Comparison of Photovoltaic Production Forecasting Methods

Mohamed Hamza Kermia *‡ (iD) , Jerome Bosche * (iD) , Dhaker Abbes ** (iD)

*Modeling, Information, and Systems Laboratory, University of Picardie Jules Verne, UFR of Sciences, 33 Rue St Leu Amiens 80000, France

**Univ. Lille, Arts et Metiers Institute of Technology, Centrale Lille, Junia, ULR 2697 – L2EP, F-59000 Lille, France

(mohamed.kermia@u-picardie.fr, jerome.bosche@u-picardie.fr, dhaker.abbes@junia.com)

‡
Corresponding Author; Mohamed Hamza KERMIA, 14 Quai de la Somme, 80080 Amiens, Tel: +33 758 394 237, mohamed.kermia@u-picardie.fr

Abstract- A new short-term photovoltaic (PV) power forecasting technique based on a polynomial model is proposed in this paper. This technique has been compared with two forecasting methods. The first method is based on deep learning and uses a recurrent neural network (RNN) to extract features from a two-dimensional matrix of PV generation data. The second method employs the Steadysun solution, which was developed by a French company and gives forecasts for up to 30 minutes. The prediction is based on data from the University of Lille "RIZOMM" plant. The main objective of this study is to show the limits of each method and to validate the proposed technique.

To select the best method, three-time levels were considered (10 min, 30 min, and 60 min). The results showed that the RNN has very high accuracy over all horizons, in particular for a 60 minutes time horizon with 6-step ahead where the forecasting accuracy can reach 97 %.

1. Introduction

The development of renewable energies has made exceptional progress due to several factors: climate change [1], big cities, and air pollution [2]. With the progressive reduction of production costs for wind power and PV energy [3], the deployment of power plants based on renewable sources is now faster and cheaper than other power generation technologies [4]. PV is the fastest growing renewable technology in the world, with the highest investment [4]. PV power is a viable supplement to the depleting fossil-fuel-based grid [47]. PV production is growing at a high rate, necessitating production forecasting. PV production, on the other hand, is very variable, owing mostly to stochastic cloud development in the sky. Accurate meteorological data and field production measurements are required to anticipate PV production and optimize power plant performance [5]. This optimization has numerous advantages, particularly for commercial-industrial (C&I) and residential installations, because accurate forecasting maximizes self-consumption and thus reduces the cost of energy produced [6, 7]. Similarly, power forecasting may be used to improve electric car charging in an energy management system (EMS) [8]. However, when it comes to larger installations, PV plant managers use forecasting to optimize plant downtime for maintenance [9]. In addition, in countries with a day-ahead electricity market, forecasting models can optimize the timing of sales, minimizing penalties and revenue losses [10,11].

Moreover, credible forecasting are also required by Transmission System Operators (TSOs) and Distribution System Operators (DSOs), to manage the unpredictability as well as the volatility of grid-connected distributed PV generators. DSOs and TSOs can manage the intermittent production of PV plants, minimize difficulties with balancing power output and load demand [12, 18], increase system stability, and decrease ancillary service costs [13, 14] with reliable forecasts.

In addition, forecasting increases dependability and lowers costs by allowing for more efficient solar energy trading [15]. Also, it reduces the number of backup units and the operating costs of the power plant [16].

In recent years, several forecasting techniques have been developed in the literature. The authors in [45] performed a comparative study of solar irradiance and solar energy forecasting approaches, using the input data, forecast results, forecast model, forecast interval, and forecast accuracy. Different solar power prediction methods were investigated in [47] based on the input data, the number of training data datasets, recording intervals and testing intervals used. Although there are no commonly accepted categorization criteria [12, 14], the following classification is common: Long term (one to ten years), medium-term (one month to one year), short term (one hour to one week), and ultra-short-term (a few minutes) [14]. Each forecasting horizon is adapted to a particular objective [8,10]. For example, very short-term models can be used for power smoothing and dispatch in real-time. Thus, short-term forecasts are used for automated control of production, unit dispatch, energy management, and load balancing, long-term forecasts are used by utilities for unit commitment, load balancing, and scheduling. In addition, TSOs and DSOs design their infrastructure using short- and medium-term predictions as well as long-term horizons [14]. A different categorization depends on the prediction method employed. This categorization divides prediction approaches into three groups: statistical methods, physical methods, and hybrid methods.

Statistical approaches use a series of measurements at different times for one or more variables. For example, we can cite (RNN) with.Long.Short-Term.Memory (LSTM).[19, 20], Support Vector Machine (SVM) [21], Naïve Bayes [46], Markov chain, Fourier Series (FS), regression method, polynomial approach [22]. These approaches depend only on previously collected data without any knowledge of the PV plant or the location of the PV plant. The most frequently used models are ANN models and regression models. The ANN-based prediction has been shown in several studies to be one of the most effective methods. ANNs can accomplish this feat because of their ability to recognize rapid changes in the input-output link due to changing environmental circumstances [14]. ANNs require a big quantity of data to train; employing a random dataset at the start may diminish the dependability of the findings. The model architecture (number of inputs, hidden layers, neurons, etc.) that is chosen can have an impact on the outcomes [12]. Auto-regressive.moving.average. (ARMA) forecasts perform well when the data is stationary, but auto-regressive.integrated.moving. average. (ARIMA) models perform better when the data is not stationary [23]. The limitation of ARIMA techniques comes from the necessity of more computing power than ARMA. According to [14], in terms of comparison, ANNs represent greater accuracy and flexibility under unpredictable weather circumstances than ARMAs and ARIMAs. In [48], the authors have developed multilayer perceptron models based on the gray wolf, lion, and whale optimization techniques for daily solar power prediction. On the other hand, a prediction based on statistical methods gives better results when the metrological conditions are classified according to day type (sunny day, cloudy day) [24,25,26]. The statistical technique is dominant when it comes to short-term and ultra-short-term forecasts [12, 20, 27, 28].

Physical approaches are composed of mathematical equations that explain the physical state and dynamic motion of the environment [29]. They are mainly used in very short-term or long-term applications. These methods are mainly based on numerical weather prediction (NWP), cloud photography, and satellite photos [30]. They are classified into global or mesoscale physical approaches depending on the size of the simulated atmosphere, which may be global or cover only a limited region [29]. Since mesoscale models have a resolution of 16-50 km, it is advisable to use only mesoscale models to anticipate the power output of solar systems [31]. Furthermore, physicals methods have a lower accuracy when weather conditions are unstable [32] and perform better when weather conditions are stable [14].

Combining two or more of the preceding methods is known as a hybrid method. To improve prediction performance, a variety of models with distinct properties are combined [32, 33, 34]. There is a general rise in computational complexity [12]. For example, the most typical examples combine ANN-based models [9, 35, 36, 37] with Seasonal Autoregressive Integrated Moving Average (SARIMA) [38]. Due to their interdependence with single-model performances, these hybrid models should be created for a given plant and region [29]. Excellent predictions have been obtained by using a combination of numerical weather predictions and historical data on weather variables. However, hybrid prediction methods have been a weakness: they perform less well when the weather is unstable [39]. LSTM-based models have been tested with other neural network PV power output techniques in [40]. These models can represent the underlying correlations between meteorological conditions and real PV energy generation on a daily, hourly, and seasonal basis.

Contrasting the performance of weather prediction methods, in general, can be difficult because of the many different aspects affecting their effectiveness. These include factors such as historical weather information availability, weather prediction accuracy, temporal horizons, resolution as well as geographic location, and installation conditions. Preprocessing data (such as eliminating the night sample when no electricity is produced) is also crucial for statistical methods in order to obtain high performance and decreased computing costs [14]. The previous results in the literature provide some insight into the effectiveness and efficiency of the various methods, but their conclusions are more qualitative than quantitative. As a result of recent reviews [12, 14, 29] that compare the work of various authors, statistical mistakes are also included in the analysis. This is not a quantitative comparison because the conditions and measurements used in each study were different.

The objective of this research is to examine three of the most successful and extensively used forecasting methodologies for PV power output. These three approaches are based respectively on RNN-LSTM, polynomial approach, and sky camera. Measurements and data from a PV installation (the RIZOMM - HEI) in Lille, France, were used to compare the different approaches. To validate the obtained results, three prediction horizons were examined.

The main objective of this study is to develop and test the accuracy of three different forecasting algorithms from three different types of families. The tests will be performed over various forecast horizons.

In contrast to the previous works in the literature, the proposed study is based on:
- A new polynomial prediction algorithm.
- Development of an RNN-LSTM model.

➢ A comparison with a commercial prediction method (Steadysun).

➢ A variety of time horizons (ten minutes, thirty minutes, and sixty minutes).

Exogenous inputs (such as air temperature, wind speed, cloud cover, etc) are not considered in our study since we only look at historical power data from the RIZOOM plant. Moreover, most of the presented models in the literature are primarily focused on hourly or daily one-step prediction, and to the best of our knowledge, very short-term prediction (a few minutes ahead) is not well considered, even though it plays a critical role in PV plant control applications and electric vehicle charging stations, which encouraged as to deal with such type of problem.

The article is structured as follows: Section 2 presents the plant and data. The prediction methods are studied in section 3. Section 4 includes a comparison of the three methods: polynomial, RNN-LSTM, and Steadysun with the results. While section 5 is dedicated to conducting and perspective elements.

2. Description of the Smart-grid Demonstrator

The demonstrator is located in the north of France, at the Catholic University of Lille. It has two rooftop solar generators of 189 kWp and 28 kWp, a 250 kWh Eaton Li-ion battery, numerous electric vehicle charging stations with a charging capacity of 22 kW, and four service buildings. This network is further connected to the distribution grid via a 15kV/0.4kV 1 MVA transformer (Figure 1).
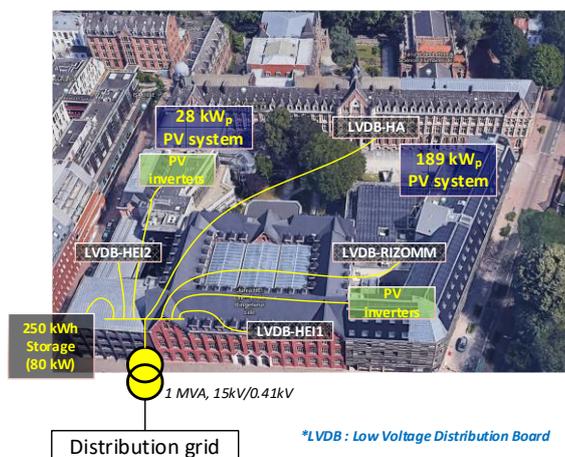
Fig. 1. Demonstrator of the Catholic University of Lille

The JUNIA engineering school owns two buildings, HEI1 and HEI2, while the Catholic University owns HA and RIZOMM. As the buildings are owned by two separate legal companies, French legislation prohibits the exchange of energy between the two. In our work, we considered data from the 189 kWp PV system of the RIZOMM building. These data are from the PV production registered at 10-minutes time intervals between September 2018 and April 2021. Figure 2a represents the PV production throughout a week in March 2021. Figure 2b shows the distribution function of the PV power.
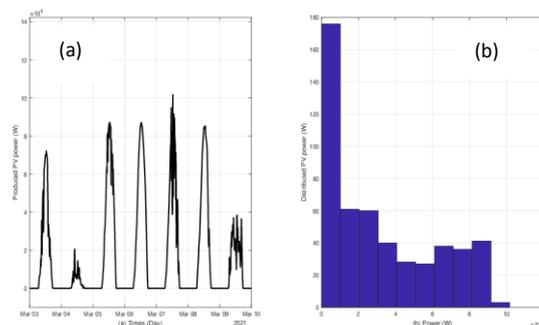
Fig 2. a) PV power data (3-9 March 2021), b) Distribution of produced PV power data (3-9 March 2021).

3. Prediction Techniques

The three forecasting approaches used on the RIZOMM data are described in detail in this section.

1.3. Steadysun

Steadysun has created a short-term solar prediction solution called Steadysun (Fig 3a). Steadysun has a camera oriented to the sky at 180° and can anticipate cloud movement over a 4 km² area with energy production predictions for up to 60 minutes using image processing techniques.

The camera takes hemispherical images-sky (Fig 3b) at regular intervals, and cloud and shadow maps are constructed by the cloud detector and motion field sensors using the method developed by Steadysun.

This data is combined with irradiance sensor data and analyzed by the Steadysun algorithm to produce PV power predictions. Local or on-site analysis and prediction are performed, and the tool may be used standalone or with an internet connection.

The Steadysun is connected to the industrial control and data acquisition system (SCADA). The SCADA operates in real-time, using a wide area network configured for an area with an Internet connection, at a download speed of 2 (Mbps). A location without a persistent internet connection will gather data from the Steadysun server and send it to the client SCADA using an interposed communication system. An X-series Super Wide (SW) lens camera, a pyranometer, a silicon cell reference, and a Steadysun data processor are among the components used.

(a)

(b)
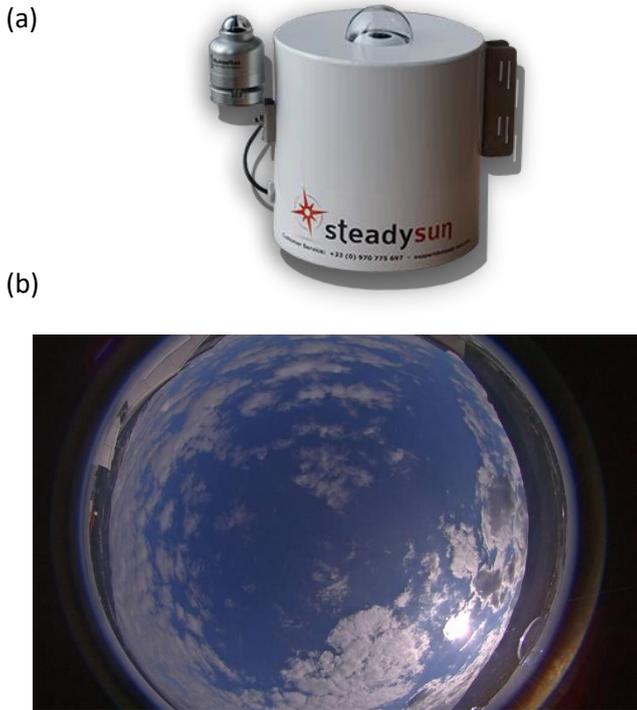
Fig. 3. a) Imager steadysun b) Steadysun sky imager.

For the data acquisition to SCADA, an IP or Modbus connection is used. Table 1 gives the details:

Table 1. Steadysun features

Description	Capability
Prediction aera	4 km²
Temporal resolution	10 s
Spatial resolution	Local area coverage
Prediction horizon	60 min
Components	Camera SW lens, silicon cell reference, pyranometer, data processor

Figure 4 depicts a block diagram for prediction generation:
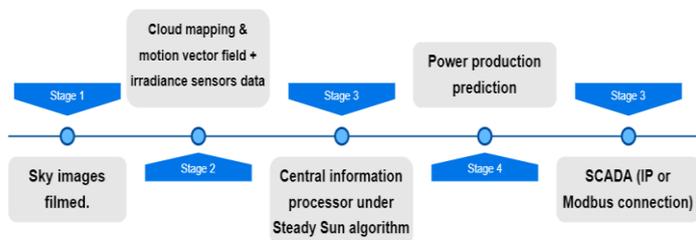
Fig. 4. Steadysun operation block diagram.

2.3. RNN-LSTM

Figure 5 shows the RNN model's structure, developed in [44]. At step t, x^t and y^t are the RNN's input and output variables. The RNN model's hidden state s^t is calculated using the input x^t at the current step t as well as the hidden state s^{t-1} at the step $t-1$. RNN's mathematical model is expressed as follows:

$$s^t = f((Ux^t + b) + Ws^{t-1}) \tag{1}$$

$$o^t = Vs^t + c \tag{2}$$

$$y^t = g(o^t) \tag{3}$$

The weight matrix between the input and hidden layers is denoted by $U \in R^{l_x \times l_s}$. The weight matrix between two hidden layers is $W \in R^{l_s \times l_s}$. The weight matrix linking the hidden and output layers is $V \in R^{l_o \times l_s}$. It should be observed that the parameter values of the weight matrices U, W and V are not modified in the different phases depicted in Fig 5. The values l_x, l_s and l_o represent the number of neurons in the input, hidden, and output layers, respectively. s^t is the hidden layer state at step t, and it serves as the RNN memory. Bias vectors are represented by the parameters b and c. o^t is a temporary variable that is only decided by the RNN model's hidden state s^t.

The hidden layer and output layer activation functions are $f = tanh$ and $g = sigmoid$, respectively.
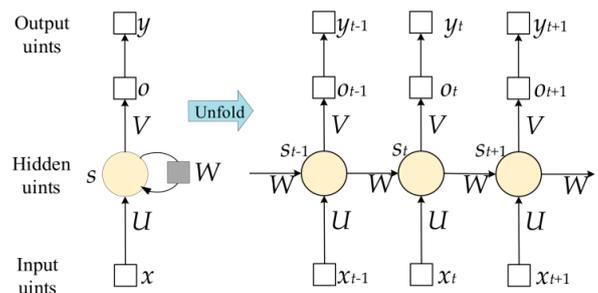
Fig. 5. Basic RNN structure.

The Back Propagation Through Time (BPTT) technique is used in this study to determine the gradient of the parameters U, V, W, b and c of the RNN model [41]. BPTT is an RNN back propagation training technique that is applicable to sequence data such as time series. The BPTT method works by unrolling all of the input time steps and computing and accumulating the sample errors of RNN at each time step. RNN's cost function may be configured as follows:

$$L = \sum_{t=1}^{\tau} \left(\frac{1}{2} \sum_{j=1}^{l_o} (y^t - y_j^t)^2 \right) \tag{4}$$

where L is the total cost of all-time sequences. According to equation (4), the total cost is just the sum of the sub-costs at each time step. The observed and anticipated values are denoted by $\hat{y}_j^t$ and y_j^t respectively. Step t is hidden state gradient is defined as:

$$\delta^t = \frac{\partial L}{\partial s^t} \quad (5)$$

according to the RNN model, δ^t is calculated by the sub-cost at the actual step t and the sub-cost at step $t+1$. As a result, t is associated with the output temporary variable δ^t and the hidden layer state s^{t+1}.

$$\delta^t = \frac{\partial L}{\partial o^t}\frac{\partial o^t}{\partial s^t} + \frac{\partial L^{t+1}}{\partial s^{t+1}}\frac{\partial s^{t+1}}{\partial s^t} = V^T(\hat{y}^t - y^t)g'(o^t) + W^T\delta^{t+1}diag(1-(s^{t+1})^2) \quad (6)$$

Where $diag()$ denotes the creation of a diagonal matrix from a specified vector. Because there are no more hidden states after the last step τ, the δ^τ is written as:

$$\delta^\tau = \frac{\partial L}{\partial o^\tau}\frac{\partial o^\tau}{\partial s^\tau} = V^T(\hat{y}^t - y^t)g'(o^t) \quad (7)$$

Back propagation is used to compute the gradient of the network parameters at step t, step by step. The gradient of U, V, W, b and c is illustrated by the following:

$$\frac{\partial L}{\partial c} = \sum_{t=1}^{\tau}\frac{\partial L^t}{\partial o^t}\frac{\partial o^t}{\partial c} = \sum_{t=1}^{\tau}(\hat{y}^t - y^t)g'(o^t) \quad (8)$$

$$\frac{\partial L}{\partial V} = \sum_{t=1}^{\tau}\frac{\partial L^t}{\partial o^t}\frac{\partial o^t}{\partial V} = \sum_{t=1}^{\tau}(\hat{y}^t - y^t)g'(o^t)(s^t)^T \quad (9)$$

$$\frac{\partial L}{\partial b} = \sum_{t=1}^{\tau}\frac{\partial L}{\partial s^t}\frac{\partial s^t}{\partial b} = \sum_{t=1}^{\tau}diag(1-(s^t)^2)\,\delta^\tau \quad (10)$$

$$\frac{\partial L}{\partial W} = \sum_{t=1}^{\tau}\frac{\partial L}{\partial s^t}\frac{\partial s^t}{\partial W} = \sum_{t=1}^{\tau}diag(1-(s^t)^2)\,\delta^\tau(s^{t-1})^T \quad (11)$$

$$\frac{\partial L}{\partial U} = \sum_{t=1}^{\tau}\frac{\partial L}{\partial s^t}\frac{\partial s^t}{\partial U} = \sum_{t=1}^{\tau}diag(1-(s^t)^2)\,\delta^\tau(x^t)^T \quad (12)$$

The final gradient of the network parameters is, the total of the subgradients at each time-step. Equations (2)-(4), (9), and (10) make the gradients of the network parameters. Therefore, the revised rule for these parameters is as follows:

$$b^{n+1} = b^n - \eta\frac{\partial L}{\partial b} \quad (13)$$

$$c^{n+1} = c^n - \eta\frac{\partial L}{\partial c} \quad (14)$$

$$V^{n+1} = V^n - \eta\frac{\partial L}{\partial V} \quad (15)$$

$$W^{n+1} = W^n - \eta\frac{\partial L}{\partial W} \quad (16)$$

$$U^{n+1} = U^n - \eta\frac{\partial L}{\partial U} \quad (17)$$

Where η is the RNN learning rate and the superscript n represents the BPTT iteration periods. Equations (14)–(16) may be used to calculate the partial derivatives of the cost function with respect to the disturbance of b, c, V, W and U.

3.3. Polynomial method

The polynomial method is a forecasting approach based on polynomial modeling. Development, implementation and use of this approach is done in three steps:

3.3.1. Data preparation

For this stage, we used the RIZOMM plant's PV production database, which was presented in the second section. Each day of production was represented by a vector denoted $\mathbf{y}^*$ with a length of 144 points, corresponding to 144 periods of 10-minutes per day. Figure 6 shows the dataset in a matrix form. Each vector $\mathbf{y}^*$ is decomposed into $\mathbf{y_{i,j}}$ sub-vectors of length T, where i is the package number and j is the day. Since there is no PV power production in the early morning and late evening, the data for these two periods have been suppressed.
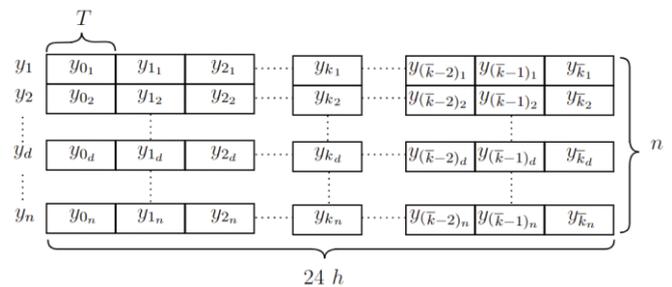
Fig. 6. Vector representation: n days and $\bar{k}$ samples.

3.1.2. A learning process

This part consists in creating a polynomial database from the PV power production data. The time period T, withe $T = 10\alpha$, where $(\alpha = 1,3\ or\ 6)$ is fixed according to the prediction horizon. The transition function between $y_{i,j}$ (the past) and $y_{i+\alpha,j}$ (the future) is calculated each time. The transition function is a polynomial function of order which allows to make the relation between the past and the future. This method has been applied on the whole database, all the transition functions have been stocked in the polynomial database.

3.3.3. Prediction PV Power

Figure 7 shows an example of PV power data. The data to be predicted is represented in green by the interval

$[t + 1, t + h]$ corresponding to the future period. Data from the past is represented in orange in the interval $[t - h, t]$. A prediction at a horizon h requires two steps.

A. Stage 1: To forecast from interval $[t + h]$ on a horizon $h = 1\ hour$, select the last package in the interval $[t - h, t]$ and noted $y_{I,j}$.

B. Stage 2: After the selection of the package, it is question to identify the package y_{I,J^*}, the most similar to $y_{I,j}$, in the sense of the least squares error, in the database. It is necessary to retrieve the transition function f^* from the polynomial database, which creates the connection between y_{I,J^*} and $y_{(I+T,J^*)}$; y_{I+T,J^*+T} with period T. In order to make a prediction, the chosen pass function will be applied to the past data.

The polynomial approach is influenced by several parameters, including the prediction horizon (h), the selection method, and the order of the transition function.
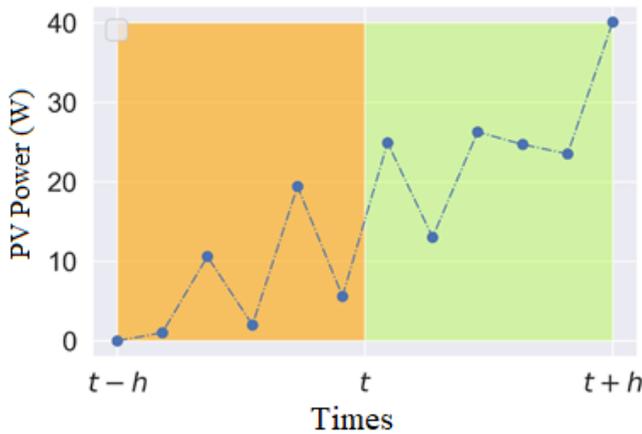
Fig. 7. The Sliding Window Modeling process.

4. Results and Discussion

In this comparison, three prediction approaches are considered, a RNN approach, a polynomial approach and an industrial Steadysun approach. A database containing PV power generation with 10-minutes samples from the RIZOMM power plant is used on all three methods. The data were divided into three parts: 80% for learning, 10% for validation, the remaining 10% for testing.

Three tests were performed, with different prediction horizons. The first test used 10 minutes, the second one 30 minutes and the third test 60 minutes. The used camera is a black box, it only gives predictions of 30 minutes. So, it will be compared to the other methods in the second test. For the RNN-LSTM, we considered the same parameters for all three tests, the number of inputs, hidden and output units (128,52,1), the same number of layers (3), output steps and database size.

In order to compare the different prediction methods, several performance indexes were used.

4.1. Performance Index

Several of the most commonly used index of error in research are used to evaluate the accuracy of forecasts. [29,42,43] were considered in this work. The e_i error is a popular error definition for evaluation, and it is defined as:

$$e_i = A_i - P_i \tag{18}$$

Where A_i is the median actual power in 10 minutes and P_i represents the forecast obtained by one of the forecasting algorithms. In order to finely analyze our algorithms, we will consider five different performances indices.

4.1.1. Mean Absolute Error (MAE)

MAE is derived by dividing the total number of absolute errors by the sample size.

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|A_i - P_i| \tag{19}$$

N is the number of steps (10 minutes) considered in the examined period in all of these definitions (i.e. 7 days).

4.1.2. Mean Absolute Percentage Error MAPE(%)

MAPE is a measure of the forecast accuracy of a forecasting method. It gives the accuracy in the form of a percentage defined by a formula:

The MAPE(%) which was normalized to the measured power:

$$MAPE_{\%} = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{A_i - P_i}{A_i}\right| \cdot 100 \tag{20}$$

if $A_i = 0$, MAPE(%) it becomes the following:

$$MAPE_{\%} = \frac{1}{N}\sum_{i=1}^{N}|P_i| \cdot 100 \tag{21}$$

4.1.3. Root.Mean.Square.Error.(RMSE)

The RMSE is defined as the square root of the second sample moment of discrepancies between anticipated and observed values, or the quadratic mean of these differences.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(A_i - P_i)^2}{N}} \tag{22}$$

4.1.4. Normalized.Root.Mean.Square.Error: nRMSE(%)

The nRMSE is used to calculate the normalized mean square error as the absolute value between the forecasted and measured values.

$$nRMSE_{\%} = \frac{RMSE}{\max(A) - \min(A)} \cdot 100 \tag{23}$$

4.1.5. Coefficient of Determination (R²)

R² is the percentage of variance in the dependent variable that can be forecast by the independent variable.

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(A_i - P_i)^2}{\sum_{i=1}^{N}(A_i - \bar{A})^2} \quad (24)$$

Where $\bar{A}$ is the mean value of the N elements of A.

4.2. Test 1: horizon 10 minutes

Two methods have been considered: Polynomial and RNN-LSTM. Referring to table 2, the coefficient of determination R^2 is in the range (0.96 % for RNN and 0.91 % for Polynomial) revealing a good correlation between the measured and forecasted PV power for the two models tested without considering night periods. When compared to the polynomial technique, the RNN-LSTM model produces the best MAE, RMSE, nRMSE and R2 results. However, in terms of MAPE the polynomial method present better results from table 2, we can easy see that. The developed RNN-LSTM model depicts the lowest errors and the smallest variations around the mean value. Moreover, the correlation between the measured and forecasted PV power with RNN-LSTM, depicted in Fig 8b, is $R^2 = 0.96$ % and with the polynomial method, shown in Fig 9b, is $R^2 = 0.91$ %.

Figures 8a and 9a show the PV power produced and predicted for the RNN-LSTM and Polynomial method respectively. It can be seen that for 10-minutes predictions the RNN-LSTM method shows good results.
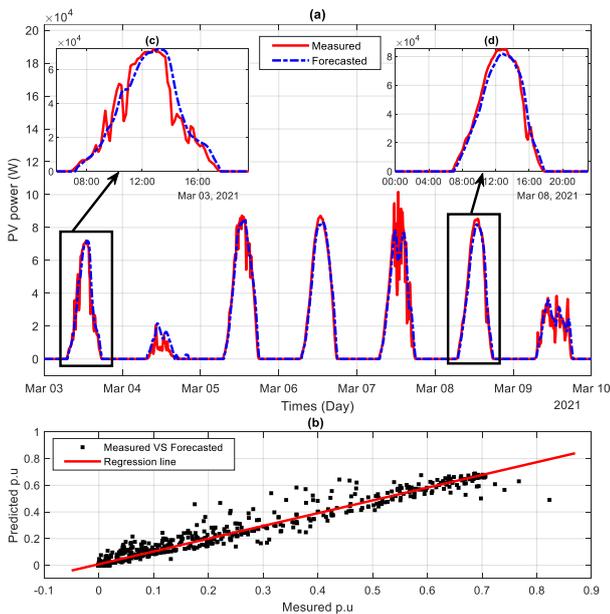
Fig. 8. a) forecasted and measured power, b) R² of measured vs forecasted.
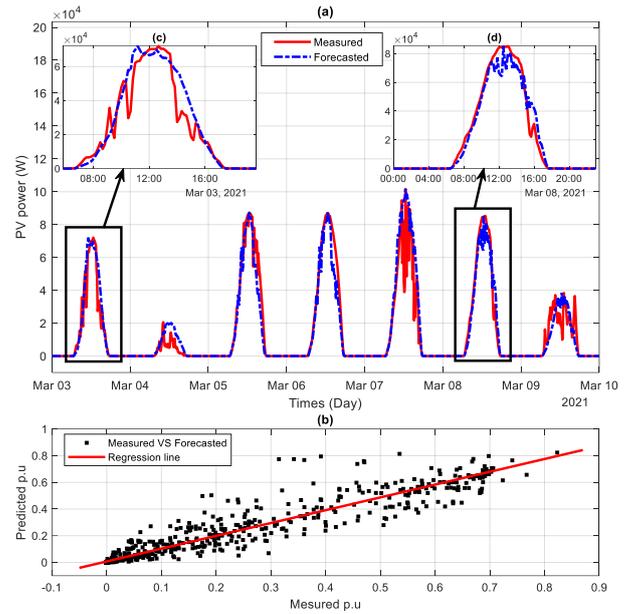
Fig. 9. a) forecasted and measured power for the Polynomial method, b) R² of measured vs forecasted.

4.3. Test 2: horizon 30 minutes

In this part, the three methods were considered. RNN-LSTM shows the best results in terms of MAE, MAPE, RMSE, nRMSE, and R² compared to the other methods. Figures (10,11 and 12) show the results obtained by Steadysun, RNN-LSTM, and the polynomial method respectively.

From the obtained results in table 2 and Figures (10,11 and 12), we can deduce that the polynomial method shows the least satisfactory results.
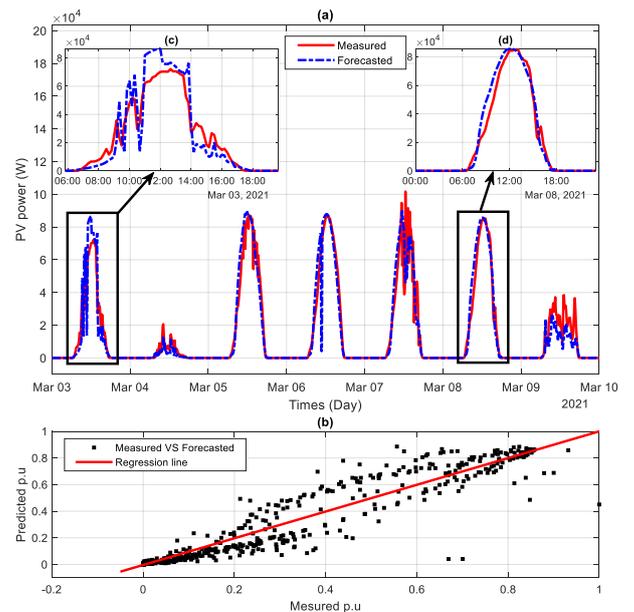
Fig. 10. a) forecasted and measured power for Steadysun, b) R² of measured vs forecasted.
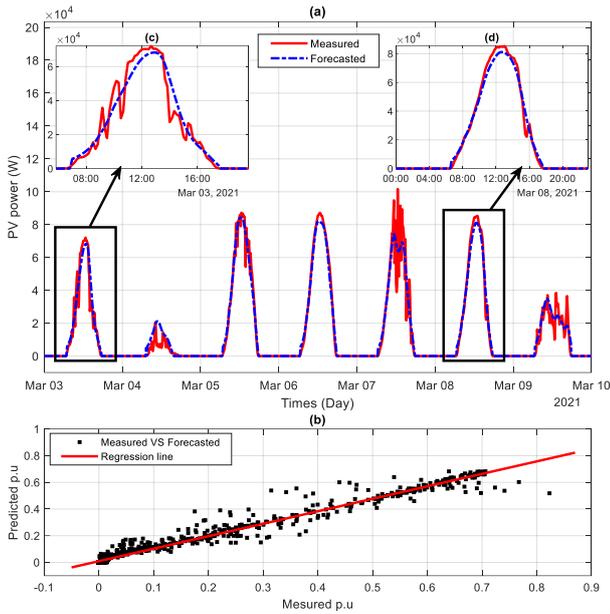
Fig. 11. a) forecasted and measured power, b) R² of measured vs forecasted.
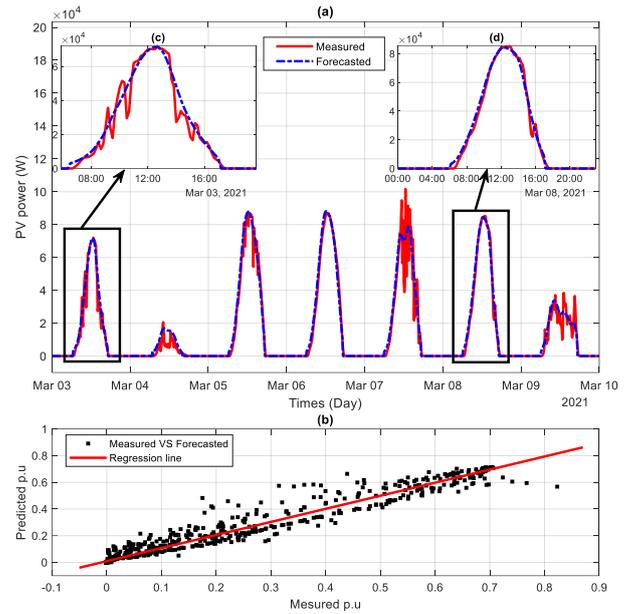
Fig. 13. a) forecasted and measured power, b) R² of measured vs forecasted.
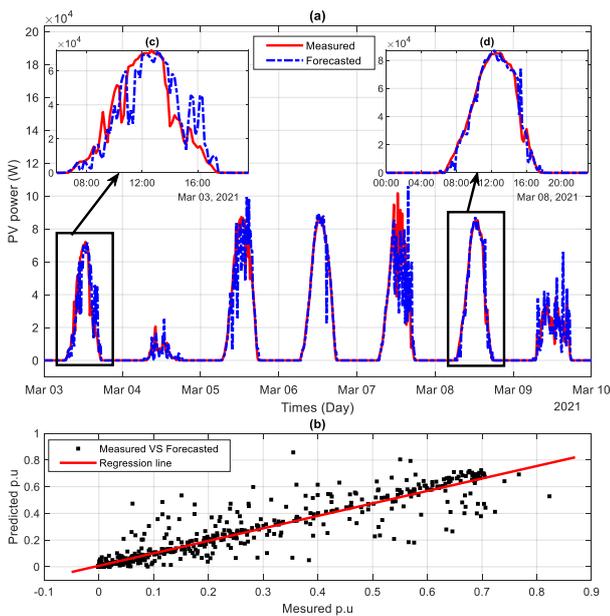
Fig. 12. a) forecasted and measured power for the Polynomial method, b) R² of measured vs forecasted.

Fig. 14. a) forecasted and measured power for the Polynomial method, b) R² of measured vs forecasted.

4.4. Test 3: horizon 60 minutes

For this test, two methods were considered: polynomial and RNN-LSTM. The RNN-LSTM has the best results of MAE, MAPE, RMSE, nRMSE, and R² compared to the polynomial method. Figs (13 and 14) show the results obtained by the RNN-LSTM and Polynomial methods respectively. In this case, for the 60-minutes prediction, the RNN-LSTM method shows good results.
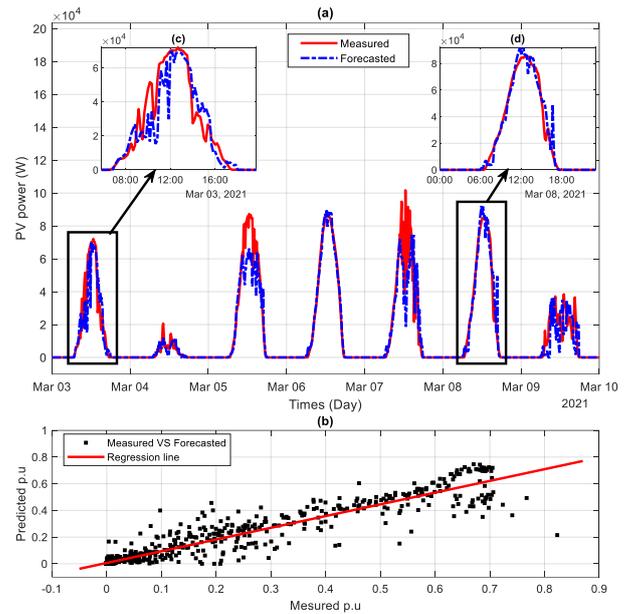
Table 2. Prediction results of the three methods.

	Steadysun	RNN-LSTM			Polynomial		
Horizon de prévision	thirty minutes	ten minutes	thirty minutes	sixty minutes	ten minutes	thirty minutes	sixty minutes
MAE (W)	3542.76	2532.02	2222.33	**2153.64**	3507,34	3378,94	3478,37
MAPE (%)	32.02	63.01	26.39	**26.40**	25,63	27,85	28,44
RMSE (W)	7537.62	5179.64	4694	**4464.05**	7672,05	8603,89	8034,31
nRMSE (%)	7.41	5.09	4.61	**4.38**	7,54	4,46	7,90
R² (pu)	0.91	0.96	0.96	**0.97**	0,91	0,89	0,90

5. Conclusions

Three approaches to forecast PV power output have been evaluated in this paper. Based on different time horizons (10 min, 30 min and 60 min), the outcomes of this study were summarized as follow:

- The RNN-LSTM model has been developed for PV forecasts, excellent results have been obtained for a 60 min time horizon $R^2 = 97\%$.
- Steadysun also gives good results for 30 min forecasts, with: R2 = 91%.
- The proposed polynomial approach is simple to implement and gives very similar results to the other two methods. For a 60-minutes forecast horizon R2 = 90%.

The proposed work has been verified by a comparative study between RNN-LSTMs, Steadysun and Polynomial. Based on the results, it can be concluded that the polynomial algorithm gives acceptable accuracy in cloudy days, but further improvements are needed to ensure effective planning and management of PV plants.

To increase the accuracy of cloudy day predictions, the strategies should incorporate a mix of weather prediction data, sky images, clarity index, etc. On the other hand, the performance of the RNN-LSTMs tested in this study is sufficient for the construction of a smart energy management system for a microgrid including a PV generator, electrical storage, and an electric car charging station. In general, we believe that the basic RNN-LSTM models represent a sufficient.

In the future, the polynomial approach will be extended to another modeling technique using state formalism. The advantage of the state-space model is that it allows expressing power data of the actual day according to power data from several other days. Moreover, in future work, the proposed strategies, used for short-term prediction, can be extended to medium and long-term prediction.

References

[1] A. Ahmadpour, E. Mokaramian, S. Anderson, The effects of the renewable energies penetration on the surplus welfare under energy policy 164 (2021) 1171–1182.

[2] D. Zhan, M.-P. Kwan, W. Zhang, X. Yu, B. Meng, Q. Liu, The driving factors of air quality index in China, Journal of Cleaner Production 197.

[3] C. Medina, C. Ana, R. M., & G. González, (2022). Transmission Grids to Foster High Penetration of Large-Scale Variable Renewable Energy Sources–A Review of Challenges, Problems, and Solutions. International Journal of Renewable Energy Research (IJRER), 12(1), 146-169.

[4] S. Adibpour, A. Raisi, B. Ghasemi, A. R. Sajadi, G. Rosengarten, Ex- perimental investigation of the performance of a sun tracking photovoltaic panel with Phase Change Material, Renewable Energy 165 (2021) 321–333.

[5] V. d'Alessandro, F. Di Napoli, P. Guerriero, S. Daliento, An automated high-granularity tool for a fast evaluation of the yield of PV plants ac- counting for shading effects, Renewable Energy 83.

[6] L. M. Putranto, T. Widodo, H. Indrawan, M. Imron, A., & S. A. Rosyadi. (2022). Grid parity analysis: The present state of PV rooftop in Indonesia. Renewable Energy Focus, 40, 23-38.

[7] M. Kamran, M. R. Fazal, M. Mudassar, S. R. Ahmed, M. Adnan, I. Abid, & H. Shams (2019). Solar photovoltaic grid parity: a review of issues, challenges and status of different PV markets. International Journal of Renewable Energy Research (IJRER), 9(1), 244-260.

[8] S. Daliento, A. Chouder, P. Guerriero, A. M. Pavan, A. Mellit, R. Moeini, P. Tricoli, Monitoring, diagnosis, and power forecasting for photovoltaic fields: A review, International Journal of Photoenergy 2017.

[9] A. M. Pavan, G. Sulligoi, V. Lughi, F. Pauli, R. Miceli, V. Di Dio, F. Viola, Leading the way toward fuel parity in photovoltaics: The utility-scale market in Sicily, Italy, in: 2016 IEEE 16th International Conference on Environment and Electrical Engineering (EEEIC), ieee, 2016, pp.1 5.

[10] A. Dolara, S. Leva, M. Mussetta, E. Ogliari, PV hourly day-ahead power forecasting in a micro grid context, in: 2016 IEEE 16th International Con- ference on Environment and Electrical Engineering (EEEIC), IEEE, 2016, pp. 1–5.

[11] L. Bird, J. Cochran, X. Wang, Wind and solar energy curtailment: experience and practices in the United States, Tech. rep., National RenewableEnergy Lab.(NREL), Golden, CO (United States) (2014).

[12] S. Sobri, S. Koohi-Kamali, N. A. Rahim, Solar photovoltaic generation forecasting methods: A review, Energy Conversion and Management 156.

[13] M. G. De Giorgi, P. M. Congedo, M. Malvoni, Photovoltaic power forecasting using statistical methods: impact of weather data, IET Science, Measurement & Technology 8 (3).

[14] M. Q. Raza, M. Nadarajah, C. Ekanayake, On recent advances in PV output power forecast, Solar Energy 136.

[15] A. Nespoli, E. Ogliari, S. Leva, A. Massi Pavan, A. Mellit, V. Lughi, A. Dolara, Day-ahead photovoltaic forecasting: A comparison of the most effective techniques, Energies 12 (9).

[16] J. Antonanzas, N. Osorio, R. Escobar, R. Urraca, F. J. Martinez-de Pison, F. Antonanzas-Torres, Review of photovoltaic power forecasting, Solar energy 136.

[17] A. Mills, Implications of wide-area geographic diversity for short-term variability of solar power.

[18] A. Mills, Integrating solar PV in utility system operations.

[19] X. Yan, D. Abbes, B. Francois, Solar radiation forecasting using artificial neural network for local power reserve, in: 2014 International Conference on Electrical Sciences and Technologies in Maghreb (CISTEM), 2014.

[20] M. H. Kermia, D. Abbes, J. Bosche, Photovoltaic power prediction using a recurrent neural network RNN (2020).

[21] A. Mellit, A. M. Pavan, M. Benghanem, Least squares support vector machine for short-term prediction of meteorological time series, Theoretical and applied climatology 111 (1).

[22] M. H. Kermia, K. Almaksour, J. Bosche, D. Abbes, Sliding prediction algorithm using polynomial approach for photovoltaic system energy man- agement, in: 2020 7th International Conference on Control, Decision and Information Technologies (CoDIT), Vol. 1, IEEE, 2020, pp. 739–744.

[23] G. Reikard, Predicting solar radiation at high resolutions: A comparison of time series forecasts, Solar Energy 83 (3).

[24] D. Naware, A. Mitra, Weather classification-based load and solar insolation forecasting for residential applications with LSTM neural networks, Electrical Engineering.

[25] R. Al-Hajj, A. Assi, & M. M. Fouad (2019, November). Stacking-based ensemble of support vector regressors for one-day ahead solar irradiance prediction. In 2019 8th International Conference on Renewable Energy Research and Applications (ICRERA) (pp. 428-433). IEEE.

[26] C. Yang, A. A. Thatte, L. Xie, Multitime-scale data-driven spatio-temporalforecast of photovoltaic generation, IEEE Transactions on Sustainable Energy 6 (1).

[27] A. Yona, T. Senjyu, T. Funabashi, C.-H. Kim, Determination method of insolation prediction with fuzzy and applying neural network for long-term ahead PV power output correction, IEEE Transactions on Sustainable Energy 4 (2).

[28] S. Pelland, G. Galanis, G. Kallos, Solar and photovoltaic forecasting through post-processing of the Global Environmental Multiscale numerical weather prediction model, Progress in photovoltaics: Research and Applications 21 (3).

[29] O. Henni and M. Belarbi, "Effect of Mathematical Models on Forecasting Analysis of Photovoltaic Power," *2021 9th International Conference on Smart Grid (icSmartGrid)*, 2021, pp. 163-166.

[30] A. Alkholidi, G., & H. Hamam, (2019). Solar energy potentials in Southeastern European countries: A case study. International Journal of Smart Grid-ijSmartGrid, 3(2), 108-119.

[31] M. Diagne, M. David, P. Lauret, J. Boland, N. Schmutz, Review of solar irradiance forecasting methods and a proposition for small-scale insular grids, Renewable and Sustainable Energy Reviews 27.

[32] S. Hanifi, X. Liu, Z. Lin, & S. Lotfian (2020). A critical review of wind power forecasting methods—past, present and future. Energies, 13(15), 3764.

[33] S. Leva, A. Dolara, F. Grimaccia, M. Mussetta, E. Ogliari, Analysis and validation of 24 hours ahead neural network forecasting of photovoltaic output power, Mathematics and computers in simulation 131.

[34] P. Mandal, S. T. S. Madhira, J. Meng, R. L. Pineda, Forecasting power output of solar photovoltaic system using wavelet transform and artificial intelligence techniques, Procedia Computer Science 12.

[35] A. Gensler, J. Henze, B. Sick, N. Raabe, Deep Learning for solar power forecasting.An approach using AutoEncoder and LSTM Neural Networks, in: 2016 IEEE international conference on systems, man, and cybernetics (SMC), IEEE, 2016, pp. 002858–002865.

[36] M. Omar, A. Dolara, G. Magistrati, M. Mussetta, E. Ogliari, F. Viola, Day-ahead forecasting for photovoltaic power using artificial neural networks ensembles, in: 2016 IEEE International Conference on Renewable Energy Research and Applications (ICRERA), IEEE, 2016, pp. 1152–1157.

[37] L. A. Fernandez-Jimenez, A. Mũnoz-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, Renewable Energy 44.

[38] A. Dolara, F. Grimaccia, S. Leva, M. Mussetta, E. Ogliari, A physical hybrid artificial neural network for short term forecasting of PV plant power490 output, Energies 8 (2).

[39] M. Bouzerdoum, A. Mellit, A. M. Pavan, A hybrid model (SARIMA–SVM) for short-term power forecasting of a small-scale grid-connected photovoltaic plant, Solar Energy 98.

[40] G. Chicco, V. Cocina, P. Di Leo, F. Spertino, A. Massi Pavan, Error assessment of solar irradiance forecasts and AC power from energy conversion model in grid-Connected photovoltaic systems, Energies 9 (1).

[41] I. Goodfellow, Y. Bengio, A. Courville, Deep learning, MIT press, 2016.

[42] R. Ulbricht, U. Fischer, W. Lehner, H. Donker, First steps towards a systematical optimized strategy for solar energy supply forecasting, in: Proceedings of the European Conference on Machine Learning and Principles and Practice

of Knowledge Discovery in Databases (ECMLPKDD 2013), Prague, Czech Republic, Vol. 2327, 2013.

[43] J. Kleissl, Solar energy forecasting and resource assessment, Academic Press, 2013.

[44] A. Alzahrani, P. Shamsi, M. Ferdowsi, & C. Dagli, (2017, November). Solar irradiance forecasting using deep recurrent neural networks. In 2017 IEEE 6th international conference on renewable energy research and applications (ICRERA) (pp. 988-994). Ieee.

[45] M. Yesilbudak, M. Colak, & R. Bayindir (2018). What are the current status and future prospects in solar irradiance and solar power forecasting. International Journal of Renewable Energy Research (IJRER).

[46] M. Colak, M. Yesilbudak & R. Bayindir (2020). Daily photovoltaic power prediction enhanced by hybrid GWO-MLP, ALO-MLP and WOA-MLP models using meteorological information. Energies.

[47] M. Yesilbudak, M. Çolak & R. Bayindir (2016, November). A review of data mining and solar power prediction. In 2016 IEEE International Conference on Renewable Energy Research and Applications (ICRERA).

[48] R. Bayindir, M. Yesilbudak, M. Colak & N. Genc. A novel application of naive bayes classifier in photovoltaic energy prediction. In 2017 16th IEEE international conference on machine learning and applications (ICMLA).