

# Forecast for Market Clearing Price with Artificial Neural Networks in Day Ahead Market

Oguz Tonyali and Duygu Bayram Kara


**Abstract**— In this study, the Market Clearing Price (MCP) is forecasted with Artificial Neural Networks (ANNs) and the modeling success is examined for different preprocessing strategies. The purpose of the study is to obtain an optimum model with a significant estimation success and to provide the best price prediction. The hour-based electricity generation data of diverse production items are assigned as inputs and the resulting MCP is modeled. The raw data are first cleaned from outliers, then normalized through 2 different strategies. After normalization, seasonal grouping is applied and 5 groups (annual and four seasons) are achieved to reveal the effect of seasonal patterns on model success. For resulted 10 datasets, ANNs having 7 different hidden layer sizes are trained. As a result, 70 different ANNs are trained and their results are presented. Finally, networks showing the optimum performance are selected. It is noted that the type of the normalization strategy and the hidden layer size are the key factors to make a decent estimation. Then, in order to test the networks with extreme cases, data for the special days (official holidays) are applied to these networks as input. The success of the networks is evaluated by comparing the MCP predictions with the actual values. It is revealed that in order to make a prediction for official holidays, a unique model, which is special to this period of year, is required.

**Index Terms**— Artificial Neural Networks, Day Ahead Market, Electricity Market, Market Clearing Price.


## I. INTRODUCTION

WITH THE increasing need of electrical energy, the cost for the generation of electricity gains importance. The renewable energy sources, as the most important alternative to depleted fossil resources, produce variable amounts of energy. In case of over production, storage cannot be applied due to the

OGUZ TONYALI, is with Artesis Innovative Industrial Maintenance Solution Corp., İstanbul Turkey, (email: [tonyali@itu.edu.tr](mailto:tonyali@itu.edu.tr))

 <https://orcid.org/0000-0002-2604-519X>

DUYGU BAYRAM KARA, is with Electrical Engineering Department of İstanbul Technical University, İstanbul, Turkey, (e-mail: [bayramd@itu.edu.tr](mailto:bayramd@itu.edu.tr))

 <https://orcid.org/0000-0001-8184-8510>

Manuscript received April 28, 2021; accepted September 10, 2021.  
DOI: [10.17694/bajece.929564](https://doi.org/10.17694/bajece.929564)

excessive cost, while in case of insufficient production, the required energy cannot be provided. In order to prevent this, some of the power plants are hold available to be engaged, if needed. However, such an application is very costly considering the limited working hours of these plants.

In order to manage all these system components and regulate costs, institutions and governments around the world have gradually liberalized electricity markets. In the liberalized markets; consumers always have access to electricity, although prices may vary. At the same time, waste of extra labor, raw materials, time and money is prevented.

By the advantages of liberalized energy market, price prediction has become essential. The price is determined according to the point where generation and consumption amounts are balanced. Purchases start to be made at the lowest price from the producers, until supply-demand balance has been achieved. In other words, consumer needs determine the price of the electrical energy, which is the market price. On the other hand, the day ahead market is also used in designation the of market price. The day ahead market is operated and organized by the operator (Electricity Markets Operating Corporation – EPIAŞ), used for electricity trading and balancing activities, using the previous day electricity and price data. Participants of this network enable the determination of the electricity reference price. These are the institutions that provide the generation, transmission, distribution and sale of energy. The day ahead market, as the name suggests, provides the prices and the productions divided into hourly based segments of the day. It covers a day that starts with midnight (00:00) and continues until the beginning of the next day. To be involved in this process, participants make separate price offers for each hour segment through their own estimations. Participants can only bid for the next day or upload their bids for the next 5 days to the system beforehand. However, they have the right to change them until 12.30 on the previous day. Participants who submit daily price offers must also afford their guarantees until 12.30 to the market operator [1].

Especially when it comes to the determination of prices and distribution of earnings, some special arrangements are adopted within the framework of bilateral agreements. These agreements are needed due to the storage problem of electricity, the unpredictable nature of renewable resources and the

variable prices for some resources such as natural gas and coal [2]. For all of these reasons, day-ahead pricing is of critical importance [3].

In order for the participants and end users in the market to buy and sell electricity at a fair price, offers and pricing from the day ahead market is made through a software. With this software, not only the price but also the amount of energy to be transferred to the network on hourly basis is determined. Market participants submit their bids via software on a block, hourly or flexible basis. Hourly bids are given with the price and quantity for the specified hour. Block bid consists of multiple hourly bids including the quantity and price. Block bids can also be submitted depending on another bid, but if the linked bid is not accepted, the second block bid is also canceled. For flexible bid, the quantity and price are just specified without any other timely manner claim. All bids are given by using lot and lot price is in TL / MWh. To determine the price, the supply-demand curve is basically taken into account. Except for flexible bids, matches are done hourly. Some mathematical models are adopted estimating the breakpoints, insufficient production (excess demand) or excess supply conditions. Later, if the bid is lower than the lowest reference price, then it is accepted as the new reference price and saved in the system. In cases of excess supply or demand, the market clearing price (MCP) is created by extra calibration to the reference price. The mathematical background of the software, which determines MCP, includes preprocessing step to clean the outliers and optimization step including two heuristic algorithms.

For the energy producers, being able to make the best estimation for the MCP is very important in order to claim a profitable bid. Thanks to ANN, a high rate of accuracy can be achieved to increase the profit margin by the best price estimation. It has been easier to make these predictions with ANNs in the last 30 years. ANNs represent a black box model of the system in case of a nonlinear relationship between input and output vectors.

As the literature is reviewed, dozens of studies are met in this field. For example, Mandal and his team predicted MCP with ANN by applying the similarity method which classifies similar consumption patterns. Their algorithm operates very well even at very high and low power levels [4]. Zolotova and Dvorkin presented a prediction study with multi-layer ANN models in accordance with the stochastic structure of the free energy market in Russia. As the input data, the price of the previous hour, day of estimation, whether it was a holiday or not, and the pricing of the desired hour in the previous week were used [5]. Nargale and Patil have conducted a MCP estimation study to increase profit margin in the Indian electricity market. One to four-layer Feed-forward ANN modeling has been performed because it can easily learn multiple input-output relationships and work with supervised learning. Three hidden layers were chosen as the optimum solution [6]. Saâdaoui has studied the model of the NORD Pool market, which is one of the largest energy markets globally and generally covers Northern European Countries. This study was needed because of the large number of variables such as; the number of countries,

climate change, hours of use, etc. A feed-forward ANN is used in the study to model nonlinear relationships met at this huge market. High prices on special days have also been checked and sorted out according to their reasons. And the success of the prediction in special values has also been observed [7]. Keleş and his team conducted an ANN based day ahead MCP prediction study for the European Electricity Market (EPEX). Fuel prices, renewable energy sources, weather conditions, electricity load and special dates (holidays, events, etc.) were used as input data in addition to the older market prices [8]. Şenocak and Kahveci estimated MCP by using the data set including the generation amounts of the major resources such as; natural gas, lignite and liquid fuels (fuel-oil) and their prices for the years 2009-2014. Unlike other studies, monthly estimates were made instead of hourly or daily estimates. The data is modeled with both ANN and ANFIS (Adaptive Neuro-Fuzzy Inference System) and their success are compared using Mean Absolute Percentage Error. ANFIS showed a better performance thanks to its adaptive nature [9]. Dalgin made an ANN based MCP prediction for special days (official holidays) by considering the effects of climate, temperature, gas prices, humidity and the pricing conditions in the previous year. This model has been executed for different hidden layer sizes and tested for a random day data [10].

As seen from the given examples in the literature, ANN trained with Back Propagation (BP) algorithm, are widely used for MCP estimation. In order to prevent overfitting and increase the accuracy, modelling the data regarding some criteria such as season, weather and holiday periods is highly recommended by Gao and his team [11]. According to Georgilakis, estimating the hourly load and the MCP with two different ANNs is another way to increase the prediction success [12]. On the other hand, for a qualified MCP and confidence interval estimation, Extended Kalman Filter [13] or Wavelet transform [14] are also employed for adaptive ANN training.

Additionally, there are some factors to be considered for the most accurate MCP forecast through ANN; these are historical MCPs, system loads and fuel prices [11, 15].

In this study, the effect of different normalization strategies on the success of ANN is investigated for the MCP estimation purposes. In order to train the network, hour based generation data and associated MCPs are used. The aim of the study is identifying the best preprocessing strategy to reduce the costs by increasing the accuracy of price estimates. In order to achieve this goal, electricity generation data and matching MCPs for 2018 were obtained through the EPIAŞ Transparency Platform.

## II. FEED FORWARD NEURAL NETWORK (FFNN)

Feed Forward Neural Network (FFNN) is the most common ANN type employed in MCP estimations. A simple FFNN topology may contain a large number of neurons, but it basically consists of three different layers. These are input, hidden and output layers. In this type of ANN, the information flow is unidirectional. The input data is processed using activation (transfer) functions in the hidden layer, and the data

is thus converted into information. The predictive success of the FFNN model is highly dependent on the hidden layer, as it models the complex relationships between input and output. Depending on the complexity of the prediction problems, different sizes of hidden layers can be considered in the FFNN topology.

The output layer receives processed information from hidden layers and transmits it to the output layers with linear combinations. Nodes between each neuron have interconnection weights that have a significant effect on the output response of the FFNN model. A simple neuron representation is presented in Fig. 1 [16].

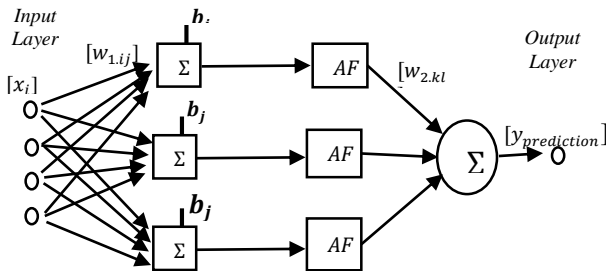


Fig.1. Simplified representation for an Artificial Neuron

The mathematical operation achieved in a neuron is given in Eq. (1). The input data (for  $i$  number of inputs and  $j$  hidden layers) is transmitted to hidden layer by weights with  $w_{ij}$  matrix. The weighted inputs are fed by certain biases  $b_j$  and applied to the activation function. Afterwards, they are re-weighted with  $w_{2,kl}$  and transmitted to the linear activation function. Following this operation, the estimation of the ANN is formed [17 - 18].

$$y_{prediction} = f_2 \left( \sum_{j=1}^m w_{2,kl} \cdot \left( f_1 \left( \sum_{i=1}^n w_{1,ij} \cdot x_i + b_j \right) \right) \right) \quad (1)$$

### III. FFNN MODEL FOR MCP FORECAST

In this study, electrical energy generation data, which was provided by EPIAŞ, is assigned as input and MCP is defined as an output. The input layer dimension is in accordance with the number energy sources (natural gas, coal, hydraulic, wind, etc.).

In order to prevent unnecessary electricity generation and to provide pointed bids, accuracy of the MCP estimation is vital. With producers retrenching their costs, consumers would have access to electric energy for more affordable prices. For this purpose, flow chart in Fig. 2 has been adopted for the study.

As first step of the pre-processing, partial deficiencies were removed in the data. For example, naphtha, which does not affect generation resources, is not included in the calculation. Solar and LNG generations were excluded also because of their low impact on hourly basis. By this way, less effective inputs are prevented from disrupting the analysis. As the second step, the different normalization strategies are used to minimize the prediction error.

The schematic for the established ANN is shown with inputs and output in Fig. 3. There are 12 input nodes representing different kinds of electric energy generation amounts. The network has single output as the MCP. The study provides a simple and handy ANN structure without any need for sophisticated preprocessing or classification operation.

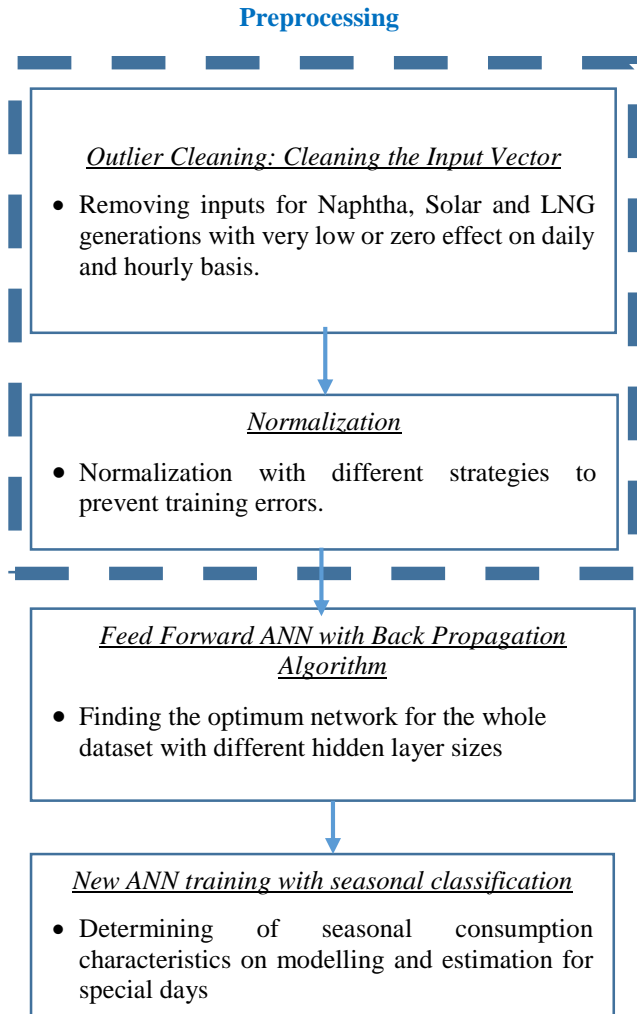


Fig.2. Flowchart for the study

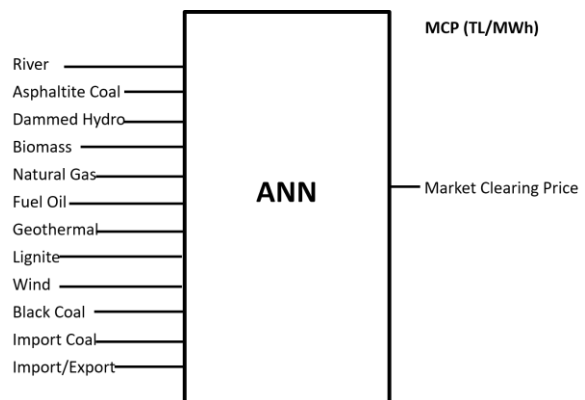


Fig.3. Established ANN for inputs and output variables

A. ANN Training with Different Normalization Strategies

Different strategies have been investigated to achieve accurate predictions. Initially, ANN was first trained using raw data and without any classification and normalization application. It was observed that the results were not satisfactory. In order to achieve decent results, new ANNs have been created with pre-processing and normalization actions.

First, the data are cleaned from outliers, then two different normalization processes (these are explained below) are determined. In order to work on the success of seasonal prediction, grouping is applied and 5 groups are achieved such as annual, winter, spring, summer and fall.

Performance tables have been arranged by reprocessing the training for different hidden layer sizes such as 2, 5, 8, 11, 14, 17 and 20. In the meantime, the learning rate is set as 0.5. For all trained networks 70% of the data set was used for training and 30% for testing. Levenberg Marquard algorithm was used in the training of networks, and the activation function is set as tangent sigmoid in the hidden layer. Each analysis is carried out with annual (whole year) and seasonal data. Thus, seasonal characteristics are modeled, as well. Fundamentally, 2 strategies are adopted for the data normalization:

1. Generation based normalization analysis: The electric energy generation data are proportioned to the highest energy generation value.

2. Time based normalization analysis: The electric energy generation data are proportioned to the total generation amount in the relevant hour. In other words, the individual contributions of the inputs are more focused to train the network.

The MCP value is also normalized in itself for both strategies. The networks are trained for different hidden layer sizes. The network is first trained for the annual data, then new networks were trained with the data allocated to the four seasons. For each normalization strategy, 35 networks were trained and training results are given in Tables 1 and 2. Performance values given in the tables are Mean Square Error (MSE) and *R*. The *R* value, which is correlation coefficient, is an indicator of the linear relationship between outputs and

targets. The closer it is to  $R = 1$ , the more linear relationship exists between outputs and targets.

For the generation based normalization analysis, all generation values are proportioned to the highest generated value. Considering the performance values, it is seen that the best network trained with annual data is the network with 17 hidden layers.

The *R* plots regarding the training, validation and test processes of this network are given in Fig. 4. In the ranking of success; networks trained with autumn, spring, winter and summer data followed the all-year data. However, the optimum hidden layer number varies in ANNs obtained by training with seasonal data.

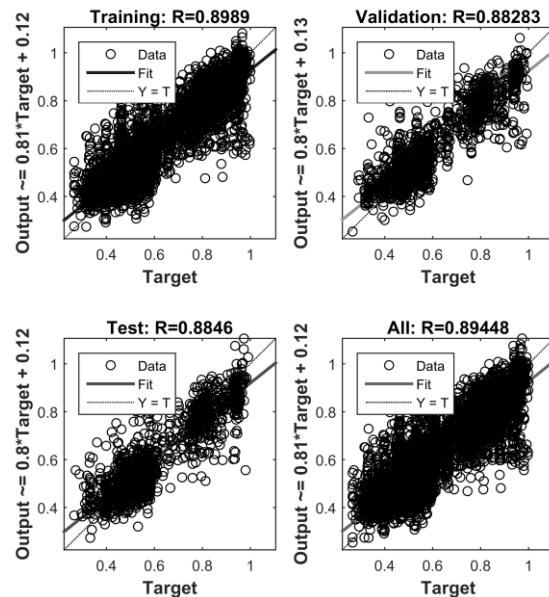


Fig.4. *R* plots of successful training for generation based normalization analysis

TABLE I  
PERFORMANCE TABLE FOR NETWORKS TRAINED WITH GENERATION BASED NORMALIZED DATA

Hidden Layer Size	Annual		Winter season		Spring		Summer		Fall	
	<i>R</i>	<i>MSE</i>	<i>R</i>	<i>MSE</i>	<i>R</i>	<i>MSE</i>	<i>R</i>	<i>MSE</i>	<i>R</i>	<i>MSE</i>
2	0.756	0.015274	0.806	0.0096721	0.745	0.0084028	0.726	0.015536	0.825	0.0075494
5	0.825	0.010923	0.851	0.0079368	0.764	0.0097491	0.811	0.014988	0.827	0.0069742
8	0.858	0.0087404	0.838	0.010267	0.791	0.008772	0.861	0.010574	0.823	0.0066041
11	0.871	0.008338	0.859	0.0087558	0.838	0.008822	0.870	0.010437	0.838	0.0076608
14	0.891	0.0073905	0.818	0.011821	0.785	0.0073203	0.892	0.010024	0.863	0.0081863
17	<b>0.898*</b>	<b>0.0068986*</b>	0.855	0.011181	0.858	0.0098968	0.891	0.010622	0.863	0.0070555
20	0.905	0.007321	0.845	0.010523	0.849	0.0071678	0.918	0.011229	0.860	0.0085836

For time-based normalization, hourly generation amounts are proportioned to the total generation value of the relevant hour. MCP value is also normalized by its highest value. The training

results can be seen in the Table II. The best training achievement is obtained by the network with 17 hidden layer again. Its *R* plots are given in Fig. 5. In seasonal terms, the

networks trained by autumn, spring, winter and summer data have been successful, respectively. However, the optimum hidden layer number varies in these ANNs, also.

When Tables I and Table II are evaluated, no outstanding success has been found in networks trained with seasonal

classification. It is clear that training with annual data is advantageous in terms of both generalization and ease of algorithm (ANN<sub>1</sub> and ANN<sub>2</sub>).

TABLE II  
PERFORMANCE TABLE FOR NETWORKS TRAINED WITH TIME BASED NORMALIZED DATA

Hidden Layer Size	Annual		Winter season		Spring		Summer		Fall	
	R	MSE	R	MSE	R	MSE	R	MSE	R	MSE
2	0.689	0.017706	0.768	0.010065	0.746	0.0081157	0.710	0.01668	0.836	0.0088811
5	0.806	0.012032	0.847	0.0081211	0.777	0.0091174	0.813	0.014284	0.847	0.0068576
8	0.842	0.0099064	0.839	0.01047	0.828	0.007903	0.852	0.012607	0.835	0.0065345
11	0.865	0.0087969	0.843	0.0092986	0.853	0.0078324	0.855	0.010531	0.844	0.0074099
14	0.867	0.0091751	0.845	0.012907	0.797	0.0066972	0.824	0.012835	0.840	0.0075855
17	<b>0.881*</b>	<b>0.0081772*</b>	0.874	0.011448	0.858	0.010501	<b>0.900</b>	<b>0.011225</b>	0.847	0.006678
20	0.881	0.0091475	0.815	0.01111	0.861	0.0078492	0.920	0.012144	0.852	0.008954

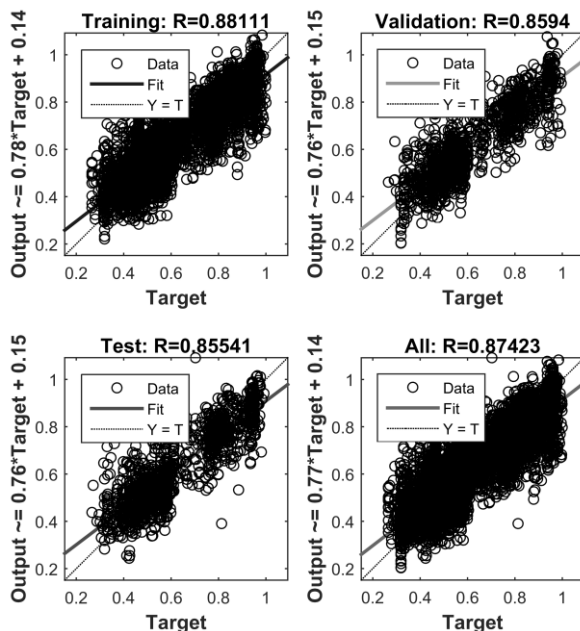


Fig. 5. R plots of successful training for time based normalization analysis

#### IV. MCP FORECAST FOR SPECIAL DAYS

Among all models, 2 ANNs (named as ANN<sub>1</sub> and ANN<sub>2</sub>) with 17 hidden layers (trained by annual data with both of the normalization strategies – marked as bold in the Tables) are tested with the energy generation values of the Ramadan and Eid al-Adha, where the special consumption conditions are valid. Since the dates of the 2018 religious holidays coincide with the summer (14 June Thursday-17 June Sunday and 20 August Monday-24 August Friday), the data is also applied in the network (named as ANN<sub>3</sub>) that is trained with time-based normalized data for summer season. This network's training

rankings can be seen in Table II, summer column written in bold.

Industrial facilities which are not working on holidays change pattern in electricity consumption. At the same time, the demand characteristics of household consumers change during holidays. This kind of energy consumption dynamic does not match with the general habits. Because of this, the success of trained ANNs for this special condition were expected low.

In order to verify this prediction and see the low success rate of the simulations in detail; results are presented with mean square error (MSE) and root mean square error (RMSE) on Table III.

TABLE III  
ESTIMATION SUCCESS OF OPTIMUM NETWORKS

	ANN <sub>1</sub>	ANN <sub>2</sub>	ANN <sub>3</sub>
MSE	0.1223	2.3856	5.0021
RMSE	0.3497	1.5445	2.2365

As a result of these error rates, it is seen that energy consumption on holidays can be modeled with ANN<sub>1</sub> and ANN<sub>2</sub> with limited success, since these networks are able to reflect the dynamics of the whole year. It has been observed that ANN<sub>3</sub> (ANN trained by time-based normalization of summer data) gives the highest error rate even with summer characteristics. The reason for this is that consumption values fluctuate and alter considerably during the holidays.

#### V. CONCLUSION

The increase in electric energy generation costs has brought along the need for liberalization of the electricity markets. It has become important for the liberalized market to estimate the MCP corresponding to generation with the highest accuracy. Accurate forecasts reduce the costs as well as manpower and raw material expenses. Today, ANNs are widely used to achieve accurate MCP estimation. The accuracy of the models

is highly dependent to the amount of well-defined data. In this study, MCP forecast is realized with ANN for the electricity generation data recorded at 2018. FFNNs trained with the back propagation algorithm are designed. The annual and the seasonal data are used. By this way 5 separate data sets (annual, spring, summer, winter, autumn) are modelled. The networks are trained with 7 different hidden layer sizes. The optimum networks are determined and MCP forecast is realized for special periods of the year (Ramadan and Eid al-Adha). As a result of the predictions, it is seen that annual networks are more successful on modelling these periods rather than seasonal network. However, the success is not sufficient to make a good MCP prediction globally. Thus it is obvious that there is a need to train another network which would make better predictions for the special days as suggested in the literature [11].

To sum up, the selection of network, preferred normalization strategy, preprocessing [14] and hidden layer size have critical importance for forecasting MCP. In addition, it is clear that for MCP estimation of special days (official holidays), an ANN trained with the data recorded for the relevant periods of previous years might be preferred. Also, considering the factors such as historical data, system loads and fuel prices might enhance the estimation accuracy [11, 15]. As the future work, authors are planning on predicting MCP with deep learning and machine learning techniques supported by adaptive algorithms [13-14] to increase the prediction success.

#### REFERENCES

- [1] Y. Biçen, "Türkiye elektrik enerjisi piyasası gelişim süreci: Gün öncesi ve dengeleme güç piyasası özellikleri." *Karaelmas Sci. Eng. J.*, vol. 6. 2, 2016, pp. 432–438.
- [2] *Electricity market Law* No: 28603, Official Gazette of the Republic of Turkey, 6446, 5 (53). March 2013.
- [3] *2017 EPIAŞ Activity report of the board*. 2 March 2018; Available from: <https://www.epias.com.tr/en/corporate/annual-reports/>.
- [4] P. Mandal, T. Senjyu, N. Urasaki, T. Funabashi, and A. K. Srivastava, "A Novel Approach to Forecast Electricity Price for PJM Using Neural Network and Similar Days Method." *IEEE Trans. Power Syst.*, vol. 22. 4, 2007, pp. 2058–2065.
- [5] I. Y. Zolotova and V. V Dvorkin, "Short-term forecasting of prices for the Russian wholesale electricity market based on neural networks." *Stud. Russ. Econ. Dev.*, vol. 28. 6, 2017, pp. 608–615.
- [6] K. K. Nargale and S. B. Patil, "Day ahead price forecasting in deregulated electricity market using Artificial Neural Network." in 2016 International Conference on Energy Efficient Technologies for Sustainability (ICEETS), India, 2016.
- [7] F. Saâdaoui, "A seasonal feedforward neural network to forecast electricity prices." *Neural Comput. Appl.*, vol. 28. 4, 2017, pp. 835–847.
- [8] D. Keles, J. Scelle, F. Paraschiv, and W. Fichtner, "Extended forecast methods for day-ahead electricity spot prices applying artificial neural networks." *Appl. Energy*, vol. 162, 2016, pp. 218–230.
- [9] F. Şenocak and H. Kahveci, "Periodic price averages forecasting of MCP in day-ahead market." in National Conference on Electrical, Electronics and Biomedical Engineering (ELECO), Turkey, 2016.
- [10] A. Dalgın, "Yapay Sinir Ağları Kullanılarak Türkiye Gün Öncesi Piyasası Elektrik Fiyat Tahmini." ITU Energy Institute, Turkey, 2017.
- [11] F. Gao, X. Guan, X. Cao and A. Papalexopoulos, "Forecasting power market clearing price and quantity using a neural network method." in 2000 Power Engineering Society Summer Meeting, WA, USA, 2000.
- [12] P.S. Georgilakis, "Market Clearing Price Forecasting in Deregulated Electricity Markets Using Adaptively Trained Neural Networks." in 4<sup>th</sup> Hellenic Conference on Artificial Intelligence, Greece, 2006.
- [13] L. Zhang and P. B. Luh, "Neural network-based market clearing price prediction and confidence interval estimation with an improved extended

- Kalman filter method." *IEEE Trans. on Power Systems*, vol. 20, no. 1, 2005, pp. 59-66.
- [14] N. M. Pindoriya, S. N. Singh and S. K. Singh, "An Adaptive Wavelet Neural Network-Based Energy Price Forecasting in Electricity Markets." *IEEE Trans. on Power Systems*, vol. 23, no. 3, 2008, pp. 1423-1432.
- [15] D. Singhal, K.S. Swarup, "Electricity price forecasting using artificial neural networks." *International Journal of Electrical Power & Energy Systems*, vol. 33, Issue 3, 2011, pp. 550-555.
- [16] I. N. Da Silva, D. H. Spatti, R. A. Flauzino, L. H. B. Liboni, and S. F. dos Reis Alves, "Artificial neural network architectures and training processes." in *Artificial neural networks*, Springer, 2017, pp. 21–28.
- [17] O. I. Abiodun, A. Jantan, A. E. Omolara, K. V. Dada, N. A. Mohamed, and H. Arshad, "State-of-the-art in artificial neural network applications: A survey." *Heliyon*, vol. 4. 11, 2018, pp. e00938.
- [18] R. Hecht-Nielsen, "Theory of the backpropagation neural network," in *Neural networks for perception*, Elsevier, 1992, pp. 65–93.

#### BIOGRAPHIES



**OĞUZ TONYALI** received the Electrical Engineering degree from Electrical and Electronics Faculty of Istanbul Technical University (ITU) in 2019. He worked as Sales & Tender Engineer in BEST for two years. He is now working for Artesis Innovative Industrial Maintenance Solution Corp. as Sales Development Manager.



**DUYGU BAYRAM KARA** received the BSc., MSc and PhD degrees from the Electrical and Electronics Faculty of Istanbul Technical University (ITU) in 2006, 2009 and 2015, respectively. She has been working at Electrical Engineering Department of ITU as Assistant Professor since February 2016. She worked at Nuclear Engineering Department, University of Tennessee, Knoxville-USA in 2013, on in-situ condition monitoring using electrical signature analysis, as a visitor researcher. Her research areas are signal processing and soft computing.