

Short Term Load Forecasting For a Micro Region Using NNs and Regression Models

Kübra Kaysal, Fatih Onur Hocaoglu, Yüksel Oğuz, Ahmet Kaysal

Department of Electrical & Electronics Engineering, Afyon Kocatepe University, Afyonkarahisar, Turkey
kkaysal@aku.edu.tr, fohocaoglu@gmail.com, yukseloguz@aku.edu.tr, akaysal@aku.edu.tr

Abstract

Thanks to today's advancing technology, some techniques have been developed to make future oriented forecast on energy generation, transmission and distribution subjects. In literature, load forecasting studies are generally realized for macro size settlement units. However, determination of parameters that effect load forecasting in macro size settlement units is very difficult. In this study, by considering a micro size settlement unit (Güney district of Denizli province), factors that may affect the load demand were examined. Then, by selecting popular methods used in the literature, load forecasting for the region was realized. Finally the results obtained from the models were compared.

1. Introduction

The electric energy is a very clean energy when compared to other generation sources. With rapid development of technology, need for electric energy has gradually increased. However, as electric energy cannot be stored, the generated energy must be consumed promptly. Therefore electric energy must be reliable.

The electric energy must be generated in needed amount in order to be able to use it efficiently. For this reason, a planning was made for electric energy. Load forecasting constitutes the basis of this planning.

The load forecasting is to make forecasts about future data by considering the past data. Load forecasting issue should be categorized under three headings; short, medium and long term load forecasting. The short-term load forecasting means hourly or daily forecast. On the other hand medium-term load forecasting is weekly and monthly forecasting. The long-term load forecasting includes yearly forecasts.

From past to today, numerous load forecasts have been performed and various methods have been developed. Eke, in his study, realized medium-term load forecasting by using artificial neural networks [1]. Erkmen and Topallı used artificial neural networks in their short-term load forecast [2]. In a different study, Erkmen and Özdoğan (1997) also used artificial neural networks in their study on short-term load forecast [3]. Eşiyok and his colleagues performed forecast by using artificial neural networks to control the load consumption in various transformer centers in İstanbul [4]. Yalçınöz and Eminoğlu made short-term load forecasting with artificial neural networks in Niğde region. Consequently, they determined that error rate in artificial neural networks was very small [5]. Topallı and his colleagues made short-term load forecasting for Turkey in their study realized in 2006 [6]. Moturi and Kioko made short-term electric load forecasting for Kenya. They used artificial neural networks in their studies [7]. Yamaçlı realized a long-term load forecasting study for Turkey in 2010. In his study, he used time series, regression and artificial neural networks and compared the results

with all methods. He decided that artificial neural networks made a good forecasting [8]. In a recent study, Wu and his colleagues made short-term load forecasting using the regression method [9]. Wang and his colleagues made load forecasting for Pekin, China. They used regression and artificial neural networks, compared the results and obtained more efficient values from artificial neural networks [10].

In this study, a small district (Denizli-Güney) was selected for examination of the load demand. After the factors affecting the load demand were researched, numerous load forecasts were performed in Güney district of Denizli region using recorded data in 2011. Artificial neural networks and regression methods were employed in forecasting. Finally, obtained results were compared.

2. Research of Factors Affecting the Load Forecast

Thanks to gradually developing technology, to make forecast in the energy sector has become very important. While forecast is rather important, reliability and correctness of the forecast is also very important. For this reason, to use only instant load data in forecast will not give efficient results.

The energy amount used in a region has many factors from physical characteristics of the region to its population. Accordingly, the factors that affect the load forecasting can be given as follows;

- Growth of population
- Density of load
- Social development
- Historical data
- Geographical factors
- Industrial factors
- The structure of the city and the land
- Seasonal factors

When these characteristics are examined, the altitude of Güney district is 830 and its surface area is 534kms. In Güney district, winters are cold and snowy and summers are hot. It locates in 70 km distance to Denizli province. According to 2011 population census, its population is 5722. The district is not much active in terms of industry. As a work opportunity, agriculture is made in the district.

In the 16 km east of the district, there is the Adıgüzel Hydroelectric Power Plant established on Menderes. The power station generates 280GW/h energy annually and it has 2 units. Power of each unit is 31MW.

3. Forecasting and Methods

Forecasting means to find an unknown value by using a known value. It is to find future values by using the past real values.

What is ideal is to make correct forecast. This means to select a correct forecasting method. In the energy sector, various methods for load forecasting have been used and the most correct forecasting methods have been selected. The mostly used methods are time series analysis, regression analysis, econometric methods and artificial neural networks. The forecasting methods used in this study are regression and artificial neural networks.

3.1. Regression Methods

Regression that is a statistical method explains the relation between two or more variables and makes forecast about future. In this method, mathematical expressions are used to explain the relation between two or more variables. It is classified as simple regression and multiple regressions depending on used variables.

Simple regression: it is a method that explains the relation between two variables. If there is a good relation between two variables, the straight line takes place on the regression line. While real values locate on the graphic in points, forecasting tries to pass from these points. When a straight line closest to points is drawn, the difference between the straight line and points gives the error; the error must be the least. The equation developed for regression line is given below.

$$y = a + bx + \varepsilon \quad (1)$$

- y : The dependent variable
- a : The y – intercept point
- b : Regression coefficient
- x : Independent variable
- ε : Random error

In multiple regressions, there is one dependent variable and more than one independent variable. Its formula is as given below;

$$y = a + b_1x_1 + b_2x_2 + \dots + b_nx_n \quad (2)$$

- y : The dependent variable
- a : The y – intercept point
- b_1 : The first variable is the forecast number of times
- b_2 : The second variable is the forecast number of times
- b_n : Variable is the forecast number of times n .
- x_1 : The first argument
- x_2 : The second argument
- x_n : n . argument

The above stated technique is the smallest squares technique. In the regression technique, generally the smallest squares technique is used. In this study, the smallest squares technique is used.

3.2. Artificial Neural Networks

The artificial neural network is like a human brain adapted to machine. Because of its non-linear characteristics, it is more preferred to other methods and has many areas of use. In Fig. 1 a simple artificial neural network is shown.

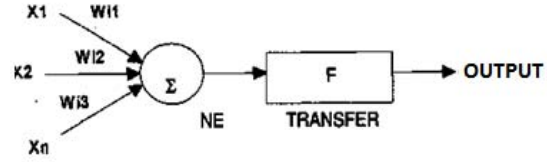


Fig. 1. Artificial Neural Network

Parts up to x_{1-n} constitute inputs of the artificial neural network. The part up to w_{1-n} constitutes the weight. Σ Function is part the input values are multiplied by weights and then added. To obtain an output, the activation function must be used.

An input incoming to network is multiplied by its own weight value; for each input, its own weight value is multiplied and so, total inputs incoming to network are found. The number of inputs is shown as follows:

$$G = \sum_{i=1}^n x_i w_i \quad (3)$$

For formation of an output, the activation function is used. Today, there are many activation functions. In this study, sigmoid function is used. The sigmoid function is as follows;

$$f = \frac{1}{1+e^{-n}} \quad (4)$$

The artificial neural networks are classified according to their flows in the network. These are feed-forward networks and feed-backward networks.

- The feed-forward networks consist of layers. The information advances from input layer towards to output layer in one direction.
- In the feed-backward networks, outputs are fed backward to previous layers.

In the artificial neural networks, some networks are feed-forward. As the feed-forward networks advance in one direction, they are fast. These networks continuously return to previous layer to find the most suitable weight value. For this reason, in solution of complex problems, the feed-backward networks are used.

4. Studies with Regression

In this study, short-term load forecasting is performed Güney district of Denizli province. During forecast regression and artificial neural network methods are used.

While forecasting by regression, simple regression and multiple -regression are used separately. In simple regression, an independent variable is used. This used simple variable is the instant data of one hour ago. To show the accuracy of the forecast, measured and predicted load values are presented in Fig. 2.

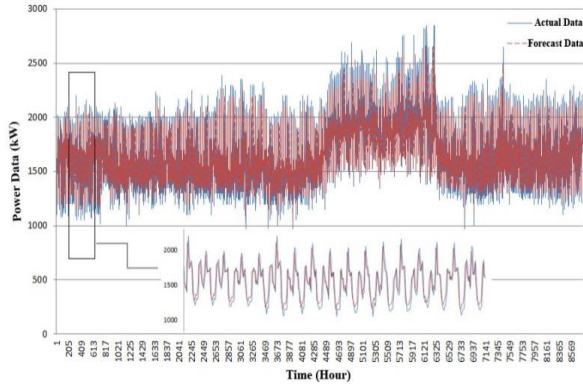


Fig. 2. Forecast of simple regression graph

As a further step relative error (RE) and root mean square errors (RMSE) are calculated from the following equations.

$$RE = \sum_{i=1}^n \left[\frac{y(i) - f(x_i)}{y(i)} \right] * 100 \quad (5)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y(i) - f(x_i))^2} \quad (6)$$

In equation (5-6); y_i = expresses the real values and $f(x_i)$ the forecasting value. The RE and RMSE values found in simple regression are given in Table 1.

Table 1. Errors of the Simple Regression

Errors of the Simple Regression	
RMSE	165,4553
RE	%7,56

In the multiple -regression, the data of one hour ago and temperature data at that hour are employed as independent variables. The obtained forecasting and measured values are plotted in Fig. 3.

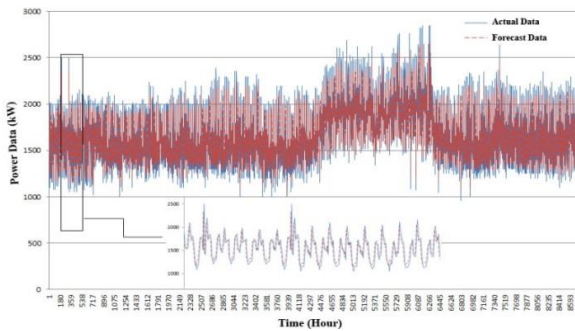


Fig. 3. Forecast of multiple regression graph

As in the simple regression, the RE and RMSE values are calculated. The table of obtained results is given in Table 2.

Table 2. Errors of the Multiple Regressions

Errors of the Multiple Regression	
RMSE	164,9646
RE	%7,38

5. Studies with Artificial Neural Network

Back propagation algorithm is employed to train the artificial neural network. The algorithm has two inputs data. These inputs are temperature data and an hour earlier data. While weights were found in this algorithm the program runs back to the very beginning to lessen the errors for fine measurement of each layer.

First of all, the real values for 2011 year were normalized in a manner recognizable by neural networks. The normalization formula is as given below.

$$v_n = 0.8 * \left[\frac{v_r - v_{min}}{v_{max} - v_{min}} \right] + 0.1 \quad (7)$$

- v_n : Values to be normalize
- v_{min} : Smallest data on the network
- v_{max} : Biggest data on the network

To select data randomly from the network, a small code is developed under C programming language. The random selection of the data is necessary to avoid the network to memorize the data pairs. Some of the selected values are determined as training data whereas the others are test data.

To train the network, many tests are performed containing various numbers of hidden layers that consist of four inputs and one output. In this study, the ideal thing is two hidden layers. The obtained total value is applied to sigmoid transfer function. A graphic that contains obtained forecasting values and real values is given below.

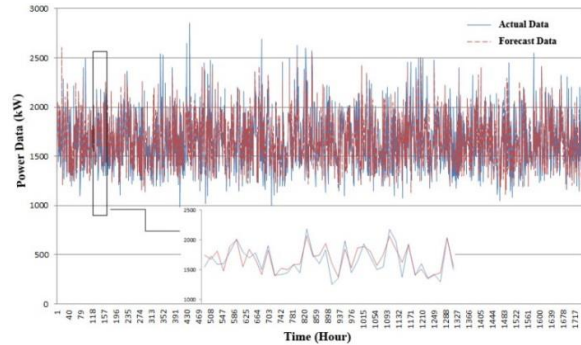


Fig. 4. The load forecast with two hidden layer of NNs

The obtained forecasting values are de-normalized again to calculate errors. The de-normalization formula is given below.

$$v_r = \frac{v_n - 0.1}{0.8} * (v_{max} - v_{min}) + v_{min} \quad (8)$$

The RE and RMSE values of de-normalized values are calculated. The RE and RMSE table is given below.

Table 3. The Load Forecast with Two Hidden of NNs

Error of NNs	
RMSE	150,59
RE	%6,65

The other study carried out in NNs is to use temperature data of one hour ago as input values. In this study, hourly temperature values are added according to previous study. The effect of temperature on consumption is examined. The graphic results obtained from the study and RE-RMSE values are given below.

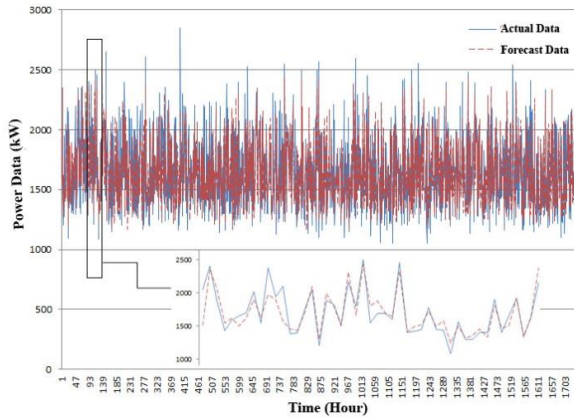


Fig. 5. Temperature effect of NNs

Table 4. Temperature Effect of NNs

Temperature Effect of NNs	
RMSE	144,0368
RE	%6,36

When the obtained results are examined, it can be seen from the results that accuracy of the forecast is improved. However the improvement of the accuracy is not considerable.

In this study, separate forecast are performed using NNs and regression methods. In both of them, instant and hourly data and temperature data are employed. The RE and RMSE results obtained from both methods are given in below tables.

Table 5. Errors of the Simple and Multiple Regressions

An hour earlier data	Results of Regression	Results of NNs
RMSE	165,4553	150,59
RE	%7,56	% 6.65
An hour earlier data and temperature data		
RMSE	164,9646	144,0368
RE	%7,38	%6,36

When the results are compared, it is determined that RSME values obtained from regression in both studies are high. But, RMSE value must be low as much as possible in forecasting. In other words, if the errors made in forecasting are less, the reliability and correctness of it will increase. Consequently, studies made in NNs gave more reliable and correct results.

6. Conclusion

Forecasting is very important in adaptation to innovations of this epoch. It is especially among the things that must be realized in energy systems. Since electricity storage is not possible for huge amount of generations, it must be consumed as soon as it is generated. To make forecasting in energy systems means to meet the demanded energy for that moment. In other words, the disadvantage of not storing the electricity can be turned in to an advantage with the ideal forecasting method. For this reason, many methods to make forecasts have been developed for years.

In this study, a small area was selected for load forecasting and first of all, factors that may affect the load forecasting in the region were examined. According to the analysis made, the temperature does not have much effect on consumption. When summer months in all around Turkey are considered, there will be increase in using of devices such as air conditioning because of the effect of high temperature. However, this is not valid for the studied region. The people immigrate to fields in summer months because of agriculture. As tourism has not developed in the region, electric consumption is not high in summer months except Mondays. On Mondays, the people return to the city to meet their needs.

When the value given in Table 5 is examined, it is seen that RMSE values are always low in NNs study. It gives more successful results than the NNs regression method as because lesser errors are very important in forecasting.

7. References

- [1] Eke, İ. (2011). Diferansiyel evrim algoritması destekli yapay sinir ağı ile orta dönem yük tahmini. *International Journal of Research and Development*, **3**: 28-32.
- [2] Erkmén, I and Topalli, A.K. (2003). Four methods for short-term load forecasting using the benefit of artificial intelligence. *Electrical Engineering*, **85**: 229- 233.
- [3] Erkmén, İ. ve Özdoğan, A. (1997). Short term load forecasting using genetically optimized neural network cascaded with a modified Kohonen clustering process. *12. IEEE Uluslararası Akıllı Kontrol Sempozyumu*, İstanbul, 16-18 Temmuz.
- [4] Eşiyok, E., Hocaoğlu, A.T. ve Dumanlı, M. (1995). Güç sistemlerinde yapay sinir ağları ile yük tahmini analizi. *Elektrik Mühendisliği 6. Ulusal Kongresi*, Bursa, 11-17 Eylül.
- [5] Yalçınöz, T. ve Eminoglu, U. (2004). Short-term and medium term power distribution load forecasting by neural networks. *Energy Conversion and Management*, **46**: 1393- 1405.
- [6] Topallı, A.K., Erkmén, İ. ve Topallı, İ. (2006). Intelligent short-term load forecasting in Turkey. *Electrical Power and Energy Systems*, **28**: 437-447.
- [7] Moturi, C.A. and Kioko, F.K. (2013). Use of artificial neural networks for short-term electricity load forecasting of Kenya national grid power system. *International Journal of Computer Application*, **63**: 25-30.
- [8] Yamaçlı, M. (2010). Türkiye'nin uzun dönem elektrik yük tahmini. Yüksek Lisans Tezi, Dumlupınar Üniversitesi, Fen Bilimleri Enstitüsü, Kütahya.
- [9] Wu, J., Wang, J., Lu, H., Dong, Y. and Lu, X. (2013). Short-term load forecasting technique based on the seasonal exponential adjustment method and the regression model. *Energy Conversion and Management*, **70**: 1-9.
- [10] Wang, J., Li, L., Niu, D. and Tan, Z. (2012). An annual load forecasting model based on support vector regression with differential evolution algorithm. *Applied Energy*, **94**: 65-70.