J. Peppanen, M. J. Reno, Robert J. Broderick and S. Grijalva, "Distribution System Model Calibration with Big Data from AMI and PV Inverters", in IEEE Transactions on Smart Grid, 2016. DOI: 10.1109/TSG.2016.2531994

Distribution System Model Calibration with Big Data from AMI and PV Inverters

Jouni Peppanen, *Student Member, IEEE*, Matthew J. Reno, *Member, IEEE*, Robert J. Broderick, *Member, IEEE, and* Santiago Grijalva, *Senior Member, IEEE*

***Abstract*— Efficient management and coordination of distributed energy resources with advanced automation schemes requires accurate distribution system modeling and monitoring. Big Data from smart meters and PV micro-inverters can be leveraged to calibrate existing utility models. This paper presents computationally efficient distribution system parameter estimation algorithms to improve the accuracy of existing utility feeder radial secondary circuit model parameters. The method is demonstrated using a real utility feeder model with AMI and PV micro-inverters, along with alternative parameter estimation approaches that can be used to improve secondary circuit models when limited measurement data is available. The parameter estimation accuracy is demonstrated for both a 3-phase test circuit with typical secondary circuit topologies and 1-phase secondary circuits in a real mixed-phase test system.**

***Index Terms*-- Load Modeling, Parameter Estimation, Power Distribution, Power System Modeling, Power System Measurements, Power System Simulation, Regression Analysis, Smart Grids**

I. Introduction

EFFICIENT management of distributed energy resources (DER), especially renewable energy sources, with advanced Volt/VAr optimization and other distribution system automation schemes requires accurate and reliable distribution system modeling, monitoring, and coordination [1], [2]. This can be realized by exploiting the large amounts of emerging data from advanced metering infrastructure (AMI) and micro-inverters.

Due to the large number of parameters, system changes, and load conditions involved in a distribution system model, there is a large degree of uncertainty with respect to the accuracy and quality of current utility models. Stored circuit model including the model parameter values may be incorrect as a result of unknown data, human errors, inaccurate manufacturing data, network changes, etc. [3]. Improving the accuracy of feeder parameters is important to allow higher penetrations of DERs and reliable control of the devices.

J. Peppanen and S. Grijalva are affiliated with Georgia Institute of Technology, North Avenue, Atlanta GA 30332 USA (e-mail: jouni.peppanen@gatech.edu).

M. J. Reno and R. J. Broderick are with the Sandia National Laboratories.

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

It is particularly important to improve the models of the distribution system secondary (low voltage) networks where a large share of the distributed energy resources are located. Moreover, the secondary networks are typically modeled with a lower level of detail compared to the well-modeled primary (medium-voltage) networks, and a significant portion of per-unit voltage drop/raise occurs over the service transformers and lines that have large impedances and low X/R-ratios. Automated parameter estimation (PE) procedures for improving the secondary circuit model parameters are necessary in order to minimize manual inspections that are costly and hard to perform in densely populated urban areas with wiring underground and in buildings.

Conventionally, PE methods have been applied to the transmission system. In that setting, PE is aimed to refine a handful of suspicious parameters and typically assume accurate, well time-synchronized and highly redundant set of measurements, which has rarely been the situation in distribution systems [3], [4]. Moreover, many conventional PE approaches for transmission systems either require a residual vector from an existing state estimator or involve some modifications to the existing state estimator algorithm [3], [5], [6]. Due to the limited deployment of state estimators in distribution systems, these methods are not readily available. The Big Data from AMI and other emerging sensors has raised the interest in new methods for distribution system parameter estimation (DSPE). A linear optimization method for topology error detection and parameter estimation has been proposed in [7] but the method only utilizes active powers. In [8], the authors utilize a quadratic formulation and a gradient-based approach to minimize the variance of voltage estimates from various smart meters. The approach makes no simplifications to the AC power flow equations but results in a non-convex optimization problem with quadratic equality constraints that with hundreds of required iterations is computationally much more intensive to solve than non-iterative linear regression based methods that require solving a simple linear system of equations. Practical methods for meter phase identification, meter-to-transformer mapping, and joint parameter and topology estimation are shown in [9].

This paper focuses on off-line estimation of time-invariant service transformer and secondary system line impedances. The local measurement redundancy is increased by utilizing a large number of historical measurement samples to reduce the impacts of the lower measurement granularity and accuracy of AMI and emerging measurement sources. This paper is a natural continuation on our previous work on the distribution system parameter estimation in [10] and the preliminary results, on refining the Georgia Tech campus distribution system model parameters in [11], and [12].

The key contribution of this paper are two-fold:

1) Present a practical and computationally efficient method for estimating 3-phase or 1-phase secondary circuit model series impedance parameters with fully available Big Data from AMI voltage and power (or current) measurements. The method is also shown for handling some meters not reporting voltage measurements. This method is validated for a 3-phase test circuit with various real secondary circuit topologies.

2) Show a novel practical and computationally efficient method for generating secondary circuit model with limited available PV microinverter data. This method is demonstrated for 1-phase real U.S. utility feeder secondary circuits.

The second method for limited available PV microinverter data is novel. The first method for fully available Big Data from AMI measurements presents several extensions over [7], [9] including the parallel branch estimation formulation (9)-(11), the upstream node voltage estimation with (12), and the linearly constrained least squares estimation (8) for bounding parameter values. This paper also presents an optimal linear regression model (13)-(16) for typical secondary circuit topologies and different levels of measurement error. The approach to handle some meters with missing voltage measurements is novel.

This paper addresses the need for utilities to improve the analytical and operational modeling accuracy for future smart distribution systems with DER. The work also provides further use cases for smart meter and DER data.

This paper has the following structure. Section II presents the utilized branch series impedance parameter estimation method, and Section III expands the method to estimating the series impedance parameters of entire radial secondary circuits. Optimal linear regression model is also discussed. Section IV presents a modified parameter estimation algorithm when only a limited set of PV inverter measurements are available. Section V discusses parameter estimation implementation in utility Big Data environment. Section VI demonstrates the performance of the parameter estimation algorithms on a three-phase test circuit, and Section VII shows parameter estimation results for utility feeder with single-phase secondary circuits. Section VIII concludes the paper.

II. Estimating Branch Series Impedance Parameters

The presented parameter estimation is based on the well-known (see e.g. [13], [14]) linear approximation of voltage drop magnitude over a series impedance shown in Figure 2 on the right

$$\Delta V = V_1 - V_2 \approx (RP + XQ)/V_2 = RI_R + XI_X, \qquad (1)$$

where V_1 and V_2 are voltage magnitudes, R and X are the series resistance and reactance between two buses (positive sequence for balanced 3-phase branches and phase impedance for 1-phase branches). The current resistive and reactive components are given with

$$I_R = P/V = I(PF) \text{ and } I_X = Q/V = I\sqrt{1-(PF)^2}, \qquad (2)$$

where (PF) is the power factor. For transformers, all values must be referred to the same voltage level. In 3-phase systems, line-line voltages and 3-phase powers are used whereas in 1-phase systems, line-to-neutral voltages are utilized.

The accuracy of the linearized voltage drop approximation equation (1) is shown to be good in most typical situations [10], [14]. However, the equation increasingly underestimates the voltage drop magnitudes for impedances with high X/R-ratios, e.g., service transformers [10]. As a result, without measurement error the resistances tend to over-estimated and reactances under-estimated especially for components with high X/R-ratios. This bias may also be present in transformer parameters estimated with noisy measurements. On the other hand, since measurement noise can be a considerable fraction of service line voltage drop, the estimation bias is less likely to be seen in service line R and X parameter estimates whose accuracy is more driven by the characteristics of the measurement error. It is possible to counteract the approximation error of equation (1) by adding higher order predictor terms to the linear regression model or by filtering parameter estimation samples with higher expected approximation error. However in practice, equation (1) approximation error is expected to be insignificant compared to measurement error and modeling simplifications and inconsistencies [10].

The goal of the branch (positive sequence) series impedance parameter estimation problem is to find the most likely parameters R and X that provide the best fit of the M measurement samples to the linear model

$$\Delta \boldsymbol{V} = \boldsymbol{V}_1 - \boldsymbol{V}_2 = R\boldsymbol{I}_R + X\boldsymbol{I}_X + \boldsymbol{\epsilon}, \qquad (3)$$

where $\boldsymbol{\epsilon} \in \mathbb{R}^M$ captures the model and measurement error and all the bold letters indicate vectors (or matrices) through time. Denoting the response vector $\mathbf{y} \in \mathbb{R}^M$, the design matrix $\boldsymbol{\mathcal{X}} \in \mathbb{R}^{M\times 2}$, and the unknown parameter vector $\boldsymbol{\beta} \in \mathbb{R}^2$

$$\boldsymbol{y} = \boldsymbol{V}_1 - \boldsymbol{V}_2,\ \boldsymbol{\mathcal{X}} = [\boldsymbol{I}_R \quad \boldsymbol{I}_X], \text{ and } \boldsymbol{\beta} = [R \quad X]^{\mathrm{T}} \qquad (4)$$

respectively, gives

$$\boldsymbol{y} = \boldsymbol{\mathcal{X}}\boldsymbol{\beta} + \boldsymbol{\epsilon}. \qquad (5)$$

An estimate for the unknown parameters, $\widehat{\boldsymbol{\beta}}$, can be obtained by, e.g., minimizing the p-norm of the model residuals over the measurement samples

$$\widehat{\boldsymbol{\beta}} = \min_{R,X} \|\boldsymbol{y} - \boldsymbol{\mathcal{X}}\boldsymbol{\beta}\|_p. \qquad (6)$$

If $p = 1$, (6) becomes a linear programming problem. With $p = 2$, (6) is a linear unconstrained least squares problem whose solution is the ordinary least squares (OLS) estimator given by

$$\widehat{\boldsymbol{\beta}} = (\boldsymbol{\mathcal{X}}^{\mathrm{T}}\boldsymbol{\mathcal{X}})^{-1}\boldsymbol{\mathcal{X}}^{\mathrm{T}}\boldsymbol{y}. \qquad (7)$$

Sometimes it is desirable to set bounds on the impedance parameter estimates. This can be done by utilizing linearly constrained least squares formulation

$$\hat{\boldsymbol{\beta}} = \min_{\boldsymbol{\beta}} \boldsymbol{\beta}^{\mathrm{T}} \boldsymbol{\mathcal{X}}^{\mathrm{T}} \boldsymbol{\mathcal{X}} \boldsymbol{\beta} \\ \text{subject to } \boldsymbol{C\beta} \le \boldsymbol{d}. \tag{8}$$

This is a quadratic programming problem that can be effectively (in polynomial time) solved to a global (but possibly not unique) optimum with any open-source or commercial solver. In this paper, the linearly constrained least squares solution is used when the OLS estimator results in non-physical, e.g., negative or too high, parameter values.

III. DSPE Algorithm

A. DSPE Problem Definition

This section generalizes the branch series impedance parameter estimation method to estimation of the series impedance parameters of entire radial secondary circuit [10]. The objective of the method is to find the most likely values of resistance (R) and reactance (X) parameters shown in red in Figure 1. The method assumes that historical voltage (V), active power (P), and reactive power (Q) measurements shown in blue in the figure are available at all the leaf nodes of the secondary circuit tree. To estimate the service transformer parameters, the method requires measured or simulated service-transformer medium-voltage values.

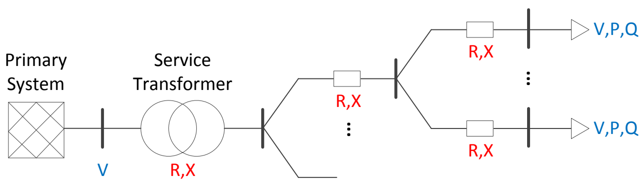

Figure 1. Secondary circuit tree for parameter estimation

The proposed method relies on the following four assumptions.

1. The secondary circuit topology is assumed to be known. If the topology is unknown, it can be estimated following the approach we have shown in [12].

2. The secondary circuit is assumed to be radial (i.e. a tree) like most real secondary circuits [13].

3. The active and reactive power (or current and power factor) and the voltage measurements are assumed to be available at all leaf nodes of the tree. In practice, this assumption is valid since most secondary circuit tree leaf nodes have either a load and/or a distributed generation (DG) unit with the respective measurements. Handling cases where some meters report no voltage measurements is discussed in [10].

4. The secondary circuit is assumed to be either balanced 3-phase or single-phase. This assumptions is often invalid since in practice many distribution system secondary circuits are split-phase, i.e., a single-phase where a center-tapped transformer connects to a triplex cable with both 120V and 240V service to the loads. Although it is possible to model the split-phase secondary circuits in detail [15], parameter estimation is limited by the available measurement data, which typically consists of the customer total power and/or current as well as voltage measurement across the 120V (or the 240V) connection. As long as the power, current and voltage measurements for both the 120V and 240V loads are not included in the MDMS, it may be desirable to model split-phase secondary circuits with single-phase transformers, lines, and loads. Using this modeling approach, typical measurement meter data can be readily utilized to estimate the secondary circuit transformer and line parameters utilizing the approach introduced below.

B. DSPE Algorithm

The distribution system secondary circuit parameter estimation (DSPE) algorithm processes one secondary circuit tree at a time, hierarchically proceeding from the tree leaf nodes towards the tree root node. At a given iteration the algorithm estimates the branch impedances for a subsection of the secondary circuit as follows. First, the algorithm searches for a circuit subsection, whose parameters shown in red in Figure 2 have not been estimated yet, that consists of either A) a branch that has known (measured or estimated at previous iteration) upstream and downstream node voltages and downstream node currents shown in blue in Figure 2 on the right, or B) a set of parallel branches with known downstream node voltages and currents shown in blue in Figure 2 on the left. Once a suitable circuit subsection has been identified, the algorithm first estimates the branch impedance parameters and then in case of the parallel branch case, estimates the upstream node voltages using the measurements and the estimated branch parameters. These steps are listed in Algorithm 1.

The linear regression formulation for circuit sections without parallel branches in Algorithm 1 is given in (4)-(5). The linear regression formulation for the case with $N \in \{2,3,\dots\}$ parallel branches and M measurement samples in Algorithm 1 is given by (5) where $\boldsymbol{\epsilon} \in \mathbb{R}^{M}$ is the error vector, $\boldsymbol{\beta} \in \mathbb{R}^{(M+2N)}$ is the parameter vector given by

$$\boldsymbol{\beta} = \left[V_{0,1}, \dots, V_{0,M}, R_1, X_1, \dots, R_N, X_N\right]^{\mathrm{T}}, \tag{9}$$

and $\boldsymbol{y} \in \mathbb{R}^{MN}$ is the response vector is given by

$$\boldsymbol{y} = \left[V_{1,1}, \dots, V_{1,M}, \dots, V_{N,1}, \dots, V_{N,M}\right]^{\mathrm{T}}. \tag{10}$$

Finally, the design matrix $\boldsymbol{\mathcal{X}} \in \mathbb{R}^{(MN)\times(M+2N)}$ is given by

$$\boldsymbol{\mathcal{X}} = \begin{bmatrix} \mathrm{I} & [-\boldsymbol{I}_{R,1} & -\boldsymbol{I}_{X,1}] & \cdots & \mathbf{0} \\ \vdots & \vdots & & \ddots & \vdots \\ \mathrm{I} & \mathbf{0} & & \cdots & [-\boldsymbol{I}_{R,N} \quad -\boldsymbol{I}_{X,N}] \end{bmatrix}, \tag{11}$$

where $\mathrm{I} \in \mathbb{R}^{M\times M}$ are identity matrices, $\boldsymbol{I}_{R,i}, \boldsymbol{I}_{X,i} \in \mathbb{R}^{M\times 1}, i \in \{1, \dots, N\}$ are the branch current measurements, and the zero submatrices have suitable sizes. This formulation has $(M + 2N)$ unknowns (excluding the error terms) and MN equations. In practice, large sample size M is used and thus, $M \gg N$.

Once the impedances, currents, and downstream node voltages of the N parallel branches are known, the voltages of the upstream node of the branches can be estimated as an average of the upstream node voltage estimates of the N branches with

$$\boldsymbol{V}_0 = \frac{1}{N}\sum_{i=1}^{N}\left\|\boldsymbol{V}_i + (R_i + jX_i)(\boldsymbol{I}_{R,i} + j\boldsymbol{I}_{X,i})\right\|. \tag{12}$$

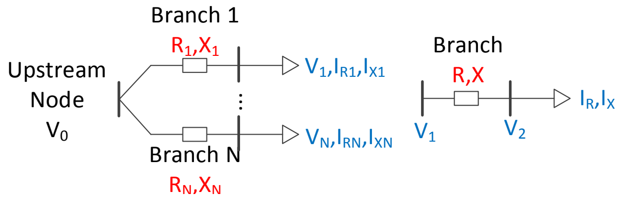

Figure 2. Secondary circuit section with N parallel branches (left) and no parallel branches (right)

Algorithm 1: DSPE Algorithm

Input: List of all secondary circuit branches, $\mathcal{L}$, with fields upstream and downstream node, # parallel branches, branch current measurements $\boldsymbol{I}_R, \boldsymbol{I}_X$, branch node voltage measurements $\boldsymbol{V}$

Output: Secondary circuit branch estimation results with fields $R_{est}, X_{est}, R_{pval}, X_{pval}, R^2$,and, MSE.

1. If $\mathcal{L}$ is empty, **STOP**.

IF $\mathcal{L}$ has only one branch, set ℓ, the list of active branches, equal to the last branch and remove the last branch from $\mathcal{L}$.

ELSE Find a branch i whose all parallel branches $j_1, \ldots, j_{N-1}$ have downstream node voltage measurements or estimates. Set $\ell = \{i, j_1, \ldots, j_{N-1}\}$.

ENDIF

IF ℓ has only one branch, estimate the impedance parameters of the branch in ℓ with the single-branch regression formulation (3)-(5).

ELSE Estimate the impedance parameters of the branches in ℓ with the parallel branch regression formulation (9)-(11) and estimate the voltages of the upstream node of the N parallel branches with (12).

ENDIF

2. Go to Step **1.**

C. Optimal Linear Regression Model

The linear regression models (3)-(5) and (9)-(11) are based on the linearized voltage drop equation (3) and thus, the predictors and unknown parameters have direct physical meanings in both formulations. However, linear regression allows models with higher order terms, cross-couplings, or any other functions of the predictor variables I_R and I_X. Unlike R and X in (4) and (11), the coefficients of other terms do not have a direct physical meaning, but including them in the regression models may better capture the intrinsic nonlinear relationship between the response variables and the predictor variables thus, leading to better estimates for R and X. The best regression model depends, among other things, on the characteristics of the data and on the values of the true parameters. More complicated models can better estimate true impedance values under conditions of low measurement noise levels. On the other hand, the higher the measurement error level is or the lower the true impedance magnitudes (and the voltage drop) are, the simpler regression models should be used. A detailed analysis of different regression models and their accuracy can be found [10].

Good parameter estimation accuracy is obtained when I_R^2, the second order term of the resistive (real power) current is added to regression problems that include one or more transformers [10]. Thus, single transformer parameters are best estimated by utilizing design matrix $\boldsymbol{\mathcal{X}}$ and parameter vector $\boldsymbol{\beta}$ are given by

$$\boldsymbol{\mathcal{X}} = [\boldsymbol{I}_{R1}, \boldsymbol{I}_{X1}, \boldsymbol{I}_{R1}^2] \tag{13}$$

$$\boldsymbol{\beta} = \left[R_1, X_1, \beta_{Rsq,1}\right]^{\mathrm{T}}, \tag{14}$$

where parameters $\beta_{Rsq,1}$ and $\beta_{Xsq,1}$ do not have a direct physical meaning. The response vector $\boldsymbol{y}$ is the same as in (4). Similarly, the parameters of N parallel branches, i^{th} of which is a transformer, are best estimated by utilizing design matrix $\boldsymbol{\mathcal{X}}$ and parameter vector $\boldsymbol{\beta}$ are given by

$$\boldsymbol{\mathcal{X}} = \begin{bmatrix} \mathrm{I} & [-\boldsymbol{I}_{R,1} & -\boldsymbol{I}_{X,1}] & \cdots & \mathbf{0} & & \mathbf{0} \\ \vdots & \vdots & & \ddots & \vdots & & \vdots \\ \mathrm{I} & \mathbf{0} & & \cdots & [-\boldsymbol{I}_{R,N} & -\boldsymbol{I}_{X,N}] & -\boldsymbol{I}_{R,i}^2 \end{bmatrix} \tag{15}$$

$$\boldsymbol{\beta} = \left[V_{0,1}, \ldots, V_{0,M}, R_1, X_1, \ldots, R_N, X_N, \beta_{Rsq,i}\right]. \tag{16}$$

The response vector $\boldsymbol{y}$ is the same as in (10).

These selected regression models are optimized for the practical setting where the measurement error dictates the parameter estimation accuracy. Without measurement error, parameters can be estimated with a smaller error by using regression models with additional higher-order terms of the predictor variables.

IV. SDSPE Algorithm With Limited PV Measurements

Ideally, secondary circuit parameters are estimated using a large set of synchronized historical voltage, active power, and reactive power measurement samples available from all the secondary circuit loads and distributed generation. In practice however, not all loads and DG units are metered and not all metered values are stored into a historical database. Moreover, some (especially older) meters may provide power (or current) measurements but no voltage measurements. A modified DSPE algorithm, that can handle some meters that do not transmit voltage measurements, is shown in [10]. Although the modified algorithm has good accuracy when some meters do not report voltage measurements, any meter without voltage measurements reduces the accuracy and observability of the (secondary circuit) parameter estimation and thus, it is desirable to have high-quality voltage measurements from all smart meters.

This section presents a simplified distribution system secondary circuit parameter estimation (SDSPE) algorithm for creating secondary circuit models when no (or very limited) AMI measurements are available but when a historical database of PV system measurements is available. Figure 3 illustrates the simplified secondary circuit type that has a customer with a PV system connected to the service transformer secondary over a service line. The secondary system also has other customers with loads (but no PV systems) connected to the service transformer with potentially several service lines.

We make the following assumptions and simplifications.

1. Each secondary circuit (of interest) has one or more PV systems (or other sensor) measuring voltages and active and reactive powers V_{PV}, P_{PV}, Q_{PV} shown in blue in Figure 3. The discussion here focuses on one PV system in each secondary circuit but generalization to multiple PV systems is trivial.

2. Total SCADA feeder power measurements are available.

3. The secondary circuit loads P_1, P_2, Q_1, Q_2 shown in green in Figure 3 are estimated from the feeder total load with load allocation, e.g., based service transformer rating and the number of customers in each secondary circuit.

4. The feeder primary (medium-voltage) system is well-modeled including the service transformer connection.

5. The customer with the PV system is assumed to be connected to the service transformer secondary over a separate service line but the rest of the secondary circuit topology, component types, and component parameters are unknown.

6. Transformer primary side voltage referred to the low-voltage side, V_0, is estimated with time-series power flow simulation. Due to primary circuit modeling inconsistencies and the simplifications of load allocation, the simulated voltages may not be very accurate.

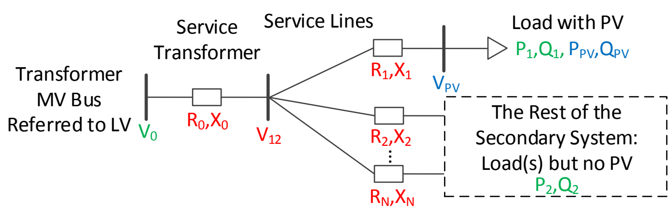

Figure 3. Simplified secondary circuit model with a PV system: available measurements are in blue, values that can be roughly estimated are in green, and unknown values and parameters are in red

Since the "rest of the secondary circuit" has no voltage measurements, it is not possible to estimate the impedances $R_2, X_2, \dots, R_N, X_N$ or any impedances in the "rest of the circuit". Thus, the circuit in Figure 3 can be simplified to the circuit shown in Figure 4.

The objective is to estimate the unknown parameters R_0, X_0, R_1, X_1 shown in red in Figure 4 by utilizing the available measurements V_{PV}, P_{PV}, Q_{PV} shown in blue in Figure 4 and the estimated measurements V_0, P_1, Q_1, P_2, Q_2 shown in green in Figure 4. This can be achieved by utilizing M synchronous measurement samples in the linear regression formulation

$$\boldsymbol{V}_0 - \boldsymbol{V}_{PV} = R_0 \boldsymbol{I}_{R0} + X_0 \boldsymbol{I}_{X0} + R_1 \boldsymbol{I}_{R1} + X_1 \boldsymbol{I}_{X1} + \boldsymbol{\epsilon}. \quad (17)$$

The currents are

$$\boldsymbol{I}_{R0} = (\delta_1 \boldsymbol{P}_{SS} - \boldsymbol{P}_{PV})/\boldsymbol{V}_{PV} + \delta_2 \boldsymbol{P}_{SS}/\boldsymbol{V}_{12}, \quad (18)$$

$$\boldsymbol{I}_{X0} = (\delta_1 \boldsymbol{Q}_{SS} - \boldsymbol{Q}_{PV})/\boldsymbol{V}_{PV} + \delta_2 \boldsymbol{Q}_{SS}/\boldsymbol{V}_{12}, \quad (19)$$

$$\boldsymbol{I}_{R1} = (\delta_1 \boldsymbol{P}_{SS} - \boldsymbol{P}_{PV})/\boldsymbol{V}_{PV}, \text{ and} \quad (20)$$

$$\boldsymbol{I}_{X1} = (\delta_1 \boldsymbol{Q}_{SS} - \boldsymbol{Q}_{PV})/\boldsymbol{V}_{PV}, \quad (21)$$

where $\boldsymbol{V}_{12}$ are the service transformer secondary voltages, $\boldsymbol{P}_{SS}, \boldsymbol{Q}_{SS}$ are the feeder total power measurements, δ_1 and δ_2 are the load allocation factors for the load at the PV and the other loads, respectively. If no reliable measurements for the feeder total reactive power are available, $\boldsymbol{Q}_{SS}$ can be estimated from the feeder total active power measurements $\boldsymbol{P}_{SS}$ with a constant power factor (PF) with

$$\boldsymbol{Q}_{SS} = \sqrt{1/(PF)^2 - 1}\,\boldsymbol{P}_{SS} := \gamma \boldsymbol{P}_{SS}. \quad (22)$$

$\boldsymbol{V}_{12}$ is unknown because none of the other loads on the secondary circuit had voltage measurements. For simplicity, the approximation of $\boldsymbol{V}_{12} \approx \boldsymbol{V}_{PV}$ is used since the voltage drop over the service line is relatively small, and in practice, the errors resulting from generic load allocation without power measurements introduces much more error. Moreover, PV systems often operate at unity power factor ($\boldsymbol{Q}_{PV} = 0$). As a result, the currents in (18)-(21) are

$$\boldsymbol{I}_{R0} = \delta_1 \boldsymbol{P}_{SS}/\boldsymbol{V}_{PV} - \boldsymbol{P}_{PV}/\boldsymbol{V}_{PV} + \delta_2 \boldsymbol{P}_{SS}/\boldsymbol{V}_{PV}, \quad (23)$$

$$\boldsymbol{I}_{X0} = \delta_1 \gamma \boldsymbol{P}_{SS}/\boldsymbol{V}_{PV} + \delta_2 \gamma \boldsymbol{P}_{SS}/\boldsymbol{V}_{PV}, \quad (24)$$

$$\boldsymbol{I}_{R1} = \delta_1 \boldsymbol{P}_{SS}/\boldsymbol{V}_{PV} - \boldsymbol{P}_{PV}/\boldsymbol{V}_{PV}, \text{ and} \quad (25)$$

$$\boldsymbol{I}_{X1} = \delta_1 \gamma \boldsymbol{P}_{SS}/\boldsymbol{V}_{PV}. \quad (26)$$

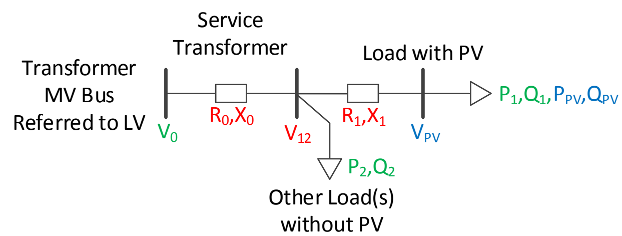

Figure 4. Simplified secondary circuit with a PV system for the SDSPE algorithm: available measurements are in blue, values that can be roughly estimated are in green, and unknown values and parameters are in red

Since no reactive power measurements $\boldsymbol{Q}_{SS}$ are available and since the PV system operates at unity power factor ($\boldsymbol{Q}_{PV} = 0$), (23)-(26) are linear combinations of two measurements vectors $\boldsymbol{P}_{SS}/\boldsymbol{V}_{PV}$, $\boldsymbol{P}_{PV}/\boldsymbol{V}_{PV}$, only two parameters can be estimated from (17), (23)-(26) as follows. If the line per-unit-length resistance r and reactance x and the transformer X/R-ratio $(X/R)_0$ are assumed to be known, transformer resistance, R_0 and the line length, l_1, can be estimated with

$$\boldsymbol{V}_0 - \boldsymbol{V}_{PV} = R_0 \boldsymbol{I}_0 + l_1 \boldsymbol{I}_1 + \boldsymbol{\epsilon}, \quad (27)$$

where the currents are given by

$$\boldsymbol{I}_0 = \boldsymbol{I}_{R0} + (X/R)_0 \boldsymbol{I}_{X0}, \quad (28)$$

$$\boldsymbol{I}_1 = r_1 \boldsymbol{I}_{R1} + x_1 \boldsymbol{I}_{X1}, \quad (29)$$

and $\boldsymbol{I}_{R0}, \boldsymbol{I}_{X0}, \boldsymbol{I}_{R1}, \boldsymbol{I}_{X1}$ are given in (23)-(26). Now, predictors $\boldsymbol{I}_0, \boldsymbol{I}_1$ are linearly independent provided that $\boldsymbol{P}_{PV} \not\equiv 0$. Once R_0 and l_1 have been estimated, the transformer reactance can be calculated with $X_0 = R_0 (X/R)_0$ and the line impedances with $R_1 + jX_1 = l(r_1 + jx_1)$.

V. Big Data Implementation for Distribution System Parameter Estimation

DSPE has an important role of validating and refining the existing utility feeder models and thus, preparing them for increased situational awareness and operational tasks in the future smart distribution systems. Figure 5 illustrates the flows of Big Data for distribution system parameter estimation.

The current model components, parameters and permanent connectivity will be fetched from GIS to build the distribution system model. SCADA will transmit the historical device measurements and states. AMI/MDMS will provide the load profiles, and DER the generation profiles, as an input to time series power flows that simulate the service transformer primary voltages. By leveraging the simulated service transformer voltages and distributed voltage and power (or current) measurements from the AMI and DER as well, the parameter estimator will estimate the (secondary system) component parameters. The estimated parameters are passed back to the distribution system model to simulate time series power flows with the estimated parameters. After passing a manual validation, the estimated component parameters are passed to GIS and the distribution system model.

The Big Data challenge is efficiently managing the data flows through advanced data analytics, optimized database queries, and rapid time series analysis. For distribution system parameter estimation to be practical, data processing and analyses must be automated as much as possible with limited human intervention to perform the manual validation of results. Moreover, to allow rapid manual validation of results, primary circuit gross modeling errors, suspected bad parameter estimates, and bad measurement data must be automatically identified.

As parameter estimation is performed offline, measurement system delays are not an issue but poorly synchronized measurement data must be re-synchronized, e.g., using a simple linear (or other) interpolation. Any inaccuracy resulting from the measurement re-synchronization can be counteracted by utilizing historical Big Data with large sample sizes, which the proposed methods can effectively handle.

Measurement data must also be preprocessed to identify and clean bad and missing data. A meter can have gross errors in some or in all of its measurement samples. Since the proposed parameter estimation methods can easily utilize thousands of samples, some measurement samples with gross errors do not have a high influence on the parameter estimation results. Some bad measurement samples of a meter can be identified with typical methods for detecting outliers in the linear regression response variable or the predictor variables [16]. On the other hand, many cases when all measurements of a meter have gross errors can be easily identified with simple checks such as the ones we have discussed in [11]. Moreover, if all loads in a secondary circuit have smart meters, it may be possible to identify meter (or model) gross errors from poor parameter estimation linear regression model fit (low R-squared values, high RMSE, insignificant parameter p-values). However as discussed in section VI, these metrics are not always effective at distinguishing between good and bad regression models. In some cases it can be very hard to identify meters with gross errors, e.g., when a load, which is small compared to the other loads in a secondary circuit, has a meter with gross errors.

Once missing and bad data samples have been identified, they must be imputed to allow the two time-series power flows that require full set of measurement data. Since bad or missing data always results in lost information, the samples during which any secondary circuit meter has missing or bad data should not be used for the actual linear regression parameter estimation but only for running the time-series power flows.

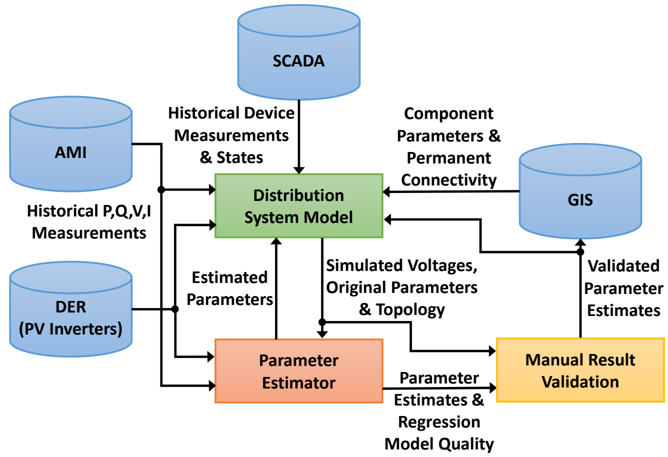

Figure 5. Big Data for distribution system parameter estimation

The high-level secondary circuit parameter estimation algorithm is shown in Figure 6. The existing utility feeder model is compiled and time series power flow is solved utilizing load active and reactive power (or current and power factor) measurements, substation voltage measurements, and PV generation as inputs. In this paper, the distribution system power flow is solved with OpenDSS, and all parameter estimation algorithms are implemented in MATLAB [17], [18]. The output from the time series power flows solutions are the service transformer MV-side voltages, which are needed to estimate the service transformer impedance.

Next, the algorithm proceeds one secondary circuit at a time, estimating the secondary circuit branch impedances with the approaches shown in Sections III-IV. After all the secondary circuits have been processed, another time series power flow simulation is run with the estimated parameters to compare measured voltages to the simulated voltages. In the manual verification step, the user needs to compare the estimated parameter values and how closely they align with physically expected values. The manual verification of the parameter estimation results is very important in order to avoid any possibilities of replacing previously accurate impedance parameters with poorer estimates. This step is also useful for detecting any data or topology problems based on, e.g., physically impossible parameter estimates or poor linear regression fits.

It should be emphasized that the presented methods do not require modifying any existing utility software. Moreover, the presented methods are computationally highly efficient since no iterative power flow solutions are required during the parameter estimation. Instead, the linear regression parameter estimation only requires solving a linear system, which can be done in a fraction of a second even when thousands of measurement samples are leveraged to counteract the accuracy, granularity, and time-synchronization issues related to AMI and DER measurements. Moreover, the presented methods allow processing each secondary circuit individually thus, making it possible to divide large and complicated feeder models to smaller sub problems. This divide and conquer approach significantly reduces the amount of input and output data that needs to be handled simultaneously and thus, allows utilizing large sample sizes for the parameter estimation. Since typical distribution feeder models consist of thousands of lines, hundreds of distribution transformers, and thousands of customers, it is very attractive to perform parameter estimation for one secondary circuit at a time. Moreover, since typical distribution system operators have hundreds or thousands of distribution feeders, the computational time per feeder model must remain modest in order for distribution system parameter estimation to be a practical and cost-effective approach for model calibration.

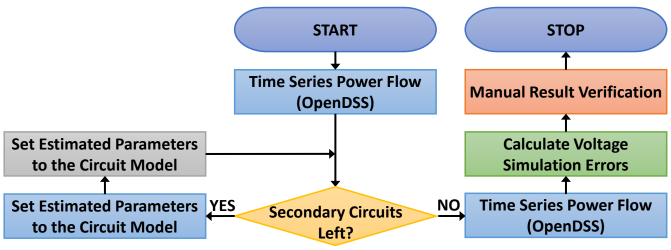

Figure 6. Parameter estimation algorithm

VI. THREE-PHASE TEST CIRCUIT RESULTS

A 66-node three-phase test circuit (3PTC) was created to demonstrate the parameter estimation performance [10]. The circuit has a 3-phase 12kV L-L backbone feeder and ten 3-phase 240V L-L secondary circuits each with a different topology. Each of the 36 loads was assigned a real AMI active power profile from [19] and a random power factor profile: (PF)~Unif(0.9,1.0). Typical primary and secondary line and transformer parameters were used. Figure 7 shows the test circuit topology with line contouring showing per-unit voltages and line widths showing current magnitudes. The service transformer MV side voltages were assumed to be accurately simulated (accurate primary system model) from a timeseries powerflow. The secondary network topologies were assumed to be known, and the hourly active power, reactive power, and voltage measurements of all loads are available from the AMI.

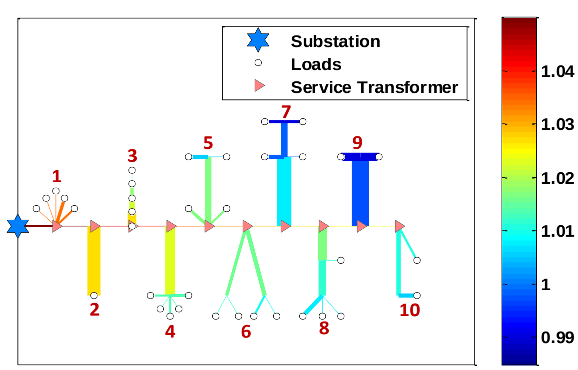

Figure 7. Three-phase 66-node test circuit contouring showing per-unit voltages and line widths showing current magnitudes

A. 3PTC Results with Full AMI Data

The 3PTC secondary circuit impedances were estimated with the DSPE algorithm both with 8759 measurement samples, first without and then with a practical level of 1% P, 1% Q, and 0.2% V measurement error. The average (longest) parameter estimation execution time for the 10 secondary circuits was 0.13 seconds (0.33 seconds). The average and maximum errors of the estimated R and X parameters without and with the measurement error are summarized in Table I. Without measurement error, all the parameters are estimated with a very low relative error. With the measurement error, the average errors of the estimated R and X parameters are still low while some parameters have higher relative errors.

TABLE I. THE AVERAGE RELATIVE ERRORS OF THE ESTIMATED R AND X

Meas. Error?	$R_{err,avg}$ [%]	$X_{err,avg}$ [%]	$R_{err,max}$ [%]	$X_{err,max}$ [%]
No	0.69	0.60	2.64	2.78
Yes	2.46	3.11	13.67	13.18

Figure 8 shows the relative errors of the line R and X parameters estimated with 8759 historical measurement samples without and with measurement error. Excluding parameters of L3-4 and L9-2, all the parameters are estimated with a reasonably good accuracy. Due to space constraints, Figure 7 omits similar transformer parameter results that can be found in [10].

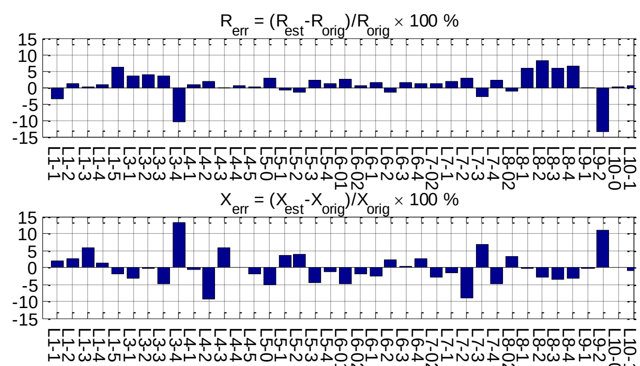

Figure 8. Relative Errors of line R and X parameters estimated with 8759 measurement samples with 1% P, 1% Q, and 0.2% V measurement error

Despite the higher relative errors of the estimated parameters of L3-4 and L9-2, the absolute errors of the estimated parameters of these components are not considerably higher than those of other parameters [10]. The higher relative errors of L3-4 and L9-2 estimated parameters can be explained by the relatively small impedance of branch L3-4 and the relatively high X/R-ratio of branch L9-2 compared to the other branches in the test circuit. The higher relative errors of these parameters may also be explained by these branches' downstream load characteristics. However as discussed in [10], there is no general metric or rule for identifying poor estimates of unknown parameters. Insignificant parameter p-values are a clear indicator of poor regression model fit but also inaccurate parameter estimates can have significant p-values. On the other hand, even highly accurate parameter estimates can have low R-squared values.

Similarly, other typical regression model quality metrics do not seem to accurately identify inaccurate parameter estimates [10].

Figure 9 shows the relative errors of the simulated per-unit voltage drops from the transformer primary winding to the load buses. The errors are calculated between the voltages simulated with the true parameters and the voltages simulated with the estimated parameters. In both cases, the voltages were simulated with the true P and Q values. All the errors are so small that in real circuits, they can be hard to distinguish from measurement noise and other modeling inconsistencies. In particular, the higher relative errors of estimated parameters of branches L3-4 and L9-2 do not result in considerably higher errors of simulated branch downstream buses 3-4 and 9-2.

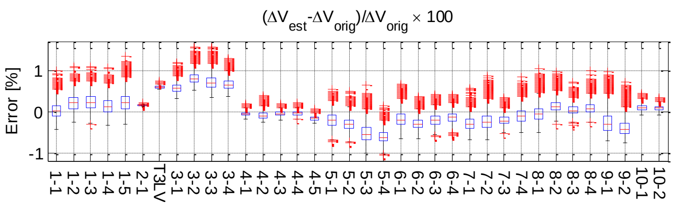

Figure 9. Boxplots of the errors of simulated voltage drops from the service transformer primary (medium-voltage) to the load buses when the parameters are estimated with 8759 measurement samples with 1% P, 1% Q, and 0.2% measurement error

VII. Utility Feeder Results

This section presents the results for a model of urban 12kV 7km-long California utility feeder that serves 3800 mainly residential customers and has the peak load of 8MW. The full OpenDSS feeder model has 5057 buses (6818 nodes), 5070 lines, 1 substation LTC, 324 service transformers, 3785 loads, and 36 PV systems each in a distinct secondary system. The feeder model was reduced using the approach shown in [20] down to 685 buses, 645 lines, 1 substation LTC, 36 service transformers, 725 loads, and 36 PV systems.

The available feeder measurement data consists of 12-months of 15-min substation SCADA active power measurements that have been allocated to the loads based on service transformer sizes. Additionally, the 36 PV systems are assumed to provide hourly voltage, active power, and reactive power measurements. The goal was to analyze the potential of improving the feeder voltage simulation accuracy at these PV systems by generating simplified secondary circuit models by utilizing the PV system measurements. Next, the accuracy of the DSPE algorithm is first shown when full AMI data is available and then, SDSPE algorithm is utilized to generate simplified secondary circuit models based on the PV system measurements.

A. Full Secondary Circuit Models with Full AMI Data

First, perfect active power, reactive power, and voltage measurements were assumed to be available at all loads and the primary system was assumed to be perfectly modeled providing an accurate estimate of the primary side voltages of the service transformers. With these assumptions, the line and transformer impedance parameters of the 36 full (not simplified) secondary circuit models with PV systems were estimated with the DSPE algorithm using 744 measurement samples without measurement error. The average (longest) parameter estimation execution time for the 36 secondary circuits was 0.17 seconds (0.37 seconds). The average and maximum absolute relative errors of the estimated line and transformer parameters are shown in TABLE II. Figure 10 shows the relative errors of the estimated line and transformer R and X parameters. Clearly, the DSPE algorithm estimates all the parameters with a very good accuracy.

TABLE II . The Average and Maximum Absolute Relative Errors of the Estimated R and X

	$R_{err,avg}$ [%]	$X_{err,avg}$ [%]	$R_{err,max}$ [%]	$X_{err,max}$ [%]
Lines	0.004	0.027	0.029	0.309
Transformers	0.132	0.283	0.551	1.052

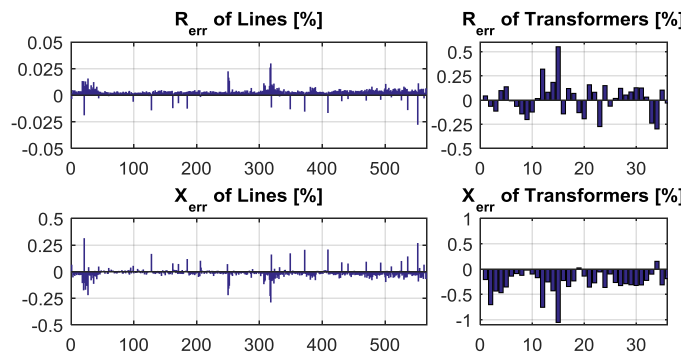

Figure 10. The relative errors of the line R (top left), line X (bottom left), transformer R (top right), and transformer X (bottom right) parameters estimated with 744 samples without measurement noise. Full secondary circuit models with fully available AMI measurements.

B. Simple Secondary Circuit Models with Full AMI Data

Next, the models of the 36 secondary circuits with PV systems were converted to the simple format shown in Figure 4 where each secondary circuit had its original transformer, one generic service drop to the PV system (and the load at the PV system), and the rest of the secondary circuit loads were lumped at the service transformer secondary. All the loads were assigned the total feeder active power profile with a constant power factor. Then, the simple secondary circuit transformer resistance R and the PV system service drop length l from (27) were estimated with the SDSPE algorithm assuming that the load allocation was perfect, i.e., all loads follow the substation profile exactly. The transformer X/R-ratio and the line per-unit-length impedances r and x were also assumed to be perfectly known. The average (longest) parameter estimation execution time for the 36 secondary circuits was 0.07 seconds (0.24 seconds). The average (maximum) error of the estimated line length and transformer R parameters were 0.343% (2.347%) and 0.530% (1.394%), respectively. The relative errors of the line lengths and transformer resistances estimated with 744 samples without measurement error are shown in Figure 11.

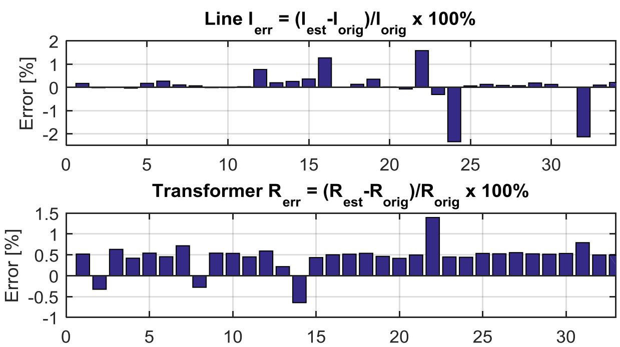

Figure 11. The relative errors of the line length (top) and transformer R (bottom) parameters estimated with 744 samples without measurement noise. Simple secondary circuits with loads modeled through perfect load allocation.

The higher overall errors in Figure 11 compared to Figure 10 can be mainly explained by the smooth profile of the allocated loads and the poorer condition of regression problems (high $\boldsymbol{I}_0/\boldsymbol{I}_1$, high correlation of $\boldsymbol{I}_0$ and $\boldsymbol{I}_1$). Since the parameter estimation predictor variables between different secondary circuits differ only by the PV current injection, the parameter estimation errors are very similar between secondary circuits. All transformer R parameters are estimated with an error less than 1% except for one secondary circuit where the difference of PV generation and its load were too small compared to the load at the transformer secondary leading to almost perfectly collinear predictor terms $\boldsymbol{I}_0$ and $\boldsymbol{I}_1$. All line parameters are estimated with an error less than 1% except for four lines that similar to the transformers, have highly correlated predictor terms $\boldsymbol{I}_0$ and $\boldsymbol{I}_1$ due to small PV generation and load relative to the other loads at the transformer secondary. Overall, these results indicate the theoretical feasibility of the SDSPE algorithm given that accurate load profiles or measurements are available.

C. Utility Feeder Results with Limited PV Measurements

Finally, the simple secondary circuit transformer resistance R and the PV system service drop length l from (27) were estimated with the SDSPE algorithm assuming that the loads are modeled through feeder total active power profile allocated to the loads based on service transformer rating. The average (longest) parameter estimation execution time for the 36 secondary circuits was 0.20 seconds (1.43 seconds). The average (maximum) error of the estimated line length and transformer R parameters were 45.94% (318.0%) and 60.63% (91.05%), respectively. The estimated line length and transformer R parameters are shown in Figure 12. In order to force the parameters to remain positive, the linearly constrained least squares estimation (8) was utilized to estimate the parameters in several secondary circuits.

These results show that such typical load allocation approach does not sufficiently capture the load characteristics in individual secondary circuits and loads to be useful for distribution system parameter estimation. Modeling secondary circuit loads through load allocation, which is based on substation SCADA measurements and service transformer rating (or similar metric), significantly simplifies the impacts on the secondary circuit level [21], [22]. In general, using load allocation tends to underestimate the voltage drops and losses in the secondary circuits. The response variable $\boldsymbol{V}_0 - \boldsymbol{V}_{PV}$ in the simplified secondary circuit parameter estimation (27)-(29) depends on the measured PV voltages $\boldsymbol{V}_{PV}$ and $\boldsymbol{V}_0$, the simulated (with allocated loads) service transformer medium-voltages referred to the secondary. Due to the load allocation, $\boldsymbol{V}_0$ tends to have a very smooth profile over time and tends to overestimate the true voltages (since the load allocation underestimates the secondary circuit voltage drops). On the other hand, $\boldsymbol{V}_{PV}$ can have a highly varying profile. As a result, the estimated voltage drop $\boldsymbol{V}_0 - \boldsymbol{V}_{PV}$ tends to vary more than it does in reality. Similarly, due to load allocation modeled loads, the predictor term $\boldsymbol{I}_0 = \boldsymbol{I}_{R0} + (X/R)_0\boldsymbol{I}_{X0}$ tends to have a relatively smooth profile compared to the predictor term $\boldsymbol{I}_1 = r_1\boldsymbol{I}_{R1} + x_1\boldsymbol{I}_{X1}$, which varies much more over time due to the varying PV generation. Since the response variable $\boldsymbol{V}_0 - \boldsymbol{V}_{PV}$ is better correlated with the $\boldsymbol{I}_1$ than with the predictor variable $\boldsymbol{I}_0$, the line length l_1 tends to be overestimated and the transformer parameter R_0 tends to be underestimated.

While previous research [10] has shown that parameter estimation can be achieved when certain meters do not report voltage measurements, these results demonstrate that generic load allocation from substation data cannot be used for parameter estimation and that all injection points in the secondary network should have meters, such as AMI, in order to estimate the impedances.

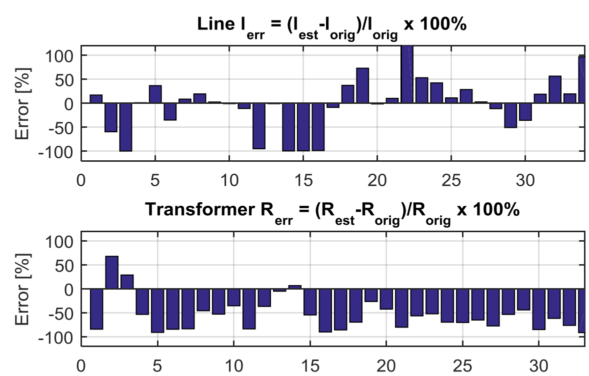

Figure 12. The relative errors of the line length (top) and transformer R (bottom) parameters estimated with 744 samples without measurement noise. Simple secondary circuits with loads modeled through imperfect load allocation.

VIII. Conclusions

This paper presents practical computationally efficient distribution system parameter estimation methods to improve the accuracy of secondary circuit parameters in existing utility feeder models. On average, both methods are executed in a fraction of a second per secondary circuit even when thousands of measurement samples are leveraged to counteract the accuracy, granularity, and time-synchronization issues related to AMI and DER measurements.

The first presented method accurately estimates the transformer and line series impedance parameters in 3-phase and 1-phase secondary circuits when full AMI active power, reactive power, and voltage measurements are available. The method can also handle some measurement error in the meters and conditions where some meters do not report voltage measurements.

The second method presented in this paper can be used for generating simplified secondary circuit models based on limited available PV system (or other sensor) power and voltage measurements. When accurate load profiles are available, the method accurately estimates the transformer impedance magnitude and the PV service drop length. However, loads modeled through conventional service transformer rating based load allocation are not sufficiently accurate to be used for parameter estimation.

The future work will utilize the proposed methods for parameter estimation of real utility feeder models with AMI and DER measurements.

IX. References

[1] "Report on Use of Distribution State Estimation Results for Distribution Network Automation Functions, D2.2.1," BU, EF, Fraunhofer IWES, INDRA, Korona, FP7 - 248135, Apr. 2011.

[2] Y.-F. Huang, S. Werner, J. Huang, N. Kashyap, and V. Gupta, "State estimation in electric power grids: meeting new challenges presented by the requirements of the future grid," *IEEE Signal Process. Mag.*, vol. 29, no. 5, pp. 33–43, Sep. 2012.

[3] A. Abur, *Power system state estimation: theory and implementation*. New York, NY: Marcel Dekker, 2004.

[4] P. Zarco and A. Gomez-Exposito, "Power system parameter estimation: a survey," *IEEE Trans. Power Syst.*, vol. 15, no. 1, pp. 216–222, Feb. 2000.

[5] M. R. M. Castillo, J. B. A. London, and N. G. Bretas, "Network branch parameter validation based on a decoupled State/Parameter Estimator and historical data," in *PowerTech, IEEE Bucharest*, 2009.

[6] M. R. M. Castillo, J. B. A. London, and N. G. Bretas, "An approach to power system branch parameter estimation," in *IEEE Electric Power Conference*, Vancouver, BC, Canada, 2008, pp. 1–5.

[7] A. J. Berrisford, "A tale of two transformers: An algorithm for estimating distribution secondary electric parameters using smart meter data," in *Annual IEEE Canadian Conference on Electrical and Computer Engineering*, Regina, SK, Canada, 2013.

[8] B. Wang and W. Luan, "Generate distribution circuit model using AMI data," in *China International Conference on Electricity Distribution*, 2014, pp. 1251–1255.

[9] T. A. Short, "Advanced metering for phase identification, transformer identification, and secondary modeling," *IEEE Trans. Smart Grid*, vol. 4, no. 2, pp. 651–658, 2013.

[10] J. Peppanen, M. J. Reno, R. Broderick, and S. Grijalva, "Distribution System Secondary Circuit Parameter Estimation for Model Calibration," Sandia National Laboratories, SAND2015-7477, Sep. 2015.

[11] J. Peppanen, M. J. Reno, M. Thakkar, S. Grijalva, and R. G. Harley, "Leveraging AMI Data for Distribution System Model Calibration and Situational Awareness," *IEEE Trans. Smart Grid*, 2015.

[12] J. Peppanen, M. J. Reno, R. Broderick, and S. Grijalva, "Distribution System Low-Voltage Circuit Topology Estimation using Smart Metering Data," in *IEEE PES T&D Conference & Exposition*, Dallas, TX, USA, 2016, pp. 1–5.

[13] T. Gönen, *Electric power distribution system engineering*, 2nd ed. Boca Raton: CRC Press, 2008.

[14] C. W. Brice, "Comparison of Approximate and Exact Voltage Drop Calculations for Distribution Lines," *IEEE Trans. Power Appar. Syst.*, vol. PAS-101, no. 11, pp. 4428–4431, 1982.

[15] W. H. Kersting, "Center tapped transformer and 120/240 volt secondary models," in *2008 IEEE Rural Electric Power Conference*, 2008, pp. A1–A1–7.

[16] M. H. Kutner, C. Nachtsheim, J. Neter, and W. Li, *Applied linear statistical models*. Boston: McGraw-Hill Irwin, 2005.

[17] M. J. Reno and K. Coogan, "Grid Integrated Distributed PV (GridPV) Version 2," Sandia National Laboratories, SAND2014-20141, 2014.

[18] "Open Distribution System Simulator," *EPRI*. [Online]. Available: http://sourceforge.net/projects/electricdss/.

[19] "Pecan Street Inc.," *Pecan Street Inc.* [Online]. Available: http://www.pecanstreet.org/about/.

[20] M. J. Reno, K. Coogan, R. Broderick, and S. Grijalva, "Reduction of distribution feeders for simplified PV impact studies," in *Photovoltaic Specialists Conference*, Tampa Bay, FL, USA, 2013, pp. 2337–2342.

[21] J. A. Taylor, T. A. Short, and B. Bushey, "Efficiency impacts of distribution secondaries," in *IEEE Transmission and Distribution Conference and Exposition*, Orlando, FL, USA, 2012, pp. 1–6.

[22] H. L. Willis, *Power distribution planning reference book*, 2nd ed., rev. and expanded. New York: M. Dekker, 2004.

X. Biographies

Jouni Peppanen (S'11) received M.Sc. degree in electrical engineering both from the Berlin University of Technology in 2010 and from the Georgia Institute of Technology in 2013 and a M.Sc. degree in operations research from Georgia Institute of Technology in 2015. He is currently working toward his Ph.D. in electrical engineering at Georgia Institute of Technology. His research includes leveraging ubiquitous sensors for improved distribution system situational awareness.

Matthew J. Reno (M'10) received his Ph.D. in electrical engineering from Georgia Institute of Technology in 2015. He is a Senior Member of Technical Staff at Sandia National Laboratories, working in the Electric Power Systems Research Department. His research includes modeling solar variability and the impact to the operation of the electric distribution system.

Robert J. Broderick Robert J. Broderick (M '03) received the M.Sc. degree in Power Systems Engineering from New Mexico State University. He is currently the Technical Lead of the Distributed Grid Integration Programs at Sandia National Labs in Albuquerque, New Mexico. Robert's research focus is on removing barriers to enable greater integration of solar energy into the electricity grid. Prior to working for Sandia, Robert worked as a consultant for TRC Engineers, Inc., as a manager of renewables at PNM (largest IOU utility in New Mexico), and as a senior power engineer in the Distribution Planning and Customer Generation Departments at PNM. Robert took a lead role in writing New Mexico's new interconnection standard utilizing industry best practices and IEEE 1547.

Santiago Grijalva (M'02 – SM'08) is Associate Professor of Electrical and Computer Engineering and Director of the Advanced Computational Electricity Systems (ACES) Laboratory at The Georgia Institute of Technology. Prof. Grijalva's research interests are on power system control, power system informatics, power system economics, and smart grids. He currently directs various leading projects on future electricity grid computation for DOE, PSERC, ARPA-E and EPRI. From 2002 to 2009 he was with PowerWorld Corporation as a senior software architect and consultant. Dr. Grijalva obtained the Electrical Engineer degree in from EPN-Ecuador (94); M.S. Certificate in Information Systems from ESPE-Ecuador (97); M.Sc.E.E. (99) and Ph.D.E.E. (02) from the University of Illinois at Urbana-Champaign. From 2003 to 2004 he was a post-doctoral fellow in Energy Systems at the University of Illinois.