INDIAN INSTITUTE OF TECHNOLOGY INDORE

UNDERGRADUATE THESIS

Estimation of Electricity Production from Photovoltaic Panel

VAMSI BULUSU
160002059

Department of Electrical Engineering

December, 2019

INDIAN INSTITUTE OF TECHNOLOGY INDORE

UNDERGRADUATE THESIS

Estimation of Electricity Production from Photovoltaic Panel

Author:
VAMSI BULUSU
ID No. 160002059

Supervisor:
Dr. KAPIL AHUJA

Thesis submitted in fulfillment of the requirements for the degree of Bachelor of Technology

in the

Department of Electrical Engineering

December, 2019

Declaration of Authorship

I, VAMSI BULUSU declare that this thesis titled, "Estimation of Electricity Production from Photovoltaic Panel" and the work presented in it is my own. I confirm that:

- This work was done wholly or mainly while in candidature for the BTP project at IIT Indore.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed:

Date:

Certificate

This is to certify that the thesis entitled, "*Estimation of Electricity Production from Photovoltaic Panel*" and submitted by Vamsi Bulusu ID No 160002059 in partial fulfillment of the requirements of EE 493 B.Tech Project embodies the work done by him under my supervision.

Supervisor

Dr.KAPIL AHUJA
Associate Professor,
Indian Institute of Technology Indore
Date:

"Don't let the sound of your own wheels drive you crazy
Lighten up while you still can
Don't even try to understand
Just find a place to make your stand and
Take it easy "

Eagles

INDIAN INSTITUTE OF TECHNOLOGY INDORE

Abstract

Department of Electrical Engineering

Bachelor of Technology

Estimation of Electricity Production from Photovoltaic Panel

The electricity grid is under a tremendous change where renewable energies are more and more integrated.The centralised management of the grid is shifting to a more distributed infrastructure where smart grids with different scales take more importance. Using renewable energies, however, requires either the use of a battery pack or to consume the electricity at the time of production. Due to the limitation of storing the energy, shifting the consumption to when the electricity is actually produced is necessary. This project deals with the estimation of solar panel production in order to forecast when and how much electricity will be available.

We propose an Articial Neural Network (ANN)-based model to predict the hourly production of photovoltaic (PV) plants. The neural network takes three different kinds of input features including the static weather forecast data, the predicted irradiance values and the processed neighbourhood minima, maxima and mean outputs for each hour. This enables the neural network to generalize better for different types of days and climates. The models are then evaluated using different standard error metrics such as the Normalized Mean Absolute Error (NMAE) and the Weighted Mean Absolute Error (WMAE) over an hourly, daily and weekly time frame. We also try to propose a qualitative method to nd the best possible timeframe for the training dataset based on the ageing of the solar panels and the overall annual change in climate and other conditions. These methods are applied to a large dataset of solar panel electricity production over a period of seven years.

Acknowledgements

I would like to thank my B.Tech Project supervisor **Dr. Kapil Ahuja** for his guidance and constant support in structuring the project and his valuable feedback throughout the course of this project. His overseeing the project meant there was a lot that I learnt while working on it. I thank him for his time and efforts. I would also like to thank him for his help in his capacity as the Dean of International Affairs to materialize this project.

I am grateful to **Dr. Yann Busnel** and **Dr. Nicolas Montavont** without whom this project would have been impossible. They provided valuable guidance with every aspect of the project and also helped me write my first scientific paper.

I am really grateful to the Institute for the opportunity to be exposed to systemic research.. Lastly, I offer my sincere thanks to everyone who helped me complete this project, whose name I might I have forgotten to mention.

Contents

List of Abbreviations

ANN	**A**rtificial **N**eural **N**etwork
PV	**P**hoto **v**oltaic
NMAE	**N**ormalized **M**ean **A**bsolute **E**rror
WMAE	**W**eighted **M**ean **A**bsolute **E**rror
CSRM	**C**lear **S**ky **R**adiation **A**lgorithm
ReLu	**R**ectified **L**inear **U**nit
MAPE	**M**ean **A**bsolute **P**ercentage **E**rror
MGmean	**M**ean **G**eometric **M**ean
MSE	**M**ean **S**quare **E**rror

Dedicated to all my friends who made this possible

Chapter 1

Introduction

1.1 Background

Different efforts worldwide are made in order to transform the traditional electricity system into smart grid [**b17**]. The notion of smart grid may refer to a large set of features, but in the scope of this thesis, we can define the smart grid as a system of distributed systems, including microgeneration units using renewable energies, and demand-side integration. Different use cases drive smart grid deployment [**b17**]: to provide electricity to isolated lands (e.g. small islands that are disconnected from the grid) [**b20**], to provide electricity in developing countries with autonomous or inter-connected systems [**b19**], or to reduce the dependency on carbon and/or nuclear energy in developed countries[**b21**, **b22**]. A lot of research is currently going on to provide efficient, reliable, resilient and flexible solutions.

Renewable energy is rapidly growing and it is important to integrate them into the electric grid and not just use them as alternate sources. Despite the obvious advantages in the use of renewable energy for the environment and sustainability, the major drawback is their instability. Most renewable energy sources are dependant on the weather or other environmental phenomena, such as tides or storms. This adds an intrinsic uncertainty in their production. These fluctuations in production make it harder to integrate them into a traditional grid and go against the aforementioned goals of reliability and resilience.

The smart grid comes with a communication framework in which components may exchange monitoring and control information [**b24**, **b25**, **b26**]. Through a unified and standardized set of protocols, all objects talk the same language, and any of them can talk with any other one. The smart meter is certainly the most popular communicating object in the smart grid [**b23**]. It allows to remotely monitor the energy consumption in real time and may also be used to broadcast price information. The innovation in this new communication framework is the bi-directional nature of the communication, which allows devices to negotiate together. Interestingly, it allows consumer devices to exchange information with production units, thus making the system flexible: the consumer devices may adapt their behavior depending on the energy availability, to some extent. For example, an electric car may advertise that it needs a given amount of energy, in a given time frame. Depending on the energy availability, the grid may choose the best time and source of energy for charging the car. To summarize, it is the combination of the three features, namely the bi-directional communication framework, the demand-response-compatible loads and the usage of renewable energy, that allow to better manage the energy service. In order to fully take advantage of the smart grid, real time monitoring and control is not enough, but predictions are needed [**b13**, **b14**, **b20**]. Accurate and reliable forecasting of the productions can further feed the optimization problem to align the consumption with the production.

Let us now focus on solar energy, as this is one of the largest and fastest-growing sources of renewable energy. Solar power plants are easy to install and require minimal maintenance making them ideal for both small scale and large scale deployment. Just like all sources of renewable energy, solar panel production also suffers from fluctuations and instability due to climatic conditions. This is one of the main reasons why small scale PV plants are used for selective purposes such as water heating and lighting or as a source of backup power but are not completely integrated into the grid. While the use of battery packs to store the produced electricity is a viable option, it is costly and cannot be considered for extended periods of time. Accurate forecasting of the production can help plan the electricity usage efficiently. Forecasts can also help monitor and maintain solar power plants. While hourly and daily forecasts can help examine the proper functioning of the plant, daily and weekly forecasts can help with scheduling maintenance and long term planning.

Solar power plants range from a few connected solar panels for individual households or small industries to larger PV parks where panels are spread over an area of a 2 - 5 square kilometers. This thesis focuses on forecasting of production for solar microgrids. The microgrids are harder to model using weather data since the weather forecasts are provided for areas accurate up to a couple of square kilometers while the microgrids vary from a few square meters to a couple of hundred square meters in area. For larger PV plants, the inaccuracies in the weather forecasts are compensated by the large area of the power plant.

In this thesis, we propose a two-step neural network based model to forecast the hourly production of solar power plants. The first step is the prediction of hourly irradiance values based on the weather forecast. The second step is the hourly production forecast using a separate model to forecast every hour of the day. The production forecast model uses a combination of the weather forecast, irradiance forecast and neighbourhood data to predict the production of a particular hour.

The remainder of the thesis is as follows. Next section deals with the related work and discuss the most recent works on solar panel production forecast and the main differences with our approach. Section III describes the data set used in this thesis, and gives an overview of our model. Section IV indicates our choices regarding the selected artificial neural network and section V finally presents our model. Section VI present the results we obtained, and section VII concludes the thesis by giving some perspectives to this work.

Chapter 2

Literature Survey

The following chapter discusses different publications and works that have similar objectives to the work described in this thesis. There has been a lot of works in the area of machine learning for solar power production forecasting due to the increase in the urgency to shift to renewable power sources. With the climate change and the price drop of solar panel, smart grids need to better integrate this source of energy into the traditional grid. However, in order to provide a smooth integration, we need to cope with the intermittent production and forecasting the production is key to address this challenge.

2.1 Time series based model

Cococcioni et al.[**b16**] propose an interesting model in which the time series nature of the data is exploited. In this model, a single neural network is trained to predict the output generation of the PV plant at 15 minute intervals. The neural network considered for this approach is a feedforward network with tapped delay lines. The neural network takes as input the irradiance and output values of the previous day corresponding to the same hour. Depending on the number and type of delay lines, the network may consider more observations from the previous day or observations from more days to learn the trend in the irradiance and output values. It then predicts the output production solely based on the time series characteristics of the data. While this approach is suitable for large PV installations in regions with fairly stable weather, smaller installations and locations with erratic weather conditions are not well modelled by this approach.

2.2 Clear Sky Irradiance based model

Dolara et al. [**b5**] show an interesting approach to the problem. The paper proposes a hybrid method to predict hourly power productions in a day-ahead manner. The proposed model consists of a neural network which takes the weather forecast and some temporal features as input. The model also takes the clear sky irradiance values that are calculated based on the deterministic Clear Sky Solar Radiation Algorithm (CSRM) [**b18**]. The neural network is trained to learn the amount of irradiance actually used by the solar plant based on the various weather forecast variables.

2.3 Average Input based model

Nespoli et al. [**b10**] propose a neural network model that takes the average temperature and average measured irradiance of the previous day to predict the irradiance of the current day. It produces 24 outputs corresponding to the irradiance forecast for 24 hours. The mean of these values is then used to decide between two separate models for “sunny” and “cloudy” days based on the mean of the predicted irradiance. The second models then use the mean of irradiance, temperature and production of the previous day to predict 24 values corresponding to the hourly production of that day.

2.4 Neuro Fuzzy model

Grimaccia et al. [**b3**] consider a neuro-fuzzy model which feeds into the neural network the weather forecast at 3 different points of the day all at once. The forecasts for 6 A.M., 12.P.M. and 6 P.M. are used together to predict the complete day's production. The weather forecast supplied to the network also includes a fuzzy logic pre-processed irradiance value for the 3 times of the day.

Our approach draws from the two main publications that use weather forecast data to forecast the PV production. The first one is an irradiance prediction based neural network model [**b10**] and the second is a neuro-fuzzy predictive model based on weather forecast [**b3**].

In this project we propose a neural network based model to predict the hourly productions. The model consists of two parts, where the first part uses hourly weather forecast data to predict the irradiance and the second part uses the forecast irradiance values along with the weather and some other processed data to predict the production. The hourly inputs were considered since a lot of the hourly variation in data is lost if the average values are provided as proposed in the first paper explained. The neural networks in the second part are specific for every hour. This means that there are separate models trained to predict the particular hour's production. This helps the neural network to learn the structural and temporal features separately. Using the same model to predict the value of different hours leads to too much variance in the data for the model to accurately learn important features.

Chapter 3

Dataset and Overview

3.0.1 PV module and Dataset

As part of this thesis, we explore the various variables provided as part of the weather forecast and study their correlation to the DC Power output of the PV plant in an attempt to select the variables that provide the best production forecast. The dataset used for the research is part of the project cited here [**b1**]. It consists of data from two separate sets of solar panels. The following are the specifications of the solar power microgrid:

- Solar panel technology - silicon mono-crystalline
- Panel Inclination - 15
- Panel rating - 960 W
- Total surface area - 5.52m^2
- Microgrid Location - Latitude of 4714′20″.76N and Longitude of 133′30″ 24W

The first dataset contains approximately 1400 days of observations spread over 6 years from 2011 to 2017. The second dataset contains 1000 days of data over 4 years from 2014 to 2017. Both datasets contain the measured irradiance values and the output DC Power. The weather forecast data used contains various variables with an hourly frequency. While the weather forecast provides various features such as the UV Index, Temperature, Wind Chill, Wind speed, precipitation, humidity, pressure, visibility and cloud cover, only some of these variables are useful in the forecasting of the hourly production.

Fig.1 is a scatter plot between the Irradiance and the DC Power variables. It can be seen that there is some kind of linear relationship between the two variables. These kinds of plots help in capturing the polynomial or periodic dependence between two variables. The Fig. 2 is a scatter plot between the humidity and DC Power. It can be seen that there is no clear relationship between the two variables. Such scatter plots where the points are scattered in a cloud show that the two variables have minimal dependence. In fact, the irradiance feature is the most correlated to the output. However, typical weather forecasts do not provide irradiance values. This problem is tackled by predicting the irradiance based on the other weather variables. This procedure is explained in detail in section V.C.

3.0.2 Overview

Fig. 3 gives a high level overview of the model. It shows the two step process involved in the prediction of the hourly production values. It also shows the input data used for the various models.

In this project, we propose a two step process to forecast hourly production values. In contrast to the first paper described in section II which uses the previous days recorded values for the forecast, we use the weather forecast for the day that is intended to be forecast. This provides

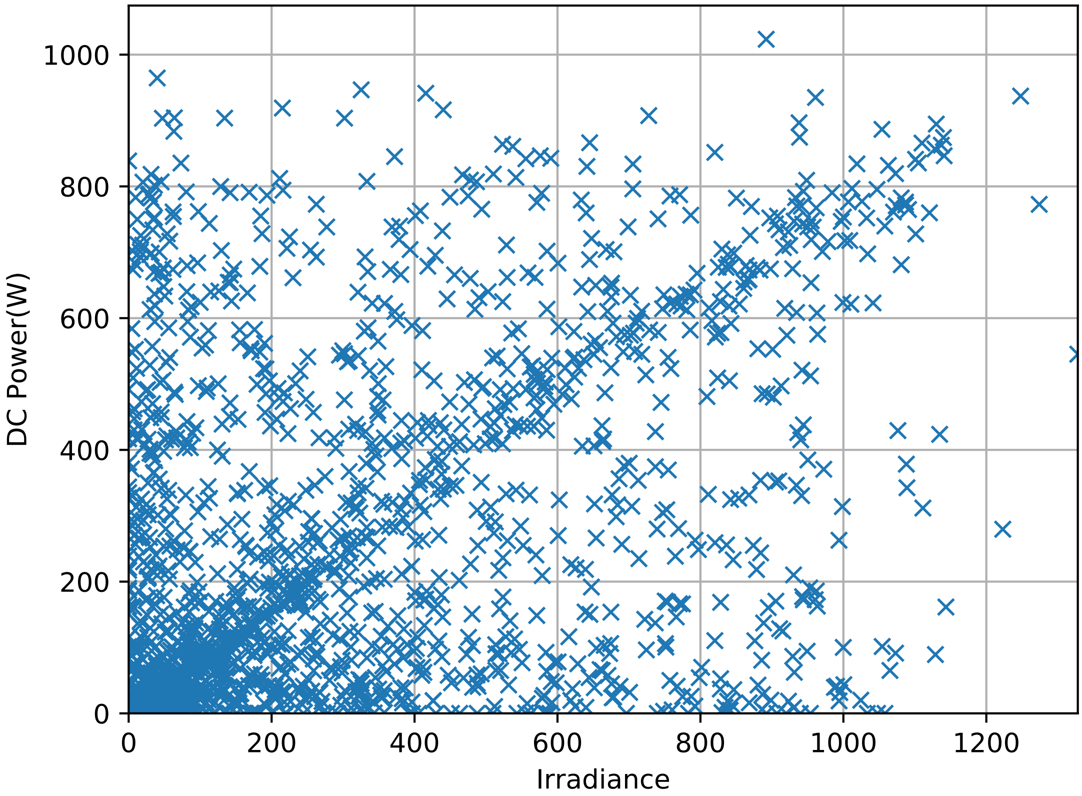

FIGURE 3.1: Scatter plot between DC Power (output) and the Irradiance(input)

a lot more correlation between the inputs and the output. The first step in the proposed method is the prediction of hourly irradiance values from the hourly weather forecast. The second step involves the feeding of the predicted irradiance values along with the weather forecast values and some processed neighbourhood statistics to the hourly models. We propose the use of 15 hourly models to predict the forecast for every hour. This is especially useful in the case of predictions for microgrids. The hourly models also help in maintaining the inputs unconnected. Providing more than one forecast at a time is recommended for predicting average values but will "confuse" the network when trying to predict the hourly production. One more reason for the hourly models is that a single model would require 15 output neurons to make hourly predictions but in general, a single neural network with multiple outputs performs better when the outputs are correlated while multiple neural networks with a single output are preferred for uncorrelated outputs.

3.1 Artificial Neural Networks

Artificial Intelligence methods have become popular in predicting outputs that are complex non-linear functions of the inputs. Artificial Neural Networks (ANNs) are a popular subclass of machine learning models, which are loosely based on the learning model of a human brain.

Artificial Neural Networks are fairly straightforward systems, which process the data with the help of artificial neurons and learn intrinsic features which help map the input to the output. The word intrinsic is used here to stress the fact that the neural network learns without any

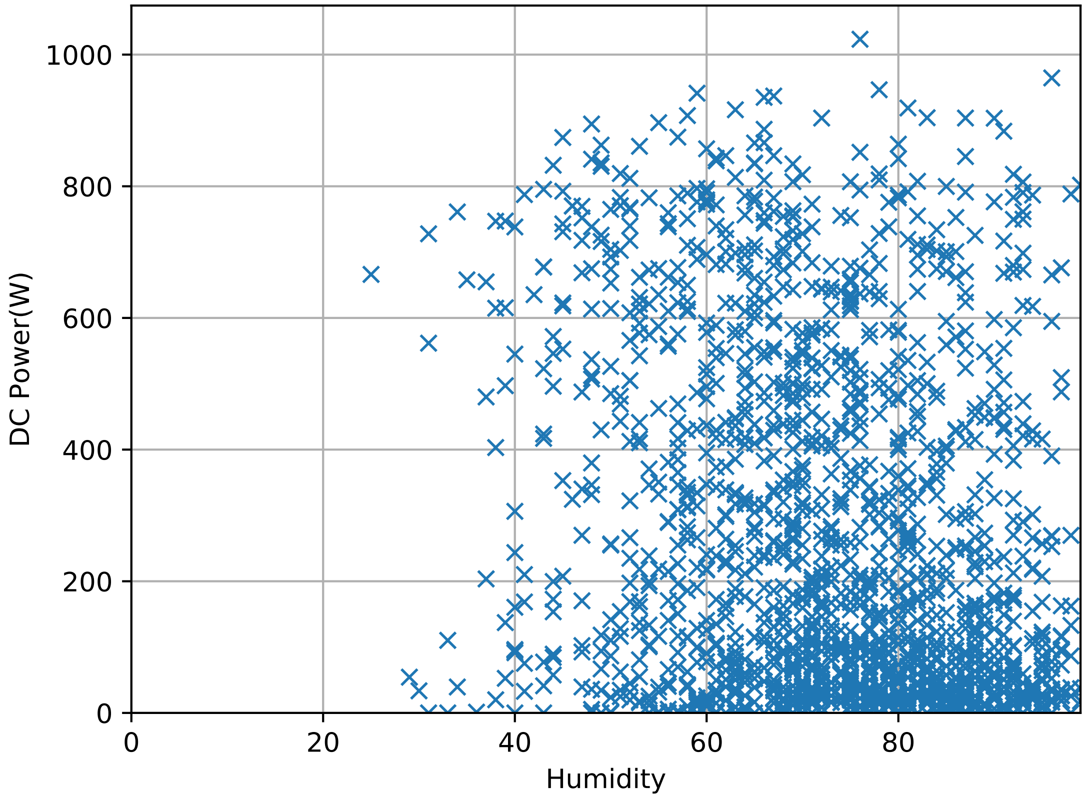

FIGURE 3.2: Scatter plot between DC Power (output) and the Humidity(input)

a priori knowledge about the features. This means that it is not possible to clearly define the features that the neural network uses to learn the problem.

The ANN method proposed in this project uses the Multi-Layer Perceptron (MLP) model. An MLP generally consists of an input layer followed by several hidden layers and an output layer. Each layer consists of a predefined number of artificial neurons depending on the type of data being modelled. Every layer in an MLP starting from the input layer is generally densely connected to the next layer meaning that every neuron in a layer is connected to every other neuron in the next layer. The connections between neurons of consecutive layers carry weights and the neurons then apply an activation function over the weighted inputs and pass it to the next layer. The usual rule of thumb is that a wider neural network (more neurons per layer) tends to memorize better and a deeper (more hidden layers) neural network is able to learn highly nonlinear features better.

Training algorithms are then used to change the weights of the connections to learn the best mapping between the inputs and the outputs. The best mapping between the input and the output is the solution to the optimization problem of minimizing or maximizing a cost function depending on the type of data. The cost function can be considered to provide some prior knowledge to the model about the kind of mapping. For regression problems, the most popular cost functions are the mean absolute error and the mean squared error. The most famous training algorithm is the error backpropagation based on gradient descent. This algorithm is computationally expensive but converges to a minima fairly easily with the proper learning rate and batch size. One drawback of the gradient descent based backpropagation is the convergence

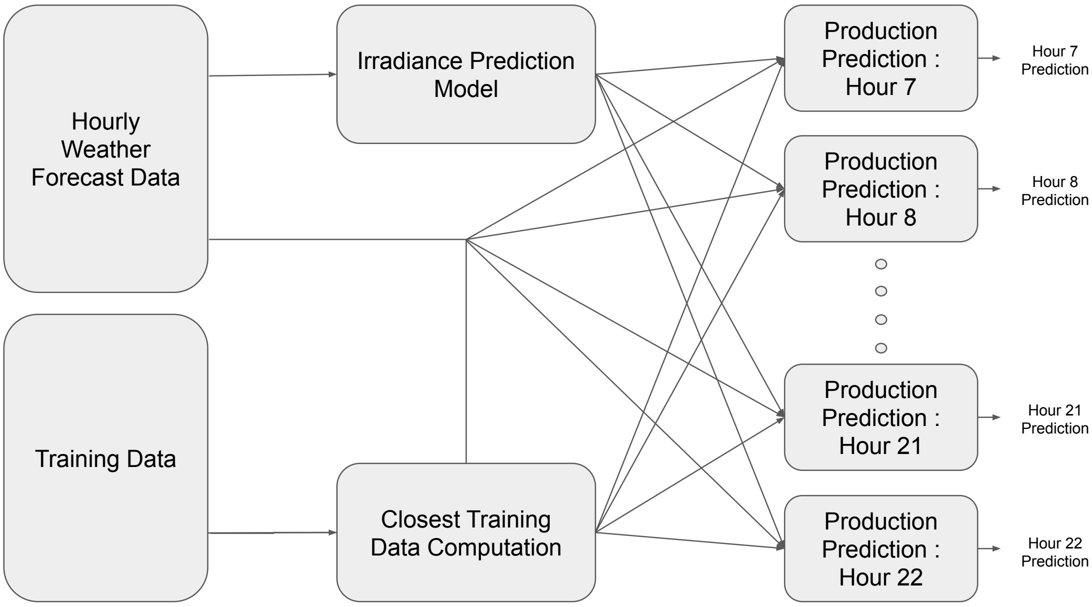

FIGURE 3.3: The proposed two step prediction model

to a local minima if the weights are not initialized properly. The backpropagation algorithm is, however, suitable for the problem described in this thesis.

Chapter 4

Data and Normalization

4.0.1 Data Processing

The data used for the forecasting of the PV power plant productions can be divided into three.

- The static weather forecast data
- The irradiance prediction
- The DC output from observations with the closest weather conditions

The static weather forecast data is the most important of the three and provides the basis for the predictions. It is mined for the particular location of the solar panels using the latitude and longitude. The weather forecast used for this research was mined using the world weather online API for python. The weather forecast consists of various variables out of which the most suitable subset has been chosen. The weather variables used for the prediction are UV Index, cloud cover, humidity, temperature and visibility.

The irradiance prediction value is obtained from a neural network model that is trained to predict the irradiance values based on the weather forecast variables and past irradiance recordings. This is a way of guiding the neural network to use the features that are known to impact the production of the panels. This intermediate step of predicting known features that affect the production helps the neural network learn more meaningful patterns. Otherwise, the network tends to learn abstract features which may not be suitable for the problem or may not be present in the testing data.

The nearest weather neighbours are obtained by processing every weather input individually to find the 50 nearest data points from the training set that have the closest weather to the forecast. The distance between two weather observations is calculated using the Euclidean distance between them which is calculated as follows

$$d(\mathbf{p}, \mathbf{q}) = d(\mathbf{q}, \mathbf{p}) = \sqrt{(q_1 - p_1)^2 + \cdots + (q_n - p_n)^2} \tag{4.1}$$

The power production of these 50 points is then retrieved and sorted. The sorted list of production values is then reduced by removing the highest and lowest 10 points. This is done to effectively remove outliers that may influence the output towards the outlier points. After the reduction of the 50 values into a list of 30 values, the minimum, maximum and the mean values of the list are taken as inputs to the final prediction model.

4.0.2 Data normalization

Neural networks tend to perform better when the inputs and outputs of the network are normalized as shown by Sola et al. [**b13**]. Normalization generally increases the speed of convergence of the error and does not affect the accuracy. Raw data tends to be in different scales and this leads to different range of weights for different connections making it hard for the gradient descent algorithm to smoothly descend into the minima.

The dataset is divided into yearly training and testing sets, and each training set is scaled to a range of 0 to 1 using a Min-Max scaler. The Min-Max Scaler subtracts the minimum values of a feature from every value and divides it by the range (max - min) of the features. The same scaling is then applied to the testing data. The training and testing data are scaled separately to avoid data leakage. If the training and testing data are scaled together, it would lead to data leakage wherein some information regarding the future data is used to predict it. This is especially true when using scaling algorithms that involve the mean and variance of the data.

$$X_{scaled} = \frac{X_i - X_{min}}{X_{max} - X_{min}} \tag{4.2}$$

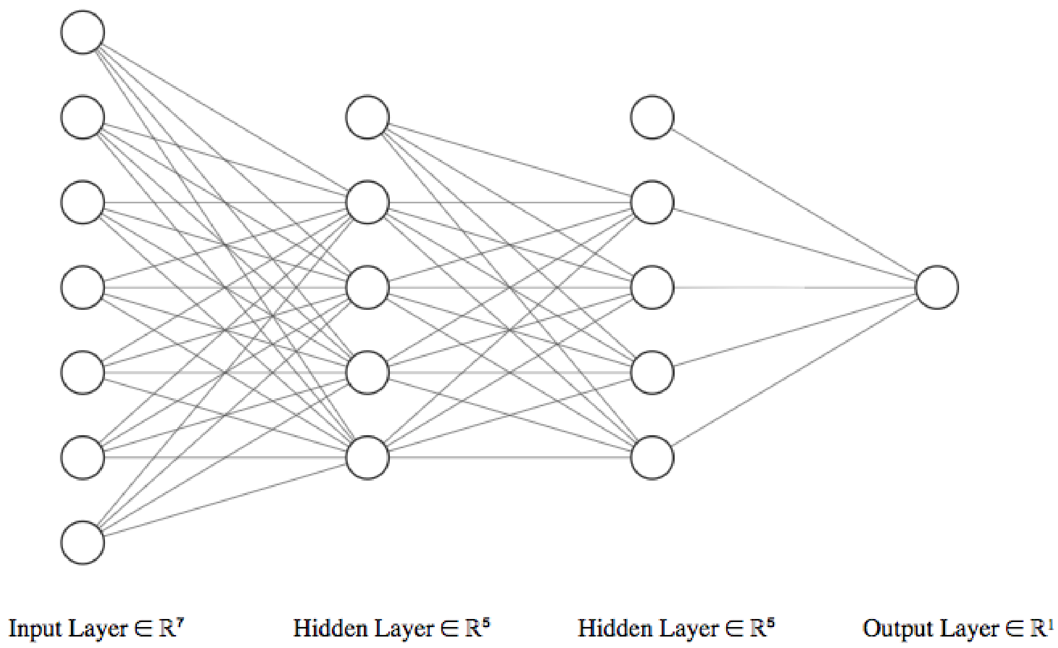

FIGURE 4.1: Irradiance Prediction Model Architecture

Chapter 5

Model Architecture

In this section, we detail our model architecture, which comprises two parts. The first part consists of the irradiance model shown in Fig. 4 which takes the weather forecast data as input and predicts the irradiance value. This model takes 7 input features including the 5 weather forecast variables stated above and the corresponding temporal data (day of the year and the hour of the day). It contains two hidden layers with 7 and 4 neurons each with the ReLU (Rectified Linear Unit) activation function. The output layer of the model contains a single neuron with a sigmoid activation function since the values are scaled between 0 and 1. This model is trained with the historic irradiance values to forecast future irradiance values as they are not provided as a part of a typical weather forecast. The second part consists of feeding the forecast irradiance values into a second neural network along with the weather forecast data, the nearest neighbours data and the previous day's production value. Separate models for every hour are not required for the irradiance prediction as we are considering the hourly models to learn the structural and temporal features more accurately, and using them for both the irradiance and power generation predictions will not improve the models accuracy.s Since the predictions are made hourly, we considered separate models for every hour of the day. The alternative was to consider a network with 24 outputs for every hour of the day but this lead to a decrease in the prediction accuracy due to the increase in the number of features. This increases in the number of features paired with densely connected MLP layers lead to the extraction of unwanted features from forecasts of different hours thereby reducing the accuracy. To obtain a similar accuracy using a single model, we had to use a much wider model which leads to an unnecessary increase in computation. The hourly models are definitely a better option especially for microgrids as they are also able to learn the periodic structural features such as shadows from buildings or obstructions which are significant only during the early and late part of days. The input layer of the hourly models

Neuron	*Variable Name*
i_1	Irradiance Predicted
i_2	UV Index
i_3	Cloud Cover
i_4	Humidity
i_5	Temperature
i_6	Minimum Neighbourhood output
i_7	Maximum Neighbourhood output
i_8	Mean Neighbourhood output
i_9	Day of the Year
i_{10}	Previous output for the hour

Table 1. Inputs to the Hourly prediction model

contains 10 features which are shown in Table 1. These hourly models contain two hidden layers with 4 neurons each and an output layer with a single neuron. The first hidden layer uses a ReLU activation function while the second layer has a Tanh (Hyperbolic Tangent) activation function. These activation functions were considered after trying different combinations and evaluating the convergence and accuracy. The output layer has a sigmoid activation function to keep in range with the scaled output values.

Chapter 6

Results

The above described model architectures were considered after experimenting with a different number of layers and neurons. The different iterations of the model were evaluated using multiple performance metrics. The performance metric that is most important for such data is the Normalized Mean Absolute Error (NMAE). The NMAE is defined as

$$NMAE = \frac{1}{N}\sum_{i=1}^{N}\frac{|P_{measured,n} - P_{forecast,n}|}{P_{rated}}.100 \tag{6.1}$$

where $P_{measured}$ is the actual power measured, $P_{forecast}$ is the predicted power for the same hour and P_{rated} is the rated power(Maximum Power) of the plant. The summation range N is chosen according to the time range considered. It will be the number of observations in a particular day if considering daily NMAE or the number of observations in a week if considering weekly NMAE.

The normalized mean absolute error calculates the error normalized to the size of the system making it suitable to compare systems of different sizes. All other absolute errors cannot be used to directly compare the performance of different systems, owing to the different sizes of solar power plants.

One can observe from Fig. 5 that the weekly NMAE for 42 weeks averages at around 15%. The weekly NMAE is calculated by carrying out the summation of the NMAE for all observations in a particular week. The Fig. 6 here shows the daily NMAE of the predictions. It can be seen that the NMAE is below 20% for most days. The NMAE gives an idea on what percent of error is observed in terms of the total rated power. While simpler models are able to predict with an NMAE of less than 10% for power plants whose production is in the range of a few 100 KW, they do not work well for smaller installations. Our model allows us to estimate the production with a maximum error of 20%. This accurate predictions are suitable for all the different applications such as unit commitment, electricity market, maintenance scheduling and electricity dispatch planning involving a microgrid.

		Mean NMAE %	
Year	*Type of Error*	**Dataset 1**	**Dataset 2**
2014	Weekly	14.5	13.5
	Daily	14.2	13.2
2015	Weekly	13.8	13.1
	Daily	13.9	13.4
2016	Weekly	16.8	14.6
	Daily	16.5	14.3

Table 2. Yearly NMAE for the 2 datasets

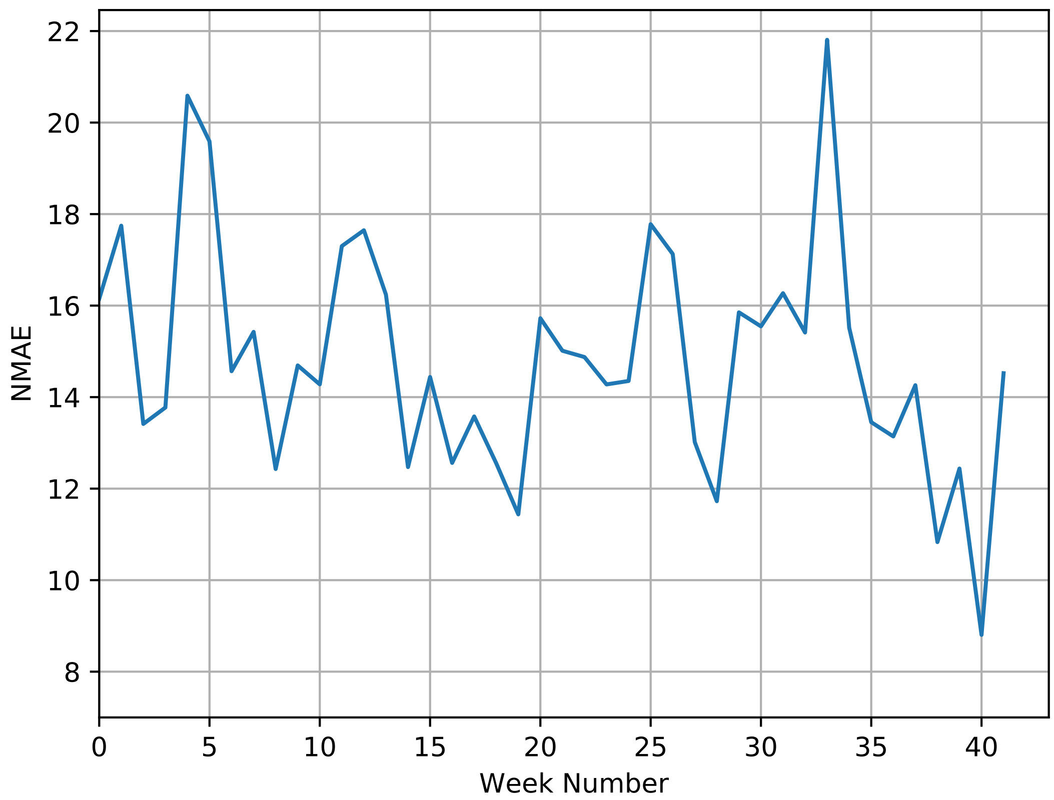

FIGURE 6.1: Weekly Normalized Mean Absolute Error

Fig. 7 shows the frequency distribution of the errors for the various hourly predictions. A positive Normalized Error% from this figure corresponds to an underestimated prediction while a negative Normalized Error% corresponds to an overestimated prediction. The fact that the model is underestimating more often than overestimating is of significance advantage as this can give an estimate on the minimum production and can help in planning distributing for the maximum production without overestimating.

Fig. 8 shows the actual and forecast daily energy productions for 100 days. The figure also shows the percentage error in green. The percentage error in the plot is the Mean Absolute Percentage Error (MAPE), which is calculated as follows

$$MAPE = \frac{X_{Actual} - X_{Predicted}}{X_{actual}}.100 \tag{6.2}$$

It can be seen that the model underestimates most of the time. This is a desirable characteristic in production prediction models as this allows the user to know the minimum production that can be expected. It should also be noted that the MAPE is considerably high for certain points. These points generally correspond to days with less production. MAPE tends to increase with lesser production because it is expressed as a ratio of the production. This however does not mean that the absolute error is high. When considering weekly or monthly productions, the days with lesser productions are insignificant when compared to the total production. In turn, the errors associated with those predictions are also not significant.

The model was also trained with different time frames of training data for a fixed testing

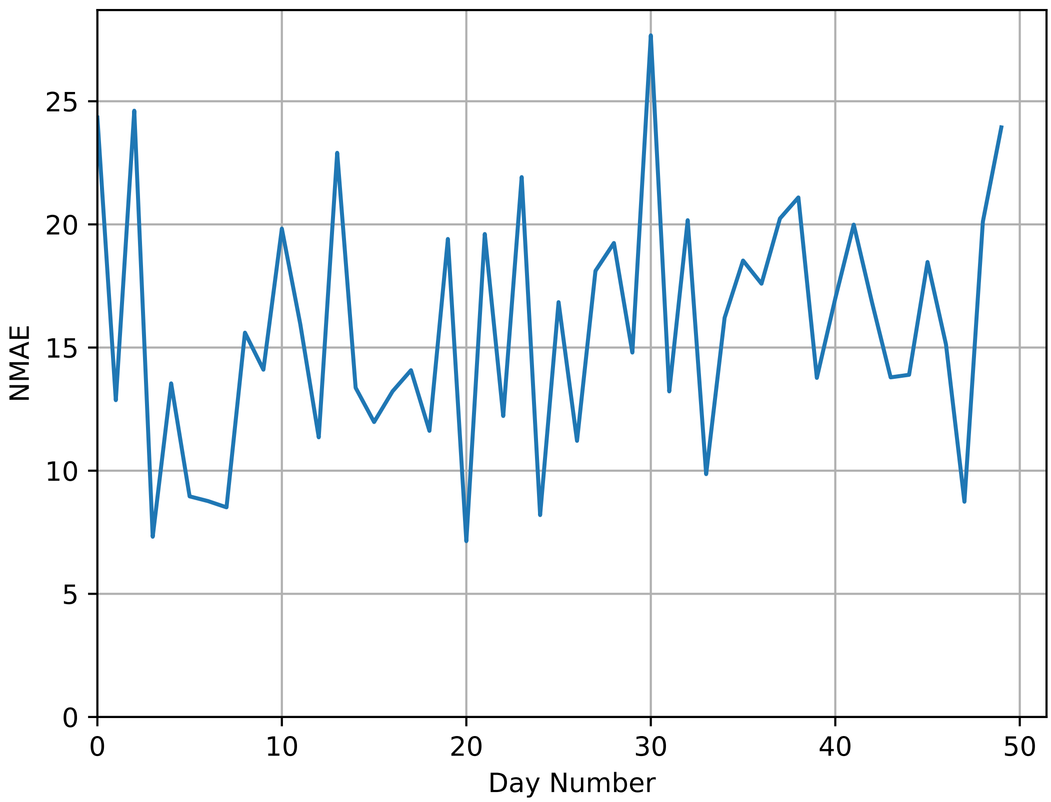

FIGURE 6.2: Daily Normalized Mean Absolute Error

data to evaluate the affect of the training data on the performance of the network. The best results were achieved when the training included the past 12-15 months for forecasting the future 6 months. It is also important to compare the hourly values for the training and the testing periods to check for consistency. Any new physical obstructions that might cast shadows on parts of the system or degradation in the infrastructure must be considered while deciding the training period. The solar panels also tend to degrade over the years due to aging and/or improper maintenance. It is therefore recommended not to train on data older than 2 years to avoid inconsistencies between the training and testing data. Fig. 9 illustrates that the model is able to accurately follow the measured power. The results are accurate considering that microgrids are highly sensitive to minute weather and climatic changes.

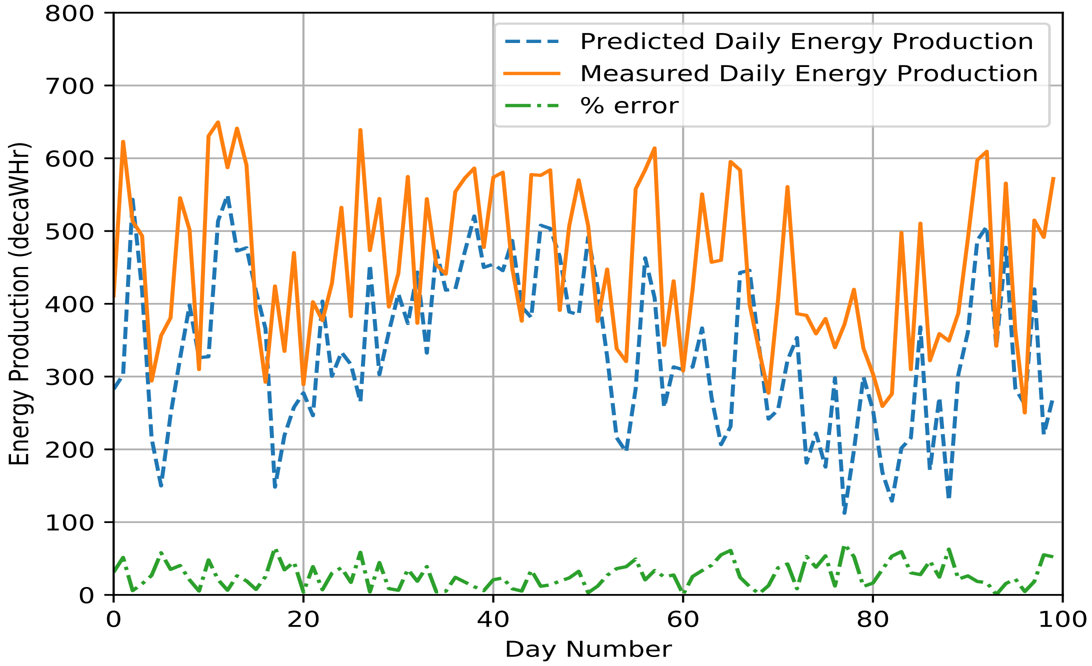

FIGURE 6.3: Daily Energy production and forecast

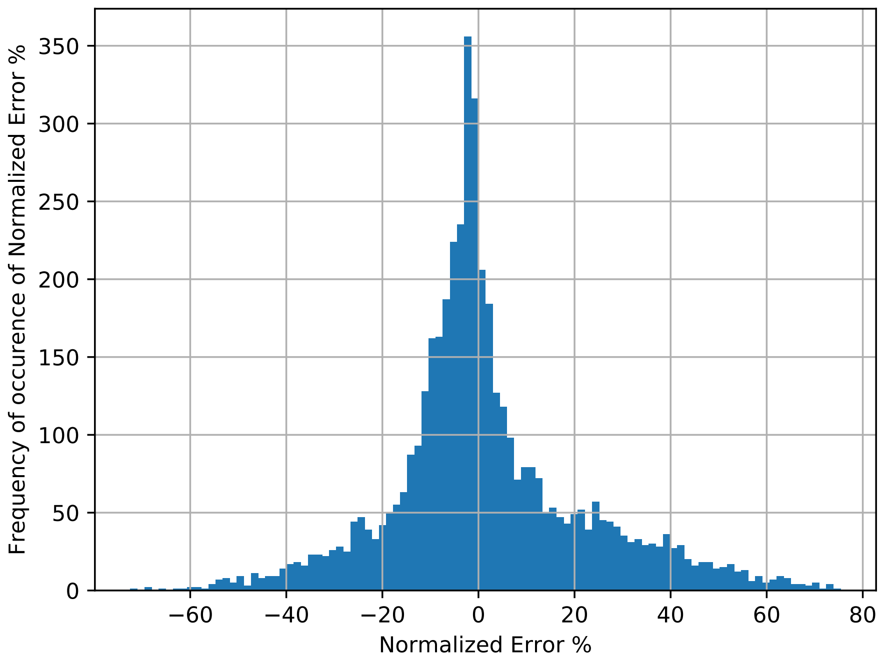

FIGURE 6.4: Frequency of occurrence of different NMAE%

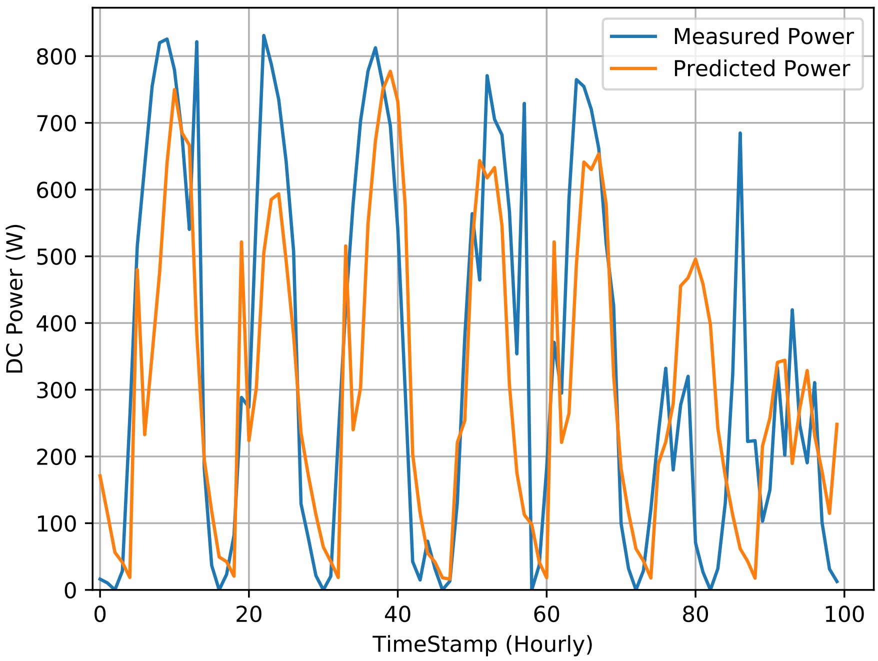

FIGURE 6.5: Hourly measured and forecast values for a week

Chapter 7

Conclusions

In this thesis we have outlined the current state of the art in the prediction of solar power production. We have described various papers and proposed a novel approach to predict hourly production of PV plants. The model proposed consists of two parts where the first part involved the use of neural networks to extract more appropriate features for the model. The second part involved hourly models which are trained to predict the production for a particular hour using the various input features as described in section V. The following sections reasoned in detail the decisions that led to the proposed architecture and model.

In section VI we have evaluated the model for different time frames and applications. It is important to note that the approach is tailored to predict the hourly production of micro and nano grids. The size of the grid is a huge factor in the prediction model, in terms of choosing the features and the model architecture. The proposed model is able to predict daily production values with an average NMAE of 15%. This average is over a testing period of 400 days. The weekly NMAE for over 42 weeks averages at 13%.

It can be concluded that the solar plant production has a certain dependency on the weather conditions but it should be noted that there is a limit to the accuracy in forecasting that can be achieved by using only the weather forecast data. This is especially true for microgrids due to the inaccuracies in the weather forecast. Nevertheless, the accuracies of forecast obtained on an hourly, daily and weekly basis are sufficient to benefit the short and medium term planning of solar power plant.

Bibliography

[1] Barreau Nicolas , Photovolta2 Project, Université de Nantes, 2011-2019. URL : http://photovolta2.univ-nantes.fr

[2] G. Habault, M. Lefrançois, F. Lemercier, N. Montavont, P. Chatzimisios and G. Z. Papadopoulos, "Monitoring Traffic Optimization in a Smart Grid," in IEEE Transactions on Industrial Informatics, vol. 13, no. 6, pp. 3246-3255, Dec. 2017.

[3] F. Grimaccia, M. Mussetta and R. Zich, "Neuro-fuzzy predictive model for PV energy production based on weather forecast," 2011 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE 2011), Taipei, 2011, pp. 2454-2457.

[4] C. Wan, J. Zhao, Y. Song, Z. Xu, J. Lin and Z. Hu, "Photovoltaic and solar power forecasting for smart grid energy management," in CSEE Journal of Power and Energy Systems, vol. 1, no. 4, pp. 38-46, Dec. 2015.

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

[6] M. Kuzlu, M. Pipattanasompom and S. Rahman, "A comprehensive review of smart grid related standards and protocols," 2017 5th International Istanbul Smart Grid and Cities Congress and Fair (ICSG), Istanbul, 2017, pp. 12-16.

[7] Alexis Bocquet, Andrea Michiorri, Arthur Bossavy, Robin Girard, Georges Kariniotakis. Assessment of probabilistic PV production forecasts performance in an operational context. 6th Solar Integration Workshop - International Workshop on Integration of Solar Power into Power Systems, Nov 2016, Vienna, Austria. pp.6 - ISBN 978-3-9816549-3-6. ffhal-01409042f

[8] L. Gigoni et al., "Day-Ahead Hourly Forecasting of Power Generation From Photovoltaic Plants," in IEEE Transactions on Sustainable Energy, vol. 9, no. 2, pp. 831-842, April 2018.

[9] Nespoli, A.; Ogliari, E.; Leva, S.; Massi Pavan, A.; Mellit, A.; Lughi, V.; Dolara, A. Day-Ahead Photovoltaic Forecasting: A Comparison of the Most Effective Techniques. Energies 2019, 12, 1621.

[10] Massucco, S. & Mosaico, Gabriele & Saviozzi, Matteo & Silvestro, Federico. (2019). 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. Energies. 12. 1298. 10.3390/en12071298.

[11] P. L. ezhnyuk, V. Komar, S. Kravchuk, V. Lesko and V. Netrebskiy, "Meteorological Parameters Analysis for hourly Forecast of Electricity Generation by Photovoltaic Power Station on the Day Ahead," 2018 IEEE 3rd International Conference on Intelligent Energy and Power Systems (IEPS), Kharkiv, 2018, pp. 235-238.

[12] J. Sola and J. Sevilla, "Importance of input data normalization for the application of neural networks to complex industrial problems," in IEEE Transactions on Nuclear Science, vol. 44, no. 3, pp. 1464-1468, June 1997.

[13] M. Marinelli, F. Sossan, G.T. Costanzo, and H.W. Bindner, Testing of a Predictive Control Strategy for Balancing Renewable Sources in a Microgrid, IEEE Transactions on Sustainable Energy, Vol. 5, No. 4, pp. 1426-1433, 2014.

[14] J.M. Morales, A.J. Conejo, H. Madsen, P. Pinson, and M. Zugno, Integrating Renewables in Electricity Markets, Springer, New York 2014.

[15] M. Cococcioni, E. D'Andrea and B. Lazzerini, "24-hour-ahead forecasting of energy production in solar PV systems," 2011 11th International Conference on Intelligent Systems Design and Applications, Cordoba, 2011, pp. 1276-1281.

[16] S. Rahman and M. Pipattanasomporn, "Smart grid information Clearinghouse: Overview of projects and deployment experience," 2011 IEEE PES CONFERENCE ON INNOVATIVE SMART GRID TECHNOLOGIES LATIN AMERICA (ISGT LA), Medellin, 2011, pp. 1-7.

[17] H. C. Hottel, "A simple model for estimating the transmittance of direct solar radiation through clear atmospheres," Solar Energy, vol. 18, pp. 129-134, 1976/01/01 1976.

[18] N. Anku, J. Abayatcye and S. Oguah, "Smart grid: An assessment of opportunities and challenges in its deployment in the ghana power system," 2013 IEEE PES Innovative Smart Grid Technologies Conference (ISGT), Washington, DC, 2013, pp. 1-5.

[19] B. Asato, T. Goya, K. Uchida, A. Yona, T. Senjyu, T. Funabashi, C-H. Kim, "Optimal operation of smart grid in isolated island," 2010 Conference Proceedings IPEC, Singapore, 2010, pp. 1100-1105.

[20] Xiaoling Jin, Yibin Zhang and Xue Wang, "Strategy and coordinated development of strong and smart grid," IEEE PES Innovative Smart Grid Technologies, Tianjin, 2012, pp. 1-4.

[21] "Technologies for a smart grid development with examples of projects in Europe," 2012 IEEE PES Innovative Smart Grid Technologies (ISGT), Washington, DC, 2012, pp. 1-1.

[22] P. Bansal and A. Singh, "Smart metering in smart grid framework: A review," 2016 Fourth International Conference on Parallel, Distributed and Grid Computing (PDGC), Waknaghat, 2016, pp. 174-176.

[23] Ren Yanming et al., "The research on communication standard framework of smart grid," CICED 2010 Proceedings, Nanjing, 2010, pp. 1-6.

[24] M. Kuzlu, M. Pipattanasompom and S. Rahman, "A comprehensive review of smart grid related standards and protocols," 2017 5th International Istanbul Smart Grid and Cities Congress and Fair (ICSG), Istanbul, 2017, pp. 12-16.

[25] M. Saleh, M. A. Abdou and M. Aboulhassan, "Assessing the use of IP network management protocols in smart grids," 2016 IEEE/ACS 13th International Conference of Computer Systems and Applications (AICCSA), Agadir, 2016, pp. 1-6.