the Figure 5.7. test results' ARE values of 1100 and 1200 are very high. Reason of these false are explained above. Figure 5.8. shows the error distribution of test values. As seen in the table %90 of the values are less than %8 ARE but rest of it, %10, is very high.

Table 5.10. Results for Istanbul Anatolian daily forecasting values

ID	Architecture	# of Weights	Fitness	Train Error	Validation Error	Test Error	AIC	Correlation	R-Squared
1	[19-1-1]	22	0,000002	459980,59	503934,87	537284,62	5823,897517	0,962972	0,926865
2	[19-2-1]	43	0,000002	383712,78	441790,25	452165,75	5696,935265	0,970276	0,941049
3	[19-3-1]	64	0,000002	377347,12	464328,03	447697,40	5723,343964	0,972238	0,944912
4	[19-4-1]	85	0,000002	370985,34	444662,84	436802,09	5749,49724	0,972447	0,945409
5	[19-5-1]	106	0,000002	354211,06	416495,09	436303,84	5748,374011	0,9779	0,956136
6	[19-6-1]	127	0,000002	376345,53	455585,78	456869,53	5846,866936	0,972394	0,945391
7	[19-7-1]	148	0,000002	358758,46	427000,68	420935,65	5844,262925	0,974319	0,948869
8	[19-8-1]	169	0,000002	361415,96	431767,46	416682,09	5893,141257	0,974321	0,949056
9	[19-9-1]	190	0,000002	363150,84	428039,96	421788,31	5939,604383	0,97668	0,9536
10	[19-10-1]	211	0,000002	355826,62	432272,81	421722,84	5962,615167	0,977155	0,95463
11	[19-11-1]	232	0,000002	327023,87	399955	421255,59	5925,944627	0,97996	0,96016
12	[19-12-1]	253	0,000002	342284,96	402148,56	426868,46	6010,453578	0,978982	0,958227
13	[19-13-1]	274	0,000002	344228,68	425601,56	412806,75	6057,731122	0,976535	0,95345
14	[19-14-1]	295	0,000002	340530,78	409110,81	412145,62	6089,664869	0,977075	0,95437
15	[19-15-1]	316	0,000002	333152,75	413240,25	413042,43	6111,249975	0,979424	0,958919
16	**[19-16-1]**	**337**	**0,000003**	**324747,28**	**385739,5**	**391243,18**	**6129,433811**	**0,980264**	**0,960835**
17	[19-17-1]	358	0,000002	351358,53	405454,09	420574,31	6244,838001	0,978487	0,957334

Table 5.10. Results for Istanbul Anatolian daily forecasting values (continued)

ID	Architecture	# of Weights	Fitness	Train Error	Validation Error	Test Error	AIC	Correlation	R-Squared
18	[19-18-1]	379	0,000003	321265,84	395546,81	398952,84	6203,388405	0,981585	0,963095
19	***[19-19-1]***	***400***	***0,000003***	***310674,40***	***384779,46***	***384544,62***	***6214,1445***	***0,982443***	***0,964894***
20	[19-20-1]	421	0,000002	345187,37	419123,15	412351,31	6354,323171	0,978183	0,956431
21	[19-21-1]	442	0,000003	335651,56	406496,09	399517,53	6370,214302	0,978203	0,956621
22	**[19-22-1]**	**463**	**0,000003**	**324961,62**	**407494,40**	**398496,65**	**6382,048784**	**0,979347**	**0,958651**
23	**[19-23-1]**	**484**	**0,000003**	**321301,281**	**392498,15**	**393863,81**	**6413,491159**	**0,979739**	**0,959729**
24	[19-24-1]	505	0,000002	348162,218	439965,68	414731,96	6530,320829	0,978156	0,956657
25	[19-25-1]	526	0,000002	336240,25	421464,53	410769,09	6539,847504	0,977502	0,955321
26	**[19-26-1]**	**547**	**0,000003**	**311368,781**	**403600,65**	**384712,81**	**6510,225244**	**0,981379**	**0,962977**
27	[19-27-1]	568	0,000002	313656,968	389154,43	403432,90	6559,049299	0,980929	0,962116
28	[19-28-1]	589	0,000002	335010,656	417185,53	400049,40	6662,433011	0,978099	0,95648
29	[19-29-1]	610	0,000002	331514,15	407486,25	416026,65	6694,65	0,978587	0,957591
30	[19-30-1]	631	0,000002	342484,68	425636,06	412006,12	6766,9	0,979573	0,959332
31	[19-31-1]	652	0,000002	330712,87	411417,31	418021,71	6776,39	0,978288	0,957026
32	[19-32-1]	673	0,000002	330628	428265,28	407157,87	6818,16	0,978598	0,957517
33	[19-33-1]	694	0,000002	334415,56	405753,43	412211,84	6870,77	0,97878	0,957884
34	[19-34-1]	715	0,000002	317300,81	408276,81	417166,34	6863,81	0,981561	0,963196

Table 5.10. Results for Istanbul Anatolian daily forecasting values (continued)

ID	Architecture	# of Weights	Fitness	Train Error	Validation Error	Test Error	AIC	Correlation	R-Squared
35	[19-35-1]	736	0,000002	332664,156	412435,75	411197,25	6949,88	0,97958	0,959333
36	[19-36-1]	757	0,000002	332705,93	425371,87	420424,37	6991,99	0,97753	0,955445
37	[19-37-1]	778	0,000002	344783,84	413326,25	415753,28	7067,23	0,97554	0,951552
38	[19-38-1]	799	0,000002	316562,62	383634,37	405914,25	7029,64	0,981656	0,963545
39	[19-39-1]	820	0,000002	328427,34	407016,56	408778,56	7105,93	0,979215	0,95878
40	[19-40-1]	841	0,000002	316524,84	403292,93	401790,65	7113,53	0,979492	0,959142
41	[19-41-1]	862	0,000003	315619,59	388115,65	399365,90	7152,86	0,980778	0,961817
42	[19-42-1]	883	0,000002	324266,56	412479,62	414111,28	7220,05	0,97943	0,958998
43	[19-43-1]	904	0,000002	320707,56	403132,53	405730,65	7251,76	0,979493	0,959301
44	[19-44-1]	925	0,000002	334076,90	431493,40	406161,53	7331,83	0,978605	0,957596
45	[19-45-1]	946	0,000002	333710,81	408961,81	405117,53	7372,80	0,978034	0,956445
46	[19-46-1]	967	0,000002	336960,78	416466,81	416275	7423,84	0,978601	0,957562
47	[19-47-1]	988	0,000002	325820,21	408611,781	417291,59	7434,50	0,978483	0,957334
48	[19-48-1]	1009	0,000002	323976,71	420064,781	416191	7471,21	0,979995	0,960272
49	[19-49-1]	1030	0,000002	328474,15	397744,031	416990,87	7526,06	0,98051	0,961052
50	[19-50-1]	1051	0,000002	311016,37	395418,06	402291,31	7517,16 3	0,981013	0,962208

Table 5.10. Results for Istanbul Anatolian daily forecasting values (continued)

Retrain	**Iters**	**Train error**	**Validation error**	**Test error**	**AIC**	**Correlation**	**R-Squared**
1	101	367560,63	446676,75	456290,84	6370,85	0,975414	0,951019
2	101	350764,55	427273,74	447989,01	6327,26	0,976596	0,953618
3	101	310674,41	384779,45	384544,62	6214,14	0,982443	0,964894
4	101	327796,50	399030,11	406472,74	6264,14	0,979031	0,958342
5	101	346500,43	422252,34	430885,96	6315,86	0,979254	0,958846
6	101	359945,47	442389,86	446303,38	6351,34	0,976268	0,952902
7	101	331157,46	409898,87	420289,55	6273,65	0,979496	0,959338
8	101	354015,95	402143,02	434810,35	6335,86	0,978784	0,957834
9	101	347827,44	423309,72	428182,32	6319,42	0,975969	0,952351
10	101	321705,65	401722,78	417457,67	6246,66	0,980352	0,961057
SUMMARY:							
Min:		310674,41	384779,45	384544,62	6214,14	0,975414	0,951019
Max:		367560,63	446676,75	456290,84	6370,85	0,982443	0,964894
Average:		341794,85	415947,66	427322,64	6301,92	0,978361	0,95702
Std. dev.:		3076201,84	3743576,48	3845957,24	56717,3	8,805247	8,613183

5.2.3. ANN Forecasting Calculations for Istanbul European Side Daily Values

In ANN daily forecasting models there are 19 input values and one output value is used again. Below graph shows the best 5 network models after all 50 hidden layer units were applied to the model one by one.

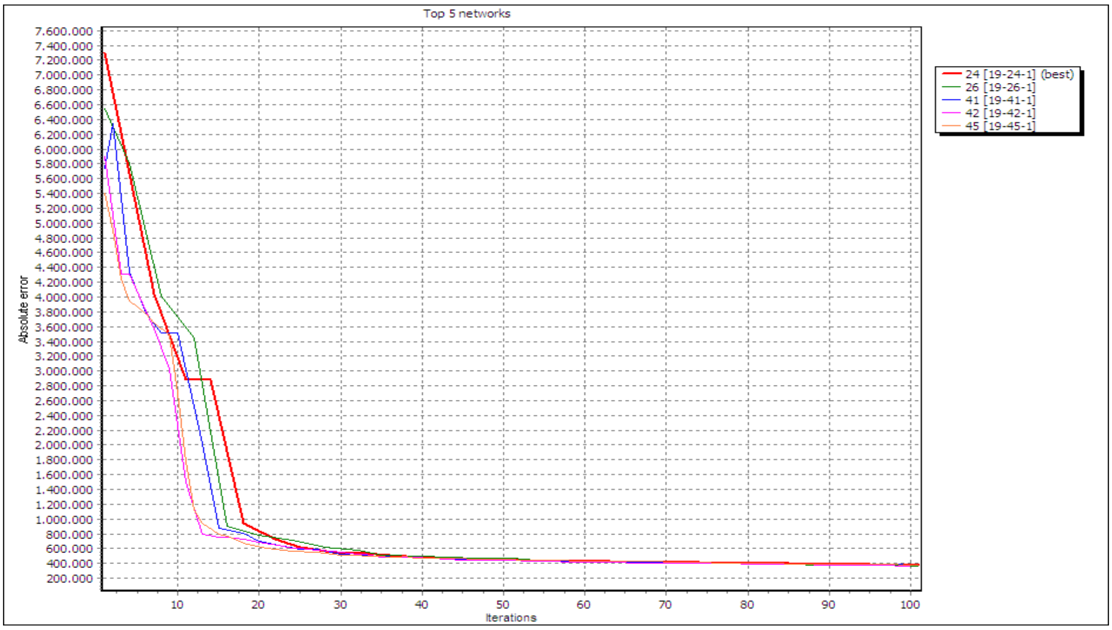


Figure 5.11. Graph of best 5 networks for Istanbul European side daily values

[19-24-1] architecture had the best fitness

The graph Figure 5.10. shows that the model with 24 hidden neuron is best for Istanbul European side daily gas consumption forecasting in neural network model.

All of the training, validation, test, correlation and R-Square values for all hidden neuron numbers are calculated for this model again. As seen in the Table 5.15. most of the R-Square values are over 98 percent.

Exhaustive search method is used for incremental of neuron number. Fitness criteria for this model is inverse test error and for verification of 50 network architectures, all of the models are retrained ten times.

Table 5.11. Importance of input factors for Istanbul European side daily forecasting

Input column name	Importance, %
• **Month**	0,111379
• **European Side Customer Number**	0,615777
• **Day**	0,131115
• **Day Factor**	1,310359
• **Max. Temp.**	29,238364
• **Min. Temp.**	2,390508
• **Pre. Days Max. Temp.**	4,94049
• **Pre. Days Min. Temp.**	0,672399
• **Year**	1,015724
• **Previous Day Consumption**	59,573885

Above Table 5.11. results shows that most important factors for daily forecasting of natural gas of Istanbul European side are “previous day consumption” and “maximum temperature” values with % 59,57 and % 29,23 importance.

The results of training, validating and test values are as shown below. Performance of the model for Istanbul European side forecasting values are better than Anatolian side values. The ARE values for training’ validating and test values are about %5 which was %15 for the Anatolian side. There may be some reason of this like the values of European side is more correct than other or gas consumption stuation of European side is more stable than Anatolian side.

Table 5.12. Istanbul European side daily training values results.

	Target	Output	AE	ARE
Mean:	6387010,87	6378298,39	312699,74	0,054177
Std Dev:	4477142,49	4431223,42	371276,16	0,050758
Min:	1355581	1635784,16	402,41	0,000133
Max:	20107135	18116164,27	3113347,98	0,510941
Correlation :	0,994117			
R-Squared :	0,988			

Table 5.13. Istanbul European side daily validating values results.

	Target	Output	AE	ARE
Mean:	6297174,53	6336784,14	313730,01	0,057681
Std Dev:	4413164,31	4373225,39	325666,80	0,055569
Min:	1647700	1781890,89	1573,03	0,000404
Max:	19545308	18149498,20	1576237,40	0,330433
Correlation :	0,994784			
R-Squared :	0,989308			

Table 5.14. Istanbul European side daily testing values results.

	Target	Output	AE	ARE
Mean:	6359658,97	6349761,49	377080,27	0,065415
Std Dev:	4590873,56	4483072,03	437491,85	0,05856
Min:	1432305	1644733,32	1978,23	0,000944
Max:	20260214	18020780,95	2731607,37	0,351157
Correlation :	0,992181			
R-Squared :	0,983402			

R-square and correlation values for training, validating and testing values are over %98.5. These values are very good results for forecasting models because this means that input values are very suitable for the model. Correlation values with over % 99 show the relation of input between each other has the best fitness.

ARE values for training values is about %5 as shown in Table 5.12., validating and testing are about %5 and %6 as in the tables 5.13 and 5.14. There is no very much extreme ARE values in input values as shown in Figures 5.11. and 5.12.

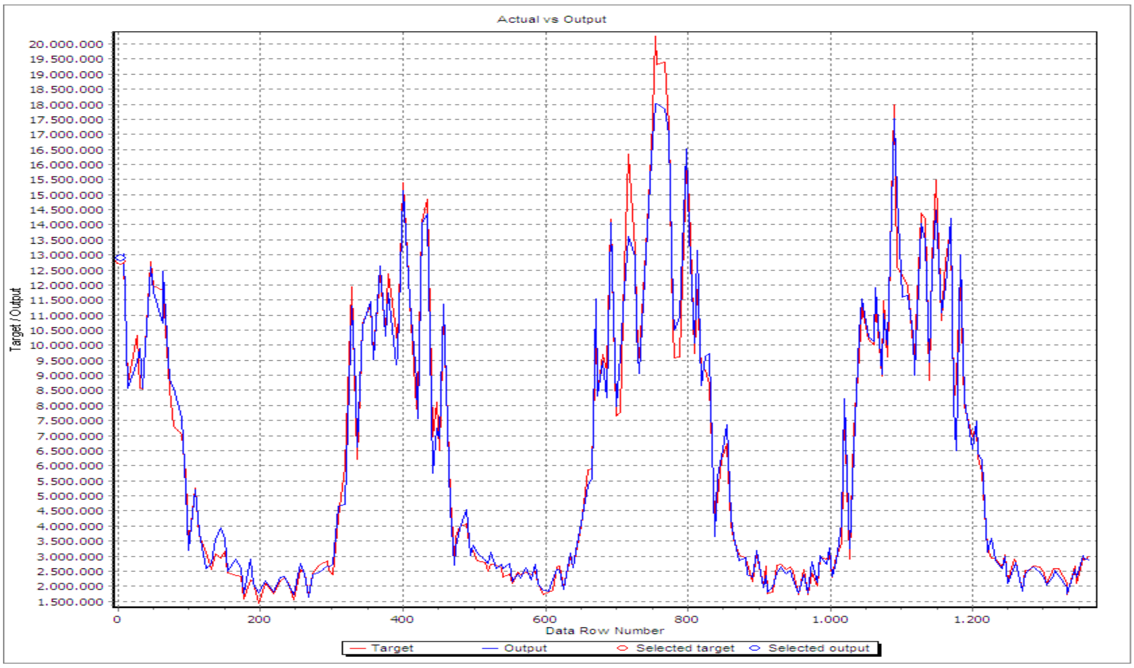


Figure 5.12. Graph of test results for Istanbul European side daily values

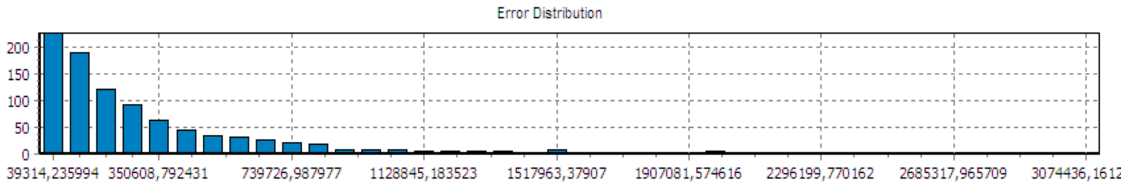


Figure 5.13. Histogram of error distribution for prediction of European side values

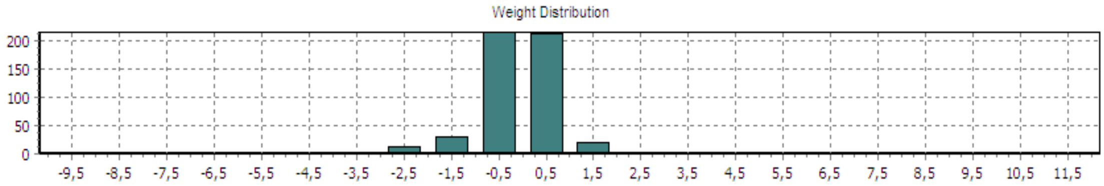


Figure 5.14. Histogram of weigth distribution for prediction of European side values

Above graphs shows test results of Istanbul European side. Figure 5.12. shows the error distribution of test values. As seen in the table %95 of the values are less than %6 ARE. Weight distribution graph shows the weights of our model that modelled by ANN training algorithm.

Table 5.15. Results for Istanbul European daily forecasting values

ID	Architecture	# of Weights	Fitness	Train Error	Validation Error	Test Error	AIC	Correlation	R-Squared
1	[19-1-1]	22	0,000002	498450,56	434265,34	573176,25	5898,75	0,986	0,972
2	[19-2-1]	43	0,000002	491420,71	441882,37	551465,06	5927,51	0,987	0,975
3	[19-3-1]	64	0,000002	431148,18	404832,03	501619,81	5847,56	0,990	0,981
4	[19-4-1]	85	0,000002	402885,718	399492,56	466574,75	5826,37	0,991	0,982
5	[19-5-1]	106	0,000002	398882,468	366472,218	466079,43	5859,07	0,991	0,982
6	[19-6-1]	127	0,000002	421777,093	397254,46	478303,12	5953,08	0,991	0,982
7	[19-7-1]	148	0,000002	396141,65	397521,37	453285,96	5936,64	0,991	0,983
8	[19-8-1]	169	0,000002	404868,75	414209,09	480912,21	5998,95	0,991	0,982
9	[19-9-1]	190	0,000002	376222,96	379592,90	442182,21	5972,56	0,992	0,984
10	[19-10-1]	211	0,000002	392135,78	374893,43	458180,56	6053,17	0,991	0,983
11	[19-11-1]	232	0,000002	406415,62	393604,06	465880,06	6128,50	0,991	0,982
12	[19-12-1]	253	0,000002	404886,34	376526,718	471039,84	6166,99	0,991	0,982
13	[19-13-1]	274	0,000002	394349,5	372419,68	454605,90	6184,41	0,991	0,983
14	[19-14-1]	295	0,000002	384827,15	358194	451157	6203,63	0,991	0,983
15	[19-15-1]	316	0,000002	381512,84	364786,62	446768,375	6237,57	0,992	0,984
16	[19-16-1]	337	0,000002	386627,15	388534,31	458217,46875	6291,98	0,992	0,984
17	[19-17-1]	358	0,000002	379573,71	381314,53	445634,78125	6316,82	0,992	0,984

Table 5.15. Results for Istanbul European daily forecasting values(continued)

ID	Architecture	# of Weights	Fitness	Train Error	Validation Error	Test Error	AIC	Correlation	R-Squared
18	[19-18-1]	379	0,000002	383093,75	372330,75	444616,625	6367,43	0,992	0,984
19	[19-19-1]	400	0,000002	388263,34	383237,71	458558,656	6421,92	0,991	0,983
20	[19-20-1]	421	0,000002	388254,84	384712,15	467769,375	6463,90	0,991	0,983
21	[19-21-1]	442	0,000002	374051,59	369107,12	442555,281	6471,16	0,992	0,984
22	[19-22-1]	463	0,000002	380895,90	365331,90	442181,625	6530,06	0,992	0,984
23	[19-23-1]	484	0,000002	372700,59	347470,25	438210,781	6551,79	0,992	0,984
24	[19-24-1]	505	0,000002	369578,25	368196,84	428057,8125	6585,95	0,992	0,984
25	[19-25-1]	526	0,000002	363429,43	375513,43	441932,59	6612,31	0,992	0,984
26	[19-26-1]	547	0,000002	365235,12	358749,09	428223,59	6658,93	0,992	0,985
27	[19-27-1]	568	0,000002	369067,56	358425,15	436580	6710,66	0,992	0,984
28	[19-28-1]	589	0,000002	374226,59	377850,15	437938	6765,60	0,992	0,984
29	[19-29-1]	610	0,000002	387090,84	378363,68	450887,93	6839,10	0,991	0,983
30	[19-30-1]	631	0,000002	370194,53	365543,53	439214,87	6839,50	0,992	0,985
31	[19-31-1]	652	0,000002	394942,75	377933,62	443091,03	6941,82	0,991	0,983
32	[19-32-1]	673	0,000002	389284,21	380460,62	451276,18	6970,37	0,991	0,983
33	[19-33-1]	694	0,000002	365703,87	362502,87	448425	6954,13	0,992	0,984
34	[19-34-1]	715	0,000002	372963,06	347166,62	439378,25	7014,45	0,992	0,984

Table 5.15. Results for Istanbul European daily forecasting values (continued)

ID	Architecture	# of Weights	Fitness	Train Error	Validation Error	Test Error	AIC	Correlation	R-Squared
35	[19-35-1]	736	0,000002	379870,43	382287,40	448658,93	7073,55	0,992	0,984
36	[19-36-1]	757	0,000002	372787,875	355280,40	447427,15	7098,01	0,992	0,985
37	[19-37-1]	778	0,000002	382582,90	356871,5	445909,37	7164,18	0,991	0,983
38	[19-38-1]	799	0,000002	371945,93	358486,03	440974,09	7179,90	0,992	0,984
39	[19-39-1]	820	0,000002	385320,781	356932,93	446125,93	7254,83	0,992	0,984
40	[19-40-1]	841	0,000002	381844,062	379039,5	444876,31	7288,38	0,992	0,984
41	[19-41-1]	862	0,000002	374921,781	364127,62	429402	7313,33	0,992	0,984
42	[19-42-1]	883	0,000002	367797,90	363356,56	434116	7337,45	0,992	0,985
43	[19-43-1]	904	0,000002	372068,62	366842,5	440809,68	7390,21	0,992	0,985
44	[19-44-1]	925	0,000002	377937,43	368287,375	444603,68	7446,80	0,992	0,984
45	[19-45-1]	946	0,000002	377648,59	357144,62	436342,96	7488,08	0,992	0,984
46	[19-46-1]	967	0,000002	386325,12	367490,90	447536,75	7551,25	0,991	0,983
47	[19-47-1]	988	0,000002	383291,21	357039,31	444133,87	7585,91	0,992	0,984
48	[19-48-1]	1009	0,000002	370988,03	359618,75	437120,03	7597,50	0,992	0,984
49	[19-49-1]	1030	0,000002	376832,28	355362,71	442742,43	7654,07	0,992	0,984
50	[19-50-1]	1051	0,000002	375620,34	371188,18	446222,15	7693,06	0,992	0,984

Table 5.15. Results for Istanbul European daily forecasting values (continued)

Retrain	**Iters**	**Train error**	**Validation error**	**Test error**	**AIC**	**Correlation**	**R-Squared**
1	101	369578,23	368196,83	428057,82	6585,95	0,992352	0,984661
2	101	390354,23	374414,918	449051,01	6636,92	0,991912	0,983773
3	101	394458,89	390545,46	471469,74	6646,67	0,991789	0,983587
4	101	406347,92	401489,57	498752,10	6674,3	0,990987	0,981973
5	101	386671,99	384712,69	457570,28	6628,09	0,992072	0,984128
6	101	413376,33	409996,58	477299,64	6690,33	0,991019	0,981826
7	101	403644,95	390831,19	467037,12	6668,13	0,991384	0,982745
8	101	377757,92	367744,01	448563,44	6606,35	0,992332	0,984541
9	101	397792,46	382779,71	474884,02	6654,52	0,991451	0,982905
10	101	394648,89	381491,49	473742	6647,12	0,991787	0,983575
			SUMMARY:				
Min:		369578,23	367744,01	428057,82	6585,95	0,990987	0,981826
Max:		413376,33	409996,58	498752,1	6690,33	0,992352	0,984661
Average:		393463,18	385220,25	464642,72	6643,84	0,991709	0,983371
Std. dev.:		3541190,72	3467006,34	4181825,78	59794,64	8,925377	8,850342

5.2.4. ANN Forecasting Calculations for Istanbul Weekly Values

Weekly values were composed from the daily values. Temperature values are designed as an average of seven days’ average temperature values. As in the daily values average maximum and minimum temperature values were taken. Seven day average number of consumers were considered for input of consumer number.

In ANN weekly forecasting models there are 8 input values like Istanbul average customer number, year, average minimum and maximum temperature values and one output value as total weekly consumption value of natural gas for that week.

Below graph shows the best 5 network models after all 50 hidden layer units were applied to the model one by one.

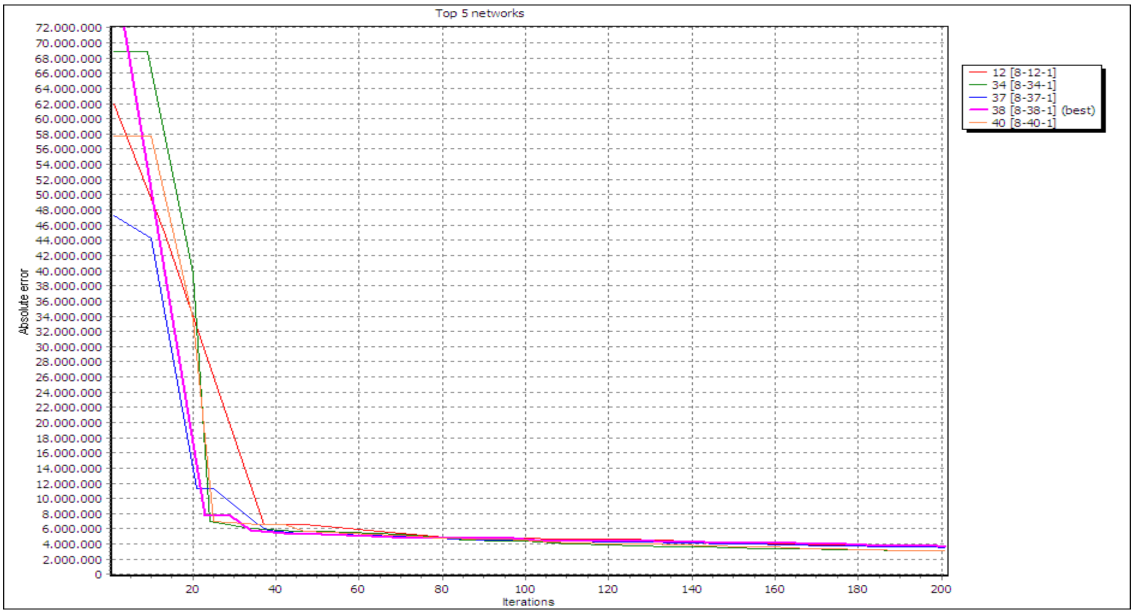


Figure 5.15. Graph of best 5 networks for Istanbul weekly values

[8-38-1] architecture had the best fitness

In weekly forecasting algorithms again exhaustive search method is used as daily models, for incremental of neuron number. Fitness criteria for weekly model is inverse test error and for verification of 50 network architectures, all of the models are retrained ten times.

The graph Figure 5.14. shows that the model with 38 hidden neuron is best for Istanbul weekly gas consumption forecasting in neural network model.

All of the training, validation, test, correlation and R-Square values for all hidden neuron numbers are calculated for this model again. Most of the R-Square values which shows that input values could explain our model or not, are over 96 percent.

Table 5.16. Importance of input factors for Istanbul side weekly forecasting

Input column name	Importance, %
• Istanbul Average Customer Number	14,353291
• Year	9,778734
• Average Minimum Temperature	9,620698
• Average Maximum Temperature	59,983923
• Week	6,263354

Above Table 5.4. results shows that most important factors for weekly forecasting of natural gas of Istanbul are “average maximum temperature” and “average customer number” values with % 59,98 and % 14,35 importance.

5.2.4.1. ANN Calculations of Istanbul Weekly Gas Prediction Training Values

After best model for our values were defined as 8 input, 38 hidden layer neuron number and 1 output number, this model is applied on all seven training algorithm that mentioned before. Best algorithm for forecasting of daily values was found as conjugate Gardient Descent Algorithm. After network is trained 200 times the target and output values of training data sets is as below graph.

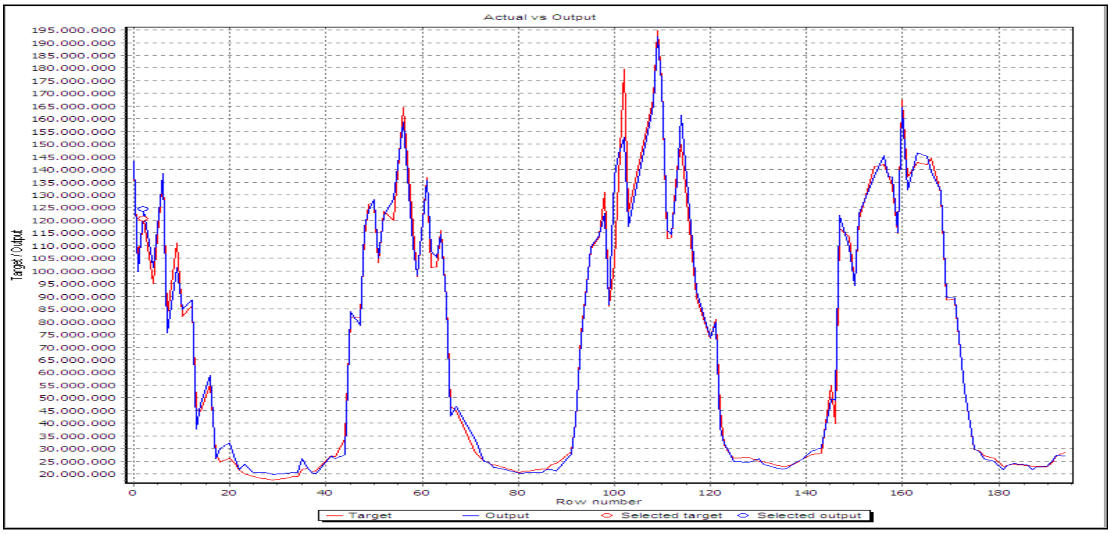


Figure 5.16. Training graph of target-output values for Istanbul weekly forecasting

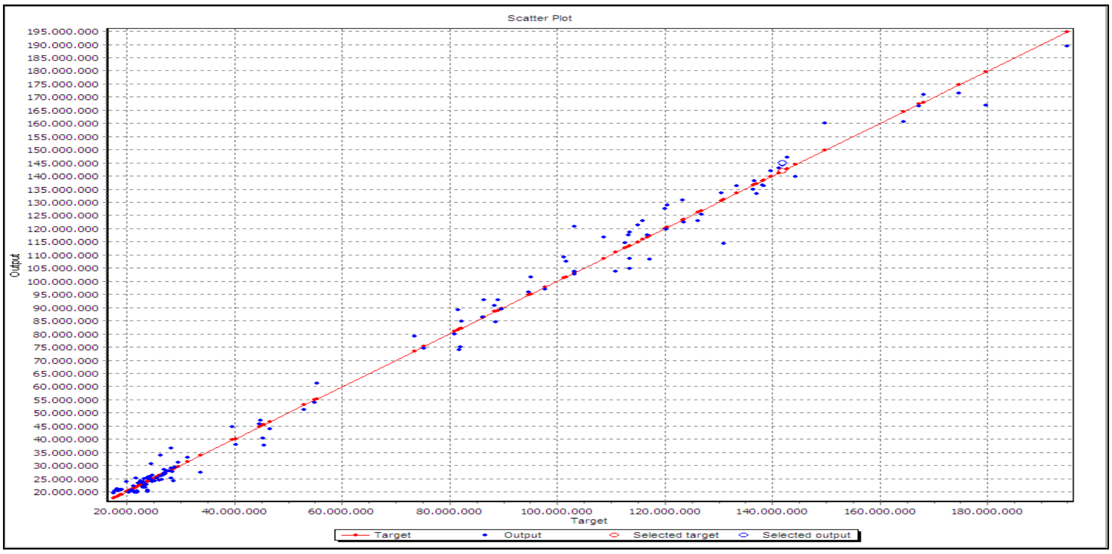


Figure 5.17. Istanbul weekly values training set target- output scatter plot graph

Above graph, Figure 5.6., also shows the target and output values. Red line shows the target actual data and blue ones shows the values that program predicted. Deviations from the actual values are clearly seen in the table. Actual numerical values are as in the Table 5.17.

Table 5.17. Istanbul weekly training values results.

	Target	Output	AE	ARE
Mean:	69408029,255	69537464,13	2867418,12	0,05
Std Dev:	49894855,7	49765248,42	4194907,86	0,056
Min:	17581208	19828318,75	6460,87	0,0002
Max:	194749953	192373832,8	34196287,82	0,33
Correlation : 0,99604				
R-Squared : 0,992014				

As it can be seen above results that for training algorithms 'correlation' and 'R-square' values are all very well with the results over 99%. ARE value is about %5 with the standard deviation value of 0,056. The ARE values are changing between % 0, 02 and % 44 with the standard deviation of % 5, 6.

Below graphs and values in the tables shows the validation and test results of Istanbul weekly forecating values.

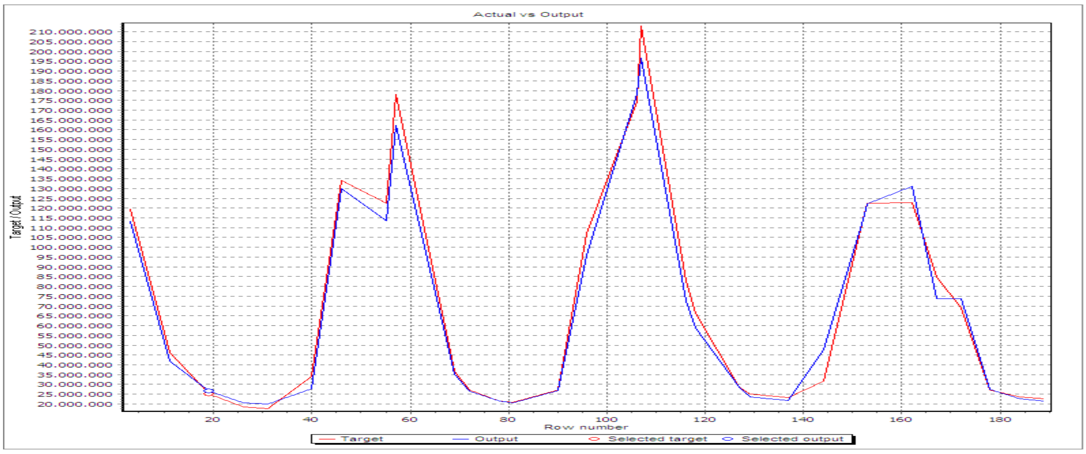


Figure 5.18. Graph of validation target-output values for Istanbul weekly forecasting

Table 5.18. Istanbul weekly validation values results.

	Target	Output	AE	ARE
Mean:	69922959,43	67639908,99	4863583,46	0,077
Std Dev:	55105731,93	52372564,84	5025552,62	0,09
Min:	17577484	20044463,24	49884,72	0,001
Max:	213037259	196489946,95	16547312,044	0,5
Correlation : 0,990888 **R-Squared :** 0,977866				

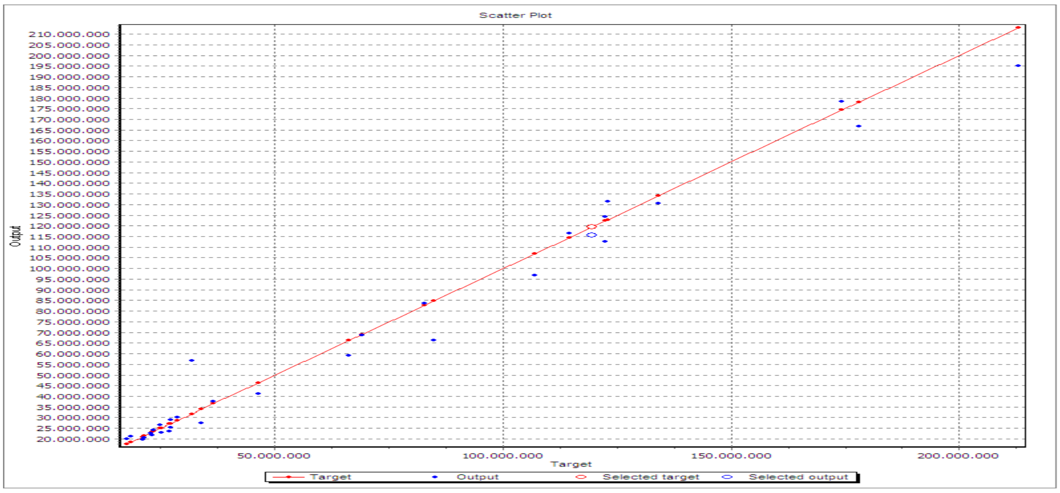


Figure 5.19. Istanbul weekly values validating set target- output scatter plot graph

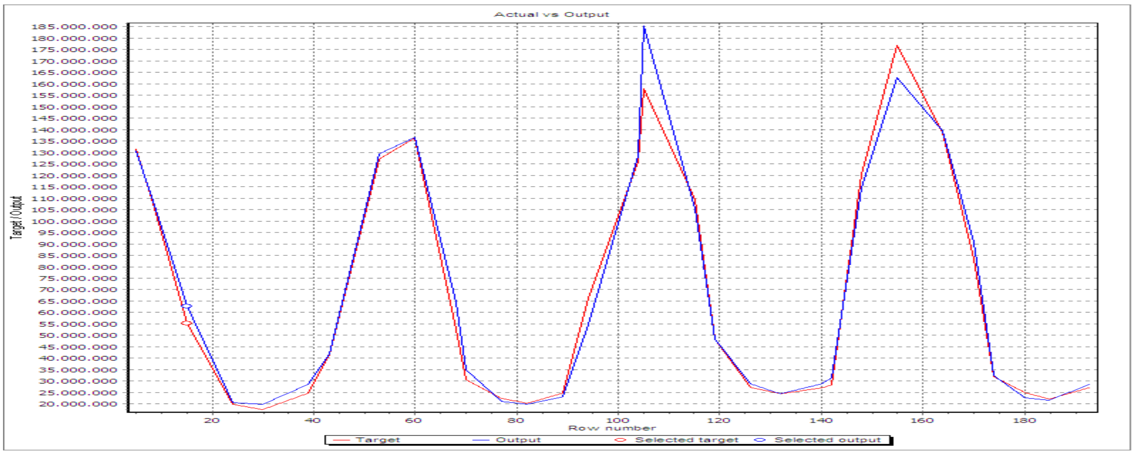


Figure 5.20. Graph of testing target-output values for Istanbul weekly forecasting

Table 5.19. Istanbul weekly testing values results.

	Target	Output	AE	ARE
Mean:	66312146,64	67527176,05	3951867,45	0,06
Std Dev:	49629549,41	50215105,43	5584685,96	0,059
Min:	17541417	19777110,76	2918,45	0,0001
Max:	176773901	185303363,87	27437332,87	0,22
Correlation :	0,990974			
R-Squared :	0,981478			

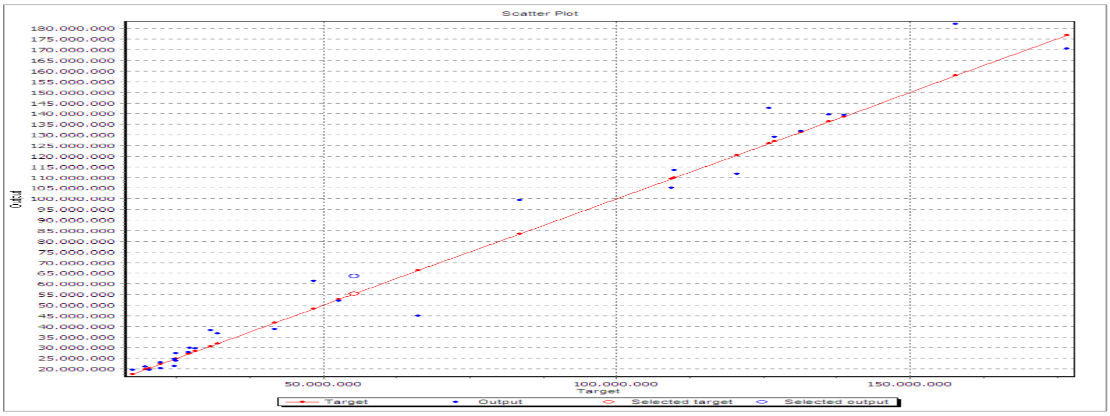


Figure 5.21. Istanbul weekly values testing set target- output scatter plot graph

In ANN forecating applications best criteria for the performance of network is the results of test values. As in weekly models the test values are randomly selected from input values. Resultant graph of target and output values are as shown in Figure 5.19. and 5.20. and Table 5.19. the performance of network is very well with the result of less than 6 % absolute relative error value.

Table 5.20. Results of ANN models for Istanbul weekly forecasting.

ID	Architecture	Weights	Fitness	Train Error	Validation Error	Test Error	AIC	Correlation	R-Squared
1	[8-1-1]	15	0,000002	460683,43	427001,15	452703,343	5790,499	0,959022	0,919218
2	[8-2-1]	29	0,000002	398293,78	366929,75	415305	5683,455	0,964656	0,930507
3	[8-3-1]	43	0,000003	391667,87	361313,81	387144,75	5695,887	0,972787	0,945847
4	[8-4-1]	57	0,000003	363147,31	361826,56	392864,031	5653,725	0,974133	0,948898
5	[8-5-1]	71	0,000003	363858,87	350792,40	362715,75	5683,542	0,976078	0,952665
6	[8-6-1]	85	0,000003	343495	325187,15625	354212,406	5658,095	0,978477	0,957401
7	[8-7-1]	99	0,000003	338361,62	324217,96875	344858,625	5672,122	0,979212	0,958764
8	[8-8-1]	113	0,000003	327390,90	311074,28125	343756,468	5669,535	0,980652	0,961646
9	[8-9-1]	127	0,000003	319287,46	309304,96875	339371,718	5674,276	0,980904	0,961597
10	[8-10-1]	141	0,000003	317795,53	334254	345241,906	5697,930	0,981921	0,964124
11	[8-11-1]	155	0,000003	316667,78	337502,90625	333379,843	5722,631	0,981471	0,96324
12	**[8-12-1]**	**169**	**0,000003**	**312695,25**	**321360,40625**	**330209,875**	**5738,916**	**0,981529**	**0,96338**
13	[8-13-1]	183	0,000003	303771,25	326398,4375	310963,406	5740,046	0,981853	0,963991
14	[8-14-1]	197	0,000003	316479,40	324488,375	330892,906	5806,079	0,981715	0,963684
15	[8-15-1]	211	0,000003	314284,81	314395,78125	328388,968	5827,621	0,981813	0,963939
16	[8-16-1]	225	0,000003	317091,03	320851,21875	330203	5863,870	0,980805	0,961908
17	[8-17-1]	239	0,000003	308866,18	307419,46875	324340,781	5867,482	0,982787	0,965772

Table 5.20. Results of ANN models for Istanbul weekly forecasting (continued)

ID	Architecture	Weights	Fitness	Train Error	Validation Error	Test Error	AIC	Correlation	R-Squared
18	[8-18-1]	253	0,000003	298393,90	301916,28125	312361,875	5863,472	0,983366	0,966997
19	[8-19-1]	267	0,000003	288550,43	306944,65625	327585,937	5860,342	0,984169	0,968573
20	[8-20-1]	281	0,000003	306911,93	315259,6875	327237,687	5945,591	0,982109	0,964514
21	[8-21-1]	295	0,000003	309690,90	303516,9375	307937,5	5981,956	0,981159	0,962599
22	[8-22-1]	309	0,000003	306961,28	300189,6875	323454,906	6001,741	0,982107	0,964426
23	*[8-23-1]*	*323*	*0,000003*	*298297,5*	*316964,09375*	*311300*	*6003,172*	*0,982827*	*0,965926*
24	[8-24-1]	337	0,000003	298915,12	306020,625	319774,375	6033,091	0,982679	0,965625
25	[8-25-1]	351	0,000003	303767,5	312760,09375	317670,812	6076,035	0,983108	0,966443
26	[8-26-1]	365	0,000003	296998,78	314148,75	321606,718	6083,123	0,983354	0,966958
27	[8-27-1]	379	0,000003	298814,90	316858,65625	319743,843	6116,780	0,983303	0,966851
28	[8-28-1]	393	0,000003	301245,09	309828,625	317277,562	6152,297	0,982632	0,965558
29	[8-29-1]	407	0,000003	304496,43	305336,0625	320839,625	6190,259	0,982831	0,96554
30	[8-30-1]	421	0,000003	288381	301203,8125	332903,5	6167,797	0,984257	0,968721
31	[8-31-1]	435	0,000003	292792,18	293674,18	318791,53	6209,88	0,983313	0,966816
32	[8-32-1]	449	0,000003	290895,56	299015,75	314373,18	6231,85	0,983802	0,967772
33	[8-33-1]	463	0,000003	291397,34	299837,5	304770,28	6261,45	0,983891	0,967911
34	**[8-34-1]**	**477**	**0,000003**	**295128,40**	**316005**	**324158,12**	**6301,26**	**0,983255**	**0,966726**

Table 5.20. Results of ANN models for Istanbul weekly forecasting (continued)

ID	Architecture	Weights	Fitness	Train Error	Validation Error	Test Error	AIC	Correlation	R-Squared
35	[8-35-1]	491	0,000003	320477,53	316815,15	313004,68	6405,72	0,980722	0,961702
36	[8-36-1]	505	0,000003	299468,15	301840,25	312035,56	6370,806	0,98281	0,965813
37	**[8-37-1]**	**519**	**0,000003**	**290177,68**	**311295,46**	**322140,53**	**6369,561**	**0,984081**	**0,968407**
38	***[8-38-1]***	***533***	***0,000003***	***311656,93***	***313872,46***	***318479,87***	***6463,829***	***0,981595***	***0,963471***
39	[8-39-1]	547	0,000003	305601,93	315071,75	317343,56	6473,622	0,982166	0,964485
40	**[8-40-1]**	**561**	**0,000003**	**277480,34**	**302315,09**	**314018,56**	**6412,039**	**0,985052**	**0,970276**
41	[8-41-1]	575	0,000003	289045,28	296211,06	313242,21	6477,932	0,983818	0,967819
42	[8-42-1]	589	0,000003	298999,31	328200,84	309262,34	6537,352	0,982786	0,965735
43	[8-43-1]	603	0,000003	310091,40	313554,40	326806,34	6599,156	0,981881	0,964001
44	[8-44-1]	617	0,000003	300392,06	322792,71	320005,78	6597,665	0,982801	0,965853
45	[8-45-1]	631	0,000003	302898,21	312741,34	322868,59	6633,375	0,982465	0,965178
46	[8-46-1]	645	0,000003	303921,84	313436,31	333074,62	6664,506	0,982196	0,964626
47	[8-47-1]	659	0,000003	297776,84	301755,68	314426,18	6673,551	0,982999	0,966194
48	[8-48-1]	673	0,000003	305820,53	312813,87	324272,90	6726,286	0,982846	0,965917
49	[8-49-1]	687	0,000003	288073,62	308014,15	322410,84	6698,808	0,983746	0,967581
50	[8-50-1]	701	0,000003	291901,06	309766,81	319473,5	6739,056	0,984143	0,968349

5.2.5 ANN Weekly Forecasting Calculations for Istanbul Anatolian Side

In ANN weekly forecasting models of Istanbul Anatolian side again same input criterias were taken like Istanbul criterias.

In weekly forecasting algorithms again exhaustive search method is used as weekly models, for incremental of neuron number. Fitness criteria for weekly model is inverse test error and for verification of 50 network architectures, all of the models are retrained ten times.Logistic activation function is selected for hidden layer activation function.

Below graph shows the best 5 network models after all 50 hidden layer units were applied to the model one by one.

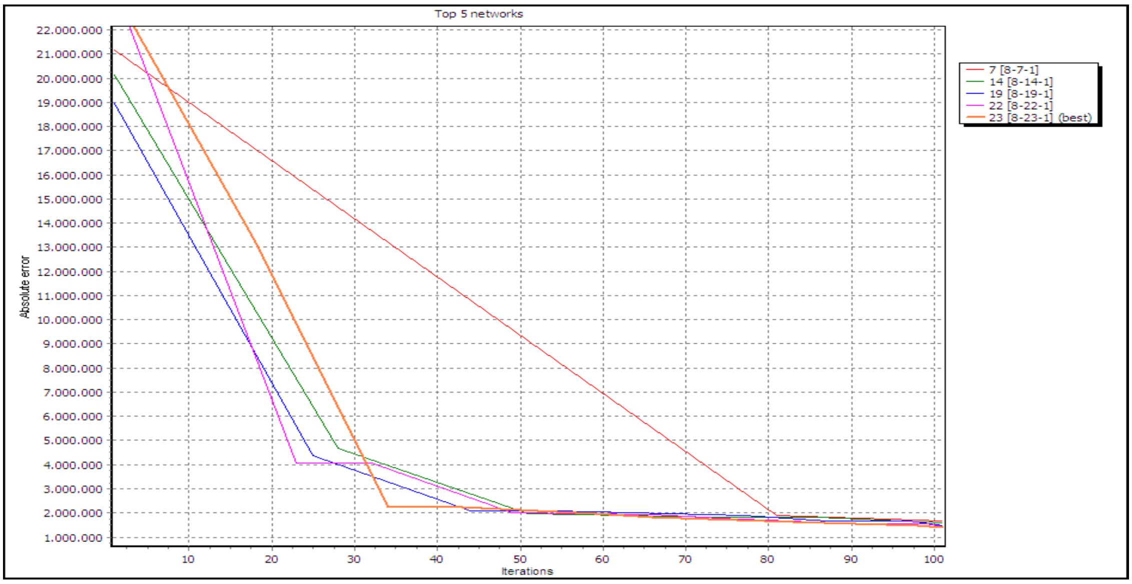


Figure 5.22. Graph of best 5 networks for Istanbul Anatolian side weekly values

[8-23-1] architecture had the best fitness

The graph Figure 5.21. shows that the model with 23 hidden neuron is best for Istanbul Anatolian side weekly gas consumption forecasting in neural network model.

All of the training, validation, test, correlation and R-Square values for all hidden neuron numbers are calculated for this model again. Most of the R-Square values which shows that input values could explain our model or not, are over 98 percent.

Table 5.21. Importance of input factors for Istanbul Anatolian side weekly forecasting

Input column name	Importance, %
• **Year**	59,258083
• **Week**	5,098838
• **Anatolian Side Average customer Number**	21,342361
• **Average Maximum Tempeature**	3,956578
• **Average Minimum Tempeature**	10,34414

Above Table 5.21. results shows that most important factors for weekly forecasting of natural gas of Istanbul Anatolian side are “year” and “average customer number” values with % 59,25 and % 21,34 importance.

Training, validating and test results are as shown below.

Table 5.22. Istanbul Anatolian side weekly training values results.

	Target	Output	AE	ARE
Mean:	24802635,77	24795808,34	1178299,85	0,070005
Std Dev:	19771992,75	19294099,88	961086,96	0,065894
Min:	4027000	5110769,29	8156,05	0,001031
Max:	72924632	73859702,24	3966918,42	0,342336
Correlation :	0,997269			
R-Squared :	0,993789			

Table 5.23. Istanbul Anatolian side weekly validating values results.

	Target	Output	AE	ARE
Mean:	23570928,74	23252957,77	2105542,26	0,120738
Std Dev:	20667501,18	19344556,28	2202615,64	0,120709
Min:	4088000	5216026,50	35704,13	0,004615
Max:	80315798	75595874,83	8314052,14	0,479378
Correlation :	0,990704			
R-Squared :	0,975188			

Table 5.24. Istanbul Anatolian side weekly testing values results.

	Target	Output	AE	ARE
Mean:	23798511,48	23153499,44	2570790,77	0,108469
Std Dev:	21591971,12	20308811,63	5439403,99	0,113561
Min:	4207000	5161742,51	20229,58	0,001955
Max:	66320108	66279696,48	29731259,18	0,448299
Correlation :	0,96108			
R-Squared :	0,912241			

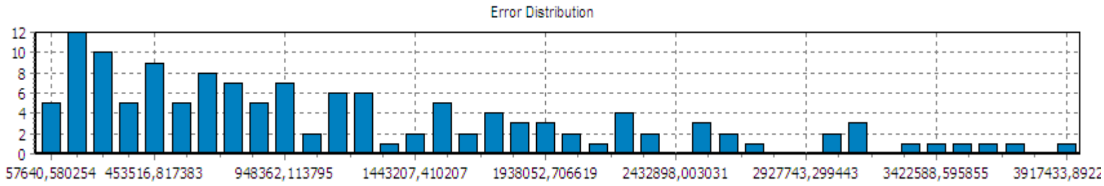


Figure 5.23. Histogram of error distribution for prediction of Anatolian side weekly values

In ANN forecasting applications for istanbul Anatolian side training ARE value is %7, validating ARE is %12 and testing ARE is % 10.

Same in the Istanbul Anatolian side daily forecasting values testing ARE value is bigger than other ones Istanbul weekly and Istanbul European side values. It is seen in error distribution graph in Figure 5.22. This shows that data collecting system and gas distribution system is not as good as European side or there are some undefined factors that effects this area.

Table 5.25. Results for Istanbul Anatolian weekly forecasting values

ID	Architecture	Weights	Fitness	Train Error	Validation Error	Test Error	AIC	Correlation	R-Squared
1	[8-1-1]	11	3,81E-07	2405251,25	2679473,25	2625190,5	1325,774503	0,983884	0,967627
2	[8-2-1]	21	4,25E-07	2213411	2450121,5	2351934,5	1334,719596	0,986954	0,973355
3	[8-3-1]	31	3,95E-07	1830864,25	2053877,5	2529778,5	1329,48329	0,990969	0,981891
4	[8-4-1]	41	4,00E-07	1930342	1985805,5	2502466,5	1356,5202	0,990179	0,980303
5	[8-5-1]	51	4,10E-07	1843336,75	2005250,375	2436081	1370,386266	0,990628	0,981
6	[8-6-1]	61	4,14E-07	1777716,75	2077019,125	2416885	1385,56534	0,992053	0,983728
7	**[8-7-1]**	**71**	**4,45E-07**	**1683889,5**	**1967043,625**	**2245744,75**	**1398,353603**	**0,99195**	**0,983877**
8	[8-8-1]	81	4,17E-07	1683103,5	1928455,25	2397461,75	1418,291513	0,992254	0,984385
9	[8-9-1]	91	4,16E-07	1694551	2220279,5	2403042,75	1439,193037	0,992194	0,984218
10	[8-10-1]	101	4,31E-07	1534295,75	2073672,75	2317985,25	1445,979981	0,993306	0,986574
11	[8-11-1]	111	4,33E-07	1633256,75	1999707	2308416	1474,293082	0,99281	0,985411
12	[8-12-1]	121	4,32E-07	1656528,75	1961577,25	2313610	1496,1748	0,991868	0,983788
13	[8-13-1]	131	4,33E-07	1521806,75	1983816,75	2311599,25	1504,892938	0,993107	0,986145
14	**[8-14-1]**	**141**	**4,40E-07**	**1580884,375**	**1906594,25**	**2272420,5**	**1529,958398**	**0,992938**	**0,985798**
15	[8-15-1]	151	4,24E-07	1592915,25	1966385,625	2359508,75	1550,966723	0,992754	0,985524
16	[8-16-1]	161	4,30E-07	1453442,375	1921581,25	2325532,25	1558,779794	0,993727	0,987423
17	[8-17-1]	171	4,21E-07	1643443,75	1886469,375	2372548,75	1595,12006	0,992404	0,984527

Table 5.25. Results for Istanbul Anatolian weekly forecasting values

ID	Architecture	Weights	Fitness	Train Error	Validation Error	Test Error	AIC	Correlation	R-Squared
18	[8-18-1]	181	4,18E-07	1766034,75	1865769,875	2389634,25	1624,68846	0,991037	0,982096
19	**[8-19-1]**	**191**	**4,52E-07**	**1578482,75**	**1965793,125**	**2213013,5**	**1629,756198**	**0,993016**	**0,986066**
20	[8-20-1]	201	4,15E-07	1679648,125	1964811,125	2406880,5	1658,018189	0,991937	0,98386
21	[8-21-1]	211	4,28E-07	1445197,375	2014424,375	2337323,75	1658,023176	0,993834	0,987647
22	**[8-22-1]**	**221**	**4,54E-07**	**1555420,75**	**1863212,875**	**2202094,5**	**1687,798698**	**0,993054**	**0,986025**
23	***[8-23-1]***	***231***	***4,58E-07***	***1453532,5***	***1780343,75***	***2184143,75***	***1698,788042***	***0,993963***	***0,98761***
24	[8-24-1]	241	4,17E-07	1589203,5	1792378,375	2399620,25	1730,656453	0,99293	0,9858
25	[8-25-1]	251	4,28E-07	1525293,75	1867927,5	2334094	1745,197348	0,993361	0,986651
26	[8-26-1]	261	4,10E-07	1587021,75	1853826,75	2436522,25	1770,47374	0,992797	0,985572
27	[8-27-1]	271	4,12E-07	1637094,375	1945951,5	2427751,25	1794,60522	0,992618	0,985238
28	[8-28-1]	281	4,36E-07	1555788,25	2007108,625	2295447,25	1807,830121	0,993096	0,9862
29	[8-29-1]	291	4,35E-07	1559140,375	1797998	2298317,75	1828,116377	0,993459	0,986786
30	[8-30-1]	301	4,15E-07	1540985,875	1798440,375	2408520,25	1846,558648	0,993308	0,986594
31	[8-31-1]	311	4,00E-07	1712723,75	2032975,625	2502963,5	1880,611767	0,991795	0,983532
32	[8-32-1]	321	4,32E-07	1430893,75	2029904,125	2314609,5	1876,700266	0,994319	0,98846
33	[8-33-1]	331	4,23E-07	1666853	1967724,125	2362623,25	1917,001145	0,991867	0,983659

Table 5.25. Results for Istanbul Anatolian weekly forecasting values (continued)

ID	Architecture	Weights	Fitness	Train Error	Validation Error	Test Error	AIC	Correlation	R-Squared
34	[8-34-1]	341	4,18E-07	1447693,75	2068786,25	2393067	1918,252716	0,993773	0,987568
35	[8-35-1]	351	4,26E-07	1641994,5	1820207,375	2348680,75	1955,002721	0,992492	0,984963
36	[8-36-1]	361	4,26E-07	1656937,125	2028944,5	2349943,5	1976,207581	0,992692	0,985367
37	[8-37-1]	371	4,32E-07	1496636,625	1770402,875	2315450,5	1982,674776	0,994014	0,987927
38	[8-38-1]	381	4,17E-07	1602313,75	1798290,5	2398624,5	2011,749146	0,992642	0,985317
39	[8-39-1]	391	4,16E-07	1581067,75	1932304,875	2405054,5	2029,973824	0,992721	0,985469
40	[8-40-1]	401	4,40E-07	1461483,625	2011774,625	2274835,5	2039,513598	0,993923	0,987649
41	[8-41-1]	411	4,29E-07	1465037,5	1986917,25	2328723,25	2059,836621	0,993887	0,987792
42	[8-42-1]	421	4,20E-07	1566648	1840418	2382784,5	2088,755268	0,993013	0,986024
43	[8-43-1]	431	4,03E-07	1655006,5	1950314,625	2481337	2116,052529	0,992055	0,984126
44	[8-44-1]	441	4,08E-07	1759203,25	1927988,25	2452983,75	2144,172983	0,991528	0,983089
45	[8-45-1]	451	4,08E-07	1887597	2053839,875	2451474,25	2173,541981	0,990459	0,980962
46	[8-46-1]	461	4,04E-07	1751973,5	1966802,75	2475323,25	2183,625267	0,991114	0,982227
47	[8-47-1]	471	4,13E-07	1518280,375	1934136,75	2422147,5	2184,584397	0,993269	0,986581
48	[8-48-1]	481	4,10E-07	1762241,5	1852937,875	2436271,25	2224,402488	0,991553	0,983077
49	[8-49-1]	491	4,13E-07	1462971,5	1990121,125	2421529	2219,648933	0,993647	0,987285
50	[8-50-1]	501	4,18E-07	1539706,375	1978696	2390174,25	2246,448175	0,993214	0,986471

Table 5.25. Results for Istanbul Anatolian weekly forecasting values (continued)

Retrain	**Iters**	**Train error**	**Validation error**	**Test error**	**AIC**	**Correlation**	**R-Squared**
1	101	1530660,253341	1985350,15	2492409,55	1705,66	0,993258	0,986507
2	101	1735977,520556	1997344,60	2461159,14	1722,40	0,991526	0,982987
3	101	1909370,583444	2014212,85	2556897,76	1735,06	0,990333	0,980537
4	101	1589968,18835	2092008,91	2567091,68	1710,72	0,993062	0,985979
5	101	1453532,470358	1780343,76	2184143,77	1698,78	0,993963	0,98761
6	101	1719173,346162	2069825,89	2629224,89	1721,11	0,991571	0,983195
7	101	1660139,547	1874218,57	2413921,14	1716,46	0,992078	0,984129
8	101	1627968,797657	2117653,46	2556713,67	1713,86	0,992777	0,985491
9	101	1884099,919015	2116518,74	2582588,63	1733,29	0,990358	0,980752
10	101	1735563,710191	1947913,62	2511068,08	1722,37	0,991683	0,983393

SUMMARY:

Min:	1453532,47	1780343,76	2184143,7	1698,788	0,990333	0,980537
Max:	1909370,58	2117653,46	2629224,89	1735,067	0,993963	0,98761
Average:	1684645,43	1999539,06	2495521,83	1717,975	0,992061	0,984058
Std. dev.:	15162424,2	17996151,5	22460015,1	15461,78	8,928546	8,856522

5.2.6 ANN Weekly Forecasting Calculations for Istanbul European Side

In ANN weekly forecasting models of Istanbul European side again same input criterias were taken like Istanbul criterias. Network models and training algorithms are same with other weekly forecasting models.

Below graph shows the best 5 network models after all 50 hidden layer units were applied to the model one by one.

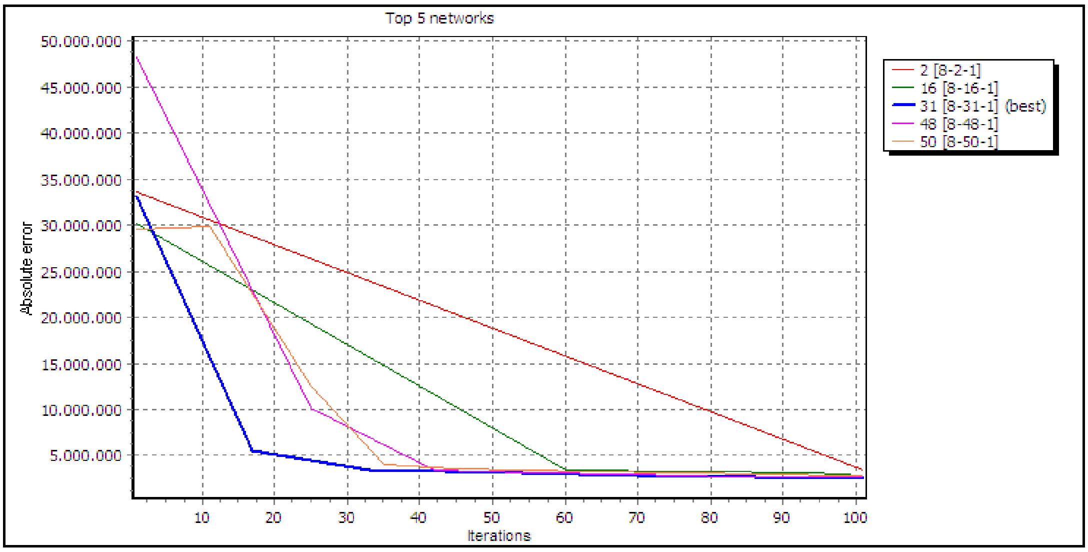


Figure 5.24. Graph of best 5 networks for Istanbul European side weekly values

[8-31-1] architecture had the best fitness

The graph Figure 5.23. shows that the model with 31 hidden neuron is best for Istanbul European side weekly gas consumption forecasting in neural network model.

Table 5.26. Importance of input factors for Istanbul European side weekly forecasting

Input column name	Importance, %
• **Year**	17,088709
• **Week**	3,758538
• **Anatolian Side Average customer Number**	23,602528
• **Average Max. Temperature**	30,626735
• **Average Min.Temperature**	24,92349

Above Table 5.26. results shows that most important factors for weekly forecasting of natural gas of Istanbul European side are “average maximum temperature” "average customer number" and "average minimum temperature" values with % 30,62 , % 23,60 and %24,92 importance.

Results of training, validating and testing are as shown below.

Table 5.27. Istanbul European side weekly training values results.

	Target	Output	AE	ARE
Mean:	43814797,49	43726749,80	1622972,90	0,042152
Std Dev:	30025753,28	29722076,59	2174909,38	0,044514
Min:	13493208	14217286,57	16411,43	0,000624
Max:	121825321	121143657,5	15675566,94	0,31187
Correlation :	0,99593			
R-Squared :	0,991664			

Table 5.28. Istanbul European side weekly validating values results.

	Target	Output	AE	ARE
Mean:	49240642,93	48521013,51	2719846,49	0,060789
Std Dev:	33215996,09	32965679,04	3567427,82	0,07272
Min:	14232641	14562685,46	30829,08	0,000299
Max:	132721461	124630833,53	14346133,29	0,261027
Correlation :	0,991076			
R-Squared :	0,981482			

Table 5.29. Istanbul European side weekly validating values results.

	Target	Output	AE	ARE
Mean:	43016934,74	43010535,42	3549519,93	0,073548
Std Dev:	29417325,95	30360331,44	4309016,07	0,06106
Min:	13470417	14285492,14	142190,31	0,001771
Max:	96028836	111836822,52	16093643,53	0,306823
Correlation :	0,98305			
R-Squared :	0,966182			

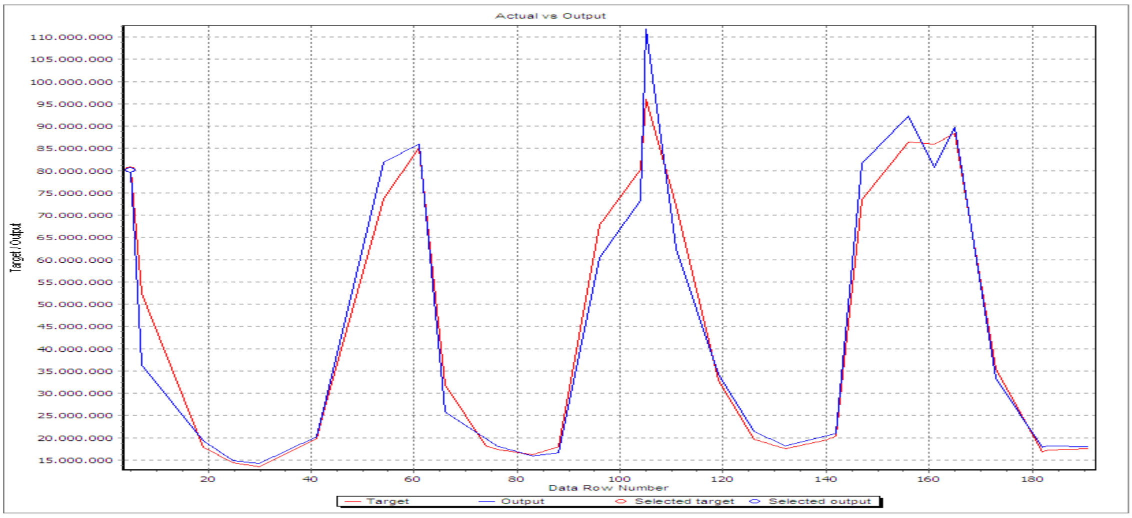


Figure 5.25. Graph of testing target-output values for Istanbul European side weekly forecasting

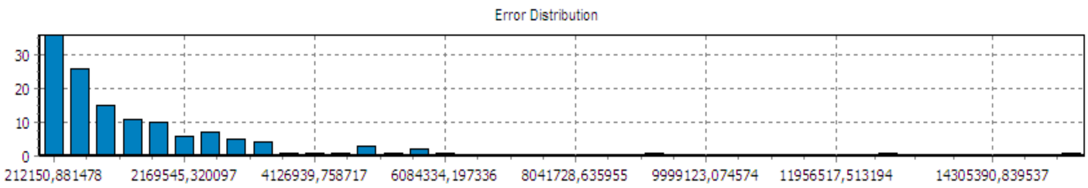


Figure 5.26. Histogram of error distribution for prediction of European side weekly values

The results of training, validating and test values are as shown above. Performance of the model for Istanbul European side weekly forecasting values are better than Anatolian side values. The ARE values for training' validating and test values are about %4, %6 and %7 for the European side weekly forecasting values which were %7, %12 and %10 for Anatolian side weekly forecasting values. There may be some reason of this like the values of European side is more correct than other or gas consumption situation of European side is more stable than Anatolian side.

As seen in Figure 5.24 the performance of the model is very good. Testing ARE value is less than %7 with the standart deviation %6.

Table 5.30. Table 5.30. Results for Istanbul European weekly forecasting values

ID	Architecture	Weights	Fitness	Train Error	Validation Error	Test Error	AIC	Correlation	R-Squared
1	[8-1-1]	11	2,62E-07	4220820,5	4283659	3812268	1400,570405	0,97932	0,958754
2	**[8-2-1]**	**21**	**3,49E-07**	**3473452,75**	**4205905,5**	**2862545**	**1394,651299**	**0,983846**	**0,967422**
3	[8-3-1]	31	3,21E-07	3399692,5	4565117,5	3113918,25	1411,796577	0,985234	0,970405
4	[8-4-1]	41	3,24E-07	3325443,75	4274027,5	3087328,5	1428,85969	0,986327	0,972595
5	[8-5-1]	51	3,03E-07	3273263,5	3988803,75	3303050,25	1446,756211	0,986378	0,972745
6	[8-6-1]	61	3,28E-07	2921537,25	4106619	3045051	1451,637095	0,989133	0,978135
7	[8-7-1]	71	3,12E-07	2960830,5	3894417,75	3202217,25	1473,413958	0,988998	0,978039
8	[8-8-1]	81	3,12E-07	3327507	4069447,75	3208453,25	1508,942182	0,985125	0,97025
9	[8-9-1]	91	2,92E-07	2902866	4118770,75	3428446,25	1510,78437	0,98889	0,977767
10	[8-10-1]	101	3,18E-07	3337355,5	3849908,5	3143126,25	1549,335239	0,986206	0,972374
11	[8-11-1]	111	3,06E-07	2787416,25	3668360,25	3271421,25	1545,386773	0,989781	0,979602
12	[8-12-1]	121	3,27E-07	3257476	4127629,75	3059122,5	1586,113177	0,986389	0,972646
13	[8-13-1]	131	3,22E-07	2993183,75	3601129	3102793,5	1594,859375	0,988329	0,976664
14	[8-14-1]	141	3,10E-07	3159448,5	3959877	3228542,5	1622,049331	0,98745	0,974808
15	[8-15-1]	151	3,14E-07	2827380,25	3724863	3185963	1627,280098	0,990024	0,979483
16	**[8-16-1]**	**161**	**3,44E-07**	**3259430,5**	**3747070,25**	**2903541**	**1666,192953**	**0,986212**	**0,972359**
17	[8-17-1]	171	3,01E-07	2963629	3463664	317950,25	1673,539607	0,989151	0,978404

Table 5.30. Results for Istanbul European weekly forecasting values

ID	Architecture	Weights	Fitness	Train Error	Validation Error	Test Error	AIC	Correlation	R-Squared
18	[8-18-1]	181	3,05E-07	3016034,25	3785577,75	3278994	1695,870864	0,988615	0,976917
19	[8-19-1]	191	3,19E-07	2925613,25	3873334	3137168	1711,822517	0,988903	0,977772
20	[8-20-1]	201	3,18E-07	2835184,75	3655095,5	3146864,75	1727,64671	0,989738	0,979438
21	[8-21-1]	211	3,11E-07	3179496,75	3740642,5	3217137	1762,890618	0,987281	0,974133
22	[8-22-1]	221	3,17E-07	3193387,75	3810561,75	3155294	1783,470419	0,986953	0,973969
23	[8-23-1]	231	3,30E-07	2829956,25	3680522	3034837,25	1787,401217	0,989288	0,978401
24	[8-24-1]	241	3,06E-07	3041977,75	3683644,75	3265766	1817,010014	0,987937	0,975953
25	[8-25-1]	251	3,05E-07	3183579	3768684,5	3273494	1843,061275	0,987133	0,974294
26	[8-26-1]	261	3,22E-07	3510521,25	3784400	3101350,75	1876,063157	0,98425	0,96865
27	[8-27-1]	271	3,19E-07	3102450,5	3746987,25	3131549,75	1879,628039	0,987561	0,975024
28	[8-28-1]	281	3,12E-07	3103455,5	3730483,25	3210011,75	1899,671116	0,987238	0,974585
29	[8-29-1]	291	3,26E-07	2669614,75	3422183	3065939,5	1899,643714	0,990263	0,980616
30	[8-30-1]	301	3,17E-07	3245621	3895825,25	3149866,75	1945,628258	0,985951	0,971944
31	***[8-31-1]***	***311***	***3,58E-07***	***3228948,5***	***3763263,75***	***2793469,75***	***1964,943289***	***0,986503***	***0,972973***
32	[8-32-1]	321	3,16E-07	3230512,5	3892907,25	3161040,75	1985,007692	0,986398	0,972897
33	[8-33-1]	331	3,09E-07	3040797,25	3610284	3234963,75	1996,958393	0,988148	0,976373

Table 5.30. Results for Istanbul European weekly forecasting values(continued)

ID	Architecture	Weights	Fitness	Train Error	Validation Error	Test Error	AIC	Correlation	R-Squared
34	[8-34-1]	341	3,26E-07	3248417,25	3754231,5	3063396,75	2025,742793	0,985999	0,972136
35	[8-35-1]	351	3,12E-07	3322413,75	3632760	3204204,25	2048,738451	0,985972	0,972077
36	[8-36-1]	361	3,13E-07	2964065,25	3486518,5	3190203,75	2053,559181	0,988615	0,977291
37	[8-37-1]	371	3,08E-07	3283827,5	3670094	3248135,5	2087,184755	0,987333	0,97476
38	[8-38-1]	381	3,28E-07	2901260,25	3505865,5	3048369,25	2090,710783	0,988902	0,977876
39	[8-39-1]	391	3,14E-07	2843186,75	3643149,75	3183004,5	2108,021562	0,989243	0,97851
40	[8-40-1]	401	3,15E-07	2969821	3691094,5	3174056,75	2133,81719	0,988751	0,977569
41	[8-41-1]	411	3,18E-07	3172108,75	3836059	3143257,5	2162,581221	0,987229	0,974604
42	[8-42-1]	421	3,22E-07	3104395,25	3532940,25	3108970,75	2179,711386	0,987507	0,975159
43	[8-43-1]	431	3,23E-07	2679773,75	3461256,75	3097921	2180,148878	0,990343	0,980627
44	[8-44-1]	441	3,12E-07	2798539	3397479,75	3202017,25	2205,916434	0,989777	0,979597
45	[8-45-1]	451	3,09E-07	3315657,25	3765162,75	3239809,75	2248,467698	0,986534	0,973132
46	[8-46-1]	461	3,25E-07	2866624,75	3185508,25	3074071,25	2249,113463	0,98908	0,978263
47	[8-47-1]	471	3,25E-07	2649055,25	3183190,75	3076565,25	2258,615478	0,991437	0,982476
48	**[8-48-1]**	**481**	**3,35E-07**	**3088405,75**	**3507733,25**	**2981435,25**	**2299,024586**	**0,988048**	**0,976221**
49	[8-49-1]	491	3,11E-07	2944710,5	3690768,25	3211137,25	2312,687868	0,988695	0,977499
50	**[8-50-1]**	**501**	**3,31E-07**	**3035684,25**	**3121413**	**3021908,5**	**2336,734569**	**0,988205**	**0,976534**

Table 5.30. Results for Istanbul European weekly forecasting values(continued)

Retrain	**Iters**	**Train error**	**Validation error**	**Test error**	**AIC**	**Correlation**	**R-Squared**
1	101	3007599,49	3727667,92	3253325,66	1955,49	0,987822	0,975767
2	101	3228948,46	3763263,82	2793469,70	1964,94	0,986503	0,972973
3	101	2983521,76	3705293,23	3114520,86	1954,42	0,988644	0,97722
4	101	3017130,37	3783497,91	3709961,17	1955,91	0,988735	0,977388
5	101	2931071,08	3914667,27	3487048,48	1952,07	0,988695	0,977449
6	101	2865678,25	3526814,98	3362546,00	1949,06	0,989451	0,978942
7	101	3003683,40	3671810,55	3343024,79	1955,32	0,988635	0,977345
8	101	2540313,33	3586524,05	3267061,12	1933,04	0,991463	0,982931
9	101	2972041,98	3886745,38	3852909,60	1953,91	0,989391	0,978812
10	101	2761124,87	3410135,14	3641729,98	1944,12	0,990187	0,980066
SUMMARY:							
Min:		2540313,33	3410135,14	2793469,70	1933,04	0,986503	0,972973
Max:		3228948,46	3914667,27	3852909,60	1964,94	0,991463	0,982931
Average:		2931111,30	3697642,02	3382559,74	1951,83	0,988953	0,977889
Std. dev.:		26380565,68	33279109,33	30444445,81	17566,50	8,900574	8,801003

5.2.7. ANN Monthly Forecasting Calculations for Istanbul

Monthly values were composed from the daily values. Temperature values are designed as an average of days' temperature values. As in the daily and weekly values average maximum and minimum temperature values were taken. Monthly average number of consumers were considered for input of consumer number.

In ANN monthly forecasting models there are 20 input values and one output value.

Below graph shows the best 5 network models after all 50 hidden layer units were applied to the model one by one.

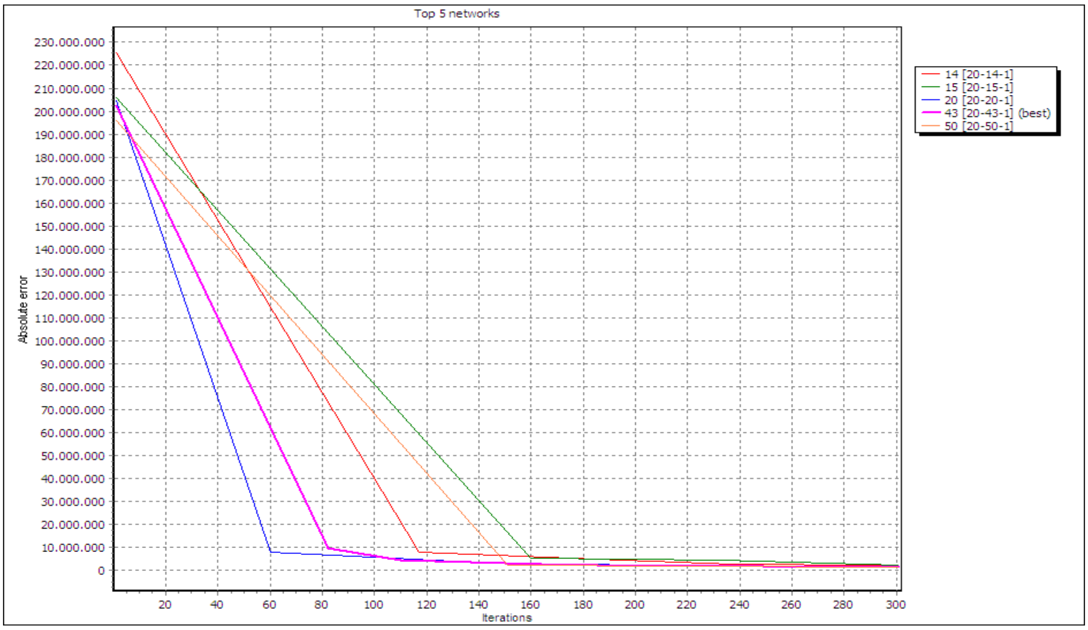


Figure 5.27. Graph of best 5 networks for Istanbul monthly values

[20-43-1] architecture had the best fitness

In monthly forecasting algorithms exhaustive search method was used as daily and weekly models, for incremental of neuron number. Fitness criteria for monthly model is inverse test error and for verification of 50 network architectures, all of the models are retrained ten times. Logistic activation function is selected for the hidden layer.

The graph Figure 5.26. shows that the model with 43 hidden neuron is best for Istanbul monthly gas consumption forecasting in neural network model.

All of the training, validation, test, correlation and R-Square values for all hidden neuron numbers are calculated for this model again. Most of the R-Square values which shows that input values could explain our model or not, are nearly 99 percent.

Table 5.31. Importance of input factors for Istanbul monthly forecasting

Input column name	Importance, %
• **Month**	83,206025
• **Max. Customer Number**	0,003156
• **Min. Customer Number**	0,026246
• **Average Customer Number**	0,224271
• **Year**	7,035271
• **Average Temp.**	9,505031

Above Table 5.4. results shows that most important factor for monthly forecasting of natural gas of Istanbul is “month” with %83,2 importance. The other factors as “year” and “average temperature” values are also important but their effects are not as much as month factor.

5.2.7.1 ANN Calcuations of Istanbul Monthly Gas Prediction Training Values

After best model for monthly values were defined as 20 input, 43 hidden layer neuron number and 1 output number, this model is applied on all seven training algorithm that mentioned before. Best algorithm for forecasting of daily values was found as Conjugate Gardient Descent Algorithm. After network is trained 200 times the target and output values of training data sets is as below graph.

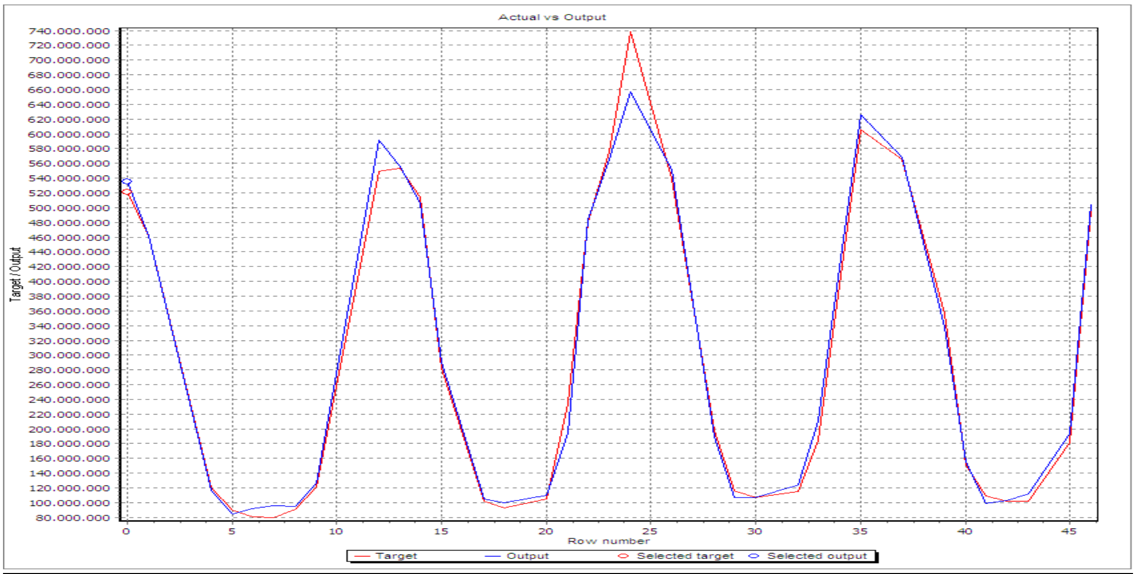


Figure 5.28. Graph of training target-output values for Istanbul monthly forecasting

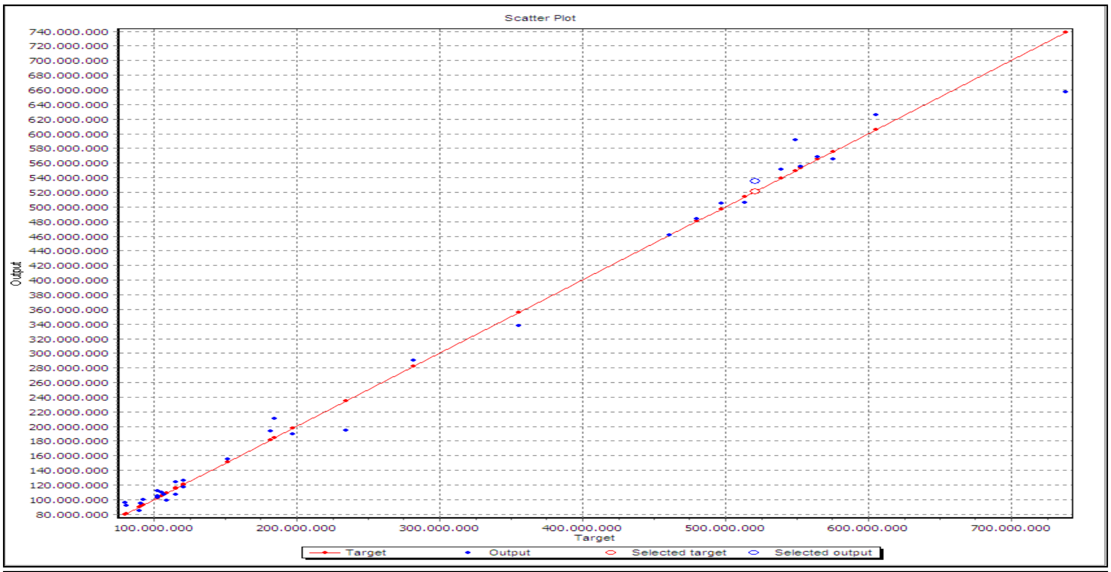


Figure 5.29. Istanbul mothly values training set target- output scatter plot graph

Table 5.32. Istanbul monthly training values results.

	Target	Output	AE	ARE
Mean:	285924665,62	286815067,66	12276786,78	0,054
Std Dev:	207049086,69	204266619,83	15455076,28	0,049
Min:	79763902	84595693,83	116467,7	0,0011
Max:	738599422	656639834,24	81959587,75	0,198
Correlation :	0,995495			
R-Squared :	0,990663			

As it can be seen above results that for training algorithms 'correlation' and 'R-square' values are all very well with the results over 99%. ARE value is about %5 with the standard deviation value of 0,049. The ARE values are changing between % 0, 01 and % 19 with the standard deviation of % 4, 9.

Below graphs and values in the tables shows the validation and test results of Istanbul weekly forecasting values.

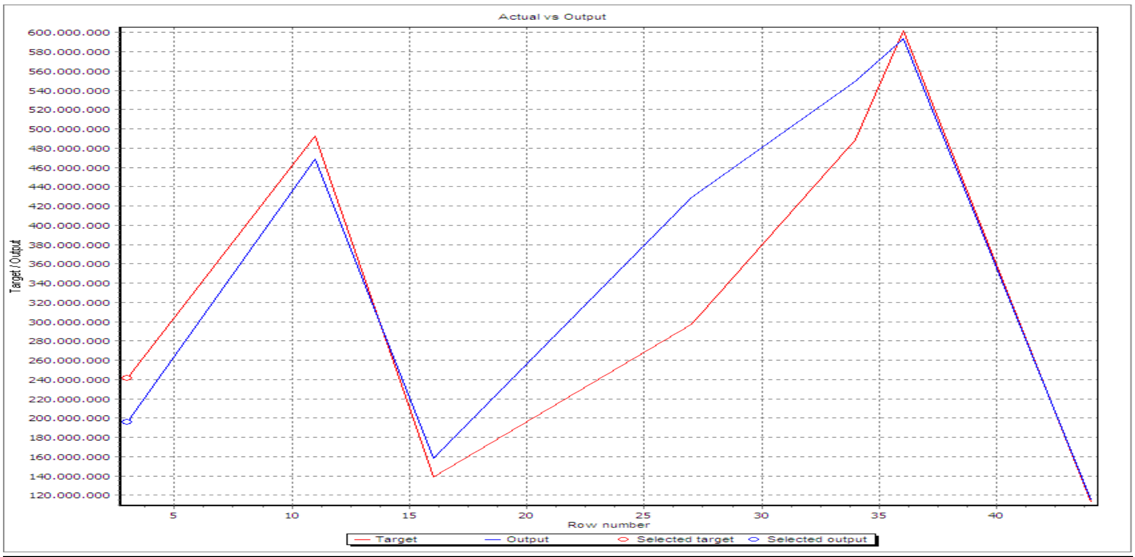


Figure 5.30. Graph of validation target-output values for Istanbul monthly forecasting

Table 5.33. Istanbul monthly validation values results.

	Target	**Output**	**AE**	**ARE**
Mean:	339092341,42	358602425,2	41643442,13	0,14
Std Dev:	176213820,97	182880948,96	41141519,91	0,13
Min:	112774512	115338789,43	2564277,43	0,013
Max:	601698304	593632511,53	131277741,3	0,44

Correlation : 0,953427

R-Squared : 0,897541

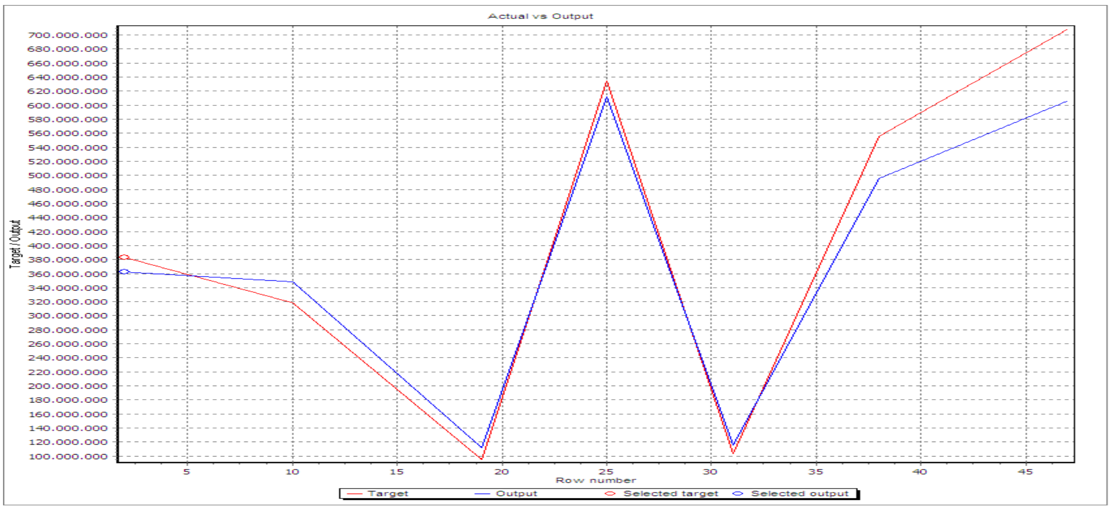


Figure 5.31. Graph of testing target-output values for Istanbul monthly forecasting

Table 5.34. Istanbul monthly testing values results.

	Target	Output	AE	ARE
Mean:	399639870	378871992,83	37792087,77	0,104
Std Dev:	227466135,87	193291364,87	29984374,34	0,044
Min:	95046390	111501393,67	12591639	0,03
Max:	708464543	612018658,47	102161696,81	0,17
Correlation :	0,99172			
R-Squared :	0,937709			

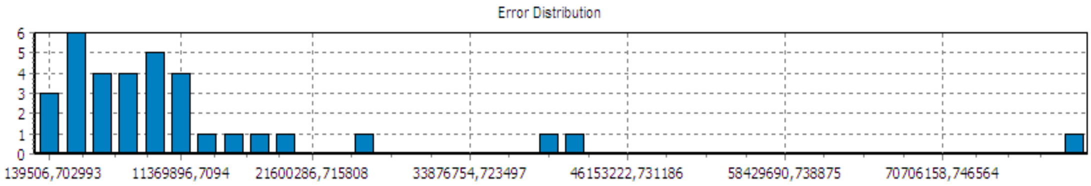


Figure 5.32. Histogram of error distribution for prediction of istanbul monthly values

Again in monthly ANN models the test values are randomly selected from input values. Resultant graph of target and output values are as shown in Figure 5.30. and Table 5.33. the performance of network is very well with the result of less than %10 absolute relative error value.

Table 5.35. Results for Istanbul monthly forecasting values

ID	Architecture	# of Weights	Fitness	Train Error	Validation Error	Test Error	AIC	Correlation	R-Squared
1	[20-1-1]	23	1,51E-08	33897840	47423360	66206584	515,62	0,979	0,949
2	[20-2-1]	45	3,44E-08	9109963	44624120	29052188	514,9	0,998	0,996
3	[20-3-1]	67	2,80E-08	13154480	62115536	35661344	571,4	0,994	0,988
4	[20-4-1]	89	2,79E-08	6935083,5	55854060	35812692	593,67	0,998	0,996
5	[20-5-1]	111	2,83E-08	12158392	44274768	35349768	656,76	0,995	0,989
6	[20-6-1]	133	2,44E-08	17018954	37300920	40993620	712,19	0,99	0,982
7	[20-7-1]	155	3,14E-08	12971503	56092088	31804786	746,96	0,995	0,990
8	[20-8-1]	177	3,32E-08	15841642	38963664	30102340	797,76	0,993	0,987
9	[20-9-1]	199	2,90E-08	7722692	52756512	34424388	817,33	0,99	0,996
10	[20-10-1]	221	2,47E-08	15959472	44389232	40511452	886,01	0,995	0,991
11	[20-11-1]	243	2,23E-08	20918458	54959904	44859552	939,21	0,991	0,98
12	[20-12-1]	265	3,56E-08	8026825	46145952	28104942	950,64	0,997	0,994
13	[20-13-1]	287	3,43E-08	13363842	44080720	29129788	1011,97	0,996	0,992
14	**[20-14-1]**	**309**	**4,40E-08**	**7977104**	**45040344**	**22725278**	**1038,43**	**0,997**	**0,994**
15	**[20-15-1]**	**331**	**4,32E-08**	**4262665**	**55853716**	**23154726**	**1061,12**	**0,999**	**0,998**
16	[20-16-1]	353	3,60E-08	8040379,5	47660048	27804130	1126,70	0,998	0,996
17	[20-17-1]	375	3,18E-08	9895977	45996752	31464422	1177,76	0,997	0,993
18	[20-18-1]	397	3,22E-08	9464190	49827768	31062910	1220,24	0,997	0,995

Table 5.35. Results for Istanbul monthly forecasting values (continued)

ID	Architecture	Weights	Fitness	Train Error	Validation Error	Test Error	AIC	Correlation	R-Squared
19	[20-19-1]	419	3,47E-08	11125753	50061316	28825302	1269,74	0,996	0,992
20	**[20-20-1]**	**441**	**4,43E-08**	**7730749**	**41180124**	**22556294**	**1301,36**	**0,997**	**0,99**
21	[20-21-1]	463	3,39E-08	11242126	47099656	29537086	1358,09	0,996	0,992
22	[20-22-1]	485	3,91E-08	13066360	35454580	25583360	1407,21	0,995	0,99
23	[20-23-1]	507	2,78E-08	12504625	48778480	35977132	1449,71	0,996	0,991
24	[20-24-1]	529	3,35E-08	11094441	41088632	29871706	1489,65	0,996	0,992
25	[20-25-1]	551	3,20E-08	10665609	46332156	31248026	1532,30	0,996	0,991
26	[20-26-1]	573	3,32E-08	9309078	47921700	30095020	1571,68	0,997	0,99
27	[20-27-1]	595	3,31E-08	7455834	46434648	30195204	1608,13	0,998	0,996
28	[20-28-1]	617	3,72E-08	11645642	47875524	26903186	1667,29	0,996	0,992
29	[20-29-1]	639	3,35E-08	4476428,5	45825096	29873926	1678,79	0,999	0,998
30	[20-30-1]	661	4,08E-08	10048379	39907852	24489548	1750,28	0,997	0,994
31	[20-31-1]	683	4,11E-08	7578468,5	49011448	24358778	1784,69	0,998	0,996
32	[20-32-1]	705	4,09E-08	16917068	49935568	24467818	1855,99	0,993	0,985
33	[20-33-1]	727	3,87E-08	13335134	37874504	25815222	1891,90	0,995	0,989
34	[20-34-1]	749	3,28E-08	6085614	42755712	30485806	1909,23	0,999	0,998
35	[20-35-1]	771	3,77E-08	11782652	41569888	26532828	1975,69	0,995	0,99
36	[20-36-1]	793	3,66E-08	9513234	42892580	27304668	2012,42	0,996	0,993

Table 5.35. Results for Istanbul monthly forecasting values(continued)

ID	Architecture	Weights	Fitness	Train Error	Validation Error	Test Error	AIC	Correlation	R-Squared
37	[20-37-1]	815	3,72E-08	1502508	44627104	26880302	1993,67	0,999	0,999
38	[20-38-1]	837	3,90E-08	7277338	48504160	25643282	2091,31	0,998	0,996
39	[20-39-1]	859	3,66E-08	8353672	38845756	27309666	2140	0,998	0,996
40	[20-40-1]	881	4,00E-08	6026919	46204260	25012844	2172,9	0,999	0,998
41	[20-41-1]	903	3,59E-08	8800250	48023844	27892004	2229,77	0,997	0,994
42	[20-42-1]	925	3,40E-08	7958479,5	48678072	29438326	2270,35	0,997	0,995
43	***[20-43-1]***	***947***	***4,69E-08***	***6593302,5***	***43458048***	***21343008***	***2307,95***	***0,999***	***0,997***
44	[20-44-1]	969	4,04E-08	9585282	50105340	24753560	2364,67	0,997	0,994
45	[20-45-1]	991	4,15E-08	5712240	37761356	24098242	2391,07	0,999	0,998
46	[20-46-1]	1013	3,94E-08	9376387	42966944	25362586	2451,92	0,996	0,993
47	[20-47-1]	1035	3,76E-08	3762508,25	44974524	26594900	2464,88	0,999	0,998
48	[20-48-1]	1057	3,47E-08	8254781,5	53313608	28800050	2535,59	0,997	0,995
49	[20-49-1]	1079	3,63E-08	11652819	49038392	27555218	2591,31	0,996	0,992
50	**[20-50-1]**	**1101**	**4,31E-08**	**7815004,5**	**45435844**	**23181820**	**2621,73**	**0,997**	**0,995**

Table 5.35. Results for Istanbul monthly forecasting values(continued)

Retrain	Iters	Train error	Validation error	Test error	AIC	Correlation	R-Squared
1	301	25309307,7	55388060,13	61654258,55	813,69	0,987578	0,970973
2	301	25309307,7	55388060,13	61654258,55	813,69	0,987578	0,970973
3	301	27050518,33	46573280,95	46736941,02	815,95	0,979757	0,959217
4	301	8655110,15	44119538,01	46268436,99	777,20	0,998726	0,996857
5	301	8655110,15	44119538,01	46268436,99	777,20	0,998726	0,996857
6	301	15841642,47	38963663,03	30102340,6	797,76	0,993807	0,987467
7	301	15841642,47	38963663,03	30102340,6	797,76	0,993807	0,987467
8	301	26022488,81	50245309,76	57173055,83	814,63	0,978289	0,951207
9	301	5336215,04	49894279,83	33400181,57	760,76	0,998852	0,997532
10	301	5336215,04	49894279,83	33400181,57	760,76	0,998852	0,997532
SUMMARY:							
Min:		5336215,04	38963663,03	30102340,6	760,76	0,97	0,951
Max:		27050518,33	55388060,13	61654258,55	815,95	0,998	0,997532
Average:		16335755,79	47354967,27	44676043,23	792,94	0,991	0,981608
Std. dev.:		147269903,2	426231387,43	402261065,6	7136,5	8,92	8,83

5.2.8. ANN Monthly Forecasting Calculations for Istanbul European Side

Same criterias and network models are applied for monthly forecasting calculations of Istanbul European side.Below graph shows the best 5 network models after all 50 hidden layer units were applied to the model one by one.

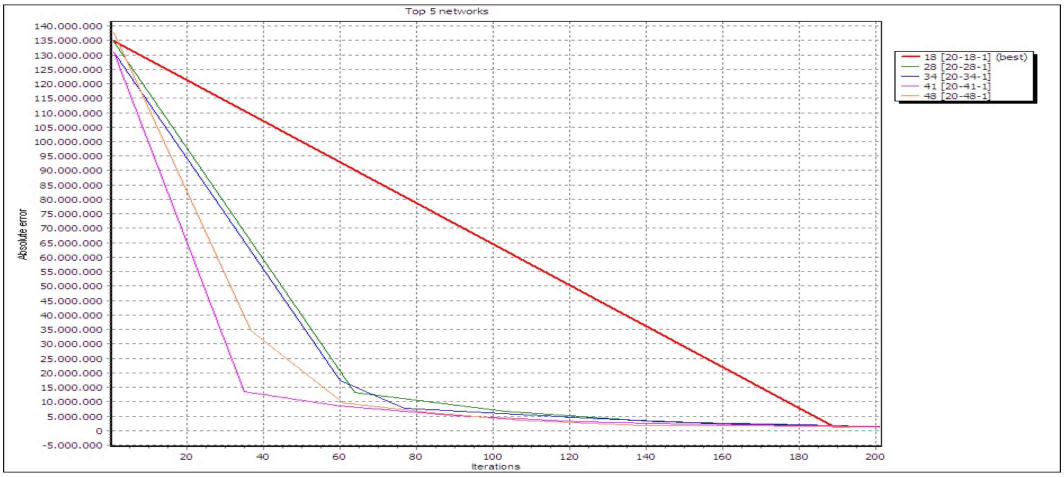


Figure 5.33. Graph of best 5 networks for Istanbul European side monthly values

[20-18-1] architecture had the best fitness

In ANN forecasting model of European side monthly values, for output parameters sum of square is used as error function and logistic function is used as an activation function. The graph Figure 5.33. shows that the model with 18 hidden neuron is best for Istanbul monthly gas consumption forecasting in neural network model.

All of the training, validation, test, correlation and R-Square values for all hidden neuron numbers are calculated for this model again. Correlation and R-Square values are over 99 % as seen in the Table 5.39.

Table 5.36. Importance of input factors for Istanbul European side monthly forecasting

Input column name	Importance, %
• **Month**	75,206025
• **Max. Customer Number**	0,003156
• **Min. Customer Number**	0,026246
• **Average Customer Number**	0,224271
• **Year**	7,035271
• **Average Temp.**	18,505031

As in the Istanbul monthly forecasting values “month” is the most important factor with 75 %. Average temperature factor is more effective in here than istanbul monthly forecastig with value of 18 %. Below graphs and tables shows the calculation values of ANN training, validating and test values for Istanbul European side forecasting.

Table 5.3. Istanbul European side monthly training values results.

	Target	Output	AE	ARE
Mean:	206178832,03	205190206,76	3308668,28	0,024
Std Dev:	132052808,2	129849154,3	2178409,84	0,020
Min:	61378902	64297591,87	92064,11	0,000553
Max:	437838111	427786654,88	10051456,11	0,086
Correlation :	0,999713			
R-Squared :	0,999069			

Table 5.38. Istanbul European side monthly validating values results.

	Target	Output	AE	ARE
Mean:	152337484,14	145316788,82	8927665,82	0,058
Std Dev:	71923339,26	69276490,49	10537874,75	0,06
Min:	78414102	78830998,73	416896,73	0,005
Max:	284247218	279701490,26	32218463,18	0,2
Correlation :	0,986508			
R-Squared :	0,960245			

Table 5.39. Istanbul European side monthly testing values results.

	Target	Output	AE	ARE
Mean:	215994820,28	202954613,07	15236483,8	0,059
Std Dev:	129651088,19	116002633,54	20951798,62	0,047
Min:	69645175	67782066,04	1863108,95	0,01
Max:	460466327	395395631,86	65070695,13	0,14
Correlation :	0,989534			
R-Squared :	0,950126			

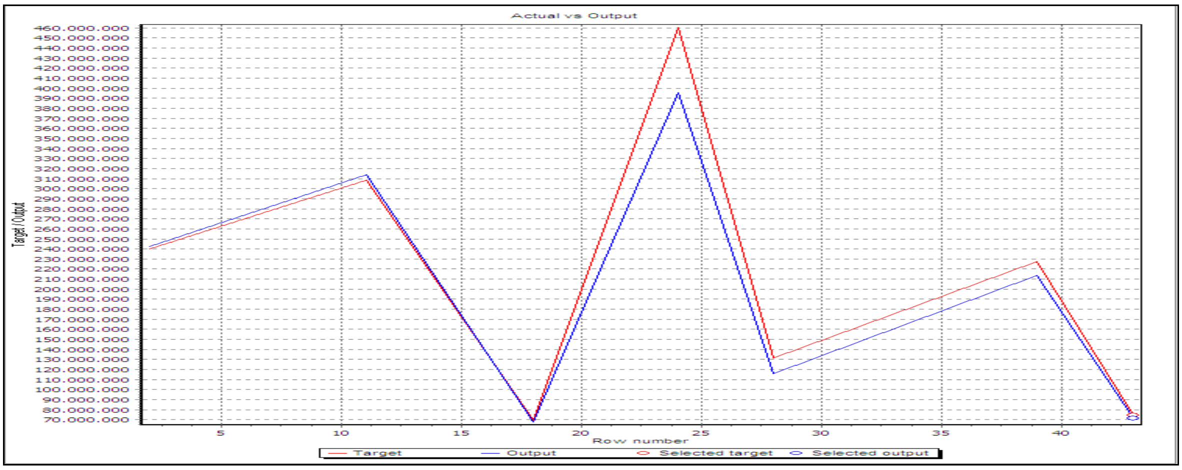


Figure 5.34. Graph of testing target-output values for Istanbul european side monthly forecasting

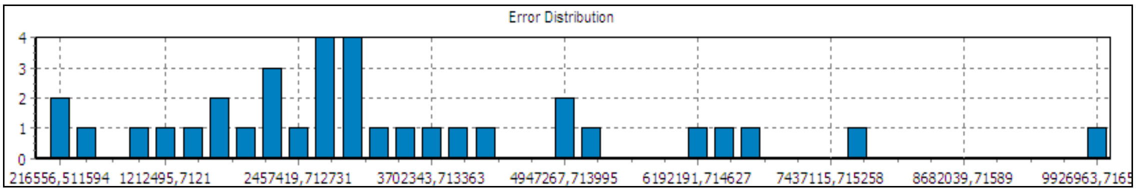


Figure 5.35. Histogram of error distribution for prediction of Istanbul European side monthly values

As it can be seen above results that for training algorithms ‘correlation’ and ‘R-square’ values are all very well with the results over 97%. Training ARE value is about %2 with the standard deviation value of 0.02. Validation and test ARE values are 5.8 % and 5.9 with the standard deviations 6 % and % 4.

Table 5.40. Results for Istanbul European side monthly forecasting values

ID	Architecture	# of Weights	Fitness	Train Error	Validation Error	Test Error	AIC	Correlation	R-Squared
1	[20-1-1]	23	2,93E-08	11862226	27558192	34126400	479,92	0,99	0,98
2	[20-2-1]	45	3,20E-08	7045456,5	19709954	31215952	506,21	0,99	0,99
3	[20-3-1]	67	3,84E-08	18394046	17795010	26068100	582,84	0,98	0,96
4	[20-4-1]	89	3,77E-08	11314789	12120837	26532452	610,31	0,99	0,98
5	[20-5-1]	111	3,82E-08	5433453,5	12782591	26194202	629,37	0,99	0,99
6	[20-6-1]	133	3,95E-08	1900823,125	17867326	25286018	637,66	0,99	0,99
7	[20-7-1]	155	5,73E-08	3303026,25	14616106	17440214	700,45	0,99	0,99
8	[20-8-1]	177	3,89E-08	3312575	12693268	25699476	744,55	0,99	0,99
9	[20-9-1]	199	5,67E-08	3317496	11675061	17639660	788,6	0,99	0,99
10	[20-10-1]	221	4,00E-08	2972132,5	9251687	24979300	828,86	0,99	0,99
11	[20-11-1]	243	4,45E-08	3353883	16421877	22478436	876,97	0,99	0,99
12	[20-12-1]	265	4,96E-08	6671680,5	16746125	20158766	944,35	0,99	0,99
13	[20-13-1]	287	4,36E-08	11576519	16364173	22931688	1007,09	0,99	0,98
14	[20-14-1]	309	5,77E-08	1438816,875	17320980	17332776	980,20	0,99	0,99
15	[20-15-1]	331	4,36E-08	13946712	17004604	22936342	1101,42	0,99	0,98
16	[20-16-1]	353	5,26E-08	1561273	14435158	19009490	1070,97	0,99	0,99
17	[20-17-1]	375	4,82E-08	9814305	18697818	20727454	1177,48	0,99	0,98

Table 5.40. Results for Istanbul European side monthly forecasting values (continued)

ID	Architecture	# of Weights	Fitness	Train Error	Validation Error	Test Error	AIC	Correlation	R-Squared
18	***[20-18-1]***	***397***	***7,54E-08***	***1680012***	***11284543***	***13265854***	***1161,47***	***0,99***	***0,99***
19	[20-19-1]	419	5,87E-08	10939924	12776837	17042574	1269,17	0,99	0,98
20	[20-20-1]	441	5,10E-08	9346530	10566891	19601570	1307,82	0,99	0,99
21	[20-21-1]	463	4,77E-08	11248286	15123941	20980100	1358,11	0,99	0,987
22	[20-22-1]	485	6,12E-08	1540227,125	13552854	16333946	1334,51	0,99	0,99
23	[20-23-1]	507	4,96E-08	20893656	14832748	20169560	1467,17	0,99	0,96
24	[20-24-1]	529	5,85E-08	1319714,875	12668440	17085078	1417,26	0,99	0,99
25	[20-25-1]	551	5,42E-08	12895957	15211523	18457590	1538,76	0,99	0,98
26	[20-26-1]	573	6,10E-08	2308119	12202838	16386561	1524,26	0,99	0,99
27	[20-27-1]	595	5,13E-08	678993,25	8294326	19479766	1526,66	0,99	0,99
28	**[20-28-1]**	**617**	**7,07E-08**	**1311055,375**	**11960085**	**14146780**	**1593,03**	**0,99**	**0,99**
29	[20-29-1]	639	5,31E-08	9857857	9970091	18820918	1705,63	0,99	0,98
30	[20-30-1]	661	5,53E-08	11425807	11785465	18083598	1754,65	0,99	0,98
31	[20-31-1]	683	6,08E-08	12748888	12957619	16457019	1802,37	0,99	0,98
32	[20-32-1]	705	5,49E-08	1891235,75	13785542	18230020	1781,49	0,99	0,99
33	[20-33-1]	727	4,88E-08	9934045	12461888	20477334	1881,89	0,99	0,98
34	**[20-34-1]**	**749**	**7,43E-08**	**1397982**	**9231191**	**13460015**	**1859,22**	**0,99**	**0,99**

Table 5.40. Results for Istanbul European side monthly forecasting values (continued)

ID	Architecture	# of Weights	Fitness	Train Error	Validation Error	Test Error	AIC	Correlation	R-Squared
35	[20-35-1]	771	5,8E-08	42790,5	13041265	17299574	1919,29	0,99	0,99
36	[20-36-1]	793	5,5E-08	18083824	17899284	19030708	2034,26	0,99	0,97
37	[20-37-1]	815	6,7E-08	51513,5	11717234	16197870	1981,53	0,99	0,99
38	[20-38-1]	837	6,14E-08	1510609	19124260	16274843	2106,90	0,993	0,98
39	[20-39-1]	859	5,33E-08	1741437	18007236	18752428	2151,57	0,99	0,98
40	[20-40-1]	881	6,23E-08	441524,125	14256906	16050856	2124,26	0,99	0,99
41	**[20-41-1]**	**903**	**6,71E-08**	**547352,375**	**11762896**	**14892665**	**2170,67**	**0,99**	**0,99**
42	[20-42-1]	925	5,97E-08	327579,875	11106115	16759273	2209,46	0,99	0,99
43	[20-43-1]	947	5,54E-08	10966776	8889982	18058026	2325,25	0,99	0,98
44	[20-44-1]	969	5,37E-08	1376907	11341722	18638160	2370,50	0,99	0,98
45	[20-45-1]	991	5,92E-08	3924317	10778124	16890292	2421,37	0,99 8	0,98
46	[20-46-1]	1013	5,15E-08	2023259	12701335	19435854	2460,38	0,99	0,98
47	[20-47-1]	1035	6,26E-08	4404110	13391544	15962249	2510,52	0,99	0,97
48	**[20-48-1]**	**1057**	**6,29E-08**	**883737**	**10715967**	**15905642**	**2485,36**	**0,99**	**0,99**
49	[20-49-1]	1079	6,25E-08	1363140	14902166	16002820	2590,46	0,99	0,98
50	[20-50-1]	1101	5,91E-08	2091870	14419984	16911754	2636,57	0,99	0,98

Table 5.40. Results for Istanbul European side monthly forecasting values (continued)

Retrain	**Iters**	**Train error**	**Validation error**	**Test error**	**AIC**	**Correlation**	**R-Squared**
1	201	1680012,06	11284542,57	13265853,67	1161,47	0,999822	0,999606
2	201	1680012,06	11284542,57	13265853,67	1161,47	0,999822	0,999606
3	201	1680012,06	11284542,57	13265853,67	1161,47	0,999822	0,999606
4	201	1680012,06	11284542,57	13265853,67	1161,47	0,999822	0,999606
5	201	3773336,82	14036396,45	23982016,50	1188,98	0,999256	0,998415
6	201	3773336,82	14036396,45	23982016,50	1188,98	0,999256	0,998415
7	201	3773336,82	14036396,45	23982016,50	1188,98	0,999256	0,998415
8	201	1472032,57	13319102,31	27187182,90	1156,97	0,999848	0,999689
9	201	1472032,57	13319102,31	27187182,90	1156,97	0,999848	0,999689
10	201	1607931,29	17277913,26	24213028,24	1159,97	0,999844	0,999636
SUMMARY:							
Min:		1472032,57	11284542,57	13265853,67	1156,97	0,999256	0,998415
Max:		3773336,82	17277913,26	27187182,9	1188,98	0,999848	0,999689
Average:		2259205,51	13116347,75	20359685,82	1168,67	0,999659	0,999268
Std. dev.:		20357138,38	118061259,91	183332306,21	10518,09	8,996934	8,993415

5.2.9. ANN Monthly Forecasting Calculations for Istanbul Anatolian Side

Below graph shows the best 5 network models after all 50 hidden layer units were applied to the model one by one.

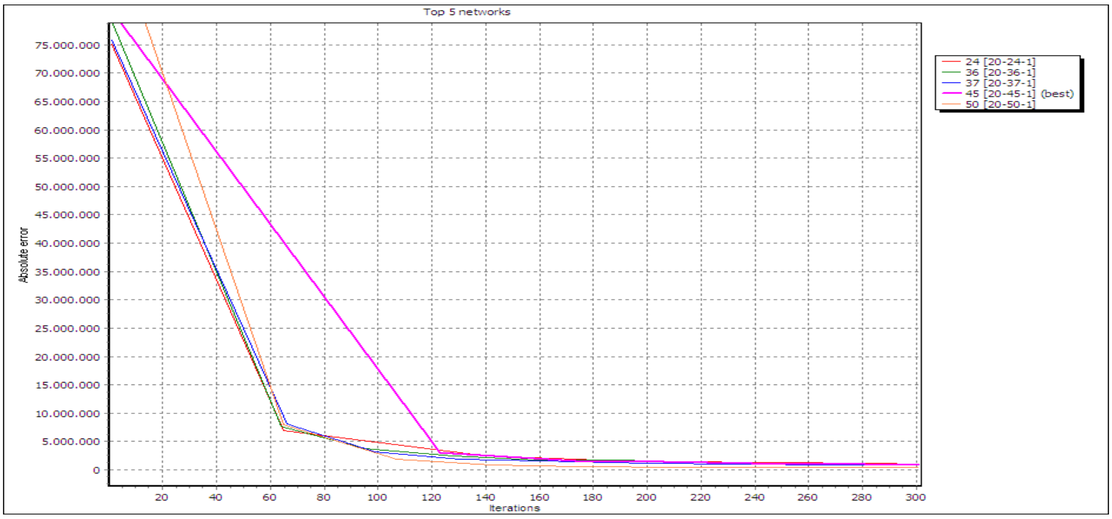


Figure 5.36. Graph of best 5 networks for Istanbul Anatolian side monthly values

20-45-1] architecture had the best fitness

Table 5.41. Importance of input factors for Istanbul monthly forecasting

Input column name	Importance, %
• **Month**	84,709
• **Max. Customer Number**	0,004156
• **Min. Customer Number**	0,0737
• **Average Customer Number**	0,368
• **Year**	7,035271
• **Average Temp.**	6,086

As in the before results “month” is the most important factor with 84 %. The other factors as “year” and “average temperature” values are also important but their effects are not as much as month factor.

Below graphs and values in the tables shows the training,validation and test results of IstanbulAnatolian side weekly forecasting values.

Table 5.42. Istanbul Anatolian side monthly training values results.

	Target	Output	AE	ARE
Mean:	111608290,2	112549417,76	5747597,004	0,074
Std Dev:	83344508,01	82297555,81	5224257,65	0,073
Min:	18385000	23613469,98	497521,04	0,0039
Max:	278133095	251610276,19	26522818,80	0,326
Correlation :	0,995747			
R-Squared :	0,991093			

Table 5.43. Istanbul Anatolian side monthly validating values results.

	Target	Output	AE	ARE
Mean:	109143962,28	105185286,25	8920802,08	0,117
Std Dev:	88963743,69	83740667,62	8720203,93	0,098
Min:	23267296	27154510,009	57952,86	0,00028
Max:	270626432	248962683,75	22638330,01	0,282
Correlation :	0,992438			
R-Squared :	0,977808			

Table 5.44. Istanbul Anatolian side monthly testing values results.

	Target	Output	AE	ARE
Mean:	106495441,14	107570686,71	10278865,9	0,11
Std Dev:	80874963,99	81613387,05	8026883,25	0,09
Min:	22003000	21005124,86	997875,13	0,035
Max:	230313399	238501092,35	25113385,38	0,33
Correlation :	0,984275			
R-Squared :	0,974464			

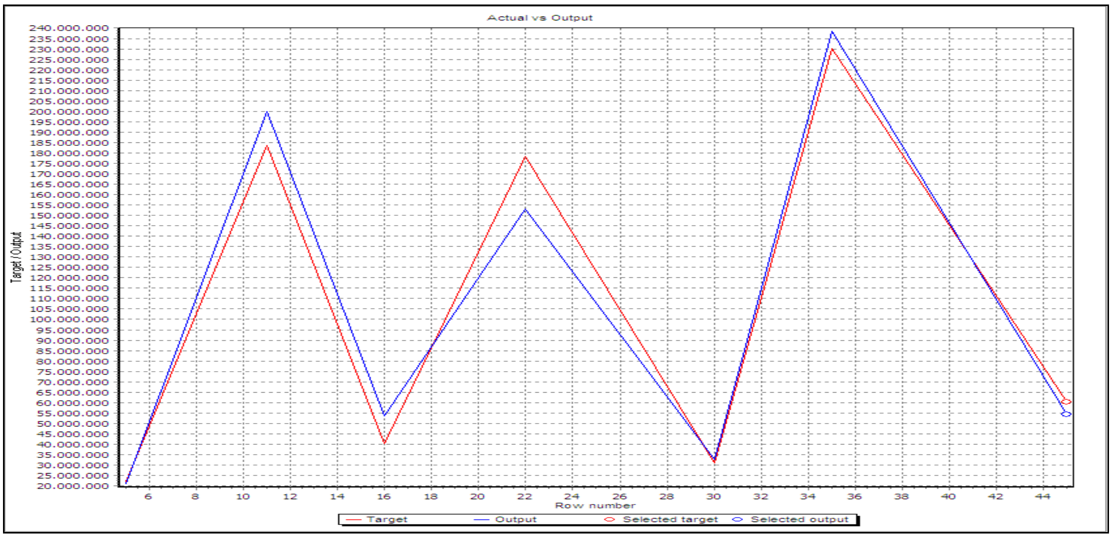


Figure 5.37. Graph of testing target-output values for Istanbul Anatolian side monthly forecasting

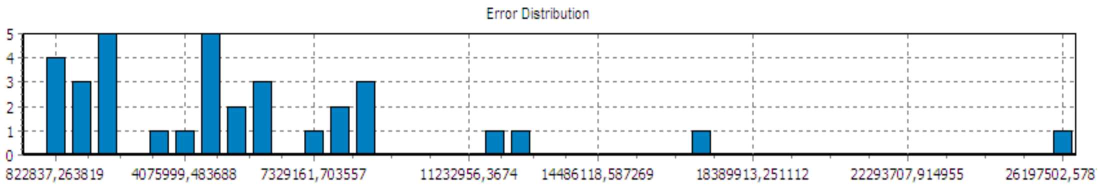


Figure 5.38. Histogram of error distribution for prediction of Istanbul Anatolian side monthly values

As it can be seen above results that for training algorithms ‘correlation’ and ‘R-square’ values are all very well with the results over 99%. Training ARE value is about %7 with the standard deviation value of 0.07. Validation and test ARE values are 11% and 11% with the standard deviations 9, 8 % and % 9.

Table 5.45. Results for Istanbul Anatolian side monthly forecasting values

ID	Architecture	# of Weights	Fitness	Train Error	Validation Error	Test Error	AIC	Correlation	R-Squared
1	[20-1-1]	23	7,74E-08	11474223	11798658	12925991	478,794	0,984	0,966
2	[20-2-1]	45	9,85E-08	4630079	9353930	10154271	491,938	0,996	0,992
3	[20-3-1]	67	1,19E-07	2764302,75	9498620	8436050	518,401	0,998	0,996
4	[20-4-1]	89	1,18E-07	4342712	9628588	8474926	577,760	0,996	0,992
5	[20-5-1]	111	1,77E-07	2539538	10195636	5663627,5	603,518	0,998	0,997
6	[20-6-1]	133	9,15E-08	6968811	9786045	10930119	681,840	0,990	0,980
7	[20-7-1]	155	9,15E-08	4074927	12017464	10923726	707,596	0,998	0,995
8	[20-8-1]	177	1,82E-07	3245756,25	12701221	5498177,5	743,860	0,997	0,995
9	[20-9-1]	199	1,29E-07	2422427,25	10061964	7757306	777,913	0,998	0,997
10	[20-10-1]	221	1,74E-07	1314786,625	7859910	5744041,5	801,136	0,999	0,999
11	[20-11-1]	243	1,00E-07	6245151	9700319	9986154	898,112	0,995	0,99
12	[20-12-1]	265	1,50E-07	4357177	8398469	6669950	929,873	0,996	0,992
13	[20-13-1]	287	1,04E-07	3698070	8180580,5	9637721	968,296	0,997	0,994
14	[20-14-1]	309	1,47E-07	6236591	11732447	6782764,5	1030,065	0,992	0,984
15	[20-15-1]	331	1,69E-07	3227592	14222185	5923072,5	1051,670	0,998	0,997
16	[20-16-1]	353	1,08E-07	3894691,5	8225221,5	9231647	1102,057	0,997	0,994
17	[20-17-1]	375	1,54E-07	824411,4375	8431994	6508402,5	1093,266	0,999	0,999

Table 5.45. Results for Istanbul Anatolian side monthly forecasting values (continued)

ID	Architecture	# of Weights	Fitness	Train Error	Validation Error	Test Error	AIC	Correlation	R-Squared
18	[20-18-1]	397	1,22E-07	898508,1875	10859022	8200031	1140,192	0,999	0,999
19	[20-19-1]	419	1,61E-07	4475925	11104321	6217195,5	1238,787	0,996	0,993
20	[20-20-1]	441	1,34E-07	4877363,5	9652818	7484134	1285,707	0,996	0,992
21	[20-21-1]	463	1,23E-07	6322081,5	10810925	8159919	1338,528	0,995	0,989
22	[20-22-1]	485	1,60E-07	625581,25	9413674	6231630,5	1303,882	0,999	0,999
23	[20-23-1]	507	1,41E-07	3182058,5	8063452,5	7113981	1403,187	0,998	0,996
24	**[20-24-1]**	**529**	**2,40E-07**	**3436322,25**	**11143317**	**4161589,25**	**1449,800**	**0,999**	**0,997**
25	[20-25-1]	551	1,29E-07	963916,4375	13162748	7755083,5	1450,581	0,999	0,999
26	[20-26-1]	573	1,26E-07	5681224,5	13564289	7948800	1554,894	0,995	0,990
27	[20-27-1]	595	2,29E-07	1258250,25	10062168	4375356	1547,641	0,999	0,999
28	[20-28-1]	617	1,68E-07	777826,9375	9765135	5954788	1575,288	0,999	0,999
29	[20-29-1]	639	1,31E-07	1086318,375	13563766	7660208	1630,646	0,999	0,999
30	[20-30-1]	661	1,27E-07	4063295,5	10164461	7901608	1719,498	0,998	0,995
31	[20-31-1]	683	1,49E-07	6173563	11118937	6716139	1777,720	0,995	0,990
32	[20-32-1]	705	1,17E-07	2663485,5	9285416	8558553	1793,138	0,998	0,997
33	[20-33-1]	727	1,72E-07	2597273,25	10575572	5814243,5	1836,282	0,998	0,997
34	[20-34-1]	749	1,70E-07	801272,9375	11782943	5891188,5	1840,298	0,999	0,999

Table 5.45. Results for Istanbul Anatolian side monthly forecasting values (continued)

ID	Architecture	# of Weights	Fitness	Train Error	Validation Error	Test Error	AIC	Correlation	R-Squared
35	[20-35-1]	771	1,84E-07	1045467,3125	10142502	5441798	1893,342	0,999	0,999
36	**[20-36-1]**	**793**	**2,35E-07**	**1985281,25**	**10687700**	**4247286**	**1959,146**	**0,999**	**0,998**
37	**[20-37-1]**	**815**	**2,57E-07**	**3657950,75**	**11617853**	**3886512**	**2023,925**	**0,998**	**0,995**
38	[20-38-1]	837	1,47E-07	2256794,75	12397588	6810133,5	2051,505	0,999	0,998
39	[20-39-1]	859	2,20E-07	3859761,5	9937978	4545094	2113,751	0,999	0,996
40	[20-40-1]	881	1,49E-07	5658478,5	9442234	6733255,5	2170,758	0,994	0,989
41	[20-41-1]	903	1,23E-07	5881624	10533713	8098464,5	2216,073	0,995	0,989
42	[20-42-1]	925	1,32E-07	4739462,5	10631496	7579812	2252,732	0,997	0,994
43	[20-43-1]	947	1,49E-07	1540553,125	11610785	6718160,5	2258,523	0,999	0,998
44	[20-44-1]	969	1,52E-07	1041851,8125	10741203	6588017	2289,225	0,999	0,999
45	***[20-45-1]***	***991***	***3,00E-07***	***2220849,5***	***12899063***	***3338323***	***2358,959***	***0,999***	***0,998***
46	[20-46-1]	1013	1,32E-07	2463062,75	11125340	7595486	2406,478	0,998	0,997
47	[20-47-1]	1035	1,70E-07	1398154,5	9473796	5868096,5	2431,226	0,999	0,999
48	[20-48-1]	1057	1,61E-07	1487282,125	10021407	6200410	2477,327	0,999	0,999
49	[20-49-1]	1079	2,17E-07	1929100,875	11470417	4599409	2530,170	0,999	0,998
50	**[20-50-1]**	**1101**	**2,30E-07**	**1776194,25**	**10249319**	**4343159,5**	**2571,363**	**0,999**	**0,998**

Table 5.45. Results for Istanbul Anatolian side monthly forecasting values (continued)

Retrain	Iters	Train error	Validation error	Test error	AIC	Correlation	R-Squared
1	301	2220849,58	12899063,011	3338322,948	2358,959	0,999	0,998
2	301	2873116,12	8623638,198	14169310,7	2367,714	0,999	0,998
3	301	6913423,28	9222681,017	11466807,8	2397,568	0,995	0,988
4	301	5171342,021	11373104,97	11205069,5	2387,697	0,995	0,991
5	301	5771002,31	10864600,501	13360988,9	2391,427	0,995	0,989
6	301	4996466,27	9407794,362	10318537,2	2386,527	0,997	0,993
7	301	1432160,52	10378567,832	6556088,4	2344,043	0,999	0,998
8	301	1585177,36	9773432,305	8050909,2	2347,494	0,999	0,998
9	301	3435419,24	10426950,928	5787824,4	2373,791	0,998	0,995
10	301	6682867,56	10976179,811	11354766,3	2396,415	0,993	0,985
SUMMARY:							
Min:		1432160,528	8623638,198	3338322,94	2344,043	0,993	0,985
Max:		6913423,285	12899063,011	14169310,7	2397,568	0,999	0,998
Average:		4108182,431	10394601,29	9560862,5	2375,164195	0,997	0,993
Std. dev.:		37025330,95	93558711,93	86111698,7	21376,486	0,974	0,944

5.3. PERFORMANCE COMPARISON OF ANN ALGORITHMS

Performance measurment of neural network algorithms is done according to their speed, completion time of training an algorithm, and test error rates. Testing is a process of estimating quality of the trained neural network. During this process a part of data that wasn't used during training is presented to the trained network case by case. Than forecasting error measured on each case and used as the estimation of network quality. Another criteria of algorithms is speed of training calculation. This changes according to the size of the problem, number of hidden layer neuron number etc.

In this study daily, weekly and monthly ANN forecasting calculations have done. After best network types are determined for all sections than selected network was applied to the seven different ANN algorithms. Fitness criteria for all models is inverse test error and all of the models are retrained 200 times. Logistic activation function is selected for the hidden layer activation function.

As all of the calculatios have shown in Chapter 5 best networks for Istanbul as shown below:

- For daily values [19-48-1] architecture had the best fitness
- For weekly values [8-38-1] architecture had the best fitness
- For monthly values [20-43-1] architecture had the best fitness

Seven different training algorithm have used for comparison. Results tables of calculations for daily, weekly and monthly values are shown below.

Table 5.46. Training, validation and test results of seven different ANN algorithms for daily forecasting

ALGORITHMS		ARE	Min ARE	Max ARE	Correlation	R-square
Quick propagation Algorithm	Training	0.058815	0.000245	0.480476	0.987356	0.989755
	Validation	0.072859	0.000087	0.398727	0.992263	0.984086
	Test	0.065697	0.000286	0.418408	0.991531	0.982588
Conjugate Gradient Descent Algorithm	Training	0.069222	0.000138	0.457778	0.994956	0.989797
	Validation	0.085055	0.000051	0.683459	0.991352	0.982654
	Test	0.078744	0.000002	0.423064	0.991026	0.982062
Quasi-Newton Algorithm	Training	0.095535	0.000147	0.765933	0.991906	0.983267
	Validation	0.100975	0.00055	0.821897	0.99079	0.981309
	Test	0.100529	0.000172	0.69036	0.98992	0.97019
Limited Memory Quasi-Newton Algorithm	Training	0.075286	0.000134	0.504101	0.994361	0.988547
	Validation	0.087333	0.000095	0.588029	0.991441	0.982806
	Test	0.082077	0.000548	0.395854	0.991134	0.982149
Levenberg-Marquardt Algorithm	Training	0.062721	0.000144	0.446264	0.996513	0.992906
	Validation	0.083382	0.000073	0.701768	0.989492	0.979056
	Test	0.077662	0.00006	0.367624	0.987739	0.945486
Incremental Back-Propagation Algorithm	Training	0.079285	0.000029	0.611337	0.993965	0.987548
	Validation	0.090609	0.000733	0.833599	0.99082	0.981512
	Test	0.081653	0.000031	0.411291	0.99122	0.982306
Batch Back-Propagation Algorithm	Training	0.325927	0.000422	1.432.005	0.972096	0.860176
	Validation	0.33178	0.004648	1.316.801	0.971556	0.860265
	Test	0.320801	0.001391	1.066.775	0.974851	0.882492

Table5.47. Training, validation and test results of seven different ANN algorithms for weekly forecasting

ALGORITHMS		ARE	Min ARE	Max ARE	Correlation	R-square
Quick propagation Algorithm	Training	0.064378	0.000154	0.454275	0.992776	0.985205
	Validation	0.0616	0.001381	0.220537	0.995441	0.990696
	Test	0.105283	0.002791	0.351786	0.983713	0.953886
Conjugate Gradient Descent Algorithm	Training	0.049523	0.000468	0.295358	0.989363	0.969738
	Validation	0.056933	0.000443	0.193444	0.993696	0.986305
	Test	0.082408	0.004108	0.243723	0.993936	0.987683
Quasi-Newton Algorithm	Training	0.088423	0.000187	0.473316	0.988796	0.977373
	Validation	0.06649	0.000218	0.234757	0.994082	0.987731
	Test	0.125167	0.011762	0.396514	0.983901	0.957557
Limited Memory Quasi-Newton Algorithm	Training	0.064378	0.000154	0.454275	0.992776	0.985205
	Validation	0.0616	0.001381	0.220537	0.995441	0.990696
	Test	0.105283	0.002791	0.351786	0.983713	0.953886
Levenberg-Marquardt Algorithm	Training	0.078658	0.000382	0.331134	0.991471	0.982646
	Validation	0.064262	0.006626	0.176007	0.99519	0.989933
	Test	0.104724	0.000429	0.42153	0.98858	0.967528
Incremental Back-Propagation Algorithm	Training	0.098084	0.000365	0.45942	0.987254	0.974046
	Validation	0.065442	0.000292	0.36862	0.995403	0.989777
	Test	0.133843	0.001621	0.5463	0.982402	0.955441
Batch Back-Propagation Algorithm	Training	0.449655	0.002911	138.776	0.943254	0.649298
	Validation	0.459412	0.003906	1.169.045	0.947751	0.64838
	Test	0.486246	0.005067	1.287.332	0.934365	0.539691

Table 5.48. Training, validation and test results of seven different ANN algorithms for monthly forecasting

ALGORITHMS		ARE	Min ARE	Max ARE	Correlation	R-Square
Quick propagation Algorithm	Training	0.038711	0.00289	0.196681	0.998216	0.996043
	Validation	0.149951	0.005277	0.250084	0.992419	0.927629
	Test	0.105091	0.003961	0.429677	0.970672	0.925284
Conjugate Gradient Descent Algorithm	Training	0.074111	0.000694	0.245997	0.991953	0.983644
	Validation	0.120994	0.011796	0.360898	0.998933	0.94875
	Test	0.090322	0.000149	0.217838	0.987336	0.964738
Quasi-Newton Algorithm	Training	0.053761	0.001836	0.147875	0.998276	0.996383
	Validation	0.147094	0.002744	0.228939	0.9384	0.930479
	Test	0.12809	0.002965	0.488405	0.961692	0.901069
Limited Memory Quasi-Newton Algorithm	Training	0.05551	0.00108	0.157847	0.99792	0.995219
	Validation	0.170377	0.00838	0.287672	0.994189	0.930772
	Test	0.11543	0.004775	0.403651	0.97401	0.932841
Levenberg-Marquardt Algorithm	Training	0.02292	0.000595	0.081729	0.999979	0.99967
	Validation	0.542942	0.115233	1.208.797	0.97846	0.592145
	Test	0.247579	0.008213	1.072.021	0.962609	0.857825
Incremental Back-Propagation Algorithm	Training	0.060445	0.000956	0.216677	0.996779	0.992697
	Validation	0.280125	0.015173	0.441624	0.990249	0.865756
	Test	0.128041	0.012995	0.336433	0.983679	0.946331
Batch Back-Propagation Algorithm	Training	0.231164	0.003341	0.675748	0.993436	0.939288
	Validation	0.633675	0.028555	1.299.133	0.987053	0.17434
	Test	0.345177	0.000587	1.103.253	0.973797	0.767443

Different neural network models for short-term natural gas consumption forecasting were studied in this work. It demonstrated which factors effects the demand of natural gas mostly. Best suitable models are listed above results. The works have shown that Batch Back-Propagation Algorithm is not suitable for this type of models**,** test ARE result are over 35%. And also not only processing time of training is very long for Levenberg-Marquardt and Quasi Newton Algorithm models but also test ARE results are not good enough as other methods. Quick Propagation algorithm method is very suitable for daily forecasting and Conjugate Gradient Descent Algorithm is good for weekly and monthly models.

The resulting gating system is capable of reliable identification of the start and end of the heating season and, combined with the statistical models, of accurate predictions of gas load.

CHAPTER 6

CONCLUSIONS

In this study, forecasting of Istanbul natural gas by using artifical neural network models is done. The forecasting models covers daily, weekly and monthly values for Istanbul general, Anatolian and European side seperately. After the data from the IGDAS and TUMAK (Turkey Metaorolgy Association) was gethered, firstly these data were designed for hourly, daily, weekly and monthly values for modeling the structure of forecasting. While forecasting the models, recent years used artificial neural network algorithms were used. The reasons behind choosing ANN are the ability of forecasting future values of more than one variable at the same time and to model the nonlinear relation in the data structure.

Before estimating the natural gas energy consumption of Istanbul, first of all general structure and working system of gas network system of Istanbul was analyzed in Chapter 2. As it is mentioned in Chapter 2, Istanbul has a very complex network system, consumption rates and number of consumers is not stable. The network system has not completed yet so consumer number is increasing day to day and consumption rates are increasing rapidly. Factors that effects forecasting values are not only depends on these but also other factors like temperature, wind speed, price, economic conditions, consumer profile have significant effect on it. As it is known geographical condition of Istanbul is very complex and temperature, the most effective factor on natural gas consumption, values are not same in all area of Istanbul. It differs from location to location. Consumer profile is also having a changeable situation. In Chapter 2 according to these factors Istanbul's natural gas consumption evaluation was done and factors could be listed as below.

As graphs in Chapter 2 analysed remarkable terms listed below:

- Climatic effects are top in winters.
- Change per average unit gas consumer alters every years (even in summers).
- Average gas consumption amount changes even in summers. This change stems from gas usage habits
- There is a close relationship between economical situations and usage habits. Gas usage decreased noteworthly especially in 2001 and 2002.
- There is a wide difference even in summers in which a linear change expected. This especially makes it hard to determine the consumption amount of unit gas consumer.
- Wide differences in summer term variations indicates that there are some factors that are independent from measurement.

All these entries as forming a real correlation in determining a real value, lead important deviations determining the real correlations because they increase vagueness of the parameters. As data reliability increases forming accurate relationships will be much more possible. First calculation related to this has been made by using the correlations that have been gained by the summation of the average monthly consumption values in yearly base. Informations about this calculation are given in Table 2.4.

After evaluating the Istanbul natural gas energy profile, before applying the prediction methods and strategies, a wide factor analyzing of data was done in Chapter 3 for hourly, daily, weekly and monthly values in order to determine the factors that affects the demand of natural gas. Data were analyzed by “analysis of variance” (ANOVA) technique to understand the effects of factors.

As mentioned in Chapter 2 and 3 outside temperature is the most important factor because natural gas is mostly used for heating. There are some other several factors that effects the consumption like wind speed, customer profile, type of combi boiler etc., but all of these factors are extremely hard to predict so including them in the model would only compromise the accuracy of the model and so results. For hourly demand forecasting the factors are not only these that mentioned above but also customers in where they live in a industry region or another normal regions effects the demand of natural gas. Second selected parameters are temperature and hour factors, because a significant difference in consumption can be observed day hours and night hours.

An example of natural gas consumption values are shown in the Table 3.1. This type of data were taken between 2004 and 2007 years. In order to determine the hour factor for different seasons, analysis of variance (ANOVA) test was applied on the values. We have calculated ANOVA test for the month of different seasons because consumption values in different hours are not same with the month of different seasons. As a result of this we have taken a monthly hour values from a winter season and a month from a summer season. We could take the rest of the month results same with these months' results because consumption behaviours are same with them but they did not be showed. Analysis of variance test is applied on consumption versus hour and temperature and results are shown in ANOVA in Chapter 3. When demand of natural gas is very high between 07 am and 09 pm in winter season, there is a decreasing trend in summer season's daily hours between 09 am and 8 pm. The results were shown in Figure 3.3 and Figure 3.4 clearly.

For daily, weekly and monthly forecasting models similar calculations were done. As a result of these effective factors for these should be listed as below:

For daily models; day, month, year, previous day maximum and minimum temperature, maximum and minimum temperature, total consumer number and previous day consumption values (shown in Table 3.2).

For weekly models; week factor, year, weekly average consumer number, weekly average maximum and minimum temperature values (shown in Table 3.3).

For monthly values; month factor, average monthly consumer number, minimum and maximum consumer number and monthly average temperature values (shown in Table 3.4),

were taken as an input factor for our ANN models.

After briefly explanation of artificial neural network was defined in Chapter 4, artificial neural network forecasting calculations was done in Chapter 5. For obtaining the best ANN forecasting models, Istanbul was examined in two different parts Anatolian and European side and Istanbul completaly.

In Chpter 5 founded network types were trained by both Mathlab Neural Network Module (nntool) and NeuroIntelligence Neural Network based program. For all models 60% of the available data is used for training, 20% for validation and 20% for testing.

In our neural network models only one hidden layer was used and for all applications number of units in each hidden layer increased until 50. Each of them replicated 10 times and all results were saved in order to find the best model for our network.

Exhausted search method is used for incremental of number of hidden layer neuron number and inverse test error is selected for fitness criteria. For output layer logistic activation function is applied. As a results of these calculations, performance of ANN models in testing values for daily, weekly and monthly are as below:

For daily values:

Location	Best Network	ARE
Istanbul	[19-48-1]	7,8 %
Istanbul Anatolian side	[19-19-1]	15 %
Istanbul European side	[19-24-1]	6, 5 %

For weekly values:

Location	Best Network	ARE
Istanbul	[8-38-1]	6 %
Istanbul Anatolian side	[5-23-1]	10, 8 %
Istanbul European side	[19-31-1]	7, 3 %

For monthly values:

Location	Best Network	ARE
Istanbul	[20-43-1]	5,4 %

Istanbul Anatolian side	[20-18-1]	5, 9 %
Istanbul European side	[19-45-1]	7, 4 %

In the last part Chapter 5, seven different learning algorithms were used and comparison of their performance is done according to the absolute relative error (ARE). ARE is error value that indicates the quality of neural network training. This index is calculated by dividing the difference between actual and desired values of output values by the modules of desired output value. After all works were done best network models for each other were found. Comparison of seven different training algorithms was done after best architectures were found for each other. For daily models best network is Quick Propagation Algorithm, weekly and monthly forecasting values' best network is Conjugate Gradient Descent Algorithm with the test ARE results for daily 6,5%, for weekly 6% and for monthly values 5,4 % .

Although it seems here, the actual cases are smaller than this value since there are several extreme errors in the data set greater than 15%. These are the ones as mentioned before in factor analyzing section, because of the complexity of Istanbul natural gas distribution system there are some other undefined reasons that effects the consumption of natural gas. After these extremes are excluded the ARE value is less than sequentially 5.5%, %5 and %4. The comparison between target and estimated consumption values for daily, weekly and monthly models are shown in Table 5.46, Table 5.47 and Table 5.48.

REFERANCES

[1] Hamzacebi C., 2006, “Forecasting of Turkey’s electricity energy consumption on sectoral bases ”, Energy Policy 35 (2007) pp. 2009-2016

[2] Wen Xie, “A new method for crude oil price forecasting based on support vector machines”, MADIS, July 2006, pp. 3

[3] Peharda D., Delimer M., Loncaric S. 2001’ “Short term hourly forecasting if gas consumption using neural networks”, 23th Int. Con. Informayion Technology Interfaces ITI 2001, June 19-22

[4] Dr. Karabay İ., Türkiye ve İstanbul’da Gaz Tüketim Analizi December-2006

[5] B. Karlık, Artificial Neural Network Lesson Powerpoint Notes

[6] Asar, A., J. R. McDonald, 1994, "A specification of neural network applications in The load forecasting problem", IEEE transactions on control systems technology, Vol. 2, No. 2, June 1994, pp. 135-141.

[7] Bakirtzis, A. G., J. B. Theocharis, S. J. Kiartzis, K. J. Satsios, 1995, "Short term load Forecasting using fuzzy neural networks", IEEE Transactions on Power Systems, Vol. 10, No. 3, August 1995, pp. 1518-1524

[8] Mohammed, O., D. Park, R. Merchant, T. Dinh, C. Tong, A. Azeem, J. Farah, C. Drake, 1995, "Practical experiences with an adaptive neural network short-term load Forecasting system", IEEE Transactions on Power Systems, Vol. 10, No. 1, February 1995, pp. 254-265.

[9] R. H. Brown, L. Matin, P. Kharouf, and L. P. Piessens, “Development of artificial neural-network models to predict daily gas consumption,” *A.G.A. Forecasting Rev.*, vol. 5, pp. 1–22, 1996.

11, No. 4, November 1996, pp. 1736-1742.

[10] A. Khotanzad, R. A. Rohani, and D. J. Maratukulam, “ANNSTLF—Artificial neural-network short-term load forecaster—Generation three,” *IEEE Trans. Power Syst.*, vol. 13, pp. 1413–1423, 1998.

[11] Alireza Khotanzad, *Member, IEEE*, Hassan Elragal, *Member, IEEE*, and Tsun-Liang Lu, *Member, IEEE* ‘Combination of Artificial Neural-Network Forecasters for Prediction of Natural Gas Consumption IEEE Transactions On Neural Networks, Vol. 11, No. 2, March 2000

[12] Nguyen Hoang Viet; Mandziuk, J. Neural and fuzzy neural networks for natural gas consumption prediction, Neural Networks for Signal Processing, 2003. NNSPapos;03. 2003 IEEE 13th Workshop on Volume , Issue , 17-19 Sept. 2003 Page(s): 759 – 768

[13] Musilek, P.; Pelikan, E.; Brabec, T.; Simunek, M., Recurrent Neural Network Based Gating for Natural Gas Load Prediction System, International Joint Conference on Neural Networks, 2006. IJCNN, Page(s):3736 – 3741

[14] Analysis of Variance, *Wikipedia*, ttp://en.wikipedia.org/wiki/Analysis_of_variance, 2007.

[17] T. Yalcinoz and U. Eminoglu, Short term and medium term power distribution load forecasting by neural networks, *Energy Conversion and Management* **46** (2005), pp. 1393–1405.

[18] C.Ch. Hsu and Ch.Y. Chen, Regional load forecasting in Taiwan—applications of artificial neural networks, *Energy Conversion and Management* **44** (2003), pp. 1941–1949.

[19] M. Beccali, M. Cellura, V. Lo Brano and A. Marvuglia, Forecasting daily urban electric load profiles using artificial neural networks, *Energy Conversion and Management* **45** (2004), pp. 2879–2900.

[20] A. Khotanzad, R.C. Hwang and A. Abaye, An adaptive modular artificial neural network hourly load forecaster and its implementation at electric utilities, *IEEE Transactions on Power Systems* **10** (3) (1995), pp. 1716–1722.

[21] A. Khotanzad, M.H. Davis, A. Abaye and D.J. Maratukulam, An artificial neural network hourly temperature forecaster with applications in load forecasting, *IEEE Transactions on Power Systems* **11** (2) (1996), pp. 870–876

[22] Azadeh, A., Ghaderi, S.F, Sohrabkhani, S., 2006a. Forecasting electrical consumption by neural network. In: Proceedings of Energex2006: The 11th International Energy Conference and Exhibition, Stavanger, Norway, 12–15 June

[23] Azadeh, A., Ghaderi, S.F., Sohrabkhani, S., 2006b. Improving neural networks output with preprocessed data in electricity consumption forecasting. In: Proceedings of the 36th International Conference on Computers and Industrial Engineering, Taipei, Taiwan, 20–23 June.

[24] A. Azadeh, S.F. Ghaderi and S. Sohrabkhani, Forecasting electrical consumption by integration of Neural Network, time series and ANOVA, *Applied Mathematics and Computation* **186** (2007), pp. 1753–1761.

[25] K.Y. Lee and J.H. Park, Short-term load forecasting using an artificial neural network, *IEEE Transactions on Power Systems* **7** (1) (1992), pp. 124–130.

[26] T.W. Chow and C.T. Leung, Neural network based short-term load forecasting using weather compensation, *IEEE Transactions on Power Systems* **11** (4) (1996), pp. 1736–1742.

[27] B.F. Hobbs, U. Helman, S. Jitprapaikulsarn, S. Konda and D. Maratukulam, Artificial neural networks for short-term energy forecasting: accuracy and economic value, *Neurocomputing* **23** (1998), pp. 71–84.

[28] K.Y. Lee and J.H. Park, Short-term load forecasting using an artificial neural network, *IEEE Transactions on Power Systems* **7** (1) (1992), pp. 124–130.

[29] O. Mohammed, D. Park, R. Merchant, T. Dinh, C. Tong and A. Azeem Farah *et al.*, Practical experiences with an adaptive neural network short-term load forecasting system, *IEEE Transactions on Power Systems* **10** (1) (1995), pp. 254–265.

[30] Werbos, P.I., 1974. Beyond Regression: new tools for prediction and analysis in the behavior sciences. Ph.D. Thesis, Harvard University, Cambridge, MA.

[31] D.E. Rumelhart and J.L. McClelland, “Parallel Distributed Processing: Explorations in the Microstructure of Cognition”, Foundations, MIT Press, Cambridge, MA (1986).

[32]S. Haykin. Neural Networks: A comprehensive foundation, Prentice- Hall, New Jersey, 1999

[33] J.P.S. Catalao, J.P.S.S. Mariano, VMF Mendes, LAFM Ferrira, “Short-term Electricity Prices Forecasting in a Competitive market: A neural Network Approach”, Science Direct, Electric Power System Research 77, 2007 ,pp1297-1304,

[34] L.B. Almeida, in: E. Fiesler, R. Beale (Eds), “Multilayer Perceptrons, Handbook of Neural Computation”, Oxford University Press, 1997

[35] Neural Network Training Algorithms, Alyuda Forecaster Help File, http://www.alyuda.com/neural-networks-software.htm