Renewable Energy 150 (2020) 797–807

Contents lists available at ScienceDirect

Renewable Energy

journal homepage: www.elsevier.com/locate/renene

Time series forecasting of solar power generation for large-scale photovoltaic plants

Hussein Sharadga, Shima Hajimirza*, Robert S. Balog

Texas A&M University, College Station, TX, 77840, USA

ARTICLE INFO

Article history:
Received 9 September 2019
Received in revised form
15 November 2019
Accepted 28 December 2019
Available online 30 December 2019

Keywords:
PV power forecasting
Grid-connected PV plant
Deep learning
Neural network
Statistical methods
Time series analysis

ABSTRACT

Accurate solar power forecasting is essential for grid-connected photovoltaic (PV) systems especially in case of fluctuating environmental conditions. The prediction of PV power output is critical to secure grid operation, scheduling and grid energy management. One of the key elements in PV output prediction is time series analysis especially in locations where the historical solar radiation measurements or other weather parameters have not been recorded. In this work, several time series prediction methods including the statistical methods and those based on artificial intelligence are introduced and compared rigorously for PV power output prediction. Moreover, the effect of prediction time horizon variation for all the algorithms is investigated. Hourly solar power forecasting is carried out to verify the effectiveness of different models. The data utilized in the current work comprises 3640 h of operation data taken from a 20 MW grid-connected PV station in China.

1. Introduction

Photovoltaic (PV) energy systems are one of the most widespread and desirable renewable energy technologies due to their high energy productivity potential. Accurate prediction of PV power is important for the integration of PV systems with the smart grids. The prediction of PV power output is essential in cases where large scale PV systems are connected to the grid or when a large number of small scale PV systems are installed on the utility end.

Studies on PV-generated power forecasting is limited [1]. Most of the published work in this field focuses on solar radiation prediction. However, the output power of a PV module has more dependencies that just the solar irradiance. Factors including the conditions of the cells, the type of solar cells, electrical circuit of the module, angle of incident, weather conditions and other paremeter all impact the electrical power generated. For example, the amount of produced power isinfluenced by the temperature of the solar cell in a PV system [2]. The cell temperature is a function of the solar irradiance power, ambient temperature, wind speed and relative humidity. PV output power forecasting based on the principles of weather classification has been studied by several authors [3–7]. Shi et al. [3] introduced a novel prediction model to estimate the power output of PV plant of 20 MW capacity for a 24 h time horizon by applying support vector machine for weather classification. Neural networks (NN) were adopted for a PV station output power short-term forecasting in Ref. [4]. The model was based on the solar radiation, seasonality and weather classification (*i.e.* sunny, cloudy or overcast). Mellit et al. [5] used two artificial neural networks to predict the power produced by a 50 W PV plant using more than one year of data. One network was trained for cloudy days one for sunny days, with three inputs to each network: cell temperature, solar radiation and PV voltage. In Ref. [6], a self-organized map was implemented to determine the day weather classifications: sunny, cloudy and rainy, thus allowing to select the appropriate architecture and the training parameters for each day ahead of time. An accurate neural network has been developed for PV cells power estimation of large-scale grid-connected photovoltaic plants in Ref. [7]. The model takes three different types of days into account: sunny, partly cloudy and overcast. The network was trained using the data of solar radiation, PV cell temperature and electric power of one-Megawatt solar plant. Deep learning NNs have also been proposed for prediction and modeling. Long short-term memory (LSTM) architectures have been implemented in PV power forecasting due to the ability to preserve past time-series information using a memory architecture [8]. Auto-Encoders and LSTM have shown to be more efficient compared to multi-layer perceptron and

* Corresponding author. J. Mike Walker '66 Department of Mechanical Engineering, Texas A&M University, College Station, TX, 77840, USA.
E-mail address: shima.hm@tamu.edu (S. Hajimirza).

https://doi.org/10.1016/j.renene.2019.12.131
0960-1481/

physical prediction methods in power prediction when applied to 21 studied PV plants [9].

Different neural network architectures and combinations along with time-series filtering techniques have been introduced and applied for PV power estimation. Wavelet transform was used for data filtering of ill-behaved PV power and then combined with neural network for one-hour-ahead power forecasting in Ref. [10]. It was claimed that the proposed combination addressed inaccuracy issues with spikes and chaotic changes in the input time series. In Ref. [11] Focused Time delay NNs and Distributed Time delay NNs were compared and the former was proven superior when applied to the estimation of the output of a 5 kW PV power plant. Support vector machine (SVM) and seasonal auto-regressive integrated moving average (SARIMA) models were combined and employed for power forecasting of 20 kW grid-connected PV system in Ref. [12]. It was demonstrated that the proposed hybrid system can capture the nonlinearity behavior of time input time-series better than both SVM and SARIMA models individually. In Ref. [13], support vector regression with the assistance of numerically predicted cloudiness was implemented to predict the PV plant power output. Numerical prediction of cloudiness was found to play an important role in the accuracy of the overall model. A hybrid system consisting of two numerical weather prediction models with different NNs of different configurations was used as an accurate short-term power prediction approach for photovoltaic plants in Ref. [14]. Support vector regression assisted with fuzzy inference was employed in Ref. [15] for one-day-ahead hourly prediction of power produced by PV systems. The data was classified based on the day type using learning vector quantization (LVQ) networks and self-organizing map prior to prediction.

In addition to those results, several other works have used NNs to predict the PV power output, though no solar radiation measurements have been used. A NARX network was used to predict the yield of grid-connected PV plant in Ref. [16]. The proposed model was based on air temperature measurements as well as the calculated clear-sky radiation incident using Hottel's radiation model. In Ref. [17] it was demonstrated that neural networks, assisted with a data of aerosol index and without any solar irradiance measurements, are more accurate than model based on conventional NNs that use solar irradiance measurements for PV power prediction.

Mellit et al. [18] proposed using recurrent neural network (RNN) to estimate the daily power production of PV power systems. The NN model was trained on 4 years of data and tested on another year. The model accepts the ambient temperature and daily global radiation (W/m^2 h) as inputs. Ashraf et al. [19] developed a NN to estimate the electricity generation of grid-connected PV systems based on the clearness index. Two multi-layer perceptron (MLP)-based NNs were used to forecast the output power generated by the PV plants [20]. The first one uses solar irradiance and ambient temperate as input while the second model is based on the solar irradiance only. The plant used a maximum power point tracking controller and is a roof-mounted 20 kW capacity facility.

The prediction algorithms are also applied to different renewable energy technologies. In Ref. [21], the hybrid neuro-fuzzy was employed to predict the electrical power output of a wind turbine. To enhance the prediction accuracy, the data was divided into four subsets based on the seasons of year. Wong et al. [22] introduced a new prediction tool to estimate the performance of a biofuel engine. The new method was called sparse Bayesian extreme learning machine. The proposed algorithm can be competitive with different neural network configurations. Kavousi-Fard [23] used a hybrid prediction tool to estimate the speed and direction of tidal currents. The proposed forecaster was based on combining one machine-learning algorithm with a well-known statistical tool: the support vector regression and ARIMA model, respectively. In Ref. [24], the artificial neural network was applied to predict the power output of a small hydropower plant power.

In general, it is understood that when used properly, NNs do a better job of predicting output PV power than analytical models, standard time series forecasting techniques and polynomial and multiple linear regression models. Nevertheless, none of the mentioned previous studies have systematically compared the performance of the different approaches.

Contributions. In the current work, different time series forecasting models are compared for PV output power prediction. The methods include both statistical (persistent) methods and those based on artificial intelligence. The statistical models used in the current work belong to the category of persistence models including: Autoregressive moving average (ARMA), autoregressive integrated moving average (ARIMA) and seasonal auto-regressive integrated moving average (SARIMA). Six different types of NN models are considered: Bi-directional long short term memory (Bi-LSTM), long short term memory (LSTM), fuzzy c-mean clustering, layer recurrent (LRNN), multi-layer perceptron (MLP) and feedforward NNs. In addition, the main novelties of the paper can be summarized as:

1. We propose a new deep learning BI-LSTM algorithm as an accurate power prediction model for large scale PV plants.
2. We assess and compare the performance of different NNs and statistical approaches for a time series forecasting of large scale PV systems.
3. We study the time horizons at which the studied predictive models work reliably.

2. Predictive models

2.1. Dataset

The data is from a grid-connected PV power plant located in south China which was used in a previous study [3]. The tilt angle of the PV arrays is fixed. The system has a maximum capability of 20 MW and the data output power is recorded within the date range from 01/13/2010 to 10/29/2010. During this date range, the PV plant power output is consistently zero between 8 p.m. and 6 a.m. The power between 7 and 8 p.m. is very low and is zero most of the times. Thus, we only consider the power between 6 a.m. and 7 p.m. in the simulations. The power data is recorded every 15 min. We converted those values to hourly average data. Averaging those values makes the signal smoother, thus easier for algorithms to learn. We found that the time series prediction of PV power on an hourly average basis is more accurate than the prediction of the PV power of 15 min ahead. The data is normalized, and the outliers and missing values are removed using Hampel filter with a window size of 14 h, which is the maximum continuous daylight timeframe. The preprocessing procedure including outlier removals is described in Section 3.4.

2.2. Description of models

2.2.1. Artificial neural network

A neural network can learn complex dependencies between inputs and outputs. Artificial neural networks have been widely employed in different research disciplines for modeling nonlinear and complex functions. In a feedforward NN, the information is fed from the input layer to the output layer through the hidden layers. Layer recurrent neural networks are used in this study because they have dynamic responses to the input time series data. A layer

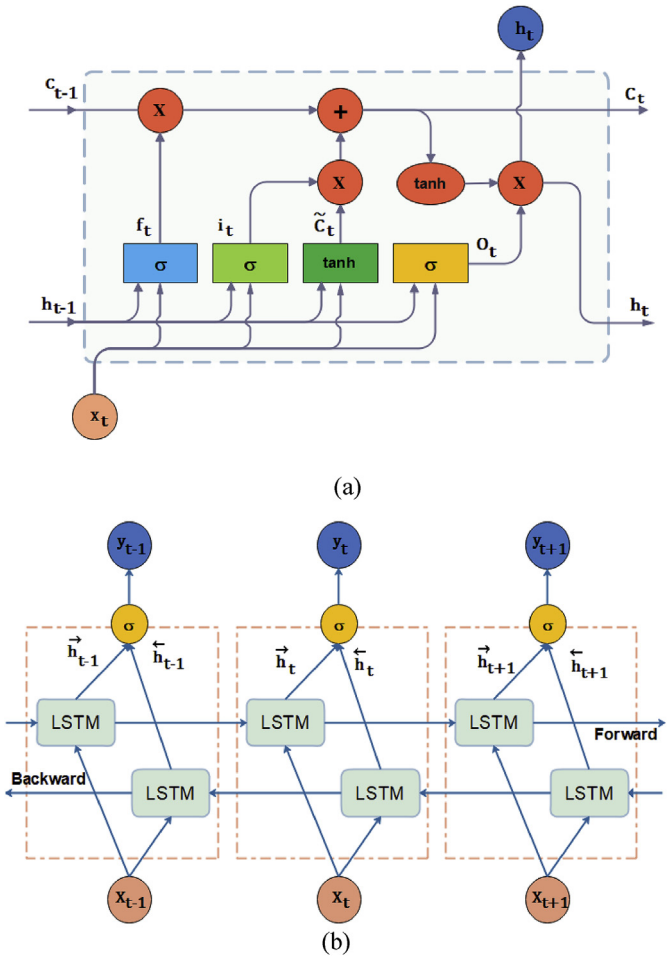


Fig. 1. Architecture of (a) LSTM cell (b) Bi-directional LSTM.

recurrent neural network has the same architecture as a feedforward network, except that the layers have recurrent connections with associated tap delays. A multilayer perceptron network (MLP) belongs to the feedforward NN category. In MLP, the input layer consists of neurons that use an activation function. LSTM is a class of recurrent neural networks developed to overcome long-term dependency problems [25]. Unlike conventional NNs, LSTM has a memory block to store the temporal information of the input data. The hidden layer of LSTM, also called LSTM cell, is shown in Fig. 1 (a). The LSTM cell connects the layer input, X_t, to the layer output, h_t, using a gate structure. The LSTM cell has three gates: input gate (i_t), output gate (O_t) and forget gate (f_t). The cell input state ($\widetilde{C}_t$) and the current cell output states (C_t), as well as cell output states of the previous irritation (C_{t-1}), are utilized in the training process. The

Table 1
Number of layers and neurons of NNs. LM and BR stand for trainlm and trainbr, respectively.

Algorithm	Training method	#layers	# neurons
Feedforward	trainlm	1	16
Feedforward	trainbr	2	[5,7]
MLP	trainlm	2	[5,8]
MLP	trainbr	2	[6,8]
Layer recurrent (LRNN)	trainlm	1	12
Layer recurrent (LRNN)	trainbr	2	[6,9]
LSTM	Adam	1	11
BI-LSTM	Adam	2	[7,8]

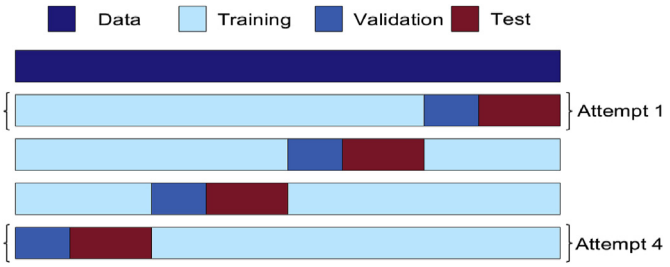


Fig. 2. Schematic of the time series cross-validation procedure.

cell input state, the cell output state and the gates can be calculated using the relations given in Ref. [26]. Bidirectional LSTM, which is shown in Fig. 1 (b), is developed using two time directions [27]. BI-LSTM is capable of learning more accurately that the unidirectional LSTM. BI-LSTM uses past and future information as inputs by connecting the neurons of two separate hidden layers to one output. The information is fed to every hidden layer from both the forward and backward layers. However, those neurons have no interactions in between. The output of BILSTM, y_t, is calculated by combining the two outputs ($\overrightarrow{h_t}$, $\overleftarrow{h_t}$) using the σ function. The σ function is the activation function in the artificial neuron that delivers an outputs based on inputs.

2.2.2. Statistical models

Autoregressive (AR) and moving average (MA) models are commonly used for time series prediction. When Autoregressive (AR) and moving average (MA) models are integrated, they result in a well-generalized time series forecasting model called ARMA. However, the ARMA model is inefficient for non-stationary time series analysis, which has led to the development of the ARIMA models which can handle non-stationarity [28]. The ARIMA model converts non-stationary time-series data to a stationary series via differencing. SARIMA model is a generalization of the ARIMA model developed by Box et al. [29] to handle the seasonal behavior of a time-series data. SARIMA model is often named with the associated parameters like SARIMA $(p, d, q) \times (P, D, Q)^m$, where p is the autoregressive order, d is the degree of integration, q is the moving average order and m is the seasonality order. For monthly and quarterly seasonal model, m equals 12 and 4, respectively. The capital letters (P, D and Q) refer to the seasonal part of the SARIMA formulation, *i.e.* P is the seasonal autoregressive order, D is the seasonal degree of integration and Q is the seasonal moving

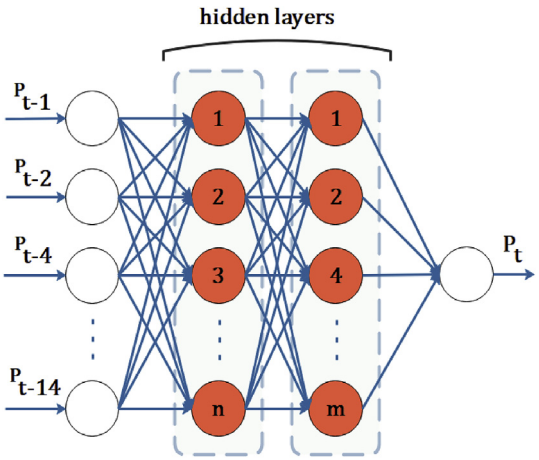


Fig. 3. Feedforward neural network structure used to predict the PV power output.

Table 2
One hour ahead forecasting using various NNs with 8 time delays (model 1), LM and BR stand for trainlm and trainbr, respectively.

	Algorithm	***R***(test data)	RMSE (test data)	Average computation time (s)
1	BI-LSTM	0.98	0.791	8.22
2	LSTM	0.965	0.841	8.61
3	LRNN-LM	0.959	0.961	2.54
4	LRNN- BR	0.953	0.995	8.21
5	FCM	0.95	1.11	22.1
6	FF- LM	0.947	1.05	5.43
7	MLP- LM	0.941	1.21	1.2
8	MLP- BR	0.936	1.16	1.85
9	FF-BR	0.932	1.27	6.11

Table 3
One hour ahead forecasting using various NNs with 4 time delays (model 2), LM and BR stand for trainlm and trainbr, respectively.

	Algorithm	***R***(test data)	RMSE (test data)	Average computation time (s)
1	LSTM	0.935	1.591	8.73
2	LRNN-LM	0.931	1.421	3.04
3	BI-LSTM	0.925	1.456	10.02
4	LRNN-BR	0.919	1.621	8.22
5	MLP-LM	0.908	1.632	1.21
6	MLP-BR	0.88	1.694	1.99
7	FF- BR	0.865	1.681	6.79
8	FCM	0.841	1.675	19.95
9	FF- LM	0.835	1.791	6.85

average order.

2.3. Models implementation

2.3.1. Artificial neural network implementation

The NNs are trained using two functions: trainbr and trainlm. trainbr (also known as Bayesian regularization) minimizes the combination of squared errors and weights. trainlm uses Levenberg-Marquardt optimization to tune biases and weights. trainlm is generally faster while trainbr is more likely to generalize the performance better and is thus more reliable out-sample. The appropriate number of layers and neurons in each hidden layer of NNs does not follow fixed theoretical guidelines [30]. In the current work, the number of layers and neurons is empirically chosen based on the mean squared error (MSE) as the performance measure [5]. The network configuration for each type of network for cross-validation attempt 5 (See Section 2.4) is listed in Table 1. LSTM and BI-LSTM are trained by Adam solver. The maximum number of epochs is set to 100. To check the validation, the performance of prediction is guided using correlation coefficient (R) and the root mean square error (RMSE).

2.3.2. Statistical models implementation

We use the following statistical models: ARMA, ARIMA and SARIMA. Akaike's Information Criterion (AIC) is widely used to optimize the model parameters in those models [31]. AIC is an estimation of the likelihood of a model. However, AIC does not have any indication of the absolute quality. A similar criterion for model selection is the Bayesian Information Criterion (BIC). The AIC and BIC of the forecasting models are determined by Ref. [32]:

$$\mathrm{AIC}(p+q) = T\log\left(r^2\right) + 2(p+q), \tag{1}$$

$$\mathrm{BIC}(M) = T\log(rss) + (p+q)\log T, \tag{2}$$

where p represents the autoregressive order, q is the moving average order, rss is residuals variance (sum squares) and T is the number of observations. Different combinations of parameters are investigated until the lowest AIC or BIC value is recorded. Then the model with the minimum (AIC, BIC) is considered in the forecasting. The plots of auto correlation function (ACF) and partial autocorrelation function (PACF) are also considered to help in making parameters choices.

2.4. Time series cross-validation

Cross-validation is a resampling procedure based on splitting the data into more than one training and testing subsets. Then, the overall performance of forecasters is basically obtained by looking at the prediction accuracy measures over all the testing subsets. Time series cross-validation is recommended as a key validation step in predictive models particularly in the case of small data samples. In the current work, the window sliding cross-validation method is implemented in the evaluation of the performance of predictors. The size of the training, validation and test sets is fixed while cross-validating: 2730 samples for training (%75), 364 samples for validation (%10) and 546 samples for testing (%15). Thus, the forecasters are tested on 2184 different samples (546×4). The performance measures are recorded for every cross-validation attempt and then used to estimate the overall prediction accuracy (Fig. 2).

3. Results and discussion

3.1. Comparison of the two NNs models

To predict the power production of PV plant, 8 time delays are applied and used as variables in the input layer of NNs, namely $t-i, i\in\{1,2,4,6,8,10,12,14\}$ hr (See Fig. 3). The results are compared with those forecasted by a model of 4 time delays $t-i, i\in\{1,2,3,4\}$ for one hour ahead power prediction. The input matrix and the corresponding output matrix used to train and test the suggested model of 8 delays are given below:

$$\text{input} = \begin{bmatrix} X_1 & X_2 & \dots & X_{3626} \\ X_2 & X_3 & \dots & X_{3627} \\ X_4 & X_5 & \dots & X_{3629} \\ \dots & \dots & \dots & \dots \\ X_{14} & X_{15} & \dots & X_{3639} \end{bmatrix},$$

$$\text{output} = [X_{15} \quad X_{16} \quad \dots \quad X_{3640}] \tag{3}$$

The simulation results summarized in Tables 2 and 3 show that the first model tracks the nonlinearity and the PV power dependency more reliably. Consequently, the first model is chosen as the model for the rest of work. In the following tables, LM and BR stand for trainlm and trainbr, respectively. The algorithms are sorted in the tables based on the correlation coefficient (R) of the testing set in the descending order. Average computational time is mentioned in the tables and will be used for comparison to the statistical algorithms. Average computational time is the average time required for training, testing and results generation for one validation attempt. Recall that correlation coefficient (R) and RMSE are defined as follows:

$$R = \frac{1}{N-1}\sum_{i=1}^{N}\left(\frac{O_i-\mu_O}{\sigma_O}\right)\left(\frac{P_i-\mu_P}{\sigma_P}\right), \tag{4}$$

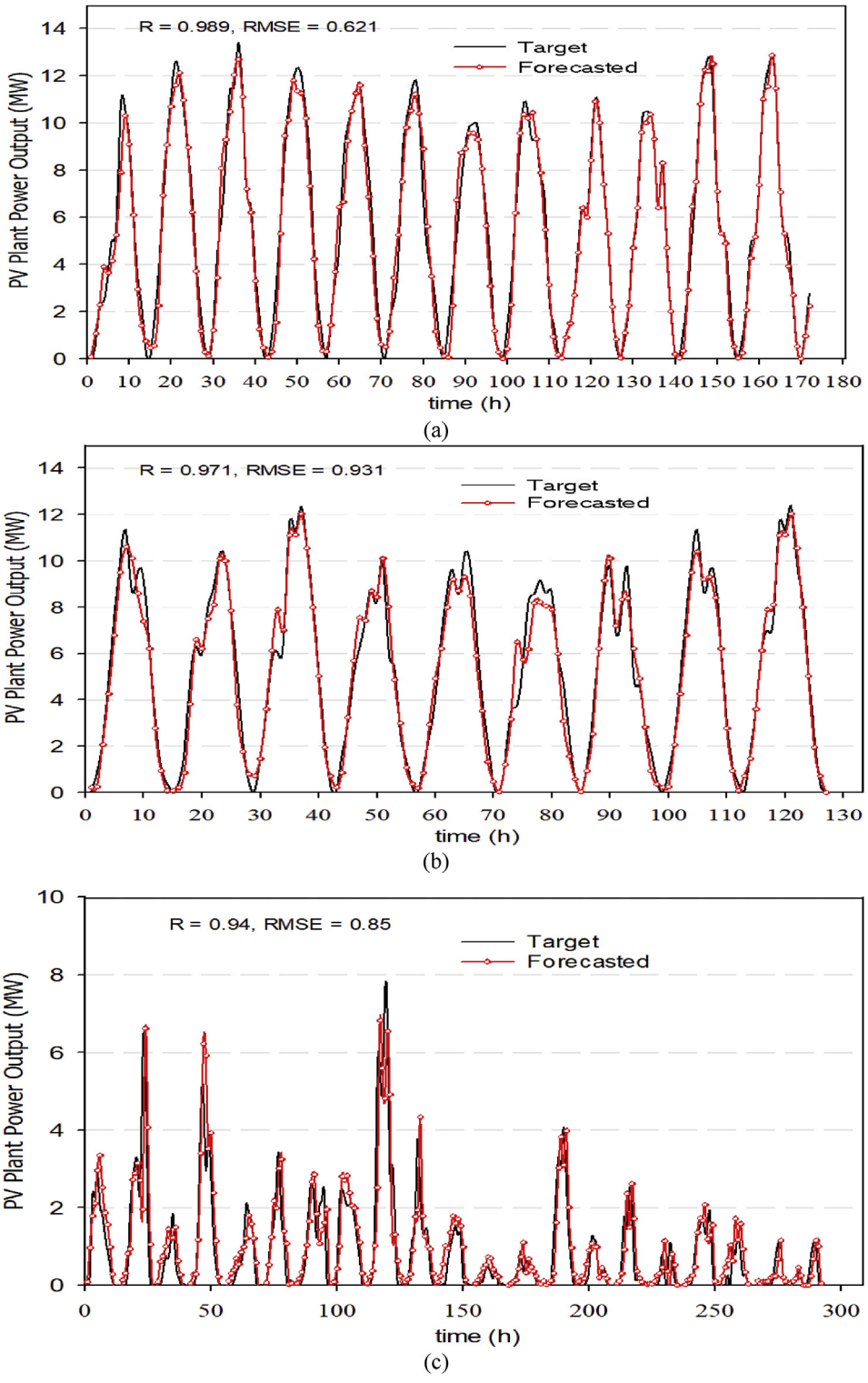


Fig. 4. Result of BI-LSTM forecast versus observed power production (a) sunny days (b) cloudy days (c) rainy days.

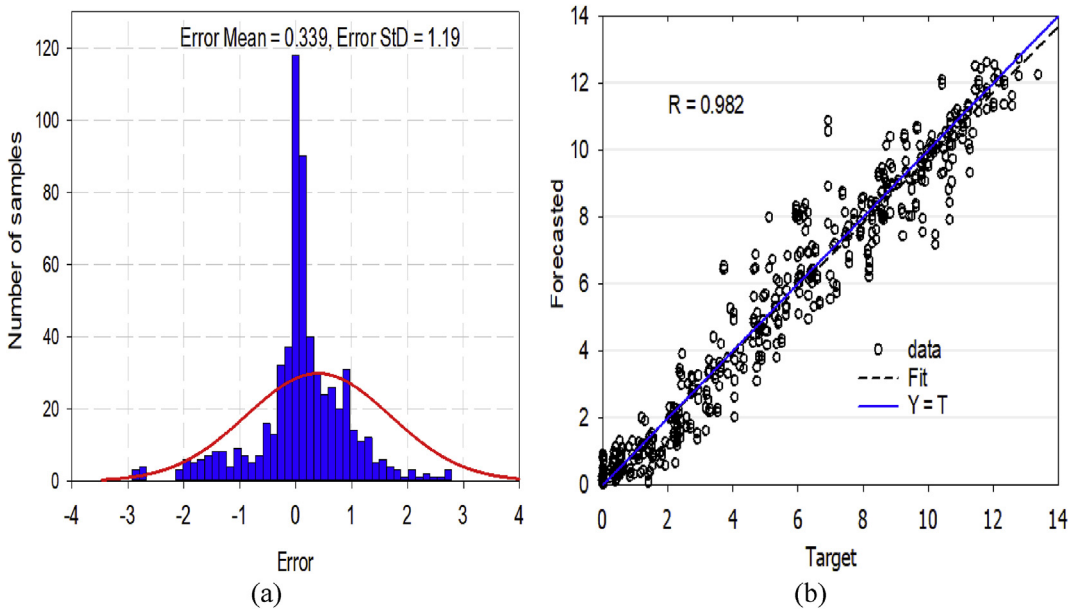


Fig. 5. Results of BI-LSTM for validation attempt number 4 for PV power prediction (a) Error histogram figure (real - forecast), (b) Regression plot.

Table 4
Two hours ahead forecasting using various NNs, 8 time delays.

	Algorithm	***R***(test data)	RMSE (test data)	Average computation time (s)
1	LSTM	0.961	1.102	8.73
2	BI-LSTM	0.952	1.331	9.02
3	LRNN- LM	0.944	1.395	3.21
4	MLP- LM	0.927	1.365	1.42
5	LRNN- BR	0.915	1.423	8.92
6	FCM	0.902	1.591	22.9
7	FF- LM	0.896	1.571	6.21
8	MLP- BR	0.889	1.792	1.9
9	FF- BR	0.785	1.821	6.8

Table 5
Three hours ahead forecasting using various NNs, 8 time delays.

	Algorithm	***R***(test data)	RMSE (test data)	Average computation time (s)
1	LRNN- LM	0.895	1.805	3.41
2	LSTM	0.894	1.824	9.41
3	BI-LSTM	0.893	1.812	8.65
4	LRNN-BR	0.881	1.822	8.95
5	MLP- BR	0.861	2.105	1.91
6	MLP- LM	0.853	2.010	1.74
7	FF- LM	0.841	2.314	6.21
8	FCM	0.801	2.521	23
9	FF- BR	0.741	2.563	6.41

$$RMSE = \sqrt{\sum_{i=1}^{N} \frac{(P_i - O_i)^2}{N}} \quad (5)$$

where N is the number of samples, P is the predicted value, O is the observed value, μ is the mean and σ is the standard deviation. The Overall RMSE and the overall R are calculated by combing the results of every validation attempt in two columns: one for observed and one for forecasted and then using equation (4) to calculate R and equation (5) to calculate RMSE. However, The Overall RMSE can also be obtained using the following expression:

$$RMSE_{overall} = \sqrt{\frac{MSE_1 + MSE_2 + MSE_3 + MSE_4 + MSE_5}{5}} \quad (6)$$

Where the subscripts 1,2 …,5 stand for the validation attempt number, and MSE is the mean square error.

BI-LSTM scored the highest correlation coefficient (R) and the lowest root mean square error (RMSE) as shown in Table 2. Therefore, we can conclude that BI-LSTM with model 1 can be the most accurate forecasting algorithm when it is used for PV power prediction of the sample system. The Bi-directional LSTM can learn faster than the one-directional LSTM. BI-LSTM is found to be the most accurate model with a correlation coefficient (R) of 98% and RMSE of 0.791. On the other hand, a multi-layer perceptron (MLP) trained by the two training functions is the fastest algorithm with an average computation time of 1–2 s. However, the correlation coefficient (R) ranges between 93.4 and 98% and RMSE ranges between 0.791 and 1.25 for different NNs with model 1.

Fig. 4 presents the predicted PV power using BI-LSTM algorithm with 8 time delays (model 1) for different day types: sunny, cloudy and rainy. It is clear that the observations and the predictions are in good agreement. The result for validation attempt number 4 (see section 2.4) is shown in Fig. 5. As seen, most of error values range between 0 and 1. The predicted value of the PV power output is usually found to be less than the observed, i.e., the error is frequently positive. In addition, in 220 of the 546 samples the error is almost zero. We can conclude that BI-LSTM shows a strong potential to predict the PV output efficiently in short term time horizon of 1 h ahead.

3.2. NNs forecasting with 2 and 3 h ahead horizon

To evaluate the performance of different NNs for mid-term forecasting, 2 and 3 h ahead forecasts are tested, and the findings are summarized in Tables 4 and 5. BI-LSTM yields the most accurate prediction. The correlation coefficient (R) of feedforward NN with trainbr changes from 93.2% to 78.5% when used for two hours

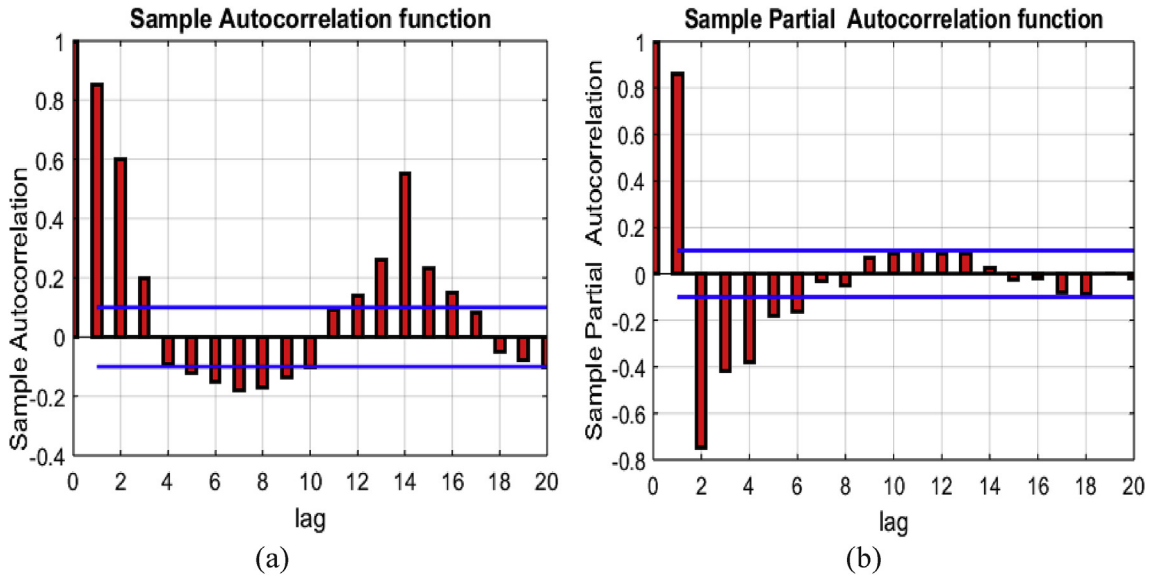


Fig. 6. (a) ACF and (b) PACF of power output.

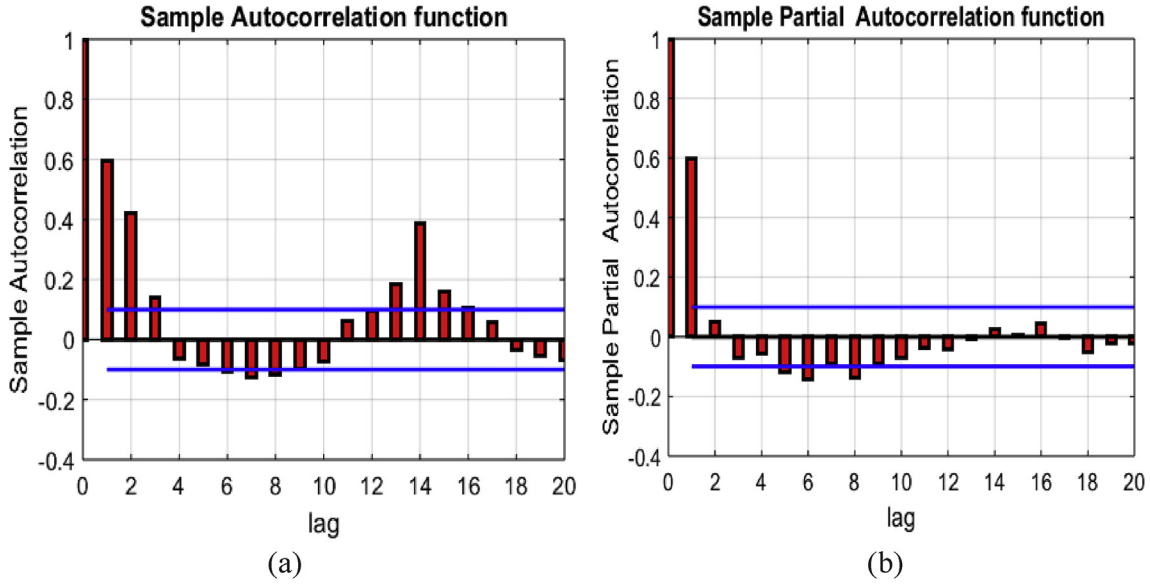


Fig. 7. (a) ACF and (b) PACF of first differenced power output.

ahead rather than one step (see Tables 2 and 4). For time horizons of more than 2 h ahead, the correlation coefficient (R) falls under 90%. Therefore, the NNs are not recommended for time series forecasting of PV system power output for more than two steps ahead (hourly basis) without additional solar irradiance measurements or weather conditions. For the average computational time, the forecasters need less than 1 s more time to complete the two and three hours ahead prediction, compared to one hour ahead forecasting.

3.3. Statistical models results

Sample autocorrelation and partial autocorrelation for produced power and for first differenced power output are plotted in Figs. 6 and 7, respectively. Based on the sample autocorrelation plot, the data has a seasonal pattern. The value of the seasonality index for

Table 6
AIC and BIC values of available ARMA models and model parameters (p, q).

p, q	AIC	BIC
1, 1	5892.8	5899.7
1, 2	5865.5	5872.9
1, 3	5850.1	5857
1, 4	5833.2	5839.6
2, 1	5877.8	5884.9
2, 2	5450.9	5363.9
2, 3	5241.2	5267.4
2, 4	5332.2	5342.2
3, 1	5458.9	5471.2
3, 2	5359.4	5366.4
3, 3	5340.6	5359.2
3, 4	**5111.5**	**5204.8**
4, 1	5400.2	5412.9
4, 2	5342.9	5354.1
4, 3	5325.6	5341.5
4, 4	5335.1	5341.9

p: AR order, q: MA order.

Table 7
Results of PV power prediction using ARMA, ARIMA, and SARIMA, one hour ahead.

	Algorithm	R (test Data)	RMSE (test data)	Average computation time (s)
1	SARIMA	0.929	1.183	84.14
2	ARIMA	0.912	1.318	37.55
3	ARMA	0.904	1.212	49.20

SARIMA model is set at 14 since the data is recorded hourly with a period of 14 h per day. Moreover, we can see from the sample autocorrelation plot that the pattern has a 14 h periodicity. To optimize the models' parameters, different combinations of parameters are investigated and the models with the lowest AIC value are used in the prediction as the most reliable choices. ACF decays after the second lag while PACF decays after the third lag as seen in Fig. 7. Therefore, we should theoretically use AR(2) of the SARIMA model. The possible candidate orders of the moving average of SARIMA model are 1 as shown in Fig. 7 (b). However, AIC values are used to select the most appropriate models. The possible ARIMA and SARIMA models usually have a degree of integration equal to 1 [33]. The Phillips-Perron test was also applied in this study to find h value for the first derivative of power output. The value of h was found to be 1 which indicated that the first difference can be used reliably to make the data stationary [34]. Thus, the degree of integration of SARIMA and ARIMA models in this study is set to 1. A one time use of SARIMA and ARIMA with a degree of integration more than 1 for forecasting was also conducted. Those models require a significant amount of computational time (See Table 7). ARMA (3, 4), ARIMA (2, 1, 3) and SARIM (2, 1, 3), (2, 0, 1)[14] have the lowest ACF values. Thus, they are considered as the best fit models in the current work. AIC and BIC values of ARMA models are given in Table 6. A similar table was generated for ARIMA and SARIMA. The residuals Q-Q plots of the SARIMA model is shown in Fig. 8. The residuals have a linear trend.

The models are listed in Tables 7 and 8 and scored based on the correlation coefficient (R) in descending order using the test set. The SARIMA model scored the highest correlation coefficient (R) and the lowest root mean square error (RMSE) for one and two hours ahead predictions. We conclude that identifying the time series pattern and accounting for the seasonality enhances the forecasting results. The predicted PV power using the SARIMA model with one hour ahead for different day types is plotted in Fig. 9. The result for validation attempt 4 (See section 2.4) is shown in Fig. 10. Note that the predicted value of PV power output using SARIMA is mostly higher than the true value. The opposite has been observed for BI-LSTM. It is concluded that the suggested statistical models consume more computation time and have an accuracy lower than most of the NNs. Artificial neural networks learn the complexity of time series data better than SARIM models. Thus, NNs are better than the suggested statistical models for PV power time series prediction. The average computation time for one hour ahead forecasting is given in Table 7. While two and three hours ahead prediction need about 1 or 2 more seconds for the forecasting to be accomplished, compared to one hour ahead forecasting.

Table 8
Results of PV power prediction using ARMA, ARIMA, and SARIMA, two and three hours ahead.

	Algorithm	R (test data)	RMSE(test data)	Average computation time (s)
Two hours ahead				
1	SARIMA	0.909	1.781	85.14
2	ARIMA	0.875	1.840	39.34
3	ARMA	0.864	1.914	50.75
Three hours ahead				
1	ARMA	0.819	2.212	85.63
2	SARIMA	0.801	2.314	39.54
3	ARIMA	0.742	2.604	51.02

3.4. Removing outliers

Outlier removal is necessary to avoid poor fitting or overfitting especially in case of ill-behaved and very noisy training data. Overfitting occurs when the model falsely tunes to the random fluctuations of the training and generalizes the concept. For outlier detection, Hampel filter is utilized to remove the extreme values. Table 9 shows the effect of removing outliers on the correlation coefficient (R) and RMSE of the BI-LSTM algorithm test results for different forecasting horizons. The enhancement in R and the reduction of RMSE are demonstrated in the same table. As demonstrated, removing the outliers improves the prediction results. The subscripts 1 and 2 refer to before and after removing the outliers, respectively.

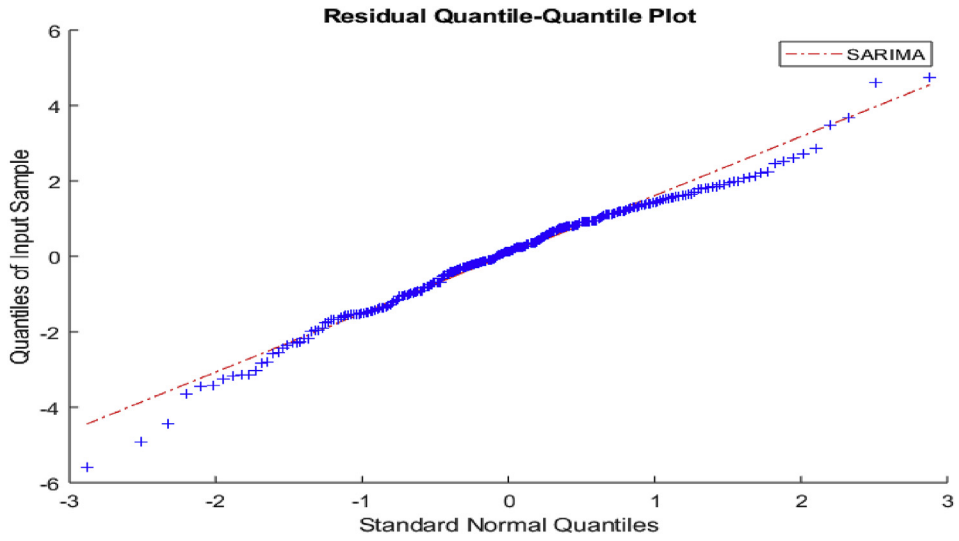
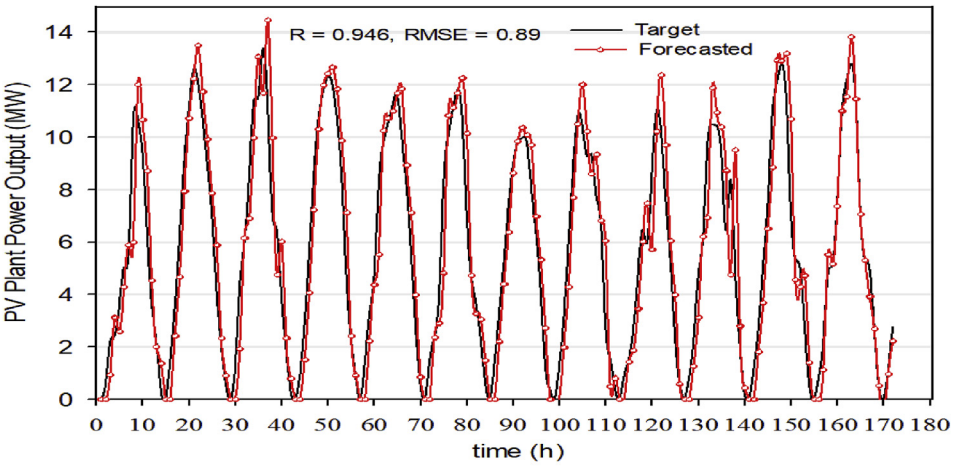
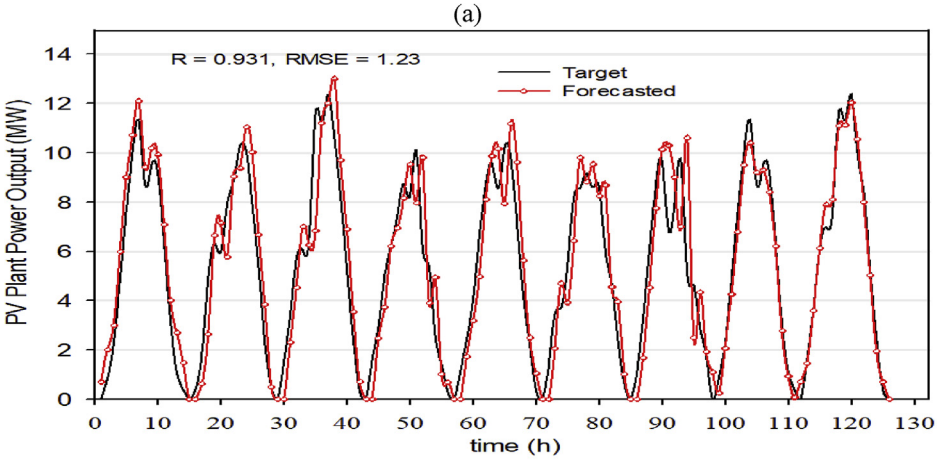


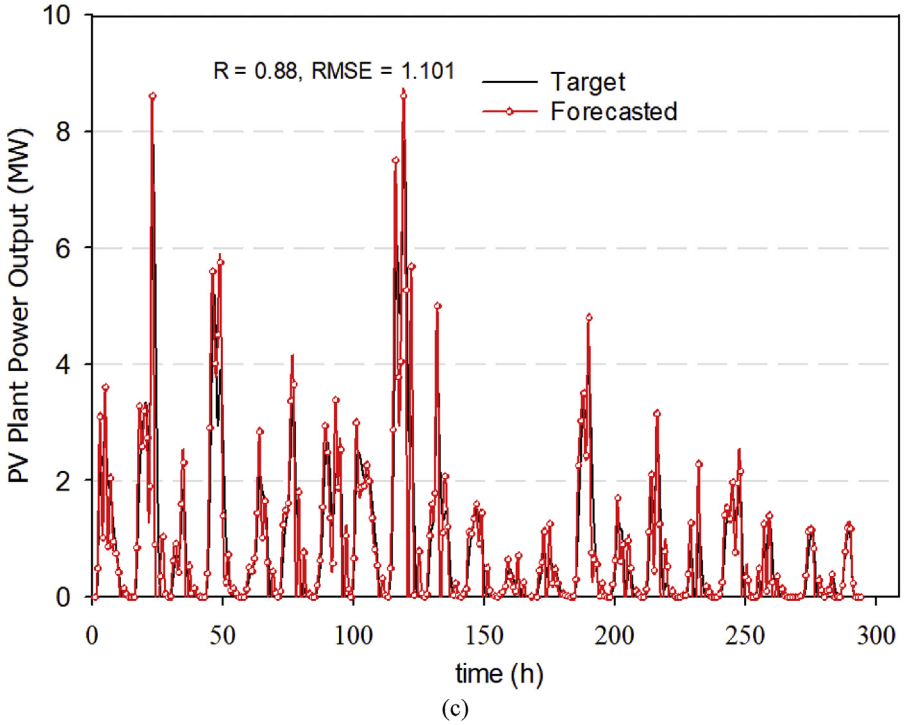
Fig. 8. Quantile-Quantile plot for SARIMA model.



(a)



(b)



(c)

Fig. 9. Results of SARIMA model for PV power prediction, one hour ahead: (a) sunny days (b) cloudy days (c) rainy days.

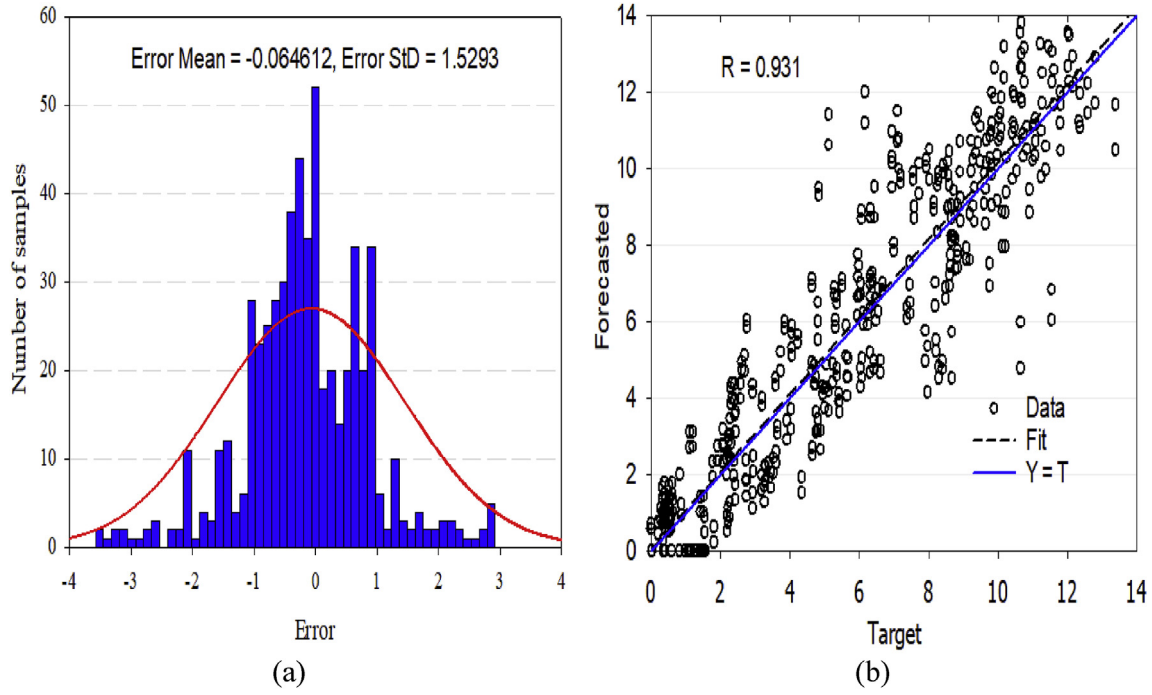


Fig. 10. Results of SARIMA model for PV power prediction, one hour ahead for the validation attempt number 4: (a) Error histogram figure (real - forecast), (b) Regression plot.

Table 9
The effect of removing outliers on BI-LSTM accuracy.

Step	R_1	R_2	$\frac{R2-R1}{R1}\times 100\%$	$MSER_1$	$MSER_2$	$\frac{MSER1-MSER2}{MSER1}\times 100\%$
1	0.961	0.98	1.98%	0.881	0.791	10.2%
2	0.937	0.965	2.99%	1.16	1.092	5.86%
3	0.86	0.882	2.56%	1.95	1.831	6.1%

4. Conclusion

Forecasting solar power is necessary for policy making, understanding the challenges and optimal integration of large-scale photovoltaic plants with the public power grid. In this paper, the performance of different NNs and simple statistical models such as ARMA, ARIMA, and SARIMA was evaluated in the time series forecasting of the power output of largescale PV plants. The comparative study shows that neural networks are more accurate than the suggested statistical models when used for time series prediction of PV power output and require less computation time. However, the NNs as well as the statistical models can be used to efficiently predict the produced power of PV plants for only one hour ahead without having access to solar irradiance measurements or any weather parameters. Thus, the time series forecasting for PV power plants is only reliable for one hour ahead prediction.

Author contribution

Hussein Sharadga: Conceptualization, Methodology, Software, Data curation, Writing - original draft preparation, Formal analysis, Investigation. **Shima Hajimirza:** Supervision, Conceptualization, Writing- Reviewing and Editing, Investigation. **Robert Balog:** Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

The authors would like to thank Dr. Shi and Dr. Lee of the University of Jinan, China and the University of Texas at Arlington for providing us with the data.

References

[1] S.K.H. Chow, E.W.M. Lee, D.H.W. Li, Short-term prediction of photovoltaic energy generation by intelligent approach, Energy Build. 55 (2012) 660–667, https://doi.org/10.1016/j.enbuild.2012.08.011.
[2] M.A. Al-Nimr, S. Kiwan, H. Sharadga, Simulation of a novel hybrid solar photovoltaic/wind system to maintain the cell surface temperature and to generate electricity, Int. J. Energy Res. 42 (2018) 985–998, https://doi.org/10.1002/er.3885.
[3] J. Shi, W.J. Lee, Y. Liu, Y. Yang, P. Wang, Forecasting power output of photovoltaic systems based on weather classification and support vector machines, IEEE Trans. Ind. Appl. 48 (2012) 1064–1069, https://doi.org/10.1109/TIA.2012.2190816.
[4] H. Jiang, L. Hong, Application of BP neural network to short-term-ahead generating power forecasting for PV system, Adv. Mater. Res. 608–609 (2013) 128–131. https://doi.org/10.4028/www.scientific.net/AMR.608-609.128.
[5] A. Mellit, S. Sağlam, S.A. Kalogirou, Artificial neural network-based model for estimating the produced power of a photovoltaic module, Renew. Energy 60 (2013) 71–78, https://doi.org/10.1016/j.renene.2013.04.011.
[6] C. Chen, S. Duan, T. Cai, B. Liu, Online 24-h solar power forecasting based on weather type classification using artificial neural network, Sol. Energy 85 (2011) 2856–2870, https://doi.org/10.1016/j.solener.2011.08.027.
[7] A. Mellit, A. Massi Pavan, V. Lughi, Short-term forecasting of power production

in a large-scale photovoltaic plant, Sol. Energy 105 (2014) 401–413, https://doi.org/10.1016/j.solener.2014.03.018.
[8] H. Kalgude, V. Jadhav, S. Joshi, M. Joshi, G. Bharambe, S. Walunj, Forecasting the output power of solar panel using Lstm-Rnn, VJER-Vishwakarma, J. Eng. Res. 2 (2018) 113–121.
[9] A. Gensler, J. Henze, B. Sick, N. Raabe, Deep learning for solar power forecasting – an approach using autoencoder and LSTM neural networks, in: IEEE Int. Conf. Syst. Man Cybern., 2016, pp. 2858–2865, https://doi.org/10.1109/SMC.2016.7844673.
[10] P. Mandal, S.T.S. Madhira, A. Ul haque, J. Meng, R.L. Pineda, Forecasting power output of solar photovoltaic system using wavelet transform and artificial intelligence techniques, in: Procedia Comput. Sci., 2012, pp. 332–337, https://doi.org/10.1016/j.procs.2012.09.080.
[11] N. Al-Messabi, Y. Li, I. El-Amin, C. Goh, Forecasting of photovoltaic power yield using dynamic neural networks, Proc. Int. Jt. Conf. Neural Networks (2012) 1–5, https://doi.org/10.1109/IJCNN.2012.6252406.
[12] M. Bouzerdoum, A. Mellit, A. Massi Pavan, A hybrid model (SARIMA-SVM) for short-term power forecasting of a small-scale grid-connected photovoltaic plant, Sol. Energy 98 (2013) 226–235, https://doi.org/10.1016/j.solener.2013.10.002.
[13] J.G. da S. Fonseca Jr., T. Oozeki, T. Takashima, G. Koshimizu, Y. Uchida, K. Ogimoto, Use of support vector regression and numerically predicted cloudiness to forecast power output of a photovoltaic power plant in Kita-kyushu, Japan, Prog. Photovolt. Res. Appl. 20 (2011) 874–882, https://doi.org/10.1002/pip.1152.
[14] L.A. Fernandez-Jimenez, A. Muñoz-Jimenez, A. Falces, M. Mendoza-Villena, E. Garcia-Garrido, P.M. Lara-Santillan, E. Zorzano-Alba, P.J. Zorzano-Santamaria, Short-term power forecasting system for photovoltaic plants, Renew. Energy 44 (2012) 311–317, https://doi.org/10.1016/j.renene.2012.01.108.
[15] H.-T. Yang, C.-M. Huang, Y.-C. Huang, Y.-S. Pai, A weather-based hybrid method for 1-day ahead hourly forecasting of PV power output, IEEE Trans. Sustain. Energy. 5 (2014) 917–926, https://doi.org/10.1109/TSTE.2014.2313600.
[16] T. Cai, S. Duan, C. Chen, Forecasting power output for grid-connected photovoltaic power system without using solar radiation measurement, in: 2nd Int. Symp. Power Electron. Distrib. Gener. Syst. PEDG 2010, IEEE, 2010, pp. 773–777, https://doi.org/10.1109/PEDG.2010.5545754.
[17] J. Liu, W. Fang, X. Zhang, C. Yang, An improved photovoltaic power forecasting model with the assistance of aerosol index data, IEEE Trans. Sustain. Energy. 6 (2015) 434–442, https://doi.org/10.1109/TSTE.2014.2381224.
[18] A. Mellit, Recurrent neural network-based forecasting of the daily electricity generation of a photovoltaic power system, Ecol. Veh. Renew. Energy. (2009) 26–29.
[19] I. Ashraf, A. Chandra, Artificial neural network based models for forecasting electricity generation of grid connected solar PV power plant, Int. J. Glob. Energy. 21 (2004) 119–130.
[20] A. Mellit, A.M. Pavan, Performance prediction of 20 kW p grid-connected photovoltaic plant at Trieste (Italy) using artificial neural network, Energy Convers. Manag. 51 (2010) 2431–2441, https://doi.org/10.1016/j.enconman.2010.05.007.
[21] A.E. Saleh, M.S. Moustafa, K.M. Abo-Al-Ez, A.A. Abdullah, A hybrid neuro-fuzzy power prediction system for wind energy generation, Electr. Power Energy Syst. 74 (2016) 384–395, https://doi.org/10.1016/j.ijepes.2015.07.039.
[22] K.I. Wong, C.M. Vong, P.K. Wong, J. Luo, Sparse Bayesian extreme learning machine and its application to biofuel engine performance prediction, Neurocomputing 149 (2015) 397–404, https://doi.org/10.1016/j.neucom.2013.09.074.
[23] A. Kavousi-fard, A hybrid accurate model for tidal current prediction, IEEE Trans. Geosci. Remote Sens. 55 (2017) 112–118, https://doi.org/10.1109/TGRS.2016.2596320.
[24] A.T. Hammid, M.H. Bin Sulaiman, A.N. Abdalla, Prediction of small hydropower plant powerproduction in Himreen Lake dam (HLD) using artificial neural network, Alexandria Eng. J. 57 (2018) 211–221, https://doi.org/10.1016/j.aej.2016.12.011.
[25] S. Hochreiter, J. Schmidhuber, Long short-term memory, Neural Comput. 9 (1997) 1735–1780, https://doi.org/10.1162/neco.1997.9.8.1735.
[26] M. Abdel-Nasser, K. Mahmoud, Accurate photovoltaic power forecasting models using deep LSTM-RNN, Neural Comput. Appl. 31 (2019) 2727–2740, https://doi.org/10.1007/s00521-017-3225-z.
[27] A. Graves, N. Jaitly, A.R. Mohamed, Hybrid speech recognition with deep bidirectional LSTM, in: 2013 IEEE Work. Autom. Speech Recognit. Understanding, ASRU 2013 - Proc, IEEE, 2013, pp. 273–278, https://doi.org/10.1109/ASRU.2013.6707742.
[28] J. Flaherty, R. Lombardo, Modelling private new housing starts in Australia, in: A Pap. Present. Pacific-Rim Real Estate Soc. Conf., 2000, pp. 24–27. http://www.prres.net/papers/Flaherty_Modelling_Private_New_Housing_Starts_In_Australia.pdf.
[29] G. Box, G. Jenkins, G. Reinsel, G. Ljung, Time Series Analysis, Forecasting and Control, John Wiley & Sons, 2015.
[30] X. Zhang, Y. Liu, M. Yang, T. Zhang, A.A. Young, X. Li, Comparative study of four time series methods in forecasting typhoid fever incidence in China, PLoS One 8 (2013).
[31] P. Chen, A. Niu, D. Liu, W. Jiang, B. Ma, Time series forecasting of temperatures using SARIMA: an example from Nanjing, IOP Conf. Ser. Mater. Sci. Eng. 394 (2018), https://doi.org/10.1088/1757-899X/394/5/052024.
[32] H. Akaike, A Bayesian analysis of the minimum AIC procedure, Sel. Pap. Hirotugu Akaike. (1998) 275–280.
[33] A.K. Mishra, V.R. Desai, Drought forecasting using stochastic models, Stoch. Environ. Res. Risk Assess. 19 (2005) 326–339, https://doi.org/10.1007/s00477-005-0238-4.
[34] M.H. Alsharif, M.K. Younes, J. Kim, Time series ARIMA model for prediction of daily and monthly Average Global Solar Radiation: the case study of Seoul, South Korea, Symmetry (Basel) vol. 11 (2019), https://doi.org/10.3390/sym11020240.