Photovoltaic power forecasting techniques: Application in Artificial Neural Networks

Yassine Benkirane[1*], *Asmae* Ouhmida[2], *Abdelhadi* Raihani[2], and *Bachir* Benhala[1]

[1]LESSI Laboratory, Faculty of Science Dhar El Mahraz, Sidi Mohamed Ben Abdellah University, Fes Morocco
[2]EEIS Laboratory, ENSET Mohammedia, Hassan II University, Mohammedia, Morocco

Abstract. Because photovoltaic energy is heavily reliant on weather and geographic regions, resulting in unpredictable fluctuations, integrating it into electric grids is complex. These discrepancies may lead to system imbalances, spikes, poor planning, and monetary losses. While forecasting methods can mitigate these problems, a number of factors must be taken into consideration, such as the forecasting horizon, input variable correlation analysis, input data processing, weather classification, network optimization, uncertainty quantification, and performance evaluation.
In this article, we highlight the main prediction techniques applied to photovoltaic production. We focus on artificial neural networks (ANN), AutoRegressive Integrated Moving Average (ARIMA), Long Short-Term Memory (LSTM) and Support Vector Machine (SVM) regarding their implementation for predicting photovoltaic energy production. Also, we cover in this study all known forecasting horizons: very short-term forecasts, short-term forecasts, medium-term forecasts, and long-term forecasts. In addition, some performance evaluation metrics such as Mean Square Error (MSE), Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and the coefficient of determination R^2 are explained. At last, we demonstrate the usefulness of ANN in energy forecasting by applying it to photovoltaic power data.
Keywords. Forecasting techniques; Photovoltaic; Solar energy; Forecasting horizons; Evaluation metrics; Artificial Neural Networks.

1 Introduction

Forecasting solar energy production is essential for maximizing energy output, enabling participation in energy markets, and effectively integrating solar energy into the electrical grid [1]. Maintaining grid stability, controlling energy storage, and implementing demand response plans all depend on these estimates [2]. In the end, they benefit power producers, grid operators, and customers by assisting with capacity planning, resilience during unfavorable weather occurrences, and lowering the environmental impact of power generation [3].

* Corresponding author: yassben204@gmail.com

But there are difficulties associated with this requirement for precise forecasts. Due to the dynamic interaction of meteorological components like cloud cover and seasonal oscillations, they necessitate models that capture both short-term subtleties and long-term trends. There is potential when combining machine learning and artificial intelligence methods with cloud forecasting. Furthermore, the complex context that solar power generation projections must traverse is highlighted by forecast uncertainty, market integration, grid stability, technology advancements, and regulatory dynamics. To maximize solar energy utilization and guarantee its smooth integration into the evolving energy system, these obstacles must be overcome [4].
Forecasts using artificial intelligence data, and more specifically with artificial neural networks (ANN), stand out for their ability to process large amounts of data and model nonlinear relationships taking on complex forms. These models are increasingly used to anticipate solar energy production at different time scales, ranging from very short term to long term.
This article proposes the most common methods for forecasting solar energy production, focusing on ANN, ARIMA, LSTM and SVM. It considers different forecasting horizons, ranging from very short term to long term, and describes the relevant metrics for performance quantification. Finally, the article demonstrates the importance of ANN by applying them to a real dataset.

2 Defining forecasting

2.1 Definition

An estimate of a future event created by carefully gathering and extrapolating historical data in a predetermined way is called a prediction. It's just a forward-looking claim. It is obvious that we need to differentiate between accurate forecasts and forecasts in general. An accurate prediction can be quite beneficial and would be very valuable. Numerous elements need to be taken into account when making long-term planning decisions, including the state of the economy generally, industry trends, potential competitor activity, the political environment generally, and more [5].
Forecasting is only possible with a history of data.
Remember that an excellent forecasting is typically unattainable. Therefore, it is much more crucial to develop the habit of regularly reviewing forecasts and learning to live with faulty projections than it is to strive for the ideal one. This is not to suggest that we shouldn't work to enhance the forecasting model or methodology; rather, it means that we should look for and employ the most effective forecasting technique that is feasible. Since our estimates are based on historical data, they will become less accurate the further into the future we project. This indicates that as the time horizon grows, forecast accuracy falls [5].

2.2 Common Procedures for Forecasting

The forecasting process consists of the following key steps [5]:

- Determine the overall need.
- Choose the Forecast Period (Time Horizon).
- Decide which forecast model to use: Understanding different forecasting models, their applicability in different scenarios, their reliability, and the kinds of data needed are necessary for this. One or more models may be chosen based on these factors.

2

- Data Collection: Gather data obtained from a variety of relevant sources that is consistent with the model or models chosen in the third step, taking into account the different indicators that have been identified. Additionally, data should go back that far in time, which satisfies the model's requirements.
- Create a forecast by applying the model to the gathered data and determining out the forecast's value.
- Evaluate: The predictions made by any of the models should not be taken at face value. It should be assessed in terms of the "confidence interval"; typically, all effective prediction models include techniques for determining the upper and lower values that the provided forecast is likely to stay within at a given degree of likelihood. Whether the value obtained is logically feasible can also be assessed from a logical perspective. Additionally, it might be assessed in relation to a linked variable or phenomenon. Therefore, it is feasible - and occasionally wise - to alter the statistically predicted value in light of evaluation.

Different forecasting techniques have been shown in Figure 1.

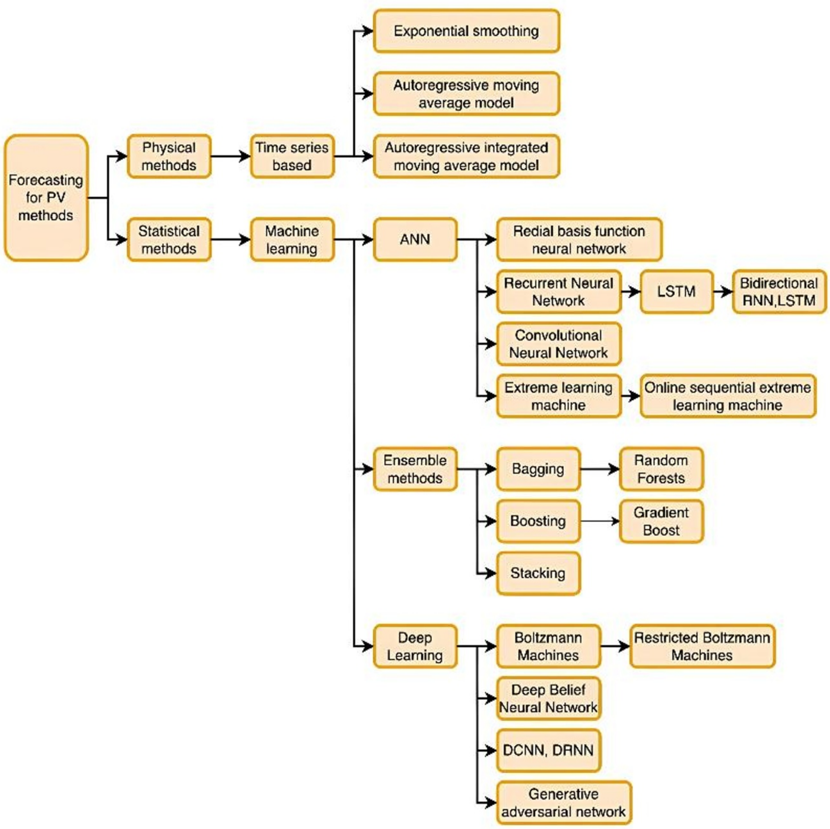

Fig. 1. PV forecasting techniques classification [4].

The following section provides explanations of some of the most popular forecasting techniques.

3 Forecasting techniques

3

3.1 Artificial Neural Network (ANN)

Another statistical model that is thought to be the most widely used and successful machine learning algorithm for modelling and forecasting PV output power is the artificial neural network (ANN), which can control the non-linear relationship between the input meteorological data and the output PV power. It has demonstrated a high rate of accuracy in forecasting PV output power [6].
Without any prior assumptions, it is better suited to managing the complicated and online interaction between the solar PV output power and the input features (meteorological data).
The basic components of an ANN are its input, hidden, and output layers, neurons, and connections, as seen in Figure 2a. The input layer receives various input data. For data analysis, the hidden layer may consist of one or more layers. The connections connect the various neurons in adjacent layers with the updated weights, while the output layer completes the input processing to deliver the network output [7]. Figure 2b provides a visual depiction of a basic ANN cell. The combination and activation functions are the two primary components of a neuronal cell, as this diagram illustrates. The combination function calculates the total of the input values. By employing the activation function to perform a weighted sum of the inputs, this network generates results. As a result, the activation function transfers the input to the output by acting like a squeeze function. Linear, Gaussian radial basis, sigmoid, hyperbolic tangent sigmoid, bipolar step, bipolar linear, unipolar step, and unipolar functions are a few of the activation functions commonly used in ANNs.
An ANN can be mathematically expressed as [8]:

$$U_J = b + \sum_{K=1}^{J} (W_k \text{ x } I_K) \quad (1)$$

where J, W_k, and I_K stand for the number of inputs, connection weight, and network input, respectively; U_J is the final network output; and b is the bias weight.

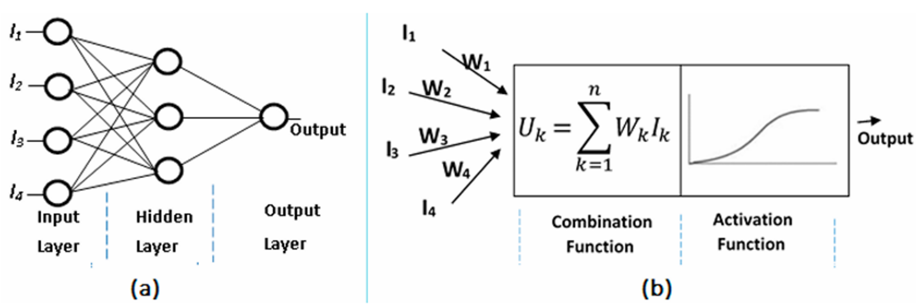

Fig. 2. (a) Schematic representation of an ANN with the input, hidden and output layers. (b) An illustration of an ANN cell's mathematical model [7].

3.2 AutoRgressive Integrated Moving Average (ARIMA)

ARIMA model is the most basic and adaptable kind of time series forecasting technique. These are predicated on the notion that the differencing process can make the time series stationary. The temporal sequence could be assumed to be stationary if all of its statistical properties hold true across time [9].
The lags of the error term can't be taken as the unbiased variables as they're not linear features of the coefficients. hence, the coefficients of ARIMA models with lagged errors must be computed by other techniques such as nonlinear optimization. The lags terms of the stationary

4

time series are called "autoregressive", while the lags of the forecasted error terms are referred as "moving common". a time series which requires to be differenced for the cause of making it stationary is said to be an "integrated" model of a stationary series. These models are called ARIMA (p, d, q), where "p" stands for the autoregressive part's order, "d" for the degree of first differencing concerned, and "q" for the moving average part's order [9]. Equation 2 represents the autoregressive component of the model with order "p".

$$Y_t = c + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_P Y_{t-P} + e_t \quad (2)$$

Equation 3 represents the moving average model of order 'q'.

$$Y_t = c + \theta_1 e_{t-1} + \theta_2 e_{t-2} + \cdots + \theta_P e_{t-P} + e_t \quad (3)$$

Where "e_t" is the error series and "Y" is the output of a time series, such as data on electricity use.

3.3 Long Short-Term Memory (LSTM)

The LSTM method is a member of the (deep) recurrent neural network (RNN) family. Figure 3 illustrates how RNNs vary from Feed Forward Neural Networks, which are among the most widely used forecasting techniques for solar energy.

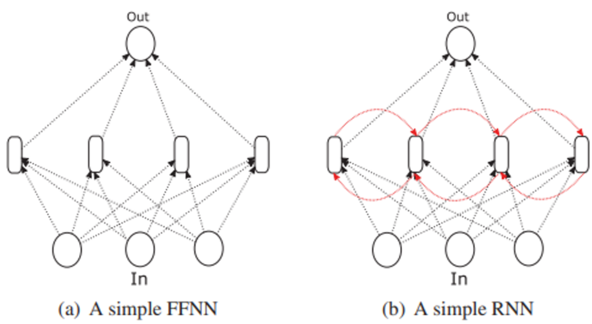

Fig. 3. Neural network bifurcations.

To apprehend the functioning of LSTM memory units, Figure 4 depicts a single localized LSTM cell in the first layer of a network at a time step t. Input (i_t), update (g_t), output (o_t) and forget (f_t) represent the output values of the 4 forms of gates. The gates obtain a enter of the same LSTM unit's output obtained at a previous time step (h_{t-1}). The gates also obtain enter data related to the modern time step (x_t). In Figure 4, we use colored arrows to depict this message waft. Equations 4 to 9 are governing equations, which we explain below. Word that θ_{xn} and θ_{hn} are weight matrices of enter and output data from previous cells, while b_n represents bias vectors with n ∈ (*i, g, f, u*), which discover the aforementioned gates. As is trendy exercise in neural network studying, the parameters are initialized via sampling from a uniform distribution and updated using back-propagation method [10].

5

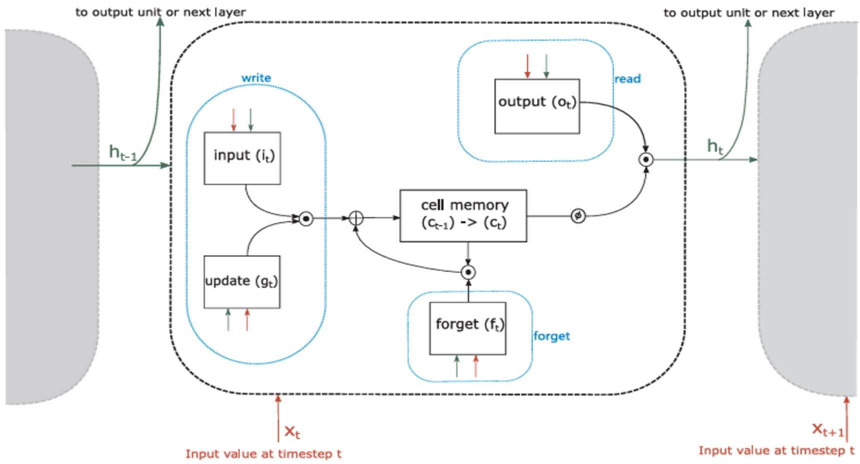

Fig. 4. An LSTM unit in the 1st layer for time step t.

Where:

- ⊙ and ⊕ stand for the sum function and point-wise scalar multiplication, respectively.
- Color-coded arrows indicate which way a system's inputs are going.
- In this experiment, the user-defined cell output activation function, denoted by ∅, inside an O, is set to ReLU.

The forget gate, an essential part of the LSTM design, chooses how much data it wants to erase from memory. It functions by means of a hard-sigmoid activation (linearly approximated standard sigmoid), which is controlled by the equation 4 and ranges from 0 (forget all) to 1 (remember all) [11].

$$f_t = sigm\,(\theta_{xf}x_t + \theta_{hf}h_{t-1} + b_f) \tag{4}$$

The write function in memory is performed by the interaction (element-wise multiplication) of the input and update gate (see Figure 4). The input gate determined which values to write using a hard sigmoid, while the update gate produced a vector of new cell values using hyperbolic tangent (tanh) activation. Their work is represented by equations (5) and (6). Since a nearly zero input activation value will stop incoming inputs from updating cell memory, stored data can be used for a very long time [11].

$$i_t = sigm\,(\theta_{xi}x_t + \theta_{hi}h_{t-1} + b_i) \tag{5}$$

$$g_t = tanh\,(\theta_{xg}x_t + \theta_{hg}h_{t-1} + b_g) \tag{6}$$

Cell memory (see Figure 4) updates itself recursively by interacting with the values of the forget and write gates and its old value (t-1). As a result, equation 7.

$$C_t = f_t \odot C_{t-1} + i_t \odot g_t \tag{7}$$

Lastly, the output of the cell memory, also known as the cell state, is calculated by combining the output gate, which implements a read function, with cell memory (element-wise multiplication). The cell memory has an activation function called ϕ that binds output values

6

and determines which values will come out of the output gate. In the cell output equation, the activation function, ϕ, is frequently *tanh*. But for this experiment, we employ Rectified Linear Unit (*ReLu*), which is a more popular option in the current literature on deep learning [12]. Equation 8 controls the output gate's read function, and equation 9 determines the final cell output.

$$o_t = sigm\ (\theta_{xo}x_t + \theta_{ho}h_{t-1} + b_0) \tag{8}$$

$$h_t = \theta_t \odot \phi(C_t) \tag{9}$$

3.4 Support Vector Machine (SVM)

SVM is a kernel-based machine learning technique for classification and regression. These models avoid overfitting of training data, compute quickly, employ kernels, and converge efficiently [13].

The SVM makes use of two key concepts. Initially, the problem is converted from the original input space into a highly dimensional one known as the feature space using kernel functions. As a result, training samples can be linearly separated into several classes. Second, the idea of maximum margin is presented in order to determine the optimal separation hyperplane. Lastly, the SVM's defining optimization problem is convex and quadratic, meaning it can be effectively solved [14].

Figure 5 illustrates the fundamental concept of Support Vector Machines (SVM). Two sets of data - one green and one blue - hat correspond to two distinct categories are included. The optimal boundary between these two classes is represented by the black line in the middle, which is the ideal hyperplane. The solid red dashed lines represent the support vectors, which are the nearest points from each class. SVM aims to improve separation and generalization to unseen data by expanding the margin provided by the support vectors.

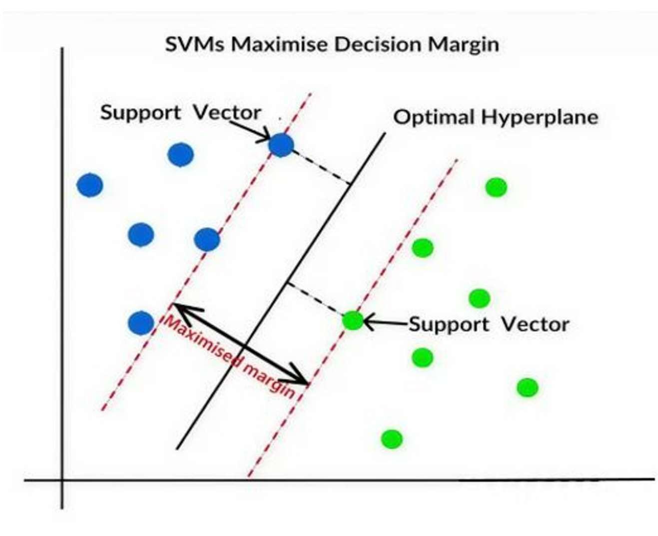

Fig. 5. Fundamental concept of Support Vector Machines (SVM).

4 Advantages and disadvantages of the forecasting techniques

7

Table 1. Advantages and disadvantages of the forecasting techniques.

Forecasting techniques	Advantages	Disadvantages
ANN	-The capacity to accurately trace the links between input and output. -Excellent performance for non-linear time series. -More robust and general.	-Requires initializing the weight values. -The local minima problem. -Overfitting and challenging to extrapolate.
ARIMA	-Uses historical data shift and lag. -Moving average regression model (improves accuracy). -Offers reliable confidence intervals for forecasts.	-Identifying the model is challenging. -Unsuitable for forecasting the long term. -Fails to adequately depict the series non-linear patterns.
LSTM	-Capturing long-term temporal dependencies. -The accuracy surpasses that of classical models. -Resilience to noise in the data. -Flexibility to different time horizons.	-High computational cost. -Need for a lot of quality data. -Difficulty of interpretation. -Risk of overfitting.
SVM	-Good for generalization and fitting. -Works effectively with long-term time series. -A kernel function handles arbitrarily organized data and introduces nonlinearity.	-The results are not transparent. -As the number of factors and dataset size grow, determining the ideal parameters may become more computationally demanding.

5 Forecasting horizons

The time period over which we are projecting PV power generation is known as the forecast horizon [15]. Depending on how far into the future we are attempting to predict, forecast accuracy can vary even with the same model parameters. Consequently, the prediction horizon needs to be taken into account while developing a forecasting model. Forecasts of PV power output can be categorized into four groups according to the forecast horizon [4].

- **Very short-term forecasting:** A prediction is made that typically lasts for a few seconds to a minute.
- **Short-term forecasting:** A forecast is created for a period of one to several days. It is crucial for managing and planning power generation and aids in preserving the security of the power system.
- **Medium-term forecasting:** A week to a month is the time frame covered by the forecast. This category serves to simplify maintenance scheduling and power system planning by forecasting future power availability.
- **Long-term forecasting:** The prediction is produced for a time frame ranging from one month to a year. It is essential to the production, distribution, transmission, and planning of power.

6 Forecasts model performance

To evaluate the effectiveness of the created forecasting benchmarks for the solar PV power output, a few pertinent assessment measures are chosen. We have chosen the main measures

8

that have been suggested by studies and recommendations in the sector. Coefficient of determination (R^2), Mean Square Error (MSE), Root Mean Square Error (RMSE), and Mean Absolute Error (MAE) are among the selected metrics [16].
RMSE is often used as an evaluation tool in research prediction because it effectively penalizes severe errors. By showing the gap between the predicted and reference DC solar power values, MAE provides a comprehensive view of how closely the forecasts and actual measurements agree. The coefficient of determination, or R^2, calculates the percentage of variation in the dependent variable that can be explained by the independent variables. A score close to 1 means that the model closely matches the observed data and has a high forecasting accuracy. Its values range from 0 to 1 (or 0% to 100%) [17].
The following are the formulas for these metrics:

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{N}(y_i - \bar{y}_i)^2} \tag{10}$$

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2 \tag{11}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2} \tag{12}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}_i| \tag{13}$$

Where:
N: The whole-time step horizon
$\hat{y}_i$: The forecast output power
y_i: The power of reference
$\bar{y}_i$: The reference's mean

7 Application in Artificial Neural Networks (ANN)

Our goal is to create an application that uses the prediction method “Artificial Neural Networks” to forecast photovoltaic powers for the future 30 days based on data that includes photovoltaic powers, dates, and hours.

7.1 Results

Figure 6 illustrates how we visualized our data on the evolution of photovoltaic power over time from July 2015 to July 2016. The seasonal variations are pronounced. There are higher production peaks compared to lower peaks from September 2015 to March 2016 due to abundant sunlight and moderate temperatures that favor the performance of solar panels. On the other hand, periods of non-production or very minimal production are often associated with winter, when the days are shorter and the sun's altitude is low, accompanied by cloudy skies. This may also be due to technical interruptions or a lack of data collection. The power fluctuates between 0 and approximately 1.8 with a high level of daily variability. This type of visualization helps to evaluate the overall performance of a PV system, assists in error detection, and supports forecasting models.

9

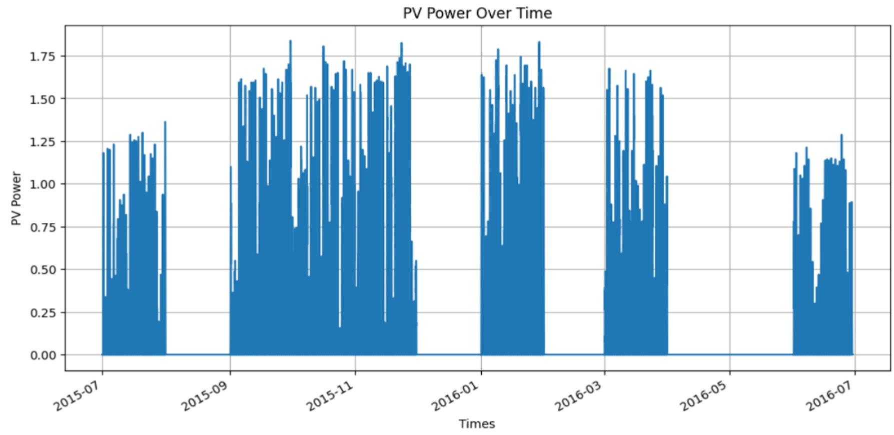

Fig. 6. Data visualization.

We will now apply the ANN to make a prediction for the next 30 days. So, the predicted photovoltaic powers are illustrated in Figure 7.

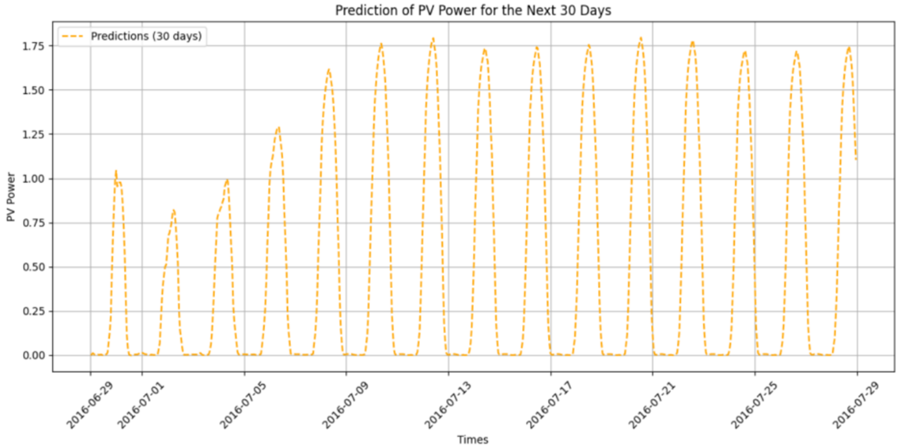

Fig. 7. Prediction of photovoltaic power for the next 30 days.

The objective of the table below, which presents the metrics obtained following the forecasting, is to evaluate them.

Table 2. Forecasts model performance of ANN.

Metrics	**Values**
R^2	0.90459 ≈ 90%
MSE	0.00309
RMSE	0.05563

10

MAE	0.02452

7.2 Discussion

The forecasted graphs for the next thirty days indicate that the predictions follow consistent and sensible patterns, with an increase during the day and a drop to zero at night. This reflects the behavior of solar energy production in sunny conditions. That said, the overly uniform and periodic aspect of the graphs suggests that the model does not sufficiently account for the natural variability related to external factors such as weather conditions or unforeseen events (disruptions, rain, shade). That said, although the somewhat elementary model offers a solid estimate of the structural trend, its use in practical contexts would require the integration of additional climatic data to enhance accuracy and precision in the face of variations.

The results demonstrate that the model possesses satisfactory predictive abilities. The observed gap MSE is minimal (0.00309), indicating that there are no significant differences between the forecasts and the actual values. Moreover, the RMSE (0.05563) obtained in the same PV power units confirms the claimed accuracy and closely aligns with previous results, proving the model's robustness. Moreover, the MAE (0.02452) supports the idea of accuracy by asserting that the mean absolute error is extremely low. To conclude, the value of the coefficient of determination R^2 stands at (0.90459), indicating that the model manages to explain more than 90% of the variations observed in the actual photovoltaic data, demonstrating remarkable predictive accuracy.

8 Conclusion

This article explored some different methods used to anticipate photovoltaic energy production. A given technique may be useful in certain case studies but not necessarily applicable to others. It is now essential to build models for forecasting photovoltaic energy generation, particularly for utility-level planning. Many studies have been conducted to develop efficient models. However, there are still unmet needs, such as the creation of extremely dependable models that are not reliant on weather predictions, as well as PV power forecasting models that are optimal, accurate, and computationally inexpensive for medium and long-term forecasts.

The study also focused on the importance of the time span (very short, short, medium, and long term) as well as the relevant evaluation indicators (MSE, MAE, RMSE, and R^2) to assess the performance of predictive models. When an ANN model was applied practically to real data, such as dates, times, and photovoltaic powers, it demonstrated good prediction potential with a low mean error and an R^2 larger than 90%, confirming the effectiveness of this approach under controlled conditions.

The work done yields promising results and presents various opportunities:

- Add weather conditions (solar irradiation, temperature, etc.) to improve the performance of the prediction model.
- Work with hybrid techniques (like ANN+LSTM, ANN+SVM, etc.) to capture the advantages of each method.
- Create a novel forecasting technique that hasn't been used in any work of literature yet.

11

References

1. Guo ZF, Zhou KL, Zhang C, Lu XH, Chen W, Yang SL. Residential electricity consumption behavior: influencing factors, related theories and intervention strategies. Renew Sustain Energy Rev Jan 2018;81:399–412.
2. Mocanu E, Nguyen PH, Gibescu M, Kling WL. Deep learning for estimating building energy consumption. Sustain Energy Grids; Networks Jun 2016;6:91–9.
3. Ssen Z. Solar energy in progress and future research trends. Prog Energy Combust Sci 2004;30(4):367–416.
4. GUPTA, Manvi, et al. A review of PV power forecasting using machine learning techniques. Progress in Engineering Science, 2025, 100058.
5. GOR, Ravi Mahendra, MOHAN, Man. Industrial statistics and operational management–forecasting techniques. Tersedia online di http://doer. col. org/handle/123456789/5107 (24 September 2017), 2014.
6. Chu,Y.; Urquhart, B.; Gohari, S.; Pedro, H.; Kleissl, J.; Coimbra, C. Short-term reforecasting of power output from a 48 MWe solar PVplant. Sol. Energy 2015, 112, 68–77.
7. Aminzadeh, F.; De Groot, P. Neural Networks and Other Soft Computing Techniques with Applications in the Oil Industry; Eage Publications: Lancaster, OH, USA, 2006.
8. Hossain, M.; Ong, Z.; Ismail, Z.; Noroozi, S.; Khoo, S. Artificial neural networks for vibration based inverse parametric identifications: A review. Appl. Soft Comput. 2017, 52, 203–219.
9. DEB, Chirag, et al. A review on time series forecasting techniques for building energy consumption. Renewable and Sustainable Energy Reviews, 2017, 74: 902-924.
10. Schmidhuber, J., 2015. Deep learning in neural networks: an overview. Neural Netw. 61, 85–117.
11. Srivastava S, Lessmann S. A comparative study of LSTM neural networks in forecasting day-ahead global horizontal irradiance with satellite data. Sol Energy 2018;162:232–47.
12. LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. Nature 521 (7553), 436–444.
13. Tongyang Xu, Tianhua Xu, Izzat Darwazeh, Deep learning for interference cancellation in non-orthogonal signal based optical communication systems, in: 2018 Progress in Electromagnetics Research Symposium (PIERS-Toyama), IEEE, 2018.
14. CHOROWSKI, Jan; WANG, Jian; ZURADA, Jacek M. Review and performance comparison of SVM-and ELM-based classifiers. Neurocomputing, 2014, 128: 507-516.
15. AHMED, Razin, et al. A review and evaluation of the state-of-the-art in PV solar power forecasting: Techniques and optimization.Renewable and sustainable energy reviews, 2020, 124: 109792.
16. Michael NE, Bansal RC, Ismail AAA, Elnady A, Hasan S. A cohesive structure of bi-directional long-short-term memory (BiLSTM)-GRU for predicting hourly solar radiation. Renew Energy 2024;222:119943.
17. MBUNGU, Nsilulu T., et al. Predictive control technique for solar photovoltaic power forecasting. Energy Conversion and Management: X, 2024, 24: 100768.

12