Forecasting Residential Consumption of Natural Gas Using Genetic Algorithms

Nil ARAS

Assist. Prof. Dr. Nil Aras, Anadolu University, Engineering and Architecture Faculty Department of Industrial Engineering, TR26470, Muttalip Mevkii, Eskisehir, Turkey

E-mail:nila@anadolu.edu.tr

ABSTRACT

This paper presents an application of genetic algorithms to forecast short-term demand of natural gas in residences. Residential demand is assumed to be a function of time, heating degree-day value, and consumer price index. A genetic algorithm is designed to estimate parameters of a multiple nonlinear regression model which mathematically represents the relationship between natural gas consumption and influential variables. Genetic algorithms have recently received attention as robust stochastic search algorithms to solve various forecasting problems since they have several significant advantages over conventional methods. Without requiring assumptions need to be made about the underlying function or model, genetic algorithms can attain proper solutions by scanning solution space from many different starting point. To show the applicability and superiority of the described approach, it is considered the monthly data of the residential sector which consumes 23% of imported natural gas in Turkey. The results have revealed that genetic algorithms can be used as an alternative solution approach to forecast the future demand of natural gas.

Keywords : genetic algorithms, nonlinear regression, forecasting, natural gas consumption

1. INTRODUCTION

Energy, with its driving force on world economies, constitutes one of the keystones of the economic growth and development. In this respect, the need for permanent and secure energy resources is continuously rising. Increasing energy demand and environmental consciousness all over the world have made it imperative that energy sources should be used more efficiently.

Natural gas is an extremely important source of energy for reducing pollution and maintaining a clean and healthy environment. It is a vital component of the world's supply of energy. While commonly grouped in with other fossil fuels and sources of energy, there are many characteristics of natural gas that make it unique. It is one of the cleanest, safest, and most useful of all energy sources. For hundred of years, natural gas has been known as a very useful substance. The Chinese discovered a very long time ago that the energy in natural gas could be harnessed, and used to heat water [NaturalGas.org, 2008]. In today, natural gas has many uses, residentially, commercially, and industrially as being used in ways never thought possible with much improved distribution channels and technological advancements.

Worldwide natural gas consumption in the International Energy Outlook (IEO2008) reference case increases from 104 trillion cubic feet in 2005 to 158 trillion cubic feet in 2030. World oil prices are expected to remain high, and as a result natural gas replaces oil wherever possible. In addition, because natural gas produces less carbon dioxide when it is burned than does either coal or petroleum, governments implementing national or regional plans to reduce greenhouse gas emissions may encourage its use to displace other fossil fuels. Natural gas is expected to remain a key energy source for various sectors throughout the projection period.

The OECD (Organisation for Economic Co-operation and Development) countries accounted for 38 percent of the World's total natural gas production and 50 percent of natural gas consumption in 2005; in 2030, they account for 27 percent of production and 42 percent of consumption. As a result, the OECD countries are projected to rely increasingly on imports to meet natural gas demand, with a growing percentage of natural gas imports coming in the form of liquefied natural gas (LNG) [IEO2008].
Increasing worldwide demand of natural gas requires development of intelligent methods and algorithm for its forecasting. The estimation of natural gas demand based on economic and non-economic indicators may be achieved by certain statistical, mathematical and simulation models. Due to fluctuations of natural gas demand indicators, the non-linear forms of the equations could estimate gas demand more effectively. The nonlinearity of these indicators and natural gas demand has lead to search for intelligent solution approach methods such as genetic algorithms, neural networks and fuzzy regression.

There has been increasing interest in recent years about new problem solving and data analysis techniques which derive their inspiration from highly abstracted models of naturally occurring processes. Genetic algorithms (GAs) are powerful and broadly applicable stochastic search and optimization techniques based on the principles of natural evolution. A genetic algorithm uses operators that are analogues of the evolutionary processes of mating, mutation and natural selection to explore multi-dimensional parameter spaces. The impressive advances in computers' capabilities make the statistical applications of the GAs be affordable in practice.

Natural gas is used across all sectors, in varying amounts. The residential sector accounts for one of the greatest proportion of natural gas use in countries. It is one of the most popular fuels for residential heating since it is one of the cheapest forms of energy available to the residential consumer. In fact, it has historically been much cheaper than electricity as a source of energy. Not only is natural gas cheap for the residential consumer, it also has a number of varied uses. The best known uses for natural gas around the home are heating and cooking.

In this study, it is focused on forecasting short-term demand for residential consumption of natural gas. Kaboudan and Liu (2004) have identified three major reason why such forecast is important. First, the supply/demand balances naturally minimize the risk of the fuel's price run-ups. Aggressive price increases typically occur when peak winter gas demand exceeds production capability plus natural gas available in underground storage. Natural gas stocks in underground storage play a critical role in meeting increasing demand. Prudent economic planning requires natural gas suppliers to optimize the amount of natural gas held underground. Overstocking is costly while under stocking is risky. Accurate forecasting of demand for the clean fuel then helps optimize these storage levels. Second, recently demand for natural gas has been climbing faster than ever. Natural gas is becoming a preferred energy source because it is cleaner and cheaper than most others. Third, periodic short-run analysis of demand increases our understanding of its short-term dynamics and helps improve energy planning.

Recently, considerable attention has been focused on the energy resources and energy policies in Turkey. Energy is one Turkey's most important development priorities. Turkey's geopolitical location made possible an important role in regional politics, while domestic energy needs required it to do so. natural gas consumption is the fastest growing primary energy source in turkey and it is projected to increase dramatically in coming years [Hacisalihoglu, 2008]. Turkey's natural gas policy has been criticized in recent years because of exaggerated demand expectations. The agreements signed with gas exporting countries are generally based on take or pay rule. The country where there are not sufficient underground storage facilities would have to pay for gas it could not use because of making contracts according to inflated demand estimations. It is evident that available forecasting method is insufficient and developing more reliable models are necessary since there is a considerable gap between official projections and reality in Turkey.

The aim of this work is to develop an appropriate forecasting model that will determine the changing form of residential natural gas consumption and help predict future demands with forecast errors as small as possible by means of genetic algorithms (GAs). To show the applicability and superiority of the proposed approach, the residential consumption values of natural gas in Turkey is used. The following sections of the paper include a short overview for the available studies of energy forecasting techniques, data analysis for the case study, general information about genetic algorithms, details of the proposed approach, and conclusions.

2. LITERATURE REVIEW

Since the petrol crisis during the 1970s, numerous researchers have studied energy issues and especially have concentrated on estimating energy demand models that will reduce the magnitude of forecasting errors for more efficient consumption of energy resources.

Bohi (1981) and Hartman (1979) have generally studied modeling energy demand in residences. Liu and Subbaryan (1996) have compared the performances of fuzzy logic, neural networks and autoregressive models. Taylor and Majithia (2000) have made combined forecasts in accordance with varying weights by taking into account three popular traditional methods. Ceylan et al. (2005), Canyurt et al. (2004, 2008), and Ozturk et al. (2006) proposes different models based on genetic algorithms for estimating energy and exergy demand in their studies. Ozcelik and Hepbasli (2006) have used a simulated annealing approach for estimating petroleum exergy production and consumption.

Other studies can be illustrated as weather sensitive approach (Campo and Ruiz, 1987), time series (Hagan and Behr, 1987), expert systems (Rahman and Bhatgar, 1988), regression (Papalexopoulos and Hesterberg, 1990; Charleson and Weber, 1993), statistical decision functions (Hubele and Cheng, 1990), econometric analysis (Lee and Singh, 1994), experimental design (Bartels and Fiebig, 2000), artificial neural networks (Hill, et al., 1996; Connor, 1995; Abdel-Aal, 2008; Azadeh et al.,2008), particle swarm optimization (Wang et al., 2008), genetic algorithms (Naggar and Rumaih, 2005; Azadeh and Tarverdian, 2007), ant colony optimization (Toksari, 2007), grey prediction with rolling mechanism (Akay and Atak, 2007), evolutionary programming (Sathyanarayan et al., 1999), and genetic programming (Lee et al., 1997).

Increase in energy demand is parallel to a rapid increase in natural gas demand. There are studies that use different methods to forecast natural gas demand. Herbert (1987) has analyzed monthly natural gas sales to the residential consumers in America. Liu and Lin (1991) have compared monthly and quarterly time series performances to forecast residential natural gas consumption in Taiwan and studied the relationships between natural gas consumption, its price and average temperature. Using two econometric models, Eltony (1996) found that natural gas demand is not elastic to price and income in short and long terms in Kuwait. Smith, Husein and Leonard (1996) have used expert systems in decision making process when forecasting short term regional gas demand. Bartels, Fiebing and Nahm (1996) have used conditional demand analysis in order to determine gas demand. In their studies, they researched the impacts of total residential income and other defining information of residences (the number of family members and rooms) on gas demand.

Hobbs, et al. (1998) and Brown (1996) have used artificial neural networks to forecast natural gas consumption. Knowles and Wirick (1998) have developed a portfolio optimization model that uses multiple weather scenarios for a natural gas distribution company. The company has annually saved $50 million using this model. Using architectural structure and physical and thermal features of structural materials, Durmayaz, Kadioglu and Sen (2000) have calculated the need for natural gas of a flat in Istanbul with different scenarios for different numbers of people by using degree hours method. Gumrah et al. (2001) have forecasted natural gas demand using a model that depends on degree-day concept consisting of the number of annual customers, average degree-day values and usage rates per customer. Gorucu et al. (2004) proposed an approach based on artificial neural network for natural gas demand. Aras and Aras (2004) have estimated residential natural gas consumption in Turkey using inferential autoregressive time series models. Kaboudan and Liu (2004) have developed regression models based on genetic algorithms for natural gas demand in U.S. A.

The following three general conclusions can be drawn based on the available studies. Firstly, the forecasting period has been mainly classified into two ranges: short and long range forecasts. The short-range forecast is useful in determining unit-commitment and economic dispatch; whereas the long-range forecasts is necessary for system expansion planning and financial analysis. Second, it has been noticed that various methods used in the studies on energy demand forecast have been generally designed for electricity demand. There are fewer studies on the other types of energy relatively. Third, many forecasting algorithms proposed in recent years can be generally divided into two broad categories: parametric methods and artificial intelligence-based methods.

3. DATA ANALYSIS

It is required that we should make an experiment on a proper data set to show the applicability of the proposed approach. The described forecasting method based on genetic algorithms has been applied by using observation data in Eskisehir, one of the cities where natural gas is supplied for residential purposes in Turkey. In this section, after general information is given about natural gas consumption in Turkey, the factors that have influence on residential natural gas demand are discussed. The data set of the study includes monthly natural gas consumption level, average daily temperatures, and consumer price index for a 140-month period from December 1996 to July 2008.

Turkey's energy demand has been rapidly increasing as a result of the factors like booming energy, high growth in population, industrialization and urbanization. There is also increase in the natural gas demand as a consequence of the increasing demand for energy. Natural gas is an imported energy source for Turkey where domestic gas reserves are rather limited. The proportion of imported natural gas in the primary energy consumption has risen more and more since gas started to be supplied from former Soviet Union in 1987. Significant progress has been made since years in the context of developing the natural gas transmission and storage infrastructure and making natural gas consumption prevalent in the country.

Turkey is located at the crossroad between both the Caspian and Middle East regions possessing rich oil and gas reserves, and the Western countries which are in need of such resources. Moreover, Turkey is one of the leading countries of OECD Europe with an annual natural gas consumption 34.5 Billion m^3 [Annual Report, 2007]. The OECD countries in Europe, natural gas consumption is projected to grow mostly as a result of increasing use for power generation. Many of the OECD Europe nations have made commitments to reduce carbon dioxide emissions, bolstering the incentive for governments to encourage the use of natural gas in place of other fossil fuels [IEO, 2008]. With renewable energy sources projected to remain more expensive than natural gas in OECD Europe, natural gas is expected to be the fuel of choice for new generating capacity.

Natural gas is one of the most popular fuels for residential heating in Turkey where the ratio of residential consumption in the national consumption reached to 23% as can be seen in Figure 1. This popularity is viewed through the almost all proportion of new homes built with natural gas heating. Figure 2 shows the continual increase in the residential natural gas demand according to years.

Natural gas is used for three reasons in residences: space heating, cooking, and water heating. Natural gas need for cooking and hot water has a stable consumption level throughout the year since it is only slightly affected by climatic conditions. On the other hand, residential heating depends very much on various climatic conditions, such as outside temperature, wind, and humidity. Since the principal aim of residential natural gas usage is space heating in Turkey, the natural gas consumption has been modeled as a function of time, weather expressed in terms of degree days and consumer price index (CPI).

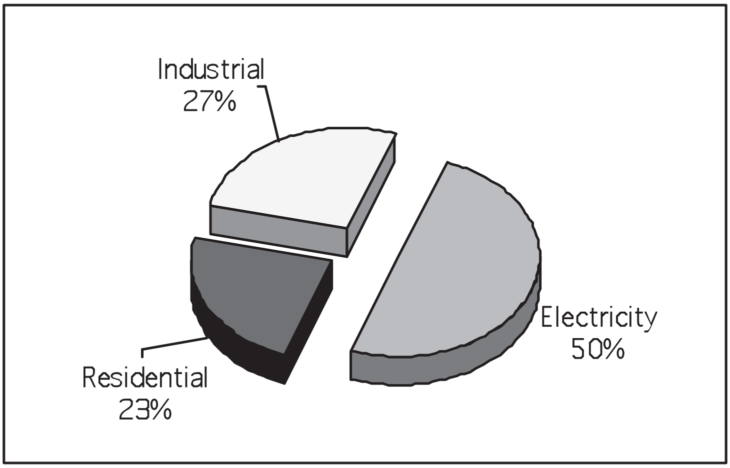

Figure 1. Breakdown of natural gas sales by sectors in 2007 in Turkey [Annual Report, 2007]

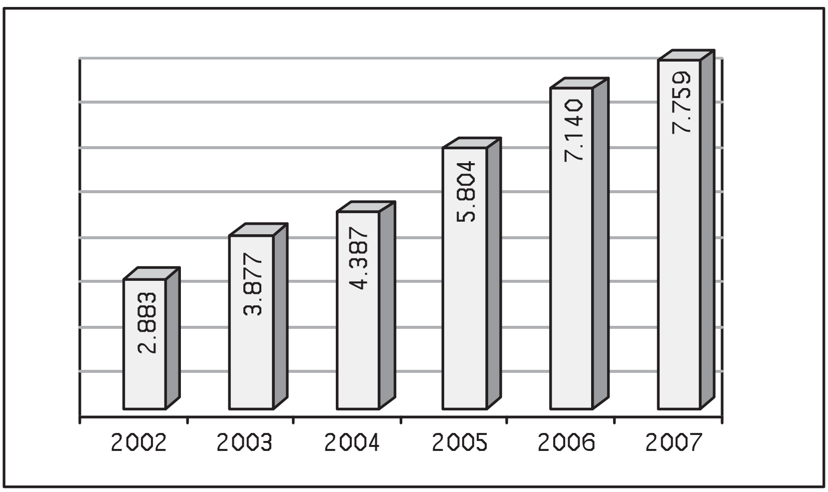

Figure 2. Natural gas sales (Million Sm^3) in residential sector by the years in Turkey [Annual Report, 2007]

The data set of the study includes monthly residential natural gas consumption levels, average temperatures and consumer price indexes for a 140-month period from December 1996 to July 2008. The residential consumption levels of natural gas were provided by ESGAZ (Natural Gas City Distribution Company) and the unit for consumption level is standard meter cube (sm^3). Average daily temperatures which are used to calculate monthly degree-day values were obtained from Turkish State Meteorological Service and the unit is centigrade (^oC). The CPI values were downloaded from the website of Turkish Statistical Institute (TURKSTAT, 2008).

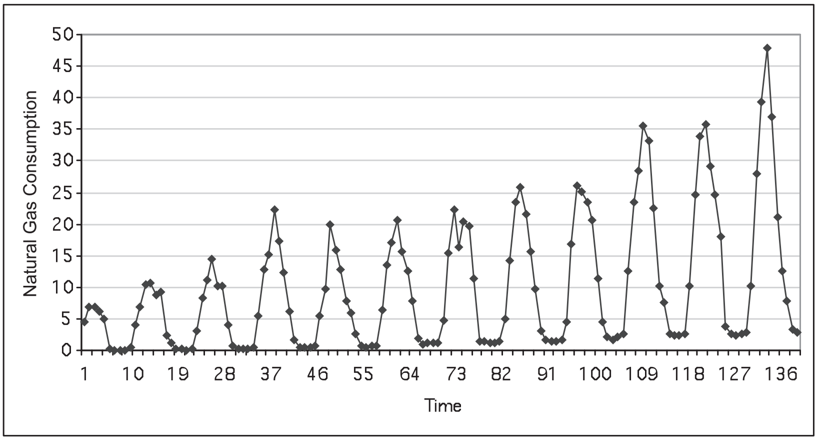

Figure 3. The monthly natural gas consumption (x10^6 sm^3) versus time (from 1996 December to 2008 July)

In Figure 3, amounts of monthly natural gas consumption in respect to time are given. It is quite clear from the consumption curve that seasonal weather changes are very effective and amplitude of the seasonal variation increases with time. It also seems that the consumption levels of natural gas are rather changeable in the months of heating period while the amounts of gas usage are almost remain stable in the non-heating periods. Figure 4 shows that the yearly consumption levels have been risen gradually by reason of increasing customer numbers in every year.

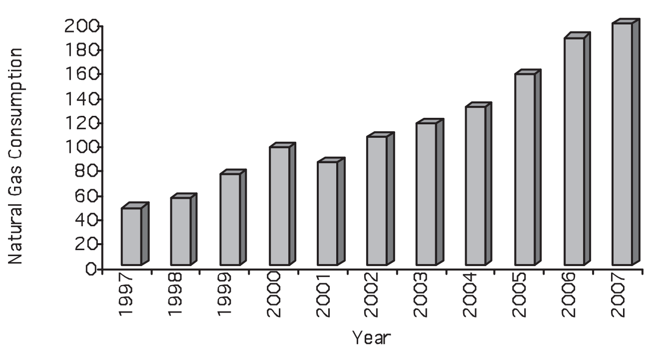

Figure 4. Natural gas consumption (x10^6 sm^3) per year

Using daily average temperatures, the monthly heating degree day numbers have been calculated for Eskisehir. Heating degree day is a quantitative index designed to reflect the demand for energy needed to heat a home or business. This index is derived from daily temperature observations, and the heating requirements for a given structure at a specific location are considered to be directly proportional to the number of heating degree days at that location.

The number of heating degree-days within a period is equal to the total of the differences of temperature of a place heated on heating days and outside temperature. The number of degree-days during heating months is calculated as

$$HDD = \sum_{j=1}^{Z} \left(T_i - T_{do} \right)_j \tag{1}$$

where Z : The length of heating month (The number of days in a month),
T_i : Temperature of a heated place / internal temperature [°C],
T_{do} : Daily average outdoors temperature [°C],
HDD : Monthly heating degree day value .

In most countries T_i = 20°C and $T_{do} \leq$ 18°C are accepted [Dagsoz, 1995]. The value of 18°C which affects the start date of heating season in buildings is called base temperature. The base temperature corresponds to human comfort requirements and varies from one location to another. If the outside temperature is over a certain degree, there is no need for heating a building. If the average outside temperature is below a limit level, the need arises. When daily average temperature is 18°C or lower, degree day is calculated for that day, while internal temperature is kept constant at 20°C. Monthly total of degree days is obtained by aggregating daily degree day values.

In Figure 5, the monthly heating degree day numbers for Eskisehir can be seen. The higher the outside temperature is, the lower degree day value is and the less energy is needed to heat buildings. The annual total HDD values are shown in Figure 6. When 11 years' average is taken, the number of annual heating degree day for Eskisehir is 3113,16, the maximum and minimum HDD values per year are 3383,80 and 2772,60. According to the results obtained during a 12 years period, the number of monthly degree days in Eskişehir ranges between 0 and 721,60.

Figure 4 shows that natural gas consumption level for 2001 is lower than the year before although the number of residences that are using natural gas is increasing each year. If we compare the figures 4 and 6, we can see that the number of degree days for that year is lower than the average. This proves that there is a strong relationship between natural gas consumption and HDD number. The higher the outside temperature is, the lower degree day value is and the less energy is needed to heat buildings.

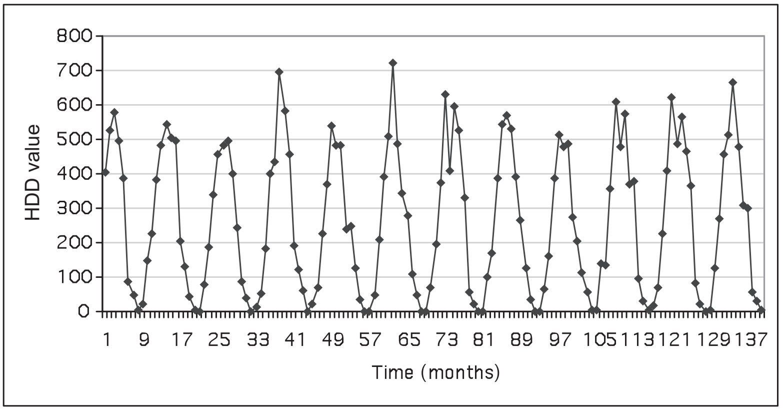

Figure 5. Monthly heating degree-day values in Eskisehir

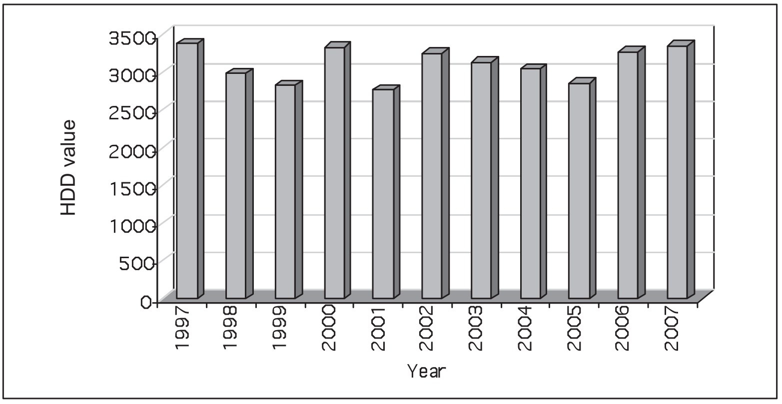

Figure 6. Total heating degree day values per year in Eskisehir

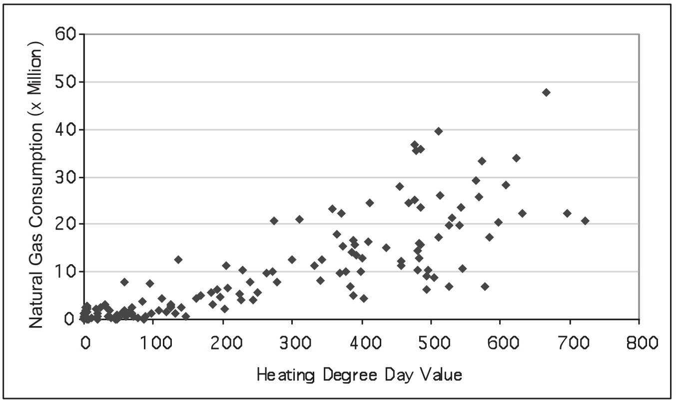

Figure 7. The monthly natural gas consumption versus the monthly HDD value

Figure 7 is a scatter plot of monthly demand versus monthly heating degree day numbers. It is obvious that there is a directly strong relationship between the two variables. Energy demand for residential heating changes in accordance with how cold the weather is. During the cool winter months, demand is inversely related to temperature; lower temperatures cause increased heating degree day values and increased usage of space heating equipment. In a similar manner, the summer months reveal a negative relationship between demand and temperature; higher temperatures yield lower heating degree days, so does natural gas consumption. It seems that there is a nonlinear relation between the natural consumption values and HDD numbers.

Temperature, as measured by heating degree days, is important because a large percentage of total residential natural gas is used in space heating. However, energy consumption levels in buildings may also be affected from purchasing power or budget limitations of the consumers. Considering this possible relationship that is inversely proportional between price and consumption, the effect of the Consumer Price Index (CPI) as price or budget variable on residential demand of natural gas has been analyzed. Natural gas sales price has been not considered in this study because natural gas is the cheapest form of energy for space heating in Eskisehir, in any case. It can be said that analyzing effect of CPI instead of the sales price is innovative according to past studies.

The CPI is a measure of the average change in prices over time in a market basket of goods and services. For example, if in November 1996 the CPI (with base year 1992 as 100) reaches in 150, the purchasing power of the November 1996 consumer dollar is $0.67 (100/150). It means that the 1996 dollar purchased only two-thirds of the goods and services that could have been purchased with a base period (1992) dollar [Hanke and Reitsch, 1998]. The purchasing power of consumers decreases as the CPI increases with respect to a base year. Since Turkey is a country that there are inflation and widely varying prices over time, CPI may be a good indicator to explain energy usage in residences. In Figure 8, the variations of this variable according to time is given. It seems that CPI values increases gradually by the years on a near-linear basis.

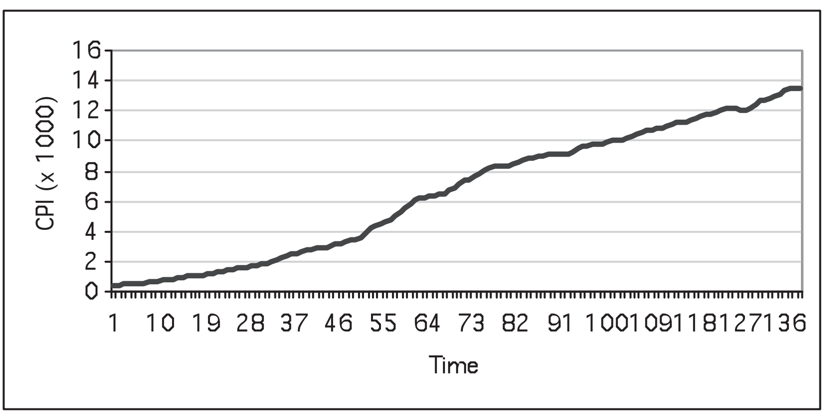

Figure 8. The consumer price index (CPI) according to time

4. FORECASTING APPROACH

Natural gas demand is defined as the rate (measured in standard cube meters) at which natural gas is delivered to customers. Since demand is recorded over time (months), the dependent variable is a time series and one approach to modeling natural gas demand is to use a multiple regression model. While a multiple regression model allows the simultaneous testing and modeling of multiple independent variables, the nonlinear time series seems to be ideal due to the nonlinearity of the data.

One of the most common situations in statistical analysis is that of data which consist of observed responses y_t known to be dependent on corresponding k-dimensional inputs $\mathbf{x}_t$. This situation may be represented by the regression equations

$$y_t = f(\boldsymbol{x}_t, \theta^*) + e_t, \quad t=1,2,..., n \tag{2}$$

where $f(\mathbf{x}_t, \theta^*)$ is the response function, θ^* is a p-dimensional vector of unknown parameters, and the e_t represent unobservable observational or experimental errors It is assumed that these errors are independently and normally distributed with mean zero and unknown variance σ^2 [Gallant, 1975].

The preceding section gives important information to understand the factors affecting natural gas demand and to develop a statistical model for forecasting. Changing weather conditions represent the primary source of variation in natural gas demand. It is typically assumed that demand consists of two components, a non-weather sensitive "base" demand that is not influenced by temperature changes and a weather-sensitive demand that is highly responsive to changes in temperature. The principal factor that affects the usage of non-weather sensitive appliances (such as cookers and water heaters) is the month of the year. Besides, the change on the consumer price index can be considered as another factor affects non-weather sensitive demand.

The single most important factor affecting the usage of weather-sensitive appliances (such as heaters) is temperature. During the winter months, as temperatures drop below comfortable levels, customers begin to operate their natural gas heating units, thereby increasing the level of demand placed on the system. Moreover, time is a independent variable in a time series and increasing customer number in every month doesn't depend on weather situation but time. To summarize, the dependent variable (Y_t) and the independent variables can be described as follows:

Dependent variable:

Y_t : Natural gas demand (in sm^3) observed on month t

Independent variables:

t : Time index, observed month, t =1, 2, ..., n

HDD_t : Heating degree day value for month t.

CPI_t : Consumer price index for month t

The following nonlinear regression model is proposed assuming that the relationship between dependent and independent variables are non-linear in the model that consist of time, heating degree day value, and consumer price index according to the above findings:

$$\begin{aligned} Y_t = {} & \beta_0 + \beta_1(t) + \beta_2(HDD)_t + \beta_3(CPI)_t \\ & + \beta_4(t)^2 + \beta_5(HDD)_t^2 + \beta_6(CPI)_t^2 \\ & + \beta_7(t)(HDD)_t + \beta_8(t)(CPI)_t + \beta_9(HDD)_t(CPI)_t \end{aligned} \tag{3}$$

where β_0 , β_1, β_2, β_3, β_4, β_5, β_6, β_7, β_8 and β_9 are unknown parameters that must be estimated.

This is a complete second order regression model with three independent variables. The complete second-order model includes the constant β_0, all linear (first-order) terms, all two-variable interactions, and all quadratic terms [Mendenhall and Sincich, 1996]. It account for interactions between each two variables. The unknown parameter values should be optimized to be able to forecast the future energy demand. The data between December 1996 and July 2008 is used as sample data for the identification and estimation of the model parameters in this paper. Many studies are made for strongly estimating the energy demand of Turkey. Since the main purpose of these efforts is to develop a model that closes the gap between the energy demand predictions and observed energy demands, the problem is typically an optimization problem which minimizes the gap. Because of the some limitations of the traditional statistical methods, genetic algorithms (GAs) has been suggested for estimation of unknown parameters as an alternative method in this paper.

Especially in current years, the development of nonlinear regression analysis provides an extensive alternative for practical applications, because of the fact that the models obtained in many research fields such are almost all nonlinear [Pan, et al., 1995]. However, the complexity of nonlinear regression models and the intractability of their parameter estimation methods strongly limit their application in practice. It cannot be guaranteed to find optimal parameter values in nonlinear regression models because of the some difficulties, while linear regression models are solved optimally. It can be said that there are two major reason why the most of the parameter estimation methods are not easy to handle for practical applications.

First, these algorithms are usually appropriate for a certain problem and may require restrictive assumptions regarding continuity, existence of derivatives, existence of a single mode and other issues. Model parameters in non-linear regression models are of nonlinear nature and they cannot be estimated through ordinary least squares method. Hence, in non-linear regression model various algorithms as to parameter estimation have been developed. Some of these algorithms are Gauss-Newton, Direct Search Method, Hooke-Jeeves Method, Nelder-Mead Method, Gradient Method and Variable Measurement Method [Nash and Walker, 1987]. For instance, much more supplementary information is necessary for the Gauss-Newton Method, one of the commonly used algorithms, to be used as it should be. Moreover, using this algorithm involves a lot of attention as unless the starting point for the algorithm is chosen correctly, the algorithm may incorrectly pinpoint local optimums and thus yields only local optimums. In this case, an inappropriate starting area may augment the orientation of the search and delay the catch of the optimal solution. Sometimes, the developed algorithms for nonlinear regression models make assumptions about the form of population distributions, assumptions that may or may not be subject to verification. Interval estimates using a regression model, for example, assume that the underlying population follows a normal distribution [Hanke and Reitsch, 1998].

There is a second reason if the data set is the time series. In business and economics many regression applications involve time-series. Economic data ordered in a time series can seldom be regarded as a random sample. An observation on price, inventory, production, stocks, and other economic variables in a given time period is usually correlated with (dependent on) the value of the same variable in the previous time period. If the residuals in a regression equation are positively autocorrelated, the use of the least squares procedure posses several problems [Hanke and Reitsch, 1998]. The problem is not bias, the problem is a large variance in the estimates. The least squares procedure is not at fault. Any other estimating procedure (such as fitting by eye) would fit the unique tilted data pattern in the same way. Other techniques that may be more efficient and may improve the estimating procedure.

GAs were developed initially by Holland and his associates at the University of Michigan in the 1960s and 1970s, and first full, systematic treatment was contained in Holland's book "Adaptation in Natural and Artificial Systems" published in 1975. The name genetic algorithm originates from the analogy between the representation of a complex structure by means of a vector of components, and the idea, familiar to biologists, of the genetic structure of a chromosome [Reeves, 1993].

Genetic algorithms differ from conventional optimization and search procedures in several fundamental ways. Goldberg [1989] has summarized this as follows:

1. GAs work with a coding of solution set, not the solution themselves.
2. GAs search from a population of solutions, not a single solution.
3. GAs use payoff information (fitness function), not derivatives or other auxiliary knowledge.
4. GAs use probabilistic transition rules, not deterministic rules.

The use of genetic algorithms (GA) is a recently developed optimization approach that can be used as an alternative to regression analysis to fit mathematical models. Chatterjee, Laudato and Lynch (1996) and Chatterjee and Laudato (1997) show that GAs may be useful for estimating the parameters of a statistical model. The statistical applications of the GAs have been discussed by Chatterjee et al. (1996), and cover optimization problems where the usual mathematical requirements, for example continuity, differentiability and convexity, do not apply. Relevant areas are the variable selection problem in regression and multivariate statistical methods such as robust estimation of model parameters.

GAs are robust search and optimization techniques which are finding application in number of practical problems. The robustness of GAs is due to their capacity to locate the global optimum in a multimodal landscape. There are two important issues in search strageties: exploiting the best solution and exploring the search space. GAs are a class of general-purpose search methods combining elements of directed and stochastic search which can make a remarkable balance between exploration and exploitation of the search space [Gen and Chang, 1997].

5. GENETIC ALGORITHMS

Genetic algorithms are stochastic search techniques based on the mechanism of natural selection and natural genetics. In this section, a brief introduction to genetic algorithms is given as search and optimization techniques, then considered genetic formulation of curve fitting is described.

GAs, differing from conventional search techniques, start with an initial set of random solutions called population. Each individual in the population is called a chromosome, representing a solution to the problem at hand. A chromosome is a string of symbols it is usually, but not necessarily, a binary bit string. The chromosomes evolve through successive iterations, called generations. During each generation, the chromosomes are evaluated, using some measures of fitness. Fitter chromosomes have higher probabilities of being selected. After several generations, the algorithms converge to the best chromosome, which hopefully represents the optimum or suboptimal solution to the problem. Let *P(t)* and *C(t)* be parents and offspring in current generation t; general structure of the genetic algorithms are illustrated in the Figures 9 and 10.

Procedure: Genetic Algorithms
  begin
    $t \leftarrow 0$;
    initialize *P(t)*;
    evaluate *P(t)*;
    while (not termination condition) do
        recombine *P(t)* to yield *C(t)*;
        evaluate *C(t)*;
        select *P(t+1)* from *P(t)* and *C(t)*;
        $t \leftarrow t+1$;
  end
end

Figure 9. Framework of a Genetic Algorithm [Gen and Cheng, 1997].

The name genetic algorithm originates from the analogy between the representation of a complex structure by means of a vector of components, and the idea, familiar to biologists, of the genetic structure of a chromosome. A comparison of the terms in genetics (biology) and computer science is displayed in Table 1.

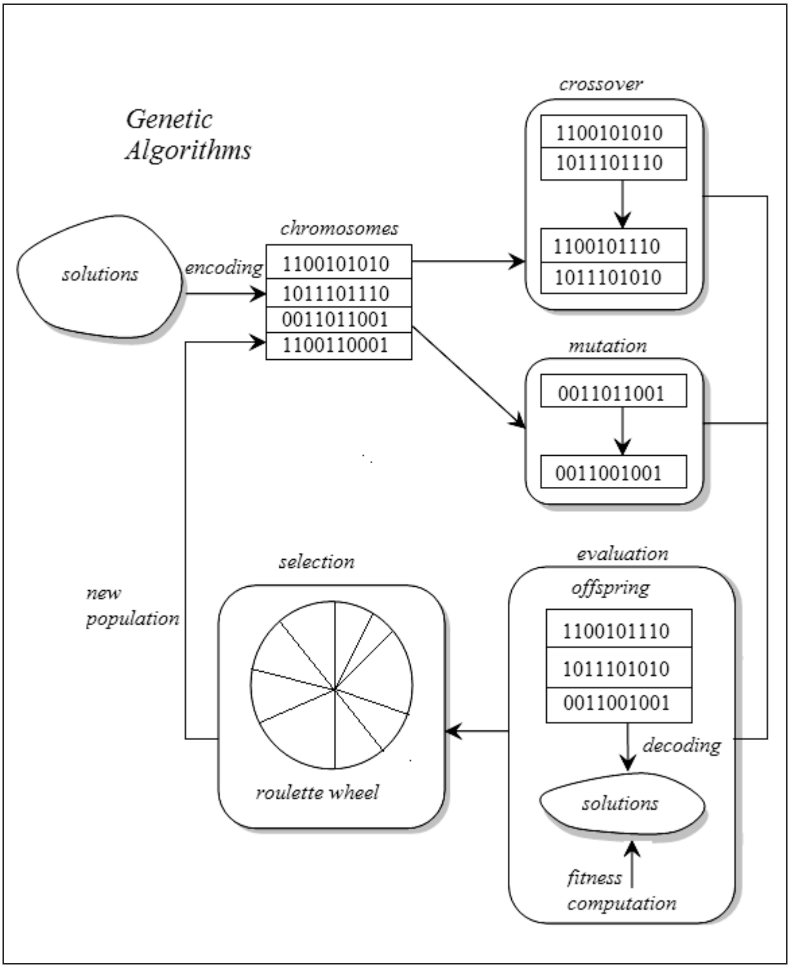

Figure 10. The general structure of genetic algorithms [Gen and Cheng, 1997].

Table 1. Comparison of terms in computer science and biology

Computing term	Biological term
Solution to a problem	Individual
Set of solutions	Population
Quality of a solution	Fitness value
Encoding of solution	Chromosome
Part of the encoding of a solution	Gene
Search operators	Crossover and Mutation
Reuse of good solutions	Natural selection

For a detailed information of GAs and their applications, one may refer to Gen and Cheng [1997] and Goldberg [1989]. Figure 11 briefly summarizes used generic search process in the study.

There are three main questions that have become in relevant topics in GAs design research: (1) encoding, (2) operators, and (3) control parameters. For encoding and operators it is used those presented in Goldberg (1989) - binary coding, roulette wheel reproduction, one point crossover and mutation.

1. **[Initial]** Start with a randomly generated population of n chromosomes (bitstrings) that each represent a solution for the problem.
2. **[Fitness]** Calculate fitness f(x) for each chromosome x in the population.
3. **[Elitism]** Copy the best chromosome as a new offspring to the population in the next generation.
4. **[New population]** Repeat ((n-1)/2) times, each time generating two individuals:
    - i. [**Roulette Wheel Selection]** select a pair of parents (put them back after selection) from a population according to their fitness (the better fitness, the bigger chance to be selected)
    - ii. **[Crossover]** With a crossover probability pc, do cross over the parents to form two new offspring (children). Otherwise just clone the parents (take an identical copy of each parent)
    - iii. **[Mutation]** With a mutation probability pm, mutate both offspring at each gene.
    - iv. **[Accepting]** Place new offspring in the new population.

    If n is odd, discard one member of new generation at random
5. **[Replace]** Replace current population with new population and use new generated population for a further run of the algorithm.
6. **[Test]** If the end condition is satisfied, stop, and return the best solution in current population.
7. **[Loop]** Go to step 2.

Figure 11. Outline of the used Genetic Algorithm

Encoding

The first step of the genetic algorithm is to encode the parameter values into an appropriate range with finite-digit digital strings (i.e. the chromosomes, usually binary strings). A widely used formula for decoding is

$$\beta = L + \frac{A}{2^B - 1} \times (U - L) \tag{4}$$

where β is the parameter value of the chromosome; U and L are the upper and the lower bounds in the parameter; A is the number in decimal form that is being represented in binary form of the chromosome; and B is the digits of the chromosome [Wu and Chang, 2002]. The binary string representation for the forecasting model is shown in the Figure 12.

$$\boxed{\underbrace{10011....1101}_{\beta_0}\ \Big|\ \underbrace{110101....001}_{\beta_1}\ \Big|\ \underbrace{00110......11}_{\beta_2}\ \Big|\ \cdots\cdots\cdots\ \Big|\ \underbrace{00101....0101}_{\beta_9}}$$

Figure 12. The representation of a solution as a chromosome

Fitness Function

Fitness is an important concept for the operation of the GA. The fitness of a string is a measure of the quality of the trial solution represented by the string with respect to the function being optimized. Thus, high fitness corresponds to a high value (in a maximization problem) or a low value (in a minimization problem) of the function.

Once a functional form and an error metric have been selected, curve fitting becomes an optimization problem over the set of given data points. The objective is to select functional coefficients (i.e. parameter values) which minimize the root mean squared error (RMSE) over the set of data points being considered. Although the problem seems as an unconstrained optimization problem, in fact, there is a sign constraint about predicted values $\hat{Y}_t$. It is required that predicted values can take nonnegative values. The decision model of the problem can be formulated as

$$\begin{aligned} &Minimum \quad RMSE(\hat{Y}_t) = \left[\left(1/n\right)\sum_{t=1}^{n}\left(\hat{Y}_t - Y_t\right)^2\right]^{1/2} \\ &\text{subject to} \\ &\hat{Y}_t \geq 0 \end{aligned} \tag{4}$$

where n is the observation number, Y_t is the observation value, and $\hat{Y}_t$ is the predicted value. This nonlinear estimation problem is an optimization problem whereby the objective function RMSE is minimized.

The fitness function consists of the two sub functions; objective function and a penalizing function. A penalty term is added to the objective function for any violation of the constraints in order to transform the constrained problem into an unconstrained problem by penalizing infeasible solutions. The fitness function can be defined as

$$\text{Fitness } (\hat{Y}_t) = F\,(\hat{Y}_t) + P(\hat{Y}_t) \tag{5}$$

where $\hat{Y}_t$ represents a chromosome, $F(\hat{Y}_t)$ is the objective function of the problem, and $P(\hat{Y}_t)$ the penalty term. The functions $F(\hat{Y}_t)$ and $P(\hat{Y}_t)$ are

$$F(\hat{Y}_t) = \left[\left(1/n\right)\sum_{t=1}^{n}\left(\hat{Y}_t - Y_t\right)^2\right]^{1/2} \tag{6}$$

$$P(\hat{Y}_t) = \begin{cases} 0 & , \quad \hat{Y}_t \geq 0 \\ c \times \left| \hat{Y}_t - Y_t \right| & , \quad \hat{Y}_t < 0 \end{cases} \tag{7}$$

where c is a constant penalty degree, 10000 for this study. Note that for the minimization problems, the fitter chromosome has the lower value of Fitness ($\hat{Y}_t$).

Roulette Wheel Selection

The evolution operation mimics the process of Darwinian evolution to create populations from generation to generation. In most practices, a Roulette Wheel approach is adopted as the selection procedure; it belongs to the fitness-proportional selection and can select a new population with respect to the probability distribution based on fitness values [Gen and Cheng, 1997].

This is a way of choosing members from the population of chromosomes in a way that is proportional to their fitness. Parents are selected according to their fitness. The better the fitness of the chromosome, the greater the chance it will be selected. The idea behind the roulette wheel selection parent selection technique is that each individual is given a chance to become a parent in proportion to its fitness evaluation. It is called roulette wheel selection as the chances of selecting a parent can be seen as spinning a roulette wheel with the size of the slot for each parent being proportional to its fitness. Obviously those with the largest fitness (and slot sizes) have more chance of being chosen.

Elitism

As replacement strategy the roulette wheel selection was applied with the elitist strategy. The best chromosome is copied to the population in the next generation. The rest are chosen in classical way. Elitism can very rapidly increase performance of GAs, because it prevents losing the best found solution to date. This technique ensures that the best members of the population are carried forward from one generation to the next.

Crossover

Crossover is the way in which ``genetic'' information from two ``parent'' strings is combined to generate ``offspring''. Crossover used here is one-cut point method. One position in the chromosome is chosen randomly and the parents are split in a right and a left part. Then two offspring are generated by exchanging the right parts of two parents to generate offspring. Consider two chromosomes as follows, and the cut-point is randomly selected the after the 15th gene:

$$\hat{Y}_t^1 = [10011011110001010110111000010101001011100]$$

$$\hat{Y}_t^2 = [11110101010001011101010101001011010101111]$$

Offspring 1 is built by taking genes from parent 1 from the left of the crossover point and genes from parent 2 from the right of crossover point. Offspring 2 is built in the same way but it takes genes from the left of the crossover point of parent 2 and genes from the right of the crossover point of parent 1. The resulting two offspring by exchanging the right parts of their parents would be as follows:

$$\hat{y}_t^{\,1} = [1001101111000101110101010100101101010111]$$

$$\hat{y}_t^{\,2} = [1111010101000101011011100001010100101100\ ]$$

If we assume that the probability of crossover is set as Pc=0.25, so we expect that, on average, 25% chromosomes undergo crossover.

Mutation

While the crossover operation leads to a mixing of genetic material in the offspring, no new genetic material is introduced, which can lead to lack of population diversity and eventually ``stagnation'' - where the population converges on the same, non-optimal solution. The mutation operator helps to increase population diversity by introducing new genetic material. This can be accomplished by making a random change to one or more randomly chosen genes in an individual. Mutation alters one or more genes with a probability equal to be mutation rate. Assume that the 21^{th} gene of the chromosome $\hat{y}_t^{\,1}$ is selected for a mutation. Since the gene is 1, it would be flipped into 0.

$$\hat{y}_t^{1} = [1001101111000101110101010100101101010111] \quad \text{before mutation}$$

$$\hat{y}_t^{1} = [1001101111000101110111010100101101010111] \quad \text{after mutation}$$

If we assume that the probability of mutation is set as Pm=0.01, so we expect that, on average, 1% of total bit of population would undergo mutation.

6. EXPERIMENTAL RESULTS

The experiments were designed to search the space of GA's defined by four control parameters, and to identify the optimal parameter settings with respect to performance measure, RMSE. The algorithm is coded with Visual C# software and run on a Pentium IV, 3.2 GHz, 1 GB RAM computer. Statistical experiments based on a full factorial design are performed in order to find the best parameter set of the genetic algorithms. Four important factors, iteration number, population size, crossover rate, mutation rate and their levels considered in experimental design are given in Table 2.

Table 2. Factors and their levels considered in the experimental design

Iteration number	Population size	Crossover rate	Mutation rate
100	50	0.50	0.001
200	100	0.70	0.01
500	200	0.90	0.05

Each simulation was repeated for 10 different random number seeds to account for the dependence of the GA on a pseudo-random number generator. After 3×3×3×3×10=810 runs are performed, the following optimal parameter values are obtained, which result in minimum RMSE between the experimental data and estimated results:

Population size (n)	*:200*
Iterations (number of generation)	*:500*
Mutation rate	*:0.01*
Crossover rate	*:0.50*

The estimation model is obtained as the following

$$
\begin{aligned}
Y_t &= 18036 + 1730\ (t) + 5099\ (HDD)_t + (CPI)_t + 57\ (t)^2 + 15\ (HDD)_t^2 \\
&\quad +356\ (t)(HDD)_t
\end{aligned} \tag{8}
$$

where

β_0 = 18036 , β_1 =$\beta 73_0$, β_2 =5099, β_3 =1, β_4 =57, β_5 =15, β_7=356, and β_6=β_8=β_9=0

In Figure 13, monthly variations of natural gas consumption amounts which were observed and obtained with the forecast model are given. Dashed lines represents the forecasted values for heating months. It seems that two curves are very close to each other, which proves that the model conforms to the observation data and can be used for short term forecasts of natural gas demand. The results of the model indicate that developed model is valid and it can be used to predict monthly consumption amounts of natural gas.

Future projections

An important advantage of a time series model is to allow us to make forecast for various scenarios. It is noted that the forecast requires estimates of reasonable heating degree day value and consumer price index in this study. The reliability of the forecast depends on correct estimations of these explanatory variables that are time series and have seasonal variations. The best way is to benefit from past investigations concerned about them. There are several studies made by observing long term historical data to determine mean degree day numbers for each city in Turkey [Dagsoz, 1995; Sen and Kadioglu, 1998]. There are also several researchers who woks on predicting future values of CPI [Yamak and Yamak, 1998; Ozturkler and Altan, 2008]. It is probably to draw necessary information from these studies that show the way to determine values of degree day and CPI.

In order to make a future projection for forecasting Eskisehir's natural gas demand in the years 2009-2013, a scenario is developed to create input variables for the model. It was assumed that the average growth rate of CPI will be 4% according to State Planning Organization's "Medium Term Programme" [SPO, 2008]. Heating degree days were produced randomly assuming they distribute uniformly between maximum and minimum values derived from past data. Figure 14 shows that the forecasted values according to this scenario.

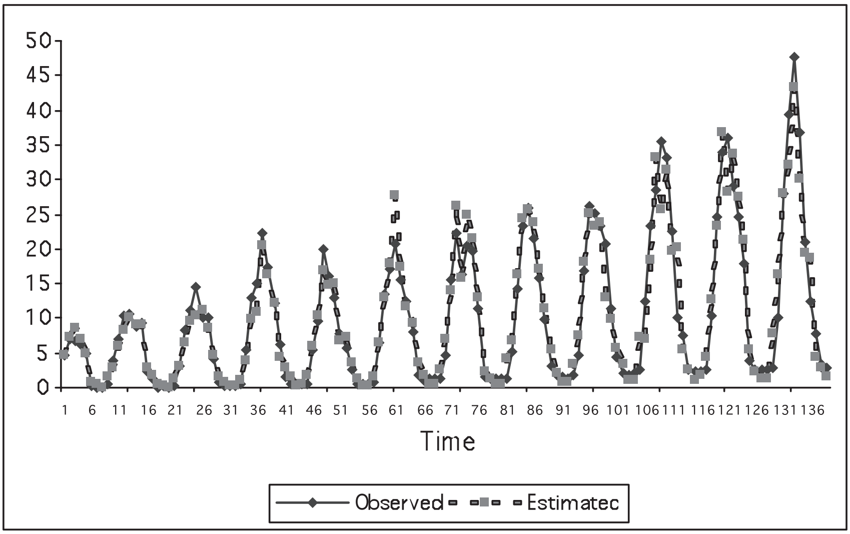

Figure 13. Actual and estimated consumption values of natural gas (x 10^6 sm^3)

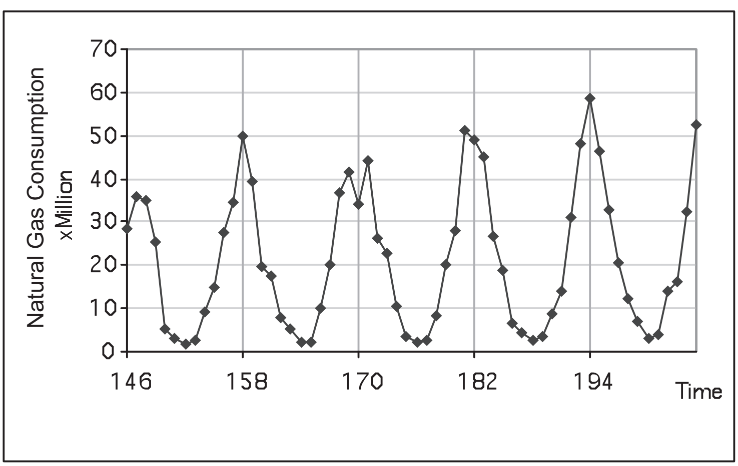

Figure 14. Future projections among the years 2009-2013 for Eskisehir

7. CONCLUSIONS AND SUGGESTIONS

Forecasting energy demand is vital for planning and investment purposes. The constructed nonlinear time series model allows us to forecast residential natural gas consumption under multiple scenarios include different values of heating degree day and consumer price index for future periods. The results of the present study are expected to give a direction to scientists and policy makers in implementing energy planning studies and in dictating the energy strategies as potential tool

Time series are analyzed to discover past patterns of variability that can be used to forecast future values and assist in the management of business operations. Time series analysis does not provide the answer to what the future holds, but it is valuable in the forecasting process and helps to reduce errors in forecasts. It is important that managers understand the past and use historical data and sound judgment to make intelligent plans to meet the demands of the future. Forecasts are made to assist management in determining alternative strategies.

Model of monthly demand serve a twofold purpose. First, the model provides short-term forecasts that will assist in the economic planning and dispatching of natural gas. Second, it provides estimates of historical peak demands under a set of alternative conditions that will support to some long term decisions such as pipeline constructions, storage capacities, and international agreements.

This analysis show that the GAs can be used as an alternative solution algorithm to estimate the future demand of energy. GAs are of particular importance when one or more of underlying assumptions in a statistical model are not satisfied. Without requiring assumptions, GAs can attain proper solutions by scanning solution space from many different starting point. Most of the existing parameter estimation methods require much auxiliary information and is not easy to handle for practical applications. In this paper, we use genetic algorithms, which are becoming an important tool for optimization problems, to estimate parameters in nonlinear regression models. The computational experiments indicate that this genetic based method is useful, effective and adaptive for nonlinear parameter estimation.

In recent years energy strategy in Turkey has focused on natural gas, which is to a large extent dependent on import and is still progressing with a limited connection and distribution system. In order to solve the imbalance between supply and demand in winter months, valid consumption forecasts are essential. It is important to use appropriate demand models which reduce forecast errors because over or under supply than necessary may lead to heavy financial losses. This benefit make a forecasting model useful and preferable in real life. In order to reinforce the growth target and social developments, it is necessary that natural gas be supplied adequately, with sustainable and competitive prices by using appropriate forecasting models. Any small reductions in forecast error may be a considerable impact on natural gas demand. Thus, it will be possible to make use of natural gas, a limited and imported energy resource, much more economically and efficiently.

REFERENCES

Abdel-Aal, R.E. (2008). "Univariate modeling and forecasting of monthly energy demand time series using abductive and neural networks" Computers&Industrial Engineering 54: 903-917.

Akay, D. and Atak, M. (2007). "Grey prediction with Rolling mechanism for electricity demand forecasting of Turkey" Energy 32: 1670-1675.

Annual Report, BOTAS (Petroleum Pipeline Corporation in Turkey). (2007). http://www.botas.gov.tr/

Aras, H. and Aras, N. (2004). "Forecasting Residential Natural Gas Demand" Energy Sources 26: 463-472

Azadeh, A. and Tarverdian, S. (2007). "Integration of genetic algorithm, computer simulation and design of experiments for forecasting electrical energy consumption" Energy Policy 35: 5229-5241.

Azadeh, A., Ghaderi, S.F. and Sohrabkhani, S. (2008). "A simulated-based neural network algorithm for forecasting electrical energy consumption in Iran" Energy Policy 36: 2637- 2644.

Bartels, R. and Fiebig, D.G. (2000). "Residential end-use electricity demand Results from a designed experiment." Energy Journal 21 (2): 51-81.

Bartels, R., Fiebig, D.G. and Nahm, D. (1996). "Regional end use gas demand in Australia." The Economic Record 72 (219): 319-331.

Bohi, D.R. (1981). Analyzing demand behavior A study of energy elasticities. Baltimore: John Hopkins Univ. Press.

Brown, R.H. (1996). "Development of artificial neural networking models to predict daily gas consumption." Am. Gas Assoc. Forecasting Rev. 5: 1-22.

Campo, R. and Ruiz, P. (1987). "Adaptive weather sensitive short term load forecast." IEEE Transactions on Power Systems 2 (3): 592-600.

Canyurt, O.E. and Ozturk, H.K. (2008). "Application of genetic algorithm (GA) technique on demand estimation of fossil fuels in Turkey" Energy Policy 36: 2562- 2569.

Canyurt, O.E., Ceylan, H., Ozturk, H.K. and Hepbasli, A. (2004) "Energy Demand Estimation Based on Two-Different Genetic Algorithm Approaches" Energy Sources 26:1313-1320.

Ceylan, H., Ozturk, H.K., Hepbasli, A. and Utlu, Z. (2005). "Estimating Energy and Exergy Production and Consumption Values Using Three Different Genetic Algorithm Approaches, Part 1: Model Development" Energy Sources 27:621-627.

Charleson, L.R. and Weber, J.E. (1993) "Energy forecasts for Western Australia 1992 - 2010." Energy Economics 15 (2): 111-122.

Chatterjee, S., Laudato, M. and Lynch, L.A. (1996). "Genetic algorithms and their statistical applications: an introduction" Computational Statistics & Data Analysis 22: 633-651.

Connor, J.T. (1996). "A robust neural network filter for electricity demand prediction." Journal of Forecasting 15 : 437-458.

Dagsoz, A.K. (1995). Degree Day Values in Turkey, National energy saving policy, Heat insulation in buildings. Istanbul: Izocam.

Dasgupta, D. and Michalewicz, Z. (1997). Evolutionary Algorithms in Engineering Applications. Springer-Verlag, Berlin.

Durmayaz, A., Kadıoglu, M. and Sen, Z. (2000). "An application of the degree-hours method to estimate the residential heating energy requirement and fuel consumption in _stanbul." Energy 25: 1245-1256.

Eltony, M.N. (1996). "Demand for natural gas in Kuwait An empirical analysis using two econometric models." International Journal of Energy Research 20 (11): 957-963.

Gallant, A.R. (1975). "Nonlinear Regression" The American Statistician. 29 (2) : 73-91.

Gen, M. and Cheng, R. (1997). Genetic Algorithms and Engineering Design. New York: John Wiley&Sons, Inc.

Goldberg, D.E. (1989). Genetic Algorithms in Search, Optimization, and Machine Learning. Boston, MA, Addison-Wesley Longman Publishing.

Gorucu, F.B., Geris, P.U. and Gumrah, F. (2004). "Artificial neural networks modeling for forecasting gas consumption" Energy Sources 26: 299-307.

Gumrah, F., Katircioglu, D., Aykan, Y., Okumus, S. and Kilincer, N. (2001). "Modeling of gas demand using degree day concept Case study for Ankara." Energy Sources 23:101-114.

Hacisalihoglu, B. (2008). "Turkey's natural gas policy" Energy policy 36 : 1867-1872.

Hagan, M.T. and Behr, S.M. (1987). "The time series approach to short term load forecasting." IEEE Transactions on Power Systems 2 (3): 785-791.

Hanke, J.E. and Reitsch, A.G. (1998). Business forecasting. Upper Saddle River, NJ: Prentice Hall Inc.

Hartman, R.S. (1979). "Frontiers in energy demand modeling." Annual Review of Energy 4: 433-466.

Herbert, J.H. (1987). "An analysis of monthly sales of natural gas to residential customers in the United States." Energy System and Policy 10 (2): 127-147.

Hill, T., O'Connor, M. and Remus, W. (1996). "Neural network models for time series forecasts." Management Science 42 (7): 1082-1092.

Hobbs, B.F., Helman, U. and Jitprapaikulsarn, S. (1998). "Artificial neural networks for short term energy forecasting Accuracy and economic value." Neurocomputing 23: 71-84.

Holland, J.H. (1975). Adaptation in Natural and Artificial Systems. Ann Arbor, MI: The University of Michigan Press.

Hubele, N.F. and Cheng, C.S. (1990). "Identification of seasonal short term forecasting models using statistical decision functions." IEEE Transactions on Power Systems 5 (1): 40-45.

International Energy Outlook (IEO2008), U.S. Department of Energy (2008). Energy Information Administration, Office of Integrated Analysis and Forecasting, http://www.eia.doe.gov/oiaf/ieo/index.html.

Kaboudan, M.A. and Liu, Q. (2004). "Forecasting Quarterly US Demand for Natural Gas" Information Technology for Economics & Management 2 (1): Paper 4.

Knowles, T.W. and Wirick, J.P. (1998). "The peoples gas light and coke company plans gas supply." Interfaces 28 (5): 1-12.

Lee, D., B. Lee and S. Chang, 1997, Genetic programming model for long-term forecasting of electric power demand, Electric Power Systems Research 40, 17-22.

Lee, R.S. and Singh, N. (1994). "Patterns in residential gas and electricity demand An econometric analysis." Journal of Business and Economic Statistics 12: 233-241.

Liu, K., Subbarayan, S. and Shoults, R.R. (1996). "Comparison of very short-term load forecasting techniques." IEEE Transactions on Power Systems 11 (2): 877-882.

Liu, L.M. and Lin, M.W. (1991). "Forecasting residential consumption of natural gas using monthly and quarterly time series." International Journal of Forecasting 7: 3-16.

Mendenhall, W. and Sincich, T.A. (1996). "Second Course in Statistics Regression Analysis." Upper Saddle River, NJ: Prentice Hall Inc.

Naggar, K. and Rumaih, K. (2005). "Electric Load Forecasting Using Genetic Based Algorithm, Optimal Filter Estimator and Least Error Squares Technique: Comparative Study" Proceedings of World Academy of Science, Engineering and Technology 6: 138-142.

Nash, J.C. and Walker-Smith, M. (1987). Nonlinear Parameter Estimation. Marcel Dekker, Inc. New York Basel

NaturalGas.org. (2008). Natural Gas Supply Association.Washington, D.C., USA. http://www.naturalgas.org/index.asp

Ozcelik, Y. and Hepbasli, A. (2006). "Estimating Petroleum Exergy Production and Consumption Using a Simulated Annealing Approach", Energy Sources Part B: Economics, Planning, and Policy 1 (3): 255 - 265.

Ozturk, H.K., Canyurt, O.E., Hepbasli, A. and Utlu, Z. (2006). "An Application of Genetic Algorithm Search Techniques to the Future Total Exergy Input / Output Estimation" Energy Sources Part A 28: 715-725.

Ozturkler, H. and Altan, S. (2008). "A genetic algorithm approach to parameter estimation in nonlinear econometric models" The journal of Social Sciences of Dumlupinar University 20: 67-76.

Pan, Z., Chen, Y. and Kang, L. (1995). "Parameter estimation by genetic algorithms for nonlinear regression" In: G.Z. Liu (Ed.)Proceedings of the International Conference on Optimization Techniques and Applications, Chengdu, China, World Scientific, Singapore 2: 946-953.

Papalexopoulos, A.D. and Hesterberg, T.C. (1990). “A regression based approach to short term system load forecasting.” IEEE Transactions on Power Systems 5 (4): 1535-1544.

Rahman, S. and Bhatnagar, R. (1988). “An expert system based algorithm for short term load forecast.” IEEE Transactions on Power Systems 3 (2): 392-399.

Reeves, C.R. (1993). Modern Heuristic Techniques for Combinatorial Problems. Blackwell Scientific Press, Oxford.

Sathyanarayan, S., Birru, H. and Chellapilla, K. (1999). “Evolving nonlinear time-series models using evolutionary programming” Proceedings of the 1999 Congress of Evolutionary Computation, July 6-9, Washington, D.C., Piscataway, NJ: IEEE, 236-243.

Sen, Z. and Kadioglu, M. (1998). “Heating degree days for arid regions.” Energy 23 (12):1089-1094.

Smith, P., Husein, S. and Leonard, D.T. (1996). “Forecasting short term regional gas demand using an expert system.” Expert Systems with Applications 10 (2): 265-273.

SPO (State Planning Office) (2008). Republic of Turkey. T.R. Prime Ministry. Medium Term Programme, 2008-2010. http://ekutup.dpt.gov.tr/program/200810i.pdf

Taylor, J.W. and Majithia, S. (2000). “Using combined forecasts with changing weights for electricity demand profiling.” Journal of the Operational Research Society 51: 72-82.

Toksari, M.D. (2007). “Ant colony optimization approach to estimate energy demand of Turkey” Energy Policy 35: 3984-3990.

TURKSTAT, Turkish Statistical Institute (2008). 1994 Based Consumer price Index Numbers. http://www.turkstat.gov.tr/

Wang, B., Tai, N., Zhai, H., Ye, J., Zhu, J. and Qi, L. (2008). “A new ARMAX model based on evolutionary algorithm and particle swarm optimization for short-term load forecasting” Electric Power Systems Research 78: 1679-1685.

Wu, B. and Chang, C. (2002). “Using genetic algorithms to parameters (d,r) estimation for threshold autoregressive models”. Computational Statistics & Data Analysis 38 :315-330.

Yamak, N. and Yamak, R. (1998). “The type and dimension of seasonality in the consumer price index.” The journal of IIBF of Uludag University 16 (2)