POLITECNICO DI MILANO

School of Industrial and Information Engineering

Department of Electrical Engineering

Forecasting output power of the Photovoltaic systems

Master of science in Electrical Engineering

Supervisors: Professor Samuele Grillo

Candidate:

Seyedkaro Khamoushi

Academic Year 2021-2022

Acknowledgement

I would like to thank everyone who supported me in completing this master thesis. Special thanks to my supervisor, Professor Samuele Grillo, for his time, support, and guidance throughout this journey.

I am also grateful to all staff and professors of the Politecnico di Milano for showing me the way during this master's degree program.

Finally, To my beloved family, dear friends and instructors who always accepted me for me and supported my dedication, drive, and ambition.

Contents

Abbreviations index

AEM	Analog Ensemble Model
AI	Artificial Intelligence
ANN	Artificial Neural Network
AR	Autoregressive Models
ARIMA	Regressive Integrated Moving Average
BP	Back Propagation
CNN	Convolutional Neural Networks
DHI	Global Horizontal Irradiation
DL	Deep Learning
DNI	Direct Normal Irradiation
DNN	Deep Neural Network
DSO	Distributed Systems Operator
EBP	Error Back Propagation
ELM	Extreme Learning Machine
EMAE	Envelope-Weighted Absolute Error
FFNN	Feed Forward Neural Network
FL	Fuzzy Logic
GHI	Global Horizontal Irradiation
GSO	Genetic Swarm Optimization
KDE	Kernel Density Estimation
KNN	K Nearest Nneighbor
LSTM	Long Short-Term Memories
MAE	Mean Absolute Error
ML	Machine Learning
MLR	Multiple Linear Regression
NAPE	Mean Absolute Percentage Error
NMAE	Normalized Mean Absolute Error
NN	Neural Network
NWF	Numerical Weather Forecast
NWP	Numerical Weather Prediction
OMAE	Objective Mean Absolute Error
PCC	Point of Common Coupling
PDF	Probability Density Function
PHANN	Physical Hybrid Neural Network Mode
PV	Photovoltaic
QRA	Quantile Regression Averaging
RMSE	Root Mean Squared Error
SCADA	Supervisory Control and Data Acquisition
SOM	Self-Organization Map
SR	Stepwise Regression
SVM	Support Vector Machine
SVR	Support Vector Regression
TCI	Thermostatically Controlled Loads
WMAE	Weighted Mean Absolute Error

Figures index

Tables index

ABSTRACT

Photovoltaic (PV) is one of the most promising forms of renewable energy. Accurate forecasting of PV power is essential for reliable operation as well as economic integration in smart grids.

Unlike transmission lines, distribution networks have a high impedance ratio, a complex network structure, and other features, which make them essential for network security analysis and economic analysis. The majority of PV plants are directly connected to the distribution network. For this reason, the method used to analyze traditional grid power flow was difficult to apply to the analysis of PV power flow.

Solar PV systems have differing power outputs because of different weather conditions. Accurate forecasting of PV power is crucial for system stability and promoting large-scale PV installations.

However, the randomness, volatility, and uncontrollability of solar energy leads to great variability when it comes to PV power output. To reduce the negative effects of PV plants on power systems, it is very important to predict PV power accurately. The challenge is to calculate and predict the output of grid-connected PV stations and the power flowing through the distribution network.

Keywords: Forecasting Photovoltaic systems, Forecasting output power of the Photovoltaic systems, power generation, Prediction, Solar irradiance

Sommario

Il fotovoltaico (PV) è una delle forme più promettenti di energia rinnovabile. Una previsione accurata dell'energia fotovoltaica è essenziale per un funzionamento affidabile e per l'integrazione economica nelle reti intelligenti.

A differenza delle linee di trasmissione, le reti di distribuzione hanno un rapporto di impedenza elevato, una struttura di rete complessa e altre caratteristiche, che le rendono essenziali per l'analisi della sicurezza della rete e l'analisi economica. La maggior parte degli impianti fotovoltaici è collegata direttamente alla rete di distribuzione. Per questo motivo, il metodo utilizzato per analizzare il flusso di energia della rete tradizionale era difficile da applicare all'analisi del flusso di energia fotovoltaica.

Gli impianti solari fotovoltaici hanno potenze diverse a causa delle diverse condizioni atmosferiche. Una previsione accurata dell'energia fotovoltaica è fondamentale per la stabilità del sistema e per la promozione di installazioni fotovoltaiche su larga scala.

Tuttavia, la casualità, la volatilità e l'incontrollabilità dell'energia solare portano a una grande variabilità quando si tratta di produzione di energia fotovoltaica. Per ridurre gli effetti negativi degli impianti fotovoltaici sui sistemi energetici, è molto importante prevedere con precisione la potenza fotovoltaica. La sfida consiste nel calcolare e prevedere la produzione di stazioni fotovoltaiche connesse alla rete e l'energia che fluisce attraverso la rete di distribuzione.

Introduction

Oil, coal, and natural gas are the principal sources of energy, and they are non-renewable as well as polluting, resulting in greenhouse effects, acid rain, and ozone depletion by combustion of fossil fuels. In addition, importance of the solar and energy is consistently growing high in these days to reduce CO2 emissions. Green energies are used to substitute fossil fuels to reduce carbon emission. Among the renewable energies having been widely used, the photovoltaic (PV) system is one of the fast emerging green power generation alternatives.

Solar power generation, both from solar photovoltaic (PV) farms and roof-top solar panel installations is on the rise and their impact is increasingly felt in traditional energy markets.

There are several reasons behind its success: (a) the cost of photovoltaic power has plummeted since PV modules, storage systems and balance of system costs have been steadily dropping. This has led to an increasing competitiveness in comparison to the conventional non-renewable resources; (b) PV peak power generation occurs during higher load periods; (c) global concerns about climate change have become more widespread, and the power sector is playing a crucial role in decarbonizing the power system. (d) the ongoing and improving efficiency of PV generation forecasting strategies. (e) most importantly, the continued quest for energy independence for most developed countries [7].

The decreasing price of photovoltaic panels is encouraging the deployment of distributed photovoltaic generation; however, the increased number of PV installations is a concern because of some reasons such as increased system reserve requirements and Line capacity violations and over voltages in low voltage and medium voltage distribution systems, which typically have a limited host capacity [2], [3], [4].

Due to the fact that distributed PV is placed behind the meter, it is invisible to retailers and the distribution system operator, thus increasing uncertainty and making forecasting more difficult. It is important to have access to real-time estimates of PV installations' power output in order to perform efficient monitoring of local distribution systems and the management of distributed energy resources. A monitoring solution that is pervasive and communication intensive is one that is based directly on PV installations. Therefore, it is unlikely. Moreover, ownership conflicts at metering infrastructures, privacy concerns, and lack of standards for monitoring, communication, and aggregating measurements may hinder the process. Using nonredundant, interconnected communication channels that are prone to failure are barriers to capturing real-time information from distributed PV systems. Researchers have considered how to estimate PV production using aggregated power flow measurements in existing literature in order to overcome issues related to the limited observability of distributed PV generation [2, 5].

Power systems with high penetrations of solar energy are frequently identified as having issues due to variability and uncertainty in solar power. Solar power forecasting can be a useful tool to address these issues. Further, in a market environment, solar power's contribution to the generation portfolio emerges as a key indicator of the daily and hourly prices, as variances in the estimated solar power will affect both energy and operating reserve prices [6].

Chapter 1 Main Statements

1.1 Problem Statement

By Using PV systems, sunlight is converted directly into electricity. PV systems are characterized by intermittent and unreliable power output.

It has become an essential tool to improve the power system's stability and alleviate the negative effects on the entire power system due to accurate forecasts of PV power [9]. A new generation of advanced electricity meters in modern power grids can be used to build more sophisticated forecasting models to generate more accurate forecasts of PV power [10].

For illustrate this problem by referring to the setup sketched in Figure 1. PV generation of the feeder might come from multiple sites with different tilt and azimuth configurations of the panels, or it may come from an unobserved distributed PV system with an unobserved demand connected to the upper level by a point of common coupling (PCC). It is a common configuration seen in low and medium voltage systems in both commercial and residential environments, so it can be considered representative of a wide range of systems and is worth studying as a case study [2].

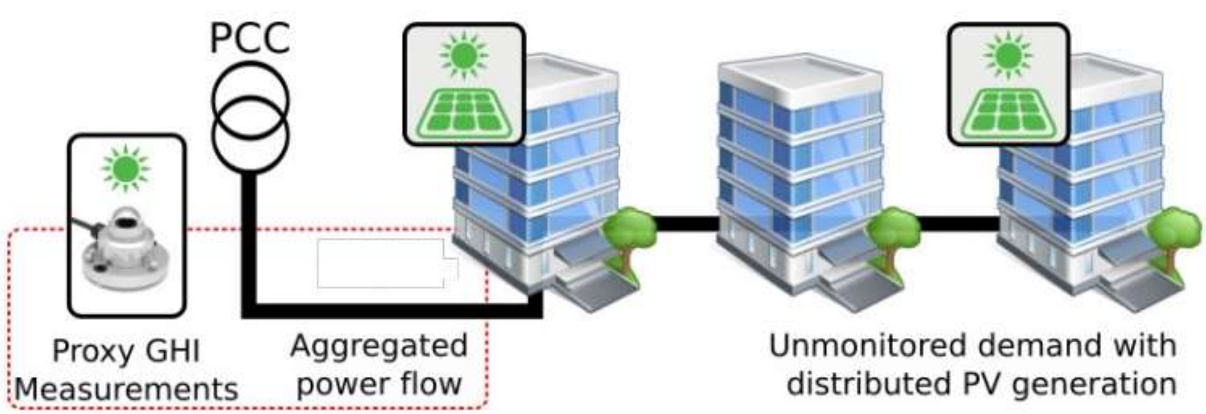

Figure 1. Electrical grids with a mass of PV arrays that are located at different azimuths/tilts and may have unmonitored demand and generation [2].

Solar PV plants connected to the grid have introduced serious issues to power networks, including unstable systems, reliability problems, and power balance issues [3]. PV grids require a forecast of solar energy generation in order to maintain a stable energy supply. In addition to improving the system's stability, accurate predictive models can reduce the cost of additional devices and eliminate the impact of solar PV output uncertainty [11-13].

1.2 Relevant Vital Factors

In order to forecast the output of photovoltaic panels, it is essential to understand the variable parameters because all forecasting methods are based on the study of these variables.

1.2.1 Solar Radiation

Solar radiation, often called the solar resource or just sunlight, is a general term for the electromagnetic radiation emitted by the sun. Every location on Earth receives sunlight at least part of the year. The amount of solar radiation that reaches any one spot on the Earth's surface varies according to: Geographic location, time of day, season, local landscape, local weather. The total radiation received from the sun, of a horizontal surface at the level of the ground for a serene day, is the sum of the direct and diffuse radiations [14].

Global horizontal irradiation (GHI): All the light that arrives on a horizontal plane, from sun, sky, clouds

• Diffuse horizontal irradiation (DHI): All the light arriving from sky and clouds, but NOT directly from the sun

• Direct normal irradiation (DNI): The light that arrives directly from the sun (and the bright patch around it) onto a plane that is normal to the direction of the sun. [14]

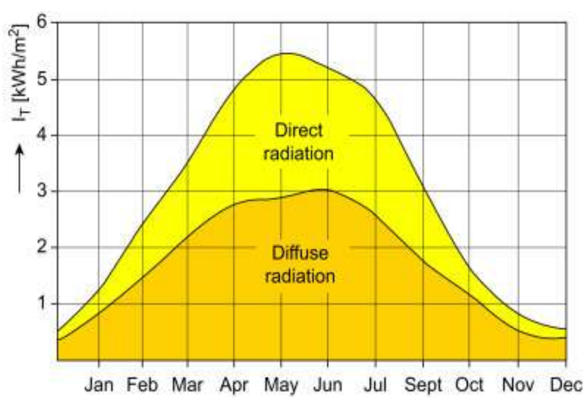

Figure 2. Represents the proportion of the diffuse radiation in total radiation.

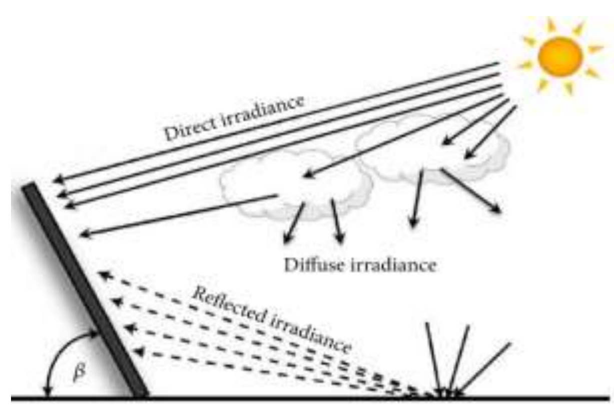

Figure 3. Solar radiation quantities.

1.2.2 Temperature and Humidity

Increases in relative humidity resulted in a significant drop in PV voltage. In addition to affecting solar radiation, relative humidity also affects PV cell output voltage [15]. Also, Solar cell performance decreases with increasing temperature, fundamentally owing to increased internal carrier recombination rates, caused by increased carrier concentrations. The operating temperature plays a key role in the photovoltaic conversion process. Aside from the weather, electricity consumption is also influenced by it. There is a significant correlation between the utility consumption pattern of the so-called thermostatically controlled loads (TCL), or the appliances used for heating and or cooling, and the outdoor temperature [16].

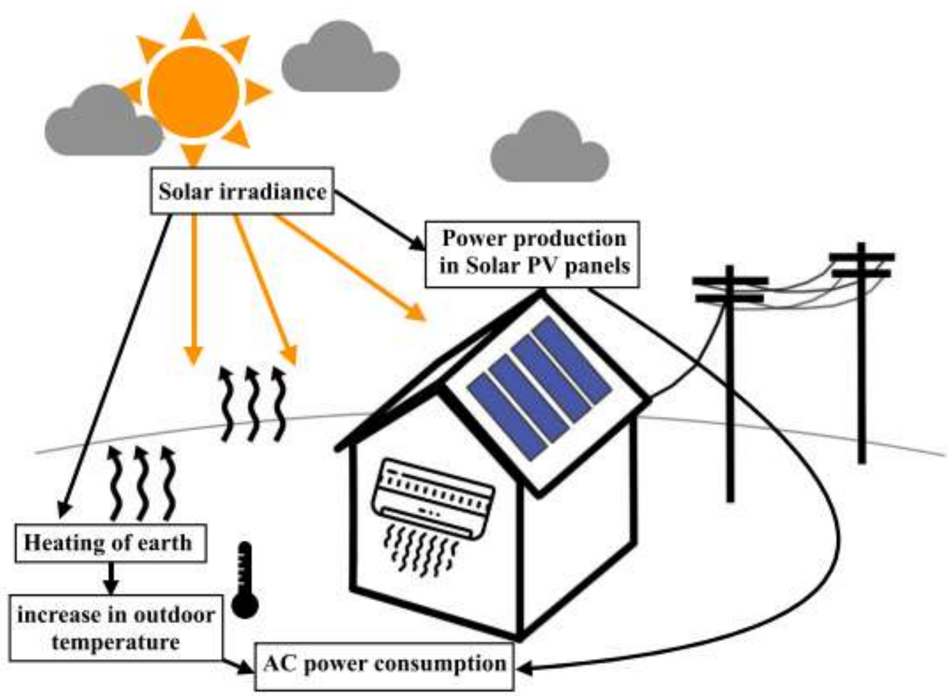

Figure 4. Relationship between solar PV power and thermostatically controlled loads.

Increases in cell temperature have the most direct effect on the open circuit voltage, which decreases linearly with increase in cell temperature. Cell voltage decreases by about 2.2 mV for every 1 °C rise in operating temperature, which reduces efficiency by about 0.5% [17].

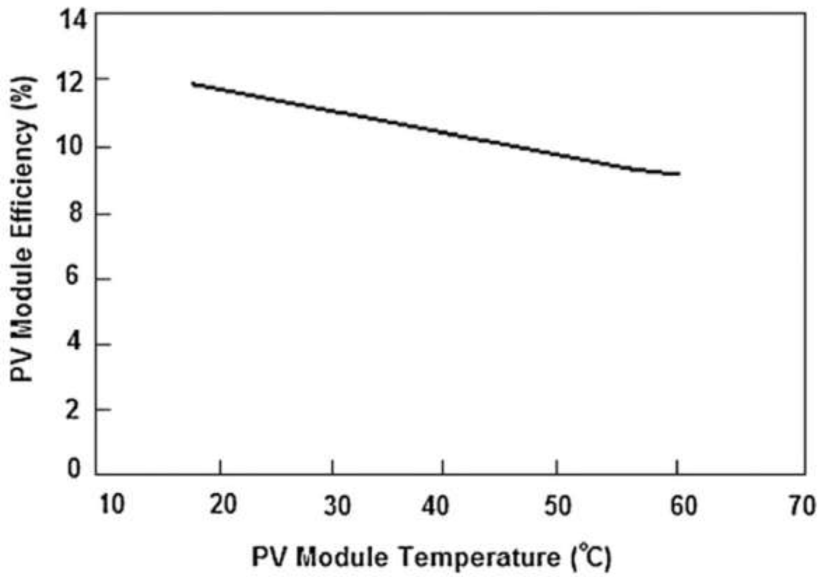

Figure 5. PV module temperature and efficiency [17].

1.2.3 Precipitation, Wind Speed and Cloud Coverage

According to the apparent movement of the sun in the sky, PV power plants generate systematic patterns of power production. Clouds have a short-term effect on this profile, causing short-term variability. PV power plants are affected by diverse geographical conditions, including atmospheric and oceanic circulation, as well as topography.

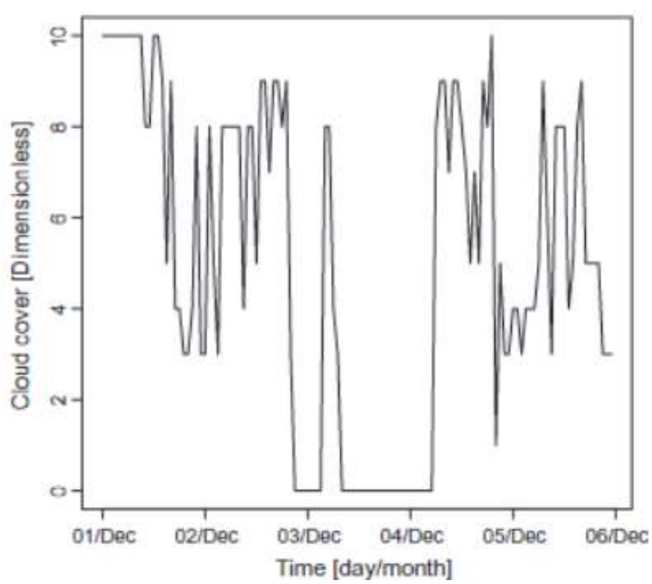

Figure 6. An example of cloud cover time series plot for Miami (2004 December 1–5) [17].

1.2.4 Dust Accumulation

Depending on the amount of dirt and dust present on the PV module, some sunlight can be blocked from reaching the panel, causing substantial losses of power generation when the irradiance of the sun is scattered across the surface. Due to dirt accumulation, a 100-watt module will typically operate at around 93 W rather than 100 * 0.93 = 100 W due to a 93% dust reduction factor [18].

1.2.5 Shading

Solar panels' power output is reduced by the shadowing effect. The shading affects not only the current flow in the shaded cells, but also the whole panel since it is normal to connect the cells in series. Among the many causes of shadows are poles, trees, and buildings. Also, leaves, birds, and bird droppings that can fall directly on the modules can cause shadows [18].

1.3 PV System Factors

Power output from a PV system is affected by several factors related to its components.

1.3.1 The I-V Characteristics of the PV Panel

The panel's rated current, rated voltage, short circuit current Is, open circuit voltage and rated power are all characteristics of the PV cell itself that affect the power generated from it. A current–voltage characteristic (I–V characteristic) of a solar cell is a plot of all possible working points in a considered range.

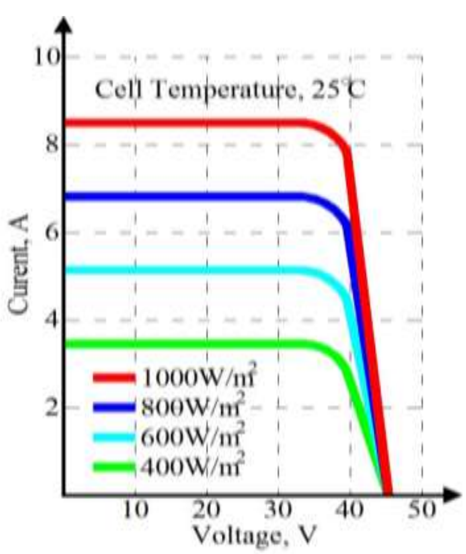

Figure 7. V-I curve for different solar irradiance

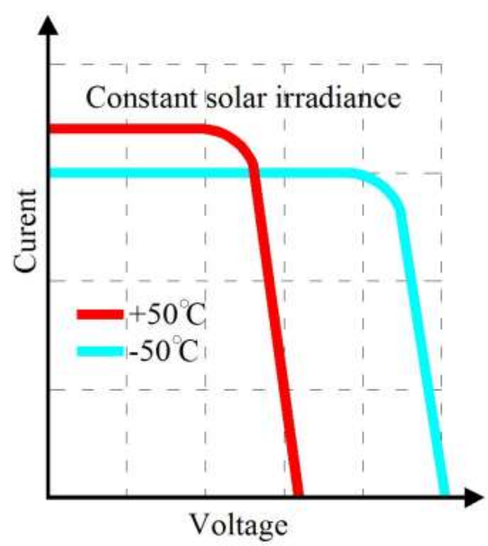

Figure 8. V-I curve for different temperatures.

1.3.2 PV Material

It is important to choose the PV material to optimize the design and performance of the system. PV materials include silicon, gallium arsenide (GaAs), copper indium diselenide (CuInSe2), cadmium telluride (CdTe), indium phosphide, and others. Moreover, PV cells can have an atomic structure that is monocrystalline, polycrystalline, amorphous or nanoscale. Monocrystalline cells typically have a high efficiency of 15% but are more expensive since their manufacturing processes are complex [13].

1.3.3 Orientation or Inclination of PV Panels

The most efficient way to utilize solar energy is when PV panels are perpendicular to the sun [19]. If the panels are installed at a fixed tilt angle, then the rule of thumb for annual optimum tilt angle states the tilt angle should be the same as the latitude of the installation location. The tilt angle deviates approximately +15° from the latitude angle in winter and about −15° of the latitude angle in summer. However, this rule of thumb does not work very well in the latitudes above 45° [20].

Therefore, the energy yield increases if the sun is tracked by the surface of the panel, so that light is also perpendicular to the panel [18]. It is possible to implement a single axis tracking or a dual axis tracking system [18].

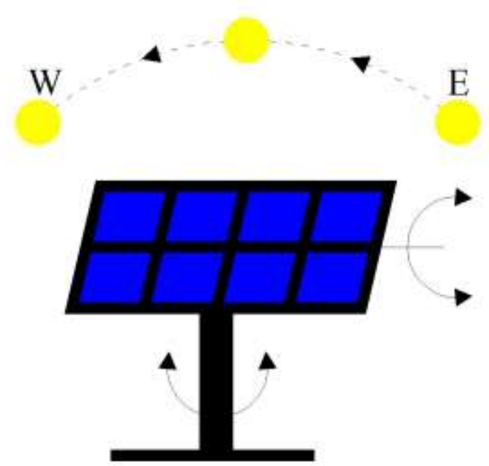

Figure 9. PV panel equipped by two axis tracking [6].

Variations in the solar irradiance level, the direction of the sun and the temperature affect the power output of a solar PV module. As tracking increases, the cost of the system, using tracking mechanism depends on the available initial investment. This method is not useful in cloudy conditions and when the solar radiation is iso-tropically distributed. The tracking process also depends on the location as its use in some locations can result in excess heating and loss of power, while its use in places with insufficient sunshine can increase power [18].

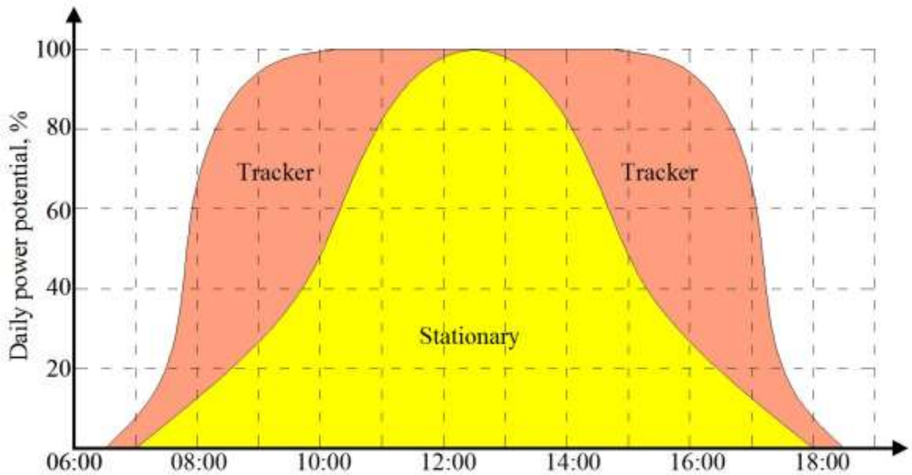

Figure 10. Efficiency of tracking and stationary PV's systems [6].

All meteorological factors that might influence the PV output are considered initially to conduct a thorough analysis of the relationship between PV forecasts and climate.

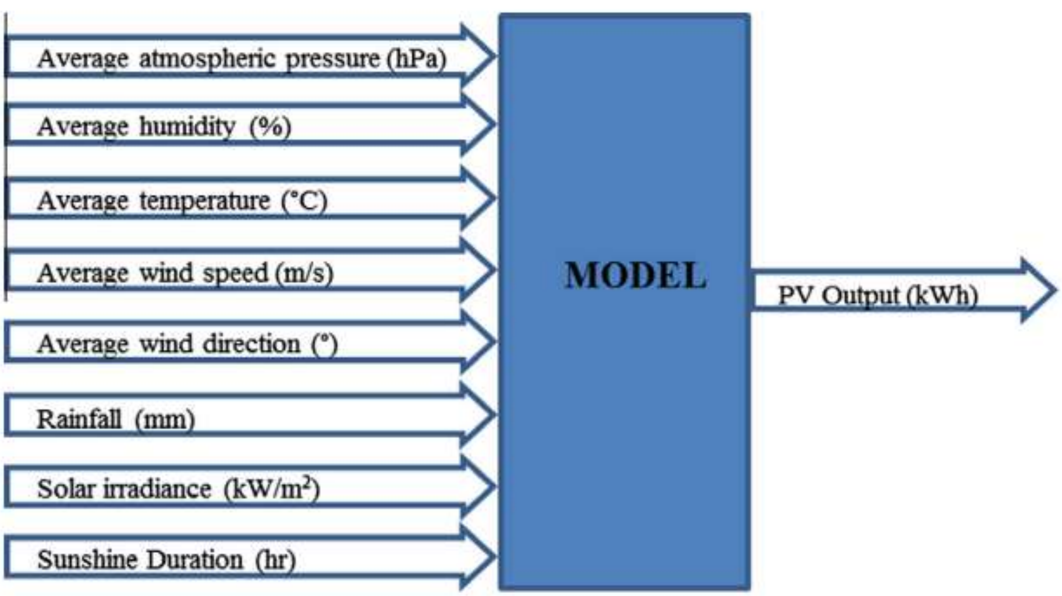

Figure 11. Effective parameters on the PV panels [31].

Measuring instruments have been installed at the site for the purpose of monitoring the meteorological conditions.

1.4 Monitoring

There are many components that contribute to the power production of a PV system, including PV panels, inverters, cabling, piping, DC circuit breakers, AC circuit breakers, leakage circuit breakers, as well as other electrical components. PV systems are available in different varieties, including rooftop mounted systems, building integrated systems, and large utility scale power systems. Sizes of PV systems range from a few kilowatts up to megawatts. In Figure 12, present the architecture of a system that can collect key electrical and meteorological information about on-grid PV systems [21].

Figure 12. A system for monitoring the on-grid PV system with an online interface [21].

The system will be monitored and forecasts generated using a variety of models will be provided. Database contains the weather conditions for the, as well as the energy generated by the PV system.

The datasets can exclude from the analysis because it is possible some data were missing due to system maintenance or failures.

1.5 Errors

In order to evaluate the accuracy of forecasting methods, the hourly error e_h is the standard error definition for the assessments, and it is defined as follows [74]:

$$e_h = P_{m,h} - P_{p.h},$$

The average actual power during an hour is $P_{m,h}$, while the prediction provided by a forecasting method is $P_{p,h}$. Following the hourly error definition for the assessment, the following error indexes are derived: Mean absolute error (MAE) [74]:

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|P_{m,h} - P_{p.h}\right|.$$

This measurement is a normalized mean absolute error (NMAE%) normalized to the net output power of the plant C. A maximum DC output power is measured over the whole period and expressed in watts [74]:

$$NMAE_{\%} = \frac{1}{NC}\sum_{i=1}^{N}\left|P_{m,h} - P_{p.h}\right|\cdot 100.$$

N is determined by the number of hours included in the calculation (i.e., 24 hours in a daily error basis calculation). When the mean absolute percentage error, MAPE%, is normalized, it is calculated as follows [74]:

$$MAPE_{\%} = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{P_{m,h} - P_{p.h}}{P_{m,h}}\right|\cdot 100.$$

Based on the total energy production, WMAE% is the weighted mean absolute error:

$$WMAE_{\%} = \frac{\sum_{i=1}^{N}\left|P_{m,h} - P_{p.h}\right|}{\sum_{i=1}^{N} P_{m,h}}\cdot 100.$$

Based on the maximum hourly power output $P_{m,h}$, the normalized root mean square error nRMSE is calculated [74]:

$$nRMSE_{\%} = \frac{\sqrt{\frac{\sum_{i=1}^{N}\left|P_{m,h} - P_{p.h}\right|^2}{N}}}{\max\left(P_{m,h}\right)}\cdot 100.$$

Envelope-weighted absolute error (EMAE%) is calculated as follows [74]:

$$EMAE_{\%} = \frac{\sum_{i=1}^{N}\left|P_{m,h} - P_{p,h}\right|}{\sum_{i=1}^{N} \max\left(P_{m,h} - P_{p,h}\right)}\cdot 100.$$

As defined by OMAE%, the Objective Mean Absolute Error is [74]:

$$OMAE_{\%} = \frac{\sum_{i=1}^{N} \left| P_{m,h} - P_{p,h} \right|}{\sum_{i=1}^{N} G_{POA,i}^{CS}} \cdot \frac{G_{STC}}{C} \cdot 100,$$

where $G_{POA,i}^{CS}$ and G_{SFC} are the irradiance under clear sky conditions and the irradiance in the standard test conditions, respectively [73].

1.6 Time frame

According to the forecast horizon, there are four types of PV power forecasting methods that have been developed [35]. The time horizon varies from a few seconds to a few minutes; the short-term forecasting ranges from 48 to 72 hours ahead; the medium-term forecasting ranges from a few days to one week ahead; and the long-term forecasting ranges from a few months to a year or more. There are different forecasting horizons used for different purposes, so for example, very short-term forecasters are used for the management of PV systems, in the electricity market, and for controlling microgrids. Power system operations, dispatching, and unit commitments are controlled with short-term horizons. Power system operations, dispatching, and unit commitments are controlled with short-term horizons. PV plants are usually maintained and planned using a medium or long-term horizon.

1.7 Overview of Different Methodologies

1.7.1 Numerical Weather Prediction (NWP)

Different types of forecasters can be selected depending on these parameters, for example physical methods which rely on using numerical weather prediction (NWP) models or satellite images which can be used for developing regional models [36, 37]. Long-term forecasting is usually based on meteorological data from satellite images [38]. NWP models can predict atmospheric conditions up to 15 days in advance without using historical data. Although these methods can be very accurate, they rely mostly on the stability of the weather conditions. It is generally difficult to implement physical models, in part because they require many parameters and expensive equipment that are not generally available. For solar forecasting, first-hour predictions are not especially useful for most of the available NWP [39].

1.7.2 Statistical Method

Probabilistic and statistical approaches are another option. There are many methods to accomplish this, including regression models [40], exponential smoothing, autoregressive models (AR), autoregressive moving integrated averages (ARIMAs), time series ensembles [41], and probabilistic approaches [42-44]. Short-term forecasting up to one day in advance is best achieved using statistical approaches. Multiple attempts were made to resolve the problem of non-linearity, for example, an extended application of seasonal time series ensembles for PV forecasts can be found in [41].

1.7.3 Artificial Intelligence (AI)

Advanced methods based on artificial intelligence (AI) techniques and machine learning (ML) [22] include artificial neural networks (ANNs), k nearest neighbor (kNN), extreme learning machine (ELM), support vector machine (SVM), etc. There are some advantages to these non-information-dependent methods that are typical of statistical approaches. Basically, they are used for short-term applications when field measurements are available [45].

An advanced method combined with a physical or statistical approach is termed a hybrid method. Due to the successful combination of two well-performing approaches, these types of forecasting techniques present a very good forecasting accuracy [46].

Models relying on weather forecast data, models using measurements taken on-site, and models combining measurements taken on-site with historical weather forecasts [8]. In recent years, hybrid models have become more prevalent [47].

1.7.4 Machine Learning

Machine Learning is the process of automating the learning process for computers based on experience (e.g., datasets) without being explicitly programmed by humans. ML algorithms fall into three categories: supervised learning, unsupervised learning, and reinforcement learning. Using input and output features as inputs, a machine learning algorithm tries to construct some relationships and dependencies between them. An unsupervised learning algorithm does not produce an output, but instead searches for patterns and rules in the available data in order to develop a better understanding of it. The reinforcement type is mainly used for visualization and analysis of high dimensional data which have been converted into lower dimensional data. There are also two types of learning problems related to this type: clustering and association.

1.7.5 Deep Learning

A deep learning system is one that generates knowledge for a computer by gathering and analyzing data, observing the world, and interacting with it. A relatively new technique in NN programming, deep learning (DL) represents a way to train deep neural networks (DNNs), since traditional NN-based methods may be affected by problems such as overfitting, diminishing gradients, etc. NNs are massively parallel-distributed processors that consist of simple processing units. NNs are naturally used for storing experiential knowledge and making it available. There are several layers of neurons in a typical NN, each one connected to those in the following layers. Data is fed into a neural network via an input layer, and the network responds to the input via an output layer. One or more hidden layers may exist between the input layer and the output layer. Essentially, any NN with more than two layers is called “deep”. On the other hand, DL is a semi-supervised training approach, suitable for DNN training. Between the input layer and the output layer may be one or more hidden layers. Deep NNs are any NNs with more than two layers. However, DL is a semi-supervised training approach, which is suitable for DNN training [48].

In the past few years, deep learning (DL) has proven to be very effective on a variety of problems, including speech and visual recognition, natural language processing, pattern recognition, automatic translations, self-driving vehicles, medical diagnosis, financial prediction, and automatic trading. In contrast, photovoltaics still has limited applications of DL. It's possible to automate the learning of complex mappings between inputs and outputs using DNNs. In DL, there are three main algorithms: convolutional neural networks (CNN), long short-term memories (LSTM), and others that are particularly useful in multistep forecasting [49].

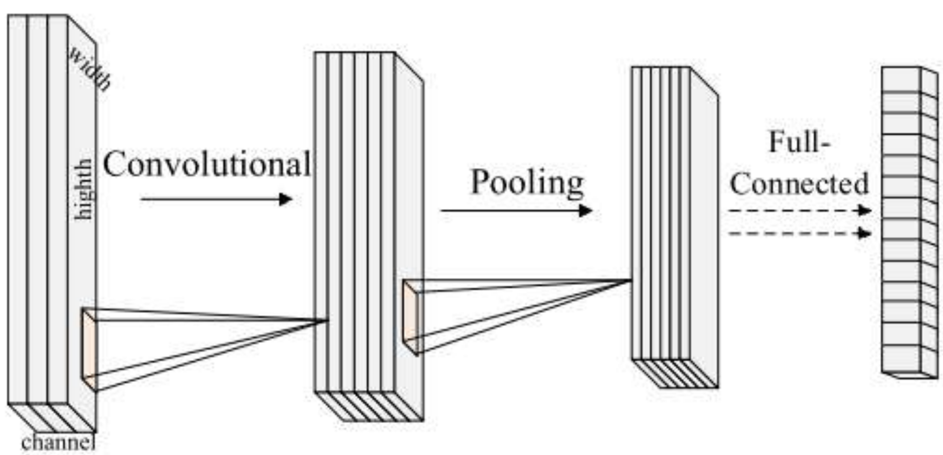

Figure 13. The structure of CNN.

Several types of layers are included, including convolutional layers, pooling layers, and fully connected layers [50, 51]. Each layer performs the following task:

The convolution layer of CNN is a fundamental component that generates new feature maps by using several convolution kernels. When all input maps are shared across the kernels, the convolution operation performs well in local feature extraction.

The pooling layer reduces the in-plane dimension of input maps, thus reducing the number of parameters that must be learned and avoiding overfitting. Pooling operations can be of various types, such as maximum pooling and average pooling.

Fully connected layer is often used to map the features from convolution layers and pooling layers to an output layer to perform high-level inference [52].

1.7.6 PV Power Output Forecasting based on Artificial Intelligence

A majority of methods (55%) use a machine learning approach, followed by hybrid approaches (36%). 48% of the DL techniques used are based on ANNs, including MLPs, RNNs, and RBFs [49].

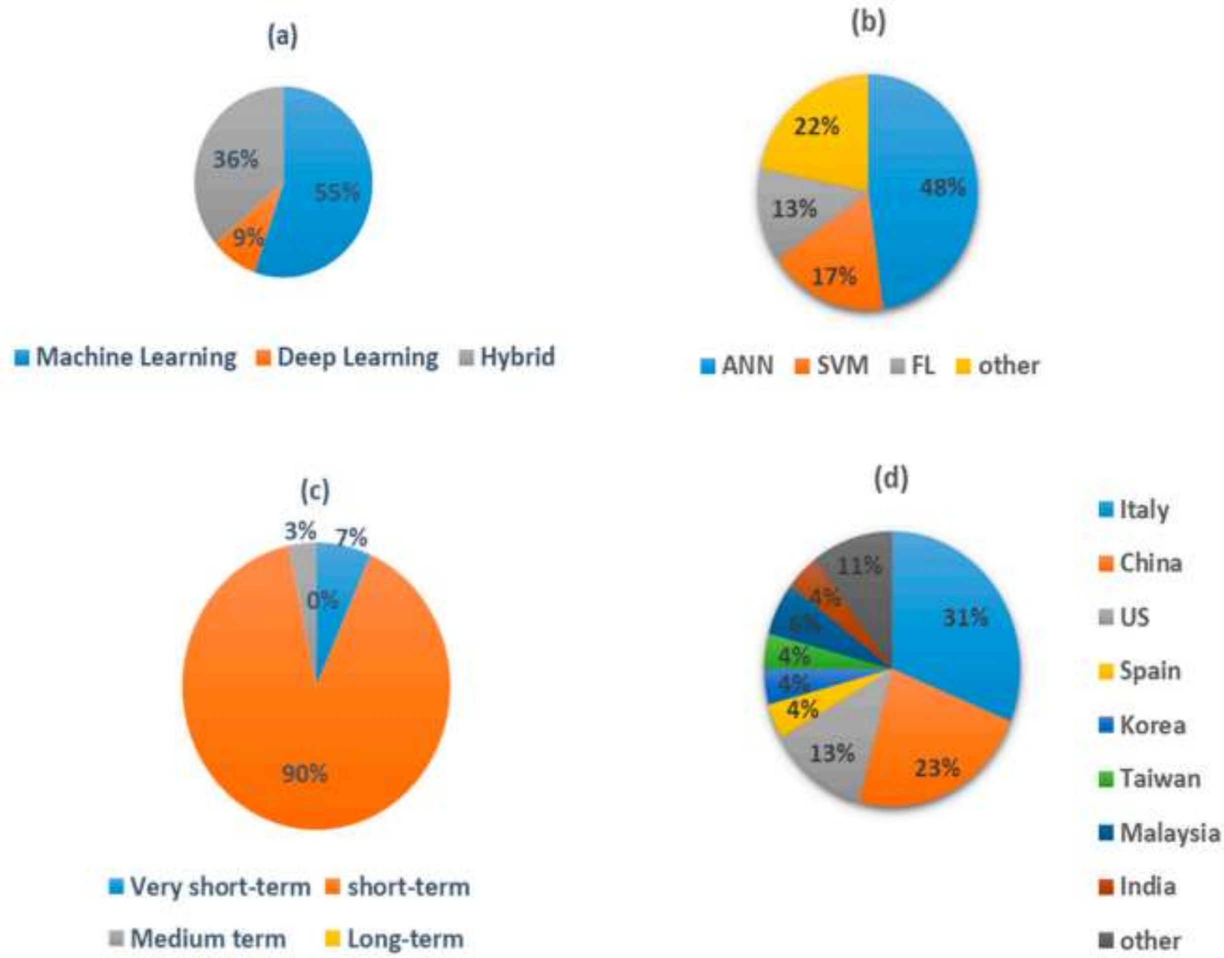

Figure 14. (a) AI-based PV power forecasting method; (b) the most common deep learning (DL)-techniques; (c) investigated time scale horizons, and (d) countries involved in PV forecasting [49].

Chapter 2 Methodologies

Forecasting methods can be categorized into different groups. Regarding the forecasted parameter, two different approaches have been implemented: direct and indirect.

2.1 Direct Methods

2.1.1 Multiple Linear Regression (MLR) Model

In the first prediction model, conventional MLR, the independent variables of the system were linearly related to the dependent variable using following equation.

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots \beta_k X_k + e$$

The Y variable is the variable to be predicted (independent), X_i represents k independent variables (where $1 \leq i \leq k$), b0 is the offset value of the predicted variable, and e represents the residual error term between the actual and predicted values. Since it was expected that it would be zero, the model did not consider it. Based on a constant level of all other independent variables, the regression coefficient for the β_0 independent variable provides an estimate of the expected change in the dependent variable per unit change in the β_i independent variable. MLR models have a number of difficulties, including their inability to extend their responses to non-central locations of explanatory variables. But despite its weaknesses, it is widely used because of its simplicity [31, 32].

2.1.2 Direct Method

Direct method is the easiest and cheapest method for forecasting output of the PV panels, it needs just weather conditions as an input also software or technique as an analyze such as Guess-sidel, MS Excel and MATLAB. [22]. Time interval in this method are different usually are used one hour and one day ahead but 30 minutes, 15 minutes and four hours are common [22, 23].

In this method cloudiness information can be found on open access websites such as www.meteoblue.com. The clear sky curve is scaled based on these data which exploit per hour. According to the season and location, PV output power generation varies. A close relationship exists between radiation, weather, sun angle, observation date, time and clouds. The PV output power generation is usually available from 8:00 am to 5:00 pm [22], [23].

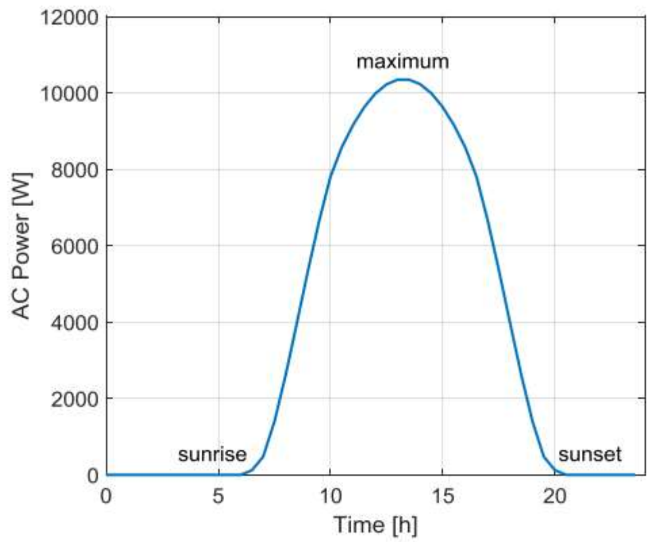

Figure 15. Exemplary clear sky curve. The curve daily using the incidence angle of the sun α

The clear sky curve is usually same for different locations and time during the year, but time categories are different, some researcher categories by time (cloudless, rainy, foggy, sunny), some others categories by the seasons.

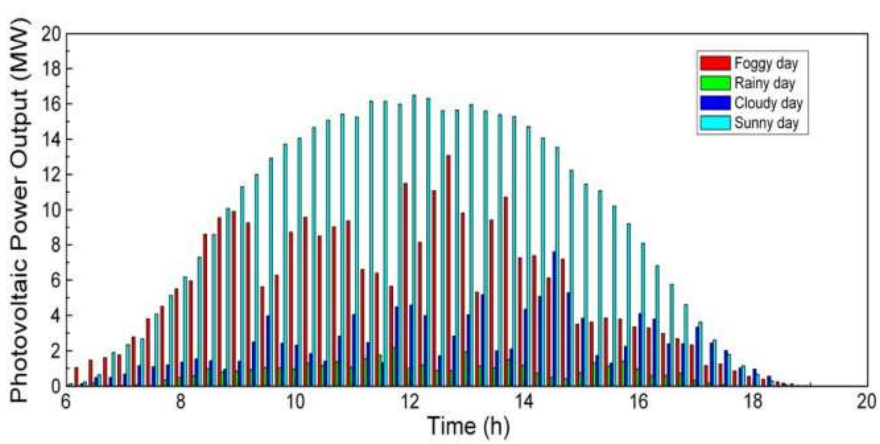

Figure 16. Photovoltaic power output in different weather conditions [23].

To assure comparability with commercial forecast data, MS Excel is automatically accessed with this weather data every day. A MATLAB loop is used to access the data and calculate the forecast or using different method for analyzing data such as such as Guess-sidel or Vector Machines [22], [23].

Usually clear sky curves are divided into six sections in order to accurately reproduce real generation profiles [22].

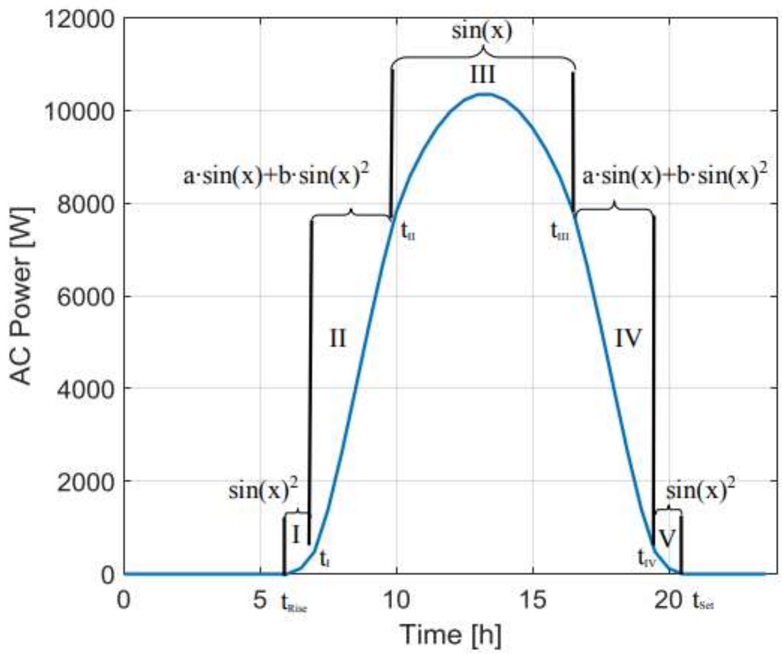

Figure 17. Sections of the clear sky curve.

$$m(t) = \begin{cases} 0 & t_{Rise} > t > t_{Set} \\ \sin(\theta)^2 & t_I > t > t_{Rise} \\ a \cdot \sin(\theta) + b \cdot \sin(\theta)^2 & t_{II} > t > t_I \\ \sin(\theta) & t_{III} > t > t_{II} \\ a \cdot \sin(\theta) + b \cdot \sin(\theta)^2 & t_{IV} > t > t_{III} \\ \sin(\theta)^2 & t_{Set} > t > t_{IV} \end{cases}$$

If classified weather in four different categories, A) Cloudless B) Lightly cloudy C) Medium cloudy D) Very cloudy, the results are presented in figure 14 [22].

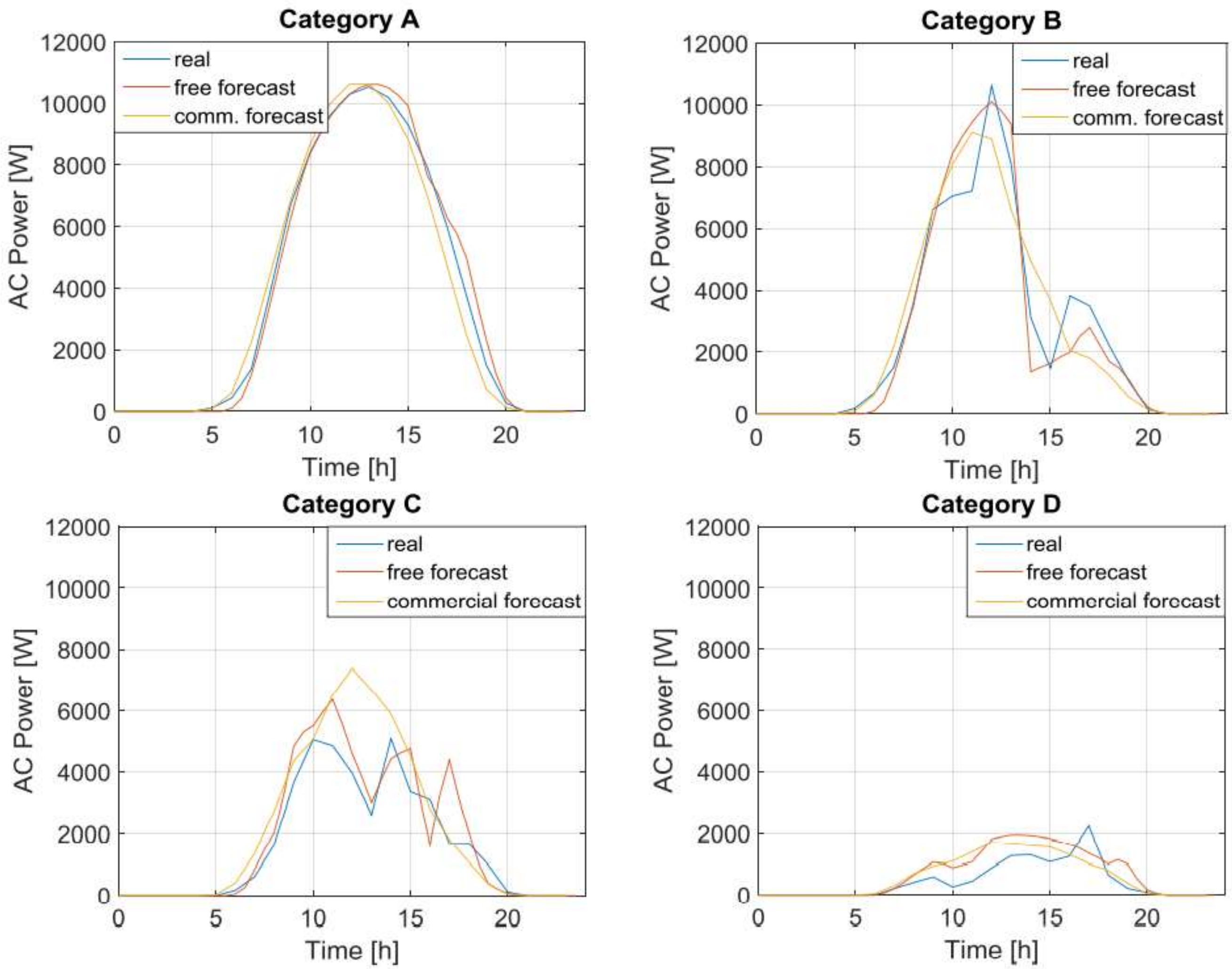

Figure 18. A comparison of predicted and measured generation curves on exemplary days

Moreover, the result for 50 days testing is presented in the figure 15. Error levels are higher at the peak. The peak errors can occur even on cloudless days. When there is a lot of wind, clouds can quickly move, even if they are merely a few and very thin. Results show the free forecast provides reasonably accurate predictions and, therefore, can be used whenever maximum accuracy is not necessary [22].

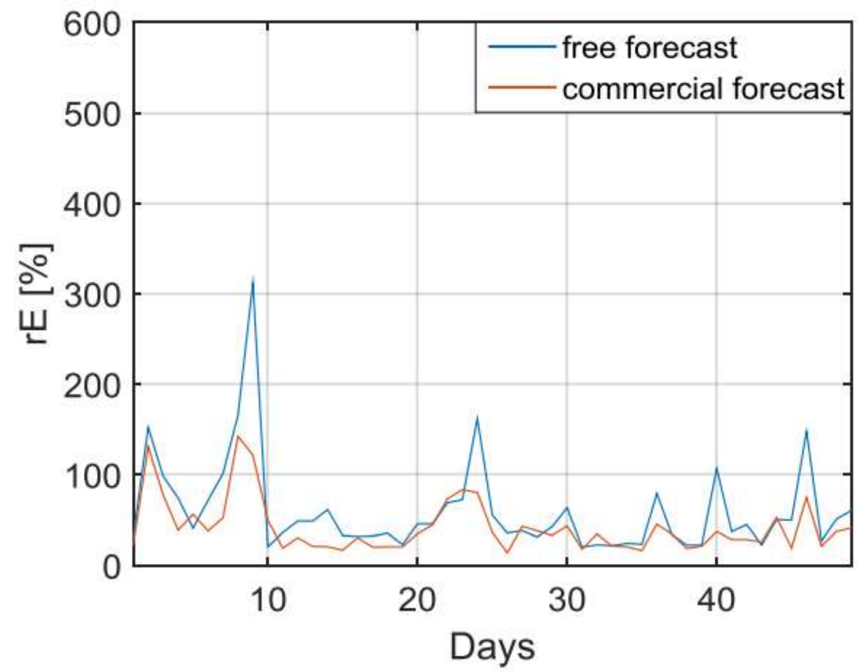

Figure 19. 50 days evaluated.

2.2 The Statistical Methods

Statistical methods depend on the given historic environmental data at the PV sites to generate their forecasting models. These models were developed for stationary time series, and thus they are not suitable to forecast PV power as solar radiation profile is non-stationary. These statistical methods are preferable for short-term and medium-term forecasting methods Statistical methods are multiple time frames. It means these methods can use for forecasting from very short time frame up to some days ahead.

2.2.1 Analog Ensemble Model

Some techniques are used in statistical forecasting, one of them is Analog Ensemble model. The AnEn model's goal; N measures make up an analog ensemble, and each analog can be viewed as a sample of the probability density function (PDF) of the PV power. From the satellite images, pixel-by-pixel estimates of GHI time series are produced [24].

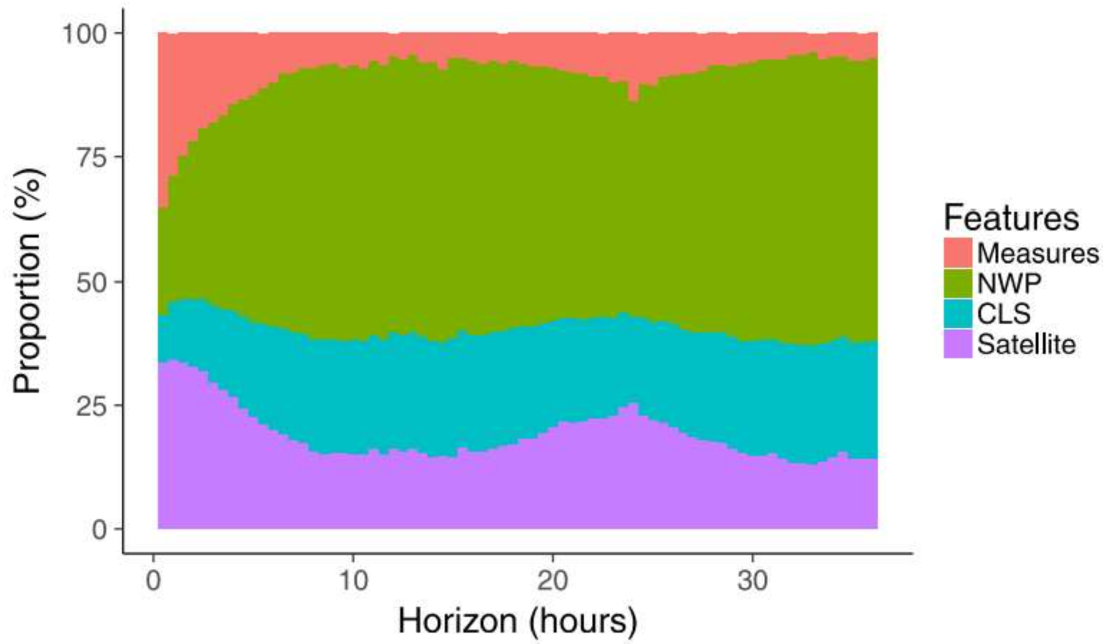

Figure 20. Time horizon and forecast for different start times. The CLS is the clear-sky profile [24].

The evaluation of probabilistic forecasts is more difficult than that of deterministic forecasts. Predictive accuracy error is between 10.1% to 12.8%. The proposed model has been shown to be an efficient alternative, both in terms of computation cost and accuracy, to the current approach, which employs different forecasting models depending on the data types and the application time horizon [24].

2.2.2 Long Short-Term Memory (LSTM) and Quantile Regression Averaging (QRA)

LSTM networks are capable of storing time series with long intervals and delays. Compared to common neural networks, LSTMs are better at forecasting and analyzing short-term and ultra-short-term data. In an LSTM network, units are connected to one another.

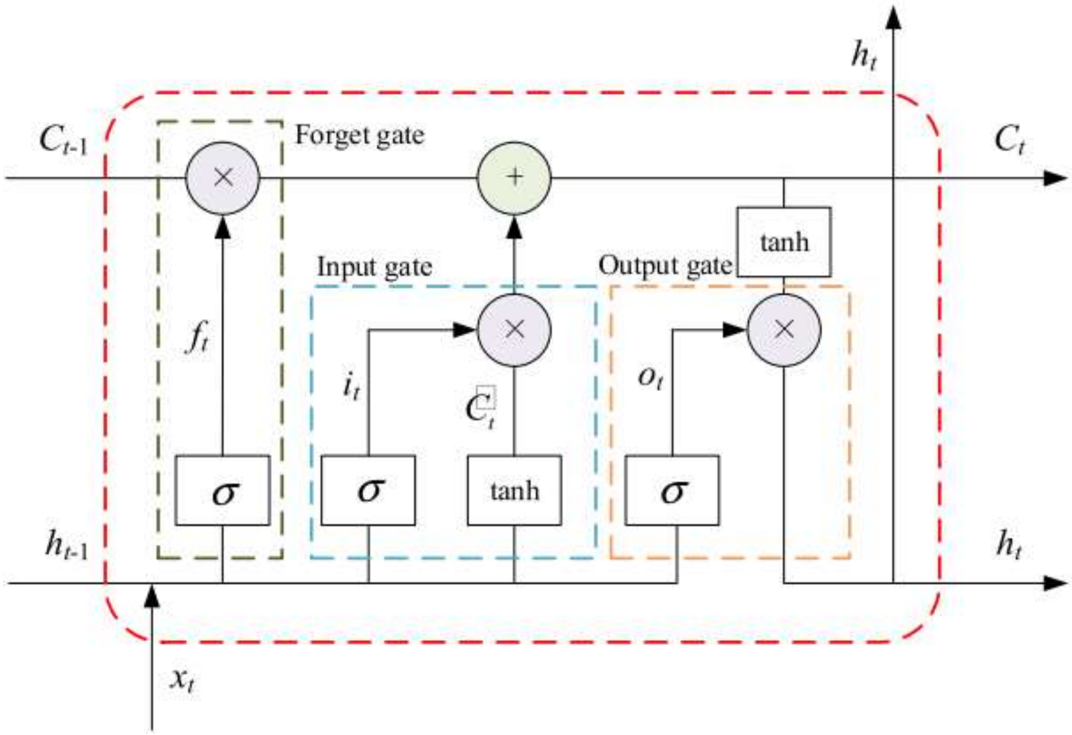

Figure 21. An overview of the LSTM structure [26].

An LSTM unit is made up of three gates: the forget gate, the input gate, and the output gate. Input x_t comes from the current series, output h_{t-1}represents the last time, and the current time is inputted by the current series C_{t-1}. h_t and C_t are stored for the next run [26].

LSTM network parameters are determined by the number of hidden layers, except for the input parameter and output parameter. A fully connected layer is usually added to the end of the output gate because the number of hidden layer units and output dimensions are usually different.

Another method that can use and combine with LSTM method is called Quantile Regression Averaging (QRA). QRA ensemble multiple independent deterministic forecasting models for a probabilistic forecasting model [27, 28].

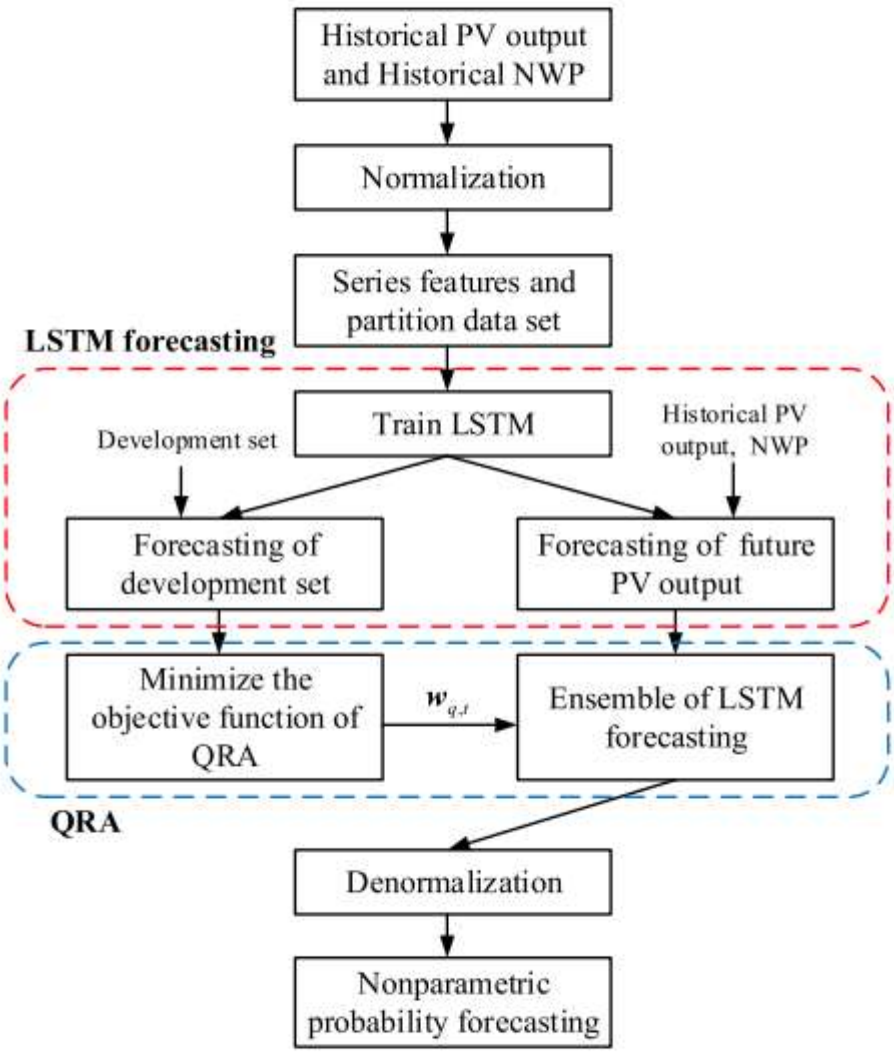

Figure 22. Process of LSTM-QRA [26].

Data leakage can be avoided by normalizing the test set by the maximum and minimum values (scaling factors) from the training and development sets. A set of independent LSTM forecasting models is constructed using ten different numbers of hidden units. In the development and test sets, these trained LSTM networks are used to predict the deterministic prediction value [26].

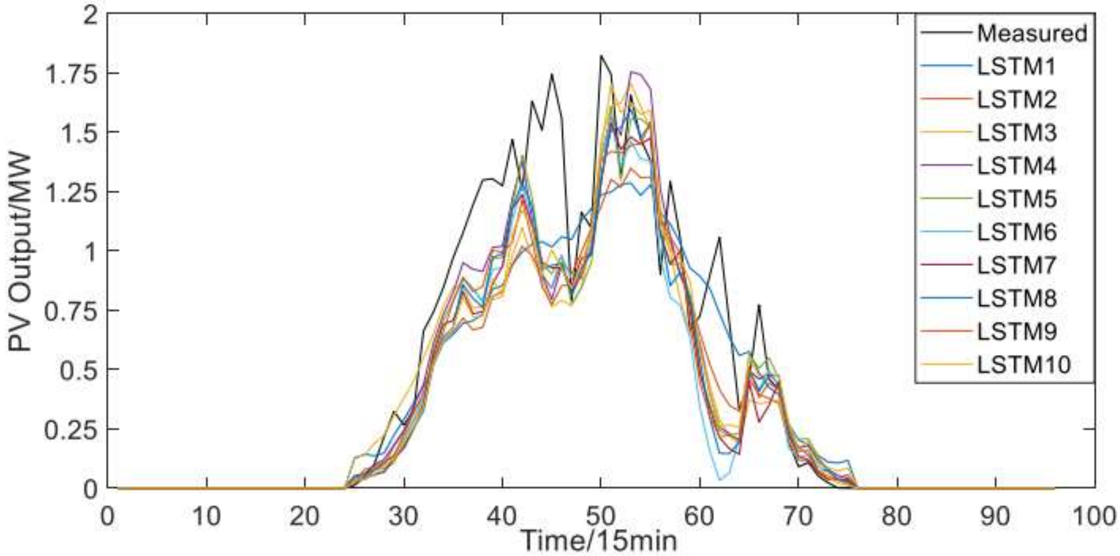

Figure 23. Each hidden value of LSTM predicts a different outcome [26].

2.2.3 Spatio-Temporal

A complete probability density function of PV production over short-term horizons (0–6 h) is provided by the model. In this method, input variables are selected automatically using quantile regression and L1 penalization. A simple modeling chain allows the model to be fast and scalable for direct on-line use.

Analysis of the correlation between stationarized production series can determine the existence of a spatial-temporal pattern. There is no effect of the sun position in these series on crosscorrelations.

using two auxiliary methods, Kernel Density Estimation (KDE) and the Extreme Learning Machine (ELM) method. The two respective models use only the data local to the PV installation of which a power forecast is made [30]. The proposed spatial-temporal probabilistic approach can therefore be evaluated by using these models. Advanced spatial-temporal methods are based on quantile regression approaches that are adapted to problems with high dimensionality. KDE is a non-parametric approach to density estimation that is significantly more accurate than parametric methods [29]. Unlike conventional density forecasting methods based on feed forward neural networks, Extreme Learning Machine does not require tuning of the hidden layer parameters (the parameters are randomly chosen, making it fast) [29], [30].

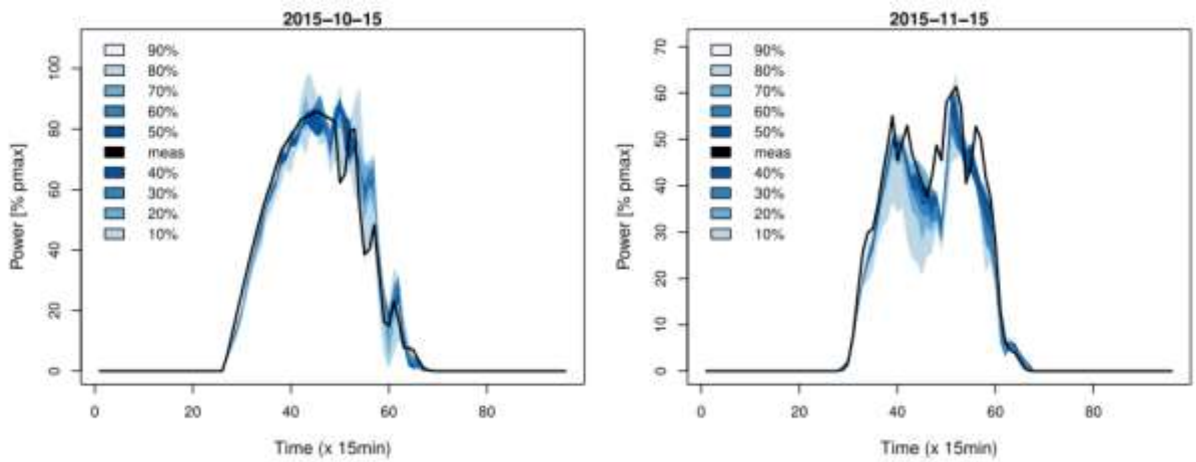

Figure 24. Predictive densities forecasted using Spatio-Temporal method

The forecast is updated every four hours and has a time resolution of three hours. They were interpolated to 15-min intervals. Predictive accuracy error is between 7% to 12.88.32%.

2.3 Supervisory Control and Data Acquisition (SCADA)

To improve the forecast, this method uses a physical approach and a few statistical techniques.

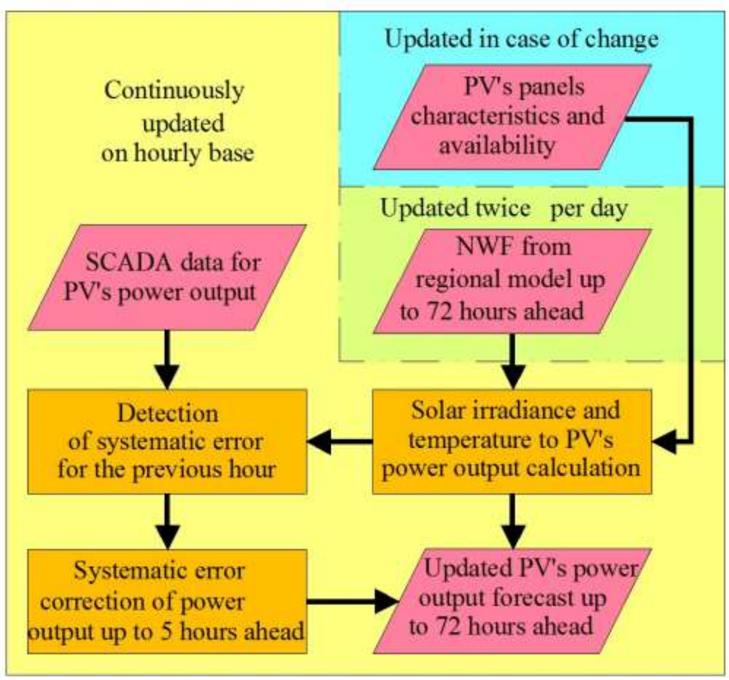

Figure 25. SCADA forecasting system functions at a basic level [6].

This method uses four different data as an inputs: a) Distributed systems operators (DSOs) and big photovoltaic producers provide live SCADA data to distribution grid operators for small photovoltaic installations. b) Numerical weather forecast (NWF) for solar irradiance and temperature, up to 72 hours ahead for PV locations. c) PV’s technical characteristic available from manufacturers. d) PV operators provide PV's operators with information regarding the number of panels available [6].

In these cases, forecasted cloud indexes in combination with daily standard profiles are suitable for forecasting power output. Aside from the tracking and temperature systems, PV's power output is affected by wind cooling, dustiness, snowbound, and shade effects, but forecasting them is very challenging.

Forecasting accuracy depends on some factors, such as: Data collected from SCADA: available panels, solar irradiance, temperature, wind speed, and cloud photo opportunities - very short time horizon; detecting shadow effects - snow, dust, terrain spatial properties; collecting data from satellite and earth photos of clouds; detecting shadow effects - snow, dust, and terrain spatial properties [6].

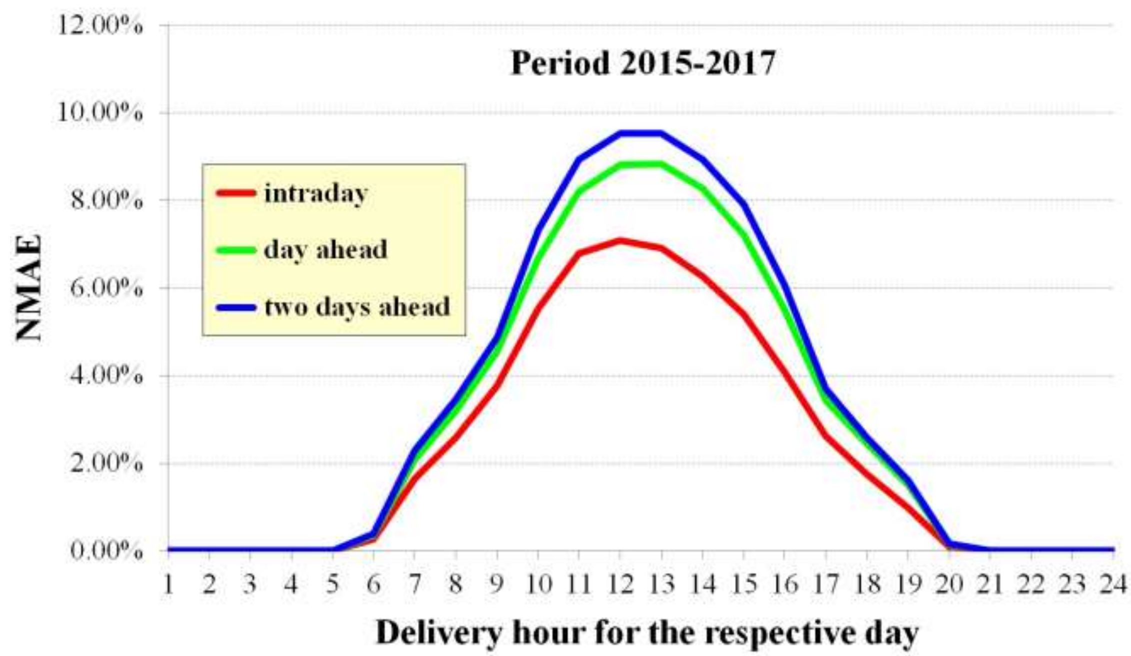

Figure 26. Forecasting accuracy by different time horizons [6].

Accuracy of this method is depends on the time horizon. Generally, it considers between 7 to 9.5%.

2.4 Hybrid Method

In order to avoid all the problems associated with large datasets, a stepwise regression analysis (SR) methodology was used to determine the relative importance of each meteorological parameter. The SR procedure is used when no predictors are provided in the model, and attempts to predict independent variables to become part of the model. The model first identifies candidates as independent variables. Basically, it calculates the F-statistic that provides a means of comparing and assessing the significance of each independent variable p-values. It evaluates whether the regression coefficient equals zero as a null hypothesis. A p-value that is below 0.05 indicates the null hypothesis might be rejected. By contrast, a high p-value indicates that the independent variable is statistically insignificant. If the independent variable has a p-value less than an established cut-off value, it can be entered into the model. F-statistics are again calculated once another independent variable is added. As a result of adding new variables, the SR procedure also determines whether any of the variables already in the model should be removed in order to maintain their significance. As long as the p-value of the largest variable exceeds the cut-off value, this variable is removed from the model. This process continues until all variables remaining in the model are significant at the cut-off level, and all variables excluded are insignificant. After identifying the most relevant input variables, the multi-stage models were used as inputs, resulting in the SR-FFNN, SR-GRNN, and SR-MLR hybrid models.

Hybrid genetic algorithms ANN-based perform better than other investigated techniques such as kNN, ANN, and ARIMA-based. In this study, a hybrid model was used to forecast short-term power from a PV system installed on the roof of a municipal building in Trieste, Italy. A seasonal auto-regressive integrated moving average method (SARIMA) is used in conjunction with a SVM approach. The result was that the designed hybrid model performs better than the SVM and SARIMA model working alone [67].

A hybrid genetic swarm optimization (GSO)/error back propagation (EBP) approach based on genetic swarm optimization was developed in order to improve the speed of convergence and the forecasting accuracy. A small-scale plant installed in Milano, Italy was tested using this method for 24-hour power forecasting. As a result, performance improved in comparison with the standard EBP alone [68].

In another hybrid methodology, SVM is combined with self-organization map (SOM) and fuzzy logic (FL) for a one-day-ahead forecast. To select the most accurate training model, the FL was used in conjunction with a SOM to classify weather types, an SVM for the training phase, and a SVM to select the most appropriate training model for the SVM. Taiwan Central Weather Bureau (TCWB) data was used to develop the method. In both SVR and ANN-based forecasts, our hybrid model performed better [69].

The physical hybrid neural network model (PHANN) is another hybrid approach that uses dual layer ANNs with the Clear Sky Solar Radiation algorithm to improve the accuracy of the day-ahead forecast compared to ANN alone. Additionally, the PHANN method has been shown to always outperform the physical five parameter equivalent model of a PV module located in Milan, Italy [70]. In this study, physical neural network models (PHANN) and network optimization (SNO) were used to improve the accuracy of some weather forecasting. By using the method, the prediction accuracy of power production by a PV plant at Milan, Italy, was improved significantly. With different time horizons, the same hybrid method had better results [71].

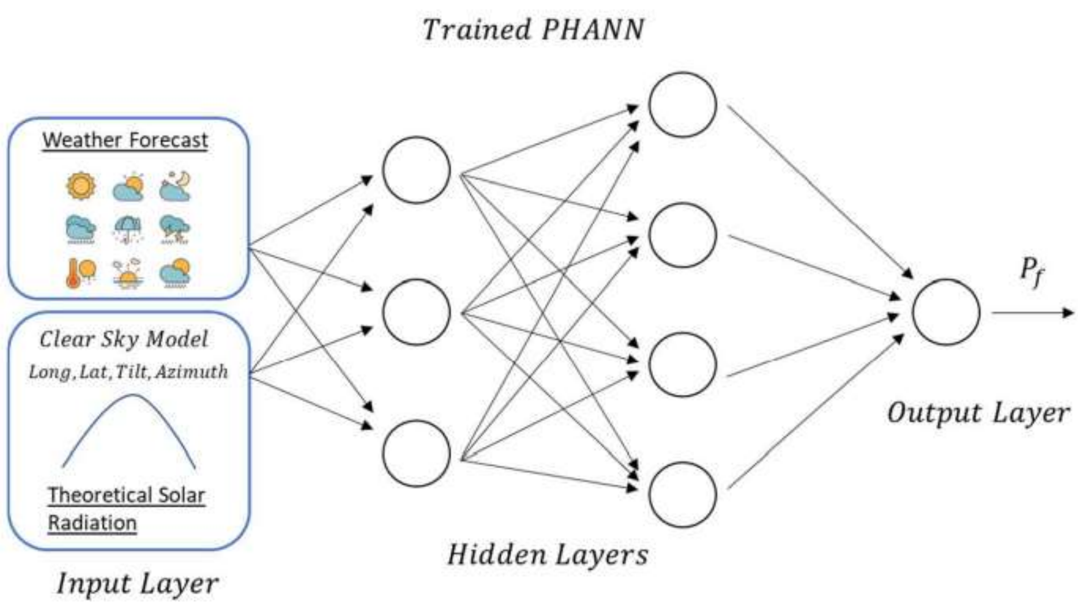

Figure 27. PHANN method [74].

The short-term power forecasting of the 100 kWp plant installed in Nanjing, China, was improved using a hybrid method consisting of a VMD together with a CNN. The PV power time series was decomposed using the VDM into different frequencies and into 2D data while the power was estimated using the SVR model. Comparing the method to other 1D VMD-based forecasting approaches, the accuracy of the prediction has been shown to be higher [72].

2.5 Artificial Neural Network Models

There are many definitions of Artificial Intelligence (AI) and one of the most used is imitating intelligent human behavior that includes systems with the following characteristics: they have a "thinking process similar to the one human beings have, they are able to act like humans, their thinking process is rational, and their acting is rational too. In the other hand, Artificial neural networks (ANNs) are models inspired by the human nervous system made up of highly connected elements called neurons. Information is processed by a neuron from a number of inputs to produce output. An input layer and an output layer are the simplest networks. It can only handle problems with linear relationships between inputs and outputs. More complex problems require the addition of an additional hidden or middle layer of neurons. Depending on the size and nature of the data set, the number of hidden layers and neurons within each layer will vary. In order to construct the ANN, the input layer, hidden layer, and output layer must be determined, as well as the connection weights, biases, activation function, and normalization function. Several network architectures may be formed depending on the arrangement of neurons and their activation function. The FFNN and GRNN architectures have been successfully applied to the prediction of PV output power [32, 33]

The FFNN data can only be moved from an input layer to an output layer through a hidden layer [34]. The FFNN performs a certain task by training with historical data from a designated period of time including an error range. Training involves adjusting the network weights so that the output matches the desired target [32].

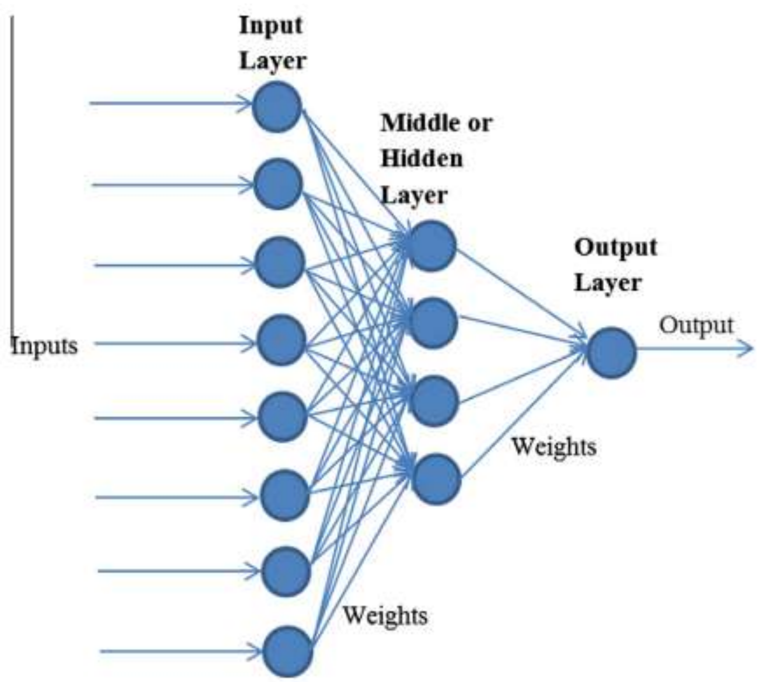

Figure 28. Four layers hidden in an FFNN [32].

A GRNN is a probabilistic network that is used to perform regression rather than classification. There are four layers in a GRNN, with neurons in the input layer and radial neurons in the output layer. Clustering takes place on the training data using the radial layer. For a given dataset, the number of neurons in the radial layer is equal to the number of samples in the dataset. There is one more neuron in the regression layer than the output layer, and the activation functions are linear. There is one more neuron in the regression layer than the output layer, and the activation functions are linear. The estimation can be done directly from the training data by estimating an approximate function between the input and output vectors. GRNN requires less time to design than FFNN [32].

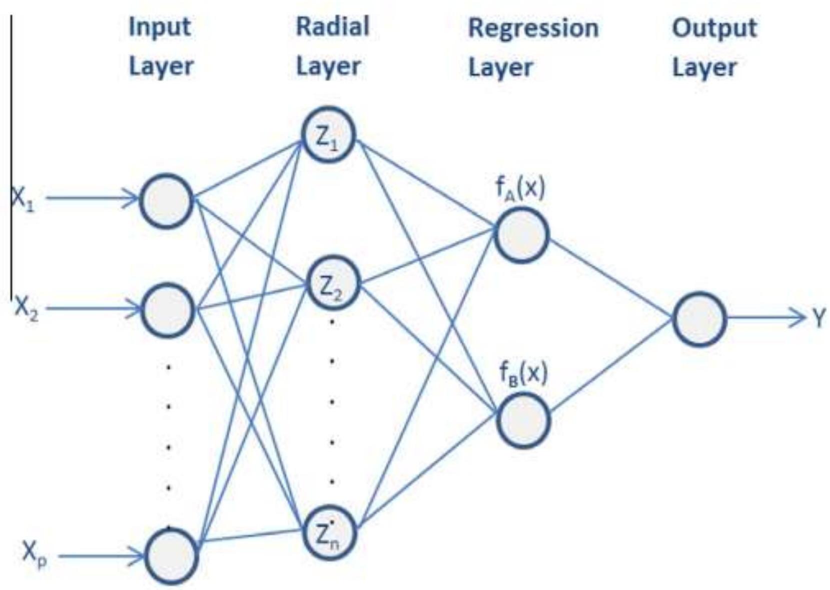

Figure 29. An overview of the GRNN network architecture [32].

Artificial Intelligence, Machine Learning and Deep Learning are linked together as a bellow figure.

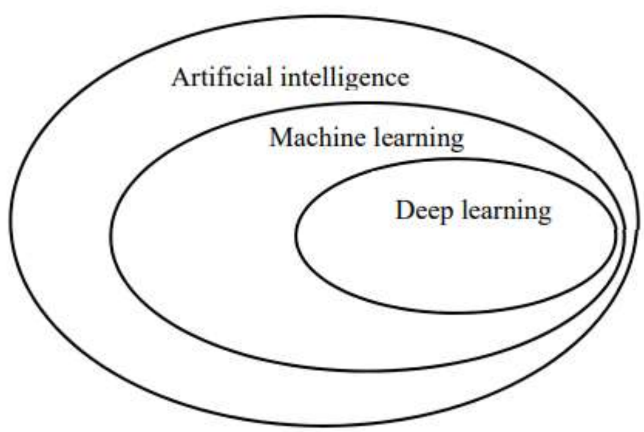

Figure 30. Relation between artificial intelligence, machine learning and deep learning.

2.6 Machine Learning

A method is presented in [50] for estimating the power produced by a PV plant 24 hours ahead in Italy. A multi-layer perceptron neural network (MLP) is used as a basis for computing the forecasted solar irradiance. There is a mean absolute error (MAE) of less than 5% across all studies examined. It was concluded that a better model accuracy could be achieved by including more input parameters such as cloud cover, pressure, and wind speed. Improved data quality and size can also improve accuracy [50].

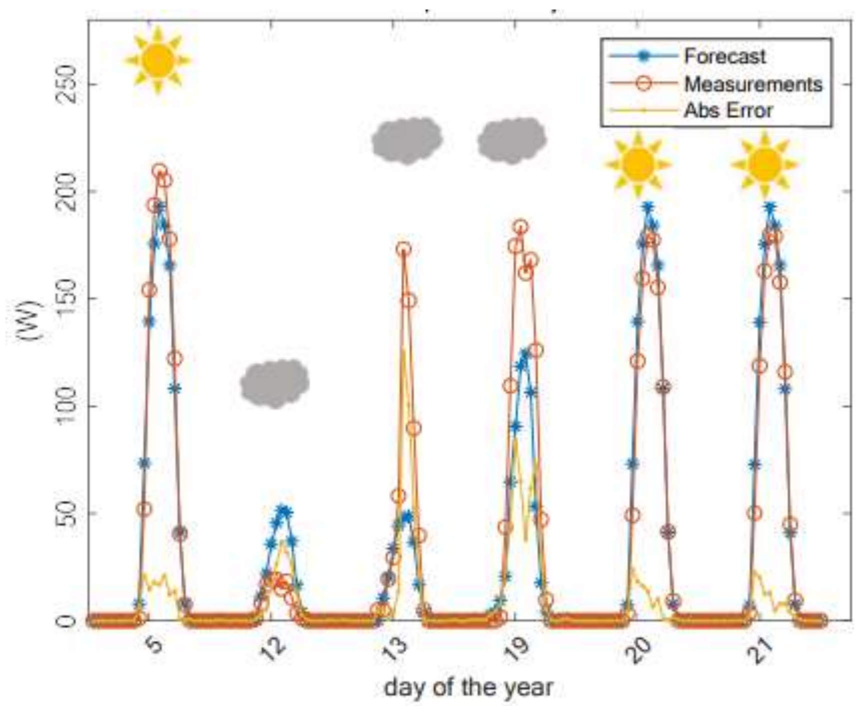

Figure 31. Comparison of measured output power and forecasted output power (sunny days and clouds)

As an alternative method, self-organized maps (SOMs) are used to classify the input variables. Solar irradiance, temperatures, wind speeds, and historic power values are inputs into the developed model. Based on the results of this model, a MAPE of 8.3% corresponds to sunny days, while a MAPE of 54.4% corresponds to rainy days. The model was tested using a 28 kWp PV plant [51].

Using the SVM technique, the method takes into account a weather classification, along with historical data of power usage and weather forecasts for the next day. Four models are developed for clouds, foggy days, cloudy days, and rainy days. In this case, the MAE of 4.85% was achieved for sunny days, while MRE was 8.56% on average [52].

In [52], describes a short-term ANN-based forecaster that is used to predict electrical energy generation and weather, based on historic data and forecasts from a NWP model. The MLP-NWP outperformed other investigated forecasters such as ARIMA and K-NN, while ANFIS and MLP-NWP based models outperformed all others. The root mean squared error (RMSE) for different forecast horizons from 16-39 h was 11.79%, while the average mean absolute error (MAE) was 6.41%.

In addition to the NN, authors developed a simple method of forecasting the power generated by a small PV plant in Istanbul, Turkey, for a short- to medium-term time frame. The model has only been designed and verified using the historical powers that work well only for a time horizon of between 5-40 minutes. Support vector regression (SVR) is used in the analysis of the power from a 1 MW plant located in Kitakyushu, Japan. Different inputs are accepted for the model, including cloudiness levels, humidity, and extraterrestrial insolation. The model has performed badly on partially cloudy days, as shown by its RMSE of 0.0948 MWh. The authors claim that relative humidity and air temperature do not need to be used as inputs to the model [54].

An SVM-based model was developed in another method for the short-term forecasting of power production from three PV plants in three different regions of the United States. Data from the National Solar Radiation Database (NSRDB) were used to develop the model. Using a sigmoid function-based normalization of solar irradiation, the study shows better results than normalizing solar irradiation with respect to transmission. An improvement in accuracy was found using the proposed technique when compared to a RBF-NN-based model. By adding information about the sky cover, the accuracy of the predictions can be improved [55].

Using a SVM-based forecasting method is another method that SVMs can also be used for forecasting one day in advance. As inputs, the model receives historical power data as well as some meteorological data. The method was tested with clear-sky, cloudy, and rainy conditions to forecast PV power generation. The proposed model's MAE value of 34.57% indicates good forecasting capabilities [56].

Using historical data and weather forecasts, a MLP-based forecaster has been trained. When the weather was sunny, the model performed better than when it was partly cloudy. All of the investigated cases had a normalized MAE below 15%. For the estimation of power produced by a photovoltaic plant installed in Italy, a new upscaling approach was developed. By using satellite and NWP data, it is possible to estimate solar generation at a regional level. An application of the method was tested on the power generation of solar panels installed in 1985 in South Tyrol, Italy. RMSE for the 4-hour estimation ranged from 5% to 7%, and for the 1-day estimation ranged from 7% to 7.5% [57, 58].

The ultra-short-term forecast for a small-scale PV plant was achieved by using SVM and MLP. The model has shown to be very efficient and particularly suitable for fog and haze conditions. Their forecasts are based on measurements of air temperature, humidity, and aerosol indexes taken on site [59].

A machine learning algorithm has been developed in order to train a MLP network that forecasts power generated by a PV plant located in Amman, Jordan 24 hours in advance. In terms of accuracy, ELM performs better than back propagation (BP). The smallest MAE (1.08%) occurred in June, and the largest (18.83%) occurred in February and March [60].

The model of echo state networks based on multiple reservoirs (MR-ESN) is presented in another forecasting algorithm. The reservoir parameters were optimized using the quasi Newton algorithm. As a result of the results, the MAPE of the model was very close to zero (0.00195%), and, in terms of one-hour forecast horizon, the model performed better than the SVM, back-propagation neural networks (BPNNs), support vector regression (SVR-ANN), and wavelet transform (WT) [61].

Another research statement proposed a multimodel ELM-based forecaster for predicting the amount of electricity produced by a 250kWp PV plant in Beijing, China. Using an annual single model, the MAE was 2.43% in spring and 1.81% in summer, while the MAE for the multi-model was 2.13% in spring and 1.7% in summer. In order to improve the accuracy, the multi-model is designed so that it can take into account the fluctuation of power output [62].

2.7 Deep Learning

There have been five LSTM-based neural networks designed to predict the hourly PV output power in [66]. In comparison to other methods, the model that relies on historical powers and doesn't include meteorological data has been shown to reduce forecasting errors [63].

The use of a deep neural network to forecast PV power from grid-connected photovoltaic systems in Seoul, Korea, has been described as a six-layer feedforward model. Based on its performance, the method, which does not rely on on-site sensors, has been improved over other models using local measurements. Nonetheless, the errors were not satisfactory during the summer and cloudy weather [64].

A study was conducted by the authors comparing different deep neural network-based one-day ahead forecasters. We analyzed two kinds of neural networks: conventional neural networks (CNNs) and linear support vector machines (LSTMs). Researchers have shown that the accuracy of the three models depends mainly on the size of the database available. Overall, the experimental results demonstrate that deep learning networks have a good effect on the prediction of photovoltaic power generation and are stable and robust [65].

Researchers in South Korea have developed a recurrent LSTM-based method to forecast the day-to-day power output of a PV plant. Among the inputs for the model are the solar irradiance, the atmospheric temperature, and the cloudiness index. Results showed that the seasonal-ARIMA model performed best in comparison to other approaches based on DNN, ANN, and auto regressive integrated moving average (ARIMA). The performance of LSTMs is particularly impressive when the output power is inherently unstable [66].

One of the newest method by high accuracy is Using Satellite Visible Images and Modified Convolutional Neural Networks. The whole algorithm process is shown in the following figure.

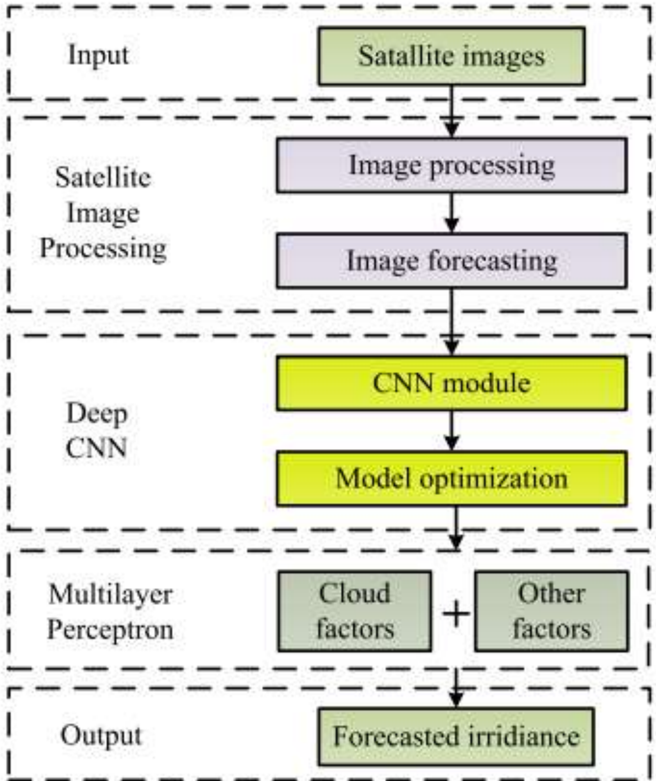

Figure 32. Framework of the method.

In order to reduce diurnal effects caused by the solar zenith angle, satellite visible images are first preprocessed. An image processing method that is more detailed is proposed to replace the traditional formula. The detailed method has the advantage of better eliminating the effects of solar zenith angle while retaining useful data. There is then a proposed method for cloud motion forecasting that can be calculated by using forecasted wind velocity and cloud top temperature. To modify the traditional model, CNN structures with varying widths, depths, and pooling methods are examined. According to the modified CNN, the results are more precise than the traditional method. Solar irradiance also depends on historical irradiance data, meteorological data, and solar zenith angle, in addition to cloud factors. To establish the mapping relationship with irradiance, factors that influence irradiance and factors that influence cloud cover are combined.

Standardizing the pixel intensity of the visible image is the first step. E represents the pixel intensity value, K represents the solar constant, represents the albedo, represents the solar zenith angle, and A represents the experimental coefficient, which depends on the satellite's relative position when viewed from the observation station. By transforming the second formula into the third one, the impact of the solar zenith angle on pixel intensity can be reduced.

$$E = K\rho(\cos(\theta))^A$$

$$E' = \frac{E}{(\cos(\theta))^A} = K\rho$$

$$E'' = \frac{E' - E'_{\min}}{E'_{\max} - E'_{\min}}.$$

As a second step, cloud information is extracted from the cloud image after it has been corrected for solar zenith angle effects. The satellite image processing framework is shown in the following figure.

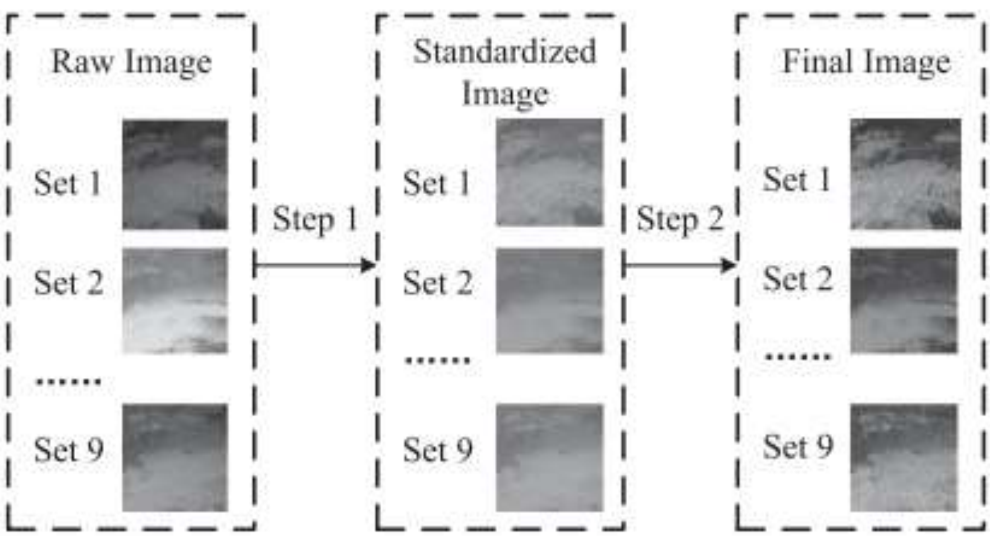

Figure 33. Structure of the satellite image processing system.

In addition to approximating complex mapping functions, ANNs have been extensively used in many areas, including solar irradiance modeling and forecasting. ANNs are generally composed of three layers: the first layer contains inputs, the last layer contains outputs, and the hidden layers are composed of hidden neurons. This article uses the cloud cover factors obtained from convolution and other influencing variables as inputs to the ANN to establish the relationship between forecasted irradiance and cloud cover factors.

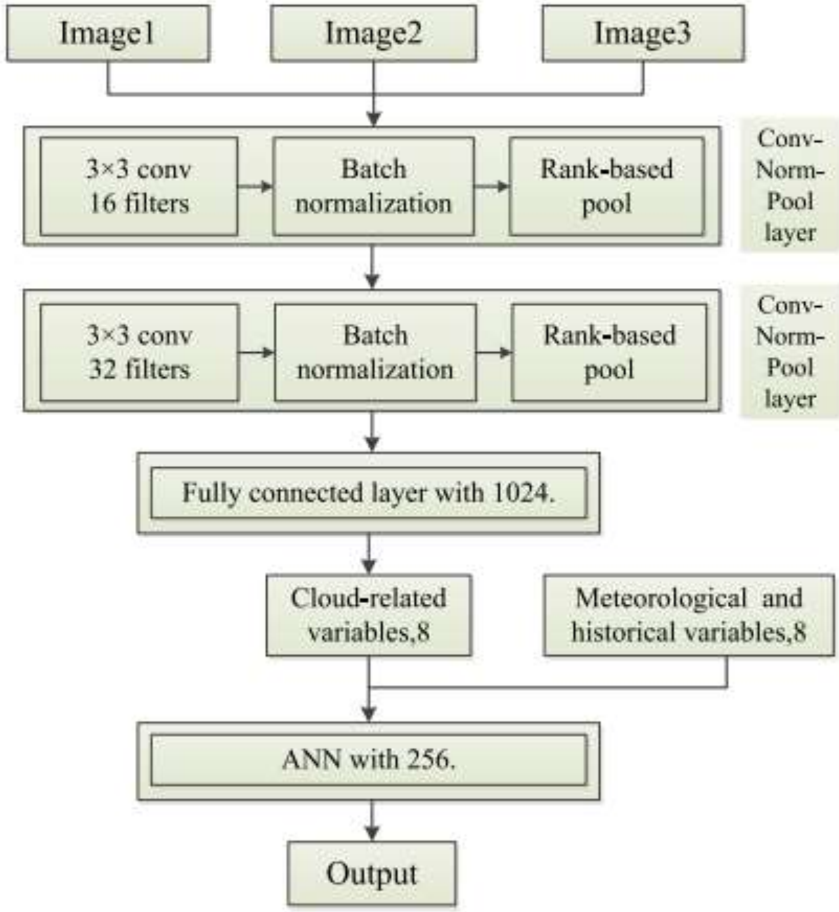

Figure 34. Forecasting model

Two convolution-pooling layers are applied to the three cloud images at three consecutive moments. In this order, each convolution pooling layer consists of a convolution layer, a normalization layer, and a pooling layer. First, we select 16 filters, and then we select 32 filters in the second convolution layer. Every filter has dimensions of 3 * 3 and steps of one with a same-value padding. Researchers have shown that the normalization layer follows the convolution layer, and the presence of the normalization layer can accelerate the training process. A rank-based average pooling algorithm [38] is used in the pooling layer, with a size of 2 * 2 and step size of two. In addition to retaining information about convergence, the average pooling layer can provide some translation invariance for inputs. Following two layers of convolutions, there are two layers of 1024 neurons each. Lastly, several cloud-variables related to cloud occlusion factors can be obtained from fully connected layers.

The results from the evaluations are shown in the following figure.

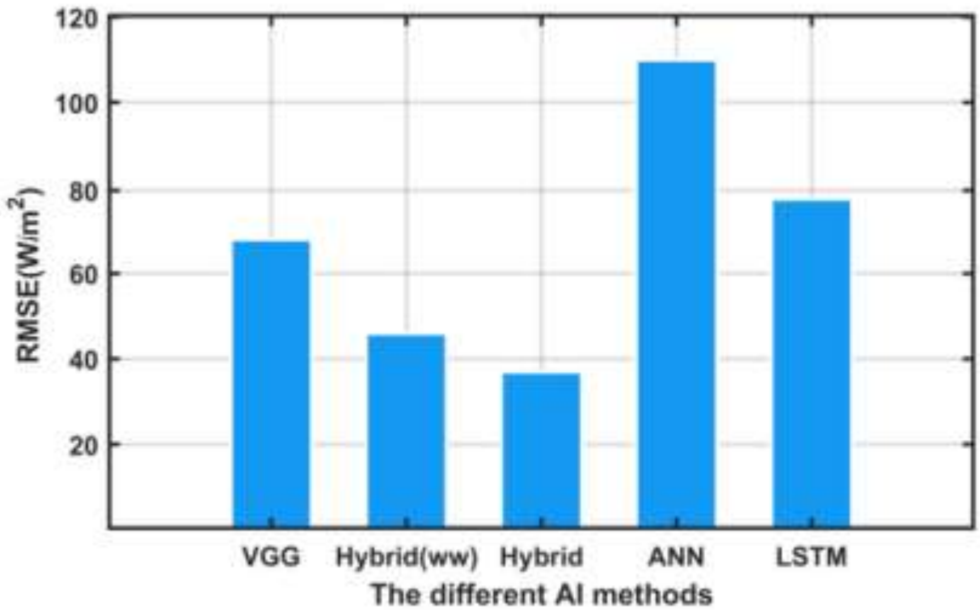

Figure 35. RMSE of the different AI methods

The Visual Geometry Group (VGG) is a popular structure of the CNN, and is well known for its superior performance. A hybrid model (ww) consists of only cloud image information without weather and history information. ANN and LSTM are utilized to build the model. The performance is worse than that of the hybrid method, which shows the strength of CNN in image feature extraction.

As it is explained in the first chapter, one of the most important criteria forecasting is GHI, in physical method the accuracy of forecasting method of GHI is shown in the bellow graph.

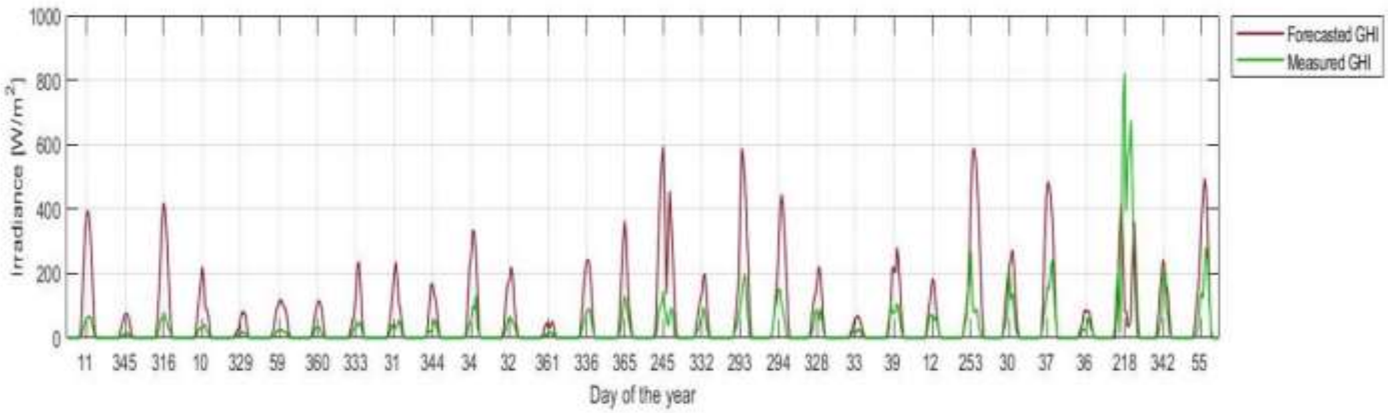

Figure 36. Physical models predict and measure GHI

According to some models, the GHI can be predicted with a mean absolute percentage error of approximately 3% on a sunny day and 23% on a cloudy day. According to the sensitivity analysis of the multiple feed-forward neural network GHI forecast model, solar angles (azimuth angle and zenith angle) and extraterrestrial irradiance can also be used to predict GHI. The GHI neural network forecast model can also be concluded to be more sensitive to the sun shine, cloudiness and temperature than to other inputs [75].

The evaluation of probabilistic forecasts is more difficult than that of deterministic forecasts. Predictive densities require many properties, whereas some aspects of forecasts cannot be identified with only proper scoring rules. The main required properties are reliability and sharpness.

In order to evaluate the reliability of a forecasting system, one must first assess its accuracy. If the levels of each quantile correspond to the frequency of observations, the forecasts are reliable. To be perfectly reliable, the empirical and nominal quantile levels should be the same, so that the reliability diagram is a diagonal line.

In the forecasting process, sharpness is measured as the spread of the distribution. Dirac distributions, for instance, have perfect sharpness, while uniform distributions have low sharpness. The goal of any probabilistic forecasting system is to maintain its reliability while still being as sharp as possible.

Direct method is cheapest and simplest method for forecasting and it usually use in low power networks. The accuracy of this method is usually acceptable but not in compare with different methods. Moreover, it is use in somewhere that high accuracy is not necessary.

On the other hand, this method successfully predicts data and is therefore applicable whenever there is no need for maximum accuracy. On all evaluated days, the commercially available forecast is slightly more accurate than the model, but its superiority is evident in days which are difficult to predict. In the field testing of the research, where only rough predictions in major time steps are needed, the direct method forecast is sufficient for use, and therefore may also be suitable in many other situations.

In spite of this, the introduced method for producing a direct method photovoltaic generation forecast still has high potential for further improvement. A further optimization of the parameters for calculating the clear sky curve and its scaling factor is possible under consideration of historical generation and weather data. Furthermore, neural networks can greatly improve the accuracy of free photovoltaic forecasts, making them more competitive with commercial ones.

One of the main parameter in accuracy of forecasting in this method is cloud cover. A sample of the error of this method in different weather condition is shown in the following figure.

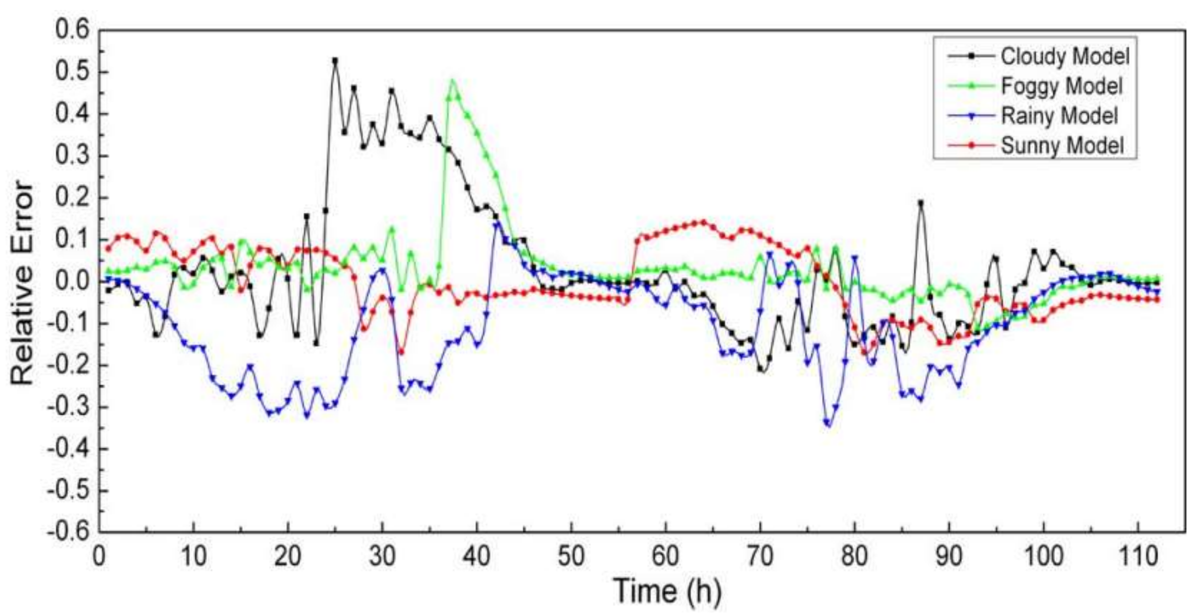

Figure 37. Testing four forecasting models and comparing their relative errors.

After the training process, the forecasting output data can be obtained based on the test data. According to the weather report for the next day, the model type can be selected based on the test data. The table below shows the relative error between measurement data and forecasting data at the same time moment (same length of data sample). Errors vary between 50% and 39% in all four models tested. In general, the Sunny Model is the most reliable, with a mean value of 4.88%, followed by Foggy, Rainy, and Cloudy Models with mean values of 8.16%, 9.12%, and 12.42%, respectively.

Model Classification	RMSE (MW)	MRE (%)
Cloudy model (Model A)	1.824	12.42
Foggy model (Model B)	2.52	8.16
Rainy model (Model C)	2.48	9.12
Sunny model (Model D)	1.57	4.85
Average Value	2.10	8.64

Table1. Showing the results of weather forecasting and measurements over the whole year are shown below.

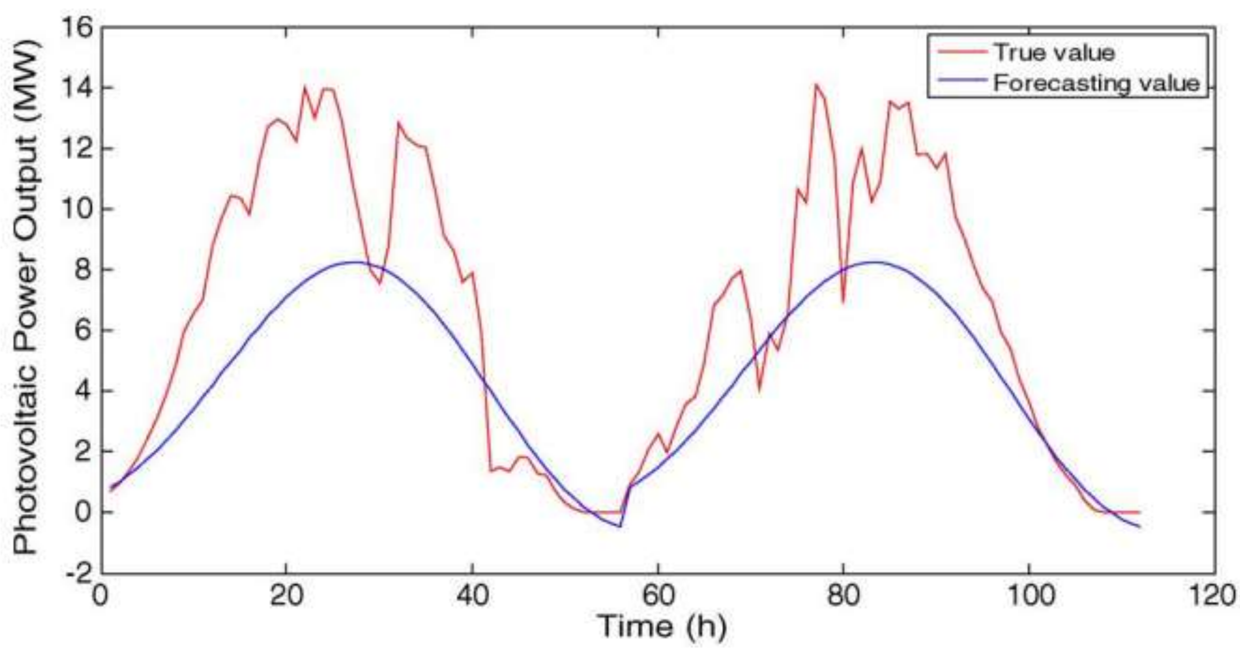

Figure 38. Results of photovoltaic power forecasting in cloudy day.

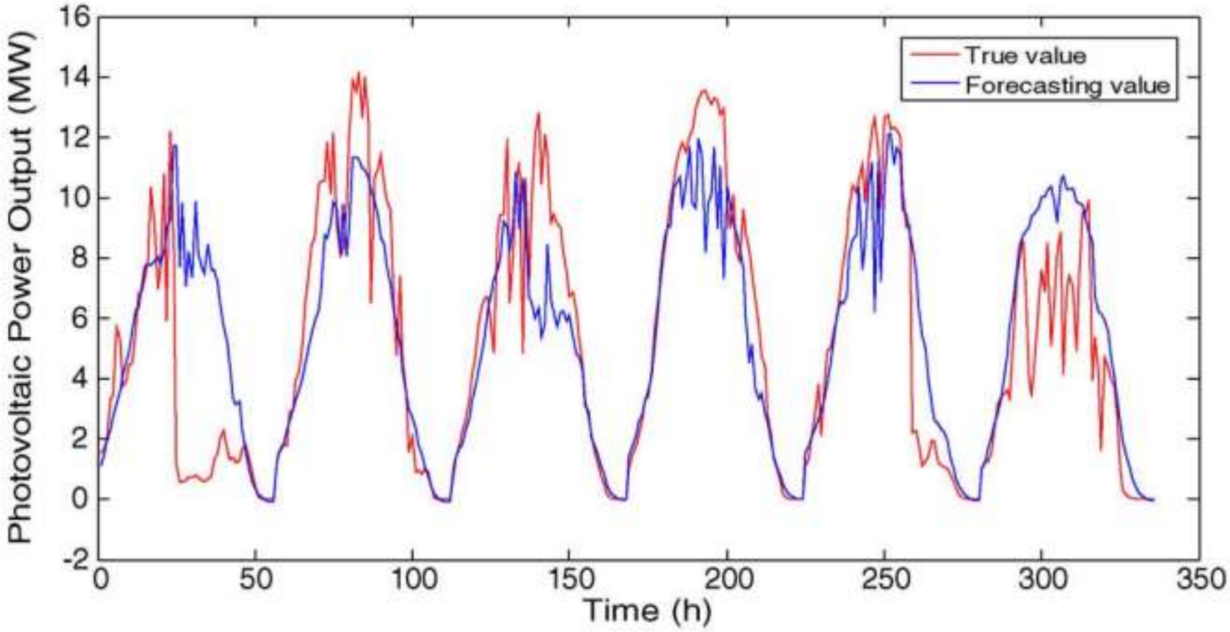

Figure 39. Results of photovoltaic power forecasting in foggy day.

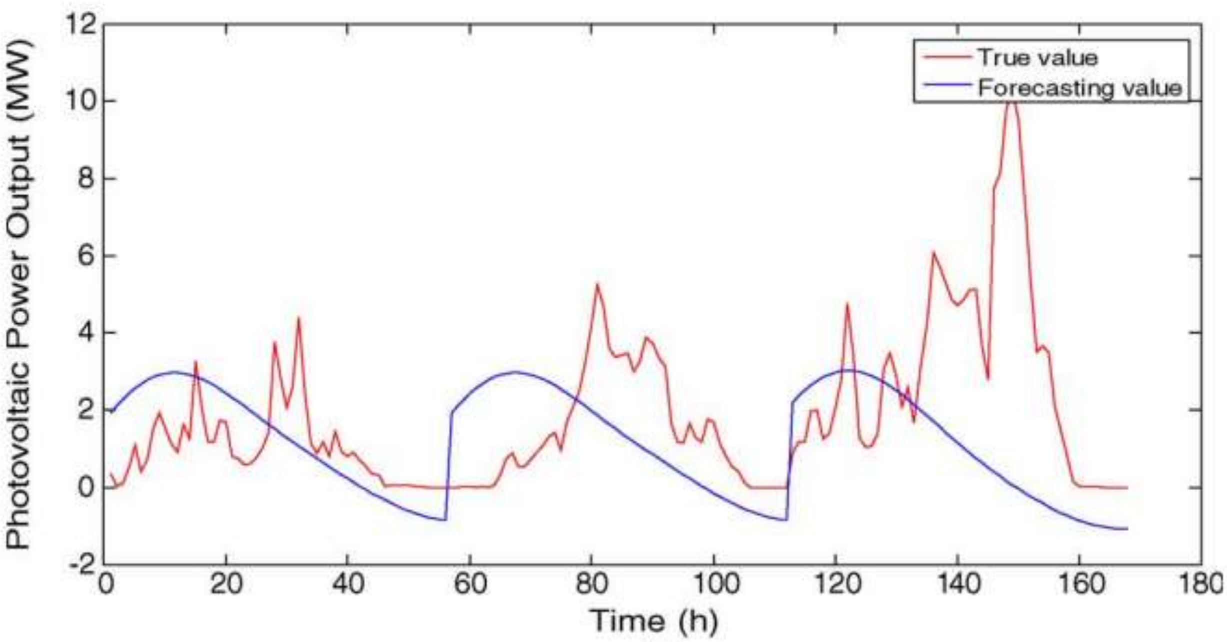

Figure 40. Results of photovoltaic power forecasting in rainy day.

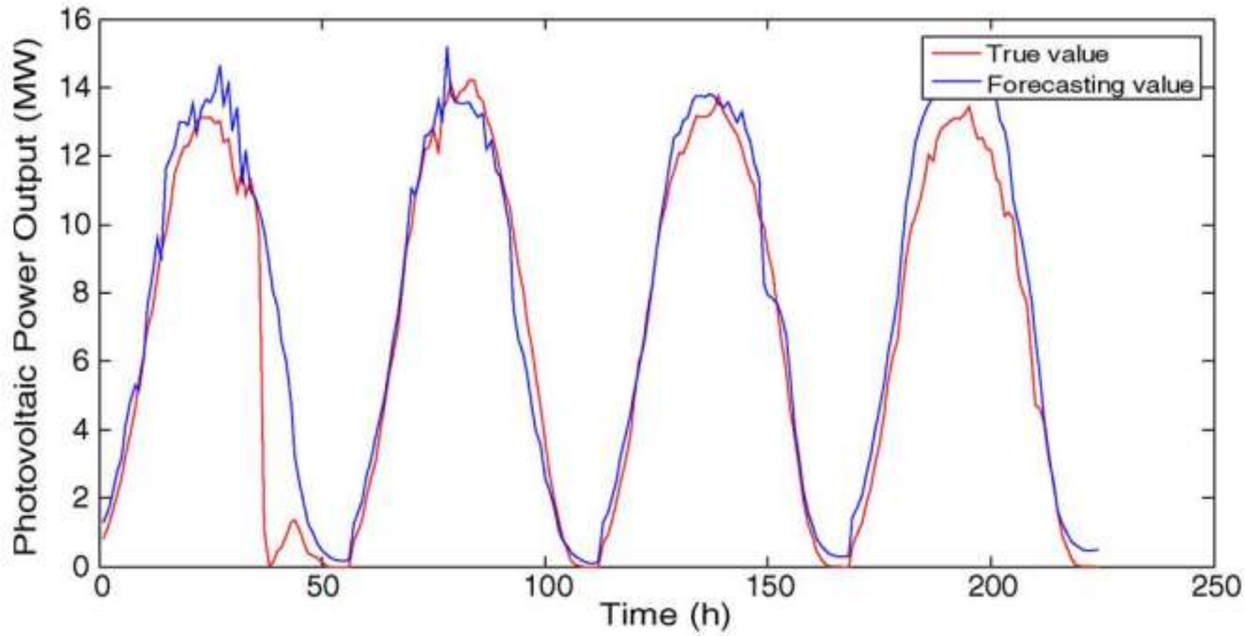

Figure 41. Results of photovoltaic power forecasting in sunny day.

Probabilistic PV power forecasting model has some benefits; 1) have the ability to predict the PV plant in a range of +5 minutes up to +36 hours; 2) automatically adapt to changes in the PV plant (environment, partial outage, soiling, etc.) without multiplying the models; and 3) start at any time of the day and any forecast horizon without multiplying the models.

This model was able to produce high performing forecasts for both intra-day and day-ahead forecasts by incorporating in situ measurements as well as satellite imagery along with NWP.

Researchers in [76] forecasted a PV system with two different methods; NWP and Statistic method. Considering the availability of NWP, we have selected March, June, October and December as spring, summer, autumn and winter respectively, to illustrate the impact of seasons on our physical models.

As can be seen from the figure, the forecast and measured results are very similar on a clear day. Consequently, in the case of frequent fluctuations in the output power, the predicted curves are only the profiles of the real values. There is no doubt that forecasting errors are stochastic [76].

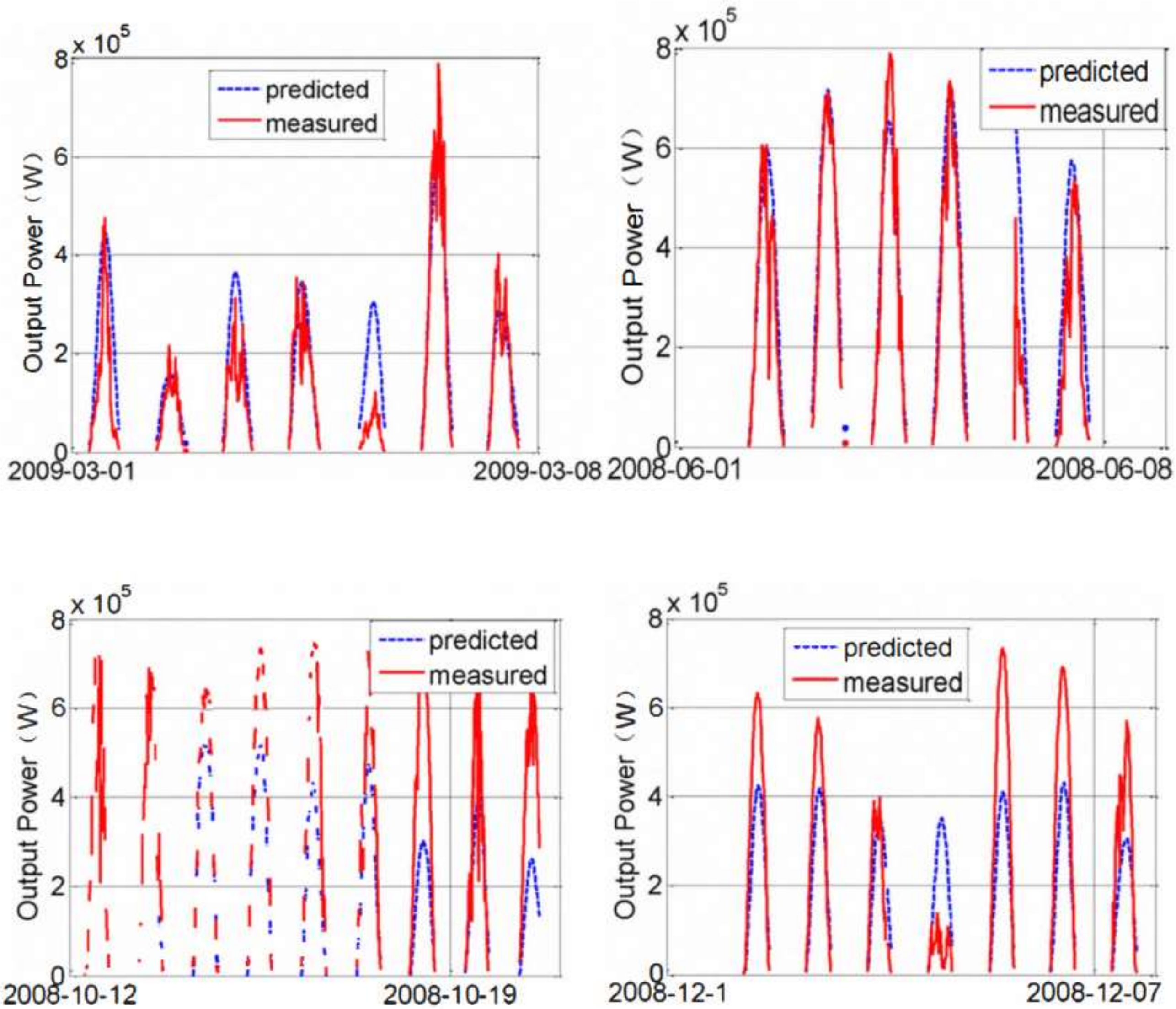

Figure 42. Physical model with NWP input

As can be seen from the table, the nRMSEs for selected months, as well as the nRMSEs for the entire year. The nRMSEs range from 11-17%. Due to the lack of good data in October's NWP, its nRMSE is the largest. The result for December is also disappointing. Measured peaks are significantly lower than forecasted peaks [76].

Table 1 Impacts of season on model accuracy (nRMSE%)

	Physical model (NWP input)	Physical model (measured R & T input)	NN model
Mar	11.48	8.78	11.05
Jun	11.54	5.17	10.45
Oct	16.58	4.93	15.49
Dec	15.33	5.61	12.62
Year	12.45	5.51	10.5

Table 2. Modelling accuracy and nRMSE% based on season

In the following figure, the forecasted results are shown on the same days as the historical results in order to compare the performance of the Neural Network (NN) model. Predicted results are remarkably similar to the real ones.

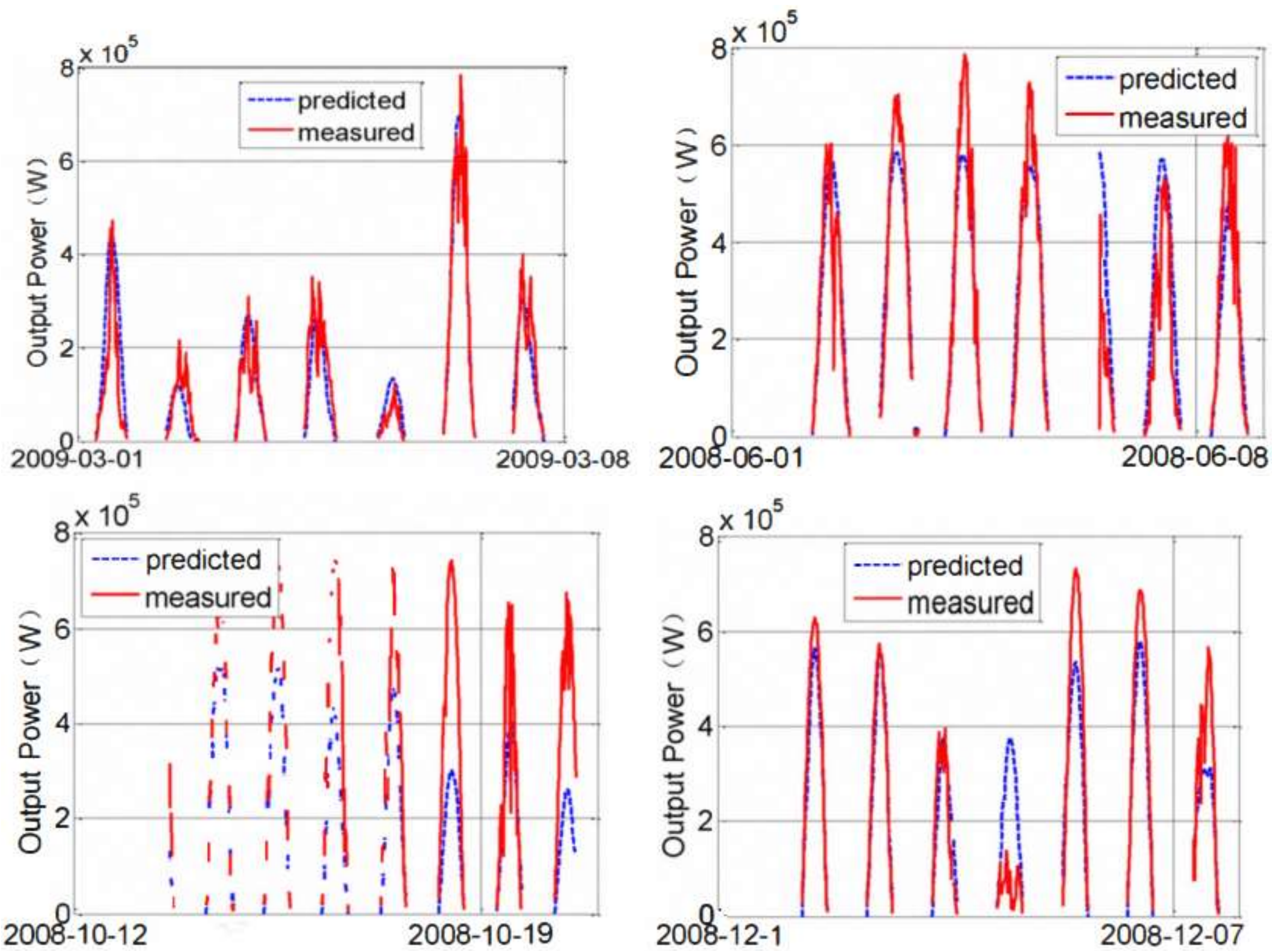

Figure 43. Statistical NN model with NWP input

Stochastic test data shows an nRMSE of about 10.5% for the whole year. When compared to the physical model, the nRMSEs tend to be lower, especially in December. As a result of the accommodation of neural networks, the NN model has improved. The network auto adaptation makes NWP a little more accurate. On the other hand, the improvement is very small when dealing with NWP uncertainty.

The bellow table shows the impacts of the NN inputs on the accuracy of the model. The temperature is represented by Tc, the cloud coefficient by C, the irradiance by R, and the position of the sun by Ps. Based on the table, the best model has Tc, C, R, H, and Ps inputs, with an nRMSE of 10.5%. In general, the nRMSE of trained models is between 10-13%. The Table demonstrates that Ps has a clear impact on the relative position of the sun since the angle is not fixed. A very good model is obtained when Ps and NWP are combined. Humidity also significantly impacts the model. Since the humidity will affect the transparency of the sky, adding it to the input will some extent reduce the uncertainty of irradiance. The result is improved performance.

Tc	C	R	H	Ps	Num of neurons	nRMSE(%)
√	√	√	√	√	11	10.5
√	√	√	√		12	11.529
	√	√	√		15	12.195
	√		√		10	12.404
√		√		√	11	12.36
√		√	√	√	13	11.467
√	√	√			13	11.894

Table 3. Inputs and their impact on the results

All of all, by comparison these two methods, four main effects are considered. 1) Power output and solar irradiance correlate very closely in PV plants. When irradiance and temperature are included as inputs to the physical model, the PV station behavior can be sufficiently described. 2) For best models, the nRMSE of the whole year is about 12.45% for physical models, and 10.5% for NN statistical models. Both models are affected by the season. Several factors affect the solar irradiance, so the variability of weather can be attributed to it. A unified model of every season is relatively difficult to construct. In contrast, NWP uncertainty varies according to the season. Thus, a variety of performances are obtained. These results can all be used in practice, though they are all forecasted with reasonable accuracy. 3) Physical models and NN models can both be used to forecast same PV station output power. Their performances are also very similar. Uncertainty in forecasted power comes mainly from uncertainty in NWP. The improvement of NWP should be resolved first in order to get better results. The NWP should be modified using real-time irradiance measurements in super-short-term forecasting. 4) Forecasting results will also be affected by humidity and sun position. NN models can be improved if they are included in the input.

In [74], researchers compere ANN and Hybrid method together and Both methods use the same dataset for network training and testing. When it comes to forecasting sunny days, both methods perform well. The hybrid approach shows excellent performance on some specific days (NMAE% by 1% in two cases; WMAE% by 4% in three cases), while the second method under study shows a consistently good performance (NMAE always between 1% and 2%; WMAE percent always between 7% and 13%). Cloudy days result in significantly reduced forecasting performance for both methods. Furthermore, while the hybrid method is better for some specific days, for at least one day the prediction is rather poor (NMAE% > 11% and WMAE% = 750). In contrast, the method that feeds solely on data from the dataset performs better on both NMAE and WMAE metrics.

The less stable performance of the hybrid model can be explained by the fact that the forecasting accuracy of the first is dependent on the quality of the weather forecast for the day. Moreover, the quality of the forecast depends greatly on the location it refers to, which usually is not the exact same location where the PV system operates. Verifying this hypothesis will require a more detailed analysis considering this possible correlation.

In [7], presents a hybrid approach that combines two artificial neural networks-based methods and it block diagram is shown in following figure.

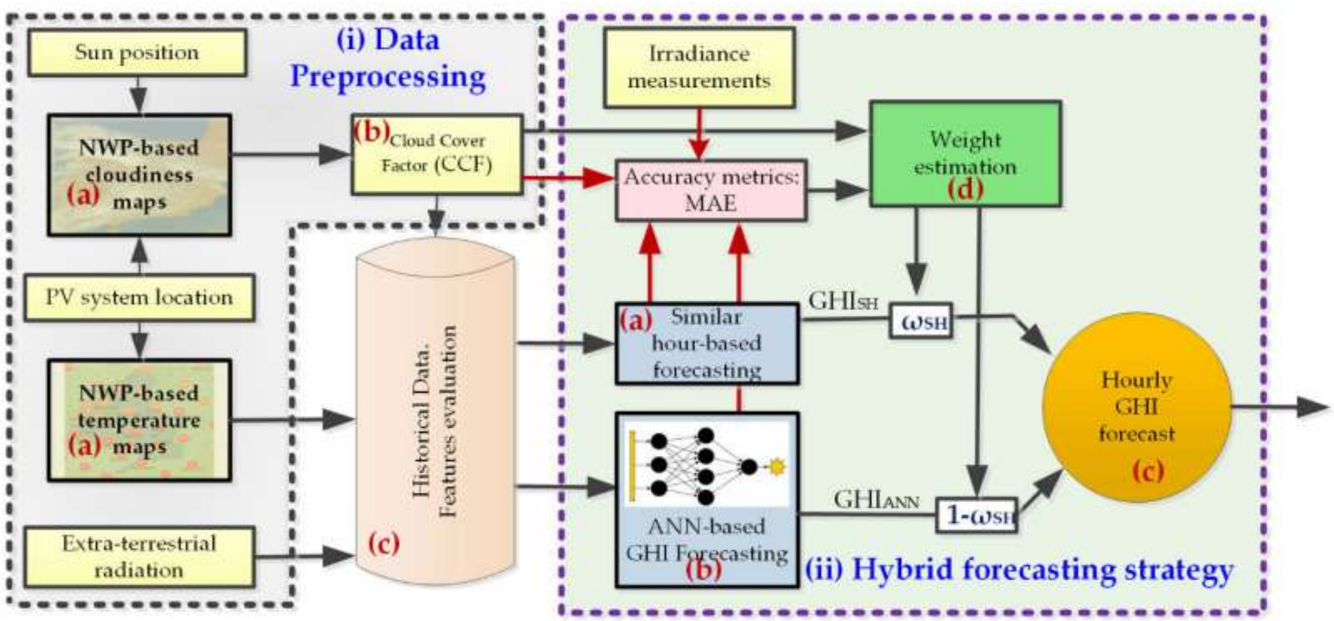

Figure 44. Schematic diagram of the forecasting method, comprised of two parts: (i) a data pre-processing stage and (ii) the hybrid forecasting method. In this forecasting strategy, artificial neural networks and an hour-based algorithm are used. Depending on the day type, the final forecast is calculated by dynamically weighting the outputs.

Based on the type of day (sunny, cloudy, overcast) and the MAE, the output of both forecasting methods is dynamically weighted. The proposed forecasting method is based on forecasts of temperature and cloudiness generated by the AEMET via NWP alongside irradiance measurements obtained from a real PV site. Based on historical data covering the last six months, a MAPE of 21.64% and a NRMSE of 31.69% are obtained for hour-based technique, and a MAPE of 21.37% and a NRMSE of 30.99% are obtained for hybrid strategy.

In comparison with ANN-based forecasts, the proposed similar hour-based method produces more accurate results with a reduced historical dataset. Mean errors are reduced by 16.32% and 9.07% for one-month and two-month datasets, respectively. Regardless of the length of the historical dataset, the proposed model yields promising results.

In [65], proposes three kinds of photovoltaic power forecasting models to solve the problem (convolutional neural networks, long short-term memory networks, and hybrid models). These three methods are including CNN network, LSTM network and the hybrid model (CLSTM). Numerous tests and verifications were performed on them, and the statistics under three statistical indicators were provided. These are the conclusion of this article:

The prediction accuracy of the three models increases with an increase in the availability of historical time series data, but as the number of data points reaches a certain point, the precision gains slow down. The time series data for the three models are from 2.5Y to 3Y, and the RMSE is generally increased by 17.00–58% while the MAE is generally increased by 36-54%. A big increase in RMSE is generally seen from 3Y to 3.5Y, the large increase in MAE is generally seen from 60 to 68% and a big increase in MAPE is generally seen from 61 to 69%). Therefore, in the present study, 3-years is recommended for predicting values.

Based on the results of the statistical analysis, all three models perform well and are accurate. The convolutional model is the most accurate. This neural network model forecasts better for seven time periods, but the short-term memory neural network model forecasts better for a certain period of time. This would indicate that spatial features are more useful than temporal features for predicting photovoltaic power.

The statistics also show higher accuracy for the hybrid model compared to the other model. The hybrid model mainly benefits from the respective advantages of convolutional neural networks.

In terms of training time, the long-short-term memory neural network requires the least amount of training, while the hybrid model requires the most training. Hybrid models train approximately 1.5-3.5 times longer than long short-term memory neural networks (depending on time series). Hybrid models have a longer prediction time than long short-term memory neural networks, while convolutional neural networks and long short-term memory neural networks do not differ much in their predictions.

As a result of this study, the hybrid model is more accurate in predicting time series data length than the single model, and we recommend using 3Y for time series data length analysis. In certain circumstances, however, convolutional neural networks and long short-term memory neural networks are also good choices due to their processing time requirements and the length of historical time series data [65].

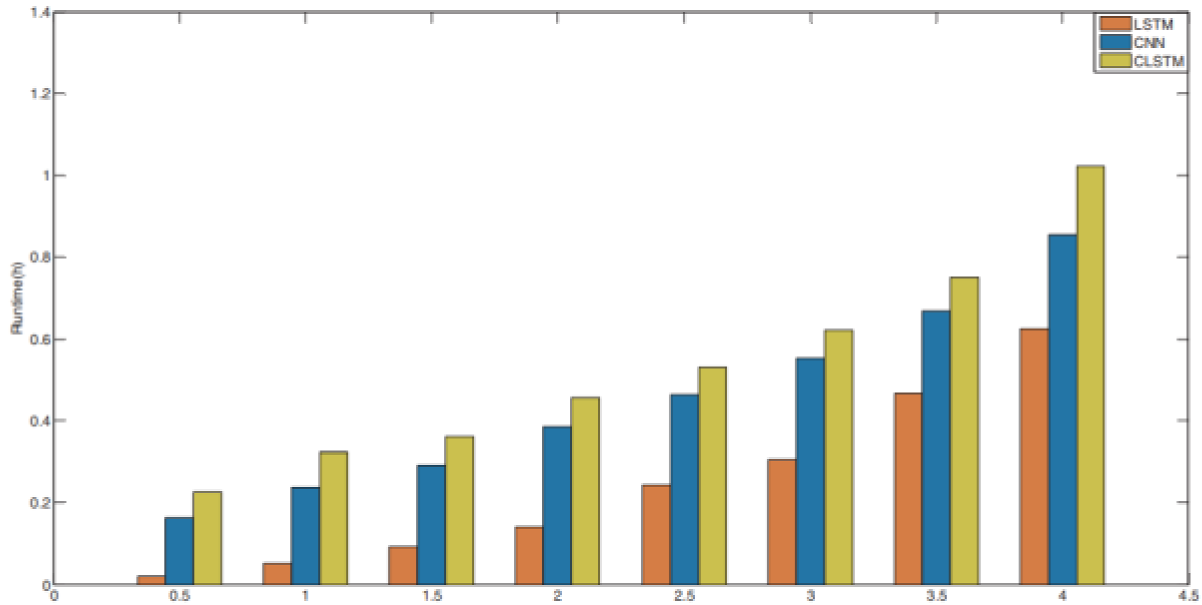

Figure 45. Each model requires a different runtime [81].

The performance of the hybrid technique was also compared with traditional linear regression and artificial neural network models (regression neural network (GRNN), feedforward neural network (FFNN), Multiple Linear Regression (MLR), and single-stage models resulting in three hybrid models, namely SR-FFNN, SR-GRNN, SR-MLR) [32].

In the following table, the six models are compared quantitatively in terms of RMSE, MBE, MAE, and R. According to the performance indicators, all six models can accurately predict future events.

Accuracy metrics	Model					
	Single-stage			Hybrid		
	MLR	FFNN	GRNN	SR-MLR	SR-FFNN	SR-GRNN
RMSE	3.22	2.65	2.99	3.15	2.74	2.94
MBE	−0.74	−0.17	0.33	−0.33	0.01	0.14
MAE	2.43	2.11	2.30	2.33	2.09	2.31
R	0.915	0.937	0.923	0.914	0.932	0.926

Table 4. Comparison of accuracy metrics for the six models

There is no difference in error indices between the hybrid and single-stage models in general. Thus, the performance of the SR-FFNN, SR-GRNN, and SR-MLR models can be compared to their corresponding single-stage models. It was determined that SR-FFNN had the best MAE and MBE values of the six investigated models, implying that this model had the least average dispersion and bias in its prediction error.

In the chart below are the measured and predicted PV energy output values for each month. When the predictions are accumulated over a long period, hybrid models perform better than single stage models. From the accuracy metrics, it can be concluded that the SR-FFNN was the best prediction model, and the hybrid models were generally better than the single-stage counterparts.

Month	Actual energy generated (kWh)	Energy generation predicted by models (kWh)					
		FFNN	GRNN	MLR	SR-FFNN	SR-GRNN	SR-MLR
June	195.5	195.7	181.3	217.2	191.6	189.7	214.7
July	619.8	586.8	569.2	665.1	608.2	584.0	650.0
August	508.5	516.1	484.2	511.7	511.0	483.9	495.1
September	296.0	322.2	305.6	311.8	309.4	292.7	295.4
October	243.1	248.5	226.8	233.5	229.5	214.8	220.7
November	152.0	159.7	160.4	182.0	149.5	179.8	175.3
December	105.2	123.9	128.8	140.3	118.7	148.4	132.8
Total (kWh)	2120.1	2152.9	2056.3	2261.7	2118.0	2093.3	2183.9
Percentage prediction error (%)		1.6	3.0	6.7	0.1	1.3	3.0

Table 5. A comparison between monthly PV output measured and predicted for six models.

While forecasters using ML have been extensively investigated in general, the use of DL for PV power prediction has been rather limited so far.

The majority of research has focused on forecasting at single locations, while little has been done on regional models; there is no accurate general regional model.

The most investigated time horizon is in the short-term regime (up to a few days)- which is also the most requested and used. ML-based forecasters are well suited for this case, particularly when combined with appropriate algorithms- such as ANN-optimized GA or PSO.

Very-short-term forecasting and long-term forecasting have been scarcely investigated.

Most AI-based models perform well for sunny days, while for cloudy days the forecasting accuracy decreases significantly.

The accuracy of AI-based on historical power output, and the use of meteorological parameters (such as air temperature, solar irradiance, relative humidity, wind speed, cloud cover), combined with an optimal learning algorithm and weather classification can improve forecasting accuracy.

One-step ahead forecasting perform best, and has been extensively investigated. Conversely, multistep-ahead predictions remain a challenging task.

Generally, in order to increase the accuracy of the forecasters based on AI techniques, the following points should be considered: 1) large datasets with good-quality data are preferable, 2) pretreatment and analysis of the database to identify outliers and missing data is required, 3) exogenous inputs should be taken into account, such as for example cloud variation; 4) combination with other physical models.

Long-term PV power forecasting could be achieved by using NWP models and DL including LSTM and CNN, as the latter (as well as RNN) have some useful characteristics, such as the ability of estimating the temporal dependencies of the investigated data, and the ability of performing a more general feature extraction. With the increasing amount of data (on-site measurements based on SCADA, satellite image data, NWP models), DL will become increasingly interesting for PV power prediction, in particular for designing accurate regional forecasters. An extensive investigation is still needed on DL-based forecasting methods. DL has been gaining popularity in particular for time series forecasting, due to the availability of increasingly large datasets and open sources codes (such as Python, TensorFlow, Keras, etc.). ML based models, on the other hand, have shown their capability for PV power forecasting as well as DL based models, while the accuracy depends mainly on the data quality, their amount and the learning algorithms. However, DL based models are more suitable in the case of large database, unlike ML based models.

To improvement of model accuracy for cloudy days is still only marginally investigated. Forecasting approaches able to estimate and classify cloud cover and to use these parameters for DL models in expected to lead to sizeable accuracy improvement.

The combination of physical models with DL-based methods has been scarcely investigated. However, such as approach could significantly improve forecasting accuracy.

In [52], A hybrid method is compared with Persistence, BPNN, and RBFNN benchmark methods. The solar power data used for these forecasting methods is gathered from the actual one. Simulated results indicate that the proposed method provides very low prediction errors compared with benchmark methods.

The study shows that no one model outperforms the other for all possible conditions. Utility-scale PV plant operators, power management system manufacturers, and distribution and transmission network operators can use this comparison to identify the most suitable forecaster for their specific situation.

References

[1] Unsupervised Disaggregation of Photovoltaic Production from Aggregated Power Flow Measurements of Heterogeneous Prosumers. Fabrizio Sossan, Member, IEEE, Lorenzo Nespoli, Member, IEEE, Vasco Medici, Member, IEEE, and Mario Paolone, Senior Member, IEEE

[2] Disaggregation of Local Photovoltaic Generation from Composite Power Flows with Direct Measuring and Satellite Estimations of the Irradiance: a Comparison,Fabrizio Sossan, Enrica Scolari, Mario Paolone Distributed Electrical Systems Laboratory (DESL) EPFL, Ecole Polytechnique F ´ ed´ erale de Lausanne, Switzerland ´ {fabrizio.sossan, enrica.scolari, mario.paolone}@epfl.ch

[3] A. Ballanti and L. F. Ochoa, "On the integrated pv hosting capacity of mv and lv distribution networks," in Innovative Smart Grid Technologies Latin America (ISGT LATAM), 2015 IEEE PES. IEEE, 2015, pp. 366–370.

[4] F. Sossan, J. Darulova, M. Paolone, A. Kahl, S. J. Bartlett, and M. Lehning, "Large scale deployment of pv units in existing distribution networks: Optimization of the installation layout," in 2016 Power Systems Computation Conference (PSCC), June 2016.

[5] Data-Driven Probabilistic Net Load Forecasting With High Penetration of Behind-the-Meter PV

Yi Wang , Student Member, IEEE, Ning Zhang , Member, IEEE, Qixin Chen , Senior Member, IEEE, Daniel S. Kirschen , Fellow, IEEE, Pan Li, Student Member, IEEE, and Qing Xia , Senior Member, IEEE

[6] Short-term power output forecasting of the photovoltaics in Bulgaria.

[7] A Day-Ahead Irradiance Forecasting Strategy for the Integration of Photovoltaic Systems in Virtual Power Plants GUILLERMO MORENO, PEDRO MARTIN , CARLOS SANTOS , FRANCISCO J. RODRÍGUEZ , (Member, IEEE), AND ENRIQUE SANTISO

[8] A Solar Time Based Analog Ensemble Method for Regional Solar Power Forecasting Xinmin Zhang , Student Member, IEEE, Yuan Li, Member, IEEE, Siyuan Lu, Hendrik F. Hamann, Member, IEEE, Bri-Mathias Hodge , Senior Member, IEEE, and Brad Lehman , Senior Member, IEEE

[9] Review of photovoltaic power forecasting, J. Antonanzas, N. Osorio, R. Escobar, R. Urraca, F. J. Martinez-de-Pison, and F. Antonanzas-Torres, Sol. Energy, vol. 136, pp. 78–111, Oct. 2016.

[10] A review of deep learning for renewable energy forecasting, Energy Convers. Manage. H. Wang, Z. Lei, X. Zhang, B. Zhou, and J. Peng, vol. 198, Oct. 2019, Art. no. 111799.

[11] Photovoltaic electricity generator dynamic modeling methods for smart grid applications: A review Renewable Sustain. Energy Rev , S. Koohi-Kamal, N. Rahim, H. Mokhlis, and V. Tyagi,., vol. 57, pp. 131–172, 2016.

[12] Photovoltaic power forecasting using statistical methods: Impact of weather data, De Giorgi, P. Congedo, and M. Malvoni, IET Sci. Meas. Technol., vol. 8, no. 3, pp. 90–97, 2014.

[13] Design of Optimal Power Point Tracking Controller Using Forecasted Photovoltaic Power and Demand AliAkbar Shafi, Hussein Sharadga , and Shima Hajimirza, IEEE TRANSACTIONS ON SUSTAINABLE ENERGY, VOL. 11, NO. 3, JULY 2020

[14] Solar Radiation, Ioan Sarbu, Calin Sebarchievici, in Solar Heating and Cooling Systems, 2017

[15] The Effect of Relative Humidity and Temperature on Polycrystalline Solar Panels Installed Close to a River, Armstrong Oreh Njok, Julie C. Ogbulezie, February 2019

[16] A Model for Joint Probabilistic Forecast of Solar Photovoltaic Power and Outdoor Temperature Raksha Ramakrishna , Student Member, IEEE, Anna Scaglione , Fellow, IEEE, Vijay Vittal , Fellow, IEEE, Emiliano Dall'Anese , Member, IEEE, and Andrey Bernstein , Member, IEEE, 2019

[17] Hourly solar irradiance time series forecasting using cloud cover index, Yang Dazhi, Panida Jirutitijaroen, Wilfred M. Walsh ,2012

[18] An integrated review of factors influencing the performance of photovoltaic panels M.M. Fouada,∗ , Lamia A. Shihatab,c , ElSayed I. Morgan, 2017

[19] Experimental and computational investigation of the thermal and electrical performance of a new building integrated photovoltaic concept, University of Thessaly, Ζώγου Ο. 2011.

[20] Energy efficient lighting systems in buildings with integrated photovoltaics, Aaltouniversity, Viitanen, J, 2015.

[21] Multiple-Input Deep Convolutional Neural Network Model for Short-Term Photovoltaic Power Forecasting CHIOU-JYE HUANG, PING-HUAN KUO, 2019

[22] Generation and Evaluation of photovoltaic Forecasts based on freely available Weather Data, Alexander Zeh Christoph Hainzinger Rolf Witzmann, 2021

[23] Forecasting Power Output of Photovoltaic Systems Based on Weather Classification and Support Vector Machines Jie Shi, Wei-Jen Lee, Fellow, IEEE, Yongqian Liu, Yongping Yang, and Peng Wang, 2012

[24] A Novel Approach for Seamless Probabilistic Photovoltaic Power Forecasting Covering Multiple Time Frames Thomas Carriere , Christophe Vernay, Sebastien Pitaval, and George Kariniotakis , Senior Member, IEEE, 2020

[25] Short-term photovoltaic power forecasting based on long short term memory neural network and attention mechanism. Zhou, Y. Zhang, L. Yang, Q. Liu, K. Yan, and Y. Du, IEEE Access, vol. 7, pp. 78063–78074, 2019.

[26] Day-Ahead Nonparametric Probabilistic Forecasting of Photovoltaic Power Generation Based on the LSTM-QRA Ensemble Model FEI MEI , (Member, IEEE), JIAQI GU1 , JIXIANG LU , JINJUN LU , JIATANG ZHANG , YUHAN JIANG , TIAN SHI , AND JIANYONG ZHENG ,2020

[27] Computing electricity spot price prediction intervals using quantile regression and forecast averaging, J. Nowotarski and R. Weron, Comput. Statist., vol. 30, no. 3, pp. 791–803, Sep. 2015.

[28] Probabilistic load forecasting via quantile regression averaging on sister forecasts, B. Liu, J. Nowotarski, T. Hong, and R. Weron IEEE Trans. Smart Grid, vol. 8, no. 2, pp. 730–737, Mar. 2017.

[29] Probabilistic Models for Spatio-Temporal Photovoltaic Power Forecasting Xwegnon Ghislain Agoua ´ , Robin Girard, and George Kariniotakis , Senior Member, IEEE, 2019

[30] Very short-term nonparametric probabilistic forecasting of renewable energy generation #x2014; with application to solar energy, F. Golestaneh, P. Pinson, and H. B. Gooi, IEEE Trans. Power Syst., vol. 31, no. 5, pp. 3850–3863, Sep. 2016.

[31] Ul-Saufie A, Yahya A, Ramli N, Hamid H. Future PM10 concentration prediction using quantile regression models. In: International conference on environmental and agriculture engineering, vol. 37. Singapore: IACSIT Press; 2010.

[32] A hybrid method for forecasting the energy output of photovoltaic systems Pamela Ramsami, Vishwamitra Oree, 2015

[33] concentration prediction using quantile regression models, Ul-Saufie A, Yahya A, Ramli N, Hamid H. Future PM10. In: International conference on environmental and agriculture engineering, vol. 37. Singapore: IACSIT Press; 2010.

[34] Patterson D. Artificial neural networks. Singapore: Prentice Hall; 1996.

[35] Photovoltaic and solar power forecasting for smart grid energy management. Wan, C.; Zhao, J.; Song, Y.; Xu, Z.; Lin, J.; Hu, Z. IEEE CSEE J. Power Energy Syst. 2015

[36] A solar time based analog ensemble method for regional solar power forecasting.Zhang, X.; Li, Y.; Lu, S.; Hamann, H.F.; Hodge, B.M.; Lehman, B. IEEE Trans. Sustain. Energy 2018,

[37] Solar power prediction based on satellite images and support vector machine. Jang, H.S.; Bae, K.Y.; Park, H.S.; Sung, D.K. IEEE Trans. Sustain. Energy 2016

[38] Photovoltaic and Solar Forecasting: State of the Art Pelland, S.; Remund, J.; Kleissl, J.; Oozeki, T.; De Brabandere, K.; IEA PVPS Task 14, Subtask 3.1. Report Iea-PVPS T14–01: 2013; International Energy Agency: Paris, France, 2013.

[39] Solar forecasting: Methods, challenges, and performance. Tuohy, A.; Zack, J.; Haupt, S.E.; Sharp, J.; Ahlstrom, M.; Dise, S.; Black, J. IEEE Power Energy Mag. 2015

[40] One-day-ahead daily power forecasting of photovoltaic systems based on partial functional linear regression models. Wang, G.; Su, Y.; Shu, L. Renew. Energy 2016

[41] Operational photovoltaics power forecasting using seasonal time series ensemble. Yang, D.; Dong, Z. Sol. Energy 2018

[42] Probabilistic Models for Spatio-Temporal Photovoltaic Power Forecasting. Agoua, X.G.; Girard, R.; Kariniotakis, G. IEEE Trans. Sustain. Energy 2018

[43] Probabilistic competitive ensemble method for short-term photovoltaic power forecasting.Bracale, A.; Carpinelli, G.; De Falco, P. A IEEE Trans. Sustain. Energy 2016

[44] Probabilistic forecast of PV power generation based on higher order Markov chain.Sanjari, M.J.; Gooi, H.B. IEEE Trans. Power Syst. 2016

[45] Artificial Intelligence: A Modern Approach, 3rd ed.; Prentice-Hall, Russell, S.J.; Norvig, P. Inc.: Upper Saddle River, NJ, USA, 2009.

[46] Physical hybrid artificial neural network for short term forecasting of PV plant power output Dolara, A.; Grimaccia, F.; Leva, S.; Mussetta, M.; Ogliari, E. A. Energies 2015

[47] Photovoltaic forecasting: A comparison of the most effective techniques Nespoli, A.; Ogliari, E.; Leva, S.; Massi Pavan, A.; Mellit, A.; Lughi, V.; Dolara, A. Day-ahead. Energies 2019

[48] Deep Learning; Goodfellow, I.; Bengio, Y.; Courville, A. MIT Press: New York, NY, USA, 2016.

[49] Advanced Methods for Photovoltaic Output Power Forecasting: A Review Adel Melli, Alessandro Massi Pavan ,*, Emanuele Ogliari, Sonia Leva and Vanni Lughi, 2019

[50] 24-h forecast of solar irradiance using artificial neural network: Application for performance prediction of a grid-connected PV plant at Trieste, Italy Mellit, A.; Massi Pavan, A. Sol. Energy 2010

[51] Chen, C.; Duan, S.; Cai, T.; Liu, B. Online 24-h solar power forecasting based on weather type classification using artificial neural network. Sol. Energy 2011 Shi, J.; Lee, W.J.; Liu, Y.; Yang, Y.; Wang, P. Forecasting power output of photovoltaic systems based on weather classification and support vector machines. IEEE Trans. Ind. Appl. 2012,

[52] Photovoltaic Power Forecasting With a Hybrid Deep Learning Approach GANGQIANG LI, SEN XIE, BOZHONG WANG, JIANTAO XIN , YUNFENG LI , AND SHENGNAN DU, 2020

[53] Fernandez-Jimenez, L.A.; Muñoz-Jimenez, A.; Falces, A.; Mendoza-Villena, M.; Garcia-Garrido, E.; Lara-Santillan, P.M.; Zorzano-Santamaria, P.J. Short-term power forecasting system for photovoltaic plants. Renew. Energy 2012

[54] . Da Silva Fonseca, J.G., Jr.; Oozeki, T.; Takashima, T.; Koshimizu, G.; Uchida, Y.; Ogimoto, K. Use of support vector regression and numerically predicted cloudiness to forecast power output of a photovoltaic power plant in Kitakyushu, Japan. Prog. Photovolt. Res. Appl. 2012

[55] Zeng, J.; Qiao, W. Short-term solar power prediction using a support vector machine. Renew. Energy 2013

[56] Das, U.; Tey, K.; Seyedmahmoudian, M.; Idna Idris, M.; Mekhilef, S.; Horan, B.; Stojcevski, A. SVR-based model to forecast PV power generation under different weather conditions. Energies 2017

[57] Leva, S.; Dolara, A.; Grimaccia, F.; Mussetta, M.; Ogliari, E. Analysis and validation of 24 hours ahead neural network forecasting of photovoltaic output power. Math. Comput. Simul. 2017

[58] Pierro, M.; De Felice, M.; Maggioni, E.; Moser, D.; Perotto, A.; Spada, F.; Cornaro, C. Data-driven upscaling methods for regional photovoltaic power estimation and forecast using satellite and numerical weather prediction data. Sol. Energy 2017

[59] Liu, W.; Liu, C.; Lin, Y.; Ma, L.; Xiong, F.; Li, J. Ultra-short-term forecast of photovoltaic output power under fog and haze weather. Energies 2018

[60] Al-Dahidi, S.; Ayadi, O.; Adeeb, J.; Alrbai, M.; Qawasmeh, B. Extreme Learning Machines for Solar Photovoltaic Power Predictions. Energies 2018

[61] Yao, X.; Wang, Z.; Zhang, H. A novel photovoltaic power forecasting model based on echo state network. Neurocomputing 2019

[62] Han, Y.; Wang, N.; Ma, M.; Zhou, H.; Dai, S.; Zhu, H. A PV power interval forecasting based on seasonal model and nonparametric estimation algorithm. Sol. Energy 2019

[63] Abdel-Nasser, M.; Mahmoud, K. Accurate photovoltaic power forecasting models using deep LSTM-RNN. Neural Comput. Appl. 2017

[64] Son, J.; Park, Y.; Lee, J.; Kim, H. Sensorless PV power forecasting in grid-connected buildings through deep learning. Sensors 2018

[65] Wang, K.; Qi, X.; Liu, H. A comparison of day-ahead photovoltaic power forecasting models based on deep learning neural network. Appl. Energy 2019

[66] Lee, D.; Kim, K. Recurrent neural network-based hourly prediction of photovoltaic power output using meteorological information. Energies 2019

[67] Bouzerdoum, M.; Mellit, A.; Massi Pavan, A. A hybrid model (SARIMA–SVM) for short-term power forecasting of a small-scale grid-connected photovoltaic plant. Sol. Energy 2013

[68] Ogliari, E.; Grimaccia, F.; Leva, S.; Mussetta, M. Hybrid predictive models for accurate forecasting in PV systems. Energies 2013

[69] Chu, Y.; Urquhart, B.; Gohari, S.M.; Pedro, H.T.; Kleissl, J.; Coimbra, C.F. Short-term reforecasting of power output from a 48 MWe solar PV plant. Sol. Energy 2019

[70] Ogliari, E.; Dolara, A.; Manzolini, G.; Leva, S. Physical and hybrid methods comparison for the day ahead PV output power forecast. Renew. Energy 2017

[71] Nespoli, A.; Ogliari, E.; Dolara, A.; Grimaccia, F.; Leva, S.; Mussetta, M. Validation of ANN training approaches for day-ahead photovoltaic forecasts. In Proceedings of the International Joint Conference on Neural Networks (IJCNN), Rio de Janeiro, Brazil, June 2018

[72] Neural Networks (IJCNN), Rio de Janeiro, Brazil, 8–13 June 2018. 81. Zang, H.; Cheng, L.; Ding, T.; Cheung, K.W.; Liang, Z.; Wei, Z.; Sun, G. Hybrid method for short-term photovoltaic power forecasting based on deep convolutional neural network. IET Gener. Transm. Distrib. 2018

[73] Leva, S.; Mussetta, M.; Ogliari, E. PV module fault diagnosis based on micro-converters and day-ahead forecast. IEEE Trans. Ind. Electron. 2018

[74] Day-Ahead Photovoltaic Forecasting: A Comparison of the Most Effective Techniques Alfredo Nespoli , Emanuele Ogliari,* , Sonia Leva , Alessandro Massi Pavan , Adel Mellit, Vanni Lughi and Alberto Dolara, April 2019

[75] Photovoltaic yield prediction using an irradiance forecast model based on multiple neural networks Saad Parvaiz DURRANI, Stefan BALLUFF, Lukas WURZER, Stefan KRAUTER, 2018

[76] Comparative Study of Power Forecasting Methods for PV Stations Yuehui HUANG, Jing LU, Chun LIU, Xiaoyan XU, Weisheng WANG, Xiaoxin ZHOU, 2010

[77] Hybrid Solar Forecasting Method Using Satellite Visible Images and Modified Convolutional Neural Networks, Zhiyuan Si , Student Member, IEEE, Yixiao Yu , Student Member, IEEE, Ming Yang , Senior Member, IEEE, and Peng Li , Student Member, IEEE, 2021

[78] Deterministic and Probabilistic Solar Power Forecasts: A Review on Forecasting Models Cheng-Liang Huang, Yuan-Kang Wu and Yuan-Yao Li, 2021

[79] Accurate photovoltaic power forecasting models using deep LSTM-RNN, Mohamed Abdel-Nasser, Karar Mahmoud, 2017

[80] Comparison of different simplistic prediction models for forecasting PV power output: Assessment with experimental measurements Meng Wang, Jinqing Peng, Yimo Luo, Zhicheng Shen, Hongxing Yang, 2021

[81] Methods for Energy Production Estimation from Photovoltaic Plants: Review and Application, Ludmil Stoyanov, Zahari Zarkov, Vladimir Dimitrov Lazarov, 2018

[82] Comparison of different physical models for PV power output prediction Alberto Dolara, Sonia Leva, Giampaolo Manzolini, 2015

[83] Humidity impact on photovoltaic cells performance: A review, Roshen Tariq Ahmed Hamdi, Sanaa Abdulhadi, Hussein A Kazem, Miqdam Tariq Chaichan, 2018

[84] Forecasting of photovoltaic power generation and model optimization: A review Utpal Kumar Dasa , Kok Soon Teya,∗ , Mehdi Seyedmahmoudian, Saad Mekhilef, Moh Yamani Idna Idris, Willem Van Deventer, Bend Horan, Alex Stojcevski, 2018

[85] Forecast Quality of Physics-Based and Data-Driven PV Performance Models for a Small-Scale PV System. Moritz Stüber, Felix Scherhag, Matthieu Deru, Alassane Ndiaye, Muhammad Moiz Sakha, Boris Brandherm, Jörg Baus and Georg Frey, 2021

[86] Hybrid PV Power Forecasting Methods: A Comparison of Different Approaches Alessandro Niccolai, Alberto Dolara and Emanuele Ogliari, 2021

[87] A Hybrid Technique for Day-Ahead PV Generation Forecasting Using Clear-Sky Models or Ensemble of Artificial Neural Networks According to a Decision Tree Approach Stefano Massucco, Gabriele Mosaico, Matteo Saviozzi and Federico Silvestro, 2021

[88] Impact of Climate Change on Combined Solar and Run-of-River Power in Northern Italy. Baptiste François, Benoit Hingray, Marco Borga, Davide Zoccatelli, Casey Brown and Jean-Dominique Creutin, 2018

[89] Methods for Energy Production Estimation from Photovoltaic Plants: Review and Application, Ludmil Stoyanov, Zahari Zarkov, Iva Draganovska, Vladimir Lazarov, 2019

[90] Photovoltaic Power Forecasting Methods in Smart Power Grid, Harendra Kumar Yadav ,Yash Pal, Tripathi, 2015

[91] Methods for Energy Production Estimation from Photovoltaic Plants: Review and Application, Ludmil Stoyanov, Zahari Zarkov, Iva Draganovska, Vladimir Lazarov, 2018

[92] An Effective Hybrid NARX-LSTM Model for Point and Interval PV Power Forecasting, MOHAMED MASSAOUDI 1,2, (Student Member, IEEE), INES CHIHI3 , LILIA SIDHOM3 , MOHAMED TRABELSI4 , (Senior Member, IEEE), SHADY S. REFAAT1 , (Senior Member, IEEE), HAITHAM ABU-RUB1 , (Fellow, IEEE), AND FAKHREDDINE S. OUESLATI2, 2021

[93] Power Generation Forecast of Hybrid PV–Wind System Mohammad Javad Sanjari , Hoay Beng Gooi , Senior Member, IEEE, and Nirmal-Kumar C. Nair, Senior Member, IEEE, 2020

[94] Exploring Key Weather Factors From Analytical Modeling Toward Improved Solar Power Forecasting Jianxiao Wang , Student Member, IEEE, Haiwang Zhong, Member, IEEE, Xiaowen Lai, Qing Xia, Senior Member, IEEE, Yang Wang, and Chongqing Kang, Fellow, IEEE, 2019

[95] Plant Considering the Incentive for Forecast Accuracy in South Korea WOONG KO, (Member, IEEE), AND MUN-KYEOM KIM, (Member, IEEE), 2021