

Analysing Effect of Solar Photovoltaic Production on Load Curves and their Forecasting

B. Sinkovics¹, B. Hartmann¹

¹ Centre for Energy Research
Hungarian Academy of Sciences

Konkoly-Thege Miklós út 29-33., 1121 Budapest (Hungary)

Phone/Fax number: +36204825310, e-mail: sinkovics.balint@energia.mta.hu, hartmann.balint@energia.mta.hu

Abstract. Increasing share of intermittent renewable energy sources has generated several issues for power system operators. One aspect of these is the unpredictability of volatile production, affecting day-ahead load forecasting on system level, which is a major challenge for transmission system operators to solve. The literature widely discusses the short- and medium-horizon forecasting methods for weather dependent renewable energy sources, and proposals have also been raised to solve the issue. The aim of present paper is to estimate the effect of solar photovoltaic generation on the daily load curve of a national power system. To achieve this, current forecasting methods of transmission system operators are reviewed and evaluated. Then an own forecasting method is designed and implemented, based on historical load data from years where share of solar photovoltaics was neglectable. A learning algorithm is used for future predictions, using installed capacity and weather data. Finally, forecasted and actual load curves are compared, and effects of solar photovoltaic generation are estimated.

Key words

Solar photovoltaic, load curve, forecasting, neural networks

1. Introduction

The electricity has very different attributions than other material products. As a special goods it can be stored at extra costs. Ideally it should be generated as soon as it is demanded. The demand of electric power consumers can be described in time with the load curve. In real life the primary aim of forecasting is to meet future requirements and reduce the cost of unexpected events. For this reason, it is an important aim especially on short-term to predict electric loads accurately. The uncertainty of predictions comes mainly from residual consumers, who have stochastic energy consumption. This tendency deepens when entities with volatile electricity production appear on the network (e.g. photovoltaic panels). In Hungary household-sized small power plants are not registered, under 0.5 MW it is not needed to have the permission of Hungarian Energy and

Public Utility Regulatory Authority. Meanwhile half of the total solar capacities are planted as household-sized plants in 2016 (165.5 MW of 325 MW). [2] This means that even 2-4% of total electricity demand covered by these entities having no schedules. In the previous 4 years the built-in capacities doubled year by year. It is seen that if the tendency continues, the flexibility of generation system won't able to compensate that. Conventional power plants have strict power generating schedules, and the uncertainty of photovoltaic technologies make quality requirements' completion more difficult. It is presumed in this paper, that regulatory environment won't change in the next years and energy storages won't appear on the electric network.

Firstly, this paper summarize the method of Hungarian TSO's load forecasting and review short load forecasting methods focusing on artificial neural networks. In section 4. a general overview is published about the trends and future challenges in view of grow of built-in photovoltaic capacities. The major aim of this paper is to verify the effect of photovoltaic generation to daily load curve in context of actual low penetration. In Section 5 and 6 an own load forecasting method was built using neural networks.

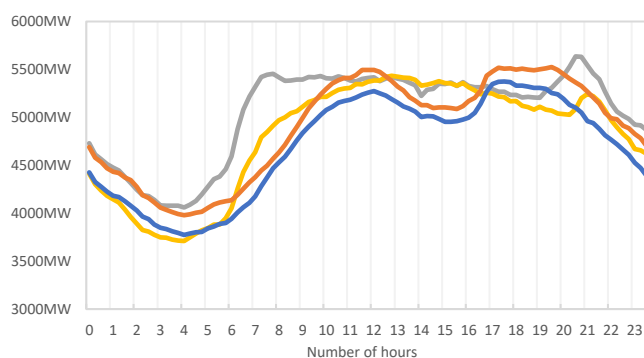


Fig. 1. Daily load curves by seasons in Hungary

2. Hungarian TSO's load forecasting

The load curve has every day a well described characteristic curve thanks to periodicity of consumer behaviours. This is distorted by many seasonal, non-quantifiable and trend effects. On Fig. 1., typical load curves of different seasons are shown. These have a global minimum at 4 a.m. Then comes a steep increase of demand until 12 a.m., where load curve usually has a local maximum. This pattern occurs again in the evening. This comparison shows well the seasonal effects like sunrise and set. Table 1. shows the measured peak electricity demand on summer and winter. In general, these two seasons occur peak loads because of heating/cooling need. As it is seen the daily electricity demand has a periodical, well predictable character and even peak values are not higher than 10-15% of mean load. When the photovoltaic penetration grows, the periodicity disappears because of the seasonal effect of solar power generation.

Among many others, the dependency from volatile generation can be avoided by transmission system operator's load forecasting system. Operators around the world have their own and often unique processes to generate day-ahead load curve forecasts as precise as possible. However, the general methodology is known among professionals, it is only partially published in the literature, since it may include sensitive information at times. The authors have decided to consult with the professionals at MAVIR Hungarian Transmission System Operator company (hereinafter referred to as MAVIR), to become familiar with their current operations in this field.

Regulations of MAVIR prescribe that error of day-ahead load curve forecast has to stay below 5%. To achieve this, the forecasting is prepared according to a two-step process, where:

- the first step is to define characteristic values of the day (e.g. morning off-peak load, noon and evening peak loads), based on personal intuition of the operators
- the second step is to finalise the shape of the load curve, using historical data of similar days

The process is repeated in a successive-approximation iteration, which largely builds on the competence of the personnel, and the data provided by a software, developed by a Hungarian company. Weather data are bought from several suppliers, including national (Hungarian Meteorological Service) and private (koponyeg.hu, idokep.hu) entities. From these weather forecasts, temperature, wind chill, luminance and cloud cover are considered in the load forecast. When determining the shape of the load curve, the operators also consider not quantifiable factors, e.g. television broadcasted sport events, changes in the law or the simultaneity of good weather and a long weekend.

Table 1. Load curve peaks in Hungary

Year	2008	2009	2010	2011	2012	2013	2014	2015
Winter peak [MW]	6388	6380	6560	6492	6463	6307	6461	6447
Summer peak [MW]	6252	5896	6232	6212	6288	6193	6050	6457

This methodology has been reliably used by MAVIR for several decades, and according to personal talks, results achieved by the operators could not be bested yet by computer-aided methods. Thus, software tools are dominantly only used to digitalise the process, and decision-making is the responsibility of the operators. The above include hidden problems, which are issued from the rigidity of the prediction method. Among the Hungarian power plants, the number of photovoltaic entities will drastically increase in the next years. This process causes many problems for MAVIR like decrease of system inertia, fluctuation in local electric power quality etc. Moreover, the presented load forecasting method largely build on the constancy of historical loads. For this reason, the load forecast prediction accuracy will expectedly decrease. The same process was observed by operators between 2008 and 2010. The effect of global financial crisis distorted the load curve shape significantly, which caused declining forecast accuracy.

3. Classification of load forecasting algorithms

One main aim of the work was to answer the questions, whether increasing share of solar photovoltaics already has affected the daily load curve – and thus the forecasting accuracy. To evaluate this question, an own forecasting method was designed and implemented, which uses historical load data of the power system from years, where solar penetration was at minimal level, thus the forecast can be considered free of the errors caused by intermittent production.

Such forecasting methods are widely discussed in the literature. [20] In general, load forecasting methods can be grouped based on their horizon, as:

- very short horizon forecasts (1-7 days)
- short horizon forecasts (1-4 weeks)
- medium horizon forecast (1-12 months)
- long horizon forecasts (1-20 years)

Long horizons are favourable when capacity expansion and power plant portfolio plans are developed, and cost-benefit analysis of future investments is performed. Econometrics and regression curves are the most widely used methods in this case, while results include peak loads and annual consumption. The best use of medium horizon forecasts is to plan maintenance works and downtime of generation units. Short horizon forecasts are used in daily operation and planning of the power system, and play a major role in maintaining the balance of generation and consumption, even with increasing penetration of intermittent resources.

The tools for short horizon forecasts include statistical, learning methods, time series and fuzzy logic. Many times, artificial networks solve a regression problem with minimization of the cost function. None of the methods is proven to be superior, but as computational capacity of computers has been increasing, efficiency and spread of artificial intelligence-based methods (ANN) is seen. Table 2. shows the most successful results using multi-layer perceptron neural networks (MLP). It is seen, that short term load forecasting has many different well-built ANN prediction models, which is followed by the authors.

Table 2.: Most successful load forecasting methods

Reference	Train method	Input data	MAPE
[6]	Backpropagation	1 year	1.4...2.8%
[7]	Backpropagation	1 year	2.07%
[8]	Backpropagation	3 years	2.8...6.4%
[9]	Backpropagation	8 weeks	5.45%
[10]	Quasi-Newton	1 year	3.26%
[11]	Backpropagation	1 year	1.73%
			1.75%
			2.06%
[12]	Backpropagation	2 years	2.7%
[13]	Backpropagation	4,5 years	15...17%
[14]	Backpropagation	30 days	1.2...1.7%
[15]	Backpropagation	4 years	1.7..1.8%
[16]	Backpropagation	4 weeks	1.94%
	generic algorithm	4 weeks	0.76%
[17]	Levenberg-Marquardt	3 years	1.83%
[18]	Levenberg-Marquardt	1 month	1.24%
[19]	Levenberg-Marquardt	4 month	3.57%

The suitability of the models is widely described with mean average percentage error (MAPE).

$$MAPE = \frac{1}{N} \cdot \sum_{i=1}^N \frac{|b_i - m_i|}{m_i}$$

where

- N is the number of the predictions
- b is the predicted load value
- m is the measured load value

Even from this indicator cannot deduct complex connections between the models, because of the different conditions, but gives a general sight.

The authors have chosen artificial neural networks to design the forecasting method, aiming to incorporate several factors, like trend effects, seasonal effects, weather effects and other special and accidental effects (human activities, pricing strategies, etc.). This model is also based on Levenberg-Marquardt backpropagation. It is the most common ANN optimisation technique, which combines the benefits of steepest descent (backpropagation) method and Gauss-Newton algorithm. The previous one has a slow convergence to the (local) optima, because the constant step size does not fit to every function gradient. Gauss-Newton can estimate proper step sizes for the directions, but the convergent result needs a reasonable quadratic approximation of error function. For this reason, Levenberg-Marquardt uses steepest descent algorithm until quadratic approximation available then change to Gauss-Newton algorithm. It is traceable on the equation of the method [22]:

$$w_{k+1} = w_k - (J_k^T J_k + \mu I)^{-1} J_k e_k$$

where

- I is the identity matrix

μ is a positive number, which express the switching between two algorithms

w is the weight vectors (with k iterations)

J is Jacobian matrix, which gives the second-order error of total error function

When μ is nearly zero, the equation approaches the Gauss-Newton method otherwise it gives the other one.

4. Future trends of power consumption and generation

First of all, it is needed to estimate the trend of the built-in solar capacities to have right conclusion about the change of load curve.

Future solar photovoltaic penetration levels were determined using historical installed capacities and curve fitting. It was presumed, that the predicted photovoltaic built-in power expansion follows exponential character, and this reason used exponential fitting to predict the trend as seen on Fig. 2. The equation of the fitted curve is the following:

$$y = a * e^{(b*x)}$$

where

$$a = 0,8469$$

$$b = 0,620$$

The value of mean squared error (MSE) is 7.63, which means an average 7.63 MW deviation from data points. In the view that 10 values were used for fitting (2007-2016) it is an adequate result. The resulted data indicates that if current speed of expansion continues, Hungarian photovoltaic capacity may reach 1600 MW by the end of the decade. Such estimations are not unrealistic, since the Hungarian Energy and Public Utility Regulatory Authority has issued permissions for 900 MW by the end of 2016. The Hungarian daily load curve usually reaches its maximum between 5500-6000 MW. In view of that the rate of investment is considerable in this sector. The expansion is pushed by two processes: global financial reasons (e.g. decline of manufacturer prices) and the change of the national regulatory environment, which guaranteed less favourable return for investors after 2016. In the Hungarian energy sector other investments or notable changes are not expected and the previous year were static too. For this reason, the change of solar capacities is well perceptible in context of historical

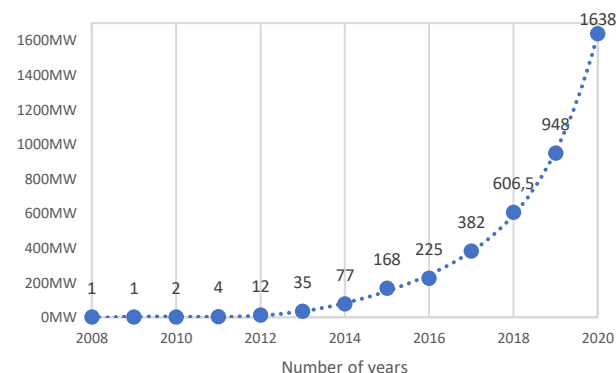


Fig. 2. Short-term estimation of built-in photovoltaic capacities

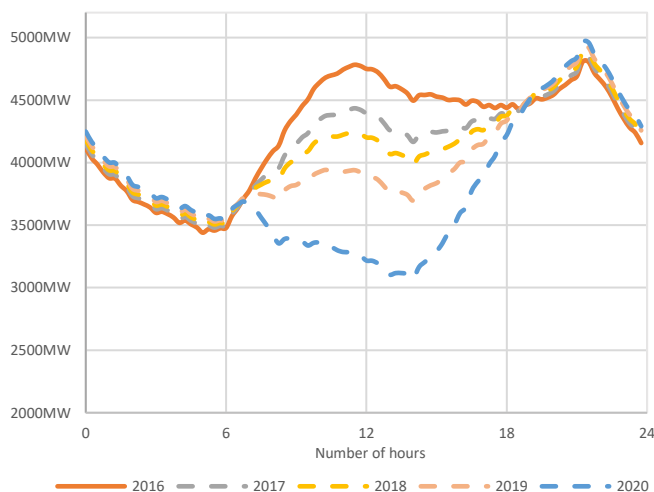


Fig. 3. Effect of solar energy production to daily load curve

information. Gross peak load of the power system is expected to increase 0.8-1.1% per annum; in this case the lower value was used for future scenarios.

To predict the impact of growing solar penetration to load curve an ideal model was set. The load curve of a household solar entity on a sunny day was scaled with the annual solar penetration. On Fig. 3. the impact of photovoltaic systems is seen, based on the load curve of a sunny summer day of 2016.

Due to the photovoltaic systems' effect is observable on daytime period, the load curve midday maximum disappears. So, the hypothesis is that increasing photovoltaic penetration primarily noticeable in the disappearing of midday peak of load curve. The problem of that it generates a steep gradient run up at 7pm, which can cause regulation problems, because conventional plants cannot control a steep load gradient change.

5. Architecture of the model

The artificial neural networks have been widely used as load forecasting method especially for short term load forecasting. These solutions are usually applied to perform non-linear curve fitting as introduced in Section 2. The full model was programmed in Matlab environment. Every architecture has an input and output layer and hidden layer(s). The logical sketch, how prediction process works is seen on Fig. 4. The learning algorithm is trained on training set. The iterative refresh of neuron weights reaches an optimum, where the prediction has a minimum error against training set. That mapping structure gives the hypothesis. Firstly, it is needed to define a vector for feeding input layer. It consists of different variables which influence the described process. In this case the estimated load values

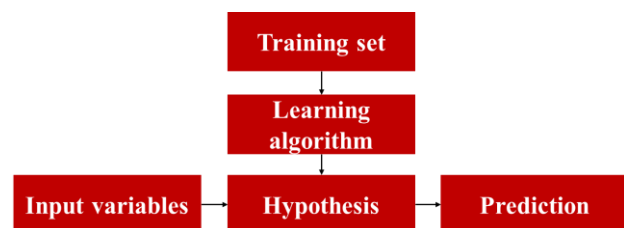


Fig. 4. Structure of prediction model

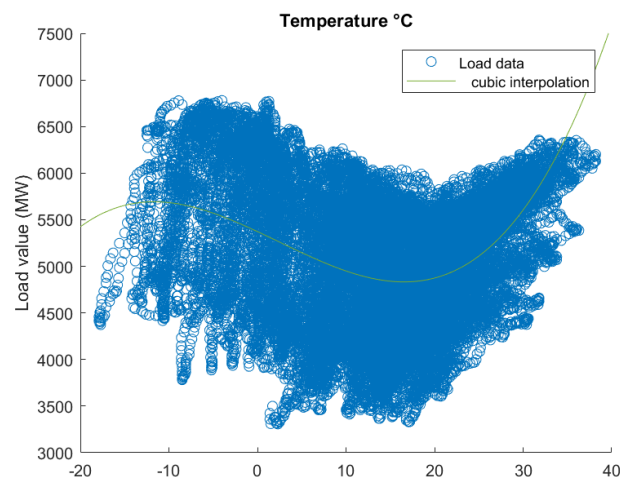


Fig.5. Connection between load and temperature variables

depend on historical loads, hour of day, weather variables (wind-speed, radiation, temperature) etc.

In the defined model the following variables are used: temperature, previous loads, day of week, special holidays.

For example, Fig. 5. shows the connection between temperature and load. It is seen that trend fitting with lowest MSE is a cubical mapping between the two variables. Temperature dependent load has a trend minimum at 20°C. The source of dataset is an amateur weather observer community who publish their private stations' recorded data on *amsz.hu*. The input database covers the whole time from 2012 January to 2017 September. The inputs have 15 minutes temporal resolution respectively.

The only problem was that the private datasets are allowed to export to Excel separately in an 24h interval. For this reason, the overall database was generated by VBA-based Excel macros which managed the conversion and combination of each downloaded tables.

An additional problem was, that the recorded values had a rare than 15 minutes resolution, and also occurred absent timetables. For this reason, to get an effective database two method was used:

- rescale the merged records to the adequate resolution
- predict the missing values from existing ones

First of all the dates were categorized by the following logic:

00:52 < X < 00:08 → 00:00

00:07 < X < 00:23 → 00:15

00:22 < X < 00:38 → 00:30

00:37 < X < 00:53 → 00:45

where X means the time of original records. This process determined the rounding of the corresponding temperature records. A simple linear interpolation was used by the following:

$$y = y_1 + \frac{(x - x_1)(y_2 - y_1)}{x_2 - x_1}$$

where

$x_1 < x < x_2$ (time variables)

$y_1 < y < y_2$ (temperature)

These were rounded up and down depending on the certain value. For this reason, the mathematical operation between

the different hours was provided. The rescaling process generated a fixed time resolution and revealed missing data points. The rescale generated deficit was solved by averaging, because the rate of changing is 0.2°C or less in the 70% of cases between neighbour values. The longer interruptions were completed manually using previous day approach. The timeframe of these cases was between 30 minutes and 24 hours. For this reason, records of a near weather station were used for replacement.

The historical loads come by public database of MAVIR, for transmission system operators it is compulsory to publish the main trends of system network. The input vector of holidays and day of week parameters were defined by Matlab built-in 'dayofweek' and 'holidays' functions.

The prediction method of the neural network based on a multiple regression problem. The general purpose of multiple regression is to learn a nonlinear mapping between several independent or predictor variables and a dependent or criterion variable. This can be described by the following equation:

$$Y = b_0 + b_1X_1 + b_2X_2 + \dots + b_pX_p$$

where

Y is the criterion variable

X₁ through X_p are predictor variables

b₀ called intercept

b₁ through b_n are partial regression coefficients

The regression coefficients represent the independent contributions of each independent variable to the prediction of the dependent variable, which called partial correlation. The cost function of the model is defined by the difference of regression model and measured data point, which can be rated by MAPE. The aim of the neural network is to learn mutual mapping between predictor and criterion variables on least error.

The neural network was designed as a feedforward one with 1 hidden layer and 20 neurons. It has sigmoid hidden neuron activation and linear outputs. The training period was 1000 epochs. Available load curve data were split by 75%, 15%, 10% for learning, validation and testing, respectively. It uses Levenberg-Marquadt optimisation, which algorithm appears to be the fastest method for training moderate-sized feedforward neural networks. The test set ensure the learning environment for the neural network. The validation set helps to estimate prediction error of the model selection. Finally, the generalization error is measured on the test set to avoid overfitting.

6. Model results

Prediction results are shown on Fig. 6. and Fig. 7., both of which are representing the operation of the neural network forecasting method, for two periods of September 2017. The model was trained on the previous years' load, when solar penetration was negligible. The mean average percentage error on the test set was 5.2%, which is the best optimized scenario of the model. The algorithm learned well to generate load curves in the context of previous loads. The days between Tuesday and Friday has the same characteristics against weekend loads. For this reason, the most effective curve fitting is noted on these days. It can be seen, that the biggest differences between forecasted and actual data occur during daytime peaks, which cannot be only the effect of the robustness of the prediction method.

The picked and plotted days were mostly dry and sunny for having unambiguous effect of photovoltaic energy production. All in all it is verified, even actual low penetration of solar entities has notable impact to daily load curve.

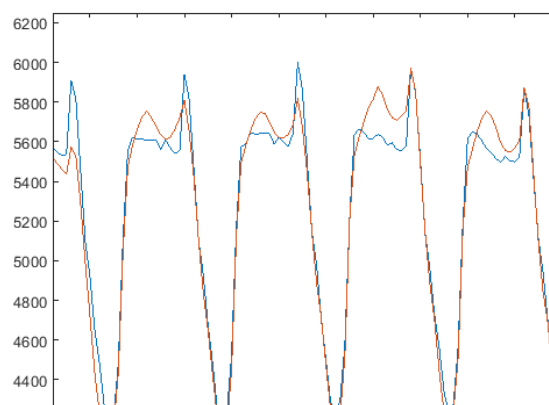


Fig. 6. Forecasted (orange) and actual (blue) load curves for the days of 26-29 September, 2017

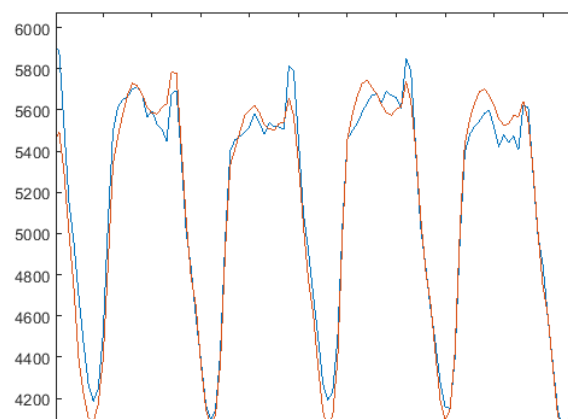


Fig. 7. Forecasted (orange) and actual (blue) load curves for the days of 19-22 September, 2017

Acknowledgement

The VEKOP-2.3.2-16-2016-00011 grant is supported by the European Structural and Investment Funds jointly financed by the European Commission and the Hungarian Government

References

- [1] Annual report of Hungarian Energy and Public Utility Regulatory Authority (2008-2016), http://www.mekh.hu/download/7/15/40000/nem_engedelykotel_es_es_hmke_%20beszamolo_2016.pdf (2017.09.22.)
- [2] IRENA, Renewable Capacity Statistics, IRENA, 2017
- [3] E. Almeshaiei, H. Soltan, A methodology for electric power load forecasting, Alexandria Engineering Journal, vol. 50., pp. 137-144., 2011
- [4] M. Markou, E. Kyriakides, M. Polcarpou, 24-Hour Ahead Short Term Load Forecasting Using Multiple MLP, in Proc. DEMSEE 08, 2008
- [5] H. Cho, Y. Goude, X. Brossat, Q. Yao, Modelling and Forecasting Daily Electricity Load via Curve Linear Regression, in Modeling and Stochastic Learning for Forecasting in High Dimensions, Springer, 2015

- [6] B. Satish, K. S. Swarup, S. Srinivas és A. H. Rao, „Effect of temperature on short term load forecasting using an integrated ANN,” *Electric Power Systems Research*, 2004.
- [7] M. Abu-Et-Magd és R. Findlay, „A new approach using artificial neural network and time series models for short term load forecasting,” 2003.
- [8] K. Methaprayoon, W. J. Lee, S. Rasmiddatta, J. Liao és R. Ross, „Multi-Stage Artificial Neural Network Short-term Load Forecasting Engine with Front-End Weather Forecast,” 2006.
- [9] J. Bao, „Short-term Load Forecasting based on Neural network and Moving Average,” 2002.
- [10] L. Xu és W. J. Chen, „Short-Term Load Forecasting Techniques Using ANN,” 2001.
- [11] I. Erkmen és A. Topalli, „Four methods for short-term load forecasting using the benefits of artificial intelligence,” *Electrical Engineering*, 2003.
- [12] Y.-K. Wu, „Short-term forecasting for distribution feeder loads with consumer classification and weather dependent regression,” 2006.
- [13] L. Hernández, C. Baladrón, J. M. Aguiar, L. Calavia, B. Carro, A. Sánchez-Esguevillas, F. Pérez, Á. Fernández és J. Lloret, „Artificial Neural Network for Short-Term Load Forecasting in Distribution Systems,” *Energies*, 2014.
- [14] S. A. Ilic, S. M. Vukmirovic, A. M. Erdeljan és F. J. Kulic, „Hybrid artificial neural network system for short-term load forecasting,” *Thermal Science*, 2012.
- [15] A. Khwaja, M. Naeem, A. Anpalagan, A. Venetsanopoulos és B. Venkatesh, „Improved short-term load forecasting using bagged neural networks,” *Electric Power Systems Research*, 2015.
- [16] P. Ray, D. P. Mishra és R. K. Lenka, „Short-Term Load Forecasting by Artificial Neural Network,” 2016.
- [17] G. L. Prakash, K. Sambasivarao, P. Kirsali és V. Singh, „Short Term Load Forecasting for Uttarakhand using Neural Network and Time Series models,” 2014.
- [18] A. Bala, N. K. Yadav és N. Hooda, „Implementation of Artificial Neural Network for Short Term Load Forecasting,” *Current Trends in Technology and Science*, 2014.
- [19] N. Liu, V. Babushkin és A. Afshari, „Short-Term Forecasting of Temperature Driven Electricity Load Using Time Series and Neural Network Model,” *Journal of Clean Energy Technologies*, 2014.
- [20] F. Elakrmi, N. Abu Shikhah “Electricity Load Forecasting – Science and Practices”
- [21] A.K. Srivastava, Ajay Shekhar Pandey and Devender Singh “Short-Term Load Forecasting Methods: A Review” *International Conference on Emerging Trends in Electrical, Electronics and Sustainable Energy Systems (ICETEESES-16)*
- [22] Hao Yu, Bogdan M. “, LevenbergMarquardt Training *Industrial Electronics Handbook*” 2011