Fault Location in Ungrounded Photovoltaic System Using Wavelets and ANN

Indra Man Karmacharya, *Member, IEEE*, and Ramakrishna Gokaraju, *Senior Member, IEEE*

***Abstract*—Identifying ground faults is a significant problem in ungrounded photovoltaic (PV) systems because such earth faults do not provide sufficient fault currents for their detection and location during system operation. If such ground faults are not cleared quickly, a subsequent ground fault on the healthy phase will create a complete short circuit in the system. This paper proposes a novel fault-location scheme in which high frequency noise patterns are used to identify the fault location. The high-frequency noise is generated due to the switching transients of converters combined with the parasitic capacitance of PV panels and cables. Discrete wavelet transform is used for the decomposition of the monitored signal (midpoint voltage of the converters) and features are extracted. Norm values of the measured waveform at different frequency bands give unique features at different fault locations and are used as the feature vectors for pattern recognition. Then, a three-layer feedforward artificial neural networks classifier, which can automatically classify the fault locations according to the extracted features, is investigated. The proposed fault-location scheme has been primarily developed for fault location in the PV farm (PV panels and dc cables). The method is tested for ground faults as well as line–line faults. These faults are simulated with a real-time digital simulator and the data are then analyzed with wavelets. Finally, the effectiveness of the designed fault locator is tested with varying system parameters. The results demonstrate that the proposed approach has accurate and robust performance even with noisy measurements and changes in operating conditions.**

***Index Terms*—Fault location, ungrounded photovoltaic (PV) system, discrete wavelet transform (DWT), multi-resolution analysis (MRA), artificial neural networks (ANNs).**

I. Introduction

VARIOUS techniques have been described in the literature with respect to fault location methods in ungrounded PV systems. The fault location approach currently employed in most industrial ungrounded PV systems uses a pulse generator to send a high frequency signal into the faulted system and traces the signal to locate the fault. This type of fault locating method for high-impedance grounded or ungrounded systems is briefly discussed in [1]. Researchers at Bender Inc., USA, have implemented this approach in the solar PV industry [2]. Generally, fault location by tracing the signal consumes a lot of time; there is also the chance for errors because it is done manually. This approach also involves tracing the fault point in live equipment enclosures and requires reaching cables at difficult locations. Reference [3] reported a protection scheme which is capable of detecting a ground fault in ungrounded DC traction power systems by applying voltage through a probe unit. The probe unit consists of a power supply that can generate voltages with arbitrary amplitude and frequency, and it is installed between the line and ground at each substation. Once the ground fault is identified, this method locates the fault by DC or swept frequency AC response analysis of the probe current returning through ground.

The traveling wave (TW)-based technique has been implemented for fault location with single-ended or double-ended principles [4]. With the advancement of communication and digital signal processing technologies, TW-based fault location methods have been implemented in recent industrial applications for more accurate and reliable fault location estimations. In recent papers [4], [5], TW fault locators were developed in industrial hardware that uses time-synchronized measurements of the TW currents at the line terminals to determine accurate fault locations. Researchers from the University of Manitoba and Manitoba HVDC Research Center utilized wavelet transform to detect the arrival times of TWs and showed fault location accuracy of ± 400 m for the conventional high voltage direct current (HVDC) systems with a mixed transmission media consisting of overhead lines and cables [6]. Although modern TW fault locators are highly accurate and reliable for AC and HVDC transmission lines, this technique would be very difficult to implement for small size distribution and distributed generation systems. This is due to a very small propagation time and the low latency between incident waves and associated reflected waves that might be very hard to individually detect in short cable lengths.

In the time domain reflectometry approach, an external voltage signal is applied into the system from one end of the line and reflection of the signal from the fault point is monitored along with the incident waveform. In [7], the experimental analysis of injected and reflected waveforms to identify and localize the common faults in a 1 MW PV plant is presented. Other papers discuss the auto-correlation plots (the incident signal with the reflected signal) generated by spread spectrum time domain reflectometry (SSTDR) to find the location of ground and arc faults in a PV system [8].

Reference [9] describes an artificial neural network (ANN)-based fault locating algorithm for three-phase transmission lines

Manuscript received August 12, 2016; revised February 18, 2017 and May 8, 2017; accepted June 10, 2017. Date of publication June 29, 2017; date of current version March 22, 2018. Paper no. TPWRD-00999-2016. *(Corresponding author: Ramakrishna Gokaraju.)*

The authors are with the University of Saskatchewan, Saskatoon, SK S7N 5C5, Canada (e-mail: indra.karmacharya@usask.ca; rama.krishna@usask.ca).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TPWRD.2017.2721903

that uses fundamental components of pre-fault and post-fault voltage and current phasors as inputs. In [10], a fault diagnosis in a PV array, especially for short-circuits, is proposed using a three-layered feed-forward neural network. Furthermore, a two-layered ANN is used in [11] to predict the expected power using temperature and insolation as the inputs. An analytical method is combined with ANN outputs to diagnose the PV string faults. However, this method only examines the occurrence and types of faults in the PV strings and the fault location approach has not been investigated. This method utilizes raw sampled data as the input to the ANN, and hence it requires long training and computation times.

Yan Pan in her Ph.D. thesis work describes an approach for locating ground faults in ungrounded shipboard DC distribution systems that utilizes the high frequency noise generated by repetitive switching of power converters interacting with parasitic elements in the system [12]. She conducted tests in an experimental hardware setup and the results verified the approach to differentiate various fault locations in a real environment [13], [14]. In [15], the noise pattern analysis approach was extended to integrate the wavelet-based multi-resolution analysis (MRA) technique and ANN for detection and classification of common faults (ground and short circuit) on ungrounded medium voltage direct current (MVDC) shipboard power systems.

In the field of distribution systems, reference [16] discusses a "compressive sensing" method to find fault location in distribution networks. Pre-fault and during fault voltages are measured along the feeder. A current vector is generated from the voltage sag matrix and impedance matrix. There is one non zero element in the sparse current matrix which corresponds to the bus at which a fault occurs. In [17], compressive sensing and l^1-norm minimization approach was extended to estimate both single and simultaneous fault location in distribution networks. Since the number of nonzero values in the estimated current vector is not necessarily identical to the number of faults, these values are examined by Fuzzy-clustering mean to estimate four possible faulted points. Furthermore, a machine learning technique based on the k-nearest neighborhood was implemented to find a single fault. Reference [18] describes a downstream graph marking algorithm for automatically locating fault area based on the statuses of fault indicators that are telemetered to the distribution control center. Reference [19] describes an algorithm to detect, classify and locate power quality disturbances and fault locations using currents and voltages recorded at substation of the a distribution system. The method described in the explores a combination of modern techniques for network analysis, signal processing (WT-MRA), and intelligent systems (Fuzzy-ANN). Reference [20] discusses a method for detecting and locating high impedance faults in multiconductor overhead distribution lines using power line communication devices. A fault is detected by monitoring the change in input impedance. Fault location is obtained by analyzing response to impulse injection by all installed PLC devices along the line. Reference [21] investigates the impact of distributed generation on voltage sag based fault location methods. The paper quantifies the impact of DG on fault location. Reference [22] discusses a PMU-based fault detection and location method for active distribution networks using real-time state estimation. The method described in the paper identifies the faulted line irrespective of the neutral connection, fault type/impedance or position along the line. Reference [23] also discusses a high-impedance fault identification method. It uses DWT to monitor high and low-frequency voltage components to find the most likely area where the disturbance has occurred. Reference [24] discusses a method for the identification and location of high impedance single-phase-earth faults on neutral compensated and isolated medium-voltage networks. It uses standard measurements available in substations.

The main objective of this research is to develop an accurate, robust fault locator for ungrounded PV systems that, unlike existing methods, can be used while the system is in operation. It uses a DWT-based MRA and a classifier based on ANNs. The main novelty of this algorithm is that it uses high frequency noise patterns induced by converters switching transients interacting with parasitic elements of PV panels and cables.

The paper is organized as follows. Section II describes the modeling of an ungrounded PV test system in a real-time electromagnetic transient (EMT) simulation platform. The basic structure of the proposed pattern recognition algorithm is provided in Section III. A brief description of the DWT-based MRA technique is presented in Section IV. Section V presents the feature extraction technique. Section VI describes the structure and training algorithm of the classifier based on ANNs. Section VII provides the fault location results. Section VIII describes the parameter sensitivity analysis of the proposed method.

II. Ungrounded PV System Model

Ungrounded DC distribution cables are used often in North America for low voltage PV systems, shipboard power systems and offshore wind farms as the system could be kept running after a first ground fault. The size of the inverters needed are also small due to transformer-less configuration and therefore the inverter cost is also less with the ungrounded PV configurations [25].

The test system shown below in Fig. 1 was modeled and simulated in a real-time digital simulator (RTDS). A multi-string ungrounded PV system was connected to the grid using a 500 kW, three-phase, two-level voltage source inverter (VSI), which converts DC power generated from the PV array into AC power. Two 250 kWp PV arrays were connected to individual DC-DC converters (buck converters) to implement the maximum power point tracking (MPPT) technique. Each PV array has 48 parallel strings, and each string consists of 17 modules connected in series. The multi-string PV arrays were connected to a 600 V common DC bus through DC underground cables, and the DC bus was integrated with the DC-link capacitor and the input terminal of the three-phase inverter. The VSI and DC-DC converters were controlled based on the sinusoidal pulse-width modulation (SPWM) technique. A harmonic filter connects the output terminals of the VSI to respective phases of the point of common coupling (PCC). The 500 kVA, 0.21/20 kV, three phase, Δ/Y, step-up transformer was used to transfer PV power to the distribution system. The station service load was 100 kW, 210 V, unity p.f., three-phase dynamic load that, was connected

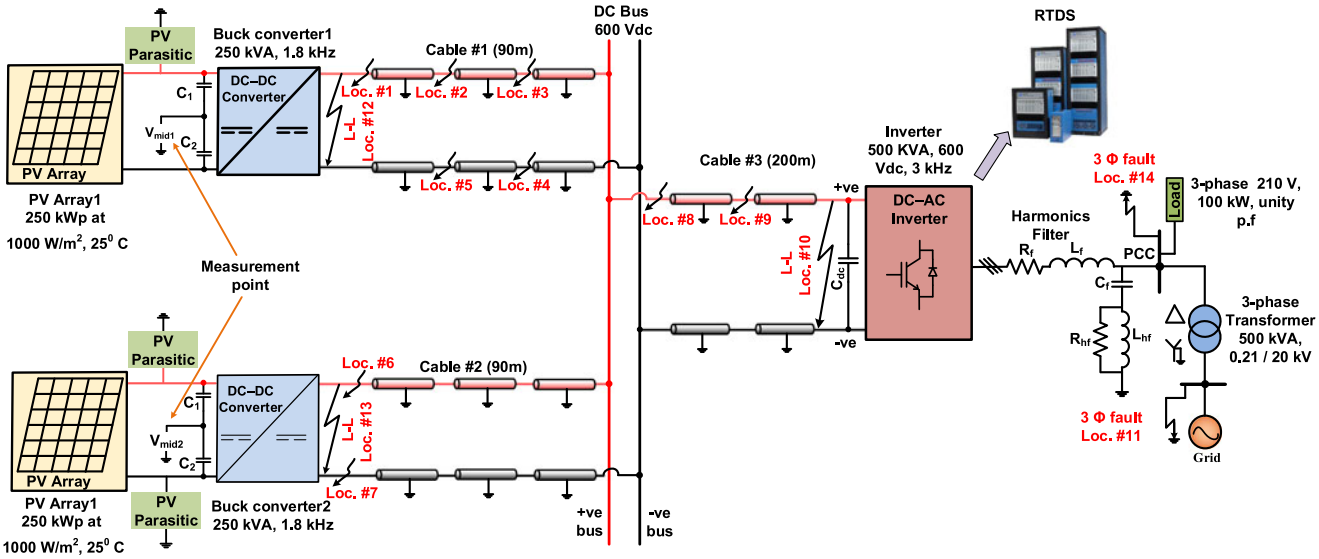


Fig. 1. 500 kWp grid integrated multi-string converter type solar PV system.

to the output of the inverter at the PCC. The mid-point voltages (V_{mid1} or V_{mid2}) of the DC-DC converters with respect to grounding were measured and analyzed using a signal processing method for detection and location of faults in the PV system. The PV data are provided in Appendix A.

The ground and line-line cable fault locations are shown in Fig. 1 and marked as loc. #1 to loc. #14. Here, the fault locations 1 to 7 simulate the ground faults on cables #1 and #2. Locations 8 and 9 are ground faults on cable #3. Locations 10 to 14 are used to simulate the line-line faults in the DC and AC parts of the PV system.

Ground, line-line, and arcing faults at different locations of the PV arrays are also simulated and analyzed, and will be discussed in Section VII-B. The fault location scheme presented in this paper is used to primarily locate the ground and line-line faults on the DC side of the PV system. A few representative results are also included with AC distribution side faults to test the effectiveness of the method.

A complete electrical model of a PV array and its insulation to ground is required to analyze ground faults in an ungrounded PV system. Therefore, a single diode model of solar cells in the PV array is used in RTDS/RSCAD, which accurately represents nonlinear current-voltage characteristics of a solar cell by including series and parallel resistance with an ideal solar cell, as discussed in [26]. An equivalent circuit model simulating the behavior of PV module insulation was proposed in [27]. In normal operating conditions, the measured values of insulation resistance and leakage capacitance in a real 68 kWp PV generator are [0.12, 4.6] MΩ and [0.7, 42] μF, respectively [27].

A classical lumped parameter cable model is not capable of representing transient behavior of the system for different frequencies. This is because the resistance and inductance of the cable are frequency dependent due to skin and proximity effects. The limitations of the lumped cable model have been minimized with various frequency dependent TW-based cable models developed in the literature [28]. TW-based frequency dependent cable models require a simulation time-step shorter than the propagation time of TW in the cable. The cables used in PV systems are very short in length, between 30 and 200 meters. As an example for a 30 m cable, travel time was calculated as 0.1 μs. Therefore, to simulate a frequency dependent EMT cable model, the simulation time-step needed would have to be less than 0.1 μs. However, a real-time simulation platform, that can simulate with time-steps less than 1.4 μs is not available. Due to the aforementioned limitations, a frequency dependent π cable model was developed, which is essentially a lumped π model with frequency dependent resistance and inductance, and EMT simulation was done in small time-steps.

The frequency dependent cable model was simulated using small-time step simulations (2.941 μs) in RTDS. For this small time-step simulation speed, the cable was modeled as 30 meter sections. Since the cable is modeled in 30 meters sections, the minimum resolution for finding the fault location using the proposed method is going to be 30 meters.

III. Pattern Recognition Algorithm

The analysis of high frequency noise patterns to identify fault locations in ungrounded PV systems is proposed. These noise signal patterns are generated due to the power electronic converter switching frequencies and parasitic elements of PV panels and cables (such as cable insulation capacitance and stray inductance). The proposed location method requires a distinguishable signal that uniquely identifies a fault position. The mid-point voltage of a DC-DC converter 1 (V_{mid1}) or converter 2 (V_{mid2}), with measurement points as indicated in Fig. 1, is monitored. The simulation results illustrate that the high frequency noise is observable in the measured voltage signal and different noise patterns are introduced due to ground faults at various locations of the PV system. Furthermore, the monitored voltage results in a significant change in noise patterns for different ground fault

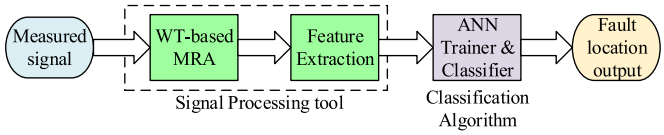


Fig. 2. Flowchart of the proposed algorithm.

locations. Hence, in this paper, the V_{mid1} voltage signal is measured and analyzed. No external signal generator is required to locate the fault using the proposed method.

Fig. 2 gives a simplistic flowchart of the proposed methodology. The measured time-domain signal is transformed into a time-frequency domain by splitting it into low and high frequency components with a DWT-based MRA technique. The energy or norm of the noise signals at each frequency band gives a unique signature for different fault locations and is used as extracted features for pattern recognition. The extracted features are then used as inputs of a multi-layer neural network classification algorithm. The outputs of the ANN classifier give the exact location of the fault.

IV. Signal Analysis

A. Wavelet Transforms and Multi-Resolution Analysis

Consider the input signal $f(t)$ that is discretized where integer variable n refers to the sample number of a discrete series. To obtain a discrete version of the WT, the scaling parameter (s) and translation parameter (u) are discretized with $s = a^j$ and $u = ku_o a^j$. Then, the DWT of the sampled input sequence, $f(n)$, can be defined as

$$DWT_{\Psi}^{f}(j,k) = \frac{1}{\sqrt{a^j}} \sum_{n=-\infty}^{\infty} f(n)\Psi\left(\frac{n - ku_o a^j}{a^j}\right) \quad (1)$$

where k, j, a, and u_o are integer numbers with $a > 1$ and $u_o \neq 0$.

The MRA introduced by Stephane G. Mallat [29] is a signal decomposition procedure that divides the signal into a set of frequency bands using a series of low-pass and high-pass filters. The time domain signal $f(t)$ can be represented in terms of scaling, $\Phi(t)$, and wavelet, $\Psi(t)$, functions as given below:

$$f(t) = \sum_{k=-\infty}^{\infty} a_N(k)\Phi_{N,k}(t) + \sum_{j=1}^{N}\sum_{k=-\infty}^{\infty} d_j(k)\Psi_{j,t}(t) \quad (2)$$

where d_j is the detail coefficient at various resolution levels and a_N is the last approximation coefficient at level N.

Thus, to analyze a discrete signal with WT-based MRA, two digital filters can be implemented: one high-pass filter, $h_1[n]$, associated with a mother wavelet $\Psi(t)$, and its low-pass mirror version, $h_0[n]$, related to the scaling function $\Phi(t)$. A schematic diagram of MRA of a signal using discrete wavelet filters is provided in Fig. 3, where three levels of decomposition are carried out. The discrete input signal is decomposed into different frequency bands, passing through a high-pass filter $h_1[n]$ and low-pass filter $h_0[n]$, and then the outputs of the filters are down-sampled by 2. The high frequency components, called detail coefficients (d_j), have high frequency and low time

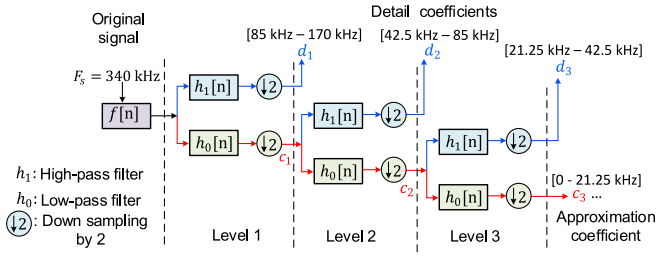


Fig. 3. Multi-resolution analysis using DWT.

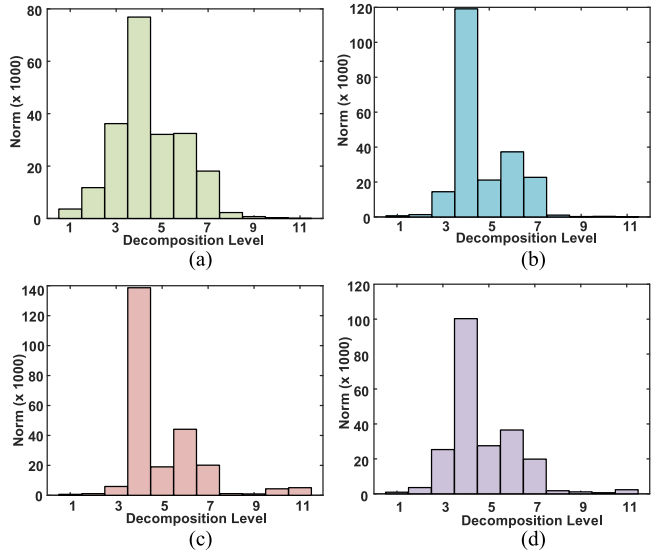


Fig. 4. Energy spectrum of V_{mid1} for ground fault at Loc. #6 with different types of wavelet functions.

resolution whereas the low frequency components, known as approximation coefficients (a_j), provide low frequency and high time resolution. Consider that the sampling frequency of an original input signal is F_s and equals 340 kHz. Instruments with a sampling frequency of 340 kHz are nowadays inexpensive and could be used practically in the field. The modern digital signal processor and analog-to-digital converters provide sampling frequency in the MHz range [5], [6], [30]. Then, the frequency bands for different decomposition levels can be found by $\left[0, F_s/2^{N+1}\right]$ for a low-pass filter and $\left[F_s/2^{N+1}, F_s/2^N\right]$ for a high-pass filter.

Based on the literature findings, Daubechies wavelets are selected as appropriate mother wavelets herein.

Fig. 4 is an energy spectrum diagram of the measured signal, V_{mid1} (refer to Fig. 1 for the fault at Loc. #6). Using db1, the energy of the measured signal is distributed over different resolution levels [in Fig. 4 (a)] because of the wide cut-off frequency of the filters. In contrast, with db10 the energy is more concentrated at the 4th resolution level and a very small part of the signal energy is leaked to neighboring resolution levels. Therefore, a db10 wavelet is selected for WT-based MRA to extract the unique features.

TABLE I
DIFFERENT DECOMPOSITION LEVELS FOR CABLE FAULT ANALYSIS

Decomposition Level	Frequency Band (Hz)
1	85 k–170 k
2	42.5 k–85 k
3	21.25 k–42.5 k
4	10.625 k–21.25 k
5	5.3125 k–10.625 k
6	2.656 k–5.3125 k
7	1.328 k–2.656 k
8	664–1.328 k
9	332–664
10	166–332
11	83–166

B. Selection of Decomposition Level

The noise signal is decomposed into 11 frequency bands to cover frequencies from 83 Hz to 170 kHz, which will accommodate the resonant frequencies in the system; these are provided in Table I. The optimum number of decomposition levels is determined as presented in [31].

V. Feature Extraction

The second norm (referred to as the norm) of the detail coefficients $\|d_j(k)\|$, which were extracted at different frequency bands, is used as an input data vector for a classifier. The norm of the decomposed signal is calculated as

$$norm_{d_j} = \|d_j\| = \sqrt{\sum_{i=1}^{n} |d_{ji}|^2} \tag{3}$$

where j represents the level of decomposition and the maximum level is N, and d_j is the detail coefficient with n elements at level j. The proposed feature vector x can be mathematically represented as

$$\begin{aligned} x &= \begin{bmatrix} x_1 & \dots & x_j & \dots & x_N \end{bmatrix}^T \\ &= \begin{bmatrix} \|d_1\| & \dots & \|d_j\| & \dots & \|d_N\| \end{bmatrix}^T \end{aligned} \tag{4}$$

Firstly, the error vector is calculated, which is the difference of the feature vector for normal and the faulted conditions. If the error vector is above a predefined threshold value (i.e., if the norm is 500), then the fault location algorithm is triggered. Secondly, to find whether the fault has occurred on PV panel, or DC side cable, or AC side cable, the features of the DC converter positive voltage could be looked at to identify if the fault has happened on the PV panels, or else at the mid-point voltage of the DC-DC converter to determine whether the fault has happened on the DC or AC side cable. Normally the AC distribution system will have its own fault detector so if the DC-DC mid-point voltage and AC fault detector confirm a fault then the fault has happened on the AC side; otherwise it must be on the DC cable.

VI. Classifier Based on ANNs

In this study, a three-layer feed-forward network was used for the fault locator. The number of neurons in the input layer (q) is the same size as the feature vector extracted from the WT-based MRA technique (i.e., $q = 11$). Similarly, the number of neurons in the output layer (m) is considered equal to the number of fault locations to be analyzed (in this paper, $m = 14$ for cable faults and $m = 7$ for panel faults). After analysis with various hidden neuron numbers, feed-forward networks with 18 hidden neurons were found to be sufficient for this fault location technique. Tan-sigmoid and log-sigmoid activation functions are utilized in the hidden layer and output layer of the network, respectively. The supervised Levenberg-Marquardt back-propagation (*trainlm*) training algorithm is selected due to its faster convergence rate and efficiency in provding good results when compared with other optimization techniques [32]. The training of the ANN is repeated until a predetermined mean squared error (MSE), 0.2%, is achieved.

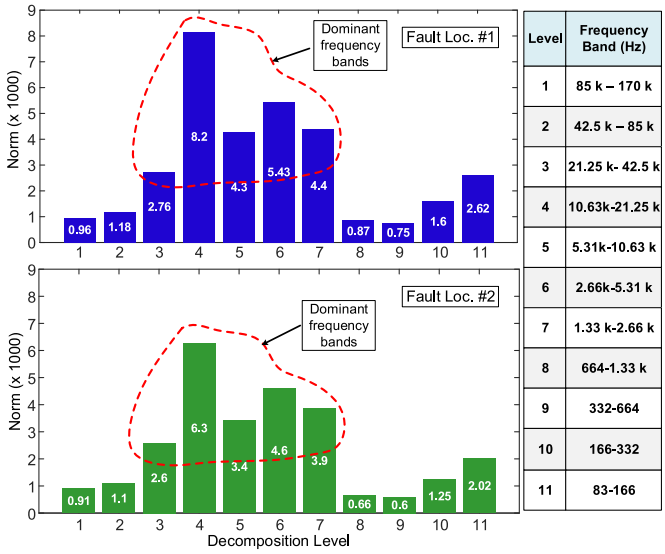
VII. Simulation Results

A. Cable Faults

The underground cables used in large PV systems are susceptible to ground faults or line-line faults. These faults are usually due to failure of cable insulation or accidental short circuits between conductors with different polarities.

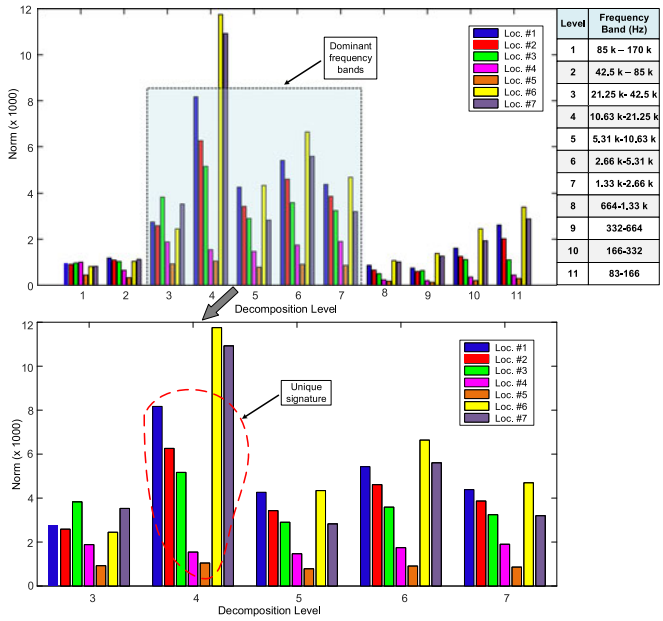
1) Results of Feature Extraction With DWT-based MRA: The fault locating method developed was first analyzed by simulating the PV system at standard test conditions (STC) with an irradiance of 1000 W/m^2, temperature of 25 °C, humidity of 50%, and fault resistance of 10 mΩ. The proposed fault location scheme is evaluated for low-resistance faults ranging from 0 to 0.5 Ω [32]. The mid-point voltage of DC-DC converter 1 with respect to grounding (V_{mid1}) was analyzed using WT-based MRA with a db10 mother wavelet. This signal is recorded with a sampling frequency of 340 kHz and decomposed into 11 resolution levels as shown in Table I. Using the wavelet MRA decomposition results of V_{mid1} for the ground faults at locations 1 and 2, norm values are calculated in each frequency band and plotted as shown in Fig. 5. A difference of approximately 1000 norms at levels 3 through 7 is evident between the two fault locations. These differences in the norms reveal that the extracted features represent unique signatures for classification of ground faults at different locations.

Fig. 6 shows the feature patterns for ground faults at locations 1 to 7 for a specific PV system operating condition. In the dominant frequency bands shown in the highlighted portion of Fig. 6, the norm values for each location are significantly different and can be used as input vectors for the ANN classifier. In particular, a large difference in the norms at decomposition level 4 is noted because the bandwidth (10.625 to 21.25 kHz) of this level covers the dominant resonant frequency of the noise signal. Moreover, resolution levels 6 and 7 also provide clear distinguishable patterns of the norms for faults at different locations. These bandwidths cover the switching frequency of the inverter (3 kHz) and DC-DC converters (1.8 kHz).



Level	Frequency Band (Hz)
1	85 k – 170 k
2	42.5 k – 85 k
3	21.25 k- 42.5 k
4	10.63k-21.25 k
5	5.31k-10.63 k
6	2.66k-5.31 k
7	1.33 k-2.66 k
8	664-1.33 k
9	332-664
10	166-332
11	83-166

Fig. 5. Feature vector extracted for ground fault at Loc. # 1 and Loc. # 2.

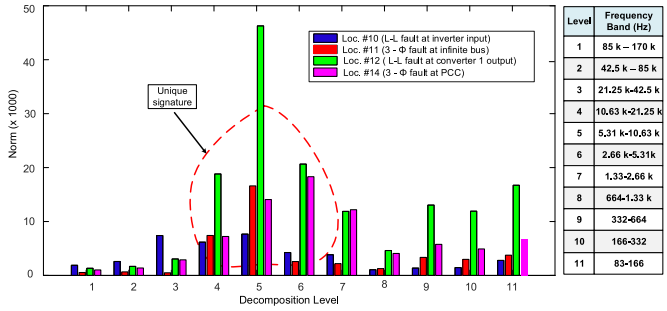


Level	Frequency Band (Hz)
1	85 k – 170 k
2	42.5 k – 85 k
3	21.25 k- 42.5 k
4	10.63 k-21.25 k
5	5.31 k-10.63 k
6	2.66 k-5.31 k
7	1.33 k-2.66 k
8	664-1.33 k
9	332-664
10	166-332
11	83-166

Fig. 6. Feature patterns for ground faults at seven different locations.

In a similar manner, a unique signature of line-line fault locations at different frequency bands can be obtained as shown in Fig. 7, and patterns corresponding to this feature can be used to determine the line-line fault locations for the DC and AC sides of the PV system. The feature patterns for 3-phase faults at the PCC and infinite bus are considerably different and can be utilized to find the location of faults on the AC side of PV system.

2) Results of the ANN Classifier: 625 test cases for fault and no fault conditions are generated by randomly varying the power output of the PV array (i.e., irradiance level), temperature, parasitic elements of PV panels (which depend on humidity),



Level	Frequency Band (Hz)
1	85 k – 170 k
2	42.5 k – 85 k
3	21.25 k-42.5 k
4	10.63 k-21.25 k
5	5.31 k-10.63 k
6	2.66 k-5.31k
7	1.33-2.66 k
8	664-1.33 k
9	332-664
10	166-332
11	83-166

Fig. 7. Feature patterns for line-line faults on the DC and AC sides of PV system.

and fault resistance. The proposed scheme was tested with five different values for each of these variables ($5\times5\times5\times5=625$ test cases). The different values of the parameters used for the simulations are given below:

Solar irradiance: 250, 550, 700, 900 and 1000 W/m^2 which varies the output power of the PV array from 200 to 500 kW. Temperature: 10, 15, 20, 25 and 30 °C

Fault resistances: 10 mΩ, 50 mΩ, 0.1 Ω, 0.2 Ω, and 0.5 Ω

PV parasitic capacitance (C_{lek}), series resistance (R_s), and parallel resistance (R_p) values are given below in parenthesis for different humidity levels [27]:

Humidity: 30 % - { 40.08 μF, 6.23 MΩ, 220.11 MΩ }
40 % - { 3.675 μF, 59.27 MΩ, 592.71 MΩ }
50 % - { 33.087 μF, 8.41 MΩ, 252.38 MΩ }
60 % - { 55.125 μF, 4.77 MΩ, 190.825 MΩ }
70 % - { 169.13 μF, 0.104 MΩ, 2.0698 MΩ }

In this paper, 13 fault locations and a no fault condition for a total of 14 possibilities, are considered. For each of these 14 possibilities, 625 test cases are simulated for a total of $14\times625=8{,}750$ different cases. The feature vectors are generated for each case to utilize as inputs to the ANN classifier. 80% of the 8,750 cases are used for ANN training and the rest of the data are used for testing and validation of the fault locating algorithm. The outputs of the ANN vary from 0 to 1 due to use of a log-sigmoid activation function in the output layer. Because the outputs of a multi-layer neural network rarely give exactly the target of 0 or 1 for each output neuron, a decision threshold output level of 10% is built into the ANN, i.e., outputs <0.1 are classified as normal conditions and outputs >0.9 are considered as faulty cases.

Table II illustrates the fault location results for cable faults in PV systems. A 99.92% correct fault location rate is obtained for ground faults, whereas a 97.32% success rate is found for line-line faults. Thus, the overall fault locating accuracy of the proposed fault location method for cable faults in PV systems is 99.2%. For line-line faults at DC cables, no circulatory loop is formed by the fault point and the parasitic elements of the PV modules and cables. Hence, only a small amount of high frequency noise due to leakage currents to ground would be produced, and this is what was used to locate the line-line fault locations. In addition, the features of the short circuit faults at locations 10 and 13 are somewhat similar to one other. Therefore,

TABLE II
FAULT LOCATION RESULTS FOR CABLE FAULTS IN PV SYSTEM

Fault location	Accuracy	Average	Overall accuracy
No fault	100 %	99.92 %	99.2 %
Loc. #1	99.2 %		
Loc. #2	100 %		
Loc. #3	100 %		
Loc. #4	100 %		
Loc. #5	100 %		
Loc. #6	100 %		
Loc. #7	100 %		
Loc. #8	100 %		
Loc. #9	100 %		
Loc. #10	94.24 %	97.32 %	
Loc. #11	100 %		
Loc. #12	100 %		
Loc. #13	95.04 %		

the fault location accuracy for DC line-line faults is marginally lower compared to the ground fault location accuracy.

The performance of the proposed method was also compared to alternate approaches. Bender's [2] method is one approach used in the industry for ungrounded PV systems. It injects a high frequency signal between the cable and the ground, and manually uses a portable current probe on the cable path to trace this high frequency signal up to the ground fault point. However, the percent accuracy for pin-pointing the fault location using this approach was not reported in [2]. A differential current based fault location scheme for finding cable faults in PV systems has a reported accuracy between 93.75 and 99.9% [33]. Using a fuzzy rule-based method for detecting faults, [34] detected 90% of fault conditions including noisy conditions. The results obtained from the proposed wavelet-ANN approach have better or comparable accuracy to these approaches.

B. PV Panel Faults

Ground and line-line type faults (intra-string and cross-string faults) are the most common type of faults in PV systems. Some of the utility engineers report that about 1% of installed PV panels fail annually. Fig. 8 shows the PV module configurations of 250 kWp PV array 1 of the PV test system shown in Fig. 1. nEach PV array consists of 48 parallel strings with 17 modules in each string. To demonstrate the fault location method in PV farms, two strings are separated and the rest are treated as a PV array of 46 strings. Each string is divided into four sections, with one group of five panels and three groups of four panels. Ground faults (Loc. #$P1$, Loc. #$P2$, and Loc. #$P3$), line-line faults (Loc. #$P4$, Loc. #$P5$, and Loc. #$P6$), and arcing faults (Loc. #$P7$ and Loc. #$P8$) are simulated as shown in Fig. 8. The arc fault model used in this paper is based on a differential equation of the arc conductance described in [35].

Due to limitation in the simulation capability of the RTDS in the lab, PV panels could not be modeled in small time-steps (2.94 μs). Therefore, the PV farms shown in Fig. 8 were simulated in large time-steps (50 μs) and the electrical information transferred to small time-step simulations. The input positive

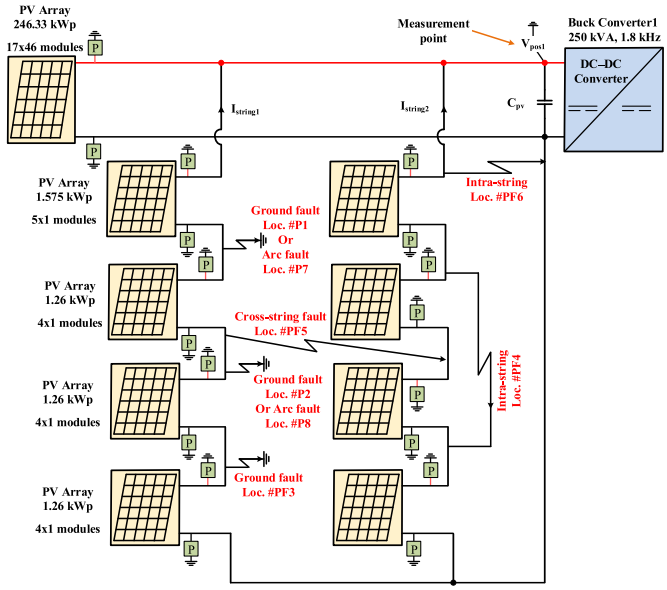


Fig. 8. Different types of faults in PV farms.

TABLE III
DIFFERENT DECOMPOSITION LEVELS FOR PV PANEL FAULTS

Decomposition Level	Frequency Band (Hz)
1	5 k–10 k
2	2.5 k–5 k
3	1.25 k–2.5 k
4	625–1.25 k
5	312.5–625
6	156.25–312.5
7	78.125–156.25
8	39–78.125
9	19.5–39

terminal voltage with respect to grounding of buck converter #1, V_{pos1}, was measured and analyzed using DWT-based MRA with a db10 mother wavelet. This signal was recorded with a sampling frequency of 20 kHz and decomposed into nine resolution levels covering frequencies from 195 Hz to 10 kHz, as shown in Table III.

For standard operating conditions of the PV system, norm values of eight fault locations are extracted for each frequency band and patterns are plotted in Fig. 9. The bar diagram shows the large difference in the norms for different fault locations for frequency bands 1 to 6, which makes the proposed fault location method practical for PV faults. Typically, decomposition levels 2 and 3 cover switching frequencies of inverter and buck converters, respectively. Also, level 6 contains the third, fourth, and fifth harmonics of the fundamental power frequency (60 Hz) and provides clear differentiation in norm values for different fault locations.

As in the case of cable faults, 625 test cases were simulated for each fault location with different system parameters. In this case, 0.5, 2.5, 5, 7.5 and 10 Ω arc fault resistances are used to simulate arcing faults in the PV array [8]. Thus, the total number

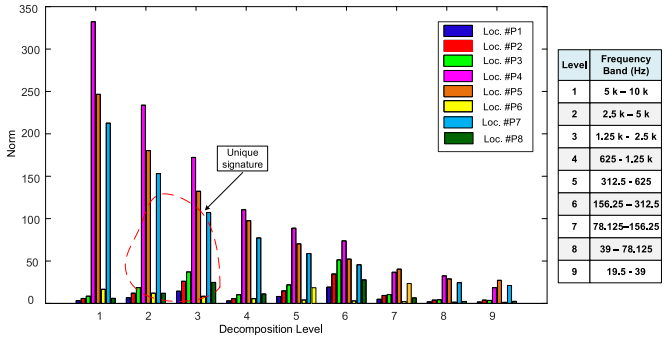


Fig. 9. Feature patterns for faults in PV modules.

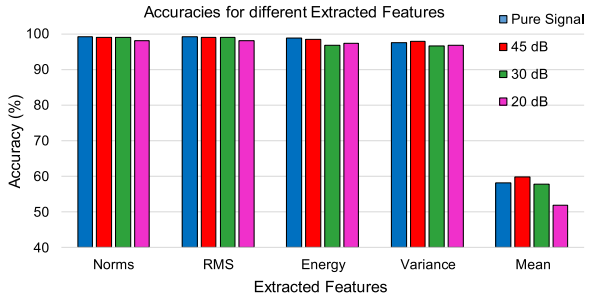


Fig. 10. Accuracy of different extracted features.

TABLE IV
FAULT LOCATION RESULTS FOR PV MODULE FAULTS

Fault location	Accuracy	Average	Overall accuracy
No fault	100 %	98.4 %	97.4 %
Loc. $\#PF1$ (ground fault)	99.7 %		
Loc. $\#PF2$ (ground fault)	94.2 %		
Loc. $\#PF3$ (ground fault)	99.7 %		
Loc. $\#PF4$ (intra-string)	95.5 %	97.4 %	
Loc. $\#PF5$ (cross-string)	96.5 %		
Loc. $\#PF6$ (intra-string)	100 %		
Loc. $\#PF7$ (arc fault)	92.2%	96.5 %	
Loc. $\#PF8$ (arc fault)	97.2 %		

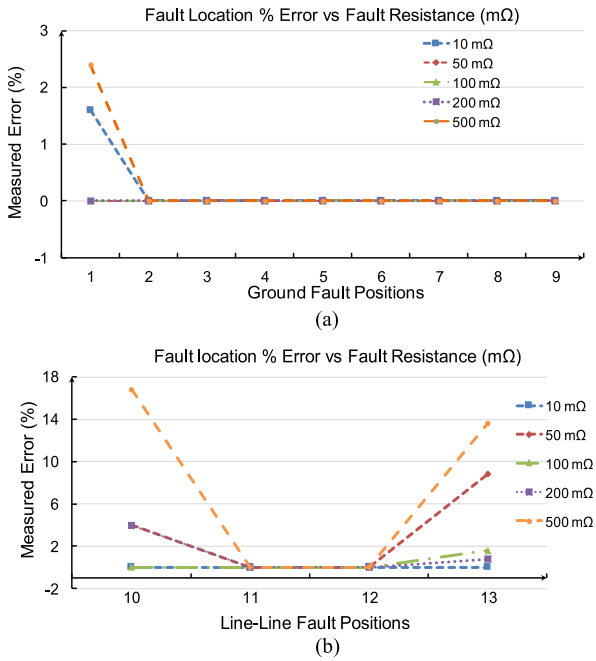


Fig. 11. Effect of fault resistance during (a) ground faults and (b) line-line faults.

of test scenarios for simulated PV panel faults was 5,625. The ANN classifier with none input, 18 hidden, and eight output neurons was trained until the MSE was $< 0.6\%$. The proposed algorithm can correctly classify 98.4% of ground faults, 97.4% of line-line faults, and 96.5% of arc faults in the PV array as shown in Table IV. An overall fault location accuracy of 97.4% can be obtained for PV panel faults with the proposed fault location method.

VIII. SENSITIVITY ANALYSIS

A. *Performance of Classifier for Different Noise Conditions*

The fault location performance was tested with input signals containing various signal to noise ratio (SNR) values between 20 and 50 dB. A typical SNR value of 30 dB is equivalent to a peak noise magnitude of nearly 3.5% of the voltage signal [36]. The performance of the fault locator for the cable and PV faults in noisy conditions was analyzed. However, only the results for cable faults are discussed. Fig. 10 shows that although the accuracy decreases with the noise level, a satisfactory fault locating success rate can be achieved even though the input signals contain different levels of noise. An overall accuracy of 98.10% can be still obtained even with a high noise level of 20 dB. A comparison of the fault location accuracies with the various features extracted from norms, energy, RMS, variance, and mean was undertaken, and the results also presented in Fig. 10. The proposed norm-based approach has the best accuracy and is robust for different noise levels.

B. *Variation of Fault Resistance*

Fault resistance is in the path of grounding oscillatory circuits, and will impact the generated noise patterns. Simulation results show that the resistance of cables for frequencies in the hundreds of kHz is approximately 300 mΩ for 30 m of 1 kV, 250 MCM (thousands of circular mils) cable [12]. For the bolted fault types (resistances on the order of several mΩ) or faults with small resistance values less than 100 mΩ, the loop resistance is still going to be dominated by the cable resistance. Therefore, fault resistances <100 mΩ do not change the noise level very much. In contrast, when the fault involves high resistance, the degree of its impact on the noise level largely depends on the specific fault resistance. Fig. 11 clearly shows that the fault location error does not exceed 2.4% in the worst case condition for the ground faults at 500 mΩ. The fault locator is more immune to changes in fault resistance (0 to 300 mΩ) for ground faults but marginally less immune for line-line faults.

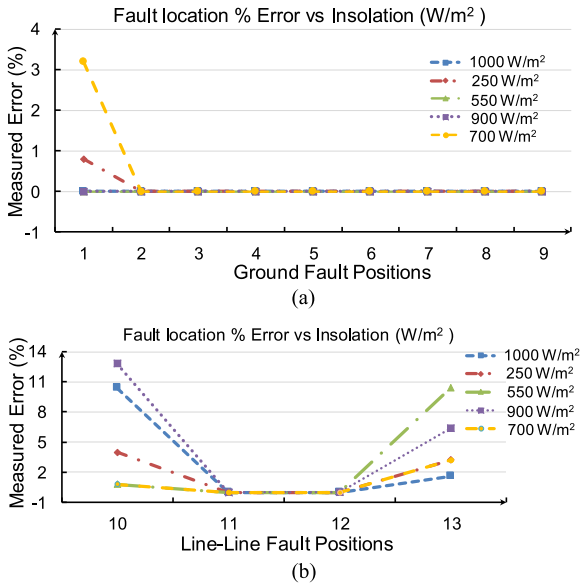


Fig. 12. Effect of insolation during (a) ground faults and (b) line-line faults.

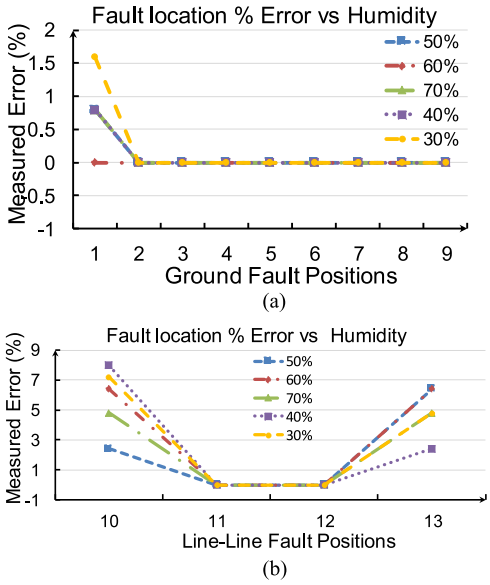


Fig. 13. Effect of humidity during (a) ground faults and (b) line-line faults.

C. Variation of Power Generation

Basically, changing the output power will change the amount of current to be switched on and off by the switching operation of power converters, which acts like a high frequency noise generator. For a fault at a given location, varying the insolation levels will only change the magnitude of the measured noise but the noise pattern will remain the same. Interestingly, the fault location error is almost 0% for ground faults in all locations, except for Loc. #1 where it is 3.2% with a change in solar insolation, as shown in Fig. 12. However, in the case of line-line faults, the fault location error increases to 12.8% at 900 W/m^2 for a line-line fault at Loc. #10.

D. Variation of PV Parasitic Elements

Variations in humidity play a significant role in parasitic capacitance. Five different values of parasitic elements were obtained by varying humidity from 30 to 70% and used to study the effect of parasitic elements on the performance of the fault locator. Fig. 13(a) demonstrates that the fault location error for ground faults does not exceed 1.6% in the worst case condition. On the other hand, the measured error at low humidity (about 30–40%) does not go beyond 8% for a short circuit fault at the inverter input terminal (Loc. #10).

E. Switching Frequency of Converters

Fig. 14 shows a case when the switching frequency of converter 1 increased from 1.8 to 2.4 kHz while that of the inverter and another converter remained the same. The norms slightly increased without changing the pattern of the features for each fault location. Hence, with varying switching frequencies, the feature vectors will provide nearly similar characteristics at different MRA levels and not affect the performance of the fault locator very much.

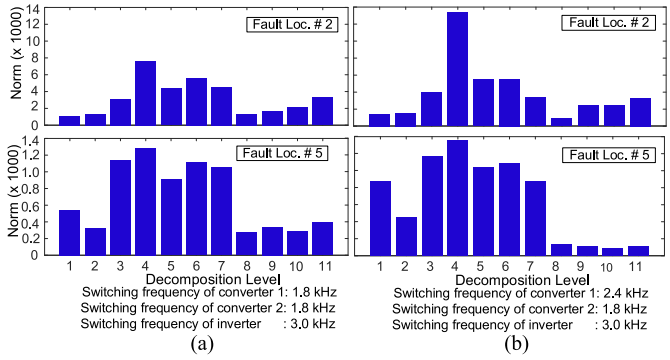


Fig. 14. Effect of switching frequencies.

IX. Conclusion

A fault location scheme for ungrounded PV systems has been presented. The ungrounded PV system was subjected to various faults (line to ground, line-line and arc faults) on both the DC and AC sides of the system. The DWT-based MRA technique was studied to extract the unique features of the noise signal measured at the mid-point of the DC-DC converter. Then, a three-layer feed-forward ANN was adopted as a classifier of the pattern recognition to identify exact fault locations for the cables and PV modules of the ungrounded PV systems. Compared to existing methods in the literature, the proposed method does not require an external high frequency signal injector, reduces the number of voltage sensors, and can locate the fault while the system remains operational. Finally, the results show the proposed method is promising, accurate, and robust with respect to changes in operating parameter values and with noisy inputs.

APPENDIX A
500 KWP UNGROUNDED PV SYSTEM DATA

PV Data (1000 W/m^2, 25 °C)	Value	
Max. Power (P_{max})	315 W	
Max. power voltage (V_{mp})	39.8 V	
Max. power current (I_{mp})	7.92 A	
Open circuit voltage (V_{oc})	49.2 V	
Short circuit current (I_{oc})	8.5 A	
Series cells/module (N_c)	72	
Parallel strings of cells (N_{cp})	1	
I_{sc} temp. coeff.J_{tmp}	0.061 %/degC	
V_{sc} temp. coeff.K_v	−0.36 %/degC	
OC series resistance (R_{so})	0.5 Ω	
SC shunt resistance (R_{sho})	100 Ω	
Series modules in a string (N_s)	17	
Parallel strings (N_p)	48	
Parasitic Elements (STC) [27]	**Value**	
	250 kWp	**1.576 kWp**
R_{iso}	0.3264 MΩ	51.77 MΩ
C_{lek}	33.0875 μF	0.208 μF
R_s	8.41 MΩ	6.68 MΩ
R_p	252.38 MΩ	200.4 MΩ

REFERENCES

[1] T. Baldwin and F. Renovich, "Analysis of fault locating signals for high-impedance grounded systems," in *Proc. IEEE Ind. Appl. Conf. 36th IAS Annu. Meeting*, Sep. 2001, vol. 3, pp. 1823–1830.

[2] Bender, "Ground fault protection photovoltaic systems," Brochure no. NAE1111041, 2012. [Online]. Available: http://www.bender-us.com/documents/Solar_brochure_NAE1111041.pdf

[3] J.-D. Park, "Ground fault detection and location for ungrounded dc traction power systems," *IEEE Trans. Veh. Technol.*, vol. 64, no. 12, pp. 5667–5676, Dec. 2015.

[4] E. O. Schweitzer, A. Guzmn, M. V. Mynam, V. Skendzic, B. Kasztenny, and S. Marx, "Locating faults by the traveling waves they launch," in *Proc. 67th Annu. Conf. Protective Relay Eng.*, Mar. 2014, pp. 95–110.

[5] S. Marx, B. K. Johnson, A. Guzmán, V. Skendzic, and M. V. Mynam, "Traveling wave fault location in protective relays: Design, testing, and results," in *Proc. 16th Annu. Georgia Tech Fault Disturbance Anal. Conf.*, Atlanta, GA, USA, May 2013, pp. 1–15.

[6] O. K. Nanayakkara, A. D. Rajapakse, and R. Wachal, "Location of dc line faults in conventional HVDC systems with segments of cables and overhead lines using terminal measurements," *IEEE Trans. Power Del.*, vol. 27, no. 1, pp. 279–288, Jan. 2012.

[7] L. Schirone, F. P. Califano, U. Moschella, and U. Rocca, "Fault finding in a 1 mw photovoltaic plant by reflectometry," in *Proc. IEEE 1st World Conf. Photovolt. Energy Convers./24th IEEE Photovolt. Spec. Conf.*, Dec. 1994, vol. 1, pp. 846–849.

[8] M. K. Alam, F. H. Khan, J. Johnson, and J. Flicker, "PV arc-fault detection using spread spectrum time domain reflectometry (SSTDR)," in *Proc. IEEE Energy Convers. Congr. Expo.*, Sep. 2014, pp. 3294–3300.

[9] M. T. Hagh, K. Razi, and H. Taghizadeh, "Fault classification and location of power transmission lines using artificial neural network," in *Proc. Int. Power Eng. Conf.*, Dec. 2007, pp. 1109–1114.

[10] Syafaruddin, E. Karatepe, and T. Hiyama, "Controlling of artificial neural network for fault diagnosis of photovoltaic array," in *Proc. 16th Int. Conf. Intell. Syst. Appl. Power Syst.*, Sep. 2011, pp. 1–6.

[11] L. L. Jiang and D. L. Maskell, "Automatic fault detection and diagnosis for photovoltaic systems using combined artificial neural network and analytical based methods," in *Proc. Int. Joint Conf. Neural Netw.*, Jul. 2015, pp. 1–8.

[12] Y. Pan, "Noise pattern analysis based ground fault location approach for ungrounded electrical distribution systems," Ph.D. dissertation, Florida State Univ., Tallahassee, FL, USA, 2009.

[13] Y. Pan, M. Steurer, and T. L. Baldwin, "Ground fault location testing of a noise-pattern-based approach on an ungrounded dc system," *IEEE Trans. Ind. Appl.*, vol. 47, no. 2, pp. 996–1002, Mar. 2011.

[14] Y. Pan, P. Silveira, M. Steurer, T. Baldwin, and P. Ribeiro, "Method for locating phase to ground faults in dc distribution systems," U.S. Patent 8 067 942, Nov. 29, 2011.

[15] W. Li, A. Monti, and F. Ponci, "Fault detection and classification in medium voltage dc shipboard power systems with wavelets and artificial neural networks," *IEEE Trans. Instrum. Meas.*, vol. 63, no. 11, pp. 2651–2665, Nov. 2014.

[16] M. Majidi, A. Arabali, and M. Etezadi-Amoli, "Fault location in distribution networks by compressive sensing," *IEEE Trans. Power Del.*, vol. 30, no. 4, pp. 1761–1769, Aug. 2015.

[17] M. Majidi, M. Etezadi-Amoli, and M. S. Fadali, "A novel method for single and simultaneous fault location in distribution networks," *IEEE Trans. Power Syst.*, vol. 30, no. 6, pp. 3368–3376, Nov. 2015.

[18] I. Dzafic, R. Jabr, S. Henselmeyer, and T. Donlagic, "Fault location in distribution networks through graph marking," *IEEE Trans. Smart Grid*, vol. pp, no. 99, Jul. 2016, doi: 10.1109/TSG.2016.2587583, to be published.

[19] A. A. P. Bscaro, R. A. F. Pereira, M. Kezunovic, and J. R. S. Mantovani, "Integrated fault location and power-quality analysis in electric power distribution systems," *IEEE Trans. Power Del.*, vol. 31, no. 2, pp. 428–436, Apr. 2016.

[20] A. N. Milioudis, G. T. Andreou, and D. P. Labridis, "Detection and location of high impedance faults in multiconductor overheadDistribution lines using power line communication devices," *IEEE Trans. Smart Grid*, vol. 6, no. 2, pp. 894–902, Mar. 2015.

[21] P.-C. Chen, V. Malbasa, Y. Dong, and M. Kezunovic, "Sensitivity analysis of voltage sag based fault location with distributed generation," *IEEE Trans. Smart Grid*, vol. 6, no. 4, pp. 2098–2106, Jul. 2015.

[22] M. Pignati, L. Zanni, P. Romano, R. Cherkaoui, and M. Paolone, "Fault detection and faulted line identification in active distribution networks using synchrophasors-based real-time state estimation," *IEEE Trans. Power Del.*, vol. 32, no. 1, pp. 381–392, Feb. 2017.

[23] W. Santos, F. Lopes, N. Brito, and B. Souza, "High-impedance fault identification on distribution networks," *IEEE Trans. Power Delivery*, vol. 32, no. 1, pp. 23–32, Feb. 2017.

[24] A. Nikander and P. Järventausta, "Identification of high-impedance earth faults in neutral isolated or compensated MV networks," *IEEE Trans. Power Del.*, vol. 32, no. 3, pp. 1187–1195, Jun. 2017.

[25] D. B. Jason Fisher, "Ungrounded PV power systems in the NEC," *Solar Pro*, vol. 5.5, pp. 1–4, Aug./Sep. 2012. [Online]. Available: http://solarprofessional.com/print-issue/august-september-2012

[26] M. Villalva, J. Gazoli, and E. Filho, "Comprehensive approach to modeling and simulation of photovoltaic arrays," *IEEE Trans. Power Electron.*, vol. 24, no. 5, pp. 1198–1208, May 2009.

[27] J. Hernández, P. Vidal, and A. Medina, "Characterization of the insulation and leakage currents of PV generators: Relevance for human safety," *Renew. Energy*, vol. 35, no. 3, pp. 593–601, 2010.

[28] A. Fernandes and W. Neves, "Phase-domain transmission line models considering frequency-dependent transformation matrices," *IEEE Trans. Power Del.*, vol. 19, no. 2, pp. 708–714, Apr. 2004.

[29] S. G. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 11, no. 7, pp. 674–693, Jul. 1989.

[30] Manitoba HVDC Research Centre, *dcLFL Line Fault Location System*, Manitoba HVDC Research Centre Brochure, Winnipeg, MB, Canada, vol. 1, Jul. 2013. [Online]. Available: https://hvdc.ca/uploads/ck/files/Brochure%20dcLFL%20Oct%2018%202013%20V7.pdf

[31] J. Antonino-Daviu, M. Riera-Guasp, J. Roger-Folch, F. Martnez-Gimnez, and A. Peris, "Application and optimization of the discrete wavelet transform for the detection of broken rotor bars in induction machines," *Appl. Comput. Harmonic Anal.*, vol. 21, no. 2, pp. 268–279, 2006.

[32] N. K. Chanda and Y. Fu, "ANN-based fault classification and location in MVDC shipboard power systems," in *Proc. North Amer. Power Symp.*, Aug. 2011, pp. 1–7.

[33] S. Dhar, R. K. Patnaik, and P. K. Dash, "Fault detection and location of photovoltaic based dc microgrid using differential protection strategy," *IEEE Trans. Smart Grid*, vol. pp, no. 99, Jan. 2017, doi: 10.1109/TSG.2017.2654267, Early access.

[34] P. Ducange, M. Fazzolari, B. Lazzerini, and F. Marcelloni, "An intelligent system for detecting faults in photovoltaic fields," in *Proc. 11th Int. Conf. Intell. Syst. Des. Appl.*, 2011, pp. 1341–1346.

[35] M. Kizilcay and T. Pniok, "Digital simulation of fault arcs in power systems," *Int. Trans. Elect. Energy Syst.*, vol. 1, no. 1, pp. 55–60, 1991.

[36] M. Uyar, S. Yildirim, and M. T. Gencoglu, "An effective wavelet-based feature extraction method for classification of power quality disturbance signals," *Elect. Power Syst. Res.*, vol. 78, no. 10, pp. 1747–1755, 2008.

Indra Man Karmacharya (S'09–M'16) received the B.E. degree with distinction in electrical engineering from Tribhuvan University, Kathmandu, Nepal, in 2009, and the M.Sc. degree in electrical engineering from the University of Saskatchewan, Saskatoon, SK, Canada, in 2016. During 2009–2013, he was an Operation and Maintenance Engineer with Bhotekoshi Power Company Pvt. Ltd., Kathmandu, Nepal. He is currently an interconnection planning Engineer at Saskatchewan Power Corporation, Regina, SK, Canada. His research interests include power system protection and renewable energy integrated systems.

Ramakrishna (Rama) Gokaraju (S'88–M'00–SM'15) received the B.E. degree with distinction in electrical and electronics engineering from the Regional Engineering College (National Institute of Technology), Trichy, India, in 1992, and the M.Sc. and Ph.D. degrees in electrical and computer engineering from the University of Calgary, Calgary, AB, Canada, in 1996 and 2000, respectively. During 1992–1994, he was a Graduate Engineer at the Larsen & Toubro-ECC, India; a Project Associate at the Indian Institute of Technology, Kanpur, India; and a Research Engineer at the Regional Engineering College, Rourkela, India. During 1999–2002, he was a Research Scientist at the Alberta Research Council and a Staff Software Engineer at IBM Toronto Lab. He joined the Department of Electrical and Computer Engineering at the University of Saskatchewan as an Assistant Professor in 2003, received Tenure and Associate Professorship in 2009, and became a Professor in 2015. His current research works are in high-speed digital relaying, renewable energy systems, wide-area-based power system protection and control, and real-time simulation approaches.