

Electrical Energy Consumption Estimation by Genetic Algorithm and Analysis of Variance

A. Azadeh, R. Tavakkoli-Moghaddam, S. Tarverdian

Research Institute of Energy Management and Planning and Department of Industrial Engineering,
Faculty of Engineering, University of Tehran, Iran
Phone number: 0098 21 66461990, e-mail: aazadeh@ut.ac.ir, sanaztn@yahoo.com

Abstract- This study presents a genetic algorithm (GA) with variable parameters to forecast electricity demand in agricultural, low energy consuming and energy intensive sectors using stochastic procedures. The economic indicators used in this paper are price, value added, number of customers and consumption in the last periods for agricultural and low energy consuming sectors and price, value added, number of customers, price of the substitute fuel and energy intensity in energy intensive sector. Three kinds of models; linear-logarithmic, exponential and quadratic are used to find which leads us to minimum error for the related sector. The GA applied in this study has been tuned for all its parameters and the best coefficients with minimum error are identified, while all parameter values are tested concurrently. The estimation errors of genetic algorithm models are less than that of estimated by regression method. Finally, analysis of variance (ANOVA) is applied to compare genetic algorithm (three models), regression and actual data. It is found that at $\alpha = 0.05$ the five treatments are not equal and therefore Duncan test is applied to see which treatment pair has led to the rejection of null hypothesis. Furthermore it is shown that genetic algorithm estimation is closer to actual data with less MAPE (Mean Absolute Percentage Error) error than that of estimated by regression. The data from 1979 to 2003 is used to forecast electricity consumption in the aforementioned sectors as the case study.

Key words

Electricity estimation, Genetic algorithm, Analysis of variance

1. INTRODUCTION

According to the increasing demand of electricity, demand function estimation is really essential. The estimation of electrical energy demand based on economic indicators may be done with different kinds of mathematical models. These equations might be linear or non-linear. Due to the fluctuations of economic indicators, the non-linear forms of the equations can estimate electrical energy demand more effectively. The non-linearity of economical indicators and electrical energy demand has led to search for different solution approach methods of genetic algorithms. Genetic algorithms are optimizing and stochastic search techniques which possess vast and powerful applications.

The estimation of Turkey's energy demand based on economic indicators using genetic algorithm was reported by Ceylan and his colleagues in 2003 [1]. Ozturk et al. estimated electricity demand of Turkey using genetic algorithm [2]. Some researches have been carried out recently to estimate the energy consumption using

genetic algorithm [3],[4],[5]. Osman et al. presented a combined GA –fuzzy logic controller technique for constrained nonlinear programming problems so that the search region is able to adapt towards the promising area [6]. Bunning and Sun presented a stochastic global algorithm for solving constrained optimization problems over a compact search domain [7]. Hashemina & Niaki have introduced a new type of genetic algorithm to find the best regression model among several alternatives and have assessed its performance by an economical case study [8]. Tang, Quek and Ng) have used a genetic algorithm based Takagi-Sugeno-Kang fuzzy neural network to tune the parameters in Takagi-Sugeno-Kang fuzzy neural network [9].

In this paper, the GA is used to estimate logarithmic-linear, exponential and quadratic models which are the function of independent variables; price, value added, number of users, consumption in the previous periods, price of the substitute fuel and electricity intensity. Electricity consumption in the agricultural, low energy consuming industries and energy intensive one sector has had an increasing pattern in the past 25 years (1979-2003). "Fig. 1.a, 1.b, 1.c" show electricity consumption in Iranian agricultural, energy intensive and low energy consuming industries sectors from 1979 to 2003.

2. GENETIC ALGORITHM

GA is similar to the natural evolution process where a population of a specific species adapts to the natural environment under consideration, a population of designs is created and then allowed to evolve in order to adapt to the design environment under consideration. These algorithms were directly described by Goldberg in 1989 [10] and have taken attention to solve optimizing problems. The most important advantage of the GAs is their ability to use accumulative information about the initial unknown search space in order to move the next searches in to useful spaces [1].

GA differs from conventional non-linear optimizing techniques as by preserving a population of the solutions, they search for better ones. The key feature of such algorithms characterized by possessing a chromosome.

This latter can be coded as a string of characters of given length l . The link between the GA and the problem at hand is provided by the fitness function (F).

The main genetic operators are applied based on fitness function evaluation [1]. A group of chromosomes is called a population. One of the genetic features is that

instead of focusing on one point of the search space of a chromosome, it works on a population of chromosomes. If the number of the chromosomes is too low the possibility of the movement operation by genetic algorithm will be also low and it only searches small part of the search space. According to the researches a suitable population size is about 20 to 30 chromosomes. Of course sometimes a population with 50 to 100 has lead to best answers [10]. The GA works with a 'population' of possible answers (e.g. sets of parameter values).

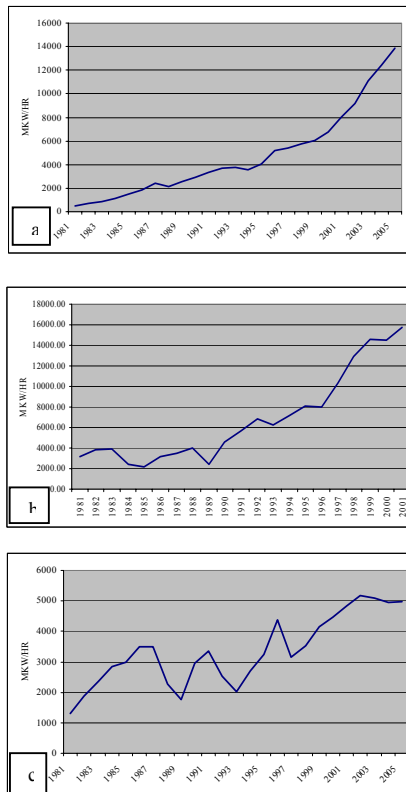


Fig. 1. (a) Electricity consumption in agricultural sector during 1979-2003. (b) Electricity consumption in energy intensive industries sector during 1979-2003. (c) Electricity consumption in low energy consuming industries sector during 1979-2003.

Because of this, it does not require initial estimates of the fitting parameters, but requires only the allowable range of each parameter [11]. GA encompasses three main operators: 1) selection, 2) crossover and 3) mutation described in the next sections.

A. Selection

Selection operator searches according to the fitness of the members based on fitness values. The fitness value of the *i*th member in the population can participate in this operation on the basis of probabilities. This probability of the *i*th member in the population is calculated as follows in which *n* is the population size. In the selection

operation the members of the population with better fitness can participate several times while the

$$P_i = F_i / \sum_{k=1}^n F_k$$

Members with worse fitness may be deleted in order to obtain a larger fitness average.

B. Crossover

Crossover operation allows an exchange of the design characteristics between two mating parents. This operation is done by selecting two mating parents on which two random places are selected on each chromosome string and the strings between these two places among the mates are exchanged. A presentation of the crossover is shown below:

```
Parent1= 1010101011
Parent 2= 1001000111
Child 1= 1010000111
Child 2= 1001101011
```

The crossover operation is applied with a probability of p_c which takes the probabilistic values from 0.2 to 0.8.

C. Mutation

Mutation operator is another essential operator in genetic algorithm process and it acts on each chromosome after crossover operator such that a random number is produced for each bite of a chromosome, if this number is smaller than p_m , mutation will occur in that bite and otherwise it does not occur. According to the latest research, p_m shows the best value while varying between 1 and 5%. Mutation operation prevents losing unexpected valuable genetic information in the population during selection and crossover operation. This operator acts at a random place of a chromosome with a low probability of p_m [2, 3].

D. The Fitness Function

To introduce fitness function, the variables should be put in the model and then the difference between estimated values and actual data for each individual should be calculated. In each generation the individual with minimum difference must be returned. Individual parameters are selected randomly and after placing them in the model the fitness function is calculated. The fitness function is shown below, where D_{actual} and $D_{estimated}$ are actual and estimated energy demand, respectively and *m* is the number of observations.

$$\min f = 1 / n \sum_j^m (D_{actual} - D_{estimated}) / D_{actual}$$

E. Tuning Genetic Algorithm

For the chosen genetic algorithm operators; crossover and mutation, and also the number of generations and the size of the algorithm population, tuning was done in which all these parameters were tested for different rates altogether. Finally, the best model was obtained yielding minimum relative error.

F. Conventional Regression

The linear–logarithmic model is estimated by OLS (conventional) method shown in equation (1) to (3). Which according to [12], [13] is the best known model in Iran to predict electrical energy consumption by conventional regression. And due to this reason all resulted model by GA in the next section is compared only to linear-logarithmic model. The estimated equations by this method are brought in equations (1) to (3) for agricultural, low energy consuming and energy intensive sectors sequentially.

$$LnEC_{agr} = 5.29 - 0.13Ln(PR_{agr}) + Ln(NC_{agr}) \quad (1)$$

$$LnEC_{low} = -1236 - 0.62Ln(PR_{low}) + 0.41Ln(NC_{low}) + 0.81Ln(VA_{low}) + [MA3] = -0.948 \quad (2)$$

$$LnEC_{high} = 0.42Ln(PR_{high}) + 1.54Ln(NC_{high}) + 0.77Ln(ET_{high}) + [MA4] = -0.83 \quad (3)$$

According to the above equations in agricultural sector, consumption is the function of the price and the number of customers while in low energy consuming industries, consumption is function of price, number of customers and value added, energy intensive sector consumption is the function of the price of substitute fuel, number of customers and electricity intensity. In the above equations MA refers to the error of the previous periods.

3. ESTIMATING PARAMETER VALUES USING GENETIC ALGORITHM

In this study linear-logarithmic, exponential and quadratic models of electrical energy demand is considered to predict electrical energy consumption in the agricultural, low energy consuming and energy intensive sectors as different kind of mathematical equations can be estimated by genetic algorithm approach. Equation (4) as an example presents the aforementioned kinds to know which model is the most suitable one to predict electricity consumption in each sector. For each sector, each equation is applied. The reason of using exponential and quadratic models is because that several papers reported lower errors using the aforementioned models [1], [2], [4], [5].

The usual way of estimating model parameters is to use data partially; one series (at least 15 years) is to estimate the parameters and saving the reminders for testing purpose. The testing procedure is to obtain minimum relative error between estimated and actual values. Model parameters are generated randomly. Electricity demand functions in agricultural, low energy consuming and energy intensive sectors of Iran are obtained for 21 years (1979-2003) [14]. To identify the best fitness, the required parameters on GA algorithm are as follows:

- Population size (n): 100
- Iterations (number of the generation): 200
- Mutation rate: 0.04

- Crossover rate: 94%

$$\begin{cases} LnEC_{agr} = \alpha_0 + \alpha_1 Ln(PR_{agr}) + \alpha_2 Ln(NC_{agr}) + \alpha_3 Ln(VA_{agr}) + \alpha_4 Ln(EC(-1)_{agr}) & (4) \\ EC_{low} = \alpha_0 + \alpha_1 (PR_{agr})^{\alpha_2} + \alpha_3 (NC_{agr})^{\alpha_4} + \alpha_5 (VA_{agr})^{\alpha_6} + \alpha_6 (EC(-1)_{agr})^{\alpha_6} \\ EC_{high} = \alpha_0 + \alpha_1 (PR_{agr}) + \alpha_2 (NC_{agr}) + \alpha_3 (VA_{agr}) + \alpha_4 (EC(-1)_{agr}) + \alpha_5 (PR_{agr})(NC_{agr}) \\ + \alpha_6 (PR_{agr})(VA_{agr}) + \alpha_7 (PR_{agr})(EC(-1)_{agr}) + \alpha_8 (NC_{agr})(VA_{agr}) + \alpha_9 (NC_{agr})(EC(-1)_{agr}) \\ + \alpha_{10} (VA_{agr})(EC(-1)_{agr}) + \alpha_{11} (PR_{agr})^2 + \alpha_{12} (NC_{agr})^2 + \alpha_{13} (VA_{agr})^2 + \alpha_{14} (EC(-1)_{agr})^2 \end{cases}$$

After applying GA on equation (4) for each sector, the resulted models which yield in less MAPE error are reported in equations (5) to (7):

$$LnEC_{agr} = 0.3 + 0.02ln(PR_{agr}) + 0.3ln(NC_{agr}) + 0.15ln(VA_{agr}) + 0.69ln(EC(-1)_{agr}) \quad (5)$$

$$EC_{low} = 1.15 - 0.81(PR_{low}) + 0.08(NC_{low}) + 0.83(EC(-1)_{low}) + 0.05(PR_{low})^2 \quad (6)$$

$$LnEC_{high} = -0.21 - 0.13(PR_{high}) + 0.94ln(VA_{high}) + 1.63ln(ET_{high}) \quad (7)$$

According to the above equations the most suitable model from the point of yielding minimum error for agricultural and energy intensive industries is the linear-logarithmic model while for low energy consuming industries is the quadratic form. In equation (6), coefficients $\alpha_4, \alpha_6, \alpha_7, \alpha_8, \alpha_9, \alpha_{10}, \alpha_{12}, \alpha_{13}, \alpha_{14}$ and in equation (7), coefficients α_3 and α_4 were found to be zero.

21 years data were used to estimate the model parameters and 4 years data were saved for testing purpose. The relative error between estimated values by GA and regression values versus actual ones are reported in Table I. a, b, c.

"Fig. 2" investigates the two approaches; genetic algorithm and conventional regression versus actual data in each sector. As shown, estimation done by GA is more fitted than regression method after the place which is circled. In other words, for the four year data in which the GA model has not been taught, it resulted in better fitness when compared with actual data. This result can be seen in Fig.2.a,b and c.

TABLE I.A- The Relative Error Between Estimated And Actual Values In Agriculture Sector

| year | Actual data | GA(agr) line-log | Relative error | Regression | Relative error |
|------|-------------|---------------------|----------------|--------------------|----------------|
| 2000 | 9147 | 9210 | 0.0069 | 9135 | 0.001 |
| 2001 | 11079 | 10821 | 0.0232 | 12207 | 0.101 |
| 2002 | 12435 | 12815 | 0.03 | 14018 | 0.127 |
| 2003 | 13859 | 14588 | 0.052 | 16788 | 0.211 |
| | | Average error=0.027 | | Average error=0.11 | |

TABLE I.B-The Relative Error Between Estimated And Actual Values In Low energy consuming Industries Sector

| year | Actual data | Regression | Relative error | GA(low) _{quad} | Relative error |
|------|-------------|--------------------|----------------|-------------------------|----------------|
| 2000 | 5164 | 5941 | 0.15 | 4807.448 | 0.069 |
| 2001 | 5099 | 5219 | 0.02 | 5154.588 | 0.010 |
| 2002 | 4950 | 2864 | 0.42 | 4971.715 | 0.004 |
| 2003 | 4984 | 3181 | 0.36 | 4828.62 | 0.031 |
| | | Average error=0.23 | | Average error=0.028 | |

TABLE I.C-The Relative Error Between Estimated And Actual Values In Energy Intensive Industries Sector

| year | Actual data | GA(high) line-log | Relative error | Regression | Relative error |
|------|-------------|----------------------|----------------|--------------------|----------------|
| 2000 | 12942 | 13247 | 0.023 | 6927 | 0.46 |
| 2001 | 14556 | 14089 | 0.032 | 6576 | 0.54 |
| 2002 | 14470 | 14984 | 0.035 | 8266 | 0.42 |
| 2003 | 15774 | 15917 | 0.009 | 5642 | 0.64 |
| | | Average error=0.025 | | Average error=0.52 | |

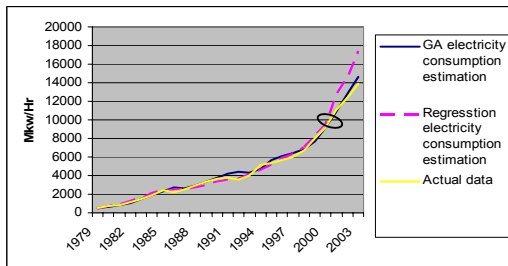


Fig. 2.a. GA and regression estimation versus actual data in Iranian agriculture sector

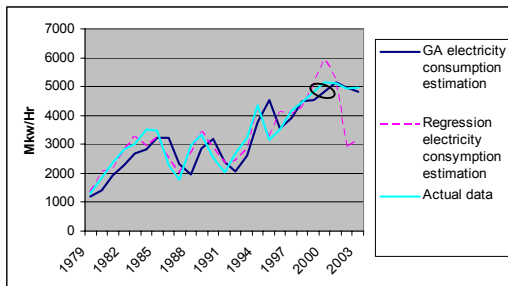


Fig. 2.b. GA and regression estimation versus actual data in Iranian low energy consuming sector

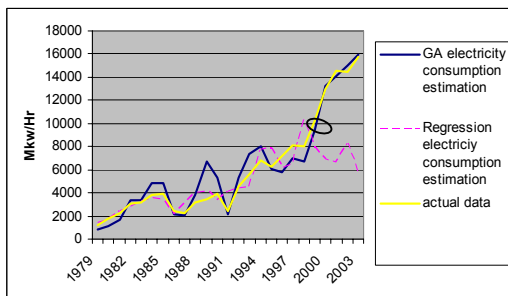


Fig. 2.c. GA and regression estimation versus actual data in Iranian energy intensive sector

4. Analysis of Variance

The estimated results of GA, regression and actual data are compared by ANOVA F-test. The experiment should be designed such that variability arising from extraneous sources can be systematically controlled. Time is the common source of variability in the experiment that can be systematically controlled through blocking [15]. Therefore a blocked design of ANOVA may be applied. The hypothesis is:

$$H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5 \quad (8)$$

$$H_1: \mu_i \neq \mu_j \quad i, j = 1, 2, 3, 4, 5 \text{ and } i \neq j$$

Where $\mu_1, \mu_2, \mu_3, \mu_4$ and μ_5 are the average values from actual data, three models estimated by GA, and regression tested at α significance level. If the null hypothesis is accepted then the preferred model is the one which has lower MAPE error. Otherwise, if the null hypothesis is rejected Duncan Multiple Range Test is used to compare treatment means and to select the preferred model.

Montgomery (2001) prescribes the use of either LSD or Duncan's Multiple Range experiments. There are $3(b - 1)$ degrees of freedom for error at α level of significance where b is number of blocks or observation for the 5 treatments (actual, three models of GA, and regression) to perform Duncan method the following equation is used where MSE is the mean square error obtained from ANOVA:

$$S_{\bar{y}_i} = \sqrt{\frac{MS(error)}{b}} \quad (9)$$

In equation (9) b is the number of blocks and R_p as calculated as follow:

$$R_p = r_\alpha(p, f) S_{\bar{y}_i} \quad (10)$$

In which $r_\alpha(p, f)$ is obtained from Duncan tables, α is the level of significance and f is the degree of freedom. Therefore a blocked design of ANOVA is applied. Table II, III and IV show the ANOVA results for agricultural, low energy consuming and energy intensive sectors respectively. For low energy consuming sector, the null hypothesis is accepted and therefore quadratic model in this sector is chosen due to its lower MAPE error rather than the exponential and linear-logarithmic and regression models, though for the two other sectors the null hypothesis is rejected and therefore Duncan test is applied to find out which treatment pair has lead to the rejection of the null hypothesis in the agricultural and energy intensive sectors.

TABLE II-ANOVA results for the comparison of
GA models with actual data in agricultural sector

| ANOVA: Blocked design SUMMARY | | | | | | |
|----------------------------------|-------------------|--------------------------|--------------------|----------|---------------------|-----------------|
| <i>Groups</i> | <i>Count</i> | <i>Sum</i> | <i>Average</i> | | | |
| Regression | 4 | 54096 | 13524 | | | |
| Actual data | 4 | 46520 | 11630 | | | |
| GA _{quad} | 4 | 67890 | 16972.5 | | | |
| GA _{exp} | 4 | 45041 | 11260.25 | | | |
| GA _{line-log} | 4 | 47434 | 11858.5 | | | |
| ANOVA | | | | | | |
| <i>Source of Variation</i> | <i>Sum Square</i> | <i>Degree of Freedom</i> | <i>Mean Square</i> | <i>F</i> | <i>F (α = 0.05)</i> | <i>p- value</i> |
| Between Groups (treatment) | 88999820 | 4 | 22249955 | 8.917012 | 3.259167 | 0.001397 |
| Blocks (months) | 1/44E+08 | 3 | 48107381 | 19.27977 | | |
| Within Groups | 29942704 | 12 | 2495225 | | | |
| Total | 2/63E+08 | 19 | | | | |

TABLE III-ANOVA results for the comparison of
GA models with actual data in low energy consuming sector

| ANOVA: Blocked design SUMMARY | | | | | | |
|----------------------------------|-------------------|--------------------------|--------------------|----------|---------------------|-----------------|
| <i>Groups</i> | <i>Count</i> | <i>Sum</i> | <i>Average</i> | | | |
| Regression | 4 | 17205 | 4301.25 | | | |
| Actual data | 4 | 20197 | 5049.25 | | | |
| GA _{quad} | 4 | 19762.371 | 4940.59 | | | |
| GA _{exp} | 4 | 17906 | 4476.5 | | | |
| GA _{line-log} | 4 | 18507 | 4626.75 | | | |
| ANOVA | | | | | | |
| <i>Source of Variation</i> | <i>Sum Square</i> | <i>Degree of Freedom</i> | <i>Mean Square</i> | <i>F</i> | <i>F (α = 0/05)</i> | <i>p- value</i> |
| Between Groups (treatment) | 1567788 | 4 | 391947.1 | 0.92 | 3.26 | 0.48 |
| Blocks (months) | 1963891 | 3 | 654630.5 | 1.54 | | |
| Within Groups | 5094278 | 12 | 424523.1 | | | |
| Total | 8625958 | 19 | | | | |

TABLE IV-ANOVA results for the comparison of GA models with actual data in energy intensive sector

| ANOVA: Blocked design SUMMARY | | | | | | |
|-------------------------------|------------|-------------------|-------------|-------|-----------------------|-------------|
| Groups | Count | Sum | Average | | | |
| Regression | 4 | 27411 | 6852.75 | | | |
| Actual data | 4 | 57742 | 14435.5 | | | |
| GA _{quad} | 4 | 20754 | 5188.5 | | | |
| GA _{exp} | 4 | 26219 | 6554.75 | | | |
| GA _{line-log} | 4 | 58237 | 14559.25 | | | |
| ANOVA | | | | | | |
| | Sum Square | Degree of Freedom | Mean Square | F | F ($\alpha = 0/05$) | p- value |
| Between Groups (treatment) | 3/37E+08 | 4 | 84224986 | 99.53 | 3.25 | 4.28201E-09 |
| Blocks (months) | 1959184 | 3 | 653061.3 | 0.77 | | |
| Within Groups | 10154559 | 12 | 846213.2 | | | |
| Total | 3/49E+08 | 19 | | | | |

For the energy intensive sector we have the following averages in the order of the highest value to the lowest:

$$\bar{y}_1 = 14435.5 \quad \bar{y}_2 = 14559.25 \quad \bar{y}_3 = 6852.7$$

$$\bar{y}_4 = 6554.75 \quad \bar{y}_5 = 5188.5$$

In the above \bar{y}_1 is the average predicted value by GA using linear-logarithmic model \bar{y}_2 is the average value of the actual data, \bar{y}_3 is the average predicted value by GA with regression \bar{y}_4 is the average predicted value by GA using exponential model, \bar{y}_5 is the average predicted value by GA using quadratic model. We have actual data as the second treatment; therefore all treatments are compared to this one. At $\alpha = 0.05$ we have:

$$S_{\bar{y}_i} = 459.9$$

$$r_{0.05}(2,12) = 3.08$$

$$r_{0.05}(3,12) = 3.23$$

$$r_{0.05}(4,12) = 3.33$$

$$r_{0.05}(5,12) = 3.36$$

$$R_2 = r_{0.05}(2,12)S_{\bar{y}_i} = 1416.49$$

$$R_3 = r_{0.05}(3,12)S_{\bar{y}_i} = 1485.47$$

$$R_4 = r_{0.05}(4,12)S_{\bar{y}_i} = 1531.46$$

$$R_5 = r_{0.05}(5,12)S_{\bar{y}_i} = 1545.26$$

Treatment 2 compared to 1: $14559 - 14435.5 = 123.75 < 1416.49 \Rightarrow \mu_1 = \mu_2$

Treatment 2 compared to 3: $14435.5 - 6852.7 = 7582.8 > 1416.49 \Rightarrow \mu_2 \neq \mu_3$

Treatment 2 compared to 4: $14435.5 - 6554.75 = 7880.75 > 1485.47 \Rightarrow \mu_2 \neq \mu_4$

Treatment 2 compared to 5: $14435.5 - 5188.5 = 9247 > 1531.46 \Rightarrow \mu_2 \neq \mu_5$

For agricultural sector we have the following averages in the order of the highest value to the lowest:

$$\bar{y}_1 = 16972 \quad \bar{y}_2 = 13524 \quad \bar{y}_3 = 11858$$

$$\bar{y}_4 = 11630 \quad \bar{y}_5 = 11260$$

In the above \bar{y}_1 is the average predicted value by GA using quadratic model \bar{y}_2 is the average predicted value by GA with regression \bar{y}_3 is the average predicted value by GA using linear-logarithmic model, \bar{y}_4 is the average value of the actual data, \bar{y}_5 is the average predicted value by GA using exponential model. We have actual data as the fourth treatment; therefore all treatments are compared to this one. At $\alpha = 0.05$ we have:

$$S_{\bar{y}_i} = 789.8$$

$$r_{0.05}(2,12) = 3.08$$

$$r_{0.05}(3,12) = 3.23$$

$$r_{0.05}(4,12) = 3.33$$

$$r_{0.05}(5,12) = 3.36$$

$$R_2 = r_{0.05}(2,12)S_{\bar{y}_i} = 2432.58$$

$$R_3 = r_{0.05}(3,12)S_{\bar{y}_i} = 2551.05$$

$$R_4 = r_{0.05}(4,12)S_{\bar{y}_i} = 2630.03$$

$$R_5 = r_{0.05}(5,12)S_{\bar{y}_i} = 2653.72$$

Treatment 4 compared to 3: $11858.5 - 11630 = 228.5 < 2432.58 \rightarrow \mu_3 = \mu_4$

Treatment 4 compared to 5: $11630 - 11260.25 = 369.75 < 2432.58 \rightarrow \mu_5 = \mu_4$

Treatment 4 compared to 2: $13524 - 11630 = 1894 < 2551.05 \rightarrow \mu_2 = \mu_4$

Treatment 4 compared to 1: $16972 - 11630 = 5342 > 2630.03 \rightarrow \mu_1 \neq \mu_4$

By comparing treatment pairs in agricultural and energy intensive sectors we find out that inequality factors in energy intensive sector are the third, fourth and fifth treatment and the only equality factor is for the first and the second treatment which are average predicted value by GA using linear-logarithmic model and average values of the actual data respectively and it is accepted at $\alpha = 0.05$ in this sector. In the agricultural sector, the only inequality factor is for the first treatment which is the average predicted value by GA using quadratic model. According to the MAPE error between the approaches in this sector the linear-logarithmic model is chosen.

5. Conclusion

This study presented a genetic algorithm (GA) with variable parameters to forecast electricity demand in agricultural, low energy consuming and energy intensive sectors in Iran using stochastic procedures. The economic indicators used in this paper are price, value added, number of customers and consumption in the last periods for agricultural and low energy consuming sectors and price, value added, number of customers, price of the substitute fuel and energy intensity in energy intensive sector. Three kinds of models; linear-logarithmic, exponential and quadratic were used to find which leads us to minimum error for the related sector. The GA applied in this study was tuned for all its parameters and the best coefficients with minimum error were identified, while all parameter values were tested concurrently. The estimation errors of genetic algorithm models were less than that of estimated by regression method. Finally, analysis of variance (ANOVA) was applied to compare

genetic algorithm (three models), regression and actual data. It was found that at $\alpha = 0.05$ the five treatments are not equal except for the low energy consuming sector, and therefore Duncan test was applied to see which treatment pair has lead to the rejection of null hypothesis. Furtherer Duncan test shows that genetic algorithm estimations are closer to the actual data with less MAPE (Mean Absolute Percentage Error) error than that of estimated by regression. And for the agricultural and energy intensive sectors, linear-logarithmic model, for low energy consuming sector quadratic model was chosen. The data from 1979 to 2003 is used to forecast electricity consumption in the aforementioned sectors as the case study.

References

- [1] H. Ceylan, H. Ozturk, "Estimating energy demand of Turkey based on economic indicators using genetic algorithm approach", *Energy Conversion and Management* (45), pp. 2525–2537, 2004.
- [2] H. Ozturk, H. Ceylan, O. E. Canyurt, A. Hepbasli, "Electricity estimation using genetic algorithm approach: a case study of Turkey", *Energy*(30), pp.1003–1012, 2003 .
- [3] H.K Ozturk, O.E.Canyurt, A.Hepbasli, Z.Utlu, "three different genetic algorithm approaches to the estimation of residential exergy input/output values", *Building and Environment*(39), pp. 807-816, 2003.
- [4] O.E, Canyurt, H.Ozturk, A. Hepbasli, "Energy demand estimation based on two-different genetic algorithm approaches", *Energy Source*, VOL. 26(14), pp. 1313-1320, 2004.
- [5] S.Haldenbilen, H.Ceylan, "Genetic algorithm approach to estimate transport energy demand in Turkey", *Energy Policy*, pp. 89-98, 2004
- [6] M.S. Osman,, M.A. Abusing, A.A. Mousa ."A combined genetic algorithm-fuzzy logic controller (GA-FLC) in non-linear programming", *Applied Mathematics & Computations*,pp. 821-840, 2005.
- [7] D. Bunning, M. Sun, , Genetic algorithm for constrained global optimization in continuous variables. *Applied Mathematics & Computation* (171), pp. 604-636, 2005.
- [8] H. Hasheminia and S.T. Akhavan Niaki, , A genetic algorithm approach to fit the best regression/econometric model among the candidates, *Applied Mathematics & Computation*, in press, 2006.
- [9] A. Tang, , C. Quek, and G. Ng, GA-TSKfnn: Parameters tuning of fuzzy neural network using genetic algorithms, *Expert Systems with Applications* (29), pp. 769-781, 2005.
- [10] Goldberg DE, *Genetic algorithm in search, optimization and machine learning*.Harlow,England:Adison-wesley; 1989.
- [11] T. J. VanderNoot , I. Abrahams, "The use of genetic algorithms in the non-linear regression of emittance data ", *Journal of Electro Analytical Chemistry*, Volume 448, Issue 1, pp. 17-23 ,1998.
- [12] M.Sadeghi, "Demand stability for energy in Iran", PhD, Faculty of Economics, University of Tehran, Iran,1999.
- [13] N. Sadeghi, "Forecasting and modeling electricity demand by an econometric model.", Ms Thesis, Faculty of economics, University of Tehran, Iran, 2003.
- [14] *Energy balances of Islamic Republic of Iran*, 2004.
- [15] Montgomery, D.C. *Design & Analyze of Experiments*, Jhon Wiley & Sons, New York, 2001.