

# Sensitivity Analysis of Voltage Sag Based Fault Location With Distributed Generation

Po-Chen Chen, *Student Member, IEEE*, Vuk Malbasa, *Member, IEEE*,  
Yimai Dong, *Member, IEEE*, and Mladen Kezunovic, *Fellow, IEEE*

**Abstract**—The presence of distributed generation (DG) in distribution networks may seriously affect accuracy of the voltage sag based fault location method. An approach toward quantifying the adverse effect of DG on the fault location calculation is described. A series of realistic scenarios is used to illustrate how DG impacts synchrophasor measurements during disturbances. Alternative Transients Program-Electromagnetic Transients Program models are used to obtain steady-state solutions in the time domain, while Sobol's approach to sensitivity analysis is used to quantify the effect of DG and imperfections of measuring instruments on fault location. Various test cases reveal that DG may adversely affect the voltage characteristic and therefore the accuracy of voltage sag based fault location.

**Index Terms**—Distributed power generation, EMTP, fault location, power distribution faults, power system measurements, power system protection, power system simulation, search methods, sensitivity analysis, smart grids.

## I. INTRODUCTION

IN DISTRIBUTION networks, outage management, and service restoration are closely related to the operator's ability to reliably identify fault location [1], [2]. The challenges of accurate fault location in distribution networks stem from the large number of components, changes in loading conditions over time, the unbalanced nature of the system, heterogeneous lines, and the presence of laterals and load taps.

Fault location approaches may be classified into the following sub-categories according to which underlying feature of the fault measurements is exploited: 1) superimposed components methods [3], [4]; 2) intelligent system methods [5], [6]; 3) impedance based methods [7]–[9]; 4) traveling wave methods [10]–[12]; and 5) voltage sag based methods [13]–[15]. A more detailed comparison of different fault location methods may be found in [7] and [15]–[17]. It was also shown that the abundance of data collected from field measurements may be leveraged for precise fault location [2]. The distinctive feature of voltage sag based fault location is the capability of identifying an unambiguous fault location,

Manuscript received May 12, 2014; revised September 22, 2014 and December 11, 2014; accepted December 24, 2014. Date of publication January 20, 2015; date of current version June 18, 2015. Paper no. TSG-00405-2014.

P.-C. Chen and M. Kezunovic are with Texas A&M University, College Station, TX 77843-3128 USA (e-mail: pchen01@tamu.edu).

V. Malbasa is with the University of Novi Sad, Novi Sad 21000, Serbia.

Y. Dong is with Electrocon, Inc., Ann Arbor, MI 48103 USA.

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/TSG.2014.2387153

whereas impedance based methods may not be able to differentiate the exact feeder/lateral.

The future distribution system is predicted to heavily rely on distributed generation (DG) for uninterrupted load support, and therefore a strategy-based around disconnecting DG would either make the system unreliable or inefficient. Because of these considerations, fault location techniques in the presence of DG [18]–[20] have been investigated. Brahma [18] used synchronized voltage and current measurements at the interconnection points of DG units to identify the faulted area and a precise fault location. Brahma and Girgis [19] developed a terminology to identify the faulted area using a superposition-based approach and precalculated short circuit results. Johnsonbaugh [20] demonstrated a detailed formulation of the variation of the apparent impedance method involving DG. Comprehensive literature reviews regarding fault location techniques in the presence of DG are available in [18]–[20]. The effects of DG on voltage sags, and therefore applications relying on this property have been recognized in [21] and [22]. The level of DG penetration in a network may also influence the voltage characteristics. Other studies in fault location under high DG penetration fail to distinctly, statistically quantify the effects of DG as well as to distinguish the impact of DG penetration from external factors such as measurement imperfections [23], fault resistance, and loading estimate.

This paper describes how to quantitatively assess the impact of DG and external measurement errors on the performance of voltage sag based fault location. This paper not only statistically analyzes the quantifiable impact of each factor influencing fault location, but also considers their total effects in combination with other factors as well. The authors have been unable to locate any systematic studies regarding this important topic in the literature.

The rest of this paper is organized as follows. Section II introduces the background of the problem and methodology. Section III indicates formulation and implementation of the studied voltage sag based fault location method. Section IV describes the formulation of sensitivity analysis. Section V demonstrates the validation results of the studied voltage sag based fault location algorithm. Section VI shows the results and related analysis. Section VII contains the conclusion.

## II. BACKGROUND

While keeping in mind how DG may aggravate fault location techniques, it seems imperative to quantify the effect of

TABLE I  
DEVELOPMENT ENVIRONMENT: "X" INDICATES USE

Software	DG Modeling	Fault Simulation	Fault Location	Sensitivity Analysis
ATP-EMTP	X	X	X	
C++	X		X	
SimLab				X
Matlab				X

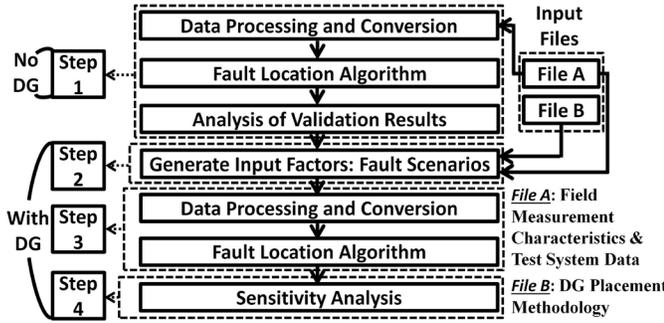


Fig. 1. Scope of study: fault location testing and evaluation function blocks.

errors of external factors on voltage sag based fault location in distribution systems. While using voltage sag based methods, voltage sag matching may become erratic in the presence of DG, where voltage sag matching at multiple locations may become similar (Section VI-B).

A. Study Design

In order to quantify the effect of so many different factors influencing fault location, it was necessary to interface multiple software tools and verify the study design and data flow. Software tools used in each step of the implementation are shown in Table I, where SimLab is a software tool suitable for global sensitivity and uncertainty analysis [24]. The studied voltage sag based fault location algorithm was implemented in C++, and interfaced with the alternative transient program (ATP) [25] software for fault simulation. In Fig. 1, the main steps of this study are illustrated including the validation at the top and experiments at the bottom.

In Step (1), the custom voltage sag based fault location software is validated by simulating faults on every receiving node of line components (Section V). In Step (2), a large number of fault scenarios are generated to form a knowledge base, ensuring reliable results from sensitivity analysis (Section IV). Then these scenarios are simulated in Step (3) (Section III) and then analyzed in Step (4) (Section VI).

B. Network Under Study

A real-world radial distribution network with underground lines was used in this paper [14] (Fig. 2). The total number of components is shown in Table II. The terminal nodes in Fig. 2 represent the locations of load and interconnection points for DG. The top ten largest loads are listed in Table III whereas indicated in Fig. 2. Due to the size of the network, only one part of it is selected for this paper, with other parts simplified as two equivalent models, so that the fault location

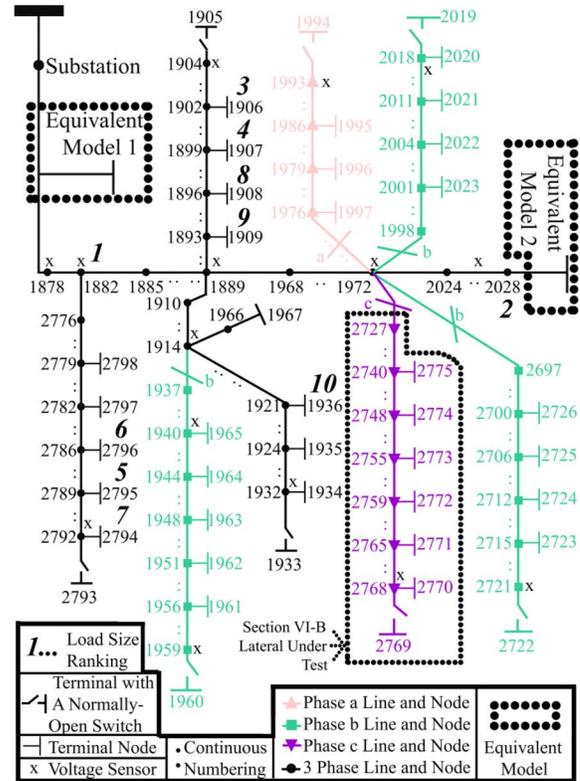


Fig. 2. Schematic of the test system [14].

TABLE II  
INFORMATION OF THE OVERALL SYSTEM

<b>Total Number of Components</b>	4352
<b>Number of Line Components</b>	1828
<b>Total length of Line Components (Foot)</b>	655617.6
<b>Total Connected Load (kVA)</b>	33606

TABLE III  
INFORMATION OF TOP TEN LOADS IN THE SYSTEM

<b>ID</b>	1878, 2028, 1906, 1907, 2795, 2796, 2794, 1908, 1909, 1936
<b>Size (kVA)</b>	25000, 6000, 500, 300, 300, 150, 150, 75, 75, 75

algorithm would focus on the area shown in Fig. 2 (detailed in Section II-D).

C. Details of Modeling and Simulation

In this study, all scenarios were simulated in the time-domain using ATP, which possesses the following advantages.

- 1) ATP is capable of simulating distribution network models realistically in the time domain: in this study, each load and DG was explicitly represented in ATP simulations.
- 2) Under- and over-voltage protection of DG can be modeled in detail: following the IEEE standard 1547 [26], the model of under- and over-voltage protections of DG was implemented as specified by local utilities [27], where the default voltage operating range for generators shall be within 88% to 110% of nominal voltage magnitude.

The loads were represented as constant impedance due to the following two reasons.

- 1) The constant impedance representation may suffice for the need of [28].
- 2) A detailed load modeling will prolong the overall simulation time [29].

The DGs were modeled as power flow control models as illustrated in [30], with unity power factor representing the inverter-based type of DG. Based on the results of the preliminary tests, the inverter output power follows the reference power commands with considerable accuracy as demonstrated in [30]. It is recognized that the voltage sag and short circuit level may vary due to different interfaces between the DG and grid [31], [32], and that the different load types may also affect the voltage magnitude at post-fault steady states in the presence of DG [29].

All the cases were simulated using a desktop having six processors which are Intel(R) Xeon(R) CPU W3670 operating at 3.20 GHz and installed memory (RAM) 12 GB. In ATP, all cases were simulated with time step of 80  $\mu$ s. The average running time for one scenario is 1.04 h.

#### D. Assumptions

1) *Meter Placement*: The optimization technique for meter placement is beyond the scope of this paper, and a generalized procedure is described in [15]. However, having more sensors may enhance the accuracy of fault location (Section VI-B). The meters (Fig. 2) are assumed to be available at 15 locations, including the intersections of subsections and the ends of feeder, where intelligent electronic devices (IEDs) are likely to be installed. The nodes with meters are: 1878, 1882, 1889, 1904, 1914, 1931, 1939, 1959, 1972, 1993, 2017, 2026, 2721, 2767, and 2791.

2) *DG Connection and Deployment*: DG was assumed to be connected at the load terminal, and each load was interconnected with a DG unit and the DG power was set to exactly the desired penetration for that scenario, as a ratio of the load power. For instance, if the DG penetration for a scenario was 10%, then the DG power was 10% of the load power at each load. The influence of probabilistic deployment of DG [33]–[35] is beyond the scope of this paper.

3) *Computational Complexity and Hardware Limitations*: Due to the large potential number of components in modern networks and the intensive computer modeling and simulation necessary for such study, the time necessary to perform fault location calculation is constrained by: 1) utilizing limited simulation and modeling hardware; 2) considering a large number of fault scenarios necessary for sensitivity analysis; and 3) simulating all types of faults at every node. Assuming the fault type is known to be single-line-ground fault, running one case with the original network takes 8 h on our computer (hardware specifications in Section II-C). Therefore, the simplified network is used and the fault type is assumed to be known. In this case, the running time is significantly reduced from 8 to 1 h. The preliminary test results reveal that the described methodology is available for the original network and capable of differentiating fault types

including single-line-ground, double-line-ground, double-line, triple-line-ground, and triple-line faults at each node.

In practice, the run time can be shortened by: 1) reducing the number of candidate fault nodes (Section III-A) based on outage mapping via utilizing big data [36] such as customer calls, asset management condition-based status [37], and automatic meter reading [38]; 2) utilizing better computational hardware; 3) separating the fault location calculation in various areas into different experiments running in parallel at the same time; and 4) applying data mining techniques to decrease computational complexity [39], [40].

### III. METHODOLOGY

#### A. Voltage Sag Based Fault Location Under Study

As described in [13], the fault location procedure consists of the following four key steps.

- 1) Voltage sag data ( $V_{\text{recorded}}$ ) is recorded at the meter locations and sent to the local distribution energy management center.
- 2) Simulated voltage sag data ( $V_{\text{calculated}}$ ) is computed, assuming in turn that the fault is located at each node and fault resistance is estimated based on the voltage match.
- 3) The node with the best match between  $V_{\text{recorded}}$  and  $V_{\text{calculated}}$  is the declared fault node.
- 4) Binary search (halving) [41] is used on the lines connected to the detected fault node to pinpoint the fault on the lines (this process will be detailed in Section III-B).

In Step 2, for cases when the fault location is close to the root node, the voltages at the physically and topologically distant meters may become similar [42], increasing the difficulty of identifying the correct fault location in Step 3. In this case, the current phase information obtained at the root can be used to differentiate the exact feeder.

To summarize the method, the findings in [13] are applied so that

$$\text{Error} = \varepsilon_{\text{amplitude}(V)}^2 + \varepsilon_{\text{phase}(I)}^2 \quad (1)$$

and

$$\text{flag} = \frac{1}{\text{Error} + \Delta} \quad (2)$$

where  $\varepsilon_{\text{amplitude}(V)}$  is the difference between the amplitude of  $V_{\text{recorded}}$  and  $V_{\text{calculated}}$ ,  $\varepsilon_{\text{phase}(I)}$  is the difference between the phase of calculated and recorded current at the root node,  $\Delta$  is a small number to prevent the division by zero, and flag is used to identify the fault node. Since voltage phase information is not always available, the difference between the phase of  $V_{\text{recorded}}$  and  $V_{\text{calculated}}$  in (1) is not taken into consideration. While obtaining the flag during the fault location process, the fault resistance may also be obtained as proposed in [13]. The method uses the current phasor at the root node while assuming the fault at node  $M$ . More details may be found in [13].

In Step 3 after the comparison, node  $M$  with the largest flag value is declared the fault node. Then the algorithm will approach the maximum value of flag on all the lines connected to node  $M$ . Step 4 will be detailed in Section III-B.

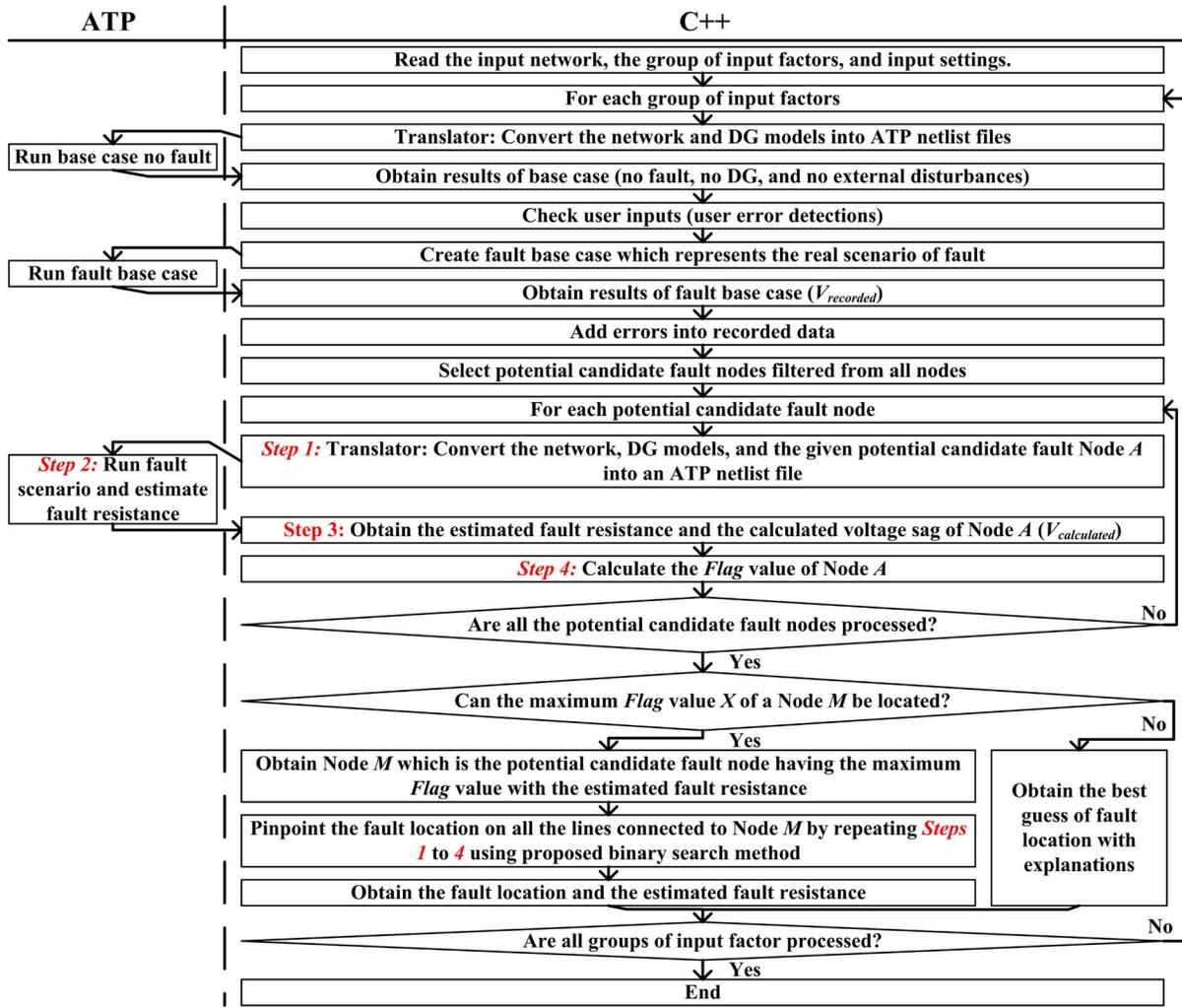


Fig. 3. Implementation details of the studied fault location algorithm.

Fig. 3 shows details of fault location algorithm implementation in C++ which interfaces with ATP so that all simulation processes are automated. The translator converts network components from the database into an ATP netlist, which is a high level descriptive language for inputs for ATP time-domain simulations [43]. After each simulation is done in ATP, the C++ code produces the results using GTPPLOT [25] and extracts information for performing the algorithm calculation. The set of candidate fault nodes in Fig. 3 are those which can exhibit the observed fault. For instance, if the fault type is single-phase-ground on phase C, then the nodes which do not have phase C cannot be the candidate fault nodes.

**B. Binary Search for Pinpointing Faults on the Line**

In this paper, a binary search method (halving) [41] is proposed to locate faults on the line (Step 4 in Section III-A). The assumption is that the true fault location produces the largest value of flag indicating the best match of voltage sags and all other locations have smaller values. Under this assumption, the fault location must be on one of the lines connected to the Node *M* (Fig. 4), which means that Node *M* is one end of the faulted line [13]. From preliminary tests, the flag value

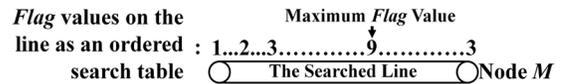


Fig. 4. Examples of flag values as ordered search tables on searched lines.

monotonically decreases around the peak at the fault location in a symmetrical fashion as the assumed fault moves more distant from the true fault location. In this case, it is proposed that the flag values of points on the line may be viewed as an ordered search table of length *N* (Fig. 4), and by halving the line the precise fault location can be found up to the specified tolerance in  $O(\log N)$  time.

At the start of the procedure, the search indices are the two values of flag from (2) at either end of the line. Then the line is halved and the flag values of new two ends are calculated for the next iteration. The total number of iterations *n* depends on how precise the results are required to be. Therefore, *n* will be the smallest nonnegative integer satisfying

$$\mu \geq \frac{l}{2^n} \tag{3}$$

where  $\mu$  is the user-given distance tolerance for pinpointing the fault on the line and  $l$  is the length of the searched line. The proposed binary search method is tested and proved to have better efficiency than the brute force search method [41].

#### IV. SOBOL'S SENSITIVITY ANALYSIS

Global variance-based sensitivity analysis, such as the method of Sobol, is a statistical tool which can ascertain how uncertainty in the inputs of an algorithm is reflected in the uncertainty of the output. Limitations in the generality of this approach lie in the necessity for: 1) input factors to be specified by garden variety distributions which are easily sampled from (such as normal and uniform); 2) input factors to be bijectively mappable to the inside of an  $n$ -dimensional  $L - 1$  sphere (or hypercube), in the sense that infinite values of inputs are not applicable; and 3) output is numerical and a measure of variance in the output can be made.

Sobol's approach to experiment design is efficient both in exploring problem space and in reducing the number of experiments which need to be performed [44], where an experiment is defined as performing one fault location scenario per iteration of the loop in Fig. 3. In the context of fault location algorithms for power systems, performing sensitivity analysis represents an examination of error made when locating faults under a set of typical operating conditions. Therefore, it may be useful for several reasons [45] to: 1) test robustness of an algorithm's accuracy in light of uncertainty in inputs; 2) understand the relationships between input and output of algorithms; and 3) identify how uncertainty in output can be reduced by suppression of uncertainty in algorithm input.

For simplicity of notation, let it be assumed that the sensitivity of a model  $f(x)$ , is the distance to fault, or relative error, and the task is to quantify the impact of each factor in this function. Let it also be assumed that the model has uncertainty in input factors,  $x \in [0, 1]^d$ , where each of the  $d$  input factors from  $\mathbb{R}$  is mapped into the interior of a  $d$ -dimensional hypercube, without loss of generality since the input factor space may be bijectively mapped to such a hypercube, and incurring the condition that the inputs  $x$  are independently and identically distributed [45].

The goal of variance-based sensitivity analysis is to distribute responsibility for uncertainty in the output  $f(x)$  among the model inputs  $x$ . The importance, or sensitivity index, of an input factor  $x_i$ , where  $x = [x_1, x_2, \dots, x_i, \dots, x_d]$ , is then directly related to the variance in output which can be attributed to it. In order to perform variance-based sensitivity analysis in more detail, it is necessary to perform the fault location procedure on a large number of input factor value combinations, sampled from the joint distribution of input factors, and then perform a decomposition of the exhibited variance. The decomposition of variance  $\sigma^2 f(x)$  is such that

$$\sigma^2 f(x) = \sum_{i=1}^d \sigma_i^2 + \sum_{i=1}^d \sum_{j < i} \sigma_{ij}^2 + \dots + \sigma_{1\dots d}^2. \quad (4)$$

Here if the main effect of a factor  $x_i$  is denoted on the output  $f(x)$  as  $\sigma_i^2 = \sigma^2 x_i (E[f(x)|x_i])$  then the compound effects of

TABLE IV  
INPUT FACTORS IN THIS STUDY

Input Factor	Range
Loading Error	uniform distribution between $\pm 1\%$
Maximum Substation Current Angle Error	uniform distribution between $\pm 1\%$
DG Penetration	between 10% and 60% in increments of 10% with same density
Maximum Voltage Magnitude Error at Meters	uniform distribution between $\pm 2\%$
System Power Factor	uniformly between 0.8 and 0.95
Fault Location	one of the three lines (near, middle, or far), chosen uniformly on the line and the percentage indicates the distance from the sending end of the line
Fault Resistance	uniformly between 1 and 50 $\Omega$

two factors  $x_i$  and  $x_j$  can be described as

$$\sigma_{ij}^2 = \sigma^2 x_i x_j (E[f(x)|x_j, x_i]) - \sigma_i^2 - \sigma_j^2. \quad (5)$$

The presented sensitivity analysis requires the coordination of software use as shown in Table I. The sensitivity analysis implemented in SimLab was used to design and evaluate the experiments and to quantify the factors impacting fault location accuracy. Based on the input factors from Table IV, sampling is first done in SimLab to create a set of specified input factors (Step 2 in Fig. 1), and then fed to fault location for fault scenario generation. The estimated fault location result is then compared with the actual location of fault to calculate the output error,  $f(x)$ . More elaborations on input factors and sampling will be demonstrated in Sections IV-A and IV-B, respectively. More details regarding how sensitivity analysis may be applied to fault location methods may be found in [14] and [46].

##### A. Input Factor

The input factors for the sensitivity analysis in this paper are shown in Table IV. The loading error represents an error between the estimated and real loading conditions. The ranges of input factor in Table IV are determined based on the results of preliminary tests. The design of measurement errors and loading errors is based on the results suggested in [23]. Note the errors chosen for different meters are different.

##### B. Sobol Index

The un-normalized Sobol indexes are reported. The first order of Sobol index represents one factor's singular contribution to the variance of output by itself. The total order of Sobol index represents the total contribution to the variance of output by one factor (first order effect plus all higher-order effects due to interactions). For instance, assume there are two input factors  $x_1$  and  $x_2$ . Suppose  $x_1$  has a low first order index, and the interaction effect index of both  $x_1$  and  $x_2$  is high. This indicates the output variance is highly dictated by both factors' interaction, even if  $x_1$  is less important. Therefore, the total effect index examines the impacts of input factors' interactions to the output variance. The combined effects of more than two input factors' interactions may be found in [47].

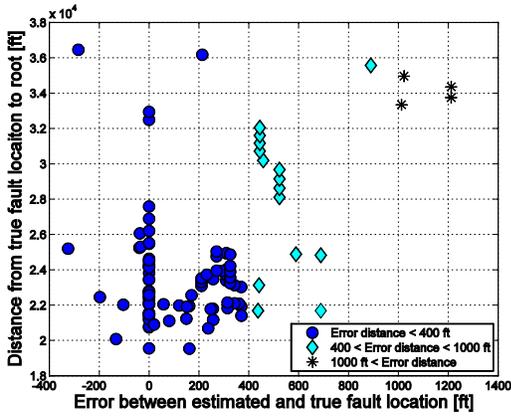


Fig. 5. Validation results.

### C. Sampling and Number of Experiment

The Sobol's sequences are within the category of quasi-random sequences which are used to produce samples of multiple parameters uniformly over the multidimensional parameter space [48]. The Sobol's sequences require a sample size (number of experiments)  $N_{\text{exp}}$  where

$$N_{\text{exp}} = 2^{N_{\text{input}}} \quad (6)$$

where  $N_{\text{input}}$  is the number of input factors. Since there are seven input factors, there should be 128 experiments. In order to obtain more accurate results, the number of experiments is doubled to be 256 in this paper.

## V. VALIDATION OF STUDIED FAULT LOCATION METHOD

Prior to the sensitivity analysis of DG impact, the studied voltage sag based fault location method was validated to ensure the correctness. For validation purposes, it was assumed that there were no measurement errors and loading estimation errors. The fault resistance was assumed to be 3  $\Omega$ . The single-phase-to-ground fault was simulated using nonerror measurement data on every receiving node of a line.

With no external disturbances, the applied voltage sag based fault location algorithm produces very accurate output as shown in Fig. 5, where positive error indicates downstream of the true fault location. In 83% of cases, the error distance is smaller than 400 ft. In 13% of cases, the error distance is between 400 and 1000 ft. In 4% of cases, the error distance is greater than 1000 ft, with the maximal error of 1212 ft. In cases where the output error is not zero, the program always detects nodes on the same lateral as the estimated location of fault. In most cases, the detected nodes are downstream of the true fault location (error in most cases are positive).

## VI. RESULTS AND RELATED ANALYSIS

In this section, two matters will be discussed: 1) the results of sensitivity analysis; and 2) detailed case studies of a fault on one lateral. The assumed 15 m locations are mentioned in Section II-D. The results of single-phase-ground fault type are presented in this section.

TABLE V  
RESULTS OF SENSITIVITY ANALYSIS:  
SOBOL FIRST ORDER INDEXES

Input Factor	Range	Sensitivity Index		
		Near	Middle	Far
Loading Error	$\pm 1\%$	0.606	0.128	0.142
Maximum Substation Current Angle Error	$\pm 1\%$	0.249	0.184	0.280
DG Penetration	10 – 60 %	0.417	0.160	0.260
Maximum Voltage Magnitude Error at Meters	$\pm 2\%$	0.417	0.160	0.260
System Power Factor	0.85 – 0.9	0.415	0.157	0.258
Percentage on the Line	0 – 100 %	0.443	0.147	0.252
Fault Resistance	1 – 50 $\Omega$	0.723	0.477	0.294

TABLE VI  
RESULTS OF SENSITIVITY ANALYSIS:  
SOBOL TOTAL ORDER INDEXES

Input Factor	Range	Sensitivity Index		
		Near	Middle	Far
Loading Error	$\pm 1\%$	0.852	0.605	0.846
Maximum Substation Current Angle Error	$\pm 1\%$	0.411	0.188	0.294
DG Penetration	10 – 60 %	0.513	0.160	0.269
Maximum Voltage Magnitude Error at Meters	$\pm 2\%$	0.513	0.161	0.269
System Power Factor	0.85 – 0.9	0.512	0.164	0.274
Percentage on the Line	0 – 100 %	0.690	0.163	0.265
Fault Resistance	1 – 50 $\Omega$	1.052	0.962	1.077

### A. Sensitivity Analysis of DG Impact

To illustrate the impact of DG, the fault locations are designated at three lines which are near, mid-way, and far away from the root node (representing distances from the source), as follows:

- 1) *Near*: The feeder segment between Nodes 1879 and 1880;
- 2) *Middle*: The feeder segment between Nodes 1911 and 1912;
- 3) *Far*: The feeder segment between Nodes 2762 and 2763.

Note that the near line is closest to the largest load node. In order to obtain global variance-based sensitivity indices, it is necessary to simulate a large number of fault scenarios. If the search space is to be efficiently explored, the conditions need to be carefully designed. In this paper, there are 256 cases (Section IV-C) conducted at each line, and the total number of cases is 768.

Tables V and VI summarize the results of sensitivity analysis, where they represent the first and total order indexes of input factors (Section IV-A), respectively. The concluding remarks can be drawn from Tables V and VI.

- 1) Fault resistance always has the largest sensitivity index among all factors and is the factor with the largest effect on the results. The reason for this is that the fault resistance is linked directly to the peak amplitude of fault current, which will cause the variation of voltage characteristic at nodes.
- 2) Loading error always has the second largest sensitivity index among all factors. This shows that the accuracy of fault location is very sensitive to the loading error estimation. It is also worth noting that results of

TABLE VII  
PERCENTAGE OF SIMULATED CASES IN DIFFERENT  
RANGES OF ERROR DISTANCE ON THREE LINES

Test Line	Percentage of Simulated Cases in Different Ranges of Error Distance in Feet		
	(0, 1000]	(1000, 2000]	Greater than 2000
Near	3.5 %	3.5 %	93.0 %
Middle	39.9 %	3.1 %	57.0 %
Far	0 %	37.9 %	62.1 %

TABLE VIII  
SCENARIOS FOR SPECIAL CASE STUDIES

Scenario Number	Total Number of Meter	Number of Meter on the Lateral	Type of DG
1	15	1	inverter-based
2	21	7	inverter-based
3	15	1	induction
4	21	7	induction

preliminary tests have shown that when the loading error is large (more than  $\pm 9\%$ ), the fault location algorithm could not locate the fault anywhere in the network.

- 3) The sensitivity index of DG penetration level reveals that while this effect is not the most important, it should be taken into account when ascertaining accuracy of fault location.
- 4) All the factors have higher sensitivity index when the fault is near the root node. This demonstrates that the match of voltage sag could become difficult when the fault is near the root node.
- 5) All the factors, except fault resistance, have the smallest sensitivity index on the middle line as shown in Table V. The reason for this is the presence of DG, which causes the system to become a multisource system. Although it is not shown in Fig. 2, by comparing the three lines where the faults are designated, it is shown the line in the middle does not have relatively large size DG compared to the line at the far side (see Section VI-B). This implies that the sensitivity indexes may be larger with significant DG closer to the fault location, where those DG protections are not tripped.

Table VII shows the percentage of cases in different ranges of distance error. The largest percentage always falls into the category of cases whose distance error is greater than 2000 feet regardless of the distance between the fault and the source, when there are external disturbances present.

### B. Special Case Studies

A more detailed example will be used to demonstrate the impact of DG on fault location. It is assumed that there are no external disturbances. The whole network is simulated but the specific focus is the lateral with the end Node 2769 (Fig. 2). The fault is set between Nodes 2766 and 2767 at one end of feeder, which is 36316.2 feet from the source.

Four scenarios in Table VIII are designed to reveal special cases of two different DG types and two different groups of meter numbers under various DG penetration levels. The induction type of DG is modeled as a power flow control model as illustrated in [30] with power factor of 0.9. The

TABLE IX  
RESULTS OF SCENARIOS

DG Penetration Level (%)	Distance between Estimated Fault and Real Fault Location (ft)			
	Scenario 1	Scenario 2	Scenario 3	Scenario 4
0 (no DG)	747.9	133.1	747.9	133.1
1	1656.7	1656.7	1656.7	1656.7
5	1656.7	1656.7	1656.7	1656.7
20	1656.7	1656.7	1656.7	1656.7
60	1656.7	1656.7	1656.7	1656.7

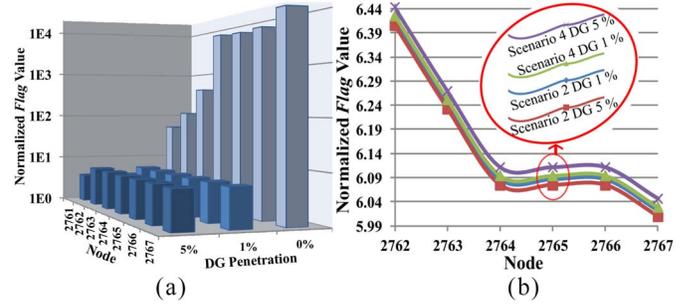


Fig. 6. (a) Flag values vary on the lateral under different DG penetration levels in scenario 2. (b) Flag values at the meters on the lateral under 1% and 5% DG penetration levels in scenarios 2 and 4.

15 m have been mentioned in Section II-B where the only meter on the lateral is at Node 2767. The 21 m include the aforementioned 15 m with 6 more meters placed along the lateral at Nodes 2751, 2755, 2757, 2759, 2761, 2763, and 2765 making 7 m in total on the lateral.

Table IX shows the distance between the real and estimated fault locations in all scenarios. Fig. 6(a) shows how the flag values vary on the lateral in scenario 2. Fig. 6(b) shows the flag values at those meters on the lateral under 1% and 5% DG penetrations in scenarios 2 and 4. Note that the flag values for other DG penetration levels were not presented in Fig. 6 since the trend of flag values remains the same. Also note that the detected fault locations in all scenarios are on the same lateral as the real fault locations.

Table IX demonstrates the summary of special case studies, and the following remarks can be observed.

- 1) While there is no DG in the system, the number of meters affects the results. More meters results in better accuracy.
- 2) In the presence of DG, the error remains the same. The estimated fault location is moving toward Node 2762 from Node 2767 regardless of DG penetration levels, DG types, or number of meters.

Fig. 6(a) and (b) further illustrates the observations from Table IX. In Fig. 6(a), it is clear that the maximum flag value under 0% DG penetration (no DG) is at Node 2767, but it is not clear where the maximum flag value is under other DG penetration levels. In Fig. 6(b), though the flag values at the nodes vary with DG penetration levels and DG types, the flag value always monotonically decreases around the peak at the fault location (Section III-B) in the same fashion and therefore the algorithm always detects the same fault location. Those DGs whose protections are not open contribute to fault current,

and therefore influence the voltage sag at the nodes close to the fault (which creates very similar flag values). These lead the algorithm toward wrong estimates in the presence of DG.

## VII. CONCLUSION

The described approach is computationally efficient under realistic DG deployment scenarios. It represents well a complex distribution grid, and it retains the ability to identify individual and compound factors severely affecting reliability of fault location calculation. The results clearly show that statistical analysis is necessary to reveal the true profile of error in fault location.

Several contributions are made in this paper.

- 1) A systematic study framework to analyze the impact of DG on the voltage sag based fault location is proposed. The specific issues relating to data flow and software components at the scale necessary to statistically measure the differences is reflected upon. A binary search approach to pinpoint the fault on the feeder within the user-specified accuracy tolerance is outlined.
- 2) Results of sensitivity analysis demonstrate the impact of measurement imperfection on the algorithm and it was concluded that the presence of DG, which makes the system become multisource, is consequential. Fault resistance is the largest sensitivity index among all factors, and it has the largest effect on the results. Loading error is the second largest sensitivity index among all factors.
- 3) The cases where the voltage sag based fault location may fail are also taken into consideration. In the special case studied, DG penetration levels, DG types, and number of meters show little (if any) impact on the output errors. Those DGs whose protections are not open contribute to fault current resulting in similar flag values around the fault nodes, and lead to wrong estimates in the presence of DG.

In addition, without the presence of DG, the input factor “distance between fault location and the root node” is defined to obtain the errors as a function of distance from the single source. Yet, in the presence of DG, the system becomes multisource, so this input factor no longer provides effective information.

## ACKNOWLEDGMENT

The authors would like to thank Dr. J. Stoupis, Dr. M. Mousavi, Dr. N. Kang, and Dr. X. Feng from the ABB Group for their valuable advice. They also would like to thank the prior work of Dr. S. Lotfifard, who worked on the initial study of the voltage sag based fault location for his Ph.D. dissertation at Texas A&M University. This paper was prepared as an account of the work sponsored by an agency of the U.S. Government. Neither the U.S. Government nor any agency thereof, nor any of their employees, makes any warranty, express, or implied, or assume any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned

rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the U.S. Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the U.S. Government or any agency thereof.

## REFERENCES

- [1] M. Kezunovic *et al.*, “Fault location,” in *Wiley Encyclopedia of Electrical and Electronics Terminology*, vol. 7. Hoboken, NJ, USA: Wiley, 1999, pp. 276–285.
- [2] M. Kezunovic, “Smart fault location for smart grids,” *IEEE Trans. Smart Grid*, vol. 2, no. 1, pp. 11–22, Mar. 2011.
- [3] R. K. Aggarwal, Y. Aslan, and A. T. Johns, “New concept in fault location for overhead distribution systems using superimposed components,” *IEE Proc. Gener. Transmiss. Distrib.*, vol. 144, no. 3, pp. 309–316, May 1997.
- [4] R. K. Aggarwal, Y. Aslan, and A. T. Johns, “An interactive approach to fault location on overhead distribution lines with load taps,” in *Proc. 6th Int. Conf. Develop. Power Syst. Protect.*, Nottingham, U.K., Mar. 1997, pp. 184–187.
- [5] J. J. Mora, G. Carrillo, and L. Perez, “Fault location in power distribution systems using ANFIS nets and current patterns,” in *Proc. IEEE/PES Transmiss. Distrib. Conf. Expo. Lat. Amer.*, Caracas, Venezuela, Aug. 2006, pp. 1–7.
- [6] J. Mora-Florez, V. Barrera-Nuez, and G. Carrillo-Caicedo, “Fault location in power distribution systems using a learning algorithm for multivariable data analysis,” *IEEE Trans. Power Del.*, vol. 22, no. 3, pp. 1715–1721, Jul. 2007.
- [7] J. Mora-Florez, J. Meléndez, and G. Carrillo-Caicedo, “Comparison of impedance based fault location methods for power distribution systems,” *Elect. Power Syst. Res.*, vol. 78, pp. 657–666, Apr. 2008.
- [8] A. A. Girgis, C. M. Fallon, and D. L. Lubkeman, “A fault location technique for rural distribution feeders,” *IEEE Trans. Ind. Appl.*, vol. 29, no. 6, pp. 1170–1175, Nov./Dec. 1993.
- [9] M. S. Choi *et al.*, “A direct three-phase circuit analysis-based fault location for line-to-line fault,” *IEEE Trans. Power Del.*, vol. 22, no. 4, pp. 2541–2547, Oct. 2007.
- [10] A. Borghetti *et al.*, “Continuous-wavelet transform for fault location in distribution power networks: Definition of mother wavelets inferred from fault originated transients,” *IEEE Trans. Power Syst.*, vol. 23, no. 2, pp. 380–388, May 2008.
- [11] Z. Q. Bo *et al.*, “Accurate fault location technique for distribution system using fault-generated high-frequency transient voltage signals,” *IEEE Proc. Gener. Transmiss. Distrib.*, vol. 146, no. 1, pp. 73–79, Jan. 1999.
- [12] F. H. Magnago and A. Abur, “Fault location using wavelets,” *IEEE Trans. Power Del.*, vol. 13, no. 4, pp. 1475–1480, Oct. 1998.
- [13] S. Lotfifard, M. Kezunovic, and M. J. Mousavi, “Voltage sag data utilization for distribution fault location,” *IEEE Trans. Power Del.*, vol. 26, no. 2, pp. 1239–1246, Apr. 2011.
- [14] P.-C. Chen, V. Malbasa, and M. Kezunovic, “Sensitivity analysis of voltage sag based fault location algorithm,” in *Proc. 18th Power Syst. Comput. Conf. (PSCC)*, Wroclaw, Poland, 2014.
- [15] S. Lotfifard, M. Kezunovic, and M. J. Mousavi, “A systematic approach for ranking distribution systems fault location algorithms and eliminating false estimates,” *IEEE Trans. Power Del.*, vol. 28, no. 1, pp. 285–293, Jan. 2013.
- [16] R. H. Diaz and T. M. Lopez, “Fault location techniques for electrical distribution networks: A literature survey,” in *Proc. Eur. Power Energy Syst.*, Benalmádena, Spain, Jun. 2005, pp. 311–318.
- [17] *IEEE Guide for Determining Fault Location on AC Transmission and Distribution Lines*, IEEE Standard C37.114-2004, Jun. 2005.
- [18] S. M. Brahma, “Fault location in power distribution system with penetration of distributed generation,” *IEEE Trans. Power Del.*, vol. 26, no. 3, pp. 1545–1553, Jul. 2011.
- [19] S. M. Brahma and A. A. Girgis, “Development of adaptive protection scheme for distribution systems with high penetration of distributed generation,” *IEEE Trans. Power Del.*, vol. 19, no. 1, pp. 56–63, Jan. 2004.

- [20] D. Johnsonbaugh, *Fault Location for Distribution Systems With Distributed Generation Using Modified Three Phase Fault Location Methods*. Clemson, SC, USA: Clemson Univ., 2004.
- [21] J. Miret *et al.*, "Control scheme for photovoltaic three-phase inverters to minimize peak currents during unbalanced grid-voltage sags," *IEEE Trans. Power Electron.*, vol. 27, no. 10, pp. 4262–4271, Oct. 2012.
- [22] M. H. J. Bollen, *Understanding Power Quality Problems: Voltage Sags and Interruptions*. Piscataway, NJ, USA: IEEE Press, 2000.
- [23] P.-C. Chen, M. Yimai Dong, V. Malbasa, and M. Kezunovic, "Uncertainty of measurement error in intelligent electronic devices," in *Proc. IEEE/PES Gen. Meeting*, National Harbor, MD, USA, 2014, pp. 1–5.
- [24] *SimLab 2.2 Reference Manual*, Inst. Protect. Security Citizen (IPSC), Joint Res. Centre (JRC), Eur. Commission, Ispra, Italy, 2004.
- [25] Alternative Transients Program. (2010). *ATP-EMTP*. [Online]. Available: <http://www.emtp.org/>
- [26] *IEEE Standard for Interconnecting Distributed Resources With Electric Power System*, IEEE Standard 1547.2-2008, Apr. 2009.
- [27] *Handbook of General Requirements for Electrical Service to Dispersed Generation Customers, Specification EO-2115 Revision 8*, Consolidated Edison Company New York, Inc., New York, NY, USA, Mar. 2006.
- [28] J. A. Martinez and J. Martin-Arnedo, "Advanced load models for voltage sag studies in distribution networks," in *Proc. IEEE/PES Gen. Meeting*, vol. 1. Denver, CO, USA, Jun. 2004, pp. 614–619.
- [29] A. Herrera-Orozco, S. Perez-Londobo, and J. Mora-Florez, "Load modeling for fault location in distribution systems with distributed generation," in *Proc. 6th IEEE/PES Transmiss. Distrib. Latin Amer. Conf. Expo. (T&D-LA)*, Montevideo, Uruguay, Sep. 2012, pp. 1–8.
- [30] S. Chakraborty, B. Kroposki, and W. Kramer, "Advanced power electronic interfaces for distributed energy systems. Part 2: Modeling, development, and experimental evaluation of advanced control functions for single-phase utility-connected inverter," Nat. Renew. Energy Lab., Washington, DC, USA, Tech. Rep. NREL/TP-550-44313, Nov. 2008.
- [31] S. Boljevic and M. F. Martin-Arnedo, "The contribution to distribution network short-circuit current level from the connection of distributed generation," in *Proc. IEEE/PES Gen. Meeting*, vol. 1. Denver, CO, USA, Jun. 2004, pp. 614–619.
- [32] E. V. Mgaya and Z. Muller, "The impact of connecting distributed generation to the distribution system," *Acta Polytech.*, vol. 47, nos. 4–5, pp. 96–101, 2007.
- [33] P.-C. Chen *et al.*, "Analysis of voltage profile problems due to the penetration of distributed generation in low-voltage secondary distribution networks," *IEEE Trans. Power Del.*, vol. 27, no. 4, pp. 2020–2028, Oct. 2012.
- [34] W. El-Khattam, Y. G. Hegazy, and M. M. A. Salama, "Investigating distributed generation systems performance using Monte Carlo simulation," *IEEE Trans. Power Syst.*, vol. 21, no. 2, pp. 524–532, May 2006.
- [35] P.-C. Chen and M. Kezunovic, "Analysis of the impact of distributed generation placement on voltage profile in distribution systems," in *Proc. IEEE/PES Gen. Meeting*, Vancouver, BC, Canada, Jul. 2013, pp. 1–5.
- [36] P.-C. Chen, T. Dokic, and M. Kezunovic, "The use of big data for outage management in distribution systems," in *Proc. Int. Conf. Elect. Distrib. (CIRED) Workshop*, Rome, Italy, 2014, Art. ID 0406.
- [37] P. Dehghanian, M. Fotuhi-Firuzabad, S. Bagheri-Shouraki, and A. A. R. Kazemi, "Critical component identification in reliability centered asset management of distribution power systems via fuzzy AHP," *IEEE Syst. J.*, vol. 6, no. 4, pp. 593–602, Dec. 2012.
- [38] H. Tram, "Technical and operation considerations in using smart metering for outage management," in *Proc. IEEE/PES Trans. Distrib. Conf. Expo. (T&D)*, Chicago, IL, USA, 2008, pp. 1–3.
- [39] Y. Dong, C. Zheng, and M. Kezunovic, "Enhancing accuracy while reducing computation complexity for voltage sag based distribution fault location," *IEEE Trans. Power Del.*, vol. 28, no. 2, pp. 1202–1212, Apr. 2013.
- [40] C. Zheng, V. Malbasa, and M. Kezunovic, "Regression tree for stability margin prediction using synchrophasor measurements," *IEEE Trans. Power Syst.*, vol. 28, no. 2, pp. 1978–1987, May 2013.
- [41] M. T. Goodrich *et al.*, *Data Structures and Algorithms in C++*, 2nd ed. Hoboken, NJ, USA: Wiley, 2011.
- [42] P.-C. Chen, V. Malbasa, and M. Kezunovic, "Locating sub-cycle faults in distribution network applying half-cycle DFT method," in *Proc. IEEE/PES Transmiss. Distrib. Conf. Expo.*, Chicago, IL, USA, 2014, pp. 1–5.
- [43] F. de León *et al.*, "Development of data translators for interfacing power-flow programs with EMTP-type programs: Challenges and lessons learned," *IEEE Trans. Power Del.*, vol. 28, no. 2, pp. 1192–1201, Apr. 2013.
- [44] I. M. Sobol', "On the distribution of points in a cube and the approximate evaluation of integrals," *Zh. Vychisl. Mat. Mat. Fiz.*, vol. 7, no. 4, pp. 784–802, 1967.
- [45] R. Zivanovic and H. Ooi, "A systematic approach for testing transmission line fault-locating techniques," in *Proc. Elect. Conf. Elect. Energy Soc. Australia (EESA)*, Melbourne VIC, Australia, Aug. 2007, pp. 1–6.
- [46] P.-C. Chen *et al.*, "Sensitivity of voltage sag based fault location in distribution network to sub-cycle faults," in *Proc. 46th North Amer. Power Symp. (NAPS)*, Pullman, WA, USA, 2014, pp. 1–6.
- [47] A. Saltelli *et al.*, *Global Sensitivity Analysis: The Primer*. Hoboken, NJ, USA: Wiley, 2008.
- [48] S. Burhenne, D. Jacob, and G. P. Henz, "Sampling based on Sobol' sequences for Monte Carlo techniques applied to building simulations," in *Proc. Int. Conf. Build. Simulat.*, Sydney, NSW, Australia, Nov. 2011, pp. 1816–1823.



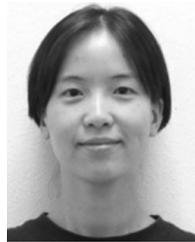
**Po-Chen Chen** (S'12) received the B.Sc. and M.Sc. degrees in electrical engineering from the Polytechnic Institute of New York University, Brooklyn, NY, USA, in 2010 and 2012, respectively. He is currently pursuing the Ph.D. degree in electrical engineering from Texas A&M University, College Station, TX, USA.

His current research interests include distributed generation, power system analysis, power system protection and control, voltage quality and stability studies, and big data application for distribution system.



**Vuk Malbasa** (M'12) received the B.S. degree in informatics from the Faculty of Sciences, University of Novi Sad, Novi Sad, Serbia, and the Ph.D. degree in computer science from Temple University, Philadelphia, PA, USA, in 2006 and 2012, respectively.

He is currently an Assistant Professor with the Faculty of Technical Sciences, University of Novi Sad. His current research interests include regularized spatial models, intelligent sensor applications, graphical models, and cognitive communications in networking.



**Yimai Dong** (S'07–M'13) received the B.S. and M.S. degrees from North China Electric Power University, Beijing, China, and the Ph.D. degree from Texas A&M University, College Station, TX, USA, in 2005, 2007, and 2013, respectively, all in electrical engineering.

She is currently a Development and Applications Engineer with Electicon International Inc., Ann Arbor, MI, USA. Her current research interests include distribution fault location, distribution outage management, detailed modeling of relays and protective devices, and simulation and analysis of protection system.

Dr. Dong is a Member of the Council on Large Electric Systems (CIGRE).



**Mladen Kezunovic** (S'77–M'80–SM'85–F'99) received the Dipl. Ing., degree from the University of Sarajevo, Sarajevo, Bosnia, and the M.S., and Ph.D. degrees from the University of Kansas, Lawrence, KS, USA, in 1974, 1977, and 1980, respectively, all in electrical engineering.

He is currently the Eugene E. Webb Professor, the Director of the Smart Grid Center, and the Site Director of the National Science Foundation Industry and University Cooperative Research Program Power Systems Engineering Research Center, Tempe, AZ, USA. His current research interests include protective relaying, automated power system disturbance analysis, computational intelligence and data analytics, and smart grids. He has published over 450 papers, given over 100 seminars, invited lectures and short courses, and consulted for over 50 companies worldwide. He is the Principal of XpertPower Associates, College Station, TX, USA.

Prof. Kezunovic is a CIGRE Fellow and a Registered Professional Engineer in Texas.