Photovoltaic Plant Yield Prediction Using Deep Learning Networks

GAURANG GUPTA[1], SANYAM AHUJA[2], VRINDA GOEL[3], DESHENDRA SIHAG[4]
[1, 2, 3, 4] *Netaji Subhas University of Technology*

Abstract- One of the most significant sources of renewable energy is solar energy. To maintain the dependability and effectiveness of solar energy conversion for PV systems, defects that arise during operation must be identified and addressed quickly. This study presents a comprehensive analysis of the predictive maintenance technology used for solar photovoltaic (PV) modules. The primary objective is to implement an algorithm based on time-series data that can identify and predict any potential faults related to the power output of PV cells. Using weather predictions as the base data, we perform a comparative analysis with RNN frameworks like Long Short Term Memory (LSTM) and Bidirectional LSTM models and Gated Recurrent Units (GRUs). The performance of the models was evaluated by comparing predicted values with actual values obtained from two test sites located in India. Our results show that GRUs outperform LSTM and Bidirectional LSTM (Bi-LSTM) networks by a margin of 4.45e-01. Our proposed framework gives a mean squared error of the order e-05 on the validation data of the solar power generation dataset.

Indexed Terms- PV Module, Maintenance, LSTM, RNN, GRU

I. INTRODUCTION

As the world looks to move on from fossil fuels to cleaner sources of energy, solar energy has become a major source of renewable energy as it contributes to 68.25% of the total renewable energy generated in India[1]. Maintenance of PV Modules plays a major role in utilizing solar energy. Maintenance in the context of PV cells ranges from straightforward inspections to extremely precise monitoring that enables the owner to identify problems or the need for cleaning procedures. Preventive maintenance activities such as cleaning dust, bird droppings, soil and other types of particles from the panels, inspecting wiring connections and replacing faulty parts should also be undertaken regularly in order to ensure optimal performance from the PV system. Through regular maintenance, the PV systems' efficiency and electricity output will improve, which could have an effect on the overall revenue. Cell defects have a negative impact on the effectiveness, dependability, and durability of solar modules. The output of PV cells is environment-dependent in terms of the sun's radiation falling on the PV panels, climate, humidity, and other environmental parameters. The inherent composition of PV cells (intrinsic inadequacies) and extrinsic process-induced defects also contribute to power output fluctuation. When unsightly impediments cover the solar panels, there is a significant power loss. The barrier causes the shaded cell to heat up and use more power since it transforms into a resistor.

Anomalies can be broadly detected in three ways:

1. by examining a specific device's parameters and performance in real-time to those of the other devices in the same group, and
2. by comparing the historical and operational data of a particular device, and
3. by comparing the operational data of a specific device to its predicted values using AI/ML Algorithms[2].

In order to prevent underutilization of photovoltaic systems, it is important to detect and classify any faults that occur in the system, rectify them as soon as they happen and maintain and optimise performance. This can help reduce the risk of fire or any other mishappening[3]. The research has been further divided into sections-Related Work, Methodology, Data Acquisition, Results and Discussion, Limitations, Conclusion and Future Scope followed by the References.

II. RELATED WORK

In this section, we highlight the work done in recent past for the detection and monitoring of faults in solar panels.

2.1 Intelligent Monitoring and Maintenance of Solar Plants

To monitor the power generation of solar plants, there is a need to integrate PV power generation systems on the grid. Authors of [4] have investigated various approaches for the automatic detection of defects by employing SVM and CNN in high-resolution EL images. They determine the best- performing pipeline as KAZE/VGG features in a linear SVM. Another system identifying various malfunctioning and possible breakdowns of such devices was developed in [5]. Once the anomaly is detected, it is immediately informed and the model uses CNN to classify each solar module as functional or dysfunctional. The authors compare each solar module with A. respect to a group using 3 Standard Deviation Method B. Comparing with respect to past patterns. [6] has focused on a laser beam focused hair-thin water jet, which uses total internal reflection to guide the laser thus attaining a higher level of cleanliness as compared to conventional dry lasers. In [7],a cleanliness monitoring system which is a combination of hardware and software has been proposed. The key to the concept lies in enforcing a servicing policy at the sensor level, which requires maintaining one of the calibrated cells as clean as possible while leaving the other one to be cleaned together with the corresponding plant that the system is monitoring.

2.2 Predictive Maintenance of Solar Plants

[8] predicts the daily electrical power curve of a PV panel based on the power curves of its adjacent panels. Two variants of CNN (CC, CUC) were implemented out of which CUC algorithm gained better accuracy with the real-life datasets. Another system proposed by authors of [9] also worked with image processing to predict the percentage of dust by using just one solar panel in their research, however, their system can be expanded to multiple solar panels by just multiplying the number of panels used with the result obtained. On the other hand, in [10], authors proposed a Multi-Layer Perceptron topology wherein two designs of ANN model were put forward. A model that monitors different parameters and uses automatic learning techniques for prediction was developed by the authors of [2]. With the aim of increasing the share of renewable energy penetration, an architectural proposal based on Edge Computing with LSTM model was included to implement the proposed model into a system. [11] has used a new solar cell model which includes power degradation of ageing modules. The predicted results are validated experimentally using an automatic data monitoring system under real outdoor conditions.

While most of the research papers include CNN and real-time datasets. The novelty behind this research lies in the application of an improved and efficient model for detecting degradation in PV modules based on bidirectional LSTM Networks and GRUs. The objective of this work is to develop an efficient algorithm to predict PV cell yield accurately and schedule maintenance-related tasks in advance. The motivation is to reduce the power losses due to unnoticed breakage and lack of maintenance and reduce downtime of PV Cells.

III. METHODOLOGY

3.1 Modelling

The semiconductor materials that make up the photovoltaic cell effectively absorb the photons that the sun emits and produce an electron flow in the external circuit. A PV cell's equivalent circuit is illustrated in Fig-1. The equivalent circuit model of a PV solar cell can be expressed using the following equation[1]:

$$I = Ipv - I \left[e^{\frac{U+IR_S}{}} - 1 \right] - \frac{U + IR_s}{R_{sh}} \qquad (1)$$

where I is the output current of the cell, V is the output voltage, Ipv is the light-generated current, I0 is the reverse saturation current of the diode, Rs is the series resistance, Rsh is the shunt resistance, VT is the thermal voltage (approximately 25 mV at room temperature), and n is the diode ideality factor (a value between 1 and 2 that characterizes the deviation of the diode from ideal behaviour). Maximum power point voltage is the value on the I-V curve at which the most power is produced. The amount of current on the I-V curve that delivers the most power is known as the maximum power point current. Efficiency is a measurement of how much solar energy is converted into peak electrical energy.

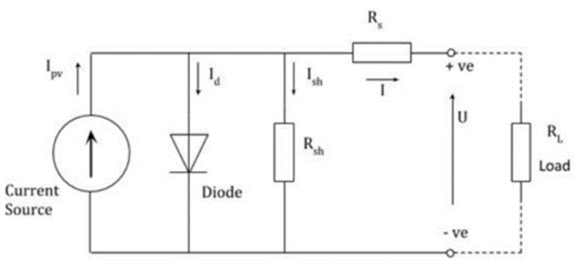

Figure 1: Equivalent Circuit of a PV Cell

Fig-2 shows the variation of current and power with voltage, there is a maxima in the power curve called the maximum power point.

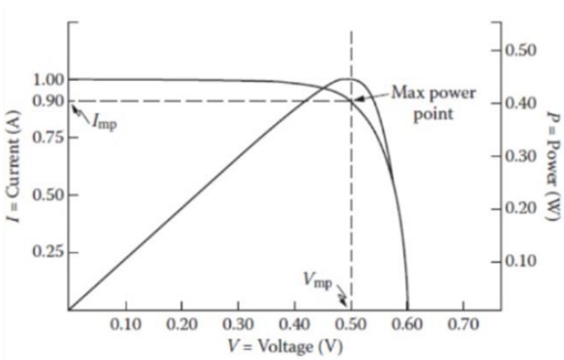

Figure 2: Variation of Current and Power with Voltage

IV. DATASET DESCRIPTION

The data set used in this research was obtained from a research undertaken by Ani Kannal[12]. The data was collected over a 34-day period with 15-minute intervals, in the months of May and June from two solar power plants in India. The data set consists of two pairs of files, each of which comprises a data set for power generation and a data set for sensor readings for two separate plants. The statistics for power generation are obtained at the inverter level since each inverter has several solar panel lines tied to it. At the plant level, a single array of strategically positioned sensors collects the sensor data. To perform an in-depth analysis various parameters were recorded along with the respective timestamps. The parameters considered were Solar Irradiance— the quantity of sunlight that is available at a certain place at a particular time is crucial for PV system design, maintenance, and operation. Solar irradiance is the amount of energy that is received from the Sun per unit area[13] (surface power density expressed in Watt per square metre (W/m2)). Pyranometers measure sun radiation by counting the photons that strike a physical or chemical component inside the sensor. In the process of converting the DC power from the PV array to AC, some power is wasted in addition to wire losses. Peak efficiency for commonly used inverters ranges from 88 to 90%[14]. To analyse how the inverter operates, the AC Power values have also been taken into account. Figure-3 below shows a block diagram depicting the operation of the PV Cell and the process of data collection. Each PV cell has a module temperature sensor attached to it which records the module temperature. An ambient temperature sensor module is also used to measure the ambient temperature. The yield generated by the PV array is also measured before and after it is converted to AC. Each inverter has a unique inverter ID which helps us identify the faulty inverter in case of substantially reduced yield after conversion to AC. The dataset also contains the solar panel id, the daily yield and the total yield. The total yield reflects the entire yield for the inverter up to that point in time, and the daily yield is a cumulative total of the power generated on that day, up to that point.

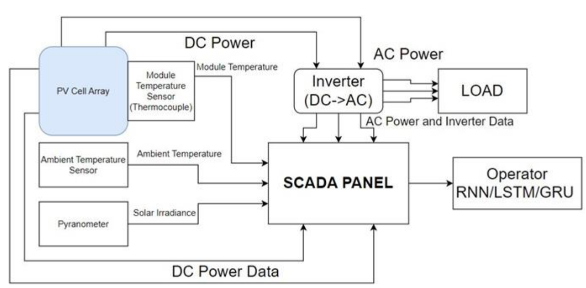

Figure 3: Operation of PV Cell and Data Acquisition of the parameters

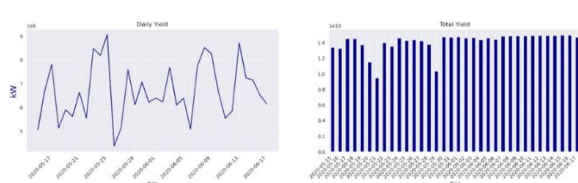

Figure 4: Daily yield and Total yield(kW) presented date wise from May to June 2020. This figure

represents a line graph plotted for Daily yield with respect to the date and a bar graph of Total Yield varying with the date. This depicts a general trend of values of daily yield and total yield. The daily yield fluctuates due to changes in climatic conditions.

V. RESULTS AND DISCUSSION

The model was trained on the aforementioned dataset of solar power generation of power plants. We evaluate the performance of the proposed framework using mean squared error as a loss function. The findings demonstrate that our model, when trained on the same dataset and validated on the same sample of test data, was able to obtain a mean squared error of 3.004e-05 on the test set, which is much better than the mean squared error of 0.4452 in comparison with Bi-LSTM network. Additionally, we added two additional dense layer with 128 and 64 neurons respectively after the GRU layer, which improved the performance of the model. Figures 5 and 5 depict the mean squared error loss of the GRU and Bi-LSTM, respectively.

The proposed configuration was finalized after performing extensive experiments with different number of hidden layers and neurons. Additionally, we also monitored overfitting by using callbacks like EarlyStopping and also by using validation data in the training process.

The implementation of GRUs has been shown to outperform LSTM and Bi-LSTM networks in our study. The results demonstrate that GRUs are capable of effectively capturing sequential information while maintaining a lower number of parameters, leading to faster training times and improved performance. It was found that GRUs regularly outperformed LSTMs and Bi-LSTMs in terms of accuracy and computing efficiency, and that they were 12 times faster than LSTMs. The results were consistent across a number of datasets and tasks. Table 1 represents the actual and GRU-predicted values of AC power.

In conclusion, our proposed model using GRU for regression task achieved the best performance. Thus, it can be a potential baseline framework for further studies on similar solar panel power output and weather data. However, it’s worth conducting further studies and testing the model on other datasets to confirm its generalizability.

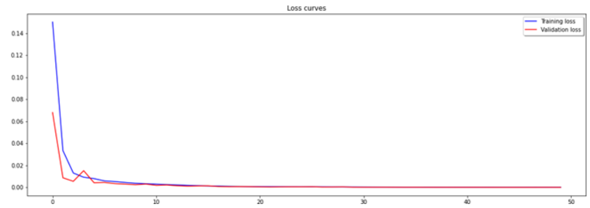

Figure 5: Mean squared error loss using GRU

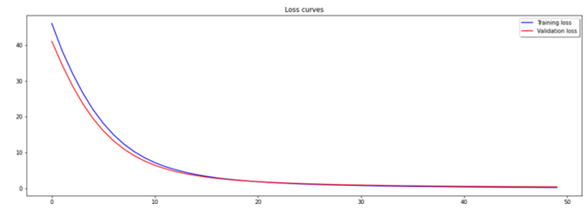

Figure 6: Mean squared error loss using Bidirectional LSTM

Actual AC Power	Predicted AC Power using GRU
184.2857143	186.8338
236.5375	239.90503
265.0857143	271.92715
277.2875	286.29535
312.3714286	318.44855
346.475	352.97797
383.8714286	390.63083
408.325	416.72726
437.1	445.6028
469.175	478.2652
511.2714286	518.09753
523.0714286	530.1792
541.65	547.93713
662.1285714	666.69617
713.1625	721.8489
665.5714286	674.32855
749.625	756.9993
859.2714286	868.309
837.9125	845.0643
727.9428571	735.4146
563.4125	563.9961

Table 1: Actual and GRU Predicted AC power

VI. LIMITATIONS

This study aimed to predict the yield of a PV power plant. While the study yielded some promising results, there were several limitations that need to be addressed.

Firstly, the sample size was relatively small and the study was conducted at two sites which may limit the generalizability of the findings to other settings. Secondly, the study was conducted over a relatively short period of time, which may not capture the difference in plant yield throughout a whole year. Lastly, the data only included a small number of climatic variables, which reduced the potential for volatility in the power plant's output.

The study's shortcomings could limit how broadly the results can be applied, but they also offer valuable information for future investigation. The current study can be expanded upon and our knowledge of forecasting a power plant's yield can be advanced by researchers by addressing these constraints.

CONCLUSION AND FUTURE SCOPE

In this study, various deep learning algorithms for sequential data were compared and evaluated. By discovering and categorizing irregularities, predicting breakdowns, and organizing repair work, the strategies enhance PV power plant operation and maintenance. The long-term financial perfor- mance of solar PV plants must be maintained, and downtime must be minimized, through predictive maintenance. A hierarchical generative model framework at the sensor level and system level is uti- lized to detect the common failure class patterns from real-time monitoring data such as power output, temperature, and weather data. In order to illustrate the proposed system for predictive maintenance, this study employed a power plant as a case study.

This study's future objectives may include the evaluation and analysis of more data samples. The results for the power yield can be substantially better and more precise when the data is logged at a greater frequency. For prediction tasks, modelling with a shorter calculation time is preferred. Datasets with more parameters like the AQI and wind speed may be used to study the effects of all such parameters. The frequency for data logging of various parameters can be increased for accurate measurements.

REFERENCES

[1] K. M. Sundaram, S. Padmanaban, J. B. Holm-Nielsen, and P. Pandiyan, *Photovoltaic Systems:Artificial Intelligence-based Fault Diagnosis and Predictive Maintenance*. CRC Press. Google- Books-ID: MJ9cEAAAQBAJ.

[2] J. Vicente-Gabriel, A.-B. Gil-Gonzâlez, A. Luis-Reboredo, P. Chamoso, and J. M. Corchado, "LSTM networks for overcoming the challenges associated with photovoltaic module mainte- nance in smart cities," vol. 10, no. 1, p. 78. Number: 1 Publisher: Multidisciplinary Digital Publishing Institute.

[3] "Methods of photovoltaic fault detection and classification: A review - ScienceDirect."

[4] Z. M. Ç ınar, A. Abdussalam Nuhu, Q. Zeeshan, O. Korhan, M. Asmael, and B. Safaei, "Machine learning in predictive maintenance towards sustainable smart manufacturing in industry 4.0," vol. 12, no. 19, p. 8211. Number: 19 Publisher: Multidisciplinary Digital Publishing Institute.

[5] M. Ejgar, B. Momin, and T. Ganu, "Intelligent monitoring and maintenance of solar plants using real-time data analysis," in *2017 IEEE International Conference on Consumer Electronics-Asia(ICCE-Asia)*, pp. 133–138.

[6] D. Perrottet, C. Boillat, S. Amorosi, and B. Richerzhagen, "PV processing: Improved PV-cell scribing using water jet guided laser," vol. 6, no. 3, pp. 36–37.

[7] M. Cendagorta, M. Friend, J. Rodrı́guez, J. Fernández, M. Padrón, G. Moncho, O. González, A. P´ıo, D. Molina, A. Linares, C. Montes, and E. Llarena, "A cleanliness monitoring system for PV installations," pp. 3951–3954. ISBN: 9783936338287 Publisher: WIP.

[8] T. Huuhtanen and A. Jung, "PREDICTIVE MAINTENANCE OF

PHOTOVOLTAIC PANELS VIA DEEP LEARNING," in *2018 IEEE Data Science Workshop (DSW)*, pp. 66–70.

[9] D. Saquib, M. N. Nasser, and S. Ramaswamy, "Image processing based dust detection and prediction of power using ANN in PV systems," in *2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT)*, pp. 1286–1292.

[10] S. A. Jumaat, F. Crocker, M. H. A. Wahab, N. H. M. Radzi, and M. F. Othman, "Prediction of photovoltaic (PV) output using artificial neutral network (ANN) based on ambient factors," vol. 1049, no. 1, p. 012088. Publisher: IOP Publishing.

[11] P. Malik, R. Chandel, and S. S. Chandel, "A power prediction model and its validation for a roof top photovoltaic power plant considering module degradation," vol. 224, pp. 184–194.

[12] A. Kannal, "Solar power generation data." https://www.kaggle.com/datasets/ef9660b4985471a8797501c8970009f36c5b3515213e2676cf40f540f0100e54.

[13] I. Haseeb, A. Armghan, W. Khan, F. Alenezi, N. Alnaim, F. Ali, F. Muhammad, F. R. Al- bogamy, and N. Ullah, "Solar power system assessments using ANN and hybrid boost converter based MPPT algorithm," vol. 11, no. 23, p. 11332. Number: 23 Publisher: Multidisciplinary Digital Publishing Institute.

[14] Z. Yuan, G. Xiong, and X. Fu, "Artificial neural network for fault diagnosis of solar photovoltaicsystems: A survey," vol. 15, no. 22, p. 8693. Number: 22 Publisher: Multidisciplinary Digital Publishing Institute.