Solar Data Faults Detection Using Machine Learning

[1]Prof. Vishal R. Shinde, [2]Mr.Vipinkumar Gautam, [3]Mr.Chinmay Gawande, [4]Miss.Pooja Mahale

[1]Asst.Professor,[2,3,4]UG Student,[1,2,3,4]Computer Engg. Dept. Shivajirao S.Jondhle College of Engineering & Technology, Asangaon, Maharshatra, India.

[1]mailme.vishalshinde@gmail.com,[2]vipinkumargautam384@gmail.com,[3]chinmaydgawande1998@gmail.com,[4]poojamahale2395@gmail.com

Abstract -The proposed system aims to detect data faults that may occur in a Solar Energy plant. Malfunctions in a solar plant are quite troublesome as they may cause monetary losses to an organization. The system detects these data faults by monitoring and analyzing the feed from the solar data-logger and any other relevant data available on the plant site. The causes of data fault can vary from tree shade to damage to cables or even inverter failure. This system detects data faults using Machine Learning[1] and reports to the user. This further enables the users to completely prevent or reduce any damage to their system.

Keywords –Solar, Inverter, Faults, Machine Learning.

I. INTRODUCTION

Malfunctions in a Solar system are more than troublesome. These faults lead to financial losses if they go unobserved for extended periods. Modern monitoring devices help to prevent losses in yields and also offer additional functions, which make operating a solar plant even more efficient. An experienced PV installer states one such case - he was recently called to a plant that was offline because of a technical fault after operating for a long period. The customer didn't notice the failure until two months later and had to undergo repairs. Due to this, the customer lost out on profits during the summer months. This example highlights the need for a more reliable system and that faulty data is not to be overlooked.

Applications of the monitoring system are the Ground mounted Solar, Rooftop Solar, Smart villages, Solar Street lights, Solar cities and Micro grids. Products like Solar home lighting systems, solar water heating, solar lanterns, solar pumps, solar mobile chargers, solar cookers, LED solar torch, , solar Inverters, solar RO plant, solar fan, etc. can be monitored through this system[2].

By monitoring the energy forecast, communities who are using solar energy can utilize their power production and expenditure during good weather.

This system detects the cause and advises on diagnostic steps onto the system supervisor.

II. AIMS & OBJECTIVES

a. Aim :

The occurrence of anomalies in solar energy data-loggers is very bothersome as it leads to wrong calculations and loss in reliability of the system. This proposed system aims to improve solar data-loggers by detecting and diagnosing the faulty data recorded by it and also inform the user if more faulty data found and this data can be used for improving and maintaining the system.

b. Objective :

The main objectives of this system are as follows:

- Faulty data detection using machine learning
- Advises on possible fixes that may not require human intervention.

III. LITERATURE SURVEY

With the dawn of Internet of Things, almost all industrial level systems have initiated automating most mechanisms of their system architecture. In this case, many researchers have come forward and brought forth prototype models to automate solar power production systems.

Paper 1: Online Fault Detection in PV Systems

Multiple researchers have come up with methodologies to detect faulty data using machine learning. In the research made by Jacques Martel, Radu Platon,, Norris Woodruff, and Tak Y. Chau, they have used Machine Learning regression algorithms to forecast the daily mean power consumption for a solar plant. This system they implemented does not necessitate any supplementary hardware. It simply utilizes existing data reading records to predict the future readings. Hence further enhancing cost efficiency while aiding preventive maintenance and profit analysis of the solar farm [1].

Paper 2: IoT based Solar Energy Monitoring System

This paper uses IoT based architecture to monitor PV systems. However, this means that additional hardware has to be implemented [5].

Paper 3: Remote Monitoring System for Solar Power Panels using Intelligent Sensors Network

This is the phase where the most number of advancements have been made. A number of systems are in place already or available for retail which establishes a remote management interface between the system and the client. In a study made by Mohamed Fezari, Ali Al-Dahoud, Fatma Thamer Al-Rawashdeh, Zohra Belhouchet, they integrate the same IoT concept, but instead of adding new hardware to each panel, they implement a wireless sensor network over the already existing architecture. Integrating the detection and categorization of faults in the sensor node is the alternative solution.

IV. EXISTING SYSTEM

The existing solar power plants have their on-site inverters equipped with multiple safety features, warnings, and etcetera. Some modern systems also include a web interface for remote monitoring of the site as well. However in most cases these remote interfaces only show alert the user after a fault occurs [2].

The prediction phase of the systems all focus on predicting the power generation and profit values while little attempt is made to predict the defects and faults occurring in the system. There are some examples which, to an extent, predict errors in energy readings, though they require additional hardware to be installed, thus compromising the system's cost efficiency[3].

There is a multitude of modern solar data-loggers available with the aforementioned capabilities built into them. Some are also installed along with the system manufacturer itself, complete with a remote web or mobile management interface.

V. COMPARATIVE STUDY

Table 5.1: Comparative Analysis

Sr. No	Paper Title	Technique	Advantages	Disadvantages
1.	Detection of transmission line faults in presence of solar PV system using stackwell's algorithm	Stackwell's algorithm, Support vector algorithm	This algorithm is used to find the fault detection in solar transmission line in presence of solar PV system.	This is only able to detect the faults in the transmission line so if the fault is not in the transmission line It will not be able to detect it
2.	Fault Detection and Localization in Solar Photovoltaic Arrays Using the Current-Voltage Sensing Framework	Current voltage sensing framework	This system is able to notice the short circuit, open circuit and hotspot fault.	This system is a hardware base model so this is costly. And also possibility of fault in the hardware as it is hardware based.
3.	A Novel Detection Algorithm for Line-to-Line Faults in PV Arrays Based on SVM	Support Vector Machine (SVM)	This system is able to sense the faults in the photovoltaic array.	This system is not able to detect fault related to other than photovoltaic array.
4.	Online Fault Detection in PV Systems	Linear Regression (Machine Learning)	This system is able to sense and validate the faults in the PV system readings.	This system uses only Irradiance to detect faults.

VI. PROBLEM STATEMENT

Even with the rapid advancements in this field of study, there is still a demand for more cost efficient and reliable alternatives. There is a multitude of modern solar data loggers available with the aforementioned capabilities built into them. Some are also installed along with the system manufacturer itself, complete with a remote web or mobile management interface. However, these systems include hardware support for detecting the fault in the system related to the solar PV system. The hardware is not cost efficient and they increase the cost of the system including additional maintenance. This proposed system aims to detect the fault in the system using software and, thus, minimizing the cost of the existing system.

VII. PROPOSED SYSTEM

The proposed system is equipped for highlighting the faulty data. The following features will be included in the proposed system. The proposed system's main objective is to remove the faulty data from the proper data set and possible reason of the data being faulty and store them on separate collection. This makes the information collected form the system clean with few or no faulty values in the collection. The proposed system have three main module one application, other is fault data detection system and the storing system. The proposed system is used for mentoring the solar inverter data with help of these data the fault in the system is able to sense the fault in system. If any error is detected, the user will be informed with the related information collected by the datalogger. The fault detection system is used to detect the fault data by passing the data through the algorithm.

For fault data detection system machine learning algorithms are used. Some of the examples are linear regression, logistic regression etc. This system can help to monitor the solar plant with providing the information related to errors and also providing the information about the information gathered with summarized details

VIII. ALGORITHM

The general idea of working of proposed system algorithm is given as follows:

Step 1: Get the data from logger then send it to the bad data detection system.
Step 2: If the data is detected as faulty data then step 3 else step 4.
Step 3: Add detected data to collection and goto step 5.
Step 4: Store the data to database and Go to step 1.
Step 5: Check for old recorded faulty data count if it is greater than 10 then goto step 6
Step 6: Notify the user.
Step 7: EXIT

Pseudo code:
detection_model(data)
Input: data from the data logger records
if(multiple_linear_regression(data)!= faulty) then set= 1;else set= 0;
if(set==0)
thencollection.add(data);
elsedatabase.insert(data);
if(count(fauly_data)>10)
then return notify;
return set;

IX. MATHEMATICAL MODEL

Under normal operation conditions, the following conditions (1-2) are true:

l1=l2=⋯=lm (1)

where, l1 is the current sensor in branch 1, l2 is the current sensor in and lm is the current sensor in branch M and soon. U11=U12=⋯=U21=⋯=Us(2) where U11 is the first voltage sensor in branch 1, U12 is the second voltage sensor in branch 1, U21 is the first voltage sensor in branch 2 and Us is the standard voltage as shown in(3):

$$Us=\left(\frac{n}{N}\right)*Um \quad (3)$$

where n is the number of PV modules covered by one voltage sensor, N is the number of PV modules connected in series in a branch, and Um is the maximum output voltage of the entire array. But under abnormal (fault) operation conditions, the above mentioned conditions are no longer valid. In this scenario, the
output current of the faulty branch is not the same as the healthy branch
As the training dataset used for developing the model, with validating the dataset it is also predicting the results. If only training dataset is used for prediction then the accuracy can be overestimated since the model is tuned to fit according to the data.

Method used for detecting the fault:
Multiple linear regression used if more than one variable are related to the output variable. As in linear regression the relationship with the variable is valid to this one.
Algorithm provides better output if two or more variable are related to the output variable. Multiple linear regression tries to model the relationship between two or more variables and a output variable by fitting a linear equation to collected data. Each value of the independent variable *x* is associated with a value of the output dependent variable *y*. The population regression line for *p* related variables $x_1, x_2, \ldots, x_p$ is defined to be

$$\mu_y = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \ldots + \theta_p x_p.$$

This line relates how the mean response μ_y changes with the related variables. The collected values for *y* vary about their means μ_y and are considered to have the same standard deviation σ. The fitted values $b_0, b_1, \ldots, b_p$ estimate the parameters $\theta_0, \theta_1, \ldots, \theta_p$ of the population regression line. Further these fitted values and the parameters used in the relationship formula and used for calculating the estimated output.

X. SYSTEM ARCHITECTURE

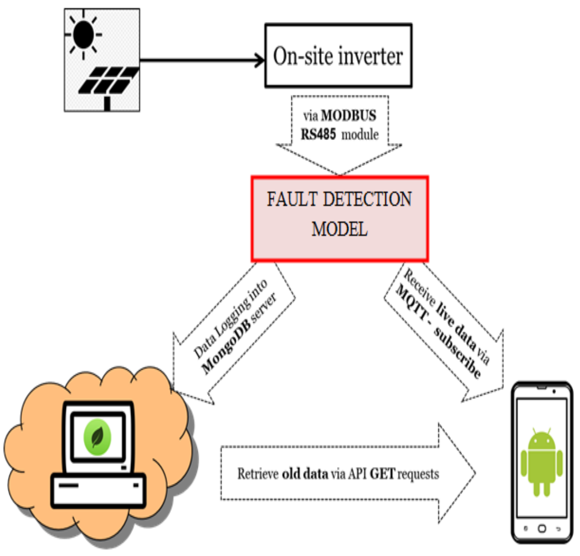

Fig 10.1: Detection model present as shown.

Description:

1. Raspberry pi or PC is used to connect via RS485 to USB converter to the inverter for getting the data.
2. This data is then send to the data detection system that detect the type of data whether the data is faulty or normal
3. The data detection system is used to return the type of data to the raspberry pi or pc
4. Data detection system is responsible for calculating the reasons of errors.
5. If the data is faulty then it transferred to MongoDB faulty collection else transferred to proper data collection.
6. The web API server is the interface between the mongo dB and android application
7. This web API server is responsible for notifying the user about the errors.
8. The mongo DB is responsible for storing the data .This contains two collection one for faulty and another for normal data.
9. The android application used for displaying the graphical chart of the data and Notifying the user about the error with proper reason.

XI. ADVANTAGES

1. The annual data average accuracy will be increase by this system.
2. The Fault detected by system is used to minimize the upcoming faults.
3. The System help to diagnose the system by collecting the fault data and there reasons.
4. This also can be used as a regular data logger application

XII. DESIGN DETAILS

The GUI is simple as a datalogger application. The GUI displays the name of the application i.e. solarDatalogervcg. This is for displaying the data collected and filtered from the inverter. This GUI is used for displaying the faulty data along with the proper data and the data is validated by the fault data detection system and then this application will display the data and also notify the user

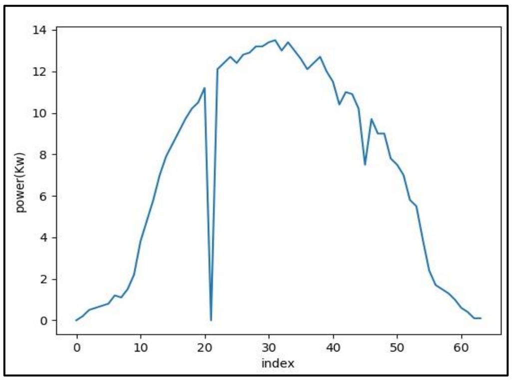

Fig 12.1: In the above diagram the portion with red colour specify the fault data collected as the graph drop this is considered as a faulty data

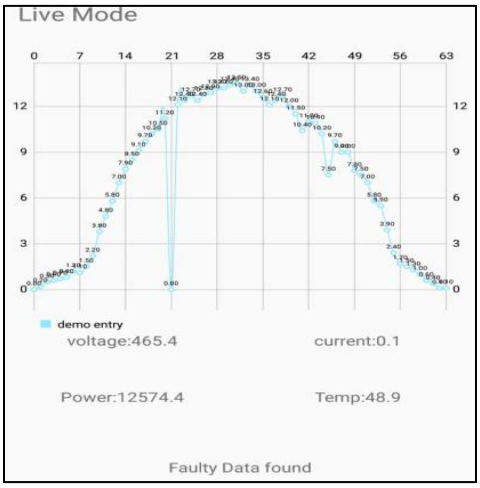

Fig 12.2: GUI snapshot 2

XIII. CONCLUSION

Thus, we have tried to implement[1] Radu Platon, Norris Woodruff, Jacques Martel, and Tak Y. Chau - “Online Fault Detection in PV Systems” IEEE 2015 [1], combining another paper Suprita M. Patil, Vijayalashmi M., Rakesh Tapaskar - “IoT based Solar Energy Monitoring System” IEEE 2017. Thus, a datalogger is designed to detect faults on a PV system. Future scope for PV systems is unbounded, enhancements like IoT, SMD, Machine Learning, etc. can profoundly increase the accuracy and efficiency of PV systems.

REFERENCES

[1] Radu Platon, Jacques Martel, Norris Woodruff, and Tak Y. Chau - “Online Fault Detection in PV Systems” IEEE 2017.

[2] Dipak Saha, Soham Adhya, “An IoT Based Smart Solar Photovoltaic Remote Monitoring and Control unit” CIEC 2016.

[3] Mohamed Fezari, Ali Al-Dahoud, “Remote Monitoring System for Solar Power Panels using Intelligent Sensors Network” IEEE Nov. 2016.

[4] Khurum Nazir Junejo, Faizan Jawaid, “Predicting Daily Mean Solar Power Using Machine Learning Regression Techniques” IEEE Aug. 2016.

[5] Suprita M. Patil, Vijayalashmi M., Rakesh Tapaskar - “IoT based Solar Energy Monitoring System” IEEE 2017.

[6] Andreas S. Spanias, “Solar Energy Management as an Internet of Things (IoT) Application” IEEE 2017.

