ELECTRIC POWER
RESEARCH INSTITUTE

STATE ESTIMATION: APPLICATION AND ALGORITHMS FOR THE DISTRIBUTION SYSTEM

November 2018

TABLE OF CONTENTS

INTRODUCTION

To a power engineer, the term *state estimation* quickly brings to mind a statistical application, central to the operation of the bulk power energy management system, that uses a redundant set of measurements and a bus-oriented network model to compute a statistical estimate of the system operating state. This concept was first formulated as a mathematical solution for transmission systems by Fred Schweppe and his team in 1969, and is now a mission-critical component of transmission system operations. State estimation importantly combines the present system topology and real-time measurements of different qualities to provide the best statistical estimate of the state of the system. In this context, what is meant by *state* are the key measurements (voltage magnitude and angles, current magnitude and angles, and power) at each bus. These estimates provide accurate situational awareness, and also support key operational applications (e.g., transmission contingency analysis, optimal network configuration).

Over the past number of years, as distribution management systems (DMSs) have evolved and become more widely deployed, the concept of *distribution system state estimation* (DSSE) has become deeply embedded in the core DMS architecture, as shown in Figure 1.

In fact, the concept of distribution system state estimation (DSSE) has become so widely discussed that engineers who are not working directly on this solution are likely to give it no more than a brief thought and assume that it consists of simply taking the statistical methods that work so well in realm of transmission systems and applying them to distribution systems. But is that true? The purpose of this paper is first to define what should be understood when someone hears or reads the term *DSSE*. Then, the state of the industry with respect to DSSE will be summarized, from the perspective of research, application requirements, and the need for future research.

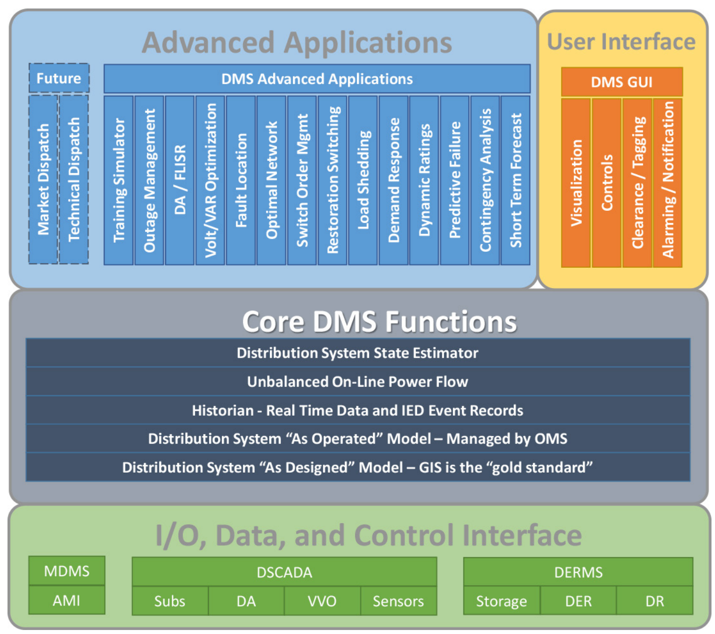

Figure 1 – Distribution management system application architecture

DSSE VS. TSSE

Again, without being deeply connected to the solution, many engineers would assume that the weighted least squares statistical methodologies long deployed on transmission energy management systems would work just as well on distribution management systems. Why would the algorithm care about the voltage level? That is a fair question. However, there are in fact a number of important differences between the distribution system and the transmission system that make the answer not so easy (see Table 1).

Table 1 – Key differences between transmission and distribution

Feature	Transmission	Distribution
Topology	Highly networked	Mostly radial
Line Impedances	High X/R ratio	Low X/R ratio
Lines	Consistent conductor sizes	Lines change size often
Observability (Sensors)	At every bus	Limited but growing
Measurement Types	P, Q, voltage, current	Current, voltage
Observability (Switch Status)	At every device	Limited monitoring of switch status
Observability (Volt/VAR Optimization Status)	Tap position and capacitor status is generally monitored	Limited monitoring of regulator and capacitor positions
Phase Load/Voltage Balance	Balanced	Unbalanced
Topology Balance	All three phases follow the same path	Phases may follow different paths; regulators tapped per phase
Number of Nodes/ Buses	Low	High
SCADA Polling	10-second analogs typical	Often uses exception reporting
Measurement Synchronicity	Entire system measured within 10-second cycles	Measurements come at a variety of cycles and latencies

Each of these differences creates a challenge to simply applying the transmission system state estimation (TSSE) algorithm directly to distribution.

DSSE VS. POWER FLOW

Distribution feeders are modeled using planning tools that perform various types of studies based on power flow, including voltage drop, protection coordination, and distributed energy resources (DER) impact analyses. The ability to calculate the voltage at any node or the current through any branch is what separates power flow models from the topological models contained in a geographic information system (GIS) or elsewhere. Typically, the physical characteristics of each device (conductor, transformer, switch, fuse, etc.) in a distribution feeder are modeled with high enough accuracy that planning engineers can use these models to answer "what if" type questions without needing to make field measurements. While the topology and physical characteristics of the distribution feeder are very accurate, the individual loads are often approximations at best. Utilities have a clear understanding of their assets and infrastructure, but the visibility of customer usage remains cloudy.

The traditional approach to customer load allocation in power flow models is based on substation supervisory control and data acquisition (SCADA) measurements and distribution transformer kVA rating. A time series measurement at the feeder head determines the peak load and load profile, and individual customers are allocated a portion of that load based on the kVA rating of the distribution transformer serving them. All customers are assumed to have a load profile that matches the feeder head. This method is accurate at reproducing the aggregate load and profile at and near the feeder head, but there is an error factor that tends to increase with distance away from the substation.

Several methods exist that can increase the accuracy of traditional load allocation. For starters, monthly energy consumption from customer billing cycles can be used to refine the load allocation to individual customers. Load profiles for different classes of customers (residential, commercial, and industrial) can also be used to reduce the error in power flow load allocation, especially in feeders with a highly mixed customer base. The most accurate form of load allocation uses individual customer load profiles derived from full deployment of advanced metering infrastructure (AMI); however, few utilities have achieved this level of deployment. Even with 100% AMI installed, the data requirements for this type of analysis can be overwhelming and possibly prohibitive for planning tools to solve in a reasonable amount of time.

Without accurate measurements along the feeder or near the feeder edge there is no way to know that errors exist in the power flow models. The traditional kVA allocation method generally results in satisfactory simulation and operation of distribution feeders, with utilities being able to maintain appropriate customer voltage levels throughout their service territories. However, modern power devices such as electronic reclosers, remotely operated switches, DER, and AMI have the ability to measure voltage and current at various locations throughout the distribution feeder, and these measurements can be used to determine errors between power flow model results and real measurements. Planning engineers and operators are now faced with differences between the power flow results and the real measurements coming in from intelligent electronic devices (IEDs) around the feeder. So what can utilities do to rectify these discrepancies?

As mentioned, transmission system operators have leveraged state estimation since the 1970s to provide the best estimate of real-time system conditions. System states can be defined as either the vector of all node voltage magnitude and angles, or of all branch currents and angles. It has been widely suggested that a similar approach can be used for the distribution system, hence the term *distribution system state estimation (DSSE)*. However, even though the formulation of traditional state estimation can be translated to the distribution system, the application of DSSE has yet to be fully investigated.

Conceptually, DSSE uses a set of measurements to provide the best estimate of the actual voltages or currents across a distribution feeder. In a sense, power flow models also use a set of measurements, in the form of load allocations, to determine the voltages and currents across the feeder. What separates DSSE from power flow models is the use of redundant measurements to detect bad data or measurements. In DSSE, the load shapes and peak loads generated by load allocation methods are considered pseudo-measurements, or approximations of the real conditions at the grid edge. Real measurements are those measurements that are streamed from IEDs, AMI, and so on. In order to identify bad data from real measurements or pseudo-measurements, there must be more measurements than there are system states.

At a high level, the proposal that there must be more measurements than states (to enable bad data identification) could be overwhelming to a planning engineer, given that traditional load allocation uses only one measurement from the substation to allocate loads. However, the allocation factors at each customer, combined with substation measurements, are considered pseudo-measurements. Therefore, incremental deployment of AMI or IEDs may provide sufficient redundant measurements to effectively address the state estimation problems, including the issue of bad data detection.

USE CASES FOR DSSE

Given the described challenges to the application of DSSE, a logical follow-up question would be “What are the application use cases for DSSE in distribution system operations?” The high-level consensus in the industry today is that DSSE will be required to manage the complex and dynamic distribution system of the (very) near future, with DER penetration levels growing to very significant levels. But is that really true? Let’s consider a comprehensive list of DSM applications one at a time (see Table 2).

The analysis presented in this paper supports the widely held perspective that some solution/algorithm is required to estimate the near real-time state of the distribution system. There can be no doubt that this is true. Even though the needed solution/algorithm for the distribution context may be completely different from the established transmission system state estimator, it seems reasonable to refer to it as distribution system state estimation. The question, then, is which DSSE solution/algorithm is best suited to provide the answers. The following section explores recent research into algorithms being studied in academia and deployed by vendors.

ALGORITHMS BEING STUDIED

State estimation is more than just the state estimation algorithm itself. As a broader process, state estimation is made up of a subset of smaller functions that can include network topology verification, observability analysis, application of the state estimation algorithm, measurement error detection, and network parameter error detection. For distribution state estimation, *network topology verification* refers to ensuring that the feeder model accurately represents the lines (e.g., three phase, single phase), connections (e.g., switch states), and other devices (e.g., capacitor bank status, voltage regulator taps). The observability analysis helps to ensure that there are enough measurements to obtain a state estimation solution, and the state estimation solution algorithm provides the state estimation solution. There are various solution algorithms that can be used (as discussed below). The state estimation solution algorithm can be just one of the many algorithms listed, or a hybrid/combination of multiple algorithms. During the solution process there is usually some sort of error detection. Unfortunately, it is typically not possible to detect both measurement and network errors at the same time. This section describes various state estimation solution algorithms and some of the terminology and names used to describe solution techniques.

WEIGHTED LEAST SQUARES

Weighted least squares is one of the most popular state estimation techniques. It assumes that the measurement errors (metering accuracy, historical data, etc.) are Gaussian in shape with a zero mean and a known variance. There are four main measurement types: line power flows, bus power injections, bus voltage magnitudes and angles, and line current flow magnitudes. The measurements themselves will be discussed in more detail later. Weighted least squares is an iterative technique, and it requires the recalculation of the Jacobian and gain matrix with each iteration, sometimes resulting in long

Table 2 – Applications of DSSE

Application Use Case	Discussion
Situational Awareness	As more and more DER are deployed on the distribution system, it will no longer be possible to estimate the operational state of a feeder based purely on a substation breaker measurement, or a few distribution automation switch measurements. With two-way power flow and high levels of intermittency of DER supply, the DMS must be able to provide a reasonably accurate near real-time picture of the distribution system.
Topology Error Detection	Due to planned expansion projects, temporary states, and restoration activities, the distribution system is in a state of constant flux. In most implementations, the as-designed topology is maintained in the GIS and then synched with the DMS on a regular basis. The Outage Management System maintains the as-operated topology based on this. With this continuous high level of change, it is common for topology errors to appear. One of the potential applications of DSSE is the detection of these errors.
Sensor Error Detection	The original reason for state estimation in the transmission system was the prevalence of measurement errors. The sensing technologies in use during the 1950s (CTs, PTs, transducers, and analog to digital converters) were quick to become uncalibrated. As a result, power flow algorithms converged to an incorrect solution, or may not have converged at all. Present-day sensing technologies deployed on the distribution system (CTs, PTs, and IED inputs) are much more accurate and consistent. The errors are perhaps different now, including inconsistent timing of data transmissions (including reporting on exception), and more likely errors in modeling to reflect the true location (and phase) of the sensor.
DA - Automated Restoration	Second- or third-generation distribution automation (DA) systems could simply rely on the measurements at the DA device (switch, sectionalizer, recloser). These automated restoration systems followed rules based on the current that was monitored immediately prior to initiating operation of the protective device. Next-generation DA systems must take into account the loss of DER supply, the location and type of DER, and other factors to model the impacts on power flow and voltage prior to initiating restoration switching.
Operator-Directed Restoration	Similarly, distribution control center (DCC) system operators must be able to model and analyze the impacts of restoration switching steps. In the past, with predictable radial feeders, this planning could be done using simple models. In the future, operators may rely on DMS modeling, supported by DSSE, to accurately predict the impacts of restoration switching steps.
Volt/VAR Optimization (VVO)	As with distribution automation, the management of voltage and VARs on the distribution system has become much more difficult due to the influx of DER. Based on a significant body of industry research by EPRI and others, it is clear that neither local controls nor centralized logic will alone be able to manage DER-induced overvoltages, while maintaining desired power factors. What will be needed is a hybrid control architecture based on a foundation of centralized modeling of the distribution system to "see" beyond the sensors.
Planned Switching	As with restoration switching, planned switching must incorporate modeling and analysis to determine the impacts of restoration switching steps. This may be more complex, because planned switching analysis is often performed days or weeks in advance. Again, in the past, with predictable radial feeders, these switching analyses could be done using simple models. In the future, operators may rely on DMS modeling, supported by DSSE, to accurately predict the impacts of restoration switching steps.
Short-Term Load Forecasting	Operation of the distribution system is completely dependent on the loads and how they change. The science of developing accurate load models and load forecasts is evolving rapidly. New measurement system approaches including AMI are providing robust data to support this application. However, it is yet unclear whether DSSE can provide an additional support.
Short-Term DER Forecasting	A new requirement for distribution system operations is the ability to develop short-term forecasts of DER production. Many of the installations (specifically low-voltage connected DER in countries/states that do not require separate meters) are essentially unmonitored. The estimates and short-term forecasts of DER production will become pseudo-measurements that serve as inputs to the DSSE algorithm. It is unclear to what extent DSSE can support these forecasts, or whether the forecasts will instead dictate how accurately DSSE can be accomplished.

solution times. However, solution times can be decreased by utilizing decoupled or dc power flow techniques where applicable. There are also other various techniques proposed in academic literature to help improve the accuracy, decrease the complexity, and reduce the solution time of the weighted least squares technique for state estimation.

The Jacobian matrix is the same as for a standard power flow, and the gain matrix is formed from both the Jacobian matrix and the error covariance matrix. The error covariance matrix is just a diagonal matrix of variance of the measurement errors when the measurements are considered independent.

LOAD ADJUSTMENT

Another state estimation technique, mainly used for distribution system state estimation, is *load adjustment* (sometimes called *load estimation*). Unlike the bulk system, the distribution system has few redundant measurements, let alone enough measurements for the number of nodes in a circuit. Therefore, the load adjustment state estimation adjusts the models' loads based on the available measurement data (real or pseudo-measurements). Load adjustment state estimation assumes that the measurements are 100% accurate, and only the model needs to be adjusted. Like the weighted least squares methodology, load adjustment state estimation techniques typically involve an iterative process using Gauss-Seidel load flow techniques instead of Newton-Raphson. Various other methodologies have been proposed as well, including more simulation-based approaches like particle swarm optimization. The issue with simulation-based techniques is that they do not guarantee the best/optimal solution. The accuracy required can greatly influence the required solution time.

DYNAMIC STATE ESTIMATION

Dynamic state estimation is the general name for any state estimation technique that utilizes measurements from different points or references in time. For example, a phasor measurement unit (PMU) may provide measurements every minute, whereas for AMI the interval is every 15 minutes. Even if all the measurements used the same time step, they would not come in at the same time. Therefore, it would be inappropriate to assume that the solution would converge utilizing the measurements alone. Dynamic state estimation takes the various measurements (at various points in time) and forecasts them to the same point in time in order for the estimation to converge. This is a recursive process, and several past measurements are required for this methodology. The forecasted measurements may be considered pseudo-measurements.

ROBUST STATE ESTIMATION

One of the purposes of state estimation is to detect and identify bad measurements. *Robust state estimation* is just a general name, like *dynamic state estimation*, for any state estimation technique that remains unaffected by major deviations in the limited number of measurements. The following are two key terms that are used to help describe robust state estimation:

- ***Breakdown point*** – the ratio of measurements that can be infinitely wrong while the state estimation remains bounded compared to the total number of measurements
- ***Leverage point/measurement*** – a measurement that lies outside the space of the rest of the measurements (an error at such a point may be difficult to detect)

Some examples of where leverage points may occur include:

- An injection or flow measurement at a bus that is incident to a large number of branches
- An injection or flow measurement at a bus that is incident to branches of very different impedances
- Using a large weight for a specific measurement, compared to others

These leverage points may appear to be bad measurements even when they may not actually be bad. A robust state estimation should be able to handle such measurements.

Techniques for robust state estimation include M-estimators (modified weighted least squares technique), which are an approach to estimating maximum likelihood, and least absolute value estimation, which can be formed as a linear program. Both techniques try to minimize the measurement error, subject to various constraints. Machine learning algorithms have also been cited in literature as a means to provide robust state estimation. However, they have one main caveat—they only work for the data they are trained on. That is, the estimator will possibly need to be retrained for any new network configuration, and should be trained for all network configurations.

DISTRIBUTED STATE ESTIMATION

Another type of state estimation is *distributed state estimation*. The idea is to split up the distribution network into multiple individual areas, allowing any state estimation technique to be applied to solve each area separately before reconciling them all into one estimation for the network. It could be possible to use different techniques depending on the area, the use case, and so on. The main advantage to this type of technique is speed. Reducing the problem variable size—by splitting the variables into multiple smaller-size variables sets—should improve the solution speed, helping improve the possibility of real-time calculations. However, this is currently not practical, because there are not enough measurements available for a single distribution system state estimation method. It would be impossible to solve more than one area with any sort of accuracy.

PROCESSED LOAD FLOW

The differences between a power flow solution and a DSSE have been discussed, but it may be possible to utilize a power flow if enough pre-processing of the models has been accomplished to ensure that the power flow solution will converge to an accurate estimate of the states. This method includes pre-processing of load allocation using feeder head and any additional sensors (DA switches, AMI, etc.), as well as pre-processed simulation of the regulator and capacitor statuses (implying knowledge of their control algorithms), and finally some methodology to automatically identify measurements that are not correct (finding them to be outside of reasonable values or inconsistent with other measurements, for example). EPRI has successfully utilized this approach for a limited set of studied feeders using OpenDSS, and this approach has also been observed in commercially available vendor solutions.

FEEDER/LOAD MODELS

Having an accurate feeder model is just as important as having accurate measurements for state estimation. Topological errors produce larger errors in state estimation, making them more easily identified. However, state estimation can only be used to detect one type of error at a time—either measurement errors or model errors, but not both. (There is current work being deployed that can do both at once, but it is for the bulk system.) While it is important to model the feeder as accurately as possible, the network model should not be overdetailed. The model should only be as detailed as it needs to be to represent the system accurately. Typical network parameters that should be considered are the following:

- Transformer models (typically, the losses can be underestimated)
- Network changes (not always updated)
- Ambient temperature
- Other network modifications (e.g., tap changes, as-operated switching)

If there are errors in the feeder model it can lead to a degradation of results, to measurements being incorrectly flagged as bad, to loss of the operator's confidence in the state estimation, or to other undesirable consequences.

Depending on the state estimation algorithm/methodology, accurate load models may be more or less critical. However, in distribution state estimation a highly accurate model may not be necessary, as most DSSE methodologies adjust the loads based on the available measurements. Reactive power is harder to estimate accurately due to load allocation algorithms and available feeder model data on capacitor banks, switching, and so on.

OBSERVABILITY

In state estimation the question of system observability often arises, but it is not clearly defined outside of the mathematical domain. In plain terms, observability is the ability to provide an estimate of the system state using the set of available measurements. The state estimation problem uses a set of equations to relate system measurements to an estimate of the system state. The problem is only solvable, or observable, if certain conditions for the set of equations and number of measurements are satisfied. The two sufficient conditions for observability are that the number of measurements be greater than or equal to the number of states, and that the rank of the state equation matrix be equal to the number of states. While these conditions may be clear from a mathematical perspective, their application to distribution systems is not.

Consider the following example, which demonstrates the observability problem as it applies to distribution systems, using a simple three-node distribution feeder as shown in Figure 2.

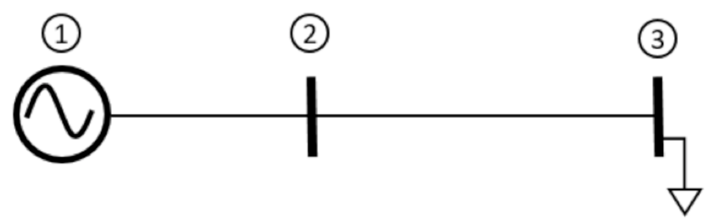

Figure 2 – A simple three-node distribution system

For the classical state estimation problem, the linearized system measurements are given by the following equation:

$$Z = Hx + \eta \qquad \text{Eq. 1}$$

Where:

Z is an m-vector of system measurements
x is an n-vector of system states
H is the linearized matrix representation of the system equations relating the system states to the measurements
η is an m-vector of measurement error

To meet the sufficient conditions for observability, the number of measurements must be greater than or equal to the number of states, and the rank of H must be equal to the number of states. Mathematically, the rank of a matrix is defined as the span of the column space of the matrix; for state estimation this condition implies that there is an equation for each state that relates to at least one measurement. Consider the case that the available measurements are the voltage magnitudes at each node and the real and reactive power between nodes 1 and 2, as shown in Figure 3.

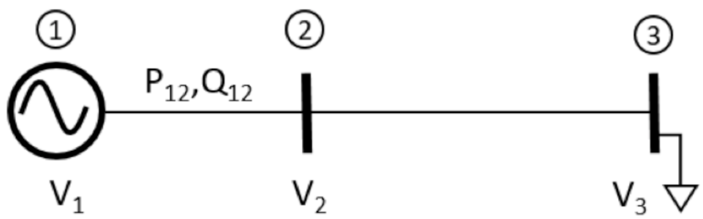

Figure 3 – The simplified three-node distribution feeder showing the available measurements of node voltages and the power flow between nodes 1 and 2

If the system state is the vector of all node magnitudes and angles the elements of Equation 1 take the following form:

$$Z = [P_{12} \quad Q_{12} \quad V_1 \quad V_2 \quad V_3]^T$$

$$x = [V_1 \quad \delta_2 \quad V_2 \quad \delta_3 \quad V_3]^T$$

$$H = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Even though the number of measurements is equal to the number of states, the sufficient conditions for observability are not met because the rank of H is not equal to the number of states. In the terms of the distribution system, this means that the measurements and state equations are not sufficient to define all of the states, namely the voltage angle at node 3. What can be done to rectify this problem and achieve full observability? There are two options.

The first option is to add another measurement to the system that results in full observability. Consider that the real power is measured between nodes 2 and 3, as shown in Figure 4.

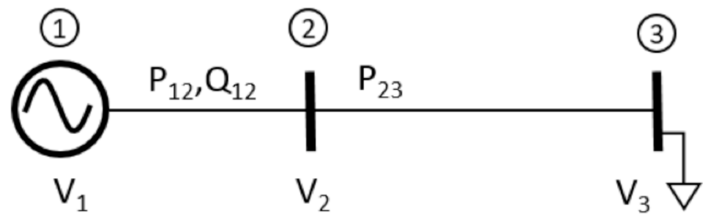

Figure 4 – The simplified three-node distribution feeder with an additional power measurement between nodes 2 and 3

With the additional measurements, the elements of Equation 1 become:

$$Z = [P_{12} \quad Q_{12} \quad P_{23} \quad V_1 \quad V_2 \quad V_3]^T$$

$$x = [V_1 \quad \delta_2 \quad V_2 \quad \delta_3 \quad V_3]^T$$

$$H = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

In this case observability is achieved because the number of measurements is greater than the number of states, and the rank of H is equal to the number of states. For each state there is an equation relating a state to a measurement. While observability has been achieved, there is also added complexity with this option in that the measurement must be collected and delivered to the state estimator. Adding measurements can be cost-prohibitive from the standpoint of the physical measurement devices required and because of the communication requirements. However, there exists another option for achieving observability.

The second option for solving the observability problem presented above is to perform a system reduction. If the state of node 2 is not a critical result of the state estimator, it is possible to reduce the distribution system model by removing the node, as shown in Figure 5.

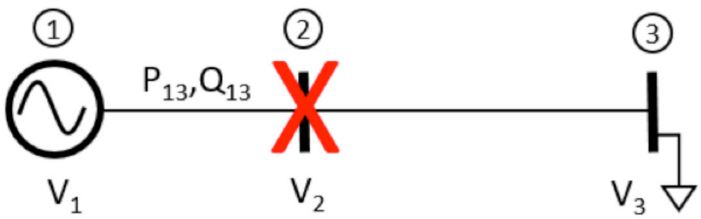

Figure 5 – The simple three-node distribution system can be reduced by eliminating non-critical nodes.

By removing node 2 the terms of Equation 1 become:

$$Z = [Q_{13} \quad Q_{13} \quad V_1 \quad V_3]^T$$

$$x = [V_1 \quad \delta_2 \quad V_2]^T$$

$$H = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

In this case observability is achieved because the number of measurements is greater than the number of states, and the rank of H is equal to the number of states. For each state there is an equation relating a state to a measurement. The system was reduced and the state at node 2 is no longer available, which may or may not be of concern to the distribution planning engineer. System reduction is an effective way of limiting the data and communication requirements of DSSE while maintaining the accuracy of the system states. This method does require a simplified system to be created and validated, which may require significant time. There are existing tools for distribution system reduction, but reduction alone may not solve all observability problems.

Observability analysis is one of the first steps to state estimation, and it determines whether an estimation of the system state can be obtained from the available measurements. Classifications of measurements for system observability include the following:

- ***Critical measurement*** – a measurement that when removed will result in an unobservable system
- ***Redundant measurement*** – a measurement that is not critical
- ***Critical pair*** – two redundant measurements that when removed simultaneously will result in an unobservable system
- ***Critical k-tuple*** – k redundant measurements that when removed simultaneously will result in an unobservable system

Observability and the associated required/critical measurements are a great concern for DSSE. The topology of the network can have a major effect on the number and location of critical measurements. As mentioned previously, the typical network configuration of a distribution system (radial) differs greatly from that of the bulk system (meshed), typically increasing the number of critical measurements required for full network observability. However, full network observability may not be necessary, depending on the use case of the state estimation.

DATA USED/REQUIRED

As mentioned earlier, there are various types and sources of measurements for state estimation. The main types of measurements include the following:

- Line power flows
- Line current magnitudes
- Line current angles
- Bus/Node power injections (P, Q)
- Bus/Node voltage magnitudes
- Bus/Node voltage angles

Bus voltage magnitudes and angles are the most common for bulk system state estimation, while power or current flows, and injections, are better for distribution system state estimation. Typically, distribution system operators are more focused on current. For DSSE, the measurements may come from equipment in substations or on the distribution feeders, or from AMIs at the customer point of connection. These measurements will most likely be at different time steps and resolutions, and may differ between an instantaneous measurements and average measurements over the time step. As discussed in the section on dynamic state estimation, steps will need to be taken to reconcile the measurements for use in state estimation. The historical meter data can also be used to create pseudo-measurements, or forecasted measurements used to make a "best guess" of sorts, when not enough real measurements are at hand. If a real-time state estimation is required, this may be a major barrier.

DATA ISSUES

The issues with measurement data on the distribution system are quite different from the issues with the transmission system. These differences involve the types of devices installed, the amount of time and effort required to add new measurements, and the types of communications infrastructure available to obtain those measurements (see Table 3).

IDENTIFYING BAD DATA

The presence of bad data negatively impacts the accuracy of the state estimate. It is critical that incorrect data be detected, identified, and ignored. This goal is achieved with information provided from redundant measurements, a typical requirement of a state estimator. This involves two requirements: first, detecting the existence of bad measurements; and second, identifying which data is bad.

There are several techniques that have been developed in association with transmission state estimation to identify the presence and location of bad sources of data, and then exclude those bad data sources from the results. The commonly applied method is the chi-squared test, which evaluates the probability that the weighted sum of squares with n degrees of freedom is less than a chi-squared distribution. If that test is successful, then the analysis can proceed with confidence that there aren't any bad measurements. If the test is not successful and the confidence level is low, then bad data exists in the measurement set. Note that the chi-squared test does not indicate which data is bad.

Identifying which data is bad is more complicated. An initial step is to identify data outside of operational norms (ratings, ranges). If that does not resolve the low confidence level, then additional, more computationally intensive statistical techniques are used. The residuals of the weighted least squares test are evaluated. Data with large residuals is excluded, and then the chi-squared test is re-executed to determine if that improves the confidence levels. A significant improvement helps identify the bad data. Obviously, this technique is reliant on redundant information.

Considering a distribution state estimation solution with a different algorithmic approach to getting a solution will likely involve a different process and algorithm for detecting and identifying bad data. Three possible approaches for data error detection and location are flagging measurements that are outside of some tolerance band around nominal, applying Kirchhoff's laws to detect anomalies, and flagging measurements that are stale, missing, or have error codes.

CONCLUSIONS AND NEXT STEPS

Considering all of the evidence and research available, it is clear that DSSE will be a necessary component of a DMS moving forward. A number of mission-critical operational applications have been identified that will require accurate near real-time estimates of the state of the system. Some insights gained in the course of developing this white paper follow.

Table 3 – Issues regarding DSSE data

Data Issue	Discussion
Measurement Latency	Distribution automation systems often rely on low-bandwidth, high-latency communications media. To economize on the resources involved, DA systems often employ exception reporting on analog measurements. This not only introduces a measurement accuracy issue, but also impacts the ability to gain measurements from different locations within a reasonable amount of time.
Measurement Types	Measurements on the distribution system may be reported as instantaneous (typical for DA) or average (typical for AMI).
Rolled Phases	The installation of three sensors on a distribution pole can be accomplished without specific drawings, and there is no way to locally verify if the measurements viewed upon testing are correct. Thus, the opportunity for phase error is not trivial.
Missing Measurements	The biggest single issue with distribution systems is the significant lack of measurements. While this situation is getting better, it is likely that there will never be enough measurements to achieve observability in the traditional sense.

Conclusions

Table 4 summarizes some key points made in the paper.

Next Steps

Commercially available DSSE processes and algorithms are being deployed. At first glance, the results seem reasonable, but the only way to determine if they are truly accurate enough to move forward with is to put them to the test. What is needed is a structured test plan to compare and statistically evaluate the accuracy of DSSE output with field-conducted, time-synchronized, accurate measurements of key state variables (current, voltage, phase angles) on feeders whose topology models have been verified. An EPRI demonstration project would be a great third-party approach to accomplish this goal across a number of utilities and implemented solutions. This project could evaluate state estimate accuracies and intentionally subject the algorithms to known issues (data errors, model errors, abnormal operating conditions) to gauge their resiliency.

BIBLIOGRAPHY

Abur, A., and A. G. Expósito, *Power System State Estimation: Theory and Implementation*. CRC Press, 2004.

Dugan, R. “Distribution State Estimation with OpenDSS.” [Online]. Available at https://sourceforge.net/p/electricdss/code/HEAD/tree/trunk/Distrib/Doc/OpenDSS%20and%20State%20Estimation.pdf. [Accessed October 4, 2018].

Primadianto, A., and C. N. Lu, “A Review on Distribution System State Estimation,” *IEEE Transactions on Power Systems*, Vol. 32, No. 5, pp. 3875–3883, September 2017.

Schweppe, F., J. Wildes, and D. Rom. “Power system static state estimation: Parts I, II, and III.” Power Industry Computer Conference (PICA), 1969.

Table 4 – Conclusions about DSSE
Distribution state estimation does not necessarily have to use the same algorithmic solution as transmission state estimation. All it really needs to do is provide some systematic, automatic, and accurate method of “estimating the state” of the distribution system, and we can continue to call it DSSE, with the understanding that although “estimating the state” is still the goal, the algorithm and process will be very different.
Distribution state estimation implies a structured process that includes accurate topology, measurements, abilities to identify/exclude bad measurements, analysis of voltage profile, analysis of power flow, and presentation of results.
Key distribution operations applications like VVO, FLISR, short-term forecasting, and planned switching are going to be more dependent on DSSE as the levels of DER penetration increase.
Some planning applications, such as load allocation and reactive planning, are also going to be dependent on an accurate assessment of the state of the distribution system during peak conditions (or other analyzed cases).
Observability requirements for DSSE must be designed with the practical limitations imposed on the ability to invest in sensors and deploy them on the distribution system.
Feeder topology reduction techniques may contribute to improving the observability requirements, but at the cost of less granular results.
Pseudo (predicted/forecasted/best guess) measurements may play an important role in distribution state estimation because there are not enough true measurements available in the distribution system. Various sources for their creation have been proposed, such as AMI data.

Disclaimer of Warranties and Limitation of Liabilities

This document was prepared by the organization(s) named below as an account of work sponsored or cosponsored by the Electric Power Research Institute, Inc. (EPRI). Neither EPRI, any member of EPRI, any cosponsor, the organization(s) below, nor any person acting on behalf of any of them:

(a) Makes any warranty or representation whatsoever, express or implied, (i) with respect to the use of any information, apparatus, method, process, or similar item disclosed in this document, including merchantability and fitness for a particular purpose, or (ii) that such use does not infringe on or interfere with privately owned rights, including any party's intellectual property, or (iii) that this document is suitable to any particular user's circumstance; or

(b) Assumes responsibility for any damages or other liability whatsoever (including any consequential damages, even if EPRI or any EPRI representative has been advised of the possibility of such damages) resulting from your selection or use of this document or any information, apparatus, method, process, or similar item disclosed in this document.

Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not neces- sarily constitute or imply its endorsement, recommendation, or favoring by EPRI.

The Electric Power Research Institute (EPRI) Prepared this Report.

This is an EPRI Technical Update report. A Technical Update report is intended as an informal report of continuing research, a meeting, or a topical study. It is not a final EPRI technical report.

Note

For further information about EPRI, call the EPRI Customer Assistance Cen- ter at 800.313.3774 or e-mail askepri@epri.com.

EPRI RESOURCES

Brian Deaver, *Senior Technical Executive*
443.910.2553, bdeaver@epri.com

Distribution Operations and Planning

Nick Heine, *Engineer/Scientist II*
865.218.8157, nheine@epri.com

Alex Melhorn, *Engineer/Scientist II*
865.218.5911, amelhorn@epri.com

Power Systems Studies

The Electric Power Research Institute, Inc. (EPRI, www.epri.com) conducts research and development relating to the generation, delivery and use of electricity for the benefit of the public. An independent, nonprofit organization, EPRI brings together its scientists and engineers as well as experts from academia and industry to help address challenges in electricity, including reliability, efficiency, affordability, health, safety and the environment. EPRI members represent 90% of the electric utility revenue in the United States with international participation in 35 countries. EPRI's principal offices and laboratories are located in Palo Alto, Calif.; Charlotte, N.C.; Knoxville, Tenn.; and Lenox, Mass.

Together . . . Shaping the Future of Electricity

Export Control Restrictions

Access to and use of EPRI Intellectual Property is granted with the specific understanding and requirement that responsibility for ensuring full compliance with all applicable U.S. and foreign export laws and regulations is being undertaken by you and your company. This includes an obligation to ensure that any individual receiving access hereunder who is not a U.S. citizen or permanent U.S. resident is permitted access under applicable U.S. and foreign export laws and regulations. In the event you are uncertain whether you or your company may lawfully obtain access to this EPRI Intellectual Property, you acknowledge that it is your obligation to consult with your company's legal counsel to determine whether this access is lawful. Although EPRI may make available on a case-by-case basis an informal assessment of the applicable U.S. export classification for specific EPRI Intellectual Property, you and your company acknowledge that this assessment is solely for informational purposes and not for reliance purposes. You and your company acknowledge that it is still the obligation of you and your company to make your own assessment of the applicable U.S. export classification and ensure compliance accordingly. You and your company understand and acknowledge your obligations to make a prompt report to EPRI and the appropriate authorities regarding any access to or use of EPRI Intellectual Property hereunder that may be in violation of applicable U.S. or foreign export laws or regulations.

3002013421 November 2018

Electric Power Research Institute
3420 Hillview Avenue, Palo Alto, California 94304-1338 • PO Box 10412, Palo Alto, California 94303-0813 USA
800.313.3774 • 650.855.2121 • askepri@epri.com • www.epri.com