

## A combined 2-dimensional fuzzy regression model to study effect of climate change on the electrical peak load

Hamed Shakouri G., Hosain Zaman

Department of Industrial engineering, College of Engineering, University of Tehran  
Tehran, Iran  
zaman@ut.ac.ir

**Abstract**—This paper studies the impact of climate change on the electricity consumption by means of a fuzzy regression approach. The climate factors which have been considered in this paper are humidity and temperature, whereas the simultaneous effect of these two climate factors is considered. The impacts of other climate variables, like the wind, with a minor effect on energy consumption are ignored. The innovation which applies in this paper is the division of the year into two parts by using the temperature-day graph in the year. To index the humidity, data of the minimum humidity per day are used. For temperature, the maximum temperature of the first part of the year (warm days) and the minimum of the second part (cold days) are used. The indicator for the consumption is the daily peak load. The model results show high sensitivity to the temperature but low sensitivity to the humidity. Moreover, it is concluded that the model structure cannot be the same and for the cold par additional variables such as gas consumption should be considered.

**Keywords**- fuzzy regression, electricity consumption, climate change

### I. INTRODUCTION

The intergovernmental panel for climate change predicted that the average temperature of our planet surface will increase by 1.4–5.8 °C by the end of 21<sup>st</sup> century. Limited resources of energy in joint to the effect of temperature on the energy consumption together attract our attention to the temperature changes phenomenon. Therefore, researches have focused on the problem, which is the effect of climate change on the energy consumption.

Zmeureanu and Renaud proposed a method for estimation of the impact of climate change on heat energy consumption in households sector [1]. Peirson and Henley consider it as a dynamic problem and also shown that by using autoregressive specification we can obtain a good explanation of present load [2]. Pardo et al. studied the relationship between weather and electricity demand in Spain. They proposed a transfer function intervention model to predict the electricity demand in the hot and cold days [3]. Bessec and Fouquau studied the relationship between temperature and electricity consumption in 15 countries of European Union. They showed the nonlinearity link between electricity consumption and temperature found in more limited geographical areas in previous studies; also they showed that the sensitivity between electricity consumption and temperature increases in summer [4]. Franco and

Sanstad studied the effect of climate change on electric power in California. They studied the relationship between temperature, electricity consumption and peak demand in that region. They also had a project to estimate the potential impacts of future temperature changes on electricity consumption and peak demand[5]. Belzer et al. used the degree-days method to survey the impact of climate change on energy consumption change [6].

The sense of the people for the weather is different. It means that at the same temperature different people may have different feelings about the weather. This fact can result different consumption of electricity for air conditioners and heaters. Linguistic variables are used to represent this concept. In this regard, researchers have also utilized fuzzy regression on energy demand and consumption. For instance, Nazarko and Zalewski applied fuzzy regression to forecast peak load in distribution systems [7]. Azadeh et al proposed an integrated fuzzy regression and time series algorithm to estimate and predict electricity demand in Iran. They used autocorrelation function to define input variables [8]. Tajvidi et al proposed an integrated fuzzy regression, computer simulation and time series framework to estimate and predict electricity demand for seasonal and monthly changes in electricity consumption in Iran [8][9][10]. Shakouri et al studied the effect of temperature changes on electricity demand by combining fuzzy regression with a TSK-FR model. They used the average of temperature per day as a climate effect and studied its effect on the average consumption time per day [10].

This paper proposes a fuzzy regression based analysis of the simultaneous effect of both temperature and humidity on the electricity demand in Tehran.

### II. PROBLEM DEFINITION

Knowing that Energy is equal to the product of time,  $T$ , and power,  $P$ :

$$E = P \times T \quad (1)$$

Previous works show that the climatic variables do change the electricity consumption by both changing the load and the time an appliance is used. We would like to study how much climate change can increase energy consumption by increasing the load connected to the power network. Clearly, the climate variables like temperature, humidity and wind can show effects on both load and time.

However, in this paper the effect of temperature and humidity on the peak load is studied. The load,  $P$ , in (1) is the average load in the time period of  $T$ , while the peak load is desired which can be obtained as:

$$P_{max} = \max p(t) \quad (2).$$

Suppose that the minimum humidity is  $h_{min}$  and the minimum and the maximum of temperature are  $\theta_{min}$  and  $\theta_{max}$  respectively. These can be treated as linguistic variables that take different fuzzy term sets, e.g.  $\theta_{min}$  can take "cold" and "warm", while  $\theta_{max}$  may take "warm" and "hot". Therefore, a fuzzy model can describe the relation:

$$P_{max} = f(\theta_{min}, \theta_{max}, h_{min}) \quad (3).$$

Now, the problem encountered is to find a proper estimate of the three-variable function of  $f$ , where there are statistical data available on all the four variables.

### III. METHODOLOGY

The electricity consumption pattern differs for the two parts of a year; i.e. in summer and winter. In other words, the climatic factors of *warm* seasons and the *cold* ones cause different behaviors of consumption. Besides the temperature, as the main factor, the effect of humidity on our feeling of weather in hot temperature and cold temperature is different. Therefore, in this research, the year is divided to two parts in order to take advantages of these two different impacts. Figure 1 shows typical fluctuations of both the maximum and the minimum temperatures in days of a year, as well as the average temperature.

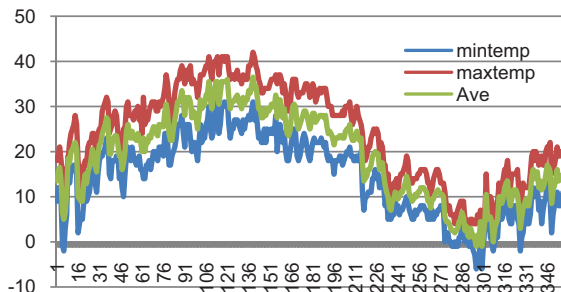


Fig.1. Daily temperature (maximum, minimum, average) in a year

Such data can be used to find a proper division for the year. Two different trends the data easily represent the *hot* part and the *cold* of the year. By simple data processing procedures including smoothing and differentiation one can find two inflection points around the beginning days of spring and autumn. To do so, the minimum temperature is applied for specifying the *cold* part and the maximum temperature for the *warm* part. As it is expected, historical data shows that the first peak load occurs in a hot day with of summer, and the second one belongs to one a cold day of winter, with the minimum temperature.

To make a deeper sense of the behavior of each climatic variable independently, it is useful to look at the graph of the

model, i.e. the peak load versus the minimum temperature (in the *cold* interval), maximum temperature (in the *warm* interval), and the minimum humidity in Figures 2 to 4. Previous research have shown that the peak load represents a polynomial shape behavior with respect to the temperature [10]. As it can be comprehended in the figures, similar interrelation exists in the two parts of the year. Therefore, a general form of second order polynomial is considered for the peak load regression function in the two parts, applying the minimum or maximum temperatures as the exogenous variables:

$$y = a_0 + a_1\theta + a_2\theta^2 \quad (4),$$

where  $\theta$  is the temperature (maximum or minimum temperature). On the other side, for the humidity effect on the peak load an irrelevance, or at last, almost linear interdependence may be assumed based on Figure 4. So, we may have:

$$y = a_0 + a_1h \quad (5),$$

where  $h$  is the independent variable representing the minimum humidity.

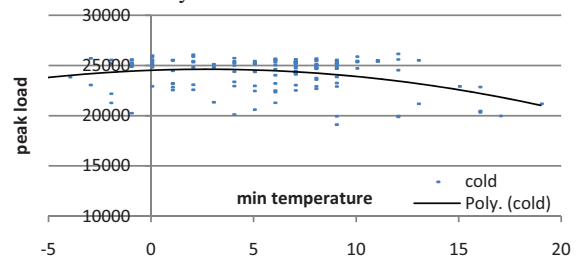


Fig.2: The peak load [MW] vs. minimum temperature [°C] in cold part

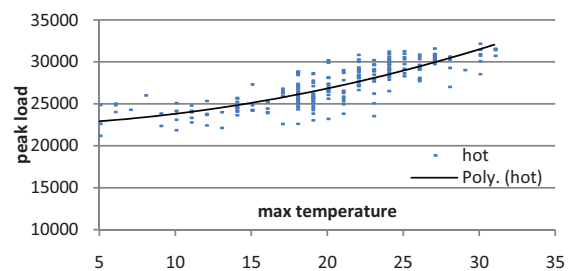


Fig.3: The peak load [MW] vs. maximum temperature [°C] in warm part

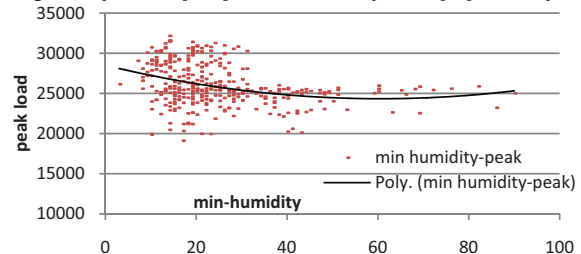


Fig.4: The peak load [MW] versus minimum humidity [%] in the year

Now, in order to construct a two-dimensional regression function, one of the following three types of combination may be chosen:

- (1) The conjunctive relationship, leading to a logical AND, which can be represented by a multiplication, or in general, by a productive operator.
- (2) The disjunctive relationship, leading to a logical OR in the logical relationships, which can be replaced by a summation in the mathematical equations, that in general appears as an additive operator.
- (3) The Mixed relationship, such as the logical XOR, which may be represented by combined relations and can be stated by perhaps complex mathematical models (commonly an Artificial Neural Net).

Based on the logics of how temperature and humidity affects electricity consumption, in this paper a disjunctive relation is used to combine the two independent variables' effect. This way, we will obtain the following general form of the proposed regression model:

$$Y = \tilde{a}_0 + \tilde{a}_1\theta + \tilde{a}_2\theta^2 + \tilde{a}_3h \quad (6)$$

where,  $Y$  stands for the fuzzy estimated peak load, and the accent  $\sim$  denotes fuzziness of the coefficients. It should be noticed that  $\theta$  is the maximum temperature in the first part of the year and the minimum for the second one, whereas  $h$  is the minimum humidity for whole the year.

#### IV. FUZZY REGRESSION MODEL

Fuzzy regression model is an extension of the classic regression, in which some elements like input or output or both are fuzzy numbers.

The purpose of a fuzzy regression model is to find the best solution with the least error, by giving more flexibility to the variables and/or the parameters. Among various kinds of regression models [11], which we have chosen several to solve (7). Tanaka (1982), Tanaka (1989), Peters (1994), Wang-Tsaur (2000), Chang-Lee (1996), Ozelkan (2000), Lai-Chang (1994), and Hojati (2005) models [12]-[19] obtained better results. However, the last one is chosen. Hojati et al. proposed a fuzzy regression model, known by HBS1, which minimizes the summation of the total deviation of the upper points of  $H$ -certain predicted and associated observed intervals and deviation of the lower points of  $H$ -certain predicted and the associated observed intervals. The proposed model is as below [19]:

$$\min \sum_{i=1}^n (d_{iu}^+ + d_{iu}^- + d_{il}^+ + d_{il}^-) \quad (7)$$

$$s. t. \sum_{j=0}^k [\alpha_j + (1-H)c_j]x_{ij} - d_{iu}^- \geq \bar{y}_i + (1-H)e_i - d_{iu}^+$$

$$\sum_{j=0}^k (\alpha_j - (1-H)c_j)x_{ij} - d_{iu}^- \leq \bar{y}_i - (1-H)e_i + d_{iu}^+$$

$$\sum_{i=1}^n \sum_{j=0}^k c_j x_{ij} \leq v$$

$$d_{iu}^+, d_{iu}^-, d_{il}^+, d_{il}^- \geq 0 \quad \begin{matrix} i = 1, \dots, n \\ j = 0, \dots, k \end{matrix}$$

$$c_j \geq 0, \quad j = 0, \dots, k$$

In this model, at most one of the  $d_{iu}^+$  and  $d_{iu}^-$  is positive, also, at most one of the  $d_{il}^+$  and  $d_{il}^-$  is a positive number and another one will be zero. Indeed,  $|d_{iu}^+ - d_{iu}^-|$  is the distance between upper point of  $H$ -certain predicted interval and the upper point of the  $H$ -certain observed interval and  $|d_{il}^+ - d_{il}^-|$  is the distance between the lower point of the  $H$ -certain predicted interval and the lower point of the  $H$ -certain observed interval. The summation of these two absolute values construct the objective function of the model. In this paper this fuzzy regression model is solve applying crisp inputs of temperature and humidity, resulting in fuzzy outputs.

#### V. COMPUTATIONAL RESULTS

The proposed method is applied to the data of the capital city of Iran, Tehran, for 2005. The daily peak load data is taken from the Dispatching Center Data Bank (Ministry of Energy, Tehran, Iran) and the humidity and temperature data are taken from the Meteorological Organization. A "mini max" normalization preprocess seems to be necessary to avoid numerical problems:

$$preprocessed(x_i) = \frac{x_i - \min(x_j)}{\max(x_j) - \min(x_j)}; j = 1, 2, \dots, i, \dots, n \quad (8)$$

where  $x_i$  is each of the input variables. This preprocess bounds the variable in the interval of [0, 1] and also results in limited regression output,  $Y$ . Thus, to compare the outputs with the real observed data, a post process should be done as an inverse process:

$$post\_processed(Y) = Y \times [\max(Y) - \min(Y)] + \min(Y) \quad (9)$$

This post process converts predicted data to the original desirable range.

The HBS1 model is implemented in LINGO8.0 and the results for the two parts of the year are shown in Tables (1) and (2).

Table (1): Parameters of the FR model for the warm part

Parameter	Fuzzy Regression	Ordinary Regression
$a_0$	0	-0.037
$c_0$	0	0.108
$a_1$	0.31	0.39
$c_1$	0.01	0.347
$a_2$	0.608	0.547
$c_2$	0.002	0.290
$a_3$	0.177	0.198
$c_3$	0.021	0.089

Table (2): Parameters of the FR model for the cold part

Parameter	Fuzzy Regression	Ordinary Regression
$a_0$	0.81	0.347
$c_0$	0.01	0.201
$a_1$	0	0.087
$c_1$	0	0.682
$a_2$	0	-1.615
$c_2$	0.12	0.677
$a_3$	0.06	-0.031
$c_3$	0	0.218

In these two tables, values of  $c_j$  in the second columns indicate the spread of each fuzzy coefficient, and  $a_j$  is the corresponding estimated center. The third columns include parameters estimated for an ordinary regression models based on (6). The values associated to  $c_j$ 's in this case indicate the 95% confidence interval of each parameter. As it is seen, the second fuzzy model has resulted in zero coefficients, while the corresponding ordinary regression is not meaningless that much, although with large confidence intervals.

To compare the model results with the observed data and assess the accuracy of the model, an error index should be calculated. Herein, the Mean Absolute Percent Error (MAPE) is used, given by:

$$MAPE = \frac{1}{N} \sum \frac{|real\ data - predicted\ data|}{real\ data} \quad (10)$$

where  $N$  is number of the data. The data is classified in two groups, train data and test data. Train data are used for training model including almost 80% of the data. Test data, including almost 20% of the whole data, is applied to test the model, and hence is used to calculate MAPE.

In the warm part there are 159 train data samples and 40 test data samples, while in the cold part these are 122 and 30 respectively. Table (3) shows MAPE of the two fuzzy regression models for the both parts, and Figures 5 and 6 show the comparison of output estimates and observed data in the relevant days of the year. According to Table (3) it can be said that the error of the model for warm part of the year is desirable but it is not desirable for the cold part.

The solid curves show the predicted polynomials and the dashed ones are the real curves in Figures 5 and 6. In Figure 5, the predicted curve is very close to the real curve, however, it is not the case for the other one. Although the model presents a good solution for the first part of the year, i.e. the warm part, not an appropriate solution is obtained for the cold part.

The results state that perhaps there is a low temperature and humidity effect in the cold days of a year. On the other hand, in the city of Tehran the air conditioners that use electricity are often applied in the first part of the year and in the second part electric devices are rarely used for the heating purposes. Instead, the heating devices are based on gas consumption. Therefore the humidity and temperature

have a little effect on the electrical load in the cold seasons of a year. As a conclusion, the two climate variables seem not appropriate factors to regress the peak load, otherwise it may lead to insignificant coefficients.

Table (3): Comparison of the models based on MAPE

Model part	MAPE (F.R.)	MAPE (O.R.)
cold	11%	8%
warm	2%	4%

F.R. : Fuzzy Regression  
O.R. : Ordinary Regression

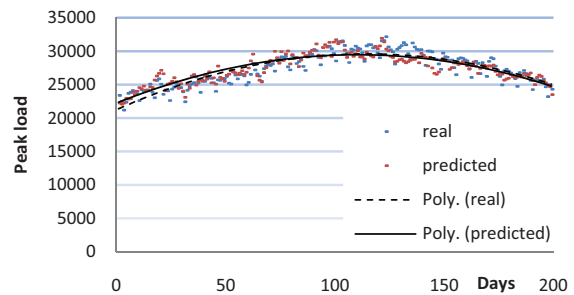


Fig.5: Comparison of the predicted load and the actual load in the warm part of the year

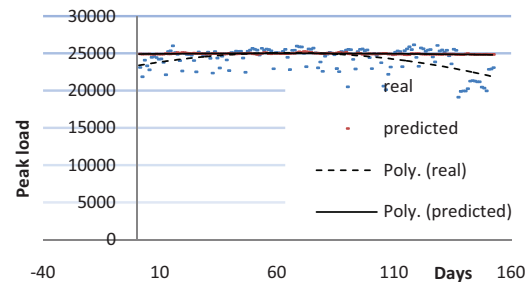


Fig.6: Comparison of the predicted load and the actual load in the cold part of the year

## VI. CONCLUSION

In this paper, a fuzzy model is applied to study the simultaneous effect of the temperature and humidity, as the most important climatic variables, on the electrical peak load. Since the electricity consumption pattern changes during a year, a rational partitioning is implemented based on the climate changes, which divides the year in two warm and cold parts. It is shown that in the first part of the year the fuzzy regression equation have positive appropriate coefficients for the maximum temperature, squared maximum temperature, and the minimum humidity. Moreover, the model has an acceptable error.

However, for the second part of the year, i.e. the cold part, the coefficients and the error of the model are not acceptable as much as the first one. It seems that the problem is originated from dependency of the energy consumption on gas, in the cold seasons of the year, and rare usage of

electricity for the heating purposes. One may conclude that if substitution of gas by electricity is augmented to the model, better results may be obtained for the second part of the regression model. Finally, a modification is possible by ignoring the holidays in which the consumer behavior defers from usual.

The most important point would be that a single equation regression cannot be a successful model to imitate the behavior of electricity consumption for the variety of different factors affecting the demand. Rather, a system of equations, describing mechanisms and dynamics of the system, is a better solution [20].

#### REFERENCES

- [1] R. Zmeureanu and G. Renaud, Estimation of potential impact of climate change on the heating energy use of existing houses, *Energy Policy*, vol. 36, Jan. 2008, pp. 303-310.
- [2] J. Peirson and A. Henley, Electricity load and temperature : Issues in dynamic specification, *Energy Economics*, vol. 16, Oct. 1994, pp. 235-243.
- [3] A. Pardo, V. Meneu, and E. Valor, Temperature and seasonality influences on Spanish electricity load, *Energy Economics*, vol. 24, Jan. 2002, pp. 55-70.
- [4] M. Bessec and J. Fouquau, The non-linear link between electricity consumption and temperature in Europe: A threshold panel approach, *Energy Economics*, vol. 30, Sep. 2008, pp. 2705-2721.
- [5] G. Franco and A. Sanstad, Climate change and electricity demand in California, *Climatic Change*, vol. 87, 2007, pp. 139-151.
- [6] D. Belzer, M. Scott, and R. Sands, Climate Change Impacts on U.S. Commercial Building Energy Consumption: An Analysis Using Sample Survey Data, *Energy Sources*, Part A: Recovery, Utilization, and Environmental Effects, vol. 18, (1996), p. 177.
- [7] Nazarko and Zalewski, The fuzzy regression approach to peak load estimation in power distribution systems, *IEEE Transactions on Power Systems*, vol. 14, (1999), pp. 809-814.
- [8] A. Azadeh, M. Saberi, and O. Seraj, An integrated fuzzy regression algorithm for energy consumption estimation with non-stationary data: A case study of Iran, *Energy*, vol. 35, Jun. 2010, pp. 2351-2366.
- [9] Tajvidi, Azadeh, Saberi, Izadbakhsh, Danesh, and Gitiforouz, "An integrated simulated-based fuzzy regression algorithm and time series for energy consumption estimation with non-stationary data and case studies," presented at the *IEEE International Conference on Digital Ecosystems and Technologies*, 2009, pp. 139-143.
- [10] H. Shakouri G., R. Nadimi, and F. Ghaderi, A hybrid TSK-FR model to study short-term variations of the electricity demand versus the temperature changes, *Expert Systems with Applications*, vol. 36, Mar.
- [11] H. Shakouri G., R. Nadimi, A novel fuzzy linear regression model based on a non-equality possibility index and optimum uncertainty, *Applied Soft Computing* 9 (2009), pp. 590-598.
- [12] H. Tanaka, S. Uegima, K. Asia, Linear regression analysis with fuzzy model, *IEEE, Trans. Syst. Man Cybernet* 12 (1982), pp. 903-907.
- [13] H. Tanaka, I. Hayashi, J. Watada, Possibilistic linear regression analysis for fuzzy data, *European Journal of Operational Research* 40 (1989), pp. 389-396.
- [14] G. Peters, Fuzzy linear regression with fuzzy intervals, *Fuzzy Sets and Systems* 63 (1994), pp. 45-55.
- [15] H.F. Wang, R.C. Tsaur, Insight of a fuzzy regression model. *Fuzzy Sets and Systems*, 112(3), (2000), pp. 355-369.
- [16] P.T. Chang, E.S. Lee, A generalized fuzzy weighted least-squares regression, *Fuzzy Sets and Systems* 82 (1996), pp. 289-298.
- [17] E. C. Ozelkan, L. Duckstein. Multi-objective fuzzy regression: a general framework, *Computers and Operations Research*, 27, (2000), 635-652.
- [18] Y.J. Lai, S.I. Chang, A fuzzy approach for multi response optimization: an off-line quality engineering problem, *Fuzzy Sets and Systems* 63 (1994), pp. 117-129.
- [19] M. Hojati, C.R. Bector, and K. Smimou, "A simple method for computation of fuzzy linear regression," *European Journal of Operational Research*, vol. 166, Oct. (2005), pp. 172-184.
- [20] Z. Esmaeeli, H. Shakouri G., A. Sedighi, Investigation of pricing impact on the electrical energy consumption behavior of the household sector by a system dynamics approach, *5th International System Dynamics Conference*, Nijmegen, The Netherlands, July 23-27, 2006.