A Survey on State Estimation Techniques and Challenges in Smart Distribution Systems

Kaveh Dehghanpour, *Member, IEEE*, Zhaoyu Wang, *Member, IEEE*, Jianhui Wang, *Senior Member, IEEE*, Yuxuan Yuan, *Student Member, IEEE*, and Fankun Bu, *Student Member, IEEE*

***Abstract*—This paper presents a review of the literature on state estimation (SE) in power systems. While covering some works related to SE in transmission systems, the main focus of this paper is distribution system SE (DSSE). This paper discusses a few critical topics of DSSE, including mathematical problem formulation, application of pseudo-measurements, metering instrument placement, network topology issues, impacts of renewable penetration, and cyber-security. Both conventional and modern data-driven and probabilistic techniques have been reviewed. This paper can provide researchers and utility engineers with insights into the technical achievements, barriers, and future research directions of DSSE.**

***Index Terms*—Distribution system state estimation, pseudo-measurements, topology, cyber-security.**

I. Introduction

DISTRIBUTION System State Estimation (DSSE) is the process of inferring the values of system's state variables using a limited number of measured data at certain locations in the system [1]. Thus, DSSE is basically a numerical process to map data measurements to state variables. While State Estimation (SE) is a well-developed and widely-used concept in transmission systems, its use at the distribution level is still the subject of active research. In recent years we have observed the rapid growth of Advanced Metering Infrastructure (AMI) in electric distribution systems (e.g., according to [2], the number of advanced meters in the U.S. was estimated to be 64.7 million devices in 2015, out of a total number of 150.8 million meters, indicating a penetration rate of 42.9%.) Hence, DSSE is expected to become a significant function in monitoring and power management of smart grids [3]. A general schematic of DSSE function is shown in Fig. 1. Extending conventional SE approaches to active distribution systems is a challenging task due to several factors that are based on the considerable differences between the transmission and distribution systems:

1) *Observability problem:* Unlike transmission systems, the distribution systems are highly unobservable, meaning that the number of metering instruments in a network is generally small compared to the huge size of the system [4].
2) *Low x/r value:* In distribution systems, we generally face low x/r levels, which render the conventional DC SE techniques in transmission systems unusable at the distribution level [5].
3) *Unbalanced operation:* Distribution systems are in practice highly unbalanced which leads to a higher level of complexity in SE problem formulation.
4) *Communication issues:* Constraints on the communication system, such as the network bandwidth and capacity also limit the accuracy and rate of data exchange [6].
5) *Network configuration problem:* Considering the huge size of the distribution network and noting that the complete data related to the topology of this network is not commonly stored an additional degree of complexity to DSSE in these networks [7].
6) *Renewable energy integration:* The higher penetration of renewable power resources introduces a higher level of uncertainty in distribution system operation and DSSE.
7) *Cyber-security issues:* The issue of cyber-security is a new concern in management and control of active distribution systems.

Despite these challenges, industrial interest in implementing DSSE is growing. Electrical energy firms such as Eaton [8], Survalent [9], ETAP [10], OSI [11], and Nexant [12] have recently devised industrial programs for promoting system monitoring and management at the distribution level for utilities using DSSE. A discussion on relevant experiences on DSSE for radial distribution networks is presented in [13], where the connections between SE implementation and practical variables, such as line lengths, switch flows, voltage regulation, and measurement areas, are elaborated. In this paper we seek to present an extensive review of the proposed solutions to different DSSE-related problems. While the main focus of this paper is DSSE, certain works on transmission system SE have also been cited and reviewed where they become relevant. In summary, this paper discusses the following issues: DSSE problem formulation, pseudo-measurement generation, uncertain network topology, integration of renewable resources,

Manuscript received May 7, 2018; revised August 16, 2018; accepted September 6, 2018. This work was supported by the Advanced Grid Modeling Program at the U.S. Department of Energy Office of Electricity under
AQ1 Grant DE-OE0000875. Paper no. TSG-00701-2018. *(Corresponding author: Zhaoyu Wang.)*

K. Dehghanpour, Z. Wang, Y. Yuan, and F. Bu are with the Department of Electrical and Computer Engineering, Iowa State University, Ames, IA 50011 USA (e-mail: kavehd@iastate.edu; wzy@iastate.edu).

J. Wang is with the Department of Electrical Engineering, Southern Methodist University, Dallas, TX 75275 USA, and also with the Energy Systems Division, Argonne National Laboratory, Argonne, IL, USA (e-mail:
AQ2 jianhui.wang@ieee.org).

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.2018.2870600

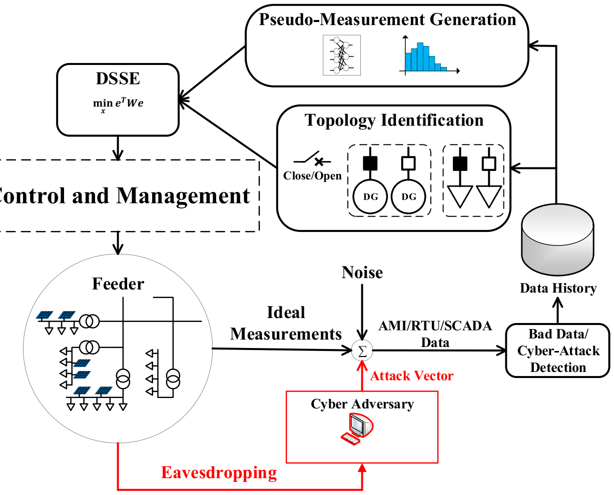


Fig. 1. DSSE function in smart grid environment.

meter placement, and DSSE cyber-security. Special attention has been given to data-driven and machine-learning-based approaches that are gaining interest to address different types of problems [14].

The reviewed works address critical aspects of DSSE shown in Fig. 1: 1) *DSSE solver module:* in Sections II and III, we summarize the fundamentals of DSSE, with respect to choice of algorithm and state variables. 2) *Pseudo-measurement generation module:* in Section IV the challenge of observability in distribution systems and proposed pseudo-measurement generation solutions in the literature are elaborated. 3) *Topology identification module:* Section V reviews the past works related to online configuration tracking, connectivity detection, and topology discovery, which are pre-requisites for obtaining accurate DSSE solutions. 4) *Feeder and instrumentation module:* The measurement units distributed across the electric power system are the main sources of the information for running the monitoring and control systems. In Section VI, the problem of optimal meter placement and potential PMU applications in distribution feeders is presented in terms of practical constraints and objective functions. Modern distribution feeders can have high penetration levels of distributed renewable resources. The impacts of penetration of renewable energy resources in distribution feeders on DSSE are analyzed in Section VII. 5) *Cyber-security module:* Reliable DSSE depends on detection and prevention of cyber-intrusions and cyber-attacks. The challenge of cyber-security when performing wide-scale distribution system measurement and monitoring is discussed in Section VIII. Furthermore, conclusions and future research directions are provided in Sections IX and IX-A.

II. Fundamentals of SE

A) Conventional Approach: Given a measurement vector $\boldsymbol{z}$ (with size $m \times 1$), and a measurement function $\boldsymbol{h}$, which connects the true state vector $\boldsymbol{x}$ (with size $n \times 1$) to the measurement vector (i.e., $\boldsymbol{z} = \boldsymbol{h}(\boldsymbol{x}) + \boldsymbol{e}$, with $\boldsymbol{e}$ denoting the measurement error vector), the state estimation problem can be formulated as a Weighted Least Square (WLS) optimization problem (with bold letters denoting vectors/matrices) [1]:

$$\hat{\boldsymbol{x}} = \arg\min_{\boldsymbol{x}} (\boldsymbol{z} - \boldsymbol{h}(\boldsymbol{x}))^T \boldsymbol{W} (\boldsymbol{z} - \boldsymbol{h}(\boldsymbol{x})) \tag{1}$$

where $\hat{\boldsymbol{x}}$ is the estimated state vector, T is the matrix transposition operation, and $\boldsymbol{W}$ denotes the weight matrix that represents the user's confidence in the measured data. A widely-used choice for the weight matrix is $\boldsymbol{W} = diag\{\sigma_1^{-2}, \ldots, \sigma_m^{-2}\}$, where σ_j^2 represents the variance of the measurement error corresponding to the j^{th} element of $\boldsymbol{z}$. This choice of the weight matrix is based on two assumptions: 1) the error vector ($\boldsymbol{e}$) has a Gaussian distribution with zero mean, and 2) the measurement errors of different elements of the measurement vector are statistically independent. Under these assumptions the WLS problem transforms to the maximum likelihood estimation. A number of papers have deviated from the conventional approach towards selecting $\boldsymbol{W}$. For instance, in [15], using active/reactive power data history, non-diagonal terms have been added to the weight matrix to obtain better WLS accuracy, by modeling the existing correlation between the different measurement samples. This problem has been analyzed in details in [16] for modeling the correlations in measurement error distributions of different variables that are measured by the same device (smart meters and PMUs.) For instance, it is shown that the non-diagonal covariance terms between different variables measured by the same device are as follows (notation: active power (P), reactive power (Q), voltage magnitude (V), current magnitude (I), power factor ($\cos\Phi$)):

$$\sigma_{V,P} = \sigma_V^2 I \cos\Phi, \quad \sigma_{V,Q} = \sigma_V^2 I \sin\Phi$$

$$\sigma_{P,Q} = \frac{1}{1}\left(\sigma_V^2 I^2 \sin 2\Phi - \sigma_\Phi^2 I^2 V^2 \sin 2\Phi + \sigma_I^2 V^2 \sin 2\Phi\right) \tag{2}$$

Through another approach, in [17] and [18], the elements of the diagonal $\boldsymbol{W}$ matrix are updated using a weight function during solution iterations to obtain robustness against bad data. The proposed weight updating mechanism for the i^{th} measurement to obtain new weight value ($\bar{w_i}$) is as follows:

$$\bar{w_i} = \begin{cases} \sigma_i^{-2}, & D_i' \le k_0 \\ \sigma_i^{-2}\zeta_i, & k_0 < D_i' \le k_0 \\ 0, & D_i' > k_1 \end{cases} \tag{3}$$

where, D_i', ζ_i k_0, and k_1 are parameters defined based on the residual level corresponding to the i^th data sample. The idea behind (3) is that as D_i' (which is a measure of low quality of the measured data sample) increases beyond the introduced thresholds (k_0 and k_1), the weight value assigned to it should decrease (with factor ζ_i), reducing the influence of unreliable or bad data samples on the outcome of the WLS.

Conventionally, Gauss-Newton method has been applied to iteratively solve the WLS problem (1) [5]. This algorithm basically finds a solution to the equation $\nabla J = 0$, where J denotes the objective function of optimization problem (1). The update rules of the algorithm at the k^{th} iteration are as follows:

$$\boldsymbol{H}(\boldsymbol{x}(k)) = \frac{\partial J}{\partial \boldsymbol{x}(k)} \tag{4}$$

$$\boldsymbol{G}(k) = \boldsymbol{H}(\boldsymbol{x}(k))^T \boldsymbol{W} \boldsymbol{H}(\boldsymbol{x}(k)) \tag{5}$$

$$\Delta\boldsymbol{x}(k) = \boldsymbol{G}(k)^{-1} \boldsymbol{H}(\boldsymbol{x}(k))^T \boldsymbol{W}(\boldsymbol{z} - \boldsymbol{h}(\boldsymbol{x}(k))) \tag{6}$$

$$\boldsymbol{x}(k+1) = \boldsymbol{x}(k) + \Delta\boldsymbol{x}(k) \tag{7}$$

where, $\boldsymbol{H}$ is the Jacobian of J with respect to the state variables, and $\boldsymbol{G}$ is the system gain matrix. Other algorithms, such as back tracking method, trust region method, and quasi-Newton techniques, have also been applied instead of the classical Gauss-Newton method, to obtain better convergence properties [19]. Noting the non-convexity of (1) and the sensitivity of Newton method to initial conditions and gain matrix ill-conditioning, in [20] and [21], a Semi-Definite Programming (SDP) approach is proposed to find a good initial guess for the Newton method. The SDP formulation is based on the convex relaxation of the original WLS problem, which also guarantees the existence of a unique global solution. The computational efficiency of SDP is shown to be superior compared to that of the original non-convex problem. To further improve the computational performance of SDP-based SE, distributed algorithms have been employed for obtaining a solution [22].

Another modification in the structure of WLS (1) is the inclusion of *virtual measurements* as equality constraints ($\boldsymbol{c}(\boldsymbol{x}) = \boldsymbol{0}$). Virtual measurements represent operator's perfect information on certain aspects of system operation (e.g., zero-power-injection at nodes without customers.) Lagrange multipliers ($\boldsymbol{\lambda}$) have been proposed as penalty factors for enforcing these equality constraints [23]. The modified WLS objective function is defined as follows:

$$\left\{\hat{\boldsymbol{x}}, \hat{\boldsymbol{\lambda}}\right\} = \arg\min_{\boldsymbol{x},\boldsymbol{\lambda}} (\boldsymbol{z} - \boldsymbol{h}(\boldsymbol{x}))^T \boldsymbol{W} (\boldsymbol{z} - \boldsymbol{h}(\boldsymbol{x})) + \boldsymbol{\lambda}^T \boldsymbol{c}(\boldsymbol{x}) \tag{8}$$

Given the above objective function, the state update step in the Gauss-Newton method (6) is changed to:

$$\begin{bmatrix} \Delta\boldsymbol{x}(k) \\ \boldsymbol{\lambda}(k) \end{bmatrix} = \begin{bmatrix} \boldsymbol{H}^T\boldsymbol{W}\boldsymbol{H} & \boldsymbol{C}(\boldsymbol{x}(k))^T \\ \boldsymbol{C}(\boldsymbol{x}(k)) & 0 \end{bmatrix}^{-1} \begin{bmatrix} \boldsymbol{H}^T\boldsymbol{W}(\boldsymbol{z} - \boldsymbol{h}(\boldsymbol{x}(k))) \\ -\boldsymbol{c}(\boldsymbol{x}(k)) \end{bmatrix} \tag{9}$$

where, $\boldsymbol{C}(\boldsymbol{x}) = \frac{\partial \boldsymbol{c}(\boldsymbol{x})}{\partial \boldsymbol{x}}$.

B) Alternative DSSE Structures: While WLS represents the conventional SE in power systems, alternative mathematical formulations have been proposed for the purpose of increasing the robustness of the state estimator when facing bad data. Noting the susceptibility of WLS to bad data, in [24], the use of Least Median of Squares (LMS) and Least Trimmed Squares (LTS) is studied, which shows improved behavior in handling outliers. Also, [25] investigates the use of Least Absolute Value (LAV) estimator, which has the property of automatic bad data rejection. Increasing the robustness of SE has also been promoted by using a Generalized Maximum-likelihood (GM) estimator instead of WLS in [26], where normalized residuals (r_n) are used through a convex score functions (denoted as $\rho(.)$) in formulating the objective function. The SE formulation for these different approaches (including pros and cons) are shown in Table I, in terms of the objective function in optimization problem (1). In this table, the residuals $\boldsymbol{r} = [r_1, \ldots, r_m]^T$ are defined as $r_i = z_i - h_i(\boldsymbol{x})$. Also, $med\{\}$ and $r_{(i)}$ define the set median and the i^{th} order statistics, respectively. Numerical comparisons of these alternative DSSE formulations in terms of robustness against system parameter uncertainties are presented in [27].

TABLE I
AVAILABLE ROBUST SE FORMULATIONS

Method	Objective Function	Pros and Cons
WLS	$\boldsymbol{r}^T\boldsymbol{W}\boldsymbol{r}$	(+) Fast, simple, widely-used, (-) Sensitive to bad data
LMS	$med\{r_1^2, ..., r_m^2\}$	(+) Robust against bad data and leverage points, (-) High computational cost, high measurement redundancy requirements
LTS	$\sum_{i=1}^{h} r_{(i)}^2$	(+) Robust against bad data, (-) High computational cost and memory requirement
LAV	$\sum_{i=1}^{m} \lvert r_i \rvert$	(+) Robust against bad data, small sensitivity to line impedance uncertainty, (-) High computational cost, sensitivity to leverage points and measurement uncertainty
GM	$\sum_{i=1}^{m} \sigma_i^{-2} \rho(r_{n_i})$	(+) Robust against bad data, (-) Parameter selection sensitivity

Other approaches towards structuring the DSSE have been presented as well. For instance, some works in the literature tend to propose estimators which relax the Gaussian uncertainty assumption inherent to WLS. This is of practical importance given that this assumption is shown, through field tests, to be largely inaccurate [28]. Using Mean Squared Estimator (MSE) an analytic SE formulation is obtained in [29] which does not depend on Gaussian uncertainty assumptions and is capable of bad data measurement detection. A similar estimator is used in [30], where a Bayesian alternative to WLS is proposed. It is shown that the Bayesian approach has specifically better performance in presence of non-Gaussian uncertainty. Unlike WLS (equation (1)), the Bayesian approach tends to estimate states as a conditional averaging operation:

$$\hat{\boldsymbol{x}} = E\{\boldsymbol{x}|\boldsymbol{z}\} = \int \boldsymbol{\alpha} f_{\boldsymbol{\alpha}|\boldsymbol{z}}(\boldsymbol{\alpha}|\boldsymbol{z}) d\boldsymbol{\alpha} \tag{10}$$

Calculating $E\{\boldsymbol{x}|\boldsymbol{z}\}$ depends on our knowledge of the distribution function $f_{\boldsymbol{x}|\boldsymbol{z}}$, which can be obtained using Bayes rule, the measurement functions, and statistical properties of the system. Citing availability of accurate knowledge of second order statistics as a shortcoming of MSE-based methods, in [31] an alternative DSSE formulation is presented as a *matrix completion* problem which can be efficiently solved for billions of entries. Using information-theoretic reasoning it is shown that the optimal performance of DSSE is bounded by the capacity of AMI communication channels in charge of transmitting measurement samples to system operator.

To reduce the size of the optimization problem and speed up the convergence of WLS for large-scale feeders, in [32], the concept of quasi-symmetric impedance matrix is employed. This is achieved by adding the following constraint to the conventional WLS:

$$\begin{aligned} &\min_{\boldsymbol{x}} (\boldsymbol{z} - \boldsymbol{h}(\boldsymbol{x}))^T \boldsymbol{W} (\boldsymbol{z} - \boldsymbol{h}(\boldsymbol{x})) \\ &s.t.\ \boldsymbol{g_0}(\boldsymbol{x}) = \boldsymbol{x} - \boldsymbol{x_0} - \boldsymbol{TRX} \cdot \boldsymbol{I}(\boldsymbol{x}) = 0 \end{aligned} \tag{11}$$

where, $\boldsymbol{x}$ and $\boldsymbol{x_0}$ represent the voltage node state vector and the substation voltage, respectively. $\boldsymbol{TRX}$ denotes the reduced impedance matrix and $\boldsymbol{I}$ is the set of nodal current injections.

TABLE II
AVAILABLE DSSE FORMULATION STRUCTURES

Reference	Approach	State Variables	Pros and Cons
[33]	Voltage-based	Magnitude and phase angle	(+) Including general measurement functions, fast convergence, (-) State dependent and impedance-dependent Jacobian, radial topology only
[35]	Voltage-based	Magnitude and phase angle	(+) Linear formulation, non-iterative direct solution, applies to meshed topology, (-) Small angle difference assumption
[36]	Voltage-based	Magnitude and phase angle	(+) Branch-based formulation and computational efficiency, low sensitivity to network impedance, (-) Radial topology only
[34]	Voltage-based	Real and imaginary parts	(+) State-independent Jacobian, (-) Need for obtaining current-based measurement functions, radial topology only
[32]	Voltage-based	Real and imaginary parts	(+) High computational efficiency, applies to meshed networks, (-) Formulation complexity, impedance-dependency
[37]	Current-based	Real and imaginary parts	(+) Formulation and computational efficiency, phase-based decoupling in radial topology, impedance-independent Jacobian, (-) State-dependent Jacobian, exclusion of voltage measurements, phase-based coupling in meshed topology
[38]	Current-based	Real and imaginary parts	(+) Phase-based and state-based decoupling for all topologies, efficient handling of current measurements, constant impedance, (-) Small angle difference assumption, exclusion of voltage measurements
[39][40]	Current-based	Real and imaginary parts	(+) inclusion of voltage measurements, phase-based decoupling for radial topology, (-) State-dependent Jacobian
[41]	Current-based	Magnitude and phase angle	(+) Efficient handling of current measurements, broad range of measurement functions, (-) State-dependent and impedance-dependent Jacobian
[42]	Current-based	Magnitude/phase and real/imaginary	(+) Efficient handling of PMU data, (-) State- and impedance-dependent Jacobian, accuracy decline for meshed topology

III. DSSE PROBLEM FORMULATION

Due to the basic differences between transmission and distribution systems, the DSSE problem formulation can have major deviations from the conventional SE. The main point of difference is the modeling of measurement function ($\boldsymbol{h}$) in DSSE, as this function reflects the power flow equations in the power system. Hence, based on the choice of state and measured variables, choice of AC versus DC Power Flow (PF), and the representation of phases in power flow equations (for application in unbalanced systems), the measurement function can have different forms. In this section, we review the two basic formulations of DSSE (in terms of choice of state variables and measurement function) provided in the literature.

A) Voltage-Based DSSE: Traditionally, bus voltage magnitude and phase angle values have been used as state variables in transmission systems [1]. This conventional approach has also been employed in DSSE [33]–[36].

B) Branch-Current-Based SE (BCSE): A notable group of works, have adopted branch current as state variables, which turns out to be a more natural way of DSSE formulation for distribution systems [37]–[42]. A summary of the properties of different DSSE formulations is shown in Table II.

IV. DISTRIBUTION SYSTEM OBSERVABILITY

"Observability" refers to the system operator's ability to solve the state estimation problem. This depends on the number and location of metering instruments in the power system. Also, the availability and quality of critical measurement data samples in real-time has a crucial impact on power system observability. Conventionally, numerical and topological methods have been used to assess the observability of transmission systems with respect to the number and location of meters, as demonstrated in [1]. Alternative observability assessment procedures have been employed at distribution level. For instance, in [43] a probabilistic approach is adopted to define an Unobservability Index (UI) as follows:

$$UI = \sum_{i=1}^{n} K_i = \sum_{i=1}^{n} \left(\sum_{j=1}^{B_i} -p\big(b_{i,j}\big) \log_2 p\big(b_{i,j}\big) \right) \tag{12}$$

where, K_i denotes the entropy of the i^{th} state (with $p(b_{i,j})$ defining the probability of the j^{th} bin for the i^{th} state.) Basically, UI represents our overall uncertainty on the distribution system state variable values. As another example, a graph-theoretic criterion for local observability assessment of distribution networks is obtained in [4].

Unlike transmission systems that enjoy a high level of data redundancy, the distribution systems are generally underdetermined with poor observability. Thus, the accuracy of DSSE can be highly affected by the quality and availability of sensor data. The distribution system can easily become unobservable in case of communication failure/delays. Hence, bad/missing measurement data is closely connected to measurement redundancy and preserving the reliability of the DSSE problem. "Bad" data refer to data measurements that have considerable deviation from the underlying actual behavior, due to meter malfunction and communication noise. Missing data can also be treated as a special case of bad data. Conventionally, at the transmission level, bad data detection

TABLE III
AVAILABLE LITERATURE ON PSEUDO-MEASUREMENT GENERATION

<table>
<tr><th>Reference</th><th>Solution Approach</th><th>Pros and Cons</th><th>Load Estimation Model Input</th><th>Verification Approach</th></tr>
<tr><td>[45]</td><td>Probability density estimation using Beta functions</td><td>(+) Accurate empirical estimation, employing temporal load correlation, (-) Slow rate of convergence, radial only</td><td>Historic power consumption</td><td rowspan="2">Chi-square goodness of fit</td></tr>
<tr><td>[46] [47]</td><td>GMM</td><td>(+) Applicable to arbitrary load distributions, captures temporal load correlation, (-) Sensitivity to number of mixture components, expensive for high-dimensional learning</td><td rowspan="2">Standard load profiles, Historic data</td></tr>
<tr><td>[48]</td><td>Statistical profile construction</td><td>(+) Captures temporal correlation, simple formulation, (-) Gaussian load distribution assumption</td><td>Peak load estimation error</td></tr>
<tr><td>[15]</td><td>Statistical power loss estimation</td><td>(+) Capturing temporal and spatial error correlations, (-) Gaussian error distribution assumption</td><td>Available measurements and estimated loss</td><td>Relative error in power loss and estimation outcomes</td></tr>
<tr><td>[6]</td><td>Statistical load variation modeling</td><td>(+) Captures load correlation at different time stamps, addresses non-synchronized measurement, (-) Gaussian estimation error distribution assumption</td><td rowspan="6">Smart meter data history, weather-related variables and load duration curve</td><td>Anderson-Darling test, Shapiro-Wilk test</td></tr>
<tr><td>[49]</td><td>Single Gaussian distribution</td><td>(+) Captures active/reactive power temporal correlation, (-) Gaussian load/estimation error distribution assumption</td><td rowspan="4">Clustering validity measures, estimation mean absolute error</td></tr>
<tr><td>[50]</td><td>PNN, clustering</td><td>(+) Applicable to arbitrary load distribution, captures correlation between consumption and economic code, (-) Sensitivity to cluster number</td></tr>
<tr><td>[17]</td><td>PDP</td><td>(+) Robustness against measurement errors, corrective closed-loop system, (-) High computational cost</td></tr>
<tr><td>[18]</td><td>NARX</td><td>(+) No a priori knowledge on load structure required, quickly adapts to changes in load pattern, (-) Gaussian load estimation error assumption</td></tr>
<tr><td>[51]</td><td>Clustering, GMM</td><td>(+) Applicable to arbitrary load distributions, captures temporal (monthly) consumption correlations, (-) Sensitivity to cluster number, high computational cost, sensitivity to number of mixture components</td><td rowspan="2">DSSE relative error (voltage magnitude and phase), Bayesian information criteria</td></tr>
<tr><td>[52]</td><td>ANN, GMM</td><td>(+) Applicable to arbitrary estimation error distributions, (-) High computational cost, sensitivity to number of mixture components</td><td>Line flow measurements (generic consumption data)</td></tr>
</table>

has been performed by inspecting the normalized measurement residuals. However, this method is subject to failure and complications in case of insufficient measurement redundancy and multiple sources of bad data [1]. Hence, alternative approaches have been employed to address this problem, along with the sub-problem of missing data, at the distribution level (refer to Section II.)

Hence, to improve the observability of distribution systems, the input measurement set needs to be artificially augmented (to compensate for missing data) or corrected (to compensate for bad data.) This can be done through employing "pseudo-measurement" samples, which are artificially-generated data-points (e.g., active/reactive power, voltage and current, etc.) based on the data history of the distribution systems [5]. A basic approach is to use standard load profiles for generating pseudo-measurements [44]. Given that these data-points are not highly accurate, they introduce high variance levels in the weight matrix ($\boldsymbol{W}$), which could even lead to ill-conditioning of the DSSE problem. Data-driven approaches are employed for generating pseudo-measurements and handling their uncertainty, including probabilistic and statistical analysis, and machine-learning-based techniques.

A) Probabilistic and Statistical Approaches: Methods based on probabilistic and statistical techniques, which employ spatial/temporal correlation and historic probability distribution data, are widely used for generating reasonable pseudo-measurements and assessing their uncertainty. This includes empirical studies [45], Gaussian Mixture Models (GMMs) and Expectation Maximization (EM) [46], [47], time-varying variance and mean modeling [44], correlation analysis (between total and individual consumption) [48], nodal active-reactive correlation analysis [15], internodal and intranodal correlation modeling [16], intertemporal correlation analysis [6], multi-variate complex Gaussian modeling [49], and constrained optimization [50].

B) Learning-Based Approaches: Machine learning algorithms have also attracted scientific attention in solving DSSE problems, including addressing the problem of active/reactive power pseudo-measurement generation and uncertainty assessment. Probabilistic Neural Networks (PNNs) [51], Artificial Neural Network (ANN) [52], clustering algorithms [53], Parallel Distributed Processing networks (PDP) [17], and Nonlinear Auto-Regressive eXogenous (NARX) [18].

A summary of the notable papers in these two categories are shown in Table III. Pseudo-measurement generation is basically a special type of load estimation at distribution level. While there is a considerable number of works done in this area, still unanswered questions remain. For instance, most of the papers, instead of using real AMI data history, rely on standard load profiles to perform numerical analysis and verification. Also, the huge amount of data in practice can cause certain learning methods to become computationally expensive. Managing this "big data" challenge in distribution systems requires further research and studies.

V. Network Topology and Configuration

The topology identification problem can be categorized into two separate, yet related, subproblems:

A) System configuration identification: The basic assumption within this set of problems is that the basic topology of the network is known to the system operator. However, due to local events (such as faults, line disconnections, switching events, etc.) the basic topology will undergo local changes over time. Limited knowledge of the operator on these changes will affect the accuracy of SE solutions. Hence, the objective is to use the system-wide measurements to update our knowledge of system configuration to avoid topology errors (i.e., state of switches, fuses, lines, DG/customer connection status.) Conventionally, generalized SE models have been used at the transmission level (with switch-related variables added to the SE formulation) to detect and correct topological errors [1], [54]. Similar classic methods have been applied to DSSE as well [55], [56]. Apart from the classical approaches, other probabilistic and data-driven methods have been applied for topology detection and identification in distribution systems. These methods are usually based on a data-driven search process in a limited topology space (i.e., topology library) defined by variations on the basic topology, as shown in Fig. 2. Probabilistic recursive Bayesian approach [7], [57], fuzzy-based pattern recognition [58], auto-encoders [59], PMU voltage time-series [60], voting technique ("vote" for the best candidate structure) [61], correlation analysis [62], and maximum likelihood estimation [63], are a few of the proposed topology search methods.

B) Topology learning: Another set of problems are based on the assumption that the system operator has very limited or no knowledge of the basic topology of the network (which is highly applicable to the secondary distribution networks.) The objective is to discover the topology of the network by relying on nodal and branch measurements. Graph-theoretic algorithms have been used widely for topology discovery and learning considering different assumptions on system operator's knowledge on topology. A sparse graph recovery model has been adopted in [64] to perform topology discovery, based on DC PF. The proposed method, which is based on nodal measurements, requires no *a priori* information on the topology of the network. Another data-driven graphical approach towards topology learning is proposed in [65]. In this work, an efficient graphical model is developed to represent the voltage magnitude dependencies (using mutual information as a measure of affinity) between neighboring buses (the basic assumption in this work is that current injections are statistically independent.) This method only depends on statistics of nodal voltage magnitude measurements (smart meter data) to reconstruct the partially or fully unknown radial or weakly-meshed topology. It is shown that for a radial feeder, the spanning tree that maximizes measures of internodal voltage mutual information corresponds to the true topology of the system. In [66], using nodal voltage measurements, the authors have been able to learn the topology of a radial feeder using mutual statistical properties of the measured variables. The proposed model is based on a linear approximation of lossless AC PF, and employs a bottom-to-top approach, in which the structure learning begins with the end nodes and moves towards the substation by choosing the proper parent nodes at each stage. The method is shown to have acceptable performance under a wide variety of assumptions, including no prior knowledge on the basic topology and missing measurement data. In [67], graph-theoretic interpretation of principal component analysis and energy conservation are employed in the context of graph theory to obtain radial distribution system topology through smart meter energy usage data. A more general approach (applicable to meshed networks even with missing PMU phase measurements) for estimating both the topology of the network and the line parameters is proposed in [68], where the line parameters and system topology are updated consecutively through an EM-based approach. Starting with an initial topology guess, at each step of the algorithm, the topology is updated by removing edges with small estimated susceptance values to improve the estimation likelihood.

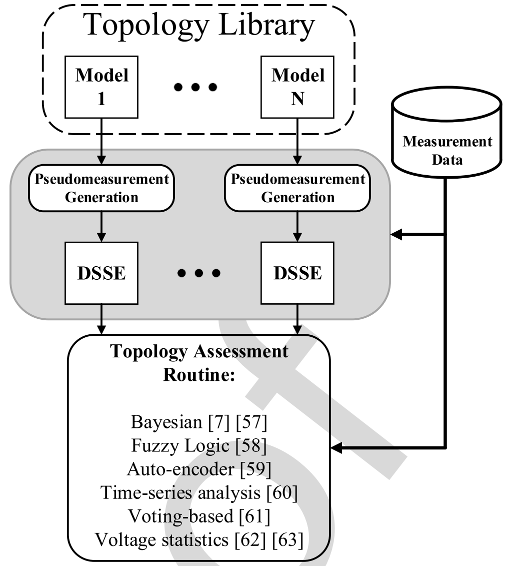


Fig. 2. Data-driven system configuration detection.

VI. Distribution Network Metering System Design and Analysis

A. Metering Instrument Placement

Optimizing the location of metering instruments in distribution systems is a significant subject for research, given the size of the system and potentially limited financial resources [69]. Different objectives have been proposed in the literature to address this problem, including improving system observability, minimizing installation/maintenance costs, bad data detection capability, and improving the DSSE

TABLE IV
METER PLACEMENT METHODS

Reference	Objective Function	Constraints	Solution Algorithm
[69]	Meter cost	Estimation accuracy	Heuristic search
[70] [71]	Estimation accuracy	X	Ordinal optimization
[72][73]	Meter cost, Estimation accuracy	Estimation accuracy	GA
[74]	Estimation accuracy	X	GA
[75]	Determinant of the overall Fisher information matrix	Meter number	Boolean-convex optimization
[76][77]	Estimation accuracy	Meter number	MISDP
[78]	Estimation accuracy	Meter number	Submodular saturation algorithm
[43]	Network observability	X	Heuristic search
[79]	Meter cost, estimation accuracy	Estimation accuracy	MOE

accuracy [41], [43], [70]–[80]. Different algorithms have been tried for solving the placement problem, including Genetic-Algorithm (GA), Mixed Integer Linear Programming (MILP), Mixed Integer Semi-Definite Programming (MISDP), and Multi-Objective Evolutionary (MOE) methods. A summary of the different meter placement approaches is given in Table IV.

B. PMU Applications and Impacts on DSSE

PMUs are able to provide synchronized voltage, power, and current measurements that enable accurate tracking of state variables and efficient control and management decisions. Also, generally the sampling frequency of PMUs (up to 30 kHz) is much higher than that of smart meters (0.277 mHz - 16.7 mHz), which leads to system observability on a higher temporal granularity. However, compared to smart meters, the use of PMUs in distribution networks is still very restricted due to their prohibitive costs. Hence, a critical research direction related to PMUs is optimizing the number and location of PMUs to enhance system observability, while limiting the measurement infrastructure costs [81] (also see Section VI-A).

In terms of application in distribution systems, PMUs have been employed for high-resolution voltage/power profiling, oscillation detection, topology identification, and event detection, as outlined in [82]. On the other hand, smart meters have been used mostly for low-resolution load forecasting and management, and connection verification [83]. In terms of algorithm design for DSSE and topology identification, one considerable difference between the methods proposed for systems with only smart meters and systems with PMUs is the "small phase angle difference assumption". Hence, due to unavailability of phase angle data in absence of PMUs many papers have assumed that the nodal voltage phase angles in a system are almost equal [65], [66], [84]. While this assumption introduces bounded inaccuracies in the final estimation/identification outcomes, it enables system operators to monitor the state of distribution systems without PMUs. Furthermore, adding the voltage phase data or flow measurements can highly improve the estimation and identification routines' performance.

VII. PENETRATION OF RENEWABLE RESOURCES

A few papers have analyzed DSSE under high penetration rates of renewable power. The main source of challenge in performing SE in presence of renewable resources is their uncertain output power [85]. Also, deep penetration of renewable power sources affect the voltage profile of distribution systems. This stresses the need for more advanced voltage monitoring capabilities [50]. In case of pseudo-measurement generation for these resources, it is believed that the non-Gaussian distribution of renewable power would adversely affect conventional WLS-based DSSE methods. Moreover, as shown in [14], fast changes in system state can result in the WLS-based DSSE to get trapped in local minima with errors as high as 10^5 times the underlying global solution. Also, given that the performance of conventional Gauss-Newton algorithm highly depends on the initial conditions, finding good initial conditions for DSSE in systems with deep renewable penetration is a difficult task [4]. To address these challenges several papers have adopted different approaches for solving the SE (in general) and DSSE (in particular) in presence of renewable-based DGs.

Probabilistic methods represent the major group of techniques for modeling the impacts of renewable uncertainty on SE. A forecasting-aided SE mechanism is proposed in [86] to capture the temporal and spatial correlation among DGs and loads for their short-term prediction (to be used as pseudo-measurements in SE), using a linear autoregressive model. In [29], another forecasting-aided SE method is proposed to manage the uncertainties of load and renewable resources based on a GMM technique for obtaining the non-Gaussian distribution of renewable power while incorporating the dynamics of the system. Moreover, this estimator shows good performance even with limited data, which makes it a promising candidate for DSSE. As an extension to [73], the effect of the uncertainty of renewable DG power profile on meter placement has been modeled in [87] using GMM. A probabilistic graphical modeling technique has been proposed in [88] for capturing short term uncertainty of SE in systems with high PV penetration. The physical governing laws of the system (i.e., PF equations) have been embedded into the SE model. A distributed belief propagation method is performed for state inference, which yields superior results compared to the conventional deterministic WLS method. Another probabilistic approach is adopted in [44] for pseudo-measurement generation in networks with high residential PV penetration using Beta distribution functions. It is speculated that the uncertainty of PV systems has the highest impact on the DSSE at mid-day time intervals (when usually the load profile is not peaking.) To model the non-Gaussian uncertainty of PV power in DSSE, pseudo-measurements are generated (with 15-minute time resolution) for roof-top dispersed PV systems employing a weather-dependent model for constructing general PV power probability density functions, considering solar radiation, temperature, number of arrays and their physical characteristics. This approach shows considerable improvements on DSSE accuracy compared to using conventional standard profiles. While in [44] the possible correlation between physically nearby renewable DGs are not modeled, it is demonstrated

TABLE V
AVAILABLE LITERATURE ON SE CYBER-SECURITY

Reference	Topic	System Model	Attacker Model	Model Constraints	Proposed Countermeasures
[89]	False data injection	DC PF	Full system knowledge	Limited resource and meter access	Spatial and temporal-based detection
[90]			Limited topology knowledge		Generalized likelihood ratio detector
[91]		AC and DC PF	Full system knowledge	Detectability constraints	Load shift factor monitoring
[94]			Local state knowledge	Limits on system loss knowledge	X
[97]	Topology Attack		Local system information	Noisy measurements	Cover-up meter protection strategy
[98]			Wide range of system knowledge	Access to the Jacobian matrix	Topology information masking
[99]		AC PF	Full system knowledge	Limited resource	Generation forecasting
[95]				Unlimited access to all meters	X
[15]	Data Privacy		X	Limited access to smart meters	Smart meter data aggregation
[100]		Linear Dynamic	Monitoring control signals	Gaussian noise assumption	Proper channel capacity assignment

in [16] that including the correlation between close DGs for pseudo-measurement generation leads to further improvements in DSSE accuracy.

VIII. CYBER-SECURITY

The vulnerability of the power system against cyber-attacks has been observed in practice. Different types of cyber-attack related to SE have been modeled and investigated in the literature: false data injection, topology attacks, and eavesdropping. In a false data injection situation, an attacker, with various degrees of knowledge on system parameters and states, alters the metered data of certain metering devices [89]–[94]. In a topology attack, the attackers tend to maliciously modify the topology model data of the system [95]–[99]. Eavesdropping defines a situation in which an unauthorized party seeks to gather system data by tapping into the communication infrastructure, compromising data privacy and confidentiality of users [15], [100]. A classification of different papers with respect to the issue of cyber-security can be seen in Table V. It can be concluded that protecting the vital automation and monitoring systems against cyber intrusion and cyber attacks requires a holistic approach to preserve the integrity, availability, and confidentiality of DSSE at all times. Different components of an effective solution include: adversary identification (in terms of knowledge and resource levels), vulnerability assessment (critical meters, communication system integration, sensitivity of DSSE to bad data), and personnel training.

IX. CONCLUSION

In this paper, we have presented an overview of the critical aspects of DSSE. Active research subjects, such as DSSE problem formulation, pseudo-measurement generation, network topology, data meter placement, renewable resource integration, and cyber-security are reviewed. Based on the survey, most recent works are more concentrated on using data-driven and machine-learning-based modifications in the conventional DSSE (for improving the accuracy, robustness, and system observability), which is a reasonable direction given the steep increase in the rate of installation of smart meters and micro-PMUs at the distribution level. Probabilistic modeling (in a data-driven context) has also attracted substantial research works, due to its capability for capturing the effects of stochastic and variable renewable resources on active distribution systems in general (and on DSSE in particular.)

A. Future Research Directions

It would be of interest to study how Demand Response (DR) programs [101] could impact the DSSE (in terms of uncertainty and variability of customer behavior and pseudo-load generation) by incorporating retail market signals into the DSSE problem formulation. In general, integrating the price-sensitivity of active distribution networks into the DSSE becomes a valid research problem in future distribution systems with deep penetration of renewable and DR resources. In a related context, optimal power management and decision making under limited distribution system observability appears to be a largely unexplored direction for research, specially in presence of emergent technologies, such as energy storage systems and networked microgrids [102]–[105]. Another very recent area of interest is topology learning. Future research is needed to discover if and how topology discovery can be performed after extreme weather events [106] as the number of data meters decreases due to communication and device failure, and the observability of the distribution system is compromised. Employing data-driven methods under extreme weather events at different stages (pre-event, during the event, and post-event) for developing system monitoring and learning techniques is another possible research direction. Thus, it would be of interest to investigate the impact of extreme events on distribution system observability and design potential solution strategies to enable effective system restoration strategies that depend on operator's real-time knowledge of system states.

ACKNOWLEDGMENT

The authors would like to thank Dr. Ravindra Singh from the Argonne National Laboratory for valuable comments that greatly improved the manuscript.

REFERENCES

[1] A. Monticelli, *State Estimation in Electric Power Systems: A Generalized Approach*. New York, NY, USA: Springer, 1999. AQ3
[2] (Dec. 2017). *FERC Staff Report: Assessment of Demand Response and Advanced Metering*. [Online]. Available: https://www.ferc.gov/legal/staff-reports/2017/DR-AM-Report2017.pdf
[3] A. Primadianto and C.-N. Lu, "A review on distribution system state estimation," *IEEE Trans. Power Syst.*, vol. 32, no. 5, pp. 3875–3883, Sep. 2017.
[4] S. Bhela, V. Kekatos, and S. Veeramachaneni, "Enhancing observability in distribution grids using smart meter data," *IEEE Trans. Smart Grid*, to be published. [Online]. Available: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7914787 AQ4
[5] F. F. Wu, "Power system state estimation: A survey," *Int. J. Elect. Power Energy Syst.*, vol. 12, no. 2, pp. 80–87, Apr. 1990.
[6] A. Alimardani, F. Therrien, D. Atanackovic, J. Jatskevich, and E. Vaahedi, "Distribution system state estimation based on nonsynchronized smart meters," *IEEE Trans. Smart Grid*, vol. 6, no. 6, pp. 2919–2928, Nov. 2015.

[7] R. Singh, E. Manitsas, B. C. Pal, and G. Strbac, "A recursive Bayesian approach for identification of network configuration changes in distribution system state estimation," *IEEE Trans. Power Syst.*, vol. 25, no. 3, pp. 1329–1336, Aug. 2010.

[8] *Eaton's CYME Distribution State Estimator*. [Online]. Available: http://www.eaton.com/us/en-us/products/utility-grid-solutions/software-modules/distribution-state-estimator-module.html

[9] *Survalent Developing Distribution State Estimation Software for Electric Utilities*. [Online]. Available: https://www.survalent.com/survalent-developing-distribution-state-estimation-software-electric-utilities/

[10] *ETAP: Distribution State Estimation*. [Online]. Available: https://etap.com/product/distribution-state-estimation

[11] *OSI: Advanced Distribution Management Systems*. [Online]. Available: http://www.osii.com/solutions/products/distribution-management.asp

[12] *Nexant Launches Grid360 Distribution Manager*. [Online]. Available: http://www.nexant.com/about/news/nexant-launches-grid360-distribution-manager

[13] D. Ablaković, I. Dzafic, R. A. Jabr, and B. C. Pal, "Experience in distribution state estimation preparation and operation in complex radial distribution networks," in *Proc. IEEE PES Gen. Meeting*, Jul. 2014, pp. 1–5.

[14] Y. Weng, R. Negi, C. Faloutsos, and M. D. Ilić, "Robust data-driven state estimation for smart grid," *IEEE Trans. Smart Grid*, vol. 8, no. 4, pp. 1956–1967, Jul. 2017.

[15] Q. Chen, D. Kaleshi, Z. Fan, and S. Armour, "Impact of smart metering data aggregation on distribution system state estimation," *IEEE Trans. Ind. Informat.*, vol. 12, no. 4, pp. 1426–1437, Aug. 2016.

[16] C. Muscas, M. Pau, P. A. Pegoraro, and S. Sulis, "Effects of measurements and pseudomeasurements correlation in distribution system state estimation," *IEEE Trans. Instrum. Meas.*, vol. 63, no. 12, pp. 2813–2823, Dec. 2014.

[17] J. Wu, Y. He, and N. Jenkins, "A robust state estimator for medium voltage distribution networks," *IEEE Trans. Power Syst.*, vol. 28, no. 2, pp. 1008–1016, May 2013.

[18] B. P. Hayes, J. K. Gruber, and M. Prodanovic, "A closed-loop state estimation tool for MV network monitoring and operation," *IEEE Trans. Smart Grid*, vol. 6, no. 4, pp. 2116–2125, Jul. 2015.

[19] Y. Nie, C. Y. Chung, and N. Z. Xu, "System state estimation considering EV penetration with unknown behavior using quasi-Newton method," *IEEE Trans. Power Syst.*, vol. 31, no. 6, pp. 4605–4615, Nov. 2016.

[20] Y. Weng, Q. Li, R. Negi, and M. Illić, "Semidefinite programming for power system state estimation," in *Proc. IEEE Power Energy Soc. Gen. Meeting*, San Diego, CA, USA, Jul. 2012, pp. 1–8.

[21] Y. Yao, X. Liu, D. Zhao, and Z. Li, "Distribution system state estimation: A semidefinite programming approach," *IEEE Trans. Smart Grid*, to be published. [Online]. Available: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8416770

[22] H. Zhu and G. B. Giannakis, "Power system nonlinear state estimation using distributed semidefinite programming," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 6, pp. 1039–1050, Dec. 2014.

[23] W.-M. Lin and J.-H. Teng, "State estimation for distribution systems with zero-injection constraints," *IEEE Trans. Power Syst.*, vol. 11, no. 1, pp. 518–524, Feb. 1996.

[24] L. Mili, M. G. Cheniae, and P. J. Rousseeuw, "Robust state estimation of electric power systems," *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, vol. 41, no. 5, pp. 349–358, May 1994.

[25] M. Göl and A. Abur, "LAV based robust state estimation for systems measured by PMUs," *IEEE Trans. Smart Grid*, vol. 5, no. 4, pp. 1808–1814, Jul. 2014.

[26] J. Zhao, G. Zhang, M. L. Scala, and Z. Wang, "Enhanced robustness of state estimator to bad data processing through multi-innovation analysis," *IEEE Trans. Ind. Informat.*, vol. 13, no. 4, pp. 1610–1619, Aug. 2017.

[27] U. Kuhar, M. Pantos, G. Kosec, and A. Švigelj, "The impact of model and measurement uncertainties on a state estimation in three-phase distribution networks," *IEEE Trans. Smart Grid*, to be published, [Online]. Available: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8331939

[28] S. Wang, J. Zhao, Z. Huang, and R. Diao, "Assessing Gaussian assumption of PMU measurement error using field data," *IEEE Trans. Power Del.*, to be published, [Online]. Available: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8067439

[29] H. Bilil and H. Gharavi, "MMSE-based analytical estimator for uncertain power system with limited number of measurements," *IEEE Trans. Power Syst.*, to be published. [Online]. Available: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8279475

[30] P. A. Pegoraro *et al.*, "Bayesian approach for distribution system state estimation with non-Gaussian uncertainty models," *IEEE Trans. Instrum. Meas.*, vol. 66, no. 11, pp. 2957–2966, Nov. 2017.

[31] C. Genes, I. Esnaola, S. M. Perlaza, L. F. Ochoa, and D. Coca, "Robust recovery of missing data in electricity distribution systems," *IEEE Trans. Smart Grid*, to be published. [Online]. Available: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8388253

[32] P. M. De Oliveira-De Jesus and A. A. R. Quintana, "Distribution system state estimation model using a reduced quasi-symmetric impedance matrix," *IEEE Trans. Power Syst.*, vol. 30, no. 6, pp. 2856–2866, Nov. 2015.

[33] M. E. Baran and A. W. Kelley, "State estimation for real-time monitoring of distribution systems," *IEEE Trans. Power Syst.*, vol. 9, no. 3, pp. 1601–1609, Aug. 1994.

[34] C. N. Lu, J. H. Tang, and W.-H. E. Liu, "Distribution system state estimation," *IEEE Trans. Power Syst.*, vol. 10, no. 1, pp. 229–240, Feb. 1995.

[35] D. A. Haughton and G. T. Heydt, "A linear state estimation formulation for smart distribution systems," *IEEE Trans. Power Syst.*, vol. 28, no. 2, pp. 1187–1195, May 2013.

[36] Y. Deng, Y. He, and B. Zhang, "A branch-estimation-based state estimation method for radial distribution systems," *IEEE Trans. Power Del.*, vol. 17, no. 4, pp. 1057–1062, Oct. 2002.

[37] M. E. Baran and A. W. Kelley, "A branch-current-based state estimation method for distribution systems," *IEEE Trans. Power Syst.*, vol. 10, no. 1, pp. 483–491, Feb. 1995.

[38] W.-M. Lin, J.-H. Teng, and S.-J. Chen, "A highly efficient algorithm in treating current measurements for the branch-current-based distribution state estimation," *IEEE Trans. Power Del.*, vol. 16, no. 4, pp. 433–439, Jul. 2001.

[39] J.-H. Teng, "Using voltage measurements to improve the results of branch-current-based state estimators for distribution systems," *IEE Proc. Gener. Transm. Distrib.*, vol. 149, no. 6, pp. 667–672, Nov. 2002.

[40] M. E. Baran, J. Jung, and T. E. McDermott, "Including voltage measurements in branch current state estimation for distribution systems," in *Proc. IEEE Power Energy Soc. Gen. Meeting*, Calgary, AB, Canada, Jul. 2009, pp. 1–5.

[41] H. Wang and N. N. Schulz, "A revised branch current-based distribution system state estimation algorithm and meter placement impact," *IEEE Trans. Power Syst.*, vol. 19, no. 1, pp. 207–213, Feb. 2004.

[42] M. Pau, P. A. Pegoraro, and S. Sulis, "Efficient branch-current-based distribution system state estimation including synchronized measurements," *IEEE Trans. Instrum. Meas.*, vol. 62, no. 9, pp. 2419–2429, Sep. 2013.

[43] B. Brinkmann and M. Negnevitsky, "A probabilistic approach to observability of distribution networks," *IEEE Trans. Power Syst.*, vol. 32, no. 2, pp. 1169–1178, Mar. 2017.

[44] A. Angioni, T. Schlosser, F. Ponci, and A. Monti, "Impact of pseudo-measurements from new power profiles on state estimation in low-voltage grids," *IEEE Trans. Instrum. Meas.*, vol. 65, no. 1, pp. 70–77, Jan. 2016.

[45] A. K. Ghosh, D. L. Lubkeman, M. J. Downey, and R. H. Jones, "Distribution circuit state estimation using a probabilistic approach," *IEEE Trans. Power Syst.*, vol. 12, no. 1, pp. 45–51, Feb. 1997.

[46] R. Singh, B. C. Pal, and R. A. Jabr, "Statistical representation of distribution system loads using Gaussian mixture model," *IEEE Trans. Power Syst.*, vol. 25, no. 1, pp. 29–37, Feb. 2010.

[47] R. Singh, B. C. Pal, and R. A. Jabr, "Distribution system state estimation through Gaussian mixture model of the load as pseudo-measurement," *IET Gener. Transm. Distrib.*, vol. 4, no. 1, pp. 50–59, Jan. 2009.

[48] D. T. Nguyen, "Modeling load uncertainty in distribution network monitoring," *IEEE Trans. Power Syst.*, vol. 30, no. 5, pp. 2321–2328, Sep. 2015.

[49] A. Arefi, G. Ledwich, and B. Behi, "An efficient DSE using conditional multivariate complex Gaussian distribution," *IEEE Trans. Smart Grid*, vol. 6, no. 4, pp. 2147–2156, Jul. 2015.

[50] I. Džafić and R. A. Jabr, "Real time multiphase state estimation in weakly meshed distribution networks with distributed generation," *IEEE Trans. Power Syst.*, vol. 32, no. 6, pp. 4560–4569, Nov. 2017.

[51] D. Gerbec, S. Gasperic, I. Smon, and F. Gubina, "Allocation of the load profiles to consumers using probabilistic neural networks," *IEEE Trans. Power Syst.*, vol. 20, no. 2, pp. 548–555, May 2005.

[52] E. Manitsas, R. Singh, B. C. Pal, and G. Strbac, "Distribution system state estimation using an artificial neural network approach for pseudo measurement modeling," *IEEE Trans. Power Syst.*, vol. 27, no. 4, pp. 1888–1896, Nov. 2012.

[53] Y. R. Gahrooei, A. Khodabakhshian, and R.-A. Hooshmand, "A new pseudo load profile determination approach in low voltage distribution networks," *IEEE Trans. Power Syst.*, vol. 33, no. 1, pp. 463–472, Jan. 2018.

[54] E. M. Lourenço, E. P. R. Coelho, and B. C. Pal, "Topology error and bad data processing in generalized state estimation," *IEEE Trans. Power Syst.*, vol. 30, no. 6, pp. 3190–3200, Nov. 2015.

[55] G. N. Korres and N. M. Manousakis, "A state estimation algorithm for monitoring topology changes in distribution systems," in *Proc. IEEE Power Energy Soc. Gen. Meeting*, San Diego, CA, USA, Jul. 2012, pp. 1–8.

[56] M. E. Baran, J. Jung, and T. E. McDermott, "Topology error identification using branch current state estimation for distribution systems," in *Proc. IEEE Transm. Distrib. Conf. Expo. Asia–Pac.*, Oct. 2009, pp. 1–4.

[57] B. Hayes, A. Escalera, and M. Prodanovic, "Event-triggered topology identification for state estimation in active distribution networks," in *Proc. IEEE PES Innov. Smart Grid Technol. Conf. Europe (ISGT-Europe)*, Oct. 2016, pp. 1–6.

[58] D. Singh, J. P. Pandey, and D. S. Chauhan, "Topology identification, bad data processing, and state estimation using fuzzy pattern matching," *IEEE Trans. Power Syst.*, vol. 20, no. 3, pp. 1570–1579, Aug. 2005.

[59] V. Miranda, J. Krstulovic, H. Keko, C. Moreira, and J. Pereira, "Reconstructing missing data in state estimation with autoencoders," *IEEE Trans. Power Syst.*, vol. 27, no. 2, pp. 604–611, May 2012.

[60] G. Cavraro and R. Arghandeh, "Power distribution network topology detection with time-series signature verification method," *IEEE Trans. Power Syst.*, vol. 33, no. 4, pp. 3500–3509, Jul. 2018.

[61] R. Arghandeh *et al.*, "Topology detection in microgrids with micro-synchrophasors," in *Proc. IEEE Power Energy Soc. Gen. Meeting*, Denver, CO, USA, Jul. 2015, pp. 1–5.

[62] W. Luan, J. Peng, M. Maras, J. Lo, and B. Harapnuk, "Smart meter data analytics for distribution network connectivity verification," *IEEE Trans. Smart Grid*, vol. 6, no. 4, pp. 1964–1971, Jul. 2015.

[63] G. Cavraro, V. Kekatos, and S. Veeramachaneni, "Voltage analytics for power distribution network topology verification," *IEEE Trans. Smart Grid*, to be published. [Online]. Available: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8054732

[64] M. Babakmehr, M. G. Simões, M. B. Wakin, and F. Harirchi, "Compressive sensing-based topology identification for smart grids," *IEEE Trans. Ind. Informat.*, vol. 12, no. 2, pp. 532–543, Apr. 2016.

[65] Y. Weng, Y. Liao, and R. Rajagopal, "Distributed energy resources topology identification via graphical modeling," *IEEE Trans. Power Syst.*, vol. 32, no. 4, pp. 2682–2694, Jul. 2017.

[66] D. Deka, S. Backhaus, and M. Chertkov, "Structure learning in power distribution networks," *IEEE Trans. Control Netw. Syst*, vol. 5, no. 3, pp. 1061–1074, Sep. 2018. [Online]. Available: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7862849

[67] S. J. Pappu, N. Bhatt, R. Pasumarthy, and A. Rajeswaran, "Identifying topology of low voltage distribution networks based on smart meter data," *IEEE Trans. Smart Grid*, vol. 9, no. 5, pp. 5113–5122, Sep. 2018. [Online]. Available: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7875102

[68] J. Yu, Y. Weng, and R. Rajagopal, "PaToPa: A data-driven parameter and topology joint estimation framework in distribution grids," *IEEE Trans. Power Syst.*, vol. 33, no. 4, pp. 4335–4347, Jul. 2018.

[69] M. E. Baran, J. Zhu, and A. W. Kelley, "Meter placement for real-time monitoring of distribution feeders," *IEEE Trans. Power Syst.*, vol. 11, no. 1, pp. 332–337, Feb. 1996.

[70] H. Liu, D. Yu, and H.-D. Chiang, "A heuristic meter placement method for load estimation," *IEEE Trans. Power Syst.*, vol. 17, no. 3, pp. 913–917, Aug. 2002.

[71] R. Singh, B. C. Pal, and R. B. Vinter, "Measurement placement in distribution system state estimation," *IEEE Trans. Power Syst.*, vol. 24, no. 2, pp. 668–675, May 2009.

[72] R. Singh, B. C. Pal, R. A. Jabr, and R. B. Vinter, "Meter placement for distribution system state estimation: an ordinal optimization approach," *IEEE Trans. Power Syst.*, vol. 26, no. 4, pp. 2328–2335, Nov. 2011.

[73] J. Liu *et al.*, "Trade-offs in PMU deployment for state estimation in active distribution grids," *IEEE Trans. Smart Grid*, vol. 3, no. 2, pp. 915–924, Jun. 2012.

[74] G. Celli, P. A. Pegoraro, F. Pilo, G. Pisano, and S. Sulis, "DMS cyber-physical simulation for assessing the impact of state estimation and communication media in smart grid operation," *IEEE Trans. Power Syst.*, vol. 29, no. 5, pp. 2436–2446, Sep. 2014.

[75] A. Kumar, B. Das, and J. Sharma, "Genetic algorithm-based meter placement for static estimation of harmonic sources," *IEEE Trans. Power Del.*, vol. 20, no. 2, pp. 1088–1096, Apr. 2005.

[76] T. C. Xygkis, G. N. Korres, and N. M. Manousakis, "Fisher information-based meter placement in distribution grids via the D-optimal experimental design," *IEEE Trans. Smart Grid*, vol. 9, no. 2, pp. 1452–1461, Mar. 2018.

[77] T. C. Xygkis and G. N. Korres, "Optimized measurement allocation for power distribution systems using mixed integer SDP," *IEEE Trans. Instrum. Meas.*, vol. 66, no. 11, pp. 2967–2976, Nov. 2017.

[78] M. G. Damavandi, V. Krishnamurthy, and J. R. Marti, "Robust meter placement for state estimation in active distribution systems," *IEEE Trans. Smart Grid*, vol. 6, no. 4, pp. 1972–1982, Jul. 2015.

[79] S. Prasad and D. M. V. Kumar, "Trade-offs in PMU and IED deployment for active distribution state estimation using multi-objective evolutionary algorithm," *IEEE Trans. Instrum. Meas.*, vol. 67, no. 6, pp. 1298–1307, Jun. 2018.

[80] Y. Yao, X. Liu, and Z. Li, "Robust measurement placement for distribution system state estimation," *IEEE Trans. Sustain. Energy*, to be published. [Online]. Available: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8115294

[81] D. Carta, A. Benigni, and C. Muscas, "Model order reduction for PMU-based state estimation in distribution grids," *IEEE Syst. J.*, vol. 12, no. 3, pp. 2711–2720, Sep. 2018. [Online]. Available: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp&arnumber7919192

[82] A. von Meier, E. Stewart, A. McEachern, M. Andersen, and L. Mehrmanesh, "Precision micro-synchrophasors for distribution systems: a summary of applications," *IEEE Trans. Smart Grid*, vol. 8, no. 6, pp. 2926–2936, Nov. 2017.

[83] Y. Wang, Q. Chen, T. Hong, and C. Kang, "Review of smart meter data analytics: Applications, methodologies, and challenges," *IEEE Trans. Smart Grid*, to be published. [Online]. Available: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8322199

[84] J. Peppanen, M. J. Reno, R. J. Broderick, and S. Grijalva, "Distribution system model calibration with big data from AMI and PV inverters," *IEEE Trans. Smart Grid*, vol. 7, no. 5, pp. 2497–2506, Sep. 2016.

[85] M. Amini and M. Almassalkhi, "Trading off robustness and performance in receding horizon control with uncertain energy resources," in *Proc. Power Syst. Comput. Conf. (PSCC)*, 2018, pp. 1–7.

[86] J. Zhao, G. Zhang, Z. Y. Dong, and M. L. Scala, "Robust forecasting aided power system state estimation considering state correlations," *IEEE Trans. Smart Grid*, vol. 9, no. 4, pp. 2658–2666, Jul. 2018.

[87] J. Liu *et al.*, "Optimal meter placement for robust measurement systems in active distribution grids," *IEEE Trans. Instrum. Meas.*, vol. 63, no. 5, pp. 1096–1105, May 2014.

[88] Y. Weng, R. Negi, and M. D. Ilić, "Probabilistic joint state estimation for operational planning," *IEEE Trans. Smart Grid*, to be published. [Online]. Available: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8026175

[89] Q. Yang *et al.*, "On false data-injection attacks against power system state estimation: Modeling and countermeasures," *IEEE Trans. Parallel Distrib. Syst*, vol. 25, no. 3, pp. 717–729, Mar. 2014.

[90] S. Li, Y. Yilmaz, and X. Wang, "Quickest detection of false data injection attack in wide-area smart grids," *IEEE Trans. Smart Grid*, vol. 6, no. 6, pp. 2725–2735, Nov. 2015.

[91] J. Liang, L. Sankar, and O. Kosut, "Vulnerability analysis and consequences of false data injection attack on power system state estimation," *IEEE Trans. Power Syst.*, vol. 31, no. 5, pp. 3864–3872, Sep. 2016.

[92] Z.-H. Yu and W.-L. Chin, "Blind false data injection attack using PCA approximation method in smart grid," *IEEE Trans. Smart Grid*, vol. 6, no. 3, pp. 1219–1226, May 2015.

[93] Y. Chakhchoukh and H. Ishii, "Enhancing robustness to cyber-attacks in power systems through multiple least trimmed squares state estimations," *IEEE Trans. Power Syst.*, vol. 31, no. 6, pp. 4395–4405, Nov. 2016.

[94] R. Deng, P. Zhuang, and H. Liang, "False data injection attacks against state estimation in power distribution systems," *IEEE Trans. Smart Grid*, to be published. [Online]. Available: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8307441

[95] G. Liang, S. R. Weller, J. Zhao, F. Luo, and Z. Y. Dong, "A framework for cyber-topology attacks: line-switching and new attack scenarios," *IEEE Trans. Smart Grid*, to be published. [Online]. Available: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8118126

[96] X. Liu and Z. Li, "Local topology attacks in smart grids," *IEEE Trans. Smart Grid*, vol. 8, no. 6, pp. 2617–2626, Nov. 2017.

[97] J. Kim and L. Tong, "On topology attack of a smart grid: Undetectable attacks and countermeasures," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 7, pp. 1294–1305, Jul. 2013.

[98] Y. Chakhchoukh and H. Ishii, "Coordinated cyber-attacks on the measurement function in hybrid state estimation," *IEEE Trans. Power Syst.*, vol. 30, no. 5, pp. 2487–2497, Sep. 2015.

[99] J. Zhang and L. Sankar, "Physical system consequences of unobservable state-and-topology cyber-physical attacks," *IEEE Trans. Smart Grid*, vol. 7, no. 4, pp. 2016–2025, Jul. 2016.

[100] H. Li, L. Lai, and W. Zhang, "Communication requirement for reliable and secure state estimation and control in smart grid," *IEEE Trans. Smart Grid*, vol. 2, no. 3, pp. 476–486, Sep. 2011.

[101] T. Lu, Z. Wang, J. Wang, Q. Ai, and C. Wang, "A data-driven Stackelberg market strategy for demand response-enabled distribution systems," *IEEE Trans. Smart Grid*, to be published. [Online]. Available: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8263241

[102] Z. Wang, B. Chen, J. Wang, M. M. Begovic, and C. Chen, "Coordinated energy management of networked microgrids in distribution systems," *IEEE Trans. Smart Grid*, vol. 6, no. 1, pp. 45–53, Jan. 2015.

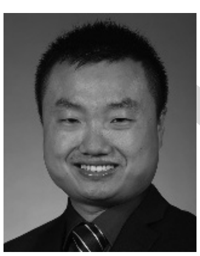
[103] Z. Wang, B. Chen, J. Wang, and C. Chen, "Networked microgrids for self-healing power systems," *IEEE Trans. Smart Grid*, vol. 7, no. 1, pp. 310–319, Jan. 2016.

[104] Y. Luo, S. Srivastava, M. Andrus, and D. Cartes, "Application of distubance metrics for reducing impacts of energy storage charging in an MVDC based IPS," in *Proc. IEEE Electr. Ship Technol. Symp. (ESTS)*, Jun. 2013, pp. 1–5.

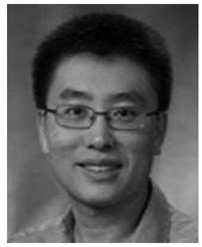
[105] Y. Luo *et al.*, "Application of generalized predictive control for charging super capacitors in microgrid power systems under input constraints," in *Proc. IEEE Int. Conf. Cyber Technol. Autom. Control Intell. Syst. (CYBER)*, Jun. 2015, pp. 1–6.

[106] S. Ma, B. Chen, and Z. Wang, "Resilience enhancement strategy for distribution systems under extreme weather events," *IEEE Trans. Smart Grid*, vol. 9, no. 2, pp. 1442–1451, Mar. 2018.

Kaveh Dehghanpour (S'14–M'17) received the B.Sc. and M.S. degrees in electrical and computer engineering from the University of Tehran in 2011 and 2013, respectively, and the Ph.D. degree in electrical engineering from Montana State University in 2017. He is currently a Post-Doctoral Research Associate with Iowa State University. His research interests include application of machine learning and data-driven techniques in power system monitoring and control.



Zhaoyu Wang (S'13–M'15) received the B.S. and M.S. degrees in electrical engineering from Shanghai Jiao Tong University in 2009 and 2012, respectively, and the M.S. and Ph.D. degrees in electrical and computer engineering from the Georgia Institute of Technology in 2012 and 2015, respectively. He is a Harpole-Pentair Assistant Professor with Iowa State University. He was a Research Aid with Argonne National Laboratory in 2013 and an Electrical Engineer Intern with Corning Inc. in 2014. His research interests include power distribution systems, microgrids, renewable integration, power system resilience, and power system modeling. He is the Principal Investigator for a multitude of projects focused on the above topics and funded by the National Science Foundation, the Department of Energy, National Laboratories, PSERC, and Iowa Energy Center. He was a recipient of the IEEE PES General Meeting Best Paper Award in 2017 and the IEEE Industrial Application Society Prize Paper Award in 2016. He is the Secretary of IEEE Power and Energy Society Award Subcommittee. He is an Editor of the IEEE TRANSACTIONS ON SMART GRID and IEEE PES LETTERS.



Jianhui Wang (M'07–SM'12) received the Ph.D. degree in electrical engineering from the Illinois Institute of Technology, Chicago, IL, USA, in 2007. He has held visiting positions in Europe, Australia, and Hong Kong including a VELUX Visiting Professorship with the Technical University of Denmark. He had an eleven-year stint with Argonne National Laboratory with the last appointment as a Section Lead Advanced Grid Modeling. He is currently an Associate Professor with the Department of Electrical Engineering, Southern Methodist University, Dallas, TX, USA. He was a recipient of the IEEE Power and Energy Society (PES) Power System Operation Committee Prize Paper Award in 2015. He is the Editor-in-Chief of the IEEE TRANSACTIONS ON SMART GRID and an IEEE PES Distinguished Lecturer. He is the Secretary of the IEEE PES Power System Operations, Planning, and Economics Committee.

Yuxuan Yuan (S'18) received the B.S. degree in electrical and computer engineering from Iowa State University, Ames, IA, in 2017, where he is currently pursuing the Ph.D. degree. His research interests include distribution system state estimation, synthetic networks, data analytics, and machine learning.

Fankun Bu (S'18) received the B.S. and M.S. degrees from North China Electric Power University, Baoding, China, in 2008 and 2013, respectively. He is currently pursuing the Ph.D. degree with the Department of Electrical and Computer Engineering, Iowa State University, Ames, IA, USA. From 2008 to 2010, he was a Commissioning Engineer with NARI Technology Company, Ltd., Nanjing, China. From 2013 to 2017, he was an Electrical Engineer for State Grid Corporation of China, Nanjing. His research interests include load modeling, load forecasting, distribution system estimation, machine learning, and power system relaying.