Solar Power Generation

Jainil Patel, Harsh Gandhi

Abstract

The world's energy consumption is outpacing supply due to population growth and technological advancements.
For future energy demands, it is critical to progress toward a dependable, cost-effective, and sustainable renewable energy source.
Solar energy, along with all other alternative energy sources, is a potential renewable resource to manage these enduring challenges in the energy crisis.
Solar power generation is expanding globally as a result of growing energy demands and depleting fossil fuel reserves, which are presently the primary sources of power generation.
In the realm of solar power generation, photovoltaic (PV) panels are used to convert solar radiation into energy.

They are subjected to the constantly changing state of the environment, resulting in a wide range of defects.
In this report we focus on certain key things:

- *Solar power prediction.*
- *Using Machine learning model for timely fault detection.*
- *Combining weather and sensor based data to draw anomaly detection insights.*
- *Consolidating various Exploratory Data Analysis Plots for better understanding of the data.*
- *Using NULL values to identify missing values in daily yield production.*

Keywords

- Solar power prediction
- Weather
- Date-Time analysis
- Sensors
- Inverters
- Machine Learning.
- Exploratory Data Analysis

Solar Power Generation

Introduction

This analysis attempts to predict the power generation for the solar plants and identify faulty and suboptimally performing equipment. Power generation is indicated by daily yield. Based on the performance of the inverter(suboptimal or faulty), this may indicate that one or more panels connected to that inverter requires maintenance.

In this analysis, the dataset is first checked for completeness. Secondly, the data for the plants are explored to indentify time-sereis trends and outliers or any other data anomalies. This analysis is restricted to plant 1 since there may be variables that affect the applicability of forecasting models to plant 2.

About Dataset

This data has been gathered at two solar power plants over a 34 day period.
The power generation datasets are gathered at the inverter level - each inverter has multiple lines of solar panels attached to it. It has two pairs of files - each pair has one power generation dataset and one sensor readings dataset. The sensor data is gathered at a plant level - single array of sensors optimally placed at the plant.
It a **Date-time** series dataset with interval of **15 minutes.**
Areas of concern:

- Grid power management and power prediction.
- Panel Maintenance.
- Faults and Defects Detection.

Solar Power Generation

	DC_POWER	AC_POWER	DAILY_YIELD	TOTAL_YIELD
count	3158	3158	3158	3158
mean	68547	6703	71782	151989182
std	88044	8603	65974	10616697
min	0	0	0	26540040
25%	0	0	90	152097608
50%	8515	823	66068	153531986
75%	140386	13750	129398	154994999
max	298937	29150	193770	156142755

Fig 1.1 : Plant 1 Power Dataset

	AMBIENT_TEMPERATURE	MODULE_TEMPERATURE	IRRADIATION
count	3182.000000	3182.000000	3182.000000
mean	25.531606	31.091015	0.228313
std	3.354856	12.261222	0.300836
min	20.398505	18.140415	0.000000
25%	22.705182	21.090553	0.000000
50%	24.613814	24.618060	0.024653
75%	27.920532	41.307840	0.449588
max	35.252486	65.545714	1.221652

Fig 1.2 : Plant 1 Weather Dataset

Insights & Discussions

Plant contains 22 inverters where each inverter are connected with several PV array. Every 15 min, each inverter records his data. So, if we want to know how many the plant has produced a power in a hour, we just compute the contribution of 22 inverters.

```
plant1_data.info() # we check if there exist missing value
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 68778 entries, 0 to 68777
Data columns (total 7 columns):
 #   Column       Non-Null Count  Dtype  
---  ------       --------------  -----  
 0   DATE_TIME    68778 non-null  object 
 1   PLANT_ID     68778 non-null  int64  
 2   SOURCE_KEY   68778 non-null  object 
 3   DC_POWER     68778 non-null  float64
 4   AC_POWER     68778 non-null  float64
 5   DAILY_YIELD  68778 non-null  float64
 6   TOTAL_YIELD  68778 non-null  float64
dtypes: float64(4), int64(1), object(2)
memory usage: 3.7+ MB
```

Fig 2.1 : Plant 1 Data Info

No Null Values in Data Power dataset.

Solar Power Generation

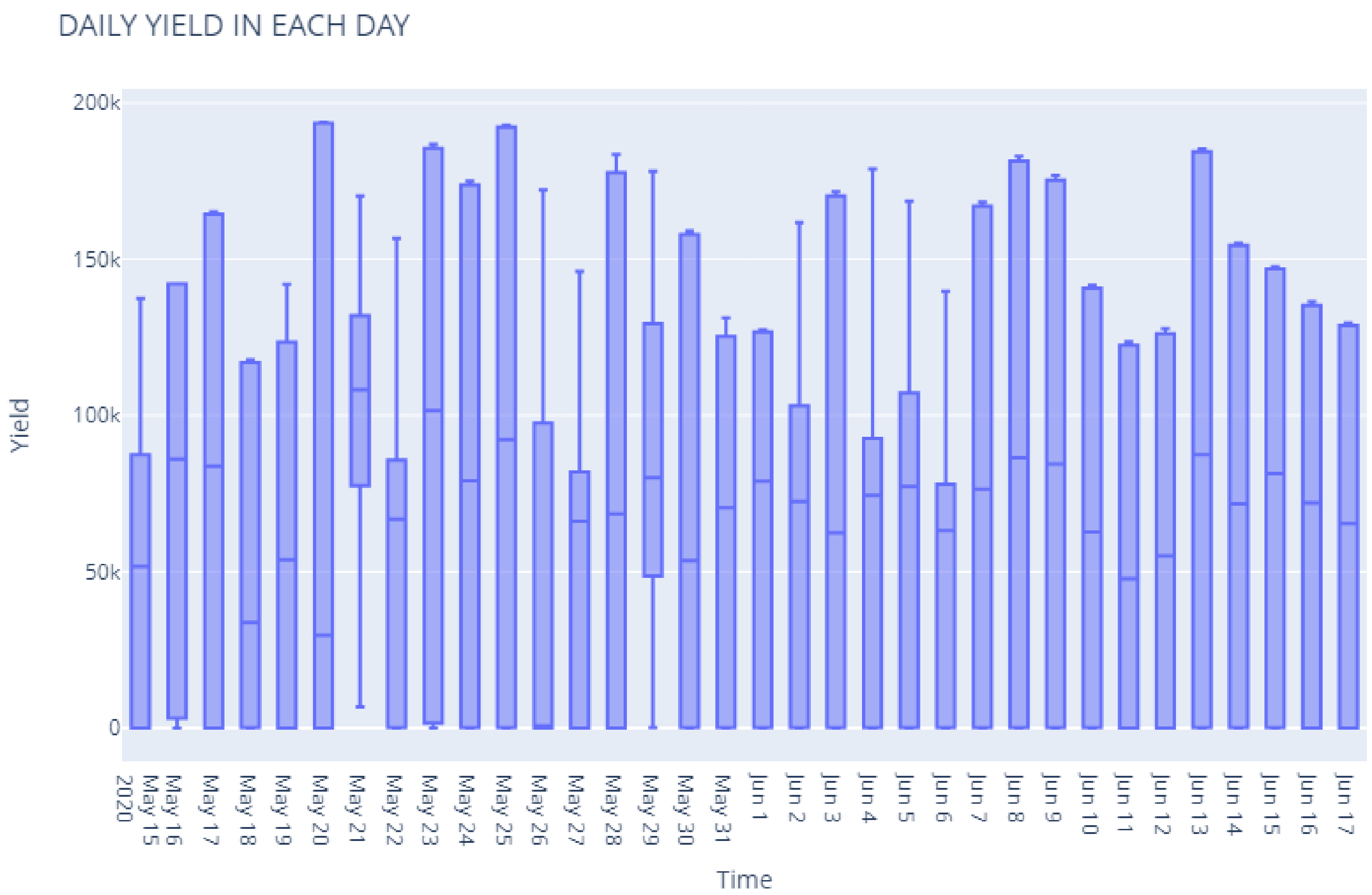


Fig 2.2 : Yield Vs Time Box Plot

- Each day, the daily yield shows a considerable difference.
- On some days the yield is high.
- From the box plots it is observed that on most of the days the PV Panels could convert majority of the DC Power to AC Power contributing to higher yields.
- Outliers do exist on certain days and necessary analysis have been carried out for anomalies and fault detection later.

Solar Power Generation

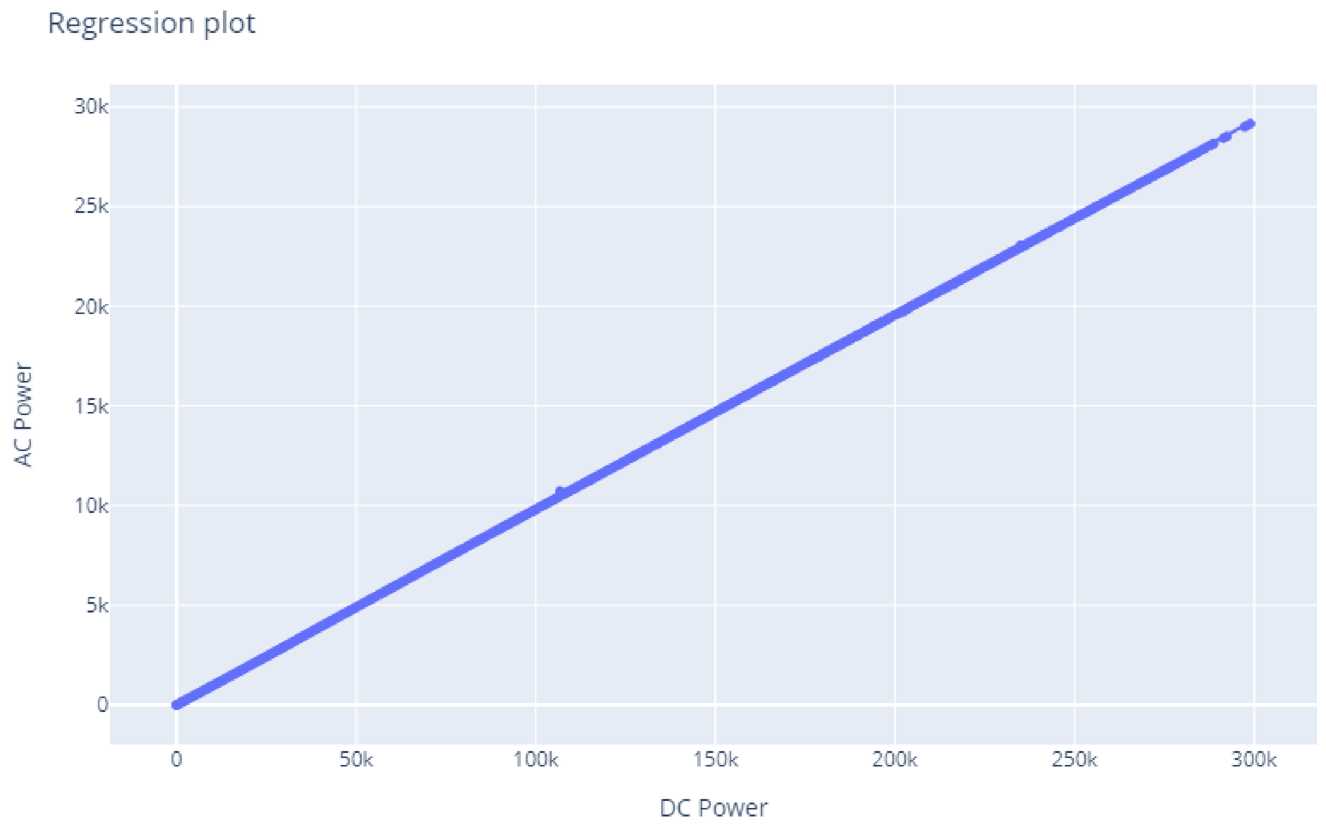
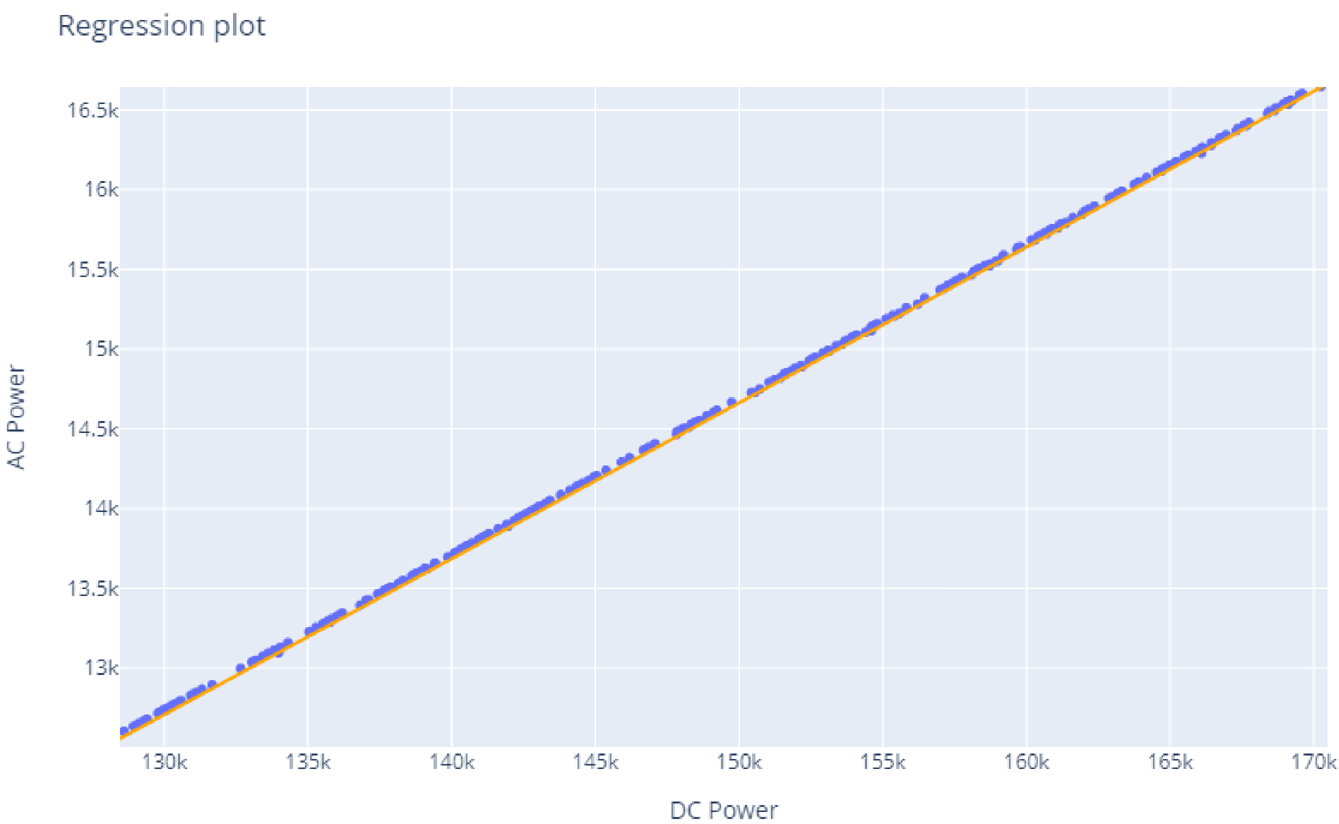


Fig 2.3 : AC Power Vs DC Power Regression Plot

This graph said that inverter convert dc power to ac power linearly.
DC Power = 10*AC Power
inverter lost 90% of their power when it convert.



Solar Power Generation

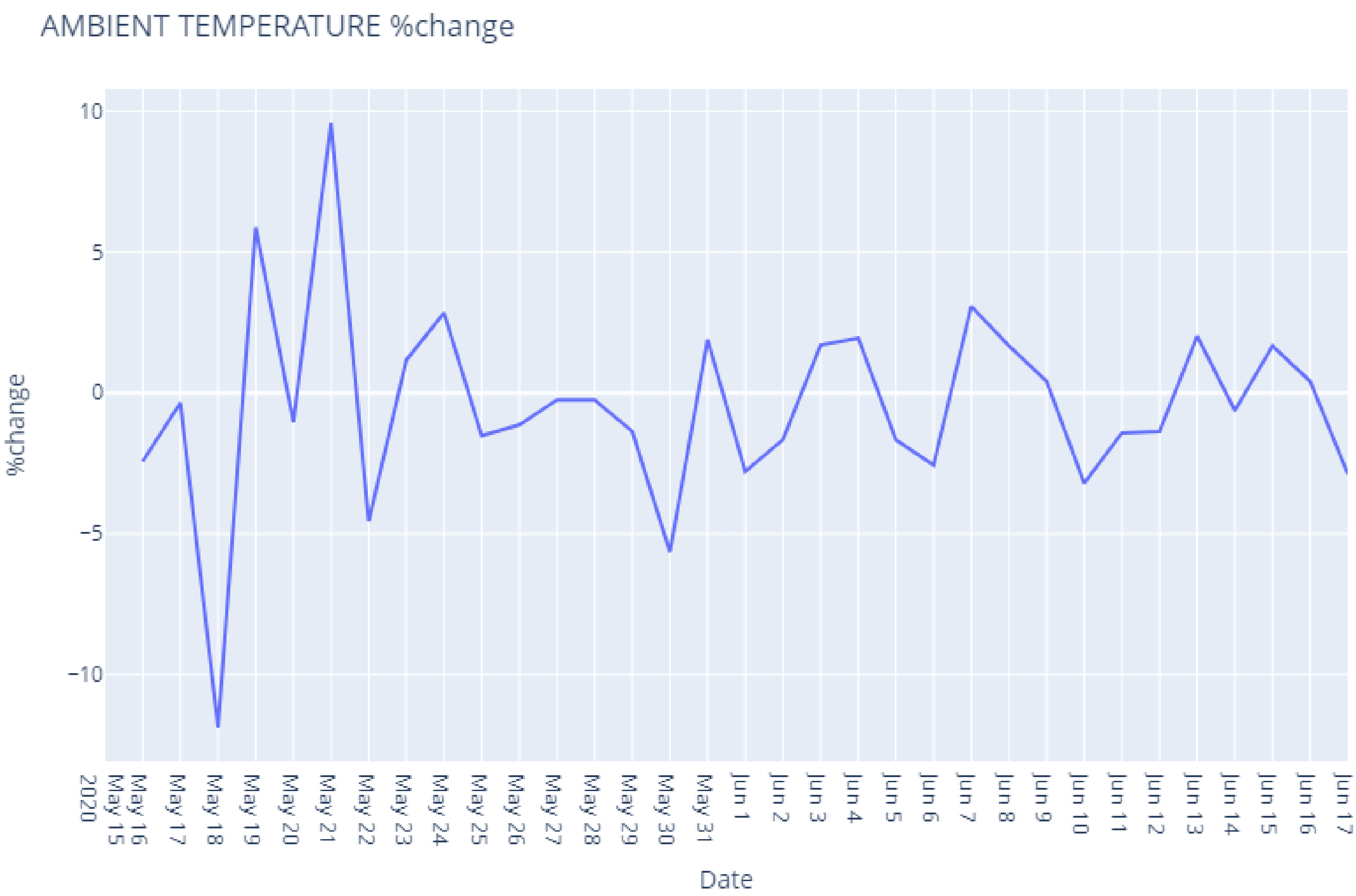


Fig 2.4 : Ambient Temperature %change Vs Date Line Graph

- Sunday 17 May 2020 to Monday 18 May 2020, the ambient Temperature decreases to 10%.
- Monday 18 May 2020 to Tuesday 19 May 2020, the ambient Temperature increases to 15% and tomorrow decreases to 5%.
- Wednesday 20 May 2020 to Thursday 21 May 2020, the ambient Temperature increases to 10% and tomorrow decreases to 15%.
- June month's, the Ambient Temperature %change stabilize between -2.5 and 2.5%.

Solar Power Generation

Anomaly Detection

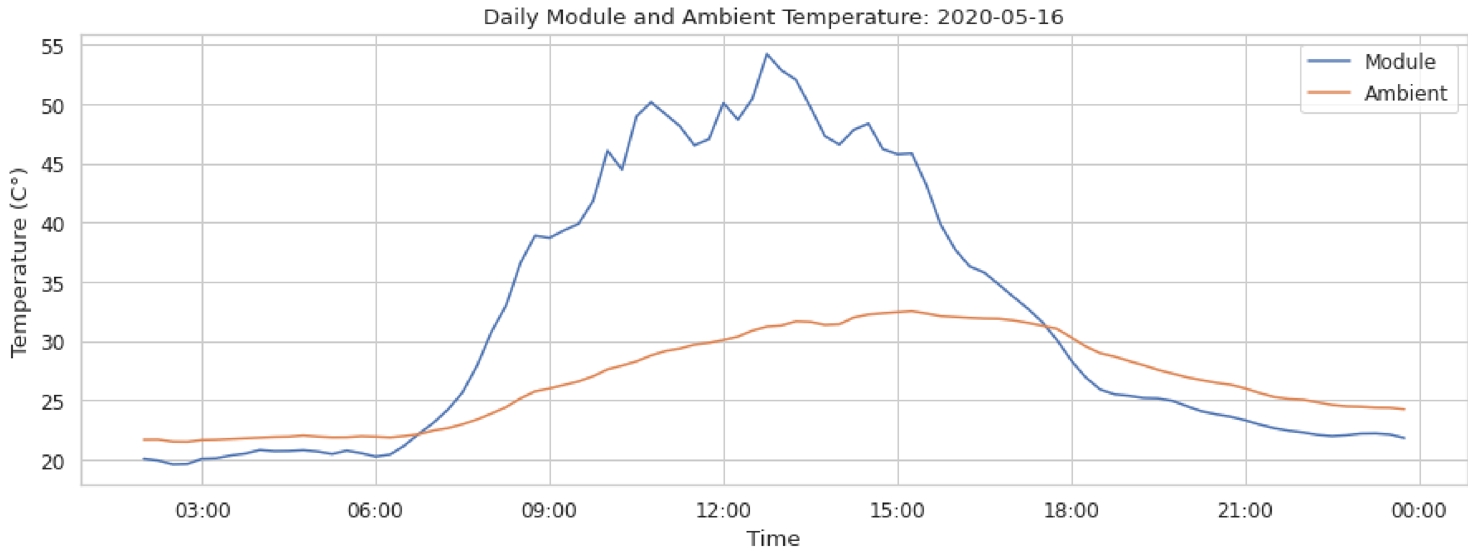


Fig 3.1 : Ambient Temperature %change Vs Date Line Graph

- Ambient temperatures range from 20 to 35°C, modules reach temperatures from 18 to 65 °C.
- Modules reach significantly higher temperatures than their ambient air during daytime.
- Ambient temperature is lagging behind daily module cooldown.
- This means the modules cool down quicker than their environment.

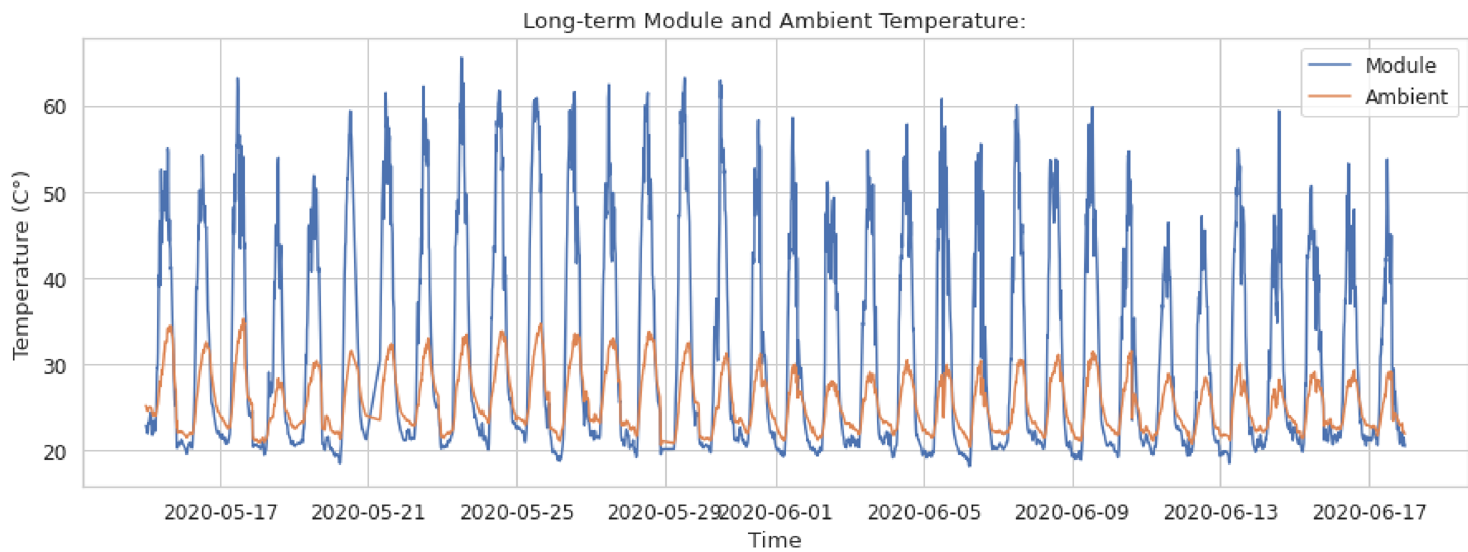


Fig 3.2 : Module Temperature Vs Time

There are two days with significantly lower temperature ("bad weather") on. Such events may be difficult to forecast without access to more weather data (air pressure, wind, humidity, cloud formation etc.) and advanced weather forecasting models.

Solar Power Generation

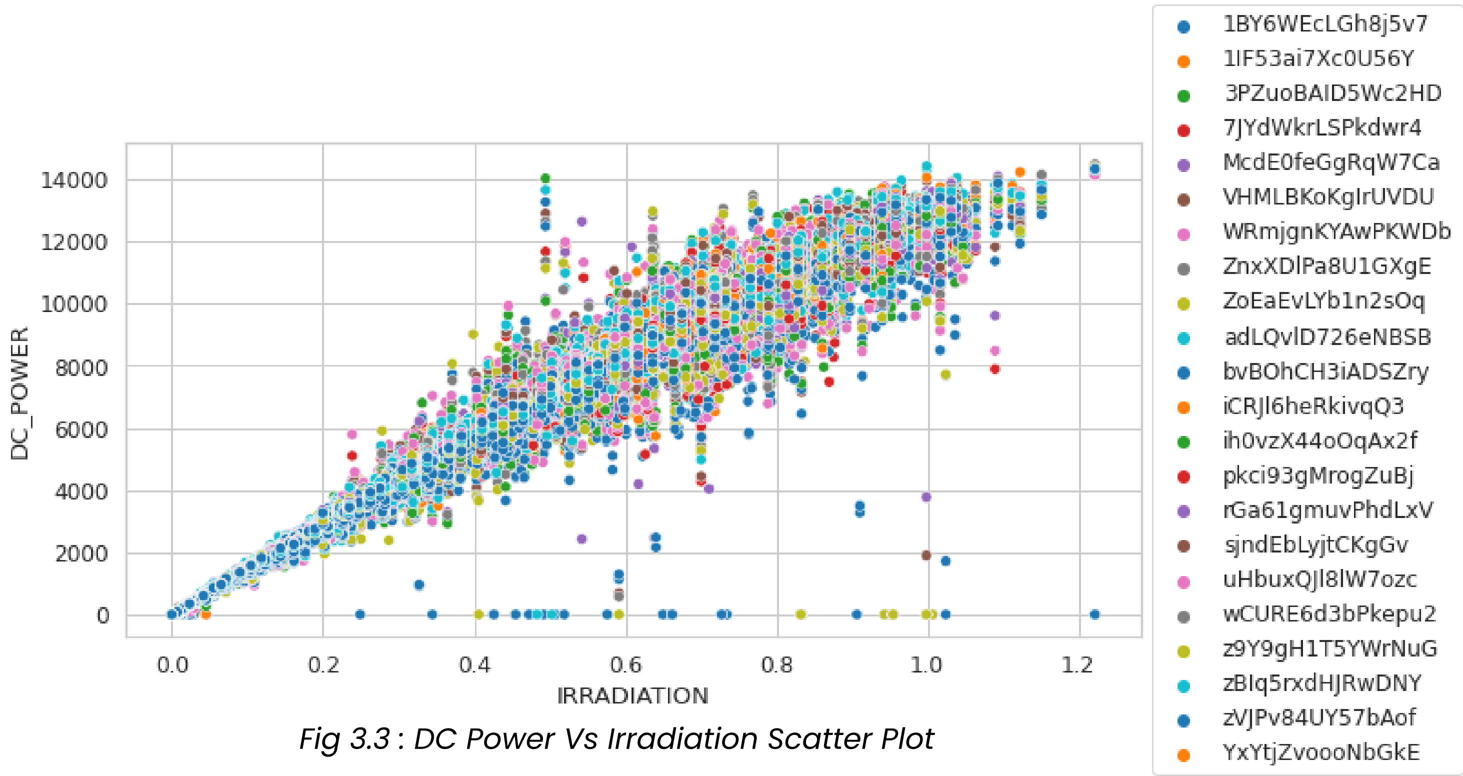
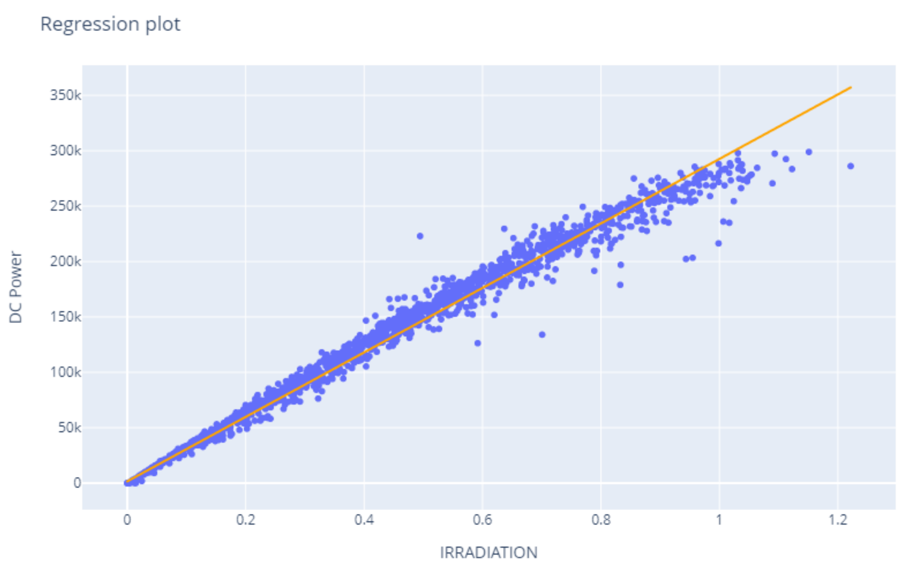


Fig 3.3 : DC Power Vs Irradiation Scatter Plot

Our data clearly shows events where some inverters received no DC power even though there was more than enough sunlight to generate power. We clearly have some equipment malfunction in our data.

To illustrate this, we can take a closer look at the daily distribution of the generated power and the measured irradiation.



Plotting the aggregate DC Power of inverters against irradiation.
This indicates that the DC Power increases with Irradiation.

Fig 3.4 : DC Power Vs Irradiation Line Regression Plot

Conclusive Summary

- Several articles have been gathered and reviewed concerning the recent advancement and research on five aspects:
- Various possible faults that occur in PV panel.
- Exploratory Analysis of the data and accordingly plotting the necessary data for deeper insights of the PV Panels.
- Online/remote supervision of PV panels.
- Weather and sensor data analysis in PV Panels.
- Exploring different data science techniques for various clustering and plotting tools to be applied on the PV panels.
- Role of machine learning techniques in fault diagnosis of PV panels.
- Various sensors used for different fault detections in PV panels and
- Benefits of fault identification in PV panels. From the fault classification point of view, various possible causes of faults, such as partial shading fault, short circuit fault, and open circuit fault, faults in diodes—blocking and bypass diodes—were discussed in detail.
- The various online techniques that are meant to monitor the errors occurring in PV panels based on the type of sensor used and the monitoring of the PV panels were discussed in this paper.
- Finally, this paper included the future and possible scope of the optimization of fault detection techniques. Through that, the cost and time incurred in the fault diagnosis of solar power panels can be reduced.

Future Scope

- In solar panels, equipment-grounding conductors are one of the crucial elements that would start as grounding wires on the array and continue with the power wires through the rest of the system. WEEB is one of the alternatives for conventional grounding. However, in the washer–electrical equipment bond, diagnosing the grounding fault is highly challenging.

- If one needs to measure fault current readings and transient state readings, it might not be possible to record the temporary state of current since the instruments have a slow sampling rate. The IoT can update the data only with a sampling rate of 30 s, which is far from the sampling rate meant to capture the fault current event. When one uses a digital storage oscilloscope, it becomes possible to record the fault event in real time.

- The IoT can be used to update the status of the panel and data rate less than the 30 s sampling rate. If a signal varies much faster than this rate, it may not be updated in the cloud since it depends on the network's speed and web updating speed. A single-shot external signal out is required for an analog oscilloscope, as well as a particular camera to record the image.