Contents list available at CBIORE journal website

**International Journal of Renewable Energy Development**

Journal homepage: https://ijred.cbiore.id

Research Article

# Evaluating and analyzing the performance of PV power output forecasting using different models of machine-learning techniques considering prediction accuracy

Abderraouf Bouakkaz [a*], Adel Lahsasna[b], Antonio J Gil Mena[c], Salim Haddad[a], Mario Luigi Ferrari[d], Rafael Jiménez-Castaneda[e]

[a] *Laboratoire de génie mécanique et matériaux, Université 20 aout 1955 Skikda, Bp26 route d'el Hadaik Skikda, Algérie*

[b] *Département d'informatique, Université 20 aout 1955 Skikda, Bp26 route d'el Hadaik Skikda, Algérie*

[c] *Department of Electrical Engineering, University of Cádiz, Escuela Politécnica Superior de Algeciras, Avda. Ramón Puyol, s/n, 11202 Algeciras (Cádiz), Spain*

[d] *Thermochemical Power Group (TPG), University of Genoa, Genoa, Italy*

[e] *Department of Electrical Engineering, Escuela Superior de ingeniería de Puerto real, Avda. de la Universidad de Cádiz, 10 / 15519 Puerto Real (Cádiz), Spain*

**Abstract**. Solar energy as a clean, renewable, and sustainable energy source has considerable potential to meet global energy needs. However, the intermittent and uncertain character of the solar energy source makes the power balance management a very challenging task. To overcome these shortcomings, providing accurate information about future energy production enables better planning, scheduling, and ensures effective strategies to meet energy demands. The present paper aims to assess the performance of PV power output forecasting in PV systems using various machine learning models, such as artificial neural networks (ANN), linear regression (LR), random forests (RF), and Support Vector Machines (SVM). These learning algorithms are trained on two different datasets with different time steps: in the first one, a historical weather forecast with a one hour time step, and in the second one, a dataset of on-site measurements with a 5-minute time step. To provide a reliable estimation of prediction accuracy for different learning algorithms, a k-fold cross-validation (CV) is applied. Through a comparison analysis, an assessment of the accuracy of these algorithms based on various metrics such as RMSE, MAE, and MRE is performed, providing a detailed evaluation of their performance. Results obtained from this study demonstrate that the random forest algorithm (RF) outperformed other algorithms in predicting PV output, achieving the smallest prediction error, where the best values for RMSE, MRE, MAE, and $R^2$ for the weather dataset were 0.856 W, 0.256%, 0.364 W, and 0.99999, respectively, while thevalues for RMSE, MRE, MAE, and $R^2$ for the on-site measurements dataset were 8.525 W, 11.163%, 3.922 W, and 0.99922, respectively.

**Keywords:** Solar energy; Power production; Energy forecasting; Machine learning; Cross-validation; Accuracy of predictions.



Received: 10th August 2024; Revised: 7th Nov 2024; Accepted: 10th Dec 2024; Available online: 23rd Dec 2024

## 1. Introduction

In the face of global warming and climate change concerns, the energy demand continues to rise worldwide. The energy transition from traditional energy sources to clean, renewable, and more sustainable energy sources is crucial and inevitable (Hoppe and van Bueren, 2015; Chen *et al.*, 2020; Henderson and Sen, 2021). Solar PV energy is a clean, renewable, and sustainable energy source, which attracted many governments' attention, and has considerable potential to meet global energy needs. Consequently, the total surface area of the installed photovoltaic panels has increased massively over the last decade, and based on statistics provided by the International Energy Agency (IEA), the generated PV power increased by 26% in 2022, reaching almost 1,300 TWh. Among all renewable technologies, it demonstrated the greatest absolute growth in 2022 (IEA, 2023).

However, solar energy is an intermittent, variable, and uncertain energy source, as it is strongly influenced by weather conditions (solar irradiance, ambient temperature, wind speed, etc.). These issues have a significant impact on network stability and make maintaining the balance of power a difficult and challenging task (Tarroja, Mueller and Samuelsen, 2013; McCormick and Suehrcke, 2018; Notton et al., 2018; Yin, Molini and Porporato, 2020). Therefore, having accurate information about future energy production is beneficial for the power balance management. It allows for better planning, better allocation of resources and provides efficient strategies to maintain grid stability based on anticipated energy production and consumption patterns (Arutyunov and Lisichkin, 2017; Zack, 2017; Iheanetu, 2022).

* Corresponding author
Email ab.bouakkaz@univ-skikda.dz (A. Bouakkaz)

In the literature, there are many techniques and methods that have been widely discussed and used to predict photovoltaic power generation, including techniques which employ physical models, probabilistic and statistical methods, as well as machine learning (ML) techniques. (Antonanzas *et al.*, 2016 ; Li, Zhou and Yang, 2018 ; Mellit *et al.*, 2020 ; Dimd *et al.*, 2022). The authors in (De Giorgi, Congedo and Malvoni, 2014) developed an algorithm combining multi-regression and ANN used to predict the electrical output of 0.96 MW in an Italian grid-connected photovoltaic plant. According to the authors, using measured weather parameters improves historical forecasting accuracy. Based on ANN, the authors in (Gandelli *et al.*, 2014) proposed a novel hybrid method called PHANN, which combines physical systems with ANN. They demonstrated that the PHANN method yields more accurate predictions compared to traditional methods. Both weather forecasts and historical data for one day ahead with one-hour time step were used to train an MLP-based forecaster in (Leva *et al.*, 2017). The authors proved that model performance was better on sunny days than partly cloudy days, as they found a normalized MAE of less than 15% in all investigated cases. (Liu *et al.*, 2018) proposed an ultra-short-term forecasting of a 1.2 kWp of PV plant installed in Beijing, China using SVM and ANN. This model was specially designed to meet specific environmental conditions, such as fog and mist. The input data includes measurements of air temperature, relative humidity and aerosol indices, collected directly from the source. (Zhou *et al.*, 2020) developed a hybrid model combining genetic algorithms and extreme learning machines, based on similar day analysis (SDA-GA-ELM) predicting a day-ahead PV power generation, where they used GA to find the optimal hidden bias and input weight values. According to the obtained results, the proposed model demonstrated its efficiency in predicting PV power output with higher accuracy and more stability. The authors in (Khadke *et al.*, 2023) proposed a model based on various machine learning techniques, including LR, RF, principal component analysis, and SVM with an RBF kernel, that accounts for the unpredictability of weather conditions when forecasting solar PV systems. The proposed model demonstrated high accuracy in solar energy prediction, achieving the highest $R^2$ value of 0.87.

The present work focuses on the evaluation of forecasting accuracy of PV power output using various machine learning models trained on two different datasets. The main innovative aspect of this study lies in evaluating the effectiveness of various machine learning models for predicting the output power in PV systems using multiple performance metrics. In addition, a k-fold cross-validation (CV) technique is applied to ensure reliable accuracy estimates for different learning algorithms.

The key contributions of this paper are outlined as follows:

- Data Source Integration: Unlike previous studies that often rely on limited or single-source datasets, this study integrates both historical weather forecasts and on-site measurements with different time steps to provide a more robust and comprehensive of model evaluation. This approach captures the impact of varying environmental conditions on PV power output forecasting, which is less commonly explored in existing literature.
- Application of various machine learning models: Although various models have been used to predict PV power output, this study demonstrates the effectiveness of the Random Forest (RF) model, which is less commonly applied in PV forecasting compared to other models.
- Use of k-fold Cross-Validation technique: To ensure robust model assessment, a k-fold cross-validation technique is employed. This method reduces bias and provides a more reliable understanding of each model's accuracy.
- Various Performance Metrics: While previous research typically focuses on a few metrics (such as RMSE or MAE) for model evaluation, in this study an assessment of the performance of algorithms across a broader range of metrics, including MAE, MRE, RMSE, and $R^2$. This comprehensive evaluation allows for a more depth analysis of each model's performance.

In the rest of the paper, the following sections are presented: Section 2 introduces Materials and Methods, which includes the proposed model's structure of PV output power forecasting while a cross-validation technique is presented in section 3. Section 4 discusses the performance metrics. Section 5 presents an analysis and comparison of the obtained results. Section 6 concludes the presented work.

## 2. Materials and Methods

The general structure of the forecasting model proposed in this paper is shown in Fig. 1 while Fig. 2 illustrates the description of the proposed PV output forecasting methodology. The following description provides a more detailed understanding of the proposed models.

### *2.1 Data description*

In this study, two different datasets were used for training the proposed algorithms. In the first dataset, a weather forecast obtained from the NSRDB ("National Solar Radiation Database") ((NREL, 2023) was used, while in the second dataset, actual solar power measurements of PV output (P, G, Tamb, and Ws) which collected from an installed PV system at the Higher School of Engineering of Puerto Real (University of Cadiz) was used.

#### *2.1.1 Weather forecast data*

The weather forecast data used in this study were obtained from the National Solar Radiation Database (NSRDB) (NREL, 2023), which is a publicly available dataset. These forecasts are based on the Physical Solar Model (PSM), which integrates multi-channel measurements from the Geostationary Operational Environmental Satellite (GOES) to predict various meteo-rological parameters, such as solar radiation, temperature, and wind speed, as well as other weather data. The dataset consists of one year of hourly data on solar irradiance (G), ambient temperature ($T_{amb}$), and wind speed (Ws), covering the geographical region of Skikda, in the northeast of Algeria [36.89; 6.90]. As a result, the dataset contains 8760 samples (24 x 365).

The target value, which is the PV power output, is calculated using the model presented in equations (1) and (2), and based on data from weather forecast data, including solar irradiance (G), ambient temperature ($T_{amb}$), and wind speed (Ws). (Tamizhmani *et al.*, 2003; Chenni *et al.*, 2007; Yona et al., 2013).

$$P_{pv} = P_{STC} \cdot \frac{G}{G_{STC}} \cdot \left(1 - \gamma(T_c - T_{c-STC})\right) \cdot N \quad (1)$$

$$T_c = 0.943 \cdot T_{amb} + 0.028 \cdot G - 1.528 \cdot W_s + 4.328 \quad (2)$$

where: $P_{STC}$, $G_{STC}$, $T_{c-STC}$ are respectively the rated output power of PV generator, the solar irradiance and the PV cell temperature at standard test conditions ($G_{STC}$=1000 W/m²; $T_{c\_STC}$ = 25 °C). $T_c$ is the PV cell temperature, $\gamma$ is the power temperature coefficient which equal to $4.3e^{-3}$ (1/°C) and $N$ present the number of PV panels.

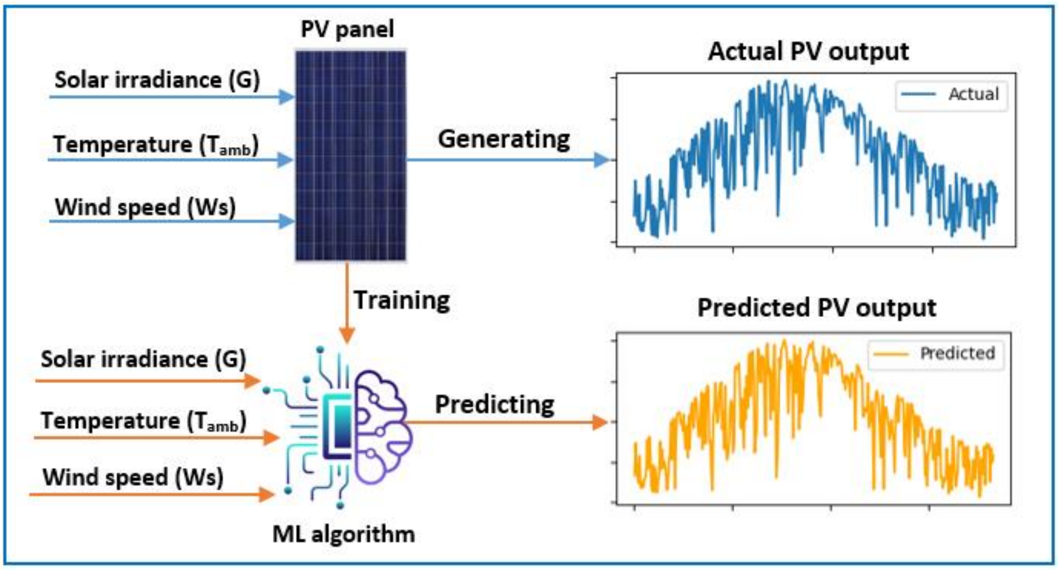


**Fig. 1** The general architecture of PV output power forecasting.

*2.1.2 On-site measurements data*

The on-site measurements were obtained using Solar-Log 1200 interface with an RS485 Modbus sensors measure solar irradiation, temperature and wind speed. The accuracy of these devices is around ±1% for irradiation and ±1°C for temperature, ensuring reliable and high-quality data. These measurements were collected from April 14-2018 to May14-2018 at the installation site of one kWp of PV system located in the College of Engineering University of Cadiz in the south of Spain (36.53, -6.20). Those measurements consist of 30 days of data with a 5-minute time step, meaning 8640 measurements (12 x 24 x 30).

*2.2 Forecasting methodology description*

In this study, a forecasting of PV power generation is performed using various supervised ML algorithms. The aim is to leverage ML algorithms to accurately predict PV power output, which enables effective energy management strategies. As a result, the forecasting methodology used in this study as presented in Fig. 3 consists of the following steps:

Step 1: Selecting training dataset: The first step involves selecting a representative set of datasets to train the machine learning models. This dataset encompasses two different datasets (weather forecast and on-site measurements) and covers most variables that influence PV power output (global horizontal irradiation, temperature and wind speed).In this study, three main variables were used as input features $X=\{G, T_{amb}, Ws\}$ and a PV power output $Y=\{P_{PV}\}$ as the target output. For this study, a total PV power output of 1 kW was used in on-site measurements. The same rated power was used to model photovoltaic production in the weather forecast dataset.

Step 2: Selecting ML algorithm: to compare the prediction accuracy, various ML algorithms were chosen for the prediction task. Therefore, a various algorithms including ANN, LR, RF, and SVM, are considered.

Step 3: Using k-fold cross-validation: compared to the single training-testing split method, the 10-CV method provides more reliable and robust estimation of predictive accuracy of the

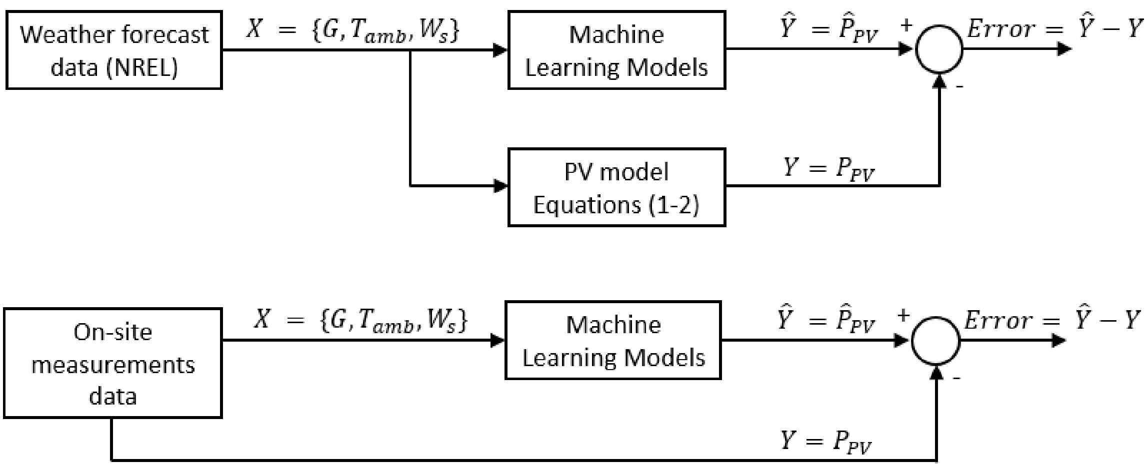


**Fig. 2** The description of the proposed forecasting methodology.

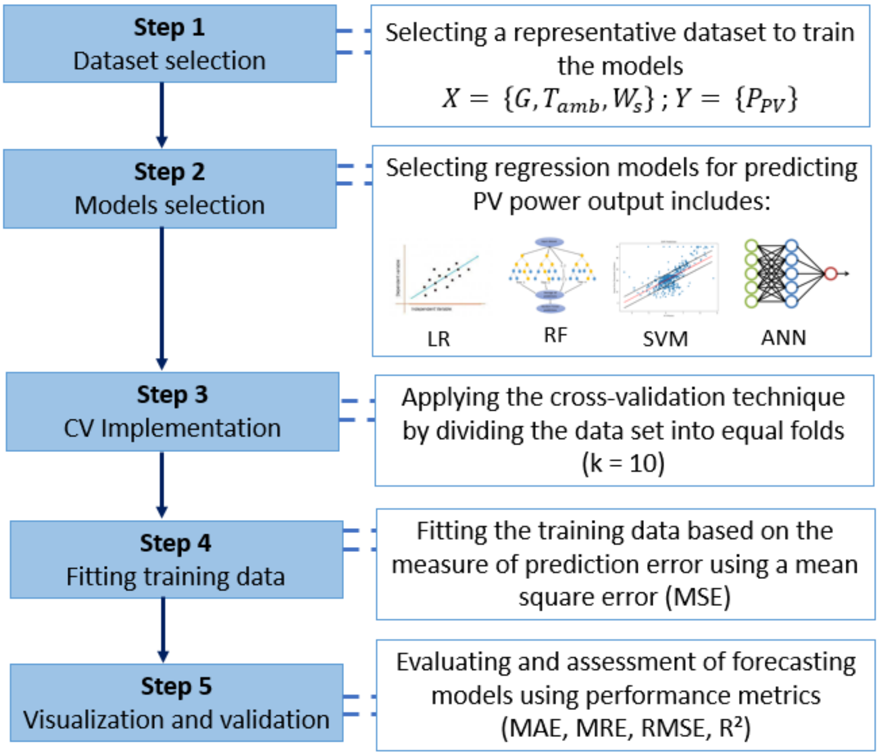


**Fig. 3** Methodology for forecasting PV output power.

algorithms under consideration. A more explanation of the method is provided in Section 3.

Step 4: Fitting training data: a MSE is used to evaluate the accuracy of the proposed models. The MSE calculates the difference between predicted and observed values on the training data, thus providing a quantitative measure of prediction error.

Step 5: Visualizing and evaluating prediction results: once the machine learning models are trained, their performance is evaluated and assessed using various metrics. In addition, a comparative analysis of the accuracy of the model is performed.

*2.3 PV power forecasting models*

For PV power output forecasting, the performance of different learning models is compared. The aim of this comparative study is to identify which model gives the most accurate results for each PV dataset. A brief presentation of each model is provided below.

*2.3.1 Linear regression model*

One of the fundamentals of statistics and machine learning is linear regression. In linear regression, it is assumed that the independent variables x (features) and the dependent variable y (target value), are linearly related as indicted by equations (3-4): (Hope, 2020)

$$\boldsymbol{x} = (\boldsymbol{x_1}, \dots, \boldsymbol{x_n}) \quad (3)$$

$$y = \beta_0 + \beta_1 x_1 + \cdots + \beta_n x_n + \varepsilon \quad (4)$$

Where: $n$ is the number of predictors, $\beta_0$, $\beta_1$,... $\beta_n$ are the regression coefficients, and $\varepsilon$ is the random error.

The objective in linear regression is to find the values of the regression coefficients that minimize the difference between the predicted and actual values of $y$. This is typically achieved by applying Ordinary Least Squares (OLS), which aims to minimize the sum of the squared differences between $\boldsymbol{y_i}$ and $\hat{\boldsymbol{y}}_i$ that represent the actual and predicted PV power outputs, respectively. The objective function of this method which called Residual Sum of Squares (RSS) is written as follows:

$$RSS = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \quad (5)$$

*2.3.2 Random Forest model*

A random forest (RF) (Breiman, 2001) is a learning method that consists of many decision trees constructed during training time to perform classification, regression, and other tasks. In regression tasks, the average prediction from all the trees is returned.

The first step consists of fitting the random forest to the dataset $D_n$ defined as:

$$D_n = \{(x_i, y_i)\}_{i=1}^{n} \quad (6)$$

Where $\boldsymbol{x_i}$ represents the input feature vector and $\boldsymbol{y_i}$ the corresponding target value (PV power output) for $i$-th observation.

The objective function minimized for Random Forest model is the Mean Square Error (MSE) for each decision tree. This measure is utilized to assess the quality of the splits made in the tree. The objective is to identify splits that give the lowest MSE in each node of the tree.

At every node, the algorithm searches for the split point that minimizes the MSE of the child nodes created by the split:

$$MSE_{node} = \frac{1}{N_{left}}\sum_{i\in left}(y_i - \hat{y}_{left})^2 + \frac{1}{N_{right}}\sum_{i\in right}(y_i - \hat{y}_{right})^2 \quad (7)$$

Where $N_{left}$ and $N_{right}$ are the number of samples in the left and right child nodes, respectively and $\hat{y}_{left}$ and $\hat{y}_{right}$ are the mean predictions of the left and right nodes.

### *2.3.3 Support vector machines*

Support Vector Machines (SVM) is a commonly learning algorithm used to perform both classification and prediction. SVM models are particularly powerful for forecasting, demonstrating an excellent performance in predicting regression, especially in time series analysis. (Thissen *et al.*, 2003; Setiawan, Koprinska and Agelidis, 2009).

Suppose $D_n$ is training time series dataset which consist of n simples of a tuple (*xi, yi*) where *xi* is the p-dimensional vector presenting features of each data point *i*, and *yi* presents the target value for the same point. Essentially, In SVM, a nonlinear mapping $\varphi$ is used to map x into a high-dimensional feature space, and linear regression is applied to find the function *f(x)* presented by the equation 8 (Müller et al., 1997; Rana, Koprinska and Agelidis, 2016).

$$f(x) = \omega \cdot \varphi(x) + b \quad (8)$$

where: $\omega$ and $b$ are the regression parameters that present the vector of weights and the threshold, respectively. These regression parameters are determined by minimizing the function presented by equation 9:

$$minimize\ \frac{\lambda}{2}\|\omega\|^2 + C\frac{1}{n}\sum_{i=1}^{n} L(y_i, f(x_i, \omega)) \quad (9)$$

The first term of the model is weight decay, which determines the model complexity. The regularization constant $\lambda$ is used to adjust the weight sizes. In the second part of the equation, a loss function $L$ used to penalize the large errors based on how the training data well fitted by the model, depending on the tolerance of the error. $C$ is a positive constant used to control the magnitude of deviations from the tolerated error.

### *2.3.3 Artificial neural network*

Artificial neural networks are highly popular prediction algorithms. They can learn complex patterns and relationships from data, making them particularly effective at forecasting solar power. (Abuella and Chowdhury, 2015; Rodríguez *et al.*, 2018; Nasuha *et al.*, 2023).

In this study, a Multi-Layer Perceptron (MLP), one of the most commonly used ANN, is utilized. The MLP used here consists of three layers (see Fig.4) and each comprising a set of neurons. A neuron is considered the fundamental building block of various neural networks. The three layers are:

Input layer: this layer comprises three features essential for PV power output prediction, namely, solar irradiance (G), ambient temperature ($T_{amb}$), and wind speed (Ws).

Hidden layer: it processes the input data to capture complex relationship between the inputs and the output. A grid search method is employed to determine the number of the neurons in this layer.

The following expression calculates the activation of each neuron in the hidden layer (Haykin, 2009):

$$h_i = f\left(\sum_{i=1}^{n} w_{ij}x_i + b_j\right) \quad (10)$$

Where:

$w_{ij}$ : Weight associated with the input $x_i$ and hidden neuron j.

$b_j$: Bias term for neuron *j*

*f*: activation function (Rectified Linear Unit (ReLU)).

Output layer: consists of a single neuron that calculates the predicted PV power output $P_{PV}$ using the following expression:

$$P_{PV} = f\left(\sum_{j=1}^{m} w_{oj}h_j + b_o\right) \quad (11)$$

Where:

$w_{oj}$: Weight that connects hidden neuron j to the output neuron.

$b_o$: Bias term for the output neuron.

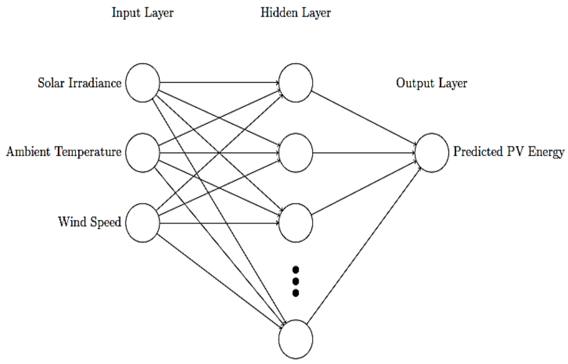


**Fig. 4** Multilayer perceptron (MLP) for PV power forecasting.

The activation function *f* in the output layer is a linear function.

The MLP model is trained using the backpropagation algorithm, which adjusts the weights of the network to minimize the error. The training process continues until either the specified number of epochs is reached or the targeted error value is achieved. The error measure used during the MLP model training is the Mean Square Error (MSE), which is calculated using the following formula:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \quad (12)$$

Where $y_i$ and $\hat{y}_i$ represent the actual and predicted PV power outputs, respectively. N: the total number of observations in data set.

## 3. Cross-validation technique

Cross-validation (CV) is one of the most widely techniques used in machine learning and statistical analysis for evaluating model generalization performance and selecting the best model based on performance metrics (Bergmeir and Benítez, 2012). Cross-validation, involves splitting the dataset several times into equal parts, with the major portion used to train models and the rest used to validate them. A k-fold CV is the common cross-validation method, where k represents the number of folds, usually 5 or 10 folds. (Deng, 2023). In 10-folds cross-validation,

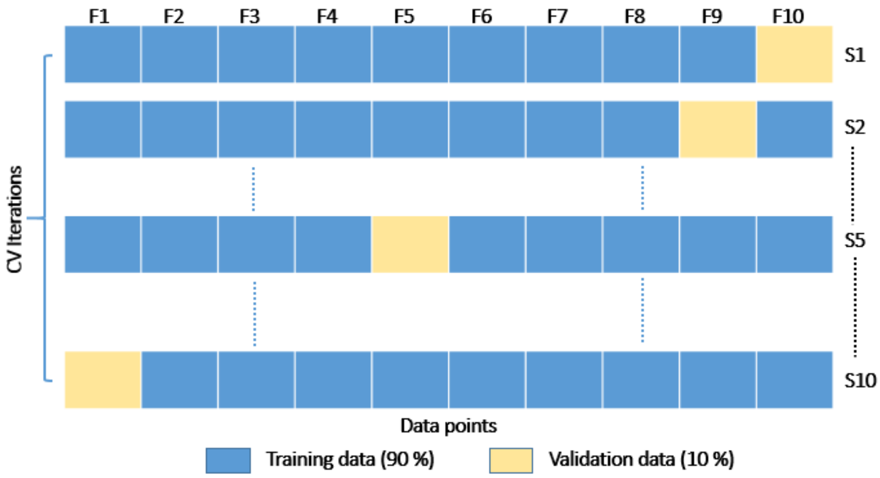


**Fig. 5** 10-fold Cross-Validation process.

the data is partitioned into ten equal sized splits, called folds (F) as presented in Fig. 5. For each model, nine folds (90%) are used to train models and one-fold (10%) is utilized for validation (the process is repeated ten times). In the next step, an evaluation of the performance of each model based on metrics such as accuracy rate or MSE is performed. A model final performance is typically reported as the average of its 10-fold performance metrics.

## 4. Model performance metrics

In this paper, an assessment of model performance based on various accuracy metrics is carried out. Metrics such as MAE and MRE are used to quantify the average magnitude of prediction errors. Additionally, metrics such as RMSE and coefficient of determination are used to calculate the linear fit of regression models.

*2.1 4.1 The prediction errors*

Prediction errors, also called residuals in a statistical or machine learning model, present the difference between a predicted and the observed value across all data points. In other words, the residual measures the error in predictions made by the model. The residual is a diagnostic metric used to assess quality of a model. This error is expressed using the equation 13 (Botchkarev, 2018).

$$Error_i = \hat{Y}_i - Y_i \quad (13)$$

*4.2 The mean absolute error*

Mean absolute error is a metric used to measure the average of absolute error values. Errors in MAE are not weighted, but scores increase linearly when errors increase. MAE is calculated using the equation (14) (Qi *et al.*, 2020) (Botchkarev, 2018):

$$MAE = \frac{1}{n} \times \sum_{i=1}^{n} |Error_i| \quad (14)$$

*4.3 The mean relative error*

The mean relative error is an extension of MAE, which calculates the percentage error of a prediction. MRE is defined as the ratio between the MAE and the amplitude of the reference E-field in the corresponding target region. The MRE is calculated by the equation (15) (Botchkarev, 2018):

$$MRE = 100 \times \sum_{i=1}^{n} \frac{|Error_i|}{Y_i} \quad (15)$$

*4.4 The root mean squared error*

Root Mean Squared Error (RMSE) is a common metric used to calculate the average squared of the prediction errors in a regression problem. It quantifies the average magnitude of the error, highlighting the discrepancy between predicted values and actual values. By squaring the error values, RMSE penalizes larger errors more than smaller errors, which has a greater impact on the overall score. The RMSE calculated using the formula given by equation (16) (Botchkarev, 2018):

$$RMSE = \sqrt{\frac{\sum (Error_i)^2}{n}} \quad (16)$$

*4.5 The coefficient of determination*

In regression analysis, the coefficient of determination $R^2$, is a key measure for assessing the quality of a regression model. $R^2$ is used to quantify the proportion of variation in the dependent variable that is explained by the predictor variables included in the regression model. It is given by the equation (17) (Ozer, 1985):

$$R^2 = 1 - \sum_{i=1}^{n} \frac{(Error_i)^2}{(Y_i - \bar{Y})^2} \quad (17)$$

## 5. Results and discussion

This section is dedicated to present the PV power output forecasting results using the learning models, namely ANN, LR, RF and SVM. Different simulations based on the two datasets were performed, in order to assess the performance of each model and to examine the impact of the input variables (features) on the model accuracy. The obtained results from the models are compared and evaluated using various performance metrics. The results of training algorithms were implemented in Python 3.11 environment and performed using a personal computer with the following characteristics: windows 10 OS

https://doi.org/10.61435/ijred.2025.60547

**Table 1**
PV power output forecasting results.

| Parameters used | Prediction algorithm | MAE (W) | MRE (%) | RMSE (W) | $R^2$ |
|---|---|---|---|---|---|
| Weather forecast (GHI, temperature, wind speed) | LR | 10.299 | 19.541 | 14.085 | 0.99683 |
| (Time step: 1 h) | RF | **0.364** | **0.256** | **0.856** | **0.99999** |
| Data points (n = 8760) | SVM | 7.411 | 9.046 | 20.056 | 0.99357 |
| | ANN | 0.651 | 0.889 | 1.310 | 0.99997 |
| On-site measurements (GHI, temperature, wind speed) | LR | 16.596 | 104.824 | 29.353 | 0.99077 |
| (Time step: 5 min) | RF | **3.922** | **11.163** | **8.525** | **0.99922** |
| Data points (n = 8640) | SVM | 14.193 | 57.446 | 26.655 | 0.99239 |
| | ANN | 10.822 | 30.326 | 22.256 | 0.99464 |

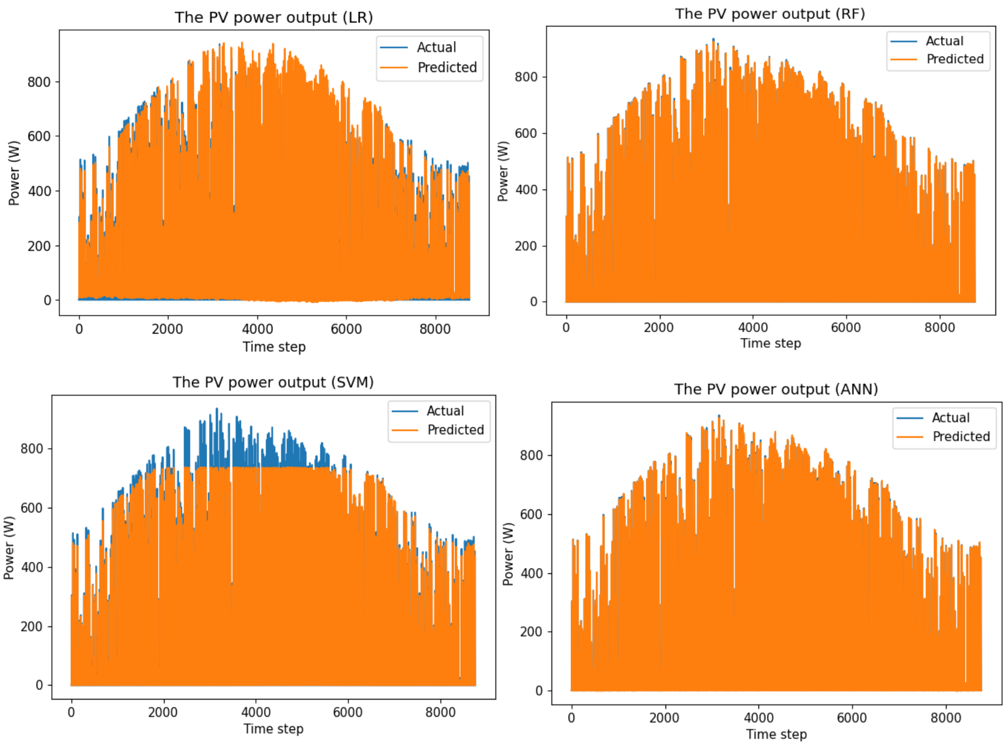


**Fig. 6** PV power output (weather forecast)

with 64 bits; a processor Intel(R) Core (TM) i5-6200U CPU @ 2.30GHz 2.40 GHz; and 8.00 GB of RAM. The forecast results obtained for both datasets are summarized in Table 1. The $R^2$ reports how well the model explains the variation in the response in both datasets, while the metrics RMSE, MRE and MAE asses how accurately a regression model can predict the values of a response variable.

Based on the results of Table 1, it is notably, when models trained on weather forecast dataset, the RF recorded the lowest values for MAE, MRE and RMSE equal to 0.364 W, 0.256% and 0.856 W, respectively. Compared to ANN, which gave MAE, MRE and RMSE values equal to 0.651 W, 0.889% and 1.310 W and LR model that recorded the values of MAE, MRE and RMSE equal to 10.299 W, 19.541% and 14.085 W, respectively. Lastly, the SVM gave the worst RMSE value at 20.056 W, even though its other MAE and MRE values were lower than those of the LR model at 7.411W and 9.046%, respectively. This was because the SVM model exhibited a large gap between the actual and predicted values (discussed above). As a result, the RMSE penalizes them more harshly than MAE and MRE. When the models were trained on on-site measurements, a significant increase in RMSE, MRE and MAE values was recorded. Furthermore, there was a convergence in the values recorded for the different models trained on on-site measurements dataset, except that the RF model, which yielded the lowest values of MAE, MRE and RMSE at 3.922 W, 11.163 % and 8.825 W, respectively. For the ANN model, there was a significant increase in MAE, MRE and RMSE values to 10.822 W, 30.326 % and 22.256 W. Similar observations were made for the MAE and RMSE indicators for the LR model, which recorded values of 16.596 W and 29.353 W, respectively, with a significant increase in MRE, reaching 104.824%, because the linear regression model struggled to capture the complex patterns present in the on-site data. In contrast, the MAE, MRE, and RMSE values of the SVM model have increased slightly and outperformed the LR model in this case with 10.822 W, 57.446%, and 26.655 W respectively.

Fig.6 shows a comparison between the actual and predicted PV power output of each model using a weather forecast

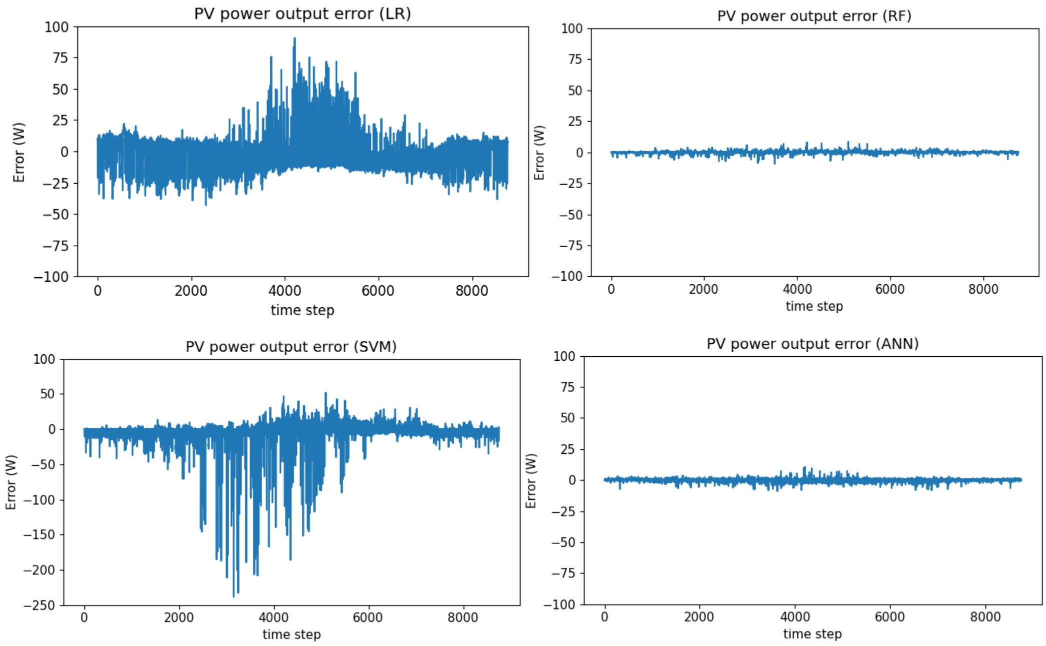


**Fig. 7** PV energy production forecast error (weather forecast).

dataset. While Fig.7 shows the comparison of PV power output forecast errors of different models based on the input meteorological dataset. Based on Fig. 6 and Fig.7, it is clear that the RF model accurately predicted PV power output with a minimum error that did not exceed 10 W compared to other models. Following RF is the ANN model, which has an absolute error range between 0 and 15 W then the LR model with an error reaching 90 W. Finally, the SVM model presents the least accuracy in PV power output forecasting with a forecasting error greater than 240 W. Moreover, forecasting errors increased between time steps 2000 and 6000. This period represents the period of high PV power output.

Fig.8 shows a comparison between the actual and predicted PV power of each model using the on-site measurements dataset. On the other hand, Fig.9 illustrates the comparison of PV power output forecast errors for the on-site measurement's dataset. According to the figures, the RF model presents the most accurate model in PV power output forecasting. Moreover, as can be seen, forecasting errors have almost the same

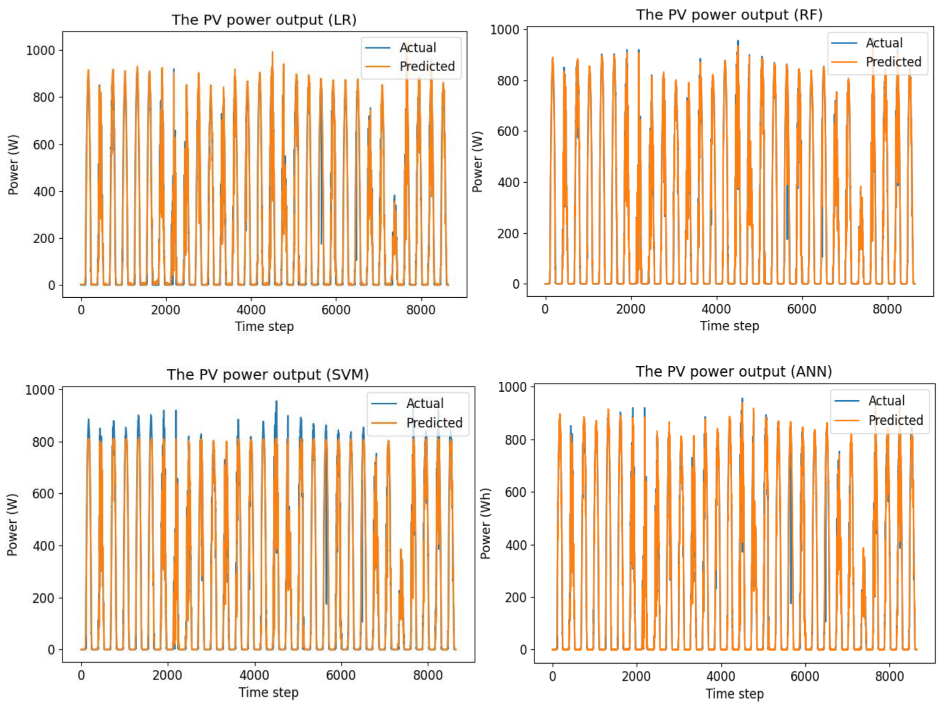


**Fig. 8** PV power output (On-site measurements).

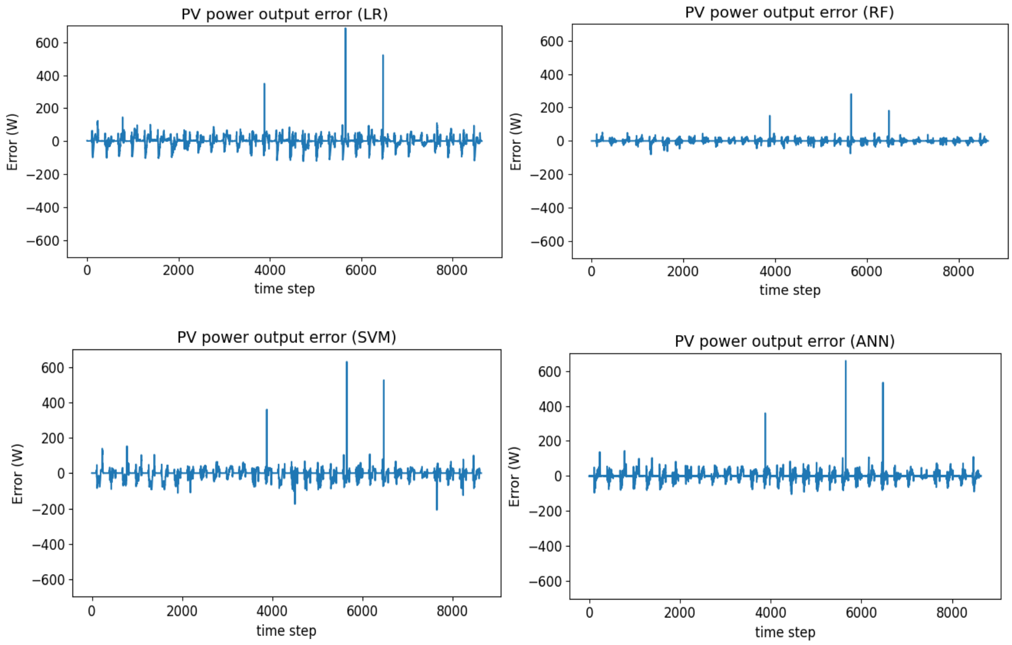


**Fig. 9** PV power output forecast error (On-site measurements).

amplitude during all time steps range between [-200,200], except for several sudden drops in power output were recorded at time steps 3882, 5656 and 6478. These errors represent a sudden change in the output power that usually occurs due to the shading effect (discussed previously. Fig.10 shows a zoom in view on the results for the three points that represent these errors in each regression model. As can be seen, the RF model predicted with high accuracy and responded well to sudden changes in output power. In contrast, the other models could not able to predict these changes in power output.

Fig.11 presents a comparison of MAE, MRE, and RMSE accuracy metrics for both datasets for each model, which aims to quantify forecasting errors, and evaluate the model's performance. These metrics showed low errors in the weather forecast dataset compared to the on-site measurement dataset. Additionally, when the models were trained on weather forecast datasets, there was a clear variation in the values of different models. Additional validation results, including detailed forecasts based on these datasets, are included in Appendix A.

According to the obtained results, the choice of data source (weather forecast or on-site measurements) and time step (5-minute or 1-hour step) visibly affects the models' performance, with the former (weather forecast with a 1-hour step) resulting in more accurate prediction models. This may be attributed to the high variability characterizing the 5-minute time step data and the challenge it poses in capturing these variations compared to the 1-hour step dataset (sensitivity to data variability). Moreover, the number of features used to train the models influence forecast accuracy. As a result, the on-site PV power output could be affected by various variables other than GHI, Tamb and Ws (e.g., shading effect, air pressure, panel degradation, etc.). In contrast, in the weather forecast dataset, the model was based on only three variables (GHI, Tamb and Ws).

In addition, across both datasets (weather forecast and on-site measurements), RF consistently outperforms other prediction models by obtaining minimal error in MAE, MRE, and RMSE and with near-perfect $R^2$ values which indicates its ability to capture complex and non-linear relationships with datasets (especially when dealing with highly variable weather data). Furthermore, ANN achieved the second-best results for both datasets, which is not surprising, giving its ability to approximate complex non-linear relationships. SVM has generally obtained comparable results to LR but the latter has clear advantage that can be crucial in certain scenarios: it is simple and understandable. In cases were understanding the reasoning behind the decision of prediction models is important, LR can play a significant role in providing interpretable predictions.

Lastly, k-fold cross-validation was critical in avoiding overfitting and providing a more reliable assessment of the generalization performance of the models, especially with the on-site measurement dataset, which is characterized by its nonlinearity, complexity, and high variability. Cross-validation allowed the model to be trained and validated across multiple subsets of data given the unpredictable nature of on-site measurements, which include sudden fluctuations in weather conditions and local environmental factors. This process helped ensure that the models did not overfit any particular part of the dataset, especially in such a complex and variable environment. As a result of splitting the data into multiple folds, bias and variance are reduced, ensuring that the model can generalize to new, unseen data, even with complex on-site data.

## 6. Conclusion

In this paper, an assessment of the accuracy of power output forecasting in PV systems using ML techniques such as LR, RF, SVM and ANN was proposed. To compare the performance of the proposed forecasting models, two different datasets with different time steps were used for training the learning algorithms; one consisted of historical forecasted weather while the other consists of on-site measurements. The performance evaluation of the proposed forecasting models was based on various metrics, namely MAE, MRE, RMSE and R-squared. After conducting a comparative analysis, and evaluating of the models' performances, the following conclusions are drawn:

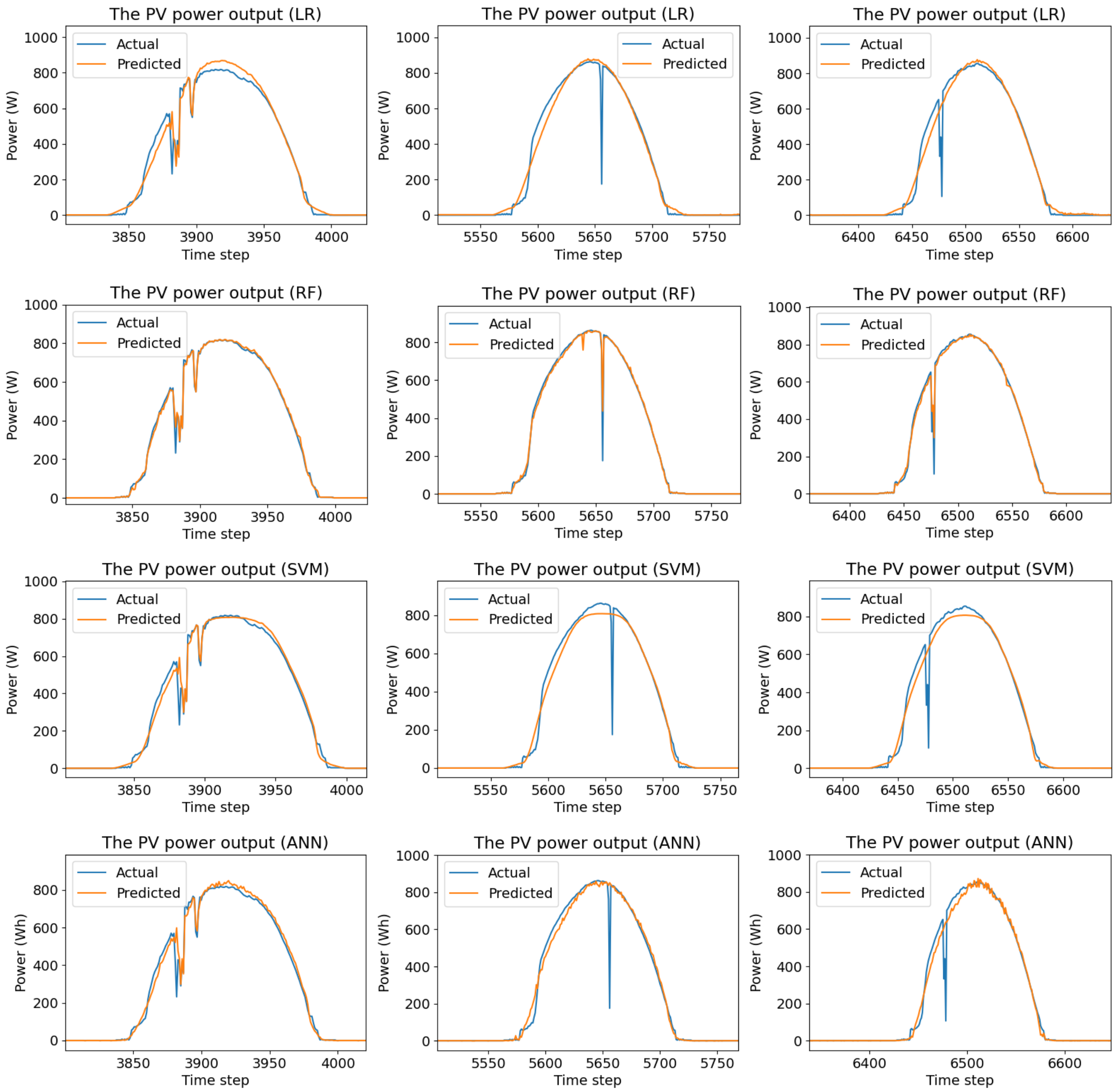

**Fig. 10** Illustration of PV power output drops in on-site measurements.

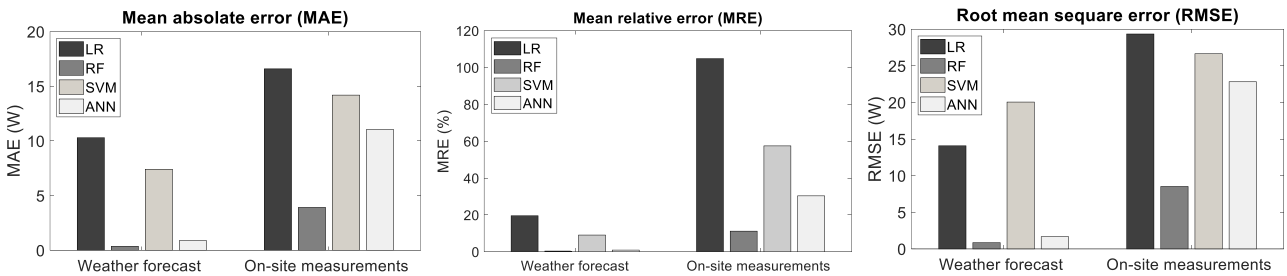

**Fig. 11** the MAE, MRE and RMSE forecasting performance metrics of models for different dataset.

– Dataset characteristics influenced model performance. As a result, models trained on historical weather forecasts performed better than models trained on on-site measurements reflected the specifics of their data source. This discrepancy may be due to inherent differences between datasets, including variations in the time intervals, data collection methods, and perhaps location-specific factors.

- In the weather forecast dataset, error values were notably higher during periods of high PV power output. Conversely, in the on-site measurements dataset, errors were more than three times higher during periods of output power drops. Thus, forecast accuracy is affected by the power output rate and the behavior of the PV system. Therefore, it is recommended to provide more information about PV systems and to take into account additional input variables (such as the shading effects, PV system degradation, etc.) to improve forecast accuracy.
- In comparison to the other proposed models, the RF algorithm demonstrated robustness in PV power output prediction accuracy by obtaining the minimal errors. In addition, the model exhibited a high degree of stability in its performance, and its ability to handle the non-linearities and complexities of the data particularly in adapting to drops in PV power output when trained on on-site measurements dataset.

**Nomenclature**

*Abbreviations*

| | |
|---|---|
| AI | Artificial intelligence |
| ANN | Artificial neural networks |
| CV | Cross validation |
| GHI | Global horizontal irradiance |
| LR | Linear regression |
| MAE | Mean absolute error |
| ML | Machine learning |
| MRE | Mean relative error |
| PHANN | Physical Hybrid Artificial Neural Network |
| RF | Random forest |
| RMSE | Root mean squared error |
| SVM | Support victor machine |

*Variables*

| | |
|---|---|
| $G$ | Solar irradiance (W/m²) |
| $T_{amb}$ | Ambient temperature (C°) |
| $W_S$ | Wind speed (m/s) |
| $P_{PV}$ | actual PV power output (kW) |
| $\hat{P}_{PV}$ | predicted PV power output (kW) |
| $X$ | input features of learning model |
| $Y$ | Output target value of learning model |
| $\hat{Y}$ | Predicted value of learning model |
| $T_c$ | panel temperature (C°) |
| $N$ | number of PV panel |
| $I$ | sample or data point index |
| $N$ | total number of samples or data points |

**Author Contributions**: A. Bouakkaz.: Conceptualization, methodology, formal analysis, writing—original draft, A. Lahsasna.; supervision, project administration, validation, A. J Gil Mena; writing—review and editing, project administration, validation, S. Haddad.; writing—review and editing, project administration, validation, M. Luigi Ferrari.: writing—review and editing, project administration, validation, R. Jiménez- Castaneda.: writing—review and editing, project administration, validation, resources. All authors have read and agreed to the published version of the manuscript.

**Funding:** The authors received no financial support for the research, authorship, and/or publication of this article.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

Abuella, M. and Chowdhury, B. (2015) Solar power forecasting using artificial neural networks, in 2015 North American Power Symposium (NAPS), pp. 1–5. https://doi.org/10.1109/NAPS.2015.7335176.

Antonanzas, J., Osorio, N., Escobar, R., Urraca, R., Martinez-de-Pison, F.J. & Antonanzas-Torres, F. (2016) Review of photovoltaic power forecasting, *Solar Energy,* 136, 78–111. https://doi.org/10.1016/j.solener.2016.06.069.

Arutyunov, V. S. and Lisichkin, G. V (2017) Energy resources of the 21st century: problems and forecasts. Can renewable energy sources replace fossil fuels?, *Russian Chemical Reviews*, 86(8), 777–804. https://doi.org/10.1070/rcr4723.

Bergmeir, C. and Benítez, J. M. (2012) On the use of cross-validation for time series predictor evaluation, *Information Sciences*, 191, 192–213. https://doi.org/10.1016/j.ins.2011.12.028.

Botchkarev, A. (2018) Evaluating Performance of Regression Machine Learning Models Using Multiple Error Metrics in Azure Machine Learning Studio, *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3177507.

Breiman, L. E. O. (2001) Random Forests, 45, pp. 5–32.

Chen, J., Xiao, G., Ferrari, M. L., Yang, T., Ni, M., & Cen, K. (2020) Dynamic simulation of a solar-hybrid microturbine system with experimental validation of main parts, *Renewable Energy,* 154, 187–200. https://doi.org/10.1016/j.renene.2019.11.022.

Chenni, R., Makhlouf, M., Kerbache, T., & Bouzid, A. (2007) A detailed modeling method for photovoltaic cells, *Energy,* 32(9), pp. 1724–1730. https://doi.org/10.1016/j.energy.2006.12.006.

Deng, A. (2023) Time series cross validation: A theoretical result and finite sample performance, *Economics Letters*, 233(January). https://doi.org/10.1016/j.econlet.2023.111369.

Dimd, B. D., Völler, S., Cali, U., & Midtgård, O.-M. (2022) A Review of Machine Learning-Based Photovoltaic Output Power Forecasting: Nordic Context, *IEEE Access,* 10, 26404–26425. https://doi.org/10.1109/ACCESS.2022.3156942.

Gandelli, A., Grimaccia, F., Leva, S., Mussetta, M., & Ogliari, E. (2014) Hybrid model analysis and validation for PV energy production forecasting, Proceedings of the International Joint Conference on Neural Networks, pp. 1957–1962. https://doi.org/10.1109/IJCNN.2014.6889786.

De Giorgi, M. G., Congedo, P. M. and Malvoni, M. (2014) Photovoltaic power forecasting using statistical methods: Impact of weather data, *IET Science, Measurement and Technology*, 8(3), 90–97. https://doi.org/10.1049/iet-smt.2013.0135.

Haykin, S. (2009) Neural Networks and Learning Machines. 3rd edn. Edited by M. J. Horton. Ontario: Pearson.

Henderson, J. and Sen, A. (2021) The Energy Transition: Key challenges for incumbent and new players in the global energy system, *The Oxford Institute for Energy Studies*. http://hdl.handle.net/10419/246577.

Hope, T. M. H. (2020) Linear regression, pp. 67–81. https://doi.org/10.1016/B978-0-12-815739-8.00004-3.

Hoppe, T. and van Bueren, E. (2015) Guest editorial: governing the challenges of climate change and energy transition in cities, *Energy, Sustainability and Society*, 5(1). https://doi.org/10.1186/s13705-015-0047-7.

IEA (2023) Tracking Clean Energy Progress 2023, IEA. Available at: https://www.iea.org/reports/tracking-clean-energy-progress-2023, (Accessed: 3 December 2023).

Iheanetu, K. J. (2022) Solar Photovoltaic Power Forecasting: A Review, *Sustainability* (Switzerland), 14(24). https://doi.org/10.3390/su142417005.

Khadke, S., Ramasubramanian, B., Paul, P., Lawaniya, R., Dawn, S., Chakraborty, A., Mandal, B., Dalapati, G. K., Kumar, A., & Ramakrishna, S. (2023) Predicting Active Solar Power with Machine Learning and Weather Data, *Materials Circular Economy*, 5(1), 15. https://doi.org/10.1007/s42824-023-00087-5.

Leva, S., Dolara, A., Grimaccia, F., Mussetta, M., & Ogliari, E. (2017) Analysis and validation of 24 hours ahead neural network forecasting of photovoltaic output power, *Mathematics and Computers in Simulation,* 131, 88–100. https://doi.org/10.1016/j.matcom.2015.05.010.

Li, P., Zhou, K. and Yang, S. (2018) Photovoltaic Power Forecasting: Models and Methods, in 2018 2nd IEEE Conference on Energy Internet and Energy System Integration (EI2), pp. 1–6. https://doi.org/10.1109/EI2.2018.8582674.

Liu, W., Liu, C., Lin, Y., Ma, L., Xiong, F., & Li, J. (2018) Ultra-short-term forecast of photovoltaic output power under fog and haze weather, *Energies,* 11(3). https://doi.org/10.3390/en11030528.

McCormick, P. G. and Suehrcke, H. (2018) The effect of intermittent solar radiation on the performance of PV systems, *Solar Energy*, 171, pp. 667–674. https://doi.org/10.1016/j.solener.2018.06.043.

Mellit, A., Massi Pavan, A., Ogliari, E., Leva, S., & Lughi, V. (2020) Advanced methods for photovoltaic output power forecasting: A review, *Applied Sciences* (Switzerland), 10(2). https://doi.org/10.3390/app10020487.

Müller, K.R., Smola, A.J., Rätsch, G., Schölkopf, B., Kohlmorgen, J., Vapnik, V. (1997). Predicting time series with support vector machines. In: Gerstner, W., Germond, A., Hasler, M., Nicoud, JD. (eds) Artificial Neural Networks — ICANN'97. ICANN 1997. Lecture Notes in Computer Science, vol 1327. Springer, Berlin, Heidelberg. https://doi.org/10.1007/BFb0020283 .

Amer, H. N., Dahlan, N. Y., Azmi, A. M., Latip, M. F. A., Onn, M. S., & Tumian, A. (2023) Solar power prediction based on Artificial Neural Network guided by feature selection for Large-scale Solar Photovoltaic Plant, *Energy Reports*, 9, 262–266. https://doi.org/10.1016/j.egyr.2023.09.141

Notton, G., Nivet, M.-L., Voyant, C., Paoli, C., Darras, C., Motte, F., & Fouilloy, A. (2018) Intermittent and stochastic character of renewable energy sources: Consequences, cost of intermittence and benefit of forecasting, *Renewable and Sustainable Energy Reviews*. 87, 96–105. https://doi.org/10.1016/j.rser.2018.02.007.

NREL (2023) NSRDB: National Solar Radiation Database. Available at: https://nsrdb.nrel.gov/data-viewer.

Ozer, D. J. (1985) Correlation and the Coefficient of Determination, *Psychological Bulletin*, 97(2), 307–315. https://doi.org/10.1037/0033-2909.97.2.307.

Qi, J., Du, J., Siniscalchi, S. M., Ma, X., & Lee, C.-H. (2020) On Mean Absolute Error for Deep Neural Network Based Vector-to-Vector Regression, *IEEE Signal Processing Letters*, 27, 1485–1489. https://doi.org/10.1109/LSP.2020.3016837.

Rana, M., Koprinska, I. and Agelidis, V. G. (2016) Univariate and multivariate methods for very short-term solar photovoltaic power forecasting, *Energy Conversion and Management*. 121, 380–390. https://doi.org/10.1016/j.enconman.2016.05.025.

Rodríguez, F., Fleetwood, A., Galarza, A., & Fontán, L. (2018) Predicting solar energy generation through artificial neural networks using weather forecasts for microgrid control, *Renewable Energy*, 126, 855–864. doi: https://doi.org/10.1016/j.renene.2018.03.070.

Setiawan, A., Koprinska, I. and Agelidis, V. G. (2009) Very short-term electricity load demand forecasting using support vector regression, in 2009 International Joint Conference on Neural Networks, pp. 2888–2894. https://doi.org/10.1109/IJCNN.2009.5179063.

TamizhMani, G., Ji, L., Tang, Y., & Petacci, L. (2003) Photovoltaic Module Thermal/Wind Performance : Long-Term Monitoring and Model Development For Energy Rating, NCPV and Solar Program Review Meeting, (May), 936–939.

Tarroja, B., Mueller, F. and Samuelsen, S. (2013) Solar power variability and spatial diversification: implications from an electric grid load balancing perspective, *International Journal of Energy Research*, 37(9), 1002–1016. https://doi.org/10.1002/er.2903.

Thissen, U., van Brakel, R., de Weijer, A. P., Melssen, W. J., & Buydens, L. M. C. (2003) Using support vector machines for time series prediction, *Chemometrics and Intelligent Laboratory Systems*, 69(1), 35–49. https://doi.org/10.1016/S0169-7439(03)00111-4.

Yin, J., Molini, A. and Porporato, A. (2020) Impacts of solar intermittency on future photovoltaic reliability, Nature Communications. Springer US, 11(1), pp. 1–9. https://doi.org/10.1038/s41467-020-18602-6.

Yona, A., Senjyu, T., Funabashi, T., & Kim, C.-H. (2013) 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), 527–533. https://doi.org/10.1109/TSTE.2013.2246591.

Zack, J. W. (2017) Wind and Solar Forecasting. USA: Springer. https://doi.org/10.1007/978-3-319-55581-2_4

Zhou, Y., Zhou, N., Gong, L., & Jiang, M. (2020) Prediction of photovoltaic power output based on similar day analysis, genetic algorithm and extreme learning machine, *Energy,* 204, 117894. https://doi.org/10.1016/j.energy.2020.117894.

**Appendix A.** PV power forecasts validation using a new dataset

This appendix provides a validation of the forecasting models used in this study, utilizing a newly acquired dataset. The dataset includes one month of weather forecast data and four days of on-site measurement data, comprising solar irradiance (G), temperature (T), and wind speed (Ws). These key variables are critical for accurately modeling the factors affecting PV power output.

Figure A.1 displays the PV power output forecasts based on the one-month weather forecast dataset, while Figure A.2 shows forecasts using the more detailed four-day on-site measurement dataset, also containing G, T, and Ws. This second figure highlights how the models respond to short-term fluctuations in weather conditions, providing insights into their ability to capture variations in power output.

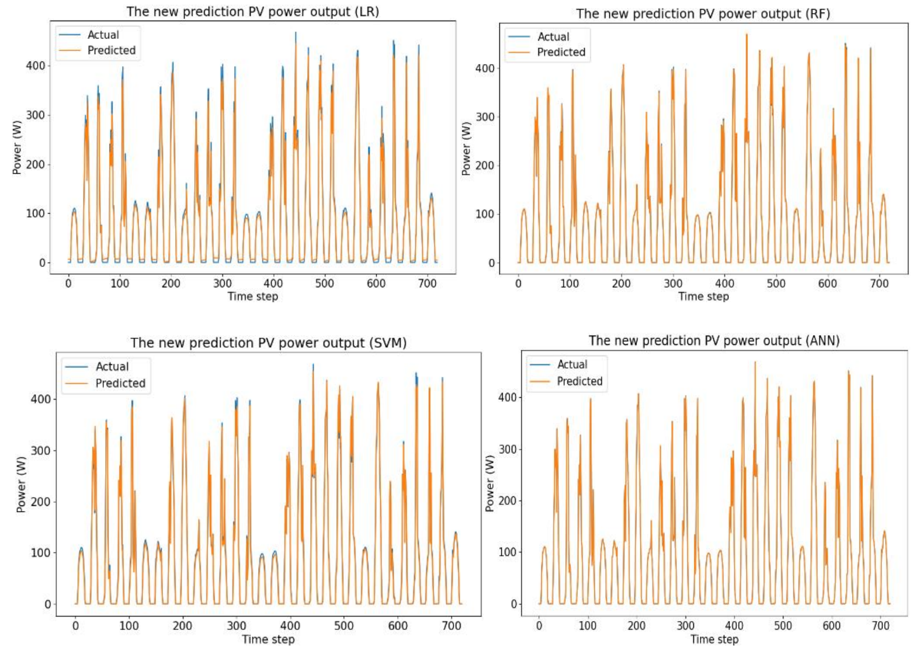


**Fig. A.1** PV power output forecast (weather forecast: one month).

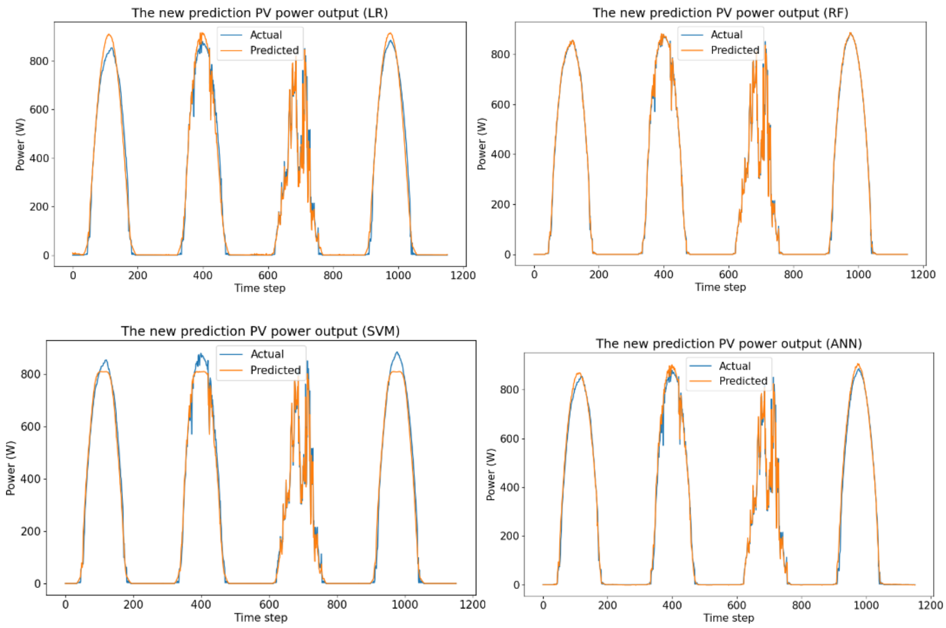


**Fig. A.2** PV power output forecast (on-site measurements: 4 days).