SANDIA REPORT
SAND2023-09849
Printed September 2023

Sandia National Laboratories

Distribution System Model Calibration for GMLC 3.3.3 “Incipient Failure Identification for Common Grid Asset Classes” – Project Summary

Matthew J. Reno
Logan Blakely

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

Issued by Sandia National Laboratories, operated for the United States Department of Energy by National Technology & Engineering Solutions of Sandia, LLC.

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@osti.gov
Online ordering:	http://www.osti.gov/scitech

Available to the public from
U.S. Department of Commerce
National Technical Information Service
5301 Shawnee Rd
Alexandria, VA 22312

Telephone:	(800) 553-6847
Facsimile:	(703) 605-6900
E-Mail:	orders@ntis.gov
Online order:	https://classic.ntis.gov/help/order-methods/

ABSTRACT

Distribution system model calibration is a key enabling task for incipient failure identification within the distribution system. This report summarizes the work and publications by Sandia National Laboratories on the GMLC project titled "Incipient Failure Identification for Common Grid Asset Classes". This project was a joint effort between Sandia National Laboratories, Lawrence Livermore National Laboratory, National Energy Technology Laboratory, and Oak Ridge National Laboratory. The included work covers distribution system topology identification, transformer groupings, phase identification, regulator and tap position estimation, and the open-source release and implementation of the developed algorithms.

ACKNOWLEDGEMENTS

Thank you to the other laboratories and participants in this project which included:

Lawrence Livermore National Laboratory: Indra Chakraborty (PI), Brenda Ng, Vaibhav Donde

Oak Ridge National Laboratory: Robert Warmack, Chris Villez, Yarom Polsky, Nance Ericson, Paul Growth, Kris Villez

National Energy Technology Laboratory: Jeffrey Wuenschell, Ruishu Wright, Michael Buric, Paul Ohodnicki

Sandia National Laboratories: Adam Summers, Rodrigo D. Trevizan, A. Furlani Bastos, Birk Jones, Jubair Yusuf, Joseph A. Azzolini, Bethany D. Peña

University of New Mexico: Aswathy R. Kurup, Ali Bidram, Manel Martínez-Ramón, Daniel R. Garcia, Binod Poudel

Georgia Institute of Technology: Santiago Grijalva, Samuel Talkington, Kavya Ashok, Deepak Divan

Electric Power Board Chattanooga

We would also like to thank Francis Therrien (Eaton), Cody Francis (Texas A&M), Vittal Rao (Texas A&M), Jouni Peppanen (EPRI)

CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

This report summarizes the work and publications produced by Sandia National Laboratories as part of the Grid Modernization Laboratory Consortium (GMLC) project titled, “Incipient Failure Identification for Common Grid Assets Classes” which covered January 2020 – January 2023. The project was led by Lawrence Livermore National Laboratory (LLNL) and included Oak Ridge National Laboratory (ORNL) and the National Energy Technology Laboratory (NETL). The project evaluated and operationalized multivariate, multimodal approaches to diagnose degradation and incipient failure of critical grid assets. The developed methodologies enable earlier detection of incipient failures, leveraging a multi-stage approached derived from healthcare analytics.

Sandia was responsible for several aspects of distribution system model calibration, including topology estimation, phase identification, and detecting changes in customer phase as they occur. Having accurate distribution system models is a critical factor in the incipient failure work as inaccuracies in the geographical information systems (GIS) model can directly impact the studies on individual component failure within the system. Analyzing the consequence of those failures depends on the accuracy of the models of the surrounding system. Distribution system models are known to have a variety of errors, [1], and are often considered to have too many errors to be trusted when running planning studies. However, there are data-driven methods to correct errors within the distribution system; details on Sandia’s prior work in this area can be found in [2], [3]. The work described in the following sections directly supports the final impacts of the incipient failure work of the larger project. In addition, an accurate distribution system model is critical for hosting capacity analysis [4].

Table I shows a breakdown of each of the areas of research pursued by Sandia over the course of this project. Work began in the left column with static distribution calibration tasks, moved to the middle column with dynamic adapting algorithms, and finally to outreach tasks designed to disseminate the work into industry. Each of the items in the left column regard the distribution system as static, a snapshot in time, and use historical data to calibrate the model for that instance. However, the distribution system is always changing and evolving through time with maintenance events, storm restoration, and new builds, necessitating algorithms that are capable of dynamically adapting to those changes as they occur (e.g. the algorithm in the middle column). That algorithm ingests data as it is obtained and can produce a dynamic model of the system through time. Finally, several of the more mature algorithms were released as open-source tools, in multiple locations, and Sandia partnered with a utility to do a field-demonstration and implementation of the algorithm within their system.

Sandia’s work within this project spanned all three years and resulted in 14 publications over the course of the work.

Table I - Breakdown of research and task areas

Grid Topology Estimation with Data Fusion	Dynamic Adapting Spatio-Temporal Models	Outreach
• Circuit Topology Identification, [5]–[10] • Transformer Groupings, [11], [12] • Regulator Controls and Tap Estimation, [13] • Phase Identification, [14]–[17]	• Online Phase Changepoint Detection, [18], [19]	• Release of Open-Source Code • Field Demonstration

2. DISTRIBUTION SYSTEM MODEL CALIBRATION

Broadly speaking the work detailed in the following sub-sections covers grid topology identification, fault identification, and phase identification in distribution systems. Both synthetic data and utility were used in the algorithm development and testing. The section is organized first by topic area and then by publications in that area.

2.1. Circuit Topology Identification & Fault Classification

Topology identification in transmission systems has historically been accomplished using supervisory control and data acquisition (SCADA) measurements. In distribution systems, however, SCADA measurements are insufficient to determine system topology. An accurate system topology is essential for distribution system monitoring and operation. Recently there has been a proliferation of Advanced Metering Infrastructure (AMI) by the electrical utilities, which improved the visibility into distribution systems. These measurements offer a unique capability for Distribution System Topology Identification (DSTI). The following publications summarize the algorithms developed during this project in the topology identification topic area.

2.1.1. "Substation-level Circuit Topology Estimation Using Machine Learning" [5]

This paper introduces an machine learning-based circuit topology estimation to be used for adaptive protection systems. An adaptive protection system relies on the communication system infrastructure to identify the latest status of the power grid (e.g., circuit topology or generation level of distributed energy resources). However, when the communication links are outaged due to physical damage or cyberattacks, the adaptive protection system may lose its awareness of the status of the system. Therefore, it is of paramount value to estimate the circuit status using the available healthy communicated data. The developed circuit topology estimation technique was verified on IEEE 123 bus test system. The case studies show that the proposed technique has a high accuracy for classifying the prevailing circuit topology. Among the utilized machine learning algorithms, a support vector machine (SVM) with a linear kernel function renders the best score. It is also shown that the impact of measurement noises on the accuracy of the proposed technique is minimal. Thus, the proposed algorithm has the potential to be valuable under real world communication failure due to abnormal events.

2.1.2. "Deep Learning Based Circuit Topology Estimation and Fault Classification in Distribution Systems" [6]

2.1.3. "Ensemble Models for Circuit Topology Estimation, Fault Detection, and Classification in Distribution Systems" [7]

This work developed a methodology to estimate the distributed circuit topology and analyze the occurrence of faults at different locations in a distribution circuit. The experiments discussed here show the performance of a convolutional neural network (CNN) based circuit topology estimation model. The algorithm uses a data-based approach to estimate the circuit configuration with faulty and normal data. The results are compared with that of a standard Support Vector Machine (SVM).

Furthermore, fault classification using an SVM is analyzed. The effectiveness of the proposed method is tested using data obtained from power simulations on the modified IEEE 123 bus system using MATLAB Simulink.
The simulated data contained four different circuit configurations and 3 fault types introduced at different locations in the system. The algorithm uses a CNN for circuit topology estimation, Figure 1, and a linear SVM for the classification of the faults. The CNN-based approach gives an average test accuracy of 96.06% for topology estimation.

Further, the results of the CNN were compared with that of a linear SVM and it was observed that the CNN outperforms the SVM by a significant margin in performance. For the fault classification, the test errors were found to be lower in the case of a three-class SVM classifier compared to that of a four-class SVM classifier. This result shows that it is good to have a fault detector before a three-class SVM for lowering the overall test error by a significant amount. In our future work, we will compare these results to conventional protection methods. Besides, we will expand this work by designing fault detection systems with lower test errors that could improve the performance of the overall fault classification system. Additionally, the sensitivity of the system against the changes in the structure and uncertainties in the system needs to be studied. The main idea would be to build a machine learning-based system that is robust to these uncertainties and changes in the structure over time.

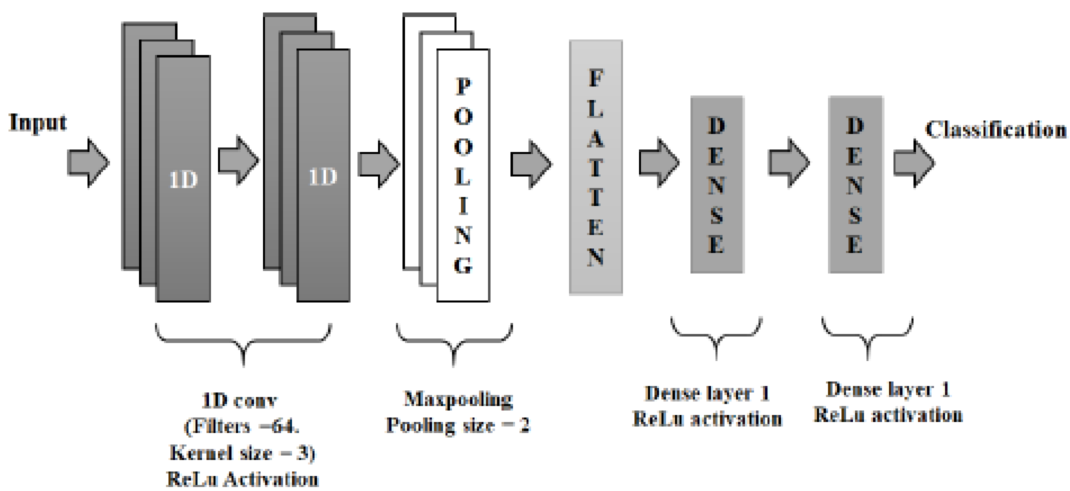


Figure 1 - CNN architecture

2.1.4. *“Enhancing Topology Identification of Distribution Power Systems Using Multi-Rate Voltage Measurement” [8]*

2.1.5. *"Topology Identification of Power Distribution Systems Using Time Series of Voltage Measurements," [9]*

An accurate system topology is essential for distribution system monitoring and operation. This work presents a novel approach to distribution system topology identification (DSTI). The proposed method identifies the topology of the distribution network in real-time by applying Linear Discriminant Analysis (LDA) and Regularized Diagonal Quadratic Discriminant Analysis (RDQDA) to the voltage magnitudes collected by distribution grid sensors. The results show that this method can utilize noisy voltage magnitude readings from load buses to identify distribution system reconfiguration between radial topologies during operation under changing loads in an accurate manner. Furthermore, the approach has shown good results for DSTI utilizing available multi-rate voltage measurements and for reduced-dimensionality input data using principal component analysis. The proposed DSTI approaches have outperformed convolution neural networks under limited training data.

To validate the performance of both DSTI algorithms presented in this paper, multiple simulations were run using both LDA and RDQDA algorithms on a test system. The simulation, and the distribution topology used for validation, are described in detail in the following section. The configuration of the IEEE 123-bus system with 8 switches was used for training and validation of the proposed algorithm for DSTI. This system has 5 radial topologies, therefore there exist 20 possible topology transitions. Additionally, there are 5 classes that represent unchanged topology between measurement scans. In total there are 25 possible switch transitions that represent the classes for the topology problem. A combination of the programs OpenDSS Version 7.6.5.91, MATLAB 2020a, and MATLAB Toolbox GridPV Version 2.2, was used to generate the synthetic data and simulate the grid operations, construct the library, and validate its successful predictions for the DSTI algorithm.

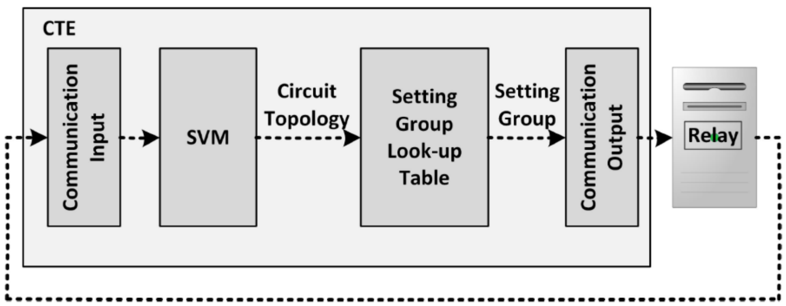
These methods were tested using the simulation data for the IEEE 123-Bus test system over a 9-month testing period with a prior 3-month training period. The accuracies of the topology identification are very good. The method has shown robustness to load variations and noise in measurements. This proposed method does not require precise knowledge of distribution system parameters but does require a dataset of voltage magnitudes correctly labeled for each topology. The proposed identification techniques were extended to accommodate the various sampling periods and multi-rate measurements using the most dominant measurements determined by using principal component analysis methods. The novel algorithms were compared to a CNN-based approach, Table 2, and the simulation results were very satisfactory.

Table 2 - Comparison between novel algorithms and a CNN-based approach

Trial #1		**Trial #2**		**Trial #3**	
% Correct	***False Positive***	***% Correct***	***False Positive***	***% Correct***	***False Positive***
CNN					
83.3	35	80.0	26	84.0	23
LDA					
98.7	2	98.0	1	98.0	4
RDQDA					
98.0	10	96.7	5	98.7	8

2.1.6. *"Circuit Topology Estimation in an Adaptive Protection System" [10]*

The goal of this work was to utilize machine learning (ML) techniques for estimating the distribution circuit topology in an adaptive protection system. In a reconfigurable distribution system with multiple tie lines, the adaptive protection system requires knowledge of the existing circuit topology to adapt the correct settings for the relay. Relays rely on the communication system to identify the latest status of remote breakers and tie lines. However, in the case of communication system failure, the performance of the adaptive protection system can be significantly impacted. To tackle this challenge, the remote circuit breakers and tie lines' status are estimated locally at a relay to identify the circuit topology in a reconfigurable distribution system. This paper utilizes Support Vector Machine (SVM) to forecast the status of remote circuit breakers and identify the circuit topology. The effectiveness of the proposed approach is verified on two sample test systems.



With the proposed circuit topology estimation (CTE) approach, the remote circuit breakers and tie lines' status are estimated locally at a relay to identify the circuit topology. An SVM was used as the ML classifier. The effectiveness of the proposed approach was verified on two sample test systems. The simulation results show that CTE's accuracy in determining the circuit topology is more than 96%, which verifies the accuracy of the proposed CTE approach.

2.2. Transformer Groupings

2.2.1. *"Identification and Correction of Errors in Pairing AMI Meters and Transformers" [11]*

2.2.2. *"Identifying Errors in Service Transformer Connections" [12]*

This work proposed a data-driven, physics-based approach for grouping residential meters downstream of the same service transformer. The proposed method involves a two-stage approach that first uses correlation coefficient analysis to identify transformers with errors in their customer grouping then applies a second stage, using a linear regression formulation, to correct the errors. This research used EPRI's Ckt 5 model containing 1379 customers and 591 transformers. The proposed method achieves >99% accuracy in identifying the transformer customer groups in the presence of measurement noise and missing data on a feeder with 1379 residential customers and 10% of those customers having incorrect transformer labels. This demonstrates algorithm robustness in the presence of common data quality issues. The method also provides a clear improvement upon two other similar methods in literature.

2.3. Regulator Controls and Tap Estimation

2.3.1. "Data-Driven Methods for Voltage Regulator Identification and Tap Estimation" [13]

In this work, data-driven methods are developed to characterize several physical parameters of voltage regulators and estimate their historical tap position states by leveraging existing measurement infrastructure on distribution grids. Specifically, the methods can differentiate if voltage regulation follows per phase or gang-operated load tap changers (LTCs) operation, estimate the spacing between tap positions, and estimate the total number of tap positions. The methods are tested on actual utility datasets from two different distribution feeders and the results are compared to the utility-verified characteristics where applicable. The impacts of different types of measurements (instantaneous vs. averaged) are also analyzed to identify the data best suited for deploying these methods by utility operators. Overall, the proposed methods are intuitive, effective, and require minimal prior knowledge of the underlying system, making them practical and useful tools for improving model fidelity.

Only the measurements of a recloser are required for the first task, while two reclosers (one on each side) are needed to estimate the tap positions. It should be noted that while the voltage measurements came from intelligent reclosers in this work, the methods apply to voltage measurements from any capable device. The method successfully determines both the tap position states and the total number of tap position changes. Also, while the methodology proposed here was applied to historic measurement data, it can be implemented on measurement streams to provide tap position estimations to inform grid operators continuously as new data arrives. The retrieval of tap changes by the proposed approach will help to calibrate the regulator settings in distribution system models and create logs of device operations that can be used to inform maintenance decisions. Lastly, the historical tap positions can be beneficial to track events such as conservation voltage reduction (CVR) events and ensure they were carried out as intended. The key takeaways from this work include:

- Downstream voltage measurements alone can be used to find the controlled phase of the LTC and to differentiate between a regulator and an LTC
- Voltage differences between upstream and downstream measurements can be utilized to estimate tap positions per phase for any line regulator
- Optimal tap positions can be extracted by maximizing the overall sum of the cosine similarities between the closest tap positions and the measurements from the reclosers
- Instantaneous measurements are more suitable to estimate the tap positions of regulators and LTCs, while averaged measurements are more suitable for identifying and distinguishing regulators and LTCs
- The spacing between tap positions remains unchanged regardless of the time resolution of the measurements. These straightforward approaches can be easily implemented to help reduce the erroneous operations occurring from incorrect modeling, misinterpreting the tap changes in real-time, and delaying the replacement of failed devices.

2.4. Phase Identification

2.4.1. "Assessment of Measurement-Based Phase Identification Methods" [14]

The task of determining the phase connection of customers, known as phase identification, is crucial to obtain accurate distribution system models, [20]. This work was conducted in collaboration with CYME, where CYME conducted testing of several phase identification methodologies, including one developed by Sandia. This work began with a thorough literature review of the existing phase identification methods, which are broadly divided into three categories: hardware-based, real power-based, and voltage-based methods. This is followed by multiple case studies assessing the accuracy of six real power- and voltage-based phase identification algorithms on four realistic distribution test systems. Synthetic load profiles along with network models are used to quantify the accuracy of each method for different scenarios: varying advanced metering infrastructure (AMI) coverage, number of initially mislabeled customer phases, number of available samples, and measurement noise. A case study using a real AMI dataset, including field verification, was also conducted

This work tested six state-of-the-art phase identification algorithms using AMI data. Four methods were based on voltage time series: ensemble spectral clustering with geographical information system (GIS) phasing (ESC-GIS) developed by Sandia, ensemble spectral clustering with SCADA measurements (ESC-SCADA), principal component analysis, and multi-tree algorithm. Two were power-based methods, LASSO and salient frequency. The ESC-GIS method performed the best in all of the testing conducted in this work, showing robustness to differing feeder configurations and data collection concerns. ESCGIS requires AMI voltage time series and existing utility phase labels where more than 50% are believed to be accurate. ESC-SCADA removes the requirement for existing phase labels but adds a requirement for SCADA data at the substation, and in that case, there is a decrease in performance. If utility labels and SCADA measurements are unavailable, ensemble spectral clustering can be conducted, and the final phase is left to manual verification; this would require only AMI voltage time series but adds a small manual verification step at the end of the method. If voltage AMI data is not available, the LASSO method performed the best out of the two power methods; it requires real power time series and real power SCADA measurements at the substation. If AMI voltage and either real power AMI or real power SCADA data are not available, then traditional phase identification methods must be used, such as manual verification or hardware-based methods. The ensemble spectral clustering methods (if voltage AMI data is available) and the LASSO method (if only real power AMI data is available) are both shown to be good choices for the distribution system phase identification task under a variety of conditions.

2.4.2. "Parameter Tuning Analysis for Phase Identification Algorithms in Distribution System Model Calibration" [15]

A challenge associated with developing algorithms for model calibration tasks is the determination of parameters for a particular algorithm. This paper presents a methodology for selecting parameters for distribution system model calibration algorithms utilizing distance matrices and demonstrated using a spectral cluster ensemble phase identification algorithm and allowing these parameters to be tuned on a per-feeder basis. This method leverages cluster analysis and the distance matrices often produced by phase identification methods. The proposed method was tested on 5 feeders from 2 different utilities to select the number of clusters used in a spectral clustering phase identification algorithm. A synthetic dataset was then used to validate the method with the phase identification algorithm performing with 100% accuracy.

The proposed methodology uses modified silhouette metrics to select parameters to optimize the clustering algorithm. Silhouette metrics are well established clustering quality metrics that were specially modified for this situation. This method was used to select the number of clusters, and window size in the spectral clustering phase identification algorithm. Phase identification with the chosen parameters was performed on five utility datasets, and the method's effectiveness was validated on a synthetic data set. Although the method was used to select parameters for the spectral clustering phase identification algorithm, this method for parameter selection can be broadly applied to other phase identification algorithms and even other tasks where a distance matrix in conjunction with clustering is applied. The proposed methodology clearly demonstrates the ability to tune algorithm parameters for datasets with unknown characteristics

2.4.3. *"Leveraging Additional Sensors for Phase Identification in Systems with Voltage Regulators" [16]*

The use of grid-edge sensing in distribution model calibration is a significant aid in reducing the time and cost associated with finding and correcting errors in the models. This work proposes a novel method for the phase identification task employing correlation coefficients on residential advanced metering infrastructure (AMI) combined with additional sensors on the medium-voltage distribution system to enable utilities to effectively calibrate the phase classification in distribution system models algorithmically. This type of data fusion leverages the power of both types of sensors, the AMI data has a lower decimal resolution while the sensor data has a much higher decimal resolution. The proposed method was tested on a real utility feeder of ~800 customers that includes 15-min voltage measurements on each phase from IntelliRupters© and 15-min AMI voltage measurements from all customers. The proposed method is compared with a standard phase identification method using voltage correlations with the substation and shows significantly improved results. The final phase predictions were verified to be correct in the field by the utility company. Although the feeder in this work uses IntelliRupter© sensors, this method generalizes to other, similar grid sensors. Correlation coefficient analysis is leveraged to provide a predicted phase for each customer along with easily interpretable confidence metrics for each prediction. This method is compared to a similar method that employs only substation data. The proposed method is shown to achieve significantly improved performance over the algorithm only using substation data. This is particularly important in the case where there are voltage regulators on the feeder. The separation between AMI meters and the substation data by the voltage regulator renders correlations with the substation unusable. Using the additional grid sensors, combined with customer AMI data and correlation coefficient analysis, provides a straightforward, interpretable solution to the phase identification task.

After the publication of the prior referenced work, collaboration with the utility partner continued. The utility partner provided field-validation for each of the customers on that first feeder predicted to be incorrect by the phase identification algorithm, a total of 6 high-confidence predictions. 100% of those customers predicted to be incorrect were shown to be accurate in the field verification. Subsequently, the algorithm was tested on three more feeders from that utility and the algorithm was again field validated to be 100% accurate in the customers predicted to have a different phase from the original utility model.

Figure 2 shows a satellite view of five customers which were incorrect in the original utility model (top view) and were corrected by the phase identification algorithm (bottom view).

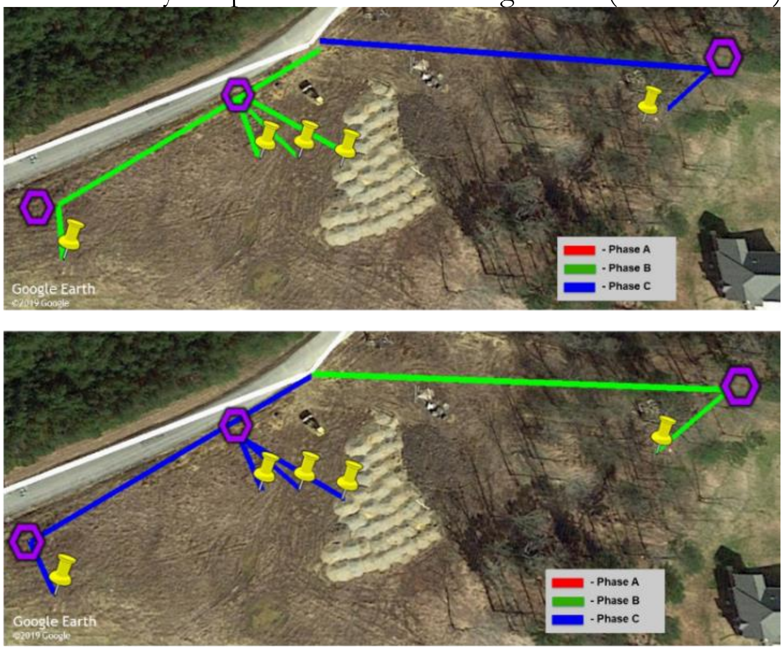


Figure 2 - Satellite image showing five customers on two laterals predicted to have incorrect phase labels. The original utility labels (top) show four Phase B customers and one Phase C customer and the predicted labels (bottom), verified with field verification

2.4.4. *"Distribution System State Estimation Sensitivity to Errors in Phase Connections" [17]*

High penetration of distributed energy resources presents challenges for monitoring and control of power distribution systems. Some of these problems might be solved through accurate monitoring of distribution systems, such as what can be achieved with distribution system state estimation (DSSE). With the recent large-scale deployment of advanced metering infrastructure associated with existing SCADA measurements, DSSE may become a reality in many utilities. This work developed a sensitivity analysis of DSSE with respect to phase mislabeling of single-phase service transformers, another class of errors distribution system operators are faced with regularly. The results show DSSE is more robust to phase label errors than a power flow-based technique, which would allow distribution engineers to more accurately capture the impacts and benefits of distributed photovoltaic (PV) systems.

The results show that DSSE is a promising tool not only for protective distribution system (PDS) monitoring but also for the calibration of PDS parameters to improve visibility into distributed energy resources (DER) operations and planning. Further, even under errors in topology, the voltage estimates by DSSE are highly accurate. These results show that DSSE can be a much more reliable tool than PF for distribution system studies. Figure 3 clearly shows the improvements in RMSE voltage error for the DSSE approach (orange) versus the standard power flow (PF) approach (yellow). Also, identification of these phase errors can be achieved through DSSE error processing. Even though the results for the identification of phase mislabeling are satisfactory, they still leave room for improvement.

DSSE can be a good tool for developing base cases for hosting capacity studies, which would allow distribution engineers to more accurately capture the impacts and benefits of distributed PV in PDS. Further, integrating DSSE into voltage control of PDS implemented by DER management systems could increase its effectiveness due to improved visibility into the grid even if there exist some errors in the labels of phases of distribution transformers.

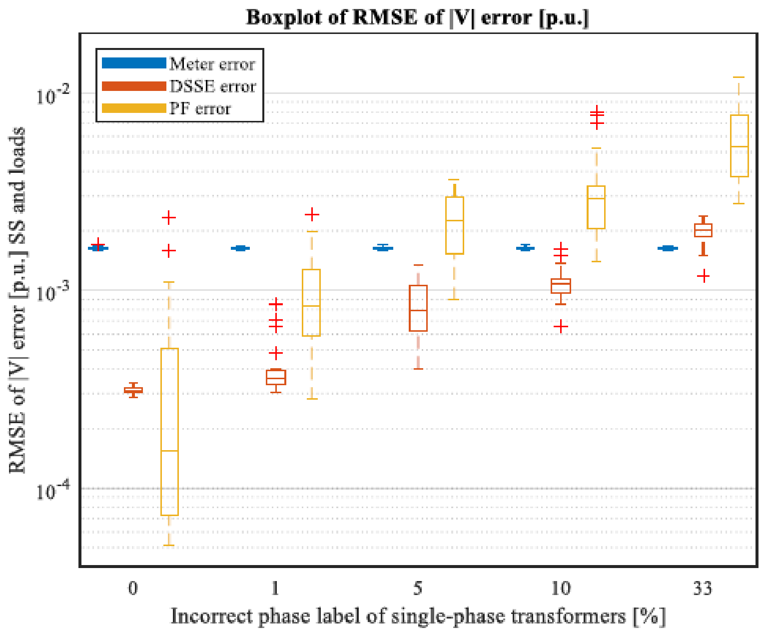


Figure 3 - Distribution of errors in RMSE voltage for a standard power flow (PF) approach versus the proposed DSSE approach

2.5. Online Phase Changepoint Detection

Developing an algorithm to detect customer phase changes as they occur was the primary focus of Sandia in year three. Two conference papers were published on this work, see below. The first focused on using historical data, and the second continued development to shift that algorithm into an online version that uses the AMI data as it becomes available. This summary focuses on that final version of the algorithm.

2.5.1. "Online Data-Driven Detection of Phase Changes in Evolving Distribution Systems" [19]

2.5.2. "Data-Driven Detection of Phase Changes in Evolving Distribution Systems" [18]

During a given year, the distribution system may experience a variety of changes and event occurrences; this poses a challenge in maintaining accurate and up-to-date models. This work proposes a method for the detection of phase change events in an online fashion and with small data requirements. The proposed algorithm uses spectral clustering to obtain predicted voltage phases as new segments of data are obtained so that the predicted phases can be observed over time; this is based on Sandia's prior phase identification work. Additionally, this work proposes a set of metrics used to evaluate the confidence of the clustering of an individual window and over time. Many phase identification algorithms require months of data, but this method demonstrates that many customers can be determined with significantly less data. 40% of customers were correctly identified within 8 days and 99% within 20 days. The proposed algorithm was tested on a synthetic dataset with simulated phase change events. The tests using the synthetic data successfully detected 100% of phase change events while incurring zero false positive events. This dramatically reduces the average time required to detect phase change events in an evolving distribution system.

The key innovation in this work is the creation of the time duration curve, Figure 4, to determine if a possible event is truly a phase change event or just a noisy, incorrect phase prediction for that window. The number of windows since a possible event is shown on the x-axis, and the cumulative confidence score is shown on the y-axis. Any possible event which falls in the red-shaded region is decided to be a true event, otherwise it is not considered a true event. This allows for high-confidence phase changes to be detected very rapidly, while less confident phase changes require more windows to be decided as a true event.

This algorithm was also tested on a set of utility data and identified an event with high confidence. A satellite view of these customers is shown in Figure 5. Ten customers were identified to have changed phase together, from Phase A to Phase C, on August 6, 2016 at 11pm. We speculated that this was due to a maintenance event. This algorithm will greatly enhance a utility's visibility into their constantly evolving distribution system.

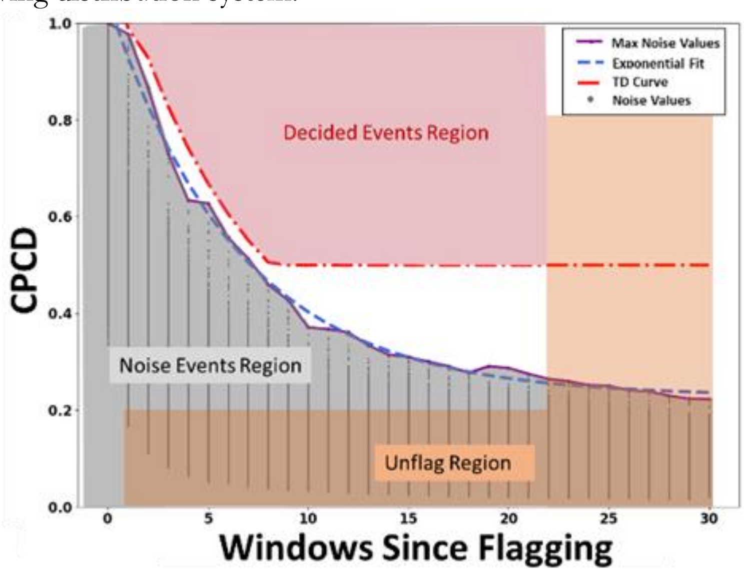


Figure 4 - Time duration curve shown in red with key regions shaded

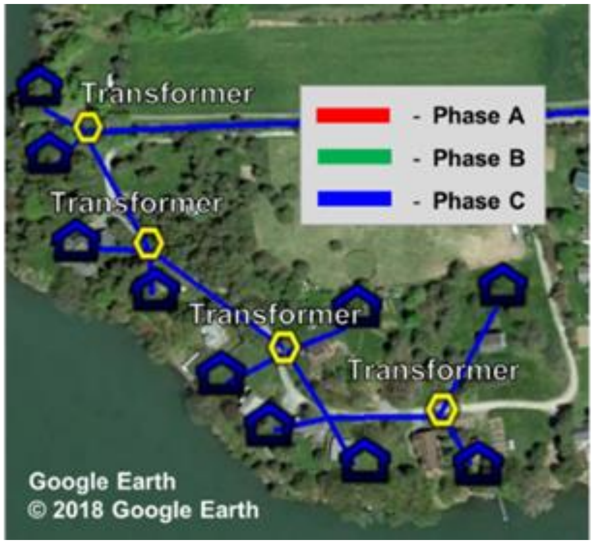


Figure 5 - 10 customers flagged for change from phase A to phase C, on 8/6/2016 at 11:00pm

This page left blank

3. OPEN-SOURCE CODE RELEASE

Several of the algorithms developed over the course of this project were released as open-source and available to the public. In partnership with National Rural Electric Cooperative Association (NRECA) one of the phase identification algorithms was also implemented within their Open Modeling Framework (OMF) tool suite. Sandia also assisted Electric Power Board (EPB) Chattanooga in implementing the phase identification algorithm within their own system.

3.1. GitHub release of Online Phase Change Point Detection Algorithm

The online phase change detection [19] code has been released to the Sandia GitHub, as part of the Distribution System Model Calibration repository. https://github.com/sandialabs/distribution-system-model-calibration. The code was released under the open-source BSD-3 license and is now publicly available for anyone to use. Figure 6 shows an example snippet from the output csv file produced by the algorithm. Events are labeled with customer ID's, event locations, confidence scores, etc. A set of sample data was also released to ease of use and testing by the users.

1		CustID	Event Location	Length Before	Cumulative TPP Before	Length After	Cumulative TPP After	Event Phase	Meets TD Req	Status	Current Window
2	0	Customer_139	6	6	0.91736227	19	0.071594878	3	FALSE	unflagged	25
3	1	Customer_139	7	7	0.093771868	18	0.889547582	2	FALSE	possible	25
4	7	Customer_182_cp	5	5	0.965236686	20	0.92121082	3	TRUE	event	25

Figure 6 - Example csv output from the phase change detection algorithm

3.2. NRECA OMF Release of Phase Identification Algorithm

Sandia collaborated with NRECA to provide access to the phase identification algorithm for electric coops. The phase identification algorithm is now live in the NRECA Open Modeling Framework (OMF) tool suite, omf.coop/. OMF provides open-source access to state-of-the-art analytics tools; the tool suite is primarily focused on granting access to coops to those types of new analytics technologies. Sandia is currently in discussion with several coops who have expressed interest in using the phase identification tool. Two ADMS vendors have also reached out to SNL regarding interest in the phase identification tools.

Having the phase identification tool available in the OMF tool suite is an excellent step toward the wide-spread usage of this research in industry. Coops serve the majority of the land mass within the United States and regularly use the OMF tool suite to further their grid modernization goals. Over 260 coops and vendors use the OMF tool. Inclusion in the OMF tool also promotes energy justice by enabling affordable solar access in persistent poverty counties, 92% of which get their energy from electric coops.
For more details on this release, please see [21].

3.3. Implementation with Electric Power Board Chattanooga

Finally, Sandia collaborated with Electric Power Board Chattanooga (EPB) to assist them in implementing the sensor-based phase identification algorithm within their own system. The algorithm was successfully implemented and tested with excellent results. EPB is currently in the process of rolling out the algorithm for wide-spread usage within their service area.

This page left blank

4. SUMMARY

Each of the data-driven distribution model calibration methodologies described above feed into the incipient failure consequence modeling and play a critical role in ensuring that those models produce accurate results. The existing distribution system models are not accurate enough to perform these (and other) types of tasks. Better models are required to move this goal forward. Data-driven model calibration algorithms are required to rapidly calibrate and update these utility models. Sandia has demonstrated multiple algorithms across several research focus areas that accomplish this task. Algorithms in Section 2.1 – 2.4 are static algorithms which use historical data to calibrate the model at a particular instance in time. In Section 2.5, this was expanded to run the algorithms online, to calibrate the utility models as the changes occur, enabling utilities to keep up with the rapid changes within their system which are happening all the time. Sandia has demonstrated the value of data fusion in these types of algorithms, combining AMI data from the low-voltage system with sensor data from the medium voltage system to achieve high-accuracy and high confidence results. The team released open-source code both to GitHub and through NRECA, enabling electric coops, universities, and the public access to this research. Sandia also partnered with a utility to conduct field demonstrations of the work within their service area and worked with them to implement the software within their own analytics system. This body of work clearly shows the applicability and the necessity of data-driven methods for distribution system model calibration. For more details on the algorithms and results, please see the referenced publications below.

REFERENCES

[1] L. Blakely, M. J. Reno, and J. Peppanen, "Identifying Common Errors in Distribution System Models," in *Photovoltaic Specialists Conference (PVSC)*, Chicago, IL, USA: Photovoltaic Specialists Conference (PVSC), Jun. 2019.

[2] K. Ashok, M. J. Reno, D. Divan, and L. Blakely, "Systematic Study of Data Requirements and AMI Capabilities for Smart Meter Connectivity Analytics," in *IEEE Smart Energy Grid Engineering (SEGE)*, Oshawa, Ontario, Canada, 2019.

[3] L. Blakely, M. J. Reno, and K. Ashok, "AMI Data Quality And Collection Method Consideration for Improving the Accuracy of Distribution System Models," in *IEEE Photovoltaic Specialists Conference (PVSC)*, Chicago, IL, USA, 2019.

[4] J. A. Azzolini, S. Talkington, M. J. Reno, S. Grijalva, L. Blakely, and D. Pinney, "Improving Behind-the-Meter PV Impact Studies with Data-Driven Modeling and Analysis," in *IEEE Photovoltaic Specialists Conference (PVSC)*, Jun. 2022.

[5] D. R. Garcia, B. Poudel, A. Bidram, and M. J. Reno, "Substation-level Circuit Topology Estimation Using Machine Learning," in *IEEE Innovative Smart Grid Technologies (ISGT)*, 2022.

[6] A. R. Kurup, M. Martínez-Ramón, A. Summers, A. Bidram, and M. J. Reno, "Deep Learning Based Circuit Topology Estimation and Fault Classification in Distribution Systems," in *Innovative Smart Grid Technologies Europe (ISGT Europe)*, IEEE, Oct. 2021.

[7] A. R. Kurup, A. Summers, A. Bidram, and M. J. Reno, "Ensemble Models for Circuit Topology Estimation, Fault Detection, and Classification in Distribution Systems," *Sustainable Energy, Grids and Networks*, Feb. 2023.

[8] C. Francis, R. D. Trevizan, M. J. Reno, and V. Rao, "Enhancing Topology Identification of Distribution Power Systems Using Mult-Rate Voltage Measurements," *IEEE Transaction on Industry Applications*, vol. 59, no. 1, pp. 412–421, Oct. 2022.

[9] C. Francis, V. Rao, R. D. Trevizan, and M. J. Reno, "Topology Identification of Power Distribution Systems Using Time Series of Voltage Measurements," in *IEEE Power and Energy Conference at Illinois (PECI)*, 2021, pp. 1–7. doi: 0.1109/PECI51586.2021.9435253.

[10] B. Poudel, D. R. Garcia, A. Bidram, M. J. Reno, and A. Summers, "Circuit Topology Estimation in an Adaptive Protection System," in *North American Power Symposium (NAPS)*, IEEE, 2021.

[11] L. Blakely and M. J. Reno, "Identification and Correction of Errors in Pairing AMI Meters and Transformers," in *IEEE Power and Energy Conference at Illinois (PECI)*, Apr. 2021.

[12] L. Blakely and M. J. Reno, "Identifying Errors in Service Transformer Connections," in *PES General Meeting*, Aug. 2020.

[13] J. Yusuf, J. A. Azzolini, and M. J. Reno, "Data-Driven Methods for Voltage Regulator Identification and Tap Estimation," in *IEEE Kansas Power and Energy Conference (KPEC)*, Apr. 2022.

[14] F. Therrien, L. Blakely, and M. J. Reno, "Assessment of Measurement-Based Phase Identification Methods," *IEEE Open Access Journal of Power and Energy*, 2021.

[15] B. D. Pena, L. Blakely, and M. J. Reno, "Parameter Tuning Analysis for Phase Identification Algorithms in Distribution System Model Calibration," in *IEEE Kansas Power and Energy Conference (KPEC)*, Apr. 2021. doi: 10.1109/KPEC51835.2021.9446218.

[16] L. Blakely, M. J. Reno, B. Jones, A. Furlani Bastos, and D. Nordy, "Leveraging Additional Sensors for Phase Identification in Systems with Voltage Regulators," in *Power and Energy Conference at Illinois (PECI)*, Apr. 2021.

[17] R. D. Trevizan and M. J. Reno, "Distribution System State Estimation Sensitivity to Errors in Phase Connections," in *IEEE 48th Photovoltaic Specialists Conference (PVSC)*, 2021. doi: 10.1109/PVSC43889.2021.9518900.
[18] B. D. Peña, L. Blakely, and M. J. Reno, "Data-Driven Detection of Phase Changes in Evolving Distribution Systems," in *Texas Power and Energy Conference (TPEC)*, IEEE, 2022.
[19] B. D. Peña, L. Blakely, and M. J. Reno, "Online Data-Driven Detection of Phase Changes in Evolving Distribution Systems," in *Innovative Smart Grid Technologies (ISGT)*, IEEE, Feb. 2023.
[20] L. Blakely and M. J. Reno, "Phase Identification Using Co-Association Matrix Ensemble Clustering," *IET Smart Grid*, no. Machine Learning Special Issue, Jun. 2020.
[21] M. J. Reno *et al.*, "IMoFi (Intelligent Model Fidelity): Physics-Based Data-Driven Grid Modeling to Accelerate Accurate PV Integration Updated Accomplishments," Sandia National Laboratories, Technical Report SAND2022-12669, Sep. 2022.

DISTRIBUTION

Email—Internal

Name	Org.	Sandia Email Address
Technical Library	1911	sanddocs@sandia.gov

This page left blank

Sandia
National
Laboratories

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.