PHOTOVOLTAIC ARRAY SYSTEMS FAULT DETECTION AND CLASSIFICATION USING MACHINE LEARNING APPROACH"

[1]Tushar G. Nakod, [2]Prof. N. N. Mandaogade

PG Student M.E.4thSemester,Electronics & Telecommunication Engineering, G H Raisoni college of engineering and management , Amravati, Maharashtra, INDIA[1], Assistant Professor, Sant Gadge Baba Amravati University, Amravati. Maharashtra, INDIA[2].

tushar.nakod@yahoo.co.in[1], nitin.mandaogade@ghru.edu.in[2]

ABSTRACT

Over the past few decades, renewable energy has garnered enormous attention which has increased the importance of photovoltaic systems significantly. However, these PV systems are susceptible to a variety of faults resulting in the variability of PV output power. In the absence of timely detection, these faults cause output power to decrease, rendering the PV array system unreliable.

Additionally, in some cases, the system falls prey to aging or wear and tear. Therefore, detection of the fault and also the identification of the type of fault are of paramount importance to ensure optimal functioning of the PV array system. In this paper, fault classification and detection in the photovoltaic array systems using machine learning techniques have been attempted.

We evaluate the performance of the classifiers based on Support Vector Machine and Random Forest algorithms. Simulation results reveal that the Random Forest classifier has the maximum accuracy (and thus the minimal mean squared error) using MATLAB software.

LITERATURE REVIEW

Researchers have been developed or implemented efficient technique for fault detection and classification using machine learning, which are explained below. The common fault occurrences in a PV array system comprise the following faults typically - ground, mismatch and line-to line [2-3]. Among these faults, the authors in [2-3] have focused their study around line-to-line faults and mismatch faults. This is primarily due to the reason that the identification of these two fault types is cumbersome vis-à-vis the ground fault by utilizing the conventional protection devices. [4]. Various studies conducted in the past have shown that certain faults like line-line faults, partial shading faults. open-circuit faults in PV arrays are difficult to be removed even by current protection devices owing to the current limiting nature and non- linear output characteristics of PV arrays [5-7]. In [8], the authors have proposed an automatic fault detection method based on power losses. This methodology detects faults that occur primarily on the DC side of the PV system. In the last few years, artificial intelligence has attracted the researchers working in this field. The primarily applied artificial intelligence techniques to identify the PV array faults include the Machine Leaming strategies. In [9 10], the authors developed a technique to detect partial shading and short circuit faults in a PV array using an artificial neural network. In [9], the authors have developed an FPGA based model and have validated the performance of ANN using the dataset generated in [21]. In [11], a fuzzy logic approach to distinguish short circuit faults from partial shading faults of a PV array has been put forth. In [12], the authors proposed a statistical method to effectively recognize different faults such as the faults in the PV module, strings, in addition to a defective MPPT unit. In [13], the researchers employed the Kalman-filter to examine the loss in the power output of a PV array. The authors of [14], devised a statistical monitoring approach.

MATLAB Environment

MATLAB is a software program that allows you to do data manipulation and visualization, calculations, math and programming. It can be used to do very simple as well as very sophisticated tasks. The name 'MATLAB' comes from two words: matrix and laboratory. According to The Math Works (producer of MATLAB), MATLAB is a technical computing language used mostly for high-performance numeric calculations and visualization. It integrates computing, programming, signal processing and graphics in easy to use environment, in which problems and solutions can be expressed with mathematical notation. Basic data element is an array, which allows for computing difficult mathematical formulas, which can be found mostly in linear algebra.. It can be widely used to analyse data, modelling, simulation and statistics. Mat-lab high-level programming language finds implementation in other fields of science like biology, chemistry, economics, medicine and many more. Most important feature of MATLAB is easy extensibility. It has evolved over many years and became a tool for research, development and analysis. MATLAB also features set of specific libraries, called toolboxes.

PV module configuration and discussion on faults:

In [17-19], the authors have discussed various models of a solar cell in order to ascertain the output power generated by it. However, the single diode model [17], has been widely accepted and used due to its accuracy and simplicity. This model views the solar cell as a diode connected in parallel with a photo-generated current source. A circuit diagram of the single diode model of a PV cell is shown in Fig 1. The series resistance R, of the solar cell is added on account recombination of carriers at the junction region, contact, and material resistance. The parallel resistance Rsh is used to account for leakage current at the edge.

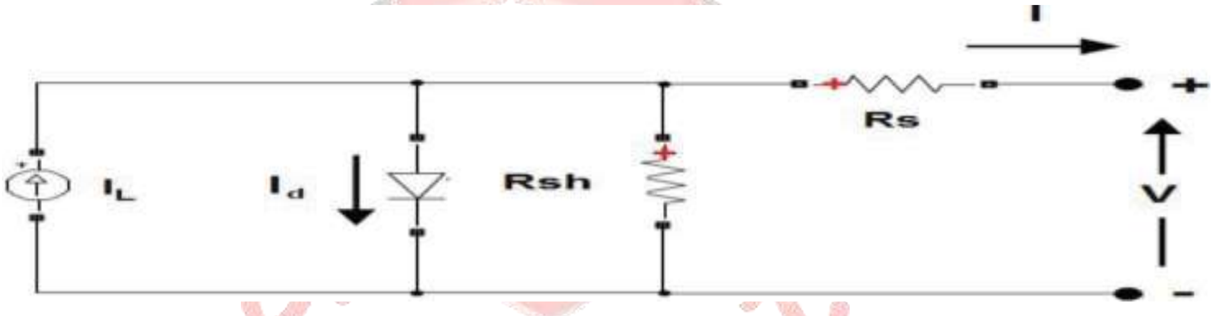


Figure : Single Diode Model of a PV Cell.

The single diode model is defined by the equations:

$$I_D = I_O \left[\exp\left(\frac{V_d}{V_t} - 1 \right) \right]$$

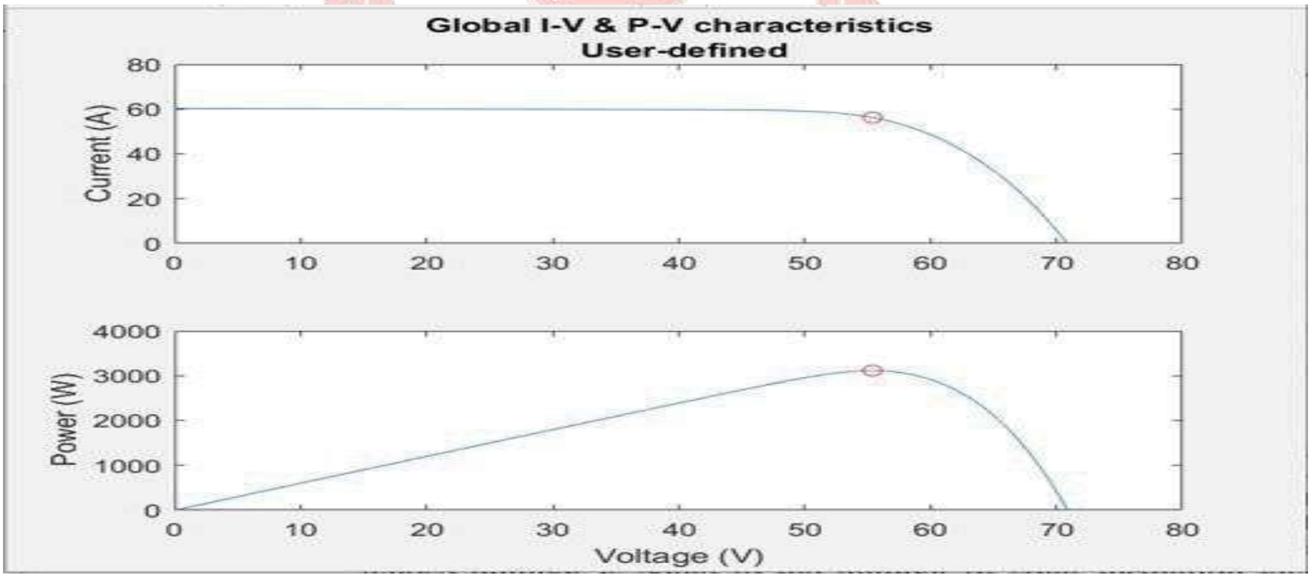
$$V_d = \frac{kT}{q} \times nI \times Ncell$$

where, Id refers to the diode current (in A), Va refers to the diode voltage, I, refers to the saturation current of the diode (in A), nl is the diode ideality factor, which is usually less than 10, K refers to the Boltzmann constant =1.3806 x 10-23 J.K-1 and Ncell refers to the number of series-connected cells present in a module. In our study, we have developed a 6x7 PV array system using the PV block available in the SimScape Electrical Library in MATLAB. This array comprises of 7 strings of modules connected in parallel (each string having 6 modules connected in a series fashion) as per the configuration defined in [1].

The various parameters for simulation that were used are tabulated in Table 1.

TABLE : PV Block Simulation and Parameters.

Input 1	Sun Irradiance in W/m^2
Input 2	Cell Temperature, in deg. C
Parallel Strings	1
Series-Connected Modules per string	1
Maximum Power (W)	83.2824
Open Circuit Voltage Voc (V)	12.64
Voltage at maximum power point Vmp (V)	10.32
Temperature Coefficient of Voc (%/deg. C)	-0.33969
Cells per module (Ncell)	20
Short Circuit Current Isc(A)	8.62
Current at Maximum Power Point Imp (A)	8.07
Temperature Coefficient of Isc (%/ deg. C)	0.063701
Light Generated Current IL (A)	8.6307
Diode Saturation Current Io (A)	$1.4176e^{-10}$
Diode Ideality Factor	0.99132
Shunt Resistance Rsh (ohms)	82.1161
Series Resistance Rs (ohms)	0.098625



WAVEFORM 1: PV Array Module Characteristics under STC

1) Partial shading fault

An obvious impact of the partial shading fault is that Impp declines. An obvious impact of the partial shading fault is that the MPP current of the PV array declines. The connection diagram for the simulation of Partial Shading Fault has been given in Figure.

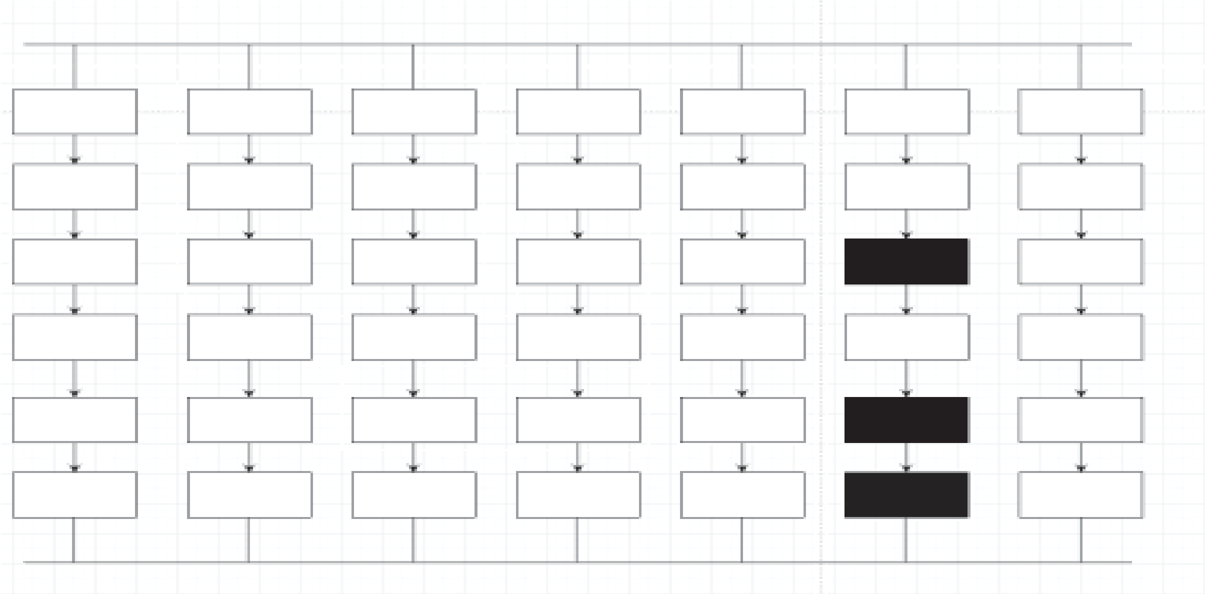


Figure : PV Array System Configuration for Partial Shading Fault

2) BYPASS DIODE FAULT

This fault occurs because of a short circuit of the bypass diode in case of an incorrect connection. If one full module is shorted by bypass diode, the maximum power, Pmax and open circuit voltage, Voc of the module drop significantly and Isc remains the same as that of other normal strings.

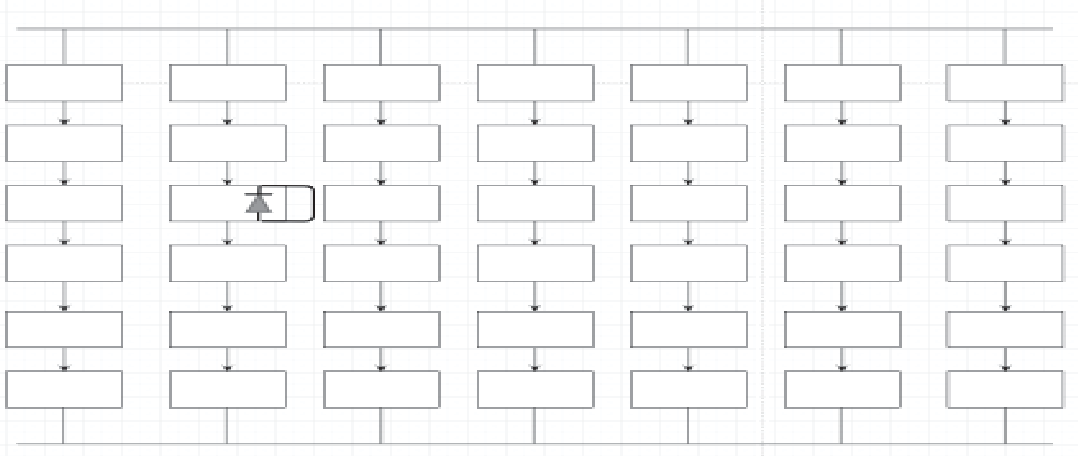


Figure : PV Array System Configuration for Bypass Diode Fault

3) BRIDGING FAULT

This fault is one of the most commonly occurring faults in photovoltaic array systems. Whenever a link is formed between two points in the string of modules (having different potentials) which generally has a low resistance, then it is categorized as a bridging fault. They usually lead to reduced array open circuit voltage (Voc) but the reduction in array short circuit current (Isc) is very less. The fault with the larger voltage difference between two fault points will lead to the larger reduction in Voc, Impp and V_1 mpp.

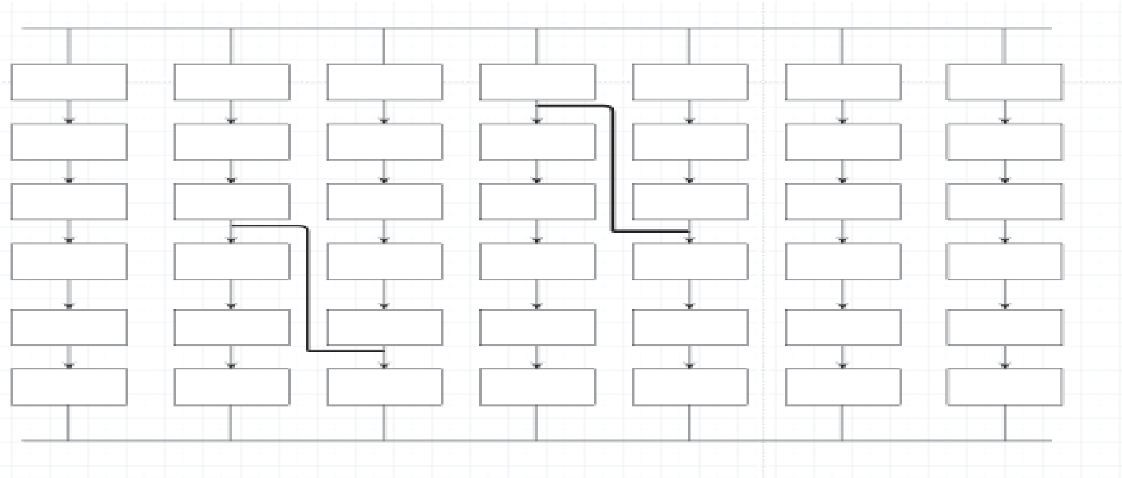


Figure: PV Array System Configuration for Bridging Fault

4) TEMPERATURE FAULT

One important factor that affects the power output of the model is the temperature condition Keeping the irradiance constant, a rise in temperature leads to decreased V. and the rate of decrease is typically 2-2.5 mV/C. However, I practically remains constant. In Fig. 6, for simulation of the temperature fault, temperature for some of the modules has been varied from the STC condition of 25°C and different data points are generated.

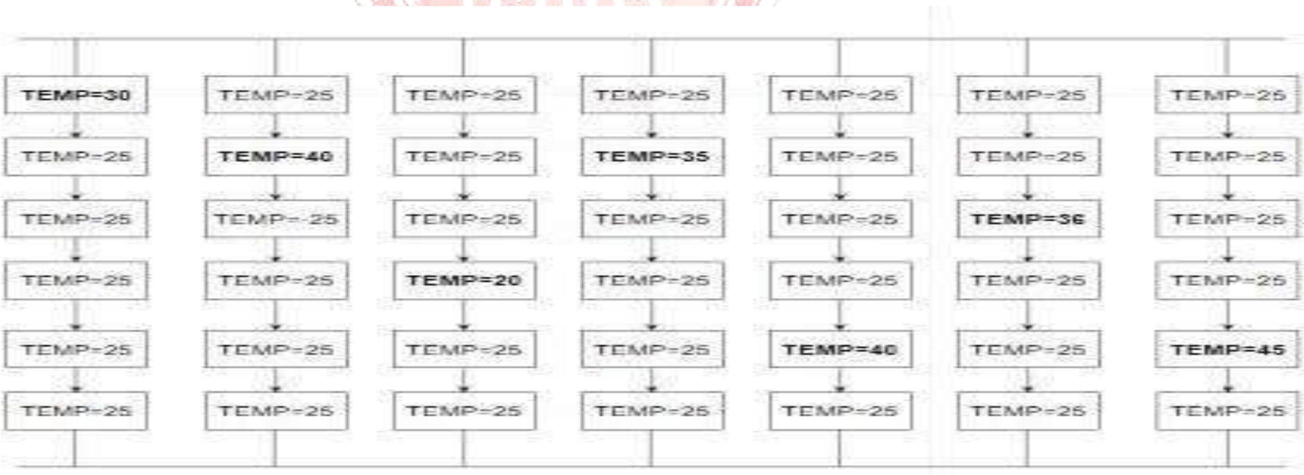


Figure: PV Array System Configuration for Temperature Fault

5) COMPLETE SHADING FAULT

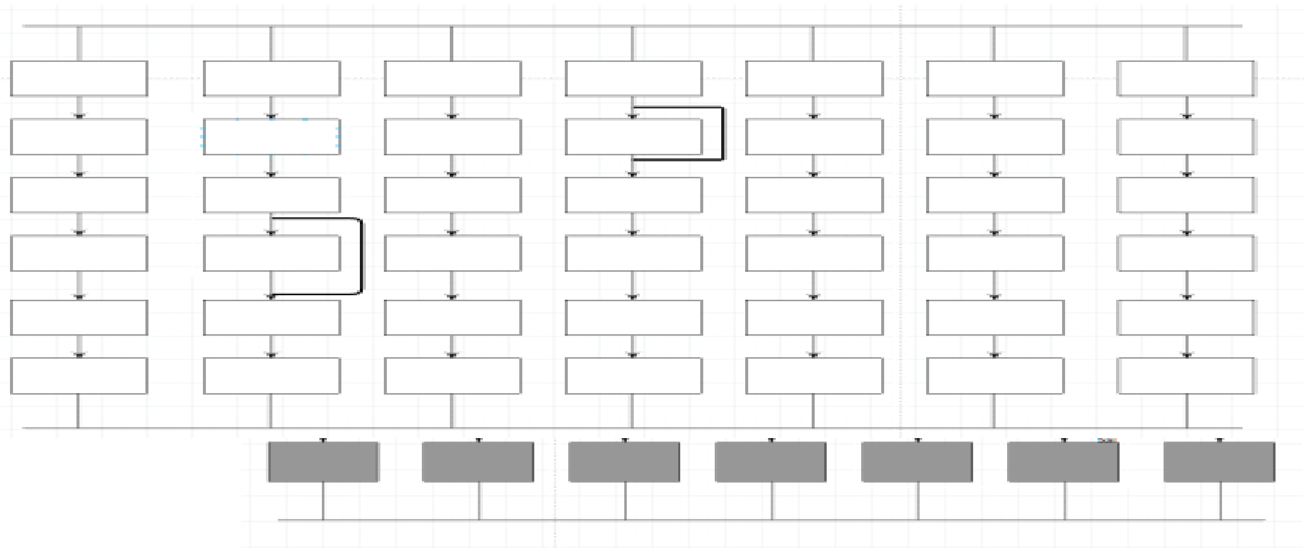
A complete shading fault occurs when all modules present in an array receive the same solar insolation, which is less than 1000 W/m² As expected, the peak point current, voltage and power decrease under this type of fault.

Figure: PV Array System Configuration for Complete Shading Fault.

6) SHORT CIRCUIT FAULT

The short circuit fault in a PV module is mainly due to incorrect wiring in the system Other factors such as aging. vibration and abrasion can also be attributed as the possible causes of such faults. In the incidence of occurrence of this fault. I remains constant but V reduces rapidly, The configuration for short circuit fault is depicted in Figure.

Figure: PV Array System Configuration for a Short Circuit Fault.



7) DATASET GENERATION AND FAULT CLASSIFICATION:

A 6 x 7 PV array system using the PV block available in the SIMSCAPE ELECTRICAL library has been developed in MATLAB Simulink. The model is then simulated for the six fault types by altering the PV model configuration in accordance with the each of the faults depicted in Fig. 3-8. For example, for generating data points for the partial shading fault, the value of insolation for some modules in the array are reduced from STC values. Similar procedures are adopted for simulating other faults. At the end, a dataset with 50 total readings for each of the six faults was generated. The eight input features of the dataset are the following parameters: Voc (Open circuit Voltage), Isc (Short circuit Current), Vmp (Maximum Peak Point Voltage), Imp (Maximum Peak Point Current), Temperature (in °C), Irradiance (in W/m2), Fill Factor (FF), and Pmax (Maximum Power Output). Fill Factor is a vital parameter that evaluates the PV array system performance. The expression for Fill Factor (FF) is given in (3).

$$FF=\frac{\text{Maximum Obtainable Power } (V_{max} \times I_{max})}{V_{oc} \times I_{sc}}$$

Fill Factor is a metric that gives a sense of how squared the I V curve is. The suitable value of FF typically varies between 0.5 and 0.82. The output label corresponds to the class of fault. After generating the dataset consisting of 200 data points for each of the six types of faults, we use Min-Max Normalization, which linearly transforms x to y = (x X-min)/(Xmax-Xmin), where Xmin and Xmax correspond to the smallest and the largest values in X where X denotes the set of observed values in x. Thus, Xmin is mapped to 0 and Xmax is mapped to

1. This means that the entire range of values of X from minimum to maximum is mapped to the range 0 to 1. For example, in our dataset, Pmax and Temperature are in very different ranges. When we do further analysis, Pmax will influence the prediction more due to its larger values. But this doesn't imply that it is more important as a feature for prediction. Thus, normalization of the databecomes important in the pre-processing stage.

RESULT AND CONCLUSION

Table : Various Test Fault Condition & its Predicted Output

Sr.no	Scenario	Actual output	Predicted output	
			RF	SVM
1	partial shading FF= Maximum fault	Fault	Fault	Fault
2	bypass diode fault	Fault	Fault	Fault
3	low temperature fault	Fault	Fault	Fault
4	high temperature fault	Fault	Fault	Fault
5	complete shading fault	Fault	Fault	Normal
6	short circuit fault	Fault	Fault	Fault
7	Partial shading with Temperature	Fault	Fault	Fault
8	Normal	Normal	Normal	Normal

8) PERFORMANCE EVALUATION FOR TESTING

Table : Performance Evaluation of Proposed System

Classifier	Events	Predicted Events		Accuracy
		Faulty	Normal	
RF	Faulty	11	1	93.33%
	Normal	0	3	
SVM	Faulty	10	1	86.66%
	Normal	1	3	

Table : Evaluation Time required for Training and Testing Phase

Sr.no	Method	Training Time	Testing Time
1	RF	0.72sec	0.12sec
2	SVM	1.16sec	0.15sec

CONCLUSION & FUTURE SCOPE

Fault detection and classification for six different faults have been studied and the generated dataset has been trained using four ML. models such as the Support Vector Machine, and Random Forest. All the four techniques give appreciable accuracies for fault classification. The best validation performance is attained by Random Forest techniques because the validation accuracy is 0.09 in terms of mean square error and the mod accuracy is 93.33%. The Random Forest algorithm works well in comparison to the other techniques used in this project as it gives the highest accuracy. The RF algorithm is thus ideally suited to detect faults primarily because it can handle complex data containing both linear and non-linear variations and converges to the optimal solution. In the future, this research work would be extended to validate these results using hardware implementation, and also catering to the power optimization aspect of PV array systems. This study and comparative analysis of the most widely used Machine Learning techniques would help researchers working in this field to inform their future research methodologies in the area classification.

REFERENCES

[1] J. Wiles, "Photovoltaic Power Systems And the 2005 National Electrical Code: Suggested Practices," Southwest Technology Development Institute, New Mexico State University November 26, 2008.

[2] S. E. Forman, "Performance of Experimental Terrestrial Photovoltaic Modules," Reliability, IEEE Transactions on, vol. R-31, pp. 235-245, 1982.

[3] E. L. Meyer and E. E. van Dyk, "Assessing the reliability and degradation of photovoltaic module performance parameters," Reliability, IEEE Transactions on, vol. 53, pp. 83-92, 2004.

[4] Y. Zhao, "Master of Science Thesis: Fault Analysis in Solar Photovoltaic Arays," Electrical and Computer Engineering, Northeastern University, Boston, 2010.

[5] Y. Zhao, B. Lehman, J.-F. d. Palma, J. Mosesian, and R. Lyons, "Fault Analysis in Solar PV Arrays under: Low Irradiance Conditions and Reverse Connections," in the 37th IEEE Photovoltaic Specialist Conference, Seattle, WA, 2011.

[6] Y. Zhao, B. Lehman, J.-F. d. Palma, J. Mosesian, and R. Lyons, "Challenges of Overcurrent Protection Devices in Photovoltaic Arrays Brought by Maximum Power Point Tracker," in the 37th IEEE Photovoltaic Specialist Conference, Seattle, WA, 2011.

[7] Y. Zhao, B. Lehman, J.-F. d. Palma, J. Mosesian, and R. Lyons, "Challenges to Overcurrent Protection Devices under Line-line Faults in Solar Photovoltaic Arrays," in the third IEEE Energy Conversion Congress & Exposition (ECCE), Phoenix, AZ, USA, 2011

[8] Il-Song K, "Fault detection algorithm of the photovoltaic system using wavelet transform," in

proceedings of the IEEE India international conference on power electronics, New Delhi, pp.1-6, 2010.

[9] Chine, W., Mellit, A.; Lughi, V. Malek, A.; Sulligoi, G.; Pavan, A.M. A novel fault diagnosis technique for photovoltaic systems based on artificial neural networks. Renew. Energy 2016, 90, 501-512.

[10] Mekki, H.; Mellit, A., Salhi, H. Artificial neural network-based modelling and fault detection of partial shaded photovoltaic modules Simul. Model. Pract. Theory 2016, 67, 1-13

[11] Dhimish, M: Holmes, V.; Mehrdadi, B.; Dales, M.; Mather, P. Photovoltaic fault detection algorithm based on theoretical curves modelling and fuzzy classification system. Energy 2017, 140, 276-290

[12] Dhimish, M.; Holmes, V. Fault detection algorithm for grid-connected photovoltaic plants. Sol Energy 2016, 137, 236-245.

[13] Kang, B.K; Kim, S.T.; Bae, S.H.; Park, J.W. Diagnosis of output power lowering in a PV amay by using the kalman-filter algorithm IEEE Trans. Energy Convers. 2012, 27, 885-894

[14] Harrou, F., Sun, Y.; Taghezouit, B.; Saidi, A.; Hamlati, M.E. Reliable fault detection and diagnosis of photovoltaic systems based on statistical monitoring approaches. Renew. Energy 2018, 116, 22-37.

[15] Silvestre, S., da Silva, M.A.; Chouder, A.; Guasch, D., Karatepe, E. New procedure for fault detection in grid connected PV systems based on the evaluation of current and voltage indicators. Energy Convers. Manag. 2014, 86, 241-249.

[16] Yahyaoui, L; Segatto, M.E. A practical technique for on-line monitoring of a photovoltaic plant connected to a single-phase grid. Energy Convers. Manag. 2017, 132, 198-206.

[17] Bonkoungou D, Koalaga Z, Njomo D. Modelling and simulation of photovoltaic module considering single diode equivalent circuit model in MATLAB. Int. J. Emerg. Technol. Adv. Eng. 2013;3:493-502.

[18] W Xiao, W. G. Dunford, and A. Capel, "A novel modeling method for photovoltaic cells," in Proc. IEEE 35th Annu Power Electron. Spec. Conf (PESC), 2004, vol. 3. pp. 1950-1956

[19] K. Ishaque, Z. Salam, H. Taheri, and Syafaruddin, "Modeling and simulation of photovoltaic (PV) system during partial shading based on a two-diode model," Simul. Modelling Pract. Theory, vol. 19, pp. 1613-1626, 2011

[20] Sunil Rao, Andreas Spanias, Cihan Tepedelenlioglu, "Solar Array Fault Detection using Neural Networks", IEEE International Conference on Industrial CyberPhysical Systems (ICPS), Taipei, May 2019