Electronics and Electrical Engineering

AIMS Electronics and Electrical Engineering, 4(1): 1–18.
DOI: 10.3934/ElectrEng.2020.1.1
Received: 09 October 2019
Accepted: 15 November 2019
Published: 26 December 2019

http://www.aimspress.com/journal/ElectrEng

Research article

Deployment of AI-based RBF network for photovoltaics fault detection procedure

Muhammad Hussain, Mahmoud Dhimish*, Violeta Holmes and Peter Mather

Department of Engineering and Technology, Laboratory of Photovoltaics, University of Huddersfield, Huddersfield, HD1 3DH, United Kingdom

* **Correspondence:** Email: mahmoud.dhimish@gmail.com; Tel: +44 1484 471236.

Abstract: In this paper, a fault detection algorithm for photovoltaic systems based on artificial neural networks (ANN) is proposed. Although, a rich amount of research is available in the field of PV fault detection using ANN, this paper presents a novel methodology based on only two inputs for the training, validating and testing of the Radial Basis Function (RBF) network achieving unprecedented detection accuracy of 98.1%. The proposed methodology goes beyond data normalisation and implements a 'mapping of inputs' approach to the data set before exposing it to the network for training. The accuracy of the proposed network is further endorsed through testing of the network in partial shading and overcast conditions.

Keywords: renewable energy; photovoltaics; fault detection; artificial intelligence; RBF network

1. Introduction

The introduction or rather re-emergence of AI has garnered interest from almost every industry. Businesses are looking at ways of implementing AI into their products/solution and a vast number of products implementing AI can be found on the present market.

Grid-Connected Photovoltaic (GCPV) systems are becoming increasingly popular for enhanced energy harvesting and reliable power production. Researchers in the field of PV systems are looking at non-conventional methods for accurate monitoring, fault detection and isolation of components in PV installations. Many research papers can be found on the topic of PV fault detection using various ANN networks.

Traditionally a distributed sensor network (DSN) would be required for the monitoring of a PV installation. This network may consist of several variables such as voltage, current, irradiance, wind speed and temperature [1,2]. The hardware cost associated with the monitoring of the above parameters through a DSN can deter enterprises from implementing fault detection in their PV systems. This barrier has been eliminated by the introduction of 'smart meters' which can provide all of the key parameters in one platform. However, in the context of ANN, this advancement does not have any impact on the efficiency of the network. The reason for this is because the accuracy of an ANN network depends on the 'quality' of data processing which takes places before exposing the network to the data set.

In order to train an ANN network to provide a high degree of accuracy in-terms of fault detection, the features mentioned above would need to go through an intense processing stage where the data sample would need to be 'cleaned' before it can effectively be used as a training sample set. Knowing the impact training data has on the overall network, researchers are looking at reducing the number of inputs required for their ANN networks.

Presently, researchers have focused their research on reducing the sample data set required for training an ANN to maximise the performance of the network. Another key factor which plays an important role in the overall accuracy is the selection of the ANN. MLP networks are very commonly used for a wide range of applications due to the option of increasing the number of hidden layers. However, by increasing the number of hidden layers in a network the overall computation time will also rise. This paper provides an in-detail analysis of MLP vs RBF with regards to the rationale for selecting an RBF network, in section 3.3.

The detection of faults in PV installations can be split into three distinct classes; visual, thermal and electrical [3]. The latter class consists of sub-sets, the first of which is methods where no climate data (irradiance, power, temperature) is required. An example of this approach is the Time-Domain Reflectometry (TDR) proposed in [4] for detection of disconnection of a PV string.

The second approach consists of methods based on the analysis of the current and voltage characteristics. S. Silvestre et al. [5] calculates Series Resistance (R_s), Fill Factor (FF) and Shunt Resistance (R_{sh}) based on the I-V characteristics leading on to performance indicators.

Another approach is based on the Maximum Power Point Tracking (MPPT). X. Li et al. [6] Suggests an automated fault detection method built on power loss analysis, leading to the identification of faults in PV module, string and faults linked to varying environment conditions such as partial shading and ageing of MPPT device.

The last method is based on Artificial Intelligence (AI) techniques. Author's in [7] analyse the effectiveness of BP neural network in comparison with Fuzzy Logic for PV fault detection. Concluding BP neural networks as more superior in PV fault detection.

Artificial neural networks (ANNs) are inspired by the human brain. Depending on the type of learning, these networks can learn from datasets, by past-experience and through the generalization of past behaviors as characteristics [8]. Consisting of an input, hidden (or many hidden layers) and an output layer, the ANN's ability to process information in a non-linear, highly-parallelism and noisy environments makes it of huge interests to researchers in many fields [9–11]. As early as 1987, A. Lapedes et al. [12] findings show backpropagation neural networks performance is significantly higher than any of the conventional linear and polynomial methods, implemented for time series data.

F. Polo et al. [13] presents ANN based models for failure mode prediction and energy harvesting in PV systems to support dynamic maintenance tasks. The author aims to remove false positive prediction of faults in a PV system by processing the data going into the network. The paper focuses

on a back-propagation network, trained on inverter data consisting of 5 years. The author categorizes the faults as equipment deterioration and useful life reduction based on geographical and operational features. The author in his concluding remarks talks about further improving the proposed methodology through a larger data for training the ANN.

Sun Yougang et al. [14] proposed an Adaptive sliding mode control of maglev system based on RBF minimum parameter learning method. The author appreciates that the convergence speed and real-time performance of RBF networks limit its implementation in the field of maglev systems. Therefore, the authors implement the RBF minimum parameter learning method. The proposed methodology containing the RBF minimum parameter learning method ensured better real-time dealing with uncertainty and disturbance. The research also highlights the limits of RBF networks and how they can be implemented with other methods to provide an effective solution for applications where real-time and fast system response if vital.

Yasuhiro Yagi et al. [15] looks at the identification of specific faults (shading effect and inverters failure) based on an expert system. A key feature of this method is its reliance on a simple and reprogrammable ANN network. However, the proposed technique is unable to identify faulty conditions occurring in PV systems such as PV short-circuit failure conditions and PV String failure. Conversely, [16] demonstrates an ANN based fault detection system for identifying faults such as faulty bypass diodes, PV modules, and faulty PV string. The paper implements two ANN networks, MLP and RBF respectively. Based on the results presented, the accuracy of the MLP (90.3%) network is far greater than that of the RBF (68.4%). MLP networks in general are used more commonly due to higher accuracy using more hidden layers.

However, MLP networks in exchange for greater accuracy, demand higher computational time. It is also important to note that the data set consisted of a modest sample set (775). In addition, the data set is generated using MATLAB and Simulink. Conversely, the data sample used in our proposal consists of 97200 samples over a 10-week period, obtained from a live installation.

The main contribution of this work is the proposal of a novel methodology based two inputs providing unprecedented accuracy in fault detection of a PV system. The selected ANN network is RBF rather than commonly used MLP networks. The rationale for the implementation of RBF over (MLP) was due to the requirement of only one hidden layer but more importantly due to RBF's significantly lower computational time discussed in detail further in the paper. As shown in section 3.1, the accuracy of the ANN with varying number of hidden layers was evaluated. According to the results an ANN architecture consisting of only a single hidden layer was the most optimal solution providing an accuracy of around 99% whilst keeping the computational time to a minimum. This reinforces the selection criteria for the ANN network was set with an emphasis to use the most effective network both in terms of accuracy and computational requirements.

The novelty of the methodology was based on obtaining the right step-size for the sample set before training the RBF network on the data. This part is the nucleus of the whole training process as it substantially impacts the results generated by the network upon completion of training and testing. The selection of a step size that could not represent the whole of the dataset would result in under-fitting whereas a step size which was considered as highly rich in representing the sample set would lead to over-fitting, further discussed in section 3.1.

As the RBF network only requires a single hidden layer, the main task involved in the setup of the network architecture was the selection of the number of neurons present within the hidden layer. This selection would have a direct impact on the overall results provided by the network in terms of

accuracy. A simulation was undertaken to find the optimal number of neurons to be added to the hidden layer, later will be discussed in section 3.

Rest of the article is organized as follows, section 2 presents the examined PV installation. In section 3 we discuss in detail the rationale for selecting RBF over MLP and look at the structure of the proposed network along with the proposed methodology to detect faults in the PV system. Section 4 looks at the results obtained from testing the proposed methodology. In section 5, we compare our developed ANN network with recent ANN-based models available in the literature. Finally, sections 6 and 7 present the conclusion and reference list, respectively.

2. Examined PV system

The PV system under consideration is shown in Figure 1, consisting of 10 PV modules (string topology), irradiance sensor, MPPT unit and DC- load. The MPPT unit is connected via Ethernet capable to a computer (PC) providing real-time data monitoring. The proposed ANN algorithm for fault detection of the PV modules is developed in MATLAB software.

The system consists of 10-polycrystalline silicon PV modules, with a nominal power of 220 Wp (per module). Electrical parameters for the PV modules under 'standard test conditions' (STC) are shown in Table 1; STC of the PV modules at solar irradiance = 1000 W/m2, module temperature = 25 °C. The Maximum Power Point Tracker (MPPT) has an output efficiency of not less than 95.0%. DC current and voltage is measured via internal sensors located within the MPPT.

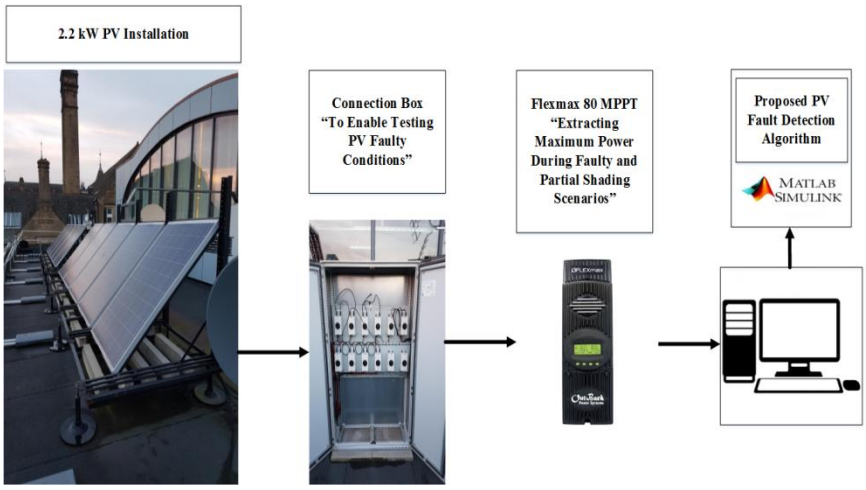


Figure 1. Overall system architecture design for the examined PV plant.

Table 1. Electrical characteristics of SMT6 (60) P PV module.

Solar Panel Electrical Characteristics	Value
Peak Power	220 W
Voltage at maximum power point (V_{mpp})	28.7 V
Current at maximum power point (I_{mpp})	7.67 A
Open Circuit Voltage (V_{oc})	36.74 V
Short Circuit Current (I_{sc})	8.24 A

Solar Irradiance is obtained through the Davis Weather Station and passed onto the monitoring unit connected to the PC for data recording and monitoring. A Hub 4 communication manager enables acquisition of modules temperature via the Davis external temperature sensor, as well as the electrical data for each photovoltaic string.

3. Methodology

Although many research papers can be found on various methodologies proposed by researchers for data normalisation, ANN training, validation and testing, to the best of our understanding, all of the methodologies found in current literature are based on data set, compromising of several input variables [17–19]. Our methodology goes a step further after data normalisation by implementing a 'mapping of inputs' concept, based on only two input variables.

3.1. ANN structure

The importance of data processing also known as 'data cleaning' plays an essential part in the development of any successful ANN network. Understanding the importance of providing quality data for the training of the network, the authors have focused their methodology on the processing of raw data set obtained directly from the PV installation shown in Figure 1.

The proposed methodology is based on 3 pre-requisites which need to be performed on the acquired data set before exposing it to the RBF network. These steps are;

1. Removal of all non-representative data from the data set.
2. The implementation of the max-min data normalisation technique.
3. The mapping of the inputs with an appropriate step size to cater for missing data points.

The first point (removal of non-representative data) is seen as standard practice in the field of machine learning. The inclusion of redundant data has an adverse effect on the training stage of the neural network having a knock-on effect on the overall performance of the network. The reason for this is because neural networks are heavily dependent on the quality of data they are being trained on, holding true to the computer science theory of 'garbage in, garbage out'.

The second point (data normalisation) is not a mandatory procedure which needs to be applied to all data sets. The need for data normalisation occurs when there is significant divergence in the inputs. This divergence was clearly visible in the values for the solar irradiance and output power, hence data normalisation was applied. The input features were normalized by implementing the max-min normalisation technique, in the range of 0 and +1 using (1).

$$y = \frac{(y_{max} - y_{min})(x - x_{min})}{(x_{max} - x_{min})} + y_{min} \quad (1)$$

where $\in \{x_{min}, x_{max}\}, y \in \{y_{min}, y_{max}\}$ and x is the original data value and y is the corresponding normalized value with $y_{min} = 0$ and $y_{max} = +1$.

Figure 2 shows the architecture of the ANN based on the Radial Basis Function (RBF). An RBF network belongs to the feedforward family of ANN consisting of three fundamental layers (input, hidden layer and output). The network can consume a vast amount of input data before transmitting it through its hidden layer. There are various activation functions which can be implemented in an RBF network but the most common is Gaussian Function [20]. For this research, a single hidden

layer was implemented, for all investigative methods, rationale provided in section 3.3.

Zhitao Zhao et al. [21] used an RBF network for Prediction of Interfacial Interactions related with Membrane Fouling in a Membrane Reactor. Although the field of study is completely different compared to our study but a common point which is given as the rationale for the selection of the RBF network is its effectiveness in the required computational needs to fulfill its objective. The author describes the network as one which can directly output the results from the input, significantly reducing the computational burden.

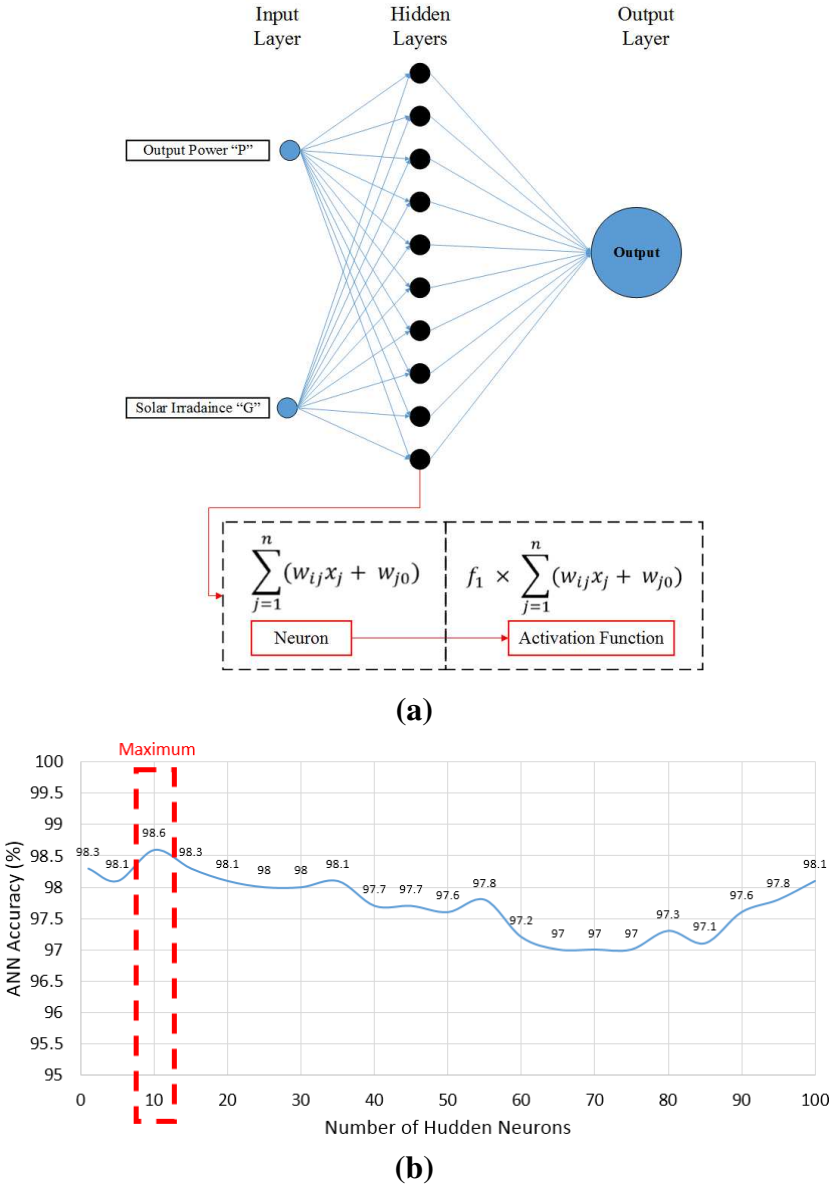


Figure 2. (a) Details of the proposed ANN network architecture, (b) ANN accuracy vs. number of hidden neurons.

The developed structure of the ANN network is shown in Figure 2(a). The RBF network with two inputs and one hidden layer containing 10 hidden neurons was selected. In fact, the selection of the inputs was obtained using the available parameters from dataset, including the solar irradiance (G) and the output power (P). The selection of the hidden layers was attained using an extensive

simulation from 1 to 100 hidden layers, as a result ten hidden layers were selected due to its optimum performance. The results of the ANN network accuracy vs. the number of hidden neurons, is shown in Figure 2(b).

3.2. ANN training and validation

For this research, 10 different conditions have been taken into consideration, including:

- Case1: Normal operation mode, where no faults were applied to the PV string
- Case 2: 1 Fault applied to the system, 1 PV module disconnected form the PV string
- Case 3: 2 Faults applied to the system; 2 PV modules disconnected form the PV string
- Case 4: 3 Faults applied to the system; 3 PV modules disconnected form the PV string
- Case 5: 4 Faults applied to the system; 4 PV modules disconnected form the PV string
- Case 6: 5 Faults applied to the system; 5 PV modules disconnected form the PV string
- Case 7: 6 Faults applied to the system; 6 PV modules disconnected form the PV string
- Case 8: 7 Faults applied to the system; 7 PV modules disconnected form the PV string
- Case 9: 8 Faults applied to the system; 8 PV modules disconnected form the PV string
- Case 10: 9 Faults applied to the system; 9 PV modules disconnected form the PV string

A flowchart of the proposed fault detection architecture is shown in Figure 3. The measured output power of the PV string was obtained via MPPT unit. If the output power was above zero, the measured power was passed into the developed RBF network. Conversely, if the output was equal to zero, the measured voltage was verified to decide if the PV string was faulty (voltage > 0), or in sleep mode (Voltage = 0).

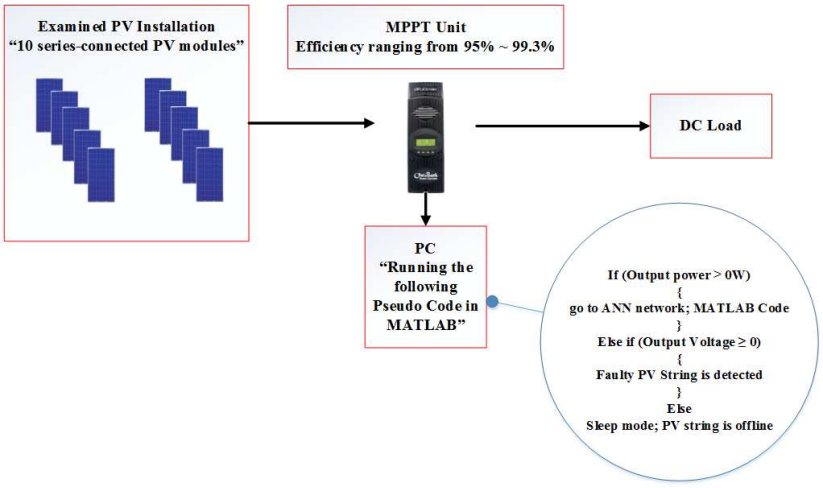
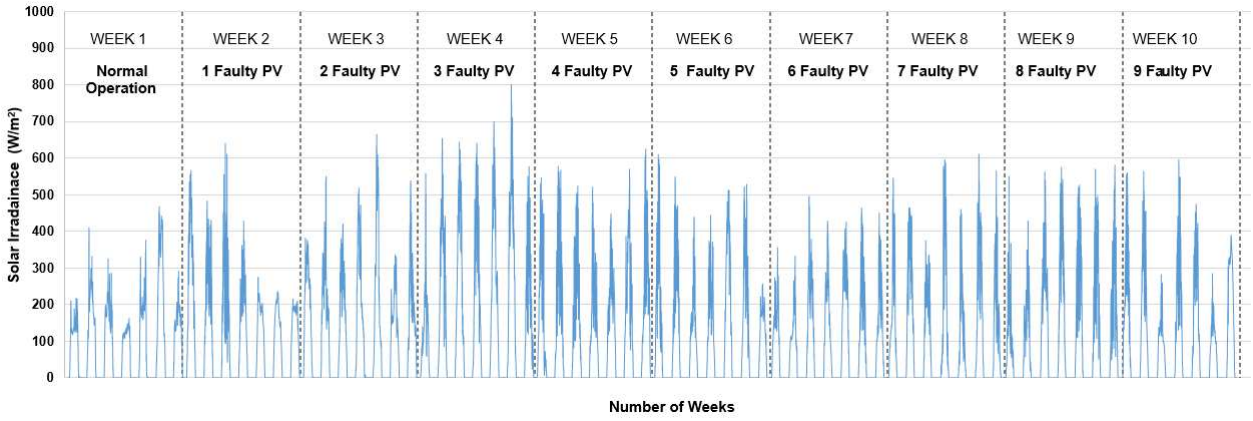
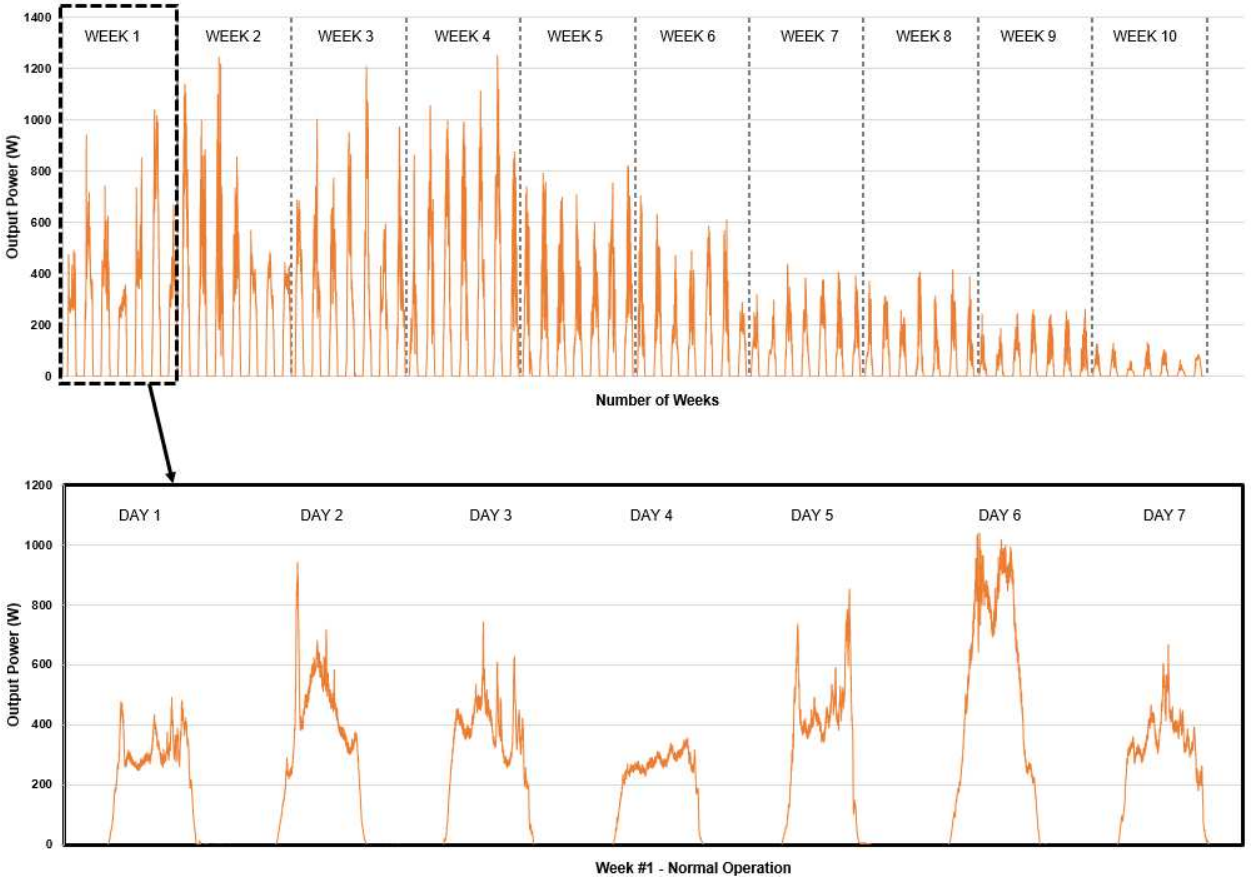


Figure 3. Flowchart of the proposed fault detection algorithm.

For this research, the solar irradiance and total power readings of the system, shown in Figure 4, were logged, the temperature of the PV modules was between 15.3–16.7 ^{o}C. However, this feature was not required as an input to the network. The data set consisted of 97200 measurements recorded over a period of 10 weeks, refer to Figure 4. A new fault case was applied on a weekly basis.



(a)



(b)

Figure 4. Data set used for training purposes; 97200 samples, each scenario has 9720 samples gathered over a period of ten weeks. (a) Solar irradiance, (b) Output PV power.

Several methodologies evaluated [17–19,22] record training samples for a period of 1 to 3 days for faulty conditions. Whereas in our research the data was recorded over a duration of 10 weeks, one week for every faulty condition. It was also noted that in recent work, 2019, conducted by [22–25] many input variables were required as inputs for the training of the ANN such as PV voltage, current, irradiance, power and ambient temperature. While our proposed RBF network was trained on only two inputs (solar irradiance and output power).

As evident from Figure 5, as the number of induced faults increased on a weekly basis, the output power of the PV system deteriorated.

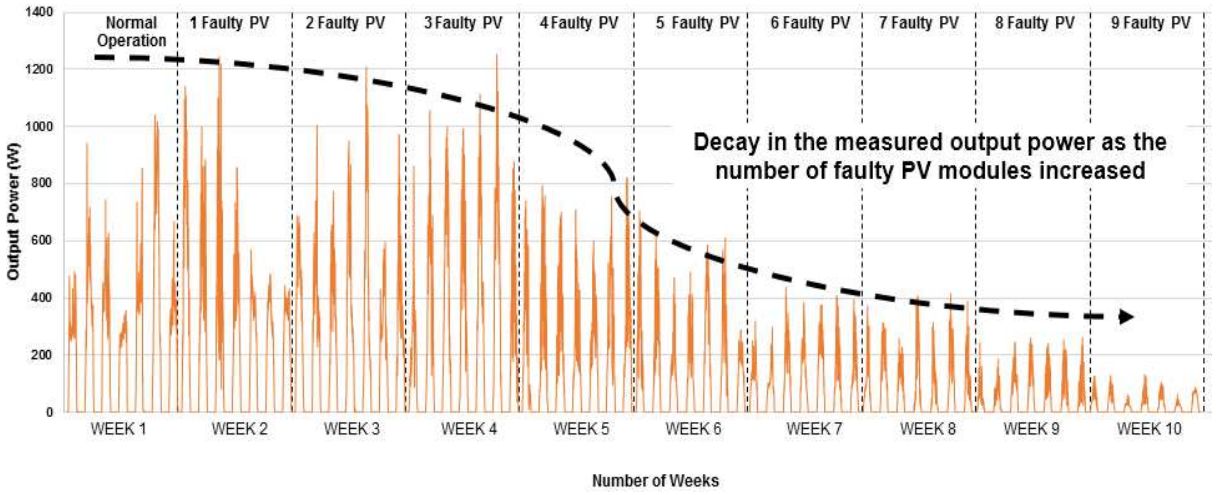


Figure 5. Flowchart of the proposed fault detection algorithm.

3.3. Proposed methodology

The proposed methodology was based on the concept of 'mapping' the inputs so that the data set was 'complete', before introducing it to the RBF network for training, validating and testing. The purpose behind this approach was to evaluate what effect the mapping of all data points within a specified range would have on the overall detection accuracy of the RBF network. For this methodology the data set was randomly selected, non-representative data was removed, max-min data normalisation technique implemented, and the mapping concept of solar irradiance against the output power was carried out as the novel part.

Figure 6 shows the solar irradiance 0–1000 W/m^2 being mapped to the corresponding output power, with a step size of 1 W/m^2. The selection of the step size was a major factor in the development of the methodology. The rationale for selecting a step size of one is explained in section 3.3.

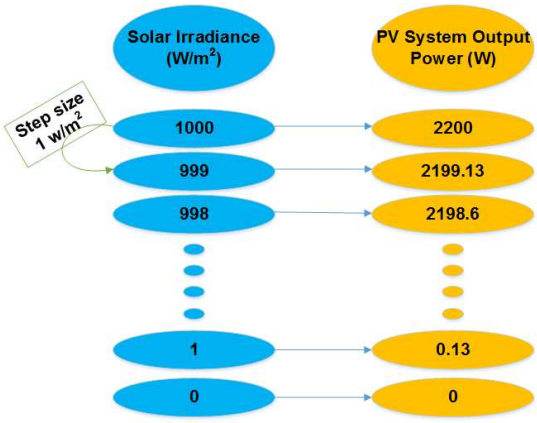


Figure 6. Mapping solar irradiance and PV system output power.

The receiver operating characteristics (ROC) provide a deeper insight into the performance of the proposed methodology, by providing the 'true-false positive' rate for each case from case 1 to case 10 (class 1 = "normal operation mode", while class 10 = "9 faulty" PV modules). From Figure 7, the high accuracy of the algorithm is evident, as every fault case is tightly aligned to the left side of the ROC plot. When reading the confusion matrix, green and red cells represent the number of correct and incorrect classifications by the ANN, respectively. Grey cells represent the total detection accuracy with respect to each row and column. The number 1, signifies 1- fault (F1) and so on, ending at 10- faults (F10). The overall accuracy of the developed algorithm as evident from the confusion matrix was 98.6%. This result is highly impressive when comparing it with present research considering its use of only two inputs which is unprecedented and the successful implementation of our methodology.

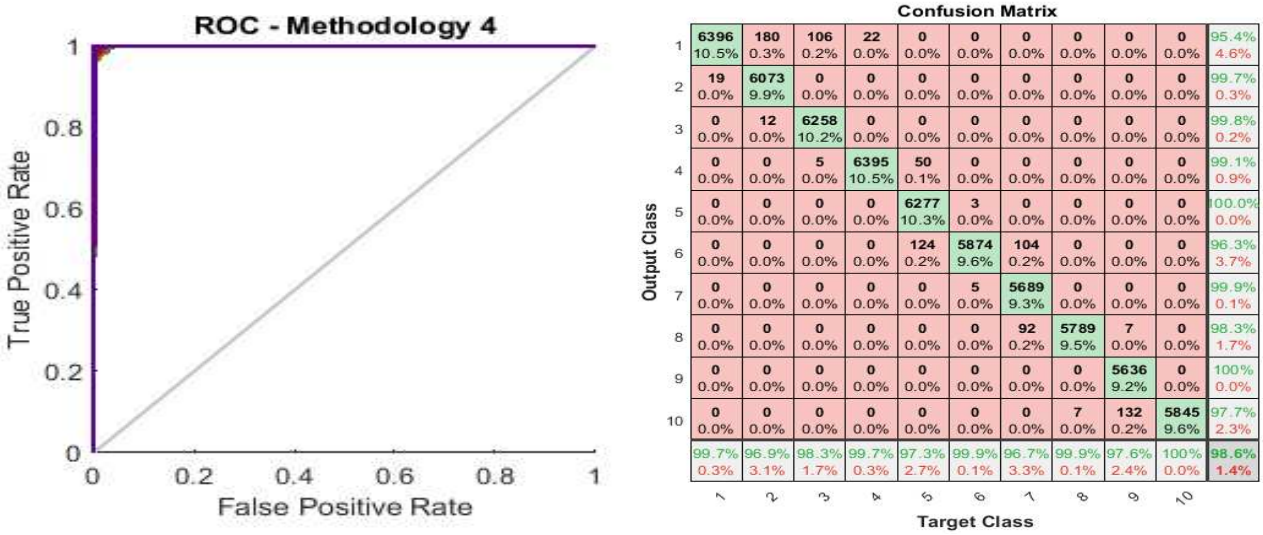


Figure 7. ROC performance & confusion matrix of proposed methodology.

3.4. Rationale for the selection of RBF network and methodology

At this point it is important to present the rationale for selecting an RBF network rather than an MLP network. The first reason for this was due to the implementation of only a single hidden layer, proven in section 3.1, providing an accuracy of almost 99%. However, MLP can also be used as a single layer with high accuracy. Therefore, the deciding factor was the computational time. As shown in Figure 8, RBF was the most effective with regards to computational time and accuracy. It provided an accuracy of 98.6% whilst requiring only 45 ms, whereas MLP improved the accuracy be a small margin of 0.5% but consumed 290 ms for this slight improvement. On the other hand, one of the novel contributions of this research is the use of only two input variables for accurately detecting faults in a PV system. The results of the network are far superior to results shared in present literature even though the latter projects use up to 5 inputs for the training of the ANN.

It is acknowledged that smart meters effectively provide a platform where readily available data can be taken for training ANN, thus removing the need for investing in more hardware for data collection, some key issues are not addressed by using smart meters. For example, an ANN network requiring 5 inputs rather than two, would result in an increased amount of time for data cleaning. As mentioned before the processing of data is the most vital part of machine learning consuming the

most amount time. Therefore, it is in the interest of the user to avoid unnecessary data from being included into the sample set in order to speed up the process and obtain high performance. This research provides an ANN network with reduced number of inputs, less time required for data censoring, reduced risk of non-representative data leaking into the training data set, resulting in unprecedented levels of accuracy. It is also important to note that the use of solar irradiance data is of wide interest to many circles within academia and industries, due to research and business specific goals such as techno-economic analysis and transportation sector. Due to this reason solar irradiance data is readily available on many online platforms for example the Met-Office in the UK. Therefore, acquiring solar irradiance data is not a major task which means effectively we only need to acquire the output power from the installation for implementation of our methodology.

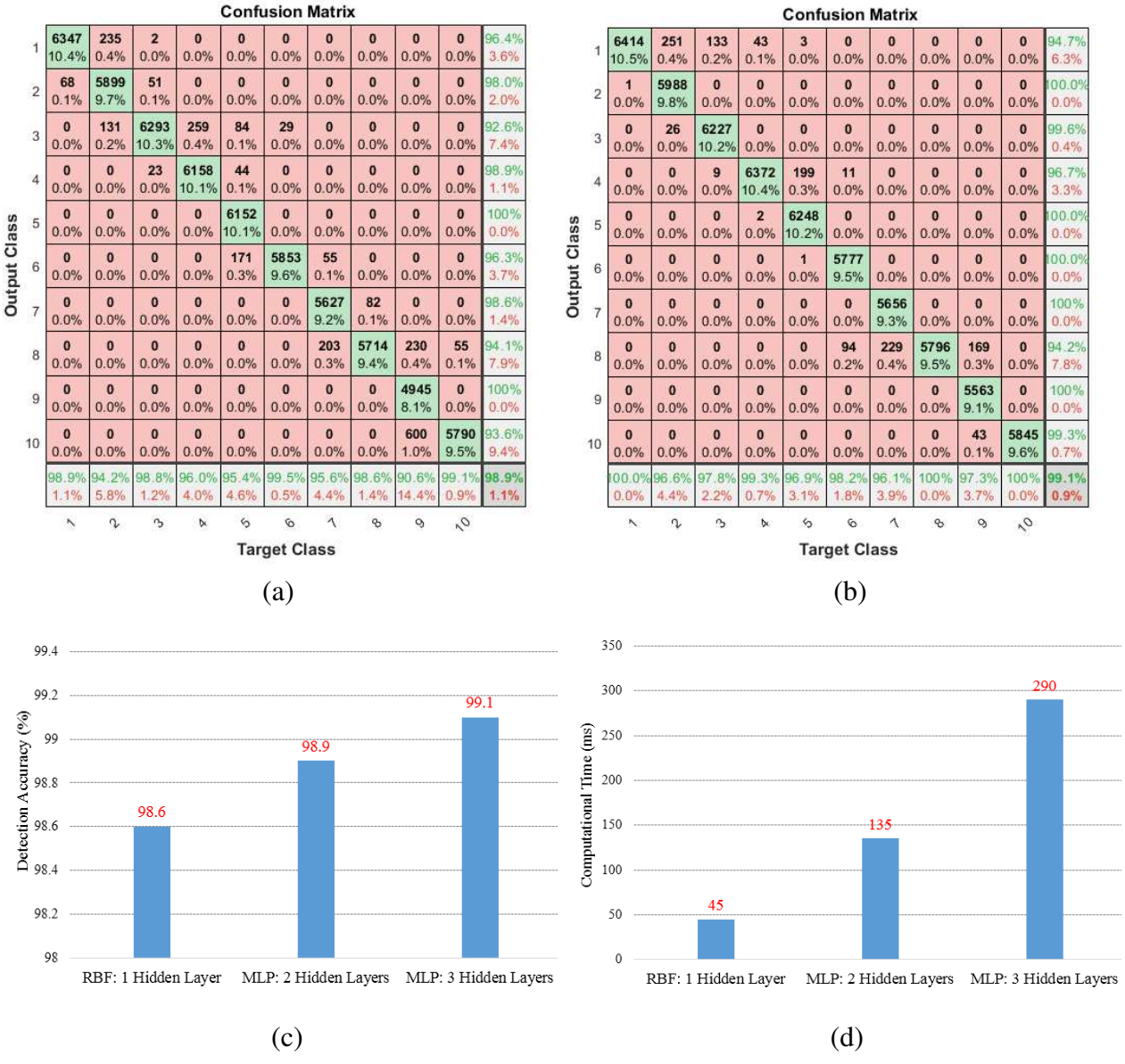


Figure 8. Output detecting accuracy of the MLP networks vs. RBF network. (a) 2 hidden layers, (b) 3 hidden layers, (c) Detection accuracy and the minimum computational time difference of each examined ANN network.

Adding to the novelty of the project is the implementation of our proposed methodology, which carries out the mapping of the inputs after removing any non-representative data and applying data normalisation. The gist of this methodology which had a direct impact on the accuracy of the network, was the selection of the 'step size'. The rationale for selecting a step size of 'one' was due to the significant impact it would have on the generalization of the network after training and most importantly testing stage. A step size of '0.5' would provide a network which falls into the category of under-fitting as the network would essentially map the result to each point in the data set and not be able to classify any new data. The selection of a step size of '10' would fall into the category of over-fitting the gap in data points would provide high impedance to the network forming any generalization.

The selection of the ANN network was based on the results obtained from testing and comparison of the two networks in terms of the overall detection accuracy and required computational time. Figures 8(a) and (b), show the overall detection accuracy of two MLP networks with 2 and 3 hidden layers respectively. There is only a small difference in the accuracy (0.2%) whereas the computational time increased by 155ms, refer to Figure 8(c). when comparing an RBF with an MLP network consisting of 3 hidden layers the detection accuracy increased by a small margin (0.5%) whilst the corresponding computational time increased considerably from 45ms to 290ms. This test formed the rationale for selecting an RBF network over MLP.

4. Results

This section presents the results obtained from further testing of the proposed methodology. The RBF network was tested in partial shading and overcast conditions, with the focus on how effectively the network would be able to adapt to the diverse conditions. It is important to note that the data set used to test the network was different to the data used for training the network, providing integrity and reliability in the results presented.

4.1. Partial shading results

The RBF network was first tested on a full-week dataset based on partial shading conditions. Figure 9 presents the solar irradiance for the first week.

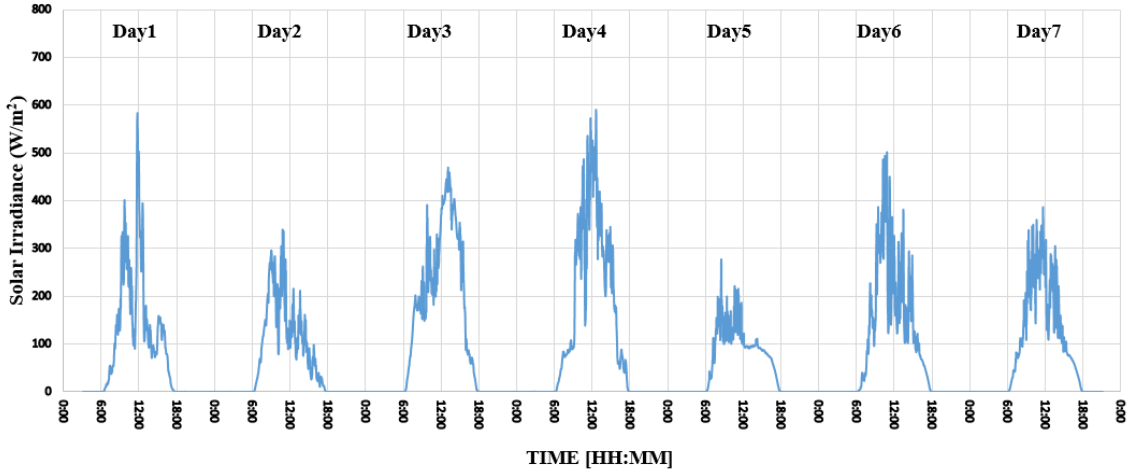


Figure 9. Solar irradiance of PV system for week one under partial shading conditions.

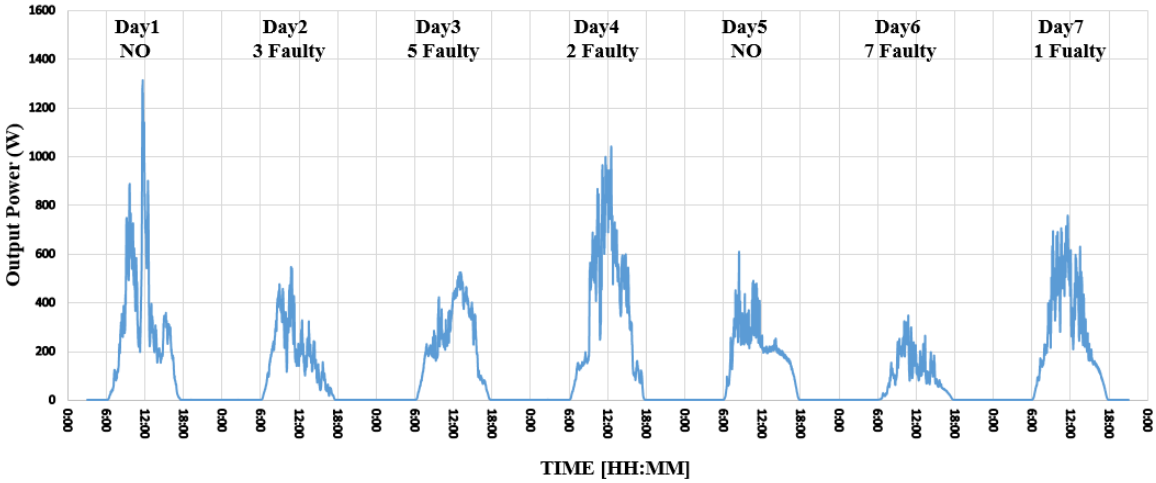


Figure 10. Total output power of PV system for week one under partial shading conditions.

Figure 10 illustrates the total output power of the system for the first week under partial shading conditions. Note, during 'day five' there is no fault applied, however the output power is at its lowest. This was due to the corresponding solar irradiance being at its lowest for that day, refer to Figure 9.

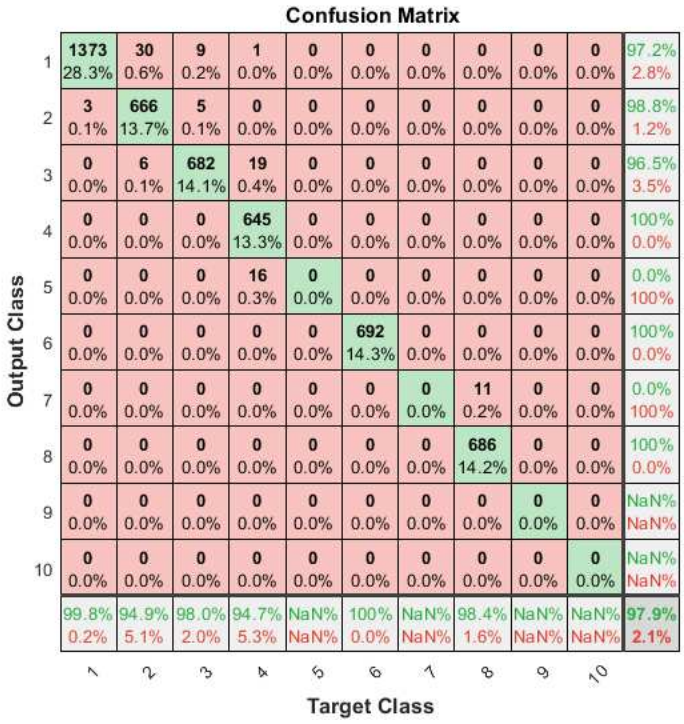


Figure 11. Confusion matrix for ANN under partial shading conditions.

Figure 11 presents the overall detection accuracy of the RBF network under partial shading conditions. An accuracy of 97.9% was recorded for the RBF network under partial shading conditions. As seen from the confusion matrix, 1373 samples for NO (normal operation), were correctly classified, corresponding to 28.3% of the overall sample set (4844 sample). In the same way, 666 samples were correctly classified as F2 (2 faulty modules), this corresponds to 13.7% of all

4844 samples. In row 1, 30 samples of F1 (1 faulty module) were incorrectly classified as normal operation "NO" corresponding to 0.4% of all samples. The overall confusion matrix for partial shading conditions is shown in Figure 11. The performance of the RBF network can be considered as highly accurate in comparison to ANN networks presented in recent publications on PV fault detection through the use of ANN. This is further discussed in the comparative study, section 5.

4.2. Overcast results

The RBF network was then tested under overcast conditions (partially cloud and overcast). Figure 12 presents the solar irradiance for the week. It is evident that the solar irradiance varies from 100 to 450 W/m^2 depends on the overcasting or shading affecting the PV system.

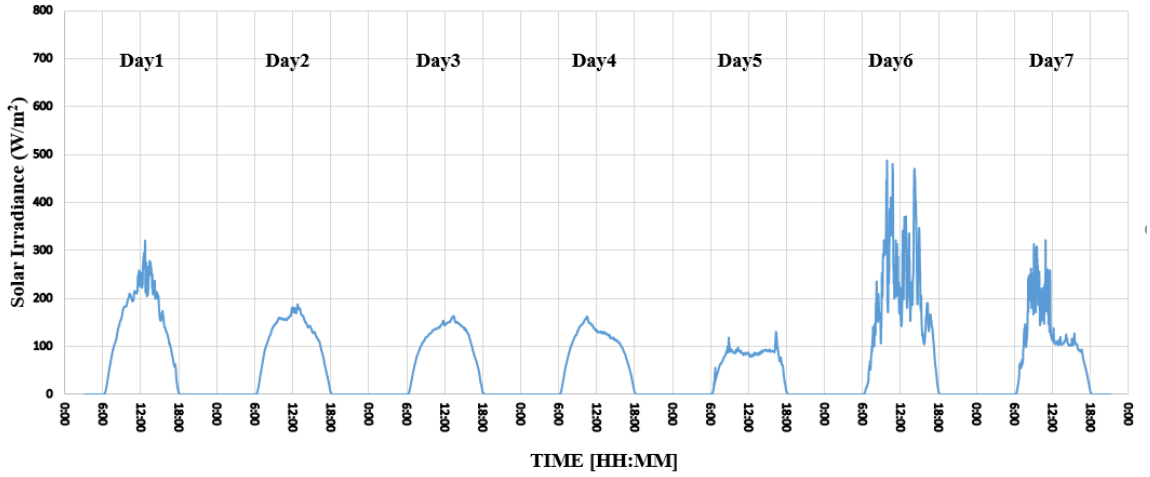


Figure 12. Solar irradiance of PV system for week two under overcast conditions.

Figure 13 illustrates the total output power of the system for the week-long dataset under overcast conditions. As expected, the total output power of the PV installation decreased as the number of faults applied to the system increased.

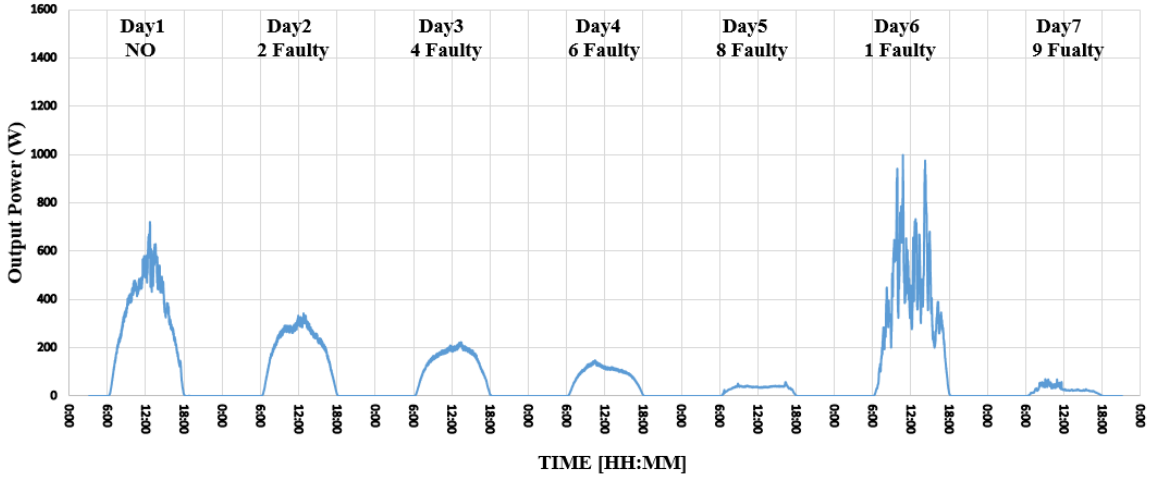


Figure 13. Total output power of PV system for week two under overcast conditions.

Figure 14 presents the overall detection accuracy of the RBF network for overcast conditions. The accuracy of the network slightly decreased to 96.5%. 710 samples were correctly classified as normal operation, corresponding to 14.2% of all samples. Likewise, 661 samples were correctly classified as F2, equating to 13.2% of the total sample set. 34 samples of NO are incorrectly classified as F1 in row 1 of the matrix, corresponding to 0.7%. This was due to the output power during normal operation mode and 1-faulty case being of similar nature, particularly for partial shading conditions.

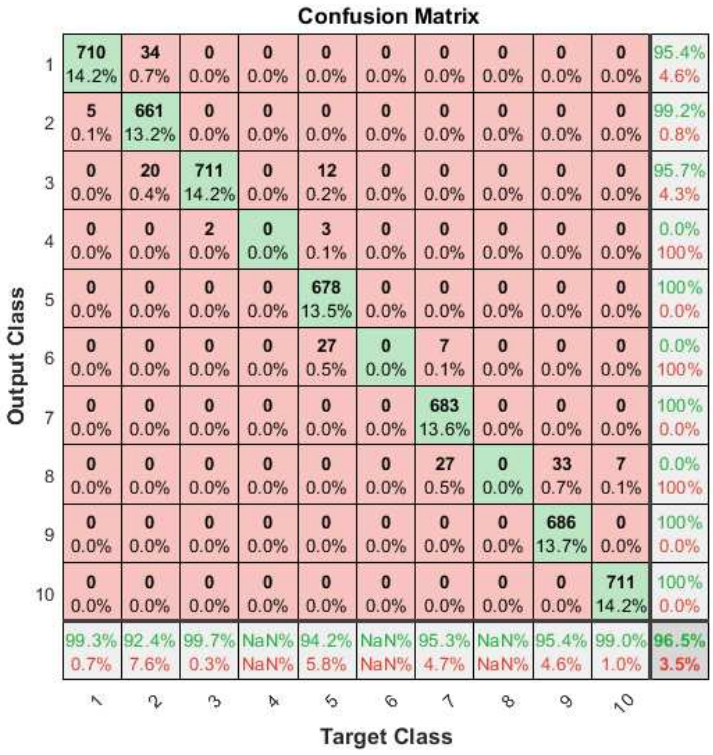


Figure 14. Confusion matrix for ANN under overcast conditions.

Although the performance of the RBF network for both scenarios was above 95% which is unprecedented based on present literature, the accuracy of the network under partial shading condition was higher than that of overcast conditions. The difference in detection accuracy of 1.4% between the two scenarios highlights the resilience, robustness of the network. Although, the quantity of data recorded in both scenarios was the same, the content differed dramatically i.e. solar irradiance was lower for overcast condition compared to partial shading. Again, this reinforces the effectiveness of the RBF network to be able to adapt and generalize based on the training data.

5. Comparative study

This section provides a comparison of the recent research outcomes [19,22–24] in the field of PV fault detection using ANN with our proposed methodology. As evident from Table 2, all papers reviewed and used for comparison with our work require a significant number of input parameters for the training of the ANN. As a result, the algorithms require a larger set of historical data for the ANN training/validation and testing purposes, a requirement which cannot always be fulfilled

depending on the availability of data. The use of an increased number of variables also places more burden on the processing of the data before it is ready for training the network. However, our proposed algorithm only requires two input parameters in order to activate, namely the solar irradiance and the output power, while there is no need for any other PV parameters such as the Vmpp and Impp, appearing in all other recent algorithms [19,22–24]. Also, the reduced number of inputs eases the process of data cleaning, paving the way for more quality data being used for training the network resulting in higher performance.

Another advantage of our proposed algorithm is that the ambient temperature has been made redundant. Thus, further reducing hardware costs along with reduced time for data processing, further enhancing the solutions proposed by [23] and [24]. According to our method, the ANN detection accuracy is ranging from 96.5%~98.1% in normal operation and partial shading conditions, respectively. So far, the obtained PV fault detection accuracy is considered the highest.

Table 2. Compative study of recent ANN-based PV fault detection algorithms and our proposed method.

Ref.	Year	ANN Methodology – required input parameters								ANN detection accruacy (%)	
		G	T_{amb}	P	I_{sc}	V_{oc}	I_{mpp}	V_{mpp}	No. of required parameters	Normal Operation	Partial Shading
[19]	2016	✓	✓	✓	✗	✗	✓	✓	5	90.3~97	max:90.3
[22]	2017	✓	✗	✗	✓	✓	✓	✓	5	96.5	n/a
[23]	2018	✗	✗	✓	✓	✓	✓	✓	4	97.1~95.3	87.3~92.1
[24]	2019	✓	✓	✓	✗	✗	✓	✓	5	87.4~98.5	max: 66.45
proposed system	2020	✓	✗	✓	✗	✗	✗	✗	2	98.1%	96.5%

6. Conclusions

In this article, we have presented a methodology to detect PV faults based on two inputs parameters; solar irradiance and output power. The accuracy of the RBF network reached an unprecedented 98.6%. This was based on the implementation our methodology consisting of data normalisation as well as mapping of solar irradiance against output PV power. The developed RBF was also tested in partial shading and overcast conditions to further endorse its success.

The whole sample set (over 97200 samples) were actual data points from a live installation rather than being simulated, further reinforcing the success of the RBF network. Results show that the developed ANN network accurately detected PV faults in the range of 97.9% during normal operational mode, where no shading/overcast is present. Whereas, during partial shading conditions, the RBF network accuracy deteriorated slightly to 96.5%. Both results are considered very high when comparing with present research, considering the reduced number of inputs.

Conflict of interest

All authors declare no conflicts of interest in this paper.

References

1. Di Piazza C, Viola F, Vitale G (2018) Evaluation of ground currents in a PV system with high frequency modeling. *International Journal of Renewable Energy Research* 8: 1770–1778.
2. Dhimish M, Holmes V, Mehrdadi B, et al. (2017) Seven indicators variations for multiple PV array configurations under partial shading and faulty PV conditions. *Renew Energ* 113: 438–460.
3. Dhimish M, Mather P, Holmes V (2018) Evaluating power loss and performance ratio of hot-spotted photovoltaic modules. *IEEE T Electron Dev* 56: 5419–5427.
4. Schirone L, Califano FP, Pastena M (1994) Fault detection in a photovoltaic plant by time domain Reflectometry. *Prog Photovoltaics* 2: 35–44.
5. Silvestre S, da Silva MA, Chouder A, et al. (2014) New procedure for fault detection in grid connected PV systems based on the evaluation of current and voltage indicators. *Energ Convers Manage* 86: 241–249.
6. Li X, Wen H, Hu Y, et al. (2019) Drift-free current sensorless MPPT algorithm in photovoltaic systems. *Sol Energy* 177: 118–126.
7. Wu Y, Lan Q, Sun Y (2009) Application of BP neural network fault diagnosis in solar photovoltaic system. *2009 International conference on Mechatronics and Automation* 2581–2585.
8. Chen Z, Chen Y, Wu L, et al. (2019) Deep residual network based fault detection and diagnosis of photovoltaic arrays using current-voltage curves and ambient conditions. *Energ Convers Manage* 198: 111793.
9. Dhimish M, Mather P, Holmes V, et al. (2018) CDF modelling for the optimum tilt and azimuth angle for PV installations: case study based on 26 different locations in region of the Yorkshire UK. *IET Renewable Power Generation* 13: 399–408.
10. Pan T, Chen J, Zhou Z, et al. (2019) A Novel Deep Learning Network via Multi-Scale Inner Product with Locally Connected Feature Extraction for Intelligent Fault Detection. *IEEE T Ind Inform* 9: 5119–5128.
11. Belaout A, Krim F, Mellit A, et al. (2018) Multiclass adaptive neuro-fuzzy classifier and feature selection techniques for photovoltaic array fault detection and classification. *Renew Energ* 127: 548–558.
12. Lapedes A (1987) Nonlinear signal processing using neural networks. Technical Report No. LA-UR-87-2662.

13. Polo O, Bermejo F, Fernández G, et al. (2015) Failure mode prediction and energy forecasting of PV plants to assist dynamic maintenance tasks by ANN based models. *Renew Energ* 81: 227–238.
14. Sun Y, Xu, J, Qiang H, et al. (2019) Adaptive sliding mode control of maglev system based on RBF neural network minimum parameter learning method. *Measurement* 141: 217–226.
15. Yagi Y, Kishi H, Hagihara R, et al. (2003) Diagnostic technology and an expert system for photovoltaic systems using the learning method. *Sol Energ Mat Sol C* 75: 655–663.
16. Chine W, Mellit A, Lughi V, et al. (2016) A novel fault diagnosis technique for photovoltaic systems based on artificial neural networks. *Renew Energ* 90: 501–512.
17. Tadj M, Benmouiza K, Cheknane A, et al. (2014) Improving the performance of PV systems by faults detection using GISTEL approach. *Energ Convers Manage* 80: 298–304.
18. Mellit A, Sağlam S, Kalogirou A (2013) Artificial neural network-based model for estimating the produced power of a photovoltaic module. *Renew Energ* 60: 71–78.
19. Chine W, Mellit A, Lughi V, et al. (2016) A novel fault diagnosis technique for photovoltaic systems based on artificial neural networks. *Renew Energ* 90: 501–512.
20. Chen Y, Yu G, Long Y, et al. (2019) Application of radial basis function artificial neural network to quantify interfacial energies related to membrane fouling in a membrane bioreactor. *Bioresource Technol* 293: 122103.
21. Zhao Z, Lou Y, Chen Y, et al. (2019) Prediction of interfacial interactions related with membrane fouling in a membrane bioreactor based on radial basis function artificial neural network (ANN). *Bioresource Technol* 282: 262–268.
22. Karmacharya M, Gokaraju R (2017) Fault location in ungrounded photovoltaic system using wavelets and ANN. *IEEE T Power Deliver* 33: 549–559.
23. Dhimish M, Holmes V, Mehrdadi B, et al. (2018) Comparing Mamdani Sugeno fuzzy logic and RBF ANN network for PV fault detection. *Renew Energ* 117: 257–274.
24. Fadhel S, Delpha C, Diallo D, et al. (2019) PV shading fault detection and classification based on IV curve using principal component analysis: Application to isolated PV system. *Sol Energy* 179: 1–10.
25. Dhimish M, Chen Z (2019) Novel Open-Circuit Photovoltaic Bypass Diode Fault Detection Algorithm. *IEEE J Photovolt* 9: 1819–1827.

AIMS Press