Estimation of Electricity Demand in Residential Sector Using Genetic Algorithm Approach

Hossein Sadeghi * & Mahdi Zolfaghari & Mohamad Heydarizade

* Hossein Sadeghi, Assistant professor, Tarbiat Modares University, Tehran, Iran  
Mahdi Zolfaghari, PhD student, Tarbiat Modares University, Tehran, Iran  
Mohamad Heydarizade, MSc of electricity restructure - Power and water University of technology

KEYWORDS
Electricity demand, Genetic algorithm
JEL: C61, C88, C41, E71

ABSTRACT
This paper aimed at estimation of the per capita consumption of electricity in residential sector based on economic indicators in Iran. The Genetic Algorithm Electricity Demand Model (GAEDM) was developed based on the past data using the genetic algorithm approach (GAA). The economic indicators used during the model development include: gross domestic product (GDP) in terms of per capita and real price of electricity and natural gas in residential sector. Three forms of GAEDM were developed to estimate the electricity demand. The developed models were validated with actual data, and the best estimated model was selected on base of evaluation criteria. The results showed that the exponential form had more precision to estimate the electricity demand than two other models. Finally, the future estimation of electricity demand was projected between 2009 and 2025 by three forms of the equations; linear, quadratic and exponential under different scenarios.


1. Introduction
Electricity energy is a vital input for the economic and social development of any country. Strong population growth and rapid urbanization in Iran have played an important role in electricity energy consumption. Iran's economy has achieved an average annual growth rate of 5% over the past 20 years. In the long term, Iran's economy has been slated for robust, albeit sometime erratic, growth and persistent inflation, which will need to be supported by steadily increasing energy supplies. Iran depends heavily on the production of electricity, especially for residential sector, having to electricity consumption in residential sector almost one thirds of its total consumption. Therefore, sufficient and secure electricity supplies are the top priority of Iran's energy policy. Countries like Iran should plan carefully about their energy demand for critical periods, such as economic crises that usually hit Iran. The electricity energy demand of Iran in residential sector has had an increasing trend in the recent years depending on the gross domestic product (GDP) and real price of electricity and natural gas in residential sector Poorazarm (2005). Estimating the future electricity energy demand is important to calculate the cost of energy investment projects and energy production. Per capita consumption of electricity in residential sector, GDP and population have increased by 1.54, 2.8 and 1.4 times in the last 20 years, respectively; when compared with the 1988 figures.
The ordinary least square (OLS) model by Lotfalipoor (2004) as well as the ARDL model by Poorazarm (2005) and Mehran et al. (2010) have been developed in order to estimate the electricity demand based on the variations of GDP and the real price of electricity and natural gas in residential sector. However, those electricity estimation models are in linear form and cannot estimate the electricity demand in non-linear form.

The estimation of electricity demand based on economic indicators may be modeled using various equations. These equations may be linear or non-linear. Non-linear form of these equations can better estimate the future electricity demand of Iran due to the fluctuations of the economic indicators (Delavae et al., 2011).

For solving the non-linear equations, a genetic algorithm (GA) approach was proposed. The GA can be used to estimate the future electricity demands under different scenarios by appropriately estimating the weighting parameters with the current data. The available data is partly used for finding the optimal, or near optimal, values of the weighting parameters and partly for testing the model. The obtained results are compared with together based on the criteria of MSE, RMSE, MAE and MAPE. GAs are a family of adaptive search procedures that are loosely based on the models of genetic changes in a population of individuals (Canuyrt and Ozturk, 2008). GAs, as elucidated by Goldberg (1989) have attracted growing interest for optimization problems. The main advantage of GAs is their ability to use the accumulating information about the initially unknown search space in order to bias the subsequent searches into useful sub-spaces (Saberi et al., 2011). GAs differ from conventional non-linear optimization techniques in that they search by maintaining a population (or database) of solutions from which better solutions are created, rather than making incremental changes to a single solution to the problem (Li and Su, 2010).

The key feature of a GA is the manipulation of a population whose individuals are characterized by possessing a chromosome. This latter can be coded as a string of characters of given length \( L \). Each string represents a feasible solution to the optimization problem. A chromosome is composed of strings of symbols called bits (in this case, binary). Each bit is attached to a position within the string representing the chromosome to which it belongs. If, for example, the strings are binary, then each bit can take the value of 0 or 1. The link between the GA and the problem at hand is provided by the fitness function (F), which establishes a mapping from the chromosomes to some set of real numbers. The greater the F, the better is the adaptation of the individual.

The procedure is generative, and makes use of three main operators: reproduction, crossover and mutation. Each generation of a GA yields a new population from an existing one. The \( P_C \) individuals are assigned allelic values to their chromosomes, where the assignment can be either deterministic or random. Reproduction is a process that selects the fit individual strings according to some selection operators. The selection operator is responsible for choosing the members that will be allowed to reproduce during the current generation. These members are selected on the basis of their fitness (F) values and the fit individuals are passed onto the future generations.

Further manipulation is conducted by crossover and mutation operators before the replacement is actually done in the view of the next cycle. Crossover is to provide a mechanism for the exchange of chromosomes between the mated parents. The mated parents may then create a child with a chromosome set, which some mix of the parent’s chromosomes. For example, Parent #1 has chromosomes ‘abcde’, while Parent #2 has chromosomes ‘ABCD’. One possible chromosome set for the child is ‘abcDE’, where the position between the ‘c’ and ‘D’ chromosomes is the crossover point. Mutation is a background operator, which produces spontaneous random changes in various chromosomes. A simple way to achieve mutation would be to alter one or more genes.

The mutation operator serves a crucial role in the genetic algorithms, either by (a) replacing the genes lost from the population during the selection process so that they can be tried in a new generation or (b) providing the those genes that are not present in the initial population. A mutation process is a small probability that (after crossover) one or more of the child’s chromosomes will be mutated, e.g. the child ends up with ‘abcDF’.

The function of this operator is to prevent the child being trapped at bad local optima over the range of the population during the generation.

Apart from the main operators described so far, the other operator used in this paper is the elitism operator, which is used to ensure that the chromosome of the best parent generated to date is carried forward into the next generation.

Upon the generation of the new population, the GA checks to see if the best parent has been replicated. If not, then a random individual is chosen and the chromosome set of the best parent is mapped into that individual. In this paper, a genetic algorithm electricity demand model (GAEDM) has been developed for estimating Iran’s electricity demand in residential sector based on the GDP, population and the real price of electricity and natural gas in residential sector. Three forms of the equations; linear, quadratic and exponential, were employed.

2. GA Application of Energy Demand

GAs are mathematical tools with a wide range of applications. They are very efficient in optimization of the problems, especially when the respective objective functions are discontinuous and exhibit many local optima. Hence, they have been becoming popular in a
field of engineering problems. The non-linearity of the economic indicators and electric energy demand leads to investigate the different form of solution approaches to a problem such as genetic algorithm. Estimation of the future energy is very important due to the make correct investments.

Many researches have recently been carried out to estimate the energy demand using GA approach. Goldberg (1989) was the first who explained the theoretical background, application and framework for GAs. Since then, many textbooks have been written on GA (Gen and et al, 1997, Michalewicz, 1999). The basic definitions of GAs can be obtained in Ceylan (2000). The main of this study is to estimate the Turkish electricity demand and is to show if the GA approach is a useful tool for this application. Therefore, various forms of the mathematical expressions are used, while genetic algorithm electricity demand model (GAELDM) and genetic algorithm industrial electricity demand model (GAINEELDM) are proposed to estimate total electricity and industrial electricity demand of Turkey. Ozturk et al. (2005) estimated industrial electricity demand using GA. Haldenbilen and Ceylan (2004) estimated transport energy demand in Turkey by GA. Stach et al. (2005) have proposed a novel learning method that is able to generate fuzzy cognitive maps models from input historical data and without human intervention, the proposed method is based on GA. Muni et al. (2006) proposed genetic programming methodology simultaneously selects a good subset of features and constructs a classifier using the selected features. Ozturk et al. (2004) have carried out some researches recently to estimate the energy input/output values using GA. Hasheminia and Akhavan Niaki (2006) have introduced a new type of GA to find the best regression model among several alternatives and have assessed its performance by an economical case study. Also, recent studies show the integration of GA and neural networks for short-term estimation and prediction of electrical energy consumption (Azadeh et al., 2007b). A.Azadeh et al. (2007) presented an Integration of artificial neural networks and genetic algorithm to predict electrical energy consumption in Iranian agriculture sector from 1981 to 2005. S. Tarverdian et al. (2007) have proposed an integrated algorithm for forecasting monthly electrical energy consumption based on genetic algorithm (GA), computer simulation and design of experiments using stochastic procedures. Canyurt and Ozturk (2008) developed some Genetic algorithm demand estimation models to estimate the future coal, oil and natural gas demand values based on population, gross national product, import and export figures. M.Saberi et al. (2011) presented an adaptive network based fuzzy inference system–genetic algorithm clustering ensemble algorithm for performance assessment, and improvement of conventional power plants. K.Li and H.Su(2010) forecasted the building energy consumption by using hybrid genetic algorithm-hierarchical adaptive network-based fuzzy inference system. The main parameters of GAs and their corresponding applications for this study can be defined as below:

2-1. Representation of the GAEDM Coefficients and Decoding

The representations of the coefficients are the main building blocks of GAs, which are made of the GAEDM parameters. The decision variables in the objective function go through an encoding process. The encoding process is executed using binary bit strings. One of the basic features of GAs is that they work on the coding space and the solution space, alternately. Genetic operations work on the coding space (chromosomes), while evaluation and selection processes work on the solution space. Natural selection is the link between the chromosomes and the performance of their decoded solutions. Let's have a glance on the GA notion is applied to the energy demand problem. The encoding and mapping of the coded regional coefficients to the real values are performed as follows:

Let $\psi = (w_1,w_2,...,w_{n})$ be the weighting parameters of the GAEDM. Then, if eight bit binary variables are selected to represent the weighting parameters of the GAEDM, the binary bit string representing the solution would be as follows (assuming that all bits in the solution are zero),

$X = \begin{bmatrix} 00000000 \mid 00000000 \mid 00000000 \mid 01010110 \end{bmatrix}$

weighting parameters, $\Phi_i$ is the integer resulting from base two-arithmetic, $z$ is the total number of parameters (i.e. coefficients) and $I_i$ is the number of binary bits per weighting parameter.

2-2. The Fitness Function

The genetic algorithm works according to the selection rules as defined by the laws of evolutionary genetics.
The model seeks the “fittest” model to the observed values. In energy demand estimation, it is the model whose parameters, when input to the source (i.e., the GDP and the real price of electricity and natural gas) model, produce the energy demand that best matches the measured demand. Given the expected range of the coefficients of the GAEDM, the algorithm randomly selects a set of models and proceeds to evolve them to produce best and fitness models.

The fitness function, $F(x)$, is calculated in the fitness subroutine, which contains the GDP and the real price of electricity and natural gas data for this problem. Then, the fitness function (i.e., the minimum sum of squared errors (SSE)) takes the following form:

$$
\text{Max}F(x) = \frac{1}{\sum_i s_i(E_{\text{obs}} - E_{\text{est}})^2}, i = 1, 2, ..., m, \quad (2)
$$

Where, $E_{\text{obs}}$ and $E_{\text{est}}$ are the observed and estimated energy demand, respectively, $m$ is the number of observations and $s_i$ is the weighting factor.

The GA picks the fittest members of the population based upon the maximum fitness value, and they effectively converts the fitness function to a continuously increasing function, such that the GA seeks the model with the largest value of $F(x)$.

The general structure of the GA optimization model is described in Fig. 1. The input parameters, which are user-specified, include the probability of crossover ($p_c$), probability of mutation ($p_m$), population size ($n$), generation number ($k$), number of decision variables for the problem ($z$) and the number of possibilities ($l$) of the decision variables.

Estimation of the coefficients for the GAEDM can be derived by the following three loops:

1. (1) The outer loop processes the mutations on the selected chromosomes and proceeds to the next generation;
2. (2) The middle loop processes the selection and crossover operators on the chromosomes; and
3. (3) The inner loop evaluates the chromosomes based on the fitness function and plays the most important role in the GA optimization model.

The inner loop, called the evaluation process, is where all the coefficients of the GAEDM are represented as binary bit strings (i.e., chromosomes). The decoding and mapping of the decision variables are described in Ceylan. The GA’s inner loop evaluates each chromosome’s fitness based on Eq. (2).

**Input:** User-specified GA parameters (e.g., probability of crossover and mutation, population size, generation number, number of decision variables and number of possibilities for each decision variable)

**Fig. 1.** The GAEDM optimization model.
The inner loop runs until the user-specified population size is completed. The set of chromosomes and the corresponding fitness values are then processed by the middle loop for the selection and crossover processes. The reproduction process is performed for each population until the user-specified population is completed.

The fittest chromosome is selected in the middle loop. The outer loop is the new generation loop, where the newly selected chromosomes (after mutation) are fed into the next generation. The process continues until the user-specified generation number is completed.

3. Development and Application of the GAEDM

The data were collected from different sources. The GDP was collected from the Central Bank of Iran (CBI). The price of electricity and natural gas were obtained from the Iran Statistics Center (ISC). Fig. 2 shows the observed data between 1969 and 2008.

![Fig. 2. Residential electricity consumption of Iran and various economic indicators in real values](image)

The three forms of the GAEDM can be fitted as follows:

The linear form:

\[ E_{\text{lin}} = w_1 X_1 + w_2 X_2 + w_3 X_3 + w_4 X_4 + \cdots \]

The quadratic form:

\[ E_{\text{quad}} = w_1 X_1^2 + w_2 X_2^2 + w_3 X_3^2 + w_4 X_4^2 + \cdots \]

And the exponential form:

Where, \( E \) is the per capita consumption of electricity in residential sector, \( X_1 \) is per capita GDP, \( X_2 \) and \( X_3 \) are the real prices of electricity and natural gas in residential sector, respectively and \( w_1, w_2, \ldots, w_n \) are the weighting parameters. Goldberg (1989) suggested the appropriate values of the user-specified GA parameters as: population size between 50 and 1000, \( p_c \) is 0.5–0.6 and \( p_m \) is 0.01–0.02. Table 1 shows the user-specified parameters, which were selected for the GAEDM model in the present work. After application of the linear, quadratic and exponential forms of the GAEDM, Eqs. (5), (6) and (7) were obtained:

\[ E_{\text{lin}} = w_1 X_1 + w_2 X_2 + w_3 X_3 + w_4 X_4 + \cdots \]

\[ E_{\text{quad}} = w_1 X_1^2 + w_2 X_2^2 + w_3 X_3^2 + w_4 X_4^2 + \cdots \]

\[ E_{\text{exponential}} = w_1 e^{X_1} + w_2 e^{X_2} + w_3 e^{X_3} + w_4 e^{X_4} + \cdots \]
The convergence of the GAEDM can be obtained in Ceylan (2000). Thirty three data (1969–2001) were used to estimate the weighting parameters of the GAEDM and the remaining data (2002–2008) were used to validate the model. The MSE\(^2\), RMSE\(^3\), MAE\(^4\) and MAPE\(^5\) criteria were applied validate the estimated and observed data of linear, quadratic and exponential equations (Table 2).

### Table 2. Comparison of the estimated and observed data of linear, quadratic and exponential equations (2002-2008)

<table>
<thead>
<tr>
<th>Model</th>
<th>MSE</th>
<th>RMSE</th>
<th>MAE</th>
<th>MAPE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear</td>
<td>11751.7</td>
<td>108.4</td>
<td>90.07</td>
<td>2.82</td>
</tr>
<tr>
<td>Quadratic</td>
<td>3295.8</td>
<td>57.40</td>
<td>44.7</td>
<td>1.40</td>
</tr>
<tr>
<td>Exponential</td>
<td>835.43</td>
<td>28.9</td>
<td>26.04</td>
<td>.79</td>
</tr>
</tbody>
</table>

The GAEDM results showed that the exponential form had more precision to estimate the electricity demand than other models on the basis of each evaluation criterion, because the exponential form showed minimum error compare with the other two forms.

### 4. The GAEDM Estimation and Scenarios

Estimation of Iran’s electricity demand in residential sector in the years 2009–2025 is based on four different scenarios as follows:

Scenario 1: It is assumed that the average growth rate of GNP is 8% (on the basis of the Iran's 20-Year Perspective Document). The population growth rate is 0.15% and the real price of electricity and natural gas are on the basis of the average growth rate during the period of 1969-2001 (4.01 and -4.03, respectively).

Scenario 2: It is assumed that the average growth rate of GNP is 6.8 % (on the basis of the average growth rate during the period of 1996-2001). The population growth rate is 0.15% and the real price of electricity and natural gas are constant.

Scenario 3: It is assumed that the average growth rate of GNP is 5.3% (on the basis of the average growth rate during the period of 1996-2001 excluding the war period).

The population growth rate is 0.15% and the real price of electricity and natural gas increase with the growth rate of 10% ( It is assumed that subsidy of electricity and natural gas would be decreased).

Scenario 4: It is assumed that the average growth rate of GNP is 2.9% (on the basis of the average growth rate during the period of 1996-2001).

The population growth rate is 0.15% and the real price of electricity and natural gas increase with the growth rate of 15% (It is assumed that subsidy of electricity and natural gas would be decreased).

The results of estimated per capita consumption under four scenarios for the period of 2009-2025 are reported in the Appendix1. Per capita consumption in all scenarios and on base of 3 forms will increase. Table 4 shows the average growth rate of electricity consumption in three periods of ‘2009-2014’, ‘2015-2020’ and ‘2021-2025’ on the basis of four scenarios. In all scenarios, the average growth rate of electricity demand increases.

This indicator for the linear form is steady in the Scenario 1 to 3 and decreases relatively in Scenario 4. This indicator for the Quadratic form is similar linear for Scenario 1 to 3 but if increases relatively in the Scenario 4. Changes for the Exponential form are similar to the Quadratic form.

It is worth noting that the pricing policy cannot be effective to control the electricity demand.

---

\(^1\) Mean Square Error  
\(^2\) Root Mean Square Error  
\(^3\) Mean Absolute Error  
\(^4\) Mean Absolute Percent Error
5. Conclusion

Knowledge about electricity consumption in each period is necessary to exact planning for main performance. This subject is more important in residential sector, because of its large portion. Therefore, examination of effective factors on electricity energy demand for the exact and right recognition of consumption behavior structure in order to accurate planning following attaining to specific aims is necessary. In this paper, we estimated the electricity demand function in residential sector in the form of Linear, Quadratic and Exponential equations using GA approach, and selected the best estimated model on the basis of evaluation criteria. Residential electricity consumption was predicted in several Scenarios until 2025. The results show that the exponential model enjoys more precision to estimate of the electricity demand compared with the other models.

References:


**Appendix 1:**

**Estimated per capita consumption of electricity in residential sector for the period of 2009-2025 (KWh)**

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Senario1</th>
<th>Senario1</th>
<th>Senario1</th>
<th>Senario1</th>
</tr>
</thead>
<tbody>
<tr>
<td>period</td>
<td>Lin</td>
<td>Quad</td>
<td>Exp</td>
<td>Lin</td>
</tr>
<tr>
<td>2009</td>
<td>2900.945</td>
<td>2958.954</td>
<td>2883.81</td>
<td>2906.932</td>
</tr>
<tr>
<td>2010</td>
<td>3106.315</td>
<td>3104.298</td>
<td>3601.466</td>
<td>3106.289</td>
</tr>
<tr>
<td>2011</td>
<td>3261.554</td>
<td>3259.512</td>
<td>3149.344</td>
<td>3261.516</td>
</tr>
<tr>
<td>2012</td>
<td>3472.133</td>
<td>3425.267</td>
<td>3308.103</td>
<td>3472.282</td>
</tr>
<tr>
<td>2013</td>
<td>3604.366</td>
<td>3602.278</td>
<td>3478.446</td>
<td>3604.305</td>
</tr>
<tr>
<td>2014</td>
<td>3793.419</td>
<td>3791.309</td>
<td>3661.127</td>
<td>3793.347</td>
</tr>
<tr>
<td>2015</td>
<td>3995.308</td>
<td>3993.177</td>
<td>3856.95</td>
<td>3995.225</td>
</tr>
<tr>
<td>2016</td>
<td>4210.903</td>
<td>4208.753</td>
<td>4066.778</td>
<td>4210.81</td>
</tr>
<tr>
<td>2017</td>
<td>4441.137</td>
<td>4438.968</td>
<td>4291.53</td>
<td>4441.034</td>
</tr>
<tr>
<td>2018</td>
<td>4687.002</td>
<td>4684.815</td>
<td>4532.192</td>
<td>4684.916</td>
</tr>
<tr>
<td>2019</td>
<td>4949.561</td>
<td>4947.357</td>
<td>4789.816</td>
<td>4949.44</td>
</tr>
<tr>
<td>2020</td>
<td>5229.946</td>
<td>5227.726</td>
<td>5065.528</td>
<td>5228.616</td>
</tr>
<tr>
<td>2021</td>
<td>5552.369</td>
<td>5557.134</td>
<td>5360.380</td>
<td>5557.231</td>
</tr>
<tr>
<td>2022</td>
<td>5840.121</td>
<td>5846.872</td>
<td>5676.408</td>
<td>5844.975</td>
</tr>
<tr>
<td>2023</td>
<td>6190.584</td>
<td>6188.321</td>
<td>6013.63</td>
<td>6188.443</td>
</tr>
<tr>
<td>2024</td>
<td>6555.231</td>
<td>6552.955</td>
<td>6374.57</td>
<td>6555.097</td>
</tr>
<tr>
<td>2025</td>
<td>6944.637</td>
<td>6942.349</td>
<td>6704.492</td>
<td>6944.469</td>
</tr>
</tbody>
</table>

International Journal of Industrial Engineering & Production Research, March 2011, Vol. 22, No. 1