SOLAR ENERGY PRODUCTION FORECASTING USING AI/ML

B.TECH SEM – VII Mini PROJECT
Dept. of Computer Science & Engineering

By
Group ID : 4

Saumya Shah **18BCP102**
Sagar Sinha **18BCP094**

<u>Under the Supervision</u>
<u>Of</u>
Dr. Rajeev Kumar Gupta

SCHOOL OF TECHNOLOGY
PANDIT DEENDAYAL ENERGY UNIVERSITY
GANDHINAGAR, GUJARAT, INDIA

July – December, 2021

CONTENTS

INTRODUCTION

Solar Panels have a fix range of output

- Generally 18-22% of incident solar radiation. This is a linear conversion rate which can be easily calculated.

To accurately forecast solar energy production we need to predict incident solar radiation

- The maximum output of incident radiation is only due to Global Horizontal Index (GHI) also known as Global Solar Insolation.
- Pyranometers are used to get accurate readings of incident solar radiations. These radiations are dependent on the weather as well as meteorological data.

LITERATURE REVIEW

To model the forecasting of Solar Irradiance, we require :

- domain knowledge
- hands-on experience on different available models

We did a comparative analysis of existing techniques, types of input parameters, and performance of various model using numerous evaluation metrics. The review paper has been accepted at an IEEE conference on Technology and Research In BEtterment of Society (TRIBES 2021).

Solar Irradiation Forecasting - Comparative Analysis Of Various Methods

Meshva Dhrumil Patel
Pandit Deendayal Energy University
Gandhinagar, India
meshva.pce18@sot.pdpu.ac.in

Saumya Shah
Pandit Deendayal Energy University
Gandhinagar, India
saumya.sce18@sot.pdpu.ac.in

Sagar Sinha
Pandit Deendayal Energy University
Gandhinagar, India
sagar.sce18@sot.pdpu.ac.in

Rajeev Kumar Gupta
Pandit Deendayal Energy University
Gandhinagar, India
Rajeev.gupta@sot.pdpu.ac.in

Santosh Kumar Bharti
Pandit Deendayal Energy University
Gandhinagar, India
Santosh.bharti@sot.pdpu.ac.in

Ved Prakash Singh
India Meteorological Dept, Ministry of Earth Sciences, Govt. of India, Bhopal
kvpssc@gmail.com

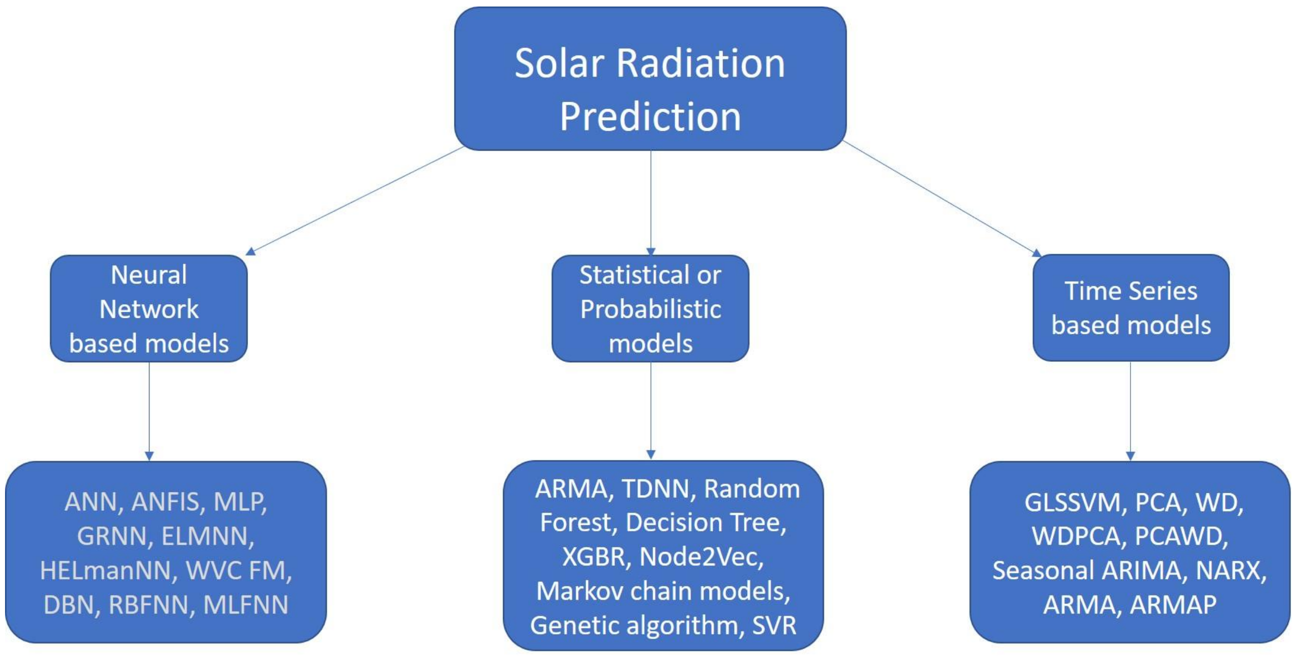
Abstract **— The global energy system is transforming with the consideration of a growing population with rising living standards that will need more energy. Simultaneously, the world must find ways to reduce greenhouse gas emissions and provide sustainable energy production. As electricity is the fastest-growing part of the energy system and thus shifting to renewable energy production systems is the need of the hour. IEA reports that the Sun could be one of the largest resources for energy production for the upcoming 'net zero goals' of the world. Rapid decrease in the cost of installation and availability of solar radiation are major advantages for adoption, although this mode of generation of energy has a variable output which is influenced by several parameters. Hence solar energy forecasting is crucial to understand and predict the output of production, the demand of the consumers, and optimizing the dispatch of the electricity from grids to users. In this paper, we have analyzed innovative methodologies for forecasting solar**

factors and other parameters it depends on. This poses issues with the grid reliability and expenses associated with operating the technical infrastructure. Moreover, peak demands and user consumption patterns are uncertain. Therefore, solar energy forecasting addresses this issue by developing models to predict the energy demand, production and irradiation from the Sun. Forecasting systems can help regulate solar energy generating systems and regulate the dispatching of the energy created.

In this paper, we will particularly focus on Photovoltaic solar systems and the comparative analysis of methodologies used for forecasting solar irradiation. Below listed are the considered parameters for data collection and model training.

1.1 Parameters

MODELS AVAILABLE



PROPOSED WORK

Since solar irradiation depends heavily on weather and meteorological data, we got our data from Indian Meteorological Dept. of Bhopal. Under the mentorship of Rajeev Gupta Sir and Ved Prakash Sir (data scientist at IMD Bhopal), we learnt time-series forecasting for Solar radiation data collected through installed pyranometers.

Global Horizontal Index (GHI) tributes majorly to solar energy production. We performed a univariate forecasting on a sample data containing readings of different types of radiation for every 10 minutes in the month of January 2020.

4319 samples were used for model training and 144 samples were used for validation purposes.

A statistical model was used for training. Reasons for using a statistical model :

- It has way better explainability as compared to other Deep Learning models.
- The data passes KPSS test as well as ADFuller test of stationarity with proves that statistical model would outperform Deep Learning models (LSTM, RNN, Facebook Prophet, …)

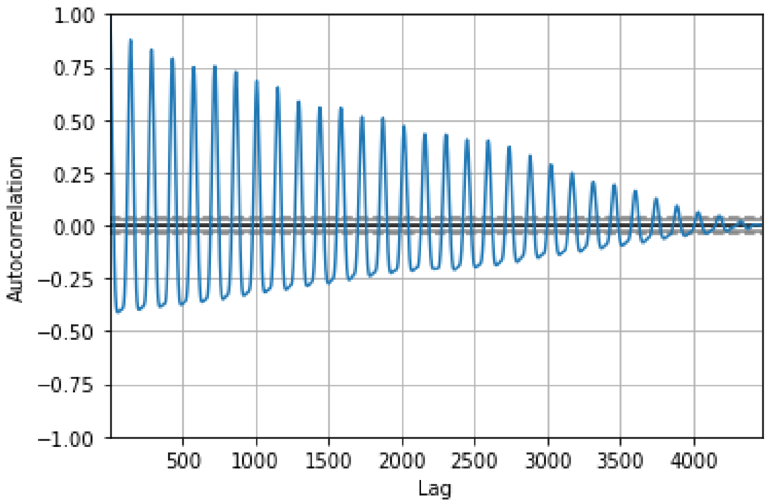
The statistical model used was Exponential moving average with exponential smoothing as well as gaussian kernels/filters.

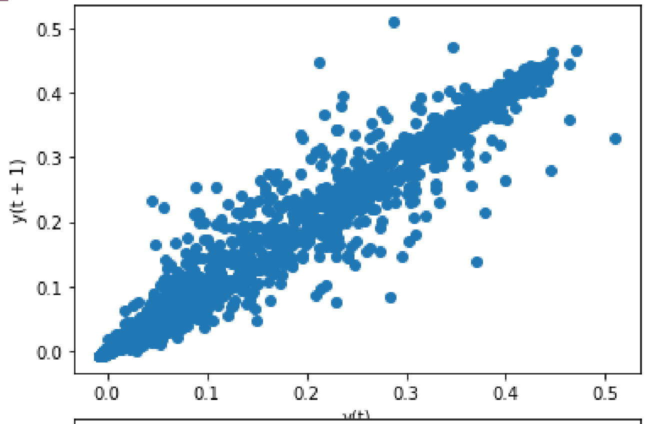
	ch	Global	diffuse	direct	terrestrial	net	reflected	uv_a	uv_b
date_time									
2020-01-01 00:10:00	Ch	-0.002	-0.005	0.035	12.563	33.432	73.741	0.0	0.524
2020-01-01 00:20:01	Ch	-0.002	-0.005	0.032	12.563	33.951	73.478	0.0	0.523
2020-01-01 00:30:00	Ch	-0.001	-0.005	0.032	12.563	33.156	73.405	0.0	0.523
2020-01-01 00:40:01	Ch	-0.002	-0.005	0.034	12.563	32.925	74.083	0.0	0.523
2020-01-01 00:50:01	Ch	-0.002	-0.005	0.035	12.563	32.704	75.122	0.0	0.522
...	...	...	...	...	...	...	...	...	...
2020-01-31 23:10:01	Ch	-0.003	-0.019	0.042	12.563	32.363	70.757	0.0	0.535
2020-01-31 23:20:00	Ch	-0.002	-0.019	0.041	12.563	31.508	70.887	0.0	0.535
2020-01-31 23:30:00	Ch	-0.002	-0.016	0.041	12.563	31.527	72.177	0.0	0.534
2020-01-31 23:40:01	Ch	-0.003	-0.018	0.042	12.563	31.659	71.082	0.0	0.535
2020-01-31 23:50:00	Ch	-0.003	-0.019	0.042	12.563	31.723	71.282	0.0	0.535

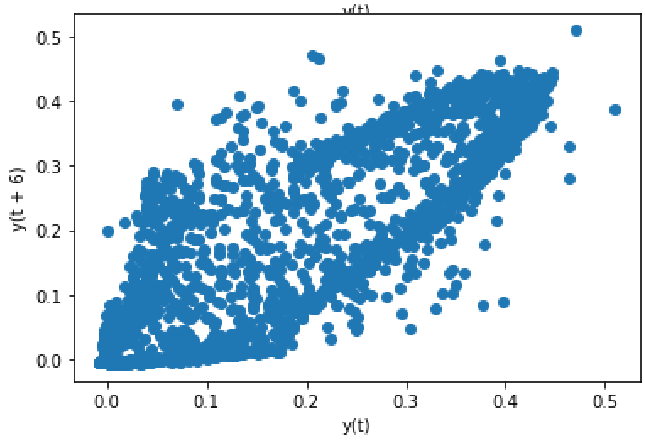
4463 rows × 9 columns

IMPLEMENTATION DETAILS

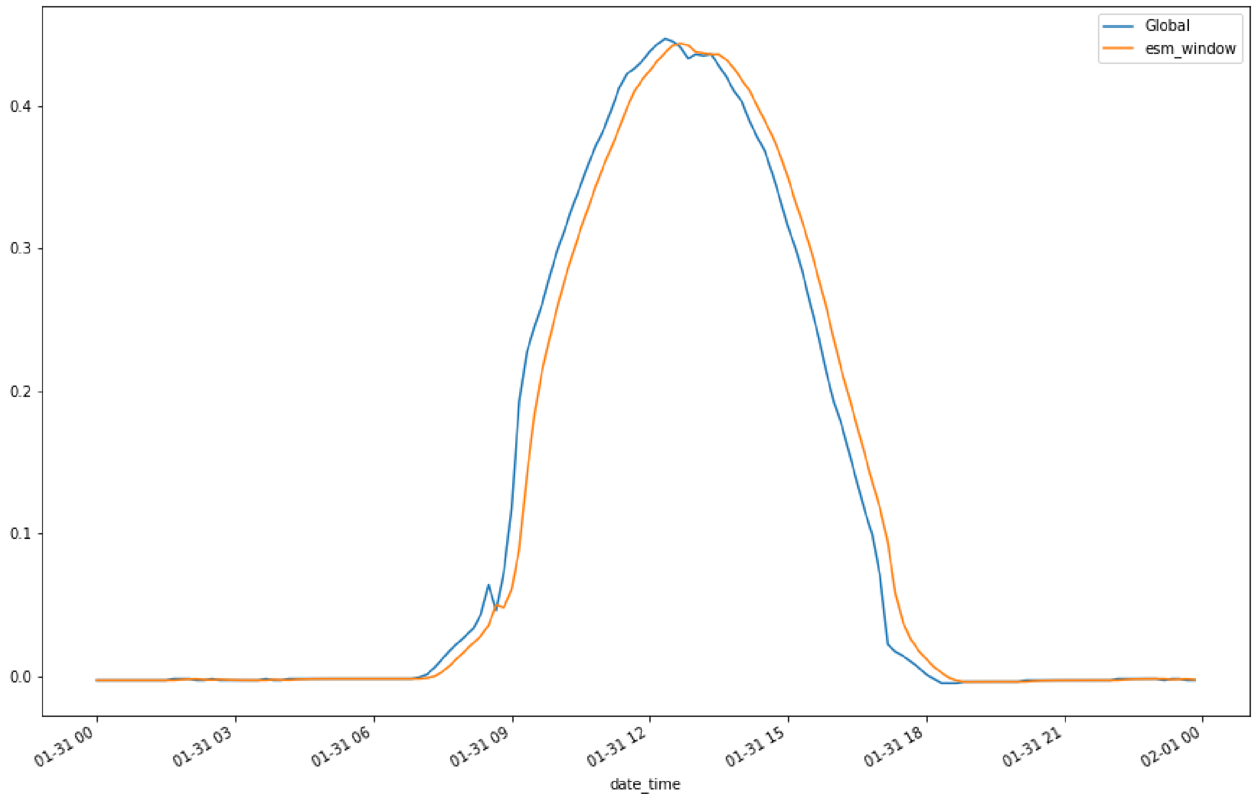
The global solar insolation data was Fat-tailed right-skewed and showed a strong collinearity in the 1st order lag plot. This proved that the time series had some level of seasonality in it. The 6th order lag plot also showcased mild collinearity. This proved that the time series data can be modelled with a good confidence interval.







After the data passed stationarity tests (KPSS and ADFuller), an Exponential Moving Average model with exponential smoothing parameter 0.2, and a window size of 2 was trained on the first 30 days data of the month of January 2020.



RESULTS

The exponential moving average model (trained on time series data of first 30 days of January) was then validated on the validation data set containing records of 10 min interval data for 31st January 2020 (144 points). The Root Mean Square Error (RMSE) and the R – Squared score was observed as follows :

```
In [44]: ((prediction_df['Global'] - prediction_df['esm_window'])**2).mean()**0.5
Out[44]: 0.02263473208681068
```

```
In [45]: from sklearn.metrics import r2_score
         print(f"R squared score : {r2_score(prediction_df['Global'], predict

         R squared score : 0.9811782383280993 | Best possible score : 1.0
```

LIMITATIONS, SUMMARY, AND FUTURE DIRECTIONS

The proposed model has been trained on a sample time-series data of the month of January 2020. It significantly captures both, the trend as well as the seasonality components of the data. It can be used to forecast solar irradiation in the Spring and Winter weathers. But this model can perform very badly on the trends and seasonality of the remaining part of the year.

As we explore data of the remaining months, we might have to switch from using a statistical model to a Deep Learning model and add multiple types of regularizations to it.