Solar Array Fault Detection using Neural Networks

Sunil Rao, Andreas Spanias and Cihan Tepedelenlioglu
{rsunilsr,spanias,cihan}@asu.edu
SenSIP Center, School of ECEE, Arizona State University.

***Abstract*— In this paper, we describe a Cyber-Physical system approach to fault detection in Photovoltaic (PV) arrays. More specifically, we explore customized neural network algorithms for fault detection from monitoring devices that sense data and actuate at each individual panel. We develop a framework for the use of feedforward neural networks for fault detection and identification. Our approach promises to improve efficiency by detecting and identifying eight different faults and commonly occurring conditions that affect power output in utility scale PV arrays.**

Keywords—Cyber-Physical Systems, Photovoltaics (PV), Fault Detection, Solar Energy, Machine Learning, Neural Networks.

I. Introduction

The efficiency of solar energy systems requires detailed analytics for each panel including voltage, current, temperature and irradiance. Solar power output is affected by factors such as cloud cover, soiling of modules, short circuits between panels, unexpected faults and varying weather conditions. In this paper we describe machine learning and neural network approaches for Fault Detection. These approaches are aimed at improving the reliability and efficiency of utility scale solar arrays. We describe theoretical, experimental, and implementation aspects of this comprehensive cyber-physical system (CPS) approach.

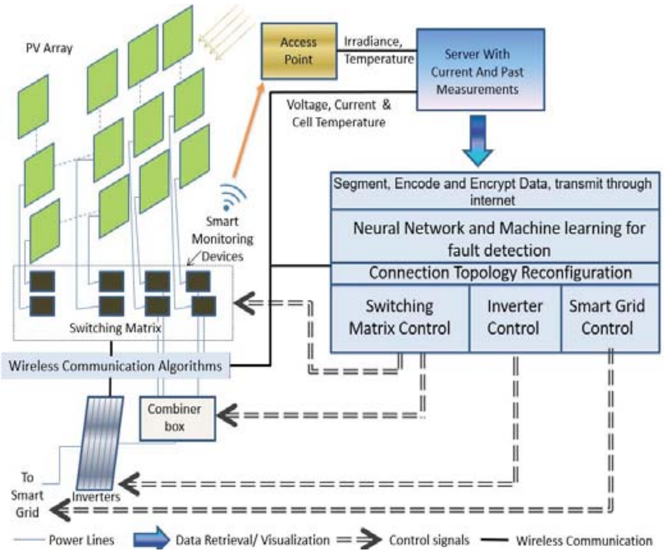

Figure 1. Overview of the CPS Solar System for monitoring.

We improve solar panel efficiency using machine learning techniques to learn and predict multiple system parameters using sensors and sensor fusion. Training and test data are acquired through cyber-physical methods including sensors and actuators. We also use machine learning and deep learning algorithms for fault detection which improve efficiency. We describe these in depth below. An overview of our research is shown in Figure 1.

Parameter sensing at each solar panel provides information for fault detection and power output optimization. Neural networks and sensor fusion enable us to implement robust shading estimation and fault detection algorithms. We have developed Smart Monitoring Devices (SMDs) with sensors that measure current, voltage and temperature. The data obtained from these sensors will be used for fault diagnosis in PV arrays. The SMDs also have relays that enable dynamical reconfiguration of connection topologies [1].

Previously, we have worked on signal processing algorithms for PV monitoring [1-4]. The methods presented in this paper will be implemented and validated on state-of-the-art PV array shown in Figure 2. This CPS system was developed by the SenSIP Center and involves an array of 104 panels.

Figure 2. The SenSIP PV experimental facility [1] used to validate our CPS algorithms for fault detection.

In our previous work [2-9,33], we have documented an efficiency improvement of 4%. We estimate an efficiency improvement of up to 10% with the use of custom clustering algorithms, customized sensor fusion and neural network algorithms for fault detection.

A utility-scale PV array consists of panels that are connected as a combination of series and parallel strings to maximize power output. Shading, weather patterns and temperature can severely affect power output. To minimize these effects, individual panel current-voltage (I-V) measurements and local weather information [11-14] are essential. We can control power output through matrix switching (i.e., real time topological changes with relay switches in each SMD [1]) of PV modules, allowing for several interconnection options. We optimize utility scale PV array systems by exploiting the measured I-V and weather data. Each SMD is connected to each PV panel that collects the individual panel metrics (current, voltage, and temperature) roughly every ten seconds.

The algorithms operate on PV array measurements.

978-1-5386-8500-6/19/$31.00 ©2019 IEEE

Parametric models are used to detect and identify faults. SMDs can perform panel switching or bypassing if necessary [1]. Reference [1] explains the use of relays in SMDs that operate on commands to reconfigure panel connection topologies.

In this paper, we propose the following:

i. Forming a unique set of custom features for fault detection and identification.

ii. Applying a customized neural network algorithm for fault classification in PV arrays.

iii. Detect and identify eight different commonly occurring PV fault cases (Section IV).

II. PROBLEM STATEMENT

Reliability is a critical factor for a PV system. Issues such as ground faults, arc faults, open circuits, short circuits, soiling, and partial shading can all reduce efficiency and need to be addressed. Some of these faults are undetected for a prolonged length of time in real-world situations. This leads to reduced and inefficient functioning of the PV array and a significantly lower power output. Unnoticed faults in PV can be dangerous and potentially life threatening. A real-world example would be the Bakersfield fire which was caused due to an undetected ground fault [32]. Although ground faults can now be detected with the use of inverters, faults such as soiling and short circuits between panels often go undetected [2].

The I-V data in a PV array can be measured at the panel-level inexpensively. I-V measurements have high correlation. This data can be used to build correlation models. Such models are useful in predicting possible ground faults, arc faults, soiling, shading, etc. [23]. The I-V curve is modeled using the single diode model shown in Figure 3, as a function of temperature, irradiance, open circuit voltage (V_{oc}) and short circuit current (I_{sc}). Each panel has a peak operating point known as the maximum power point (MPP). Fault detection using I-V data can be accomplished by measuring MPPs and observing the variation of the measured MPP from the actual MPP. Various parameter estimation methods for PV systems have been proposed in [36,37] and the references there in.

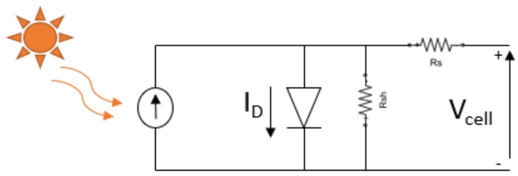

Figure 3. Typical circuit model for a PV module.

A. Existing Approaches for Fault Detection

Fault detection in PV arrays have been studied extensively. While [18-22] use statistical methods to identify faults, [10, 28-30] use graph and neural network architectures to detect and identify faults. However, to our knowledge, none of these techniques addresses specifically precise identification of the type of fault. There is a need for a holistic method which can detect and identify and localize faults. In [6-8], we show how to detect faults in individual modules using clustering algorithms. In this paper, we develop an algorithm to identify and localize underperforming modules using neural network architectures.

B. Using Machine Learning for Fault Detection

Human operators are currently required to manually perform fault detection and identification. Studies have [22-24] showed that the current method for mean time to repair (MTTR) is at approximately 19 days. There is a significant need to reduce MTTR to reduce power losses from the PV array. We use machine learning methods to reduce the MTTR for PV arrays.

Fault identification and localization problems pose several challenges and research opportunities. A system must first accurately classify the PV array condition and then react to unseen data to correctly classify the condition of operation of the PV array. Considering these challenges, we explore the use of machine learning techniques [15]. Semi-supervised learning can be used to label many realistic faults from few measured examples.

III. MACHINE LEARNING RESULTS

Signal and image processing applications use machine learning algorithms in several applications [27, 34-35]. The utility of machine learning tools has also been shown in various Internet-of-things (IoT) applications [16-17]. In the following section, we explore the use of clustering algorithms and the need for neural networks in solar energy systems.

A. Results using K-means algorithm:

The K-means algorithm is a clustering-based approach in machine learning that can be used for fault detection. Given a dataset, K-means clustering partitions n observations into k clusters. Each observation belongs to the cluster with the nearest mean. The mean serves as a representative of the cluster. Simulated data to generate MPPs was obtained using MATLAB's Simulink model as shown in Figure 4. The K-means algorithm was applied to simulated data as shown below. While generating MPPs, we consider a variance of ±5V for V_{mp} and a variance of ±1A for I_{mp} to account for variability in real time scenarios [29].

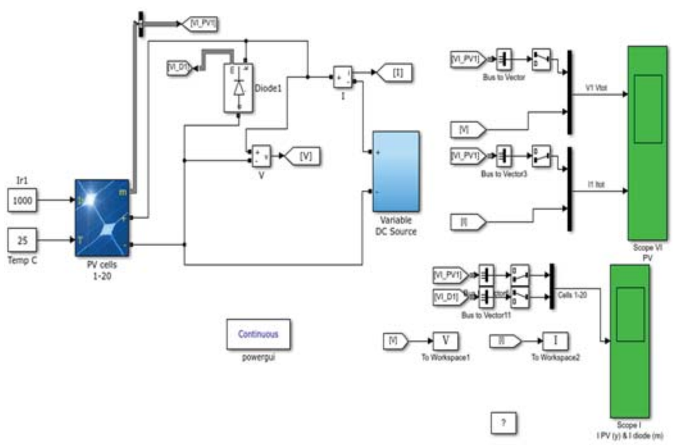

Figure 4: Simulink Model used to create the dataset for the eight different types of faults considered.

To simulate a varying temperature panel, the simulated panel was assigned a higher temperature value. The data was obtained and trained with the K-means algorithm.

The results obtained are shown in Figure 5. Each set of data points represent one condition associated with the PV Array. Using K-means with voltage, current and temperature as our three axes, we successfully identify ground faults (Gnd), arc faults (Arc), standard test conditions with irradiance at 1000 W/m^2 and a module temperature of 25ºC (STC), shaded conditions (Shading) and varying temperature conditions (Varying Temp).

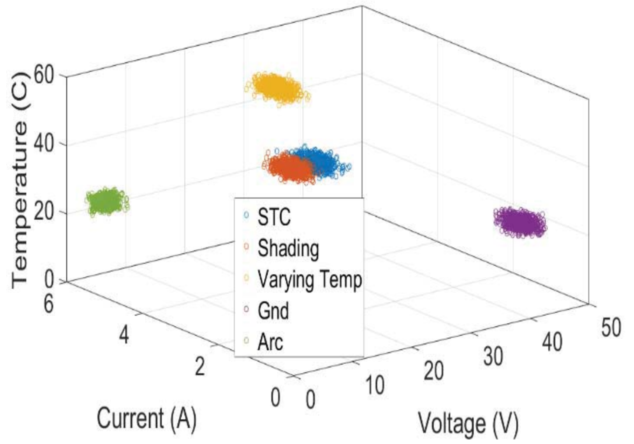

Figure 5: Clustering using the K-means algorithm. Training forms clusters of normal and simulated faults for PV data.

However, certain other conditions such as soiling and short circuits are not identified using this method due to the lack of labels in the dataset. Soiling and short circuits conditions have MPPs which lie in similar areas in the two dimensional I-V curve space. The K-means algorithm also does not identify partial shading versus complete shading of modules. The differentiator for these cases is described in the next section. There is a need for the use of neural network algorithms to detect and identify faults in PV arrays.

IV. Use of Neural Nets for Fault Detection

Various signal processing and statistical methods have been developed for detection and identification of faults in utility scale PV arrays. However, there is a need for a comprehensive algorithm which captures a wide variety of faults. In our data simulation, we use the model described earlier in Figure 4. While several methods have been proposed in the past for fault detection, this method aims to detect and identify the type of fault occurring in PV arrays. We emphasize that, successful classification of faults may lead to a significant reduction of the meantime to repair (MTTR) in utility scale PV arrays [18-21].

Figure 6 shows the I-V curve for the multi-class classification problem we have considered in this paper. While traditional signal processing algorithms use the statistical properties of a single I-V-curve of a given module, most methods do not cover multiple cases. To do this, we used neural networks for fault detection and identification. With the use of unsupervised machine learning algorithms, a fault could be detected but not identified as discussed previously. In section III, we demonstrated that unsupervised algorithms could not classify the type of fault (ground fault, arc fault, shading, etc.). We need an algorithm which uses partially labelled data to classify unlabeled data. Using neural networks allows not only detection but identification of the fault type with a high accuracy. Previous studies that used neural nets have been used to make binary decisions on fault detection, i.e., detect faults but not classify the type of faults [28-31]. In our study, we identify eight different cases of faults and shading effects.

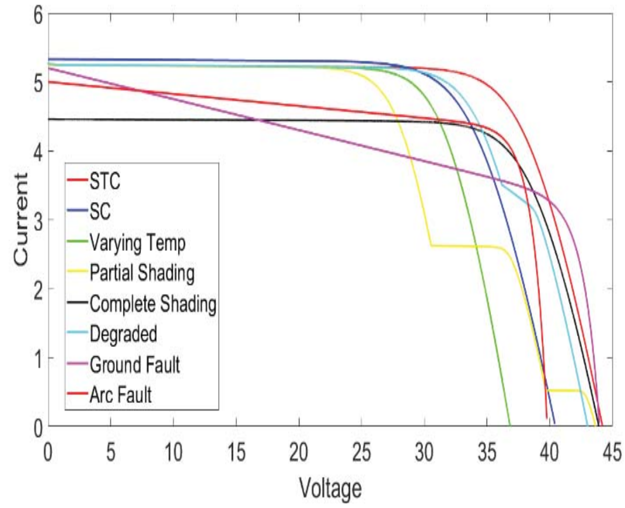

Figure 6: I-V curves for the cases considered.

To do this, we use a multi-layer feedforward neural network with multiple inputs as features. A set of unique features is selected as inputs to the neural network and are critical in identifying the type of fault.

The maximum voltage (V_{mp}) and maximum current (I_{mp}) lie at the knee of the I-V curve. These two features help identify the power produced by the PV array. We chose power as a third feature to help classify shading.

The next set of features include irradiance and temperature. Irradiance and temperature are critical features which help identify shading conditions from varying temperature conditions. V_{mp} and I_{mp} for shading and varying temperature conditions lie at similar points along the I-V curve, making it difficult to classify the two cases. With these two critical features, along with those previously mentioned, we can separate shading from temperature conditions.

We considered other features such as Gamma (γ) - the ratio of power over irradiance, and Fill Factor - a ratio of the product of the short circuit current (I_{sc}) and open circuit voltage (V_{oc}) over product of V_{mp} and I_{mp}. These two features capture the area of the I-V curve along different dimensions which help classify multiple shading conditions. Multiple shading conditions include partial shading versus complete shading of the module.

Additionally, we considered features V_{oc} and I_{sc} which helps in classifying shading versus soiling. Shading and soiling often have overlapping data points and hence it is difficult to identify one versus the other. However, the difference between the two is captured in open circuit voltage and short circuit current causing these two features to serve as distinguishing parameters to identify shading versus soiling.

In our most recent neural network algorithm, we use nine inputs namely V_{oc}, I_{sc}, V_{mp}, I_{mp}, temperature of module, irradiance of module, fill factor, gamma and power, to classify eight different faults. The eight faults classified are ground fault (Gnd), arc fault (Arc), complete module shading (Fully Shaded), partial module shading (Partial Shading), varying temperatures of module (Varying Temp), soiling (Degraded), short circuits (SC) and standard test conditions with irradiance at 1000 W/m^2 with module temperature of 25ºC (STC). Research at this stage also involves obtaining data in real time. Figure 7 gives an overview of describing the process for real time scenarios.

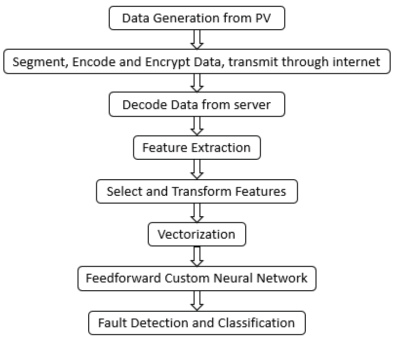

Figure 7: Block diagram describing the process of fault detection and identification for a real time scenario.

Using the features mentioned, we apply them as inputs to a multilayer feedforward neural network, popularly called as the multilayer perceptron (MLP). We use a 5 layered neural network with backpropagation to optimize the weights used in each layer. Each layer uses 6 neurons.

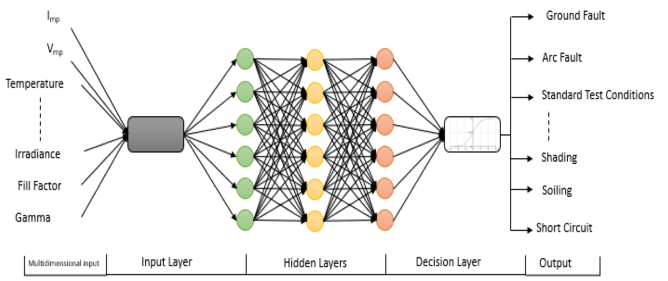

Figure 8: Neural Network Architecture used for Fault Detection and Classification.

Information flows through the neural networks in two ways: (i) In *forward propagation* the MLP model predicts the output for the given data and (ii) In *backpropagation* the model adjusts its parameters considering the error in the prediction. The *activation* function used in each neuron allows the MLP to learn a complex function mapping. The MLP architecture used for Fault Classification is shown in Figure 8. Input to the model is the feature vector $\boldsymbol{x}$, the output of the first and consecutive hidden layer is given by

$$\boldsymbol{h}_1 = \sigma(\boldsymbol{W}_1 \cdot \boldsymbol{x} + \boldsymbol{b}_1) \quad (1)$$

$$\boldsymbol{h}_i = \sigma(\boldsymbol{W}_i \cdot \boldsymbol{h}_{i-1} + \boldsymbol{b}_i) \quad (2)$$

Where i is the layer index and σ is the activation function. $\boldsymbol{x}$, has a dimension of 48000x9. Each column represents a feature of the neural network mentioned earlier. The output of the MLP is obtained as:

$$\hat{y} = \phi_{softmax}\,(\boldsymbol{h}_{out}) \quad (3)$$

Weights of each neuron are trained using a scaled gradient backpropagation algorithm. Each layer is assigned a *tanh* (hyperbolic tangent) activation function. From our experiments, we see that the *tanh* decision boundary gives the best accuracy. The output layer uses the SoftMax activation function to categorize the type of fault in the PV array.

We simulate each fault type versus shading versus standard conditions so as to have the same number of datapoints and avoid bias in the training of the neural network. For the training of the neural network, we use 70% of labelled data for training, 15% of data for validation and the remaining 15% data as a test dataset, allowing the algorithm to classify the "unknown" testing datapoints. The results of the algorithm are shown in the form of a confusion matrix in Figure 9. We obtained an accuracy of over 99% for noiseless measurements. More experiments are underway to explore the use of neural networks for real time measurements including noisy data. This is a significant improvement from previous fault detection and identification methods.

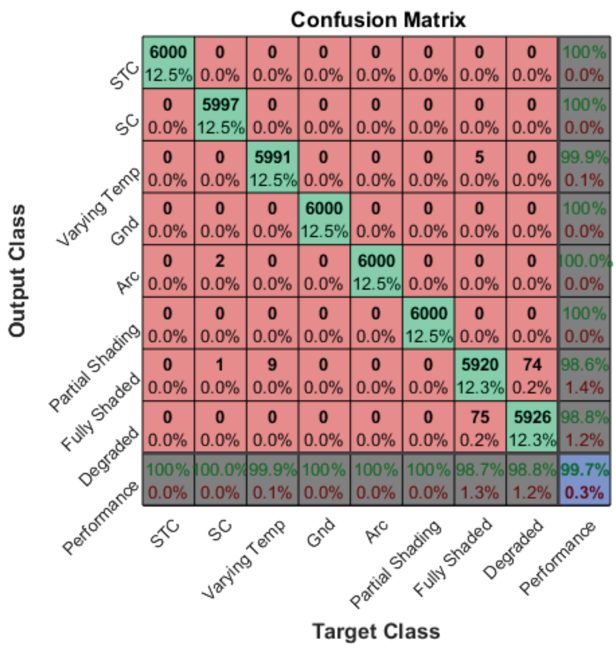

Figure 9: Confusion matrix for fault identification.

V. Conclusion

We address the problem of PV array monitoring and control using advanced neural network algorithms. We propose the use of neural networks for real time monitoring of PV arrays. We consider nine input features for the neural network to identify faults in PV arrays. Simulation results using neural networks demonstrated successfully detecting and identifying commonly occurring faults and shading conditions including soiling, short circuits, ground faults, and partial shading in utility scale PV arrays. We show a significant improvement in accuracy of detection and identification of faults compared to traditional and existing methods using noiseless synthetic data. Experiments are underway for fault detection using real-time data including examination of the algorithms using noisy observations.

ACKNOWLEDGEMENT

This research is supported in part by the NSF CPS Award number 1646542 and by the ASU SenSIP center.

REFERENCES

[1] A. Spanias, C. Tepedelenlioglu, E. Kyriakides, D. Ramirez, S. Rao, H. Braun, J. Lee, D. Srinivasan, J. Frye, S. Koizumi, Y. Morimoto, "An 18 kW Solar Array Research Facility for Fault Detection Experiments," *Proc. 18th MELECON,* April 2016.

[2] H. Braun, S. T. Buddha, V. Krishnan, A. Spanias, C. Tepedelenlioglu, T. Takehara, S. Takada, T. Yeider, and M. Banavar, Signal Processing for Solar Array Monitoring, Fault Detection, and Optimization, Synthesis Lect. Power Electronics, J. Hudgins, Ed. Morgan & Claypool, vol. 3, Sep. 2012.

[3] Braun, H.; Buddha, S.T.; Krishnan, V.; Spanias, A.; Tepedelenlioglu, C.; Yeider, T.; Takehara, T.; "Signal processing for fault detection in photovoltaic arrays," *2012 IEEE International Conference on, Acoustics, Speech and Signal Processing (ICASSP),* pp.1681-1684, 25-March 2012.

[4] Buddha, S.; Braun, H.; Krishnan, V.; Tepedelenlioglu, C.; Spanias, A.; Yeider, T.; Takehara, T.; "*Signal processing for photovoltaic applications*," *IEEE ESPA*, vol., no., pp.115-118, 12-14 Jan. 2012.

[5] H. Braun, S. T. Buddha, V. Krishnan, C. Tepedelenlioglu, A. Spanias, M. Banavar, and D. Srinivasan, "Topology reconfiguration for optimization of photovoltaic array output," *Elsevier Sustainable Energy, Grids and Networks* (*SEGAN),* pp. 58-69, Vol. 6, June 2016.

[6] S. Rao, S. Katoch, P. Turaga, A. Spanias, C. Tepedelenlioglu, R. Ayyanar, H. Braun, J. Lee, U. Shanthamallu, M. Banavar, and D. Srinivasan, "A Cyber-Physical System Approach for Photovoltaic Array Monitoring and Control," in Proc. IEEE IISA, Larnaca, Cyprus, 2017.

[7] S. Katoch, G. Muniraju, S. Rao, A. Spanias, P. Turaga, C. Tepedelenlioglu, M. Banavar, D. Srinivasan, "Shading Prediction, Fault Detection, and Consensus Estimation for Solar Array Control*", 1st IEEE International Conference on Industrial Cyber-Physical Systems (ICPS-2018)*, Saint. Petersburg, Russia, May 2018.

[8] G. Muniraju, S. Rao, S. Katoch, A. Spanias, P. Turaga, C. Tepedelenlioglu, M. Banavar, D. Srinivasan, "A Cyber-Physical Photovoltaic Array Monitoring and Control System", *International Journal of Monitoring and Surveillance Technologies Research*, vol., issue 3, May 2018.

[9] Farib Khondoker, Sunil Rao, Andreas Spanias, Cihan Tepedelenlioglu, "Photovoltaic Array Simulation and Fault Prediction via Multilayer Perceptron Models", *Proc. IEEE IISA-2018*, Zakynthos, July 2018.

[10] Y. Zhao, R. Ball, J. Mosesian, J.-F. de Palma, and B. Lehman, "Graphbased semi-supervised learning for fault detection and classification in solar photovoltaic arrays," *IEEE Transactions on Power Electronics*, vol. 30, no. 5, pp. 2848–2858, 2015.

[11] D. Nguyen and B. Lehman, "Modeling and simulation of solar pv arrays under changing illumination conditions," in Computers in Power Electronics, COMPEL '06. IEEE Workshops on, pp. 295 –299, 2006.

[12] V. Quaschning and R. Hanitsch, "Numerical simulation of current-voltage characteristics of photovoltaic systems with shaded solar cells," *Solar Energy,* vol. 56, no. 6, 1996.

[13] G.D. Gregory and G.W. Scott, "The arc-fault circuit interrupter: an emerging product," *Industry Applications, IEEE Transactions on*, vol. 34, no. 5, pp. 928-933, 1998.

[14] M.G. Villalva, J.R. Gazoli, and E.R. Filho, "Comprehensive Approach to Modeling and Simulation of Photovoltaic Arrays," *IEEE Transactions on Power Electronics,* vol. 24, no. 5, pp. 1198-1208, 2009.

[15] U. Shanthamallu, A. Spanias, C. Tepedelenlioglu, M. Stanley, "A Brief Survey of Machine Learning Methods and their Sensor and IoT Applications," *Proceedings 8th International Conference on Information, Intelligence, Systems and Applications (IEEE IISA 2017)*, Larnaca, August 2017.

[16] A. Spanias, "Solar Energy Management as an Internet of Things (IoT) Application," *Proceedings 8th International Conference on Information, Intelligence, Systems and Applications (IEEE IISA 2017)*, Larnaca, August 2017.

[17] G. Du et al, "Based on the Internet of things a self-cleaning solar power system of the household micro-grid," *2016 Chinese Control and Decision Conference (CCDC)*, May 2016.

[18] H. Haeberlin and M. Kaempfer, "Measurement of Damages at Bypass Diodes by Induced Voltages and Currents in PV Modules Caused by Nearby Lightning Currents with Standard Waveform," in *23rd European Photovoltaic Solar Energy Conference*, 2008.

[19] N. Bosco, "Reliability concerns associated with PV technologies," *National Renewable Energy Laboratory*, Albuquerque, 2010.

[20] L. L. Jiang and D. L. Maskell, "Automatic fault detection and diagnosis for photovoltaic systems using combined artificial neural network and analytical based methods," 2015 *International Joint Conference on Neural Networks (IJCNN), Killarney*, pp. 1-8, 2015.

[21] M. N. Akram and S. Lotfifard, "Modeling and Health Monitoring of DC Side of Photovoltaic Array," in *IEEE Transactions on Sustainable Energy*, vol. 6, no. 4, pp. 1245-1253, Oct. 2015.

[22] A. Maish, C. Atcitty, S. Hester, D. Greenberg, D. Osborn, D. Collier, and M. Brine, "Photovoltaic system reliability," *Photovoltaic Specialists Conf. Conf. Rec 26th IEEE*, 29, pp. 1049 –1054, 1997.

[23] D.L. King, J.A. Kratochvil, and W.E. Boyson, Photovoltaic array performance model, United States, Dept. of Energy, 2004.

[24] P. J. Rousseeuw and K. v. Driessen, "A fast algorithm for the minimum covariance determinant estimator," *Technometrics*, pp. 212–223, 1999.

[25] Sunil Rao, Andreas Spanias, Cihan Tepedelenlioglu, "Machine Learning and Neural Nets for Solar Array Fault Detection," Patent Pre-disclosure, M19-102P, Provisional Patent submitted, Skysong Innovations, 2019.

[26] Vivek Sivaraman Narayanaswamy, Raja, Ayyanar, Andreas Spanias, Cihan Tepedelenlioglu, "Systems And Methods For Connection Topology Optimization In Photovoltaic Arrays Using Neural Networks," Provisional Patent US 62/808,677, 2019.

[27] A. Spanias, T. Painter, V. Atti, Audio Signal Processing and Coding, ISBN: 9780471791478, Wiley, March 2007.

[28] Chine W, Mellit A, Lughi V, Malek A, Sulligoi G, Pavan AM, "A novel fault diagnosis technique for photovoltaic systems based on artificial neural networks", *Renew Energy*, 2016.

[29] Mekki, H., Mellit, A., Salhi, H., "Artificial neural network-based modelling and fault detection of partial shaded photovoltaic modules". *Simul. Model. Pract. Theory* 67, 1–3., 2016.

[30] Zhicong Chen, Lijun Wu, Shuying Chen, Peijie Lin, Yue Wu, Wencheng Lin, "Intelligent fault diagnosis of photovoltaic arrays based on optimized kernel extreme learning machine and I-V characteristics", in *Applied Energy*, Elsevier, May 2017.

[31] R. Hariharan, M. Chakkarapani, G. Saravana Ilango, and C. Nagamani, "A Method to Detect Photovoltaic Array Faults and Partial Shading in PV Systems", *IEEE Journal of Photovoltaics,* Vol 6. No 5. Sept. 2016.

[32] Bill Brooks, "The Bakersfield Fire: A Lesson in Ground-Fault Protection", *Solar Pro*, Issue 4.2 March 2011.

[33] Vivek Narayanaswamy, Raja Ayyanar, Andreas Spanias, Cihan Tepedelenlioglu and Devarajan Srinivasan, "Connection Topology Optimization in PV Arrays using Neural Networks'," *IEEE International Conference on Industrial Cyber-Physical Systems (ICPS*), Taipei, May 2019.

[34] H. Song, J. Thiagarajan, P. Sattigeri, A. Spanias, "Optimizing Kernel Machines using Deep Learning" Accepted *IEEE Transactions on Neural Networks and Learning Systems,* NLS-2017-P-8053.R1, pp. 5528 – 5540, Feb. 2018.

[35] McCann, M. T., Jin, K. H., & Unser, M. (2017). Convolutional neural networks for inverse problems in imaging: A review.*IEEE Signal Processing Magazine*, *34*(6), 85-95

[36] R. Ramakrishna, A. Scaglione, A. Spanias, C. Tepedelenlioglu, "Distributed Bayesian Estimation With Low-Rank Data: Application To Solar Array Processing," *IEEE ICASSP 2019*, Brighton, UK, May 2019.

[37] Oliva, D., El Aziz, M. A., & Hassanien, A. E. (2017). Parameter estimation of photovoltaic cells using an improved chaotic whale optimization algorithm. *Applied energy*, *200*, 141-154.