Fazıl Gökgöz (Turkey), Fahrettin Filiz (Turkey)

Electricity price forecasting in Turkey with artificial neural network models

Abstract

The electricity market has experienced significant changes towards deregulation with the aim of improving economic efficiency. The electricity pricing is a major consideration for consumers and generation companies in deregulated electric markets, so that offering the right price for electricity has become more important. Various methods and ideas have been tried for electricity price forecasting. Artificial neural networks have received much attention with its nonlinear property and many papers have reported successful experiments with them. This paper introduces artificial neural network models for day-ahead electricity market in Turkey. Using gradient descent, gradient descent with momentum, Broydan, Fletcher, Goldfarb and Shanno (BFGS) and Levenberg-Marquardt algorithm with different number of neuron and transfer functions, 400 different models are created. Performances of different models are compared according to their Mean Absolute Percentage (MAPE) values; the most successful models MAPE value is observed as 9.76%.

Keywords: electricity price forecasting, neural networks, day-ahead electricity market, Turkey.
JEL Classification: C02, C13, C45, C53.

Introduction

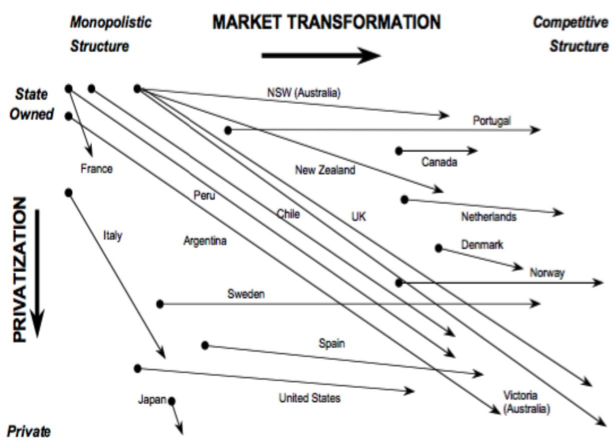
Deregulations in electricity markets have been reshaping the monopolistic power sector since the early 1990s. Many countries have recently deregulated their power system. Electricity is now traded under market rules using spot and derivative contacts. The costs of under or over contract in deregulated market can result in huge financial loss. A generator or consumer can use electricity price forecast to adjust its offered price in order to maximize profits. Likewise, regulatory agencies need price-forecasting tools to supervize or monitor the market.

Energy price forecasting is an interdisciplinary field. Energy price forecasting involves various communities, such as statistical inference, artificial intelligence, meteorological science, finance and electrical engineering. Because pricing is a complex strategy, forecasting price of commodities becomes more difficult task. Electricity is different from other commodities. It is non-storable and, for power system stability, there should be a constant balance between electricity consumption and production (Yamin, 2004).

The purpose of this study is by using artificial neural network models with different optimization algorithms to find the most successful day-ahead electrical price-forecasting model for day-ahead electricity market in Turkey. This paper is organized as follows: section 1 gives a brief literature review of the general electricity price forecasting methods. Section 2 explains paradigms of the neural networks, especially those used for the forecasting purpose. In section 3, different models are created and tested with using Turkish electricity market data. Finally, results and directions of future research are discussed in conclusion.

1. Literature review

Power market liberalization was started in Chile. The reform, which began in 1982, was based on the idea of separate generation and distribution companies. British electricity sector in 1990s followed the Chilean reforms. Nordic market which is established by integration of Norway, Sweden, Denmark and Finland, is first international electricity market opened in 1990. The study conducted by the Hagler Bailly firm in the transformation and privatization of the electricity market in the selected countries over the past two decades is shown in Figure 1. Focusing on introduction of competition and introduction of private participation, arrows shows progress of countries over last two decades. United States of America (USA) and Japan have mainly private companies. France, Peru, Chile have more vertical integration structure (Müller-Jentsch, 2001).



Source: Hagler Baily (as cited in Müller-Jentsch, 2001).

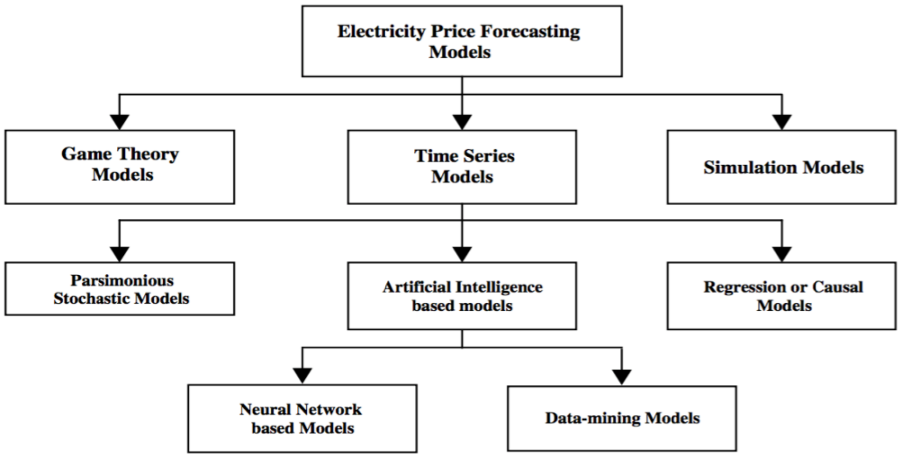
Fig. 1. Electricity market changes

Fazıl Gökgöz, Professor, Dr., Quantitative Methods Division, Department of Management, Faculty of Political Sciences, Ankara University, Turkey.
Fahrettin Filiz, MBA - MSc., Ankara University, Turkey.

Several methods have been proposed for forecasting electricity price, which are agent based modeling, time series, artificial neural networks, generalized autoregressive conditional heteroskedasticity (GARCH) models, ARIMA models and wavelet models. Among these methods, artificial neural networks have received more attention due to its simplicity, easy implementation and performance (Shahidehpour, 2002).

Rafal Weron classifies price forecasting models as regression methods, similar days, ARMA models and artificial intelligence methods (neural networks, support vector and fuzzy logic) (Veron, 2006).

Aggarwal et al. proposed a classification for electricity price forecasting models; game theory models, time series models and simulation models (Aggarwal, 2009). Nash equilibrium, Bertrand models and Cournout models are investigated under game theory models. In simulation models, major problem is the lack of adequate data for simulation. Time series models grouped in regression models, stochastic models and artificial intelligent models. Regression type forecast models is based on relationship between dependent variable and a number of independent variables. Stochastic models are Autoregressive (AR), Moving Average (MA), Autoregressive Moving Average (ARMA), Autoregressive Integrated Moving Average (ARIMA) and Generalized Autoregressive Conditional Heteroskedastic (GARCH). The artificial intelligent models are classified as data mining techniques and neural network models. Aggarwal et al. study electricity price forecasting models as shown in Figure 2.



Source: Aggarwal (2009).

Fig. 2. Electricity price forecasting models

Hybrid models can be applied to electricity price forecasting models. For example, ARIMA and neural networks are used together for the electricity price estimation of Brazil electricity market and it gives successful a result compared to GARCH, ARIMA or neural network models (Filho, 2014).

2. Methodology

2.1. Artificial neural networks. Inspired by the current knowledge of brain structure, artificial neural networks are mathematical nonparametric and nonlinear models that map the input and output relationship without exploring the underlying process. Neural networks learn from experience and produce input output relationships, which are unknown or hard to recognize. In this section, a short introduction to neural networks is provided.

The basic unit in neural networks is artificial neurons. Artificial neurons process information and produce output. The connections between neurons are characterized by weight coefficient. A single neuron's mathematical model is represented in Figure 3.

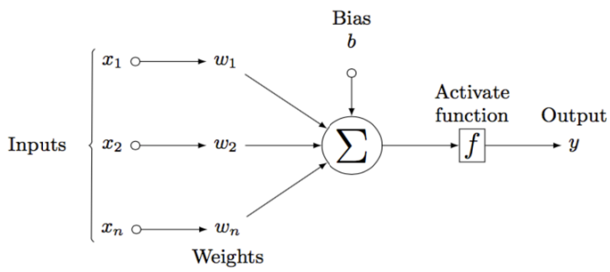


Fig. 3. Mathematical model of a nonlinear neuron

In mathematical terms, a single neuron can be described by the following equation:

$$y = f\left(\sum_{i=1}^{n} x_i w_i + b\right), \tag{1}$$

where x_i is input data, w_i is weight of neuron, b is bias and f is activate function and y is the output of neuron.

Sigmoid function is generally used in neural networks as activate functions to produce output in a specified range. Another important feature of sigmoid function is differentiability. Differentiability is an important feature in network learning algorithms, because, generally, algorithms uses first order or second order derivatives. Sigmoid function is defined by the following formula

$$f(x)=\frac{1}{1+e^{-\beta x}}. \quad (2)$$

In network learning, the weights are adjusted in accordance to a learning algorithm with the help of training inputs. A common network learning technique is working on calculation error, starting from output layer down through hidden layer. It is named as back propagation of error with modified delta rule (Rumelhart, 1986). The main aim of all algorithms is aim minimizing the error. A frequently used error function can be defined as follows.

$$\varepsilon=\frac{1}{2}\sum_{k=1}^{n}(o_k-t_k)^2, \quad (3)$$

where n is the total number of output nodes, o_k is the network output at the k th output node and t_k is the target output at the kth output node. Training algorithms attempt to reduce the global error by adjusting weights and biases. Various training algorithms have been proposed to improve the neural network learning procedures such as gradient descent, conjugate gradients, Quasi-Newton and Levenberg-Marquardt.

Gradient descent is a standard back-propagation algorithm in which the network weights are moved along the negative of the gradient of the performance function. The back propagation algorithm with gradient descent is given as follows:

$$\Delta w_k=-\alpha_k g_k, \quad (4)$$

where Δw_k is a vector of weight changes, g_k is the current gradient, α_k is the learning rate that determines the length of the weight. The major disadvantages of standard back propagation are its relatively slow convergence rate and being trapped at the local minima (Azar, 2013). In order to avoid oscillations and to reduce the sensitivity of the network, there is a momentum term added to gradient descent algorithm as shown in the following formula:

$$\Delta w_k=-\alpha_k g_k+p\Delta w_{k-1}, \quad (5)$$

where p is the momentum parameter. Furthermore, the momentum allows escaping from small local minima. The gradient descent and gradient descent with momentum do not produce the fastest convergence and they are even too slow to converge.

There has been considerable research on training methods to speed up the convergence of neural networks. These techniques include such ideas as varying the learning rate, using momentum and rescaling variables. There are four types of algorithms that are commonly used to minimize the network error. These methods are steepest descent, conjugate gradients, Quasi-Newton and Levenberg-Marquardt (Fine, 1999).

If the error function is truly quadratic, Newton's method can be used to minimizing weight vector in a single step (Fine, 1999). While the gradient descent algorithm requires only information of the first partial derivatives, Quasi-Newton algorithm produces improved convergence with second derivatives.

The Newton step size w_k for a second order Taylor series approximation of $f(x)$ at any current point w_i is obtained from the following equation

$$\Delta w H(w_i)=-\nabla f(w_i). \quad (6)$$

Assuming that the Hessian matrix (second derivative matrix) is non-singular, Eq. (6) can be written as

$$\Delta w=-\nabla f(w_i)H(w_i)^{-1}. \quad (7)$$

Broyden, Fletcher, Goldfarb and Shanno (BFGS) algorithm has been the most successful algorithm in Quasi - Newton method studies (Prasad, 2011).

The Levenberg-Marquardt (LM) algorithm is the most widely used optimization algorithm. It outperforms gradient descent and conjugate gradient methods. Levenberg-Marquardt is a popular alternative to Quasi - Newton technique. Unlike the Quasi-Newton technique, Levenberg-Marquardt algorithm uses approximate Hessian matrix with the following equation:

$$H=J^tJ+\mu I. \quad (8)$$

J is the Jacobian matrix that contains first derivatives of the network errors with respect to weights.

The Levenberg-Marquardt (LM) uses this approximation to get updated weighted matrix as in the following formula:

$$w_{k+1}=w_k-\left[J^tJ+\mu I\right]^{-1}J^te, \quad (9)$$

where w is the weight vector matrix, I is the unit matrix, μ is the combination coefficient, J is Jacobean matrix, e refers to the error vector. If μ is zero, this becomes the same as the Newton method, when μ is large, this equation becomes a gradient descent with small step size.

There are many different learning algorithms that are not explained here such as the resilient methods and conjugate gradient. Table 1 lists commonly used neural network learning algorithms that are going to be used in this study.

Table 1. Training algorithms used in models

Algorithm	Formula
Gradient descent	$w_{k+1} = w_k - \alpha_k g_k$
Gradient descent momentum	$w_{k+1} = w_k - \alpha_k g_k + p\Delta w_{k-1}$
Broydan, Fletcher, Goldfarb and Shanno (BFGS)	$w_{k+1} = w_k - \nabla f(w_i) H(w_i)^{-1}$
Levenberg-Marquardt	$w_{k+1} = w_k - [J^t J + \mu I]^{-1} J^t e$

Artificial neural networks in forecasting. Forecasting in economy (electricity load and prices, stock price, rate of inflation) is a complex task because of interconnected and interdependent parameters such as human reaction and social events. There is no complete model that covers all potentials of forecasting. Neural networks' nonlinear property and complex input output mapping facility make it a popular forecasting tool.

The idea of neural networks for forecasting starts in 1964. Hu (1964) in his thesis uses Widrows adaptive linear network for weather forecasting (Guoqiang, 1998).

Neural networks have been employed in volatility forecasting, risk rating of bonds, stock market predictions, option pricing and inflation forecasting (Habib, 2014). In addition to electricity price forecasting with neural networks in the deregulated markets has been achieved with reasonable accuracy (Singhal, 2011).

2.1.1. Artificial neural network design. Despite of the many satisfactory characteristics of neural networks, building a neural network for a particular forecasting problem is a challenging task. Generally, artificial neural network design procedures include the following steps:

- The selection of architecture.
- The selection of input, hidden and output nodes.
- The selection of layers.
- The selection of activation functions.
- Data preprocessing methods.
- Training and test sets.
- The selection of training algorithm.
- The selection of performance measures.

Design decisions on the items listed will affect the network performance. The important question here is how to develop a specialized structure in neural networks since there are no well-defined rules, but rather ad-hoc procedures yield useful results (Haykin, 1999).

Preprocessing makes forecasting problem more manageable. By preprocessing the data, it can be simplified before the actual calculations. Getting rid of the noise of the input data will affect the performance of the network positively.

Based on architecture preference, artificial neural networks generally grouped into two categories as feed-forward networks and recurrent networks (Jain, 1996). In feed forward networks, the output of one layer is used as the input to the following layer. In recurrent networks, every layer can take input from another layer. The feed forward networks are generally preferred for forecasting.

Generally, neural networks consist of input, hidden and output layers. For the number of hidden layers, Zhang experiments with networks with more than two hidden layers, but it does not provide significant improvement (Guoqiang, 1998). Also the number of neurons in layers has to be determined by heuristic way. The most common way to determine the number of neurons in layers is via experiments with trial and error.

Another design factor is the sample size to train and test the model. The larger the sample size, the more accurate the results will be calculated. In reality, the sample size is defined also by the availability of data.

There is no training algorithm currently available to guarantee the optimal solution for a general nonlinear optimization problem. The most popular training method is the back propagation algorithms.

Transfer functions limit the output of neurons. These functions must be differentiable and non-decreasing. Most papers use either logsig or the tansig functions.

Mean absolute percentage error (MAPE), the weighted mean absolute percentage error (WMAPE), the mean absolute error (MAE) and root mean square error (RMSE) are widely used in network performance. Commonly used network performance measurement methods are shown in Table 2.

Table 2. Performance measurement methods

Method	Algorithm
MAE	$\frac{1}{n}\sum_{t=1}^{n} \lvert A_t - F_t \rvert$
MAPE	$\frac{1}{n}\sum_{t=1}^{n} \lvert \frac{A_t - F_t}{A_t} \rvert$
WMAPE	$\frac{1}{n}\sum_{t=1}^{n} \lvert w * \frac{A_t - F_t}{A_t} \rvert$
RMSE	$\sqrt{\frac{1}{n}\sum_{t=1}^{n} \lvert A_t - F_t \rvert}$

A_t indicates the real value and F_t indicates the estimated value.

An artificial neural network is used in many different areas. It has different designs, since there is absence of certain assumptions about the network design. To determine the correct input and outputs, design decision criteria's should be analyzing the problem deeply. Algorithms to be used, neuron number, layer number, and performance evaluation methods should all be considered to find a suitable model.

3. Data and empirical findings

3.1. Data. In an actual electricity market, price curve exhibits considerably volatile structure and has the following characteristics: high frequency, nonconstant mean and variance, multiple seasonality, calendar effect, high level of volatility and high percentage of unusual price movements (Aggarwal, 2009).

Electricity price data for Turkish Electricity Market is acquired from the Energy Exchange Operations Authority of Turkey (EPIAS). Electricity price data between 01.01.2012 and 31.12.2014(26,304 hours) are selected as input data. The input data are divided into the two categories as follows:

- 01.01.2012-03.03.2014 is used (19,008 hours) is used as training data.
- 03.03.2014-31.12.2014 (7,296 hours) is used as testing data.

Each one of the neural network models created in this study is trained and tested with this same data.

After applied 19,008 hours training data and 7,296 hours testing data, the actual data and model results are compared according to MAPE values. Average of the 7,296 MAPE value accepted as model's MAPE value.

3.2. Factors considered in price forecasting. Appropriate selection of input factors is the keypoint in electricity price forecasting. Historical load and weather temperature represent the most important inputs. Proposed neural network architecture in this study is designed based on previous loads, type of day, hours of a day, previous day's temperature and natural gas prices.

Hourly weather temperature data usage would increase the reliability of the model, but there is no hourly weather temperature data available between dates 01.01.2012 and 31.12.2014. Due to this reason, daily temperature data are used as input to the forecasting model and these data are obtained from www.wunderground.com.

Electrical load consumption of cities changes for each other according to population and industrialization level. Therefore, temperature of the city should not be considered alone, the amount of electricity consumed by the city with the temperature factor should also be considered. Figure 4 shows Turkey's electricity load demand percentage by cities.

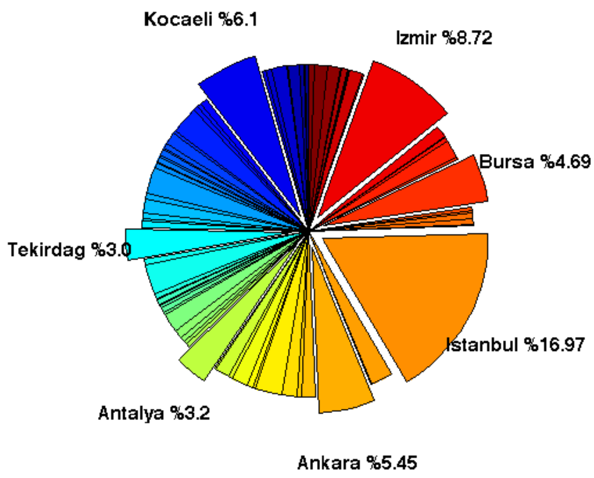


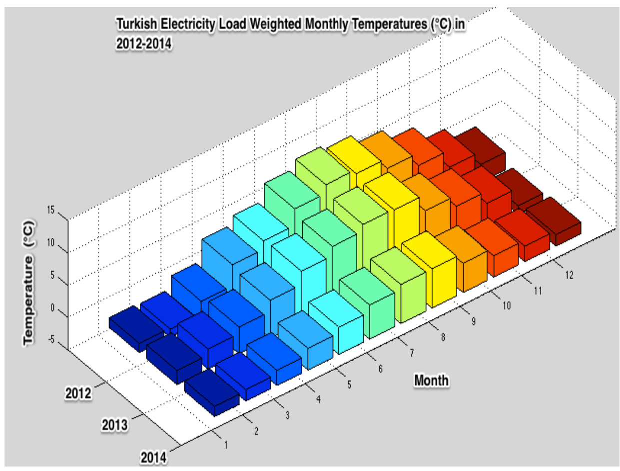
Fig. 4. Electricity load demand percentages by cities in 2011

Temperature is used as input to the model using province-based electricity consumption. Equation 10 shows the temperature formula weighted by the electricity consumption of the city

$$\tau_m = \sum_{i=1}^{81} w_i t_i, \tag{10}$$

i value in the formula shows city in Turkey which has 81 city, w_i is the percentage of electricity consumption in the city, t_i the temperature of the city and m shows the day, τ_m represents the electricity consumption weighted calculated temperature.

Electricity consumption weighted calculated temperature according to months is shown in Figure 5. Monthly average temperatures have increased in June, July and August 2012. A temperature value weighted by electricity consumption is lowest in 2014.



Source: author's calculation.

Fig. 5. Turkey electric load weighted monthly temperatures

Big deal of the electricity used in Turkey is produced from natural gas. Due to the dependence on natural gas in electricity generation, natural gas prices are affecting electricity prices in Turkey. Natural gas prices at the global level generally vary with Japanese LNG prices, the average price of gas imported from Germany and the Henry Hub in United States prices accepted (Cangüzel, 2012). Due to the lack of day-ahead gas market in Turkey, electricity price forecasting model uses data from United States Henry Hub Natural Gas Spot.

3.3. Implementation of the price forecasting model. Forecasting models for commonly used models of artificial neural networks feed forward multilayer networks (Zhang, 2004). In this study, the amount of price forecasting models for electric day-ahead market, the three-layer feed-forward neural network is optimized using different algorithms. After using 4 different learning algorithms, two different transfer functions and 50 different number of hidden layer neurons, total of 400 different models are created. Electricity price forecasting design parameters are presented in Table 3.

Table 3. Price forecasting implementations

Training algorithm	Transfer function	Number of neuron in hidden layer
Gradient descent	Tansig/Logsig	1,2...50
Gradient descent momentum	Tansig/Logsig	1,2...50
Levenberg-Marquardt	Tansig/Logsig	1,2...50
BFGS	Tansig/Logsig	1,2...50

In the proposed architecture, neural network is designed based on previous load, type of day, hours of a day, load consumption weighted average temperature, previous day prices. Figure 6 shows the model entries, the hidden layer neuron number and model outputs.

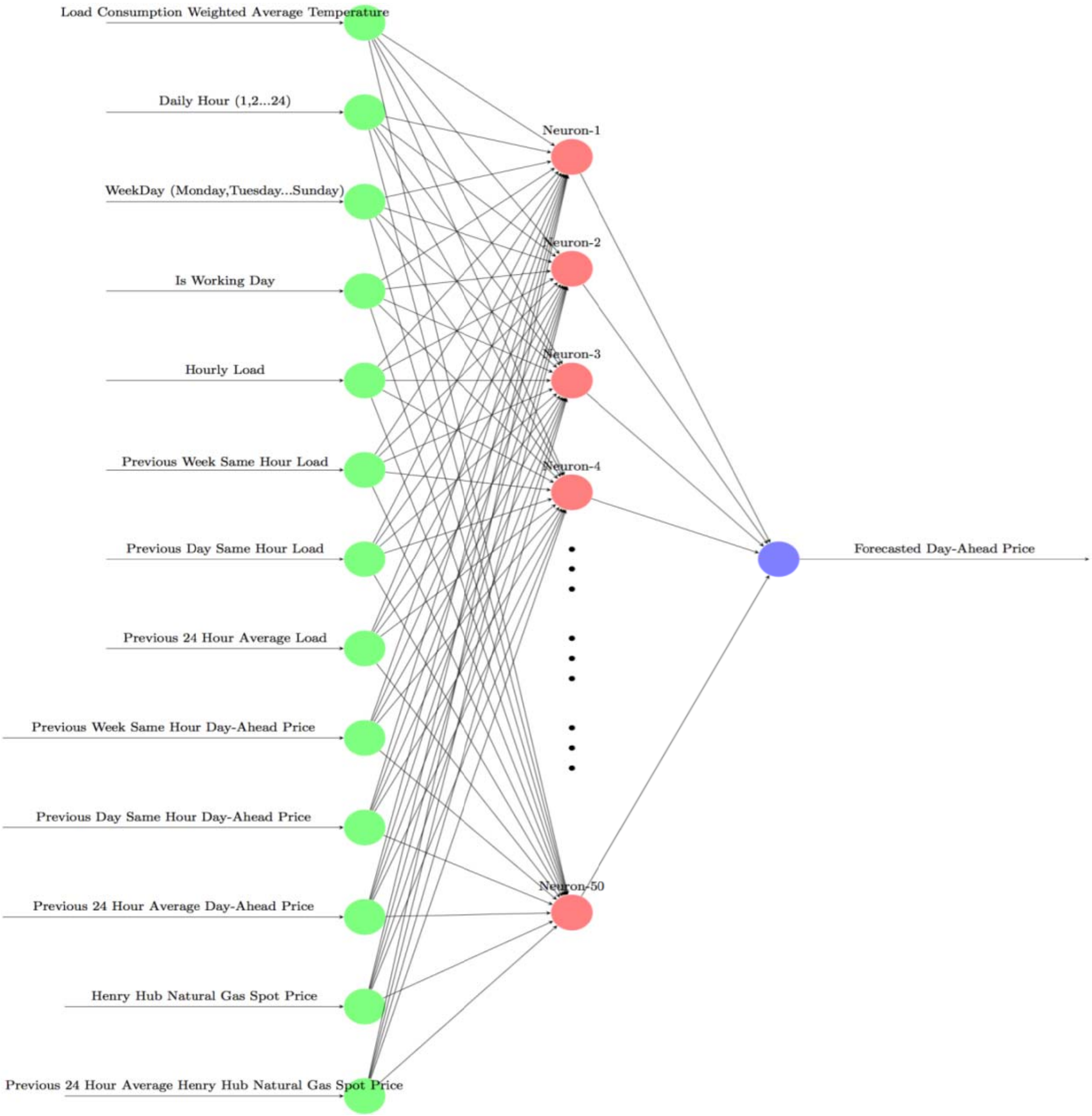


Fig. 6. Electricity price forecast neural networks models

Input data is divided into two categories, 72% hours of total data trains the models and %28 hours of total data tests the model. Figure 7 shows data with testing and learning parts.

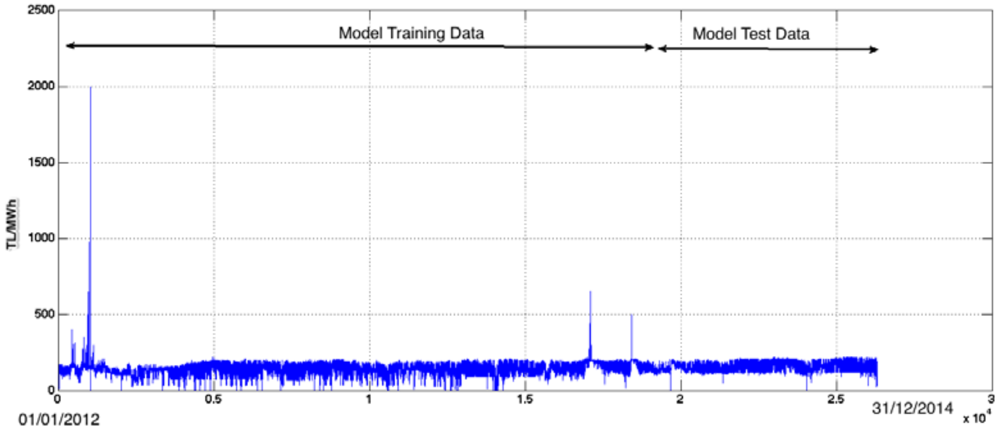


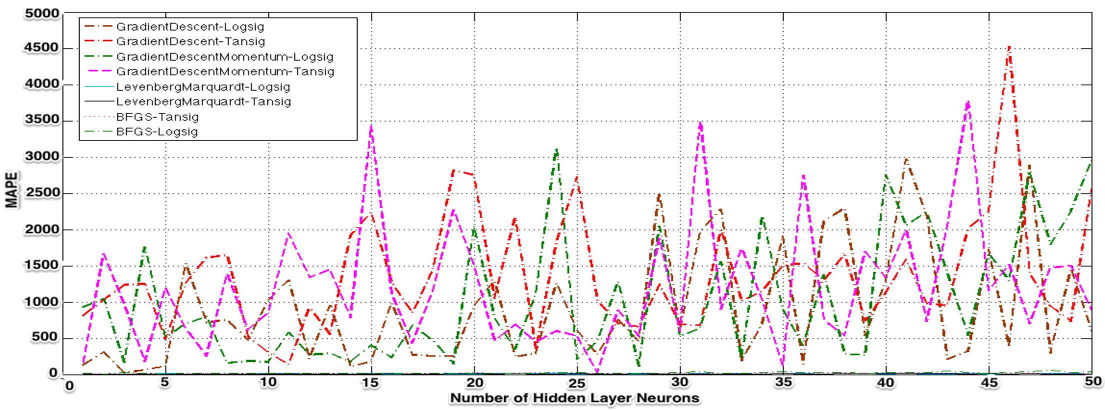
Fig. 7. Model training and test data

3.4. Numerical results in price forecasting. The lowest MAPE value of models is 9.76% by the Levenberg-Marquardt algorithm. This value is achieved by number of 17 hidden layer neuron and tansig transfer function. When we look at the average MAPE values, then, we can see again Levenberg-Marquardt algorithm, but with logsig transfer function. Best average MAPE value is 10.57. All MAPE results are given in Table 4.

Table 4. Price forecasting algorithms MAPE results

	Minimum MAPE %	Maximum MAPE %	Average MAPE %
GDM Logsig	113,9158	3,1354e+03	1,0274e+03
GDM Tansig	23,5656	3,7922e+03	1,2027e+03
GD Logsig	23,1152	2,9839e+03	908,6795
GD Tansig	137,4356	4,544e+03	1,3635e+03
LM Logsig	9,8478	11,6847	10,5766
LM Tansig	9,7619	11,7166	10,6384
BFGS Logsig	9,8605	30,1433	14,8745
BFGS Tansig	13,100	39,6000	21,5320

MAPE result shows that day-ahead price forecasting in deregulated Turkish market give a reasonable range. MAPE values of all algorithm used in models depending on number of hidden layer neuron are shown in Figure 6.

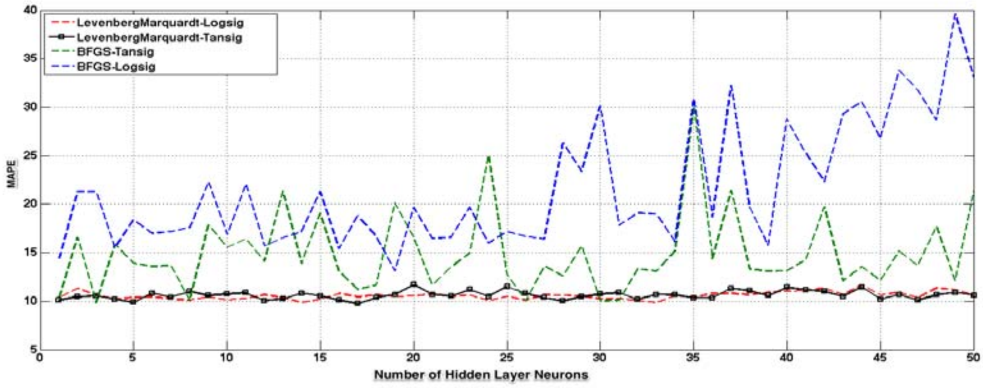


Source: author's calculation.

Fig. 8. Number of hidden layer neurons and transfer functions price forecast results

Because Levenberg-Marquardt and BFGS algorithms have more successful MAPE value, they are shown separately in Figure 9. When compared to the variation of the hidden layer neurons BFGS and Levenberg-Marquardt algorithm Levenberg-Marquardt MAPE results are analyzed algorithm seems to change values in a fixed range. BFGS algorithms increase in

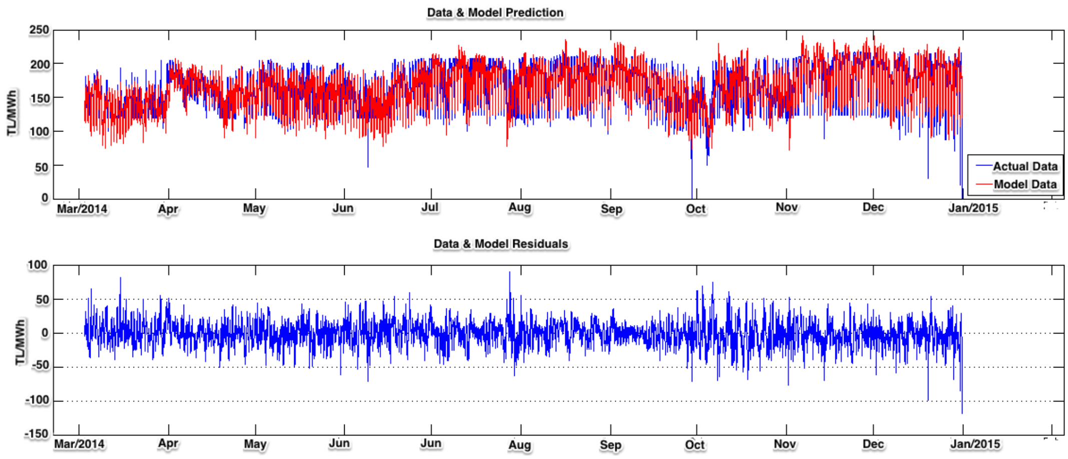
MAPE values increased by the number of hidden layer neurons is observed. Levenberg-Marquardt algorithms generally seem to be more successful than BFGS algorithms.



Source: author's calculation.

Fig. 9. Number of hidden layer neurons and transfer functions price forecast results

Models before 7,296 hours on the day of test data produced by the data in real time data on electricity price forecasts are shown in Figure 8. With model data resulting from differences between actual average 7,296 hours of test data "-1.12 TL / MWh" calculated.



Source: author's calculation.

Fig. 10. Actual data & model prediction data comparisons

Error rate on 29.09.2014 exceeds the rate of 1000%. We limit the graphics to 1000% rate for better explanation. Model is trying to estimate day-ahead price on 29.09.2014 that its actual value is 0.92 TL / MWh. The models estimated value is about 73.13 TL / MWh on 29.09.2014. Another lowest price, which is 0.79 TL / MWh, has occurred on 31.12.2014 at 23.00. Model estimation for corresponding time is 119.88 TL/ MWh. So the abnormal day-ahead price of electricity makes estimation negatively. Between 03.03.2014-31.12.2014 days (7,296 hours of test data) the histogram is shown in Figure 11. Day-ahead electricity price data generally changes between 100 TL / MWh to 250 TL / MWh.

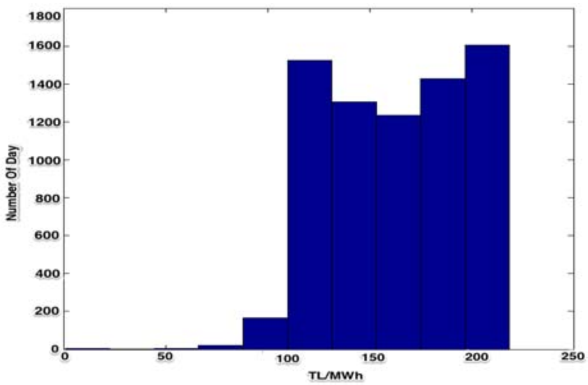


Fig. 11. Histogram of day-ahead electricity prices

Conclusion

This study investigates different artificial feed-forward neural networks models for achieving the best forecasting results for Turkish day-ahead electricity market. We employed 13 factors with 26,304 hours historical data and created 400 different neural network models. Besides, 26,304 hours data for each model have been used to reach most successful artificial feed forward model. Models success is calculated according to the models MAPE values.

In this study, the most successful MAPE values obtained for the electrical day-ahead price is 9.76 %. Aggarwal et al. review on electricity price forecasting in deregulated market and compare neural network performance. In different price forecasting studies, MAPE values changes between 2.18% and 25.77% (Aggarwal, 2009). In addition to these studies, artificial neural networks would be a significant alternative for financial decision makers in forecasting the day-ahead electricity price.

Gradient descent, gradient descent momentum algorithms in neural network models are more sensitive to weight changes or abnormalities in their input to calculate the first order derivatives. Levenberg-Marquardt and BFGS methods uses second order derivative that produce better results against the data anomalies.

Selection of input variables for a particular model is still an open area of research. Further research can include selection of more appropriate data classification.

References

1. Anil, K., Jain, J.M. (1996). Artificial Neural Networks: A Tutorial, *Computer*.
2. Azar, A.T. (2013). Fast neural network learning algorithms for medical applications, *Neural Computing and Applications Applications*.
3. B.D.C.N. Prasad, P.K. (2011). A Study on Backpropogation Models through Hidden Units, *Advances In Computational Sciences and Technology*.
4. Cangüzel, T.A. (2012). European gas pricing system and differences in global gas prices (*Avrupadogalgazfiyatlandırmasistemiveküreselgazfiyatlarıfarklılıkları*), Turkish Chamber of Physics Engineers.
5. David, E., Rumelhart, R.G. (1986). Learning representations by back-propagating errors, *Natura*, 323, p.533.
6. Filho, J.R. (2014). Energy price prediction multi-step ahead using hybrid model in the Brazilian market, *Electric Power Systems Research*.
7. Fine, T.L. (1999). *Feedforward Neural Network Methodology.* Springer.
8. Guoqiang, Z.P. (1998). Forecasting with artificial neural networks: The state of the art, *International Journal of Forecasting*.
9. H.Y. Yamin, S.S. (2004). Adaptive short-term electricity price forecasting using artificial neural networks in the restructured power markets, *International Journal of Electrical Power and Energy Systems*.
10. Habib, A.H.S. (2014). Financial forecasting with neural networks, *Academy of Accounting and Financial Studies Journal*.
11. Haykin, S. (1999). *Neural Networks: A Comprehensive Foundation*. 2nd. Edition, Prentice Hall.
12. Lahmiri, S. (2011). A Comparative Study of Backpropagation Algorithms in Financial Prediction, *International Journal of Computer Science, Engineering and Applications (IJCSEA)*, 1(4).
13. Müller-Jentsch, D. (2001). *The Development of Electricity Markets in the Euro-Mediterranean Area.* The World Bank.
14. Mohammed Shahidehpour, H.Y. (2002). *Market Operations in Electric Power Systems.* Willey-Interscience.
15. Sanjeev Kumar Aggarwal, L.M. (2009). Electricity price forecasting in deregulated markets: A review and evaluation, *International Journal of Electrical Power and Energy Systems*.
16. Singhal, D. (2011). Electricity price forecasting using artificial neural networks, *International Journal of Electrical Power and Energy Systems*.
17. Veron, R. (2006). *Modeling and Forecasting Electricity Loads and Price.* John Wiley & Sons Ltd.
18. Zhang, G.P. (2004). *Neural Networks in Business Forecasting.* Idea Group Inc.