EPIC Final Report

Program	Electric Program Investment Charge (EPIC)
Administrator	San Diego Gas & Electric Company
Project Number	EPIC-3, Project 3, Module 1
Project Name	Application of Advanced Metering Infrastructure Data to Advanced Utility System Operations
Date	December 31, 2021

Attribution

This comprehensive final report documents the work done in Electric Program Investment Charge (EPIC) 3, Project 3, Module 1. The project team that contributed to the project definition, execution, and reporting included the following individuals:

San Diego Gas and Electric (SDG&E)

- Tom Bialek
- Jay Bick
- Frank Goodman
- Chippy Impreso
- Julian Jones
- Kyle Kewley
- Gina Lindsay
- William O'Brien
- Amin Salmani
- Subburaman Sankaran
- Matt Smith
- Tyson Swetek
- Catarino Vargas
- Stacy Williams
- William Wood

National Renewable Energy Laboratory (NREL)

- Santosh Veda
- Harsha Padullaparti
- Valerie Rose
- Murali Baggu
- Martha Symko-Davies
- Jiyu Wang
- Jing Wang
- Jun Hao
- Marcos Netto

Executive Summary

San Diego Gas & Electric (SDG&E) and National Renewable Energy Laboratory (NREL) collaborated on a project funded through the California Public Utility Commission's (CPUC) Electric Program Investment Charge (EPIC) and U.S. Department of Energy (DOE). The purpose of EPIC-3, Project 3 is to demonstrate capabilities for leveraging SDG&E's AMI system with its 1.4 million electric meter endpoints to provide actionable secondary voltage data and analysis to SDG&E staff and other prospective users. The project focus included two modules. Module 1 focused on using Advanced Metering Infrastructure (AMI) data for a voltage sensor network, while Module 2 focused on using AMI data to identify endpoint phasing and meter-to-transformer mapping. This report addresses Module 1 only. The comprehensive final report for Module 2 is provided in a separate document.

SDG&E is looking to leverage its existing (AMI) to provide a foundational, pervasive secondary voltage monitoring network and a phase identification system. Through this pre-commercial demonstration project, NREL assisted SDG&E in demonstrating the tools and strategies for using the AMI-based field measurements for distribution system monitoring and planning. Specifically, NREL configured the tools to estimate primary network voltages, identify planning network model discrepancies, and automate phase mapping and meter-to-transformer mapping using the AMI measurement data.

The AMI data-based grid operation is an alternative to conventional model-based approaches where non-validated equivalent circuit models are used for determining the operational strategies, such as controlling voltage regulation devices and smart inverters to achieve voltage regulation. Further, as the penetration of photovoltaic (PV) systems increases, SDG&E also desired to study the associated impacts on the distribution system and the effectiveness of the integration measures. The PV smart inverter study in this project examined the impacts of PV on the distribution circuit voltage profile and traditional mechanical voltage regulation equipment operations. The efficacy of various PV smart inverter settings was quantified using the SDG&E feeder models.

Key Findings

The following are the key findings of the pre-commercial demonstrations using NREL developed tools:

- The utility feeder models used in this project have inaccuracies due to incorrect phasing, line parameter issues, and approximations used in the load and PV profiles.
- The machine learning models of the service transformer secondaries can provide reasonable estimation of the primary voltages. However, training data is essential for building these machine learning models. The primary voltage estimates from the physics-based method can be used as the training data for building the machine learning models of the secondaries.
- Circuit characteristics, measurement data available, PV penetration levels, and quality of geographic information system (GIS) data impact the selection of phase identification algorithms. The phase identification algorithms that work well in one feeder may not show similar performance in a different feeder depending on the feeder characteristics.

The algorithms based on supervised learning showed higher accuracy levels compared to those based on unsupervised learning for phase identification. However, supervised learning requires training data. Accuracy levels of nearly 90% and 94% are obtained on the selected two feeders in this study using a supervised learning algorithm with 30% training data, where a feeder is defined as a three-phase set of conductors (power lines) emanating from a substation circuit breaker serving customers in a defined local distribution area.

- Existing high PV penetration levels in the selected SDG&E feeder create voltage issues both on the primary and secondary networks. Enabling PV smart inverter settings significantly reduced the voltage exceedances in the simulations. There are minor variations in the voltage improvement among different smart inverter settings, but generally the voltage profile is better when smart inverters are enabled compared to when they are disabled.
- The volt/var curve slope is a key parameter influencing the voltage improvement.

Recommendations

The use cases presented in this project demonstrate the accuracy, feasibility, and rationality of using AMI data for greatly improving the planning and operations activities in the near-term, especially for feeders with high levels of PV adoption. It is recommended that specific tools (Utility Planning Network Model Anomaly Detection Tool, AMI Meter-to-Transformer Mapping, and Phase Identification Using AMI Data) be applied by the SDG&E team for other feeders. The evaluation of data-driven controls using realistic emulation capabilities of the ADMS Test bed provides a feasible demonstration for real-time data-driven control of high-PV feeders for consideration and implementation in the medium-term. Such an approach could reduce the reliance on planning models and make the operations resilient to the ubiquitous problem of poor model quality.

SDG&E will need to identify a stakeholder group within the company to lead this commercial adoption process.

Table of Contents

List of Illustrations

List of Tables

List of Acronyms

Acronym	Acronym Description
ADMS	Advanced Distribution Monitoring System
AMI	Advanced Metering Infrastructure
ANSI	American National Standards Institute
CPUC	California Public Utilities Commission
CRADA	Cooperative Research and Development Agreement
DER	Distributed Energy Resources
DERMS	Distributed Energy Resource Management System
DOE	Department of Energy
EPIC	Electric Program Investment Charge
ESIF	Energy Systems Integration Facility
GIS	Geographic Information System
IAB	Industry Advisory Board
HPC	High Performance Computing
MAPE	Mean Absolute Percentage Error
NREL	National Renewable Energy Laboratory
PHIL	Power Hardware in the Loop
PLM	Peak Load Management
PU	Per Unit
QSTS	Quasi Static Time Series
SCADA	Supervisory Control and Data Acquisition
VE	Voltage Exceedance
VUI	Voltage Unbalance Index
VVO	Volt-var Optimization

1.0 Introduction

The purpose of EPIC-3, Project 3 is to demonstrate capabilities for leveraging SDG&E's AMI system with its 1.4 million electric meter endpoints to provide actionable secondary voltage data and analysis to SDG&E staff and other prospective users. The project focus included two modules. Module 1 focused on using Advanced Metering Infrastructure (AMI) data for a voltage sensor network, while Module 2 focused on using AMI data to identify endpoint phasing and meter-to-transformer mapping. This report addresses Module 1 only.

As the penetration of photovoltaic systems increases, the voltage on the service transformer secondary increases, and when coupled with power production from renewable generation sources, there is potential for voltage to exceed American National Standards Institute (ANSI) C84.1 Range A limits. SDG&E is looking to leverage its existing AMI infrastructure to provide a foundational, pervasive secondary voltage monitoring network and a phase identification system to address these and other issues associated with monitoring and managing their distribution system. Through this project, NREL assisted SDG&E in evaluating the challenges and mitigation strategies associated with high penetration PV-in the distribution circuits. The proposed approach is an alternative to a conventional model-based approach where equivalent circuit models are used for determining operational strategies. NREL leveraged its Energy Systems Integration Facility (ESIF) high performance computing (HPC) simulation and advanced distribution monitoring system (ADMS) test bed capabilities for pre-commercialization and evaluation of the approach and solutions envisioned by SDG&E.

2.0 Project Objectives

The primary objective of the proposed scope was pre-commercial demonstration and evaluation of NREL-developed algorithms and tools for leveraging secondary network voltage measurement data made available through the AMI. The algorithms, tailored for use in this project, use the AMI data provided by SDG&E, collected from selected customer smart meters on selected feeders for data-centric grid planning and operations. A feeder (also known as a circuit) is defined as a three-phase set of conductors (power lines) emanating from a substation circuit breaker serving customers in a defined local distribution area. In addition, some of the challenges of using a non-modeled control i.e., without relying on the network model for computing the control decisions, were identified. The lessons learned will be disseminated to the broader utility community and other stakeholders through this comprehensive final report filed with the CPUC and released on SDG&E's EPIC public website and through submission to peer-reviewed publications and workshops, either co-authored by SDG&E and NREL or with the approval of SDG&E. The project demonstrated the value of AMI data for performing the following:

- Identification of anomalies in planning models
- Identification of customer phasing information at the meter-level
- Demonstration of effectiveness of different smart inverter settings on customer voltages

3.0 Issues and Policies Addressed

This project module endeavored to understand the impact of high PV penetration in the SDG&E service territory and determine the effectiveness of technology solutions such as energy storage, smart inverters, flexible loads, and other solutions to mitigate issues. The proposed approach is an alternative to a conventional model-based approach where non-validated equivalent circuit models are used for determining operational strategies. The demonstration leveraged NREL's ESIF HPC simulation and PHIL capabilities to evaluate the approach and solutions envisioned by SDG&E.

The project team collaborated through a Cooperative Research and Development Agreement (CRADA) on pre-commercial demonstration of tools, models, and algorithms for leveraging the SDG&E AMI infrastructure and data for advanced grid monitoring and planning. Through this collaboration, the project team developed a framework for generating synthetic AMI data for different PV penetration scenarios; developed, demonstrated, and integrated a new technique for identifying discrepancies in primary network models; and developed methods for analyzing and visualizing AMI datasets for an SDG&E feeder. This framework is shown in Figure 1 below. The framework allows for the generation of multiple scenarios of interest by applying different load/PV characteristics to the validated planning model, and the application of data processing techniques that mimic the data processing in AMI meters to generate synthetic AMI data. The machine learning-based data analytics algorithms can be applied on this synthetic AMI data to study the performance of the algorithms in these planning scenarios that do not exist in the field today but may be anticipated in the future. The work performed in this demonstration project was an extension of this CRADA with specific focus on analyzing the impact of different PV penetration scenarios through AMI data, phase identification, and meter to transformer mapping.

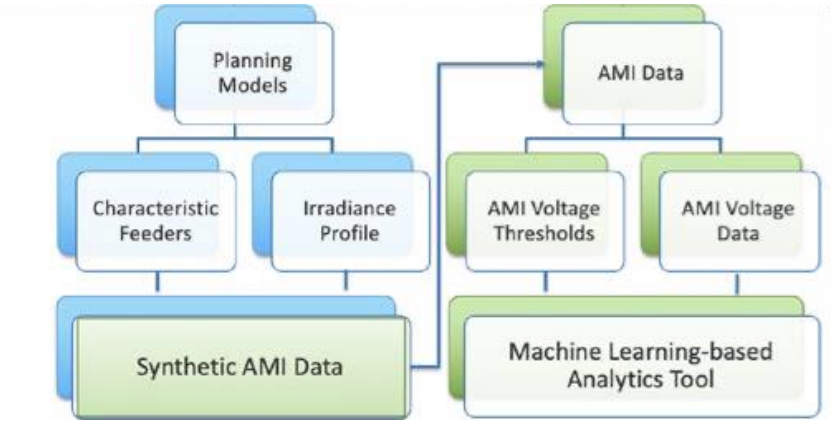

Figure 1. Framework for generating synthetic AMI data

4.0 Project Focus

The focus of this project module was to demonstrate advanced AMI data-based techniques for improving utility planning and operations. Some of the focus areas included identification of anomalies in the network planning model, demonstrating the impact of different PV inverter settings on customer voltage profiles as observed through AMI data, and demonstrating an advanced AMI data-based voltage control using a realistic emulation capability at NREL called the ADMS Test Bed.

5.0 Project Scope Summary

The scope of the work included using data from SDG&E's existing AMI infrastructure. AMI data from selected SDG&E feeders was used to identify the individual phases of the customer meter on the feeder and map the customer meters to the corresponding service transformers. These selected feeders are Feeders A and B. Feeder A has 325 connected transformers and 5,173 connected meters. It serves a relatively dense suburban neighborhood with a mix of overhead and underground wiring and a relatively even mix of line-line (L-L) and (L-N) phasing on the transformers. Feeder B has 649 connected transformers and 2,393 connected meters. It serves a spread-out suburban neighborhood with predominantly underground wiring and predominantly line to neutral (L-N) phasing on the transformers. The results from the analysis were validated through field verification conducted by SDG&E field personnel. The team evaluated the algorithms on the ADMS test bed located in the ESIF at NREL using realistic feeder parameters and distributed energy resources (DER). Further, NREL worked with SDG&E personnel to deploy the algorithms on SDG&E's analytics platform for further consideration for operationalization.

The primary objective of the project module was pre-commercial demonstration and evaluation of NREL-developed algorithms and tools for leveraging secondary network voltage measurement data made available through the AMI. The algorithms used the AMI data provided by SDG&E, collected from selected customer smart meters on selected feeders for data-centric grid planning and operations. In addition, some of the challenges of using a non-modeled control i.e., without relying on the network model for computing the control decisions, were identified. The lessons learned will be disseminated to the broader utility community through this comprehensive final report that will be filed with the CPUC and released on SDG&E's EPIC public website and through peer-reviewed publications and workshops with the consent of SDG&E. The project team embarked on the following use cases:

- PV Smart Inverter Study
- Utility Planning Network Model Anomaly Detection Tool
- Phase Identification Tool
- Meter to Transformer Mapping
- Data-Centric Grid Operations

PV Smart Inverter Study

High PV penetration levels in distribution feeders can cause operational challenges including voltage issues, reverse power flow, and protection issues. Standards recommend using PV smart inverters to support the distribution grid services, specifically voltage regulation. However, there are many smart inverter settings recommended by the standards and their performance on the SDG&E feeders has not been reported in literature. In this study, the impact of various smart inverter settings including California Rule 21 (CA 21) [6], Hawaii Rule 14 [7], IEEE 1547 [8] with no deadband, CA 21 with reactive power compensation at only high or low voltage range (referred to as "hockey stick" in this report), and volt-var-watt control on the selected SDG&E feeders was examined.

Utility Planning Network Model Anomaly Detection Tool

An automated tool was demonstrated that uses the AMI measurement data to identify the inaccuracies in the network model used for distribution planning. Numerous distribution network analysis, monitoring, and control applications; including volt/var optimization, state estimation, and distribution automation, require accurate distribution network models. The GIS maintained by utilities can be inaccurate because of a significant amount of missing data, restoration activities, and network reconfiguration. This can lead to network model inaccuracies. The utility planning network model anomaly detection tool used the AMI data to identify network model issues. It accomplished this by building the approximated secondary network models from the AMI data and using them to estimate the primary voltages. The estimated primary voltages were then compared with the primary voltages obtained from the simulations of the utility planning network model to identify model anomalies.

Phase Identification Tool

The phase identification tool performs automated phase mapping of the AMI meters based on AMI data. The GIS database maintained by the utility is known to have phase connectivity errors due to restoration activities, network reconfiguration, human error, and missing data. Traditionally, the phase connectivity database is periodically updated by field verification which is expensive and time-consuming. With the availability of AMI data, the phase connectivity can be identified through data analytics. The existing phase identification techniques work well in distribution feeders that have low or no PV generation; however, they fail to identify the phases accurately when considerable PV generation is present. The phase identification tool demonstrated in this project uses supervised learning to determine the phase connectivity accurately even when significant PV generation is present.

Meter-to-Transformer Mapping

Utilities generally have a meter-to-transformer connectivity mapping database. However, the records in this database do not always reflect the latest field conditions due to routine meter field changes and occasional human data entry errors. Accurate meter-to-transformer mapping information is needed for load balancing, service order work, and transformer load management. A solution that can check and correct the service transformer and meter mapping records is required to address this need. The goal of this use case was to demonstrate a proof-of-concept AMI meter-to-service transformer mapping solution that identified incorrect records based on AMI measurement data.

Data-centric Grid Operations

The integration of ADMS and AMI measurements offers a unique opportunity to further modernize grid control. In this use case, an AMI-based, data-driven, volt/var control algorithm, and its synergies with ADMS for distribution grid operations, were evaluated using SDG&E feeder and AMI data. The inputs of this algorithm were AMI power and voltage measurements. The algorithm controls the substation transformer load tap changer (LTC) tap position, capacitor bank switch positions, and PV inverter setpoints to ensure voltage regulation. This new paradigm for grid operations was demonstrated using NREL's ADMS Test Bed capability wherein the feeders and the controls were implemented and evaluated in a realistic utility environment.

6.0 Initial Benefits Analysis and Value Proposition

The initial benefit estimate and value proposition focused on improved distribution network reliability, reduced cost, increased safety, and enhanced environmental sustainability.

Improved Reliability

The utility planning network model anomaly detection tool provides distribution network operators greater visibility into their systems and helps them understand the voltage dynamics on the primary network. This can help in taking corrective actions in terms of updating device control settings or installing additional regulation devices to ensure voltages are maintained within the desirable limits and reliable network operation. As DER penetration levels increase, the traditional assumption that voltage drops network-wide is no longer valid. Virtual sensors on the primary network that provide voltage estimates are desirable to monitor voltage dynamics on the primary network. Further, identification of the inaccuracies in the utility planning network model is a key step in the correction process. The corrected models result in improved planning and operation decisions, leading to greater network reliability.

The phase identification tool further enhances reliability by maintaining the phase connectivity information in the model based ADMS and the Distributed Energy Resource Management System (DERMS). Phase identification will further support complex phase balancing. The PV smart inverter study quantified the impact of different smart inverter settings recommended in the standards on the network voltage profiles. These insights guide utilities to recommend and configure appropriate smart inverter settings for the PV systems within their service areas to ensure desired voltage levels across their networks.

Reduced Costs

The estimation of the primary voltages, based on the AMI data by the utility planning network model anomaly detection tool, removes the need for installing physical voltage sensors for planning purposes. Thus, it is more economical. Similarly, the phase identification tool determines the customer phase connectivity analytically based on the AMI data. Traditionally, the utilities undertake expensive field verification activities from time to time to keep their phase connectivity up to date. Alternatively, special equipment is used to detect customer phasing. When the AMI data is available, relying on the phase

identification tool for the automatic phase mapping results in lower costs as it does not require additional equipment or field checks.

As the PV penetration levels increase, high voltage volatility is anticipated on the distribution networks. Traditionally, network upgrades or installation of advanced grid-edge devices are required for improving the network voltage profiles. As the grid standards now mandate PV smart inverters to participate in voltage regulation, configuring the appropriate smart inverter settings effectively supports the voltage regulation without having to resort to expensive traditional network upgrades. Further, the fast voltage regulation capability of the PV smart inverters reduces the wear and tear on the traditional mechanical voltage regulation devices. Thus, appropriate PV smart inverter settings can lead to lower investment and maintenance costs.

Increased Safety and/or Enhanced Environmental Sustainability

The phase identification tool promotes increased safety by reducing the need for personnel to travel and work in the field for the phase connectivity verification activities. This also results in reduced greenhouse gas (GHG) emissions from the vehicles used for this travel. The PV smart inverter study demonstrated that the PV smart inverters can help improve the voltage profile of the distribution networks and support grid integration of higher amounts of renewable energy sources. This promotes the environmental sustainability by accelerating the use of clean energy resources.

7.0 Use Cases

7.1 PV Smart Inverter Study

In recent years, the penetration of residential and commercial rooftop PV systems has been increasing rapidly [1]. The PV systems, however, can cause issues with the distribution system when penetration is high [2] [3]. The power outputs from PV systems peak at noon on a sunny day but the load consumption for residential customers at that time is typically low. These inconsistent profiles of PV output and load can cause overvoltage and reverse power flow problems [4].

To solve issues resulting from renewable energy generation, standards are now requiring utilization of PV smart inverters. Smart inverters can absorb or supply reactive power, or automatically curtail power output to maintain voltage levels [5]. Additionally, smart inverters have the ability to compensate for the voltage fluctuations in the grid via reactive power control even when the generation is not operating. There are many recommended smart inverter settings, and their performance on SDG&E feeders has not yet been fully studied. This demonstration simulated and studied the feeder operation with all distributed PV systems equipped with smart inverters. Different rules were tested including California Rule 21 (CA 21), Hawaii Rule 14, IEEE 1547 with no deadband, hockey stick, and volt-var-watt control. The performance of each rule was evaluated by using a set of metrics.

7.1.1 Configuration and Methodology

Load disaggregation was performed to extract the load and PV profiles from the AMI net load measurements. A custom function was created in Python to emulate different volt-var-watt

functionalities and settings suggested in standards as such a function is not available as needed in OpenDSS. The OpenDSS is an electric power Distribution System Simulator (DSS) for supporting distributed resource integration and grid modernization efforts.

Load Disaggregation

The distribution feeder model used in this study, for the purpose of discussion, is Feeder A. The AMI data of this feeder was provided for the period between October 1, 2018, to January 15, 2019 (107 days). AMI load measurements from SDG&E included the net load consumption of each customer, therefore, a disaggregation was required to extract the PV profile and load profile for each load location. From the load definition in the feeder model and peak power generation of each PV system, the determination was made that the PV penetration for this feeder is around 70% relative to peak load. The irradiance profile of the feeder area during the selected period of 107 days was downloaded from the National Solar Radiation Database (NSRDB) [9]. By using the ratings of each PV system, the irradiance profile, and the net load profile of each load node, the PV profile and load profile at each load location were disaggregated. After the disaggregation, the scenario of 100% PV penetration was modeled with the load and PV profiles. This was the case used in the following simulations.

Volt-VAR-Watt Smart Inverter Function

In OpenDSS, the volt-var-watt smart inverter control function is not yet fully operational, therefore a Python function to implement the volt-var-watt control was created. The inputs of the function include inverter rated kVA, solar irradiance at current time step, and measured per-unit (PU) voltage at the previous time step. First, the volt-var and volt-watt curves were predefined. Then based on the voltage and volt-watt curve, the function determined the required real power output and the maximum available reactive power. After that, the maximum available reactive power, and volt-var curve, and the reactive power output was calculated based on the measured voltage. The outputs of this function were the real and reactive power outputs of the PV system. These outputs were used to update the PV system output in OpenDSS.

PV Smart Inverter Curves

The smart inverter curves for all cases are summarized in this section. Several smart inverter curves, both from the standards and the custom curves of interest were studied in this work. These curves are depicted in Figure 2.

- *California Rule 21 (CA 21):* The maximum and minimum percentage of available reactive power is +/-30%. This percentage is zero when the voltage is within 0.967-1.033 PU and reaches maximum/minimum when the voltage is below/over 0.92/1.07 PU
- *Hawaii Rule 14 (HI 14):* The maximum and minimum percentage of available reactive power is +/-44%. This percentage is zero when the voltage is within 0.97-1.03 PU and reaches maximum/minimum when the voltage is below/over 0.94/1.06 PU
- *IEEE 1547*: The maximum and minimum percentage of available reactive power is +/-44%. This percentage is zero when the voltage is within 0.98-1.02 PU and reaches maximum/minimum when the voltage is below/over 0.92/1.08 PU

- *California Rule 21 without deadband:* The maximum and minimum percentage of available reactive power is +/-30%. This percentage reaches maximum/minimum when the voltage is below/over 0.92/1.07 PU
- Hockey stick curve without compensation in low voltage region: The minimum percentage of available reactive power is -30%. This percentage is zero when the voltage is below 1.033 PU and reaches maximum when the voltage is above 1.07 PU
- Hockey stick curve with deeper Q absorption: The minimum percentage of available reactive power is -75%. This percentage is zero when the voltage is below 1.033 PU and reaches maximum when the voltage is above 1.07 PU
- *Volt-VAR-Watt:* The volt-var curve is the same as California Rule 21. For its volt-watt curve, the maximum available real power starts to decrease from 100% when the voltage is above 1.06 PU and reaches zero when the voltage is above 1.1 PU

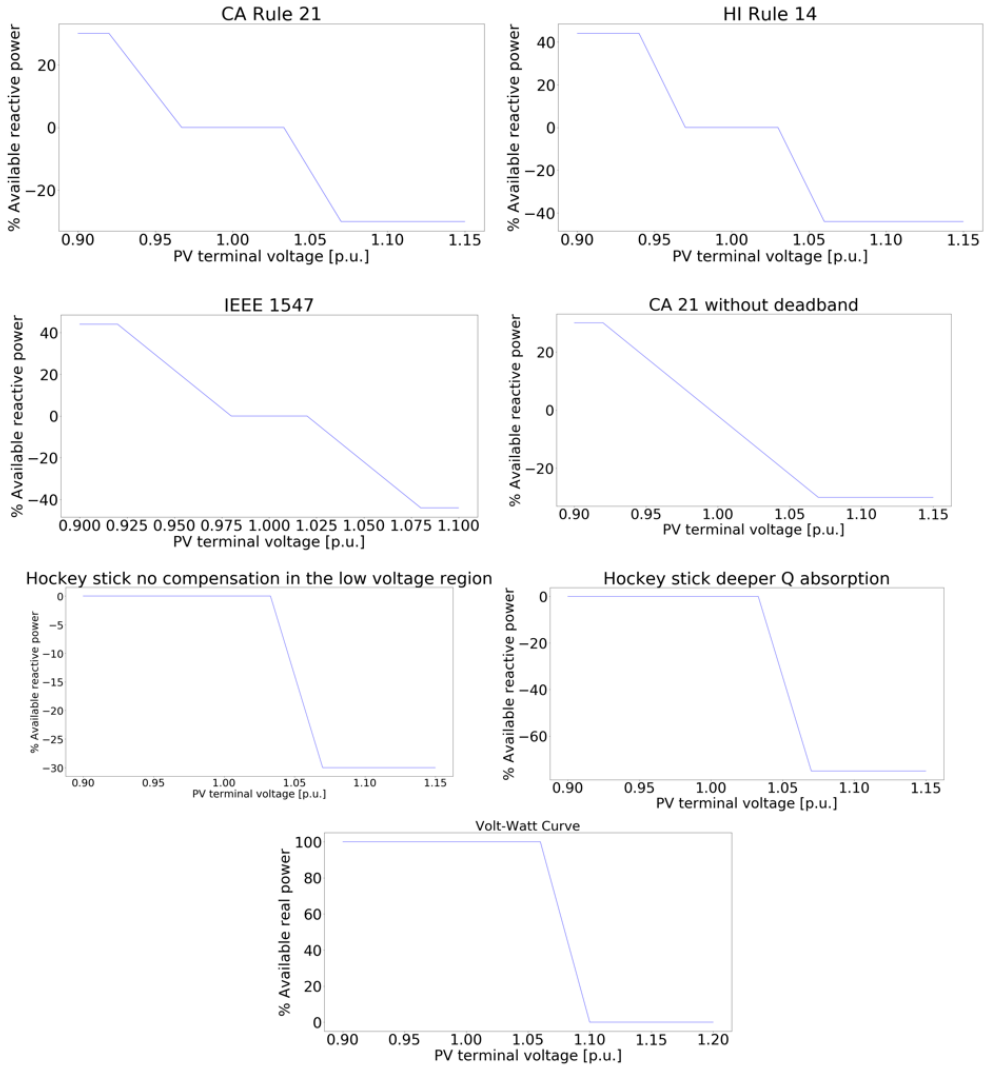

Figure 2. Volt-var and volt-watt curves for PV smart inverter

Metrics

The performance of each smart inverter function was evaluated by using multiple metrics: number of capacitor changes, number of LTC operations, average voltage, voltage fluctuation index, voltage unbalance index, and number of voltage exceedances nodes. Voltage fluctuations are described as repetitive or random variations of the voltage envelope due to sudden changes in the real and reactive power drawn by the load.

Let the T stand for the total time steps in the simulation and N stand for the total number of nodes in the feeder, the average voltage is calculated by:

$$V^{mean} = \frac{1}{N} \times (\frac{1}{T} \sum_{i=1}^{N} \sum_{t=1}^{T} V^i(t))$$

The voltage fluctuation index (VFI) measures how the nodal voltage is changing between time steps, i.e., voltage fluctuations across the circuit. It is calculated by:

$$VFI = \frac{1}{N} \times (\frac{1}{T} \sum_{i=1}^{N} \sum_{t=1}^{T} \left| V^i(t+1) - V^i(t) \right|)$$

The voltage unbalance index (VUI) measures the unbalance level of nodal phase voltages across the circuit. It is calculated by:

$$VUI = \frac{1}{N} \times (\frac{1}{T} \sum_{i=1}^{N} \sum_{t=1}^{T} V_{imb}^i(t))$$

where $V_{imb}^i(t)$ is calculated by using the maximum deviation from average voltage over the average voltage.

The voltage exceedance is defined as voltage out of range 0.94-1.06. The exceedance node is defined as a node with more than 12 hours exceedance in the three-month period.

7.1.2 Results and Discussion

The PV and load profiles were interpolated to five-minute resolution and the simulation was run using the data from October 2018 to January 2019 (107 days, 30,816 data points in total). There were 1,560 PV systems in the model in total, all with ratings between 5-10 kVA. For the baseline, the power outputs of PV systems were determined by the irradiance and inverter rating. For the case with smart inverter function enabled, the power outputs of these PV systems follow the corresponding curves. The voltage plots in this section are presented for one selected day i.e., October 1, 2018. The metrics were computed for the three-month period and summarized in Table 1 and 2.

California Rule 21

The voltage profiles for the selected day are shown in **Error! Reference source not found.**3. Note that the bus voltages are reduced during the day due to the LTC at the feeder head lowering the tap position to regulate the voltage rise due to PV.

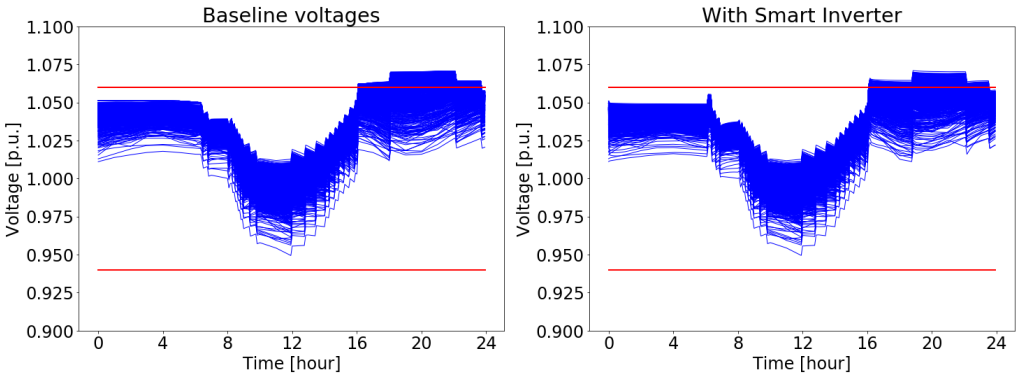

Figure 3. Daily voltage profile with/without CA 21 smart inverter

Hawaii Rule 14

The voltage profiles for the selected day are shown in **Error! Reference source not found.**4.

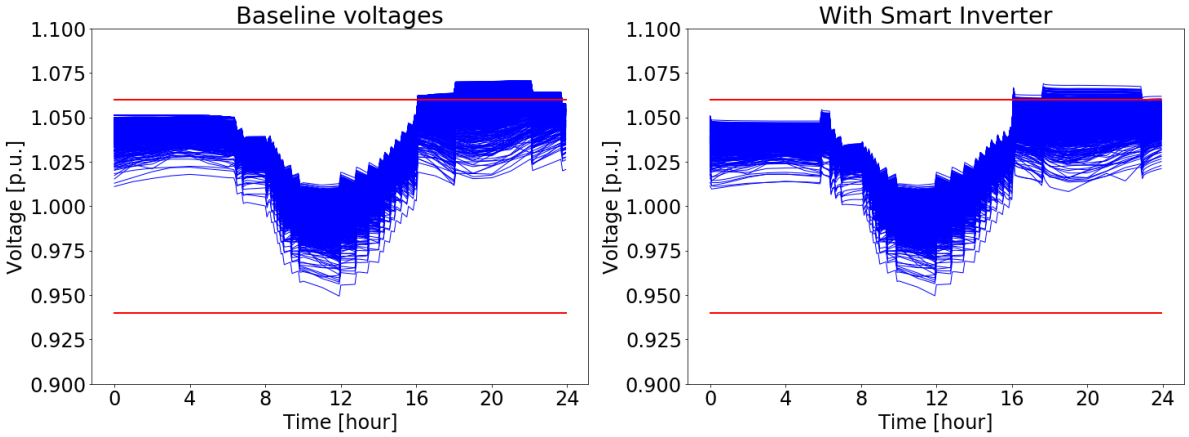

Figure 4. Daily voltage profile with/without HI 21 smart inverter

IEEE 1547

The voltage profiles for the selected day are shown in

5.

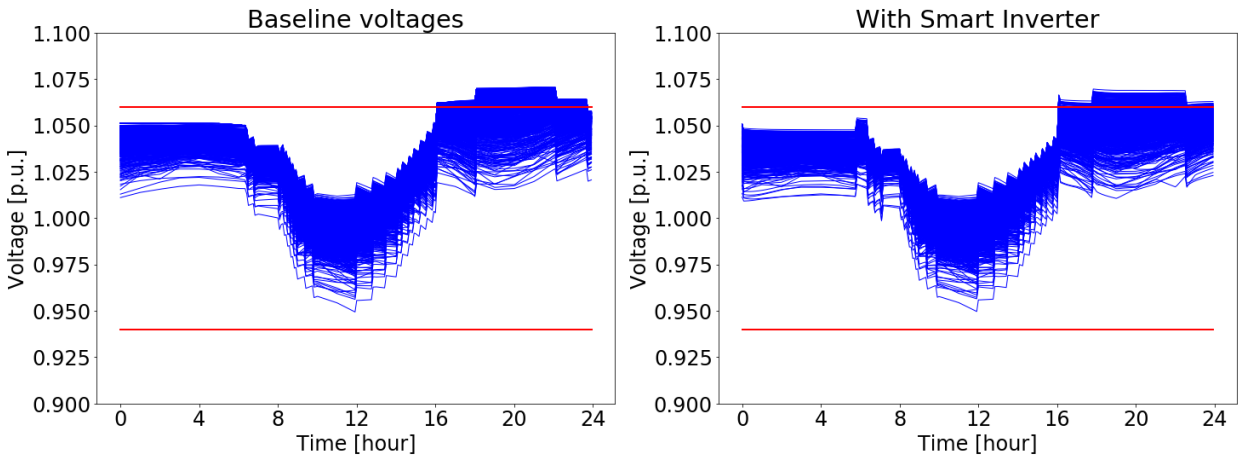

Figure 5. Daily voltage profile with/without IEEE 1547 smart inverter

California Rule 21 without Deadband

The voltage profile for the selected days is shown in **Error! Reference source not found.**6.

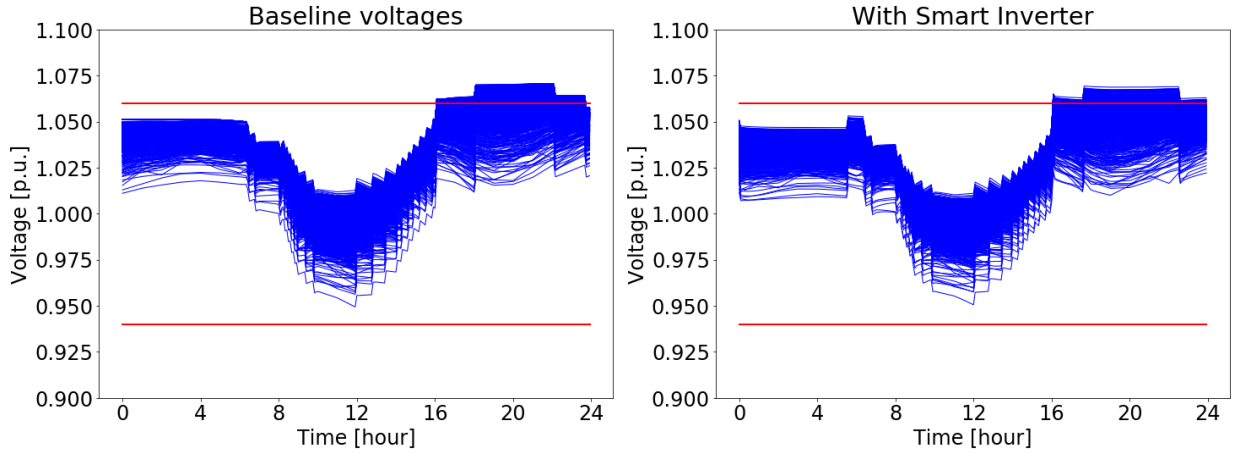

Figure 6. Daily voltage profile with/without CA 21 without deadband smart inverter

Hockey Stick Without any Compensation in the Low Voltage Region

The voltage profiles for the selected day are shown in **Error! Reference source not found.**7.

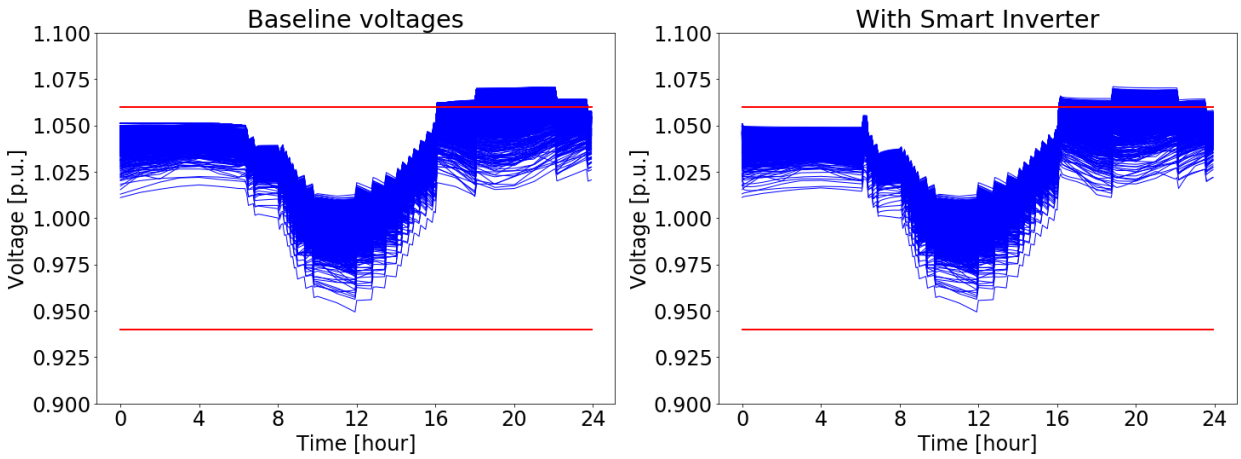

Figure 7. Daily voltage profile with/without hockey stick without compensation smart inverter

Hockey Stick with Deeper Reactive Power Absorption

The voltage profiles for the selected day are shown in **Error! Reference source not found.**8.

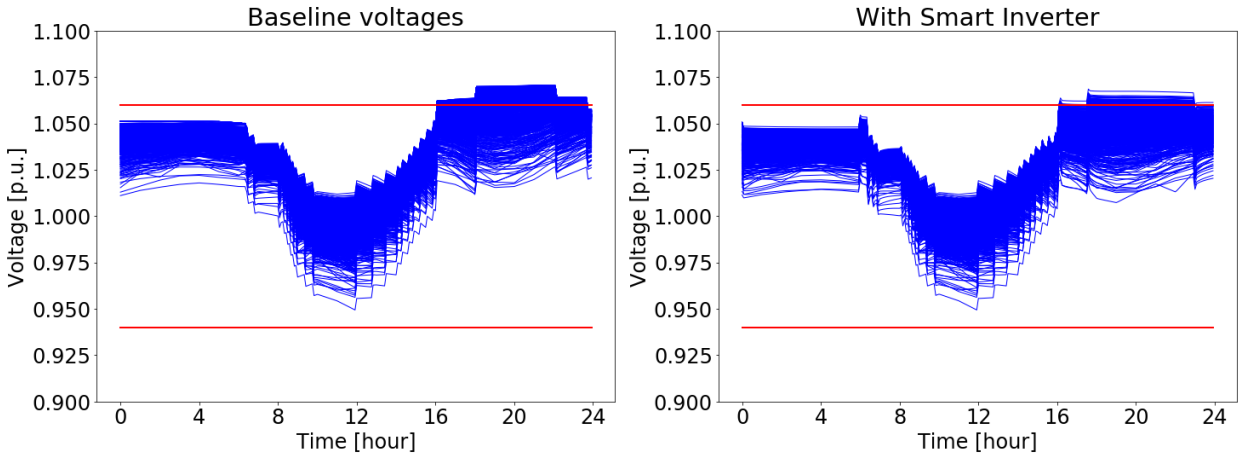

Figure 8. Daily voltage profile with/without hockey stick with deeper Q absorption smart inverter

Volt-Var-Watt Mode

The voltage profiles for one example day are shown in **Error! Reference source not found.**9.

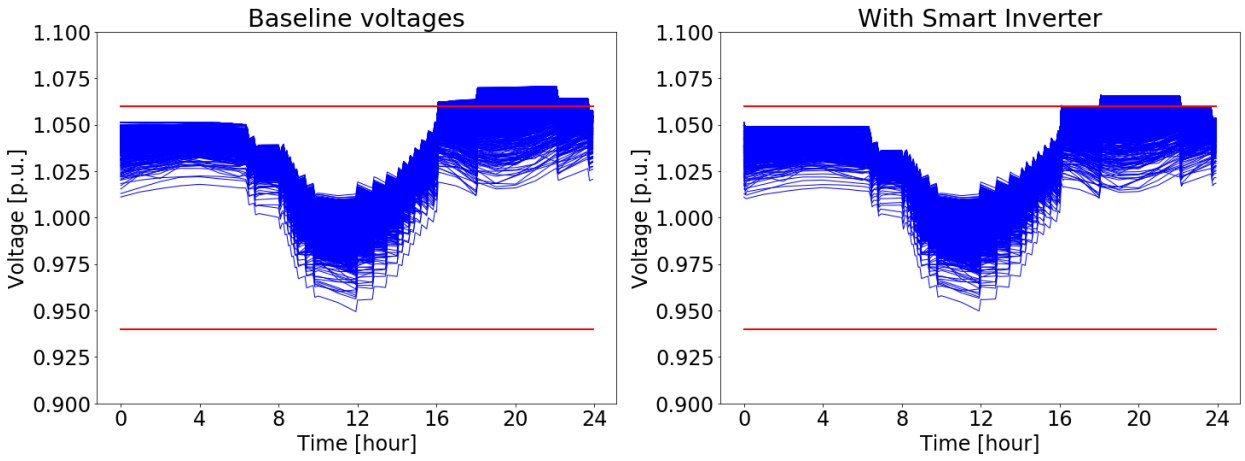

Figure 9. Daily voltage profile with/without volt-var-watt smart inverter

Summary for Three-Month Simulation

The summary of the three-month simulation for each smart inverter function is provided in Table 1 and Table 2.

Table 1. Summary of Feeder Operations

	Capacitor bank status changes	LTC tap changes	Average Voltage (V)	V fluctuation index score	V unbalance index score
Baseline	562	1291	249.93	9.67	9.90
CA 21	646	1335	249.20	9.68	9.86
HI 14	538	1471	248.60	9.67	9.82
IEEE 1547	580	1478	248.66	9.66	9.81
No Deadband	606	1394	248.41	9.67	9.77
HS-no compensation	646	1335	249.20	9.68	9.86
HS-deeper Q	624	1506	248.60	9.62	9.83
Volt-Var-Watt	158	784	250.27	9.61	9.81

Table 2. Summary of Voltage Exceedances

	Secondary		Primary	
	Voltage exceedances hours per node	**Number of voltage exceedances nodes**	**Voltage exceedance-hours per node**	**Number of voltage exceedance nodes**
Baseline	23.52	752	42.83	481
CA 21	0.55	16	0.61	0
HI 14	0.21	9	0.76	12
IEEE 1547	0.47	28	0.96	14
No Deadband	1.05	37	2.84	42
HS-no compensation	0.55	16	0.61	0
HS-deeper Q	0.09	3	0.91	12
Volt-Var-Watt	4.45	110	2.95	53

The results show the implementation of smart inverter settings improves the feeder voltage profile by reducing the voltage exceedances. Based on the results from the demonstration, the Rule 21 curve showed superior results in terms of the number of voltage regulation device actions and eliminating the primary voltage exceedances. The Rule 14 curve showed superior results in terms of eliminating the secondary voltage exceedances. The voltage exceedances for the volt-var-watt function are higher than the others, but it has the lowest number of voltage regulation device actions. The numbers of voltage regulation device changes are similar for all other smart inverter functions, and the average voltages are all near 249 V. Based on different purposes of controlling the feeder, the corresponding smart inverter functions can be selected by using the results from this demonstration. For example, if the utility wants to minimize the action times of the voltage regulation devices, Rule 21 can be set for the smart inverters on this feeder.

7.2 Utility Planning Network Model Anomaly Detection Tool

Numerous distribution network analysis, monitoring, and control applications, including volt/VAR control, state estimation, and distribution automation, require accurate distribution network models. The GIS maintained by utilities can be inaccurate because of a significant amount of missing data, restoration activities, and network reconfiguration. The voltage transfer software tool uses AMI data recorded on the distribution secondary network, static primary network connectivity data derived from the utility planning model, and voltage data from the time-series simulations of the utility planning model to identify the model anomalies. It accomplishes this by building the approximated secondary network models from AMI measurements and using them to estimate the primary voltages. The estimated voltages are then compared with primary voltages obtained from utility planning network

model simulations to identify the mismatches. A list of primary buses having high mismatches is saved along with other mismatch statistics.

The software tool uses AMI data recorded on the secondary distribution network, static primary network connectivity data derived from utility planning model, and voltage data from time-series simulations of the planning model to identify the model anomalies as shown in **Error! Reference source not found.**.

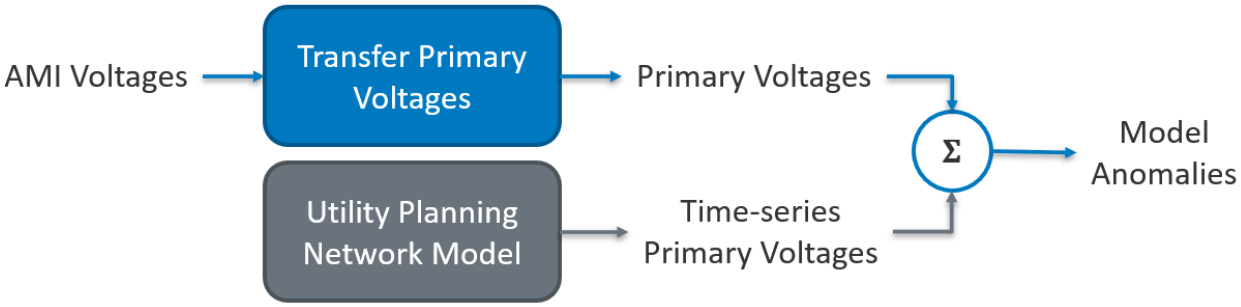

Figure 10. Illustration of utility planning network model anomaly detection tool

7.2.1 Configuration and Methodology

The software tool uses the combination of a physics-based method and a machine learning method to estimate the primary bus voltages from the AMI voltage measurements on the distribution secondary network.

Physics-based Method

In the physics-based primary voltage estimation method, the voltage magnitudes on the primary side of the distribution service transformers were estimated using only two smart meters per secondary network. The smart meters were strategically placed on the closest and farthest load (in the electrical sense) from the transformer, in the electrical sense. This method relies exclusively on smart meter data, and therefore it is fully data driven.

The physics-based primary voltage estimation method has two stages:

First Stage – Linear Regression: The first stage performed a linear regression on the latest data window available at the control center. A data window of 288 points was used, which is equivalent to a day for five-minute sampling resolution. The first stage was executed only once.

The equivalent circuit shown in **Error! Reference source not found.** was used for each service transformer. In this circuit, r_p and r_s denote the losses in the primary and the secondary winding of the service transformer, respectively. The variables, v_p' and v_s' denote the voltage magnitudes at the primary and the secondary of an ideal transformer, respectively and n_t is the transformer's turns ratio. The variables, v_1 and p_1 denote the voltage magnitude and the active power measured at the closest load from the service transformer while v_2 and p_2 denote the voltage magnitude and the active power measured at the farthest load from the transformer. The variables r_1 and r_2 account for cable impedance; and v_u and p_u are unknown. A constrained linear least-squares minimization problem was

solved to estimate the resistance r_2 and the equivalent resistance between the first meter and the primary bus.

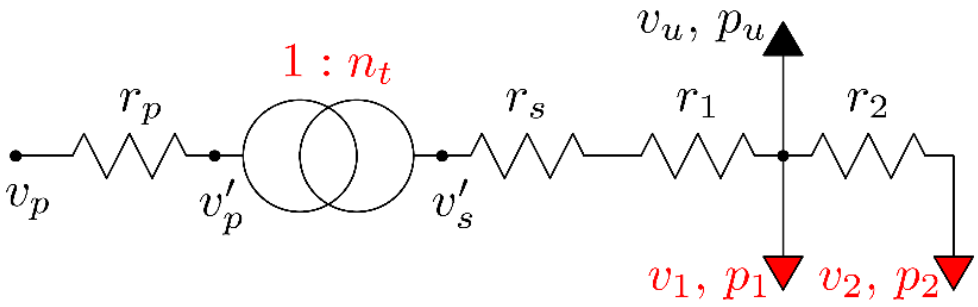

Figure 11. Equivalent circuit used for each secondary in physics-based voltage estimation method

Second Stage – Kalman Filtering: The second stage used a Kalman filter to update the primary voltage magnitude estimates continuously, based on new data points. The processing steps in both stages are discussed in [20].

Machine Learning Method

Machine learning approaches typically require a training data set that contains the features to be estimated. In this application, the inputs included AMI measured power and voltages of two customers under each service transformer and the total power consumption of all customers under the service transformer. The output is the transformer primary-side voltage. Therefore, the transformer primary voltage data must be included in the training data set in addition to the other specified feature data. However, because no primary-side measurements were available for this feeder, except the RTU voltage, time-series voltage data recorded from the simulations in OpenDSS were used to form the required training data set in the algorithm development stage. The quasi-static time-series (QSTS) simulation of Feeder A was performed in OpenDSS for the period between October 1, 2018, to January 15, 2019 (107 days) to obtain primary-side voltages. In the QSTS simulation, time resolution was set to hourly to follow AMI load time resolution. The load profile of each secondary-side measured load was set to be the AMI measured total power under that transformer. The simulated primary-side synthetic voltages from the QSTS simulation and the actual measured secondary-side voltages at the two AMI measured loads were used to train the machine learning model. The machine learning method for this application is discussed in [21].

After validating the performance of the machine learning-based algorithm using simulation data, it was applied to actual AMI measurement data recorded in the field. In this stage, the machine learning models were trained using the primary voltages estimated by the physics-based method instead of the simulation data.

Combined Method

Both the physics-based and the machine learning-based methods have limitations in estimating the primary-side voltages when applied individually. The physics-based method can use the available AMI data to conduct the estimation, but the accuracy is lower than desired. The machine learning-based

method can have a higher estimation accuracy, but requires measurements of service transformer primary voltages, which was not included in the dataset provided. A combined method was developed to leverage the advantages of both these methods. In the combined method, primary voltages were estimated by the physics-based method for a given time duration, a primary voltage correction was applied, and corrected primary voltages were used to train a machine learning model. The trained machine learning models can be used to estimate primary voltages for any time duration.

Voltage drop across a service transformer typically varies from two to 13 V. The average voltage drop between the two AMI meters in all the secondaries of SDG&E Feeder A is 0.58 V. This implies that the voltage drop on a service transformer would be four to 20 times of the voltage drop between the AMI meters in the corresponding secondary. Accordingly, a correction factor was applied to the primary-side voltages estimated by the physics-based method before using them as training data for the machine learning models.

7.2.2 Results and Discussion

Physics-based Method

Five service transformer locations as shown in Figure 12 were selected to perform the validation of the physics-based method. The corresponding secondaries were modeled in detail with the realistic topology and load data. Each secondary model comprises a few loads including the two loads for which the AMI load consumption data was available. The load profiles of the two loads were set to be the same as AMI load consumption data for those two loads. The aggregated power consumption data at the service transformer level (minus the sum of the two loads) were distributed evenly across the rest of the loads of that secondary. The primary voltages of the selected secondaries were estimated using the physics-based method. The estimation mismatch results are summarized in Table 3. It can be observed that the estimation mismatches are all around 4%, which is larger than expected. The physics-based method usually will overestimate the primary voltages. Therefore, a machine-learning based method and a combined method are developed to improve the estimation accuracy.

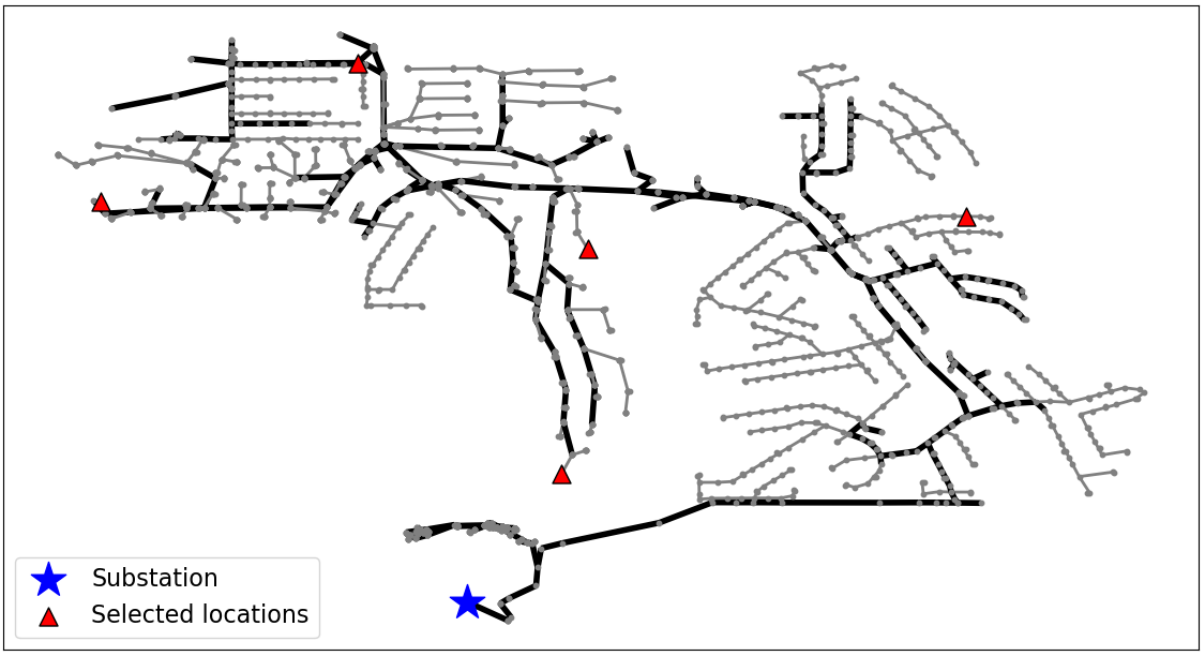

Figure 12. Five selected locations for the validation of physics-based method.

Table 3. Physics-based method validation results

	Estimation Mismatch
Secondary 1	3.74%
Secondary 2	4.06%
Secondary 3	3.53%
Secondary 4	4.28%
Secondary 5	3.84%

Machine Learning Method

Multiple machine learning algorithms: Random Forest, Adaptive Boosting, and Gradient Boosting [10] [12] were tested to find the relationship between the primary-side voltages and the AMI measurements under each service transformer. The data from each service transformer (341 in total) were trained separately to account for their unique characteristics, i.e., separate models were constructed for each service transformer. The input of each model was the hourly load measurement from two AMI meters under that service transformer, the average hourly AMI voltage measurement, and the total load of that service transformer. The output of the model was the voltage on the primary side of the service transformer.

The data from the first month were selected to compare the estimation accuracy of different algorithms. K-fold cross validation was used to validate the machine learning models, and the validation was repeated 30 times. In each test, 80% of the monthly data was randomly drawn from the data set to train the model, and the remaining 20% was used for testing. The mean absolute percentage error (MAPE) and maximum absolute percentage error between the synthetic primary voltage and estimated primary voltage was used to evaluate the performance of each machine learning method. The performance comparison is shown in Table 4.

Table 4. Performance of Different Methods

	Machine Learning Method		
	Random Forest	**Adaptive Boost**	**Gradient Boost**
MAPE	0.12%	0.75%	0.48%
Maximum	0.46%	1.08%	0.95%

The results summarized in Table 4 show that the Random Forest model performs better than the other two models in the selected performance criteria; therefore, it was selected to estimate the primary-side voltages in this study. Another advantage of using the Random Forest algorithm (as Random Forest is an ensemble learning method that integrates multiple decision trees) is that it will combine these decision trees and use average, or voting, schemes to calculate the results. Therefore, any outliers in the AMI measurements can be handled with this algorithm. Further, an exhaustive search was conducted to determine the model parameters (number of decision trees and maximum depth). These two parameters are varied from one to 500 and one to 30, respectively, to test the estimation performance. Considering both estimation accuracy and training time, the number of decision trees were selected to be 80 and the maximum depth to be 10. The time to build the machine learning model for each service transformer was around five seconds, and the total time for building the models for all service transformers was 30 minutes. As the process of training the model is usually developed for the distribution system planning studies, it meets the run-time requirement.

The performance of the machine learning-based approach was validated by the synthetic primary-side voltages generated from the QSTS simulation of the feeder model in OpenDSS. A secondary model was built for each service transformer in OpenDSS. Each secondary model included the two loads with voltage measurements and a load without voltage measurement. The load profiles of each secondary measured load were set to be one AMI measured power for one meter under that transformer. The load profile of the unmeasured load was set to be the AMI measured total power at that service transformer minus the two measured loads. The primary-side and secondary-side voltages at the two AMI measured loads recorded from the QSTS simulation were used to train the machine learning model. The data from the first 1,000 hours were used as a training data set to train the model for each service transformer, and the next 1,568-hour data were used to test the performance of each machine learning model.

The MAPEs for the estimation of all service transformer primary-side voltages are shown in Figure 13. All of them are less than 0.07%. Although the largest estimation error is around 0.65%, the number of such occurrences is very small. For most estimations, the error is less than 0.02%. Overall, the MAPE for all predictions in the feeder is 0.012%, and the MAPE for the service transformer with maximum error is 0.056%. Comparing with the 4% error from the physics-based method, the estimation accuracy improved a lot, however, this method requires some primary voltage data to train the machine learning model for each node. The comparison between estimated and actual voltages (synthetic voltage, in this case) for one example service transformer is shown in the two subplots of Figure 14. The first subplot shows the voltage comparison, and the second subplot shows the estimation absolute percentage error at each time step. Generally, the shape of the estimated voltages follows the actual voltages. The mismatch between the estimated and actual voltages is within 0.2%, which is very small. The model was also tested when using the first 2,000-hour data as the training dataset and tested with the remaining 568-hour data. The performance is similar to the previous case, which means the over-fitting problem does not exist for the model.

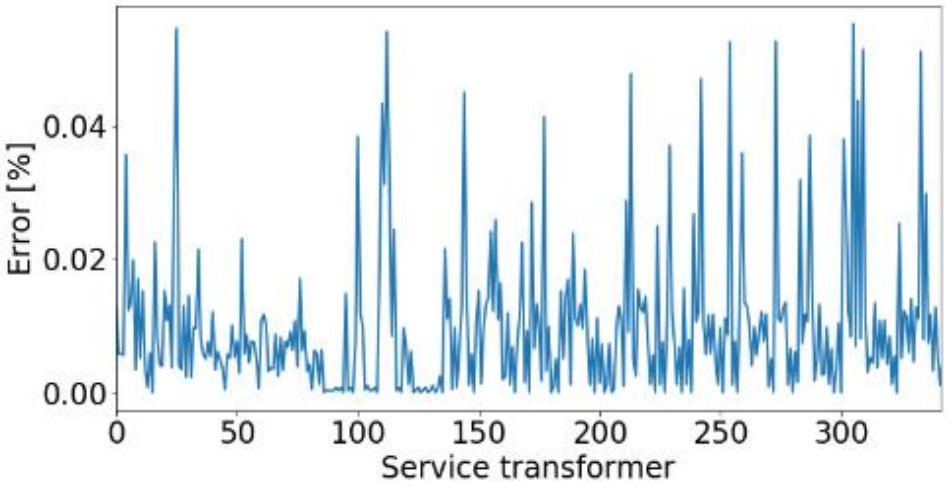

Figure 13. MAPE for the voltage estimation of each service transformer

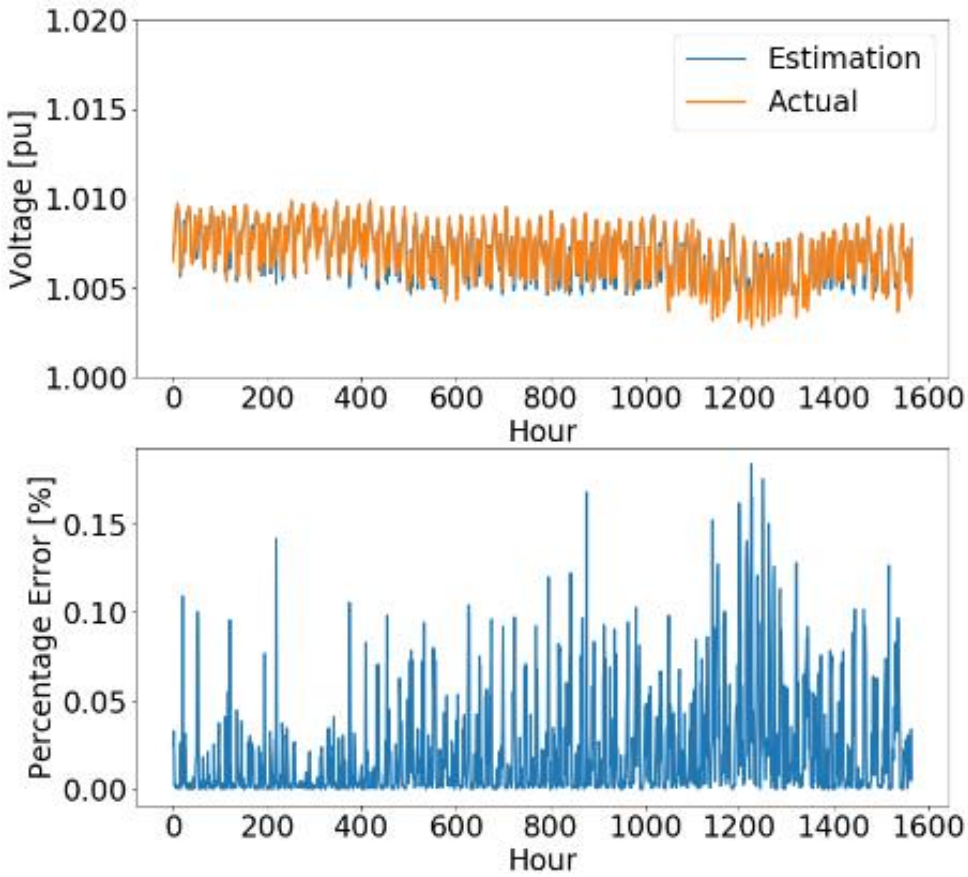

Figure 14. Comparison between estimated and synthetic voltages for one example service transformer

Combined Method

The performance of the combined method was validated by the synthetic primary-side voltages generated from the QSTS simulation of the OpenDSS model, since the primary-side voltages were not recorded in the field. The QSTS simulation was performed and the primary-side and secondary-side voltages at the two AMI measured loads for the five selected secondaries were recorded. These measurements from the secondary were used as AMI data to validate the combined methods. The estimation results are summarized in Table 5. From the results it can be observed that the estimation error decreased from 4% to 1% in the 2,000-hour testing period. Most of the time the errors are within 1%. If some other information is available, for example, some secondary topologies or the reactive power measurement, we can integrate them in the existing method to improve the estimation accuracy. This combined method was developed as a tool to estimate the primary voltages by using secondary AIM measurements.

Table 5. Combined method validation results

	Estimation Mismatch	
	Physics-based	**Combined**
Secondary 1	3.74%	0.90%
Secondary 2	4.06%	0.59%
Secondary 3	3.53%	0.43%
Secondary 4	4.28%	0.79%
Secondary 5	3.84%	1.40%

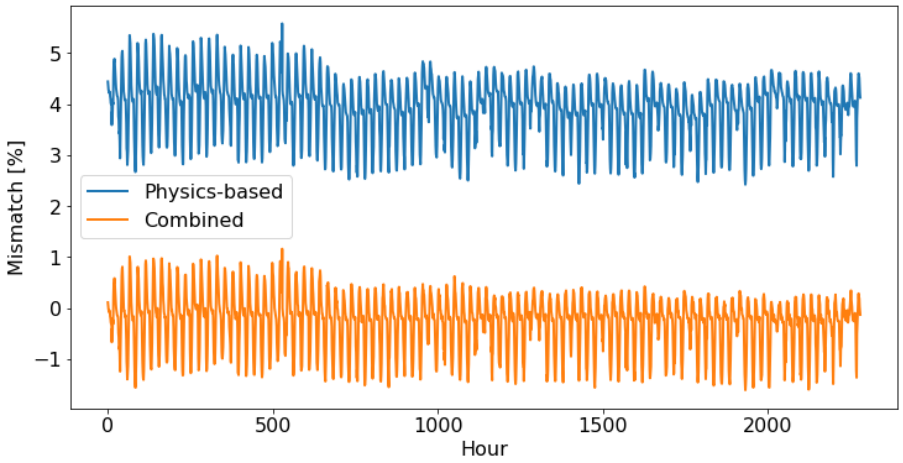

Figure 15. Estimation mismatch from one example secondary model

Identifying Planning Model Anomalies

The combined method was used to identify the anomalies in the distribution network planning model of Feeder A first. For this, the primary voltages estimated by the combined method for a selected duration were compared with those obtained from the time-series simulation of the distribution network planning model for the same duration. The peak load and minimum load days in December 2018 and January 2019 (four days) were selected for this process. The average estimation mismatches for all primary buses are shown in Figure 16 and the histogram of all estimation mismatches is shown in Figure 17. The geographic plot is shown in Figure 18.

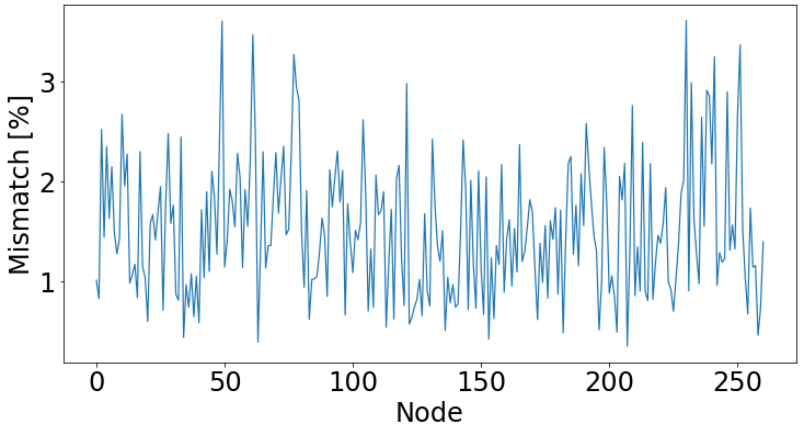

Figure 16. Estimation mismatch for all primary buses on Feeder A

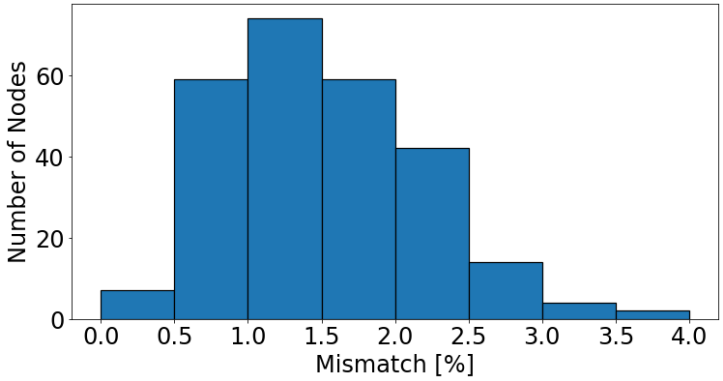

Figure 17. Histogram of all estimation mismatches in Feeder A

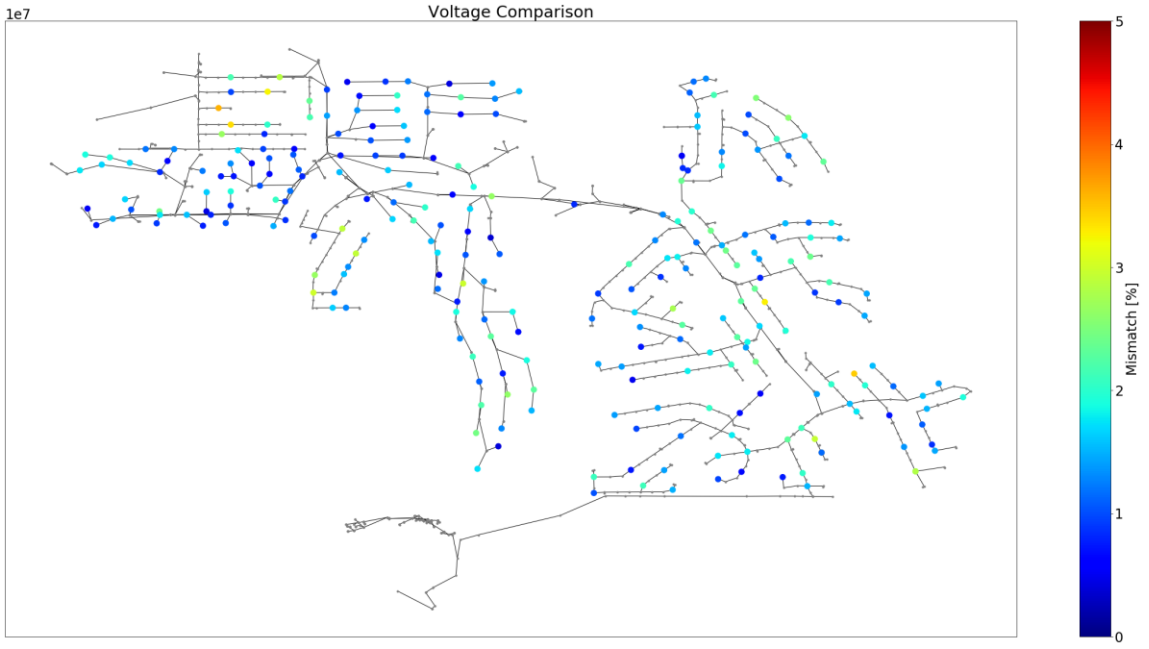

Figure 18. Geographic plot with the mismatch distribution on Feeder A

The combined method was also used to identify the anomalies in the distribution network planning model of Feeder B first. For this, the primary voltages estimated by the combined method for a selected duration were compared with those obtained from the time-series simulation of the distribution network planning model for the same duration. The peak load and minimum load days in August and September 2019 (four days) were selected for this process. The average estimation mismatches for all primary buses are shown in Figure 19 and the histogram of all estimation mismatches is shown in Figure 20. The geographic plot is shown in Figure 21.

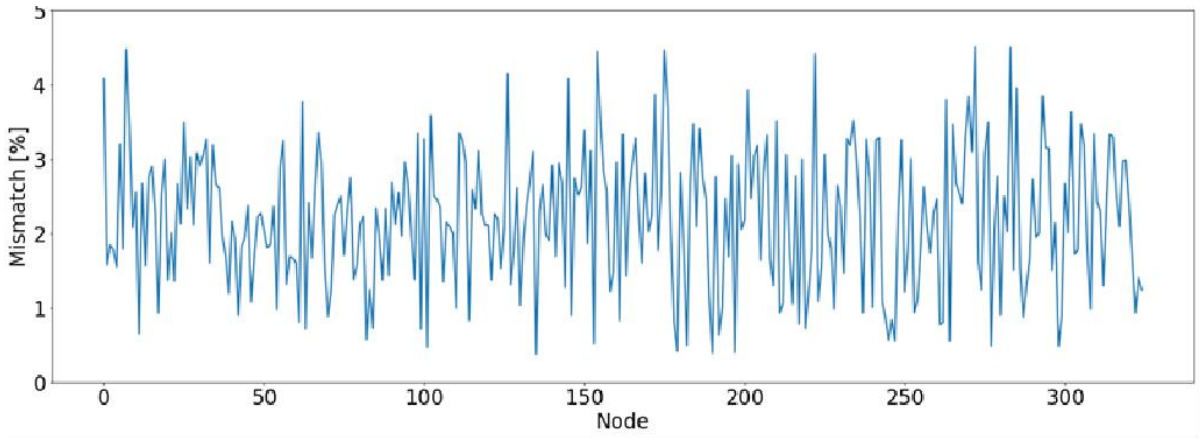

Figure 19. Estimation mismatch for all nodes on Feeder B

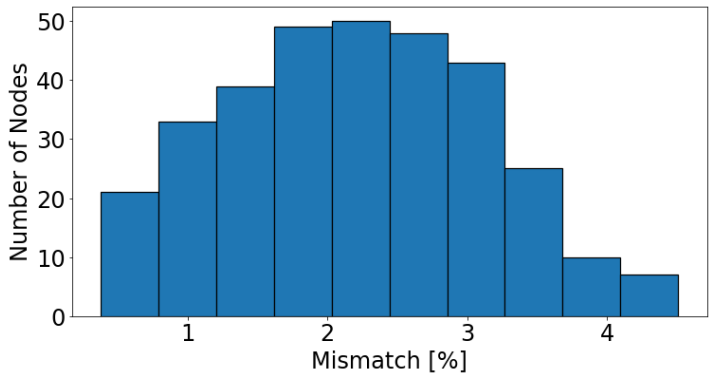

Figure 20. Histogram for all estimation mismatches for Feeder B

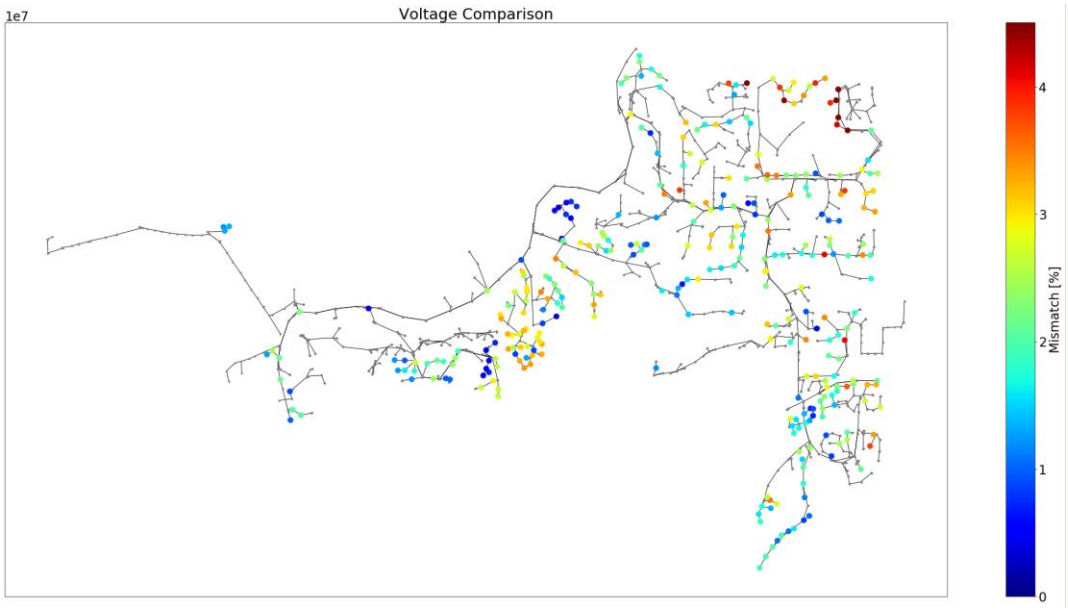

Figure 21. Geographic plot with the mismatch distribution on Feeder B

7.2.3 Challenges

One of the key challenges was the lack of measurement data. While the existing methods in the literature assume that every load in the secondary circuit is monitored by a smart meter, this study assumes that only two loads in the secondary circuit are monitored by smart meters. Lack of measurement data on the primary caused difficulties in the validation process. If the relevant network data, such as line and service transformer impedance parameters were available, that information could have been incorporated into the methods to achieve better accuracy levels.

7.3 AMI Meter-to-Transformer Mapping

With emerging technologies and distributed energy resources, the nature of power distribution systems has changed dramatically. Utility power engineers need more measurement data at the distribution level to monitor and keep the system stable. The AMI enables collection of a tremendous amount of data at the distribution level and can be used to design, test, and implement sophisticated distribution planning and control strategies. Real-time data recorded by AMI meters, along with the information recorded and provided by the GIS, improve the observability of distribution power systems.

7.3.1 Configuration and Methodology

The methodology for meter-to-transformer mapping solely uses the voltage data recorded by the AMI. There is no requirement for the length of the data. Additionally, the methodology accommodates missing data, which can be observed frequently in the voltage dataset.

The main idea is that the voltage measurements from the AMI meters connected to the same service transformer secondary should be highly correlated and have a high correlation coefficient. Therefore, the key is to find a threshold to identify the potential incorrect records. If the correlation coefficient is

lower than the threshold, that means records are incorrect. The AMI meter to service transformer mapping procedure consists of the following steps:

1) Calculate correlation coefficient between meters connected to the same service transformer for the records in the existing database.
2) Rank the calculated correlation coefficients and pick a threshold.
3) Loop through all the correlation coefficients and select the AMI records whose correlation coefficients are lower than the threshold.
4) Calculate the correlation coefficients with the rest of the dataset and choose the service transformer with the highest score.
5) Perform Step 4 for each meter in the selected set of records in Step 3. If the score is higher than the threshold, then correct the record to the new service transformer and meter pair, otherwise keep the record the same.

For instance, consider that the correlation coefficient of AMI1 and AMI2 are lower than threshold τ in Figure 22, and they are connected to the service Transformer 1 in the original record. Then the algorithm will check the rest of the dataset to find a transformer with the highest correlation coefficient for both AMI1 and AMI2. If the score is higher than τ , then the algorithm will update the records. In this case, the mapping of AMI1 is changed to associate with Transformer 2 after running the algorithm.

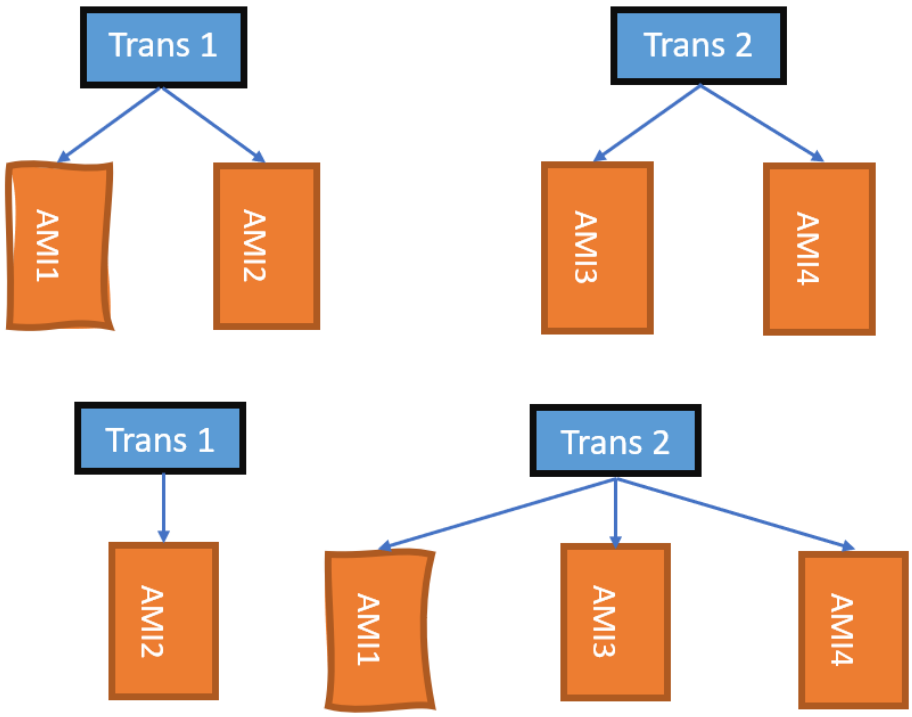

Figure 22. Illustration of the AMI meter to service transformer mapping algorithm

7.3.2 Results and Discussion

Recall and precision are used to evaluate the performance of the proposed methodology. Recall evaluates the overall accuracy of the algorithm, and precision evaluates the accuracy of the correction. The equations below show the definition of recall and precision in mathematical formulas. In our case, the true positive is the right correction, the false positive is the wrong correction, and the false negative means the records remain the same as before.

$$R(recall) = \frac{TP(True\ Positive)}{TP + FN(False\ Negative)}$$

$$P(precision) = \frac{TP(True\ Positive)}{TP + FP(False\ Positive)}$$

Three correlation coefficients namely Pearson [13], Kendall's rank [14], and Spearman's rank [15] are used in this work. These are defined as follows:

The Pearson correlation coefficient is a measure of linear correlation between two sets of data. It is the ratio between the covariance of two variables and the product of their standard deviations. The Kendall rank correlation coefficient is a statistic used to measure the ordinal association between measured quantities. It is a measure of rank correlation and the similarity of the orderings of the data when ranked by each of the quantities. Spearman's rank correlation coefficient is a nonparametric measure of rank correlation. It assesses how well the relationship between two variables can be described using a monotonic function.

Tables 6, 7, and 8 below, show the results of the methodology with three different correlation coefficient calculation methods.

Table 9 indicates that the Spearman correlation coefficient has the highest score in both recall and precision, and therefore fits the methodology best. All methods detect 97% of the incorrect records.

Table 9 summarizes the overall comparison results. The category, "# of swap" means the total number of incorrect records which are randomly created to test the methodology.

Table 6. Pearson correlation coefficient

Test	# of Swap	Detection	Correction	Right Correction
1	10	10	7	5
2	15	15	8	8
3	20	20	11	10
4	25	24	15	12
5	30	28	22	16
Summary	100	97%	63%	51%

Table 7. Spearman correlation coefficient

Test	# of Swap	Detection	Correction	Right Correction
1	10	10	7	5
2	15	15	10	9
3	20	20	12	10
4	25	24	15	13
5	30	28	23	17
Summary	100	97%	67%	54%

Table 8. Kendall correlation coefficient

Test	# of Swap	Detection	Correction	Right Correction
1	10	10	7	5
2	15	15	11	10
3	20	20	12	10
4	25	24	16	13
5	30	28	23	18
Summary	100	97%	69%	56%

Table 9. Method comparison

Method	Recall	Precision
Pearson	63%	80.95%
Kendall	67%	80.60%
Spearman	69%	81.16%

7.3.3 Challenges

In this exploratory task, our proposed algorithm detects almost all the incorrect records, achieves 81.16% of precision, and 69% of recall with the least information. However, there is 30% of the detection that cannot be assigned with correct service transformers. This part needs to be addressed and will further improve the overall performance.

7.4 Phase Identification using AMI Data

An accurate phase connectivity database is needed for the utility distribution networks for efficient grid operations [16]. This requires periodic updates by field verification to keep the phase connectivity database accurate, which is an expensive and time-consuming process. A faster and less expensive method of estimating the phase connectivity is necessary.

The widespread deployment of AMI presents opportunities to develop applications for grid planning and operations using measurement data. The AMI data can be used for identifying the phase connectivity of each AMI meter. This process is referred to as phase identification. Existing phase identification techniques that estimate phase connectivity work well in distribution feeders that have low or no PV generation; however, they fail to identify the phases accurately when considerable PV generation is present [17]. Further, some existing phase identification approaches can perform phase identification for phase-to-neutral or phase-to-phase connection only. They cannot be applied to the distribution feeders having both phase-to-neutral and phase-to-phase connections. In this demonstration, a phase identification algorithm was used that can be applied to distribution feeders having a mix of phase-to-neutral and phase-to-phase AMI meter connections. The algorithm used supervised machine learning to detect the phase connectivity accurately even in the presence of high PV penetration. The phase identification performance was validated on two feeders. The first feeder has a mix of phase-to-neutral and phase-to-phase AMI meter connectivity and nearly 70% PV level relative to the peak load. The second feeder has predominantly phase-to-neutral AMI meter connectivity, with very few service transformers having phase-to-phase connectivity. This feeder also has approximately 24% PV penetration relative to the peak load.

7.4.1 Configuration and Methodology

The key assumption in this study is that voltage profiles from AMI meters pertaining to each phase connectivity are highly correlated with each other. Thus, the voltage magnitude time series of the AMI meters that are on the same phase tend to exhibit similar variations in the voltage measurements which are different from the meters on the other phases.

Phase Identification using Supervised Learning

Phase identification was performed using the random forest classifier [18] [19]. The random forest classifier is a supervised machine learning model. In the phase identification process, first the voltage magnitude time series from each meter in the AMI dataset was obtained for a selected duration of time. Next, a preset percentage of meters were selected for each phase connectivity as a training dataset for the supervised machine learning algorithm. The phase connectivity of these meters must be accurate since this is part of the training process for the machine learning algorithm. Then a random forest classifier was constructed, which is a function that predicts the phase connectivity of each meter in the training dataset based on the voltage magnitude time series data. Finally, the trained random forest classifier was used to identify the phase connectivity of the rest of the meters in the AMI dataset based on their voltage magnitude time series.

The phase identification algorithm steps are given below:

1) *Data preprocessing:* Load the AMI dataset with the voltage magnitude time series data and perform data standardization. A small number of meters with consistently reported bad data or empty data were removed from the AMI dataset in this step.
2) *Training the random forest classifier:* Select 30% of the AMI meters for each phase connectivity for training the random forest classifier. The phase connectivity of these AMI meters, obtained

through field validation or some other means, was supplied to the random forest classifier in this step.

3) *Phase identification:* Input the voltage magnitude time series data of the rest of the meters to the random forest classifier model trained in Step 2 to identify the phase connectivity of the rest of the meters in the AMI dataset.

Data Requirements

The phase identification algorithm in this demonstration used the voltage magnitude time series data of the AMI meters and the validated phase connectivity information for 30% of the meters in the AMI dataset for training. The AMI dataset had average, maximum, and minimum voltages for each meter at five-minute intervals. The five-minute average voltage magnitude data was used for phase identification. The inputs and output of this algorithm are summarized below:

Inputs:

- Average voltage magnitude time series data (preferably three months or more) for each AMI meter at five-minute resolution
- Accurate phase connectivity information for 30% of the AMI meters for each type of phase connectivity

Output:

- A table with AMI meter ID and associated phase connectivity for all the AMI meters

7.4.2 Results and Discussion

The phase identification was performed on two of SDG&E's feeders, namely Feeder A and Feeder B. Two AMI datasets are used for Feeder A. The first dataset has the AMI data of the three-month period between October 1, 2018 to December 31, 2018 (2018 dataset) and the second dataset has the AMI data for the entire 2019-year period. For Feeder B, the AMI data for the entire year of 2019 was used. The phase identification results of the two feeders are documented in this section.

Distribution Feeder Details

The first feeder used for the phase identification is Feeder A. This is a 12-kV feeder with a peak load of 10.3 MW. The topology of the feeder is shown in Figure 23. The substation transformer is equipped with a load tap changer. Three capacitor banks are available on the feeder for reactive power support. The feeder serves more than 5,000 customers using 341 service transformers. Solar generation of approximately 70% relative to the peak load is present in this feeder.

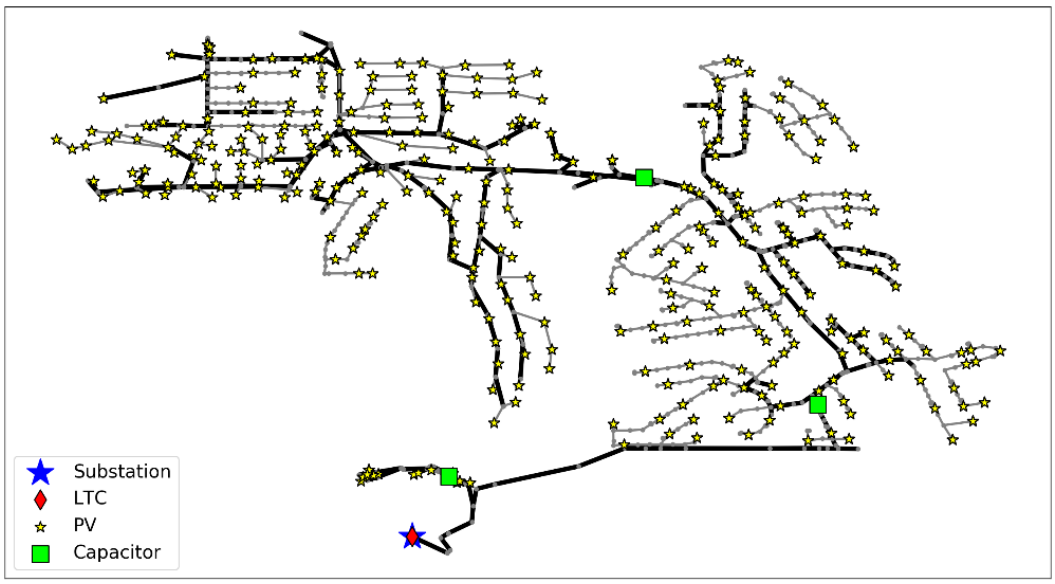

Figure 23. Topology of Feeder A

The second feeder used for the phase identification is Feeder B. This is also a 12-kV feeder with a peak load of 13.29 MW. The topology of the feeder is shown in **Error! Reference source not found.**4. The substation transformer is equipped with a load tap changer. Two capacitor banks are available on the feeder for reactive power support and there are no line voltage regulators. This feeder has 657 service transformers. Solar generation of 3.14 MW is present in this feeder which is approximately 24% relative to the peak load.

Figure 24. Topology of Feeder B

Phase Identification in Feeder A using 2018 AMI Dataset

Phase identification was performed on Feeder A first using the 2018 AMI dataset. This dataset has the average voltage magnitude time series data for 561 AMI meters for the three-month period between October 1, 2018 to December 31, 2018. The AMI data for two meters per service transformer were available in this dataset. Additionally, the field validated phasing information was also available for all the meters. Based on this information, the distribution of the phasing for the AMI meters is shown in Figure 25.

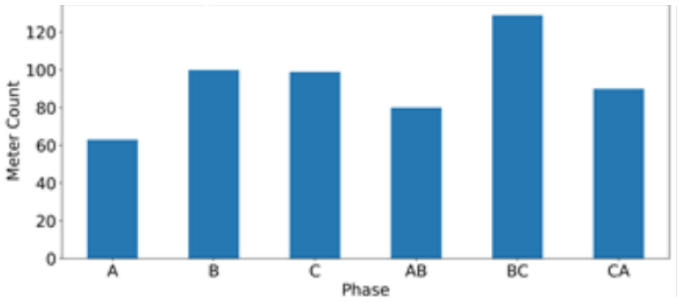

Figure 25. AMI meter phasing distribution in 2018 AMI dataset of Feeder A

The phase identification results are shown in Figure 26. The field validated phasing information is considered as the ground truth. For each type of phase connectivity, the number of meters the algorithm identified as pertaining to that connectivity is shown against the ground truth. The results show the phase identification algorithm can identify all the types of phase connectivity accurately.

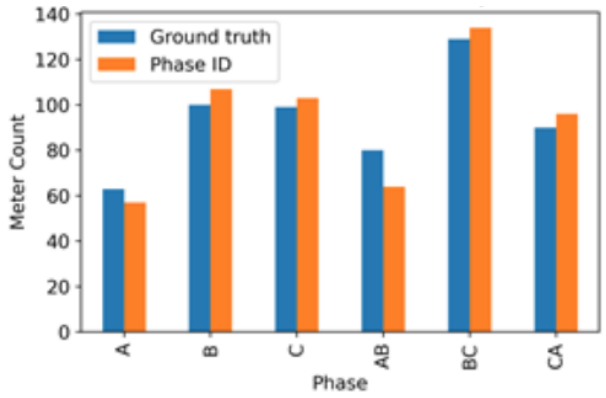

Figure 26. Phase identification results of Feeder A using 2018 AMI dataset

The detailed breakdown of the AMI meter counts in each of the training, testing, and full datasets of phase identification are shown in Table 10. For each type of phase connectivity, 30% of the meters are selected randomly along with their ground truth phase connectivity for the training dataset. The full dataset includes both the training and testing datasets together. With the phase connectivity identified accurately for 335 out of 391 meters in the testing set alone, the phase identification accuracy is 85.7% on the testing set. The phase identification accuracy on the training and full datasets are 100% and 90%, respectively.

Table 10. Summary of phase identification results of Feeder A using 2018 AMI dataset

Dataset		Phase Connectivity						Total	Accuracy
		A	B	C	AB	BC	CA		
Full	Ground truth	63	100	99	80	129	90	561	90%
	Phase identification	53	96	96	60	122	78	505	
Testing	Ground truth	43	67	70	57	92	62	391	85.7%
	Phase identification	33	63	67	37	85	50	335	
Training	Ground truth	20	33	29	23	37	28	170	100%
	Phase identification	20	33	29	23	37	28	170	

The geographic distribution of the AMI meters for which the phase connectivity identified by the algorithm matched the ground truth is shown in Figure 27. The meters are distributed all over the feeder; thus, the algorithm can detect the correct phase connectivity in all the feeder neighborhoods.

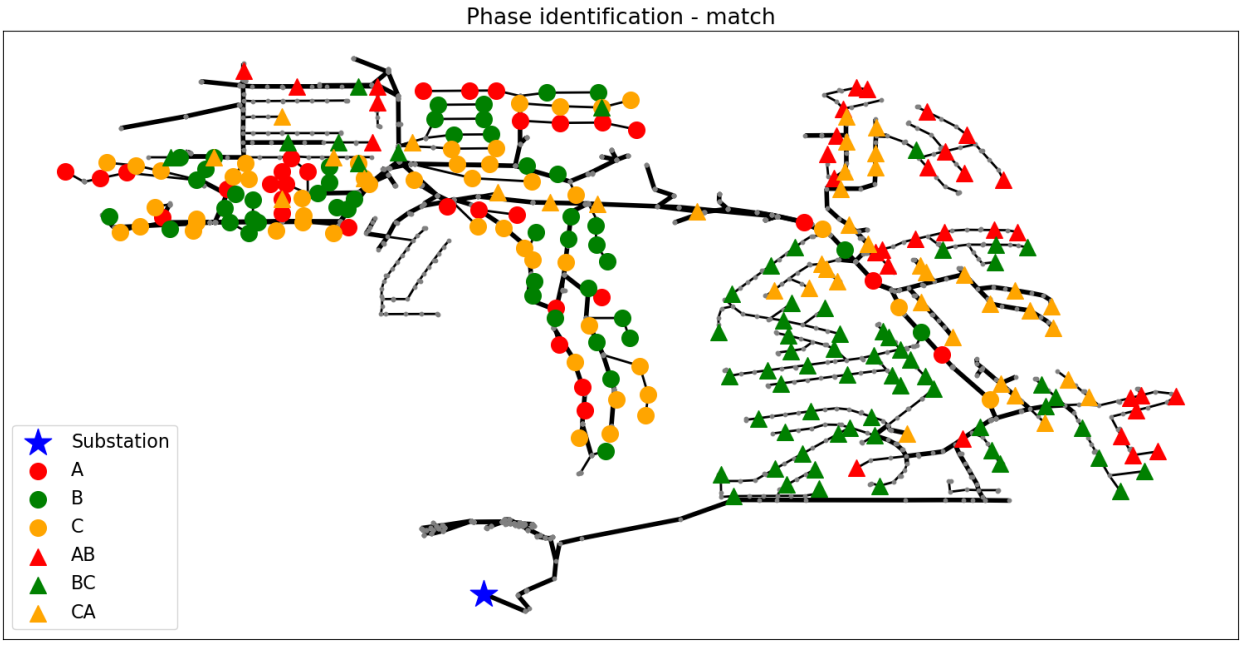

Figure 27. Locations of AMI meters for which the phase connectivity is identified correctly

The locations of the AMI meters where the identified phase connectivity does not match the ground truth are shown in Figure 28. The correct phase connectivity according to the ground truth is shown in this figure at these locations. The mismatches are generally not clustered or constrained to any specific feeder neighborhoods.

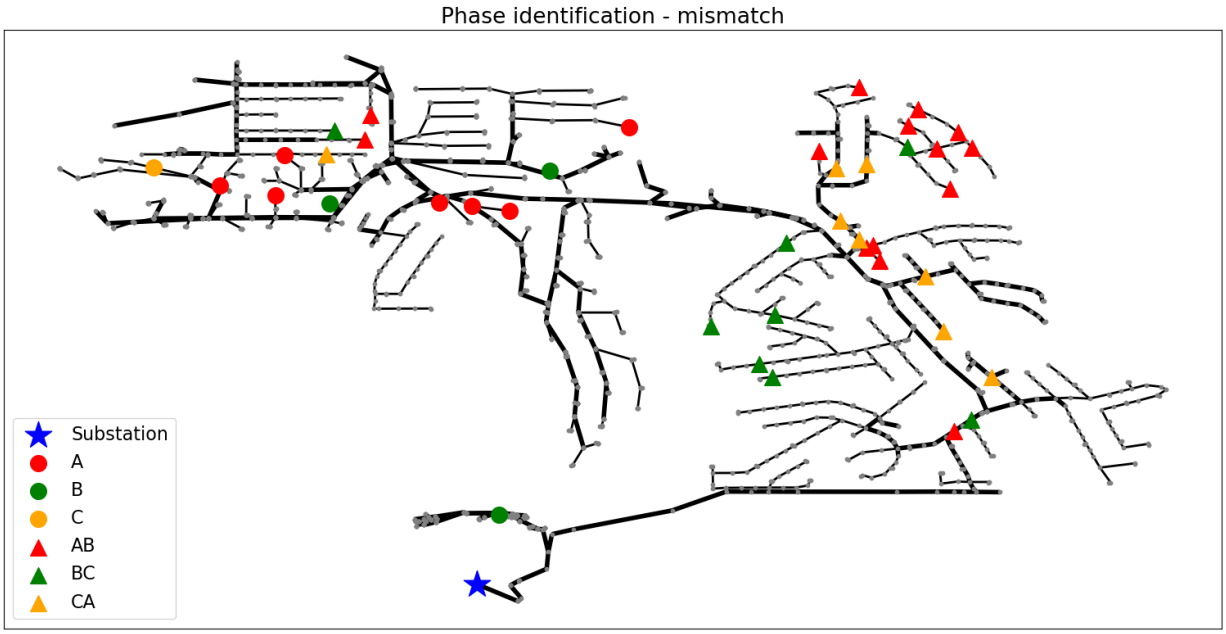

Figure 28. Locations of AMI meters for which the phase connectivity is identified incorrectly

Phase Identification in Feeder A using 2019 AMI Dataset

The phase identification was also performed on Feeder A using the 2019 AMI dataset. This dataset has the average voltage magnitude time series data for 568 AMI meters for the full 2019-year period. The AMI data for two meters per service transformer were available in this dataset in addition to the field validated phasing information. The phase identification results are shown in Figure 29. The results are similar to those obtained using the 2018 AMI dataset and show that the phase identification algorithm can identify all the types of phase connectivity accurately.

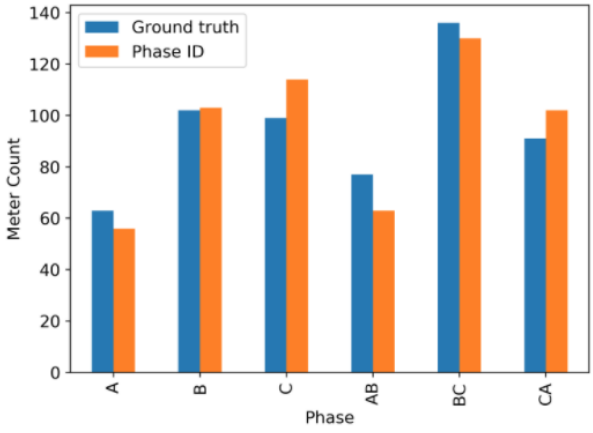

Figure 29. Phase identification results of Feeder A using 2019 AMI dataset

The detailed breakdown of the AMI meter counts in each of the training, testing, and full datasets of phase identification are shown in Table**Error! Reference source not found.** 11. For each type of phase

connectivity, 30% of the meters were selected randomly along with their ground truth phase connectivity for the training dataset. The phase identification accuracies on the testing, training, and full datasets are 86.5%, 100%, and 90.5%, respectively.

Table 11. Summary of phase identification results of Feeder A using 2019 AMI dataset.

Dataset		Phase Connectivity						Total	Accuracy
		A	B	C	AB	BC	CA		
Full	**Ground truth**	63	102	99	77	136	91	568	90.5%
	Phase identification	55	98	98	56	126	81	514	
Testing	**Ground truth**	45	72	70	54	96	64	401	86.5%
	Phase identification	37	68	69	33	86	54	347	
Training	**Ground truth**	18	30	29	23	40	27	167	100%
	Phase identification	18	30	29	23	40	27	167	

The geographic distribution of the AMI meters for which the phase connectivity identified by the algorithm match the ground truth is shown in Figure 30. The meters are distributed all over the feeder; thus, the algorithm can detect the correct phase connectivity in all the feeder neighborhoods.

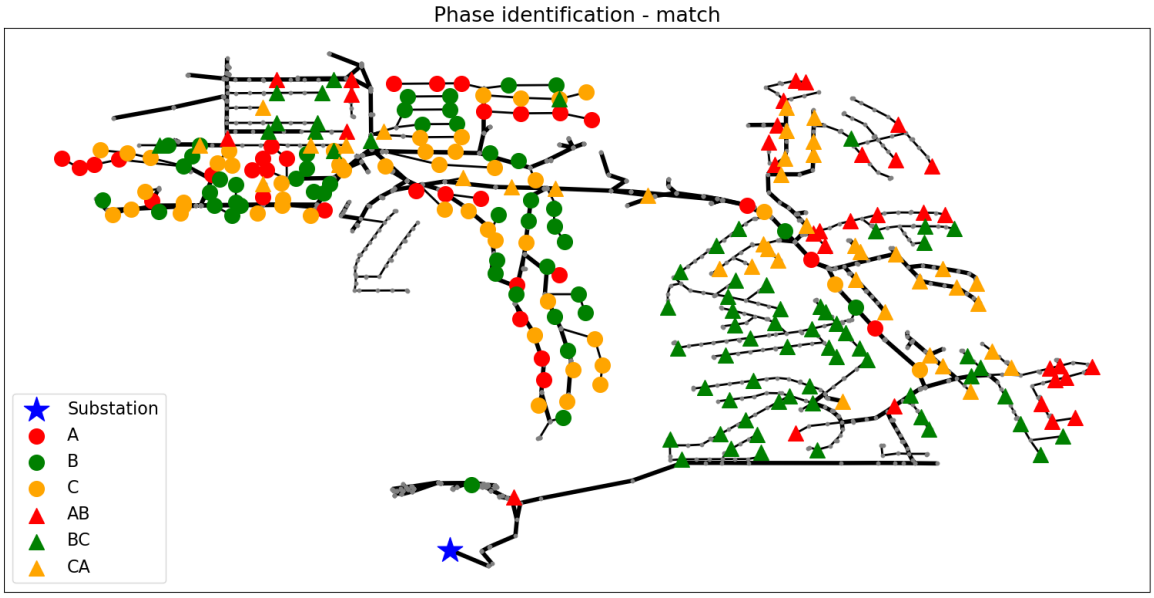

Figure 30. Locations of AMI meters for which the phase connectivity is identified correctly

The locations of the AMI meters where the identified phase connectivity does not match the ground truth are shown in Figure 31. The correct phase connectivity according to the ground truth is shown in this figure at these locations.

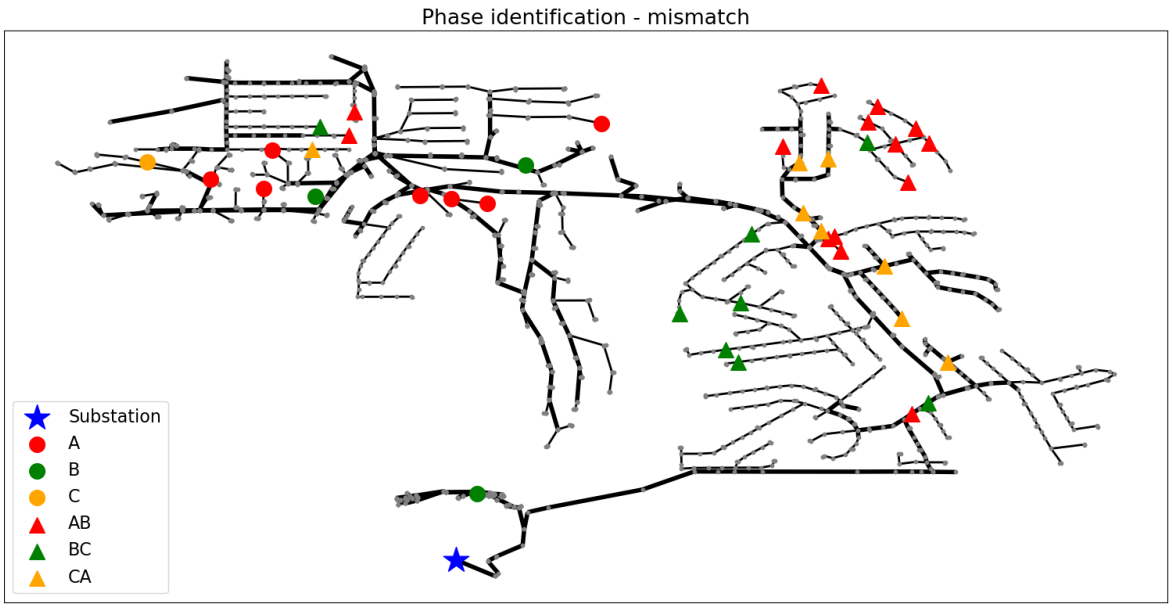

Figure 31. Locations of AMI meters for which the phase connectivity is identified incorrectly

Phase Identification in Feeder B using 2019 AMI Dataset

The phase identification algorithm was applied to the 2019 AMI dataset of Feeder B. This dataset has the average voltage magnitude time series data for 857 AMI meters for the full 2019-year period. The AMI data for two meters per service transformer were available in this dataset in addition to the field validated phasing information for these meters. The field validated phasing information was considered the ground truth. The phase identification results are shown in Figure 32. The ground truth phasing distribution in this figure indicates this feeder primarily has phase-to-neutral AMI phase connectivity. A small number of meters are connected to phase-to-phase.

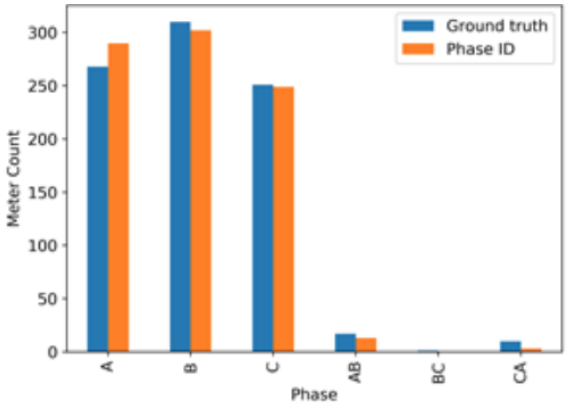

Figure 32. Phase identification results of Feeder B using 2019 AMI dataset

The detailed breakdown of the AMI meter counts in each of the training, testing, and full datasets of phase identification are shown in Table 12. For each type of phase connectivity, 30% of the meters are selected randomly along with their ground truth phase connectivity for the training dataset. With the

phase connectivity identified accurately for 809 out of 857 meters in the testing set alone, the phase identification accuracy is 94.4% on the testing set. The phase identification accuracy on the training and full datasets are 100% and 92%, respectively.

Table 12. Summary of phase identification results of Feeder B using 2019 AMI dataset

Dataset		Phase Connectivity						Total	Accuracy
		A	B	C	AB	BC	CA		
Full	**Ground truth**	268	310	251	17	1	10	857	94.4%
	Phase identification	260	293	241	12	0	3	809	
Testing	**Ground truth**	188	217	176	12	1	7	601	92%
	Phase identification	180	200	166	7	0	0	553	
Training	**Ground truth**	80	93	75	5	0	3	256	100%
	Phase identification	80	93	75	5	0	3	256	

The geographic distribution of the AMI meters for which the predicted phase connectivity matches the ground truth is shown in Figure 33. The meters whose phase connectivity is identified correctly are distributed all over the feeder. This indicates the algorithm can detect the correct phase connectivity in all the feeder neighborhoods.

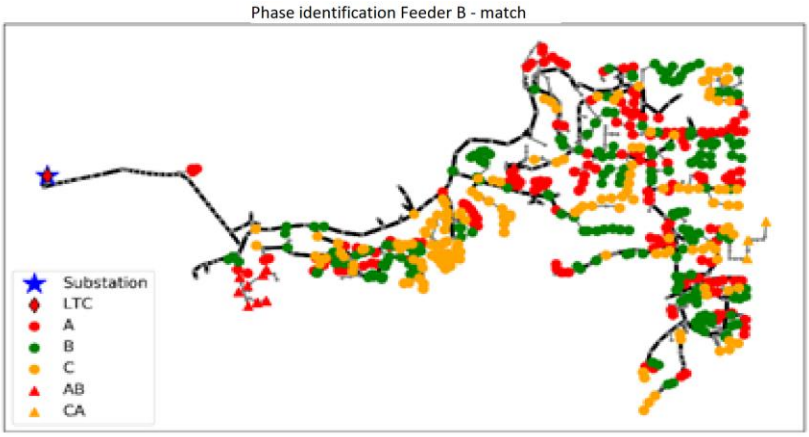

Figure 33. Locations of AMI meters for which the phase connectivity is identified correctly

The locations of the AMI meters where the identified phase connectivity does not match the ground truth are shown in Figure 34. The correct phase connectivity according to the ground truth is shown in this figure at these locations. The mismatches are generally not clustered or constrained to any specific feeder neighborhoods.

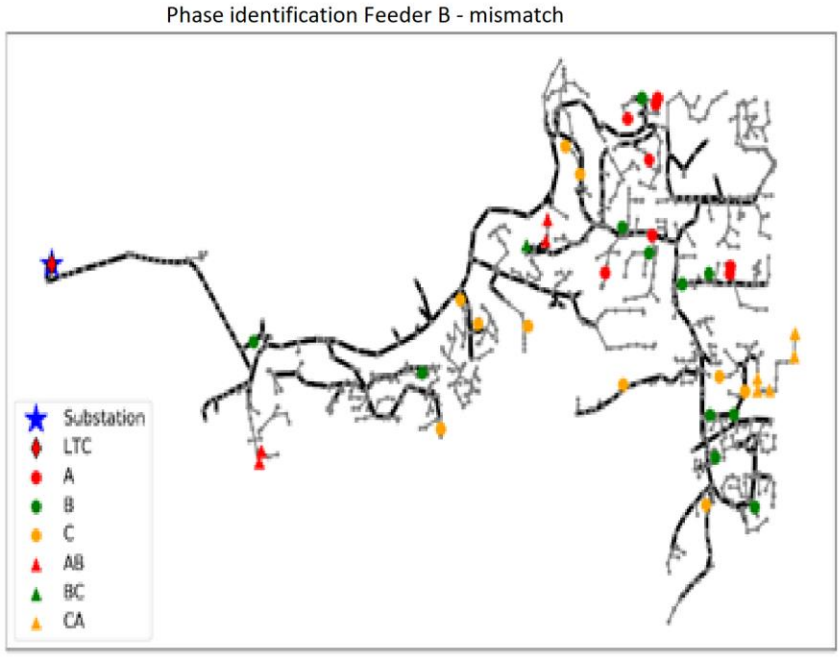

Figure 34. Locations of AMI meters for which the phase connectivity is identified incorrectly

7.4.3 Challenges

The lack of reasonably accurate data on the phase connectivity created issues in the algorithm configuration. Before the field verification, the algorithm configuration efforts primarily relied on the planning network model for the phase connectivity information. This consistently resulted in poor accuracies in any algorithm that was applied for phase identification. However, the field validated phasing information showed that the phase connectivity information in the planning network model is 60% incorrect in the case of Feeder A. Thus, the poor performance of the algorithms attempted during this project phase could have been due to using the incorrect phase connectivity database as reference.

7.5 Data-Centric Grid Operations

AMI data analytics help develop insights into distribution network operation including service quality, power consumption patterns, presence of DERs, etc. In this task, interactive 3D and 2D tools useful for visualizing essential information from vast amounts of AMI data were developed. Figure 35 below shows sample 2D visualizations on the type and duration of the voltage exceedances and the potential EV locations on the SDG&E feeder, based on the AMI data.

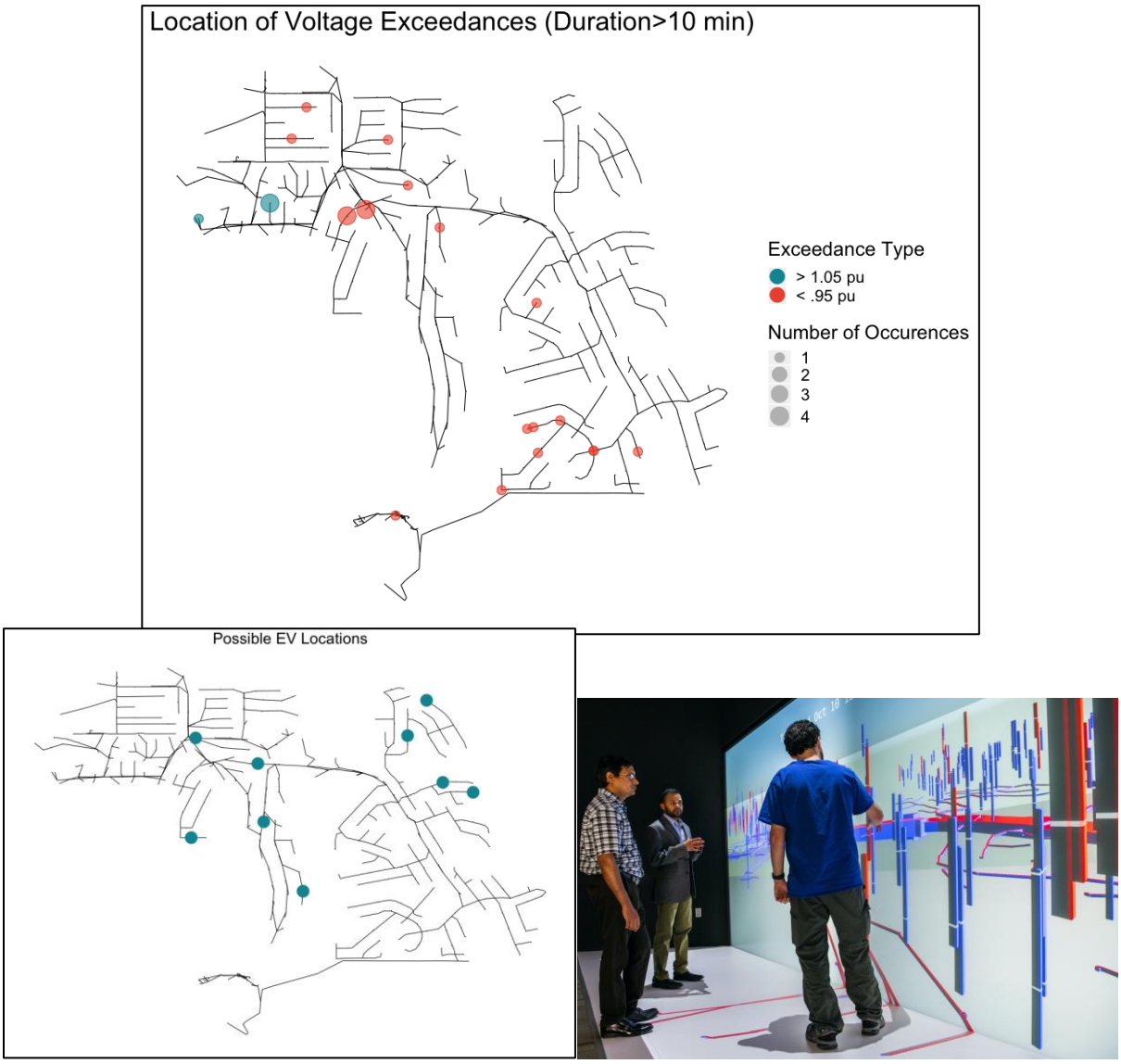

Figure 35. Interactive 2D and 3D visualization tools

AMI provides a new paradigm for utility planning, operations, and controls. The data captured by AMI can provide insights into the system dynamics at the grid edge. Utilities need tools to leverage this capability. In addition to AMI, utilities are also investing in the deployment of ADMS to prepare for future distribution operations with high levels of DERs. The ADMS is an integrated platform that combines the functionalities of distribution management systems, outage management systems, and SCADA systems for optimized distribution grid operations. Traditionally, the ADMS used the limited primary sensor measurements available on the distribution network through the SCADA system for network control decisions. With the availability of AMI measurements, there is significantly higher visibility, which can be leveraged for improved network monitoring and control.

7.5.1 Configuration and Methodology

In the voltage prediction task that attempts to predict the voltage issues in the distribution network, an AMI voltage forecasting model was prototyped for the distribution feeder. The secondaries were modeled as having three AMI data points, with the closest and furthest to the transformer as individual data points and a third one that is an aggregate of all other power consumed on the secondary. Synthetic average hourly voltage data was simulated for three and a half months. Two machine learning algorithms were used in modeling voltage time series data which can be used for forecasting. The models are learned globally and simultaneously process all AMI time series data. Simulations with various scenarios of available historical data (60, 30, and 15 days) were performed which were explicitly incorporated into the model and evaluated for performance. The performance of a model hyper parameter set when forecasted 24-hours ahead is shown in **Error! Reference source not found.** 36. The plot represents an averaging over five-folds in the validation set. The model is evaluated globally on all simulated meters on the distribution feeder.

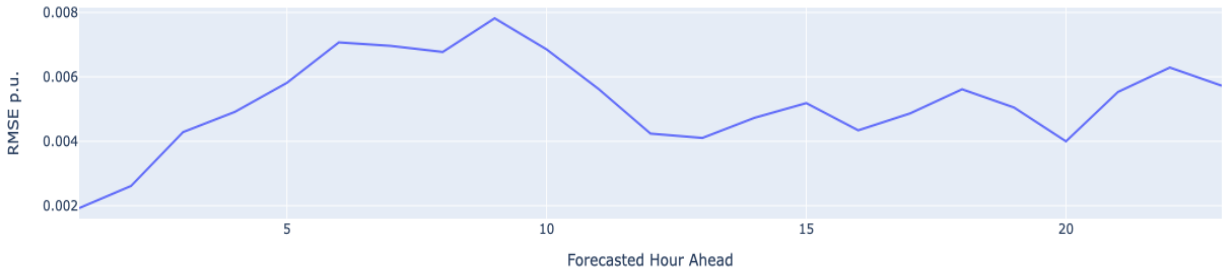

Figure 36. Twenty-four-hour forecast performance of a hyperparameter run averaged over five folds

The integration of ADMS and AMI measurements offers a unique opportunity to further modernize distribution system control. In this use case, an AMI-based data-driven volt/var control algorithm and its synergies with ADMS for distribution grid operations were evaluated using the SDG&E feeder and the AMI data. The inputs of this algorithm were AMI power and voltage measurements. The algorithm controls the LTC tap position, capacitor banks switches, and PV inverter setpoints to ensure voltage regulation.

7.5.2 Results and Discussion

Figure 37 and Figure 38 show the results from the evaluation of this algorithm. In the base case, the LTC and capacitor banks follow their local controllers, and the PV smart inverters inject power at unity power factor. In the unity power factor operation, the PV smart inverters inject active power only and no reactive power is injected or absorbed. As observed in **Error! Reference source not found.** 37, many customer voltages on the secondary are experiencing high voltage exceedances in the base case. In the next scenario in which the data-driven control algorithm is enabled, the voltage exceedances are significantly reduced, and the average voltages are closer to 1.0 PU Once the voltage deviates from the preset voltage regulation set point (selected 1.0 PU in this case), the algorithm primarily raises or lowers the LTC tap position to regulate the voltages as observed in Figure 38.

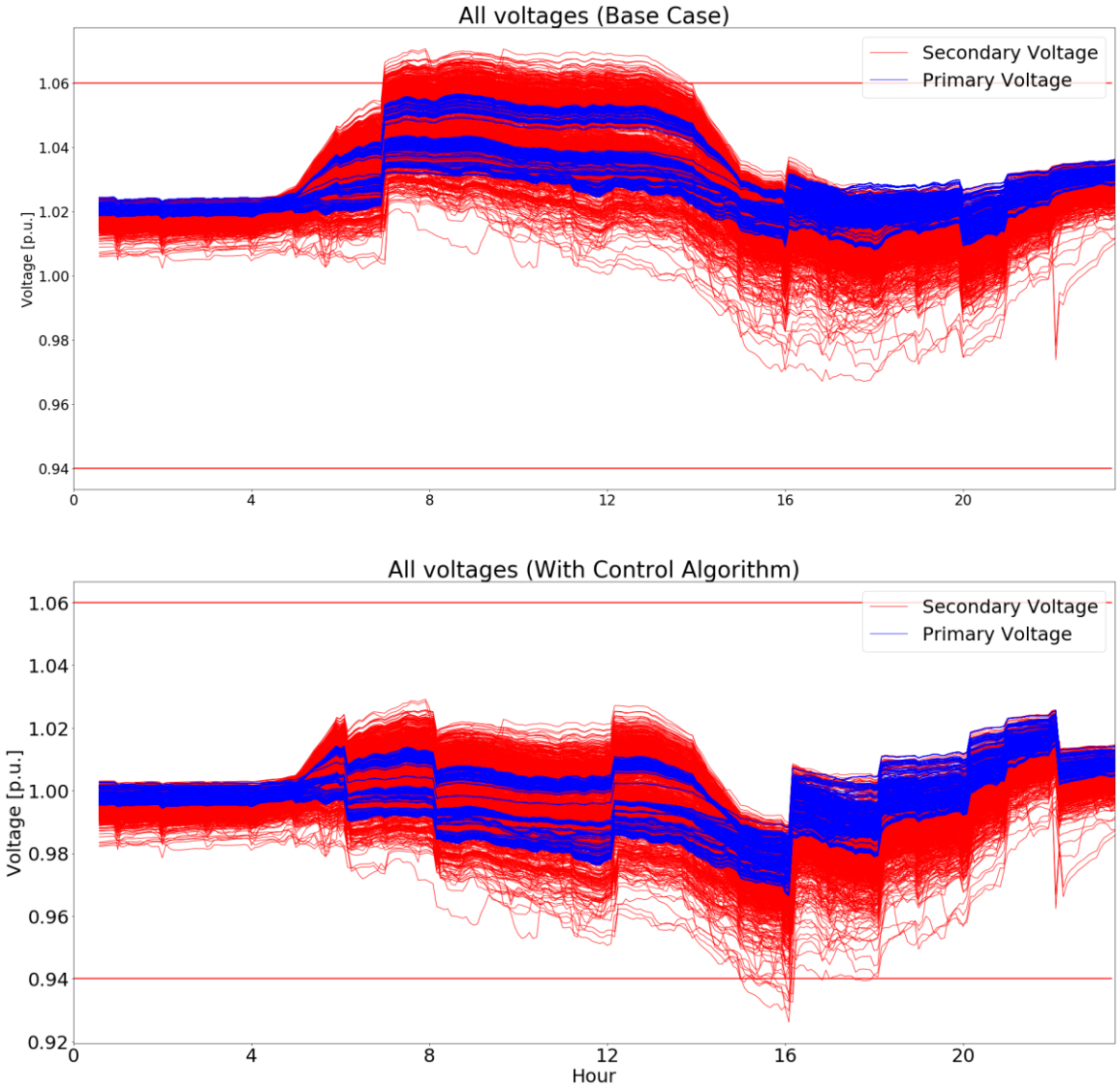

Figure 37. Comparison of Bus Voltages

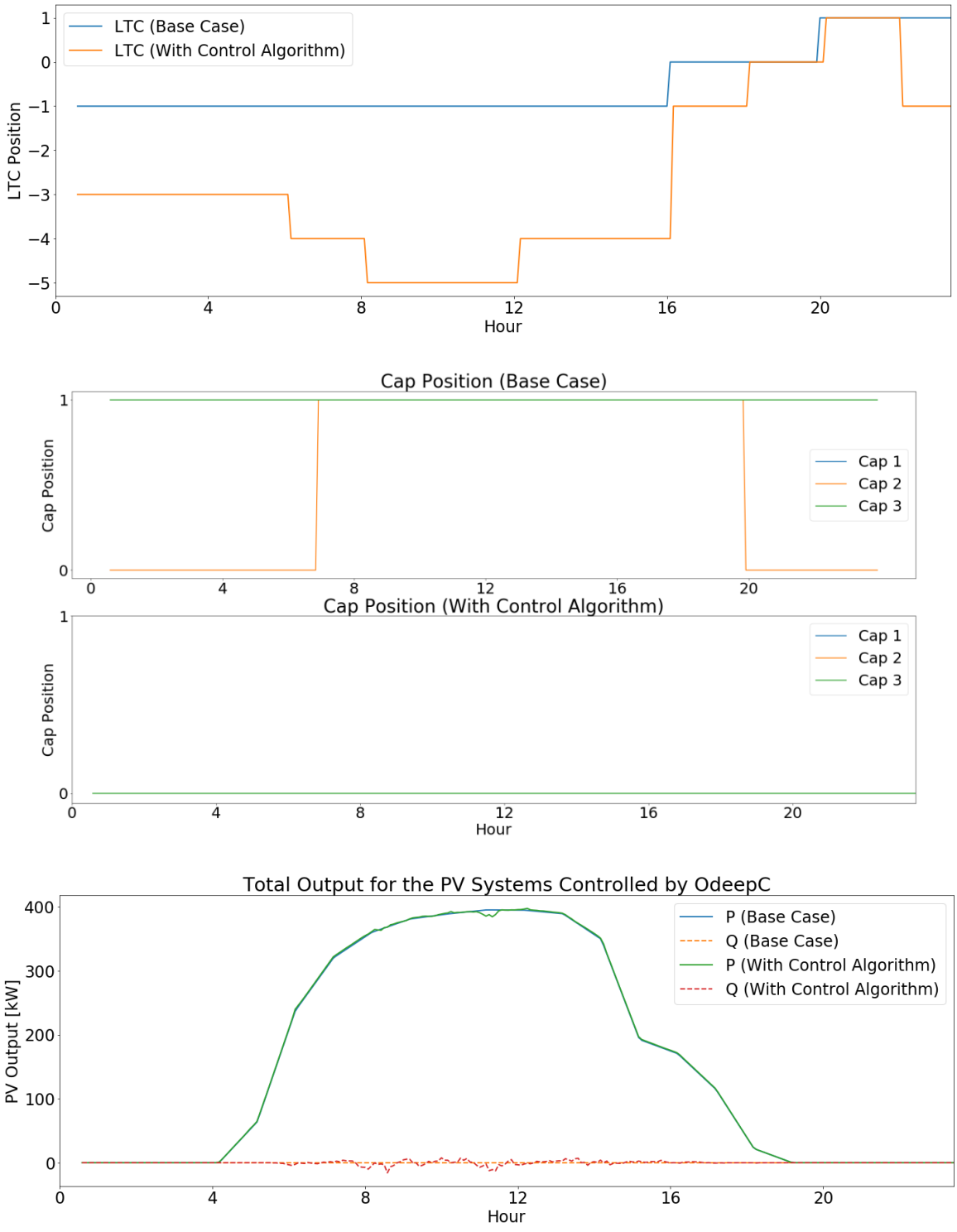

Figure 38. Comparison of LTC, capacitor bank statuses, and total PV generation

7.5.3 Challenges

The data-driven control demonstrated in this use case requires data from bellwether AMI meters every five minutes. The feasibility and scalability of this faster data reporting rate for real-time controls for all the feeders has not been explored. The utility control architecture to enable this new paradigm of grid operations is yet to be addressed.

8.0 Findings

8.1 Findings Discussion

The key findings for each of the use cases are presented in the section below.

8.1.1 PV Smart Inverter Study Key Findings

The PV smart inverter study considered various smart inverter settings from the standards. The primary objective of applied advanced inverter functionality is to support system voltage compliance. The quasi-static time-series simulation was performed for a period of 107 days with the different PV smart inverter settings enabled. When a node (bus) voltage magnitude exceeds the desirable range of 0.94 PU - 1.06 PU, it is considered a voltage exceedance (VE). The voltage exceedance node (VEN) refers to the node that has more than 12 hours of cumulative voltage exceedance in the 107-day period.

The voltage exceedance results are summarized in Table 13. Summary of voltage exceedances

The results show when the PV smart inverters are not utilized for the voltage regulation in the baseline, a significant number of nodes on the primary and secondary can experience voltages beyond the desirable voltage range for prolonged periods of time. However, these voltage exceedances significantly dropped when the smart inverter settings were enabled. While there were minor differences in the resulting voltage improvement among the different smart inverter settings, generally superior voltage regulation was experienced when the PV smart inverter settings were enabled compared to when they were disabled.

The volt-var curve slope is a key parameter influencing the voltage improvement. The Hockey Stick 2 curve setting that has the most aggressive curve slope in the high voltage region resulted in the lowest VE hours per node on the secondary where the PV systems are installed. As the PV penetration is very high (70%) on this feeder, voltage rise during the peak PV generation was the major voltage regulation issue on this feeder. The HS2 curve, with the highest curve slope, forced the PV smart inverters to absorb the reactive power more intensely compared to the other smart inverter settings and resulted in better voltage regulation in such cases.

Table 13. Summary of voltage exceedances

PV Smart Inverter Setting	Secondary		Primary	
	VE hours per node	Number of VEN	VE hours per node	Number of VEN
Baseline	23.52	752	42.83	481
CA 21	0.55	16	0.61	0
HI 14	0.21	9	0.76	12
IEEE 1547	0.47	28	0.96	14
No Deadband	1.05	37	2.84	42
HS1	0.55	16	0.61	0
HS2	0.09	3	0.91	12
Volt-Var-Watt	4.45	110	2.95	53

The voltage regulation device operations per day are summarized in Table 14. The legacy device operation counts did not vary significantly with different PV smart inverter settings, except in the case of volt-var-watt setting. However, the volt-var-watt setting resulted in the highest number of voltage exceedances as shown in Table 14 compared to the other cases when the smart inverter settings were enabled.

Table 14. Summary of legacy voltage regulation device operations

PV Smart Inverter Setting	LTC tap changes per day	Capacitor bank status changes per day
Baseline	12.07	5.25
CA 21	12.47	6.04
HI 14	13.75	5.03
IEEE 1547	13.81	5.42
No Deadband	13.03	5.66
HS-no compensation	12.47	6.04
HS-deeper Q	14.07	5.83
Volt-Var-Watt	7.32	1.47

8.1.2 Utility Planning Network Model Anomaly Detection Key Findings

Distribution planning and operations rely on accurate and reliable network models. There is a high degree of uncertainty involved in the distribution network modeling due to the changes in the network and loading conditions, lack of sufficient data, assumptions, and incorrect data, which impact the quality of network models. Further, traditional practice ignores modeling of secondary networks. As DER penetration levels increase, improving network model quality, including secondary becomes important. With the installation of AMI, more sensor measurement data is available which can be used for improving distribution network models and keeping them current. The utility planning network model anomaly detection tool developed in this project is a key step towards realizing this goal.

The utility planning network model anomaly detection tool used both physics-based and machine learning-based methods. The physics-based method estimates the secondary network parameters using limited AMI measurement data. It then estimated the primary network voltages where the sensor data were not available. The machine learning-based method used the estimated primary voltages to build machine learning models for each service transformer secondary which can be used for subsequent estimations of the primary network voltages. Below are the key findings from this task:

Physics-based Method

The physics-based method used constrained optimization to estimate the secondary network parameters based on the AMI measurement data [20]. Existing approaches for this application assume that secondary topology is known, or the AMI data for all the customers on a given secondary is available. The fact that these two assumptions do not hold true introduced some initial difficulties in the algorithm configuration. Further, it was found that the AMI data used in this project do not include any information such as reactive power or the voltage angle needed to estimate the reactance of the secondary lines. It was found that the primary voltage estimates from the physics-based algorithm were consistently higher than the actual primary voltages from the simulations, likely due to the approximations made for the data that was unavailable.

Machine Learning-based Method

The machine learning-based method used random forest regression to build the equivalent machine learning models for each service transformer secondary [21]. While the primary voltage estimates from the machine learning models were fairly accurate, building these models requires primary voltage data as training data which is typically not available. Therefore, the primary voltage data from the physics-based method was used as the training data for building the machine learning models. It was observed that the voltage estimates from the combined method that uses both the physics-based and machine learning-based methods accurately matched the primary voltages obtained from the simulations.

Lack of Primary Sensor Data

The primary network voltage data from selected primary buses was supposed to be collected according to the original plan. However, the primary meters were not installed due to some practical constraints. The tool validation process relied more on the simulations due to the lack of primary measurement data from the field.

8.1.3 Phase Identification Key Findings

The phase identification algorithm identified and configured in this study focused on achieving high phase identification accuracy levels on two distribution feeders. A few key parameters were influential in the selection of the algorithm and the accuracy of the phase identification results. These parameters are discussed in this section.

Circuit Connectivity

A distribution feeder can have phase-to-neutral connections AN, BN, and CN (also known or referred to as A, B, and C), and/or phase-to-phase connections, AB, BC, and CA. They also consist of single-phase and two-phase branches and connections. Some of the existing phase identification algorithms, specifically those based on the correlation and linear regression analysis, work only for the feeders having phase-to-neutral connections. When a feeder has a mix of phase-to-neutral and phase-to-phase connections, the voltage dependencies among the phases can be a challenge for the phase identification algorithm. For example, some level of voltage dynamics occurring on phase A can be seen by the meters connected to AB. When these interdependencies are high, the algorithm can incorrectly represent the phase connectivity.

The phase connectivity details of the two feeders in this study and the associated phase identification results are summarized in Table 15. The numbers represent the AMI meter counts associated with a given type of phase connectivity. The first feeder has high meter counts for all six phase connections. The second feeder has all six phase connections, most of the meters have phase-to-neutral connection. In all instances the phase identification results show high accuracy.

Table 15. Phase connectivity details and phase identification results of two feeders

Feeder		Phase Connectivity						Total	Accuracy
		A	B	C	AB	BC	CA		
Feeder 1	Ground truth	63	102	99	77	136	91	568	90.5%
	Phase identification result	55	98	98	56	126	81	514	
Feeder 2	Ground truth	268	310	251	17	1	10	857	94.4%
	Phase identification result	260	293	241	12	0	3	809	

Data Availability

The type of data available significantly influences the type of phase identification algorithm that is chosen for use. Some phase identification algorithms based on correlation analysis require voltage measurements from the substation SCADA. Specifically, phase-to-neutral and/or phase-to-phase voltage measurements at the feeder head are required depending on the meter connections present in the feeder. However, typically utilities record either phase-to-neutral or phase-to-phase voltage only at the substation. If only phase-to-phase voltage is collected for a feeder having a mix of six phase connections, these algorithms cannot be applied. In this project, only the phase-to-phase voltage measurement data at the feeder head was available in the substation SCADA data, while the studied feeders have all the six

possible combinations of phase connections. As such, the correlation-based algorithms that depend on voltage measurement data from the substation SCADA could not be applied in this case.

The electrical quantities that the AMI meters measure also influence the choice of algorithm. Some regression-based algorithms require availability of both voltage magnitude and power consumption data from all the AMI meters connected to the service transformers. These algorithms cannot be applied if all customers do not have AMI meters because the voltage data from some customers would be missing.

PV Penetration Level

Most of the existing phase identification algorithms use voltage time series data. In particular, the voltage variations during a selected time period are the basis for determining the phase connectivity for these algorithms. Since the voltages are significantly influenced by the PV power generation, these algorithms may fail to identify the phase connectivity accurately. While this may not be an issue for some feeders having low PV penetration levels, this parameter should be considered for the feeders with high PV penetration.

Both the feeders used in this project have significant levels of PV generation. The PV penetration levels are 70% and 24% relative to the peak load for these feeders. The phase identification algorithm configured and tailored for the application in this project worked well for both these feeders, signifying its robustness to the PV penetration levels.

GIS Data Quality

It is a common practice to provide the known information about the distribution network to the phase identification algorithms to accomplish high accuracy levels. Specifically, the expected number of phase connections is supplied to the clustering algorithms as an input and the network topology information is used as constraints to obtain better phase mapping. However, the phase connectivity information in the GIS can sometimes be significantly inaccurate. More than 50% of the AMI meters were observed to be on different phases in the field validated data as compared to the GIS database in this project. When such high phase connectivity inaccuracies exist in the GIS, using such data as an input can result in poor phase identification results.

Bad Data

AMI measurement data is not perfect. In addition to the standard measurement errors, some AMI meters can report completely unreasonable data. The bad data should be identified and removed from the phase identification process. Otherwise, the bad data can lower the phase identification accuracy depending on the algorithm. Some phase identification algorithms, specifically those based on clustering, can identify the bad data without requiring additional processing.

Training Data

The type of algorithm used depends on many parameters including the type of phase connections present, data availability, and DER penetration levels as mentioned earlier. While unsupervised learning techniques are known to provide good phase identification results, such performance requires specific feeder characteristics. When the phase identification algorithm must work well for a wide variety of feeders, supervised learning is a better option. The training data provided to the supervised learning

algorithm can supply information specific to each feeder to make it robust. As a result, the supervised learning algorithms can provide consistently high accuracy levels in the phase identification for many feeders. Note that obtaining the required training data may involve some level of field verification.

8.1.4 AMI Data-Centric Distribution System Operations:

Model-free controls

The demonstration of a data-driven method for voltage control for a feeder with high PV penetration with data from less than 10% of the meters on the feeder, proved feasible and practicable as a method for future adoption. The model-free controls also demonstrate that these methods can perform without an impedance model, and hence are more resilient to model quality errors including missing feeder data.

Pervasive Secondary Monitoring

This use case demonstrated how even if the voltages on the primary network are within the ANSI voltage limits (+/- 5%), the secondary networks might see voltages that are beyond these desired limits due to the presence of PV and the lack of accurate models to represent the secondary networks. By using AMI voltages as input for voltage control, these methods can maintain desired voltage on the secondary networks as well, thereby enabling pervasive secondary network monitoring.

Uncertainties from PV output

The method demonstrated in this use case does not require explicit modeling of PV systems or their outputs, but only the voltages as measured by the AMI meters. This obviates the need for detailed PV modeling for this voltage control algorithm, making it suitable for feeders with presence of high PV adoptions.

8.1.5 Meter-to-Transformer Mapping

In this exploratory task, the project team identified some key methods for mapping AMI meters to the correct service transformer. The project team identified recall and precision as two metrics for validating the performance of these methods. The Spearman correlation coefficient [13] had the highest score in both recall and precision. Therefore, it fits the proposed meter-to-transformer methodology best. All methods detected 97% of the incorrect records.

8.2 Updated Value Proposition

In this project, tools to benefit utilities in upgrading the distribution planning and operation practices were demonstrated. Specifically, the utility planning network model anomaly detection tool helps estimate the primary voltage based on secondary AMI data. The estimated primary voltage provides visibility of the primary network where the physical voltage sensors do not exist. Further, the estimated primary voltage based on AMI data was used to identify inaccuracies in the planning network models. The demonstration of the phase identification tool indicated that customer phase connectivity can be identified based on AMI data. Additionally, numerous simulations were performed using distribution feeder models developed based on AMI data from the field to develop insights into PV smart inverter settings.

The tools and demonstrations performed in this project promote and support the initial benefits and value proposition of greater reliability, lower costs, and increased safety as described in Section 6.0. Further, this project promotes the following additional benefits:

Societal Benefits

This project helps correct the distribution network models used for planning and operations. The improved network models lead to more efficient control decisions, and reduction in losses and outages. The smart inverter study lowers the cost of network operations by deferring the investments to maintain the desired electric service quality. Overall, this project shows potential to reduce the overall cost of electric service to customers.

Greenhouse Gas (GHG) Emissions Reduction

The efficient PV smart inverter settings may increase the PV hosting capacity of the distribution feeders, allowing more renewable generation. Improved phase balancing, volt/var optimization and other network controls based on the accurate network models further support the higher levels of renewable generation. The tools developed in this project are collectively geared toward reducing the dependence on traditional fossil fuels for our energy needs, thus lowering the associated GHG emissions. As noted in the initial benefits analysis, reducing travel associated with field visits/field verification, would also reduce GHG emissions in commercial use of the demonstrated tools.

Economic Development

This project demonstrated that the AMI data can be used for phase identification and distribution network model improvements. By developing use cases for the AMI, this project promotes the deployment of AMI and the associated communication infrastructure. Thus, it influences the market for the development of advanced sensing and communication capabilities. Further, the PV smart inverter study enables higher levels of DER adoption through efficient voltage control. With the higher levels of DER and availability of communication networks the DERs can participate in grid services with suitable incentives to customers. This project paves the way for future AMI-based distribution network operations that promote efficient use of renewable energy, where both the utility and customers economically benefit from the services provided.

Efficient Use of Ratepayer Funds

This project developed a phase identification tool that performs automated customer phase mapping based on AMI data. Traditionally, utilities perform the same task manually by sending a crew to the field for the identification of the customer phasing. This manual process, performed periodically, is expensive and time-consuming. The phase identification tool greatly simplifies this process and determines the customer phase connectivity more economically, thus providing savings for the ratepayers. The PV smart inverter results may help achieve the desired power quality using existing smart inverters without having to invest in network upgrades. This approach has the potential to reduce the cost of electric service which in turn benefits ratepayers through lower electricity bills.

9.0 Conclusions

This project demonstrated a utility planning network model anomaly detection tool, a phase identification tool, a meter to transformer mapping algorithm proof-of-concept and analyzed the impacts of PV smart inverter settings. The tools used AMI measurement data to estimate the primary voltages, identify planning model inaccuracies, and automate phase mapping. With the tool's promising results, the SDG&E team is currently examining deployment opportunities.

Several issues in the GIS data planning network models, and the AMI measurement data were detected during the algorithm development. These issues are related to the incorrect phase connectivity information; approximated load and PV profiles used in the planning network models; lack of reactive power measurements in the AMI data; and desirable voltages from SCADA. Additionally, using the AMI data from only two AMI meters per transformer and the lack of primary voltage measurement data from the field, created challenges in the algorithm development and validation. The selected feeders have significantly different characteristics in terms of phase connectivity, high PV penetration levels, and presence of underground cables. This provided an opportunity to develop algorithms that are sufficiently robust to the feeder characteristics. In the PV penetration study, it was observed that enabling the PV smart inverter settings is desirable for improving the network voltage profile. The selection of feeders with different characteristics also highlighted the disconnect between the voltage dynamics on the primary and secondary networks.

As the AMI deployments increase and more measurement data becomes available, leveraging the AMI data for the distribution system planning and operations is desirable. Relying on the automated AMI data analytics for this purpose instead of the network models is more economical and efficient as it reduces the time and effort in the manual periodic field verifications and database updates. This project's results indicate the desirability of shifting from the traditional model-based grid operations to data-driven grid operations.

10.0 Transfer Plan

10.1 Project Results Dissemination

This report is the primary documentation of this project work. It will be posted on the SDG&E's EPIC public website and filed with the California Public Utilities Commission.

The results from the project were presented in multiple conferences through peer-reviewed technical papers. The following is the list of papers developed under this project:

Published Papers:

1) J. Wang, H. Padullaparti, S. Veda, M. Baggu, M. Symko-Davies, A. Salmani, and T. Bialek, "A Machine Learning-based Method to Estimate Transformer Primary-Side Voltages with Limited Customer-Side AMI Measurements," in *IEEE Power & Energy Society General Meeting (PESGM)*, 2021.

2) M. Netto, J. Hao, H. Padullaparti, and V. Krishnan, "On the Use of Smart Meter Data to Estimate the Voltage Magnitude on the Primary Side of Distribution Service Transformers," in *IEEE Power & Energy Society General Meeting (PESGM)*, 2021.
3) H. Padullaparti, S. Veda, S. Dhulipala, M. Baggu, T. Bialek and M. Symko-Davies, "Considerations for AMI-Based Operations for Distribution Feeders," in *IEEE Power & Energy Society General Meeting (PESGM)*, Atlanta, GA, USA, 2019, pp. 1-5.

The project results were also disseminated to the industry through NREL's ADMS Test Bed and DERMS Applications Industry Advisory Board (IAB) through quarterly meetings, webinars and a workshop held in October 2020. The following are the meetings where the project results were discussed.

- NREL hosted two virtual workshops that were held on November 9th and 10th, 2020. The title of the event was Advanced Distribution Management System Test Bed and Architectures for Grid-Edge Management Workshops. The event brought over 70 external participants from 45 organizations. This was the best attended ADMS workshop to date.
- NREL hosted a virtual IAB meeting on April 29, 2021. The focus was on Peak Load Management (PLM) and AMI for operations use cases. We also launched a series of webinars beginning with the Peak Load Management use case in June 2021.
- NREL hosted a virtual IAB meeting on July 22, 2021. We focused on the AMI-based, data-centric grid operations and ADMS network model quality impact on VVO use cases. We also presented the process for identifying future use cases, including an overview of the RFI.
- NREL continued the series of webinars with a presentation by Dr. Santosh Veda focused on AMI data data-centric planning and operations that showcased results from this project on October 6, 2021. This webinar had over 75 participants.

10.2 Transition for Commercial Use

The demonstrations performed in this project were supported by the Electric Program Investment Charge, a public purpose program funded by ratepayers of California's investor-owned utilities, and by the US Department of Energy funding to NREL. These tools were then tested, validated, and demonstrated on real-world utility feeders provided by SDG&E. As discussed in Section 6, some of these tools were further validated using field verification. These tools are being examined for opportunities for deployment on SDG&E's AMI data system. The architecture for implementing these tools is shown in Figure 39 below.

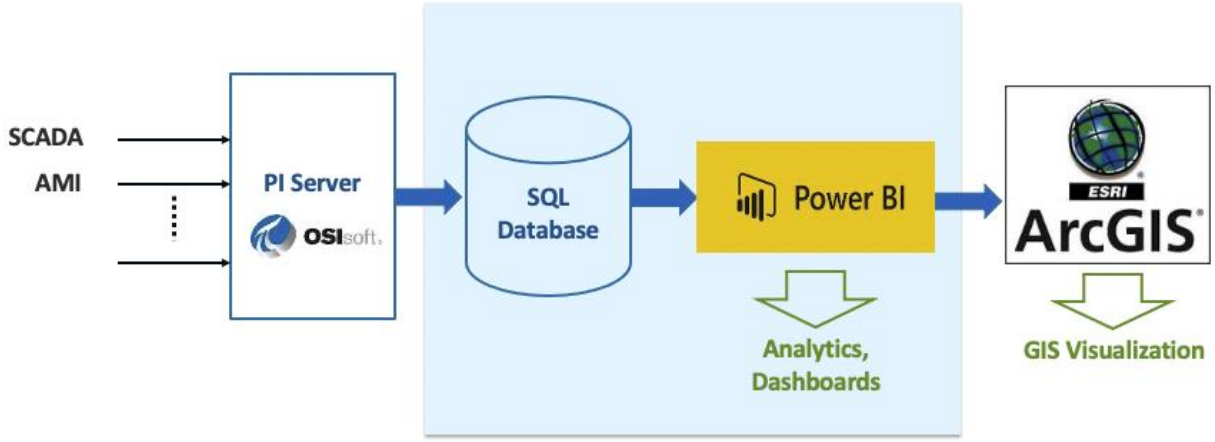

Figure 39. Architecture for potential deployment on SDG&E AMI data collection system

In addition, a subset of these tools will also be released as open-source tools for use by the industry and research for further improvement by the community at large.

11.0 Recommendations

The use cases presented in this project demonstrate the accuracy, feasibility, and rationality of using AMI data for greatly improving the planning and operations activities in the near-term, especially for feeders with high levels of PV adoption. It is recommended that specific tools (Utility Planning Network Model Anomaly Detection Tool, AMI Meter-to-Transformer Mapping, and Phase Identification Using AMI Data) be applied by the SDG&E team for other feeders. The evaluation of data-driven controls using realistic emulation capabilities of the ADMS Test bed provides a feasible demonstration for real-time data-driven control of high-PV feeders for consideration and implementation in the medium-term. Such an approach could reduce the reliance on planning models and make the operations resilient to the ubiquitous problem of poor model quality.

SDG&E will need to identify a stakeholder group within the company to lead this commercial adoption process.

12.0 References

Number	Reference Description
1	B. Uzum, A. Onen, H.M. Hasanien, and S.M. Muyeen. "Rooftop Solar PV Penetration Impacts on Distribution Network and Further Growth Factors—A Comprehensive Review," in *Electronics,* vol. 10, no. 1, pp. 55, 2021.
2	X. Zhu, J. Wang, N. Lu, N. Samaan, R. Huang, and X. Ke, (2018). A hierarchical vlsm-based demand response strategy for coordinative voltage control between transmission and distribution systems, in *IEEE Transactions on Smart Grid* vol. 10, no. 5, pp. 48338-4847, 2021.
3	J. Wang, X. Zhu, D. Lubkeman, N. Lu and N. Samaan, "Continuation power flow analysis for PV integration studies at distribution feeders,"*2017 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT)*, pp. 1-5, 2017.
4	Wang, Y. Dong, L. Wu and B. Yan, "Interval Overvoltage Risk Based PV Hosting Capacity Evaluation Considering PV and Load Uncertainties," in *IEEE Transactions on Smart Grid*, vol. 11, no. 3, pp. 2709-2721, 2020.
5	H. V. Padullaparti, Q. Nguyen and S. Santoso, "Advances in volt-var control approaches in utility distribution systems," *2016 IEEE Power and Energy Society General Meeting (PESGM)*, 2016, pp. 1-5.
6	Rule 21 Interconnection, California Public Utilities Commission, CA, available at https://www.cpuc.ca.gov/Rule21/
7	Rule no. 14, Service Connections and Facilities on Customer's Premises, Hawaiian Electric, available at https://www.hawaiianelectric.com/documents/billing_ and_payment/rates/hawaiian_electric_rules/14.pdf.
8	IEEE Standard for Interconnection and Interoperability of Distributed Energy Resources with Associated Electric Power Systems Interfaces, IEEE Std 1547-2018, April 2018.
9	Y. Xie, M. Sengupta, Y. Liu, H. Long and A. Habte, "Progress on the National Solar Radiation Data Base (NSRDB): A new DNI computation," *2020 47th IEEE Photovoltaic Specialists Conference (PVSC)*, pp. 0330-0332, 2020.
10	A. Liaw, and M. Wiener. "Classification and regression by random forest. "*R news*" 2.3 (2002): 18-22.
11	Y. Freund and RE. Schapire. "Experiments with a new boosting algorithm."*icml".* Vol. 96. 1996.

Number	Reference Description
12	L. Breiman. “Arcing the edge.”, *Technical Report 486*, Statistics Department, University of California at Berkeley, 1997.
13	Benesty, Jacob, et al. "Pearson correlation coefficient. "Noise reduction in speech processing”. Springer, Berlin, Heidelberg, 2009. 1-4.
14	McLeod, A. Ian. "Kendall rank correlation and Mann-Kendall trend test. "R Package” Kendall (2005).
15	Mukaka, Mavuto M. "A guide to appropriate use of correlation coefficient in medical research." Malawi medical journal 24.3 (2012): 69-71.
16	H. Padullaparti, C. Pisitpol, and S. Surya. "Grid-Edge Voltage Control in Utility Distribution Circuits.” Power Engineering: Advances and Challenges Part B: Electrical Power, 251, 2016.
17	H. Padullaparti, S. Veda, S. Dhulipala, M. Baggu, T. Bialek and M. Symko-Davies, "Considerations for AMI-Based Operations for Distribution Feeders", 2019 IEEE Power & Energy Society General Meeting (PESGM), 2019, pp. 1-5, doi: 10.1109/PESGM40551.2019.8973888.
18	G. G. Moisen, “Classification and regression trees,” Encyclopedia Ecology, vol. 1, pp. 582–588, 2008.
19	Y. Ren, L. Zhang, and P. N. Suganthan, “Ensemble classification and regression-recent developments, applications and future directions [review article],” IEEE Comput. Intell. Mag., vol. 11, no. 1, pp. 41–53, Feb. 2016.
20	M. Netto, J. Hao, H. Padullaparti, and V. Krishnan, “On the Use of Smart Meter Data to Estimate the Voltage Magnitude on the Primary Side of Distribution Service Transformers,” presented at the IEEE Power and Energy Society General Meeting, 2021.
21	J. Wang et al., “A Machine Learning-based Method to Estimate Transformer Primary-Side Voltages with Limited Customer-Side AMI Measurements,” presented at the IEEE Power and Energy Society General Meeting, 2021.