March - April 2020
ISSN: 0193-4120 Page No. 132 - 139

# Solar Forecasting Performance using Artificial Intelligence Algorithms

Jagadeesh.V[1]
*Mechanical Engineering Dept,*
*GMRIT - Rajam*

Dr. Venkata Subbaiah .K[1]
*Prof, Mechanical Dept, Andhra University*
*Visakhapatnam, India*

***Article Info***
***Volume 83***
***Page Number: 132 - 139***
***Publication Issue:***
***March - April 2020***

***Article History***
***Article Received**: 24 July 2019*
***Revised**: 12 September 2019*
***Accepted**: 15 February 2020*
***Publication**: 12 March 2020*

***Abstract***

Renewable energy sources, especially Solar Energy, are to play a larger role in Hybrid Generation in the upcoming and future energy and utility requirements. The importance of renewable energies has been growing at a fast pace, both because of the need to solve problems related to environmental issues and as a way of helping the increasingly difficult management of electricity grids. Solar power forecasting can solve number of the equality issues, in the concern of accurate forecasts of solar output from design and equipments. The techniques of Artificial Intelligence have already shown their effectiveness in tasks of high complexity, namely, Regression, Classification, and Forecasting. Also in the field of Renewable energies these tools can be extremely useful, in particular in the prediction of Solar Irradiance.we developed two algorithms in Python of prediction of Solar Irradiance based on two methods of Artificial Intelligence, which are the Artificial Neural Networks and the K-Nearest Neighbors Method. In the forecasting process, the models are trained with subsets of the one-year Solar Irradiance register in the city of Lisbon and then the next hour's forecast is carried out. In order to understand the best method to perform predictions of solar irradiance among those studied, a comparative study between the models was carried out, taking into consideration the prediction errors and the simulation times of both models in the simulations made in different situations.

***Keywords:*** *ANN, KNN, Solar Energy. Forecasting, renewable energy, machine learning.*

## 1. introduction

Renewable Energy is becoming a technology and an increasingly viable alternative to the Conventional as Non-Renewable Sources of Energy. The constant threat of the Climate Changes thus requires the mankind to look for better and more efficient ways of producing electricity. Especially when all our basic needs are based on this type of energy. The total amount of Solar Energy delivered by the Sun to the Earth is unimaginable. The San Francisco earthquake of 1906 reached a 7:8 Richter's magnitude equal to $10^{17}$J of released energy, which is the same amount of energy delivered by the Sun per second. The total World Oil Reserves (about1.7 _ $10^{22}$J) are exactly the same amount of energy received by the Earth in a day and a half. During an hour, the sun delivers to the Earth the same amount used in the human activities during a year4.6 x $10^{20}$J).After the big increase in the use of the Hydro and Wind Power, now it is the Solar Power that is showing some signs of increasing its use and it is one in which it has been invested most in order to improve its efficiency. The technology associated with the Solar Power is improving at a good pace as we can see by the appearance of new types of cell candidates to the

March - April 2020
ISSN: 0193-4120 Page No. 132 - 139

replacement of the traditional and still expensive Crystalline Silicon Cells, like the Thin-Film Technologies, the Multifunction Cells or the Emerging Photovoltaic Technologies (i.e. Organic Cells, Dye-sensitized Cells or Perovskite Cells). These facts, coupled with the global need to reduce Greenhouse Gas emissions, make the Solar Power one of the most promising Renewable Energies[1-5].

NOx emission control is becoming a public concern and an alternative source of energy generation has become alternative solution to fulfill the energy demand. Renewable energy (Solar and wind energy) is free and abundant in nature [6].

The solar PV systems to the electricity supply is increasing continuously. An efficient use of the fluctuating solar power production will be highly useful from the forecast data on the expected power generation [7].

The increase in the amount of installed capacity of renewable power sources is an important driving factor toward sustainable power generation Penetration of such renewable sources, along with the intermittency and randomness of their generated power, has posed challenges to power engineers in recent years Power generation is generally planned for real-time, daily, weekly, monthly, and yearly for reliable operation and economic dispatch [8].

Renewable Sources of PV plant collected for the solar forecasting [11].

## 2. Motivation

Currently, Solar Power is a well-known Renewable Energy Source (RES) with a great potential. But his dependence on external meteorological conditions turns this technology somewhat volatile and sometimes unattractive both at the energy and at the economic level. Currently the use of Artificial Intelligence Tools is becoming more and more common due to their capacity of solving highly complex problems. The improvement in the computers and in the algorithms performance also helped at solving problems, not only in engineering, but also in many areas like medicine, finances and literature. That is why these tools are becoming more popular and can still make some progress and can be applied in more areas or problems.

In the case of RES these artificial intelligence tools are used to perform forecasts of the energy produced in a Power Station or even forecasts of the behavior of weather conditions. These forecasts plays a vital role in the Energy Market. The issue of uncertainty of the Solar Irradiance that comes to the Earth surface is well-known. Thus, an forecasts are supposed to be accurate and this can be further useful for planning and operation and power supply and demand at economic level or energy production level, either by alternative sources for traditional power and overall the right quantity of energy resources and reserves in order to reduce the power system operational costs. In addition to these advantages a good forecast can help to reduce the impact of PV output uncertainty on the grid, improve the system reliability, maintain the Power Quality and increase the penetration level of PV Systems. The forecasts generated by these tools can also be useful for studying the viability of solar production projects in a given location. There were several relevant reasons for choosing this theme of the thesis. The evaluation of the importance that this kind of technology has but also the potential that it can bring to the world necessities, the understanding of Portugal's current situation on the Energy Sector, the principal characteristics of the resource used in this kind of technology and the main advantages of forecasting Solar Irradiance became key factors in the theme framework [1-5].

*March - April 2020*
*ISSN: 0193-4120 Page No. 132 - 139*

In Table 1.1 some examples of application of Artificial Intelligence Tools in different study fields.

Table 1.1: Fields of Application of AI Tools

| Field | Description |
|---|---|
| Automotive | Automatic guidance systems, automatic braking systems, misfire detection, virtual emission sensors |
| Banking | Credit application evaluators, cashforecasting,firm classification,exchange rate forecasting |
| Electronics | Code sequence prediction, process control,chip failureanalysis,non linear modelling |
| Medical | Breast cancer cell analysis, EEG and ECG analysis, optimization of transplant times, emergency room test advisement |
| Speech | Speech recognition, speech compression, vowelclassification, text to speech synthesis |
| Telecommunications | Image and data compression, Automated information services, real-time translationof spoken language,customerpayment processing |

## 3. The Importance of the Solar Power and Global Overview

Since the Industrial Revolution that the use of Fossil Fuels was of big importance for the Energy Supplying every sectors (e.g. Industry, Transport, Domestic use). After several studies about the effects of the use of Fossil Fuels, Humanity became aware of the harmful effects on the environment and on living things caused by burning this kind of fuels to produce energy. The energy production process using Fossil Fuels releases large quantities of Carbon Dioxide (CO2 is the most anthropogenic GHG) into the atmosphere, which increases the Global Earth Temperature As a consequence, it causes the melting of polar ice caps, and increases the average level of the sea water .With this evidence of the danger of overuse of Fossil Fuels, Governments and Societies have taken a more environmentalist position by adopting measures and developing new technologies to reduce the dependence on Fossil Fuels and the damage to the Environment. The principal objective of this agreement is to limit the global warming to well below $2^{\circ}C$ through several actions in different areas like Mitigation of Emissions (i.e. undertaking rapid reductions in accordance with the best available science), Transparency and Global Stock take [9-10]

## 4. Forecasting Models

Starting with the Persistence Model, the reference model used to evaluate the performance of the forecasts, followed by the definition of the models studied and compared in this work, the Artificial Neural Networks Modeland the K Nearest Neighbors Model.

### 4.1 Persistence Model

Th Model is the simplest forecasting model and is usually used as baseline or comparison for other models. During the study of new forecasting models, the performance of these models will be considered if they perform better than the this Model. This Model calculates future time series values with the premise that all the effecting condition does not change between time t and time $t + \Delta t$ .For this work, this premise is applied to the evolution of irradiance and the model considers that their radiance for time $t + \Delta t$, or the next irradiance value, is equal to irradiance for time t, or the actual irradiance value (example in Figure 4.1). Mathematically this can be represented by the equation (1):

$$I(t + \Delta t) = I(t) \quad \text{Equation 1}$$

Where $I(t + \Delta t)$ is the irradiance for future time (where $\Delta t$ can be any time interval) and I(t) the current

value of irradiance.

### 4.2 The Importance of Artificial Neural Networks (ANN)

The Artificial Neural Networks (ANN) Model is an Artificial Intelligence method based on the human capacity of learning and adapt his way of thinking through his obtained life time experience This method is capable of compute nonlinear modeling without knowing in the beginning the relation between input and output variables thanks to his nonlinear data-driven structures.

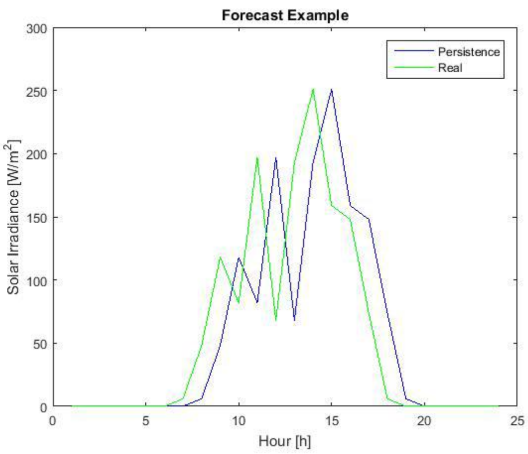


Figure 1: Solar Forecast example to explain the Persistence Model Forecast

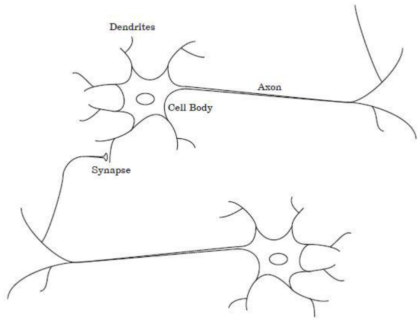


Figure 4.2: Biological Neuron

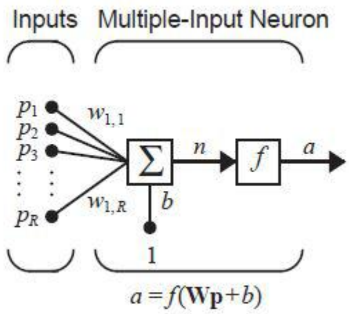


Figure 4.3: Artificial Neuron

Comparing both biological and artificial neural networks (Figure 4.2) we can observe two key similarities between them. First, the building blocks in each case are simple computational devices and highly interconnected. And for second, the function of the network is, in both cases, determined by the connections between neurons. Their structures have also characteristics that can be considered similar. The Biological Neuron, who has been taken as inspiration for this model, can be divided in three main parts, the Dendrites, a tree-shaped network of nerve fibers that receive from other neurons and conducts the electric signal to the Cell Body, the Axon, a single long fiber that conducts the signal from the Cell Body to other Neurons, and the Cell Body, who sums and integrate the synaptic information before sends it through the axon in addition to his capacity to perform a variety of biochemical processes in order to keep the neuron functioning properly. The point of connection between two different neurons is called Synapse. Depending on the arrangement of the different neurons and their

March - April 2020
ISSN: 0193-4120 Page No. 132 - 139

strengths of each synapse the neural network will perform different functions

The Artificial Neurons are composed by a Sum Blockconnected to a Transfer Function. The inputs of the Sum Block are the Input Vector multiplied by the correspondent Weight Vector and the Bias. After the sum the result, called Net Input, goes into a Transfer Function, and the result produced by it is called the Neuron Output. The Transfer Function depends on the specification of the problem that the neuron is trying to resolve and can be Linear or Non-Linear(Figure 4.3). Relating with the biological model we can correspond the Weights Vector to the strength of the synapses, to the set composed by the summation and the transfer function we can correspond to the cell body, and the result output from the artificial neural network can be equated to the electrical signal on the axon

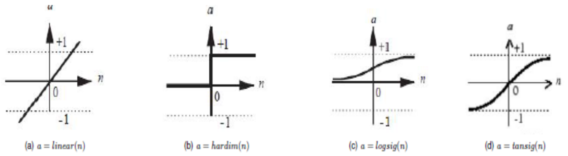


Figure 4.4: Example of Transfer Functions

Like a brain or a biological neural network, an ANN is composed by a lot of Neurons and the way that these Neurons are connected between them defines the type of architecture we are using. Usually an ANN is composed by different Layers, a group of neurons receiving the same inputs (Figure 3.4).In this case, each neuron has his Weight Vector and produce a determined Output. Simplifying the mathematical notation and gathering all weights, bias and outputs, each layer can be represented bya Weight Matrix, a Bias Vector and an Output Vector. All the basic elements can be represented by mathematical expression. The Net Input for the i neuron of the k + 1 layer can be expressed by:

$$n^{k+1}(i) = \sum_{j=1}^{S^k} W^{k+1}(i,j) \times a^k(j) + b^{k+1}(i) \quad \text{Equation}$$
$$k = 0,1,2,.......M-1$$

2

where Wk+1(i; j) is the Weight Vector for the that neuron, ak(j) is the input from the output of the previous layer and bk+1(i) is the Bias orrespondent of that neuron.When k = 0 the input of that neuron is equal to the input of the Neural Network. The Neuron Out-put for the same neuron and layer can be expressed by:

$$n^{k+1}(i) = f^{k+1}(n^{k+1}(i) \quad \text{--------Eq4.1}$$
$$k = 0,1,2,.......M-1$$

Where fk+1(…) is the Transfer Function for the k + 1 layer.

The layer that is responsible to produce the problem’s output is the Output Layer. The remaining network ‘slayers are considered the Hidden Layers. The number of elements in an ANN (i.e. number ofinputs or outputs, number of layers, number of neurons per layer) is quite arbitrary but must obey some rules. The number of inputs and outputs of the network are defined by the problem specifications, the

number of inputs of the network is equal to the number of external variables to be used as inputs and the number of neurons in the Output Layer is equal to the number of outputs of the problem. The number of neurons in the remaining layers will only influence the complexity of the network and this will depend on the problem taken, but the number of neurons of a layer must be equal to the number of inputs of the next layer.

The training algorithm used in this ANN is the Levenberg-Marquardt Algorithm. This algorithm is anvariation on Back propagation Learning Algorithm which is an Supervised Learning Algorithm. The Back propagation algorithm starts with the Propagation of Input through the network with the equation 4.1. After that the algorithm


Published by: The Mattingley Publishing Co., Inc.

March - April 2020
ISSN: 0193-4120 Page No. 132 - 139

calculate and propagates the sensitivities back using the next recurrence relation:

$$\underline{\delta}^M = -\dot{F}^M(\underline{n}^M)(\underline{t}_q - \underline{a}_q)$$

$$\underline{\delta}^k = \dot{F}^k(\underline{n}^k)W^{k+1^T}\underline{\delta}^{k+1}, \quad k = 0,1,2,...,M-1$$

$$\dot{F}(\underline{n}^k) = \begin{bmatrix} \dot{f}^k(n_1^k) & 0 & \cdots & 0 \\ 0 & \dot{f}^k(n_2^k) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \dot{f}^k(n_{S_k}^k) \end{bmatrix}$$

Equation 3 is the matrix composed by the derivatives of the transfer functions, and $t_{-q}$ and $a_{-q}$ are, respectively, thevector of the desired targets and the vector of the output resulting from equation 4.1. The recurrence relation always initializes at the final layer using the equation 4.1 and propagates the sensitivities through equation 4.2. The last step of each it eration is the update of the weights and offsets of the networks. For that, and using the chain rule, we can obtain the update values for the weights and bias of the network through the equations:

$$\Delta\omega^k(i,j) = -\alpha\frac{\partial\hat{V}}{\partial\omega^k(i,j)} = -\alpha\partial^k(i)a^{k-1}(j)$$

$$\Delta b^k(i) = -\alpha\frac{\partial\hat{V}}{\partial b^k(i)} = -\alpha\partial^k(i)$$

Equation 4

Where $\alpha$ is the learning rate and ^ V is the performance index. This process is repeated with different input vectors, until the performance error is small enough or the number of iterations reached the maximum number. The Leven berg-Marquardt Algorithm, in addition to having derived from Back propagation Algorithm, was based on the Newton Method

**4.3 K-Nearest Neighbors (KNN)**

The K-Nearest Neighbor (KNN) Model is, like the ANN method, an Artificial Intelligence method, beingalso considered a System of Pattern Recognition This type of systems are able to identify a random object knowing previously the nature of one set of similar objects.

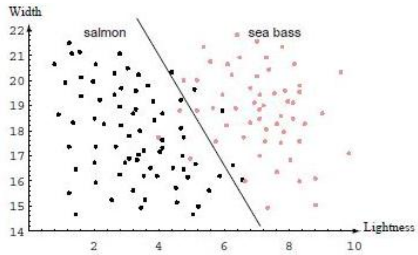


Figure 5.1: Example of KNN Classification Problem

This method started to be used in Classification problems. The main objective of these problems is to assign a classification to a chosen point based on the class of the training set points. This process can be compared to the human sensory capacity like recognizing a face, identify our accessories in our bag by feeling their shape, understanding different words in different languages by listening or decide if an apple is ripe only by its smell. This Model can also be classified according to his training algorithm and characteristics. Just like the ANN, the KNN Learning Process is classified as a Supervised Learning Process because it attributes teach input object an target value. These systems can be divided into Parametric or Nonparametric Techniques. In the first case is known the expression for Probability Density or other Discriminate Functions(e.g.Maximum Likelihood Method). The Nonparametric Techniques does not possess that information (Discriminant Functions) a priori. In this case, the estimation problem is formulated with great generality in a space of non-negative functions, thus not restricting the probability density function to belong to are stricted class of functions that depend on a parameter vector. It is in this class of techniques that the KNN Model is inserted.

The KNN method is an application of the Maximum a Posteriori Probability (MAP) Classification with the Probability Density

TEST Engineering & Management

*March - April 2020*
*ISSN: 0193-4120 Page No. 132 - 139*

Functions obtained by the Parzen Method with Adaptive Window. Mathematically lets consider X a set composed by N classified training patterns and Xi a subset of X generated by the i-th class. By the MAP classifier, the classification of a pattern x is given by

$$\hat{\omega} = \arg\max_{\omega} \hat{p}(x/\omega)\hat{P}(\omega)$$

where ^p(x=!) is the probability density function estimation and ^ P(!) is the a priori probability estimationfor class !. The estimation ^p(x=!) can be obtain through the Parzen Method. The window used in pointx is chosen so that only the k nearest patterns to x are inside the same window and thus contributing to obtain the probability density functions:

$$w(x-\tilde{x}) = \begin{cases} c, & \tilde{x} \in V_\delta(x) \\ 0, & \text{otherwise} \end{cases}$$

Equation 5

Where c is normalization constant and:

$$V_\delta(x) = \{\tilde{x} : \|\tilde{x} - x\| < \delta\}$$

Equation 6

is the circle region of radius $\delta$ centered in x, composed of points whose distance to x is less than $\delta$. As the aim is to ensure that the window possess only the k-nearest training patterns to x, we can get this bydefining the value of $\delta$ as the distance of the (k +1)-th nearest training pattern to x. From the probability calculus and from the MAP classification we know

$$\hat{P}(\omega_i) = \frac{N_i}{N}, \quad i = 1, \ldots, t_c$$

$$\hat{p}(x) = \frac{1}{N} \sum_{\tilde{x} \in X} w(x - \tilde{x})$$

Equation 7

The KNN Method can also be used to solve Regression Problems. In this case, instead of classify, the main objective is to determine the numerical value of a variable of a unknown case, which means, each training pattern is not associated to a class but to a numerical value of a specific variable. To solve problem like this, the algorithm of KNN Regression is in all similar to the one of KNN Classification except for the last steps. It starts the same way calculating the distance between all the training patterns and the point to be studied, x. After this, and after discovering the k-nearest patterns, the variable numerical value of x is equal to the average of all the variable numerical values of the k-Nearest Neighbors.

## 5 Results and Discussions

Statistical Modeling : **Table 5.1**

```
In [135]: df.describe()
Out[135]:
```

| | electricity | irradiance_direct | irradiance_diffuse | temperature |
|---|---|---|---|---|
| count | 8760.000000 | 8760.000000 | 8760.000000 | 8760.000000 |
| mean | 0.191942 | 0.176694 | 0.073752 | 25.421458 |
| std | 0.251086 | 0.263364 | 0.091673 | 9.655811 |
| min | 0.000000 | 0.000000 | 0.000000 | 1.999000 |
| 25% | 0.000000 | 0.000000 | 0.000000 | 17.956000 |
| 50% | 0.003000 | 0.000000 | 0.009000 | 26.503000 |
| 75% | 0.414250 | 0.336250 | 0.137000 | 32.810750 |
| max | 0.806000 | 0.990000 | 0.420000 | 48.089000 |

**Understanding the predictive analytics and behavior pattern:**

Most values are lie in between the first and second quartile, there is no much variance in these graphs,

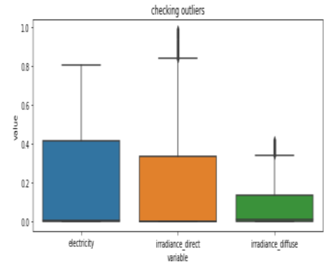


Figure 5.2: Checking four outliers:

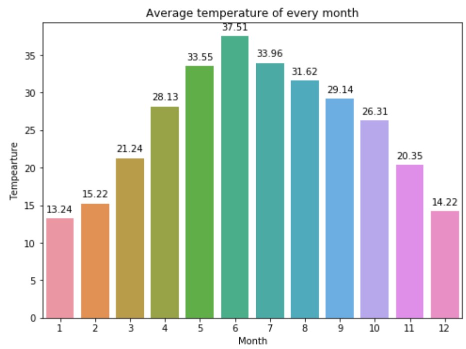


Figure 5.3: Average temperature of every month

March - April 2020
ISSN: 0193-4120 Page No. 132 - 139

## Conclusions

Throughout this work it was possible to learn more about the Artificial Intelligence Techniques, in particular about the ANN and the KNN models and how these models could be applied to the Solar Irradiance Forecast. In the beginning of the process we could understand the sustainability reasons that make the commitment to RES a global need and the work that has already been made earlier as well. This ends with a summary of the work already been done in the Solar Irradiance Forecast area. Then, it was presented a theoretical view of both models, where was discussed the points that served as inspiration for the creation of these models, the mathematical foundations that support them and the particularities of the models studied in this work.After this, it was developed two different algorithms using the studied methods ANN and KNN to perform Solar Irradiance Forecast and before the final tests, the models were tuned in order to perform the best forecasts.After these small optimizations, the models were subjected to two tests. The first one to understand the performance of the models on seasonal days, training with elements of one season of the year to predict another day of the same season. The second test was intended to evaluate the performance of both methods for a typical clear day profile and a typical cloudy day profile.The results on the performance tests showed us that both methods are better alternatives to the simple Persistence Model. In terms of results, both methods obtained average forecast errors below 30% of MAPE. The only exception was for the Cloudy Day tests where they got errors between 50% and 30%.The KNN showed the best results by surpassing ANN performance in every tests except for the Cloudy Day Test where he obtained an average of 46.16% and the ANN obtained an average of 35.49%. It is worth noting by the positive the great forecasts of the Models in the Spring and Summer forecast(ANN predicted with errors below 10% in both days and KNN also predicted below the same result in.

## References

[1] R. Castro, Uma Introduc¸ ˜ao `as EnergiasRenov´aveis: E´ olica, Fotovoltaica e Mini-H´ıdrica, 1st ed.IST Press, 2011.

[2] (2017, February) Facts: Vital Signs - Global Temperature. NASA.[Online]. Available: https://climate.nasa.gov/vital-signs/global-temperature/

[3] (2017, February) Facts: Vital Signs - Sea Level. NASA.[Online]. Available: https: //climate.nasa.gov/vital-signs/sea-level/

[4] (2017, January) Kyoto Protocol. United Nations Framework Convention on Climate Change. [Online]. Available: http://unfccc.int/kyoto protocol/items/2830.php

[5] (2017, April) 2020 climate & energy package. European Comission.[Online]. Available:

[6] Hossein Sangrody , Morteza Sarailoo, Ning Zhou, Nhu Tran, Mahdi Motalleb, Elham Foruzan "Weather forecasting error in solar energy forecasting" *IET Renew. Power Gener.*, 2017, Vol. 11 Iss. 10, pp. 1274-1280

[7] Elke Lorenz, johannes Hurka,Detlev Heinemann,and Hans Georg Beyer "Irradiance Forecasting for the power prediction of Gride-Connected Photovoltaic Systems" IEEE journal of selected topics in applied earth observations and remote sensing.vol.2,No.1,March2009.

[8] P. Bacher, H. Madsen, and H. A. Nielsen, "Online short-term solar power forecasting," Solar Energy, vol. 83, no. 10, pp. 1772–1783, October2009.

[9] (2017, April) Paris Agreement. European Comission. [Online]. Available: https://ec.europa.eu/clima/policies/international/negotiations/paris en

[10] G. Carr. (2012, November) Sunny uplands: Alternative energy will no longerbe alternative. The Economist.[Online]. Available: http://www.economist.com/news/2156641

[11] https://www.renewables.ninja/