Chapter

Photovoltaic Power Forecasting Methods

Ismail Kaaya and Julián Ascencio-Vásquez

Abstract

The rapid growth in grid penetration of photovoltaic (PV) calls for more accurate methods to forecast the performance and reliability of PV. Several methods have been proposed to forecast the PV power generation at different temporal horizons. In this chapter the different methods used in PV power forecasting are described with an example on their applications and related uncertainty. The methods discussed include physical, heuristic, statistical and machine learning methods. When benchmarked, it is shown that physical method showed the highest uncertainties compared to other methods. In the chapter, the effect of degradation on lifetime PV power and energy forecast is also assessed using linear and non-linear degradation scenarios. It is shown that the relative difference in lifetime yield prediction is over 5% between linear and non-linear scenarios.

Keywords: Degradation, Lifetime, Photovoltaic, Power, Forecasting

1. Introduction

The current trends in photovoltaic (PV) deployments worldwide are a clear indication that PV energy will play a big role in the near future energy mix. For example, the global solar photovoltaic (PV) capacity is projected to increase by 37.5% from 2019 to 2030 (i.e from 593.9 GW in 2019 to 1,582.9 in 2030) [1]. This rapid growth in grid penetration of PV calls for more accurate methods to forecast the performance and reliability of PV. Additionally, PV projects policy and investment decisions rely on PV performance and reliability forecasts. Therefore, to reduce the risks of PV investments, more reliable methods to forecast the performance and reliability of PV power are a prerequisite.

Different methods have been proposed for PV power forecasting. These methods can be classified as: physical, heuristic, statistical and machine learning methods [2, 3]. Each method might have different conceptual design, implementation, application and accuracy. In this chapter, the application and accuracy of the different methods are assessed using measured PV module power and weather data.

PV power forecasting can range from different temporal horizons depending on ones need. At a moment there is no standard classification criterion of the temporal horizon. General classification can be made as: very short term (Intra-hour: 15 minutes to 2 hours ahead), short-term (hour ahead: 1 to 6 hours ahead, day ahead: 1–3 days ahead), Medium-term (week to months ahead), long-term (one to several years) and lifetime forecast (until PV expected lifetime).

Achieving high-accuracy forecasts at each of these temporal horizons is influenced by different variables such as: solar radiation models, PV performance

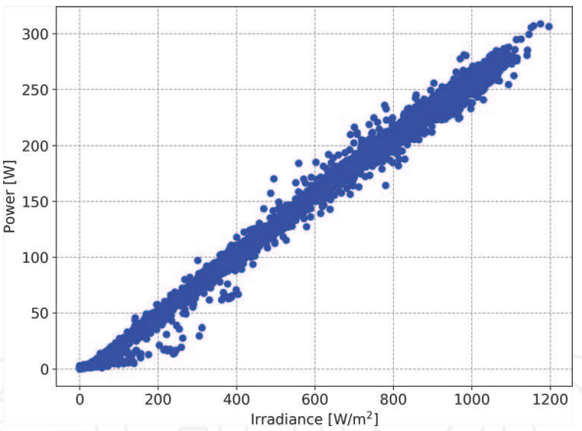

Figure 1.
Measured irradiance versus measured power. Data corresponds to six month measurements of irradiance and module power in Gran Canaria (Spain).

models, data availability, data quality and forecasting methods. The accuracy highly deteriorates with increasing forecasting temporal horizon. This is because, more input parameters such as: seasonal variations, soiling effects, degradation and many other performance reducing effects need to be taken into consideration [4]. These factors are not easy to evaluate precisely, therefore, they are simply approximated which increase the uncertainties in long-term and lifetime PV forecast.

The main influencing factor of PV production is the amount of global solar irradiation incident on the PV panels. As shown in **Figure 1**, there is a quasi-linear correlation of power and irradiance. This property means that the accuracy of power prediction is highly determined by the accuracy of the solar irradiation forecast.

In this chapter, the different methods used in PV power forecasting are presented. The chapter is organised as follows: In Section 1, a brief introduction on power forecasting is presented. In Section 2, different PV forecasting methods are presented, for some methods a practical example of their application and their accuracy are evaluated using real measured PV module power. Section 3, is dedicated to lifetime PV power forecasting. In this section, several effects affecting lifetime PV power forecasting are stated and a more elaborative discussion of the degradation effect on lifetime PV power forecasting is presented. Lastly, in Section 4, a summary of the different aspects within the chapter is presented.

2. PV power forecasting methods

Different methods: physical, heuristic, statistical and machine learning are commonly used in PV power forecasting [2, 3]. The methods are based on two main approaches to generate the PV power forecasting. One is the physical approach, which requires prior knowledge of PV material properties and the metadata of a PV system, together with the need of weather data. The second ones is the data-driven approach, which requires operational data to train/calibrate coefficients of the models which are then used to generate the predictions. This means that, a data-driven approach can only be applied after a given PV module or system has been exposed and enough data is available to train/calibrate the models. On-contrary, a physical approach can be applied even when the PV system is not yet commissioned. Hence PV power forecasting methods based on a physical approach are the mostly used methods by PV stakeholders to evaluate the economic viability of PV projects during the initial phases.

Figure 2 illustrates the required inputs and general steps to generate PV power forecast by the two approaches. What is clear is that both approaches require weather data (mainly solar irradiation) as input. Therefore, solar irradiation forecast is highly essential step for PV power forecast using both approaches. Unlike data-driven approach, physical approach is based on physical assumptions and therefore, knowledge of the physical parameters influencing PV generation is required.

2.1 Physical method

Physical models calculate the PV power using the equivalent electrical circuits [5]. The equivalent circuit model developed for a single cell can be used to derive equivalent circuit models for a PV module as well as a PV system [6]. They are the commonly applied models in the PV power forecasting commercial software packages (such as PVSyst [7] PVWATTS [8]).

2.1.1 PV cell and module equivalent circuit models

To build the physical model one need to know the basic photo-to-voltage theory. The diode model is used to develop the PV cell model to calculate the PV output power. The diode model can be characterized as: one-, two- and three-diode models [5, 7] (see **Figure 3**). The choice of the model selection depend on charge carrier recombination mechanisms one need to take into consideration. Because of its simplest, the one-diode model is the most commonly used to model the operating principles of a PV cell. The one-diode model can consist of three, four or five parameters (see **Table 1**) .

The three parameter (3-P) one-diode model is only used to demonstrate the basic working principles of a PV cell but not as a representative of the real operating condition. To simulate the actual working conditions of a PV cell, the five parameter (5-P) one-diode model is commonly used because it takes into account both the series and shunt resistive losses. It is expressed as [9, 10]:

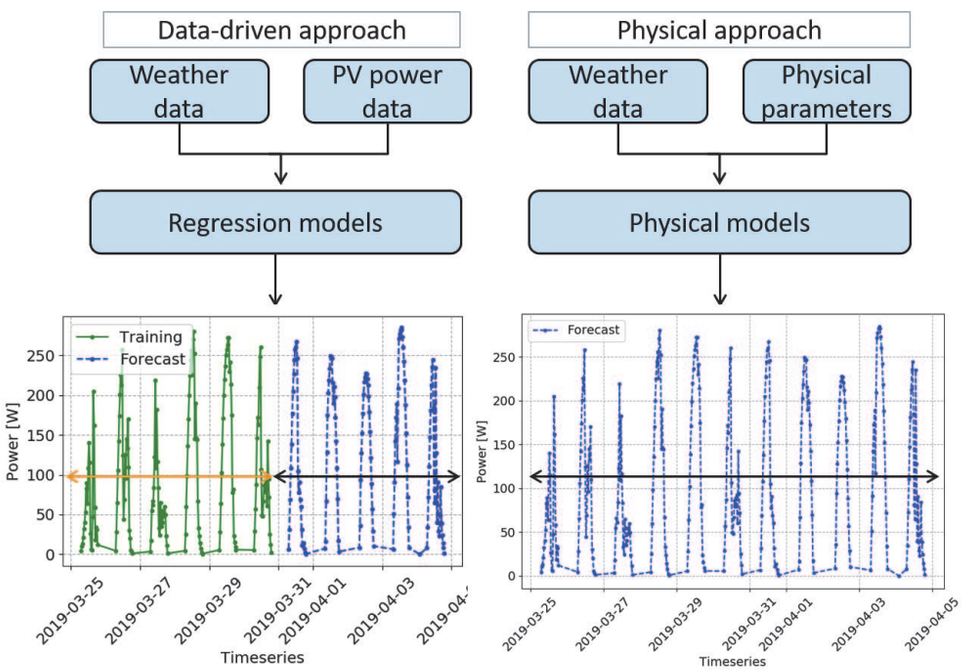

Figure 2.
Schematic of data-driven and physical approaches for PV power prediction. Data-driven approach requires historical data (in green) to train and physical approach requires physical parameters as inputs.

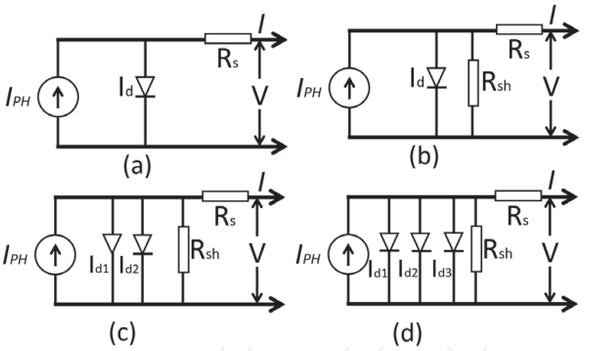

Figure 3.
a, four parameters (4-p) one-diode model. b, five parameters (5-p) one-diode model. c, two-diode model (seven parameters) and d, three diode model (nine parameters).

Model	Parameters	Characteristic
3-p One-diode model	I_{PH} [a], I_{01} [b], n_1 [c]	Basic model No series and shunt resistive losses
4-p One-diode model	I_{PH}, I_{01}, n_1, R_s [d]	Includes series losses No shunt losses
5-p One-diode model	I_{PH}, I_{01}, n_1, R_s, R_sh [e]	Includes both series and shunt shunt resistive losses
Two-diode model	I_{PH}, I_{01}, I_{02}, n_1 ,n_2, R_s, R_sh	Two diodes to represent junction recombination Relevant at low irradiance operation of a PV cell
Three-diode model	I_{PH}, I_{01}, I_{02},I_{03}, n_1,n_2,n_3, R_s, R_sh	Takes into account grain boundaries and leakage current

[a] *Photocurrent.*
[b] *Reverse saturation current, the subscripts (1, 2 and 3) represents the diode number respectively.*
[c] *diode ideality factor the subscripts (1, 2 and 3) represents the diode number respectively.*
[d] *Series resistance.*
[e] *Shunt resistance.*

Table 1.
Parameters and characteristics of different diode models.

$$I = I_{PH} - I_0\left[\exp\left(\frac{V + R_sI}{nV_t}\right) - 1 - \left(\frac{V + R_sI}{R_{sh}}\right)\right] \quad (1)$$

where I, I_o and I_{PH} are the generated solar cell current, reverse saturation current and photo-generated current respectively. V, R_s and R_{sh} are the solar cell voltage, series resistance and shunt resistance respectively. n is the ideality factor of the diode, and V_t is the thermal voltage.

To derive the equivalent circuit model for a PV module, the basic assumption that a PV module comprises of identical PV cells in series is usually taken [6]. This assumption implies that under similar conditions (irradiance and temperature), all the PV cells should generate equal current and voltage. According to [6], the PV module equivalent model is derived from a PV cell diode model (Eq. (1)) as:

$$I_M = I_{PH} - I_0\left[\exp\left(\frac{V_M + R_sN_sI_M}{nN_sV_t}\right) - 1 - \left(\frac{V_M + R_sN_sI_M}{N_sR_{sh}}\right)\right] \quad (2)$$

where N_s is the total number of cells in series I_M and V_M are the current and voltage of a PV module respectively.

The photo-generated current I_{PH} has a direct relation with solar irradiance and operating solar cell temperature (see Eq. (3) It can be evaluated as [10]:

$$I_{PH} = I_{SC}\frac{G}{G_{STC}} + k_i(T_c - T_{STC}) \tag{3}$$

where G is the given irradiance level, T_c is the cell temperature. G_{STC} and T_{STC} are the irradiance and temperature at standard test conditions (STC) respectively. k_i is the temperature coefficient of the current in (A^0C) and I_{SC} is the short-circuit current at STC.

The reverse saturation current (I_0) can be evaluated as a function of short-circuit current (I_{SC}), open-circuit voltage (V_{oc}), temperature and energy bandgap of a semiconductor (E_g) as:

$$I_o = \left(\frac{T_c}{T_{STC}}\right)^3 . \frac{I_{SC}\exp\left(E_{go}/V_{to} - E_g/V_t\right)}{\exp\left(V_{oc}/nN_sV_{to} - 1\right)} \tag{4}$$

where V_{to} is the thermal voltage at STC and E_{go} is the energy bandgap at T = 0 K.

Readers are referred to [6, 10–12] for extended knowledge on how to derive and evaluate the different PV cell and module model parameters. The functions are also implement is freely available pvlib simulation packages [13].

2.1.2 Temperature dependence of I-V characteristics

Addition to solar irradiation, the I-V curve characteristics also depend on the operating temperatures of the cell T_c. According to IEC60891 standards [14], the temperature and irradiance correction of I-V characteristics are given as:

$$I_{sc} = I_{SC-STC}.\left(\frac{G}{G_{STC}}\right).(1 + \alpha_{sc}(T_c - T_{STC}) \tag{5}$$

$$V_{OC} = V_{OC-STC}.\left(1 + \beta_{oc}(T_c - T_{STC}) + n.N_s.V_t.\ln\left(\frac{G}{G_{STC}}\right)\right) \tag{6}$$

where α_{sc} and β_{oc} are the temperature coefficients for short-circuit current and open-circuit voltage respectively. n, N_s and V_t are the ideality factor, total number of cells and thermal voltage respectively.

Figure 4 shows the effect of irradiance and module temperature on I_{sc} and V_{oc}. The irradiance has a greater impact on the short-circuit current and the temperature has a greater impact on the open circuit voltage.

According to module mounting (e.g Open rack, close to the roof, insulated rack) and module construction (e.g glass–glass or glass-backsheet), the cell temperature might be some degrees hotter than module temperature [15]. In [15] the cell temperature (T_c) is derived from the module temperature measured at the surface of the module(T_m) and the irradiance G by a simple relations as:

$$T_c = T_m + \frac{G}{G_{STC}}.\Delta T \tag{7}$$

where ΔT is the temperature difference between the cell and the module back surface at an irradiance level of 1000 W/m^2. In [15], ΔT was found to be around 3 ^0C for open-rack mount, 1 ^0C roof mount and 0 ^0C for insulated back.

The module temperature is calculated from solar irradiation, ambient temperatures and/or wind speed using different methods [16]. The commonly used models are: Standard NOCT model (Eq. (8)), the Faiman model (Eq. (9)) [17] and the Kings model also known as Sandia model (Eq. (10)) [15].

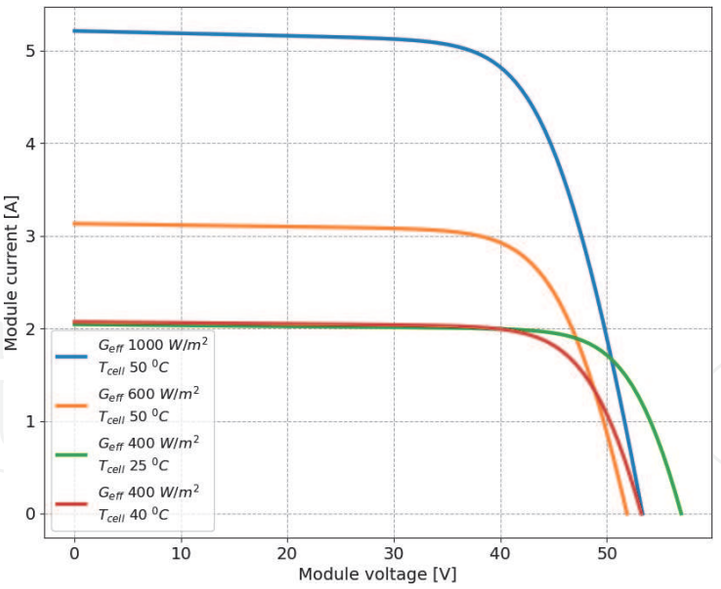

Figure 4.
Effect of irradiance and cell temperature on I-V characteristics. (I-V curve simulated using Pvlib one-diode model).

$$T_m = T_{amb} + \frac{(NOCT - T_{NOCT})}{G_{NOCT}}.G \tag{8}$$

$$T_m = T_{amb} + \frac{G}{U_0 + U_1.WS} \tag{9}$$

$$T_m = T_{amb} + G.\exp(a + b.WS) \tag{10}$$

where NOCT is the nominal operating cell temperature at given conditions (T_{NOCT} = 20 ^0C, G_{NOCT} = 800 W/m^2, wind speed (WS) = 1 m/s), NOCT is in the range of 40–50^0C [9]. T_{amb} $[^0C]$ is ambient temperature, G $[W/m^2]$ is the incident solar irradiance on the module, and WS $[m/s]$ is the wind speed. U_0 $\left[W/^0Cm^2\right]$ and U_1 $\left[Ws/^0Cm^2\right]$ are the coefficients describing the effect of the radiation on the module temperature and the cooling by the wind, [18] respectively. a and b are parameters that depend on the module construction, materials and on the mounting configuration of the module [15].

2.1.3 Example of PV power prediction using physical method

In this example, we demonstrate a practical application of the described physical model to predict four days PV module power using measured irradiance. The predicted power is then compared with real measured power of the PV module. The properties, electrical and thermal parameters of the PV module are presented in **Table 2**. The module is exposed in Gran Canaria, Spain at tilt angle of 23^0 and azimuth angle of 169^0. The module is installed as open rack-fixed configuration. The power and module temperatures are recorded every 5 minutes. The module temperature is recorded using a Pt100 sensor attached at center-back of the module. In addition, weather data (ambient temperature, wind speed and global horizontal and in plane of array solar irradiation) are also recorded with a 1 minute resolution.

The PV module specific parameters and the weather variables (irradiance and cell temperature) are used as input variable in Pvlib to simulate the five parameter (5-P) one-diode model (Eq. (2)). **Figure 5** shows the simulated I-V curves at

Module properties	
Design and cell technology	Glass–Glass and Poly-crystalline silicon cells
Number of cells	80
Electrical parameters	**Values**[a]
Maximum power rating (P_{mpp})	283 [Wp]
Rated current (I_{mpp})	7.2 [A]
Rated voltage (V_{mpp})	39.3 [V]
Short-circuit current (I_{SC})	7.8 [A]
Open-circuit voltage (V_{OC})	48.9 [V]
Thhermal parameters	**Quantity**
Temperature coefficient of power (γ)	-0.47 [$\%/K$]
Temperature coefficient of short-circuit current (α)	2.39 [mA/K]
Temperature coefficient of open-circuit voltage (β)	-161 [mV/K]
Normal operating cell temperature T_{NOCT}	48 [^0C]

[a]*The values are measured at STC (i.e. 1000 W/m^2 irradiance, 1.5 air mass and at 25°C temperature).*

Table 2.
Module properties and manufacturer datasheet electrical and thermal parameters.

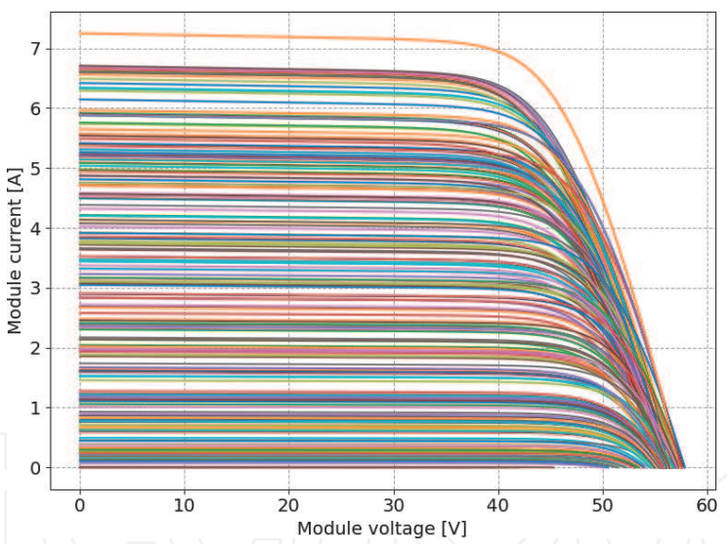

Figure 5.
I-V curves of simulated solar module at different incident irradiance and temperature levels. Irradiance range from 0 w/m² to 1168 w/m² and module temperature range from 10 °C to 50.3 °C.

different irradiance and temperature levels for a period of four days. 15 minutes aggregated data of temperature and irradiance are used hence the curves are evaluated every 15 minutes. From each I-V curve the power at maximum power point can be computed using:

$$P_{mpp} = I_{mpp} \times V_{mpp} \tag{11}$$

In this example, the uncertainty of power prediction due to temperature models are evaluated. The three commonly used temperature models presented in (Eqs. (8), (9) & (10)) are used to model the module temperature. The parameters of the models are: a = -3.87 & b = -0.0594 for Kings model and $U_0 = 25.6$ &

$U_1 = 25.6$ for Faiman model. These parameters are extracted from literature values in [15, 18] with small modifications to minimize the uncertainty.

To evaluate the uncertainty in prediction, the normalized root mean square error (NRMSE) (Eq. (12)) is used. Additionally, the normalized mean bias error (NMBE) (Eq. (13)) is also evaluated as a metric to capture the average bias in the prediction (i.e, to check whether the predictions are overestimated or underestimated).

$$NRMSE = 100 \times \frac{\sqrt{\frac{1}{N}\sum_{j=1}^{N}\left(p_j - m_j\right)^2}}{\overline{m}} \tag{12}$$

$$NMBE = 100 \times \frac{\frac{1}{N}\sum_{j=1}^{N}\left(p_j - m_j\right)}{\overline{m}}, \quad \overline{m} \neq 0 \tag{13}$$

Where p is the predicted data, m measured data, $\overline{m}$ is the mean of the measured data.

Figure 6 shows the plot of measured and modeled temperature (a), measured and predicted power (b). The uncertainty in temperature models as well as the corresponding uncertainty in power predictions are presented in (c). It is clearly visible that, for each temperature model, the prediction is different and hence the uncertainty value. Generally, the Kings model showed the best performance both in temperature modeling as well as power prediction depending on the NRMSE and R^2 values. The standard NOCT model showed the least performance which is not surprising since the model doesn't take into account the impact of wind speed. All

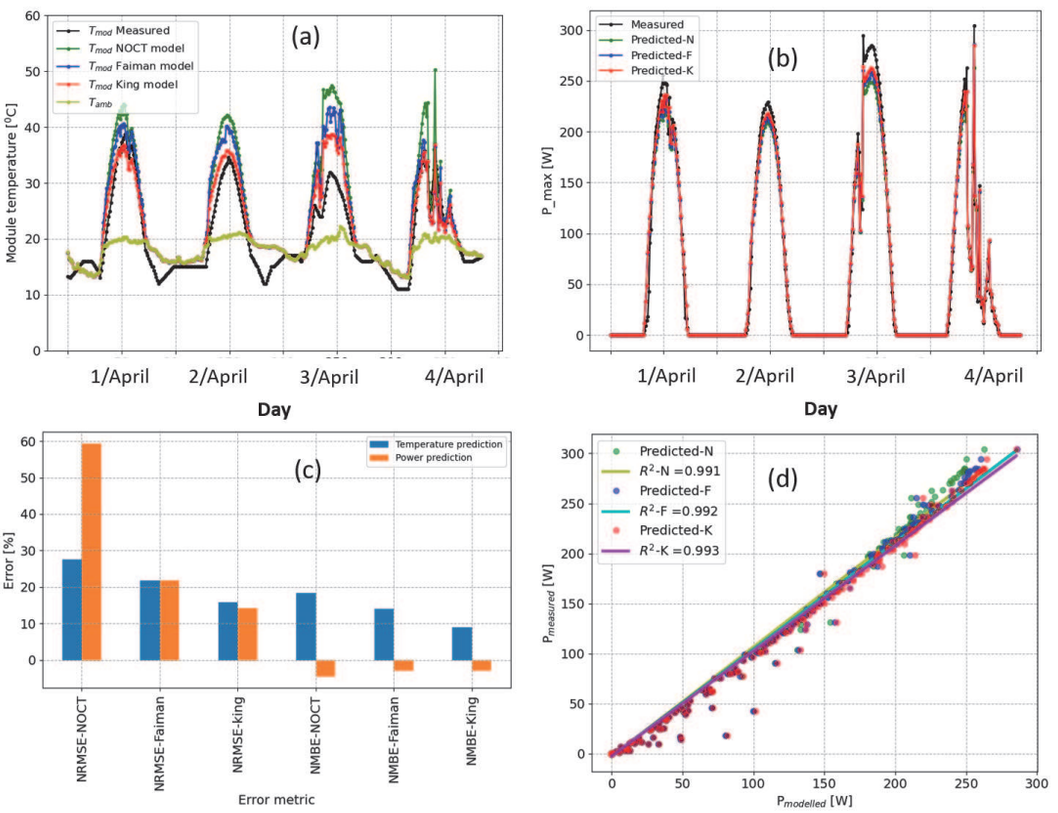

Figure 6.
a, Measured module (black) and ambient (yellow) temperatures, modeled temperature with standard NOCT (green), Faiman (blue) and Kings (red) models. b, Measured (black) and predicted power with module temperature calculated using NOCT (green), Faiman (blue) and Kings (red) models. c, Evaluated NRMSE and NMBE for module temperature modeling (blue) and power prediction (orange). d, Measured versus predicted power with module temperature calculated using NOCT, Faiman and Kings models.

the models overestimate the module temperature which correlates with the underestimation of the predicted power (see **Figure 6c**). It should be noted that, although the Kings model showed good predictions based on this example, it is not enough to guarantee that this will always be the case when the model is applied on different dataset. This is because the accuracy of the temperature models has been found to be influenced by; geographical locations, model design and mounting conditions [18].

2.2 Data-driven methods

Data-driven methods can be categorized into: data-driven heuristic methods, statistical methods and machine learning methods.

2.2.1 Data-driven heuristic methods

The physical models described in Section 2.1 have one big drawback that they require too many input variables which are not usually directly available. In this case, heuristic models are proposed to reduce the number of required inputs. They are heuristic models because they are not developed from physical assumptions/theories. Therefore, they have no physical dependencies/interpretations. They are classified as data-driven models because they are derived from correlation between weather and power output data. In [19], several heuristic models are presented and compared. They are developed on similar principles of generating power forecast from irradiance and module temperature but only differs in the numbers of fitting model parameters. The basic advantage of heuristic models is their simplicity to derive the model parameters from PV power historic data. Here we present the two- (Eq. (14)) and three- (Eq. (15)) parameter models described in [19, 20] respectively.

$$P_{mpp} = \left[\left(1 + x.\ln\left(\frac{G}{G_{STC}}\right) + y.\ln^2\left(\frac{G}{G_{STC}}\right)\right).P_{STC}.\frac{G}{G_{STC}}\right].(1 + \gamma(T_c - T_{STC}) \tag{14}$$

$$P_{mpp} = \left[a.G + b.G + c.G^2.\ln^2\left(\frac{G}{G_{STC}}\right)\right].(1 + \gamma(T_c - T_{STC}) \tag{15}$$

Where P_{mpp}, is the generated power at maximum power point by a PV module, G is the simulated or measured irradiance T_c is the cell temperature evaluated using equation (Eq. (7)), a, b, c, x and y are the models fitting parameters γ is the temperature coefficient of power in (%/^{0}C). P_{STC} and $G_{STC} = 1000w/m^2$ are the power and the irradiance at STC.

2.2.2 Example of PV power prediction by heuristic methods

In this example, the described heuristic models are applied to predict the power of the same PV module described in subSection **2.1.3**. To calibrate the models, six days historical data (25/March −31/March) are used. The extracted parameters are presented in **Table 3**. To demonstrate the usefulness of temperature correction term, the second term in (Eq. (15)) is removed and the model is re-calibrated.

In **Table 3** 3-parameter model corresponds to calibration without a temperature term and 3-parameter-T_{corr} model corresponds to calibration with a temperature correction term.

Model	Parameter				
	a	b	c	x	y
1. parameter model (Eq. (14))	—	—	—	0.0255	−0.03016
2. parameter model (Eq. (15))	0.2842	−2.935e-5	−0.0106	—	—
3. parameter-T_{corr} [a] model (Eq. (15))	0.2802	−8.985e-7	−0.0111	—	—

[a] *T_{corr} is the temperature correction.*

Table 3.
Extracted model parameters of the heuristic models.

Figure 7a shows a plot of the four days measured power and the predict power using 2-parameters and the 3-parameters model. **Figure 7b** shows the uncertainty in model calibration and the corresponding uncertainty in prediction using the different models. According to the NRMSE and NMBE values, the 3-parameter model with temperature correction term showed the best performance. From the same figure it can also be concluded that it is important to include a temperature correction in power prediction since the same model showed the least performance when applied without a temperature correction.

2.2.3 Statistical and Machine learning methods

Like heuristic models, statistical and machine learning (ML) methods are also based on historical data to generate PV power forecast. While heuristic models focus on an in-deep formulation of mathematical operations, statistical models require selecting a model that considers previous knowledge of the system. ML methods require the selection of a predictive algorithm by relying on its empirical capabilities. Statistical models aim to "inference" the outcome of a model, while ML approaches aim to find generalizable predictive patterns [21]. Both statistical and ML methods are data-driven approaches that rely on the availability and accuracy of existing operational data to generate the forecasting. Usually, the larger the historical data, the better the PV system can be understood in terms of operational behaviour under different weather conditions and hence the better the forecasting accuracy.

The list of published methods is extensive and a case-to-case benchmark is usually needed. Statistical methods such as Naive method, ARIMA (Autoregressive

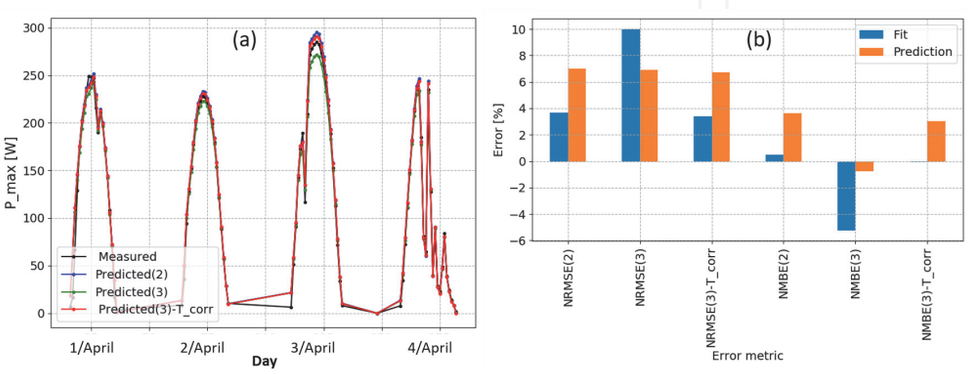

Figure 7.
a, Measured power (black), predicted power using 2-parameter model (blue) and with 3-parameter model without (green) and with (red) temperature correction. b, Evaluated NRMSE and NMBE for the models during calibration (blue) and prediction (orange).

Integrated Moving Average), SARIMA (seasonal ARIMA), Ordinary Least Squares (OLS) or Facebook Prophet (FbP) are typically applied for PV forecasting with and without regressors [22–24]. Below is a basic description of the commonly applied statistical methods. A detailed description of each method can be found on the respective cited reference.

- **Autoregressive Integrated Moving Average (ARIMA)**: This method is composed of three main elements: the auto-regression order (p), differencing order (d) and moving average order (q). The statistical formulation of this method can be found in [25].

- **Seasonal ARIMA (SARIMA)**: This model is an extension of the ARIMA approach, which adds the seasonal behaviour of the dataset analysed. This feature is of interest for PV applications due to the high seasonality on daily and annual basis observed in PV systems [26].

- **Ordinary Least Squares (OLS)**: This method analyses the system by fitting linear relationships between one or more input variables, and by minimizing the sum of square errors of a continuous or at least interval outcome variable (actual versus predicted values) [27].

- **Facebook Prophet (FbP)**: This methodology has been developed to allows non-experts in data science to adapt and configure the model to their needs. The FbP method is based on a decomposable time series model including trend, seasonality, holidays (not important for our application), and an error term. It is also possible to define the type of evolution: linear or logistic. For PV forecasting, it facilitates the modelling for short- and long-term by enabling features such as time resolutions and temporal seasonalities [28].

More advanced methods, the so-called machine learning (ML) methods, can provide better results [29, 30], however, in most cases they require more computational efforts. Some examples of machine learning methods used in PV power forecasting include [2, 29–31]: k-nearest neighbors (k-NN), artificial neural networks (ANN), support vector machine (SVM), random forests (RF) and light gradient boost machines (LightGBM). The basic description of these methods is presented below with the references for a detailed description.

- **k-Nearest Neighbors (k-NN)**: Is classified as one of the simplest and straightforward ML method. The K-NN algorithm compares the current state's Euclidean distances with training samples in feature space to select the "k" nearest neighbours used in predictions. Detailed description and application in PV power forecasting are presented in [32, 33]

- **Artificial Neural Networks (ANN)**: Inspired by biological neural networks, this algorithm is composed by neurons (mathematical units) and weights (the link between mathematical units). Using gradient-based optimization techniques, the ANNs learn a specific task (e.g., prediction) by the optimization of the "weights". This method is widely used in PV power forecasting [2] described in [29, 34, 35]

- **Support Vector Machine (SVM)**: The method separates the data linearly and transforms it into a higher dimensional feature space through a specific kernel function. The linear separation is performed with the so-called "hyperplanes".

The SVMs can be used for regression as well as classification tasks. In [32] the model is presented and applied for short-term PV power forecasting.

- **Random Forests (RF)**: The RF algorithm is based on an ensemble learning. A set of decision/regression trees is created and the final result is voted. For a regression problem, a set of regression trees are trained and the forecast will equal to the mean of individual regression tree results [36].

- **Light gradient boosting machine (LightGBM)**: This algorithm is an advanced gradient boosting decision tree (GBDT), which combines two techniques to find more effectively (higher accuracy and high processing speed) the optimal split point in the GBDT: (1) Gradient-based one-side sampling (GOSS), to reduce the number of data instances, and (2) exclusive feature bundling (EFB) to reduce the feature space. In [30, 37], the method is presented and applied for PV power prediction.

ML algorithms can be classified as supervised and unsupervised. A supervised ML algorithm uses labelled training data. It is related to a standard fitting procedure to find the unknown function/relationship between the input and output variables. The unsupervised ML algorithm uses unlabelled training data to find the data patterns (e.g., in the samples' clustering). For PV power prediction, the supervised algorithms are commonly used, due to weather forecasting availability. In general, the procedure to run a ML algorithm can be composed of the following stages:

- Data collection: the available historical data (weather and PV operational) are collected and filtered. The collection of weather forecasting data is also considered.

- Feature selection: identification of the most relevant variables with regard to the PV power output selected for further analysis.

- Data augmentation: in this stage, the enhancement of the initial dataset is expected by typically applying mathematical operations (e.g., physical relationships) to one or more relevant input variables.

- Dataset split: the input dataset is divided into a training and testing dataset. Also, a validation dataset is recommended. This task is typically applied over the sorted or random timestamps.

- Accuracy improvement: statistical indicators such as the MBE, RMSE or R^2 are used to quantify the accuracy of any forecasting model. Cross-validation techniques are recommended.

A simple ANN network architecture is shown in **Figure 8**, where the layers of a multilayer perceptron (MLP) for PV power forecasting are presented. The input data for training can be the historical weather and PV power output, while for testing and forecasting, expected weather variables are the input to the expected PV power output in the future.

2.2.3.1 Example of PV power prediction by statistical and machine learning approaches

The statistical and ML approaches are applied to the same dataset used for physical and heuristic models. To train the models, the regressors selected are

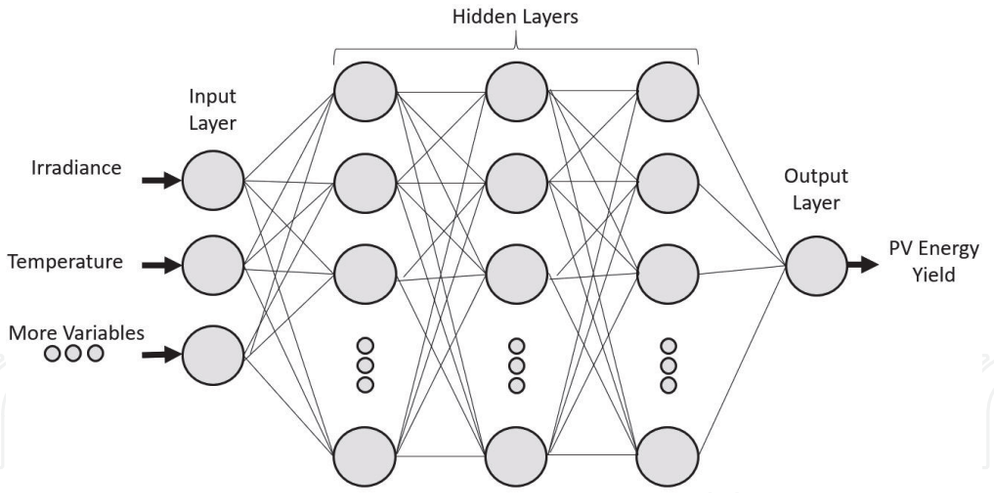

Figure 8.
Multilayer perceptron (MLP) for PV power forecasting, where the input layer includes at least irradiance and temperature, while the output layer comprises the PV energy yield.

Model		Parameter	
	Setup 1	**Setup 2**	**Setup 3**
ARIMA[a]	p: 0	q:0	d: 0
SARIMA[a]	p: 1	q:1	d: 1
OLS[a]	—	—	—
Prophet[b]	Daily Seasonality	Changepoint prior scale: 0.01	—
DNN[c]	Hidden layer size: (40,20,10)	relu activation	adam solver
LightGBM[d]	Number of leaves: 10	Min. data in leaf: 5	Number iterations: 100
SVM[c]	Kernel: rbf	Degree: 3	Regularization: 1e6

[a]*Statsmodels.*
[b]*Prophet.*
[c]*Sklearn.*
[d]*Lightgbm.*

Table 4.
Setup parameters for different statistical and machine learning approaches.

limited to the plane-of-array irradiance and PV module temperature, while the target variable is the PV output power. Most of these models are already implemented in open-source software packages (i.e., statsmodels [38], prophet [28], sklearn [39] and lightgbm [39]) and executed with Python scripts. Setting parameters for each statistical and machine learning method are given in **Table 4**.

Figure 9a shows a plot of the four-day measured power and the predicted power using statistical and machine learning models. **Figure 9b** shows the uncertainty in model calibration and the corresponding uncertainty in prediction using the different models. According to the NRMSE, LightGBM model shows the best performance, followed by the SVM and Facebook Prophet.

Generally, comparing the uncertainty values among the physical, heuristic, statistical and machine learning methods (see **Figures 6c, 7b** and **9b**), physical method has the worst performance in comparison to other methods. This is not surprising given that physical models generate prediction without preliminary performance data, unlike the counterpart models that base their predictions from historical data. Also, the different assumptions and the too many input parameters increase the uncertainty range in physical models.

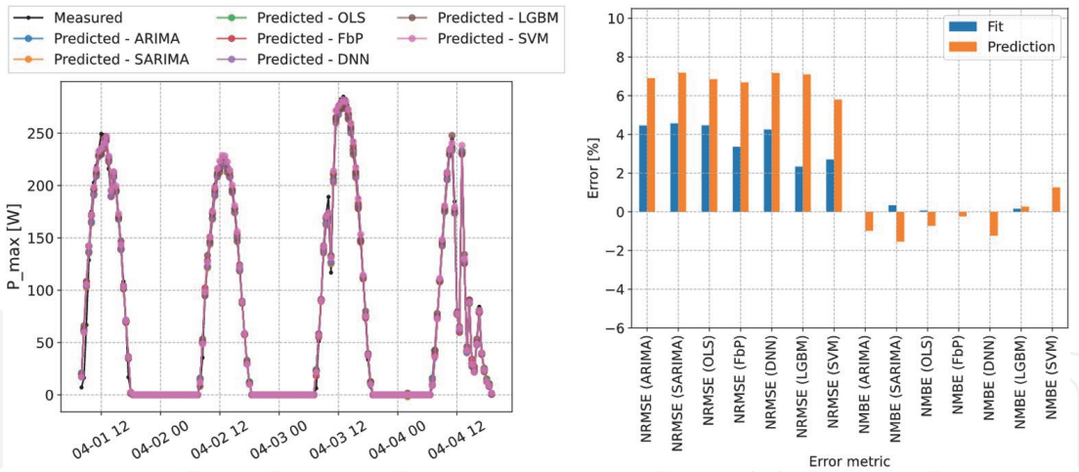

Figure 9.
A, Measured power (black), predicted power using statistical and machine learning approaches. B, Evaluated NRMSE and NMBE of statistical and machine learning approaches at calibration and predictions.

3. Lifetime PV power forecasting

To begin with, it is important to understand how the lifetime of a PV module is defined. Unlike other electrotechnical devices where the term lifetime is clearly defined [40], the definition of a PV module "lifetime" is somewhat complex. This is because, despite the catastrophic events (such as fire) it is unlikely that a PV module drops its power generation to zero. However, even though a PV module is still generating power, its output might be too low to be economically viable to continue its operation. Therefore, in general terms PV lifetime is defined in economical rather than technical terms.

For economical viability of PV projects, most PV module manufacturers guarantee a power reduction of less than 20% within 25–30 years of operation. The 20% power reduction is usually referenced at standard test conditions (STC) (modules tested under 25^0C temperatures, 1000 W/m^2,irradiance and air mass 1.5). Therefore, in this context the lifetime of a PV modules is defined as the time required for a PV module to loss it's STC power by 20% .

The actual performance of a PV module throughout its lifetime is very uncertain and difficult to accurately forecast. This is because many factors can influence the performance of a PV module. Some of these factors may include: solar resource, the quality of the PV components and the long-term variations in system performance (degradation). All these factors increases the uncertainty in PV lifetime forecast. The Internation Energy Agency- Photovoltaic Power Systems Programme (IEA-PVPS) -task 13 report [41] provides a detailed overview of the uncertainties in lifetime yield predictions. To improve the accuracy and to achieve reliable lifetime PV forecast, all these effects must be explored separately. Here, we asses the effect of degradation on lifetime PV power forecasting.

3.1 Effect of degradation on lifetime PV power forecasting

Degradation is defined as the gradual and non-reversible decrease in PV performance over time. Degradation is a crucial influencing factor to be taken into account during lifetime PV power forecast. This is because over time the PV components are ageing and deteriorating in their normal operation. Understanding how PV degrades is a very widely studied topic in the PV community but it is also among the not well understood topics. This can be explained by the numerous factors

influencing PV degradation. These factors include: PV technology, bill of materials (BoM), climatic conditions [4], transportation, installation and operational conditions. More-so, since new materials are being proposed frequently, it increases the complexity to correlate the factors that influence PV degradation. Usually, different materials have different degradation kinetics and are influenced by different stress factors differently.

In lifetime power prediction models, degradation effect is included in a number of ways. For example, according to the survey carried out among PVPS-task 13 experts regarding degradation effects in lifetime energy yield prediction, the following assumptions are taken [41]: (a). A variable degradation during the first five year of operation and a fixed degradation from 5 to 30 years of operation. A degradation of 1–2% is assumed in the first year, 0.7% to 0.5% to year 5 and0.3% to 0.5% up to year 30. (b). Initial degradation of 0.3% to 1.0% in the first years to include the effects of initial degradation modes such as light induced degradation (LID). (c). Constant degradation over the years with the exception of the first year to take into account technology specific behaviour.

Generally, a constant degradation rate with linear performance loss is considered (see Eq. (16)). However, some authors [42–44], have evaluated and modeled the non-linearity in degradation rates and performance. For example in [43] a non-linear power degradation model (Eq. (17)) was proposed and applied in [44] with a time-dependent degradation rates to predict PV performance lifetime.

$$P(t) = P_{mpp}.(1 - k.t) \tag{16}$$

$$P(t) = P_{mpp} \times \left[1 - exp\left(-\left(\frac{B}{k.t}\right)^{\mu}\right)\right] \tag{17}$$

Where P_{mpp} is the power calculated using either the physical model or the heuristic models described in the previous sections. k [%/year] is the degradation rate B and μ are model and shape parameters respectively.

3.2 Example of lifetime PV power forecasting

In this example, the effect of degradation rate on lifetime power forecast is presented. Using, 30 years of historical weather data (global irradiance, ambient temperature and wind speed from ERA 5 reanalysis [45]), three different degradation scenarios are presented and their impact on lifetime power and energy yield prediction. The first scenario is using a non-linear performance degradation with a shape parameter ($\mu = 1.2$), the second scenario is using a linear performance degradation and the third scenario is using a non-linear model with a different shape parameter ($\mu = 0.2$). It should be noted that in all the three cases, a constant degradation rate of 0.8%/year corresponding to a lifetime of 25 years (to have a $-$ 20% power loss) is used. In all the cases P_{mpp} is calculated using a 3-parameter heuristic model (Eq. (15)). Although it is not usually the case, the prediction without degradation effect and its impact on lifetime energy prediction is also shown in this example.

Figure 10 shows the PV lifetime power and yield predictions using different scenarios. It can be seen that depending on the degradation scenario, the lifetime yield is significantly different. In numbers, when compared to the usually used linear scenario, a relative difference of over 5.0% is evaluated in respect with the non-linear scenarios.

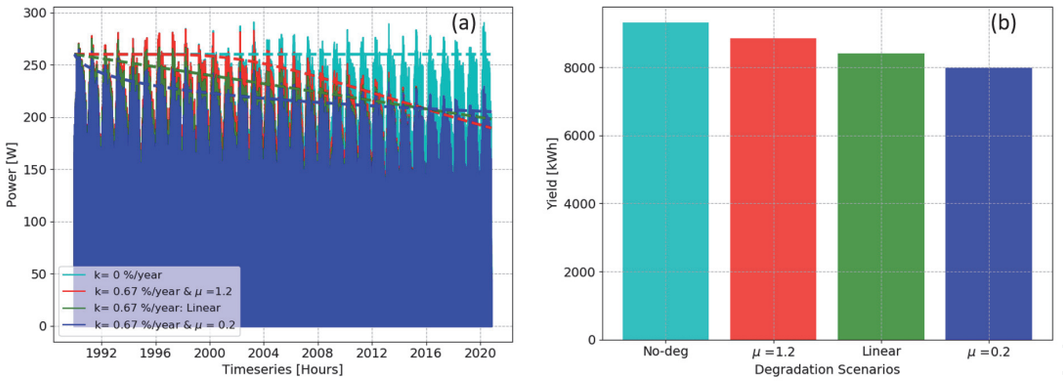

Figure 10.
A, Lifetime power prediction using different degradation scenarios: non-linear with shape parameter = 1.2 (red), linear (green) and non-linear with shape parameter = 0.2 (blue) as well as a no degradation scenario (cyan). B, Corresponding lifetime yield for all the respective scenarios.

4. Conclusions

PV power forecasting is important to stabilize the electrical grids, financing PV projects and also to plan operational and maintenance activities. In this regard, different methods are proposed to forecast the PV power generation. In this chapter, the different methods used in PV power forecasting are presented, applied and their accuracy in PV forecasting is evaluated using measured PV module and weather data. The degradation effect on lifetime PV power forecasting is also assessed using two main scenarios; linear degradation scenario and non-linear degradation scenario. The key observations in the chapter are:

- The uncertainties in PV module temperature modelling affect the forecasting accuracy. In the chapter, three temperature models: Standard NOCT, Faiman and King's are compared, the King model showed the best performance among the three models. The standard NOCT model, that neglects the impact of wind speed displayed huge uncertainty.

- Data-driven models outperforms physical models in prediction accuracy. This can be explained by the fact that physical models are derived from to many assumptions and that they need too many input parameters that are usually approximated.

- For lifetime PV power forecasting, a relative difference of over 5% is evaluated between the linear and non-linear degradation scenarios.

Conflict of interest

The authors declare no conflict of interest.

DOI: http://dx.doi.org/10.5772/intechopen.97049

Author details

Ismail Kaaya[1]* and Julián Ascencio-Vásquez[2,3]

1 Fraunhofer ISE, Freiburg, Germany

2 Faculty of Electrical Engineering, University of Ljubljana, Ljubljana, Slovenia

3 3E sa, Brussels, Belgium

*Address all correspondence to: ismail.kaaya@ise.fraunhofer.de

References

[1] Global solar photovoltaic capacity [Internet]. Available from: https://www.globaldata.com/global-solar-photovoltaic-capacity-expected-to-exceed-1500gw-by-2030-says-globaldata/ [Accessed: 28-October-2020].

[2] Antonanzasa J, Osoriob N, Escobar R, Urraca R, Martinez-de-Pisona F.J, Antonanzas-Torresa F: Review of photovoltaic power forecasting. Solar energy. 2016; 136(15): 78–111. https://www.sciencedirect.com/science/article/abs/pii/S0038092X1630250X [Accessed: 10 November 2020]

[3] Pelland, Sophie, Remund, Jan, Kleissl, Jan, Oozeki, Takashi and De Brabandere, Karel Photovoltaic and Solar Forecasting: State of the Art. (IEA-PVPS T14-01: 2013) , International Energy Agency Photovoltaic Power Systems Programme (2013). [Online]: https://iea-pvps.org/wp-content/uploads/2013/10/Photovoltaic_and_Solar_Forecasting_State_of_the_Art_REPORT_PVPS__T14_01_2013.pdf [Accessed: 5-January-2021].

[4] Ascencio-Vasquez, J., Kaaya, I., Brecl, K., Weiss, K.-A., & Topic, M., Global Climate Data Processing and Mapping of Degradation Mechanisms and Degradation Rates of PV Modules. Energies, 2019, 12, 4749. doi:10.3390/en12244749.

[5] Nazmul Islam Sarkar Md: Effect of various model parameters on solar photovoltaic cell simulation: a SPICE analysis. Renewables. 2016; 3–13. DOI: 10.1186/s40807-016-0035-3

[6] Tian, H., et al.: A cell-to-module-to-array detailed model for photovoltaic panels. Solar Energy, 2012. 86(9): 2695–2706. DOI:10.1016/j.solener.2012.06.004

[7] PVsyst [Internet]. Available from: https://www.pvsyst.com/features/ [Accessed: 16-November -2020]

[8] PVWATTs [Internet]. Available from: https://pvwatts.nrel.gov/ [Accessed: 16-November -2020]

[9] Alberto Dolara, Sonia Leva, Giampaolo Manzolini:Comparison of different physical models for PV power output prediction. Solar energy. 2015; 119(-): 83–99. DOI:10.1016/j.solener.2015.06.017

[10] One-Diode Model [Internet]. Available from: https://www.sciencedirect.com/topics/engineering/one-diode-model [Accessed: 20-November -2020]

[11] De Soto W, Klein S.A, Beckman W. A: "Improvement and validation of a model for photovoltaic array performance", Solar Energy, 2006; 80 (1): 78–88. DOI:10.1016/j.solener.2005.06.01

[12] Single Diode Equivalent Circuit Models [Internet]. https://pvpmc.sandia.gov/modeling-steps/2-dc-module-iv/diode-equivalent-circuit-models/ [Accessed: 20-November -2020]

[13] Pvlib python [Internet]. https://pvlib-python.readthedocs.io/en/stable/index.html [Accessed: 20-November -2020]

[14] International Electrotechnical Vocabulary. Chapter 191: Dependability and Quality of Service, IEC60050-191, International Electrotechnical Comission, Geneva, CH, Standard, 1990.

[15] Kratochvil, Jay A, Boyson, William Earl, and King, David L. Photovoltaic array performance model.. United States: N. p., 2004. Web. doi:10.2172/919131

[16] Segado P.M , Carretero J & Sidrach-de-Cardona, M: Models to predict the operating temperature of different

photovoltaic modules in outdoor conditions. Progress in Photovoltaics: Research and Applications, 2015, V23, 1267–1282. DOI:10.1002/pip.2549.

[17] Faiman, D. Assessing the outdoor operating temperature of photovoltaic modules. Progress in Photovoltaics: Research and Applications, 2008; 16(4), 307–315. https://doi.org/10.1002/pip.813

[18] Koehl, M., Heck, M., & Wiesmeier, S: Categorization of weathering stresses for photovoltaic modules. Energy Science & Engineering, 2018,V6, 93–111. DOI:10.1002/ese3.622189.

[19] Ding K, Ye Z, Reindl T. Comparison of Parameterisation Models for the Estimation of the Maximum Power Output of PV Modules. Energy Procedia, 2012; V25, 101–107. DOI: 10.1016/j.egypro.2012.07.014

[20] Performance Evaluation and Prediction of BIPV Systems under Partial Shading Conditions Using Normalized Efficiency; Energies 2019; V12, 3777; doi:10.3390/en12193777

[21] Bzdok D, Altman N & Krzywinski N, Statistics versus machine learning, Nat Methods, 2018 15, 233–234. DOI: urlhttps://doi.org/10.1038/nmeth.4642

[22] Subhra Das, Short term forecasting of solar radiation and power output of 89.6kWp solar PV power plant, Materials Today: Proceedings, 2020, DOI: 10.1016/j.matpr.2020.08.449.

[23] David P. Larson, Lukas Nonnenmacher, Carlos F.M. Coimbra, Day-ahead forecasting of solar power output from photovoltaic plants in the American Southwest, Renewable Energy, 2016 V.91, 11–20, DOI: 10.1016/j.renene.2016.01.039.

[24] M. Bouzerdoum, A. Mellit, A. Massi Pavan, A hybrid model (SARIMA–SVM) for short-term power forecasting of a small-scale grid-connected photovoltaic plant, Solar Energy, 2013, 98, 226-235, DOI: 10.1016/j.solener.2013.10.002.

[25] Pasari S, Shah A, Time Series Auto-Regressive Integrated Moving Average Model for Renewable Energy Forecasting. In: Sangwan K., Herrmann C. (eds) Enhancing Future Skills and Entrepreneurship. Sustainable Production, Life Cycle Engineering and Management: Springer; 2020. p.71–77. DOI: https://doi.org/10.1007/978-3-030-44248-4_7

[26] Dimri T, Ahmad S & Sharif M, Time series analysis of climate variables using seasonal ARIMA approach. J Earth Syst Sci,2020, 129, 149. DOI: https://doi.org/10.1007/s12040-020-01408-x

[27] Zdaniuk B, Ordinary Least-Squares (OLS) Model. In: Michalos A.C. (eds) Encyclopedia of Quality of Life and Well-Being Research. Springer, Dordrecht, 2014, DOI: https://doi.org/10.1007/978-94-007-0753-5_2008

[28] Taylor SJ, Letham B. Forecasting at scale. PeerJ Preprints 5:e3190v2, 2017; [Online] https://doi.org/10.7287/peerj.preprints.3190v2

[29] Spyros Theocharides, George Makrides, Andreas Livera, Marios Theristis, Paris Kaimakis, George E. Georghiou, Day-ahead photovoltaic power production forecasting methodology based on machine learning and statistical post-processing, Applied Energy, 2020; V 268, 115023, DOI: 10.1016/j.apenergy.2020.115023.

[30] Julian Ascencio-Vasquez, Jakob Bevc, Kristjan Reba, Kristijan Brecl, Marko Jankovec and Marko Topic, Advanced PV Performance Modelling Based on Different Levels of Irradiance Data Accuracy, Energies 2020, 13(9), 2166; DIO, 10.3390/en13092166

[31] Suresh V, Janik P, Rezmer J, Leonowicz Z. Forecasting Solar PV Output Using Convolutional Neural Networks with a Sliding Window Algorithm. Energies. 2020; 13(3):723.

[32] Fei Wang, Zhao Zhen, Bo Wang, and Zengqiang, Comparative Study on KNN and SVM Based Weather Classification Models for Day Ahead Short Term Solar PV Power Forecasting, applied science, 2018, 8, 28; DOI: 10.3390/app8010028

[33] Zhang, Y.; Wang, J. GEFCom2014 Probabilistic Solar Power Forecasting based on k-Nearest Neighbor and Kernel Density Estimator. In Proceedings of the 2015 IEEE Power & Energy Society General Meeting; 26–30 July 2015; Denver,CO, USA; DOI: 10.1109/PESGM.2015.7285696

[34] Utpal Kumar Das, Kok Soon Tey, Mehdi Seyedmahmoudian, Saad Mekhilef, Moh Yamani Idna Idris, Willem Van Deventer, Bend Horan, Alex Stojcevski, Forecasting of photovoltaic power generation and model optimization: A review, Renewable and Sustainable Energy Reviews, 2018; 81(1), p.912–928 DOI: https://doi.org/10.1016/j.rser.2017.08.017.

[35] Guido Cervone, Laura Clemente-Harding, Stefano Alessandrini, Luca Delle Monache, Short-term photovoltaic power forecasting using Artificial Neural Networks and an Analog Ensemble,Renewable Energy,2017, v108, p.274–286, DOI:https://doi.org/10.1016/j.renene.2017.02.052.

[36] Huertas Tato, J.; Centeno Brito, M. Using Smart Persistence and Random Forests to Predict Photovoltaic Energy Production. Energies 2019, 12, 100. DOI: https://doi.org/10.3390/en12010100

[37] Caroline Persson, Peder Bacher, Takahiro Shiga, Henrik Madsen, Multi-site solar power forecasting using gradient boosted regression trees, Solar Energy, 2017; v.150, p.423–436, DOI: https://doi.org/10.1016/j.solener.2017.04.066.

[38] Seabold, Skipper, and Josef Perktold. "Statsmodels: Econometric and statistical modeling with python." Proceedings of the 9th Python in Science Conference. 2010.

[39] Scikit-learn: Machine Learning in Python, Pedregosa et al., Journal of Machine Learning Research 12 2011; 12, 2825–2830.[Online] http://scikit-learn.sourceforge.net.

[40] IEC60891, 2010. Photovoltaic Devices. Procedures for Temperature and Irradiance Corrections to Measured IV Characteristics. IEC60891. [Internet] https://standards.globalspec.com/std/1207301/IEC%2060891 [Accessed: 23-November-2020].

[41] Reise C, Müller B, Moser D, and et al, IEA PVPS Task 13: Uncertainties in PV System Yield Predictions and Assessments (2018). [Online]: https://iea-pvps.org/key-topics/uncertainties-in-pv-system-yield-predictions-and-assessments/ [Accessed: 16-November-2020].

[42] Marios Theristis, Andreas Livera, C. Birk Jones, George Makrides, George E. Georghiou, and Joshua S. Stein, Nonlinear Photovoltaic Degradation Rates: Modelingand Comparison Against Conventional Methods, IEEE Journal of Photovoltaics, 2020; 10(4):1112–1118, DOI: 10.1109/JPHOTOV.2020.2992432

[43] Kaaya I., Koehl M.,Mehilli A., Mariano S. d. C., & Weiss K. A. 'Modeling Outdoor Service Lifetime Prediction of PV Modules: Effects of Combined Climatic583Stressors on PV Module Power Degradation', IEEE Journal of Photovoltaics, 2019; 9, 1105–1112. doi:10.1109/JPHOTOV.2019.291619

[44] Kaaya, I., Lindig, S., Weiss, K.-A., Virtuani, A., Ortin, M. S. d. C., & Moser, D., Photovoltaic lifetime forecast model based on degradation patterns. Progress in Photovoltaics: Research and Applications, 2020; 28, 979–992. doi: 10.1002/pip.3280

[45] Copernicus Climate Change Service (C3S) (2017): ERA5: Fifth generation of ECMWF atmospheric reanalyses of the global climate. Copernicus Climate Change Service Climate Data Store (CDS), [Internet] Available from: https://cds.climate.copernicus.eu/cdsapp#!/home, [Accessed: 16-November -2020]