A Predictive Model for Solar Photovoltaic Power using the Levenberg-Marquardt and Bayesian Regularization Algorithms and Real-Time Weather Data

Mohammad H. Alomari*, Ola Younis† and Sofyan M. A. Hayajneh‡
*Electrical Engineering Department
Applied Science Private University, Amman, Jordan
†School of Electrical Engineering, Electronics and Computer Science
University of Liverpool, Liverpool, United Kingdom
‡Electrical and Computer Engineering Department
Isra University, Amman, Jordan

Abstract—**The stability of power production in photovoltaics (PV) power plants is an important issue for large-scale grid-connected systems. This is because it affects the control and operation of the electrical grid. An efficient forecasting model is proposed in this paper to predict the next-day solar photovoltaic power using the Levenberg-Marquardt (LM) and Bayesian Regularization (BR) algorithms and real-time weather data. The correlations between the global solar irradiance, temperature, solar photovoltaic power, and the time of the year were studied to extract the knowledge from the available historical data for the purpose of developing a real-time prediction system. The solar PV generated power data were extracted from the power plant installed on-top of the faculty of engineering building at Applied Science Private University (ASU), Amman, Jordan and weather data with real-time records were measured by ASU weather station at the same university campus. Huge amounts of training, validation, and testing experiments were carried out on the available records to optimize the Neural Networks (NN) configurations and compare the performance of the LM and BR algorithms with different sets and combinations of weather data. Promising results were obtained with an excellent real-time overall performance for next-day forecasting with a Root Mean Square Error (RMSE) value of 0.0706 using the Bayesian regularization algorithm with 28 hidden layers and all weather inputs. The Levenberg-Marquardt algorithm provided a 0.0753 RMSE using 23 hidden layers for the same set of learning inputs. This research shows that the Bayesian regularization algorithm outperforms the reported real-time prediction systems for the PV power production.**

Keywords—*Solar photovoltaic; solar irradiance; PV power forecasting; machine learning; artificial neural networks; Levenberg-Marquardt; Bayesian regularization*

I. INTRODUCTION

The rising fuel costs and increasing energy demands with the ongoing industrial growth and environmental awareness have engaged to the importance of new renewable energy sources such as the solar Photovoltaic (PV) systems [1, 2]. As one of the most important renewable energy sources, PV energy is becoming the dominant clean and reliable energy source that is widely used around the world without causing any damage to the environment. Mentioning the light-electricity process, the term "Photovoltaic" is first used by Alfred [3], as the light conversion process into electricity. There are two modes of installation for solar PV power plants: grid-tied and off-grid systems [4]. The first mode is widely used and proven to be hugely beneficial. It depends on the variable weather conditions according to the geographical area of the system which is the reason why it was known as uncertain, uncontrollable, and non-scheduling power source [5]. The second mode, off-grid systems, is used for isolated or remote areas that are normally on a smaller scale.

Many studies were reported in the literature suggesting different modeling, simulation, and prediction methods for the expected power production of solar PV plants for the purpose of improving the investment feasibility and maintaining a stable power quality and scheduling [6, 7]. Fonseca [8], compared the accuracy of one-day ahead prediction for the power produced by 1MW PV System using two methods: Support Vector Machines (SVM) and Multilayer Perceptron (MP) Artificial Neural Networks (ANNs). It was found that the two algorithms approximately obtained almost the same accuracy with 0.07 KWh/m^2 and 0.11 KWh/m^2 Mean Absolute Error (MAE) and Root Mean Square Error (RMSE), respectively.

Various forecasting methods of PV power output were reviewed in [9]. It was demonstrated that any model uses numerically predicted weather data will not take into account the effect of cloud cover and cloud formation when initializing, therefore sky imaging and satellite data methods used to predict the PV power output with higher accuracy. The article also outlined some key factors affecting the accuracy of prediction, such as forecast horizon, forecasting interval width, system size and PV panels mounting method (fixed or tracking). A model using multilayer perceptron-based ANN was proposed in [5] for one day ahead forecasting. The daily solar power output and atmospheric temperature for 70 days used for training the ANN. For the different settings of the ANN model (number of hidden layers, activation function, and learning rule), the minimum MAPE achieved was 0.855%.

The aim of the work published in [10] was to study the effect of forecast horizon on the accuracy of the method used to predict the PV power production, which was Support Vector Regression (SVR) using numerically predicted weather data. Two forecast horizons studied: up to 2 and 25 hours ahead. As expected, the forecasting of up to 2 hours ahead was more accurate with RMSE and MAE increased 13% and 17%, respectively, when the forecast horizon was up to 25 hours ahead. Cococcioni [11], developed and validated a model that adapted an ANN with tapped delay lines and built for one day ahead forecasting. The inputs were the irradiation and the sampling hours. The model achieved seasonal MAE ranging from 12.2% to 26% in spring and autumn, respectively. Monteiro [12], compared two short-term forecasting models: the analytical PV power forecasting model (APVF) and the MP PV forecasting model (MPVF), with both of the models using numerically predicted weather data and past hourly values for PV electric power production. The two models achieved similar results (RMSE varying between 11.95% and 12.10%) with forecast horizons covering all daylight hours of one day ahead, thus the models demonstrated their applicability for PV electric power prediction.

Leva [13], proposed a new Physical Hybrid ANN (PHANN) method to improve the accuracy of the standard ANN method. The hybrid method is based on ANN and clear sky curves for a PV plant. The PHANN method reduced the Normalized MAE (NMAE) and the Weighted MAE (WMAE) by almost 50% in many days compared to the standard ANN method. In [14], the PV energy production for the next day with 15-minutes intervals was accurately predicted with an SVM model that uses historical data for solar irradiance, ambient temperature, and past energy production. The method demonstrated very good accuracy with R^2 correlation coefficients of more than 90%, and the coefficient was strongly dependent on the quality of the weather forecast.

In our previous work [15], we proposed an initial real-time forecasting model for the PV power production using ANNs based on the available solar irradiation records for the last few days. In this research work, ANNs were optimized comparing the Levenberg-Marquardt (LM) and Bayesian Regularization (BR) algorithms to analyze and correlate the available data of temperature, solar irradiance, timing, and the generated solar PV power. The suggested system provides real-time PV energy forecasts for the next 24 hours based on real-time weather data for the last week.

II. PV AND WEATHER DATA

A. PV Systems

There are four separate PV systems installed at the university campus for a total generation capacity of 550KWp:

- PV ASU00 (The Test Field): This system was installed in 2013 with a capacity of 56.4KWp including a CPV tracker, a Polycrystalline tracker, Poly-crystalline and Mono-crystalline panels (South and East/West oriented), and thin film panels.

- PV ASU08 (The Library): A rooftop-mounted 130.1KWp system of Yingli Solar panels and SMA sunny tripower inverters.

Fig. 1. Rooftop mounted Solar panels on top of the engineering building.

- PV ASU09 (Faculty of Engineering): This is the largest rooftop-mounted PV system at ASU that is installed on top of the faculty of engineering building with a capacity of 264KWp [16]. It consists of 14 SMA sunny tripower inverters (17KW and 10KW) connected with Yingli Solar (YL 245P-29b-PC) panels that are tilted by 11°and oriented 36°(S to E) (see Fig. 1).

- PV ASU10 (Deanship of Student Affairs): A rooftop mounted 117.4KWp system of Yingli Solar panels and SMA sunny tripower inverters.

B. ASU Weather Station

ASU's weather station was installed in 2015 to be the first of its kind in Jordan providing a wide range of weather data measured by the latest sensors and devices [17]. It is located about 175m from the engineering building as shown on the map of Fig. 2. The station is equipped with many instruments to measure:

- The wind speed and direction (4 altitudes between 10-36m).

- Ambient temperature (3 altitudes between 1-35m).

- Relative humidity (at altitudes 1m and 35m).

- Global solar irradiance (including separate direct radiation and diffuse radiation records).

- Subsoil and soil surface temperature.

- Barometric pressure.

- Precipitation amounts.

The real-time weather measurements obtained by the station are updated continuously and published on the website: http://energy.asu.edu.jo as depicted in Fig. 3.

In this research work, we created a two years dataset using the available hourly records of weather and PV energy data for the duration between 16 May 2015 and 15 May 2017. This dataset for 731 days includes 17544 weather records and 17544 PV power values. PV power records were obtained only from the PV ASU09 system as it is the largest and most stable system on campus.

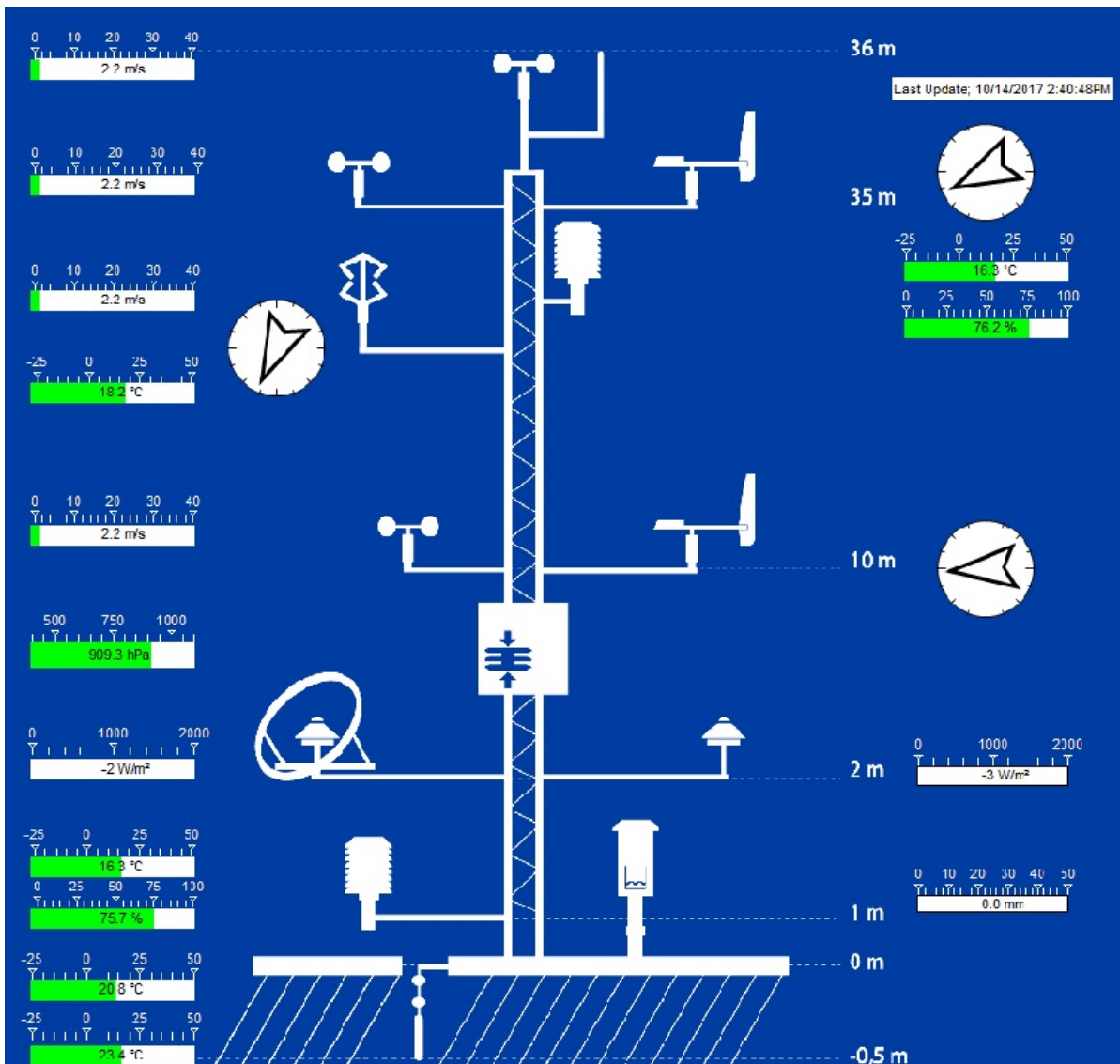Fig. 2.   A map showing part of ASU's campus.

Fig. 3.   Real-time weather station data available at [17].

III. Forecasting System

The suggested forecasting system can be represented by the block diagram of Fig. 4 and is described in the following subsections.

A. Data Filtering and Association

The collected weather and PV power data are first filtered, as shown in Fig. 4, to filter out any missing records which guarantee the consistency of the dataset. This includes any weather information with no PV power values associated at the same time or any PV power records with a missing weather data.

Based on the timing data, associations were found by matching solar PV power records with weather records including the record time, temperature, and global solar irradiance. To obtain homogeneous data and reliable machine learning experiments, the final dataset was normalized between 0 and 1.

B. LM and BR ANNs

Compared with similar algorithms [18, 19, 20], ANNs are known as one of the most powerful machine learning techniques with a wide range of applications [21, 22, 23]. ANNs map non-linear inputs through adjustable weights into the desired targets. The network is created by three layers: the input, hidden, and output layers [24] as illustrated in the example of Fig. 5 for a network of 11 inputs, 23 hidden layers, and one output.

ANNs showed excellent learning and classification performances while dealing with real-world sensor data [25, 26, 27].
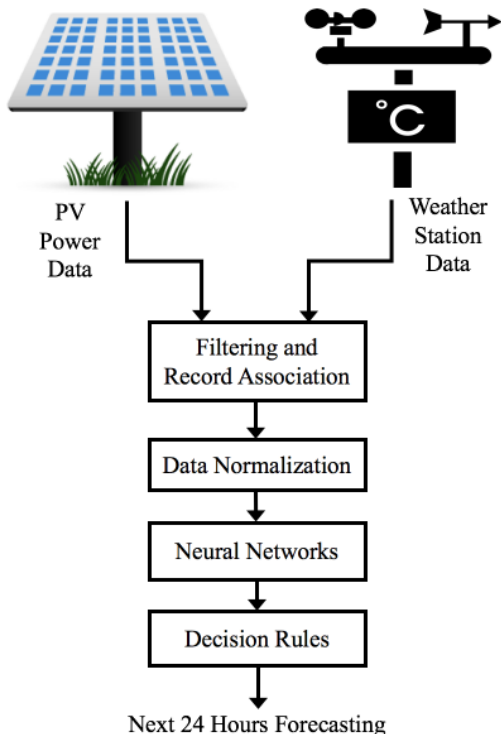

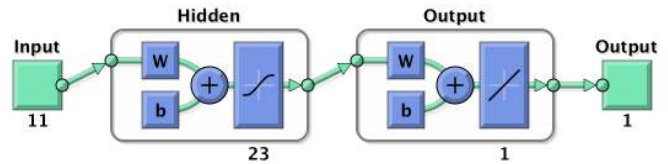Fig. 4.   Implementation flowchart for the proposed forecasting system.

Fig. 5.   An example for the structure of ANNs with 23 hidden layers.

In this work, we applied the LM and BR to neural networks and compared the learning performances using different training configurations.

The LM backpropagation optimization algorithm has been initially reported in [28] and has been applied later to neural networks in [29, 30]. The LM ANNs algorithm is implemented in MATLAB and it is known as the fastest backpropagation supervised algorithm especially while training feedforward ANNs with moderate sizes [31]. The BR backpropagation algorithm has been introduced in [32] and [33] and it is implemented in MATLAB [34]. Both of the LM and BR ANNs calculate the neural network errors' derivative functions with respect to weights and biases to obtain a Jacobian matrix that is used for calculations which means that the performance can only be measured by the mean squared errors [29].

C. Training and Testing Experiments

Data of the global solar irradiance ($Rad_d(t)$) and the temperature ($Temp_d(t)$) at the altitude of 1m are used to form the weather information vector $W_d(t)$ at time t of day d. Two neural network models were created based on the LM and BR algorithms with the target function of the mean PV power $P_d(t)$. The inputs to these models are the current time stamp from the beginning of the current year ($T_d(t)$) and the available weather vectors $W_d(t)$ at the same time t over the previous five days before day d. So, as depicted in Fig. 6, each input sample of the training dataset consists of:

$$T_d(t), W_{d-1}(t), W_{d-2}(t), W_{d-3}(t), W_{d-4}(t), W_{d-5}(t) \quad (1)$$

and one output value $P_d(t)$ with

$$W_{d-i}(t) = \{Rad_{d-i}(t), Temp_{d-i}(t)\}, \quad i = 1, 2, ..., 5 \quad (2)$$

In this work, the MATLAB Neural Networks toolbox was used for huge amounts of training, validation, and testing experiments while using different input combinations and varying the number of hidden layers from 1 to 30. The used set of inputs are:

- ALL inputs (1 time value, 5 radiation values, and 5 temperature values).

- Only 5 radiation values.

- Only 5 temperature values.

- 6 inputs (1 time value and 5 radiation values).

- 6 inputs (1 time value and 5 temperature values).

- 3 inputs (1 time value, 1 radiation value, and 1 temperature value).
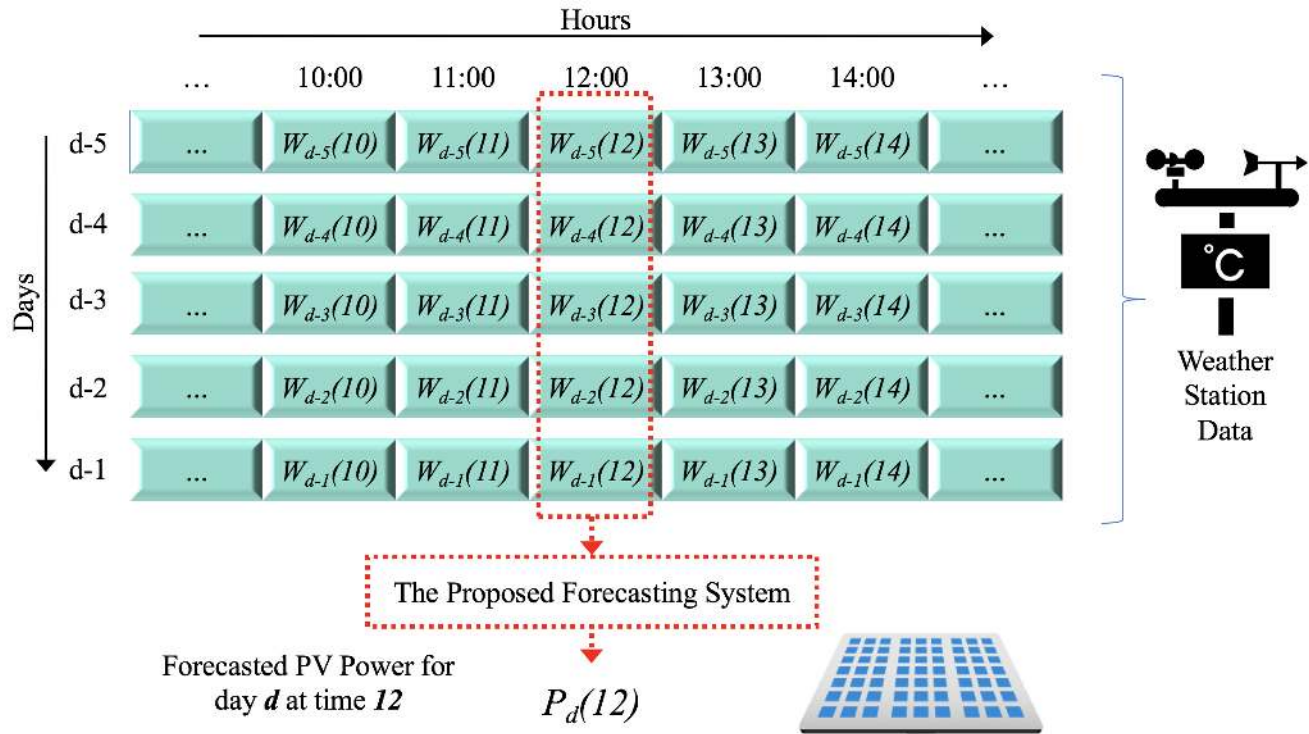
Fig. 6. Next-day PV forecasting based on the weather data of the previous five consecutive days.

- 10 inputs (5 radiation values and 5 temperature values).

Ten experiments were handled at each value for the number of hidden layers. At each experiment, the samples of the dataset were randomly mixed to generate the sub-datasets: 80% for training, 5% for validation, and 15% for testing. Then, the performance was evaluated by calculating the average RMSE for each of ten experiments using:

$$RMSE = \sqrt{\sum_{n=1}^{N} D_n^2} \qquad (3)$$

The network configurations that provided the best performances are listed in Table I. It can be concluded from the results that the best training/testing experiments provided an average RMSE of 0.0706 and a best testing correlation coefficient of R=0.9660 and mean square error of 0.00485 while using all inputs to the BR ANNs with 28 hidden layers for the testing performance illustrated in Fig. 7 and 8.

A histogram of 20 Bins is depicted in Fig. 9 for the overall errors. These results are very low compared to the methods and measures reported in the literature and related to the current research as summarized in Table II.

IV. CONCLUSION

In this research, a predictive forecasting model is proposed by applying the Levenberg-Marquardt and Bayesian Regularization algorithms to neural networks for the purpose of correlating historical weather data to photovoltaic outputs. Two years of hourly data were processed to associate the available
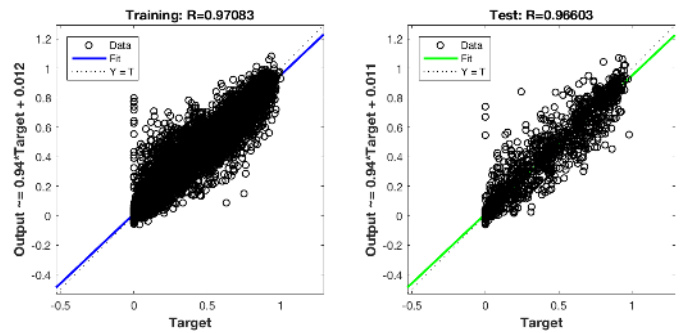

Fig. 7. Correlation coefficients calculations for the best performance.
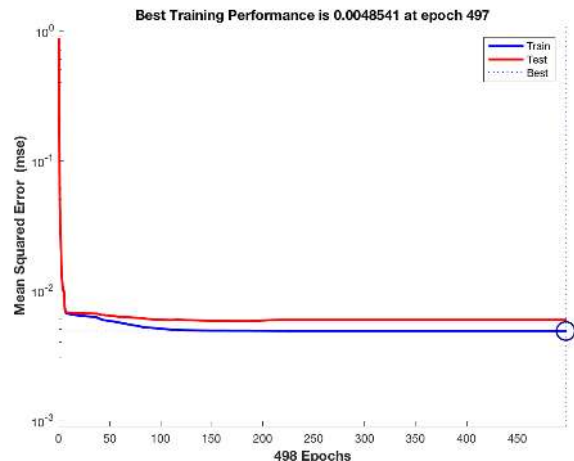

Fig. 8. Best ANN training performance using 28 hidden layers.

TABLE I. RESULTS OBTAINED USING DIFFERENT SETS OF INPUTS

Inputs	Levenberg-Marquardt			Bayesian Regularization		
	Hidden Neurons	Average RMSE	Testing R	Hidden Neurons	Average RMSE	Testing R
ALL	23	0.0753	0.9532	28	0.0706	0.9660
Rad	23	0.0828	0.9417	30	0.0797	0.9514
Temp	23	0.2340	0.9034	28	0.2285	0.8911
Time, Rad	24	0.0819	0.9420	26	0.0774	0.9501
Time, Temp	27	0.2107	0.9041	30	0.2061	0.8921
Time, Rad, Temp	23	0.0935	0.9271	14	0.0941	0.9384
Rad, Temp	27	0.0969	0.9385	15	0.0966	0.9295

TABLE II. A COMPARISON BETWEEN THE FORECASTING PERFORMANCE FOR DIFFERENT METHODS AND MEASURES RELATED TO THE CURRENT RESEARCH

| Reference	Forecasts	Method	Measure	Result
[5]	One day ahead	MP-ANNs	MAPE	0.855%
[8]	One day ahead	SVM	MAE, RMSE	0.07 KWh/m^2, 0.12 KWh/m^2
[8]	One day ahead	MP-ANNs	MAE, RMSE	0.11 KWh/m^2, 0.12 KWh/m^2
[10]	2 hours ahead	SVR	MAE	0.065 MWh
[10]	25 hours ahead	SVR	MAE	0.076 MWh
[11]	One day ahead	ANN	MAE	0.122
[12]	One day ahead	APVF	RMSE, MAE	0.121, 0.0597
[12]	One day ahead	MPVF	RMSE, MAE	0.1195, 0.0646
[13]	One day ahead	PHANN	NMAE, WMAE	50% error reduction
[14]	One day ahead	SVM	R^2 correlation coefficients	90%
[15]	One day ahead	ANN	RMSE, R^2 correlation coefficients	0.0721, 96.7%
This work	One day ahead	LM-ANN	RMSE, R^2 correlation coefficients	0.0753, 95.32%
This work	One day ahead	BR-ANN	RMSE, R^2 correlation coefficients	0.0706, 96.60%
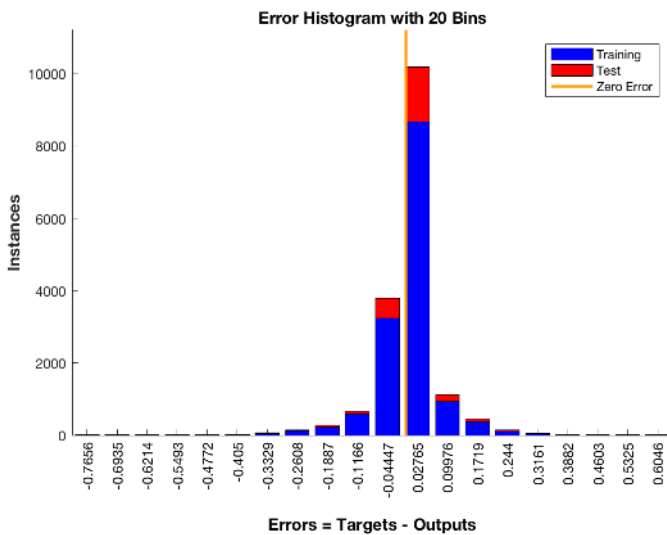
Fig. 9. Error histogram of 20 Bins.

temperature and global radiation records to the generated PV power. The associated datasets were used as a source of learning for a neural network model that use real-time weather data to provide PV power forecasts for the next 24 hours.

After a vast amount of training/testing experiments, excellent prediction results were obtained using the BR ANNs based on time, temperature, and radiation inputs. These predictions can be used by many energy management systems and power control systems of grid-tied PV plants. The proposed model is being developed into a real-time online application in our near future work.

ACKNOWLEDGEMENT

REFERENCES

[1] P. Ramsami and V. Oree, "A hybrid method for forecasting the energy output of photovoltaic systems," *Energy Conversion and Management*, vol. 95, no. Supplement C, pp. 406 – 413, 2015.

[2] I. E. Agency, "Technology roadmap: Solar photovoltaic energy 2014 edition," Paris, France, 2014, last accessed: 2017-10-2. [Online]. Available: https://goo.gl/6opzZ8

[3] A. Smee, *Elements of electro-biology,: or the voltaic mechanism of man; of electro-pathology, especially of the nervous system; and of electro-therapeutics*. London, UK: Longman, Brown, Green, and Longmans, 1849.

[4] V. Karthikeyan, S. Rajasekar, V. Das, P. Karuppanan, and A. K. Singh, *Grid-Connected and Off-Grid Solar Photovoltaic System*. Cham: Springer International Publishing, 2017, pp. 125–157.

[5] R. Muhammad Ehsan, S. P. Simon, and P. R. Venkateswaran, "Day-ahead forecasting of solar photovoltaic output power using multilayer perceptron," *Neural Computing and Applications*, vol. 28, no. 12, pp. 3981–3992, Dec 2017.

[6] M. Elyaqouti, L. Bouhouch, and A. Ihlal, "Modelling and predicting of the characteristics of a photovoltaic generator on a horizontal and tilted surface," *International Journal of Electrical and Computer Engineering*, vol. 6, no. 6, pp. 2557–2576, 2016.

[7] S. Mankour, A. Belarbi, and M. Benmessaoud, "Modeling and simulation of a photovoltaic field for 13 kw," *International Journal of Electrical and Computer Engineering*, vol. 7, no. 6, pp. 3271–3281, 2017.

[8] J. G. da Silva Fonseca Junior, T. Oozekia, T. Takashimaa, G. Koshimizub, Y. Uchidab, and K. Ogimoto, "Forecast of power production of a photovoltaic power plant in japan with multilayer perceptron artificial neural networks and support vector machines," in *26th European Photovoltaic Solar Energy Conference and Exhibition*. WIP-Renewable Energies, Sep. 2011, pp. 4237–4240.

[9] A. Tuohy, J. Zack, S. Haupt, J. Sharp, M. Ahlstrom, S. Dise, E. Grimit, C. Moehrlen, M. Lange, M. Garcia Casado, J. Black, M. Marquis, and C. Collier, "Solar forecasting: Methods, challenges, and performance," *IEEE Power and Energy Magazine*, vol. 13, pp. 50–59, 11 2015.

[10] J. G. da Silva Fonseca, T. Oozeki, T. Takashima, G. Koshimizu, Y. Uchida, and K. Ogimoto, "Photovoltaic power production forecasts with support vector regression: A study on the forecast horizon," in *2011 37th IEEE Photovoltaic Specialists Conference*, June 2011, pp. 002 579–002 583.

[11] M. Cococcioni, E. D'Andrea, and B. Lazzerini, "One day-ahead forecasting of energy production in solar photovoltaic installations: An empirical study," *Intelligent Decision Technologies*, vol. 6, no. 3, pp. 197–210, Aug. 2012.

[12] C. Monteiro, L. A. Fernandez-Jimenez, I. J. Ramirez-Rosado, and P. M. Lara-Santillan, "Short-term forecasting models for photovoltaic plants: Analytical versus soft-computing techniques," *Mathematical Problems in Engineering*, vol. 2013, pp. 1–9, 2013.

[13] A. Dolara, F. Grimaccia, S. Leva, M. Mussetta, and E. Ogliari, "A physical hybrid artificial neural network for short term forecasting of pv plant power output," *Energies*, vol. 8, no. 2, pp. 1138–1153, 2015.

[14] R. Leone, M. Pietrini, and A. Giovannelli, "Photovoltaic energy production forecast using support vector regression," *Neural Computing and Applications*, vol. 26, no. 8, pp. 1955–1962, Nov. 2015.

[15] M. H. Alomari, J. Adeeb, and O. Younis, "Solar photovoltaic power forecasting in jordan using artificial neural networks," *International Journal of Electrical and Computer Engineering*, vol. 8, no. 1, pp. 497–504, February 2018.

[16] ASU, "Pv system asu09: Faculty of engineering," 2017, last accessed: 2017-10-14. [Online]. Available: https://goo.gl/cxGYVb

[17] U. ASU, "Weather station," 2017, last accessed: 2017-10-14. [Online]. Available: http://web.asu.edu.jo/wsp

[18] R. Qahwaji, M. Al-Omari, T. Colak, and S. Ipson, "Using the real, gentle and modest adaboost learning algorithms to investigate the computerised associations between coronal mass ejections and filaments," in *Mosharaka International Conference on Communications, Computers and Applications*, 2008, pp. 37–42.

[19] M. AL-Omari, R. Qahwaji, T. Colak, S. Ipson, and C. Balch, "Next-day prediction of sunspots area and mcintosh classifications using hidden markov models," in *2009 International Conference on CyberWorlds*, Sept 2009, pp. 253–256.

[20] M. Al-Omari, R. Qahwaji, T. Colak, and S. Ipson, "Machine leaning-based investigation of the associations between cmes and filaments," *Solar Physics*, vol. 262, no. 2, pp. 511–539, Apr 2010.

[21] M. AL-Omari, R. Qahwaji, T. Colak, S. Ipson, and C. Balch, "Next-day prediction of sunspots area and mcintosh classifications using hidden markov models," in *2009 International Conference on CyberWorlds*, Sept 2009, pp. 253–256.

[22] M. H. Alomari, R. S. Qahwaji, and S. S. Ipson, *Applied Machine Learning for Solar Data Processing: Developing Automated Technologies for Knowledge Extraction and Prediction of Solar Activities Using Machine Learning*. Germany: LAP Lambert Academic Publishing, 2011.

[23] M. H. Alomari, "Mind controlled universal tv remote control," in *2016 International Conference on Image Processing, Production and Computer Science*, London (UK), Mar 2016, pp. 85–91.

[24] N. J. Nilsson, "Introduction to machine learning. an early draft of a proposed textbook," 1996.

[25] T. M. Mitchell, *Machine Learning*, 1st ed. New York, NY, USA: McGraw-Hill, Inc., 1997.

[26] M. Alomari, E. Awada, and O. Younis, "Subject-independent eeg-based discrimination between imagined and executed, right and left fists movements," *European Journal of Scientific Research*, vol. 118, no. 3, pp. 364–373, 02 2014.

[27] R. Qahwaji, T. Colak, M. Al-Omari, and S. Ipson, "Automated prediction of cmes using machine learning of cme – flare associations," *Solar Physics*, vol. 248, no. 2, pp. 471–483, Apr 2008.

[28] D. W. Marquardt, "An algorithm for least-squares estimation of non-linear parameters," *Journal of the Society for Industrial and Applied Mathematics*, vol. 11, no. 2, pp. 431–441, 1963.

[29] M. T. Hagan and M. B. Menhaj, "Training feedforward networks with the marquardt algorithm," *IEEE Transactions on Neural Networks*, vol. 5, no. 6, pp. 989–993, Nov 1994.

[30] M. Hagan, H. Demuth, and M. Beale, *Neural Network Design*. Boston, MA, USA: PWS Publishing, 1996.

[31] MathWorks, "trainlm: Levenberg-marquardt backpropagation," 2017, last accessed: 2017-12-14. [Online]. Available: https://www.mathworks.com/help/nnet/ref/trainlm.html

[32] D. MacKay, "Bayesian interpolation," *Neural computation*, vol. 4, no. 3, pp. 415–447, 1992.

[33] F. D. Foresee and M. T. Hagan, "Gauss-newton approximation to bayesian learning," in *Proceedings of the International Joint Conference on Neural Networks*, June 1997, pp. 1930–1935.

[34] MathWorks, "trainbr: Bayesian regularization backpropagation," 2017, last accessed: 2017-12-14. [Online]. Available: https://www.mathworks.com/help/nnet/ref/trainbr.html