State Estimation Techniques for Electric Power Distribution Systems

Barry Hayes and Milan Prodanović
IMDEA Energy Institute, Madrid, Spain.
E-mail: barry.hayes@imdea.org

Abstract—This paper provides a survey of techniques for state estimation in electric power distribution systems. While state estimation has been applied in the monitoring and control of electricity transmission systems for several decades, it has not been widely implemented in distribution grids to date. However, with the recent drive towards more actively-managed, intelligent power distribution networks (“smart grids”) and the improvements in monitoring and communications infrastructure, Distribution System State Estimation (DSSE) has been receiving significant research interest. DSSE presents a number of unique challenges due to the characteristics of distribution grids, and many of the well-established methods used in transmission systems cannot be applied directly. This paper provides a detailed survey of the available methods for DSSE, reviewing around 70 papers from the major journals. In addition, it discusses the potential for applying Advanced Metering Infrastructure (AMI) data and computational intelligence methods in DSSE.

Keywords—state estimation; distribution networks; distributed energy resources; smart grids

I. INTRODUCTION

Since the initial development of the concept in the early 1970’s [1], power system State Estimation (SE) has become a critical part of the operation and management of transmission systems worldwide. Until recently, the application of SE at the distribution level, i.e. Distribution System State Estimation (DSSE), has not been of significant interest. This is largely because distribution networks have traditionally been designed and operated as passive systems, where power flows are unidirectional and relatively easy to predict and manage. However, distribution networks are seeing increasing penetrations of distributed energy resources, such as small to medium-sized Distributed Generation (DG), demand-responsive loads, electric vehicles and devices with storage capability. This has led to a requirement for improved observability in distribution systems, and the need for Distribution System Operators (DSOs) to take a more active role in monitoring and controlling the operation of the networks. DSSE has a crucial importance in this context.

Since distribution networks have different characteristics to transmission networks (e.g. radial construction, high \( R/X \) ratios, phase imbalances, and a much lower quantity and quality of available measurement data), many of the methods and approaches developed for “conventional” transmission-level SE cannot be applied directly to DSSE. Hence, a number of SEs specifically designed for application at the distribution level have been proposed in the literature in recent years. However, despite the growing importance of DSSE, the authors were unable to find a relevant survey paper in the literature, summarising the current state of the art, and discussing research trends and future directions in the area of DSSE (one conference paper was found [2], but the literature survey in this paper is not comprehensive, and focuses mainly on Chinese-language publications). While there have been several survey papers and books with literature reviews in the general field of power systems SE [3]–[11], these deal primarily with techniques and methods applied to transmission systems, and there are none which focus specifically on the developments and applications of DSSE. This paper aims to fill this gap by providing a survey of the most important techniques and algorithms currently available for DSSE.

This paper will also discuss the application of Advanced Metering Infrastructure (AMI) data, such as smart meter measurements, as inputs to the DSSE algorithms. Additionally, the use of novel computational intelligence methods and machine learning approaches and their potential benefits in this context of DSSE will be explored. The paper is structured as follows: Section II describes the main techniques and applications of DSSE, Section III outlines the current state of the art, and highlights some of the most advanced methods currently available. Section IV discusses the use of AMI data and computational intelligence methods in DSSE. Finally, conclusions are drawn in Section V.

II. DSSE TECHNIQUES AND APPLICATIONS

A. Conventional Power System State Estimation

SE is used to improve system observability, check for and detect errors in both system measurements and network parameters, and to mitigate against measurement and communication system noise. Detailed summaries of the main techniques and applications of conventional power systems SE can be found in [3]–[5]. Fig. 1 shows a graphical overview of the main processes and information flows. First, a topology processor verifies that the network parameters (e.g. line and switch statuses) provided to the estimator are correct, ensuring that the network model is accurate and up
to date [12]–[15]. Next, observability analysis is establishes that sufficient measurement data is available for the SE. The observability can be quickly determined by examining the null space of the Jacobian matrix [16]. If the network, or parts of it, are not observable, estimated values of network inputs (often referred to as pseudo-measurements) need to be provided. SE uses the available measurement data to find a unique solution for the system state. Finally, bad data processing is used to identify and remove data affected by gross errors and noise, e.g. due to measurement or communication system failures [17]–[21].

The network state is expressed as the vector $x$, containing the voltages and power angles at each node in the system. To estimate $x$, the set of measurements from the network, $z$, is applied. The values in $z$ can comprise of measurements of power/current injections or voltage magnitudes at system buses, measurements of active and reactive power flows in system branches, pseudo-measurements (i.e. estimates) of network quantities, or any combination of the above. This forms a set of over-determined, non-linear equations:

$$z = h(x) - e$$  \hspace{1cm} (1)

where $h(x)$ are the power flow functions corresponding to each measurement in $z$, and $e$ is the vector of measurement errors. The most commonly-used approach to minimise the objective function $J(x)$ is the “conventional” Weighted Least Squares (WLS) method:

$$\min_x J(x) = W(z - h(x))^2$$  \hspace{1cm} (2)

$$\Rightarrow \min_x (z - h(x))^T \tilde{W} (z - h(x))$$  \hspace{1cm} (3)

where: $\tilde{W}$ is the measurement weight matrix. Each of the weights in $\tilde{W}$ are set according to the inverse of the variance of the corresponding metered system measurement. This allows the weights in $\tilde{W}$ to be adjusted so that the estimator gives more weight to input data points which are known to have greater accuracy. The minimisation in (1) is solved iteratively as follows:

$$\Delta z_n = z - h(x_n)$$  \hspace{1cm} (4)

$$\Delta x_n = (H^T \tilde{W} H)^{-1} H^T \tilde{W} \Delta z_n$$  \hspace{1cm} (5)

$$\Delta x_{n+1} = x_n + \Delta x_n$$  \hspace{1cm} (6)

where the Jacobian matrix, $H = \delta h(x)/\delta x$, and $n$ is total number of SE iterations.

The presence of bad data in the system measurement data set can be detected by applying statistical tests to the objective function $J(\hat{x})$, and to the normalised residual vector given by $r = z - h(\hat{x})$, which is normalised by $r_n = \rho_{jj}^{-1} r$, where $\rho_{jj}$ is the diagonal of the covariance matrix:

$$C_r = W^{-1} - H(\hat{x}) G^{-1} H^T(\hat{x})$$  \hspace{1cm} (7)

Bad data is detected and identified (provided there is sufficient redundancy in the measurement data set) through statistical testing, where the $J(\hat{x})$ Performance Index and Largest Normalised Residual Tests [7] are most commonly used in conventional SE. Further studies into bad data processing and removal are given in [17]–[19], [21], [22]. There are also alternative SEs discussed in the literature, such as the Weighted Least Average Value estimator and the Schweppe-Huber generalised M-estimator [3]–[8]. These replace $J(x)$ in (2) with a different objective function, but otherwise the overall approach to SE remains the same. One of the problems encountered in general SE is the computational complexity of solving (2). In order to reduce the computational burden, some authors have proposed a fast decoupled SE [22], or carrying out a Direct Current (DC)-only SE by neglecting all branch resistances and shunt elements [5]. However, the methods and assumptions used in transmission-level SE described above are often not valid when considering distribution systems, and many of well-established techniques used in “conventional” SE cannot be applied directly [23]. This has motivated research into DSSE, i.e. state estimators designed specifically for use in distribution networks.

B. Distribution System State Estimation

Initial research into DSSE began in the 1990’s [24]–[27]. DSSE presents a number of new challenges, since the characteristics of distribution networks differ fundamentally from transmission networks in the following ways:
Construction: Most distribution networks have a radial construction (whereas transmission systems are more meshed), often with high R/X ratios.

Redundancy: For technical and economic reasons the number of measurement points in distribution networks is much lower than in transmission networks. Systems are under-determined, rather than over-determined.

Measurement types: Most of the available input data at the distribution level are measurements (or pseudo-measurements) of power or current injections. Direct measurements of voltages and power flows are rare.

Scale and complexity: Distribution systems are diverse (e.g. networks in rural areas are very different from those in urban areas) and have very large numbers of components. This means that the methods developed for DSSE need to be scalable, have a relatively low computational burden, and be applicable across a range of different network types.

Phase imbalances: Conventional SE techniques assume that the network is a balanced system. However, distribution systems, can have significant phase imbalances, requiring the use of full three-phase system models.

Some of the techniques developed in order to overcome these issues are discussed below.

1) Adapting Conventional WLS Techniques to DSSE:
Many of the earlier research papers on DSSE focussed on adapting conventional WLS techniques to distribution networks [24], [26]–[28]. However, there are significant limitations to adapting approaches from transmission-level SE to DSSE, particularly in dealing with noisy input data and “robustness”, i.e. the ability of the estimator to reach a unique solution of the minimisation described in (2) and (3) in the presence of gross input errors [29]. In addition, due to the construction of distribution systems (radial feeders and high R/X ratios), the fast decoupled methods and DC approximations often applied in conventional SE simply do not work when applied to DSSE [30].

2) Load Estimation for DSSE: In DSSE, the number of telemetered devices that can provide system measurements is often very limited, and not sufficient to allow observability of the entire network, or bad data identification. In many cases, DSSE relies on pseudo-measurements of the demands at each load point in the network, based on historical data or load forecasts, which have significantly lower accuracy than actual measurements. Load estimation techniques and their application to DSSE are discussed in [31]–[33].

3) DSSE in Unbalanced Networks: In [25], a branch-current-based SE methodology was developed, in which network branch currents, rather than node voltages are used to represent the system state x. This has the advantage that the Jacobian matrix H can be decoupled on a per-phase basis, allowing conventional SE methods to be applied to distribution systems which are unbalanced, or have single-phase or two-phase lateral feeders. A number of papers have built on this approach in order to develop robust and accurate three-phase DSSE techniques [34], [35].

III. State of the Art

This section briefly describes some of the most advanced techniques and applications of DSSE in the literature.

A. Forecast-Aided State Estimation

The SE methods discussed in Section II are static in nature, in that the estimation of the system state is only dependent on the current “snapshot” of input measurements, and not on previous input data values. There are also SE techniques which are designed to recursively update the state estimate in order to track changes during normal operation. This approach has been called “dynamic” SE [36], however the term “Forecast-Aided State Estimation” (FASE) is preferred by many authors to avoid confusion since the word “dynamics” in power systems is strongly associated with transient stability studies. An excellent summary of the main concepts in FASE is given in [9]. Most FASE approaches model the system using the state-space form introduced in [37] and the Extended Kalman Filter (EKF) [38]:

\[
x_{k+1} = F_k x_k + g_k + w_k
\]

\[
z_k = h_k(x_k) + v_k
\]

where \( F_k \) is the state transition matrix, \( g_k \) is a vector representing the trend behaviour of the state trajectory, and \( w_k \) and \( v_k \) representing the process and observation noises which are assumed to correspond to white Gaussian noise with zero mean. At each time step \( k \), the Jacobian matrix \( H_k \) is evaluated with the current predicted states, and used in the EKF equations. A short-term forecast (e.g. several seconds/minutes ahead) of the state variables is made, and each time a new set of measurements becomes available, an “innovation analysis” can be used to determine if the new measurements are significantly different from the predicted values [39]–[41]. This analysis filters the new input data using the EKF equations, allowing detection of anomalies such as bad input data, and network configuration or parameter errors [37]. While most of the FASE techniques and applications proposed to date are focused at the transmission level, FASE approaches are also interesting for DSSE, particularly if high-resolution data is available from synchronised metering devices, such as Phasor Measurement Units (PMUs) [42], [43].

B. Multi-area and Hierarchical DSSE Techniques

Since distribution systems are typically very large and dense, (e.g. comprising of many thousands of individual nodes) one of the most challenging aspects of implementing DSSE is the computational complexity. In conventional
SE methods, all measurements are typically processed in one centralised SE. However, a better solution for large distribution systems may be to split the networks into a number of smaller sub-networks, or “measurement areas”, in which the SE is solved locally [44]–[46]. The multi-area SE can be expressed as [47]:

\[ z_m = h_m(x_m), \quad m = 1, \ldots, M, \]  

(10)

where \( x_m = [x_{im}, x_{om}] \) is the local state vector of measurement area \( m \), which contains the internal state variables \( x_{im} \), border state variables \( x_{om} \), for the total number of measurement areas, \( M \). In the multi-area approach, the SE is solved locally within each measurement area, and data is exchanged between areas only where they border each other. In [46], a multi-area method has been developed for DSSE, which meets the performance requirements for real-time applications in very large networks. Traditionally, transmission-level and distribution-level SE have been developed separately. However, with the increased requirements for communication and interaction between transmission and distribution network management systems, and several authors have investigated the development of multi-level, or hierarchical SEs, designed to integrate transmission SE and DSSE [48], [49].

C. Advanced Distribution Management Systems

Due to the need for better situational awareness and more active system support, there has been much interest in adapting operational techniques, previously only used at the transmission level, to distribution systems [50], [51]. A number of studies have investigated “Advanced Distribution Management Systems”, systems designed to optimise energy management in distribution networks, where DSSE is an important part of the methodology [51]–[55].

IV. Future Research Areas for DSSE

Application of Advanced Metering Infrastructure Data

The widespread introduction of smart meters means that an unprecedented amount of detailed historical data on user loads is becoming available. This data can be used to better understand and model the behaviour of distribution network loads, allowing to improve load estimation techniques, and ultimately, DSSE accuracy. Some initial studies have been made into the incorporation of smart meter data into DSSE in [56], [57]. The use of smart meter data to estimate flows, voltages, and losses in the low voltage distribution network is demonstrated in [58]. In [59], the DSSE is carried out using “compressed” measurements from smart meters and DGs. While there are significant opportunities for further work in this area, the scope for using smart meter data as an input to DSSE for real-time applications is limited for a number of reasons, including: data privacy issues, low smart meter data rates (e.g. 15-minute or hourly intervals), low data reliability, and the lack of smart meter data time-synchronisation. However, there are potentially significant benefits in terms of applying smart meter data for providing insight into the behaviour of end-user loads, and in providing a certain degree of visibility in parts of the distribution system which were previously unobservable. A key feature of future DSSEs will be the flexibility to incorporate multiple types of input data, e.g. estimators capable of integrating both analogue and digital inputs (e.g. power/voltage measurements and switch/breaker statuses [60]) and also measurement data from a range of diverse sources, e.g. SCADA, PMU, smart metering, DG units.

Computational Intelligence Methods in DSSE

Computational intelligence methods have been proposed for a range of power systems and smart grids applications. Machine learning techniques are particularly attractive for DSSE. Artificial neural networks have already been used extensively in the power systems research literature for load estimation and forecasting [61]. In [62], a neural network approach is used to create pseudo-measurements of load point power injections for use in the DSSE algorithm. In [63], a machine learning approach is used to develop load estimates for DSSE. The advantage of this approach is that load models are designed to recursively re-train themselves as more measurement data becomes available, leading to improvement of the performance of the DSSE over time. Further research is required on the implementation of “closed-loop” DSSE methods. While almost all of the SE methods proposed in the literature have an open-loop information flow, a closed-loop DSSE allows the predictive database used to estimate loads (and DG outputs) to be continuously updated and improved based on feedback from the SE [64], [65]. An “autonomous” approach to DSSE is presented in [66], in which the DSSE is designed to automatically detect new connections (e.g. from DG units) to the distribution network and update the system model accordingly. The further application of machine learning techniques will be crucial in allowing DSSE to be implemented on a large scale. Without the automation of the majority of control and network model management functions, the implementation of DSSE across all of the distribution networks in a DSO region would lead to an unreasonable increase in operator workload.

It is likely that future DSSEs will make extensive use of event-triggered approaches, particularly with regard to carrying out advanced DSSE functions requiring significant computational effort, such as model identification [67], [68]. In the event-triggered approach, the relevant DSSE is function carried out locally (i.e. in the network area of interest), only when the measurements received indicate that there is a potential issue, e.g. a suspicion of a network topology error. This approach is particularly important in the context of DSSE model identification, where the aim is to identify
the correct network model from a range of possibilities. Model identification is particularly relevant in distribution systems, where the switch statuses throughout the network are typically unmonitored. The authors in [67] propose the development of a model bank, in which all of the network configurations that are critical to the DSSE are stored. Based on the results of the DSSE applied to each network model, the model which has the highest probability is selected as the true network configuration using a Recursive Bayesian approach. The a posteriori probability of the model \( \eta_i \) being correct is given by:

\[
p(\eta_i \mid \epsilon^k) = \frac{p(\epsilon_i^k \mid \eta_i) p(\eta_i \mid \epsilon^{k-1})}{\sum_{j=1}^{N} p(\epsilon_j^k \mid \eta_j) p(\eta_j \mid \epsilon^{k-1})}
\]

where \( N \) is the total number of models, \( \epsilon^k = \{\epsilon_1^k, \epsilon_2^k, ..., \epsilon_N^k\} \) is the set of error vectors of each model, \( p(\eta_i \mid \epsilon^{k-1}) \) is the prior probability, and \( p(\epsilon_i^k \mid \eta_i) \) is the conditional probability of the error in the \( i \)th model, given model \( \eta_i \) is correct. The a priori probabilities, \( p(\eta_i) \), can be calculated using a Gaussian distribution as in [67]. The algorithm works by assigning an equal initial probability to each model, and carrying out a Monte-Carlo simulation, updating (11) at each iteration. It is demonstrated in [67], [68], that this approach can converge on the correct model quickly and reliably.

V. CONCLUSIONS

This paper provides a detailed survey of techniques and applications related to state estimation in electric power distribution networks. It also discusses the potential for applying AMI data in order to improve estimation performance, and the application of novel computational intelligence methods to DSSE. Despite the high level of research interest in the last two decades, DSSE has not been widely implemented in distribution networks to date. This has mainly been due to a lack of investment in information and communications technology infrastructure at the distribution level. However, it is becoming more and more evident that DSSE will play an important role in the operation of future distribution networks, as these systems become much more active and complex due to increasing penetrations of variable, highly-dispersed resources, e.g. small-scale renewable energy, demand-responsive loads, electric vehicles, electricity storage, and microgrids. Further work will examine the adaptable and intelligent DSSE approaches discussed in this paper in detail, in order to recommend the best approaches for future large-scale implementation.

REFERENCES
