SANDIA REPORT

SAND2015-7477
Unlimited Release
Printed September 2015

Distribution System Secondary Circuit Parameter Estimation for Model Calibration

Jouni Peppanen, Matthew J. Reno, Robert J. Broderick, Santiago Grijalva

Prepared by
Sandia National Laboratories
Albuquerque, New Mexico 87185 and Livermore, California 94550

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.

Sandia National Laboratories

Issued by Sandia National Laboratories, operated for the United States Department of Energy by Sandia Corporation.

Available to DOE and DOE contractors from
U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831

Telephone: (865) 576-8401
Facsimile: (865) 576-5728
E-Mail: reports@adonis.osti.gov
Online ordering: http://www.osti.gov/bridge

Available to the public from
U.S. Department of Commerce
National Technical Information Service
5285 Port Royal Rd.
Springfield, VA 22161

Telephone: (800) 553-6847
Facsimile: (703) 605-6900
E-Mail: orders@ntis.fedworld.gov
Online order: http://www.ntis.gov/help/ordermethods.asp?loc=7-4-0#online

SAND2015-7477
Unlimited Release
Printed September 2015

Distribution System Secondary Circuit Parameter Estimation for Model Calibration

Jouni Peppanen, Santiago Grijalva
School of Electrical and Computer Engineering
Georgia Institute of Technology
777 Atlantic Drive NW
Atlanta, GA 30332-0250

Matthew J. Reno, Robert J. Broderick
Photovoltaics and Distributed Systems Integration
Sandia National Laboratories
P.O. Box 5800
Albuquerque, New Mexico 87185-1033

Abstract

To analyze and coordinate the operation of distribution systems with rapidly increasing amounts of PV, more accurate distribution system models are required, especially for the distribution system secondary (low-voltage) circuits down to the point of common coupling for distributed PV. There is a growing need for automated procedures to calibrate the distribution system secondary circuit models that are typically either not modeled at all or are modeled with a lower level of detail than the better modeled medium-voltage systems. This report presents an accurate, flexible, and computationally efficient method to use measurement data to estimate secondary circuit series impedance parameters in existing utility feeder models. The parameter estimation method assumes well-modeled primary circuit models, known secondary circuit topologies, and AMI active power, and reactive power measurements at all the loads in the secondary circuit. The method also requires AMI voltage measurement at most of the loads in the secondary circuit but can handle loads that do not have voltage measurements. No existing secondary circuit model information is needed, except for topology. The method is based on the well-known linearized voltage drop approximation and linear regression. The performance of the method is demonstrated on a three-phase test circuit with ten different secondary circuit topologies and on the Georgia Tech campus distribution system with AMI data. The developed method can be utilized to improve existing utility feeder models for more accurate analysis and operation with ubiquitous distributed PV interconnected on the low-voltage circuits.

CONTENTS

FIGURES

TABLES

NOMENCLATURE

ANSI	American National Standards Institute
AMI	advanced metering infrastructure
CVR	conservation voltage reduction
DER	distributed energy resources
DOE	Department of Energy
DR	demand response
DSPE	distribution system parameter estimation
DSSE	distribution system state estimator
FLISR	fault location, isolation and service restoration
GIS	geographical information system
GW	gigawatt (AC)
IEA	International Energy Agency
kW	kilowatts (AC)
kVAr	kilovolt-amperes reactive (AC)
LAV	least absolute value estimator
LP	linear programming problem
MAE	mean absolute error
MBE	mean bias error
MDMS	meter data management system
MVUE	minimum variance unbiased estimator
OLS	ordinary least squares estimator
OpenDSS	Open Distribution System Simulator™
PE	parameter estimation
PF	power factor
PSO	particle swarm optimization
pu	per unit
PV	photovoltaic
TSPE	transmission system parameter estimation

1. INTRODUCTION

1.1. Enhanced Distribution System Models for High PV Penetrations

Driven by falling costs and government incentive programs, solar photovoltaics (PV) has experienced exponential growth rates, as shown by the International Energy Agency (IEA) statistics on the top of Figure 1. Assuming strong PV growth rates continue, IEA high renewable (hi-Ren) scenario expects PV to make up to 16% of the global electricity supply by 2050 [1].

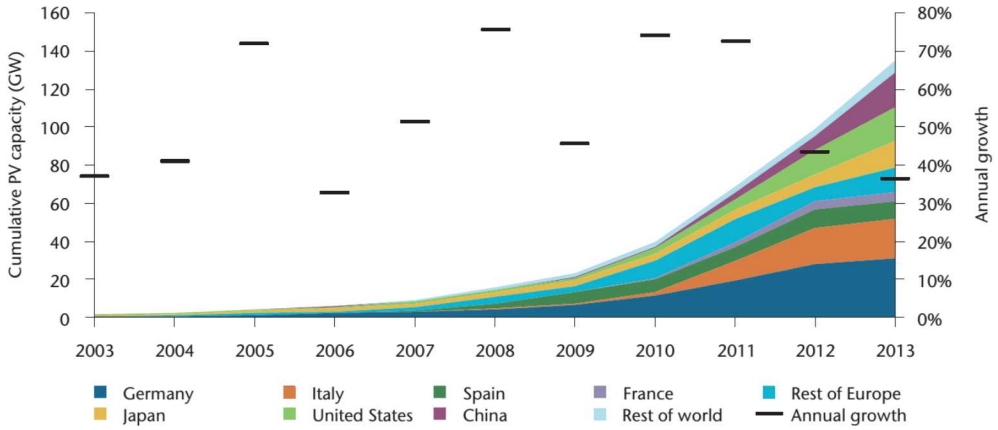

Figure 1. Rapid increase of distributed PV: global cumulative growth of PV capacity [1]

In February 2014, the total U.S. solar electric capacity was 12.1 GW. In 2014 , the U.S. PV capacity grew by 6.2 GW, growth 30% higher than in 2013 and over 12 times higher than five years earlier [2]. With the expected annual growth rate of 6.8% between 2013 and 2040, PV is expected to expand faster than any other source of renewable energy [3, p. 81]. The experienced past growth and expected future growth are shown in Figure 2.

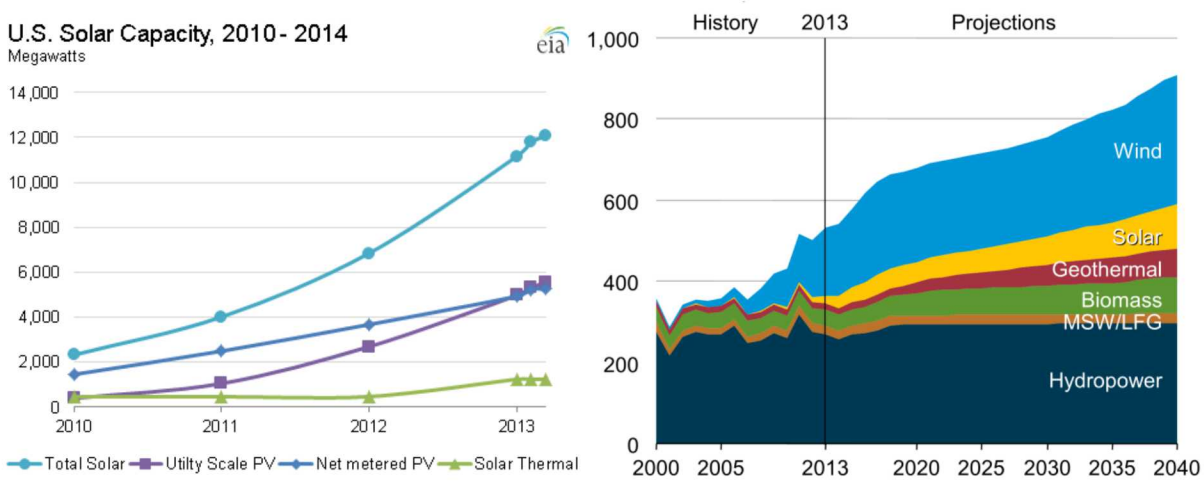

Figure 2. Past and future PV growth in the U.S.[4]

High PV penetration levels are already seen in some geographic areas with high solar resources and supportive politics. In Hawaii, the national leader of customer PV penetration, there are already circuits where the installed PV capacity exceeds 75% of the daytime peak load and 250% the daytime minimum load, as shown in Figure 3. California leads the nation in both number of PV installations with over 230,000 and with total installed PV capacity at almost 10 GW [5], [6].

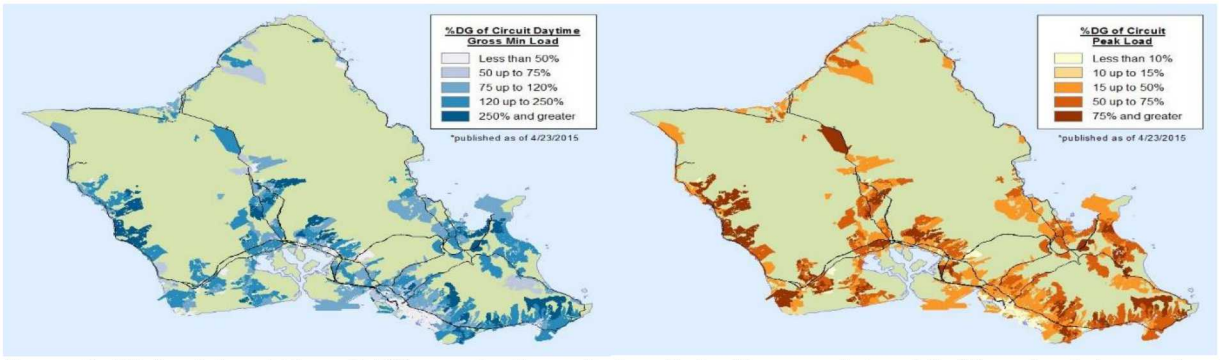

Figure 3. O'ahu island Hawaii PV penetration of circuit daytime peak load (left) and minimum load (right) [4]

An increasing share of PV will be located in distribution systems where it raises concerns of maintaining feeder operation within component loading and voltage standard limits [7], [8]. In order to maintain economic, high-quality, reliable, and safe distribution system operation under pervasive PV, faster and more accurate monitoring, coordination and control is imperative [9]. Much of distribution system operation is based on the assumption that the system analysis is correct, which means that the system models and the input data to those models must be very accurate. Currently, neither is. Instead, the models are often outdated and inaccurate, and measurement data is typically not properly integrated to be fully leveraged. However, emerging data and new sensors have the potential to provide enough information to support the new operational needs.

Simultaneously, efficiency, quality, reliability, and safety requirements drive smart distribution automation schemes, including advanced Volt/VAr optimization, conservation voltage reduction (CVR), demand response (DR), and fault location, isolation and service restoration (FLISR). Smart distribution automation schemes require more accurate and reliable situational awareness [9]–[11], which is increasingly being provided by modern distribution system measurement sources such as smart meters and PV micro inverters [12]. Accurate and robust use of all available measurement information, as well as accurate distribution system models, will be essential for future distribution system state estimation (DSSE). DSSE is envisioned to become the cornerstone of the monitoring and coordination of future smart distribution system with ubiquitous PV and advanced distribution automation functions [13], [14].

Much of the distribution system analysis and operation is based on the assumption that the models used to run steady-state simulation are accurate. Circuit models, including the parameter values, may be incorrect as a result of data entry errors, inaccurate equipment data, network changes (e.g. phase balancing), incorrect tap information, etc. [15]. The most commonly encountered errors in the distribution system Geographical Information System (GIS) and power flow models include incorrect component parameters, customers modelled connected to the

wrong distribution transformer, and distribution transformers modelled on the wrong phases of feeders [16]. Improving the accuracy of feeder parameters becomes critical as the numbers of DERs increases. DERs make it more challenging to operate the feeder within ANSI limits while optimally managing advanced applications such as Volt/VAr optimization, protections, and CVR schemes.

It is particularly important to improve the models of the distribution system secondary (low voltage) networks where a large share of the new controllable devices, such as electric vehicles, PV with smart inverters, and demand response, are located. The secondary networks are typically either not modeled at all or modeled with a low level of detail although a significant portion of per-unit voltage drop/raise occurs over the high impedance service transformers and low voltage lines with high losses.

The typical approaches to correct circuit errors, such as performing physical inspections or utilizing added measurements, require considerable man hours and additional resources and thus, are not cost-effective [16]. Physical inspections can also be hard to perform in densely populated urban areas with wiring underground and in buildings. There is a growing need for automated procedures to improve the accuracy of distribution system including secondary circuit models with minimal physical inspections.

1.2. Parameter Estimation

The objective of parameter estimation (PE) is to find the most likely component parameters that are typically known with varying levels of accuracy [15], [17], [18]. The parameter estimation problem is closely related to topology estimation, which has the goal of identifying the most likely system topology. Because the number of possible topologies and parameter value combinations can be very large, parameter and topology estimation should not be seen as an optional process to obtain a good initial system modeling [18], but rather as a necessary step to calibrate and verify the accuracy of existing utility models. It is not advisable to attempt to estimate parameters using a tolerance accuracy smaller than the average measurement error. The estimated parameter errors are proportional to the average measurement error, and in the worst case, the presence of measurement noise can result in replacing rather accurate original parameter values by less accurate estimated values [17].

Line and transformer parameters can be assumed to be time invariant and can be estimated off-line, whereas load tap changing transformer tap positions change over time and require online PE [17], [19]. The local measurement redundancy and robustness of offline parameter estimation is increased by utilizing historical databases of measurement data that can be selected free of gross and topological errors [17], [19]. Additionally, offline PE requires no modifications to the existing online algorithms [17].

1.2.1. Transmission System Parameter Estimation

Before describing distribution system parameter estimation, let us briefly review methods for transmission system parameter estimation, which have been studied since the 1970s. These algorithms are typically integrated with the state estimation algorithm and are based either on

residual sensitivity analysis or augmented state vectors [15], [19]. In the former type, PE is performed after state estimation by utilizing linear sensitivities between the parameter errors and measurement residuals. In the latter, the typical state vector is augmented with additional variables that represent suspicious parameters. The augmented state vector methods apply either normal equations or Kalman filter theory. The augmented state vector methods have surpassed the residual methods, which however, are important for identifying suspicious parameters [17], [19].

Parameter and measurement gross errors are often harder to identify in state estimation than topology errors and thus, may go unnoticed for longer periods of time [15], [17], [19]. Topology errors cause several normalized residuals to violate a specified threshold in state estimation algorithm. These residuals correspond to measurements close to the topology error. Similar phenomena is observed with gross measurement errors because of the so-called "smearing effect" that can make it challenging to distinguish between topology errors and gross measurement errors [20]. Since erroneous network parameters have a relatively local impact on the state estimation results, parameter estimation can be performed in a local manner. Accurate measurements typically help in identifying parameter errors [17].

1.2.2. Distribution System Parameter Estimation

Compared to the well-established transmission system parameter estimation, distribution system parameter estimation (DSPE) is subject to a number of different challenges. First, multi-phase asymmetric, radial distribution systems with unbalanced loads, low X/R-ratios, and various connections of transformers and loads make distribution system models complicated and different from transmission system models [21]. Second, most utilities do not have existing distribution system state estimators and thus, most conventional transmission system parameter estimation approaches that are integrated in the state estimator are not directly applicable. Finally, the low number and quality of measurements in distribution systems results in a low measurement redundancy or even the lack of observability in certain circuit sections. For these reasons, DSPE has been studied less than transmission system parameter estimation but is becoming more possible to implement using advanced metering infrastructure (AMI) and other modern distribution system measurement sources [22]–[28]. However, compared to transmission system parameter estimation, DSPE algorithm needs to operate without existing state estimator and manage complicated distribution system models and the lower redundancy, reliability, accuracy, and granularity of the modern distribution system measurements.

There has been some previous work on DSPE. A linear optimization-based method for topology error detection, parameter estimation, and theft identification has been proposed in [29]. The authors did not estimate the reactances or leveraged the reactive power measurements. Topology error detection regarding smart meter placement in GIS system is introduced in [16]. In [30], the author presents a method for meter phase identification and meter-to-transformer mapping by applying a voltage drop equation and linear regression with AMI energy and voltage measurements. In [31], the authors assume a known radial network topology and derive a quadratic equation between the smart meter measurements and upstream bus voltage. Then, utilizing this equation, the authors estimated branch parameters using a gradient-based approach with the objective of minimizing the variance of voltage estimates from various smart meters.

The approach makes no simplifications to the AC power flow equations but results in an optimization problem with quadratic equality constraints that is computationally much more intensive to solve than linearized approaches.

This report builds on the approach presented in [30] but presents more detailed analysis and a number of refinements to increase the performance of the method proposed in [30]. This report focuses on impedance parameter estimation of radial secondary circuits with known topologies. In future work, the method will be expanded for estimating impedance parameters of radial circuits with unknown topologies.

1.3. The Structure of This Report

Following Chapter 1, which provides the background and motivation for distribution system secondary circuit parameter estimation, Chapter 2 presents a 2-bus test case and a 3-phase 66-node test circuit that are used throughout this report to test and compare algorithms and methods. Chapter 3 first presents two heuristic methods to estimate series impedance parameters and, motivated by the disadvantages of these methods, proposes an alternative linear regression parameter estimation method. The chapter also introduces three alternative estimators and presents the method that was chosen for this report. Chapter 4 expands the single branch parameter estimation to estimation of all the parameters in a radial secondary circuit. The chapter begins by discussing the problem formulation, assumptions, and principle after which two alternative approaches are presented and compared to estimate the parameters of a secondary circuit subsection with N parallel branches. Chapter 5 provides a detailed comparison of alternative linear regression models and their performance for both tests circuits without and with measurement error. The chapter also presents the method that was chosen for this report. Chapter 6 discusses the implementation of the distribution system secondary circuit algorithm in a practical utility setting. Chapter 7 includes the parameter estimation results for the 3-phase 66-node test circuit with the chosen parameter estimation method with practical levels of measurement error. Chapter 8 extends the proposed parameter estimation method to handle cases where some meters do not report voltage measurements. Finally, Chapter 9 concludes the report.

2. TEST CIRCUITS FOR PARAMETER ESTIMATION

This chapter introduces the circuits that are used throughout the rest of the paper to illustrate parameter estimation methods. A two-bus test circuit and a larger 66-node 3-phase test circuit were generated to develop, test, and compare different parameter estimation methods. While the ultimate parameter estimation objective is to improve the voltage (drop) simulation accuracy in a given utility feeder model, in these test circuits it was possible to measure the parameter estimation accuracy directly as the relative percentage error of the estimated parameters with respect to the known parameters with: $\left(Par_{est} - Par_{orig}\right)/Par_{orig} \times 100\%$. Finally, the real distribution system on Georgia Tech campus is introduced and is used later to demonstrate parameter estimation for an actual system with unknown parameters.

2.1.1. Two-bus Test Circuit

A simple two-bus test circuit was generated to demonstrate the behavior of the different parameter estimation algorithms at a conceptual level. The circuit consists of a single load that is connected to the fixed voltage sources at the upstream bus over a service transformer and a service line. The circuit is shown in Figure 4.

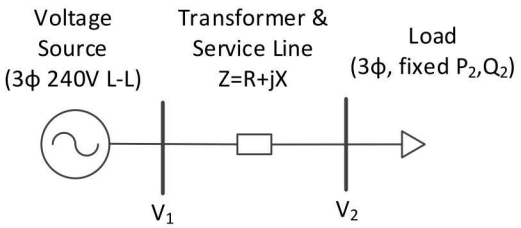

Figure 4. Simple two-bus test circuit

2.1.2. 66-Node Three Phase Test Circuit

A three-phase balanced test circuit model was implemented with a single backbone feeder and ten secondary circuits with different topologies. In each of the topologies, there is a single MV/LV transformer with arbitrarily (but realistically) selected parameters. Each load is connected to a point in the secondary system over a service line as listed in Table 1.

Table 1. Secondary circuits of the 66-node three-phase test circuit

Secondary Circuit Number (Order from the Substation)	Load Connections
1	5 loads connected to the transformer
2	1 large load connected to a pedestal
3	5 loads connected in series on a service line (without separate service drops)
4	5 loads connected to a pedestal
5	2 loads connected to the transformer and 2 loads connected to a pedestal
6	2 separate pedestals each with two loads
7	2 pedestals in series each with 2 loads
8	2 pedestals in series: first with one load, second with 3 loads
9	1 pedestal with two loads
10	1 load connected to the transformer, 1 pedestal with 1 load

An example of the circuit line plot and voltage profile with given load active and reactive power values are given in Figure 5 and Figure 6, respectively.

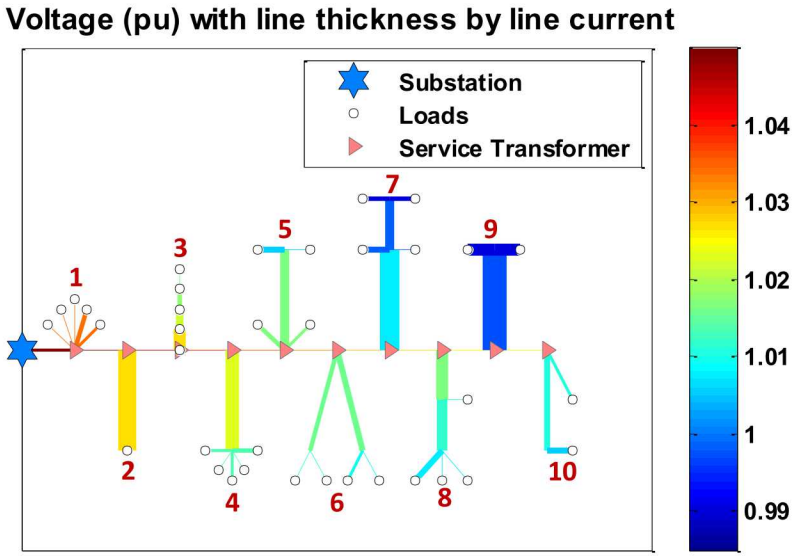

Figure 5. Topology of the 66-node three-phase test circuit (secondary circuit numbers in red)

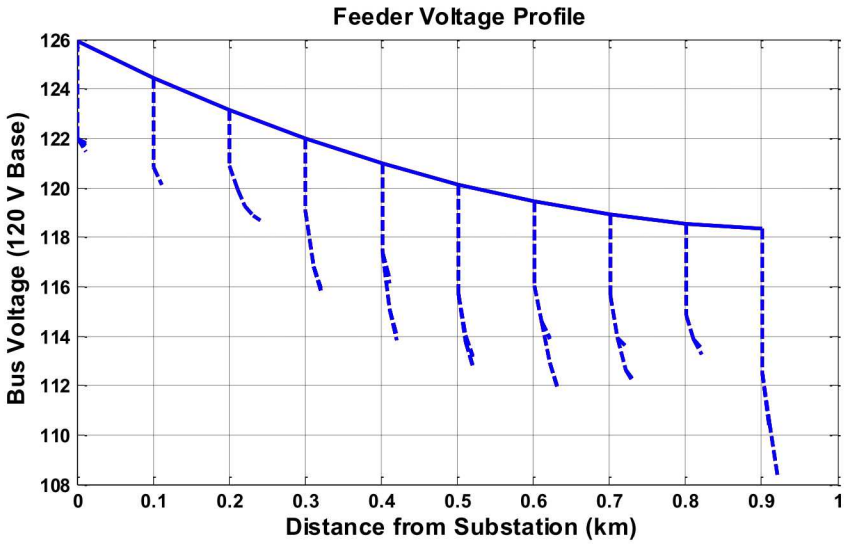

Figure 6. Voltage profile of the 66-node three-phase test circuit

In order to create the profile for each load in this test circuit, a year of 36 distinct residential active power hourly measurement profiles were acquired from Pecan Street Inc. [32]. The total building consumptions (table *Use,* [32]) in year 2014 are utilized. First, each circuit load was assigned a peak kW and kVAr depending on the number of loads in the given secondary circuit and the service transformer kVA rating. Then, each circuit load was randomly assigned one of the Pecan Street load profiles. The load profiles were scaled to selected average load kW. All values exceeding a selected peak load kW were set randomly to 60-100% of the load kW and all negative or zero load values were set to random values 5-15% of load kW. Then, reactive power consumption profiles were created based on the active power profiles utilizing a different

random power factor in the range of 0.9-1.0 for each measurement. The active power, reactive power, and power factor profiles over the first week of data are visualized in Figure 7.

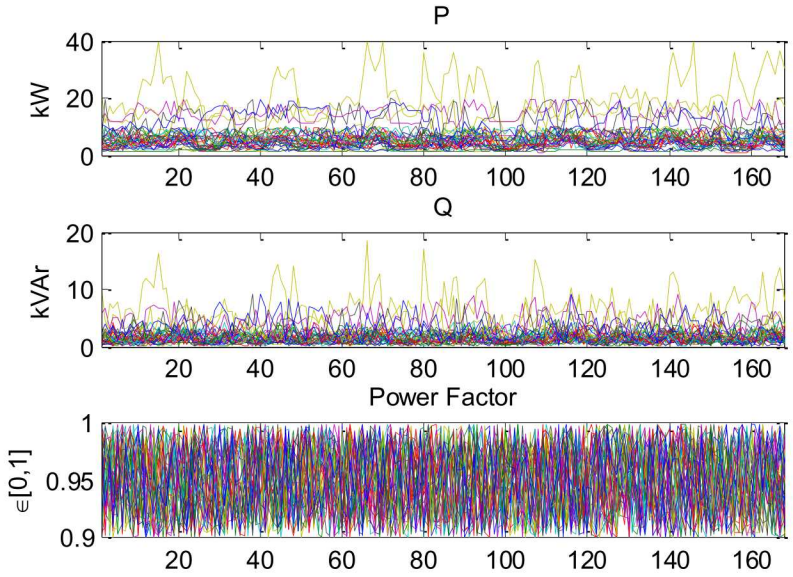

Figure 7. First week of the active power, reactive power and power factor profiles

The load voltage "measurements" were acquired by solving the time series power flow simulation with the loads varying according to their real and reactive power profiles. The resulting voltage "measurements" for the first week in 2014 are shown in Figure 8.

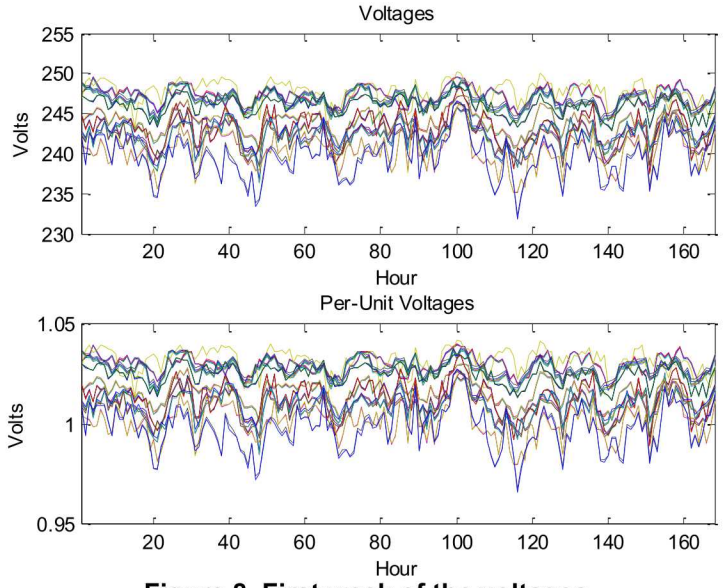

Figure 8. First week of the voltages

2.1.3. Georgia Tech Distribution System Feeder

The distribution system parameter estimation is also demonstrated on one of the feeders of the Georgia Tech campus electricity distribution system. Georgia Tech system is a perfect case study for distribution system parameter estimation since it is well-balanced 3-phase system that has an

extensive AMI including a 15-min historical measurement database with several years of measurements [33], [34]. The 19.8 kV feeder shown in Figure 9 is an underground system that is approximately slightly over 3.5 km long and has a peak load of 0.90 MW. In the base case, standard manufacturer parameters were used for the service transformers, and the unknown secondary cable lengths were assumed to be 100 feet long. The feeder circuit line plot and voltage profile with the base case parameters are shown in Figure 9.

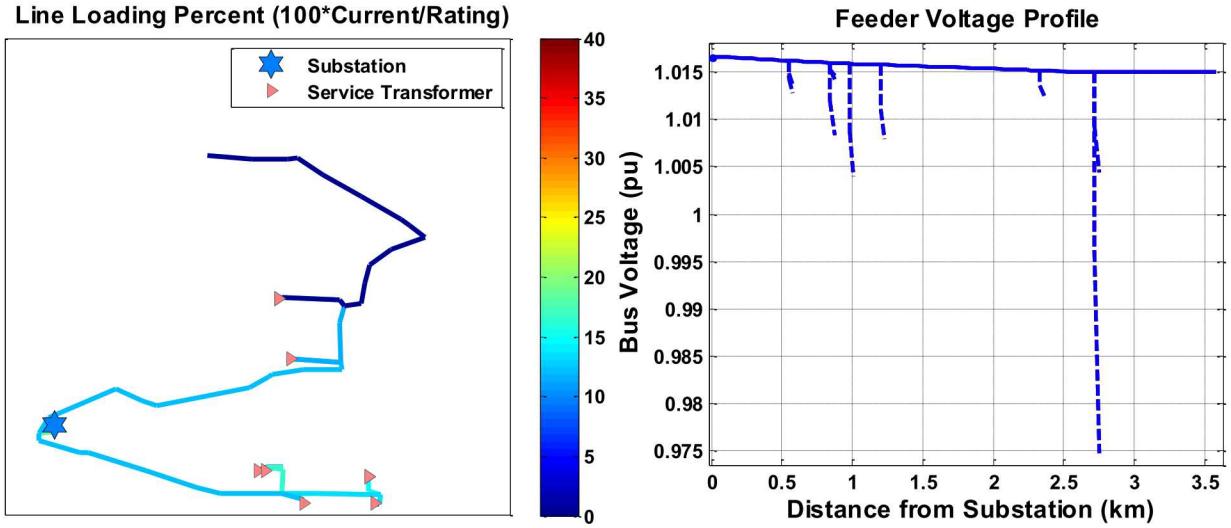

Figure 9. Georgia Tech feeder circuit line plot (left) and voltage profile (right)

3. BRANCH SERIES IMPEDANCE PARAMETER ESTIMATION

This chapter begins by presenting two simple, flexible heuristic approaches to estimate line and transformer series impedance parameters. Motivated by the disadvantages of the heuristic approaches and the challenges of distribution parameter estimation, the rest of the chapter discusses an approach for estimating line and transformer series impedances based on linearized voltage drop approximation and linear regression. Alternative linear regression estimators are also discussed.

3.1. Heuristic Parameter Estimation Approaches

Heuristic parameter estimation approaches do not leverage any parameter estimation problem structure to search for the best parameter estimates. Such non-model based approaches are very simple and flexible to implement, but they tend to be computationally highly demanding since they perform some sort of exhaustive search over the solution space. Two simple heuristic parameter estimation methods, brute force and particle swarm optimization, were implemented and tested on the two-bus test circuit shown in Figure 4.

3.1.1. Brute Force

The brute force algorithm goes through all possible user-set (realistic) parameter (R and X) combinations, and for each parameter combination, it performs a time series power flow analysis to simulate the circuit voltages for that set of parameters. The objective is to find the parameter combination with the smallest mean absolute error of the simulated voltages compared with the measured voltages. The brute force approach requires running a large number of time series power flows, which is computationally very expensive. Moreover, due to the curse of dimensionality, the computational requirements grow exponentially with the number of parameters to be estimated. For example, if a sparse grid of ten R and X values is evaluated, this approach would require evaluating one million (10^6) time series power flows for a simple circuit with three branches and six parameters. If one of the time series power flows was executed every second, it would take over 11 days to evaluate all the combinations. Clearly, the brute force approach is only useful for validating more sophisticated methods in example networks with one to four parameters.

The brute force algorithm was tested by estimating the three-phase test circuit secondary circuit number 2 parameters using one week of 1-hour load data (168 data points). A linearly spaced grid of 100×100 R and X combinations were evaluated with $R \in [0.1,5]$ and $X \in [1,10]$. The algorithm executed the resulting 10,000 time series power flows over the 168 time instances in 290 seconds (4 min, 50s) and returned the parameters $R_{Est} = 0.9919\%$, $X_{Est} = 5.727\%$. Compared to the original parameters $R_{Orig} = 1\%$ and $X_{Orig} = 5.7\%$, the parameter estimation errors were for $R_{err} = -0.81\%$ and $X_{err} = 0.47\%$. The mean absolute load voltage simulation error with the optimal parameters over the time period was $MAE_{Opt} = 0.000598$ V. If no model or measurement error was present, an arbitrary parameter estimation accuracy could be reached by further refining the grid. The mean absolute voltage simulation error dependency on the parameters is visualized in Figure 10. As expected, the voltage simulation error seems to be a

relatively smooth convex function (with an optimal value) with respect to the transformer parameters. However, the objective function is quite flat around the optimum solution, which makes it harder to find the optimal solution.

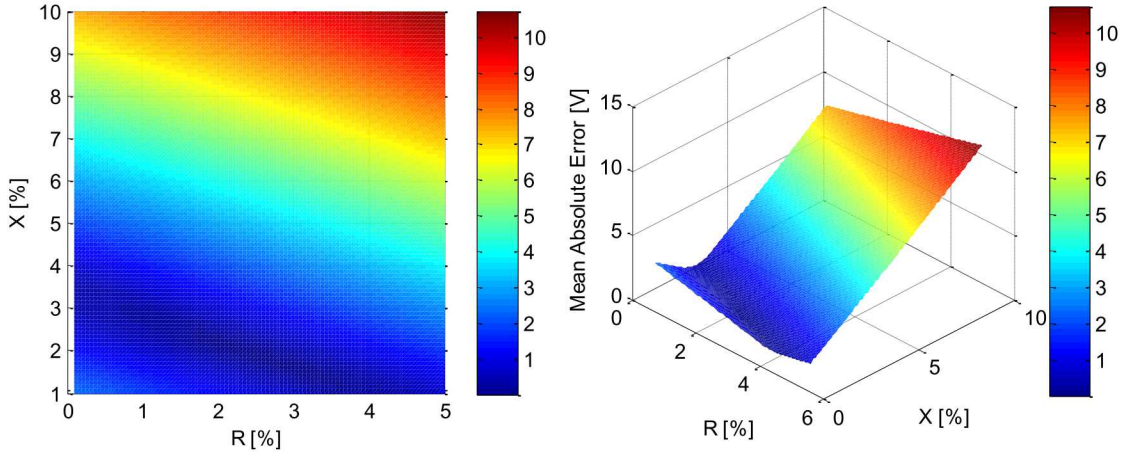

Figure 10. Brute force estimated secondary circuit no. 2 transformer parameters: mean absolute error dependency on R and X

Figure 11 shows the small difference between the measured voltages and the voltages simulated with the estimated parameters.

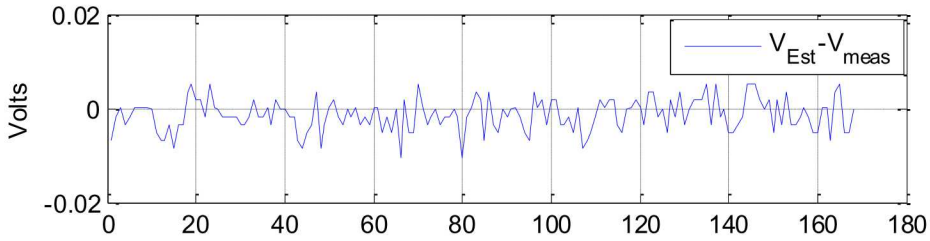

Figure 11. Load voltage simulation accuracy with brute force estimated parameters

3.1.2. Particle Swarm Optimization

Compared to the brute force approach, particle swarm optimization (PSO) is a more intelligent meta-heuristic approach that can be used to find good solutions for nonlinear (mixed-integer) optimization problems [35]–[38]. PSO searches for good solutions by iteratively trying to improve a candidate (i.e. particle) solution with respect to a given measure of quality that is called the particle fitness. Typically, a particle movement is calculated based on the particle best solution, the best solution of all particles, and the current particle movement velocity. PSO is very simple to implement since it does not require knowing any problem-specific structure apart from a way of evaluating the fitness of a given particle solution. On the other hand, since PSO does not leverage problem-specific structure, it is also unable to provide guarantees of converging, even to a locally optimal solution. Moreover, contrary to convex optimization methods, PSO does not tell anything about the quality of the returned best solution compared to the theoretically best solution (i.e., no lower/upper bounds are returned).

A PSO algorithm was implemented to search for the series impedance values for a selected transformer to minimize the mean absolute difference between the measured voltages and the simulated voltages of a selected load. The voltages are simulated with OpenDSS. The parameters estimated in OpenDSS are the winding resistances R_1 and R_2, and the high-to-low inter-winding reactance X_{hl}. The algorithm was tested by estimating the transformer parameters of the 3-phase test circuit secondary circuit number 2. In 60 seconds and 100 PSO iterations, the globally best particle solution was: $R_1 = 0.1\%$, $R_2 = 0.901917\%$ $(R_1 + R_2 = 1.01917\%)$, and $X_{hl} = 5.68623\%$. Compared to the original parameters $R_{Orig} = 1\%$ and $X_{Orig} = 5.7\%$, the parameter estimation errors were for $R_{err} = 0.0192\%$ and $X_{err} = -0.0024\%$. PSO estimated the total transformer winding resistance $R_1 + R_2$ with a good accuracy but was unable to divide the resistance properly to the winding resistances R_1 and R_2. This applies in general that it is impossible to distinguish parameters of series branches from each other since series branches have the same impact on the voltage drop. Therefore, all series parameters must be estimated jointly and manually divided to the respective components by applying some engineering judgment.

Figure 12 illustrates the quick convergence of the PSO algorithm close to the final estimated parameters. Typically, PSO is executed repeatedly several times with different initial particle locations to achieve better exploration of the solution space [35].

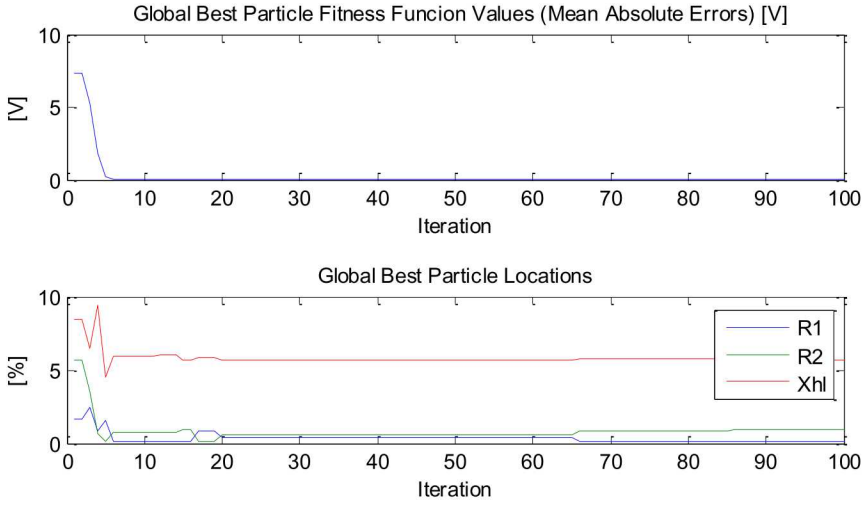

Figure 12. Particle swarm optimization progress over the iterations

Figure 13 shows the small difference between the measured voltages and the voltages simulated with the estimated parameters.

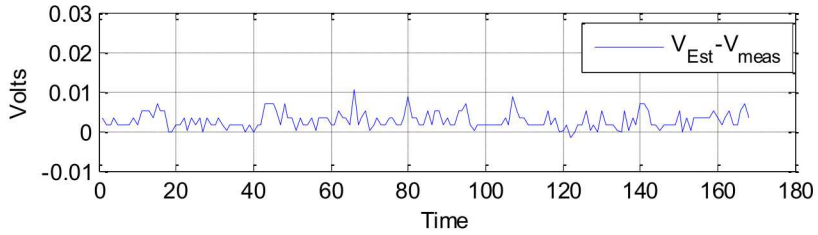

Figure 13. Load voltage simulation accuracy with PSO estimated parameters

3.2. Linearized Voltage Drop Parameter Estimation Approach

Motivated by the disadvantages of non-model based parameter estimation approaches underlined above, this section presents a model-based method to estimate all the distribution system secondary circuit positive sequence series impedance parameters. As introduced in section 1.2, the absence of distribution system state estimator (DSSE), as well as the poor time synchronization, low accuracy, low granularity, and low measurement redundancy of distribution system measurements, makes it challenging to directly apply the conventional transmission system parameter estimation approaches. To overcome these limitations and challenges, a different parameter estimation approach is presented.

The proposed parameter estimation approach leverages the well-known (see e.g. [39]–[41]) linear approximation of voltage drop magnitude over a series impedance

$$V_{drop} = |V_1| - |V_2| \approx (RP + XQ)/V_2 = RI_R + XI_X \tag{1}$$

where R and X are the (positive sequence) series resistance and reactance between bus 1 and bus 2 as shown in Figure 14. The current resistive and reactive components are given with $I_R = P/V_2$ and $I_X = Q/V_2$. In case actual values are used, all the values must be referred to the same voltage level. In 3-phase systems, line-line voltages and 3-phase powers are used.

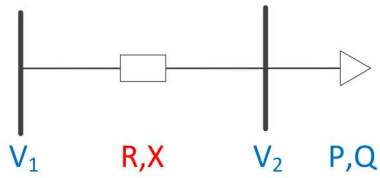

Figure 14. Linearized voltage drop over a series impedance

The goal of the parameter estimation problem is to find the parameters R and X (shown in red in Figure 14) that provide the best fit of the available measurement samples of V_1, V_2, P, and Q (shown in blue in Figure 14) to one of the linear models

$$\boldsymbol{V'} = (\boldsymbol{V}_1 - \boldsymbol{V}_2)\boldsymbol{V}_2 = R\boldsymbol{P} + X\boldsymbol{Q} + \boldsymbol{\epsilon} \tag{2}$$

or

$$\Delta\boldsymbol{V} = \boldsymbol{V}_1 - \boldsymbol{V}_2 = R\boldsymbol{I}_R + X\boldsymbol{I}_X + \boldsymbol{\epsilon}. \tag{3}$$

where the bold letters indicate vectors of K measurement samples, i.e., $\boldsymbol{V}, \boldsymbol{V_1}, \boldsymbol{V_2}, \boldsymbol{P}, \boldsymbol{Q}, \Delta\boldsymbol{V}, \boldsymbol{I_R}, \boldsymbol{I_X} \in \mathbb{R}^K$, and $\boldsymbol{\epsilon} \in \mathbb{R}^K$ represents the joint measurement and model error. In this report, (3) will be preferred over (2) since (3) does not involve multiplying by $\boldsymbol{V}_2$, which itself is an estimate when estimating R and X of the upstream components in the hierarchical radial circuit parameter estimation discussed in sub section 4. It is desirable to avoid dividing by the estimated values of V_2 since this would cause any errors to propagate in a multiplicative manner as opposed to additive.

By denoting $\boldsymbol{y} = \Delta\boldsymbol{V} = \boldsymbol{V}_1 - \boldsymbol{V}_2$, $\boldsymbol{\mathcal{X}} = [\boldsymbol{I}_R \quad \boldsymbol{I}_X]$, and $\boldsymbol{\beta} = [R \quad X]^T$, (3) becomes the linear equation

$$\boldsymbol{y} = \boldsymbol{\mathcal{X}}\boldsymbol{\beta} + \boldsymbol{\epsilon}. \tag{4}$$

An estimate of the parameters $\widehat{\boldsymbol{\beta}}$ can be found by utilizing statistical estimation techniques. The estimated parameter vector $\widehat{\boldsymbol{\beta}}$ will strongly depend on the defined loss function. If $\boldsymbol{L}_p$-norm loss is used, $\widehat{\boldsymbol{\beta}}$ can be found by solving the optimization problem

$$\widehat{\boldsymbol{\beta}} = \min_{\boldsymbol{\beta}} \|\boldsymbol{y} - \boldsymbol{\mathcal{X}}\boldsymbol{\beta}\|_p. \tag{5}$$

Typically, either $\boldsymbol{L}_2$ or $\boldsymbol{L}_1$ are used yielding two different estimators each with different advantages and disadvantages. Next, three typical approaches (ordinary least squares, linearly constrained least squares, and least absolute value) are introduced for solving (5).

3.2.1. Ordinary Least Squares Estimator

Solving (5) using p=2 (i.e. $\boldsymbol{L}_2$-norm) yields the (convex) quadratic programming problem

$$\widehat{\boldsymbol{\beta}} = \min_{\boldsymbol{\beta}} \|\boldsymbol{y} - \boldsymbol{\mathcal{X}}\boldsymbol{\beta}\|_2. \tag{6}$$

Since square preserves convexity, the solution to (6) can be equally obtained by solving

$$\widehat{\boldsymbol{\beta}} = \min_{\boldsymbol{\beta}} \|\boldsymbol{y} - \boldsymbol{\mathcal{X}}\boldsymbol{\beta}\|_2^2 = \min_{\boldsymbol{\beta}} \sum_{t=1}^{T} (y_t - \boldsymbol{x}_t\boldsymbol{\beta})^2 = \min_{\boldsymbol{\beta}} \boldsymbol{\beta}^T\boldsymbol{\mathcal{X}}^T\boldsymbol{\mathcal{X}}\boldsymbol{\beta}. \tag{7}$$

Problem (7) (without any constraints) is a linear regression problem that, assuming $\boldsymbol{\mathcal{X}}$ has full column rank, has the closed-form solution dubbed as the ordinary least squares estimator (OLS)

$$\widehat{\boldsymbol{\beta}} = [\hat{R} \quad \hat{X}]^T = (\boldsymbol{\mathcal{X}}^T\boldsymbol{\mathcal{X}})^{-1}\boldsymbol{\mathcal{X}}^T\boldsymbol{y}. \tag{8}$$

The linear regression problem (7) can be solved with any open source or commercial linear regression package. Next to the parameter estimates, these packages typically provide the user with various quantities such as parameter confidence intervals. Under a handful of conditions, the OLS has very attractive properties. First, OLS is a consistent estimator, i.e., as the sample size grows, the estimated parameters approach the true parameters. OLS is also unbiased (i.e. $\mathbb{E}[\boldsymbol{x}] = \boldsymbol{x}$) and has the minimum variance among all unbiased estimators (minimum variance unbiased estimator, MVUE) [42]. Under the further assumption of independent, normally distributed errors, OLS is equal to the maximum likelihood estimator [42]. One of the disadvantages of OLS is that the results tend to be sensitive to outliers, which is why an effective outlier detection and removal is essential.

3.2.2. Linearly Constrained Least Squares Estimator

If problem (7) is solved subject to linear constraints such as parameter bounds, it becomes a quadratic programming problem

$$\begin{aligned} \widehat{\boldsymbol{\beta}} &= \min_{\boldsymbol{\beta}} \boldsymbol{\beta}^T\boldsymbol{\mathcal{X}}^T\boldsymbol{\mathcal{X}}\boldsymbol{\beta} \\ &\text{s.t. } \boldsymbol{C}\boldsymbol{\beta} \leq \boldsymbol{d}. \end{aligned} \tag{9}$$

A large variety of open-source and commercial solvers exist for solving the problem.

3.2.3. Least Absolute Value Estimator

Motivated by the OLS sensitivity to outliers, sometimes (5) is solved using $\boldsymbol{L}_1$-norm yielding the least absolute value estimator (LAV). LAV is found by solving the linear programming problem (LP)

$$\widehat{\boldsymbol{\beta}} = \min_{\boldsymbol{\beta}} \|\boldsymbol{y} - \boldsymbol{X\beta}\|_1 = \min_{\boldsymbol{\beta}} \sum_{t=1}^{T} |y_t - \boldsymbol{x}_t\boldsymbol{\beta}|. \quad (10)$$

While LAV is much less sensitive to outliers than OLS, (10) has no closed-form solution and instead it must be solved with an LP solver.

3.3. Selected Branch Parameter Estimation Approach

Motivated by the disadvantages of the heuristic parameter estimation approaches underlined above, this report will utilize the linearized voltage drop approach. Each of the three estimators has unique advantages and disadvantages. Even though OLS is sensitive to outliers, there are several ways of detecting and removing measurement outliers before including them as inputs to the regression solution. Since OLS has a closed form solution, it is computationally more attractive than the least absolute value estimator and thus, attractive for parameter estimation with large data sets. Contrary to the OLS estimator, the linearly constrained least squares estimator allows setting bounds on the parameters. However, it was observed in the Georgia Tech distribution feeder that the parameter bounds may not be particularly useful. Parameter solutions at the boundary mean that the optimal parameters would be beyond the boundary. This typically happens if there is something wrong with the regression problem formulation or the measurement data. Thus, manual inspection and correction will be required in any case. For these reasons, the remainder of this report focuses on parameter estimation with the OLS estimator.

4. RADIAL SECONDARY CIRCUIT PARAMETER ESTIMATION

This chapter generalizes the method to estimate series impedance parameters of an individual branch discussed in chapter 3 to estimation of the series impedance parameters of entire radial secondary circuit. The method, which builds upon the method shown in [30], is illustrated in Figure 15. The objective of the method is to find the most likely values of resistance (R) and reactance (X) parameters shown in red in the figure. 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 at the tree root node.

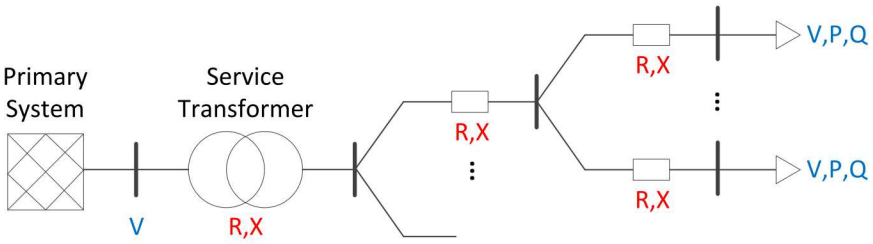

Figure 15. Secondary circuit tree for parameter estimation

The proposed method relies on the following assumptions:

1. Known circuit topology
2. Fully radial circuit topology (i.e. a tree)
3. Each of the leaf nodes (buses) of the tree has a smart meter or PV micro inverter measuring the voltage and either the active and reactive power or the current and the power factor.
4. Balanced three-phase or a single-phase circuit

If the circuit topology is unknown, it can be estimated following the approach discussed in [30]. In this report the topologies are assumed to be known, and the unknown topology case is addressed in future work. The second and the third assumptions are valid in most secondary circuits [39]. The fourth assumption 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 [43], 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. The method applies for both three-phase and single-phase circuits.

4.1. Principle

The proposed parameter estimation method operates hierarchically proceeding from the leaf nodes (customer buses) of the secondary circuit tree towards the tree root node (upstream bus). Figure 16 illustrates the estimated parameters in red and the available measurements in blue. At a given iteration, the algorithm estimates the parallel branch impedances $R_1, X_1, \dots, R_N, X_N$ of a subsection of the secondary circuit shown in Figure 16. First, the algorithm searches for a new "Upstream Bus" whose immediate downstream bus voltages and downstream branch currents (in blue Figure 16) are known (measured or estimated at previous iterations) but whose downstream branch impedances (in red in Figure 16) have not been estimated yet. Once a suitable bus and the corresponding downstream branches have been identified, the algorithm first estimates the branch impedance parameters using the available downstream bus measurements and then, estimates the upstream bus voltages using the measurements and the estimated branch parameters. These steps are explained in detail in the following subsections.

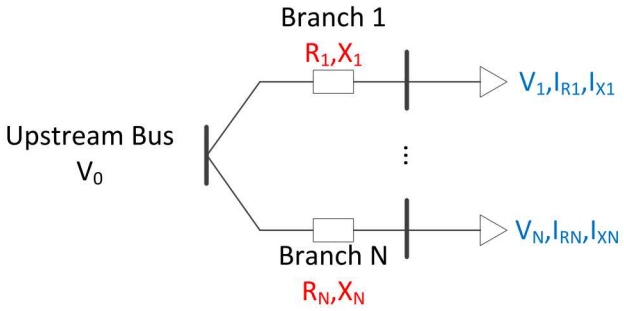

Figure 16. Radial circuit parameter estimation

4.2. Estimating Series Impedance of N Parallel Branches

There are several ways to estimate the parallel branch impedances $R_1, X_1, \dots, R_N, X_N$. In this report, two approaches are compared. The first approach estimates the parameters from a single regression problem. The second approach estimates parameters for each branch pair and merges the estimated parameters. Next, the approaches are introduced and compared.

4.2.1. Simultaneous Parallel Branch Estimation

All the parameters of the parallel branches can be estimated simultaneously by formulating a single regression problem as follows. Using the linearized voltage drop approximation (1), each of the N branches provides an approximate estimate of the upstream bus voltage V_0

$$\begin{cases} V_0 = V_1 + R_1 I_{R1} + X_1 I_{X1} + \epsilon_1 \\ V_0 = V_2 + R_2 I_{R2} + X_2 I_{X2} + \epsilon_2 \\ \vdots \\ V_0 = V_N + R_N I_{RN} + X_N I_{XN} + \epsilon_N \end{cases}. \tag{11}$$

By using M synchronous measurement samples $\boldsymbol{V}_i, \boldsymbol{I}_{Ri}, \boldsymbol{I}_{Xi} \in \mathbb{R}^M, i \in \{1, \dots, N\}$, the parameters can be estimated with one of the approaches introduced in section 3.2. with the linear regression problem

$$\boldsymbol{y} = \boldsymbol{\mathcal{X}}\boldsymbol{\beta} + \boldsymbol{\epsilon}, \tag{12}$$

where $\boldsymbol{\epsilon} \in \mathbb{R}^M$ is the error vector, $\boldsymbol{\beta} \in \mathbb{R}^{(M+2N)}$ is the parameter vector given by

$$\boldsymbol{\beta} = \left[V_{0,1}, \dots, V_{0,M}, R_1, X_1, \dots, R_N, X_N\right]^T, \tag{13}$$

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

$$\boldsymbol{y} = \left[V_{1,1}, \dots, V_{1,M}, \dots, V_{N,1}, \dots, V_{N,M}\right]^T. \tag{14}$$

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

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

where $\mathbf{I} \in \mathbb{R}^{M\times M}$ 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 and MN equations. In practice, $M \gg N$ and thus, there are many more equations than unknowns.

4.2.2. Pairwise Parallel Branch Estimation

Alternatively to the simultaneous parameter estimation shown above, the parallel branch parameters can be estimated by formulating multiple smaller regression problems and merging the resulting parameter estimates. A pairwise branch parameter estimation was considered here by formulating a regression problem for each branch pair of the N branches. This results in multiple estimates of each parameter that can be merged in several ways. In this report, the parameter estimates were simply averaged since this approach was observed to perform better than alternative approaches such as selection of best parameters based on the regression problem R-squared values. In general, there is no single metric that describes the quality of a regression model.

4.2.3. Comparison

The two approaches for estimating parallel branch impedances were utilized to estimate the 3-phase test circuit parameters without and with 1% P, 1% Q, and 0.2% V measurement error, shown in Figure 17 and Figure 18 respectively. There is no considerable trend between the two approaches between the two cases. Some parameters are estimated better with one of the approaches while others are estimated better with the other.

The pair-wise approach requires solving multiple smaller regression problems for the N parallel branches instead of a single larger regression model. For the pair-wise approach, the number of required regression models is given by the number of combinations of two branches from the N branches, i.e., $\binom{N}{2}$ (or “N choose 2”). The resulting number of regression models are 1, 3, 6, 10,

15, etc. for 2, 3, 4, 5, 6, etc. parallel branches, respectively. Since secondary circuits seldom have several parallel branches, the number of required regression problems remains small. Thus, the pairwise approach can be computationally more attractive with small number of parallel branches since it does not require building the design matrix in (15) that can be very large with large sample sizes. On the other hand, at large parallel branch numbers, the single regression problem formulation discussed in section 4.2.1 can become computationally preferable.

Since there is no clear difference in accuracy between the two methods and since the test circuits used in this report does not have large numbers of parallel branches, the remainder of this report utilizes the pairwise parallel branch estimation approach.

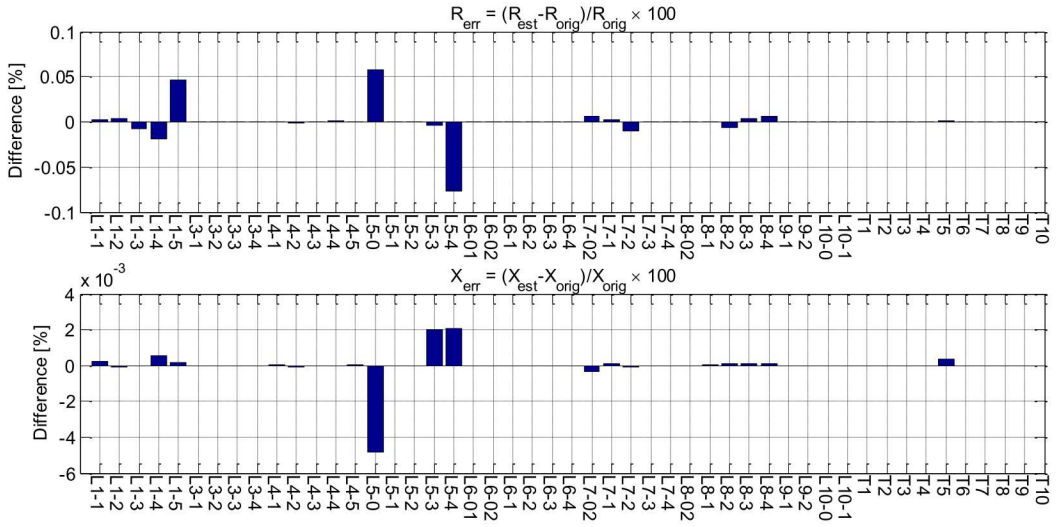

Figure 17. Average relative error difference of R and X parameters estimated with 8759 samples without measurement error (pairwise method error – single regression model error)

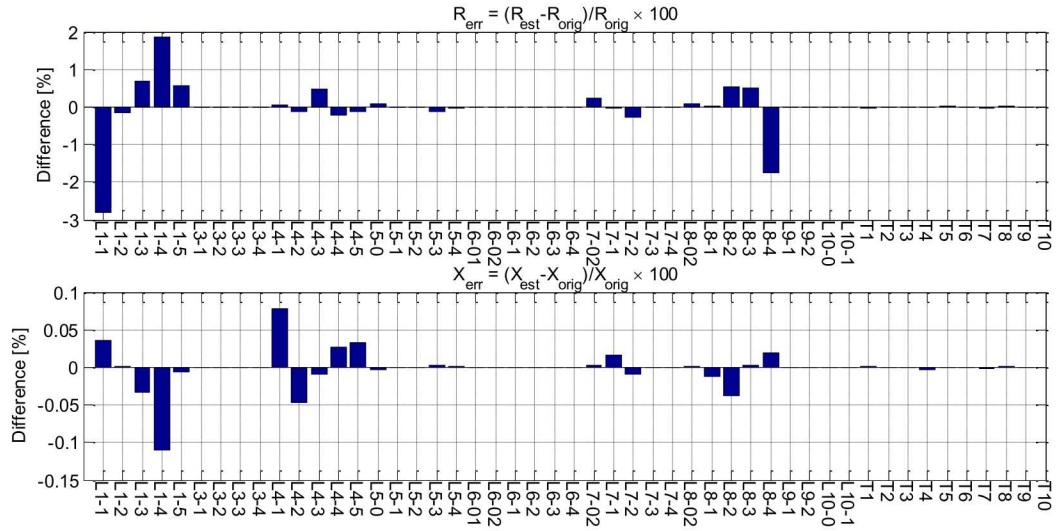

Figure 18. Average relative error differences of R and X parameters estimated with 8759 samples without measurement error 1% P, 1% Q, and 0.2% V measurement error (pairwise method error – single regression model error)

4.3. Estimating Upstream Voltages

Once the parallel branch impedances $R_1, X_1, \ldots, R_N, X_N$ have been estimated with the method shown above, the upstream bus voltages $\boldsymbol{V}_0$ can be estimated as an average of the individual branch voltage estimates

$$\boldsymbol{V}_0 = \frac{1}{N}\sum_{i=1}^{N} \|\boldsymbol{V}_i + (R_i + jX_i)(\boldsymbol{I}_{Ri} + j\boldsymbol{I}_{Xi})\| \tag{16}$$

where $\boldsymbol{V}_i$ is bus i measurement vector, $\|\cdot\|$ refers to taking the magnitude of the complex number. It should be noted that even though each upstream node voltage is calculated with the full voltage drop equation (16), the voltage parameters are estimated using the linearized voltage approximation.

4.4. Hierarchical Estimation vs. Entire Circuit Estimation

It is worth noting that the linear voltage drop approximation requires performing the parameter estimation in the hierarchical fashion shown above. It is possible to formulate a large linear regression problem for estimating all the branch impedances at once but unfortunately, the resulting design matrix $\boldsymbol{X}$ is perfectly collinear and has no unique solution. The reason for this is that all upstream branch predictors are linear combinations of the downstream predictors. The regression problem can still be solved, but it is unclear how to set sufficient additional conditions to get a unique solution that would provide the branch parameter estimates. Alternatively, the predictor linear dependency could be avoided by utilizing a nonlinear relationship between the voltage drop and the downstream predictors (such as the AC power flow). However, the resulting problem would not be linear with respect to the parameters, and iterative nonlinear optimization algorithms would be needed.

4.5. Data Selection for Parameter Estimation

Since the parameter estimation algorithm is run off-line with historical data, it is possible to selectively pick a subset of the available measurement samples. Many bad data types can be detected with conventional approaches such as checking for unrealistically high or low values based on historical data [44]. Typical distribution system secondary circuits have 5-15 customers (meters) and thus, when any of the necessary meters has missing or bad data, all measurement time stamps should be ignored. In statistical literature this is referred to as row-wise deletion.

The linear regression based branch parameter estimation presented above utilizes the voltage drop magnitude approximation (1) that is well-known to be quite accurate for typical P, Q, R, and X values [41]. The largest error occurs under heavy load (current) and leading power factor [41]. The relative linearization error with respect to P and Q for a line with an X/R=1 is shown in Figure 19. With typical P and Q combinations, the error is below 1-2%, but it can be significantly higher with either 1) large positive P and small negative Q or 2) large negative P and small positive Q. While Case 1 is very untypical in distribution secondary systems where most loads are inductive, Case 2 can occur in secondary systems when a large injection from

distributed generation at unity power factor leads to reverse active power flow while the inductive loads consume VArs.

The relative linearization error with respect to R and X for a 50kVA load at power factors $(PF) = \{0.9, 0.95, 0.98, 1.0\}$ are shown in Figure 20. Typical distribution system secondary circuit X/R ratio is in the range of 1 to 2, resulting in a linearization error below 2%. However, for circuits with high X/R ratio, the accuracy can be considerably worse.

Provided that sufficient data is available, the figures suggest that the following data be filtered before parameter estimation: 1) Measurement samples that have both high P demand and $(PF) \geq 0.95$ and 2) Measurement samples that have both high P generation (reverse power flow) and Q consumption. Filtering data has the disadvantage of reducing the number of available measurement samples, which can reduce the parameter estimation accuracy. Therefore, the filtering should only be considered for samples that have a considerable negative impact on the estimation accuracy. Whether the sample filtering is advantageous or not may depend on the characteristics of the load data at hand. The next chapter presents and compares alternative linear regression models that can be used to partially compensate the error in the linearized voltage drop equation and to handle error in practical measurement data sets.

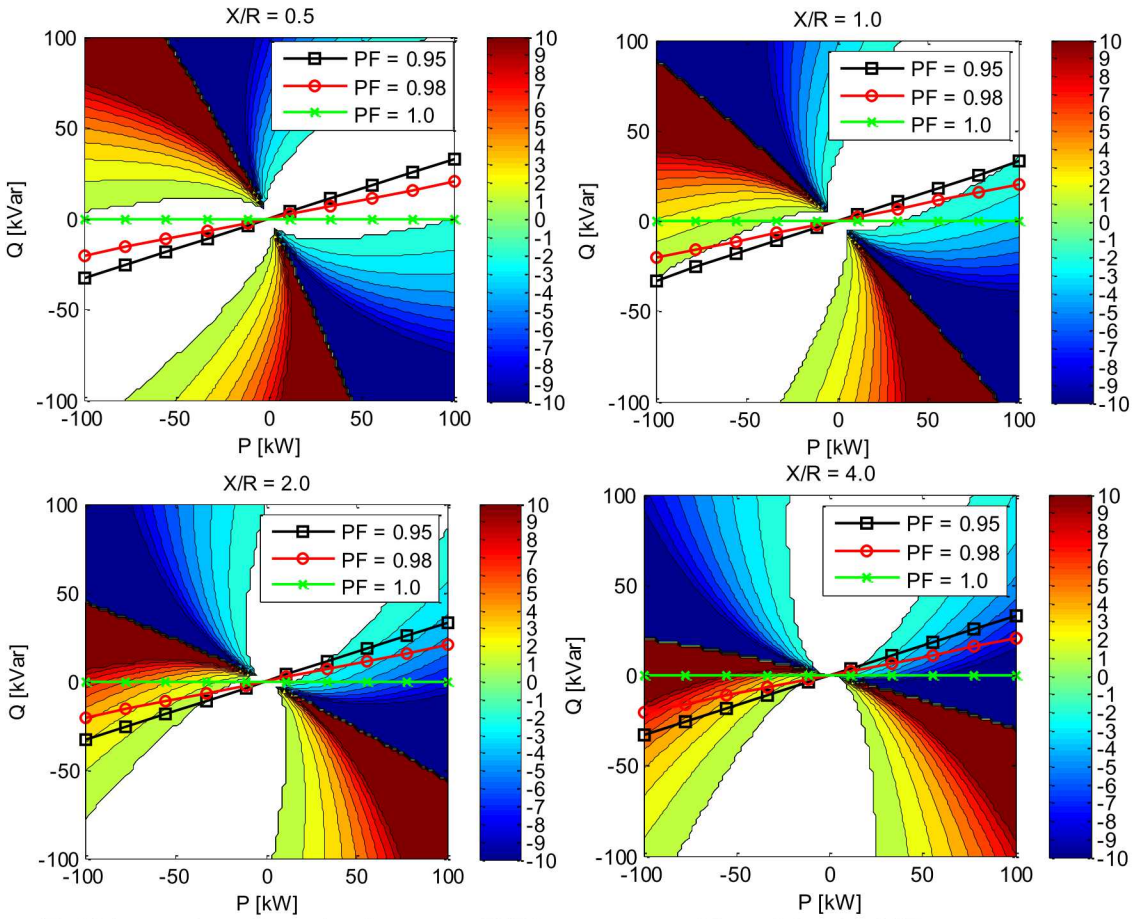

Figure 19. Voltage drop linearization error [%] for a range of P and Q with X/R $= \{0.5, 1.0, 2.0, 4.0\}$, white areas have error ≤1%, error magnitudes ≥10% are set to 10%.

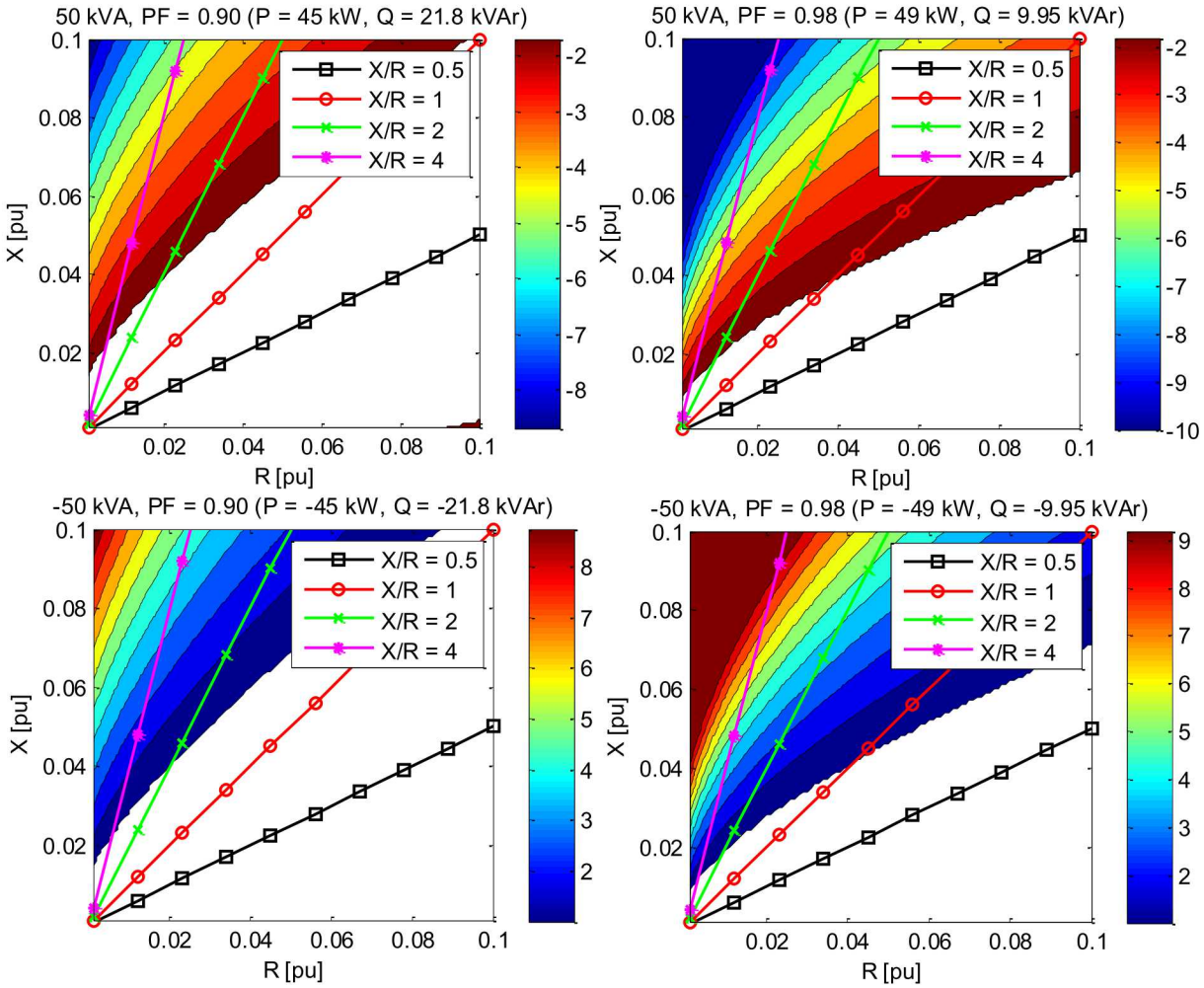

Figure 20. Voltage drop linearization error [%] for a range of R and X with S = ±50kVA, (PF) = $\{0.95, 0.98\}$, white areas have error ≤1%, error magnitudes ≥10% are set to 10%.

5. LINEAR REGRESSION MODEL SELECTION

Chapter 4 introduces a linear regression approach for estimating series impedance parameters R and X of all branches of a radial secondary circuit. The method presented so far utilizes the linear voltage drop approximation (3), which is an obvious candidate for building a parameter estimation linear regression model. 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 (3), 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 $\Delta V = V_1 - V_2$ and the predictor variables I_R and I_X thus, leading to better estimates for R and X.

To add additional terms to the radial secondary circuit parameter estimation method discussed in chapter 4, the design matrix $\boldsymbol{X}$ and the parameter vector $\boldsymbol{\beta}$ need to be modified accordingly. The response variable $\boldsymbol{y}$ is not changed by adding predictor terms. For example, if the second order current terms are added to the parallel branch regression problem in section 4.2, the measurement design matrix becomes

$$\begin{aligned}\boldsymbol{X} = [&-(N-1)\boldsymbol{I}_{R1}, -(N-1)\boldsymbol{I}_{X1}, \boldsymbol{I}_{R2}, \boldsymbol{I}_{X2}, \dots, \boldsymbol{I}_{RN}, \boldsymbol{I}_{XN}, \\ &-(N-1)\boldsymbol{I}_{R1}^2, -(N-1)\boldsymbol{I}_{X1}^2, \boldsymbol{I}_{R2}^2, \boldsymbol{I}_{X2}^2, \dots, \boldsymbol{I}_{RN}^2, \boldsymbol{I}_{XN}^2]\end{aligned} \tag{17}$$

and the parameter vector becomes

$$\boldsymbol{\beta} = \left[R_1, X_1, \dots, R_N, X_N, \beta_{Rsq,1}, \beta_{Xsq,1}, \dots, \beta_{Rsq,N}, \beta_{Xsq,N}\right], \tag{18}$$

where parameters $\beta_{Rsq,1}, \beta_{Xsq,1}, \dots, \beta_{Rsq,N}, \beta_{Xsq,N}$ do not have a direct physical meaning. The response vector $\boldsymbol{y}$ remains unchanged. Sections 5.1 and 5.2 compare the performance of different regression models on the two-bus test case from Figure 4 and the 66-node test case from Figure 5. Finally, Section 5.3 presents an adaptive approach that utilizes two different regression models.

5.1. Linear Regression Model Selection on the Two-Bus Test Case

5.1.1. Regression Model Comparison

To compare the parameter estimation accuracy with different regression models, the (known) parameters of a two-bus test circuit (Figure 4) were estimated with a set of 168 samples of real residential AMI active power measurements and randomly generated power factors. The relative errors of the estimated parameter, $R_{err} = (R_{est} - R_{true})/R_{true} \times 100\%$ and $X_{err} = (X_{est} - X_{true})/X_{true} \times 100\%$, are listed in Table 2.

The results indicate that regression models that only utilize the first order terms of I_R and I_X (models 9 and 10 shown in orange) are not the best ones. Instead, the parameter estimation error can be reduced over 8% by adding the second order terms of both of the predictor variables (models 1 and 2 shown in blue). Adding the intercept term (in table the absence of "-1") or the power factor term (in table "(PF)") to a given model lead to no considerable improvement. The best performance is obtained with regression model $\Delta V \sim I_R + I_X + I_R^2 + I_X^2$, which includes the

first and second order terms of I_R and I_X as well as the intercept term as predictors and has $\Delta V = V_1 - V_2$ as the response variable. The model does not include the power factor term.

Table 2. Estimated parameter errors with different linear regression models

Model Rank	Regression Model	R_{err} [%]	X_{err} [%]	$\lvert R_{err}\rvert + \lvert X_{err}\rvert$ [%]
1	$\Delta V \sim I_R + I_X + I_R^2 + I_X^2$	0.42	-0.07	0.49
2	$\Delta V \sim I_R + I_X + I_R^2 + I_X^2 - 1$	0.43	-0.07	0.50
3	$\Delta V \sim I_R + I_X + I_R^2 + I_X^2 + (PF) - 1$	0.49	-0.07	0.56
4	$\Delta V \sim I_R + I_X + I_R^2 + I_X^2 + (PF)$	1.01	-0.39	1.40
5	$\Delta V \sim I_R + I_X + I_R^2 - 1$	1.04	-0.41	1.44
6	$\Delta V \sim I_R + I_X + I_R^2$	1.06	-0.41	1.47
7	$\Delta V \sim I_R + I_X + I_R^2 + (PF) - 1$	1.21	-0.41	1.62
8	$\Delta V \sim I_R + I_X + I_R^2 + (PF)$	1.59	-0.71	2.29
9	$\Delta V \sim I_R + I_X - 1$	8.35	-0.44	8.79
10	$\Delta V \sim I_R + I_X$	13.62	-0.42	14.04
11	$\Delta V \sim I_R + I_X + (PF) - 1$	14.06	-0.66	14.72
12	$\Delta V \sim I_R + I_X + I_X^2$	13.52	-1.27	14.79
13	$\Delta V \sim I_R + I_X + (PF)$	14.11	-0.69	14.80
14	$\Delta V \sim I_R + I_X + I_X^2 + (PF) - 1$	13.89	-1.57	15.46
15	$\Delta V \sim I_R + I_X + I_X^2 - 1$	11.17	-4.53	15.70
16	$\Delta V \sim I_R + I_X + I_X^2 + (PF)$	15.07	-2.63	17.70

5.1.2. Parameter Error Dependency on R and X

As shown in subsection 4.5, the true circuit branch series impedance has a significant impact on the linearized voltage drop approximation accuracy. The impact of the linear voltage drop equation accuracy on the parameter estimation accuracy was studied by estimating the series impedance parameters of the two-bus circuit (Figure 4) with different series impedance parameters R and X. The parameters were estimated with regression models $\Delta V \sim I_R + I_X - 1$ (model rank 1 in Table 2) and $\Delta V \sim I_R + I_X + I_R^2 + I_X^2 - 1$ (model rank 2 in Table 2) using one week (168 samples) of measurement data of the 3-phase test feeder secondary circuit number 2 load (Figure 5). For each R and X pair, first the load voltages were simulated and then, based on the simulated load voltages and P and Q values, the branch parameters were estimated and compared to the true parameters.

The parameter estimation errors and the average simulated load voltages are summarized for regression model $\Delta V \sim I_R + I_X - 1$ in Figure 21. The R estimation errors are higher with lower X/R-ratios, and the X estimation errors are higher with very high X/R-ratios or a band of intermediate X/R-ratios. Typically, distribution transformer series impedance Z is between 1.5 % and 6 % and X/R-ratios between 1.5 and 5. Thus, service transformer R and X estimation errors below 2 % are expected. Since secondary circuit line series impedances and X/R ratios are considerably smaller, somewhat higher relative estimation error may occur. However, due to the smaller absolute impedance of lines, the resulting absolute impedance error is not expected to be higher.

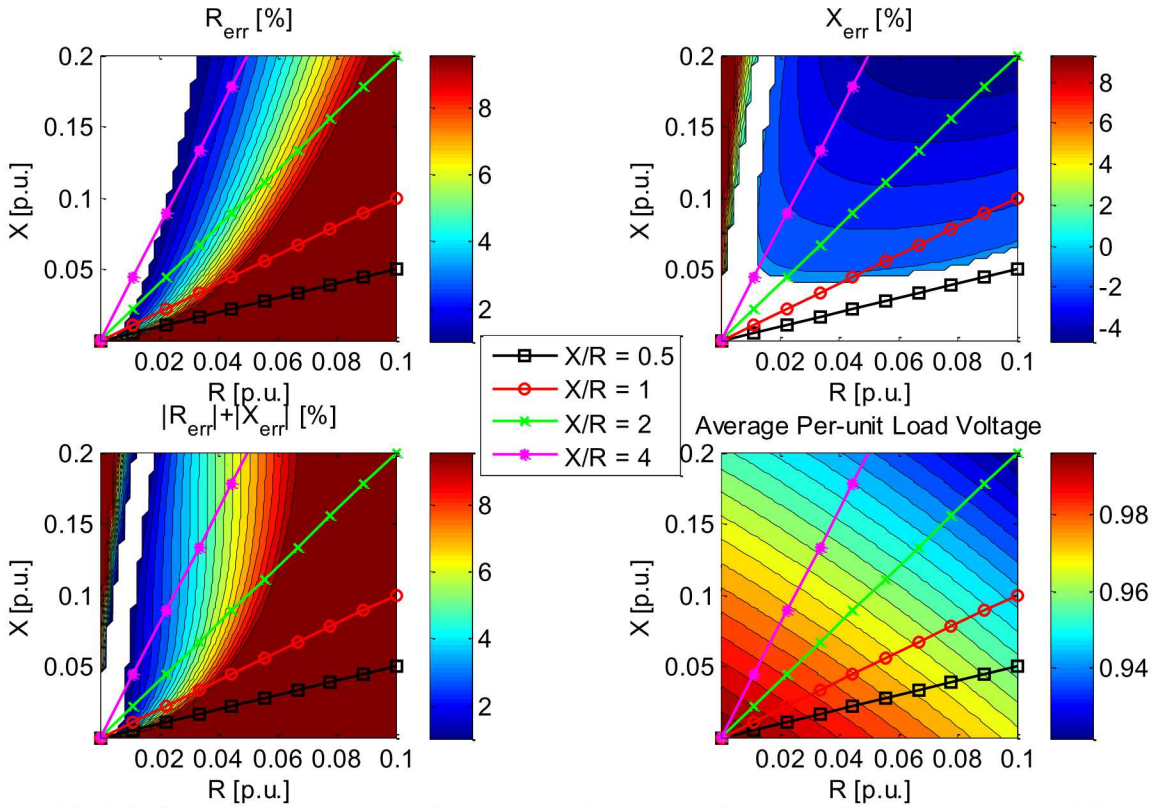

Figure 21. Relative parameter estimation errors with regression model $\Delta V \sim I_R + I_X - 1$, white areas have error ≤1%, error magnitudes ≥10% are set to 10%.

Figure 22 shows the results for the regression model $\Delta V \sim I_R + I_X + I_R^2 + I_X^2 - 1$. The model is clearly superior to model $\Delta V \sim I_R + I_X - 1$, especially at higher branch impedances.

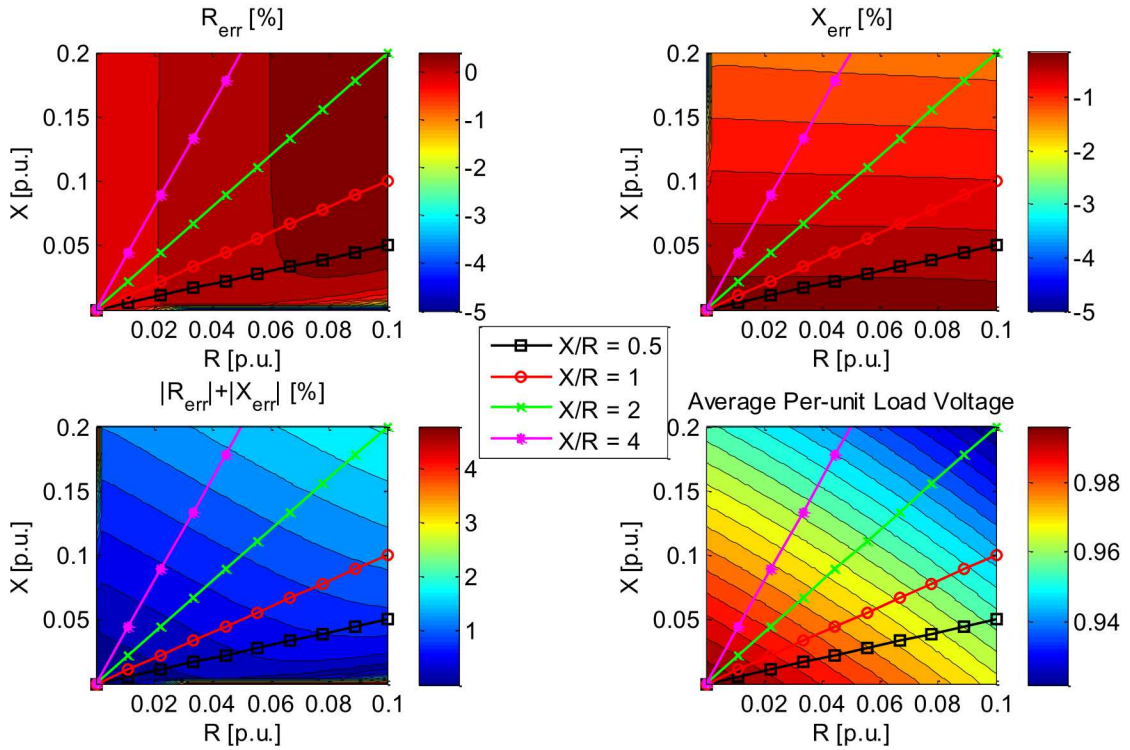

Figure 22. Relative parameter estimation errors with regression model $\Delta V \sim I_R + I_X + I_R^2 + I_X^2 - 1$, error magnitudes ≥5% are set to 5%.

5.1.3. Parameter Error Dependency on Reverse Powerflow from PV

As shown in subsection 4.5, reverse power flow can significantly deteriorate the linearized voltage drop accuracy. To see how reverse power flows influence the parameter estimation accuracy, the branch series impedance parameters of the two-bus circuit (Figure 4) were estimated with load power factor values $(\mathrm{PF}) = [0.8, 1.0]$ and PV penetration levels $E_{PV}/E_{Load} = [0, 200]$ % based on energy. The PV generation values shown in Figure 23 were scaled so that PV generated weekly energy was equal to a given fraction (PV penetration) of the weekly load consumption.

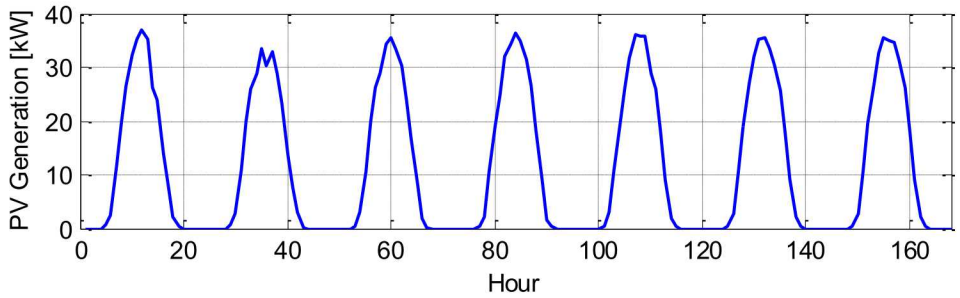

Figure 23. PV generation profile for the test data week before scaling

Figure 24 and Figure 25 display the parameter estimation accuracies for model $\Delta V \sim I_R + I_X - 1$ and for model $\Delta V \sim I_R + I_X + I_R^2 + I_X^2 - 1$, respectively. Once again, model $\Delta V \sim I_R + I_X + I_R^2 +$

$I_X^2 - 1$ clearly outperforms model $\Delta V \sim I_R + I_X - 1$. Figure 25 indicates that model $\Delta V \sim I_R + I_X + I_R^2 + I_X^2 - 1$ R estimation errors are slightly higher at high PV penetration levels and low power factor values while the X estimation error are very accurate independent of the PV penetration and the power factor.

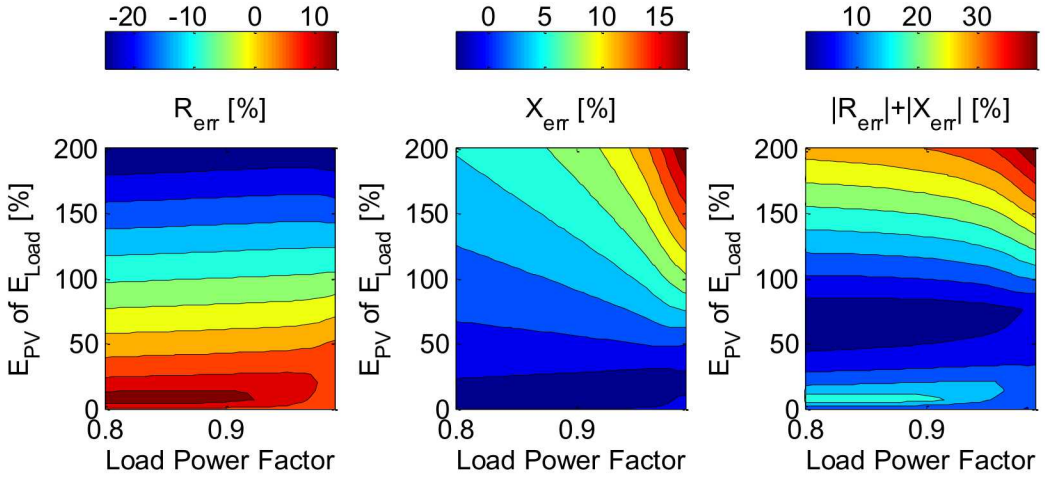

Figure 24. Relative parameter estimation errors with regression model $\Delta V \sim I_R + I_X - 1$ for a range of load power factor and PV penetration

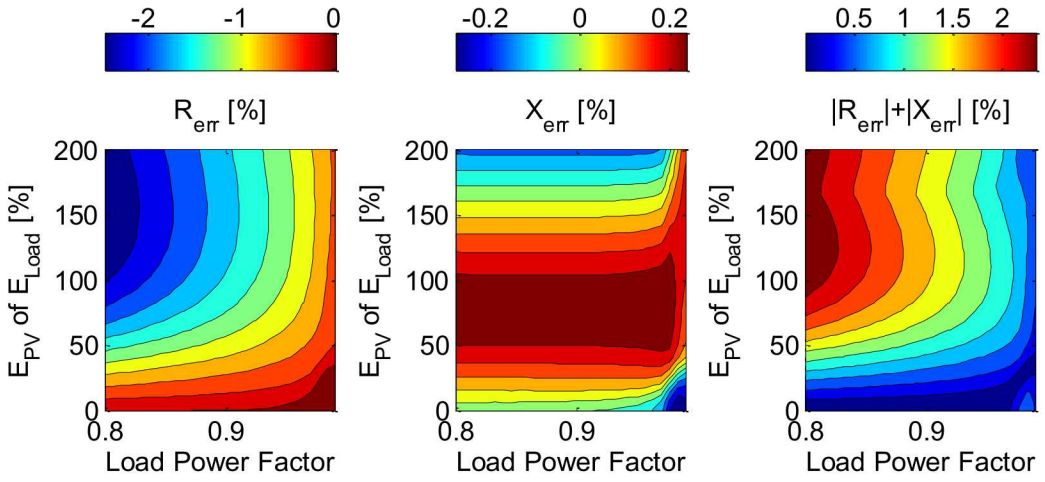

Figure 25. Relative parameter estimation errors with regression model $\Delta V \sim I_R + I_X + I_R^2 + I_X^2 - 1$ for a range of load power factor and PV penetration

5.2. Regression Model Selection on the 66-Node Test Circuit

5.2.1. Regression Model Comparison without Measurement Error

The accuracy of the hierarchical radial circuit parameter estimation accuracy presented in section 4 was analyzed by estimating all the secondary circuit parameters in the 66-node circuit and calculating the average absolute error of the estimated R and X parameters.

Table 3 compares the errors of the parameters estimated with different regression models using 8759 measurement samples (one year). The meters are assumed to be perfectly accurate and able to record voltage and current without any measurement error. The results are listed in the order from the simplest regression model to the most complicated. The best results (in terms of the absolute average $|R_{err}| + |X_{err}|$) are obtained with regression model $\Delta V \sim I_R + I_X + I_R \times I_X + I_R^2 + I_X^2 - 1$ (linear, quadratic and cross-coupling terms of the current components but no intercept or power factor terms). In all cases, regression models with intercept performed slightly worse than the respective models without the intercept.

Table 3. Relative parameter estimation errors for different linear regression models without measurement error

Included Predictor Variables (All Models Include I_R and I_X)					Avg. Abs. R_{err} [%]			Avg. Abs. X_{err} [%]			Avg. Abs. $\lvert R_{err}\rvert + \lvert X_{err}\rvert$ [%]	Max. R_{err} [%]	Max. X_{err} [%]	Model Order (Best to Worst)
Inter-cept	$I_R \times I_X$	I_R^2	I_X^2	(PF)	Lines	Trafos	All	Lines	Trafos	All				
					0.517	7.833	2.010	0.349	1.471	0.578	2.588	18.548	2.671	11
X					0.588	12.712	3.063	0.349	1.553	0.594	3.657	29.941	2.739	12
		X			0.151	1.502	0.426	0.350	1.593	0.604	1.030	2.328	2.781	8
X		X			0.169	1.584	0.458	0.350	1.592	0.604	1.062	3.969	2.781	9
		X	X		0.088	0.522	0.176	0.213	0.603	0.293	0.469	0.839	1.283	5
X		X	X		0.106	0.861	0.260	0.213	0.607	0.294	0.554	2.516	1.283	6
	X	X	X		0.035	0.165	0.062	0.026	0.137	0.049	0.111	0.570	0.408	1
X	X	X	X		0.056	0.618	0.171	0.026	0.144	0.050	0.221	1.994	0.413	2
		X	X	X	0.184	0.900	0.330	0.476	0.553	0.492	0.822	2.654	1.948	7
X		X	X	X	0.209	1.755	0.524	0.508	1.867	0.785	1.310	4.177	3.313	10
	X	X	X	X	0.066	0.657	0.187	0.060	0.168	0.082	0.269	2.168	0.411	3
X	X	X	X	X	0.074	0.980	0.259	0.085	0.483	0.167	0.425	2.965	1.021	4

The errors of parameters R and X estimated with regression model $\Delta V \sim I_R + I_X - 1$ and $\Delta V \sim I_R + I_X + I_R \times I_X + I_R^2 + I_X^2 - 1$ are shown in Figure 26 and Figure 27, respectively. Each bar represents a low-voltage branch in the 66-node test circuit. Branch names that start with L are lines, and branch names that start with T include service transformers. The errors of the estimated impedance magnitude and X/R-ratio for model $\Delta V \sim I_R + I_X - 1$ and $\Delta V \sim I_R + I_X + I_R \times I_X + I_R^2 + I_X^2 - 1$ are shown in Figure 28 and Figure 29, respectively. Without measurement error, regression model $\Delta V \sim I_R + I_X + I_R \times I_X + I_R^2 + I_X^2 - 1$ estimates all the parameters with a very high accuracy. Regression model $\Delta V \sim I_R + I_X - 1$ estimates the line parameters with a relatively good accuracy but does poorly especially in estimating the service transformer resistances and X/R-ratios. The transformer R parameters are clearly over-estimated while the transformer X parameters are clearly under-estimated. An explanation for this is the linearized voltage drop approximation illustrated in Figure 20, where the higher the X/R-ratios and the impedance magnitudes are, the more the linearized voltage drop equation underestimates the voltage drop. As a result, the transformer resistances will be over-estimated and the reactances under-estimated in the linear regression parameter estimation. This is the direction where the voltage drop approximation error reduces the fastest.

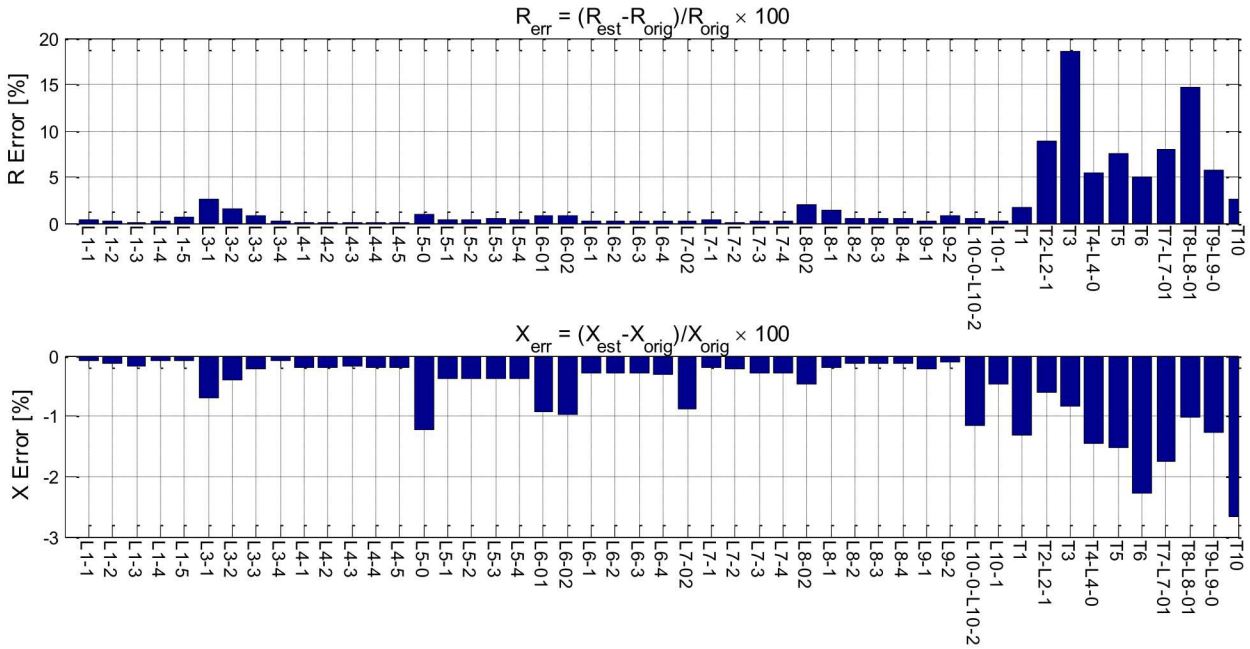

Figure 26. Relative errors of estimated R and X with regression model $\Delta V \sim I_R + I_X - 1$ without measurement error

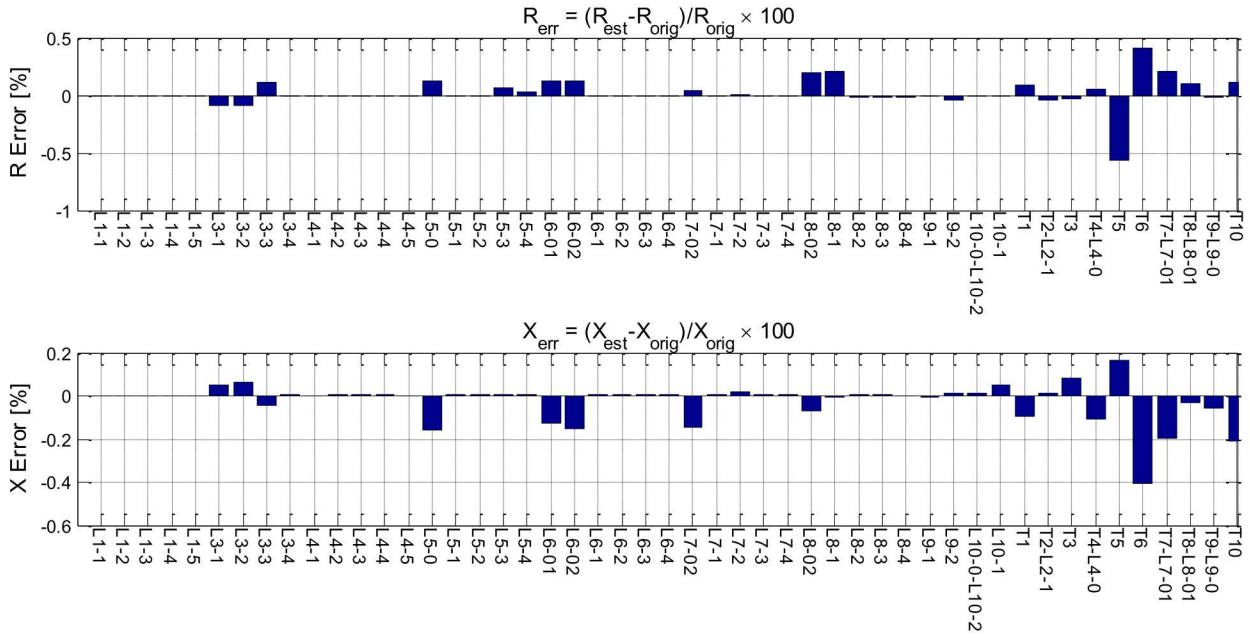

Figure 27. Relative errors of estimated R and X with regression model $\Delta V \sim I_R + I_X + I_R I_X + I_R^2 + I_X^2 - 1$ without measurement error

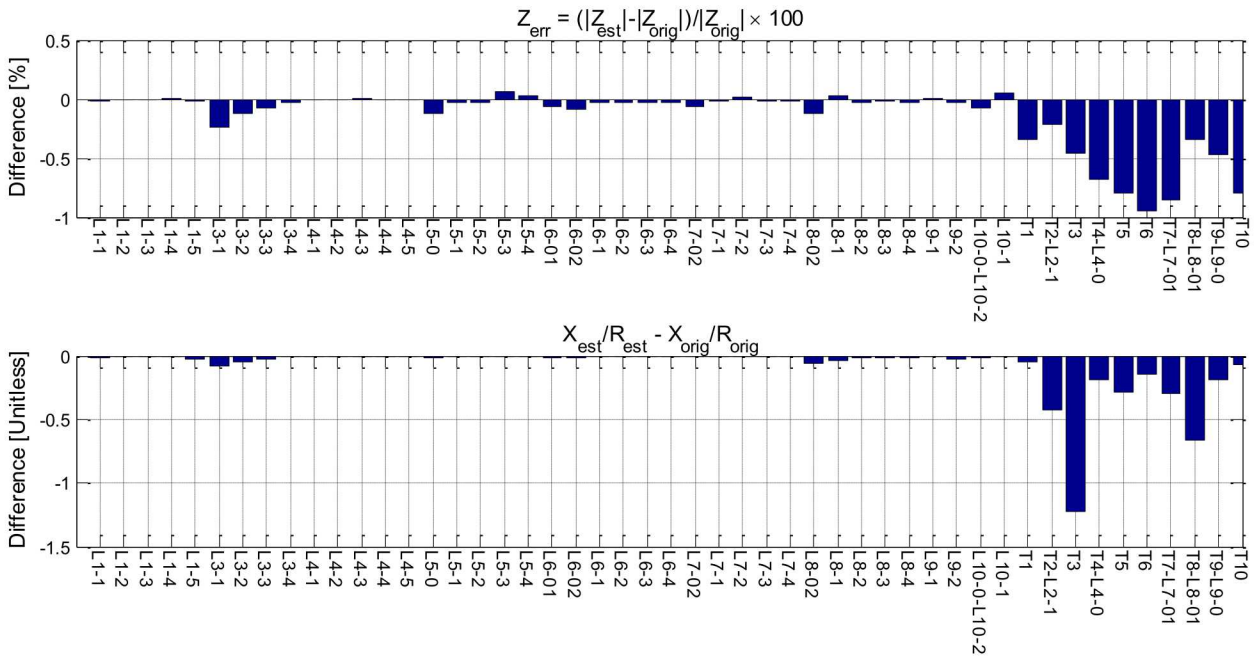

Figure 28. Relative errors of estimated Z and X/R-ratio with regression model $\Delta V \sim I_R + I_X - 1$ without measurement error

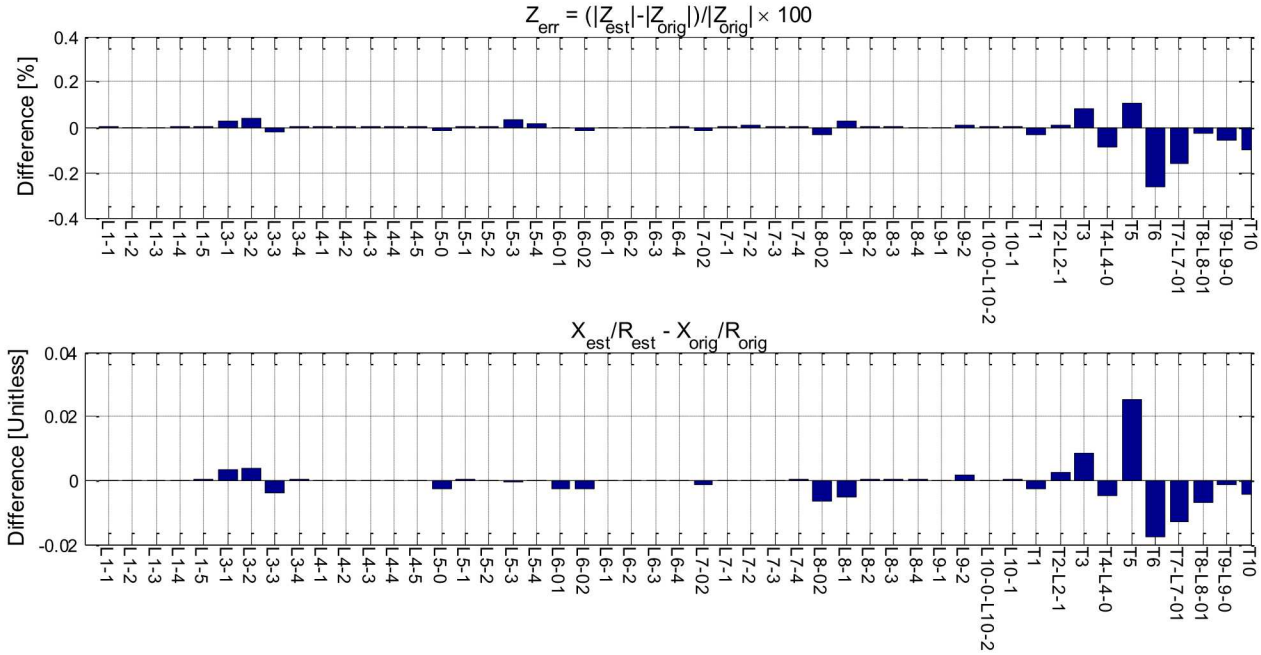

Figure 29. Relative errors of estimated Z and X/R-ratio with regression model $\Delta V \sim I_R + I_X + I_R I_X + I_R^2 + I_X^2 - 1$ without measurement error

5.2.2. Regression Model Comparison with Measurement Error

While the previous section assumes that all voltages and currents are perfectly known, generally any meter has some measurement noise that introduces error. Next, the regression model parameter estimation errors were analyzed with 1% P, 1% Q, and 0.2% V random uniform measurement error. The 0.2% voltage error correspond to the ANSI .2 accuracy class meters [45]. The 1% active and reactive power measurement error level is at or above ANSI .2 and .5 accuracy class meters [45]. This measurement error does not include calibration problems or large bias, and instead is only focused on the stochastic noise common to measurement devices. Table 4 lists the results for all secondary circuit branch R and X parameters estimated with 8759 measurement samples (one year). With measurement error, simpler models perform better than complicated ones. The best overall parameter estimates are obtained with the simplest model $\Delta V \sim I_R + I_X - 1$. Line parameters are estimated best with $\Delta V \sim I_R + I_X - 1$ while transformer resistances are estimated best with model $\Delta V \sim I_R + I_X + I_R^2 - 1$.

Table 4. Relative parameter estimation errors for different linear regression models with 1% P, 1% Q, and 0.2% V measurement error

Included Predictor Variables (All Models Include I_R and I_X)					Avg. Abs. R_{err} [%]			Avg. Abs. X_{err} [%]			Avg. Abs. $\lvert R_{err}\rvert + \lvert X_{err}\rvert$ [%]	Max. R_{err} [%]	Max. X_{err} [%]	Model Order (Best to Worst)
Inter-cept	$I_R \times I_X$	I_R^2	I_X^2	(PF)	Lines	Trafos	All	Lines	Trafos	All				
					2.83	7.38	3.76	3.50	1.51	3.10	6.86	15.30	13.18	1
X					3.08	12.18	4.94	3.48	1.57	3.09	8.03	26.09	13.20	3
		X			5.44	1.18	4.57	3.49	1.60	3.10	7.68	32.61	13.18	2
X		X			14.14	6.42	12.56	3.52	1.60	3.13	15.69	86.10	13.14	4
		X	X		10.40	2.76	8.84	15.80	2.02	12.98	21.82	54.00	43.07	5
X		X	X		16.27	7.95	14.57	15.85	2.01	13.02	27.59	108.13	43.57	8
	X	X	X		10.13	3.85	8.85	20.16	4.01	16.87	25.72	47.77	54.19	6
X	X	X	X		17.19	8.52	15.42	20.23	4.08	16.93	32.35	101.94	54.30	9
		X	X	X	16.25	8.24	14.62	14.80	1.96	12.18	26.79	116.15	60.25	7
X		X	X	X	18.35	9.93	16.63	23.31	5.34	19.64	36.27	118.64	66.16	10
	X	X	X	X	19.68	9.67	17.64	27.54	4.54	22.85	40.48	96.29	85.72	11
X	X	X	X	X	32.84	21.02	30.42	65.05	17.91	55.43	85.86	187.20	225.31	12

The errors of the R and X parameters that are estimated with the regression models $\Delta V \sim I_R + I_X + I_R^2 + I_X^2 + I_R \times I_X - 1$ (best without measurement error) and $\Delta V \sim I_R + I_X - 1$ (best with measurement error) are shown in Figure 30 and Figure 31, respectively. With measurement error, model $\Delta V \sim I_R + I_X + I_R^2 + I_X^2 + I_R \times I_X - 1$ estimates some parameters with considerably higher error than the model $\Delta V \sim I_R + I_X - 1$. This is likely caused by the measurement errors that can be large for the squared and cross-coupling terms. Model $\Delta V \sim I_R + I_X - 1$ estimates most of the line R and X parameters and the transformer X parameters with an acceptable accuracy but does worse in estimating the transformer (branch names that start with a T) resistances.

The errors of the impedance magnitude and X/R-ratio parameters that are estimated with the regression models $\Delta V \sim I_R + I_X - 1$ and $\Delta V \sim I_R + I_X + I_R^2 + I_X^2 + I_R \times I_X - 1$ are shown in Figure 32 and Figure 33, respectively. Excluding the parameters L3-4 and L9-2, $\Delta V \sim I_R + I_X - 1$ estimates all the impedance magnitudes with a good accuracy but performs worse in estimating

the transformer X/R-ratios. $\Delta V \sim I_R + I_X + I_R^2 + I_X^2 + I_R \times I_X - 1$ estimates many impedance magnitudes with considerable error.

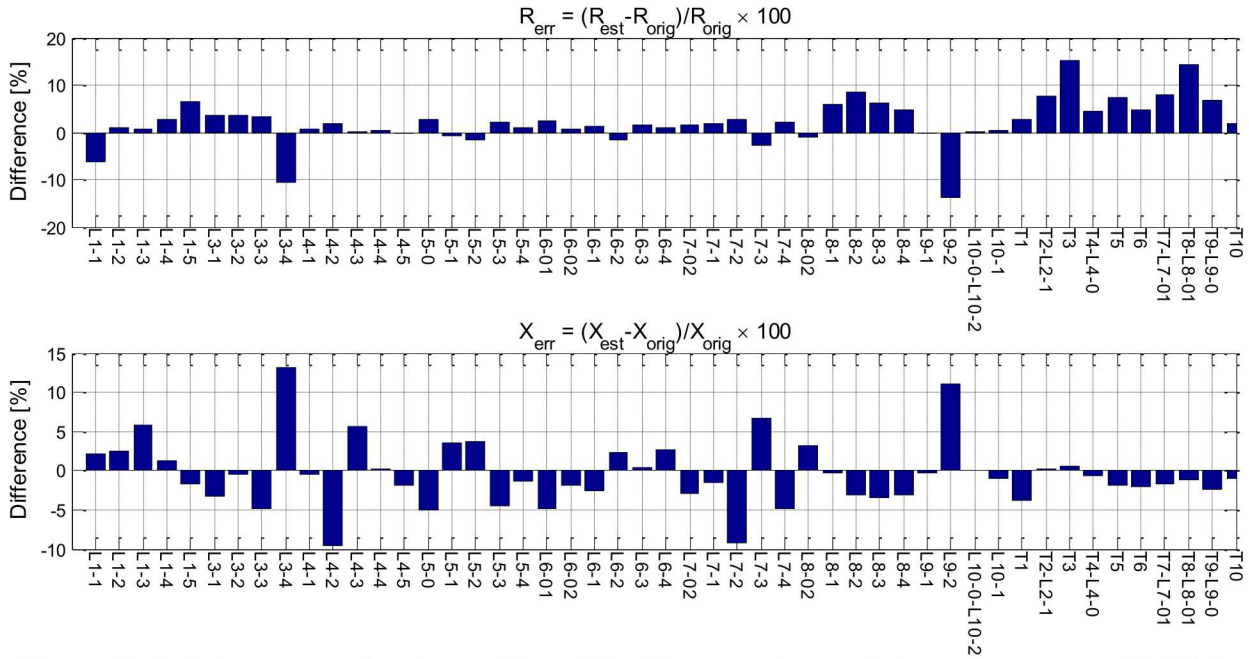

Figure 30. Relative errors of estimated R and X with regression model $\Delta V \sim I_R + I_X - 1$ with 1% P, 1% Q, and 0.2% V measurement error

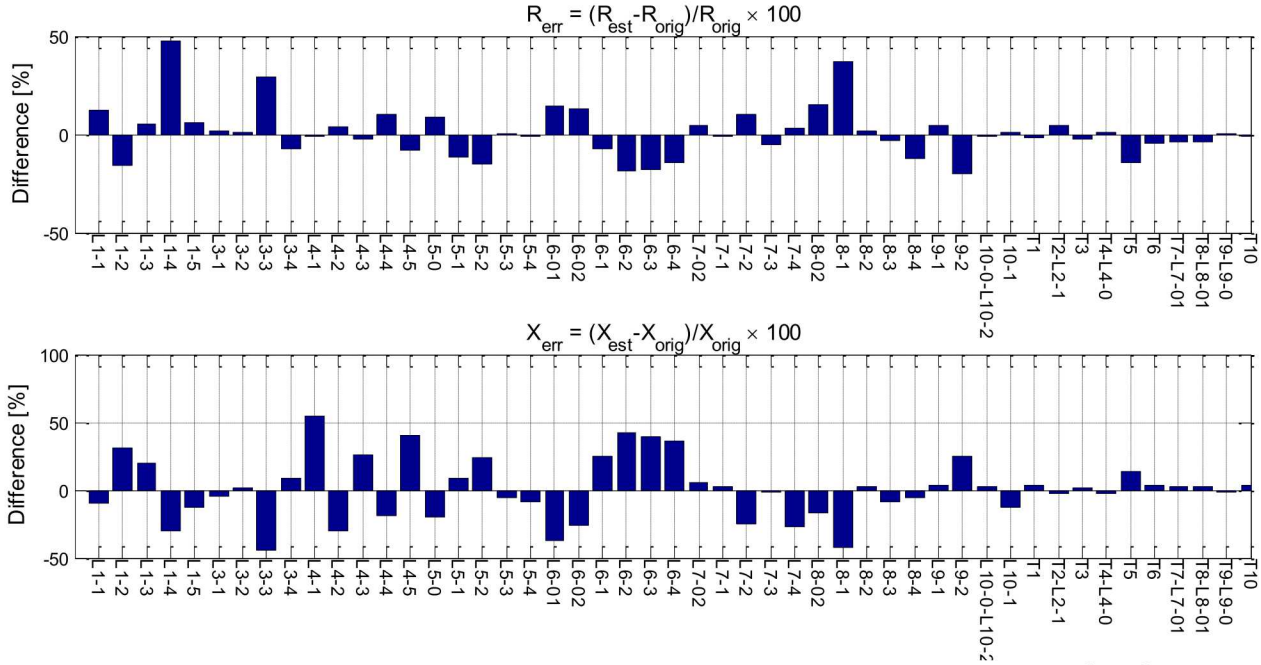

Figure 31. Relative errors of estimated R and X with regression model $\Delta V \sim I_R + I_X + I_R^2 + I_X^2 + I_R \times I_X - 1$ with 1% P, 1% Q, and 0.2% V measurement error

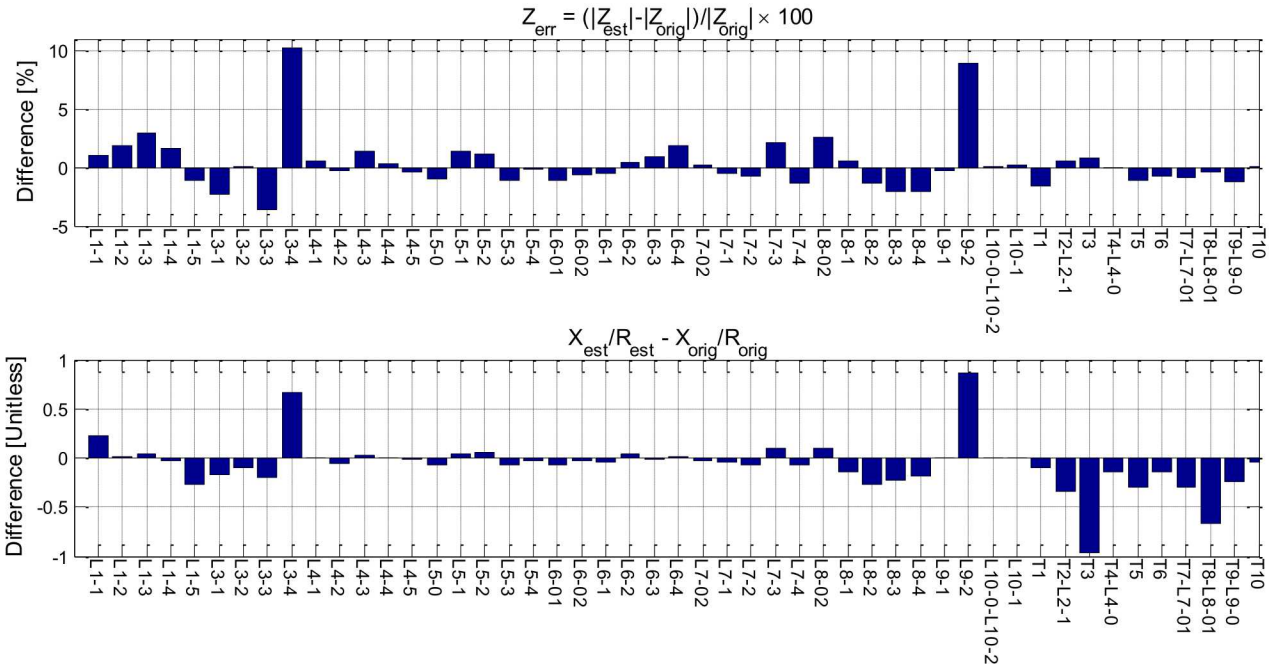

Figure 32. Relative errors of estimated Z and X/R-ratio with regression model $\Delta V \sim I_R + I_X - 1$ with 1% P, 1% Q, and 0.2% V measurement error

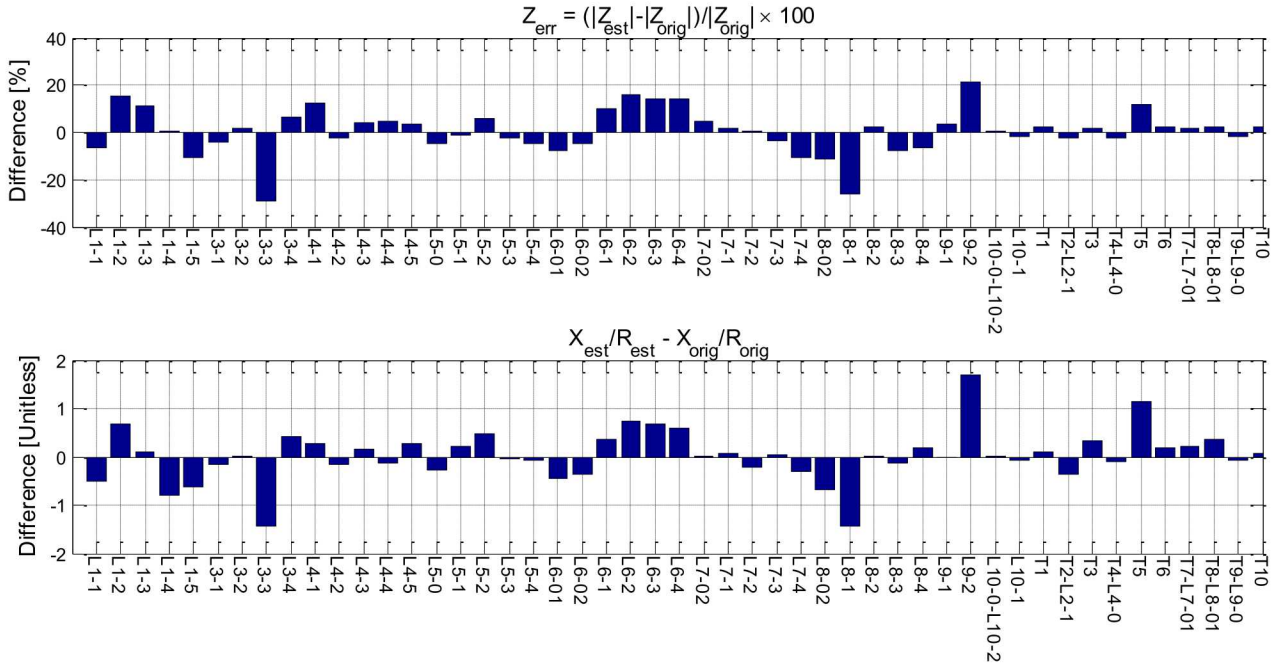

Figure 33. Relative errors of estimated Z and X/R-ratio with regression model $\Delta V \sim I_R + I_X + I_R^2 + I_X^2 + I_R \times I_X - 1$ with 1% P, 1% Q, and 0.2% V measurement error

5.3. Adaptive Regression Model Selection

Next, an adaptive regression model selection approach is discussed. Subsection 5.2 results show that the best overall parameter estimation accuracy in the presence of measurement error is obtained with model $\Delta V \sim I_R + I_X - 1$. However, other models estimate better service transformer parameters, because the transformer X/R-ratios and impedance magnitudes are higher than those of lines, which results in higher errors in the linearized voltage drop approximation as discussed in Subsection 5.2.1. Based on this insight, this section discusses an adaptive approach where regression problems consisting solely of line parameters are estimated with model $\Delta V \sim I_R + I_X - 1$ and regression problems involving transformer parameters are estimated with a regression model that includes other terms. For this work, the topology is assumed to be known, so each branch is known as either a line or transformer, and all connections between branches are known.

The adaptive approach was analyzed by estimating the transformer R and X parameters in the 66-node test case with 8759 measurement samples. Six different regression models for transformer regression problems were analyzed, and the line parameters were estimated using model $\Delta V \sim I_R + I_X - 1$. The results without measurement error are listed in Table 5, and the results with 1% P, 1% Q and 0.2% V measurement error are listed in Table 6. Similar to section 5.1 and 5.2, complicated models perform better without measurement error and simple models perform better in the presence of measurement error. The best overall parameter estimates were obtained with model $\Delta V \sim I_R + I_X + I_R^2 - 1$, i.e., a model that includes the squared current term I_R^2.

Table 5. Relative parameter estimation errors of the adaptive approach with different regression models for regression problems including transformers without measurement Error

Included Predictor Variables for Regression Problems with Transformer Parameters (All Models Include I_R and I_X)					Avg. Abs. R_{err} [%]			Avg. Abs. X_{err} [%]			Avg. Abs. $\lvert R_{err}\rvert + \lvert X_{err}\rvert$ [%]	Max. R_{err} [%]	Max. X_{err} [%]	Model Order (Best to Worst)
Inter-cept	$I_R \times I_X$	I_R^2	I_X^2	(PF)	Lines	Trafos	All	Lines	Trafos	All				
					0.517	7.833	2.010	0.349	1.471	0.578	2.588	18.548	2.671	3
		X			0.517	1.384	0.694	0.349	1.593	0.603	1.297	2.638	2.782	4
			X		0.517	12.371	2.937	0.349	9.506	2.218	5.154	31.464	14.792	6
		X	X		0.517	0.439	0.501	0.349	0.575	0.395	0.896	2.638	1.218	2
	X				0.517	7.903	2.025	0.349	6.657	1.636	3.661	18.615	9.563	5
	X	X	X		0.517	0.186	0.450	0.349	0.100	0.298	0.748	2.638	1.218	1

Table 6. Relative parameter estimation errors of the with different regression models for regression problems including transformers with 1% P, 1% Q, and 0.2% V measurement error

Included Predictor Variables for Regression Problems with Transformer Parameters (All Models Include I_R and I_X)					Avg. Abs. R_{err} [%]			Avg. Abs. X_{err} [%]			Avg. Abs. $\lvert R_{err}\rvert + \lvert X_{err}\rvert$ [%]	Max. R_{err} [%]	Max. X_{err} [%]	Model Order (Best to Worst)
Inter-cept	$I_R \times I_X$	I_R^2	I_X^2	(PF)	Lines	Trafos	All	Lines	Trafos	All				
					2.834	7.380	3.761	3.500	1.514	3.095	6.857	15.299	13.178	4
		X			2.834	1.009	2.461	3.500	1.603	3.113	5.574	13.667	13.178	1
			X		2.834	11.809	4.665	3.500	9.293	4.683	9.348	26.657	15.161	6
		X	X		2.834	2.878	2.843	3.500	2.108	3.216	6.059	13.667	13.178	2
	X				2.834	7.444	3.774	3.500	6.547	4.122	7.897	15.339	13.178	5
	X	X	X		2.834	3.274	2.924	3.500	3.205	3.440	6.364	15.846	13.178	3

Figure 34 and Figure 35 compare the performance of the adaptive approach with (transformer regression problem) models $\Delta V \sim I_R + I_X - 1$, $\Delta V \sim I_R + I_X + I_R^2 - 1$, and $\Delta V \sim I_R + I_X + I_R^2 + I_X^2 - 1$ with different P, Q, and V error levels and sample sizes. Model $\Delta V \sim I_R + I_X - 1$ is equal to the nonadaptive approach with regression model $\Delta V \sim I_R + I_X - 1$. Model $\Delta V \sim I_R + I_X + I_R^2 - 1$ beats model $\Delta V \sim I_R + I_X - 1$ as long as sufficiently large sample size is used. As shown in Figure 35, model $\Delta V \sim I_R + I_X + I_R^2 - 1$ outperforms model $\Delta V \sim I_R + I_X + I_R^2 + I_X^2 - 1$ independent of the error level and sample size. Based on these results, the rest of this report utilizes the adaptive parameter estimation approach where regression models $\Delta V \sim I_R + I_X - 1$ and $\Delta V \sim I_R + I_X + I_R^2 - 1$ are used for regression problems without and with transformer parameters, respectively.

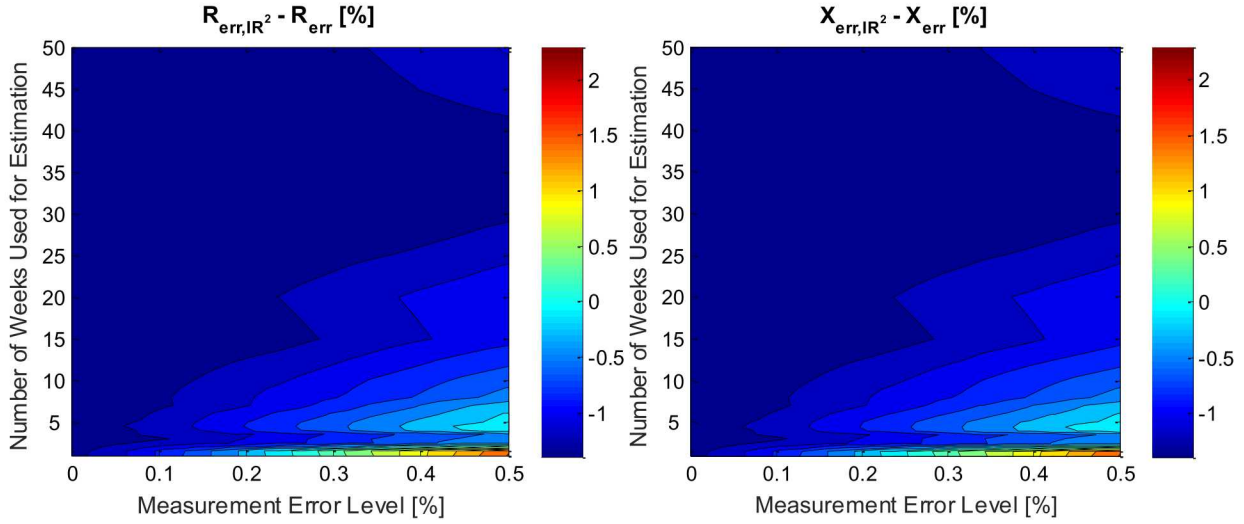

Figure 34. Difference of relative errors of estimated R and X between adaptive regression model $\Delta V \sim I_R + I_X + I_R^2 - 1$ and the non-adaptive model $\Delta V \sim I_R + I_X - 1$

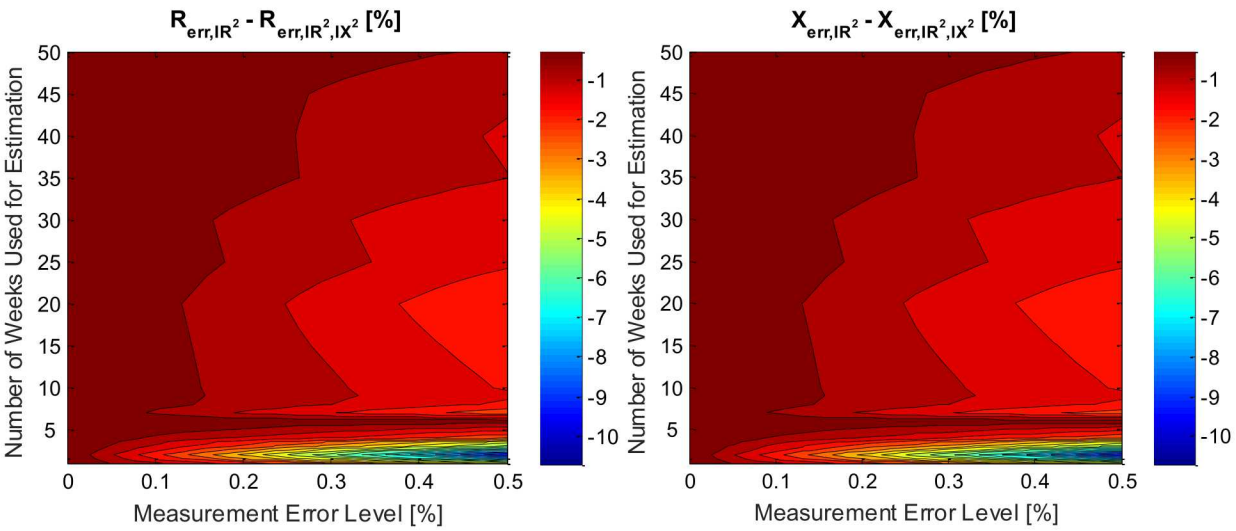

Figure 35. Difference of relative errors of estimated R and X between regression models $\Delta V \sim I_R + I_X + I_R^2 - 1$ and $\Delta V \sim I_R + I_X + I_R^2 + I_X^2 - 1$

6. IMPLEMENTATION

This chapter discusses the implementation of the distribution system secondary circuit parameter estimation.

6.1. Data Flows 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. It may be beneficial to integrate the offline DSPE closely with the current distribution system simulator, since DSPE requires moving and managing large quantities of data, much of which is shared with the current distribution system model/simulator as shown in Figure 36.

The current model components, parameters and connectivity will be fetched from GIS. SCADA will transmit the measurements and device 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 distributed voltage measurements from the AMI and DER as well, the parameter and topology estimator will calculate the secondary system component parameters and pass the refined component parameters and permanent connectivity (in case topology estimation is also performed) to GIS. The Big Data challenge is efficiently managing these data flows through advanced data analytics, optimized database queries, and rapid time series analysis.

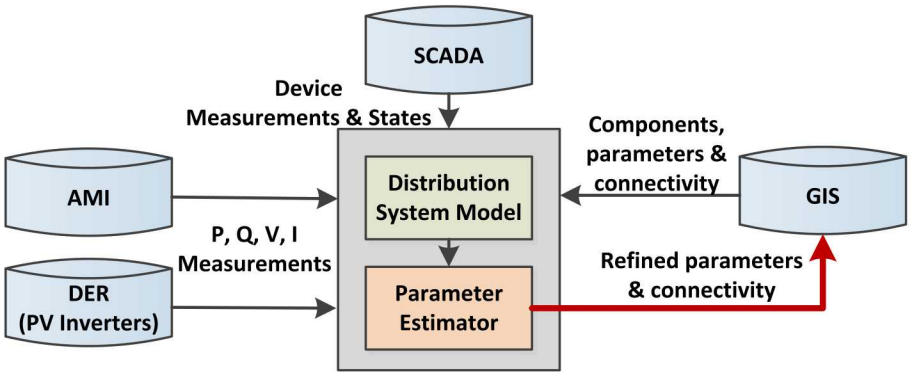

Figure 36. Data flows for distribution system secondary circuit parameter estimation

6.2. Algorithm

The high-level secondary circuit parameter estimation algorithm is shown in Figure 37. The algorithm starts by compiling the existing utility feeder model in the distribution system simulator for time series power flow analysis. For all analysis in this report, the distribution system power flow is solved by OpenDSS, and all parameter estimation algorithms are implemented in MATLAB [46], [47]. The input data required for the time series analysis includes load active and reactive power (or current and power factor) measurements, substation

voltage measurements, and PV generation. The output from the time series power flows solutions is the service transformer MV-side voltages.

Once the time series power flow has finished, the parameter estimation algorithm processes one secondary circuit at a time, estimating the secondary circuit branch impedances with the hierarchical approach introduced in section 4. For each secondary circuit the algorithm does the following. First, the measurement samples are selected following the principles discussed in section 4.5.Then, at each iteration the algorithm identifies and selects a suitable circuit subsections with N parallel branches as explained in section 4. Next, the branch parameters are estimated using one of the two parallel branch estimation approaches discussed in section 4.2, the ordinary least squares estimator introduced in section 3.2.1., and the adaptive regression model selection introduced in section 5.3. With small number of parallel branches, either of the parallel branch parameter estimation approaches can be used while at large number of parallel branches, the simultaneous parallel branch estimation approach discussed in section 4.2.1 becomes preferable. When the ordinary least squares estimator results in non-physical (negative or too large) parameters, the linearly constrained least squares estimator introduced in section 3.2.2 can be used instead. The adaptive regression model selection includes the second order terms of the resistive (real power current) I_R^2 to all the regression problems that include transformer parameters in order to compensate for the larger error of the linearized voltage drop equation over these branches. Alternative regression models can be considered depending on the error level in the available measurements.

A key advantage of the parameter estimation algorithm described is that it does not require any power flow solutions during the linear regression optimization of the parameters. Validation and error analysis is done by running another time series power flow simulation with the estimated parameters and comparing the simulated and measured voltages at the ends of the secondary system.

The manual verification of the parameter estimation results is very important to avoid any possibilities of replacing previously accurate impedance parameters with poorer estimates. In the manual verification step, the user needs to compare the estimated parameter values and how closely they align with physically expected values. This step is also useful for detecting any data or topology problems that become evident in the form of physically impossible parameter estimates or insignificant linear regression p values of some parameters. Insignificant p value of an estimated parameter (regression coefficient) indicates that we do not have sufficient evidence to reject the null hypothesis that the parameter is a given value (here zero). In other words, we are not confident rejecting that the parameter has no role in the regression. In impedance parameter estimation, this would likely imply that either the regression model is incorrect (missing measurements, incorrect topology, etc.) or the measurement data is inaccurate due to gross errors. However, insignificant parameters (regression coefficients) may also be caused by some of the linear regression assumptions being invalid. For detailed discussion on these assumptions, the reader is referred to e.g. [48].

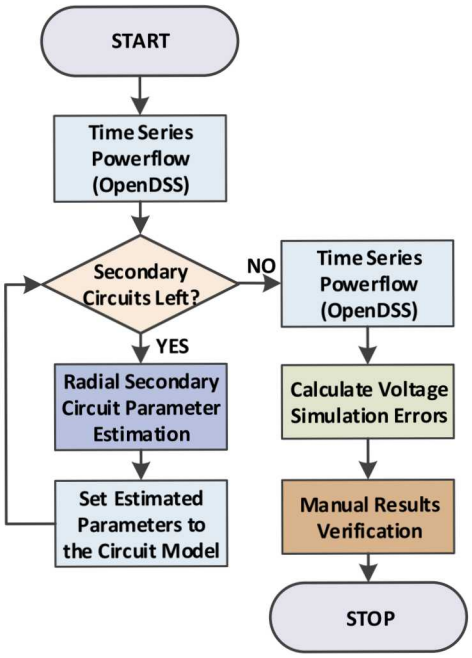

Figure 37. High-level algorithm for distribution system secondary circuit parameter estimation

7. RESULTS

7.1. Three-Phase Test Circuit

The hierarchical linear regression parameter estimation algorithm was used to estimate the three-phase test system secondary circuit parameters. Before estimating the parameters of a secondary circuit, the algorithm merges all the series branches that have no measurements in between, e.g., the service transformer and the service drop of secondary circuit no. 2 (Figure 5). Figure 38 shows the merged secondary circuit tree topologies that the algorithm has processed based on the OpenDSS circuit model. The node names and the node upstream branch names are shown on the lower left and upper right sides of the nodes, respectively. The transformer medium voltage and low voltage side nodes are abbreviated with "HV" and "LV', respectively. The circuits where the transformer is merged with its downstream branch do not have a node with "LV". Branches that include a transformer have "T" and branches that include a line have "L" and in their label.

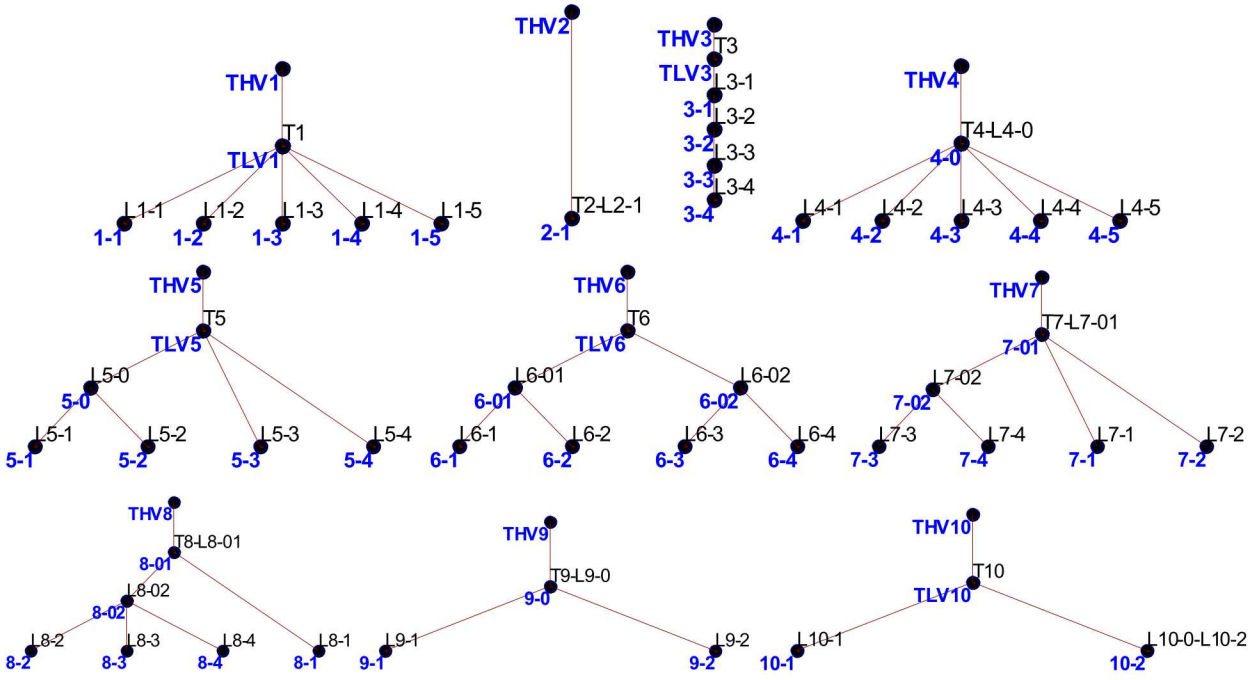

Figure 38. Merged secondary circuit trees with bus names shown in blue and bus upstream branch names shown in black

7.1.1. Parameter Estimation Accuracy without Measurement Error

Typically, the utilities have abundant AMI measurement data that can be filtered for the offline parameter estimation. However, in order to keep the parameter estimation algorithm computationally and data efficient, the necessary number of measurement samples must be determined. A sufficiently large measurement sample should be selected to achieve the highest possible accuracy, but at some point adding more samples is expected to have diminishing

marginal returns. The parameter estimation error dependency on the number of measurement samples was studied on the 66-node test circuit. Figure 39 shows the relative errors of the estimated R and X parameters with measurement sample lengths 1-53 weeks when no measurement error is present. The measurement samples were selected in a random order from the available set of 53 weeks of load data. Each plot contains 49 lines, one for each R or X parameter. The average error of all parameters is shown with the red bold line. As the figure indicates, when no measurement error is present, there is no need to utilize large numbers of measurement samples. It should also be noted that utilizing a larger number of measurement samples can also reduce parameter estimation accuracy depending on the characteristics of the additional samples. Generally, parameter estimation results depend on the characteristics of the utilized measurement data and measurement error.

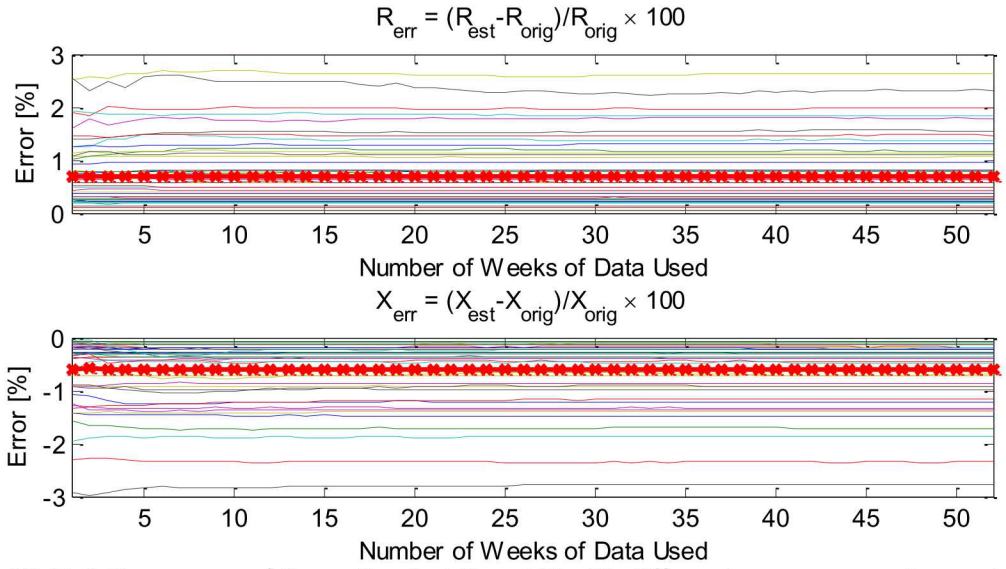

Figure 39. Relative errors of the estimated R and X with different measurement sample sizes without measurement error

The detailed parameter estimation accuracies without measurement error are shown in Figure 40. All the parameters are estimated with an error less than 3%. As discussed in chapter 5, considerably better parameter estimation accuracies without measurement error are achieved with alternative regression models. However, the proposed adaptive regression model approach achieves the best performance in the practical setting with noisy measurements.

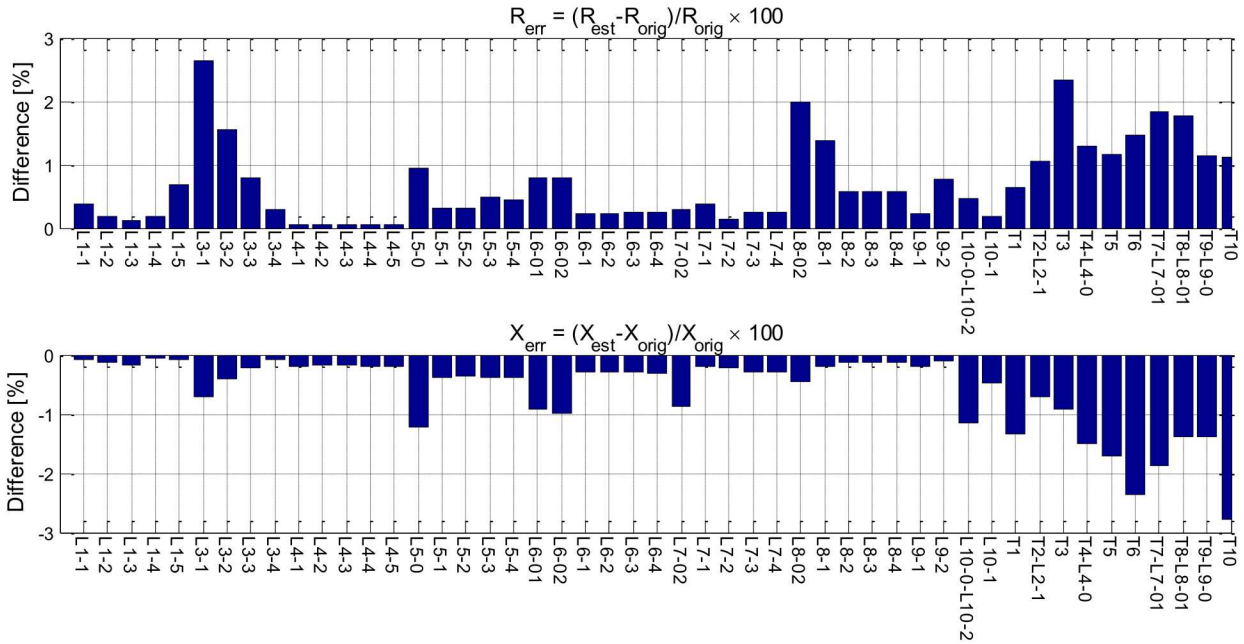

Figure 40. Relative errors of the estimated R and X with 53 weeks of measurement data and without measurement error

7.1.2. Parameter Estimation Accuracy with Power Measurement Error

Next, the parameter estimation error dependency on P and Q measurement error and sample size was studied. Measurement error was added to each active and reactive power measurement sample i with $P_i = (1+\delta)P_i$ and $Q_i = (1+\delta)Q_i$, where the measurement error magnitude was set to $\delta \in \text{Uniform}(-0.05, 0.05)$. Perfect voltage measurements were assumed.

Figure 41 shows the average absolute R and X estimation errors of all the 66-node test circuit parameters with measurement sample lengths from 1 through 50 weeks with different P and Q measurement error levels and sample sizes. With reasonably small measurement error levels and sufficient sample sizes, the average parameter estimation errors are well below 1-2%. In the presence of measurement error, increasing the measurement sample size considerably improves the accuracy of the estimated parameters up to the sample sizes of around 10 weeks (1680 samples) after which adding further samples has only small if any improvement. The estimation accuracy does not improve monotonically with the sample size due to the randomness of the load data and the sample selection. If Figure 41 was repeatedly plotted over a randomly drawn order of the load data, the average error of the repetitions is expected to reduce monotonically as the sample size grows.

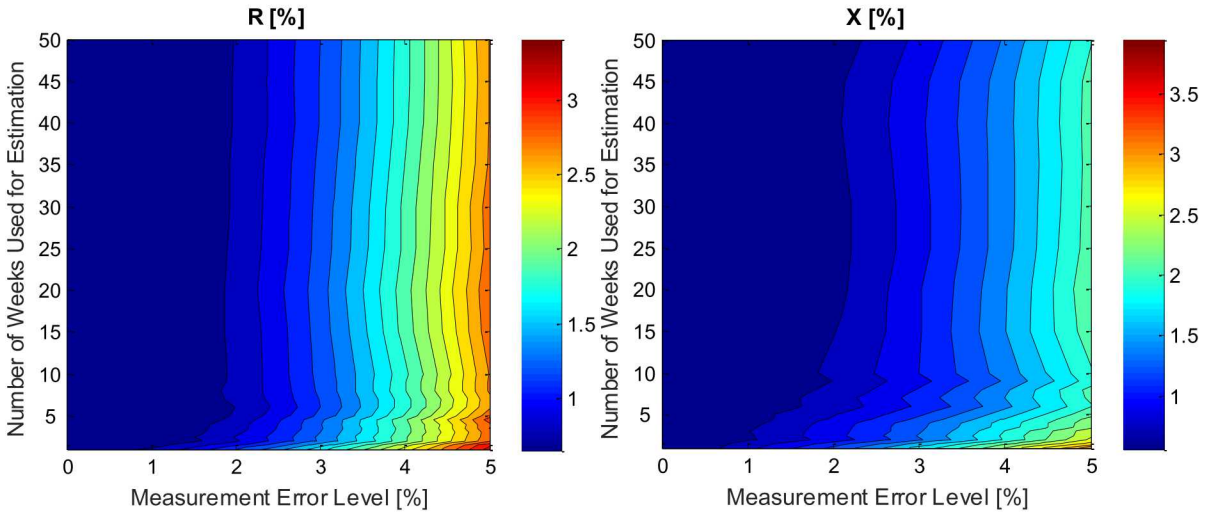

Figure 41. Average relative errors of R (left) and X (right) estimated with 1-50 weeks of load data and 0-5% of P and Q measurement error

7.1.3. Parameter Estimation Accuracy with Voltage Measurement Error

Next, parameter estimation performance was studied in the presence of voltage measurement error. The same principle was used as with the power measurement error above, but now voltage measurement error up to 0.5% (e.g. Class 0.5 smart meter) was added to the voltage measurements. Figure 42 shows the average absolute R and X estimation errors of all the 66-node test circuit parameters with measurement sample lengths from 1 through 50 weeks with different voltage measurement error levels and sample sizes. Again the errors of the estimated parameters reduce (although not monotonically) as the sample size is grown. Clearly, voltage measurement error has a much larger influence on the parameter estimation accuracy than the power measurement error. Therefore, it is an imperative to have high quality voltage measurements.

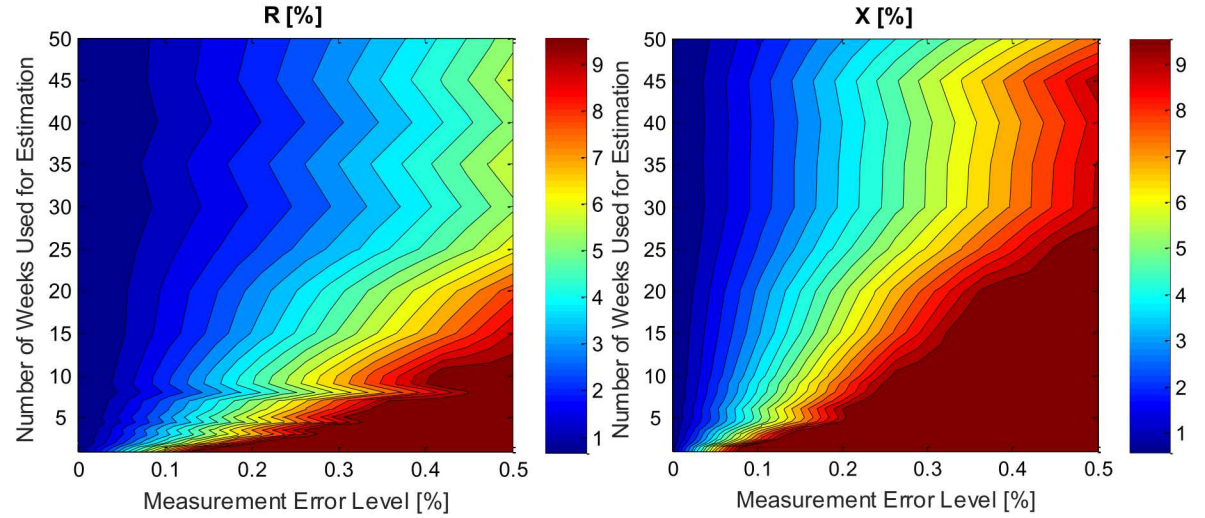

Figure 42. Average relative errors of R (left) and X (right) estimated with 1-50 weeks of load data and 0-0.5% of V measurement error, error magnitudes >10% are set to 10%

It is relevant to point out the reason for the parameter estimation to be very sensitive for the voltage measurement error. Figure 43 shows the boxplots of load voltage measurement errors at error level 0.2%, voltage drops over the service lines, and voltage drops over the service transformer for the entire year of load data (8760 samples). Each boxplot on the top of Figure 43, represents the load voltage measurement errors. Each boxplot in the bottom of Figure 43 represents the voltage drops over a given branch or transformer. The voltage drops over some lines are on the same order or smaller than the measurement error. Since the proposed parameter estimation utilizes the branch voltage drop as the linear regression response variable, it is not possible to estimate effectively impedance parameters for branches over which the voltage drop is less or equal than the voltage measurement error level. Therefore, it is imperative to have high quality voltage measurements.

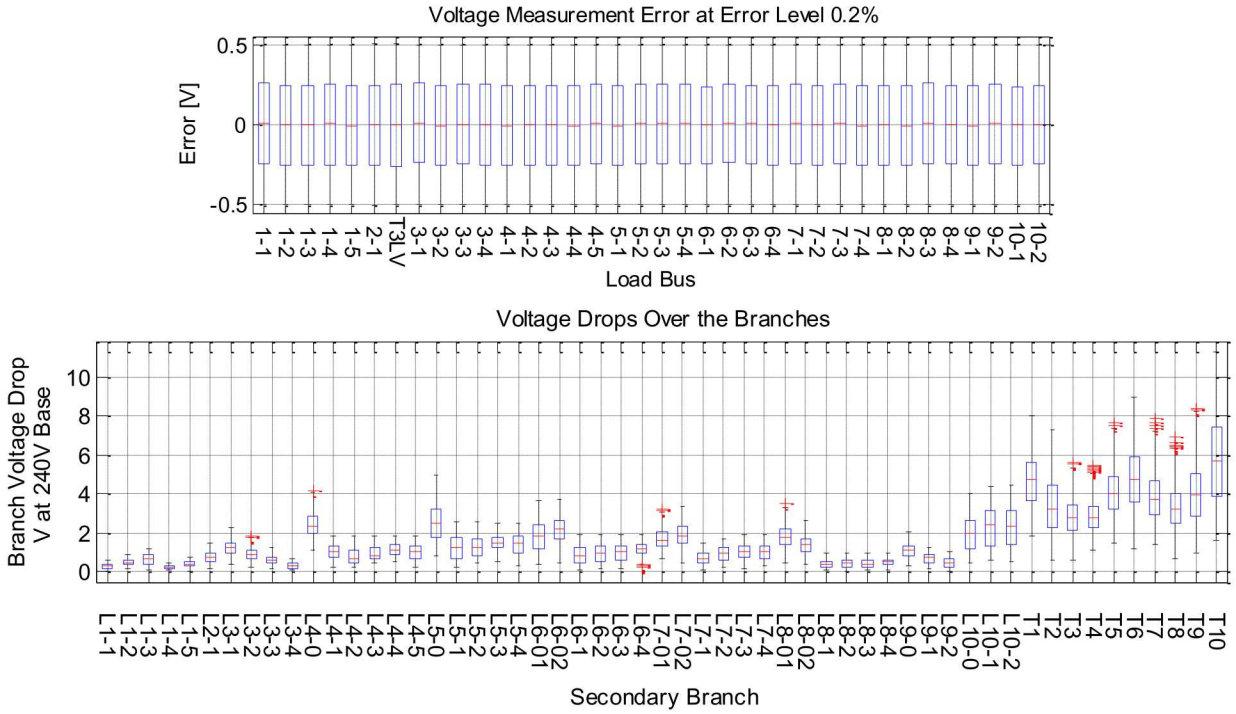

Figure 43. Load voltage measurement errors at error level 0.2% (top) compared to voltage drops over the 240V base secondary circuit transformers and lines (bottom)

7.1.4. Parameter Estimation Accuracy with Power and Voltage Measurement Error

Finally, parameter estimation performance is shown in the presence of both power and voltage measurement error. Similarly to the previous subsections, measurement error up to 0.5% (e.g. Class 0.5 smart meter) was added to the P, Q, and V measurements. Figure 42 shows the average absolute R and X estimation errors of all the 66-node test circuit parameters with measurement sample lengths from 1 through 50 weeks with different voltage measurement error levels and sample sizes. Again the errors of the estimated parameters decrease (although not monotonically) as the sample size is grown. Even at relatively high measurement error levels of 0.5%, the

average absolute error or the estimated R and X can be brought down to around 6% and 9%, respectively by utilizing sufficiently large sample sizes. However, adding more samples does not completely remove the influence of measurement error. Therefore, it is necessary to have high-quality (especially voltage) measurements in order to accurately estimate the parameters.

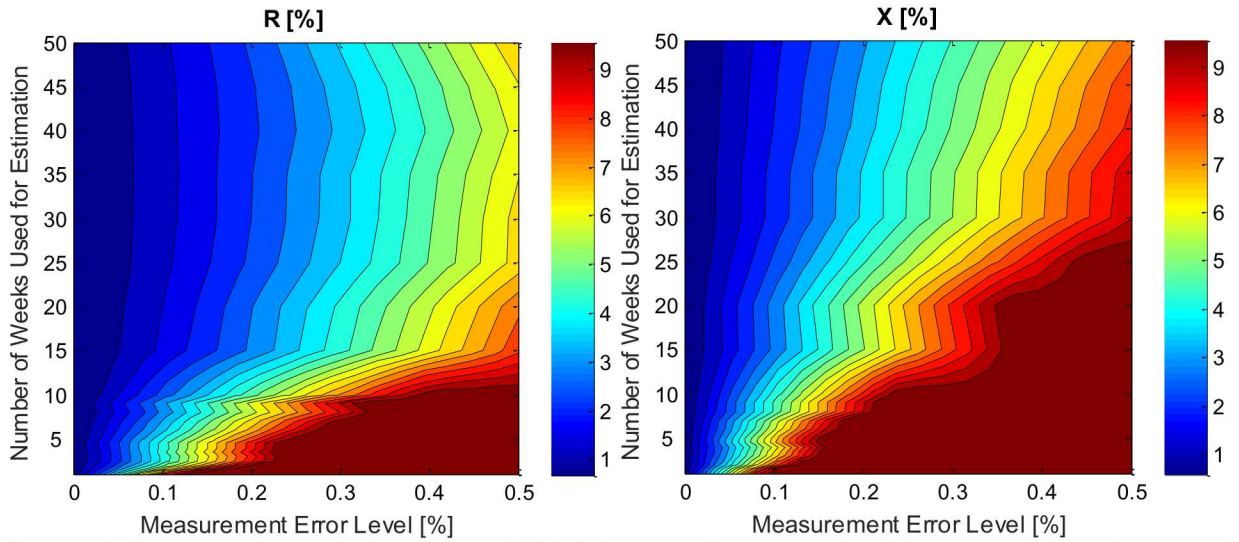

Figure 44. Average absolute R (left) and X (right) estimation errors for 1-50 weeks of load data and 0-0.5% of P, Q, and V measurement error, error magnitudes >10% are set to 10%

The relative errors of the estimated R and X parameters with 0.2% measurement error level and 8759 measurement samples are shown in Figure 45. Excluding parameters of L3-4 and L9-2, all the parameters are estimated with a reasonably good accuracy with mean (maximum) relative error of R and X at 2.05% (8.67%) and 2.73% (9.50%), respectively. The relative errors of the estimated Z and absolute errors of the estimated X/R-ratios are shown in Figure 46. Again, excluding parameters of L3-4 and L9-2, the mean (maximum) relative Z and absolute X/R-ratio errors were 1.10% (3.61%) and 0.08 (0.28), respectively.

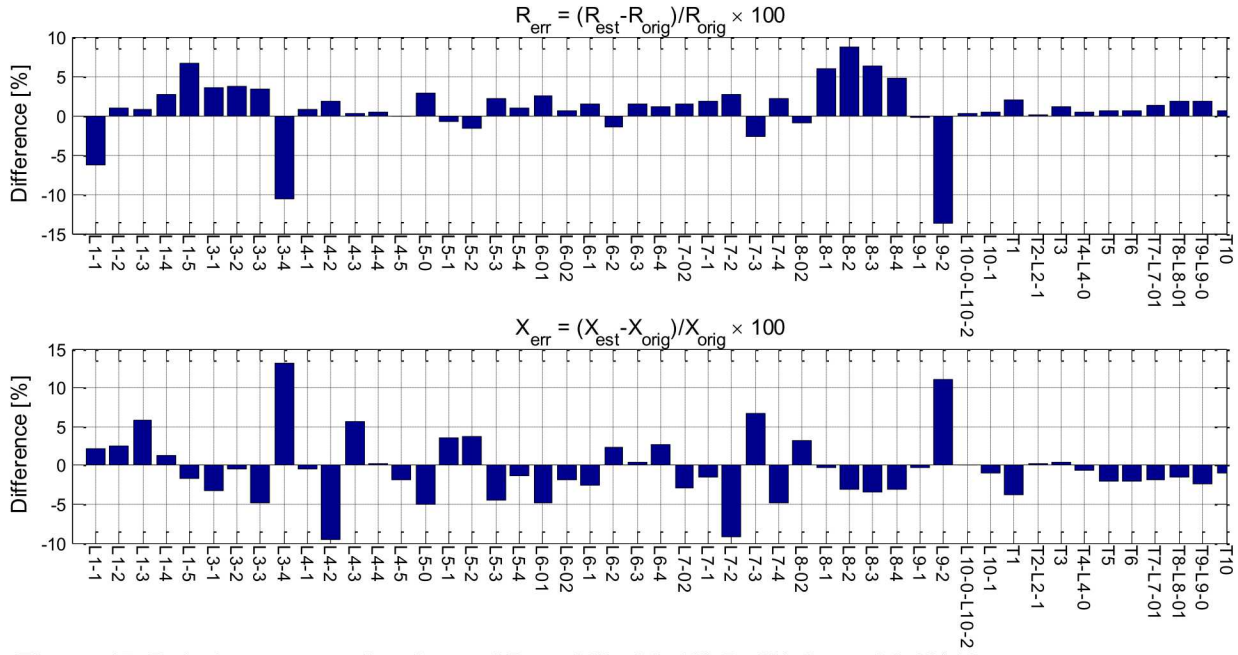

Figure 45. Relative errors of estimated R and X with 1% P, 1% Q, and 0.2% V measurement error when the parameters are estimated with the adaptive approach

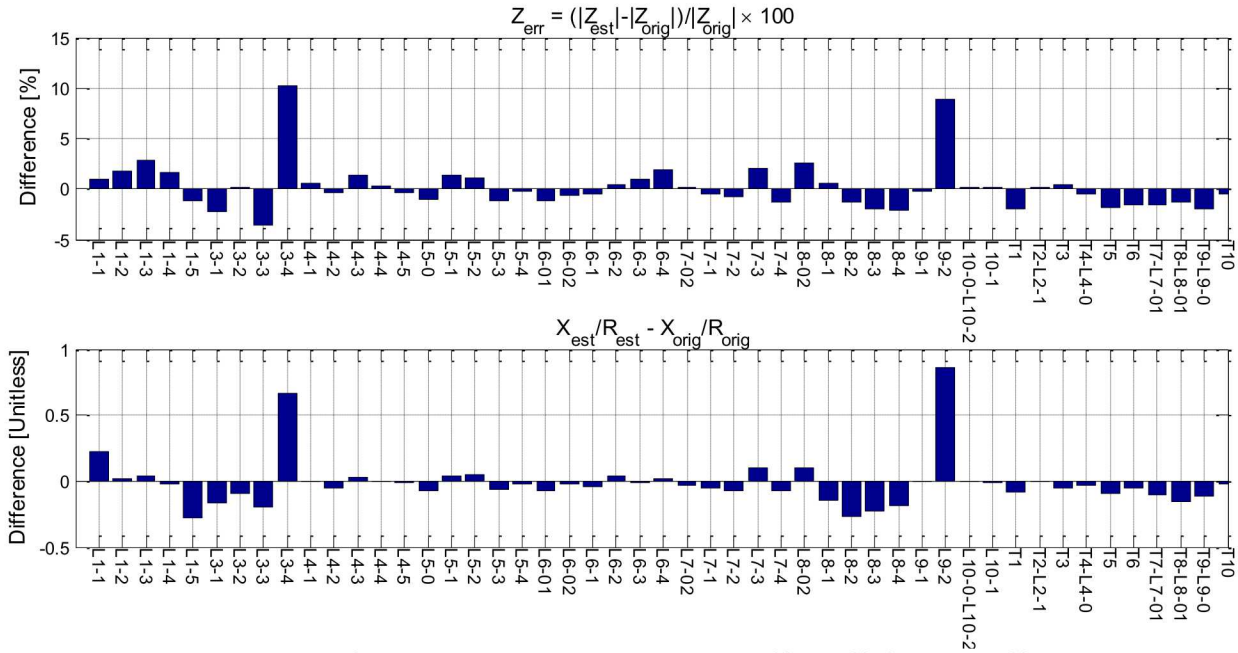

Figure 46. Relative errors of estimated Z and X/R-ratio with 1% P, 1% Q, and 0.2% V measurement error when the parameters are estimated with the adaptive approach

No straightforward reason was found for the low quality of L3-4 and L9-2 parameter estimates. The voltage drop over branch L3-4 is quite small but not uniquely small in the test circuit. Branch L9-2 has a relatively high X/R-ratio compared to the other branches in the test circuit. However, neither of these factors fully explains the considerably poorer quality of the estimated parameters. Additionally as shown in Figure 47 the two regression problems that contain L3-4

and L9-2 have very low R-squared values, which indicates a low quality fit. However, there are other well-estimated parameters with similar R-squared values and thus, R-squared cannot be directly used to describe the quality of a regression model. Figure 47 also shows that the sum of squared errors of these two regression problems are not considerably higher and that the parameter p-values are significant. Finally, as Figure 48 and Figure 49 indicate, the regression problems that include L3-4 and L9-2 do not have particularly high or low means and/or standard deviations of the response and/or predictor variables. Potentially, other characteristics in the load data of L3-4 and L9-2 would explain the lower estimation quality of these parameters. To conclude, more work is needed to determine the approach to detect and fix parameter estimation regression problems with low quality.

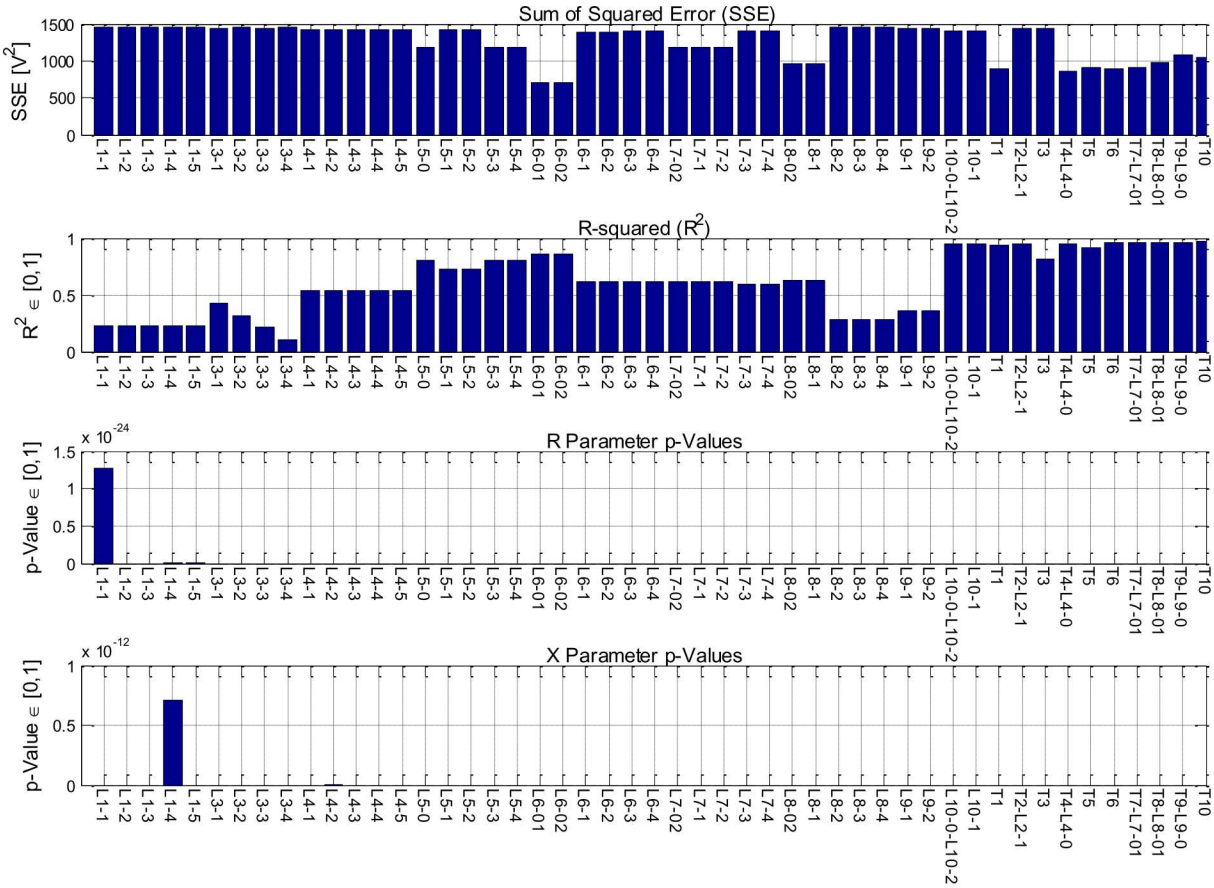

Figure 47. Sum of squared errors, R-squared values, R and X p-values with 1% P, 1% Q, and 0.2% V measurement error

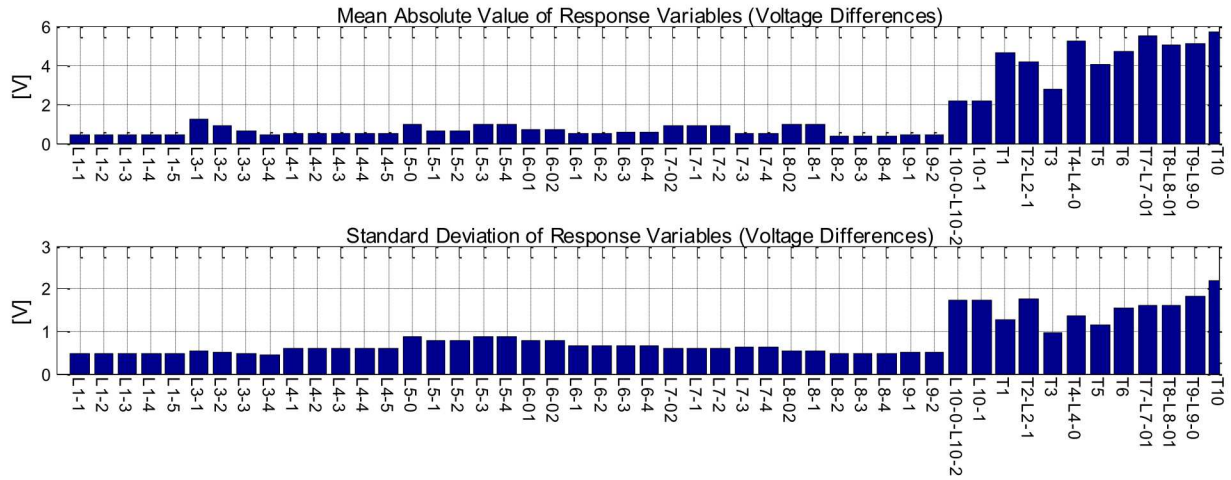

Figure 48. Means and standard deviations of the response variables (voltage drops)

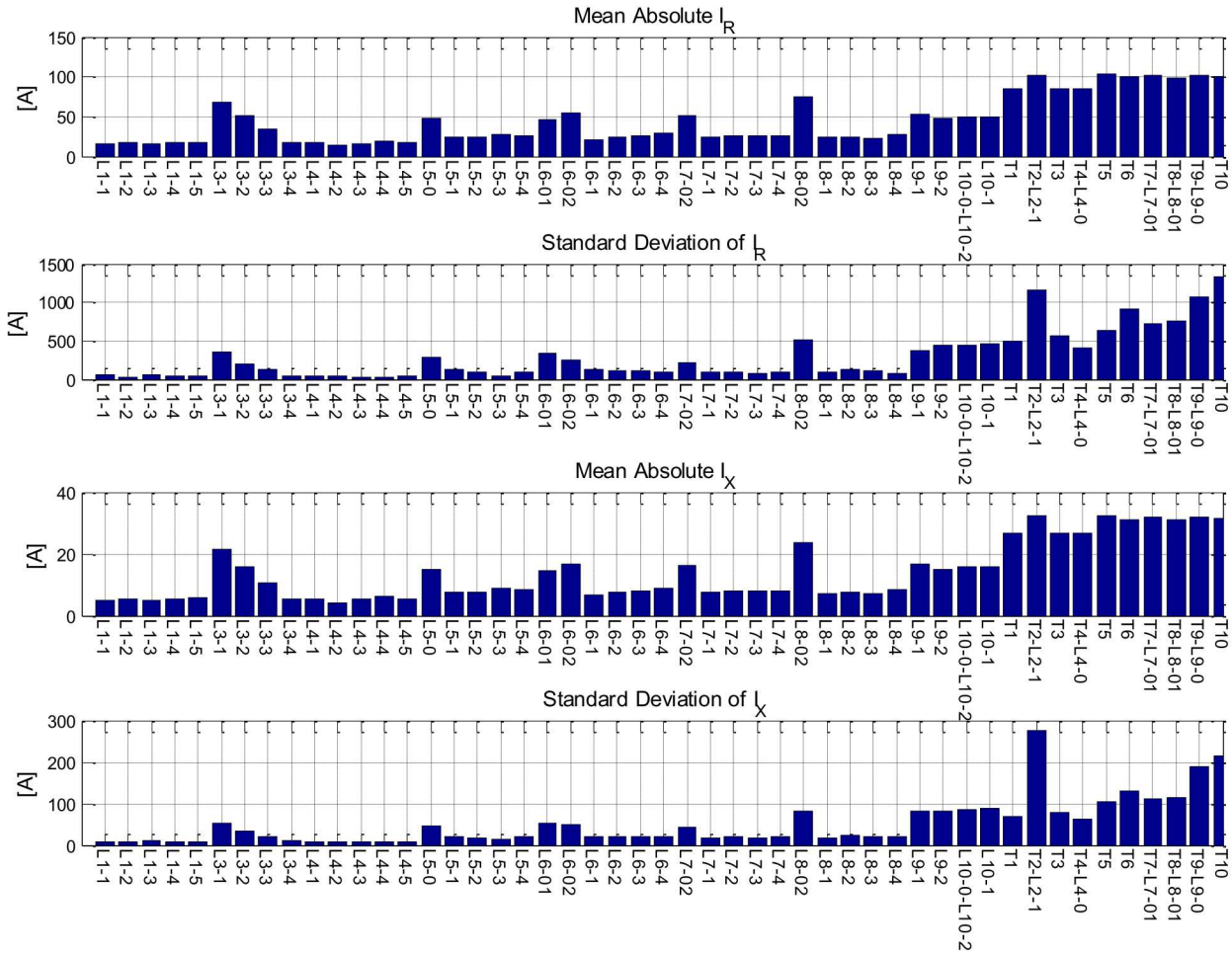

Figure 49. Means and standard deviations of the predictor variables I_R and I_X

Figure 50 shows the absolute and 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.

It should be noted that since parameter estimation results are data-driven, different results are obtained with different characteristics of load data and measurement error. As shown in Section 5.3, without measurement error, parameters can be estimated with a very small error especially if regression models with additional terms are used. The presented parameter estimation approach is optimized for the practical setting where the measurement error dictates the parameter estimation accuracy.

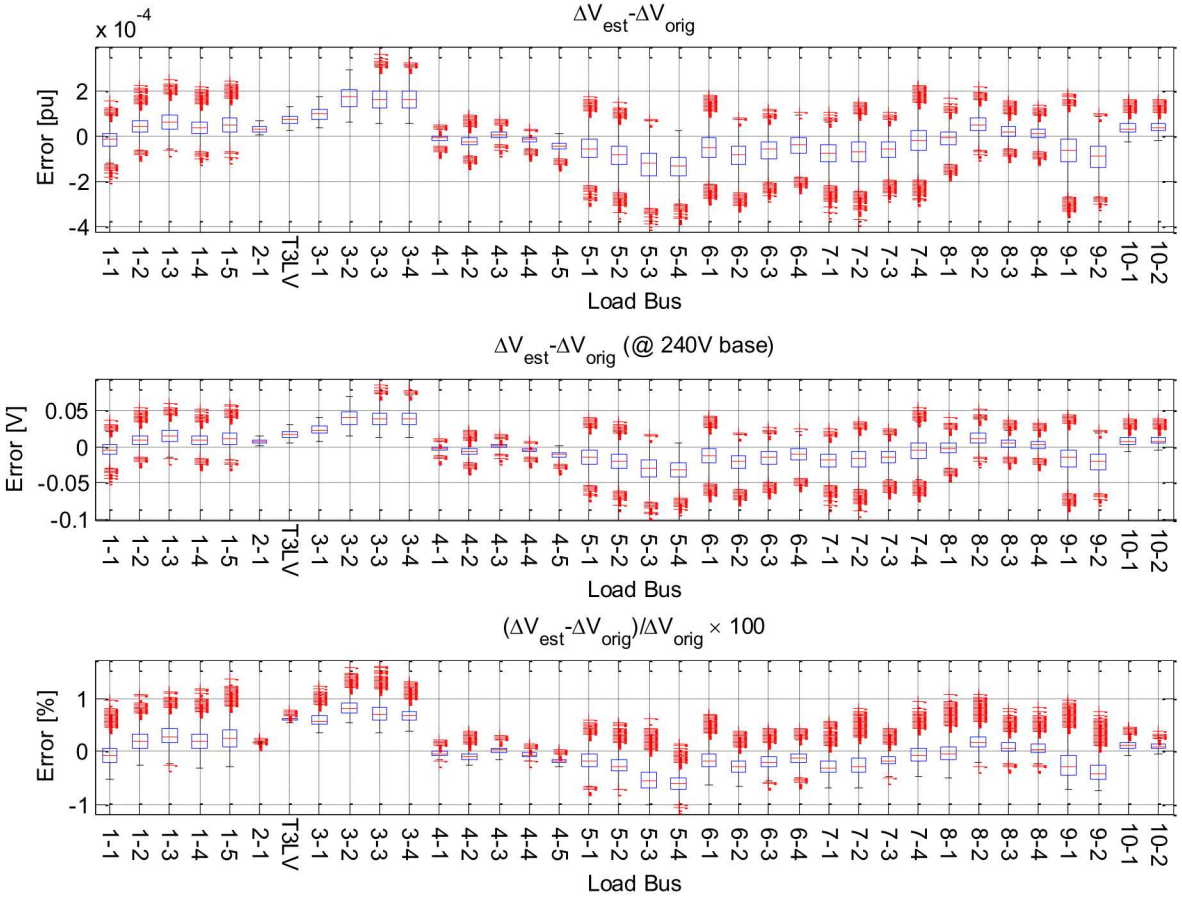

Figure 50. Errors of simulated voltage drops from the service transformer primary to the load buses when the parameters are estimated with the adaptive approach

7.2. Georgia Tech Feeder

The proposed parameter estimation algorithm was utilized to calibrate the secondary circuit parameters of one of the Georgia Tech feeders. Since the true parameter values are unknown, the parameter estimation accuracy was measured with mean bias errors (MBE $= \frac{1}{n}\sum_{i=1}^{n}\frac{(V_{sim}-V_{meas})}{V_{meas}}$) of the voltage drops simulated with the basecase parameters and the estimated parameters. The results are shown in Figure 51. The parameter estimation very effectively reduced the bias of the voltage drop simulation error.

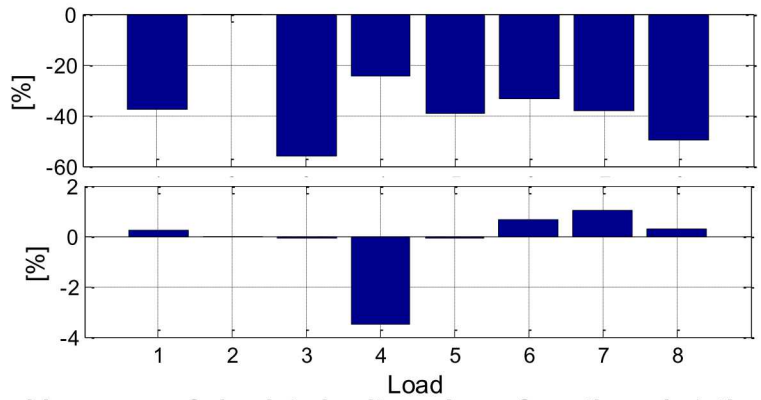

Figure 51. Mean bias errors of simulated voltage drops from the substation with basecase transformer parameters (top) and estimated transformer parameters (bottom)

Typically distribution system secondary circuits are fed by a single service transformer whose upstream and downstream bus voltages are rarely measured. In order to estimate the transformer impedance, voltage estimates of both of the transformer buses are needed. The downstream bus voltages can be estimated from the secondary circuit measurements with the hierarchical radial circuit parameter estimation approach shown above. However, the same approach cannot be used to estimate the transformer upstream bus voltages. Instead, the transformer upstream bus voltages can be estimated from time series power flow analysis by assuming that the distribution system primary circuit is well-modeled and any secondary system impedance inaccuracies have only a small impact on the transformer medium-voltage side voltages.

It is challenging to verify the Georgia Tech primary system model accuracy due to the absence of measurements between the loads in the secondary circuits and the substation. There is uncertainty with respect to the primary system underground cable capacitances and in some cases the exact primary system topology. Moreover, the Georgia Tech AMI historical database has errors, many of which have already been detected and removed but further undetected problems are likely to exist [33], [34].

Figure 52 shows the boxplots of the relative voltage simulation errors for all the 10 secondary circuit loads in the studied Georgia Tech feeder. The larger mean and range of errors are explained next. Meter 136E_ML1 records instantaneous measurements, which are subject to considerable variation in a given 15-min measurement time period. Since the other feeder measurements are 15-min averages, these instantaneous measurements do not synchronize well, so simulated voltages have more variation. Due to large research equipment in the building, meter B149E_MH2 has an abnormal load shape that is only at a few kW most of the time and occasionally jumps quickly up to 200-300kW for a while. Filtering out these jumps from the parameter estimation improves the estimated parameters and reduces the errors for test data periods without the jumps. However, voltage simulation errors during these jumps cannot be eliminated. Building B199E_MH1 has a lot of PV with negative power injections present during the daytimes. Parameter estimation effectively reduced the means and variations of the other loads.

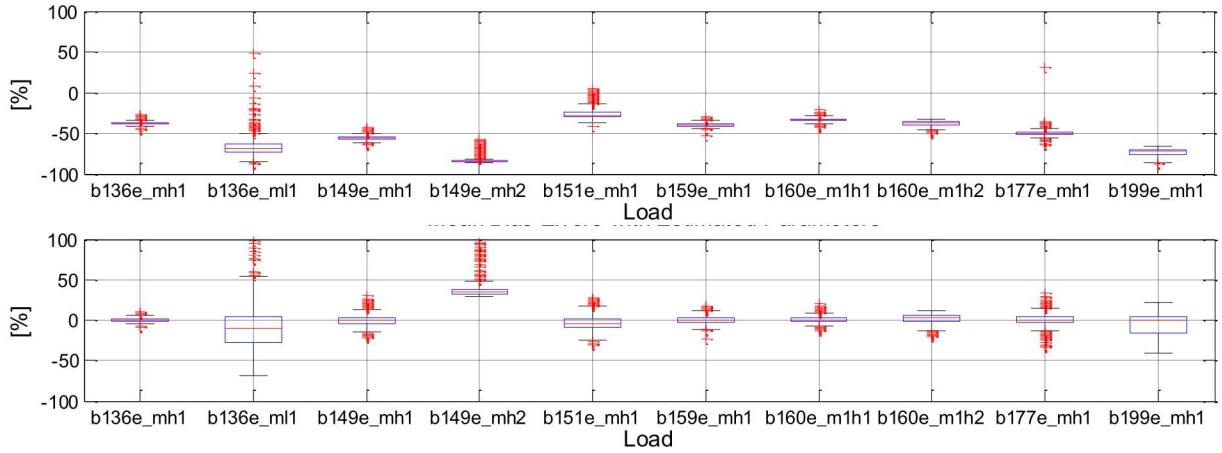

Figure 52. Relative voltage drop simulation errors with the basecase transformer parameters (top) and estimated transformer parameters (bottom)

8. HANDLING METERS WITHOUT VOLTAGE MEASUREMENTS

Smart meters and PV microinverters measure both voltages and currents to derive power measurements from them. Modern smart meters can allow firmware upgrades [49] to be able to transmit the voltage measurements to the utility database. However, older smart meters may not have either of these capabilities and thus, in practice some meters may provide only power (or current) measurements. This chapter presents a modified secondary circuit parameter estimation approach that can handle some meters that do not transmit voltage measurements. It should be noted that 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.

The modified approach presented in this chapter relies on the assumptions listed in chapter 4 except that all leaf nodes without voltage measurements must have current and power factor measurements. If a meter without voltage measurements has only power measurements, these power measurements must be converted to current measurements (I_R and I_X) by utilizing estimated (e.g. nominal or simulated) voltages. Once this has been done, the algorithm proceeds as follows to estimate the parameters of each secondary circuit:

1. Remove the upstream branches of leaf nodes without voltage measurements (their parameters cannot be estimated). Add the currents of such leaf nodes to the currents of the immediate upstream nodes. Set the list of other branches as the set of active branches.
2. While the set of active branches is not empty
    - If the list of active branches has only one branch, select the branch and go to 3.
    - If the list of active branches has a branch with both upstream and downstream voltage measurements, select the branch and go to 3.
    - If the list of active branches has a comprehensive (no other branches with the same upstream node) set of parallel branches that each have downstream voltage (and current) measurements, select the branches and go to 3.
    - If the list of active branches has a set of two or more parallel branches that each have downstream voltage (and current) measurements, select the branches and go to 3.
    - If the list of active branches has a branch with downstream voltage measurements, select the branch and go to 4.
    - If no suitable branch(es) found, return an error.
3. Estimate the parameters of the selected branch(es) either with "single branch model" or "pairwise parallel branch approach" shown in section 4. Go to 5.
4. Estimate the parameters of the selected branch by forming a regression problem between the branch downstream node and the closest upstream node with voltage measurements as shown in section 8.1. Go to 5.
5. Remove the selected branches from the set of active branches, add the currents of each branch to the currents of the immediate upstream branch and go to 2.

8.1. Estimating Series Branch Impedances

The regression problem for a set of N series branches can be formulated based on the voltage drop over the branches:

$$V_{Up} - V_{Down} = \sum_{i=1}^{N}(R_i I_{Ri} + X_i I_{Xi}), \tag{19}$$

where V_{Up} is the known voltage of the upstream node of the highest branch in the set and V_{Down} is the known voltage of the downstream node of the lowest branch in the set. The current components of branch i, I_{Ri} and I_{Xi}, can be calculated as a sum of the downstream branch currents of branch i. Consider T synchronous measurement samples $\boldsymbol{V}_{Up}, \boldsymbol{V}_{Down}, \boldsymbol{I}_{Ri}, \boldsymbol{I}_{Xi} \in \mathbb{R}^T, i \in \{1, \dots, N\}$ and define the response vector

$$\boldsymbol{y} = \boldsymbol{V}_{Up} - \boldsymbol{V}_{Down}, \tag{20}$$

the measurement (design) matrix

$$\boldsymbol{\mathcal{X}} = [\boldsymbol{I}_{R1}, \boldsymbol{I}_{X1}, \boldsymbol{I}_{R2}, \boldsymbol{I}_{X2}, \dots, \boldsymbol{I}_{RN}, \boldsymbol{I}_{XN}], \tag{21}$$

and the parameter vector

$$\boldsymbol{\beta} = [R_1, X_1, \dots, R_N, X_N]^T . \tag{22}$$

Then, the parameters $R_i, X_i, i \in \{1, \dots, N\}$ can be estimated from

$$\boldsymbol{y} = \boldsymbol{\mathcal{X}}\boldsymbol{\beta} + \boldsymbol{\epsilon} \tag{23}$$

with one of the approaches introduced in section 3.2. If the parameters of an upstream branch $i \in \{1, \dots, N\}$ are known (or estimated previously), branch i can be removed from the regression problem by doing the following steps.

1) Remove $\Delta\boldsymbol{V}_i$, the vector of voltage drops over branch i, from the response vector (20)

$$\boldsymbol{y} = \boldsymbol{V}_{Up} - \boldsymbol{V}_{Down} - \Delta\boldsymbol{V}_i. \tag{24}$$

2) Remove $\boldsymbol{I}_{Ri}$ and $\boldsymbol{I}_{Xi}$, the predictors of branch i, from the design matrix (21)

$$\boldsymbol{\mathcal{X}} = \left[\boldsymbol{I}_{R1}, \boldsymbol{I}_{X1}, \boldsymbol{I}_{R2}, \boldsymbol{I}_{X2}, \dots, \boldsymbol{I}_{R(i-1)}, \boldsymbol{I}_{X(i-1)}, \dots, \boldsymbol{I}_{R(i+1)}, \boldsymbol{I}_{X(i+1)}, \dots, \boldsymbol{I}_{RN}, \boldsymbol{I}_{XN}\right]. \tag{25}$$

3) Remove R_i and X_i, the coefficients of branch i, from the parameter vector (22)

$$\boldsymbol{\beta} = [R_1, X_1, \dots, R_{i-1}, X_{i-1}, R_{i+1}, X_{i+1}, \dots, R_N, X_N]^T. \tag{26}$$

If the parameters of multiple upstream branches are known, steps 1) – 3) are repeated for each branch with known parameters.

8.2. Results for the 66-Node Test Circuit

First, the operation of the algorithm is illustrated with different meters without voltage measurements in secondary circuit 6 of the 66-node test circuit (Figure 5). Figure 53 shows both the true parameters and the estimated parameters when all loads have voltage measurements. Figure 54 shows the estimated parameters when load 6-1 has no voltage measurements. In this case, branch L6-1 parameters are not estimated (and thus, are not show in the figure), branch L6-2 parameters are estimated with the approach described in section 8.1, and the rest of the parameters are estimated normally utilizing voltage measurements. Figure 55 shows the

estimated parameters when loads 6-1 and 6-4 have no voltage measurements. In this case, branch L6-1 and L6-4 impedances are not estimated (and thus, are not shown in the figure), branch L6-2, L6-3 impedances are estimated with the approach from section 8.1, and the rest of the branches are estimated normally. Finally, Figure 56 shows the estimated parameters when loads 6-1, 6-2, and 6-4 have no voltage measurements. In this case, branch L6_1, L6_2, L6_4, and L6_01 impedances are not estimated (and thus, not show in the figure), branch 6-3 impedance is estimated with the approach in section 8.1 and the rest of the parameters are estimated normally.

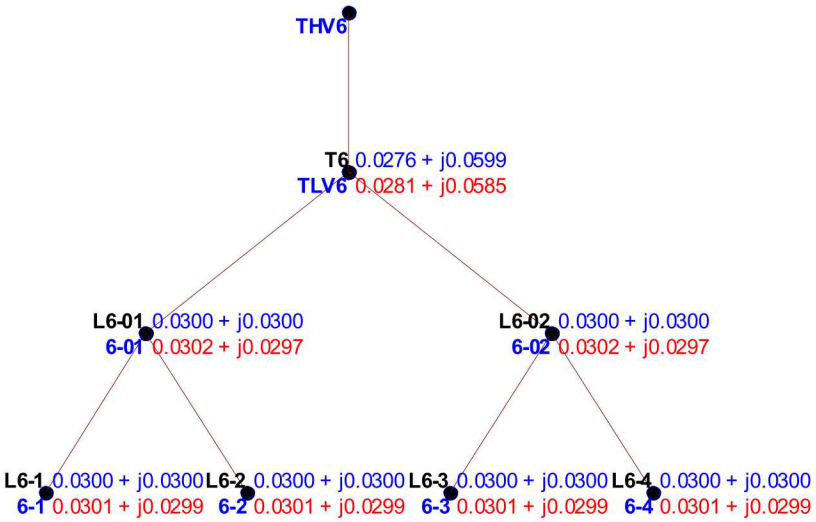

Figure 53. Original secondary circuit 6: node name (black bold), node upstream branch name (bold blue), branch true impedance (blue), and branch estimated impedance (red) , branches whose parameters are not estimated are not shown

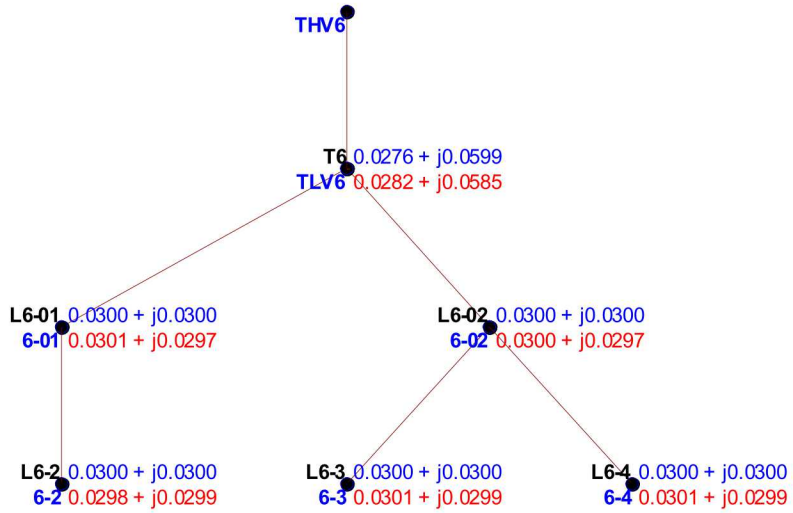

Figure 54. Estimated secondary circuit 6 parameters when load 6-1 has no voltage measurements: node name (black bold), node upstream branch name (bold blue), branch true impedance (blue), and branch estimated impedance (red) , branches whose parameters are not estimated are not shown

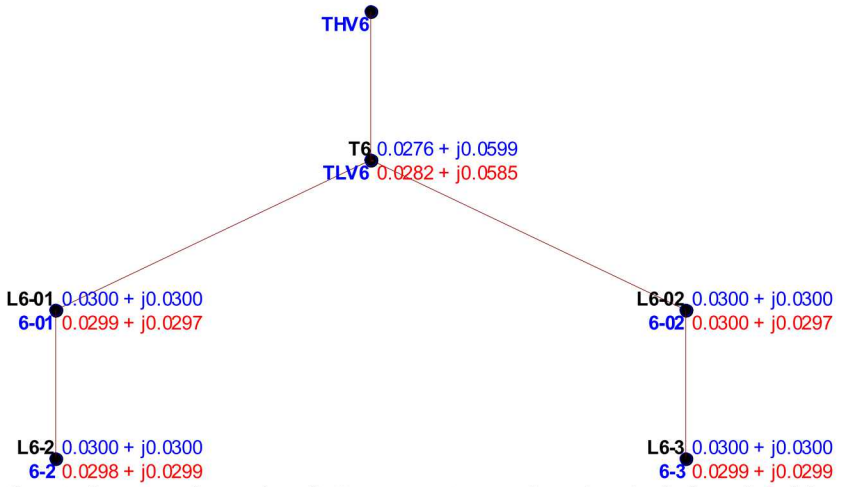

Figure 55. Estimated secondary circuit 6 parameters when loads 6-1 and 6-4 have no voltage measurements: node name (black bold), node upstream branch name (bold blue), branch true impedance (blue), and branch estimated impedance (red), branches whose parameters are not estimated are not shown

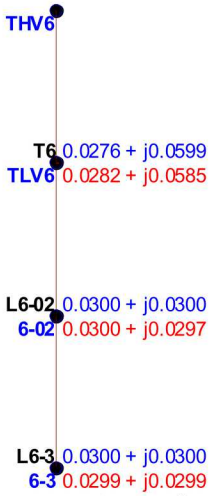

Figure 56. Estimated secondary circuit 6 parameters when loads 6-1, 6-2, and 6-4 have no voltage measurements: node name (black bold), node upstream branch name (bold blue), branch true impedance (blue), and branch estimated impedance (red) , branches whose parameters are not estimated are not shown

The modified algorithm was validated on the 66-node test circuit by first removing voltage measurements from a given number of randomly selected meters in each secondary circuit and then estimating the parameters. This was repeated for 50 times for each secondary circuit. All the meters were assumed to have current measurements, i.e., no conversion from powers to currents with estimated/simulated voltages was necessary. The parameters were estimated with a set of 8760 samples of perfect measurements (no measurement error). Figure 57 and Figure 58 show the average (over the 50 repetitions) absolute relative errors of the estimated R and X, respectively. The more meters without voltage measurements a given secondary circuit has, the higher the errors of the estimated parameters become. In some secondary circuit topologies, the errors increase more than on others. This is clearly illustrated in Figure 59 and Figure 60 that

show how much the average (over the 50 repetitions) absolute errors of the estimated R and X increase as the number of meters with no voltage measurements increases. It is also interesting to observe that in some cases, e.g., branches "L5_2" and "T3", the average parameter estimation error decreases when a meter is removed. This could potentially be explained by the particular load characteristics of the removed meter.

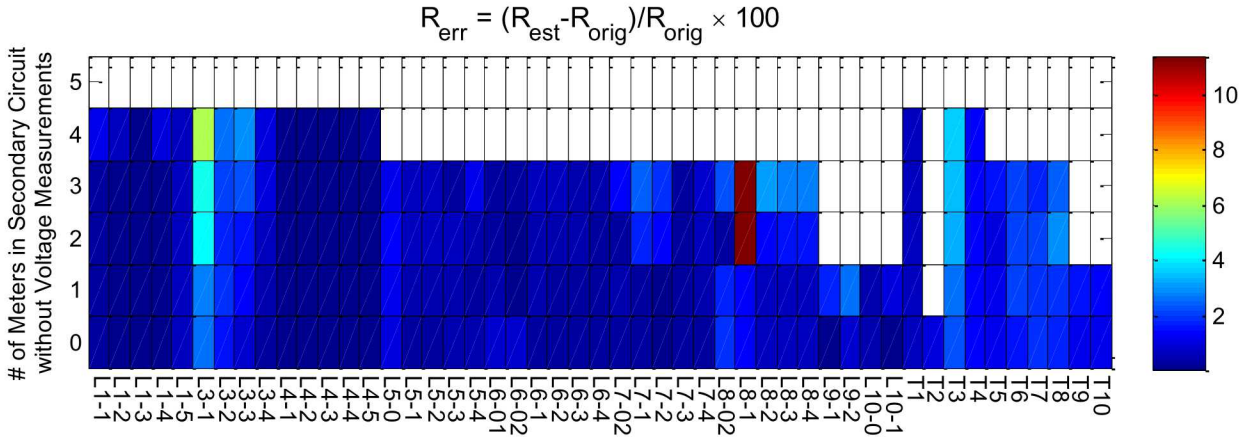

Figure 57. The average errors of the estimated R parameters over 50 repetitions where at each repetition a given number of randomly selected meters had no voltage measurements. In white areas, the parameter was not estimated in any of the repetition.

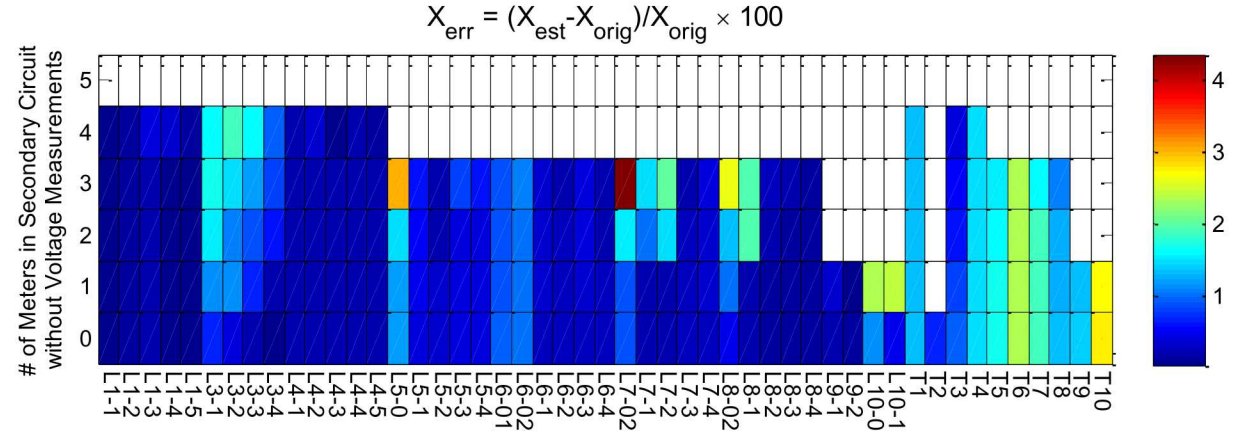

Figure 58. The average errors of the estimated X parameters over 50 repetitions where at each repetition a given number of randomly selected meters had no voltage measurements. In white areas, the parameter was not estimated in any of the repetition.

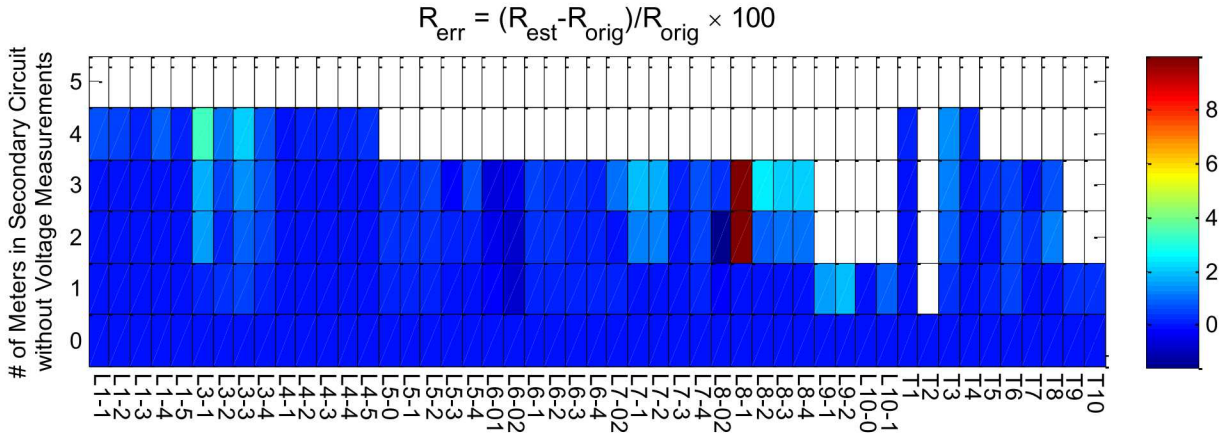

Figure 59. The impact of the number of meters with missing voltage measurements to the R estimation error (the difference of the results in Figure 57 compared to the case when all the meters have voltage measurements)

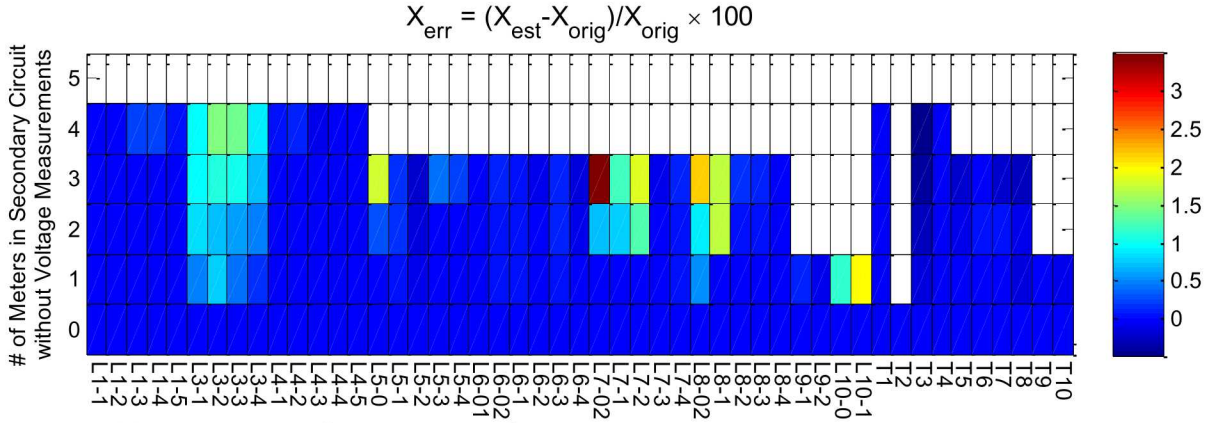

Figure 60. The impact of the number of meters with missing voltage measurements to the X estimation error (the difference of the results in Figure 58 to the case when all the meters have voltage measurements)

9. CONCLUSION

To analyze and coordinate distribution systems with rapidly increasing amount of PV, more accurate distribution system models are required, especially for the distribution system secondary circuits down to the point of common coupling for distributed PV. There is a growing need for automated procedures to calibrate the distribution system secondary circuit models that are typically modeled with a lower level of detail than the well-modeled medium-voltage systems.

This report presents an accurate, flexible, and computationally efficient method to use measurement data to estimate secondary circuit series impedance parameters in existing utility feeder models. The parameter estimation method assumes well-modeled primary circuit models, known secondary circuit topology, and AMI active power and reactive power measurements at all the loads in the secondary circuit. The method also requires AMI voltage measurements at most of the loads in the secondary circuit but can handle some meters without voltage measurements. The method is based on the well-known linearized voltage drop approximation and linear regression. The algorithm accuracy is studied with respect to various circuit parameters and operation conditions, including a range of circuit impedances, load power factor, the presence of reverse power flows, power and voltage measurement error, and sample size.

The performance of the developed method is demonstrated on a three-phase test circuit without and with power and voltage measurement error. The optimal regression model is shown to depend on the measurement noise and an optimal regression model selection is shown for the practical case with noisy measurement. The optimal regression model for the ideal case without measurement error estimates all the branch parameters with relative errors less than 1.5%. In the presence of 1% active power, 1% reactive power, and 0.2% voltage measurement error, the proposed optimal approach for noisy measurements estimates all the parameters with an average relative errors 2.46% and 3.11% for R and X, respectively. The average errors of the voltage drops over the secondary circuits, simulated with the estimated parameters, are all within 1% percent error of the actual voltage drops. These high levels of accuracy demonstrate the efficiency of the proposed methodology in improving secondary circuit models. The parameter estimation accuracy is shown to be much more sensitive to voltage measurement errors than power measurement errors, underlying the importance of having high-quality voltage measurement data.

The algorithm performance is also demonstrated on one the Georgia Tech campus distribution system feeders with AMI data. In this feeder, the algorithm very effectively reduced the bias of the voltage drop simulation error. Challenges related to real feeder models and AMI data are discussed.

In the future work, the proposed method is extended to single-phase secondary circuits and unknown secondary circuit topologies and applied on real utility feeder models.

REFERENCES

[1] "Technology Roadmap Solar Photovoltaic Energy 2014 Edition," International Energy Agency, 2014.

[2] "SEIA Solar Energy Industries Association," 2013. [Online]. Available: http://www.seia.org/. [Accessed: 10-Dec-2013].

[3] "Annual Energy Outlook 2015 with projections to 2040," U.S. Energy Information Administration, Apr. 2015.

[4] "HECO PV Penetration Maps," *Hawaiian Electric Maui Electric Hawai'i Electric Light*. [Online]. Available: http://www.hawaiianelectric.com/portal/site/heco/lvmsearch.

[5] "California Solar Statistics." [Online]. Available: https://www.californiasolarstatistics.ca.gov/.

[6] "California Solar," *SEIA*. [Online]. Available: http://www.seia.org/state-solar-policy/california.

[7] Y. Liu, J. Bebic, B. Kroposki, J. de Bedout, and W. Ren, "Distribution System Voltage Performance Analysis for High-Penetration PV," in *IEEE Energy 2030 Conference, 2008. ENERGY 2008*, 2008, pp. 1–8.

[8] J. Quiroz and M. Reno, "Detailed grid integration analysis of distributed PV," in *2012 38th IEEE Photovoltaic Specialists Conference (PVSC)*, 2012, pp. 000596–000601.

[9] 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.

[10] X. Feng, F. Yang, and W. Peterson, "A practical multi-phase distribution state estimation solution incorporating smart meter and sensor data," in *IEEE Power and Energy Society General Meeting*, San Diego, CA, 2012.

[11] T. Taylor and H. Kazemzadeh, "Integrated SCADA/DMS/OMS: Increasing Distribution Operations Efficiency," *Electric Energy Online*, Apr-2009.

[12] J. S. John, "Can Microinverters Stabilize Hawaii's Shaky Grid? : Greentech Media," 02-Feb-2015. [Online]. Available: http://www.greentechmedia.com/articles/read/enphase-to-help-hawaii-ride-its-solar-energy-wave.

[13] "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.

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

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

[16] W. Luan, J. Peng, M. Maras, and J. Lo, "Distribution network topology error correction using smart meter data analytics," in *IEEE Power and Energy Society General Meeting*, Vancouver, BC, Canada, 2013.

[17] 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.

[18] A. S. Debs, "Parameter estimation for power systems in the steady-state," in *1974 IEEE Conference on Decision and Control including the 13th Symposium on Adaptive Processes*, 1974, vol. 13, pp. 587–592.

[19] 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.
[20] J. C. S. Souza, A. M. Leite da Silva, and A. P. Alves da Silva, “Online topology determination and bad data suppression in power system operation using artificial neural networks,” *IEEE Trans. Power Syst.*, vol. 13, no. 3, pp. 796–803, Aug. 1998.
[21] X. Feng, F. Yang, and W. Peterson, “A practical multi-phase distribution state estimation solution incorporating smart meter and sensor data,” in *2012 IEEE Power and Energy Society General Meeting*, 2012, pp. 1–6.
[22] G. N. Korres and N. M. Manousakis, “A state estimation algorithm for monitoring topology changes in distribution systems,” in *2012 IEEE Power and Energy Society General Meeting*, 2012, pp. 1–8.
[23] A. Alimardani, S. Zadkhast, J. Jatskevich, and E. Vaahedi, “Using smart meters in state estimation of distribution networks,” in *2014 IEEE PES General Meeting | Conference Exposition*, 2014, pp. 1–5.
[24] A. Abdel-Majeed and M. Braun, “Low voltage system state estimation using smart meters,” in *Universities Power Engineering Conference (UPEC), 2012 47th International*, 2012, pp. 1–6.
[25] D. Ablakovic, I. Dzafic, R. A. Jabr, and B. C. Pal, “Experience in distribution state estimation preparation and operation in complex radial distribution networks,” in *2014 IEEE PES General Meeting | Conference Exposition*, 2014, pp. 1–5.
[26] M. Baran and T. E. McDermott, “Distribution system state estimation using AMI data,” in *IEEE Power Systems Conference and Exposition*, Seattle, WA, 2009.
[27] E. Manitsas, R. Singh, B. Pal, and G. Strbac, “Modelling of pseudo-measurements for distribution system state estimation,” in *IET-CIRED SmartGrids for Distribution*, Frankfurt, Germany, 2008.
[28] K. Samarakoon, J. Wu, J. Ekanayake, and N. Jenkins, “Use of delayed smart meter measurements for distribution state estimation,” in *2011 IEEE Power and Energy Society General Meeting*, 2011, pp. 1–6.
[29] 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.
[30] 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.
[31] B. Wang and W. Luan, “Generate distribution circuit model using AMI data,” in *2014 China International Conference on Electricity Distribution (CICED)*, 2014, pp. 1251–1255.
[32] “Pecan Street Inc.,” *Pecan Street Inc.* [Online]. Available: http://www.pecanstreet.org/about/.
[33] J. Peppanen, J. Grimaldo, M. J. Reno, S. Grijalva, and R. G. Harley, “Increasing distribution system model accuracy with extensive deployment of smart meters,” in *IEEE Power and Energy Society General Meeting*, Washington D.C., 2014.
[34] 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.

[35] J. Peppanen and S. Grijalva, “Neighborhood electric vehicle charging scheduling using particle swarm optimization,” in *2014 IEEE PES General Meeting | Conference Exposition*, 2014, pp. 1–5.
[36] J. A. Peppanen, T. Alquthami, D. Molina, and R. Harley, “Optimal PMU placement with Binary PSO,” in *IEEE ECCE*, 2012.
[37] J. Kennedy and R. Eberhart, “Particle swarm optimization,” presented at the IEEE International Conference on Neural Networks, 1995. Proceedings, vol. 4.
[38] J. Kennedy and R. C. Eberhart, “A discrete binary version of the particle swarm algorithm,” in *Systems, Man, and Cybernetics, 1997. Computational Cybernetics and Simulation., 1997 IEEE International Conference on*, 1997, vol. 5, pp. 4104 –4108 vol.5.
[39] T. Gönen, *Electric power distribution system engineering*, 2nd ed. Boca Raton: CRC Press, 2008.
[40] E. Lakervi and E. J. Holmes, *Electricity distribution network design*, 2nd ed. Stevenage, United Kingdom: Peter Peregrinus Ltd., 2003.
[41] 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.
[42] M. H. Kutner, C. Nachtsheim, J. Neter, and W. Li, *Applied linear statistical models*. Boston: McGraw-Hill Irwin, 2005.
[43] W. H. Kersting, “Center tapped transformer and 120/240 volt secondary models,” in *2008 IEEE Rural Electric Power Conference*, 2008, pp. A1–A1–7.
[44] “Uniform Business Practices for Unbundled Electricity Metering - Volume 2,” Edision Electric Institute, Dec. 2000.
[45] “ANSI C12.1 - Electric Meters Code for Electricity Metering.” American National Standards Institute (ANSI), 2008.
[46] M. J. Reno and K. Coogan, “Grid Integrated Distributed PV (GridPV) Version 2,” Sandia National Laboratories, SAND2014-20141, 2014.
[47] “Open Distribution System Simulator,” *EPRI*. [Online]. Available: http://sourceforge.net/projects/electricdss/.
[48] M. H. Kutner, C. J. Nachtsheim, J. Neter, and W. Li, *Applied Linear Statistical Models*, 5th ed. New York, NY: McGraw-Hill Irwin, 2005.
[49] “NEMA Smart Grid Standards Publication SG-AMI 1-2009 - Requirements for Smart Meter Upgradeability,” National Electrical Manufacturers Association, 2009.

10. DISTRIBUTION

1	MS1033	Robert J. Broderick	6112
1	MS1033	Abraham Ellis	6112
1	MS1140	Matthew J. Reno	6113
1	MS1140	Ross Guttromson	6113
1	MS0899	Technical Library	9536 (electronic copy)

Sandia National Laboratories