e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
Volume:03/Issue:06/June-2021 Impact Factor- 5.354 www.irjmets.com

A NOVEL APPROACH TO DEVELOP ENERGY PREDICTION MODEL FOR LARGE INDUSTRIES USING ARTIFICIAL INTELLIGENCE

Kshitij Kotasthane[*1], Medha Kasture[*2], Manas Satpute[*3], Dr. Santosh Vashney[*4]

[*1,2,3]Student, Department Of Computer Science Engineering, Acropolis Institute Of Technology And Research, Indore, Madhya Pradesh, India.

[*4]Associate Professor, Department Of Computer Science Engineering, Acropolis Institute Of Technology And Research, Indore, Madhya Pradesh, India.

ABSTRACT

In recent years clean energy has gained massive popularity and the leading industry is the solar energy industry. The energy generation estimation is essential to understand and manage the output variability ,and is of interest for several factors in the energy market. In the short term, a transmission system operator is interested in the energy output from PV panels to find the adequate balance for the whole grid, since over and under producing electricity often results in penalty fees. On the other side of the spectrum, electricity traders are interested in long time horizons, ordinarily, day ahead forecasts since most electricity is traded on the day ahead market. We attempt to face this challenge with the help of Machine learning algorithms. The higher availability of data and increased computation power have enabled machine learning algorithms to perform improved production. Solar energy is highly dependent on the weather and because of this the energy output of the PV panels is unstable , so implementing a time series model can help us approach a solution in an efficient manner. We make use of various time series models such as ARIMA and Prophet , and check which one suits our data better and help us in generating meaningful results.

Keywords: Renewable Energy, Photovoltaic Panel, Energy Spectrum, Forecasting, Machine Learning, Prediction.

I. INTRODUCTION

Almost 80% of the world's population has access to electricity. This enormous need is met mostly with fossil fuels, which are now causing global concerns like climate change. But with the changing time we are observing a paradigm shift in the energy industry. Diversifying energy supply and reducing dependence on carbon based fuel is the top priority. Solar plants are helping us meet our top priority.

Solar energy has seen astounding growth in recent years, the cost of Photovoltaic panels has reduced by 70%, which has made it a cost efficient alternative to carbon based fuels. This growth has made solar energy a competitive industry. Countries are increasingly becoming dependent on solar energy. This dependence and competition has increased the need for suitable means for forecasting solar energy output.

In the energy industry the need for electricity fluctuates according to the situation. Forecasting is important because if solar energy is to replace conventional energy, it needs to match the requirement. It also helps large industries by giving them an estimation before time and making plans accordingly. Forecasting solar energy output is a difficult task because it is linked with the weather forecasting challenge. But we can approach this problem with the help of Machine learning techniques. The increased computational capacity and the abundant quality data have made the techniques useful for forecasting.

II. PROBLEM FORMULATION AND REVIEW OF LITERATURE

The project is based on analyzing the dataset for a given period of a solar power plant and then predicting the power generation in the coming days. A Machine Learning model is made in such a way that it predicts the power generated by a single solar panel. Along with counting the prediction , it is also able to detect a faulty panel or if it needs maintenance.

Background of the problem

Solar energy is an inexhaustible and renewable form of energy which is boon to the ever increasing energy requirements of today's world. This energy can be harnessed through Solar Power Plant and then can be used for various domestic or industrial needs. We know that fossil fuels are limited and the power extraction process has a high carbon footprint, so efficient use of solar energy is the need of time. In the solar power industry, prediction of the production of solar energy for short periods (day, several days, week) does not have the well-established and tested technology and is often associated with large errors, which can be 60-65%. We intend to deal with this problem in this paper.

Literature Review

During our research we came across many general articles related to energy forecasting. Significant findings for time series and machine learning techniques for solar forecasting were also examined. The ones important to our research are listed below:

1. ARIMA (Auto Regressive Integrated Moving Average)

ARIMA is a model which is used for predicting future trends on a time series data. It is a model that forms a regression analysis.

a. **AR (Autoregressive):** Model that shows a changing variable that regresses on its own lagged/prior values.
b. **I (Integrated):** Differencing of raw observations to allow for the time series to become stationary.
c. **MA (Moving average):** Dependency between an observation and a residual error from a moving average model. (Dabakoglu 2019)

2. Prophet

Prophet is a procedure for forecasting time series data based on an additive model where non-linear trends are fit with yearly, weekly, and daily seasonality, plus holiday effects. It works best with time series that have strong seasonal effects and several seasons of historical data. Prophet is robust to missing data and shifts in the trend, and typically handles outliers well. (Taylor and Letham 2017)

III. METHODOLOGY

For the problem of predicting the future energy generation, the system is designed in such a way so as to mimic the pattern previously generated by the plant.

Working

When a photon hits the surface of the photovoltaic cell, its energy is transferred to the electrons present on the silicon cell. These electrons are "excited" and begin to flow in the circuit producing electric current. A solar panel produces Direct Current energy(DC power).

Then, it's up to the inverter to convert it into alternating current to transport it and use it in our distribution networks. In fact, domestic and industrial buildings are designed for the transport and use of alternating current.

Every solar power plant consists at least of two basic components:

1. Modules - that convert sunlight into electricity.
2. One or more inverters- devices that convert direct current to alternating current.

Data Visualization of daily yield and AC power

The daily yield and AC power generated are two of the main attributes of our model. Their graphical representation gives us an insight to the general pattern made by them and will help in making decisions for the model. It also helps us to recognize the pattern- when most of the power is generated in a day.

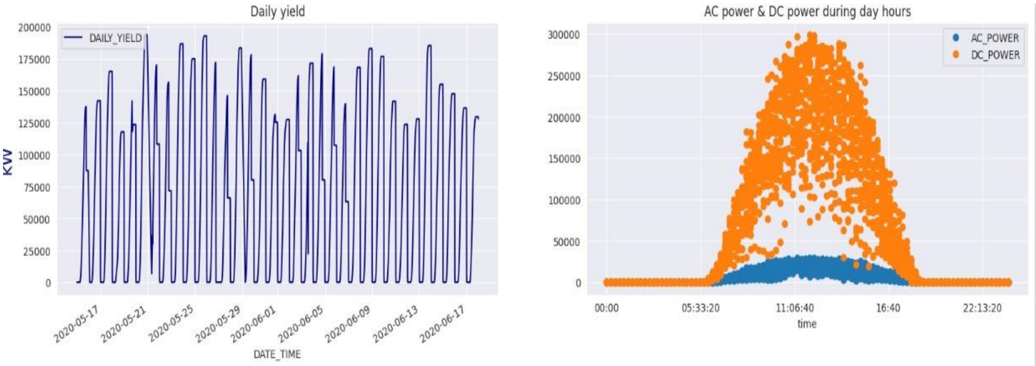

Figure 1.1 (a) Shows the power generation plotted against each day for 10 days. (b) The AC and DC power generated during a day.

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
Volume:03/Issue:06/June-2021 Impact Factor- 5.354 www.irjmets.com

Data visualization of daily yield and AC power for 34 days

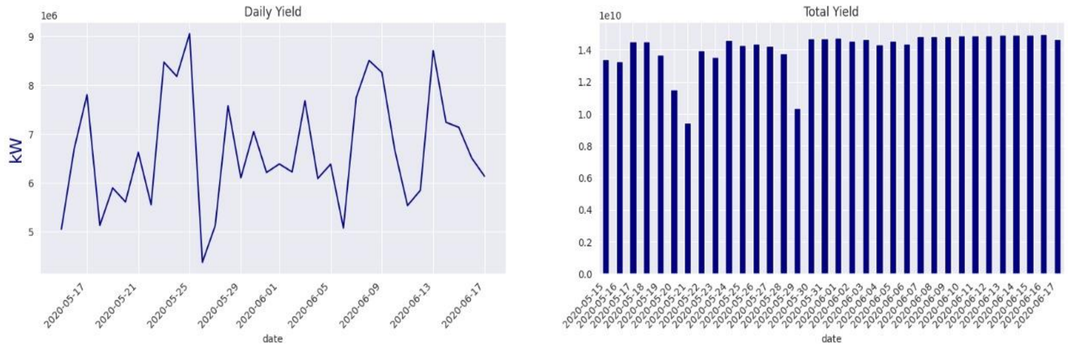

Figure 1.2 The plots for the (a) Daily yield and (b) Total yield

To identify the sub-optimally performing equipment

We know how all inverters work during day hours, because we know from the quick exploratory data analysis that something went wrong with dc power generation. We can try to identify inefficient equipment by plotting how the inverters work during the day and then compare their efficiencies. We will only focus on the DC POWER which is converted to AC POWER for domestic use.

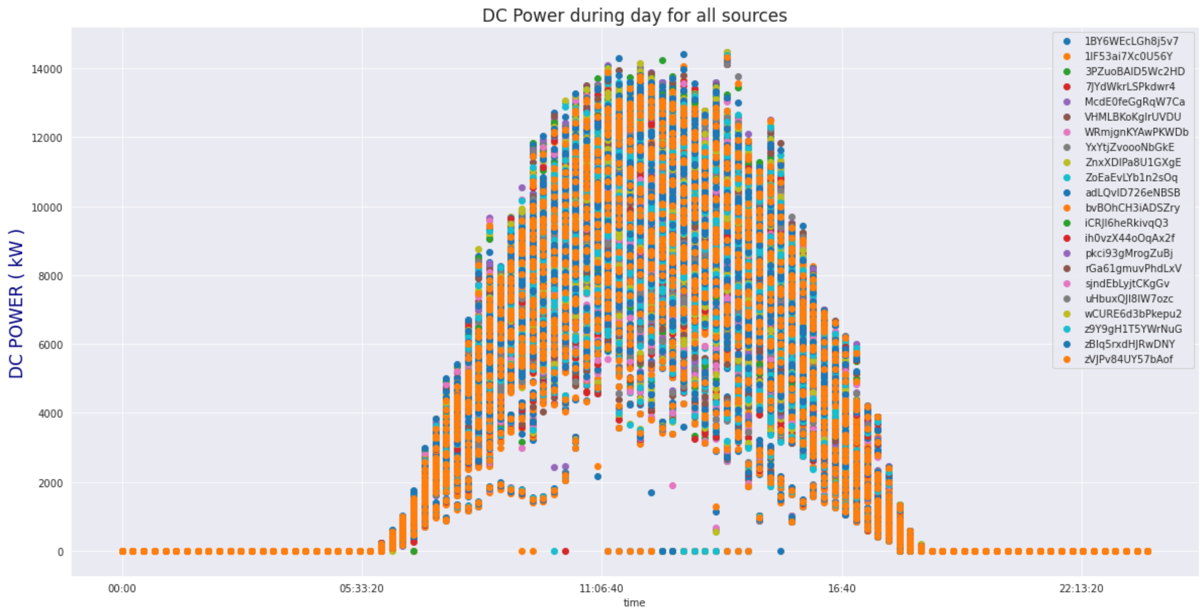

Figure 1.3 The DC power generated by all sources in the plant

Here we are interested in the solar power that is generated by the inverters which get input from the solar module. Although this plot shows what might be wrong with the equipment, there are still two things that we can conclude from this:

1. Not all of the energy that should be converted from DC POWER from solar modules to AC POWER is actually converted and there is a loss of energy in between.
2. Inverters (from solar power to electrical energy) are inefficient and these are the suboptimally performing equipment we are looking for.
3. Only about 9-10% of the overall power is converted. Still this does not mean that all inverters are faulty and contribute to losses.

We compare the DC power generated by different panels and we can see clearly what is going wrong with the power plant.

From Fig. 4.4 it is clear that inverters **1BY6WEcLGh8j5v7** & **bvBOhCH3iADSZry** are underperforming compared to other inverters, maybe these inverters require maintenance or require to be replaced. But before going into depth about underperforming inverters, let's look at which are the common problems for the entire plant, so let's see DC power generation during day hours for all 34 days.

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
Volume:03/Issue:06/June-2021 Impact Factor- 5.354 www.irjmets.com

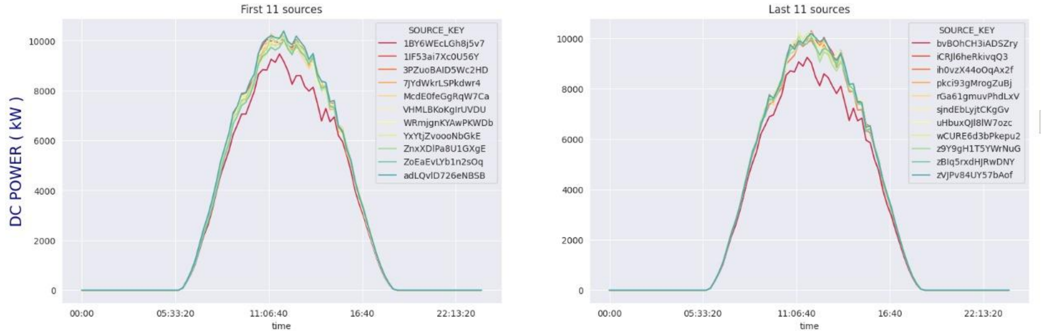

Figure 1.4 The performances of the 22 energy sources used in the power plant. The curves in red highlight the sub-optimally performing equipment that our model identifies.

Forecasting

We want to predict the yield generated by plant_1 for the next two days.

1. ARIMA model

The first model that we implemented was the ARIMA model. It's a model used in statistics and econometrics to measure events that happen over a period of time. The model is used to understand past data or predict future data in a series. It's used when a metric is recorded in regular intervals, from fractions of a second to daily, weekly or monthly periods.

This model fits all the parameters for our project as our project is also based on making predictions from analyzing past data.

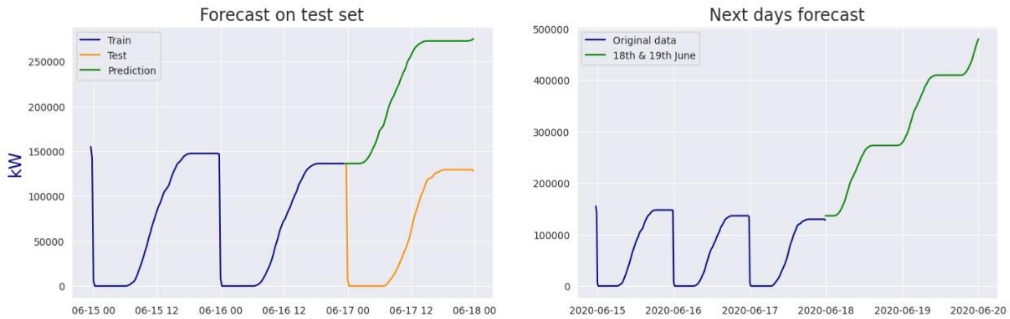

Figure 1.5 Forecast of power generation obtained by ARIMA(SARIMAX) model on the (a)Test set and (b) Next day forecast.

As is apparent in Fig. 1.5, even though the pattern is similar, it would not be able to predict the generated energy.

2. Prophet

The next model that we implemented was the Prophet model. Prophet is a procedure for forecasting time series data based on an additive model where non-linear trends are fit with yearly, weekly, and daily seasonality, plus holiday effects. It works best with time series that have strong seasonal effects and several seasons of historical data. Fig 4.6 shows the predictions made by the **Prophet model.**

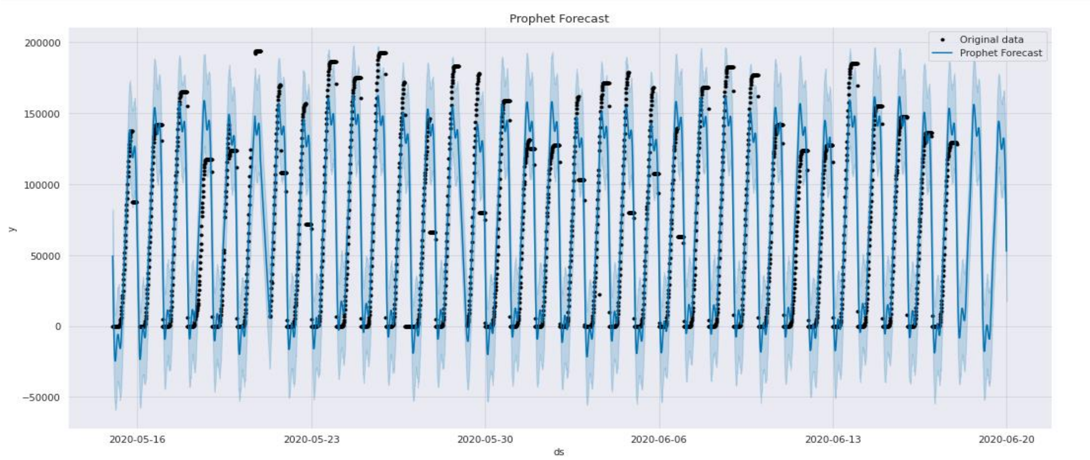

Fig 1.6 The prediction on the input data by **Prophet** (shown by the continuous curve in blue) and the original data (shown by the discrete black points). The prediction resembles the actual data closely. fects and several seasons of historical data. Prophet is robust to missing data and shifts in the trend, and typically handles outliers well.

IV. METHODOLOGY

In this work, we have created an application which takes the solar weather data and uses it to forecast solar energy for days in the future. Also it is able to identify sub-optimally performing solar panels which need cleaning or maintenance. In the section using data analysis techniques we were able to identify the equipment. Below is a comparison of the **SARIMAX** against the **Prophet** model based upon three metrics, namely their,

1. R2 score, the proportion of the variance.
2. MAE score, the average magnitude of the error
3. RMSE (root mean squared error)

```
SARIMAX R2 Score: -7.340983
Prophet R2 Score: 0.894944
---------------
SARIMAX MAE Score: 157388.955212
Prophet MAE Score: 13487.890100
---------------
SARIMAX RMSE Score: 160753.728442
Prophet RMSE Score: 18486.021026
```

Figure 1.7 Comparison of the performance of SARIMAX and Prophet models

outliers well.

V. CONCLUSION

In this work, we find that employing time series models is a complex procedure due to the non-stationary energy time series. In contrast, machine learning techniques were more straightforward to implement. We find that with a *higher R2 score (variation proportion)* and *lower MAE and RMSE (error measures)* the ***Prophet model performs much better than the SARIMAX*** to predict solar energy generation rates.

ACKNOWLEDGEMENT

We as the authors would like to extend a special thanks of vote to the reviewers of this paper for their valuable suggestions to improve this paper. The paper is supported by Acropolis Institute of Technology and Research, Indore (M.P.).

VI. REFERENCES

[1] Dabakoglu, Caner. 2019. "Time Series Forecasting — ARIMA, LSTM, Prophet with Python." medium.com. https://medium.com/@cdabakoglu/time-series-forecasting-arima-lstm-prophet-with-python-e73a750a9887.

[2] Taylor, Sean J., and Benjamin Letham. 2017. "Prophet | Forecasting at scale." Facebook. https://facebook.github.io/prophet/ .

[3] KTH RITSES, Solar Power Forecasting with Machine Learning Techniques in 2018, by EMIL ISAKSSON and MIKAEL KARPE CONDE.

[4] IRENA. Renewable power generation costs in 2017. Technical report, International Renewable Energy Agency, Abu Dhabi, January 2018

[5] Jose R. Andrade and Ricardo J. Bessa. Improving renewable energy forecasting with a grid of numerical weather predictions. IEEE Transactions on Sustainable Energy, 8(4):1571–1580, October 2017.

[6] Rich H. Inman, Hugo T.C. Pedro, and Carlos F.M. Coimbra. Solar forecasting methods for renewable energy integration. Progress in Energy and Combustion Science, 39(6):535 – 576, 2013.

[7] J. Antonanzas, N. Osorio, R. Escobar, R. Urraca, F.J. Martinez-de Pison, and F. Antonanzas-Torres. Review of photovoltaic power forecasting. Solar Energy, 136:78–111, October 2016.