Neural Networks Application for Modelling of RES production

Metody Georgiev
Faculty of Automatics
Technical University of Sofia
Sofia, Bulgaria
georgievmg@tu-sofia.bg

Alexandra Georgieva
Faculty of Automatics
Technical University of Sofia
Sofia, Bulgaria
aleksa.georgieva@tu-sofia.bg

Dilyana Gospodinova
Faculty of Electrical Engineering
Technical University of Sofia
Sofia, Bulgaria
dilianag@tu-sofia.bg

Kostadin Milanov
Faculty of Electrical Engineering
Technical University of Sofia
Sofia, Bulgaria
kmilanov@tu-sofia.bg

Abstract: **The article is focused on modeling and analyses of electricity production from renewable energy sources in a single-family house using artificial intelligence. The neural networks are one of the main instruments for modeling and forecasting of dynamic and stochastic processes. In the current research, they are implemented to modeling of the electricity production of a photovoltaic station of a single-family house with the aim to be able to analyze and predict the PV production in short and long term periods. This is very important in the current electricity systems in respect to ensure optimal energy utilization in the house and optimal facilities exploitation. In the article real data from an existing SCADA are used which makes the results close to real exploitation.**

Keywords: neural networks, forecasting, solar energy, photovoltaic system, electricity production

I. INTRODUCTION

Neural Networks (NN) suggest a different way of analyzing and investigating situations compared to the classical ones like regression analyses, statistical analyses, etc.

This way is characterized by the building of a complex nonlinear function with an iterative way of parameters adaptation (teaching). This way is very similar to the Kalman filtering, [1].

The advantages of the traditional neural networks comparing the “classical” methods can be summarized in the following points:

- They are self-adapting systems based on expert suggestions about the data model;
- They can generalize. As they are based on previous data they are the perfect tool for time series analyses with or without additional variables.
- Neural networks are universal approximation tools;
- Neural networks are nonlinear.

These NN advantages make them the perfect instrument for analyzing the time series related to power production and especially to the production of electricity from solar sources [2].

The power production with photovoltaic installation characterizes by the following properties:

1) It has stochastic behavior related to changes in weather.

2) It is strongly related to electricity consumption (especially in the case of prohibition to return energy to the grid). In this case the PV station measures the energy consumed and the energy produced and adapt the production in relation with the consumption. The schematic is given at Fig.1 where two smart meters are connected to the PV in respect to ensure all necessary data for the PV control system. In the family house the problem with the stochastic character is even stronger because of the habitant's life dynamics. For improvement of the electricity production models for the family houses, models of household energy profiles [3, 4] could be used.

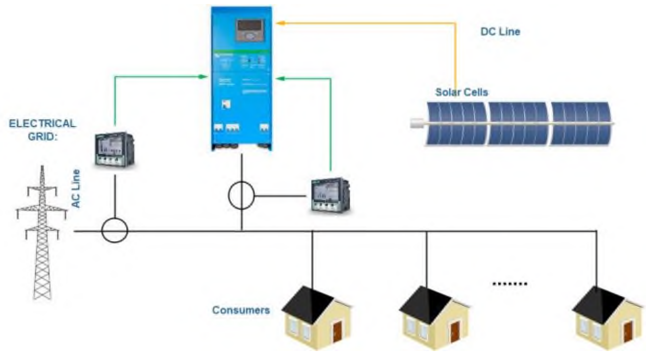

Fig. 1. Photovoltaic installation.

3) It has seasonal characteristics related to day/night change and related to the change of solar radiation during the year.

These properties make the goal to model and forecast the PV production very complex. Different authors deal with the problem using different network structures, deep learning and other technics like: RNN (Recurrent Neural Network), LSTM (GRU (Gated Recurrent Units) [5,6];LSTM(Long-Short Term Memory) [5,7], FFNN (Feed Forward Neural Network) [8,9]; MLP (Multilayer Perceptron), Elman Network [10]; NAR (Nonlinear Autoregressive); NARX (Nonlinear Autoregressive with exogenous inputs), [10]. The researchers report good results with the proposed algorithms and suggest that the neural network structures should be selected depending on the input and the output data which are available and the specific condition for any use case.

The main weather component with an influence on the PV production is the solar radiation. The level of the solar radiation depends on geographic location, time of day, seasons, local landscape and local weather, [11]. That’s why the local weather parameters like temperature [12, 13, 14] and humidity [15] could be used to estimate the solar radiation in a specific region. As the cloud coverage has strong impact on the level of solar radiation, some authors use more complex input satellite data and sky images, [16, 17].

The goal of this paper is to present an algorithm for PV production modeling using Neural Network in case of lack of information about the consumption profile and solar radiation

978-1-6654-4192-6/21/$31.00 ©2021 IEEE

and to present an algorithm for selecting the best ANN (artificial neural network) structure.

II. DESCRIPTION OF THE PROBLEM

The neural networks has several parameters which have to be fixed before starting the network training: number of inputs, number of outputs, number of the hidden layers, the activation functions of the neurons of the hidden and output layer which form the neural network structure. In more complex case feedbacks between hidden layers could be stated. The research works show many different ways to construct the neural networks to deal with stochastic uncertainty but usually, it is done separately depending on the exact use case and the data which are obtainable for analysis and there is no common algorithm to calculate the mentioned above neural network parameters, [18].

This article presents a practical approach for the implementation of neural networks for modeling the PV production in a single-family house. The use case presents a test facility with a photovoltaic installation that is connected to the grid but produce energy only for self-consumption (based on consumption measurements the photovoltaic system produce only if loads are presented in the consumer network (Fig.1). The level of production is strongly related to the level of consumption). Photovoltaic installation can produce if the necessary energy is available and if loads are connected to the system. If the load is less than the available energy the PV installation reduces the production to the level of the presented loads. For this purpose measurements of the energy production and the consumed energy are available in the system. The presented measurement system Fig.1 is connected with a SCADA system to record energy data for monitoring and analyses. The data for consumed and produced active energy in kWh, produced and consumed active power in kW are recorded with step of 1 minute Fig.2. From the records timestamps an information about the hour of the day (range 1-24), day of the week (range 1-7), month of the year (range 1-12), day of the year (range 1-366) and week of the year (range 1-52) are extracted. Additionally from the specialized websites a weather information is integrated for air temperature, humidity, pressure, and wind speed. The weather data is collected hourly. To be able to synchronize the records from the PV system and the records from the weather stations hourly data for the electricity production are extracted for the database using the equation (1)

$$E_n = E_n^{end} - E_n^{start} \quad (1)$$

Where:

- E_n is the energy in kWh for the n-th hour,
- E_n^{end} is the active energy totalizer of the smart meter at the end of the n-th hour,
- E_n^{start} is the active energy totalizer of the smart meter at the end of the n-th hour.

To be able to use time in the model a representation of the timestamp as a real value is used where the whole part of the number is the number of the day starting from 01.01.1900 and the fractal part is the time of the day.

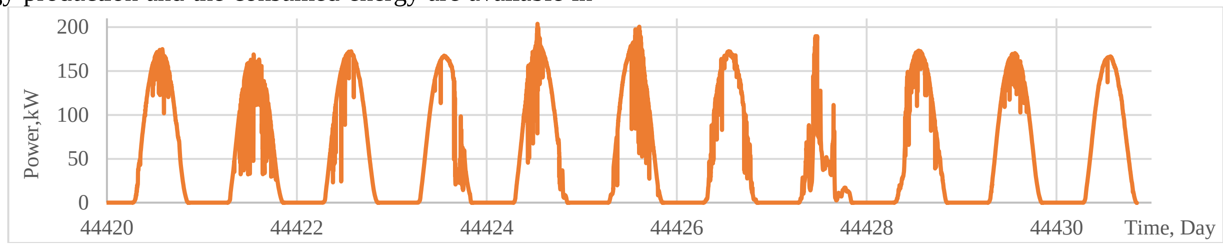

Fig. 2. Output PV power curve.

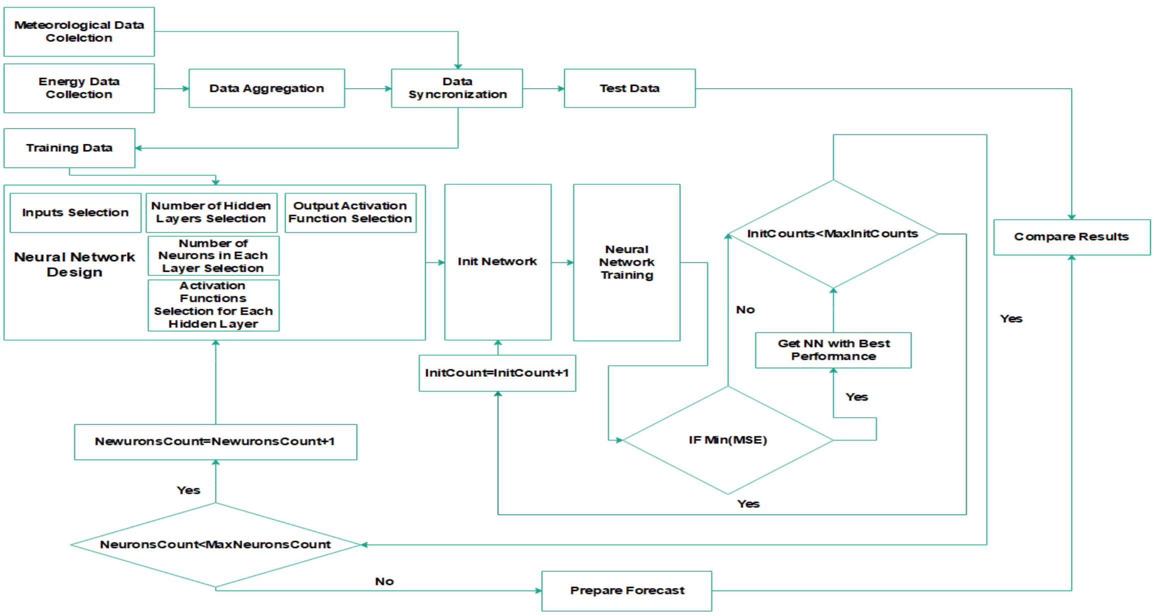

Fig. 3. Algorithm for modeling of PV production.

III. METHODOLOGY OF RESEARCH

Figure 3 depicts the investigation's methodology and developed algorithm. The proposed algorithm consists of a comparison between networks behavior based on minimal Mean Squared Error (MSE). To be able to avoid getting stuck in a local extremum it is suggested to make several pieces of training with different initial neurons weights. The method used to estimate the initial weights and biases of the neuron is the Nguyen-Widrow initialization algorithm [8]. Also to be able to choose the number of neurons in hidden layer a scan is done changing the number of neurons between preliminary defined minimal value and preliminary defined maximal. The network with the best performances is stored and used for data prediction. To be able to estimate the NN prediction ability the predicted data are compared with the real ones. From the recorded data hourly energy consumption is extracted and the data are shown at Fig.4. The main parameters for training the neural networks are:

- Maximum number of training iterations: 100000;
- Performance: Mean-Squared Error;
- Training Algorithm: Levenberg-Marquardt [19];
- Minimum gradient. 1.10^{-7}.As the energy data are collected every minute and the meteorological data are collected hourly the algorithm starts with energy data aggregation to extract the hourly averages Fig. 4.

The proposed neural network structure with the corresponded input pattern is shown at Fig.5.

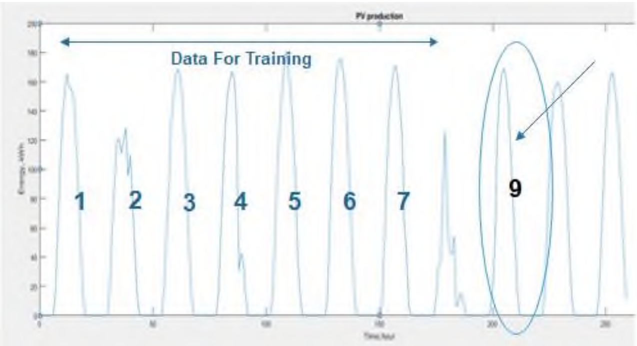

Fig. 4. Hourly power generation.

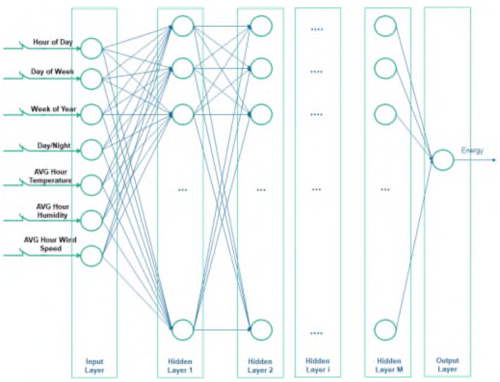

Fig. 5. Neural Network Structure.

The input vector is given as equation (2):

$$u = [u_1, u_2, \dots, u_6,] \tag{2}$$

where:

u_1- is the vector of an hour of the day (hourly)
u_2- is the vector of a day of the week (hourly)
u_3- is the vector of average temperature (hourly)
u_4- is the vector of average humidity (hourly)
u_5- is the vector of average wind speed (hourly)
u_6- is the vector of Day/Night index (0 – the dark part of the day-night; 1 – light part of the day-night)

The input data consists of 168 samples. The training periods are given with numbers from 1 to 7 at Fig.4 and as test data the values for the 9-th period given at Fig.4 are used The FFNN is configured with one hidden layer with tangent-sigmoidal activation function. The output neuron has a linear activation function. The algorithm makes several runs with a different number of neurons in the hidden layer (between 6 and 15).

IV. EXPERIMENTAL RESULTS

During the inquiry, it was discovered that the outcome of network training is influenced by the initial values of the neurons, thus numerous runs were carried out to compare the results of these calculations while commencing training with various training start points. The first input pattern consists only hour of the day and the average temperature. Sample graphical results are given in Fig.6 for NN with 11 neurons. The first graphic is the sample series, the second one is the Network behavior after training and the third one is the error. The MSE values for networks with different neurons are given in Table I. MSE1 is obtained when we zeroes as initialization values of weights and the biases are applied and MSE2 values are obtained when pseudo-random numbers are used for initial weights and biases. It is visible that there is no big difference between the results but the best result are obtained with 11 neurons in the hidden layer but it is visible from the data in Table 1 that there is no significant change in MSE for the networks with number of neurons in hidden layer greater than 9.

TABLE I. MEAN SQUARED ERRORS FOR THE FIRST EXPERIMENT

Number of neurons	MSE1	MSE2
6	118.1840	118.3321
7	118.1834	118.1823
8	118.1965	118.1809
9-15	118.1809	118.1809

Fig. 6. Neural Network Prediction Results for 2 inputs NN.

The comparison of the best-performing neural network model data with real data on a day of the week where there is no dip in production due to no consumption demands. It is visible that this network predicts almost perfectly the average behavior of the PV. In this case, it is visible that the model eliminates the disturbance (drop down in the production) in the second PV training period (Period 9 shown on Fig.4) of the PV which is a result of load Fig.7.

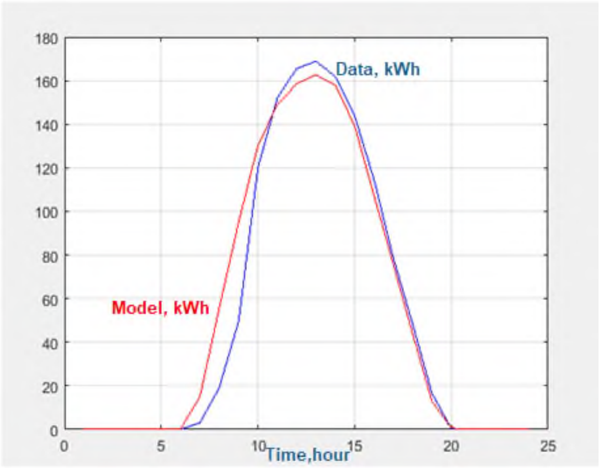

Fig. 7. Neural Network Training Results for 2 inputs NN

It is obvious that this network will not give good results in a long-term period but it will be a good tool for short-term prediction. The changes in radiation due to the sun's seasonal activity might be followed in this scenario by retraining the network using the schematic shown in Fig. 8.

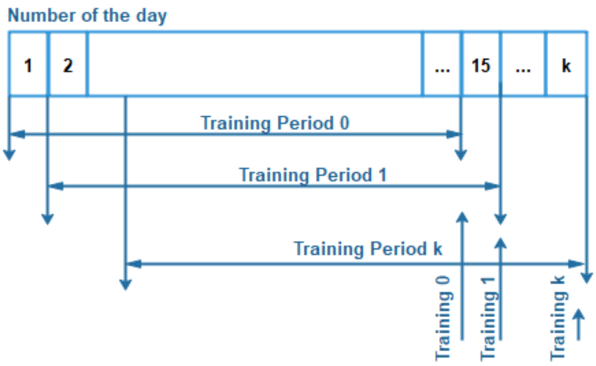

Fig. 8. Network retraining schedule.

Also it is visible from the figures that this network model estimate average electricity production for a specific hour of the day but does not estimate the change of the load related to the specific day (Period 2, Fig.4). This could be avoided using more complex input pattern of the network. In the next input pattern the number of the day of the week is added and day/night index which shows the transition between the day and the night so $u = [u_1, u_2, u_3, u_6]$.

The results for the mean squared training errors are given in Table II.

The graphical results for training are shown in Fig. 9 and the comparison with the test data is shown in Fig.10.

TABLE II. MEAN SQUARED ERRORS FOR THE SECOND EXPERIMENT

Number of neurons	MSE1	MSE2
6	29.3708	30.6185
7	29.3708	30.6185
8	14.3708	15.4756
9	14.3708	14.0331
10	8.6707	13.7080
11	7.5695	13.5565
12	5.8096	8.3064
13	5.8096	6.0527
14	5.8096	4.1163
15	5.8096	4.1163

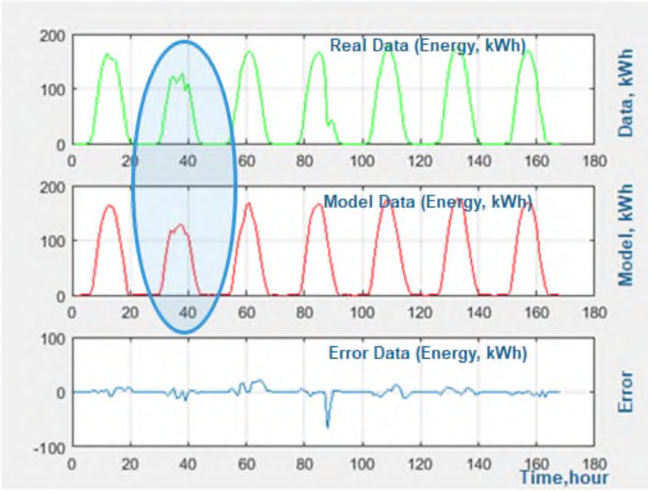

Fig. 9. Neural Network Prediction Results for 4 inputs NN.

It is visible that in the first case the prediction is not sensitive concerning the load profile related to the specific day of the week, but it fits perfectly for the prediction of the PV curve when enough loads are presented in the system to cover the total production capacity of the PV. It is clear that in the first scenario, the prediction is unaffected by the load profile associated with a single day of the week, but it is ideal for predicting the PV curve when the system has enough loads to cover the PV's whole output capacity.

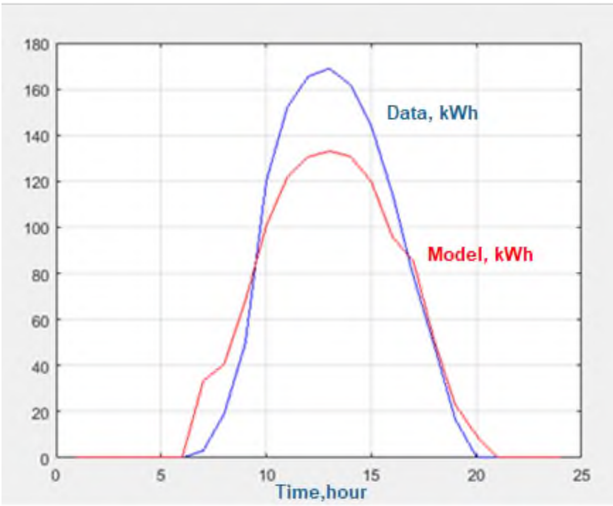

Fig. 10. Test results (Predicted and Real Data for period No 9).

It is obvious that the test data and the predicted data are different and the reason for this is that the trained network in this case estimates also the load behavior and the meteorological data related to that day of the week Fig.9. It is visible that the period 2 Fig.4 differs from the period 9 Fig.4 because of the change of the load behavior. As the network model estimates the load behavior, the predicted data differs from the real ones Fig.1.

The third experiment is made using six variables in the vector of the inputs where the input vector is defined as (2). The neural network training results are shown in Table III and Fig.11. The comparison between the predicted and the test data are shown at Fig.12.

TABLE III. MEAN SQUARED ERRORS FOR THE THIRD EXPERIMENT

Number of neurons	MSE1	MSE2
6	19.1975	17.6339
7	11.0448	16.4395
8	7.2404	7.3080
9	5.8176	5.9675
10	3.5155	4.8076
11	3.5155	3.8558
12	3.0038	2.8008
13	2.6921	2.8008
14	2.6921	1.9086
15	1.3807	1.1123

The results of the second and third experiments are similar and the difference is in small decrease of the MSE against increasing the model complexity which lead to increasing time for retraining. In this case the model is more flexible concerning the weather conditions (in the second and the forth periods the model fits better to the real data) but from the practical point of view both will have one and the same success.

The summarized results concerning the networks training performance are given in Fig.13 where Exp.1 represents the results of the network training with two inputs, Exp.2 is the result from the network performance with 4 inputs and Exp.3 is the results of using 6 inputs networks.

The network with 6 inputs and 15 neurons in the hidden layer has the best performance and its results take into account the weather condition and the usual consumer's behavior.

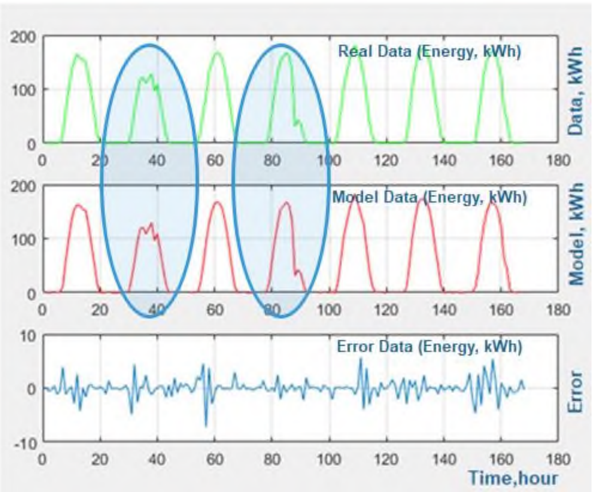

Fig. 11. Test data results for 6 inputs neural network.

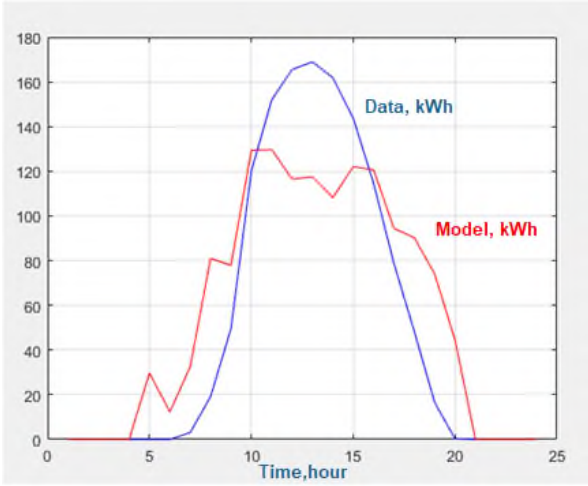

Fig. 12. Test data results for 6 inputs neural network.

V. IV.CONCLUSION

This research presents some experimental data related to modeling of PV production using neural networks. In the research a feed-forward neural network structure with one hidden layer is used with different number the neurons and inputs. An algorithm is proposed for searching the best number of the hidden layer length based on the mean squared error estimation.

During the investigation following conclusions are made:

- • A neural network with two inputs can perfectly fits the hourly average PV production. The predicted data could be used in estimation of the PV efficiency in comparison to the real production.
- • A neural network with two inputs can be used for forecasting the weekly profile of the PV production.
- • In both cases (Exp.2 and Exp.3), an additional vector is utilized to represent the light and dark parts of the day, which improves network performance and reduces network training time.
- • Extending the networks with more inputs like temperature, humidity and wind speed improve the results from the network training.
- • The proposed algorithm of the best trained structure show that the utilization pseudo-random initialization of the neural network weights do not have large impact on the results.

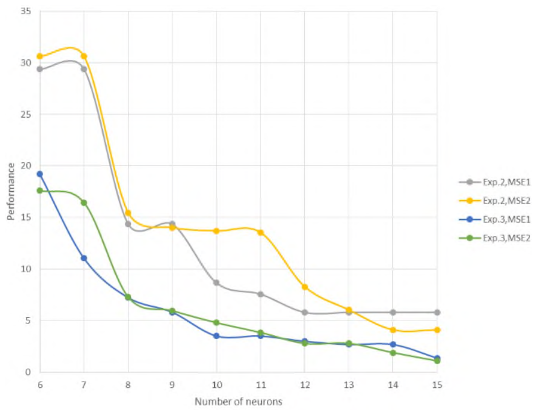

Fig. 13. Performance vs Number of Neurons.

The research reveals that feed-forward neural networks can be successfully employed for forecasting PV production, with the issue in this case being forecasting customer behavior, which has a significant impact on production. As can be seen from the results, increasing the number of neurons improves forecasting results while also increasing the time required for network training. Future work will focus on developing a practical approach for implementing the proposed neural networks in real-world systems and developing a faster practical method for identifying the best neural network structure.

Acknowledgment

The authors would like to thank the National Scientific Program NNP "Low Carbon Energy for Transport and Households - Eplus", D01-214/28.11.2018 for the financial support

References

[1] S. Theodoridis, “Machine learning : a Bayesian and optimization perspective,” p. 1031.

[2] A. Olawoyin and Y. Chen, “Predicting the Future with Artificial Neural Network,” Procedia Comput. Sci., vol. 140, pp. 383–392, Jan. 2018, doi: 10.1016/J.PROCS.2018.10.300.

[3] [A. Marszal-Pomianowska, P. Heiselberg, and O. Kalyanova Larsen, “Household electricity demand profiles – A high-resolution load model to facilitate modelling of energy flexible buildings,” Energy, vol. 103, pp. 487–501, May 2016, doi: 10.1016/J.ENERGY.2016.02.159.

[4] T. Csoknyai, J. Legardeur, A. A. Akle, and M. Horváth, “Analysis of energy consumption profiles in residential buildings and impact assessment of a serious game on occupants’ behavior,” Energy Build., vol. 196, pp. 1–20, Aug. 2019, doi: 10.1016/J.ENBUILD.2019.05.009.

[5] I. Jebli, F.-Z. Belouadha, M. I. Kabbaj, and A. Tilioua, “Deep Learning based Models for Solar Energy Prediction,” Technol. Eng. Syst. J., vol. 6, no. 1, pp. 349–355, 2021, doi: 10.25046/aj060140.

[6] S. Al-Dahidi, O. Ayadi, M. Alrbai, and J. Adeeb, “Ensemble approach of optimized artificial neural networks for solar photovoltaic power prediction,” in IEEE Access, 2019, vol. 7, pp. 81741–81758, doi: 10.1109/ACCESS.2019.2923905.

[7] C. Hua, E. Zhu, L. Kuang, and D. Pi, “Short-term power prediction of photovoltaic power station based on long short-term memory-back-propagation:,” https://doi.org/10.1177/1550147719883134, vol. 15, no. 10, Oct. 2019, doi: 10.1177/1550147719883134.

[8] J. M. Barrera, A. Reina, A. Maté, and J. C. Trujillo, “Solar Energy Prediction Model Based on Artificial Neural Networks and Open Data,” Sustain. 2020, Vol. 12, Page 6915, vol. 12, no. 17, p. 6915, Aug. 2020, doi: 10.3390/SU12176915.

[9] C.-D. Dumitru, A. Gligor, and C. Enachescu, “Solar Photovoltaic Energy Production Forecast Using Neural Networks,” Procedia Technol., vol. 22, pp. 808–815, Jan. 2016, doi: 10.1016/J.PROTCY.2016.01.053.

[10] J. L. Gómez, A. O. Martínez, F. T. Pastoriza, L. F. Garrido, E. G. Álvarez, and J. A. O. García, “Photovoltaic Power Prediction Using Artificial Neural Networks and Numerical Weather Data,” Sustain. 2020, Vol. 12, Page 10295, vol. 12, no. 24, p. 10295, Dec. 2020, doi: 10.3390/SU122410295.

[11] I. Sarbu and C. Sebarchievici, Solar Heating and Cooling Systems | ScienceDirect. Elsevier Inc., 2017.

[12] J. T. Hardinata, M. Zarlis, E. B. Nababan, M. R. Wayahdi, M. Zarlis, and P. H. Putra, “Initialization of the Nguyen-widrow and Kohonen Algorithm on the Backpropagation Method in the Classifying Process of Temperature Data in Medan,” in Journal of Physics: Conference Series, Jun. 2019, vol. 1235, no. 1, p. 012031, doi: 10.1088/1742-6596/1235/1/012031.

[13] O. Dupré, R. Vaillon, and M. A. Green, Thermal Behavior of Photovoltaic Devices: Physics and Engineering. Springer International Publishing, 2017.

[14] A. Saberian, H. Hizam, M. A. M. Radzi, M. Z. A. Ab Kadir, and M. Mirzaei, “Modelling and prediction of photovoltaic power output using artificial neural networks,” Int. J. Photoenergy, vol. 2014, 2014, doi: 10.1155/2014/469701.

[15] N. N. Tasie, C. Israel-Cookey, and L. J. Banyie, “The Effect of Relative Humidity on the Solar Radiation Intensity in Port Harcourt, Nigeria,” Int. J. Res., vol. 5, no. 21, pp. 128–136, 2018, Accessed: Dec. 06, 2021. [Online]. Available: https://journals.pen2print.org/index.php/ijr/article/view/16369.

[16] S. D. Miller, M. A. Rogers, J. M. Haynes, M. Sengupta, and A. K. Heidinger, “Short-term solar irradiance forecasting via satellite/model coupling,” Sol. Energy, vol. 168, pp. 102–117, Jul. 2018, doi: 10.1016/J.SOLENER.2017.11.049.

[17] L. M. Halabi, S. Mekhilef, and M. Hossain, “Performance evaluation of hybrid adaptive neuro-fuzzy inference system models for predicting monthly global solar radiation,” Appl. Energy, vol. 213, pp. 247–261, Mar. 2018, doi: 10.1016/J.APENERGY.2018.01.035.

[18] K. Yotov, E. Hadzhikolev, and S. Hadzhikoleva, “Determining the Number of Neurons in Artificial Neural Networks for Approximation, Trained with Algorithms Using the Jacobi Matrix,” TEM J., vol. 9, no. 4, pp. 1320–1329, Nov. 2020, doi: 10.18421/TEM94-02.

[19] B. M. Wilamowski and J. D. Irwin, Eds., Intelligent Systems. Boca Raton, FL, USA: CRC Press, 2018.