ABSTRACT

ZAMANI FARAHANI, VAHRAZ. State Estimation for Volt/VAR Control on Active Power Distribution Systems. (Under the direction of Dr. Mesut E. Baran).

This dissertation focuses on real time monitoring of Power Distribution Systems, to improve the efficiency and reliability of the distribution systems operation. First of all, the requirements or implementing the proper real time monitoring for the distribution system were reviewed. Then, the necessity of the state estimator for Active Distribution System Monitoring and Control is investigated. Main requirement for distribution systems operation is the voltage regulation. Various Volt/VAR Control (VVC) schemes are developed to provide the proper voltage profile along the feeders with presence of Photovoltaic (PV) panels. Implementing VVC schemes need to have accurate voltage estimation along the feeder at distribution level. Here, the SE has been deployed to find the operating condition of the feeder. It was shown that effective management and control of distribution systems required estimating the state of the system properly by State Estimation (SE). Implementing SE on distribution systems had its own challenges which were thoroughly discussed. Afterwards, two main approaches for Distribution State Estimation (DSE) have been discussed. Finally, Branch Current-based State Estimation (BCSE) has been introduced in details for further studies. The performance of BCSE method in the presence of measurement noises through Monte Carlo simulations due to the statistical nature of pseudo measurements has been assessed. One load estimation method based on the availability of AMI data has been implemented and tested by actual load data. The goal of this method is to improve the load estimation of the distribution transformers for distribution automation applications, such as VVC.

The Meter Placement problem on the distribution feeders is addressed in Chapter 5. First, available literature regarding this problem is reviewed, after which the problem is formulated for VVC application, especially CVR application. The heuristic approaches have been proposed, based on the extensive studies and sensitivity analyses, to obtain guidelines for placing the different types of measurements on the system. The proposed schemes have two stages. First stage starts by placing the meters by the rules. At the second stage, an efficient search scheme has been implemented to find the minimal set of measurement for VVC. These schemes are flexible in that it allows incorporation of different metering options and robustness measures. Finally, the performances of these methods are assessed on the IEEE prototype test feeder.

The last chapter deals with the concept of the robust state estimation, by detecting and identifying the topology changes on the distribution feeders. BCSE relies on the basic assumption that the topology of the system is given correctly. However, in most real world situations, the state of some switching devices is unknown, especially in the case of blown fuses. Two common topology errors taking place at the distribution feeder are explained. Following to that, the proposed algorithm, which is based on the analysis of the sum of the weighted residual for each zone of the switches and fuses, has been applied to detect and identify these topology changes. In addition, impact of topology changes on voltage estimation for VVC has been investigated.

State Estimation for Volt/VAR Control on Active Power Distribution Systems

by
Vahraz Zamani Farahani

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Electrical Engineering

Raleigh, North Carolina

2014

APPROVED BY:

Dr. Mesut E. Baran Committee Chair	Dr. Subhashish Bhattacharya
Dr. Aranya Chakrabortty	Dr. Sujit Ghosh
Dr. David Lubkeman	

DEDICATION

TO MY PARENTS
and
MY FAMILY

BIOGRAPHY

The author, Vahraz Zamani Farahani was born in Tehran, Iran. He received B.Sc. in Electrical Engineering from Sharif University of Technology in 2004, and M.Sc. in Electrical Engineering specialized in Power Systems from Iran University of Science and Technology in 2007. He worked as an Electrical Engineer in Mapna Group for almost three years. Then, he started to pursue the Ph.D. degree in North Carolina State University at FREEDM System Center in 2009. He received his Ph.D. degree from NC State University in spring 2014. His research interests include power system analysis and control, distribution automation, distribution system state estimation, statistical solutions for power and energy systems, and volt/var control.

ACKNOWLEDGMENTS

First and foremost, I would like to express my sincere appreciation to my advisor, Professor Mesut E. Baran, for his persistent guidance as well as teaching how to conduct the practical research. Prof. Baran's deep knowledge in power systems and his unique approach to solve the problems have been very inspirational to my research and development at NC State. I hope to be privileged to benefit from his mentorship and collaboration throughout my future career.

I am grateful to my committee members, Professor Sujit K. Ghosh from Statistics Department, Dr. Subhashish Bhattacharya, and Dr. Aranya Chakrabortty for their valuable suggestions and helps. I would also like to thank Professors Iqbal Husain and Alex Huang, by providing the great research environment at Future Renewable Electric Energy Delivery and Management (FREEDM) Systems Center to conduct my research.

It has been a great pleasure to work with EPS group, Dr. Hossein Hooshyar, Dr. Zhan Shen, Mr. Urvir Singh, Mr. Moyeen Kazi, Mr. Sanujit Sahoo, Mrs. Yue Shi, Mr. Kyle Barth, Mr. Bharadwaj Vasudevan, and Mr. Travis Tippens at the FREEDM Systems Center.

During my graduate studies, I had an opportunity to work with the Varentec, Inc. I would like to thank my mentors for their support, Professor Deepak Divan, Mr. Mehrdad Hamedani, Mr. Andrew Dillon, and Dr. Soumen Ghosh.

I appreciate the assistance from the staff members of the FREEDM Systems Center and the graduate office of the ECE department, including Dr. Pam Carpenter, Dr. Michael

Devetsikiotis, Mr. Rogelio Sullivan, Ms. Karen Autry, Ms. Colleen Reid, Ms. Cailan Mang, and Ms. Elaine Hardin.

I would like to thank my colleagues and friends at NC State University, especially in FREEDM Systems Dr. Babak Parkhideh, Mr. Hesam Mirzaee, Dr. Saman Babaei, Mr. Behzad Nabavi, Dr. I. Safak Bayram, Dr. Habib Rahimi, Mr. Sina Parhizi, Dr. Gangyao Wang, Mr. Daniel Fregosi, Mr. Sumit Dutta, Mr. Arun Kadavelugu, Dr. Edward van Brunt, Mrs. Ghazal Fallahi, Mr. Mohammad Ali Rezaei, Mr. Maziar Vanouni, Mr. Nima Yousefpoor, Mr. Tom Nuddel, Mr. Anken De, and Mr. Mohammad Etemadrezaei.

My deepest appreciation goes toward my parents and family, father and mother: Mr. Mojtaba Zamani Farahani and Mrs. Vila Gharagouzlou, my dear sisters: Mrs. Vishpar Zamani, and Mrs. Roshanak Zamani, who have always encouraged and supported me to pursue my goals.

Last but not least, I would like to thank my partner and friend for her endless support and concern. Thank you, Marwa El-Sayed!

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1: Introduction

1-1 Background

To improve efficiency and reliability of the operation of the power distribution systems, implementing an effective real-time monitoring structure is needed to manage and control the distribution systems effectively. Literature survey indicates that utilities have been improving their means of monitoring their distribution systems to advance service reliability as well as energy efficiency thorough their Smart Grid projects [1-6, 10, 12, 131-133]. More recent energy efficiency solutions on the distribution systems are based on the optimization of the voltage profiles to control.

Effective management of distribution systems requires analysis tools that can estimate the state of the system (the operating condition) and predict the response of the system to changing load and weather conditions. The main tool used for system analysis is power flow analysis. But this tool is not very suitable for real-time monitoring as it requires accurate load and system data. Real-time monitoring of the distribution networks has its own challenges, due to radial topology, three-phase unbalanced system, high resistance to reactance ratio, limited real-time measurements, integrating the distributed and renewable energy resources, keep changing topology, and etc. of the distribution systems

Some utilities have deployed large scale Advanced Metering Infrastructures (AMI) for enabling the two-way communication between utilities and customers [57, 58]. Smart meters of AMI can provide more accurate data about the load conditions and load patterns to operate the distribution systems more efficiently. Additionally, in order to monitor the system operating conditions, some utilities have begun installation of limited Supervisory Control And Data Acquisition (SCADA) systems at the distribution level. These two approaches provide the opportunity to deploy available information to find the state of the distribution network in a more accurate and reliable manner to effectively manage the distribution networks.

To estimate the state of the system at the distribution level, different approaches have been proposed in the literature. One of the approaches is power flow based [17, 19, 20] and the

others [25-32] are extensions of the conventional state estimation (SE) method for three-phase analysis. SE based solutions are preferred over the power flow approach by considering the uncertainty of the input data and the system, despite of its computational complexity. In this dissertation, the Branch Current based State Estimation (BCSE) method [30] is considered for further studies, it is computationally more efficient and tailored for distribution system characteristics, such as unbalancing in loading and the structure. Furthermore, some statistical techniques are applied for assessing the BCSE performance in this study.

One of the main inputs of SE is the estimated of distribution transformer loads. Previous studies show that the quality of state estimation depends on the load estimation accuracy. Normally, historical data of the customers are used to estimate the loads. By adopting more AMI implementation on the distribution networks, AMI can then provide more up-to-data information about customer loads. To address this challenge, one load estimation method has been implemented to utilize the near real-time data from AMI for enhancing the accuracy of load estimation.

Voltage regulation is a fundamental operating requirement for operation of the distribution systems. Voltage profile must be maintained under all operating conditions of the system at the range of 114 V to 126 V, in 120 V scale. Different Volt/VAR Control (VVC) schemes have been developed for the distribution networks using capacitor banks and voltage regulators. Recently, Conservation Voltage Reduction (CVR) has become one of the most demanded VVC schemes to improve the system efficiency by lowering the voltage levels along the feeder. In order to reach the goals of the CVR program, the voltage profile should be estimated within the accuracy of +/- 1 V in scale of 120 V. In this study, the BCSE has been adopted to estimate the voltage within the desired accuracy. It is shown in this dissertation that using load estimation only cannot provide enough data to estimate the voltage with this requirement. To improve the performance of the SE, some field real-time measurements are added to the measurement set of the SE. Therefore, we need to acquire the proper approaches to find the best locations for these real-time measurements along the feeder as well as their number. This problem has been addressed thoroughly in this dissertation by considering different operational conditions of the distribution feeder. Initially, the proposed schemes placed the current and

voltage measurements on the feeder by following a set of rules. In the second stage, one efficient sorting procedure is used to select the minimal set of meters. To amend the final set of the measurement, some practical concerns, due to the operation of the distribution systems and measurement loss, are considered.

It is usually assumed that the SE is running in a way that the topology of the system is given without any doubts [17-32]. However, in most of real world conditions, as mentioned before, the state of some switching devices is unknown, for some reasons, or one un-monitored devices, like fuses, are operated. In this case, the current status of the switches plus the given topology are not reliable to run the SE. It is common in the distribution systems that some switches are not monitored by SCADA, and their open/close position in the database of the topology is manually updated by the system operators. The result of these conditions is a topology error in the network. Model topology errors can also occur when the telemetered circuit breaker ON/OFF status is incorrect. Correct network model is crucial to estimate the operating conditions of a system. That is the reason why topology error identification and detection on the distribution feeders by using BCSE method were investigated in this dissertation.

1-2 Dissertation Objective

The main objective of the dissertation at hand is to propose the algorithms for real-time monitoring of distribution systems. The following materials are presented and the main goals of each chapter are explained in the following section.

- Challenges for real-time monitoring of the distribution systems
- Assessment of the BCSE method
- Load estimation of distribution transformers by AMI data
- Meter placement for Volt/VAR Control on distribution networks, and
- Topology changes detection and identification by BCSE.

1-3 Dissertation Outline

- **Chapter 2**

Basic concepts of the voltage regulation on the distribution system are explained in this chapter. In addition, impacts of distributed energy resources, especially PV solar panels, on the voltage profile of the system are reviewed. As a result, the motivation for adopting SE for monitoring and control of the distribution systems are studied.
Special characteristics of the power distribution systems in comparison to power transmission systems are mentioned and listed, in order to be considered in the implementation of the effective real-time monitoring solutions for distribution systems.

- **Chapter 3**

After presenting the needs for the SE on distribution systems, available methods in literature are considered. Afterwards, the Branch Current based State Estimation method is chosen for further studies. The performance of the BCSE is assessed through statistical measures. Some statistical measures quantified in terms of bias, consistency, and quality are adopted for assessment in this chapter by adopting the statistical tests. For statistical analysis, 10,000 Monte Carlo simulations are performed.

Due to the application of the BCSE for voltage profile estimation, voltage standard deviation calculation has been considered. Here, standard deviation is a measure for quality of the voltage estimation from SE. At the end, the voltage standard deviations from output of BCSE is formulated and calculated. The calculated voltage standard deviations are verified by Monte Carlo simulations.

- **Chapter 4**

As indicated in the previous chapter, the accuracy of load estimation has a significant impact on the performance of the SE. In this chapter, the load estimation approach is implemented

using the available AMI. This method has two stages. In the initial stage, the given customers are categorized by k-means clustering approach to figure out similar load patterns. First, harmonic-based time series models are developed using historical load data from AMI. When new load data is available from each cluster, the load estimation model can be modified by regression.

This chapter presents a method to improve the historical-based model of load estimation using real time load data from AMI. This method has been tested with actual load data. In addition to present the model for load estimation, impact of the load estimation error on the performance of the BCSE has been investigated too.

- **Chapter 5**

To estimate the voltage profile accurately, more real-time measurements are needed to be placed on the feeder for advanced VVC. This chapter addresses this challenge by proposing solutions based on defining the meter placement problem. Sensitivity analyses are performed for different conditions and situations. Based on these studies, a set of guidelines is developed to determine an initial set of measurements which are reasonably small and are redundant enough to provide the desired level of accuracy. At the second stage, a meter sorting procedure is used to identify the minimal set of meters. Three approaches namely: mixed, low cost, and robust meter placement schemes. The performance of the methods is tested using a prototype distribution feeder.

- **Chapter 6**

The detection and identification of the topology changes on the distribution systems by Branch Current State Estimation (BCSE) program is investigated in this chapter. BCSE relies on the basic assumption that the topology of the system is given without any doubts. However, in most real world situations, the state of some switching devices is unknown, especially for the case of blown fuses. Therefore, this chapter focuses on addressing the topology identification

problem in the scope of state estimation for real-time monitoring of distribution systems for two common topology changes.

- **Chapter 7**

The overall conclusions of this dissertation at hand are presented in this chapter. In addition, the proposed future work is also outlined within the chapter.

1-4 Abbreviations

In what follows is a list of the abbreviations and their corresponding definitions presented in this dissertation.

Table (1-1): Abbreviations

Abbreviation	Definition
AMI	Automated Metering Infrastructure
ANN	Artificial Neural Network
BCSE	Branch Current State Estimation
BW	Band Width
CAPs	Capacitor Bank
CB	Circuit Breaker
CES	Community Energy Storage
CM	Current Measurement
CT	Current Transformer
CVR	Conservation Voltage Reduction
DER	Distributed Energy Resource
DG	Distributed Generation

Table (1-1): Continued

DMS	Distribution Management System
DOE	Department of Energy
DR	Demand Response
DS	Distributed Storage
DSE	Distribution State Estimation
DVC	Dynamic VAR Compensator
EPRI	Electric Power Research Institute
EV	Electrical Vehicles
FLISR	Fault Location, Isolation, and Service Restoration
GMM	Gaussian Mixture Model
IEEE	Institute of Electrical and Electronics Engineering
LAV	Least Absolute Value
LDC	Line Drop Compensation
LF	Load Flow
LS	Least Squares
LTC	Load Tap Changing
MC	Monte Carlo
NETL	National Energy Technology Laboratory
NMC	Number of Monte Carlo simulations
NP-hard	Non-deterministic Polynomial-time hard
NTP	Network Topology Processing
OMS	Outage Management System
ONR	Optimal Network Reconfiguration
OO	Ordinal Optimization
P.U.	Per Unit
PE	Power Electronic

Table (1-1): Continued

PF	Power Flow
PHEV	Plug-in Hybrid Electrical Vehicles
PM	Power Measurement
PNNL	Pacific Northwest National Laboratory
PT	Potential Transformer
PV	Photovoltaic
RPS	Renewable Portfolio Standards
RTU	Remote Terminal Unit
SCADA	Supervisory Control And Data Acquisition
SE	State Estimation
SOC	State of the Charge
SST	Solid State Transformer
Std Dev	Standard Deviation
SVM	Support Vector Machine
SW	Switch
TP	Topology Processor
VM	Voltage Measurement
VR	Voltage Regulator
VVC	Volt/VAR Control
VVO	Volt/VAR Optimization
WLS	Weighted Least Square

Chapter 2: Real-time Monitoring for Power Distribution Systems

In this chapter, the necessity of the state estimator for active distribution system monitoring and control is investigated. Utilities are required to provide voltage within a certain band for all of the customers. Hence, Volt/VAR Control (VVC) is one of the smart grid applications will be briefly described. Consequently, VVC challenges for distribution systems with Distributed Energy Resources (DER), especially Photovoltaic (PV) panels, will be reviewed. To understand the current situation of the distribution systems, we need to put all the available measurements and information into the State Estimation (SE) to obtain the system state as the input of the VVC schemes. Implementing SE on distribution systems has its own challenges which will be discussed in both “data” and “system” categorizes below.

2-1 Volt/VAR Control on Distribution Systems

2-1-1 Volt/VAR Control Overview

The Modern Grid Strategy project of the Department of Energy (DOE) and the National Energy Technology Laboratory (NETL) has identified the Volt/VAR Control as one of the main applications for smart grid [1, 2]. So, we explain VVC on distribution networks here.

Voltage regulation is a fundamental operating requirement of all electric distribution systems. The primary objective of Volt/VAR Control is maintaining voltage (120 V plus or minus 5%) within acceptable limits at the service entrance of all customers served by the feeder under all possible operating conditions. In conventional networks, the most common VVC methods are direct voltage regulation and reactive power compensation. Voltage Regulators (VRs) and Capacitor Bank (CAPs) are conventional devices used in the distribution systems. ANSI C84.1 [6] has determined the range of both service voltage and utilization voltage for utilities to keep voltages within limits. Table (2-1) clearly shows the recommended service and utilization voltage range of service and utilization in 120V scale. Once the voltage falls into range B,

corrective measures shall be undertaken within a reasonable time to improve voltages to go back to Range A [6].

Table (2-1): ANSI C84.1Voltage range for 120V [6].

	Service		Utilization	
	Min	Max	Min	Max
Range A (Normal)	-5%	5%	-8.30%	4.20%
Range B (Emergency)	-8.30%	5.80%	-11.70%	5.80%

2-1-2 Conventional VVC on Distribution Systems

In a traditional distribution system, as shown in Figure (2-1), without Volt-VAR control devices, the typical voltage profile under peak load decreases gradually along the feeder. Under heavy load conditions, the node farthest from the substation may have low-voltage violation, Figure (2-2). In order to fix this problem, we need to manually raise the source voltage, Figure (2-3), though this will cause other problems at light load, see Figure (2-4). In other words, Volt-VAR control is needed to deal with all possible normal operation conditions [7].

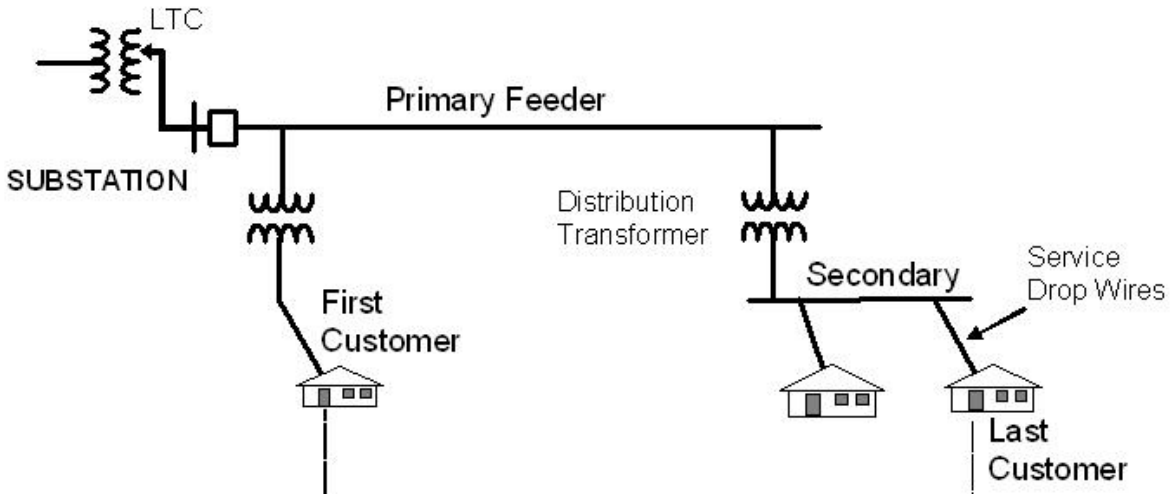


Figure (2-1): Traditional Distribution System [7].

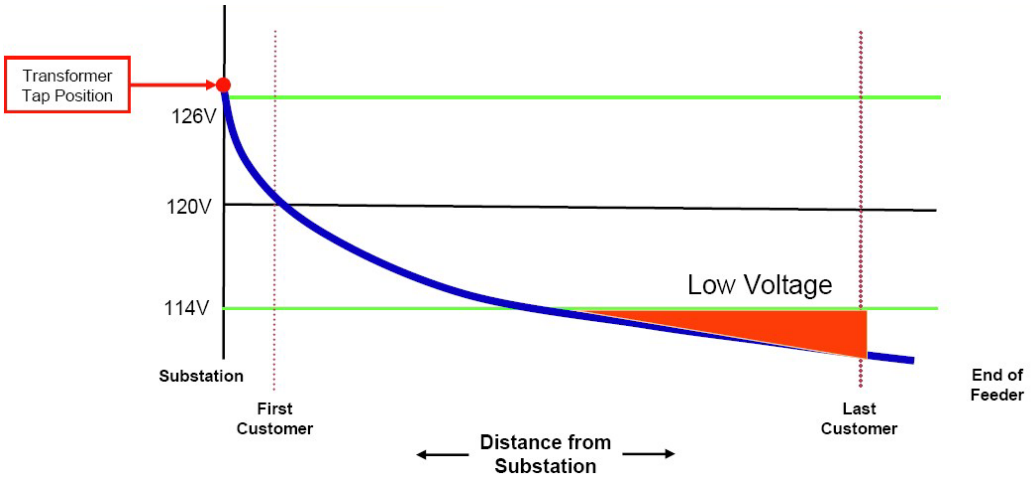


Figure (2-2): Voltage profile without Volt-VAR Control under peak load [14].

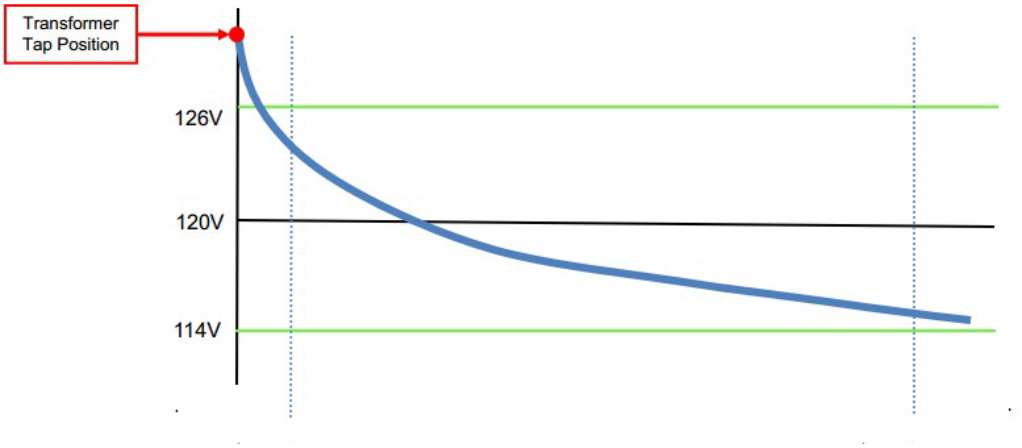


Figure (2-3): Voltage profile after raising the source voltage under peak load [14].

The most common devices used in conventional Volt-VAR control are voltage regulators (VRs) and capacity banks (CAP). On one hand, VRs adjust the voltage at the substation or along the feeder. They raise and drop the voltage profile of the whole feeder. On the other hand, CAPs are used to further raise the feeder voltage by offsetting the reactive power demand of the load when the voltage gets too low below the feeder. Conventional control schemes of both VRs and CAPs are simple local measurement-based schemes.

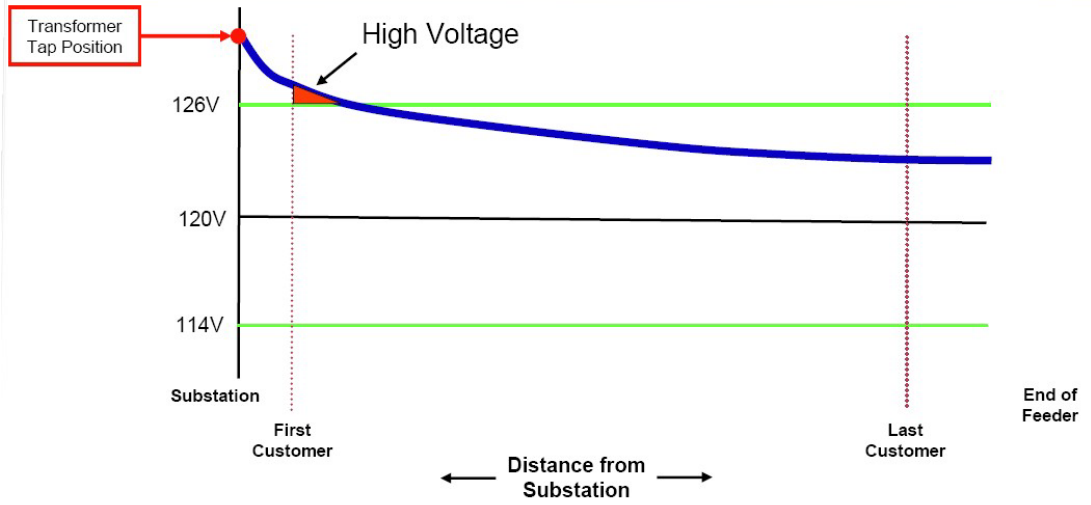


Figure (2-4): Voltage profile without Volt-VAR Control under light load after raising setting voltage [14].

Coordination of CAP and VR: One of the main challenges of local control schemes is the difficulty of coordinating the control between VRs and CAPs. With recent efforts towards extending SCADA at distribution feeder level, it is now becoming possible to coordinate the operation of these devices [8].

2-2 DERs on Distribution Systems

Recently, Renewable Portfolio Standards (RPS) has been proposed in several countries [10]. In America for instance, many states have passed RPS programs with various different targets. California's target is to reach 33% of total power generation by 2020 and North Carolina's target is 12.5% by 2021 [10].

This kind of bills incentivizes the market for solar photovoltaics (PVs), especially for grid-connected applications. Furthermore, most of the new PVs are installed in distribution network as solar panels on the roofs and working as Distributed Energy Resources (DERs). It should be noted that PVs may adversely affect the existing distribution systems. Utilities and power

system operators need to consider the potential impact of high penetration levels of PVs on traditional distribution power systems and prepare robust measures to mitigate these impacts. Next part, PV impacts on distribution networks has been reviewed.

2-2-1 PV Impacts on Distribution Systems

An increased amount of DERs may have a significant impact on a distribution system [10-14]. The main concerns identified involve reverse power flow, voltage rise, voltage unbalance, voltage fluctuation, improper VR operation, and increased power loss. In what follows, we will give a detailed description of voltage issues from [14]:

- *Voltage Rise*

PV integration can modify feeder voltage profile and raise the voltage close to the location of PVs. When high penetration level of PV is connected in a lightly loaded system, the voltage rise will be significant. If switched capacitor banks are on when the output of PV is maximal or if there are many fixed capacitor banks in the system, this voltage rise will lead to voltage violations on utility planning limits and industry standards. Voltage rise leads to high voltage violation [13].

- *Voltage Fluctuations*

PV is an intermittent resource, therefore its varied output power leads to voltage variations which may cause power quality issues and complaints from customers. The severity of these voltage fluctuations must be assed to ensure that the system will not have voltage violation under any circumstances [14].

- *Interaction with voltage-controlled capacitor banks, LTCs, and line voltage regulators*

Voltage rise and fluctuations can both cause frequent operation of LTCs, line VRs and voltage-controlled capacitor banks. At noon, PV systems provide a large amount of power into the system, which boosts the voltage letting the voltage regulators move down the tap position to maintain all of the voltage limits. When the sun sets, no power is supplied by the PV system,

the voltage regulators need to move up the tap positions to keep the voltage within limits. The higher the number of operations, the more maintenance required and the shorter the life-cycle of the equipment is. In addition, frequent operations in turn can augment voltage fluctuations and affect power quality. Furthermore, voltage fluctuations may affect the implementation of advanced Volt/VAR Control and Optimization schemes and Conservation Voltage Reduction (CVR) approaches [14].

2-3 SE for Active Power Distribution Systems

2-3-1 Purpose of State Estimation (SE)

Fred Schweppe introduced state estimation to power systems in 1970's and defined the state estimators as "a data processing algorithm for converting redundant meter readings and other available information into an estimate of the state of an electric power system" [3] for real-time monitoring. A state estimation algorithm which is a result from a combination of two fields, load flow and statistical estimation theory [1], fit measurements made on the system to a mathematical model in order to provide a reliable data base for other monitoring, security assessment and control functions [3, 15]. Inputs of the SE are measurements, system parameters, and structural (topology) information. Then state estimator gives the reliable estimation of the system states, $\hat{x}$, to other EMS applications such as contingency analysis, optimal power flow and etc. This process has been shown in Figure (2-5), state estimator is the first block which gathers the all kind of available information about system such as all measurements, parameter values and topology (structural) information and provides the state of the system for other applications.

Today, state estimation is an essential part in almost every Energy Management System (EMS) to insuring secure operation of power systems on transmission level throughout the world [15].

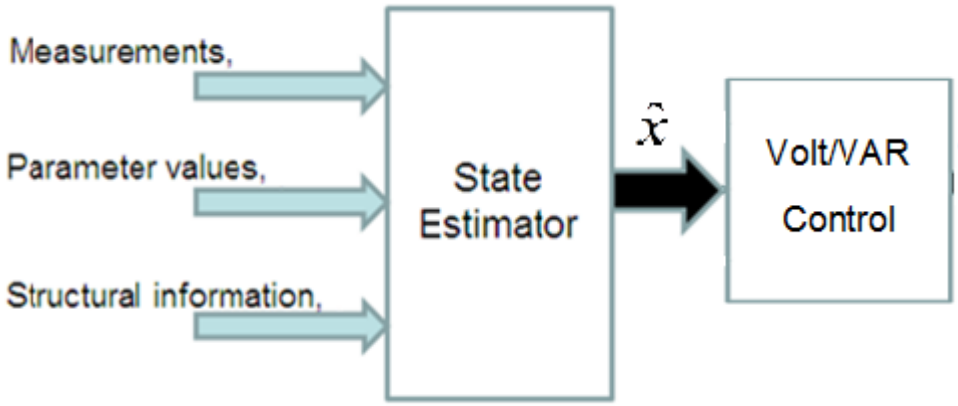


Figure (2-5): State Estimation framework for Volt/VAR Control

Due to radial topology, three-phase unbalanced system, and high resistance to reactance ration and limited number of real-time measurements, transmission state estimation techniques cannot be applied directly to distribution systems. Therefore, SE in distribution systems is not similar to the one in transmission systems, which is a routine task and a host of established methodologies exist [15]. In addition to smart grid initiatives, some new features, issues, and specifications have been added to the distribution systems by integrating the Distributed Energy Resources (DERs), AMR/AMI, new PE components, etc. These systems are named "Active Distribution Systems" or "Smart Distribution Systems." A simple overview of an active distribution system is depicted in Figure (2-6).

The requirements for an optimized Volt/VAR profile at the feeder are controllable volt regulating and VAR control devices with appropriate communication infrastructure. To control and manage these devices beside one optimization tool, the accurate state of the feeder, such as voltage profiles, load estimates, and estimated and scheduled generation from DERs is needed. Here, SE provides this accurate state of the system for VVC application which has been shown in Figure (2-5).

Traditionally, electric distribution utilities have relied primarily on manual, paper driven process for electric distribution operations. Managing the electric distribution system was handled in mostly a manual fashion with voice communications between responsible parties supported by a collection of independent (stand-alone) computer systems, communication facilities, and device controllers. Distribution Management System (DMS) is a concept that integrates these mostly independent facilities so that the distribution system can be operated in a well-coordinated, highly efficient manner. “A DMS is defined by EPRI as a decision support system to assist the control room and field operating personnel with the monitoring and control of the electric distribution system [1]”. The main part of this energy management system is state estimation and load estimation at the customer, feeder, and substation levels.

Furthermore, SE is considered to include any analysis of available data and information that assists the distribution operator to better understand the real-time topology and three-phase electrical state of a distribution network. Due to this fact, SE was listed as one of eight non-prioritized requirements for modeling and simulation at 2008 DOE meeting hosted by Pacific Northwest National Laboratory (PNNL) [5].

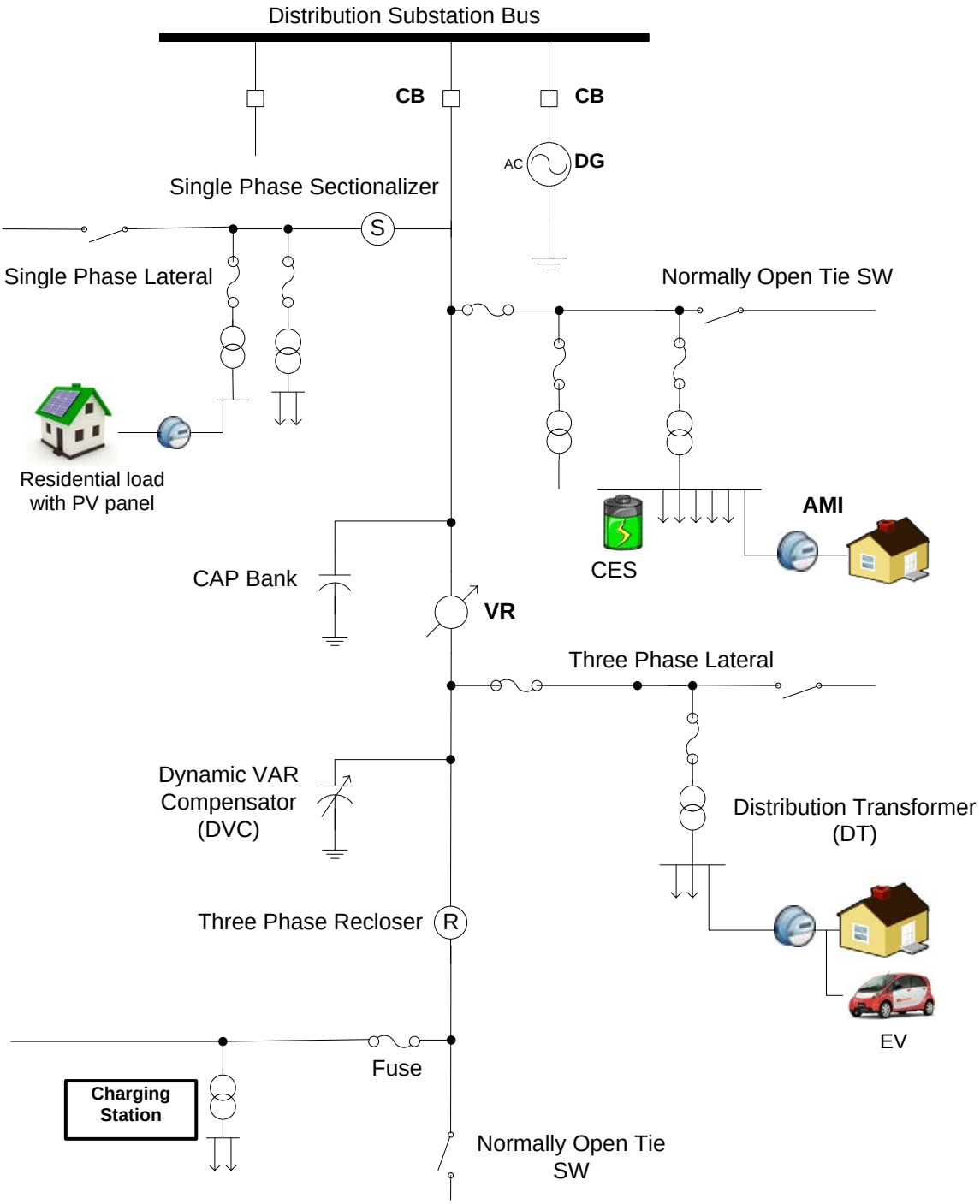


Figure (2-6): Overview of the active distribution system.

Moreover, according to [4], the EDF[1] (Electricity of France) R&D developed and implemented a DMS package offering these three functions: Event Synthesis Function (ESF), Fault Location Function (FLF), and Network Restoration Function (NRF). By January 2011, 90% of EDF control centers for distribution network in France have these functions established and nowadays EDF R&D is implementing other three advanced distribution automation functions, which are:

- Distribution State Estimation (DSE).
- Volt/VAR Control, and
- Steady-state network reconfiguration.

To provide the optimized voltage profile along the feeders of distribution systems, SE is needed to make the system state available to VVC schemes.

These new infrastructures, components, and devices add more challenges to develop the SE for active distribution networks. These challenges are listed in following part.

2-3-2 Challenges of Distribution State Estimation (DSE)

The main challenges involving these specifications, features, and issues can be categorized as follows:

a- System

Distribution systems consist mainly of feeders. Feeders are mainly radial, but have laterals that can be single or two phase, rather than three phase. Furthermore, loads on the feeders are more distributed than that of the transmission, and these loads can be single, two-phase (for residential services) or three phases (for commercial and industrial services). Therefore, distribution systems are unbalanced in nature and many switches (circuit breakers, reclosers, line switches etc.) have been installed and been operated [16].

[1] EDF is the largest electric utility company in the world [Wikipedia].

In addition, feeder line sections are usually short, un-transposed and have high r/x ratio [25]. These features and specifications are listed below:

- Keep changing topology [16, 17, 21, 24, 54, 56, 60, 62, 65-68]:

Since distribution systems involve many protective devices (reclosers, sectionalizers, fuses, etc.) and switches which are normally open during the normal operation and become close when the topology of the serving load changes due to faults or other reasons, therefore keeping the system topology up-to-date has been a major challenge [16]. Some protective devices, such as fuses, are not monitored, but some of them can be monitored. Thus consideration of the effect of these situations is inevitable. Most of the proposed solutions for SE in distribution system considered that the topology is a given accurate and fixed amount [18-47]. A few papers tried to address this challenge in distribution networks, for instance M. Baran *et al*. [54] presented the algorithm for topology error identification by using the branch current state estimation. R. Singh *et al*. [66] proposed a recursive Bayesian approach for identification of the network configuration changes in distribution networks. Moreover, Y. Sharon *et al.* presented a statistical method to detect the changes in the status of switching devices [64]. Therefore, this challenge must be addressed in future solutions for distribution system monitoring and control, by utilizing the outage signal which is available from AMR/AMI and developing the tailored solutions for topology error detection and identification procedures for distribution systems, in both priori and posteriori analyses for topology analysis.

- Presence of three-phase and one-phase electronically coupled distributed generation (DG) and distributed storage (DS) or three-phase rotating machine with various control strategies and operational modes [41-47, 53, 55, 60-70]:

Integration of distributed energy resources (DERs), such as DGs and DSs, to the power grid has become popular, therefore consideration of them in SE of the distribution system is inevitable. In addition, they can provide additional real-time data for SE [41]. One of the solutions for adding DS into the system is Community Energy Storage (CES) which can be installed at the customer side of the distribution transformers. The main characteristic of these

DERs is that they are ‘dispatchable’ and predictable. Different type of DGs are shown in Figure (2-6).

- Presence of ‘non-dispatchable’ DG units, e.g. photovoltaic and wind units [72-73]

These DERs are mostly installed at the customer side, like PV panels at the roof of the houses. The type of these power generations is intermittent, non-dispatchable, and heavily depends on weather conditions, such as sunny, windy and other weather conditions. DERs will change the typical load profiles of the customers. Therefore, these new types of generation must be predicted at the accepted level to have a reasonable picture of the distribution system. Different types of DERs are shown in Figure (2-6).

- Integration of Electrical Vehicles (EV) and Plug-in Hybrid Electrical Vehicles (PHEV) to the power grid [72-75]

These new types of electrical loads will add new kind of uncertainty to the distribution system which must be considered in the operation of the system, which is shown in Figure (2-6). They can be charged in houses or in charging stations. Because of their size as well as their number, the typical load profile of the feeder has to be changed. Therefore, we need to have an acceptable estimation of their loads on the network and their effects on the system.

- Integration of dynamic and static compensator to the power systems at distribution level, such as Dynamic VAR Compensator (DVC), Solid State Transformer (SST), etc. [14, 72, 76, 79, 115]

These new components which are based on power electronic (PE) devices are designed to be added to the current distribution systems to improve the quality of the service by different control strategies and various operational modes, e.g. providing enough reactive power to maintain the voltage profiles in the main feeders [72]. The models of these components must be considered in the system modeling and must utilize their characteristics to find the status of the system e.g. fixed output voltage or zero real power injection and so on. These components

can provide data of the system such as bounded voltage at the connected node [79]. One DVC is shown in Figure (2-6).

- Three-phase unbalances in loading and structure [16, 29, 20, 25-32, 39-40, 43-44, 53-56]

This feature of distribution systems has led to developing the three-phase state estimation to consider all details of the system. It was found that a common single phase SE for transmission systems cannot be used for distribution networks.

- Radial and weekly meshed topology [16, 17, 19, 20, 22, 25-32, 39-40, 43-44, 53-56]

This characteristic of the distribution network can help to find the system state with simpler methods in comparison to common approaches for transmission power systems which mostly have a meshed structure.

- High r/x ratios [16, 19, 25-32, 39-40, 53-56]

“Fast decoupled state estimation has found wide acceptance in industry and various versions have been implemented in control centers all over the world [15]”. As observed in power flow problem, sensitivity of real (reactive) power equations to changes in the magnitude (phase angle) of bus voltages is very low, especially for high voltage transmission systems with low r/x ratio. On the other hand, these assumptions cannot be applied for distribution systems which have short lines with high r/x ratio [26]. Hence, a full coupled solution must be applied for distribution system SE and it should not be sensitive to this ratio.

b- Data

In conventional distribution networks, the most common measurement point is the substation. Very few points on the feeders are measured and monitored in real-time [16]. The most common measurement type is the current magnitude, especially on the feeder. Power measurements are available only at the substation. Here, typical available real-time measurements in distribution feeders are given:

a) Voltage measurements at substation transformer or voltage regulator [16, 17, 25, 39];

b) Power measurements at substation [16, 17, 25, 39]; and

c) Limited current measurements along the main feeder or laterals [17, 19, 39, 40].

The first question typically asked at this point is: "*How many measurements do real distribution systems have*?" The answer is: "Many in absolute numbers, but a few percentage-wise." The average size of the distribution utility we usually work with is about 100-300 subsystems. One subsystem includes from a few hundred branches up to five or six thousand (distribution transformers directly supplying loads are not included in this count). Approximately 2.5% to 5% of branches have either power or current measurements. All LTC transformers and voltage regulators, whose number in a subsystem can reach 50-60, have voltage measurements. At any given time, at least a few subsystems are connected in parallel, creating a large subsystem of up to ten thousand branches [17].

Therefore, we can list the issues, features, and characteristics of the available data in distribution system for SE as follows.

1. Very few measurements are available, sometimes only the voltage and current at the substation, i.e. low redundancy in real-time measurements, e.g. around "0.2-0.3 [20]" [16-23, 37-46, 62-63, 64-69]

This limitation makes the state estimation of the distribution system impossible by just using the real-time data. The usual remedy for this issue is the introduction of pseudo measurement. The main source for these measurements is the estimated loads from historical data such as monthly billing data, yearly random sample, or other ways. Nowadays, by installing the AMR/AMI systems, these load estimation can be more accurate and less delayed from the real-time application.

2. Limited accuracy (poor quality) of load data obtained from load curves [17, 32, 40, 62]

Load estimation is obtained from historical data and not from real-time measurements; hence the accuracy of these estimated loads cannot be very high. And the most operating companies do not update this data often enough. This limitation can be improved by obtaining data from AMR/AMI and updating load profile data.

3. Many of the feeder measurements are current, rather than power magnitude (active and reactive) [16-20, 25-32, 39, 40, 44]

This kind of measurement needs new formulation for measurement functions, which. Another problem is that the direction of the power is not clear, resulting in the availability of only the current magnitude in comparison to power (P and Q) flow measurement in transmission systems. Some methods have been tailored to consider this fact in order to design the state estimator and improve the performance of the estimation by these kind of measurements.

4. Discrepancy between real-time measurements, load data, and static network data (e.g. line impedances, or transformer rated values) [3, 26, 31, 48], and limited reliability of real-time measurements (analog and digital) [26, 32, 48]

Due to measuring errors of meters and problems occurring during transmission of the measurements from field to the control center, there are some devices in the system which are not completely monitored, such as: status of the fuses and switches.

5. Switch states, capacitor bank states and transformer/regulator taps may not be directly monitored, as they typically are monitored on transmission systems [16, 17, 39, 40, 65, 66]

Distribution network conditions change time by time by changing the loads in feeders and by disturbances in the system which causes changes in the network configuration or direction of power. “Undetected switching device errors during estimation show up as analog measurement errors in the solution, which are difficult to distinguish from actual analog measurement errors [67]. Hence, reliable and prompt detection of the switching device statuses is crucial for accurate state estimation [65]”.

6. AMR/AMI data [35, 36, 38, 39, 43]

- Both AMR and AMI systems provide the outages and online restoration verification. By assessing real-time information along all points of the electrical distribution system, outages can be located and during the outage time the topology structure can be corrected and the load estimation can also be modified using this new kind of data which is shown in Figure (2-6).

- Energy consumption with power factor: they provide the energy consumption at different time intervals for real and reactive power on a customer level or big commercial loads. Therefore, the accuracy of the load estimation (pseudo-measurement) will be improved by utilizing these new data.

- Voltage measurements on the customer level (secondary network) which can be converted to the node voltages. Utilizing these new provided data will result in accurate state estimation.

7. Real and reactive power measurements at nodes where the DGs have been connected to the primary distribution feeder and telemetered the generation data and connection status to the control centers in most of the active distribution systems [40-47, 53, 60, 68, 70]

CES can provide voltage at the secondary of the distribution transformer, by detecting the amount of real power supplied or absorbed by the battery pack and state of the charge (SOC) [77]. These measurements can be helpful in having a better estimation.

8. New dynamic and static VAR compensator,

These power electronic based components have been installed in the distribution networks to provide the voltage regulations and reactive power compensation [14, 72, 73, 76, 115], in order to have a regulated voltage at the connection point. Therefore, this information can help us to enhance the estimated system state and have a better picture of the network.

Chapter 3: State Estimation for Active Power Distribution Systems

3-1 Overview

First of all, two available methods for DSE have been introduced from available literature. Then, Branch Current based State Estimation has been chosen for further studies and has been explained in details. Because of the statistical nature of pseudo measurements, the performance of the BCSE needs to be assessed through statistical measurements. Thus, this chapter focuses on the performance of the BCSE method in the presence of measurement noises using Monte Carlo simulations. Some statistical measures were quantified in terms of bias, consistency, and quality through Monte Carlo simulations by Singh *et al.* in [42, 84]. So, Monte Carlo method will be explained briefly in the first parts. Following, these performance measures are adopted and calculated to evaluate the BCSE method [82]. At the end, the application of the standard deviations of system states, branch currents, and voltage magnitudes are mathematically calculated and verified through Monte Carlo simulations for VVC applications.

3-2 Distribution System State Estimation Methods

To address the issues and features mentioned in section 2-3, several algorithms and methods have been proposed. Some of them are discussed here.

3-2-1 Probabilistic Approach for Distribution State Estimation [22-24]

This paper is one of the three companion studies that focus on probabilistic distribution state estimation algorithm [22], other papers discuss the proposed load modeling technique for distribution circuit state estimation [23], as well as field results for these proposed methods implemented on Rochester Gas & Electric Corp [24].

First, Ghosh *et al.* proposed a probabilistic approach to the distribution circuit state estimation problem in 1997 [22]. Recognizing that there is both, a severe limitation to the number of telemetered variables as well as a large degree of uncertainty associated with pseudo-measurements (load demand estimates), the problem was thereby formulated as a radial power flow with telemetered variables acting as solution constraints. The statistics of the states are calculated using a probabilistic formulation of the equations where the states are modeled as random variables. In a sense, one may describe the algorithm as a probabilistic distribution circuit power flow which takes advantage of telemetered variables and the radial nature of distribution circuits.

Estimation Algorithm

The proposed algorithm can be conceptually separated into two distinct parts: Deterministic and probabilistic.

a) *Deterministic section*: this section calculates the expected values of the states resulting from a modification of radial power flow approach by incorporation of real-time measurements. This consists of a series of *backward* and *forward sweeps* until convergence is reached. To account for additional telemetered values besides the substation measurements, the telemetered values were treated as solution constraints.
b) *Probabilistic section*: it calculates higher order moments of the states, specifically variances, based on the results of the deterministic section. The calculation of state statistics converts the application of probability theory into a linearized set of radial power flow equations where all states are modeled as random variable (r.v.) 's, except for voltage angles. For example, X, Y, and, Z are r.v. and they follow this relationship:

$$Z = XY$$

By linearizing around the expected values of X and Y; respectively $E[X]$ and $E[Y]$:

$$Z = E[X]Y + E[Y]X - E[X]E[Y]$$

The higher order moments of Z can be calculated as shown by [37]:

$$Var(Z) = E[X]^2 Var(X) + E[Y]^2 Var(Y) + 2E[X]E[Y]\rho_{XY}\sqrt{Var(X)Var(Y)}$$

where: ρ_{XY} is the correlation coefficient between X and Y. By applying the results of the above relationship among expected values, one can obtain a linearized form of the radial power flow equations. Accordingly, unknown state variances can be calculated from other known state variances.

Estimator Evaluation [22, 14]

To evaluate the probabilistic state estimator for distribution system, a Monte Carlo simulation was performed where the loads were modeled as beta r.v's and the voltage measurements as normal r.v's [22-24]. Five thousand trials were used because no changes were observed in the simulation results for a higher number of trials [22]. The results indicate that there is close agreement between the proposed algorithm and Monte Carlo simulations. It was indicated that the confidence of the voltage estimates is extremely dependent on the confidence associated with the load demand estimates.

Based on the field results, the authors came up with the following comments [24]:

- One of the problems faced in the evaluation of the state estimator algorithm was the potential for error in the customer connectivity database. Another source of error when attempting to build load models, based on customer and transformer data, was the constant change in the circuit. There was some load growth on this circuit and new transformers as well as customers were constantly added.
- The customer load curves used for the state estimator evaluation were derived from the feeder measurements. They were based on determining the exact mix of customer load types below a certain meter point. A more reliable approach to construct these curves would have been to make use of individual customer load survey data.

- The measured voltages on the circuit did not match up well to the calculated estimator values as was hoped. The measured values were rather high, leading to the belief that they could have been in error. One would expect that on a circuit branch with no reactive compensation, line voltages would drop.
- Another problem encountered with the use of the measurements was the tendency for bad values to show up in the database. Sometimes these counter values would overflow, resulting in abnormal values. In one case, abnormal values were showing on a phase due to a loose connection caused by a traffic accident. Some type of bad data detection and compensation should be added to future implementation of a measurement database used for state estimation.

3-2-2 Branch Current Based Three-Phase State Estimation (BCSE) [30, 31, 38, 39]

BCSE is tailored to perform state estimation on distribution networks. There are a number of significant differences in the characteristics of typical distribution networks compared to typical transmission networks. First, Baran and Kelly proposed this algorithm for current and power measurements in1995 [30]. Later in 2009, Baran *et al.* considered the incorporation of voltage measurements in BCSE with adoption of large scale AMI technologies in distribution networks [38].

This method was found to be computationally more efficient and more insensitive to line parameters than the conventional node-voltage-based SE methods. The method had superior performance in terms of computational speed as well as memory requirements. Furthermore, the method was insensitive to line parameters, which improved both its convergence and bad data handling performance.

The BCSE method is based on the Weighted Least Square (WLS) approach. It is very efficient method in handling line-flow and power-injection measurements for radial networks.

However, handling voltage measurements increases the complexity of the algorithm, since using the branch currents as state variables makes the treatment of voltage measurements difficult.

Power system state estimation relies on topological model as well as measurement data obtained from substations. The SE method is based on the WLS approach. WLS state estimation works primarily on finding a system state, represented by $\hat{x}$, by solving the following optimization problem:

$$f = \min_{x} \quad J(x) = \sum_{i=1}^{m} w_i (z_i - h_i(x))^2 = [z - h(x)]^T W [z - h(x)]$$

where: w_i and h_i represent the weight and the measurement function associated with measurement z_i; respectively. By solving this optimization problem, we obtain the estimated state $\hat{x}$ which must satisfy the following optimality condition:

$$\frac{\partial f}{\partial x_i} = 0 \Rightarrow \frac{\partial f}{\partial x_i} = 2 \cdot \sum_{i=1}^{m} w_i (z_i - h_i(x)) \cdot \frac{\partial h_i(x)}{\partial x_i} = 0$$

$$\sum_{i=1}^{m} w_i (z_i - h_i(x)) \cdot \frac{\partial h_i(x)}{\partial x_i} = 0 \Rightarrow H^T W \left[z - h(x) \right] = 0$$

where: $H(x) = \dfrac{\partial h(x)}{\partial x}$ is the Jacobian matrix of the measurement function $h(x)$. Since $h(x)$ is usually non-linear, the solution is obtained by an iterative method. The iterative method involves solving the linear equation of the following type at each iteration to compute the correction $x^{k+1} = x^k + \Delta x^k$.

$$\Rightarrow \quad G(\hat{x}^{(k+1)} - \hat{x}^{(k)}) = H^T W[z - h(x^{(k)})]$$

where: $\frac{\partial d(\hat{x})}{\partial x} = -H^T WH = -G$ is the Jacobian of the optimality condition equation called Gain matrix:

$$d(\hat{x}) = H^T W[z - h(x)]$$

BCSE uses branch current as a system state rather than voltage which is a system state in conventional SE. Hence, the state vector in BCSE becomes:

$$x = [I_r, I_x]$$

where: I_r is the current real part and I_x is the current imaginary part.

Feeder Representation

One of the main challenges in implementing this approach for SE in distribution feeders is incorporating the unbalanced nature of distribution feeders into the problem. The most important of these issues is the representation of feeders which will be discussed in this subsection. In general, main feeders are three-phase; however some laterals can be two-phase or single-phase. The lines are usually short and un-transposed. Loads can be three-phase, two-phase or single-phase, such as: residential customers. Therefore, it is desirable to use a three phase model as recommended for power flow analysis of feeders. A three-phase line model takes into account the magnetic coupling between the phases in lines, which for a line section $l, l = 1 \cdots b,$ such as the one shown in Figure (3-1), is in the following form:

$$\begin{bmatrix} V_{r,a} \\ V_{r,b} \\ V_{r,c} \end{bmatrix} = \begin{bmatrix} V_{s,a} \\ V_{s,b} \\ V_{s,c} \end{bmatrix} - l \begin{bmatrix} z_{aa} & z_{ab} & z_{ac} \\ z_{ba} & z_{bb} & z_{bc} \\ z_{ca} & z_{cb} & z_{cc} \end{bmatrix} \begin{bmatrix} I_{l,a} \\ I_{l,b} \\ I_{l,c} \end{bmatrix}$$

or $$V_r = V_s - Z_l I_l$$

where: $Z_l = g_l Z$ is the line impedance matrix and g_l is the line length. Note that this equation is written for the assumed branch current direction shown in Figure (3-1), and the phases are labeled as: $\varphi = a, b, c$.

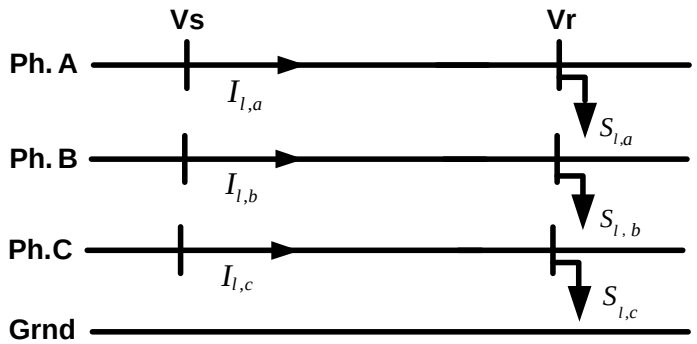


Figure (3-1): A three-phase line section.

- Hence, the only difference between the node voltage based SE and BCSE is the measurement functions associated with the type of measurements to be processed. To illustrate these functions for BCSE, consider two cases.
- *Case 1*: power flow (*P*, *Q*) or current magnitude (*I*) measurements on a line section of a feeder
- *Case 2*: voltage measurement (*V*) at a node of a feeder.

Case 1 – power flow (*P, Q*) or current magnitude (*I*) measurements [30, 31, 38, 80]

Power measurements in BCSE are converted to equivalent complex current measurement by using the current estimate of the node voltage:

$$I_r^{\,m} = \frac{P^m V_r + Q^m V_x}{V_r^2 + V_x^2} \;,\; I_x^{\,m} = \frac{P^m V_x - Q^m V_r}{V_r^2 + V_x^2}$$

In addition, we can rewrite the previous equations in the following form:

$$I_l^{\,m} = \left[\frac{P^m + jQ^m}{V}\right]^* = I_{r,l}^m + jI_{x,l}^m$$

Hence, the resulting measurement functions are linear as the state variables are the complex branch currents,

$$I_l = I_{r,l} + jI_{x,l}$$

By assuming $V = 1\angle 0$ at the substation, we have $S = P + jQ = I^*$. The phase angles are the same for both the complex power and current with a sign change at the substation, assuming small voltage angles, $angle(I) \cong -angle(S)$ [31]. In this condition the Jacobian terms become "+1" and "-1".

The current magnitude measurements, on the other hand are non-linear, as

$$\left|I_l\right| = \sqrt{I_{r,l}^2 + I_{x,l}^2}$$

The current magnitude measurements introduce coupling terms between the real and imaginary parts. For example, the current measurement I_l^m introduces the following non-zero elements into the measurement Jacobian H:

$$\frac{\partial h_I^m}{\partial I_r} = \cos\phi \ , \ \frac{\partial h_I^m}{\partial I_x} = \sin\phi$$

where: $\phi_{l,\varphi} = Tan^{-1}(I_{x,l,\varphi} / I_{r,l,\varphi})$.

Due to the lack of phase information, the linearity is lost. To address this non-linearity issue, current measurement could be processed by multiplying the measured magnitude, $\left|I_l^m\right|$ with the ratio of the phasor I_l^{cal}, calculated branch current of branch, to its absolute value becomes [31]:

$$I_l^m = \left|I_l^m\right| \left(\frac{I_l^{cal}}{\left|I_l^{cal}\right|} \right) = I_{r,l}^m + jI_{x,l}^m$$

In this situation, the Jacobian terms with respect to the branch currents are "1". So far, by these treatments, a constant Jacobian matrix has been developed for each phase and decoupled on real and imaginary parts. By this linearization, the Jacobian matrix for current and power measurements, $H_{PM,CM}$, becomes [31]:

$$H_{PM,CM} = \begin{bmatrix} H_{aa}^r & 0 & 0 & 0 & 0 & 0 \\ 0 & H_{aa}^i & 0 & 0 & 0 & 0 \\ 0 & 0 & H_{bb}^r & 0 & 0 & 0 \\ 0 & 0 & 0 & H_{bb}^i & 0 & 0 \\ 0 & 0 & 0 & 0 & H_{cc}^r & 0 \\ 0 & 0 & 0 & 0 & 0 & H_{cc}^i \end{bmatrix}$$

where: H_{aa}^r is the sub-Jacobian matrix of the real part of phase a and H_{aa}^i is the sub-Jacobian matrix of the imaginary part of phase ; respectively and similar for other phases. The non-zero terms of the sub-Jacobian matrices are +1 and -1 values only.

Case 2–Voltage magnitude (*V*) measurements [38, 80]

A voltage at the node t of a radial feeder V_t is the voltage at the substation minus the voltage drop on the line sections between the substation and this node, hence, the measurement function for the voltage measurement V_t can be written in terms of the branch currents as:

$$V_t^m = V_s - \sum_{l \in \Omega_t} Z_l I_l$$

where: Ω_t is the branch set from substation to the bus t. The voltage magnitude measurements introduce coupling terms between the phases of branch currents and both the real and

imaginary parts of branch currents. The voltage measurement V_t introduces the following non-zero elements into the measurement, Jacobian H.

$$\frac{\partial h_{V,\varphi}^{m}}{\partial I_{r,l}} = X_l \sin\phi_{l,\varphi} - R_l \cos\phi_{l,\varphi} \ , \ \frac{\partial h_{V,\varphi}^{m}}{\partial I_{x,l}} = -R_l \sin\phi_{l,\varphi} - X_l \cos\phi_{l,\varphi}$$

where: $Z_l = R_l + jX_l$ is line impedance and $V_s = V_{r,s} + jV_{x,s}$ is substation voltage [38] :

$$\phi = Tan^{-1}\left[\frac{V_{r,s} - real(\sum_{j=1}^{m} Z_j I_j)}{V_{x,s} - imag(\sum_{j=1}^{m} Z_j I_j)}\right]$$

Hence, both the Jacobian H and the gain matrix G must be revised to include voltage measurements in BCSE. In addition, these matrixes are coupled in the sense of real and imaginary parts as well as phases. We can rewrite the previous equations in the format of real and imaginary parts of V_t, by considering phase φ :

$$V_{r,t,\varphi} = V_{r,1,\varphi} - \sum_{l\in\Omega_t} (R_{l,\varphi} I_{r,l,\varphi} - X_{l,\varphi} I_{x,l,\varphi})$$

$$V_{x,t,\varphi} = V_{x,1,\varphi} - \sum_{l\in\Omega_t} (R_{l,\varphi} I_{x,l,\varphi} + X_{l,\varphi} I_{r,l,\varphi})$$

where: V_1 is the substation voltage. Using the calculated voltage angle, the real and imaginary parts of the equivalent voltage measurement at bus t can be written as [80]:

$$V_t^{m-eqv} = \left|V_t^m\right|\left(\frac{V_t^{cal}}{\left|V_t^{cal}\right|}\right) = V_{r,t}^{m-eqv} + jV_{x,t}^{m-eqv}$$

Here, equivalent voltage measurement, V_t^{m-eqv}, has been formulated from the available measurements, $\left|V_t^m\right|$, and the calculated voltage from first iteration of load flow calculation. The mismatched for voltage measurements are:

$$\Delta V_{r,t} = V_{r,t}^{m-eqv} - V_{r,t}^{cal},\ \Delta V_{x,t} = V_{x,t}^{m-eqv} - V_{x,t}^{cal}$$

where: $\Delta V_{r,t}$ and $\Delta V_{x,t}$ are the real and imaginary parts of mismatch vectors for voltage measurements; respectively. The differential terms of voltage measurements with respect to branch currents can be developed. Then, the general expression of the sub-Jacobian matrix for voltage measurements with phase quantities can be expressed as [80]:

$$\begin{bmatrix} -R_{aa} & -R_{ab} & -R_{ac} & \vdots & X_{aa} & X_{ab} & X_{ac} \\ -R_{ba} & -R_{bb} & -R_{bc} & \vdots & X_{ba} & X_{bb} & X_{bc} \\ -R_{ca} & -R_{cb} & -R_{cc} & \vdots & X_{ca} & X_{cb} & X_{cc} \\ \cdots & \cdots & \cdots & \vdots & \cdots & \cdots & \cdots \\ -X_{aa} & -X_{ab} & -X_{ac} & \vdots & -R_{aa} & -R_{ab} & -R_{ac} \\ -X_{ba} & -X_{bb} & -X_{bc} & \vdots & -R_{ba} & -R_{bb} & -R_{bc} \\ -X_{ca} & -X_{cb} & -X_{cc} & \vdots & -R_{ca} & -R_{cb} & -R_{cc} \end{bmatrix} \begin{bmatrix} \Delta I_{r,a} \\ \Delta I_{r,b} \\ \Delta I_{r,c} \\ \cdots \\ \Delta I_{x,a} \\ \Delta I_{x,b} \\ \Delta I_{x,c} \end{bmatrix} = \begin{bmatrix} \Delta V_{r,a} \\ \Delta V_{r,b} \\ \Delta V_{r,c} \\ \cdots \\ \Delta V_{x,a} \\ \Delta V_{x,b} \\ \Delta V_{x,c} \end{bmatrix}$$

where: matrix R_{aa} contains the values of 0 and the line resistance; while matrix X_{aa} contains the values of 0 and the line reactance. The previous equation can be rewritten as follows:

$$\begin{bmatrix} -R_{aa} & X_{aa} & \vdots & -R_{ab} & X_{ab} & \vdots & -R_{ac} & X_{ac} \\ -X_{aa} & -R_{aa} & \vdots & -X_{ab} & -R_{ab} & \vdots & -X_{ac} & -R_{ac} \\ \cdots & \cdots & \vdots & \cdots & \cdots & \vdots & \cdots & \cdots \\ -R_{ba} & X_{ba} & \vdots & -R_{bb} & X_{bb} & \vdots & -R_{bc} & X_{bc} \\ -X_{ba} & -R_{ba} & \vdots & -X_{bb} & -R_{bb} & \vdots & -X_{bc} & -R_{bc} \\ \cdots & \cdots & \vdots & \cdots & \cdots & \vdots & \cdots & \cdots \\ -R_{ca} & X_{ca} & \vdots & -R_{cb} & X_{cb} & \vdots & -R_{cc} & X_{cc} \\ -X_{ca} & -R_{ca} & \vdots & -X_{cb} & -R_{cb} & \vdots & -X_{cc} & -R_{cc} \end{bmatrix} \begin{bmatrix} \Delta I_{r,a} \\ \Delta I_{x,a} \\ \cdots \\ \Delta I_{r,b} \\ \Delta I_{x,b} \\ \cdots \\ \Delta I_{r,c} \\ \Delta I_{x,c} \end{bmatrix} = \begin{bmatrix} \Delta V_{r,a} \\ \Delta V_{x,a} \\ \cdots \\ \Delta V_{r,b} \\ \Delta V_{x,b} \\ \cdots \\ \Delta V_{r,c} \\ \Delta V_{x,c} \end{bmatrix}$$

Now, we can write the whole system in the form of a Jacobian matrix for the measurement set composed by bus injection, line flow, and voltage measurements, i.e. PMs, CMs, and VMs:

$$H = \begin{bmatrix} H_{PM,CM} \\ \cdots \\ H_{VM} \end{bmatrix}$$

where: H_{VM} is the sub-Jacobian matrix for voltage measurements. Note that, in the real world no matter how closely the conductor is bundled, the mutual coupling terms are always smaller than the self impedance [80].Since, the self impedance is significantly greater than the mutual-coupling terms, the off-diagonal blocks in the sub-Jacobian matrix of voltage measurements can be neglected. Then, the Jacobian matrix for phase a, H_a ,can be expressed as [80]:

$$H_a = \begin{bmatrix} H_{aa}^r & 0 \\ 0 & H_{aa}^i \\ -R_{aa} & X_{aa} \\ -X_{aa} & -R_{aa} \end{bmatrix}$$

Here, mutual-coupling terms are neglected in the WLS-solving process, not in the modeling phase. Therefore, the constant-gain matrix including current, power, and voltage measurements for phase a can be expressed as:

$$G_a = \begin{bmatrix} H_{aa}^r & 0 \\ 0 & H_{aa}^i \\ -R_{aa} & X_{aa} \\ -X_{aa} & -R_{aa} \end{bmatrix}^T \begin{bmatrix} W_{r,a} & 0 & 0 & 0 \\ 0 & W_{i,a} & 0 & 0 \\ 0 & 0 & W_{v,a} & 0 \\ 0 & 0 & 0 & W_{v,a} \end{bmatrix} \begin{bmatrix} H_{aa}^r & 0 \\ 0 & H_{aa}^i \\ -R_{aa} & X_{aa} \\ -X_{aa} & -R_{aa} \end{bmatrix}$$

where: $W_{r,a}$, $W_{i,a}$, and $W_{v,a}$ are weighting matrices for real and imaginary parts of PMs and CMs as well as VMs; respectively for phase a. The previous equation can be rewritten for real and imaginary parts as [80]:

$$G_a = \begin{bmatrix} G_a^r & -R_{aa}^T W_{v,a} X_{aa} + X_{aa}^T W_{v,a} R_{aa} \\ -X_{aa}^T W_{v,a} R_{aa} + R_{aa}^T W_{v,a} X_{aa} & G_a^r \end{bmatrix}$$

$$G_a^r = (H_{aa}^r)^T W_{r,a} H_{aa}^r + R_{aa}^T W_{v,a} R_{aa} + X_{aa}^T W_{v,a} X_{aa}$$

$$G_a^i = (H_{aa}^i)^T W_{i,a} H_{aa}^i + R_{aa}^T W_{v,a} R_{aa} + X_{aa}^T W_{v,a} X_{aa}$$

The off-diagonal blocks of the gain matrix are very likely to be smaller than the corresponding terms in the diagonal blocks owing to the partial cancellation. Now, we have H and G matrices for phase a. The formulation for phases b and c can be developed similarly. Therefore, all matrices needed are available to run the BCSE algorithm.

BCSE algorithm

BCSE constructs both the Jacobian H and Gain G matrices and solves the updated equations iteratively [30, 38, 80]. The algorithm involves the following steps at each iteration k:

Step 1: Given the node voltage V^{k-1}, convert power measurements into equivalent current measurements.

Step 2: Use current measurements to obtain an estimate of branch currents $\hat{x}_\varphi^k = [\hat{I}_{r,\varphi}^k \quad \hat{I}_{x,\varphi}^k]$ by solving the update equations (1) for each phase $\varphi = a, b, c$.

Step 3: Given the branch currents, update the node voltages V^k by the forward sweep procedure.

Step 4: Check for convergence; if two successive updates of branch currents are less than a convergence tolerance then stop, otherwise go to step 1.

Generally, the performance of the BCSE can be summarized as follows:

- Computationally efficient in terms of computation speed and memory requirement.

- Insensitive to line parameter which improves both its convergence and bad data handling.
- In [38, 80] voltage measurements are added to the measurement set and its performance was proven.

This BCSE method has been improved by Lin, Teng, and Chen in years 2001 [31] and 2002 [80]. All of the previous sections are come from these references: [25], [27], [30], and [80]. Other improvements were conducted by Wang and Schulz in year 2004 [32].The latter proposed another revised branch current-based SE algorithm for distribution system by defining the state vector in polar form rather than the Cartesian form proposed in other works [32]. The performance of this SE method will be assessed in the following chapter.

3-3 Performance of BCSE Method

3-3-1 Monte Carlo (MC) Simulation

Monte Carlo simulation is used to estimate expected values of random variables when it is infeasible or impossible to compute an exact result with a deterministic algorithm as well as to verify the mathematical calculation. Monte Carlo methods are a class of computational algorithms that rely on repeated random sampling to compute their results. These methods are often used when simulating physical and mathematical systems. Because of their reliance on repeated computation and random or pseudo-random numbers, Monte Carlo methods are most suited to calculation by a computer. In addition, they tend to be used when it is infeasible or impossible to compute an exact result with a deterministic algorithm [81-83, 89].

In general, the Monte Carlo simulation involves the following series of steps [83]:

- *Step 1* – Construction of a simulated “universe” of some randomizing mechanism whose composition is similar to the universe whose behavior we wish to describe and investigate. The term “universe” refers to the system that is relevant for a single simple event.

- *Step 2* – Specification of the procedure that produces a pseudo-sample which simulates the real-life sample in which we are interested. That is, specification of the procedural rules by which the sample is drawn from the simulated universe. These rules must correspond to the behavior of the real universe in which we are interested. To put it in another way, the simulation procedure must produce simple experimental events with the same probabilities that the simple events have in real world.
- *Step 3* – If several simple events must be combined into a composite event, and if the composite event was not described in the procedure in step 2, this is the time to describe it.
- *Step 4* – Calculation of the probability of interest from the tabulation of outcomes of the resampling trials.

3-3-2 Monte Carlo Simulation Sample Size

Determining the sample size in a Monte Carlo study is no different than for other types of studies. Basically, it is key to determine the acceptable error of estimates or power of tests and invert analytical functions of sample size set equal to those values. Monte Carlo simulation sample size can be defined depending on which quantities seem interesting [83]. Here, the interesting quantities are voltage magnitude standard deviations (variances) which are chosen to see the variation of voltage profile along the feeder for VVC application. Hence, we utilized the appropriate approaches which are available in [81-82] for VVC. At the end, we chose the number of MC by looking at the changes of interesting quantities' variances in regard of simulation numbers.

- Based on Power Estimation [81, 83]

For a new test procedure of the form, "reject the null hypothesis if $T > c_\alpha$", the power at a particular alternative by:

$$\widehat{pow} = \frac{1}{N}\sum_{i=1}^{N} I(T_i > c_\alpha)$$

where: T_i is the test statistic for the i-th Monte Carlo sample, c_α is a given critical value, and I is the indicator function having value 1 if $T_i > c_\alpha$ and 0 otherwise. This is binomial sampling, and the worst variance of our estimate (occurring at power ½) is given by $1/(4N)$. Setting $d = 1/(2\sqrt{N})$ yields:

$$NMC = \frac{1}{4d^2}$$

For VVC application d should be around $2/720$. This yields NMC to become around 32,400 when using this approach.

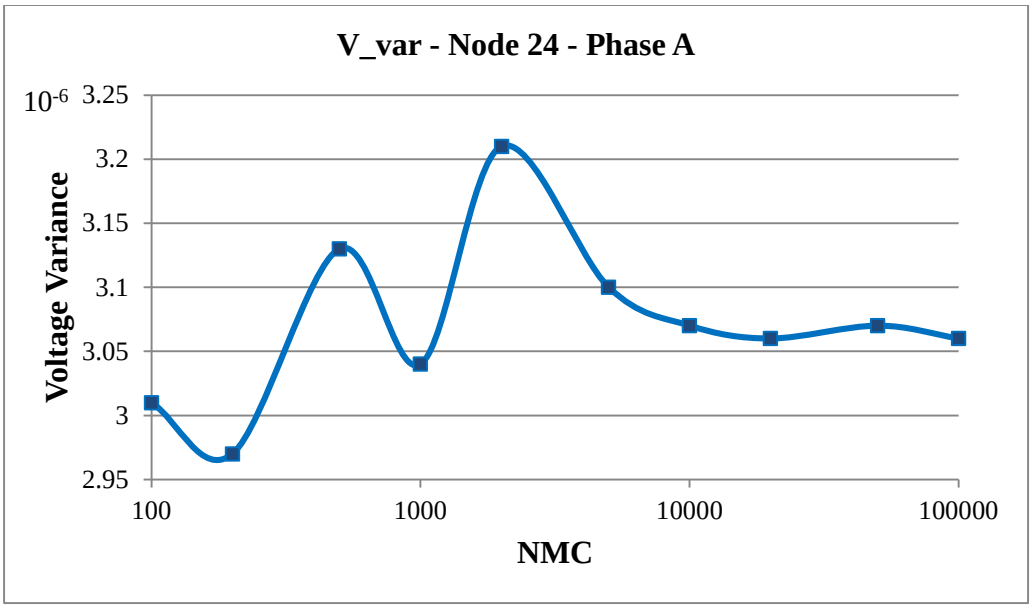
- Based on Confidence Intervals [81, 83]

Coverage probability and average confidence interval length are both important quantities that should be reported whenever studying confidence intervals. Obviously, it is desirable to keep intervals that achieve the nominal $1-\alpha$ coverage (like 95% which is common in statistical analysis) and those intervals which are short on average. For sample size considerations, we need a preliminary estimate $\tilde{\sigma}$ of the standard deviation of the lengths and an acceptable d value for the standard error of our estimate of average length. For coverage estimation or one-sided error estimation, it is inverted to $d = \sqrt{\alpha(1-\alpha)/N}$ where d is the acceptable standard deviation for coverage estimate, to get:

$$NMC = \frac{\alpha(1-\alpha)}{d^2}$$

For VVC application d is around 0.0028. Therefore, for 95% coverage and using this approach, NMC becomes around 6,156.

By now, we have different values for required number of simulations from 6,000 to 32,000. To have a better understanding of variance changes regarding different number of MC simulations, variance of voltage magnitude (i.e. interesting quantity) at nodes 24 and 34 for different number of simulations are shown in the following Figure (3-2).



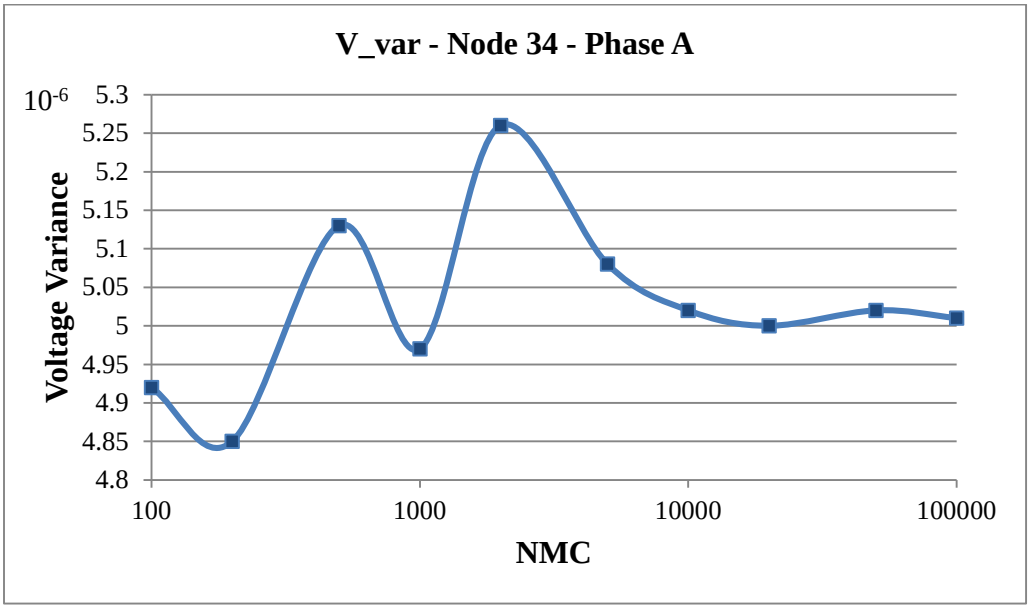


Figure (3-2): Voltage variance for different number of Monte Carlo simulations.

It is worth to be noted that the same behavior at those two nodes has been observed at different nodes. For different number of simulations, ranging from 100 to 100,000, it is obvious that after 10,000 simulations the changes of variance become stable for these two voltages at nodes

24 and 34. In addition, this pattern has been observed for other voltages along the feeder. To add up these simulation results with previous approaches in order to find the acceptable number of trials, it has been found that running 10,000 Monte Carlo simulations is sufficient to estimate the variance (standard deviation) of the voltages.

The performance measures used are bias, consistency and overall quality. These concepts of statistical techniques are illustrated and used extensively in this chapter [42, 84].

3-3-3 Monte Carlo Simulations on BCSE Method

Here, Monte Carlo simulation procedure is described using BCSE Method below [82]:

Step 1 – BCSE method tries to find a system state, represented by c, branch current magnitude, and voltage magnitude by minimizing the weighted sum of the squares of the measurements errors.

Step 2 – First, the actual measurements were obtained by running a power flow for the given load. Then measurement error obtained from random generator based on Normal distribution was added to the actual measurements.

$$Z = Z^a \pm e_Z$$

where Z^a is actual data and e_Z is the measurement error. The forecasted load data is created by perturbing the actual load data by adding error of 50%. The power and current magnitude measurement errors are selected from Normal distribution with 3% accuracy.

Step 3 – The sample size is defined in order to achieve acceptable results. In this study, 10,000 Monte Carlo simulations were chosen.

Step 4 – To calculate the probability of the interest, the estimated states ($x = [I_r, I_x]$) of branch currents are obtained. Consequently, the performance measures which are bias, consistency and overall quality are examined.

For testing the BCSE, 34 node IEEE test feeder is used. The test feeder is a 34 bus, 23kV, 3-phase radial feeder [87]. A one-line diagram of the feeder is given in Figure (3-3) with

renumbering the nodes to make the illustration of the results easier. The feeder is predominantly three-phase with some single-phase laterals and has both spot and distributed loads. For test purposes, distributed line section loads are lumped equally at terminal nodes of the line section. The nominal load data are taken as the actual load and the power flow results are used to determine the correct measurements for this load.

For SE, the available measurements assumed are given in Figure (3-3) as well: voltage and power flow at the substation, current measurements (CMs), and voltage measurements (VMs).

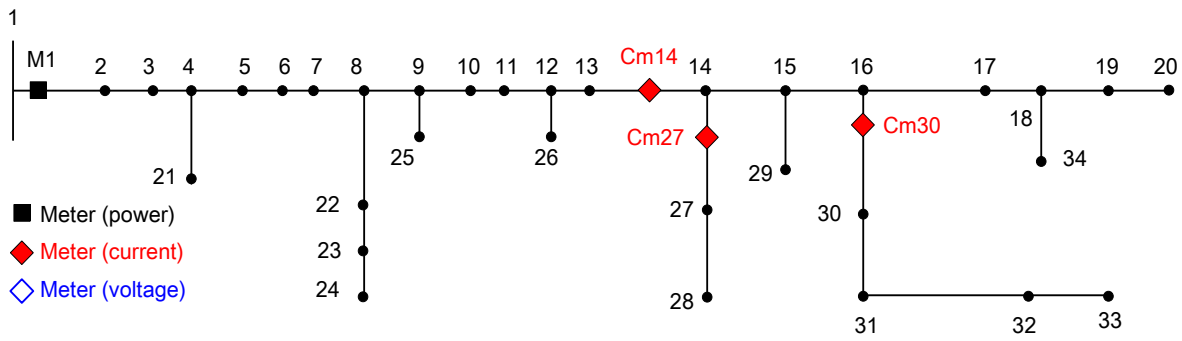


Figure (3-3): Single-line diagram of IEEE 34 node feeder [87].

The algorithm was implemented and the result was analyzed using MATLAB R2011b. Convergence tolerance is 10^{-3} and maximum iteration is 10. For testing, the following cases were considered:

Case 1: Power and voltage measurement (M1) at the substation and forecasted load data.

Case 2: Case 1 plus three CMs at branches: Cm14, Cm27, and Cm30.

Case 3: Case 1 plus four VMs at nodes: Vm14, Vm18, Vm28, and Vm33.

These Monte Carlo simulations were run and have been compared with the PF results in following Figures (3-4) and (3-5). Figure (3-4) shows the estimated voltage profile from BCSE outputs for the prototype feeder in error bars, $\bar{V} \pm z_{\alpha/2}.\sigma_{\bar{V}}$ at a 95% confidence level, versus the actual voltage profile from PF output for phase A. As shown in this same figure, the true values, PF output, are within the estimated range from BCSE. For branch current magnitude, the output of the BCSE versus the true value is illustrated in Figure (3-5). Estimated branch currents for each branch are shown in error bars at a 95% confidence level with true values from PF. Here, the true values of branch currents are within the 95% confidence band.

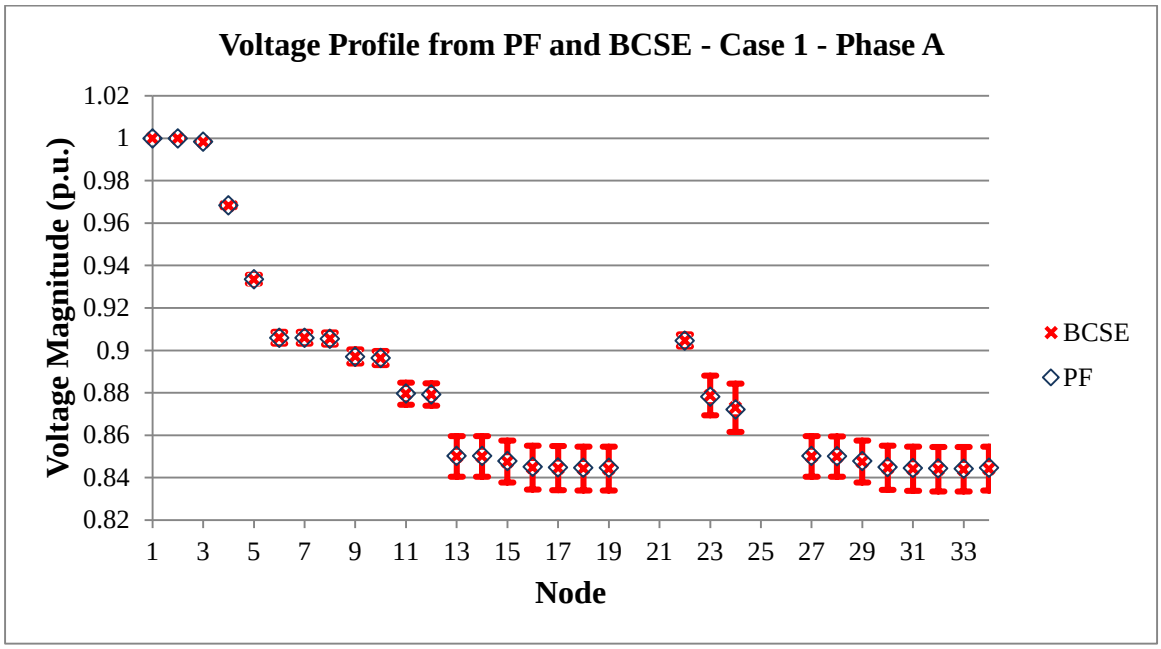


Figure (3-4): Voltage profile for prototype feeder from PF and BCSE – Case 1 and Phase A.

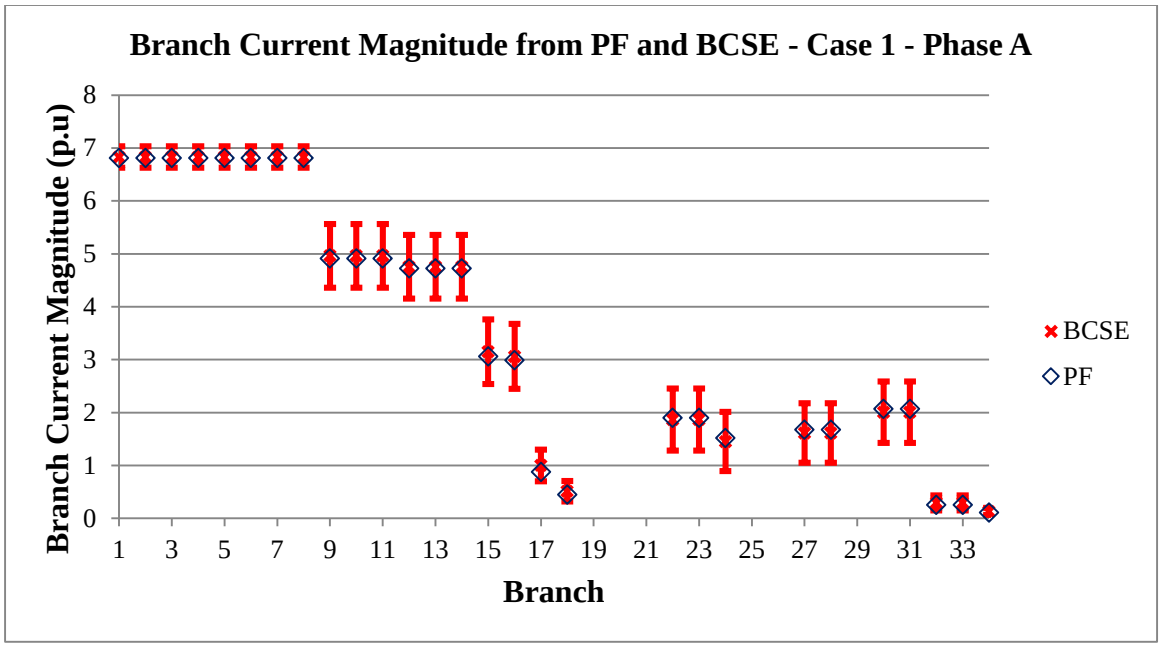


Figure (3-5): Branch current magnitude for prototype feeder from PF and BCSE – Case 1 and Phase A.

For system states, both real and imaginary part of branch currents, the outputs of the BCSE versus the true values are illustrated in Figure (3-6) and Appendix 1. Estimated system states, $\hat{x}$, for each branch are shown in error bars, $\overline{x} \pm z_{\alpha/2}.\sigma_{\overline{x}}$ at a 95% confidence level, for real and imaginary part with the true values from PF. Again, the true values of branch currents are within the 95% confidence band. More results for phases B and C as well as different cases are compiled in the Appendix 1. These results provide the good sense of BCSE output with associated error in the load estimation and real-time measurements.

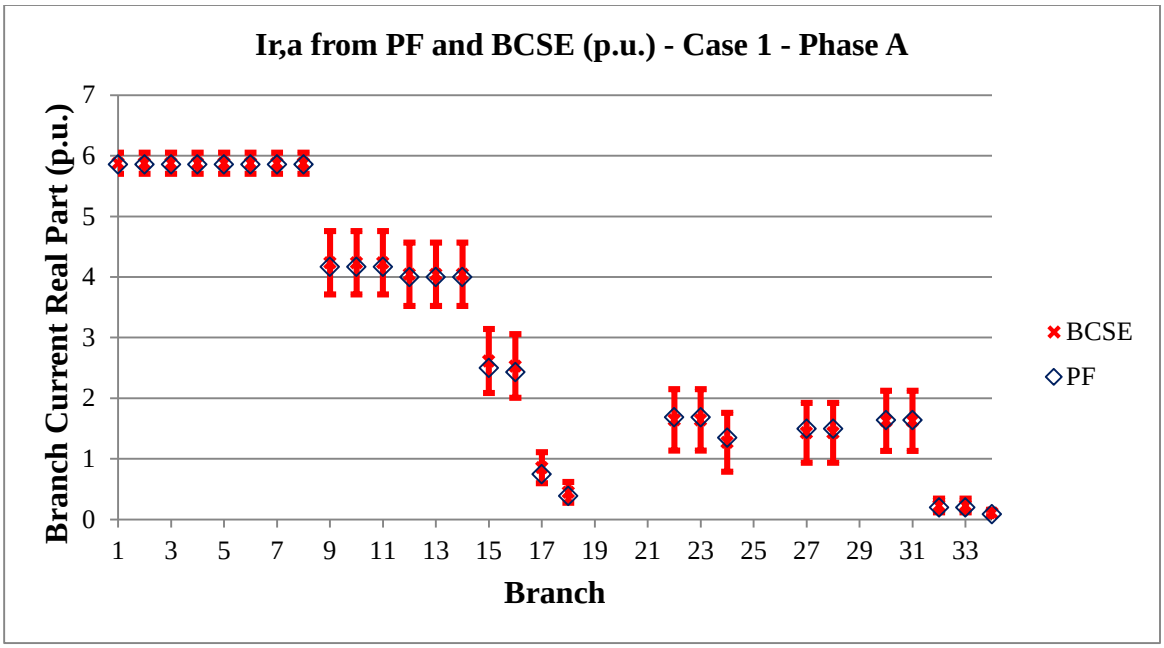


Figure (3-6): Estimated system states and true values, $x = [I_r, I_x]$, from BCSE and PF; respectively for prototype feeder, real part – Case 1 and Phase A.

Now, we will assess the performance of BCSE with metrics in regard to the bias, consistency, and quality. These studies are examined in the next section. Because of the statistical nature of pseudo measurements, the performance of the BCSE needs to be assessed through statistical measures. Singh *et al.* have proposed some statistic measures quantified in terms of bias, consistency, and quality through Monte Carlo simulation in [42, 82, 84] for assessing the quality of state estimation. In this section, these performance measures are explained and adopted to evaluate the BCSE method.

3-3-4 Bias

State vector, $\hat{x}$ is an unbiased estimate of true state vector, x_t if the expected value of $\hat{x}$ is equal to x_t [42] .

$$E(\hat{x}) = x_t$$

This is equivalent to saying that the mean of the probability distribution of $\hat{x}$ is equal to x_t .Since we have N state ($x = [I_r, I_x]$) of branch currents, we need to perform multivariate tests in order to test the biasness of the BCSE method. Oja and Randles have examined a number of hypothesis testing problem settings for multivariate data in [82, 85]. In this section, hypothesis testing is adopted to test the bias in estimated state variables obtained from BCSE method.

3-3-4-1 Multivariate Statistical Tests

One of the most important problems in the area of multivariate analysis is to obtain the mean vector of the given sample. After that, we can get rough information about the population where the sample is surveyed. Often people hope to test the hypothesis of whether the sample mean vector equals to the specified value in advance. Generally, there are two typical approaches, parametric and nonparametric tests. Most hypothesis tests are based on the assumption that random samples are from normal populations which we have assumed here. This is called a parametric test because it is based on a particular parametric family of distributions. Alternately, these procedures are not distribution-free because they depend on the assumption of normality. The primary advantage of the parametric test is that it has greater statistical power to detect differences. For parametric tests, Hotelling's T^2 test is usually used [85].

3-3-4-2 Hotelling's T^2 Test [82, 85]

Let $x_1, x_2, \cdots, x_N$ be independent and identically distributed from $F(x-\theta)$, where: $F(\cdot)$ represents a continuous p-dimensional distribution "located" at the vector parameter $\theta = (\theta_1, \theta_2, \cdots, \theta_p)^T$. The hypothesis of whether the sample mean vector equals to the vector specified in advance is :

$$H_0 : \theta = \mu \qquad H_a : \theta \neq \mu$$

Note the above hypothesis is equivalent to :

$$H_0 : \theta - \mu = 0 \qquad H_a : \theta - \mu \neq 0$$

Hotelling's T^2 test statistics involves the following calculations.

$$T^2 = N(\overline{x} - \mu)^T S^{-1}(\overline{x} - \mu)$$

where:

$\overline{x}$: the mean vector of sample, $ave\{x_i\}$

N : the sample size.

S : the sample covariance matrix, $ave\{(x_i - \overline{x})(x_i - \overline{x})^T\}$

μ : the vector given in the hypothesis.

$\frac{N-1}{N-p}T^2$ has F distributions with degrees of freedom, p and $N-p$. Thus, for a given significant level α, the null hypothesis is rejected when :

$$T^2 \geq \frac{N-1}{N-p}F_{p,N-p}(\alpha)$$

where: $F_{v1,v2}(\alpha)$ is the upper α the quantile of an F distribution with $v1$ and $v2$ degrees of freedom.

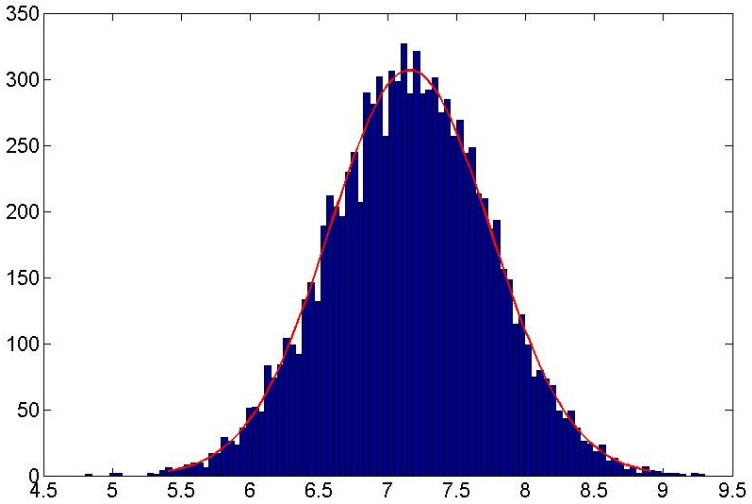
Moreover, the p-value for this statistics test is:

$$p-value = 1 - F\left(\frac{N-p}{(N-1)p}\cdot T^2, p, N-p\right)$$

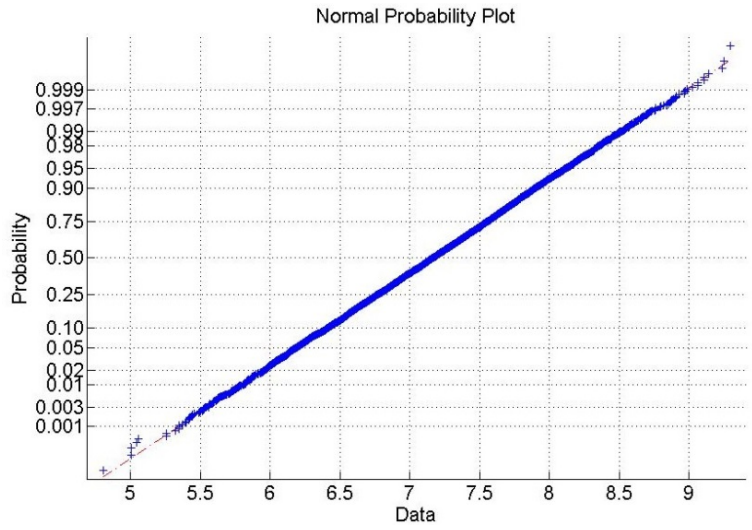
3-3-4-3 Multivariate Statistical Test Results

Many bias tests depend on the assumption of normality. Thus, we checked the normality of the result given in Figure (3-7) before the bias test. To check normality, histogram and normal probability plot using MATLAB is given in Figures (3-7), (3-8) and (3-9). The purpose of a normal probability plot is to graphically assess whether the data in x could come from a normal distribution or not. If the data are normal, the probability plot should be linear. Other

distribution types will introduce curvature in the plot. Current magnitude for branch 1 at phases A and B, as well as current magnitude for branch 32 at phase C are shown in the aforementioned figures. These figures show the histogram and normal probability plot of the result obtained from 10,000 Monte Carlo simulations.



(a)



(b)

Figure (3-7): Current Magnitude of Branch 1, Phase A; (a) is histogram and (b) is normal probability plot.

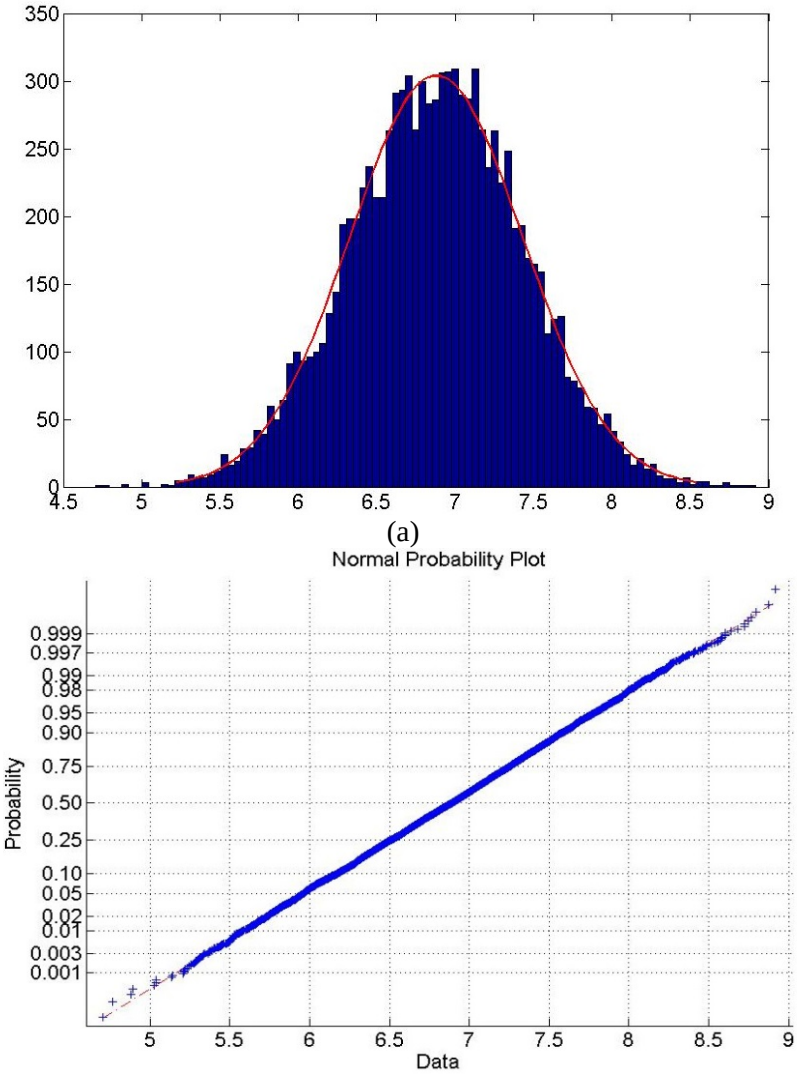
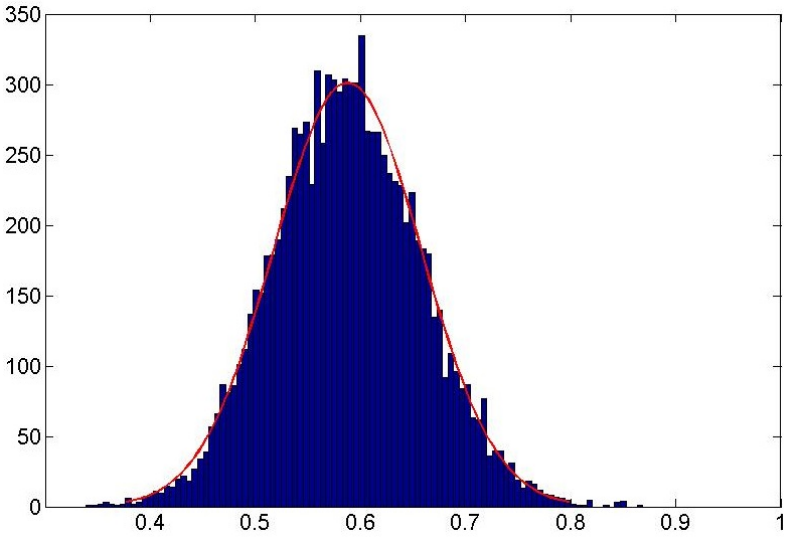
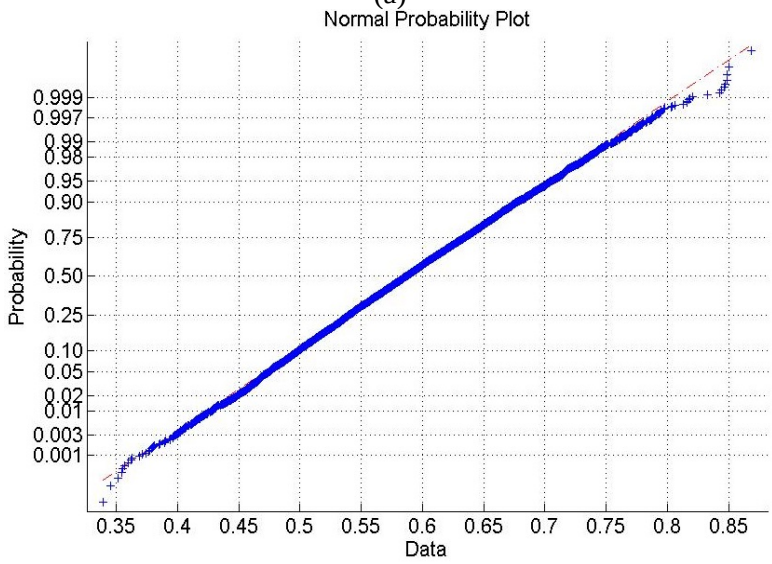


Figure (3-8): Current Magnitude of Branch 1, Phase B; (a) is histogram and (b) is normal probability plot.

If the data are normal, the plot will be close to the red line which indicates data extracted from normal distribution. The three previous figures indicate that the results are close to red line, thus, the results have this normality. Furthermore, we have N state ($x=[I_r, I_x]$) of branch currents, we need to check all states to ascertain the normality of BCSE method. However, we have the sample of 10,000 observations and 166 states for the whole system.



(a)



(b)

Figure (3-9): Current Magnitude of Branch 32, Phase C; (a) is histogram and (b) is normal probability plot.

Therefore, checking the normality of all states is an extremely huge work, thus, only the interesting quantities were chosen to define the normality of BCSE method. In addition to these normality check tests, other measures such as the third and fourth standardized moments (skewness and kurtosis) are considered for the earliest tests for normality [81]. Target values

for these measures are 0 and 3; respectively. These measures are easier to check and do not need to be checked visually. Figures indicate that the results are close to red line, thus, the results have the normality. For the bias test, both parametric and nonparametric tests are performed. Given the sample of 10,000 observations and 166 parameters, we can now use the Hotelling's T^2 test. We calculated the mean vector using MATLAB. The last output of the 'princomp' function in PCA toolbox for multivariate statistics is Hotelling's T^2 [88]. Since we found that T^2 value using this sample was 171.45 which is greater than $9,999/9,834 F_{166,9834}(0.05) = 1.2091$, we can reach to the conclusion of rejecting the null hypothesis since the significant level is 0.05. That is, the mean vector is significantly different than the alternative. It is worth mentioning that the p-value for this test is close to zero, which supported our conclusion. Here, we can conclude that the population mean vector is not equal to the true vector which had been observed before by J. Jung in his M.Sc. thesis [82].

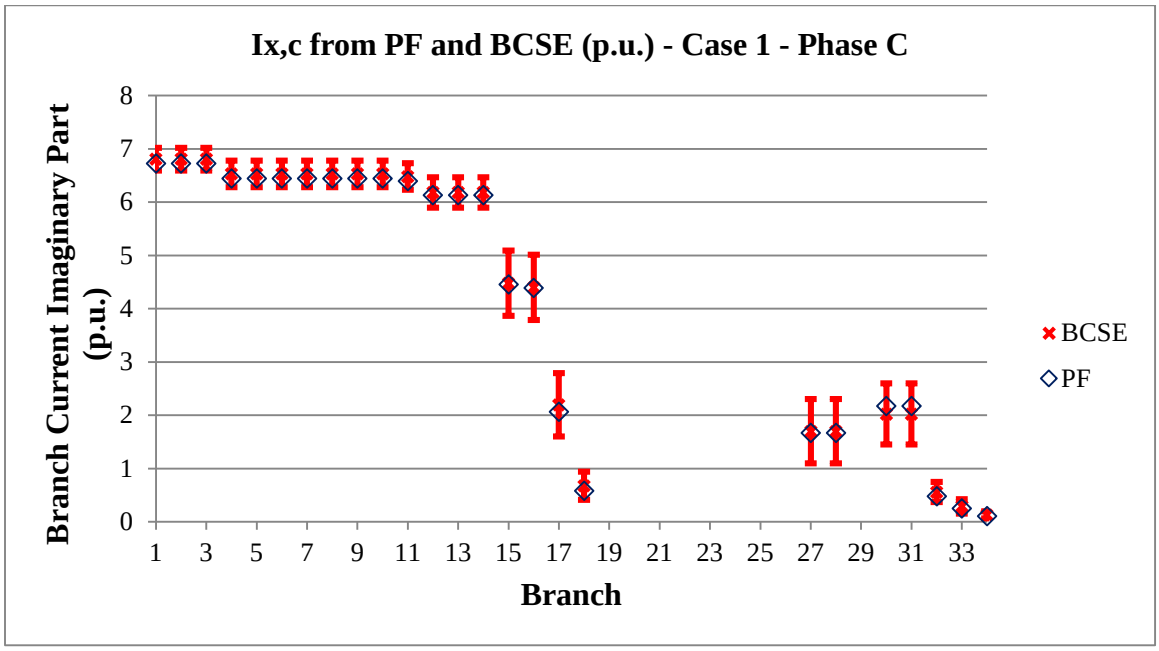


Figure (3-10): Estimated system states and true values, I_x , from BCSE and PF; respectively for phase B – Case 1.

Furthermore, Figure (3-10) shows the range of the SE output and the true value for phase B for imaginary part. More results can be found in the Appendix 1. It is clear that the means of the BCSE outputs do not fit with true values from PF. In fact, for some branches the mean values of the state estimation outputs is equal to the true values, however for some other branches these mean values are different from the true value. Having said this, in all cases the true values are always placed in the range of the BCSE output.

3-3-5 Consistency

If $\hat{x}_N$ is an estimator of x_t based on a random sample of N observations, $\hat{x}_N$ is consistent for x_t in this condition [42, 82, 84]:

$$\lim_{N\to\infty} P(|\hat{x}_N - x_t| < \varepsilon) = 1$$

Thus, consistency is a large-sample property, describing the limiting behavior of $\hat{x}_N$ as N tends to infinity. In other words, if the error in an estimate statistically corresponds to the corresponding covariance matrix then the estimate is said to be consistent. One measure of consistency is the normalized state error squared variable:

$$\varepsilon = (x_t - \hat{x}_N)^T \hat{P}_x^{-1} (x_t - \hat{x}_N)$$

where: $\hat{P}_x = E[(x_t - \hat{x})(x_t - \hat{x})^T]$ denotes the estimated state error covariance matrix. This value must be of a decreasing order by increasing the number of MC simulations.

Here, ε_N values have been depicted for different number of MC simulations. Figure (2-10) shows the average of ε_N, expected values for each set of simulations. This measure is calculated for Case 1 with 3% error on real-time measurements and 50% error on pseudo measurements. As it is shown in Figure (3-11)., this measure decreases by increasing the number of simulations.

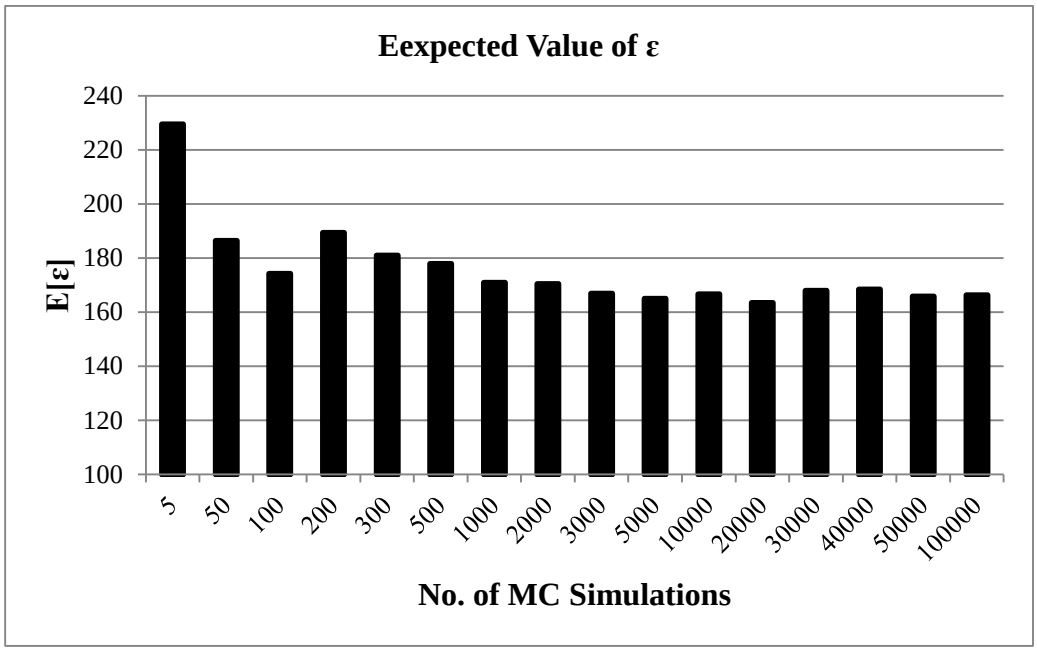


Figure (3-11): Average of ε_N for different MC simulations.

3-3-6 Quality

In the previous two measures, the mean of the estimate is the interesting quantity. On the other hand, quality measures the degree of its variance. In other words, if the variance is large, it corresponds to poor quality. Therefore, quality of an estimate is inversely related to its variance. For the multivariate case, the square root of the determinant of the error covariance matrix measures the total variance of an estimate. Hence, the quality of the estimate can be defined as [42, 82, 84]:

$$Q_{\det} = \log\left(\frac{1}{\sqrt{\det(P_x)}}\right)$$

where: P_x is the state error covariance matrix. Sometimes, in large networks, it becomes difficult to compute the determinant of the error covariance matrix numerically because of the

precision limits of the solver. In this situation, an alternate way to define the quality is to use the trace of the error covariance matrix. However, this ignores the off-diagonal information. The quality as function of the trace of the error covariance matrix can be written as:

$$Q_{trace} = \log(\frac{1}{tr(P_x)})$$

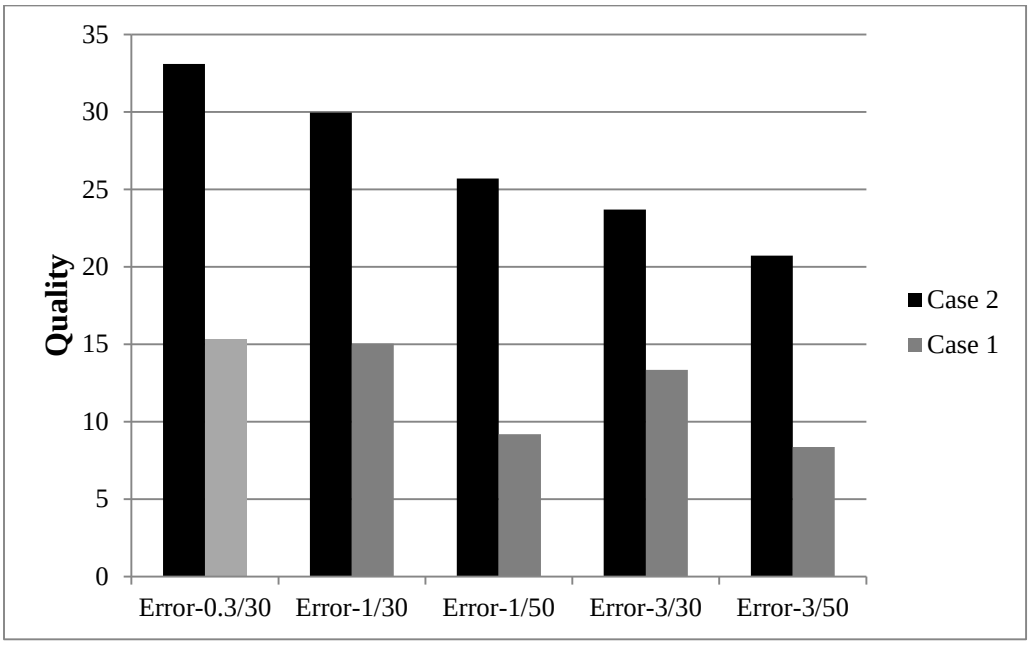


Figure (3-12): Quality for different measurement error scenarios in Cases 1 and 2.

The trace of P_x is the numerical sum of the variance of estimates. To study this measure for BCSE, Case 1 and Case 2 have been considered with different set of measurements. Case 1 only has power measurements at substation, as for Case 2 has power measurements in addition to three current measurements along the feeder. Here, Q_{trace} has been adopted for different scenarios which are depicted in Figure (3-12). Different scenarios have different load

estimation and real-time measurement errors which are 0.3%, 1%, and 3% as well as 30% and 50% error.

It can be deduced from Figure (3-12) that adding real-time measurements, ultimately increases the quality in different set of errors for real-time and pseudo measurements. In other words, when more real-time measurements are added to the system, the variance decreases in general. In the same situation (e.g. 1% error for real-time measurements and 30% error for load estimations), the average quality of the estimation was doubled from 15 to 30 when three current measurement were integrated to the measurement set.

3-4 Standard Deviation

Here, we want to calculate the standard deviation (variance) of the interesting quantities and system states. For VVC application, voltage standard deviation (variance) is the measure of voltage variation, i.e. this metric change during the load changes. It is thus recommended to calculate these standard deviations for the nodal voltages profile. One way is using mathematical models and another is performing Monte Carlo simulations. Actually MC simulations are utilized for verifying the calculations.

3-4-1 Formulation

In BCSE, branch currents have been considered for system state. On the other hand, for VVC, the estimate of the voltages and voltage violations are needed. Therefore, we need to have a method to estimate the voltages with their variances or standard deviations. A voltage at the node t of a radial feeder V_t is the voltage at the substation minus the voltage drop on the line sections between the substation and this node, hence, the measurement function for the voltage measurement V_t can be written in terms of the branch currents in the form:

$$V_t = V_s - \sum_{l \in \Omega_t} Z_l I_l$$

where: Ω_t is the branch set from substation to the bus t. The voltage magnitude measurements introduce coupling terms between the phases of branch currents and both the real and imaginary parts of branch currents. After solving the branch currents, a forward computation procedure can be used to find the node voltages of the circuit. Now, we write the function to relate the nodal voltage magnitudes to the system state:

$$y = f(x)$$

where: x is the system state vector y is the voltage magnitude vector and function $f(.)$ has been defined before. The level of accuracy of the estimated voltage must be determined by its variance calculation. For VVC application, voltage magnitude is an interesting quantity. So we can write the voltage deviation from a sending node to receiving one as:

$$\Delta V_i = V_i - V_{i-1} = Z_i I_i = \begin{bmatrix} r_{aa} + jx_{aa} & r_{ab} + jx_{ab} & r_{ac} + jx_{ac} \\ r_{ab} + jx_{ab} & r_{bb} + jx_{bb} & r_{bc} + jx_{bc} \\ r_{ac} + jx_{ac} & r_{bc} + jx_{bc} & r_{cc} + jx_{cc} \end{bmatrix} \begin{bmatrix} I_{r,a} + jI_{x,a} \\ I_{r,b} + jI_{x,b} \\ I_{r,c} + jI_{x,c} \end{bmatrix}$$

Now, voltage magnitude drop in one phase, e.g. phase a, at node i can be written:

$$\begin{aligned} \left|\Delta V_{i,a}\right| = [&(r_{aa}I_{r,a} - x_{aa}I_{x,a} + r_{ab}I_{r,b} - x_{ab}I_{x,b} + r_{ac}I_{r,c} - x_{ac}I_{x,c})^2 \\ &+ (r_{aa}I_{x,a} + x_{aa}I_{r,a} + r_{ab}I_{x,b} + x_{ab}I_{r,b} + r_{ac}I_{x,c} + x_{ac}I_{r,c})^2]^{1/2} \end{aligned}$$

Then, the covariance matrix of interesting quantities can be calculated from [33, 42]:

$$R_y = F(H^T R^{-1} H)^{-1} F^T$$

where: is H the Jacobian of the measurement functions, F is the Jacobian of the functions of interesting quantities, $f(.)$, and R is a diagonal covariance matrix containing the variances of the measurements, $W = R^{-1}$. Therefore, it is needed to calculate the partial differential of $\left|\Delta V_{i,a}\right|$ in regard of the real and imaginary part of the branch current at each phases, $I_{r,a}$, $I_{x,a}$, $I_{r,b}$, $I_{x,b}$, $I_{r,c}$ and $I_{x,c}$, here $\partial\left|\Delta V_{i,a}\right| / \partial I_{r,a}$ has been written:

$$\partial\left|\Delta V_{i,a}\right|/\partial I_{r,a} = \frac{1}{\left|\Delta V_{i,a}\right|}(r_{aa}^2 I_{r,a} + r_{aa}r_{ab}I_{r,b} - r_{aa}x_{ab}I_{x,b} + r_{aa}r_{ac}I_{r,c} - r_{aa}x_{ac}I_{x,c}$$
$$+ x_{aa}^2 I_{r,a} + x_{aa}r_{ab}I_{x,b} + x_{aa}x_{ab}I_{r,b} + r_{aa}r_{ac}I_{x,c} + x_{aa}x_{ac}I_{r,c})$$

Variances for each system variables have been calculated by the product of the BCSE, e.g. variance of the branch currents for phase a are the diagonal of this matrix: $(H_a^T R_a^{-1} H_a)^{-1} = (H_a^T W_a H_a)^{-1}$.

Standard deviation, σ_{zi} ,of measurement, z_i are calculated based on the considered accuracy, a_{zi} , for each type of measurements by:

$$\sigma_{zi} = \frac{z_i \times a_{zi}(\%)}{3\times 100}$$

For real-time measurements, corresponding accuracy, a_{zi} is about 1%-3% error and for pseudo measurements it is about 20%-50% error for simulations and calculations. In this study, zero injections, with a very low variance (10^{-8}), are modeled as the virtual measurements.

3-4-2 Test Results

In this section, we want to compare the calculated standard deviation from mathematical formulation to those from Monte Carlo simulations. Case 1 has been considered for this study with 1% error for real-time measurements and 30% error for pseudo measurements. For zero injection nodes, the associated variance is 10^{-8}. Following results of standard deviation of system state, branch current magnitude and voltage magnitude are shown in Figures (3-13) through (3-17). More results are shown in Appendix 2.

Based on the available results of calculated standard deviation for system states, branch current magnitudes, and voltage profile magnitudes from mathematical formulation are similar to those results obtained using Monte Carlo simulation. These comparisons that were made to

Monte Carlo simulations verify indeed the effectiveness of the approach. Any minor discrepancies are mainly due to the linearization of the equations in BCSE.

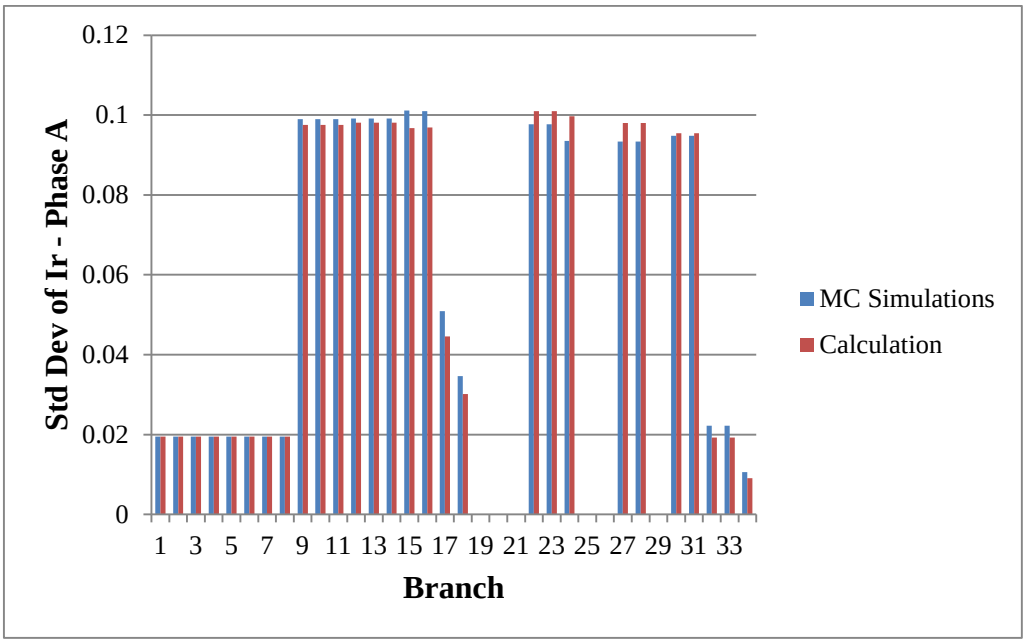


Figure (3-13): MC simulation vs Calculation for Ir at Phase A.

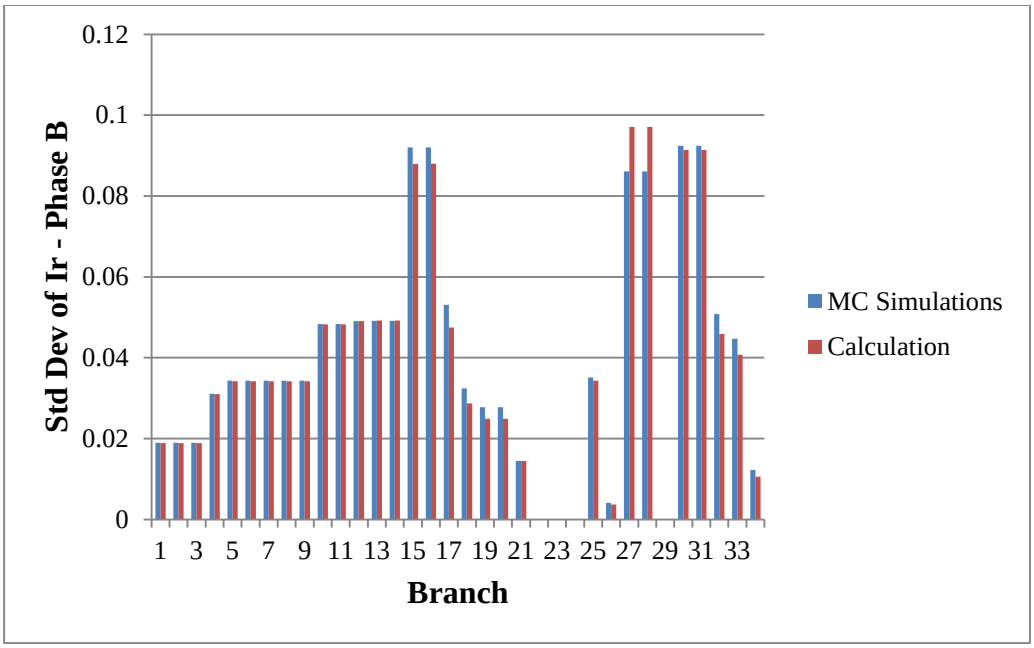


Figure (3-14): MC simulation vs Calculation for Ir at Phase B.

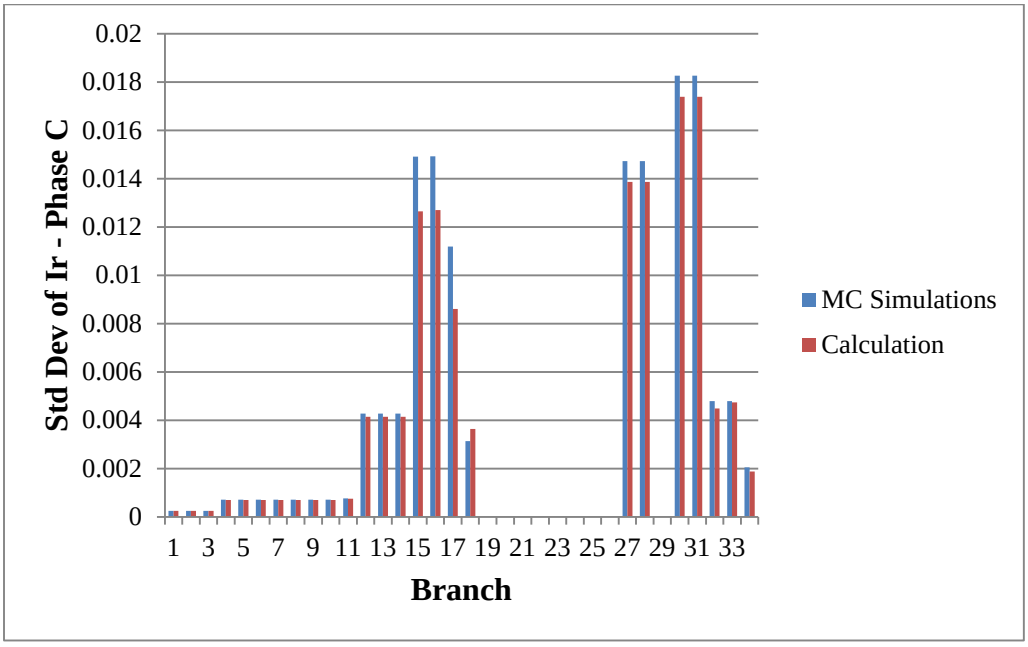


Figure (3-15): MC simulation vs Calculation for Ir at Phase C.

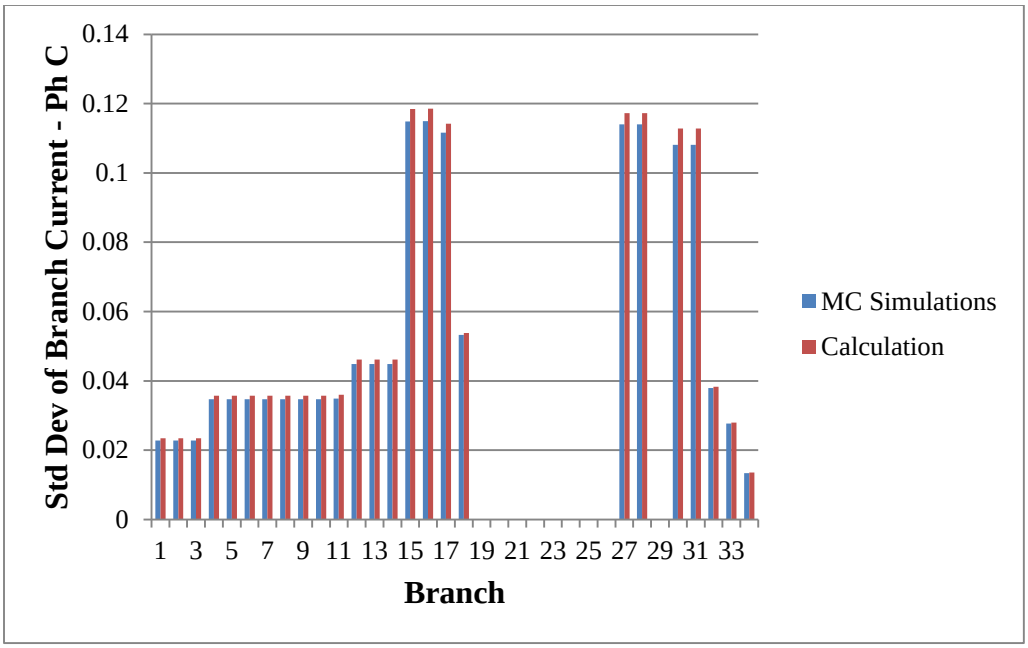


Figure (3-16): MC simulation vs Calculation for standard deviation of branch current at Phase C.

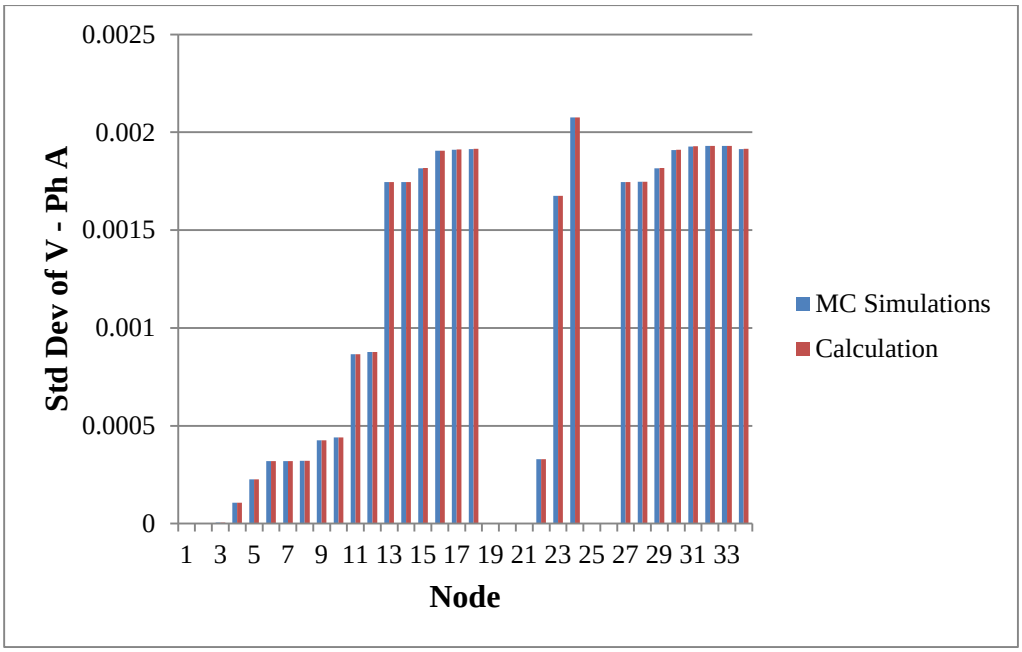


Figure (3-17): MC simulation vs Calculation for voltage magnitude at Phase A.

3-5 Summary and Conclusion

First some different methods for SE on distribution systems are introduced. Then, Branch Current bases SE method has chosen for further studies and analysis. Due to statistical characteristic of the load estimation, statistical techniques are adopted to assess the BCSE performance such as bias, consistency, and quality. For statistical analysis, 10,000 Monte Carlo simulations are performed. The initial tests indicate that the error in the estimated state ($x = [I_r, I_x]$) is not zero. Hence, we have adapted the bias test to test if this error was significant. Multivariate tests were performed to make statistical decision using BCSE and it was found that the population mean vector was significantly different from the actual mean vector from power flow. In addition, the overall performance of the state estimation was quantified in terms of bias, consistency, and quality. It was shown that adding more real-time measurements improved the quality of the SE output as well as lowered the associated error for pseudo measurements. For VVC application, finding voltage violation was necessary. Standard deviations of the voltage profile calculations were formulated mathematically as well as through Monte Carlo simulations. Both calculations were verified using the BCSE method.

Chapter 4: Load Estimation for Distribution Transformers by AMI Data

4-1 Overview

In this chapter, estimating the distribution transformer load by AMI data has been investigated. Then, impact of the load estimation error on the performance of the BCSE has been studied at last section of this chapter.

Load Estimation is an indispensable tool for distribution system studies, since knowledge of load profile along a feeder is the main data needed for system monitoring, analysis, and control [131-138]. Traditionally, it is considered that the loads on a distribution feeder usually follow quite predictable daily and seasonal cycles. Hence, loads at distribution level are not monitored, and most of the load estimation methods are aimed for estimating the peak loading conditions and load growth for planning purposes [139-152].

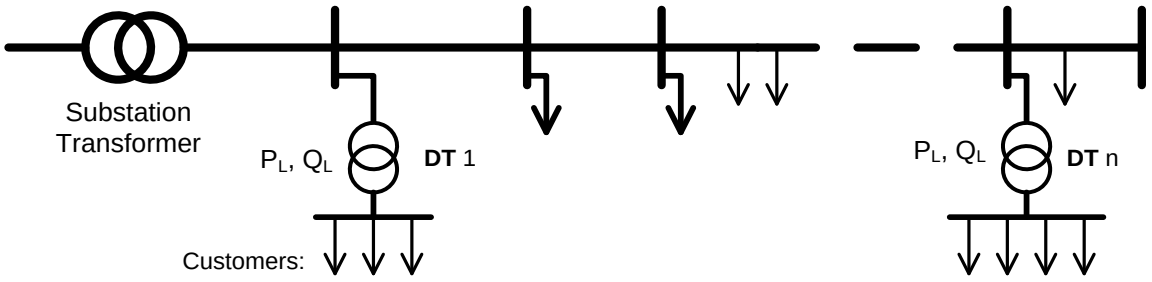


Figure (4-1): Model of a main feeder showing Distribution Transformers (DTs), loads and customers.

However, recent efforts for distribution automation, such as real time Volt/VAR Control for Conservation Voltage Reduction (CVR), requires real-time monitoring of the system [16, 133]. Another motivation for close monitoring of loads on a feeder is the integration of distributed generation, as they can alter the system loading considerably especially the intermittent energy resources [133, 134]. Furthermore, with the increase adoption of electric vehicles, distribution

systems are expected to have quite dynamic load profiles that need to be monitored in real time in order to provide proper control and protection of the system.

The diagram below represents a model of the main (primary) feeder of a distribution system. The main feeder provides a path for power to flow from the substation to the distribution transformers (DTs) which are shown in Figure (4-1). A typical distribution system may have several such feeders emanating radially from the substation. The feeder is divided into *n* line sections which are segments of the feeder that have the same conductor size, material, configuration, and number of phases. For our study, the goal is the estimation of the DT loads. Along each section there are branching points which are the connection nodes for distribution transformers and laterals which serve primaries of DT's. For our studies, when load estimation is called which means the estimation of the DT loads.

For analysis purposes, the load on a section can be considered as being uniformly distributed along the section (as shown in the diagram above), lumped at the beginning of every section, or lumped at the end. Secondaries and services distribute power to the consumers at utilization voltages. There may be several or few customers served from one DT which is shown in Figure (4-1). Here, loads of distribution transformers are considered fixed real and reactive load, P_L and Q_L. Most of the times, DT loads are estimated based on the rating of their transformers or from historical billing information of connected customers to that DT. Hence, these load estimations are associated with errors and uncertainties. In coming sections, we want to investigate more in modeling of these uncertainties.

4-2 Literature Review

The main challenge in real-time monitoring of loading conditions on a distribution feeder is the lack of measurements at the feeder level. This limits the direct adoption of state estimation at distribution level [39]. An alternative approach has been emerging, which is a two level approach; the first uses mainly historical data about the loads to estimate the loads on the feeder, while the second uses these estimates together with actual measurements to obtain an improved estimate using a state estimation [134], [39]. In this chapter this approach is assumed,

and thus, we hereby focus on developing methods for the short term estimation of loads, that are used to update the load estimates based on the available real-time measurement in 5-30 minutes intervals.

The major difficulties in the load modeling result from the random behavior of loads, diverse load shapes, limitation as well as the uncertainty in the information on loads. To achieve the goal of real-time monitoring and control, a distribution circuit state estimator tool, providing real-time estimates of the state (operating point) of the system, is required. Due to the limitation on the availability of the real-time measurements on distribution systems, a load modeling technique is necessary, to provide real-time estimates of customer load demands needed for state estimation [131]. In addition, utility engineers need to perform the analysis for feeder loading, feeder voltage profiles, feeder unbalance, and substation transformer loading all of which need to have a good estimate and model of the loads in the system.

Load forecasts can be classified into three categories: short-term forecasting which result usually from one hour to one week, medium forecasts which are usually from a week to a year, and long-term forecasts based on longer than one year data. The forecasts for different time horizons are important for different applications within a utility company [132, 151].

Moreover, load estimation can be utilized in off-line and on-line applications. Few methods have been developed for on-line applications using real-time data [134] and developed for distribution automation (DA). In [4], the load for next time interval has been estimated by utilizing the real-time data through advanced metering infrastructure (AMI) in the form of combined model of the Auto Regressive Integrated Moving Average (ARIMA) model and the load curve fitting with received real-time data to increase the level of estimation accuracy. On the other hand, most of the other methods have been utilized for off-line studies such as system planning and analysis using historical data [135-148], which will be reviewed in the next paragraphs.

A variety of methods including the so-called similar day approach, depend on finding the similar characteristics in historical data to estimate the load in the next day/next time step [136, 137]. Various regression models incorporate deterministic influences such as weather data, average load, and other related and available data, time series. These models are all based on

the existence of internal structure such as autocorrelation, trend, or seasonal variation [134], neural networks which are essentially non-linear circuits that have demonstrated the capability to perform non-linear curve fitting while considering the wavelet for data decomposition [135], statistical learning algorithms. This is done using different methods as well as trimming the data to achieve better estimation [139], fuzzy logic [140], and expert systems which incorporate rules and procedures used by human experts in the field of interest into software that is then able to automatically make forecasts without human intervention, have been developed for short-term forecasting [132].

A major drawback of traditional load modeling procedures has been their inability to provide a measure of uncertainty regarding their estimates. Lubkeman *et al.* proposed a probabilistic load modeling technique, based on daily load curves, illustrating the necessity of the time-of-day dependency [141, 146]. Time-of-day variation is incorporated by building daily load curves. Charytoniuk and Chen in [142] discussed the application of non-parametric probability density estimation to the customer demand forecasting problem using data available at utilities [142]. In their study, they used the demand survey information (energy data of a sample number of customers) and temperature conditions to build a probabilistic model, which denotes both the random nature of demand and its temperature dependence. The main input in their work was the energy usage and outside temperature. They were able to come up with the conclusion that the accuracy of forecast depends not only upon the quality of customer classification, size of sampling populations and composition but also on the size of the groups. In a similar study, Chen and his colleagues proposed a load survey system to determine the load characteristics of various classes, served by a utility company, followed by a statistical analysis on the acquired data to build a power consumption model of each class [139].

The modeling techniques, found in literature, extend from being simple approaches (based on assumptions such as load (power) which is directly proportional to kWhr consumption / transformer kVA ratings) to more statistically intensive oriented methods. Most of the current load modeling techniques rely on just the historical kWhr data or static transformer kVA and billing kWh, and the more sophisticated approaches for operation studies take advantage of statistical analysis techniques, power flow tools and available SCADA information [131-138].

Current load forecasting methods at distribution level are aimed for off-line studies and they use mainly historical data. A common approach, the so-called similar day approach, makes use of the fact that most of the loads follow a very similar daily load profile.

A few methods have been developed for on-line load monitoring using real-time data [132, 136, 137]. In [136], real-time load has been estimated by utilizing the real-time data through Advanced Metering Infrastructure (AMI). Recognizing the long latency of obtaining real time data from all the loads, which can be up to one day with the current AMI implementation, [137] proposes a load model which can be used to forecast the loads while waiting for the new measurement to become available. The load model is a combined model of the ARIMA and curve fitting.

This chapter proposes a load estimation method and its performance has been tested for the actual load data. This method can estimate the load profile on a feeder for real-time monitoring and control purposes, and thus the load estimation is aimed to be performed at a frequency of 15 minutes. The proposed method consists of two modules. Firstly, load clustering and load model, which are constructed based on the historical customer load data from AMI. Secondly, the load model is used to estimate the load based on the limited real-time data available from AMI. The method will be introduced in section II, the test results for actual load data are shown in Section III and finally the conclusions are provided in Section IV.

4-3 What is “AMI”?

AMI has been defined by FERC as the “metering system that records customer consumption [57, 58] hourly or more frequently and provides for daily or more frequent transmittal of measurements over a communication network to a central collection point.” The key concept reflected in this definition is that advanced metering involves more than a meter than can measure consumption (kWh) in frequent intervals. Advanced metering refers to the full measurement and collection system, and includes customer meters, communication networks, and data management systems. This full measurement and collection system is commonly referred to as AMI.

“AMI can provide additional value to utilities by enhancing customer service, reducing theft, improving load forecasting, monitoring power quality, managing outages, and supporting price-responsive demand response programs [57]”. There are communication networks such as broadband over power line, power line communications, fixed Radio Frequency (RF) networks, systems utilizing public networks, and etc. available to implement the AMI. Figure (4-2) depicts the building blocks of the advanced metering from FERC point of view [58, 164]. The frequency of collecting and transmitting smart meter measurements varies from utility to another. It is a common practice in the US that meters are currently to collect measurements with data intervals of an hour or less, and a data retrieval frequency of at least daily. In the USA, a survey conducted by the FERC showed that 20.2% of existing meters (average in 2012) had automated metering infrastructure (AMI). As this report mentions advanced metering systems can also open up new ways of monitoring voltage throughout an electric distribution system; which, besides improving billing applications can also improve operational control and efficiency.

This AMI brings the opportunity to estimate the customer loads more accurately and make more on-line applications of DA possible. This chapter aims to develop the model to enhance the quality of the load estimation by using more frequent AMI data.

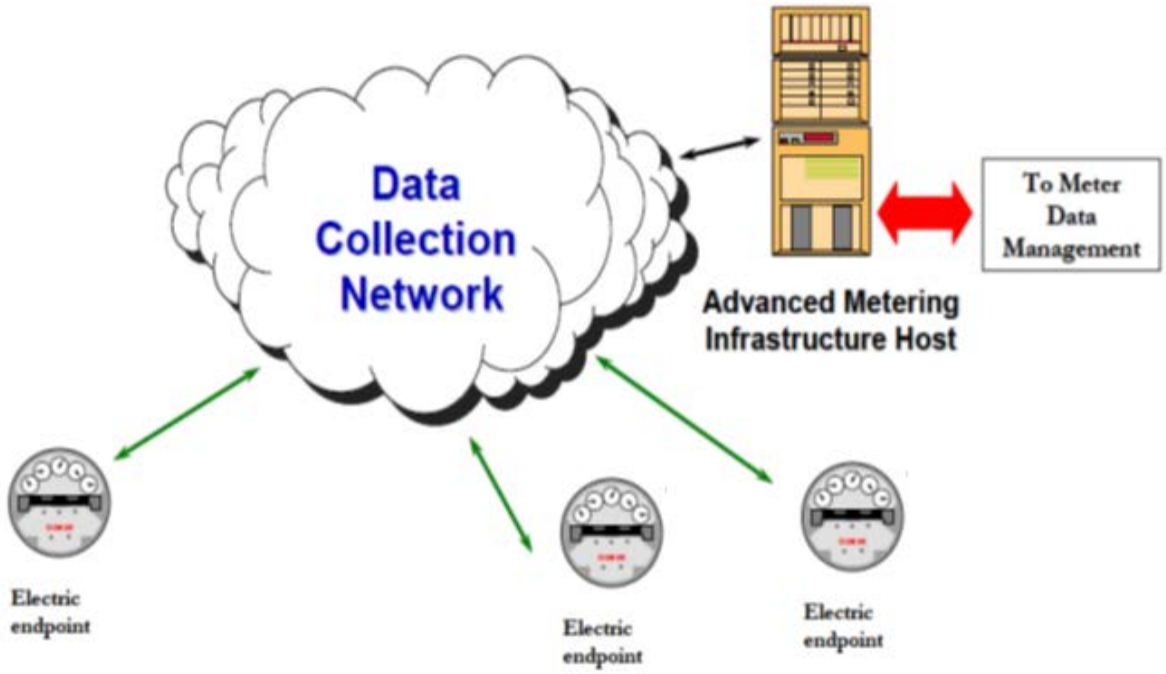


Figure (4-2): Building blocks of advanced metering from FERC, source: UtiliPoint International [57, 58].

4-4 Actual AMI Data Observation

The main data for real-time monitoring of loads is through AMI. In current AMI implementations [136, 152, 164], each AMI meter samples the load and monitors 15 minute interval demand, P_k, and stores it locally. Since there are in the order of thousands of meters on a feeder, each meter is polled typically once or twice a day to get the load profile stored on the meter. Hence, with the current AMI implementations, the data latency is quite long – up to a day or twice polling in a day. Thus the main challenge in load monitoring involves obtaining more frequent update and estimation of the load data.

To explore the real load profiles, Figure (4-3) shows the normalized load profile of three residential customers, i.e. LD676, LD700, and LD870, in one week during spring. Load profiles have been depicted in 15 minutes interval bases. These profiles illustrate the repetitive daily load behaviors in most of the days of the week. First observation from these data is the

periodic behavior of each load. This daily pattern of electricity consumption must be considered in the load modeling. In addition, the proposed method must exploit this periodic behavior of the load which occurs in daily behavior of electricity consumption. Therefore, to estimate the load effectively, the model should find these time series, i.e. periodic, based behaviors.

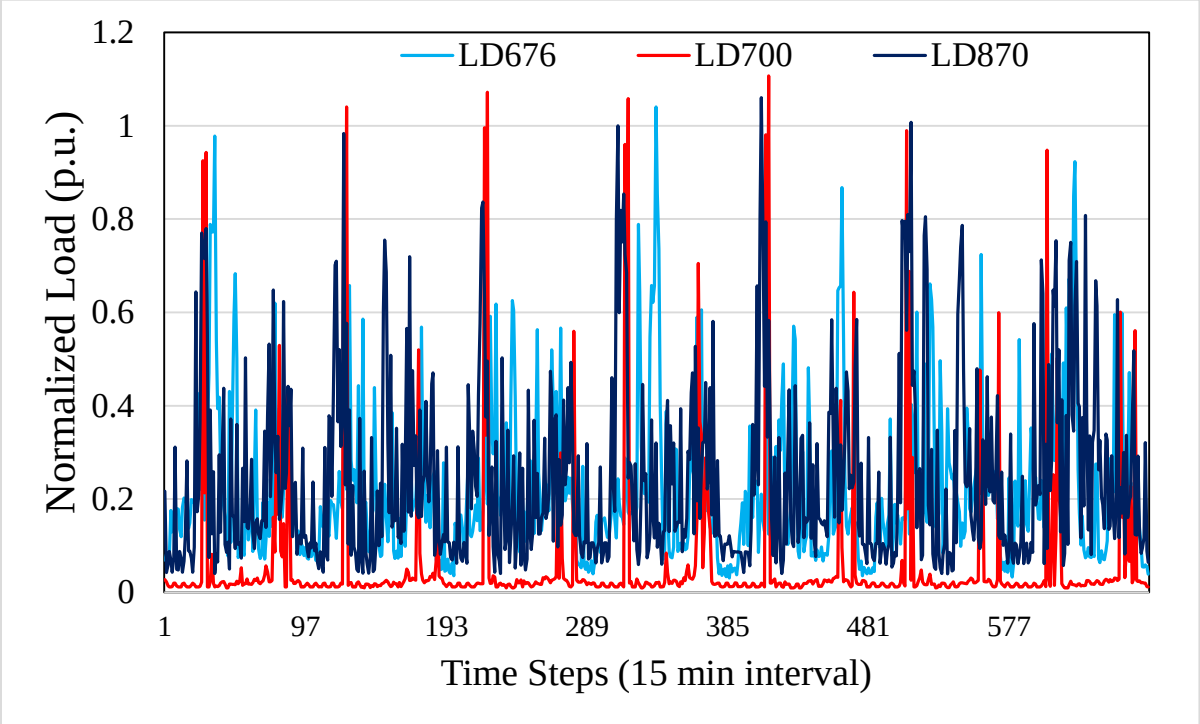


Figure (4-3): One week normalized load profiles for three customers: LD676, LD700, and LD870 [152].

Furthermore, the normalized load profile for LD700 shows a totally different pattern in comparison with other normalized load profiles, i.e. LD676 and LD870. LD700 has spiky load in the early morning at 7:00 am. One the other hand, the load profiles of LD676 and LD870 have some similarities.

Both of these loads show variation during the day and they have some peak loads in the morning as well as in the evening. When the peak value is compared with average load, it is clear that there is huge difference between the maximum and average load in LD700, but this ratio is not very high in LD676 and LD870. Based on this observation, electricity consumption of different customers can be either similar to each other or they can be different. Therefore, we need to set a tool to distinguish the different load patterns from each other. Here, the clustering method has been utilized to recognize the similar groups of the loads.

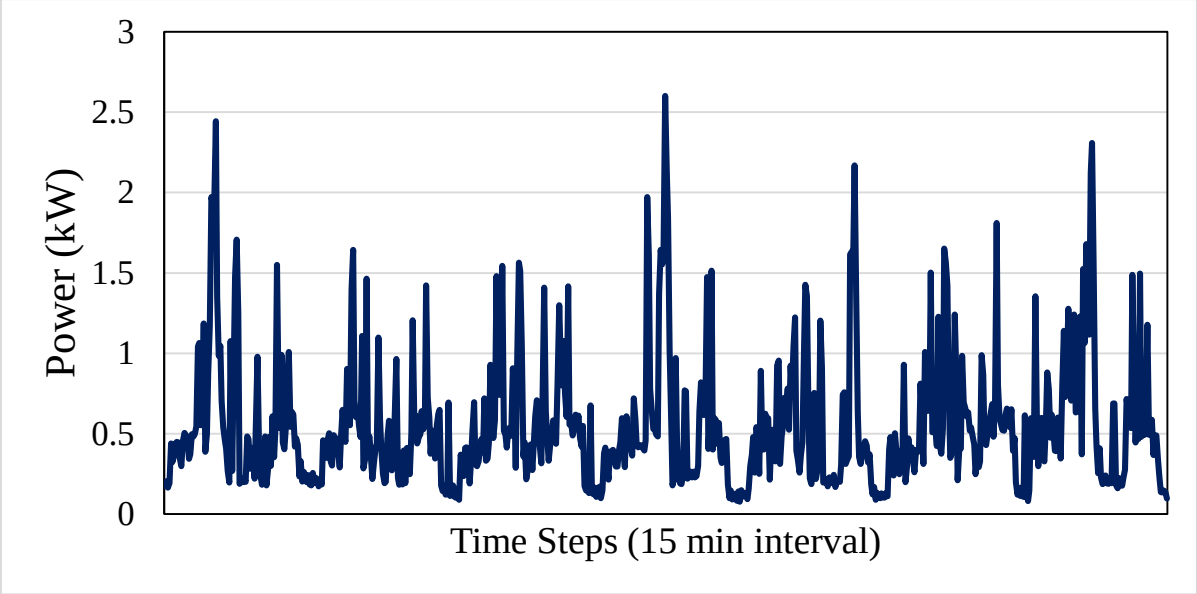


Figure (4-4): Actual load profile for LD676 for one week in spring [152].

Now, the time frame for load analysis has been changed from weekly to daily basis. Observing actual load profile during one day depicts how electricity consumption changes minute by minute in a very stochastic pattern. For instance, Figure (4-5) depicts the daily load profile of LD676, which shows the high variation of the load. For on-line applications, these high variations must be considered in the load modeling. Here, it is assumed that real-time data from

AMI enters the load estimation procedure and helps detect these load variation and by regression this new data can be modeled in the load modeling process.

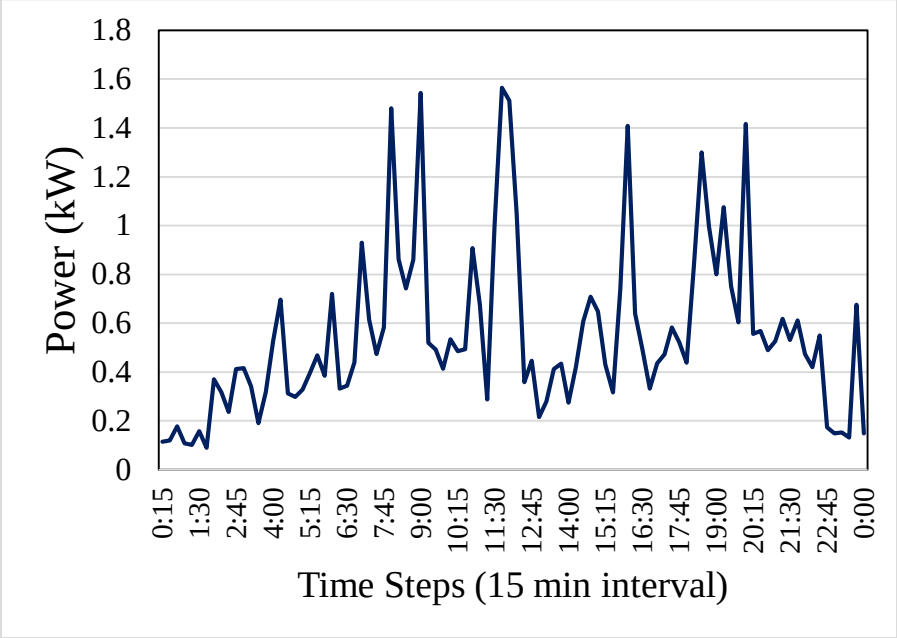


Figure (4-5): Actual load profile for LD676 for one day in spring.

4-5 Load Estimation Approach

To address these challenges for load estimation at Distribution Transformers (DT), the following scheme has been proposed. Step 1: Load Clustering by using historical load profiles, identifying the loads which have similar profiles and grouping them into clusters. Step 2: Load Estimation by polling only a subset of real-time data from each cluster at each update interval (5-30 minutes) and exploiting these real-time data to estimate the other loads which are not polled and increase the accuracy of the load estimation. Details about these two steps are given next.

4-5-1 Step 1: Load Clustering

Cluster analysis is a technique used for combining observations into various clusters such that; each cluster is homogenous with respect to certain characteristics, however, it should be different from other groups in regards to the same characteristics [156]. There are two general methods for clustering the data [154-160]:

1) Hierarchical algorithms find successive clusters using previously established clusters. These algorithms are usually either agglomerative ("bottom-up") or divisive ("top-down"). Agglomerative algorithms begin with each element as a separate cluster and merge them into successively larger clusters. Divisive algorithms begin with the whole set and proceed to divide it into successively smaller clusters.
2) Non-hierarchical clustering is partitioning of the sample. Each cluster has a seed point and all objects within a prescribed distance are included in that cluster. Another way of non-hierarchical clustering is to loop through the sample, assigning each case to the seed point to which it is closest.

Usually quite a few of the customers on a feeder have quite similar daily load patterns [132, 137, 142, 166], i.e. some of the load profiles are highly correlated. Therefore, we can use analytical tools for clustering to identify these loads.

All the load modeling methods reported in literature (chapter 3) and the method explained in chapter 3 are not enough to conduct overall analysis on the recorded data. The models obtained by implementing those techniques have limited ability to only perform curve-fitting for several groups of data. As it is unfeasible to install the load measurement units in each substation and it is unnecessary to construct a model for each load nodal point, 'load clustering' technique provides an effective approach to handle the aforementioned problem, hence increasing the credibility of the modeling technique.

In literature, Zalewski [140] applied fuzzy inference approach to cluster various substations and he then used fuzzy regression models to predict load consumption of the substation clusters. Wang and Li [160] explained a fuzzy approach to choose cluster centers (cluster means). The research conducted in [135] and [154-160] explains the k-means method of

clustering, the most widely used method in clustering analysis. Based on the similarities or distances (dissimilarities), objects are grouped together into groups (no assumptions on the number of groups). All the customers can be classified into several clusters, using a clustering technique, and in case of partial available AMI field data from the customers, data for the remaining customers in the clusters can be estimated using their historical data and real-time AMI data of those in their clusters.

K-means Clustering

K-means method of clustering is one of the most widely used techniques to find the homogenous groups of observations (which are load profiles in this application) [160-162]. In k-means clustering, the cluster centers are derived from the means of observations assigned to each cluster. Each iteration reduces the variation within the clusters and maximizes the difference between the distinct clusters until convergence is achieved.

One of the main challenges in k-means clustering is determining a-priori the number of clusters. Recently, statistics-based approaches have been developed for determining the cluster size [156-160]. In our study, we used the SAS Enterprise program for clustering customer load profiles. While there is no unique way to determine the optimum number of clusters, macro FASTCLUS in SAS uses some statistics (Cubic Clustering Criterion, pseudo F statistic, and pseudo T^2 statistic) to determine the optimum number of clusters [159]. These statistics are plotted against the number of clusters and the point where a jump occurs is selected as a good number of clusters. Different measures have been proposed in literature to find the optimal number of clustering and each measure quantifies one aspect of the clustering quality. To summarize the proposed methods and measures, these steps are proposed to find the optimal number of clusters for load estimation:

1) Run k-means clustering for each candidate number and calculate the clustering quality measures for each candidate. SAS recommendation is that the number of clusters cannot be greater than one-fifth of the number of observations (one-fifth rule) [156-160].
2) Look to matrix of measures for different number of clusters to find the optimal

number of clusters for a required application.

The following measures can be used to help quantify the goodness of clustering:

1. *Overall R-square*: It is the ratio of sum of the squares of distances between clusters to the total sum of the squares of all clusters [154]. R-square is a measure of the extent to which the clusters are homogeneous. It is the ratio of sum of the squares of distances between clusters (SS_b) to the total sum of the squares of all clusters (SS_t):

$$R^2 = \frac{SS_b}{SS_t} = \frac{SS_b}{SS_b + SS_w}$$

   where: SS_w is the sum of the squares of distances within clusters from the cluster centroid, which is a measure of the extent to which the clusters are homogeneous, for details please see [149].

2. *RMS-STD of the clusters*: The RMS Standard Deviation of a cluster i is:

$$RMS - STD_i = \sqrt{\frac{1}{p}\sum_{j=1}^{p} \hat{s}_j^2}$$

where p indicates the number of variables, and $\hat{s}_j^2$ is the variance of the j th variable. $RMS\text{-}STD_i$ gives an idea of how homogenous the cluster is. A smaller value suggests better homogeneity.

3. *Centroid distance between nearest clusters*: The overall distance between two nearest clusters (considering all variables) is a measure of the goodness of the clustering.

Visual inspection of data and the clustering results can give a good idea as to how good the clustering is, i.e. how far apart the clusters are, and also some idea about the homogeneity of the clusters, i.e. quality of the clustering [153], [159-162]. In the next section an example is given to illustrate the use of these metrics in assessing the quality of clustering.

4-5-2 Step 2: Load Estimation Models

Our goal here is to estimate the load for on-line applications of distribution systems, the proposed model considered the limited time frame of the load availability. From [153] and Figures (4-3), (4-4), (4-5) it brings the idea of observing the trend of changes occurring in a day and seasonal changes, an implemented model using harmonic components based on the concept of Fourier series. Fourier series is an infinite series whose terms are constants multiplied by sine and cosine functions and that can, if uniformly convergent, approximate a wide variety of functions. Therefore, the power consumption would be broken down into its harmonic parts. Based on the weakly and daily data available, similarly the load consumption of any random day, shows a distorted sinusoidal variation, which can be modeled by harmonics. Figure (4-4) illustrates the load profile for one residential customer in one week during the spring.

4-5-2-1 Load Model with Daily Harmonics

As illustrated in Figure (4-3) by the load profiles, the daily load profile are cyclic, and thus they can be regarded as a distorted sinusoid. Load profiles can then be modeled by using Fourier series [29]. In [132, 153], load has been modeled for yearly and daily variation. Here, because of the short-term application of the load estimation, the daily variation of the load has been considered. We can hypothesize the model with N_h harmonics for daily variation in the form of a general regression model as follows:

$$y(t) = \beta_0 + \sum_{i=1}^{N_h} \beta_i \cos(\frac{2\pi ti}{n}) + \sum_{j=1}^{N_h} \beta_j \sin(\frac{2\pi tj}{n}) + \varepsilon$$

where: y is power (response variable in the regressive model), and β_0, β_i, and β_j are unknown parameters for the harmonic components representing the daily variation, ε is a random error term, n is the number of samples in one day, and N_h is the number of harmonics considered for the model. To obtain the number of harmonics N_h needed for the model, four different cases (three, five, seven, and ten harmonics) have been investigated on one sample

data set. It was found that after five harmonics, the performance of the statistical model did not significantly improve [132, 153]. If the load profile is truly stationary (i.e., strictly periodic) then this model will be a good model and the error terms ε associated with different observations would be uncorrelated. To improve the model, we added a new term to take into account the correlation between errors - an autoregressive model R_t - to the model:

$$y(t) = \beta_0 + \sum_{i=1}^{N_h} \beta_i \cos(\frac{2\pi ti}{n}) + \sum_{j=1}^{N_h} \beta_j \sin(\frac{2\pi tj}{n}) + R_t$$

where: R_t captures the correlations between the errors by the following linear regression model:

$$R_t = \varphi_1 R_{t-1} + \varphi_2 R_{t-2} + \varepsilon$$

where: φ_1 is the first order lag, φ_2 is the second order lag and R_t is the residue. With this correction the new error terms ε became uncorrelated. Durbin-Watson test has been applied to find the order of lag for model construction. More details about the Durbin-Watson test can be found in [153].

4-5-2-2 Load Model with Real-time Data

Note that this model so far makes use of the periodic nature of the load variation from historical load data which are stored in AMI. If we have an actual measurement from another load with a similar profile (belongs to the same cluster) then we can incorporate this data in the model in order to further improve the model. For example, Figure (4-6) shows that DT1 has three customers from different clusters, i.e. blue, black and red arrows representing different load clusters. DT203 serves six different customers with three different clusters as well. In case of a similar cluster for LD1 and LD3 from a blue cluster, by having the real-time measurement from one of them, there is an opportunity to improve the load estimation at the green point for DT203. A similar pattern can be considered for LD5 and LD6 which are taken from a red cluster. In other words, a one real-time measurement from AMI for one of these loads from a

similar cluster will enhance the quality of the load estimation. Moreover, LD2 and LD4 are form other clusters that they can be estimated from historical data.

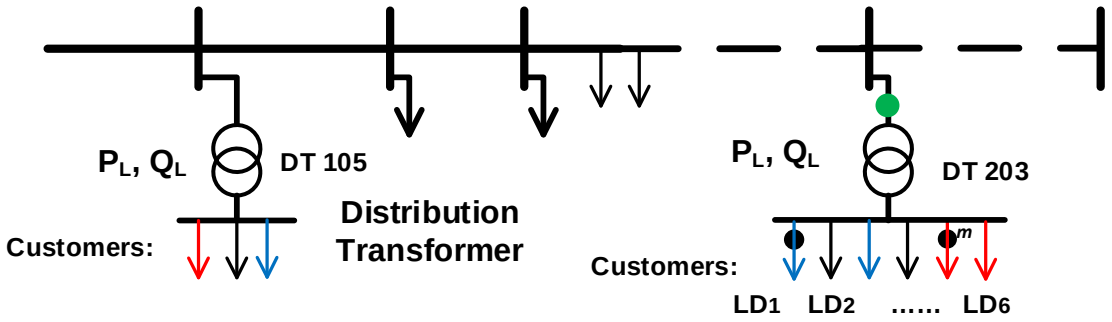


Figure (4-6): Distribution transformer load estimation from different load clusters and available measurements.

For addressing the availability of the real-time data, which is necessary for improving the load estimation quality, these data must be regressed into the harmonic-based load models developed in the previous section. Hence, this data can be modeled as another predictor and the modified model will be as follows:

$$y(t) = \beta_0 + \sum_{i=1}^{N_h} \beta_i \cos(\frac{2\pi ti}{n}) + \sum_{j=1}^{N_h} \beta_j \sin(\frac{2\pi tj}{n}) + \beta_{t0} P_t + R_t$$

where: P_t is the real-time AMI data from another load of the same cluster, and β_{t0} is the unknown parameter for this new predictor. In case of having more real-time measurements from AMI for different clusters, then we can incorporate these data in the model as shown below:

$$y(t) = \beta_0 + \sum_{i=1}^{N_h} \beta_i \cos(\frac{2\pi ti}{n}) + \sum_{j=1}^{N_h} \beta_j \sin(\frac{2\pi tj}{n}) + \sum_{k=1}^{N_C} \beta_{t,C_k} P_{t,C_k} + R_t$$

where: P_{t,C_k} captures the real-time AMI data from the cluster, i.e. C_k, N_C is the number of the clusters and β_{t,C_k} is the unknown parameter for this new predictor.

4-6 Test Results

To test the effectiveness of the proposed method, the 34 node IEEE test feeder has been considered [34]. The actual customer load profiles is used to present the load profiles on this test feeder. Figure (3) shows the one-line diagram of the feeder. In this feeder, it is assumed that the loads are monitored through AMI and we have historical load profile for each customer load. It is also assumed that AMI can provide demand interval data from a selected subset of loads at each update interval of 15 minutes [152]. In this study, these proposed two steps - load clustering and load modeling- have been applied to the actual load profiles.

4-6-1 Load Clustering

It is assumed that there are a total of 22 customers served in gray rectangular area of the feeder in Figure (4-7). The load profiles of each customer have been provided from AMI. These loads are numbered as: LD676, LD700, and etc. based on the actual numbering of the loads in AMI database.

For clustering, we have used a fixed time frame of load profile data as the historical load. This data is imported into the SAS Enterprise Miner and Macro FASTCLUS is used for clustering [149]. FASTCLUS performs “k-means algorithm” and disjoints cluster analysis on the basis of distances computed from the data. In this case, each sample consists of a load profile made up of 480 data points.

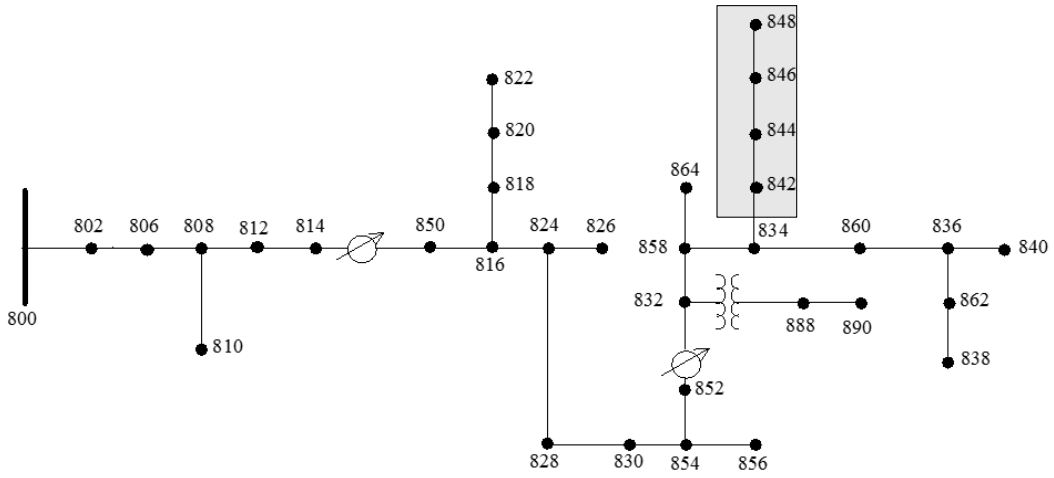


Figure (4-7): IEEE 34 Node Test Feeder (taken directly from [87]).

To determine the proper number of clusters for this data set, first per SAS guideline of one-fifth rule, the maximum number of clusters for these 22 loads is four. Hence, four cases are considered: five, four, three, and two clusters. Table (4-1) shows the main metrics obtained from running the clustering analysis in SAS Enterprise for these four cases. The metrics measured the quality of the clusters: m1: overall R-square, m2: RMS-STD of the clusters (min, max), and m3: centroid distance between the closest cluster. These results show that there is significant difference in statistical measures associated with these four cases. For instance, and as expected, the overall R-square for m1, has been increased by increasing the number of clusters. Moreover, the minimum RMS-STD has been reduced by increasing the number of clusters from 1.66 for two clusters to 1.37 for five clusters. In addition, the maximum centroid distances among the clusters has been increased with the number of clusters, i.e. more homogenous data have been grouped in similar clusters in comparison to the case with only two clusters. These studies show that more clustering provide more homogenous groups of load data, however, the optimal number of needed clustering is still not clear. In this case, the five clusters case can be removed from the list of the candidate number of the clusters by using the one-fifth rule. Then, we can focus more on cases of four, three, and two clusters. Measures:

m1, m2, and m3 have significantly been improved from two to three clusters, however, the improvement from three to four is not significant like the case from two to three clustering.

Table (4-1): Summary of Performance Metrics for Clustering.

Metrics	5 Clusters	4 Clusters	3 Clusters	2 Clusters
m1	-	0.259	0.171	0.076
m2	(1.37, 2.01)	(1.30, 1.97)	(1.68, 2.08)	(1.66, 1.95)
m3	(24.6, 63.6)	(37.1, 63.5)	(24.1, 40.9)	24.8

In the end of these analyses, there are two final candidates which have a similar performance. These candidates are the four and three clusters cases.

In addition to the above, clustering analysis can provide more information about the electricity consumption of the different customers in homogenous groups. To further investigate this fact, the average of each cluster for different number of clusters has been illustrated in Figure (4-8) where part (a), (b), part (c) represent the average of two, three, and four clusters cases; respectively.

It is clear from Figure (4-8 (a)) that in the case of two clusters, there is no distinguishable difference between the load patterns of customers. The three clusters case categorizes all of the customers in three groups, in this case we can distinguish more load characteristics in comparison with two clusters case. In following, four cluster case has been shown in Figure (4-8(c)) which demonstrates the different load behavior of this set of customers in a better way. Visually, there are some spiky loads among customers which are not identified well in the two clusters case, because they are covered and averaged by other loads. Otherwise, the four clusters case has distinguished this type of the load better in cluster 3, i.e. C3. There is another type of load which its peak takes place in the evening but not in the large comparison with spiky loads, this type is illustrated in C2 in the four cluster case. Another type of the load can be recognized in the C4 of the four clusters case which has a flat electricity consumption that changes smoothly during the hours of the day.

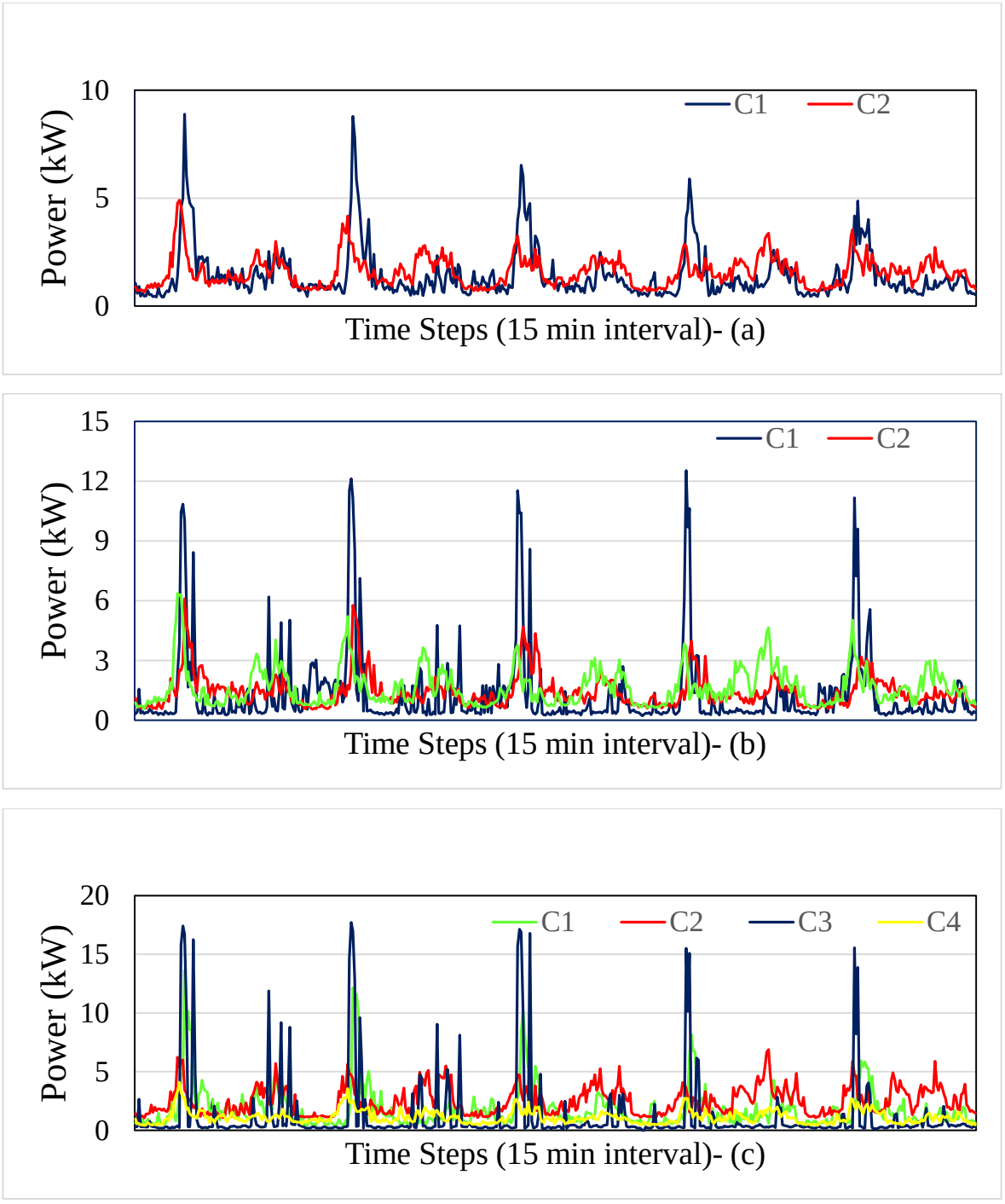


Figure (4-8): Average of clusters for 22 customers of two (a), three (b) and four clusters (c) cases.

Moreover, they are some loads whose peak times happens during the morning, i.e. when the people start their day by preparing their breakfast, taking showers, and etc. On the other hand, there are some loads whose peak times occur during the evening, i.e. when people come back from their work.

Observing the output of different number of clustering by the k-means method, it is clear that the customers can be categorized in these groups in average:

1- Morning peak customers: the maximum load of this group of customers happens in the morning, when people wake up to start their working days, e.g. C1 in k=2.

2- Evening peak customers: the maximum load of this group of customers happens in the evening, when this group of people come back to start their daily activities at home, e.g. C2 in k=4.

3- Both morning and evening peak customers: customers of this type have the peak electricity consumption during both mornings and evenings, e.g. C2 in k=2.

4- Odd customers: customers in this group have an odd behavior in electricity consumption which is different with other customers at these range of days, e.g. C3 in k=4.

This type of analysis will help the operator of the distribution feeder to find out the various patterns of electricity consumption along the distribution feeder.

4-6-2 Distribution Transformer (DT) Load Estimation

The second step of load estimation involves development of load models that can be used to estimate the loads which are not monitored. For this step, we have considered two possibilities, as noted that while the raw data we have from AMI is the customer load profile, the load we want to estimate for feeder level analysis is the total load at each distribution transformer, or even aggregation of a few transformers (for example serving a small subdivision of a feeder). For this case, one DT has been chosen for further study. This DT has been placed at node 848 on the test feeder with six customers from all of the 22 customers in the gray area of the feeder. These six customers are LD700, LD867, LD684, LD697, LD688, and LD841. We still have a choice to select the number of clusters. One way to estimate the DT load is by using historical data only, another way is by using historical data plus updated AMI data which is assumed

updates twice in a day. Since real-time measurements from AMI has become available, we can update the load estimation for the next time frame (next day or next 12 hours).

We used regression to fit the real-time measurement in load modeling which is explained in section two.

Table (4-2): ANOVA Table from SAS.

Coefficient	Estimate	Std-Error	t Value	p-value
Intercept	6.8588	0.1582	43.35	<0.0001
1st Harmonic S-D	1.9690	0.2237	8.80	<0.0001
2nd Harmonic S-D	-2.1529	0.2237	-9.62	<0.0001
3rd Harmonic S-D	-1.1142	0.2237	-4.98	<0.0001
4th Harmonic S-D	1.4129	0.2237	6.31	<0.0001
5th Harmonic S-D	-0.2700	0.2237	-1.21	0.2279
1st Harmonic C-D	-0.5248	0.2237	-2.35	0.0192
2nd Harmonic C-D	-2.7198	0.2237	-12.16	<0.0001
3rd Harmonic C-D	1.3980	0.2237	6.25	<0.0001
4th Harmonic C-D	0.6972	0.2237	3.12	0.0019
5th Harmonic C-D	-0.9425	0.2237	-4.21	<0.0001

* S-D: Sine-Daily, C-D: Cosine-Daily, Std-Error: Standard Error

For historical modeling only, there are eleven coefficients, one for intercept, five coefficients for the cosine part of the daily harmonics, and the other five coefficients are for the sine part of the daily harmonics. Analysis of Variance (ANOVA) test has been applied to this data set

and Table (4-2) shows the SAS output for estimated coefficients on 95% level of confidence [151-154].

The "p-value" obtained during the regression analysis in SAS dictate which components are to be included in the model. The "p-value" means the probability of getting a "t Value" greater than the threshold "t Value". Hence, with a 95% level of confidence, any variable with "p-value > .05" would fail to make a place in the model. The "*t-test*" used for testing the individual test of each coefficient in the model. These quantities are simply the parameter estimates divided by their standard errors. The column labeled "p-value", gives the p-values for the two-tailed test. For example, a p-value of 0.2279 for 5st harmonic seasonal component for Sin part from the ANOVA table (Table I) means that if we reject the null hypothesis that the coefficient value is zero, there is a 0.2279 chance of erroneous rejection. Hence, in this case, where the null hypothesis that all the coefficients are zero, and at the level of 95% confidence, the variables with "p-value > 0.05" should not be included in the model.

For example, the power consumption model of the DT with historical method, (estimated value: $\hat{y}(t)$), becomes:

$$\hat{y}(t) = 6.8588 + 1.969\sin(\frac{2\pi t}{96}) - 2.1529\sin(\frac{4\pi t}{96}) - 1.1142\sin(\frac{6\pi t}{96}) + 1.4129\sin(\frac{8\pi t}{96})$$
$$-0.5248\cos(\frac{2\pi t}{96}) - 2.7198\cos(\frac{4\pi t}{96}) + 1.398\cos(\frac{6\pi t}{96}) + 0.6972\cos(\frac{8\pi t}{96}) - 0.9425\cos(\frac{10\pi t}{96}$$
$$-0.6129R_{t-1} - 0.0495R_{t-2}$$

With AMI data from the four different clusters, the same ANOVA analysis has been conducted yielding the Table (4-3). After that, analysis of variance (ANOVA) has been applied for the selection of the model parameters within 95% level of confidence [154, 155], the coefficients with higher p-value in 0.05 are rejected to be considered in the model.

In this condition, the power consumption model of the DT with real-time data has augmented by elements of the clusters data from most recent AMI real-time data.

Table (4-3): ANOVA Table from SAS.

Coefficient	Estimate	Std-Error	t Value	p-value
Intercept	1.7495	0.2709	6.46	<0.0001
1st Harmonic S-D	0.5791	0.2549	2.27	0.0234
2nd Harmonic S-D	-0.2150	0.2513	-0.86	0.3924
3rd Harmonic S-D	-0.3653	0.2315	-1.58	0.1149
4th Harmonic S-D	-0.1370	0.2213	-0.62	0.5361
5th Harmonic S-D	0.1581	0.2028	0.78	0.4360
1st Harmonic C-D	-0.5617	0.2527	-2.22	0.0265
2nd Harmonic C-D	-1.4173	0.2617	-5.42	<0.0001
3rd Harmonic C-D	0.3262	0.2331	1.40	0.1621
4th Harmonic C-D	0.6528	0.2247	2.91	0.0038
5th Harmonic C-D	-0.1039	0.2040	-0.51	0.6107
$\hat{\beta}_{t,C1}$	0.9819	0.0246	39.97	<0.0001
$\hat{\beta}_{t,C2}$	0.0954	0.0519	1.84	0.0662
$\hat{\beta}_{t,C3}$	1.0396	0.0367	28.02	<0.0001
$\hat{\beta}_{t,C4}$	1.0256	0.0444	23.10	<0.0001

* S-D: Sine-Daily, C-D: Cosine-Daily, Std-Error: Standard Error

By ANOVA, the model of the power consumption has been constructed:

$$\hat{y}(t) = 1.7495 + 0.5791\sin(\frac{2\pi t}{96}) - 0.5617\cos(\frac{2\pi t}{96}) - 1.4173\cos(\frac{4\pi t}{96}) + 0.6528\cos(\frac{8\pi t}{96})$$
$$+0.9819P_{C1,t} + 1.0396P_{C3,t} + 1.0256P_{C4,t} - 0.659761R_{t-1} - 0.003016R_{t-2}$$

Statistical models for load estimation are now ready to predict the load profile for next day. The first model uses historical data only, the second model uses historical data with AMI data from two clusters, the third model uses historical data plus AMI data from three clusters, and the last method has been constructed by historical data with AMI data from four clusters.

Figure (4-9) illustrates the load estimation for the first 12 hours of the day (0:15 am – 12:00 am). Historical model (red dotted profile) in comparison with actual load profile (dark blue profile) follows the load change pattern along the day where the electricity consumption has been increased from 5:30 am and has reached to the peak at 7:00 am. This model cannot predict the peak load and it cannot predict the sudden load variation during the morning such as at 8:00 am and 10:00 am. On the other hand, real-time data from AMI, which was used in the other three models, helps to predict the load variation more accurately, such as the peak load estimation at 7:00 am which can provide 30 kW in comparison with 22 kW from historical model. Consecutively, Figure (4-10) shows the rest of the day load estimation from 12:00 am to 11:45 pm. As it is clear from this figure, the historical model (red dotted profile) illustrates the average load variation in comparison with actual load profile (dark blue profile) and it cannot predict any sudden variation. On the other hand, models with real-time data from AMI can predict the load variation in a better way, such as at 19:15 pm.

Visual checking shows that more clustering does not significantly improve the estimation. We should then quantize the load estimation error. Root Mean-Squared Error (RMSE) is a common measure of the differences between the predicted values interpreted by a model and those actual values [139, 164]. RMSE has been calculated here by:

$$RMSE = \sqrt{\frac{\sum_{t=1}^{n} (y(t) - \hat{y}(t))^2}{n}}$$

where: $y(t)$ is the actual value, $\hat{y}(t)$ is the estimated value, and n is the number of values which have been estimated.

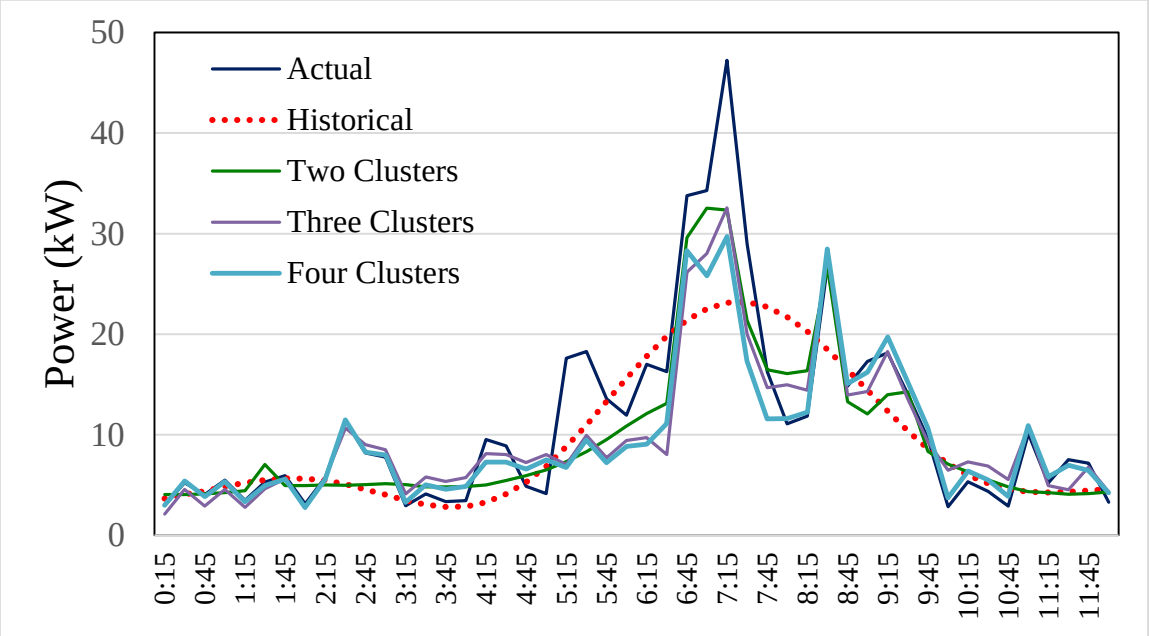


Figure (4-9): Comparison of different methods for DT 848 load estimation, first step (00:15 am- 12:00 pm).

Table (4-4): Clustering Assignment for Each Individual Load of The DT.

Loads	4 Clusters	3 Clusters	2 Clusters
LD700	C1	C1	C1
LD867	C2	C2	C2
LD684	C3	C2	C2
LD697	C4	C3	C2
LD688	C4	C3	C2
LD841	C4	C3	C2

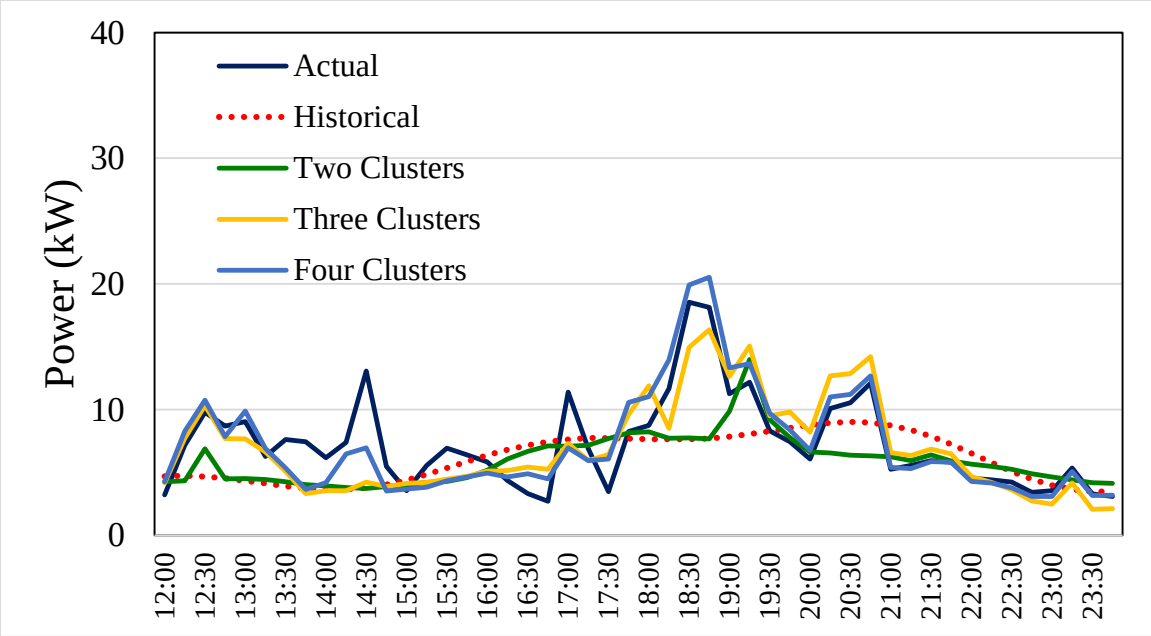


Figure (4-10): Comparison of different method for DT 848, second step (12:00 am – 23:45 pm).

RMSE for historical model is calculated to be 4.87. In addition, RMSEs for two, three, and four clusters models are 3.88, 3.41, and 3.38; respectively. Based on this measure, the prediction quality has been improved by having more real-time measurement from AMI. Three clusters model can predict the load behavior in a better manner than two cluster and historical models. As for four clusters model estimates the load for a day in better way than three clusters model. Consecutively, more AMI data from more clusters improve the load estimation quality significantly. Table (4-4) tabulates the apportionment of each individual load to each clustering case when these six loads have been categorized into two, three, and four clusters.

4-6-3 Performance Test

Five consecutive days have been chosen for further assessment of the proposed method. The historical window for model construction is two week and real-time load data has been polled from AMI twice a day. For this study, based on the result from the previous section, the three

cluster model has been chosen and the error of the load prediction has been calculated. Figure (4-11) illustrates the actual load estimation error (actual measurement-predicted value) for 480 data points. From Figure (4-11) it is shown that the maximum and average error of the estimation has been dropped significantly, solid dark blue line shows the actual error with historical data and the dashed red line shows the actual error with real-time data.

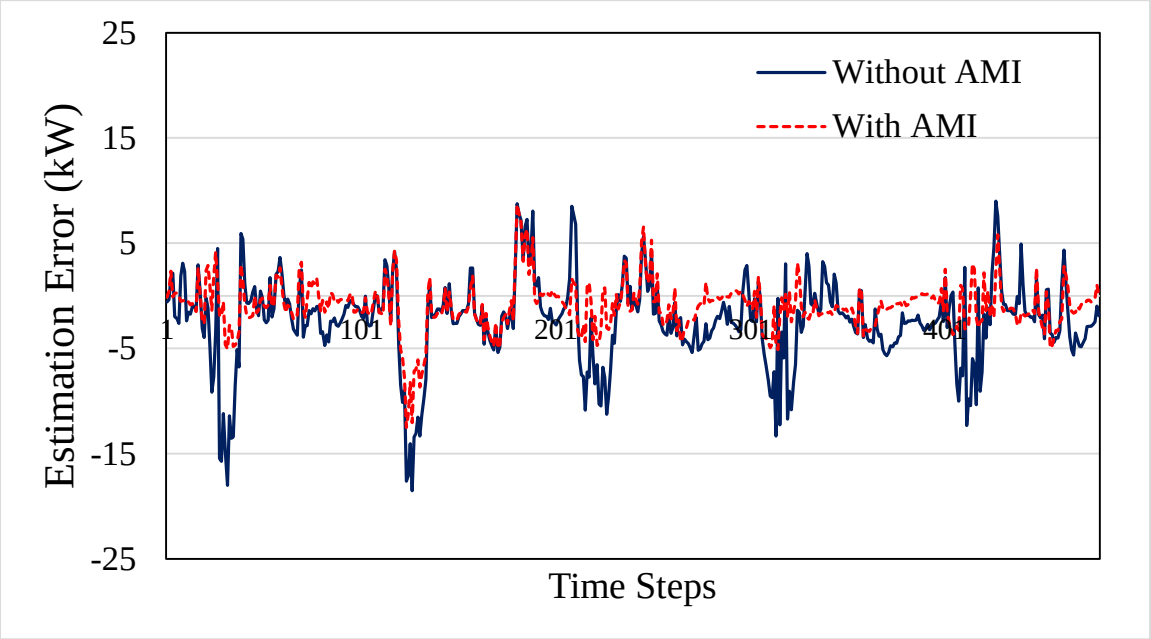


Figure (4-11): Actual estimation error (kW) for five consecutive days with and without AMI.

The averages of the error for with and without AMI are 2.45 and 1.03 kW. In addition, RMSE for these five days with and without real-time load data from AMI are 4.82 and 2.59, respectively. In other words, the accuracy of load estimation has been improved 46%. This result shows how real-time data from AMI improves the quality of the load estimation. In addition, the maximum error with AMI happens on Tuesday, second day, based on the investigation in customer load profiles, peak load happens both in the morning and the evening for this DT. Otherwise, the proposed model predicts the load within the range of the +/- 5 kW, in comparison of the load estimation error with historical data within the range of -20 to +15

kW, as it is shown in Figure (4-11).

In addition, Mean Absolute Percentage Error (MAPE) has been calculated to compare the performance of the proposed model [139, 164]. MAPE is a measure of accuracy of a method for constructing the prediction models in statistics. It usually expresses accuracy as a percentage, and is defined by the following formula:

$$MAPE = \frac{100\%}{n} \sum_{t=1}^{n} \left| \frac{y(t) - \hat{y}(t)}{y(t)} \right|$$

MAPEs for with and without AMI cases are 86.35% and 46.71%; respectively. This measure presents the 48.2% improvement in load estimation quality.

Up to now, the performance of the method has been assessed for one week, i.e. five consecutive days. Here, RMSE and MAPE for these five days are calculated and illustrated in Table 5. RMSEs for these five days are in this range: from 4.262 to 5.742 for the without AMI case and from 1.848 to 4.115 for the with AMI case.

Generally, the quality of the load estimation has been improved significantly by using the real-time data from AMI. In one special case, the proposed model cannot improve the estimation quality on Tuesday. As mentioned before, in this particular day the late evening peak happened, this was not possible to predict properly with this method. Likewise, MAPEs for these five days are in this range: from 69.84% to 104.6% for the historical data case and from 68.53% to 23.76% for the real-time data case. On Thursday, the proposed method has a significant performance by reducing the load estimation error from 104.6% to 28.64%. By this measure, i.e. MAPE, the proposed method does not have a significant performance on load estimation on Tuesday. The results of these measures: MAPE and RMSE are conforming to each other. The success of the proposed method can enhance the load estimation by adding more real-time data in the load models.

Figure (4-11) and Table (4-5) present real-time load data from AMI, the accuracy of load estimation has been improved by 48% in average, from 0.81 without AMI to 0.41 with AMI. All in all, based on these results, adding AMI data in the on-line load estimation process

improves the accuracy of the estimation at DTs.

Table (4-5): Daily RMSE and MAPE for with and without AMI cases.

Day	RMSE		MAPE	
	without AMI	with AMI	without AMI	with AMI
Monday	5.032	1.848	69.84	23.76
Tuesday	5.742	4.115	86.16	68.53
Wednesday	4.481	2.206	87.08	38.74
Thursday	4.415	2.002	104.6	28.64
Friday	4.262	2.082	96.70	44.27

4-7 Load Uncertainty Modeling for SE

The first step to analyze the impact of load estimation error on SE performance is load estimation error modeling. In this study, three different models have been implemented. In the first model, the error was proportional to the estimated load which had a fixed accuracy which has been used in lot of studies. The second model has considered constant load estimation error. The third focused on modeling the load estimation errors based on the number of the customers which were connected to one of the distribution transformer.

The main particularity of distribution system state estimation is the lack of real-time measurements. For the purpose of establishing state estimation function, pseudo-measurements need to be considered. As there is generalized uncertainty in the power demand, the load characteristics can be utilized in order to appropriately model the pseudo-measurements.

It is natural to model pseudo-measurements through normal distribution due to its compatibility to weighted least squares (WLS) estimation based on the maximum likelihood theory. In [98], Billinton *et al.* stated that “It is difficult to obtain sufficient historical data to determine the distribution type and the most common practice is to describe the epistemic

uncertainty by a normal distribution with a given standard deviation." In order to study the load estimation uncertainty effects on reliability indexes calculation, historical weather and load data both provide a good picture to estimate uncertainty in future loads. However, it does not depict the entire picture. Hence, Ghosh *et al.* in 1997 have investigated this issue further through load correlation coefficients using diversity factors and validated various models such as normal, log-normal, and beta distribution through testing the goodness of fit [23]. Seppala [99] has suggested log-normal distribution models which were verified from hourly load measurement data obtained from the Finnish load research project. In industry, PJM assigned normal distribution to model the load estimation uncertainty [100]. PJM claimed that 21-point normal curves to describe load forecast uncertainty was a reasonable and justifiable assumption, and they recommended what needed further investigation on using means and standard deviations.

In addition, it is a common practice in the distribution system SE literature [19, 22-37, 42, 47, 62-63] to consider normal distribution for load estimation errors. Here in our work, three different models for load estimation uncertainty have been investigated and proposed by considering the distribution networks nature.

4-7-1 Model 1: Proportional Error Approach

In this model, the estimated load is considered as the mean of the load distribution and standard deviation is calculated from empirical rule (three-sigma rule) by knowing the accuracy [33, 42]:

$$\sigma_{zi} = \frac{z_i \times a_{zi}(\%)}{3 \times 100}$$

where: z_i is the estimated load, a_{zi} is the accuracy in percentage (%), and σ_{zi} is standard deviation calculated from empirical rule. Empirical rule states that for a normal distribution, nearly all values lie within three standard deviations of the mean. This model assumes that the accuracy of the load estimation is fixed and standard deviation of load estimation is linearly

proportional to the value of the estimated load. In other words, the ratio of the load estimated and the standard is constant. The accuracy of the load estimation must be known. This factor has a significant impact on accuracy of the SE output. Case 1 was considered for this study with two different accuracies. The prototype feeder with M1 measurement in the beginning of the feeder is shown in Figure (2-2). As the results show, by increasing the level of load estimation error the standard deviation of voltage magnitude increase as well. Figure (4-12) show the changes of standard deviation of voltages along the network in two levels of load estimation error (i.e. 30% and 50%) for each phase (i.e. phases A, B, and C). Case 1 was chosen to study the variation of the load estimation error with constant real time measurement error (1%) at substation.

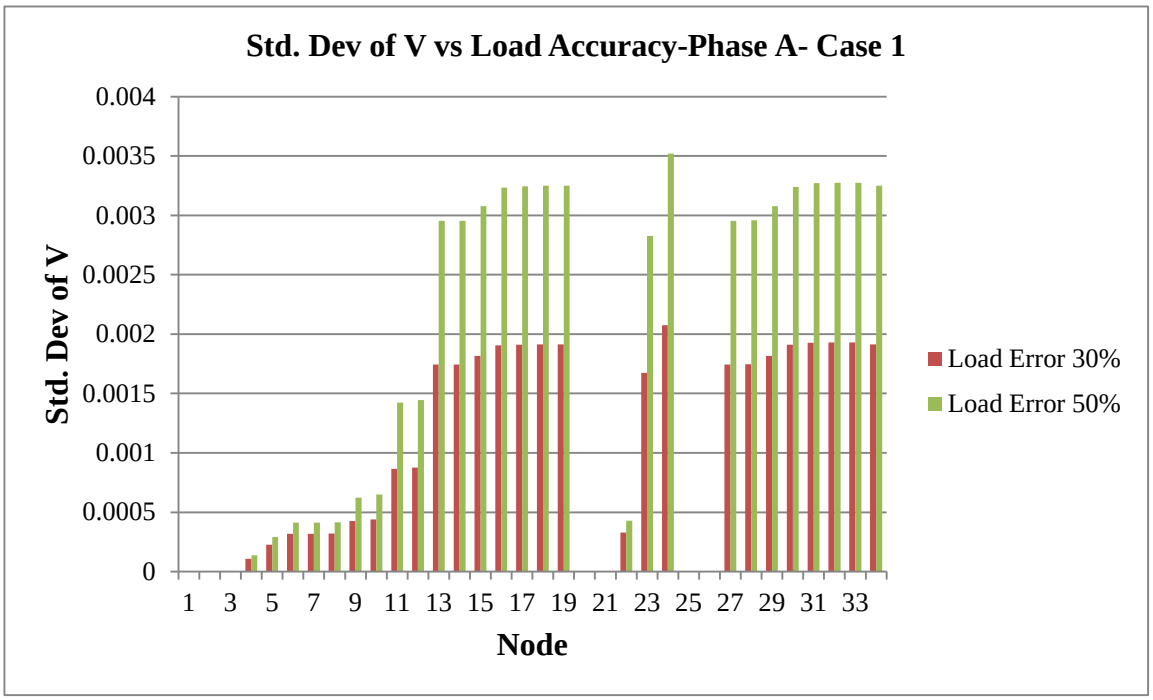


Figure (4-12): Standard deviation of voltage magnitude with different load accuracies (Model 1) for phase A.

From Figure (4-12), it can be deduced that the standard deviation of voltage magnitude increases from the substation (node 1) to end of the feeder (node 19). In addition, standard deviation of voltage magnitude along each lateral increases when it reaches to the end. For example, at phase A, standard deviation of voltage increases from 0.00043 (at node 22) to 0.00352 (at node 24), this lateral is just the one phase lateral. For further investigation, standard deviation of voltage magnitude increases from 0.00234 (at node 30) to 0.00237 (at node 33) for phase B, this lateral is the three-phase lateral at the end of the feeder.

In addition to the standard deviation of voltage magnitude, standard deviation of system state has been analyzed. Figure (4-13) shows the standard deviation of the real part of branch current at phase A. It also shows the correlation between high accuracy in system state and accuracy of the load estimation. It is clear from figure that the accuracy of the estimated interesting quantities is proportional to the accuracy of the system states. In other words, when system states are calculated more accurately, interesting quantities (e.g. voltage magnitude) are estimated in a better way.

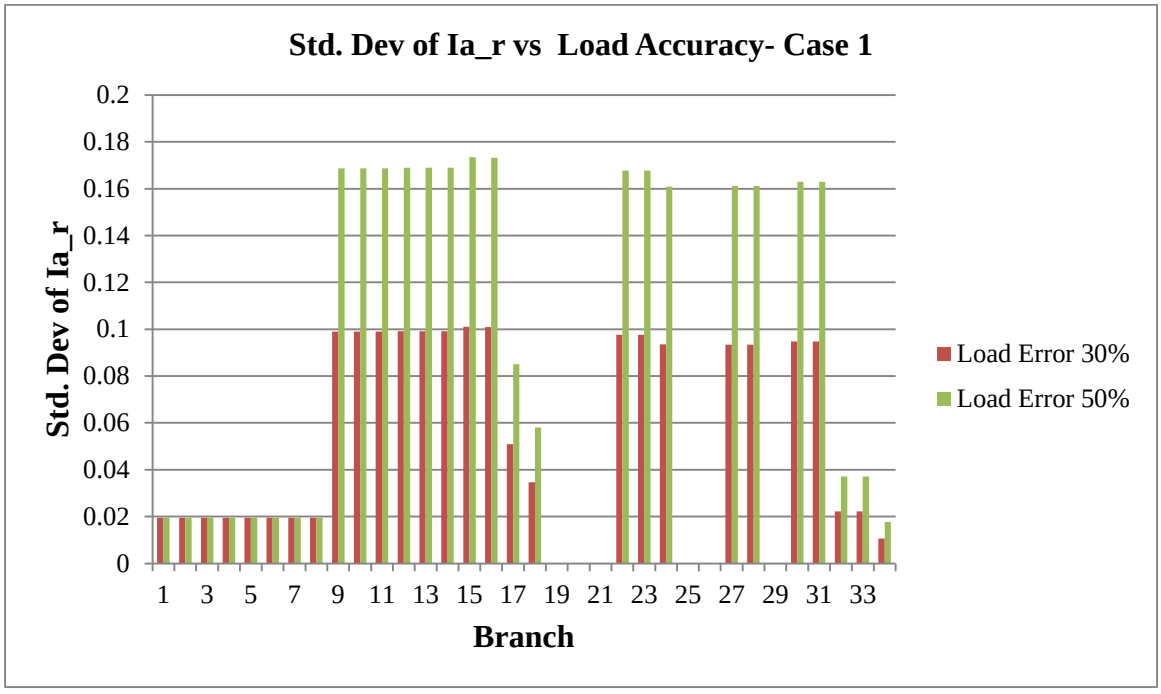


Figure (4-13): Standard deviation of real part of branch current with different load errors (Model 1) for phase A.

4-7-2 Model 2: Fixed Error Approach

Load estimation of DTs is totally related to this fact that the customers are connected to the DT or not. Furthermore, it is possible that the DT connection to the feeder has been lost and the DT is not in the service. Some other issues are mentioned too in [24]. In other words, the error of load estimation at distribution feeder is really high and it does not seem practical to consider the fixed accuracy for load estimation. In some conditions, one or a few customers are connected to one distribution transformer. Therefore, the error of the load estimation can be a fixed value and independent from the size of the transformer. To model this condition the load error has been considered with constant value:

$$\sigma_{zi} = \frac{Load\ Error}{3}$$

where: *Load Error* is the constant load estimation error for all kinds of load from small to big ones. Here, the distribution of this error has been considered as a normal distribution. For this approach standard deviation of the voltage magnitudes for phase A have been calculated and is shown in Figure (4-14) .In this part, the effect of the associated error for load estimation has been investigated. Case 1 has been chosen to study the variation of the load estimation error with constant 1% error for available measurements at the substation.

As it is expected, by increasing the level of load estimation error the standard deviation of voltage magnitude increase as well. By looking at the results, for two different level of load errors (i.e. 50 kW and 25 kW), it can be deduced that the standard deviation of voltage magnitude increases from the substation (node 1) to end of the feeder (node 19). In addition, standard deviation of voltage magnitude along each lateral increases when it reaches the end. For example, at phase A, standard deviation of voltage increases from 0.0031 (at node 22) to 0.0069 (at node 24) at 50 kW error in load estimation, this lateral is just the one phase lateral. When the error level decreases to 25 kW, the standard deviation of voltage magnitudes decreases to roughly half of the values at 50 kW error in load estimation. For further investigations, standard deviation of voltage magnitude increases from 0.00881 (at node 30)

to 0.00888 (at node 33) for phase B at 50 kW, and from 0.0044 (at node 30) to 0.0044 (at node 33) for phase B at 25 kW, this lateral is the three phase lateral at the end of the feeder. This study shows that the level of the load estimation has a significant effect on voltage estimation. As it was shown, by doubling the load error from 25kW to 50kW standard deviation of the voltage magnitude has also doubled. To further analyze, Figure (4-15) shows the standard deviation of the state system (e.g. real part of branch current in phase A). This implies that by doubling the load errors the standard deviation of the state system have also doubled similar to the voltage magnitude. For example, at node 28 for phase B, standard deviation of voltage magnitude at this node becomes 0.008 from 0.004.

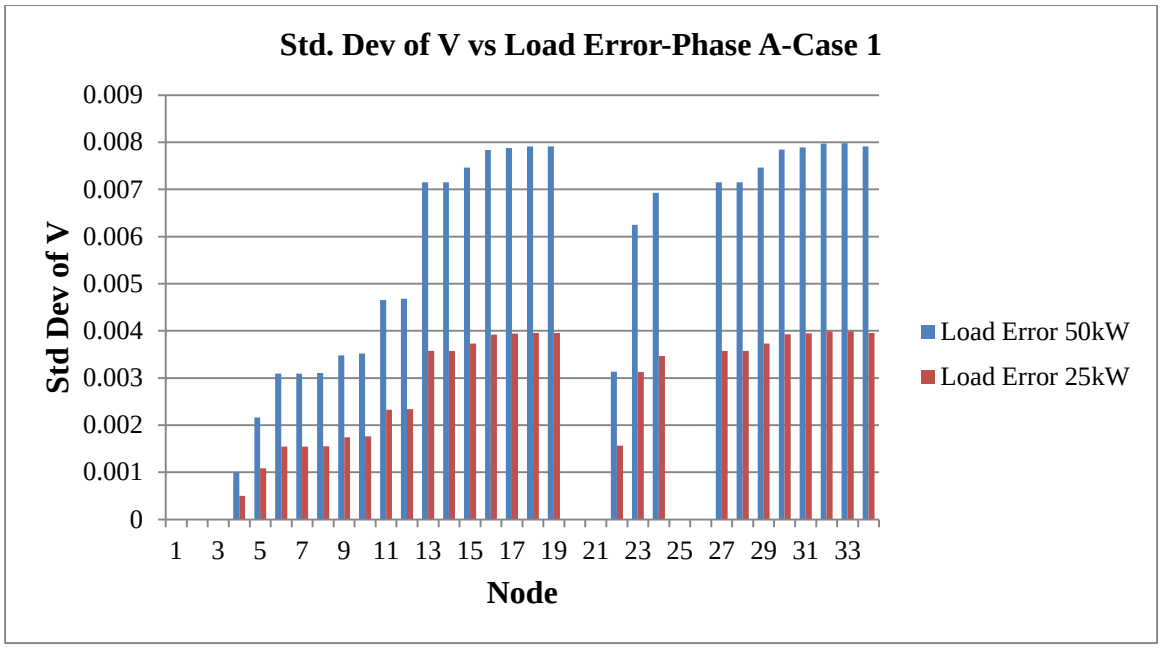


Figure (4-14): Standard deviation of voltage magnitude with different load errors (Model 2) for phase A.

Figure (4-15): Standard deviation of real part of branch current with different load errors (Model 2) for phase A.

4-7-3 Model 3: Customer Based Approach

As mentioned in model 2, load estimation for DTs is related to the number of customers which are connected to the distribution transformers. By considering this fact, a fixed load error for each customer has been considered. For example, average load for each customer is 25 kW and an associated error for each customer is 10 kW. This approach seems more practical when compared to other approaches due to the fact that load estimation uncertainty has been linked to number of customers at each transformer. In this model, normal distribution has been considered for error, e.g. 10 kW. Standard deviation of the estimated load, z_i, is calculated by:

$$\sigma_{zi} = \frac{Load\ Error_{per\ customer}}{3} \left\lfloor \frac{z_i}{Load_{\max}} \right\rfloor$$

where: $Load_{\max}$ is the maximum load of each customer, $Load\ Error_{per\ customer}$ is the associated error for each customer, $\lfloor\ \rfloor$ rounds the elements of bracket if greater than or equal to the nearest integers.

Standard deviations of voltage magnitudes as well as system states have been shown for two different conditions. One is with 10 kW load error for each customer and another is with 15 kW for each customer. It is clear from Figure (4-16) that when the load error for each customer increases from 10 kW to 15kW, all voltage profile standard deviation have increased.

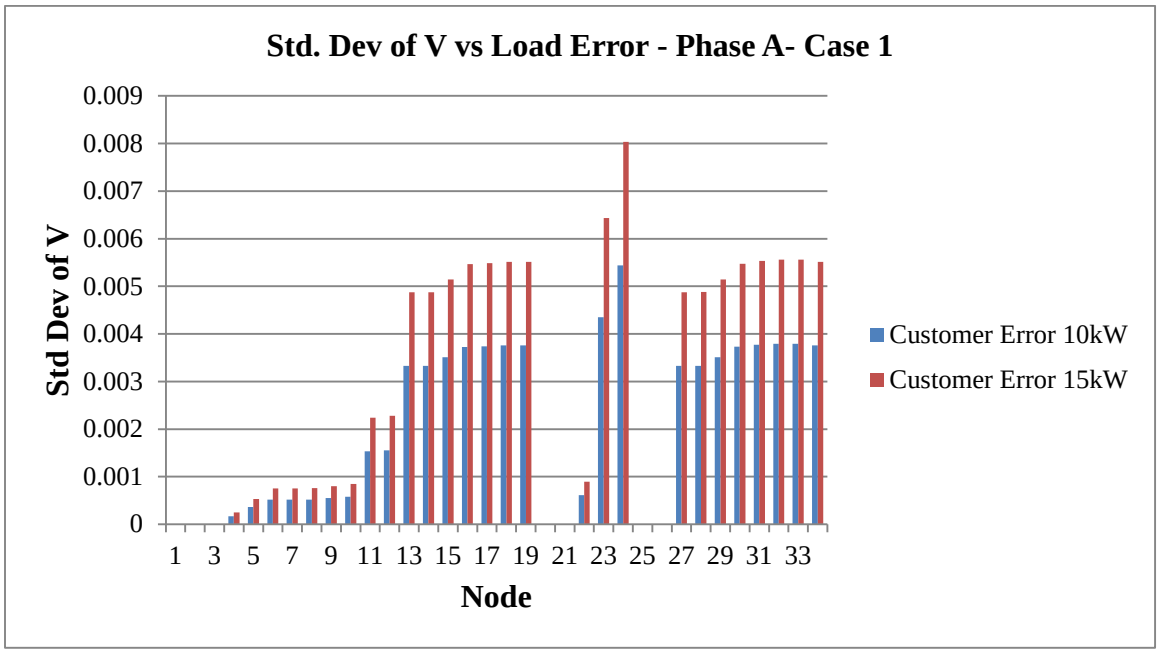


Figure (4-16): Standard deviation of voltage magnitude with different customer load errors (Model 3) for phase A.

4-7-4 Comparison of Models

Same pattern for voltage standard deviation was observed for different load uncertainty modeling. However, the levels of the standard deviation values are different for each model. Model 2 (fixed load error) has the highest level of values in voltage standard deviations in compare with other models. On the other hand, load estimation errors were limited to the

accuracy factor in model 1; therefore model 1 provided the most conservative estimation. As for model 3, it was developed to consider the realistic situations of the distribution feeders by adding the number of customers connected to the distribution transformers. This model provided more practical estimation for standard deviation of voltage magnitudes. The problem of the model 3 is that there is now information about the number of connected customers to one DT and it is only based on the assumptions. Voltage standard deviations of Phases A and B for the three models are shown in Figures (4-17) and (4-18).

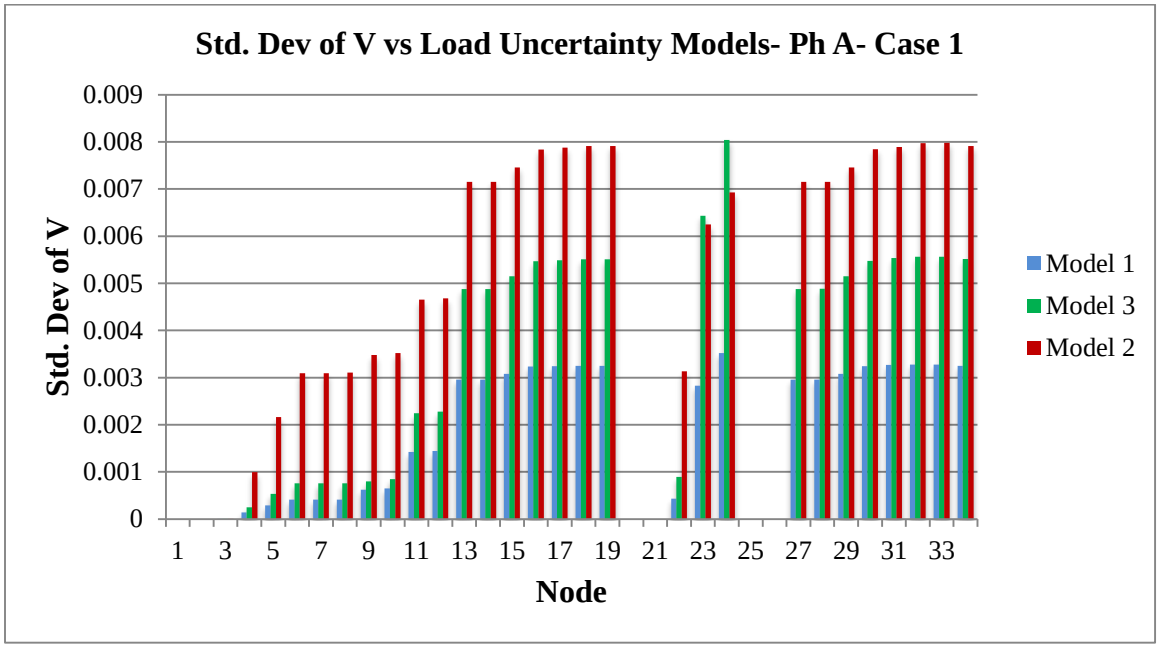


Figure (4-17): Standard deviation of voltage magnitude for the different Models 1-3, for phase A.

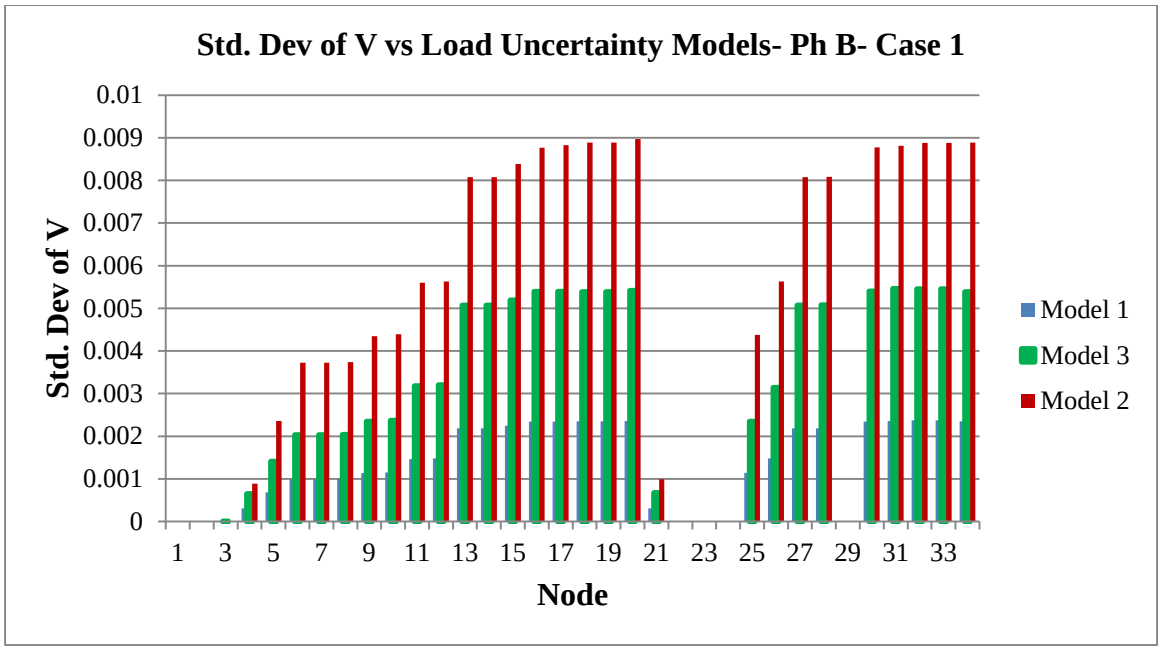


Figure (4-18): Standard deviation of voltage magnitude for the different Models 1-3, for phase B.

It is clear from figures that voltage standard deviation increases in the beginning of the feeder to the end of the feeder, this pattern consequently happens in the laterals (one phase or three phases). At the main feeder, when the loads or laterals are added to the circuit, standard deviation of voltages increases as well. For example, some step changes can be seen at nodes 9, 13, and 15. For future studies, we have chosen to concentrate on model 2 due to the fact that implementing of this model is easier and showing the voltage standard deviation changes clearly. Furthermore, this model considers the nature of uncertainty of DT load which is needed for future studies.

4-8 Summary and Conclusion

A new statistical approach to model and predict the power consumption using an auto-regressive model, based on harmonic decomposition of the power consumption, is proposed

and implemented in this chapter. To build a model capable of predicting power consumption, one parameter-power is used. In addition, a clustering technique was applied to cluster customers into homogenous groups and utilize these results to improve the accuracy of estimation by AMI data. Following, DT loads have been estimated by historical and real-time data. As clear from the results, using real-time data of customers from AMI along with the available historical data helps to increase the accuracy of estimation of the load. This can provide the one venue to utilize the AMI data in the operation of power distribution systems. As shown, using this modified load estimation method by taking advantage of AMI data provides the possibility to implement more advanced DA applications which require high accuracy load estimation. As indicated earlier, the need for real-time load data on distribution system is mainly dictated by two applications: real-time Volt/VAR Control (VVC), and energy management of Distributed Energy Resources (DERs). VVC for conservation voltage reduction requires the most frequent update, as it requires accurate estimation node voltages all the time. For DER energy management, the estimation is needed for up to a day ahead. It is worth to mention, there is a remarkable potential to improve and enhance the load estimation at DT level on distribution systems.

In the section 4-7, impact of the load estimation accuracy on the BCSE performance has been analyzed by using different models for load estimation errors. Results show that more accurate load estimation provides better input for the SE to estimate the nodal voltages along the feeder accurately.

Chapter 5: Meter Placement on Distribution Feeders for Volt/VAR Control

5-1 Overview

This chapter focuses on the meter placement problem on the distribution feeders. Available literature about this topic has been investigated after which the problem has been formulated for Volt/VAR Control (VVC) application. One heuristic approach for solving the problem will be proposed in what follows based on the extensive observations in two steps and finally its effectiveness has been assessed by another approach.

The performance of the SE depends on these factors:

- Real-Time measurements:
    - Number of measurements and their location
    - Type of the measurements, i.e. voltage, current, and power measurements
    - Metering accuracy
- Accuracy of load estimation

The main goal of meter placement in distribution systems becomes supplementing the forecasted load data with real-time measurements such that the SE with these measurements will satisfy the performance requirements. Therefore, the meter placement problem for VVC is:

Determine the number, place, and type of meters that needs to be placed on a given feeder such that the SE with these measurements can estimate the voltages with desired accuracy and cost effectively.

After developing a set of guidelines to place the initial measurements which are reasonably small and are redundant enough to provide the desired level of accuracy, different schemes have been tailored to reduce the number of measurements. The search scheme has been implemented to identify the minimal set of the meters needed to estimate the node voltage with

a desired accuracy. In addition, this search scheme is flexible to incorporate of different metering options and robustness measures.

5-2 Literature Review

Meter placement is a complex problem, not only due to the size of problem (number of choices available), but also often due to the conflicting requirements between the SE performance and the cost of the measurement system necessary to achieve the desired performance. The problem has been considered widely in the literature for transmission systems into two main categories: 1. observing the network and 2. minimizing the error of estimation. Recently, by utilizing the accurate measurements, Phasor Measurement Units (PMUs), the placement of these measurement devices at transmission level has been studied extensively [15, 21, 47].

On the contrary, the number of meters that has been placed on distribution feeders; usually below the minimum needed for state estimation (i.e. the system won't be observable by using the actual meters alone). To overcome this observability problem, forecasted load data needs to be added as pseudo-measurements to the measurement set. Therefore, the main goal of meter placement in distribution systems becomes supplementing the forecasted load data with real-time measurements such that the SE with these measurements will satisfy the performance requirements for different applications such as load estimation, voltage profile estimation, monitoring of the network and state estimation.

For monitoring of distribution feeders, Baran *et al.* [33] proposed a rule based meter placement scheme with three specific rules from different observations. If these three rules alone are followed, a large number of meters are placed on the system. To overcome this difficulty, Koglin's method [33] is utilized to eliminate the measurements with low impact on the accuracy of the quantities of interest. To consider the accuracy of the interesting quantities, a system accuracy index was defined as the sum of variances of the interesting quantities, i.e. some branch current variances. Goal of this work is to place the meters on the feeder to monitor of switches through SE by focusing on current magnitudes changes. Following to this work, Schulz *et al.* [32] have found three valuable observations to place the different meters on the

feeders by running three test cases. Branch power, current magnitude, and voltage measurements were considered in this research. Firstly, they found out power and current measurements could both lead to better results in comparison to voltage measurements. Secondly, it was found that the BCSE would have better performance when branch power and current measurements were placed near the source. Finally, the results were better when meters were placed at different locations [32].

For load estimation, Liu and Chiang presented a heuristic method in [90]. The objective of this work is to find the meter placement strategy to put the power measurements on the feeder for reducing the error of load estimation to the acceptable range. This approach had two stages. The first stage started by traversing the tree backward from the leaves to the root, candidate locations for both main and backup monitoring were found based on proper load estimation. Confidence intervals for load estimation could then be calculated by adding the certain error to the load estimation. In the second stage, if the placed meters would not give a satisfactory confidence interval, the first stage would be repeated with stricter concerns. These two stages were iterated when a confidence interval was reached completely. Some practical considerations such as meter failure, space availability, and unbalanced distribution systems were reviewed in this study.

More researches have been conducted in *voltage profile estimation*, some of which are listed here. The goal of these works is to estimate nodal voltages within the certain accuracy. First, Leou and Lu worked on improving feeder voltage profile estimation with telemetered data in [91]. They modeled load uncertainty as fuzzy numbers and all plausible system states encountered as a result of the expected uncertainties through Least Squares method for state estimation (fixed and same weights for all kind of measurements). Technical guidelines from system dispatchers at Taiwan Power Company were used to develop four rules for meter placement at the feeder to determine tentative meter locations. The number of measurements was determined in this research and placed on the different zones. The Dynamic Programming approach was also considered to find the best meter locations in each load zone to estimate the voltages by minimizing the uncertainty [91].

Another example of voltage profile estimation is the work performed by Strbac, Jenkins and Shafiu in [92], finding measurement locations for SE of distribution networks. This paper predetermined the number of the voltage measurement set. The authors proposed another heuristic approach to identify potential points for voltage measurements. Measurements were placed primarily at the heavily loaded nodes with a good cover of network topology. Standard deviations of voltage magnitudes were calculated through changing the load estimation error and running the power flow for the designated network. As a next step, measurements were moved to the nodes with higher standard deviations. Measurements were added until the required performance was reached. Most recently in 2011, R. Singh *et al.* have proposed two probabilistic approaches to meter placement on distribution system for state estimation, based on Monte Carlo simulations [47] and [93].

Table (5-1): Comparison of Different Approaches for Meter Placement on Distribution Feeders.

Ref.	Application	Approach	Type of Measurements	Test System	Optimality	No. of Measurements
[32]	SE	Heuristic	VM,CM, PM	3-phase	No	Not fixed
[33]	Monitoring	Heuristic	VM, CM, PM	1-phase	No	Not fixed
[90]	Load Est.	Heuristic	PM	1-phase	No	Not fixed
[91]	Voltage Est.	Heuristic	VM	1-phase	No	Fixed
[92]	Voltage Est.	Heuristic	VM	1-phase	No	Fixed
[47]	Voltage Est.	Probabilistic	VM, PM	1-phase	No	Not fixed
[93]	Voltage Est.	Probabilistic	VM, PM	1-phase	No	Not fixed

Their objective was to bring down the relative errors in the voltage and angle estimates, at all buses, below some predefined thresholds in more than 95% of the simulated cases coincidently. In these approaches, the voltage and line power flow measurements were progressively moved until relative errors became mostly below the targets. First [47] formulated the meter placement problem as a feasibility problem. Subsequently the authors have modified the proposed method

by utilizing Ordinal Optimization (OO) to reduce the search space of potential combinations of measurement locations in [93]. This modified algorithm was used to find the smaller set of meter locations. In addition, in [93], meter placement problem has formulated as a stochastic optimization problem unlike the approach undertaken in an earlier work where it was looked at the problem as a feasibility problem like in the earlier work. Both approaches did not provide optimal solution for meter placement problem, however, the latter approach lead to a better solution in comparison to the former approach [47, 93]. All of the aforementioned methods are listed in Table (5-1).

5-3 Meter Placement Problem

To completely as well as ideally monitor the distribution feeders, it is required to put one voltage measurement at each node and one current measurement on each branch for each phases in case of unbalanced feeders. This measurement set provides a redundant measurement set to monitor the network. Mostly the number of the nodes and branches for real feeder is large, therefore the cost of this implementation is expected to be very high. From an economic point of view, it is not possible to place these numbers of measurements on the feeder. Based on this fact, it is crucial to define the meter placement problem to find the locations, numbers, and types of the measurement in order to provide a sufficient measurement for the SE to estimate the voltage profile within the desired accuracy. It is desirable to place as few measurements as possible under the constraints of automation requirements (e.g. Volt/Var Control application), which is an optimization problem. As previously mentioned, the problem statement of “Meter Placement for VVC” is determining the number, place, and type of meters needed to be placed on a given feeder such that the SE with these measurements can estimate the voltages with desired accuracy in a cost effective way. This problem could be presented as follows:

$$\min\ f_O = \sum_{i=1}^{n} C_i(d_i)$$

$$s.t. \quad \max\left\{\hat{\boldsymbol{\sigma}}_V\right\} \le \tilde{\sigma}_V$$

$$\mathbf{M}_1(\hat{\mathbf{x}}, \mathbf{z}) = 0$$

$$\hat{\boldsymbol{\sigma}}_V = \mathbf{M}_2(\hat{\mathbf{x}}, \mathbf{z})$$

$$d_i \in \{0,1\}$$

where:

d_i corresponds to the decision that measurement, i, is placed,

C_i is the cost function of the measurement, i, is placed on the designated location,

$\hat{\mathbf{x}}$ is the vector of the estimated system state from SE,

$\mathbf{z}$ is the vector of measurement for SE,

$\hat{\boldsymbol{\sigma}}_V$ is the estimated standard deviation of the voltages on the feeder,

$\mathbf{M}_1$ is the non-linear function of SE to calculate $\hat{\mathbf{x}}$ from $\mathbf{z}$,

$\mathbf{M}_2$ is the non-linear function for $\hat{\boldsymbol{\sigma}}_V$ calculation from $\hat{\mathbf{x}}$ and $\mathbf{z}$,

$\tilde{\sigma}_V$ is the target value of voltage standard deviation for VVC application, and

n is the number of possible measurements.

The above is a non-linear, integer programming problem. Decision variables, d_i, are to be used to choose the measurements along the feeder. For VVC application, voltages must be estimated within certain accuracy. Here, the measure of the accuracy is standard deviation where a higher standard deviation means lower accuracy. Therefore, the maximum of standard deviation of the estimated voltages, $\max\left\{\hat{\boldsymbol{\sigma}}_V\right\}$, must be less than the target value. Target value, $\tilde{\sigma}_V$, is the maximum acceptable standard deviation for voltage estimation which comes from the VVC goal (voltage bandwidth).

A very general problem-solving technique consists of systematically enumerating all possibilities for the solution and checking whether each candidate satisfies the problem's statement and requirements. In addition, solving this problem needs to run an exhaustive search (Brute-force search) to check and calculate the standard deviations for each different scenarios of meter configuration then choose the best one accordingly. For instance, if we want to place

three voltage measurements on three node feeder, there is $2^3 = 8$ possibilities to search. If the size of the feeder increases we need to run another search to find three possible locations which means $\binom{n}{3}$ combinatorial searches. Therefore, the needed searches for placing k meter at n possible location is $\binom{n}{k}.2^k$.

Furthermore, choices are made regarding the locations, types, and number of measurements. For example, to monitor one feeder with 34 nodes with balanced structure, it needs 34 voltage measurements and 33 current measurements. Totally 67 measurements must be considered in the network which is not economically possible. So, the search size to find the optimal number and place becomes $2^{67} \cong 1.4757 \times 10^{20}$ which needs an exhaustive search and this search size increases exponentially. By considering unbalancing of the distribution feeders, three phases, the size of the search would increase dramatically.

5-4 Solution Approach

From computer science point of view, this problem is a NP-hard (non-deterministic polynomial-time hard) problem. “Results about NP-hardness in theoretical computer science make heuristics the only viable option for a variety of complex optimization problems that need to be routinely solved in real-world applications [94]”. Heuristic methods are categorized in two main groups [94-97]:

(a) Constructive heuristic methods: Using simple minded greedy functions and rules for the evaluation of various options (choices) to build a reasonable/feasible solution [97].

(b) Perturbative heuristic methods: These methods are also called local search (neighborhood search) methods. They start searching from some initial solution and move to neighboring solutions. Generally, it continues until it reaches at a local optimum, one that does not have a better neighbor [97], such as Simulated Annealing (SA), Tabu Search, Genetic Algorithm (GA), and others [95].

A very general problem-solving technique consists of systematically enumerating all possibilities for the solution and checking whether each candidate satisfies the problem statement and its requirements. Due to the size of the possible search space, it is needed to reduce the search space by using some heuristics specific of the problem. By looking back to the meter placement problem, most of the references approach the problem through the constructive heuristic methods. They have proposed some rules to reduce this search space, for instance Baran *et al*. [33] have adopted three rules to put the meters on the feeder for on-line monitoring. Moreover, Schulz *et al.* [32] have developed other three rules to place measurement to improve the performance of the SE. Another heuristic technique has been applied to identify locations to place a certain number of voltage measurements to provide the desired accuracy for voltage profile estimation by Shafiu *et al.* [92] where a predetermined number of meters was chosen arbitrarily in this research.

Some references have considered the fixed number of measurements to place at the network, [92] is a good example. The problem lies in finding the appropriate places to put the meters. On the contrary, there are some methods which focus initially on finding the appropriate location for a designated type of meters; as a next step they look to eliminate the meters with low contribution on the objective of the optimization problem. For instance, the meters are ranked based on their impact on the accuracy index of the system [33, 92]. Finally, low impact meters will be eliminated from the measurement set.

Here the problem was approached by exploring optimum solutions and finding the "guidelines" for meter placement on distribution feeders. Guidelines result from the extensive observations and sensitivity analyses to find out the behavior of the standard deviation of voltages in different conditions and situations. These sensitivity analyses were about load estimation uncertainty (this study has been done in previous chapter), real-time measuring accuracy classes, number of the measurements, type of the measurement (i.e. voltage and current measurements), better locations for different type of measurements such as voltage and current, and cost analysis of the current and voltage measurement for distribution feeders. Based on these observations, we can adopt the guidelines for meter placement problem to find the near-optimal solution.

In these studies, it is assumed that power measurements (active and reactive) are available at the substation.

5-5 Cost of Measurements (CM and VM)[2]

To monitor the feeder, we need to set a measurement set on it. Energy consumption of the customer monitors by power meter at the secondaries. In case of Automated Metering Infrastructure (AMI) [57], the voltage can also be measured and recorded at the secondaries. It is a common behavior to install, the power as well as the voltage and current measurements at primary circuits to measure the power, voltage and current, along the feeder. These measurements are available with various technical specifications at different costs.

These measurements must be collected from the field and transmitted to the control center to utilize them as the input of the SE. Consequently, communication cost will be added to the total cost, a cost that depends on the technology, such as cellular, fiber optic, and others. A communication cost has two parts, one that accounts for the measurement device and the other accounts for the master point at the substation. Sometimes, remote terminal units (RTUs) can be installed at the field to collect the measurements at one location from different measurements then transmit them to the control center. For one example of these architectures, the COOPER communication architecture [103] is shown in Figure (5-2).

In addition to the communication costs, some other instruments such as CTs, PTs, enclosures, and other devices and components are needed to be installed in the field to facilitate the physical connection between the measurement meters and communication devices.

Different technologies have been developed to provide the accurate devices to measure the voltage and current and send it to the control center. For example, one company added the metering part with communication part together and developed one device, like SHARK 200-S of GE Inc. [102, 107] and S.T.A.R. technology of Cooper Inc. [103]. Another way to optimize the network was considering a one RTU for several meters in a way that they are connected together, using one communication gateway they will connect to the control center.

[2]This market research has been performed by the guidance of Dr. Lubkeman.

However, one concern regarding the current sensor is the presence of a battery as a backup power needed for power supply. Fortunately, GridSense can be used to solve this issue by adding a small solar panel on the measurement set to provide the different energy supply in LineIQ [104].

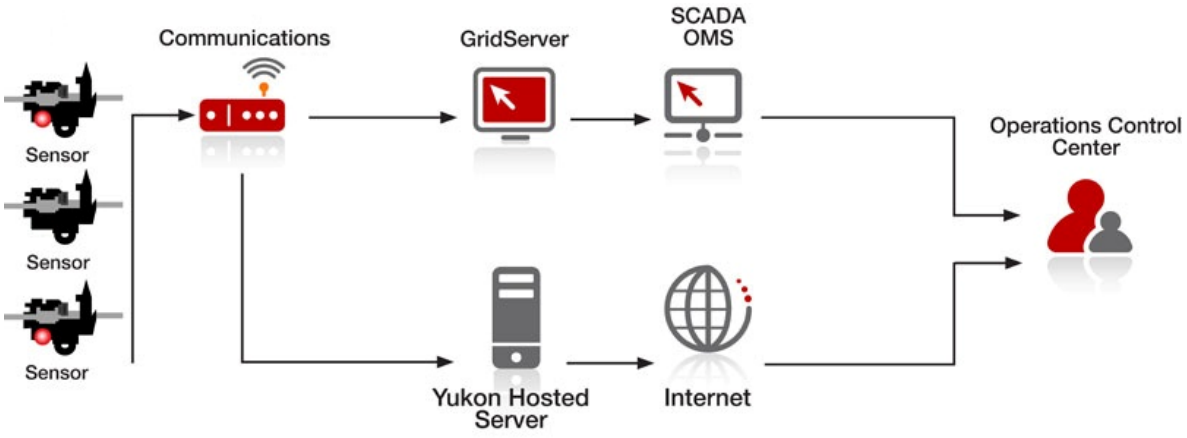


Figure (5-2) Solution Architecture Diagram from COOPER [103].

To measure the voltage along the feeder, end-of-line (EOL) voltage sensors are utilized. For example, Landis+Gyr FOCUS AX [106] smart meters residential customers and SATEC EM920 eXpertmeters [105] for industrial and commercial customers could be considered for measuring voltages.

Table (5-2): Cost comparison of voltage and current sensors [102-107].

Type of measurement	w/ integrated Communication	w/o integrated Communication
VM	~ 670 USD/per phase	~ 350 USD/per phase
CM	~ 1350 USD/per phase	~ 500 USD/per phase

The cost of current and voltage measurements is summarized in the Table (5-2). There are two types of sensors, one is integrated with communication capability and the other is implemented without communication capability. Generally speaking and based on the available technologies

and information, the cost of three voltage measurements is roughly similar to that cost of the two current measurements.

The total cost can be summarized in the following items:

- Cost of meters (sensors),
- Cost of communication,
    - This includes meter side, medium, and control center side communication for gathering data from field and saving them in the data base, e.g. master radio at the substation and etc.,
- Cost of electrical facilities (CTs, PTs, enclosures, etc.), and
- Cost of installation and testing.

By interviewing, it was found that the total cost of this technology implementation is proportional to the number of meters (end points). Hence, less measurement points on the network provides the cheaper measuring system architecture. In case of considering one type of the measurement (voltage or current) with a certain number of the sensors, the total cost for implementation of the measurement system with voltage sensors is cheaper than with the current sensors.

5-5-1 Metering Accuracy

Metering accuracy is related to the accuracy measurement class of meters which are installed along the feeder with different standard [101]. Normally, power measurements are installed at the substation and at the secondaries on the customer side. Common measurements along the feeder are current and voltage measurements which are analyzed as follows.

5-5-1-1 Current Measurement (CM) Accuracy

In this part, the effect of the associated error for current measurements was investigated in different conditions. Case 2 with three current measurements (CM) in the prototype system was chosen and is shown in Figure (3-3). As mentioned, this case has three CMs, namely: Cm14, Cm30, and Cm 27; respectively along the feeder. In this study, associated errors for

current measurement are considered in the following levels: 0.3%, 1%, 3%, and 5%, for each phase in Figure (5-3). In the work at hand, the load estimation has been considered constant at the level of 30% in model 1.

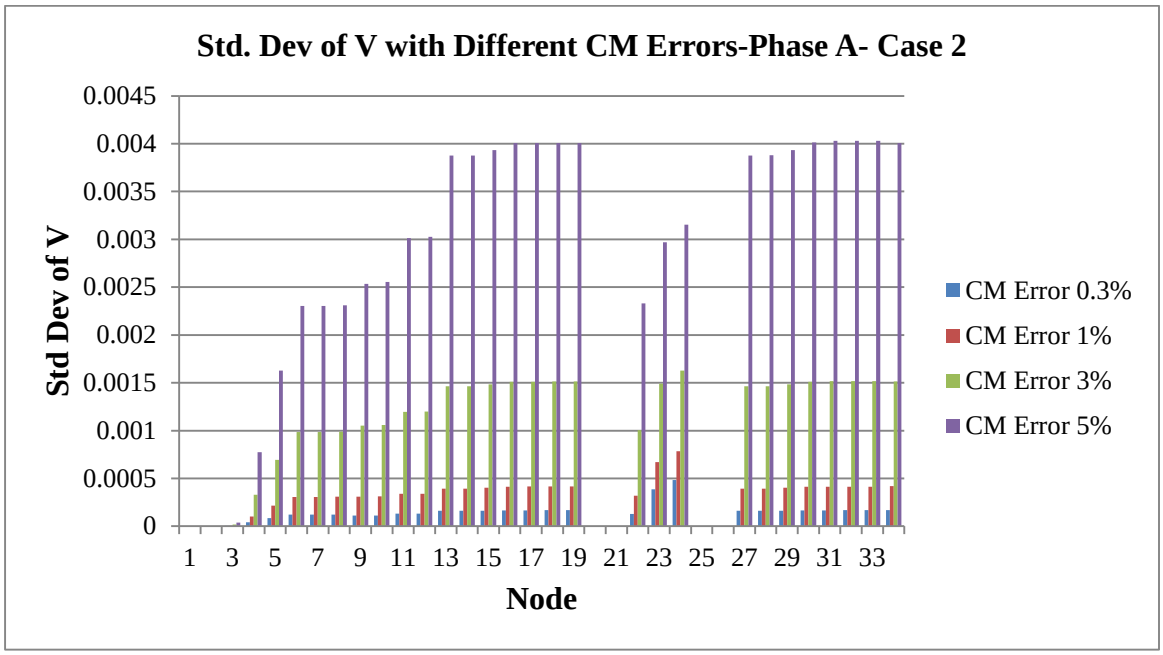


Figure (5-3): Standard deviation of voltage magnitude with different CM errors (Model 1) for phase A.

As observed from Figure (5-3), the level of associated error of real-time measurements has a significant impact on the voltage estimation along the network. For further studies and more investigation to see the pattern of voltage standard deviations, Model 2 has been considered for studies. Here, two different metering accuracy classes, which are 0.3% and 1.2%, have been considered where the load estimation has been considered constant at the level of 50 kW. As can be clearly observed from Figure (5-4), the level of the associated error of real-time measurements, CM, has an impact on the voltage estimation along the network. To demonstrate the changes of the standard deviation of system state, $I_{r,a}$ has been chosen for this

purpose and shown in Figure (5-5). As expected the standard deviation of the system states depended on the metering accuracy class implemented in the system.

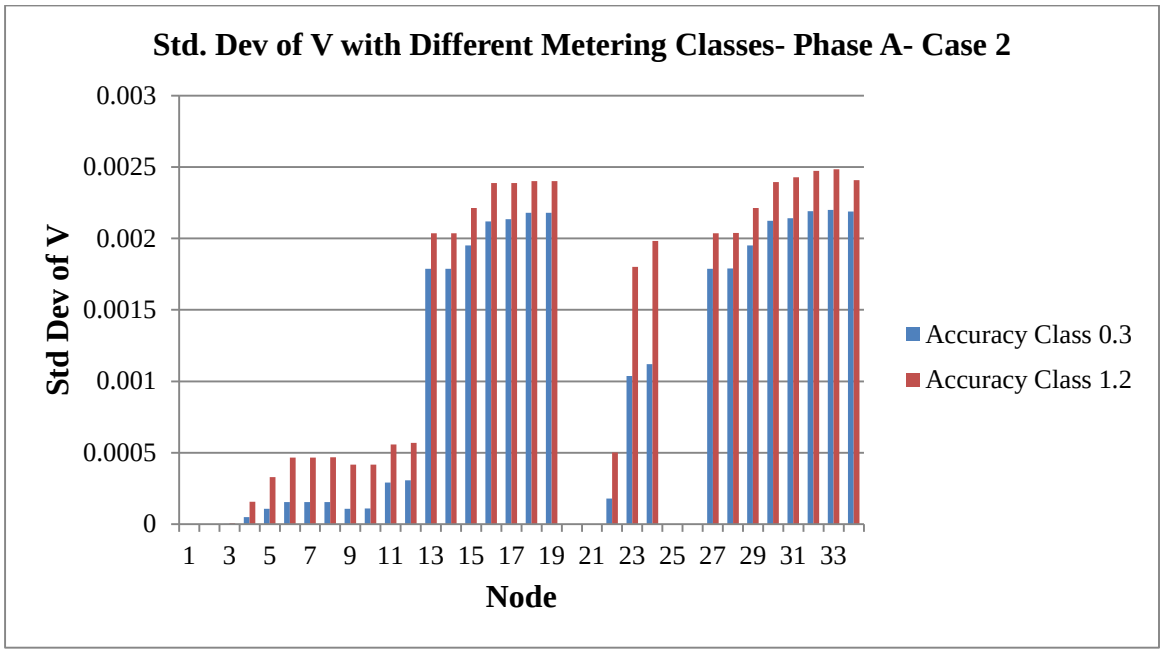


Figure (5-4): Standard deviation of voltage magnitude with different CM errors (Model 2) for phase A.

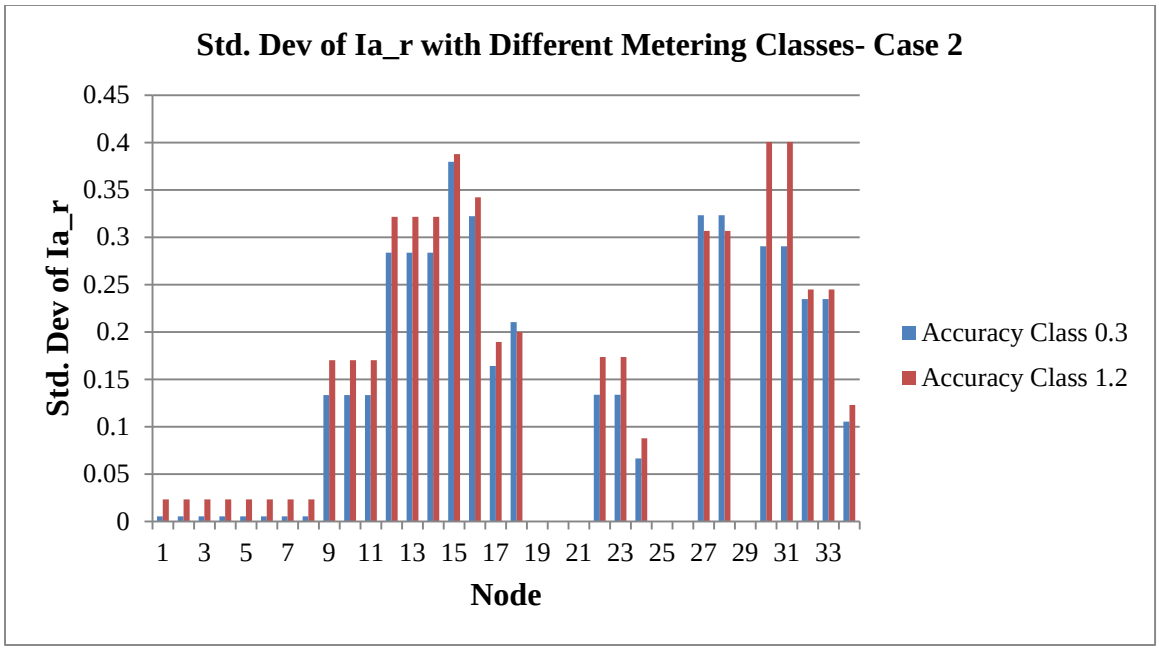


Figure (5-5): Standard deviation of real part of branch current with metering classes (Model 2) for phase A.

5-5-1-2 Voltage Measurement (VM) Accuracy

In this part, the effect of the associated error for voltage measurements was investigated in different conditions. Case 3 with four voltage measurements (VM) at the prototype system was chosen for this study with associated error in levels: 0.3%, 0.6%, 0.9%, and 1.2%, for each voltage measurement [101]. Four VMs are allocated at nodes 18, 33, 28 and 14 and shown in Figure (5-6). Voltage standard deviation profile with 30% error on load estimation and 1% error for M1 are depicted in Figure (5-7) in regard of different VM errors.

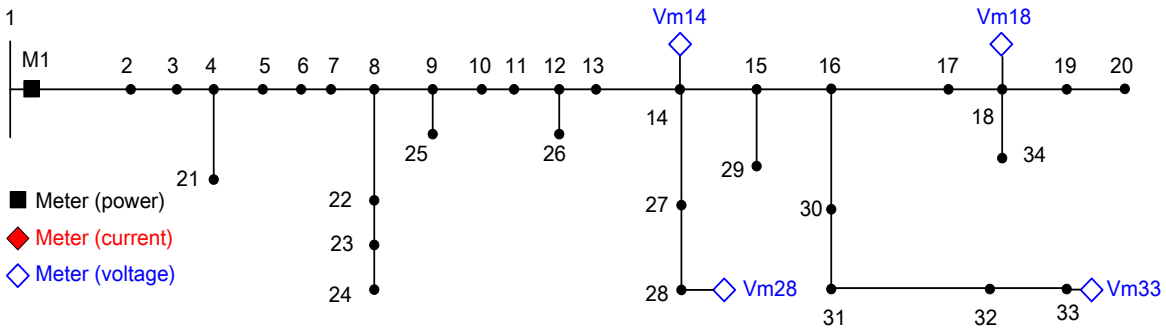


Figure (5-6): IEEE 34 node test feeder with M1 at substation and four other VMs (Case 3).

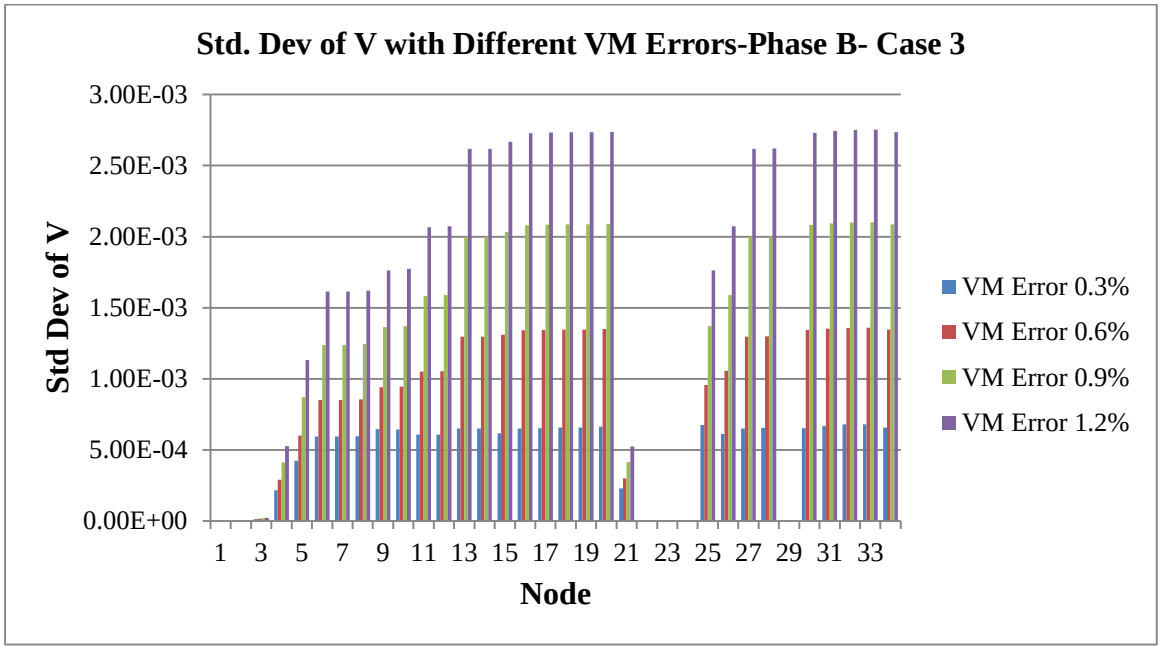


Figure (5-7): Standard deviation of voltage magnitude with different VM errors (Model 1) for phase B.

Based on the previous observation, voltage measurements were allocated at the end of the main feeder (node 18), the end of the three phase laterals (node 33 and node 28), and after the voltage regulator at node 14 in the middle of the feeder near the main loads.

For further studies and more investigation to identify the pattern of voltage standard deviations, Model 2 was considered for further studies. Here, two different metering accuracy classes were considered with a constant load error of 50 kW. Figure (5-8) shows the voltage standard deviation profile along the feeder in regard to different metering accuracy classes 0.3 and 1.2. In addition, more accurate metering devices provide more accurate measurement for SE; hence, SE can provide a more accurate estimation.

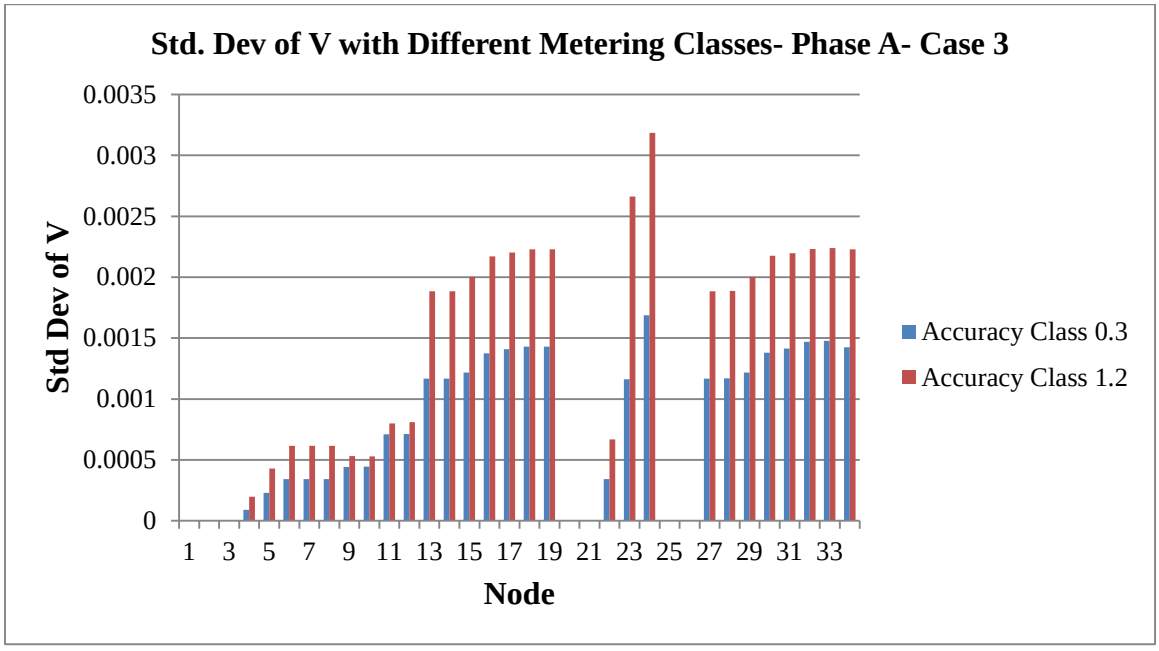


Figure (5-8): Standard deviation of voltage magnitude with different VM errors (Model 2) for phase A.

Based on the previous observations, the accuracy level of the real-time measurements has a significant impact on both the voltage and system state estimations. In other words, more accurate measuring devices provide better estimation. On the other hand, the price of more accurate measuring devices is higher than the price of less accurate measuring devices. For our future studies in meter placement problem, we considered fixed class for measuring devices.

5-6 Sensitivity Analyses

To find the scheme of putting the meters along the feeder, these sensitivity analyses need to be investigated.

- Number of measurements,
- Meter locations, and
- Impact of voltage measurements versus current measurements (VM vs CM).

5-6-1 Meter Locations

In this part, current and voltage measurement locations have been studied in two different load estimation error models, i.e. Models 1 and 2.

5-6-1-1 Study of CM Location

To investigate the impact of the location of the current sensor in the network, two different setups have been considered. The first one is Case 2 which has power measurements at the substation and three current measurements along the network, namely: Cm14, Cm27, and Cm30. The second setup has power measurements at the substation and three current measurements along the network, namely: Cm18, Cm27, and Cm30. In other words, the current sensor at placement Cm14 has been moved to placement Cm18 to examine the impact of the current sensor location on the feeder. To visualize this movement, a new set of measurements with Cm18 has been depicted in Figure (5-9).

The load estimation and real-time measurement error have been considered constant at 30% and 1%; respectively for this study. Standard deviation of voltages and system states along the feeder are shown in the following figures.

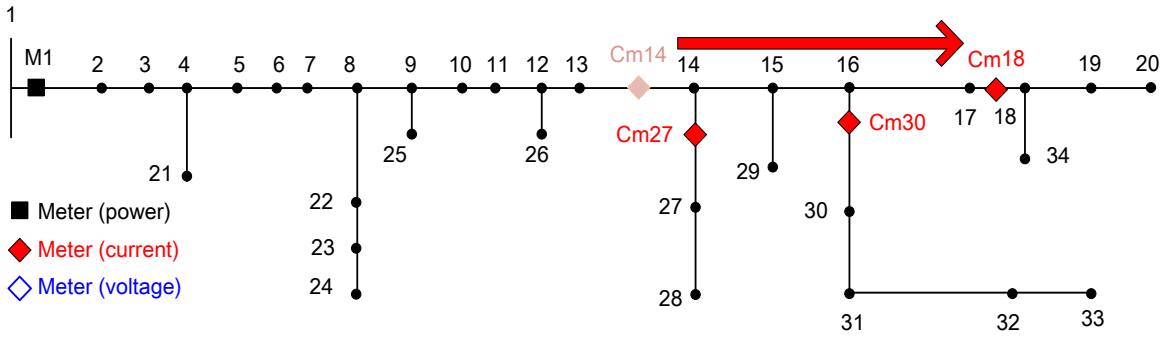


Figure (5-9): Moving one CM along the feeder from Cm14 to Cm18.

In addition, the impact of the CM locations on the voltage estimation quality has been analytically investigated on a sample feeder with same loading conditions for each branches. In addition, all branches have the same the line characteristics. This study is provided in Appendix 3.

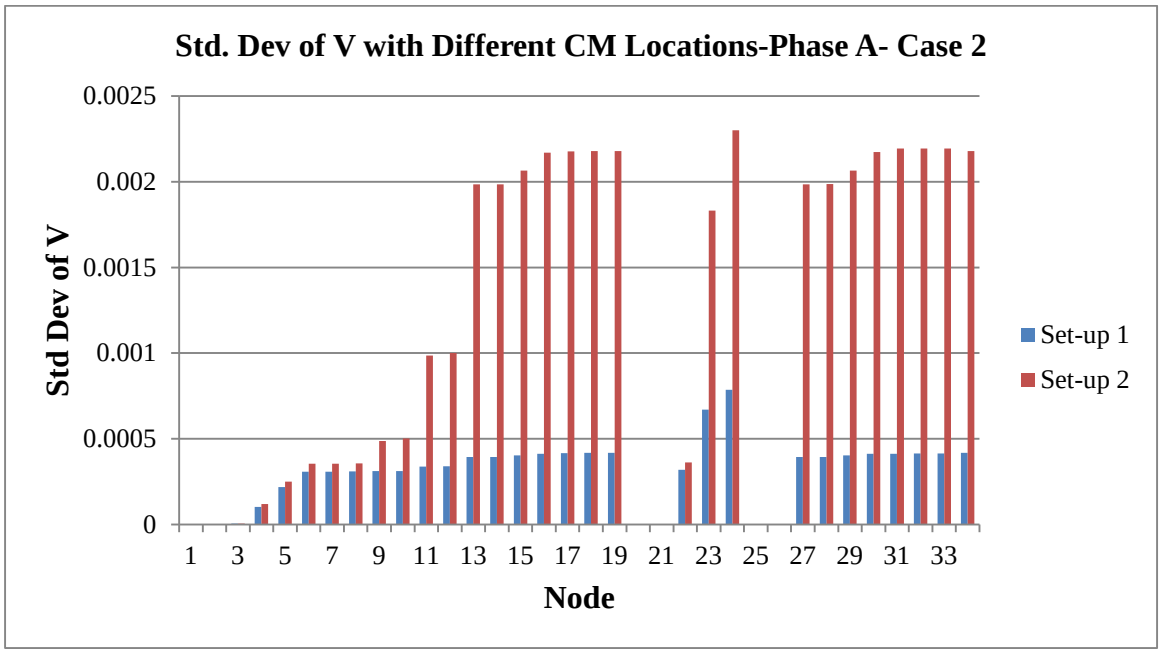


Figure (5-10): Standard deviation of voltage magnitude with different CM locations (Model 1) for phase A.

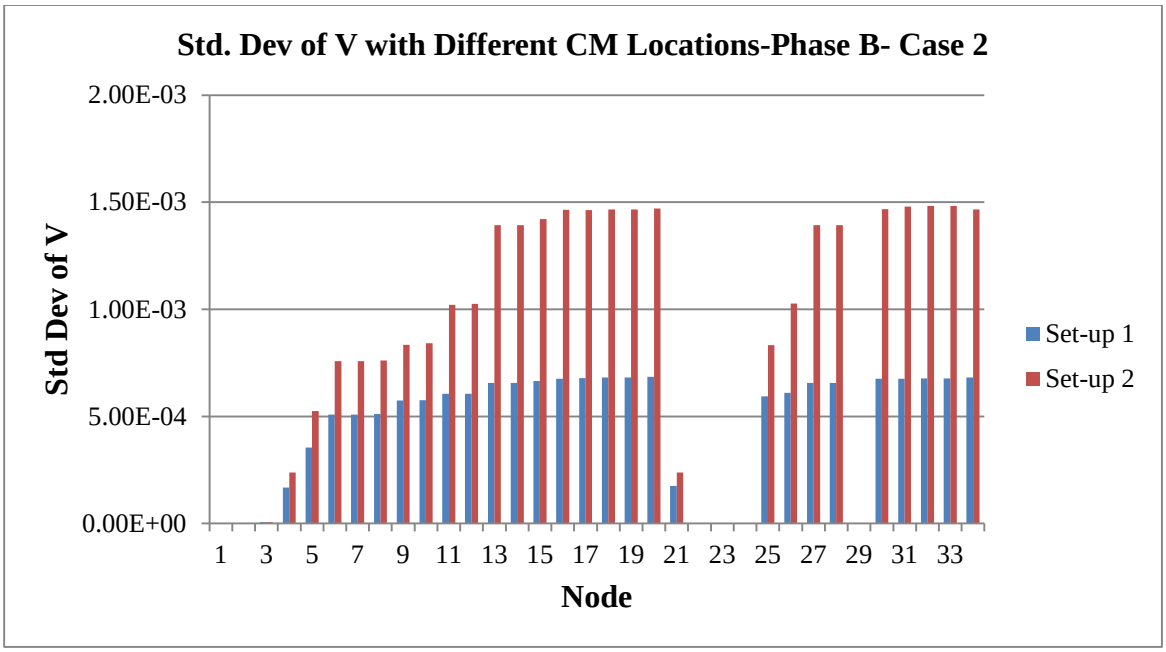


Figure (5-11): Standard deviation of voltage magnitude with different CM locations (Model 1) for phase B.

As shown in Figures (5-10) and (5-11), the location of the current sensor has a significant impact on the standard deviation of the estimated voltage magnitudes at different nodes. For example, standard deviation of estimated voltage at node 33 for phases A, B, and C become 0.0022, 0.0015, and 0.0016; respectively from 0.0004, 0.0007, and 0.0007; respectively by replacing the Cm14 by Cm18 in the system. Hence, this study showed that the impact of the current sensor on the voltage estimation depends on the location, i.e. if the sensor is placed near the beginning of the substation or the beginning of the lateral, the state estimator would provide a better estimated voltage magnitudes and system states.

For further studies and more investigation to see the pattern of voltage standard deviations, Model 2 has been considered for studies too. In this study, the load error is 50 kW with 0.3 accuracy class of measurement. Standard deviation of voltages along the feeder is shown in the Figure (5-12).

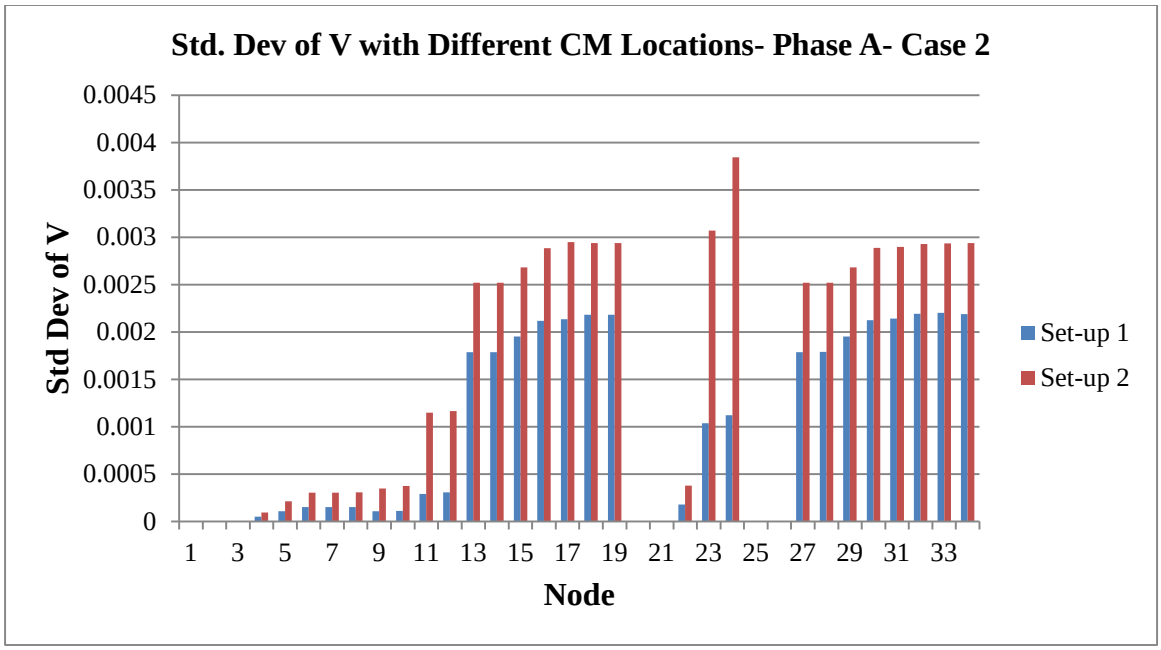


Figure (5-12): Standard deviation of voltage magnitude with different CM locations (Model 2) for phase A.

From Figure (5-12), by replacing the Cm14 by Cm18 in the system, the standard deviation of estimated voltage at node 33 for phases A, B, and C becomes 0.0029, 0.0045, and 0.0032; respectively from 0.0022, 0.0021, and 0.0012; respectively. This investigation shows that the voltage estimation depends on the location of the CM, i.e. if the sensor is placed close to the beginning of the substation or the beginning of the lateral, the state estimator would then provide a better estimated voltage magnitudes and system states.

To investigate more in regard of the CM locations on the feeder, it is considered that there is one CM. This CM has been moved from different branch consecutively to see the difference on the voltage estimation accuracy. For instance, the CM has placed in branch 18 first. Then this measurement has moved to branch 17, branch 16, branch 15, branch 14, and branch 13, as shown in Figure (5-13). For each of these cases, the voltage standard deviation profiles have been calculated and illustrated in Figure (5-14). As this figure shows, the maximum of the voltage standard deviation has been decreased when the CM moves from the end of the feeder

to the middle of the feeder, where this measurement can monitor the loads after itself significantly. In other words, putting the CM after the load is not practical and it does not have an impact on the performance of the SE for voltage estimation.

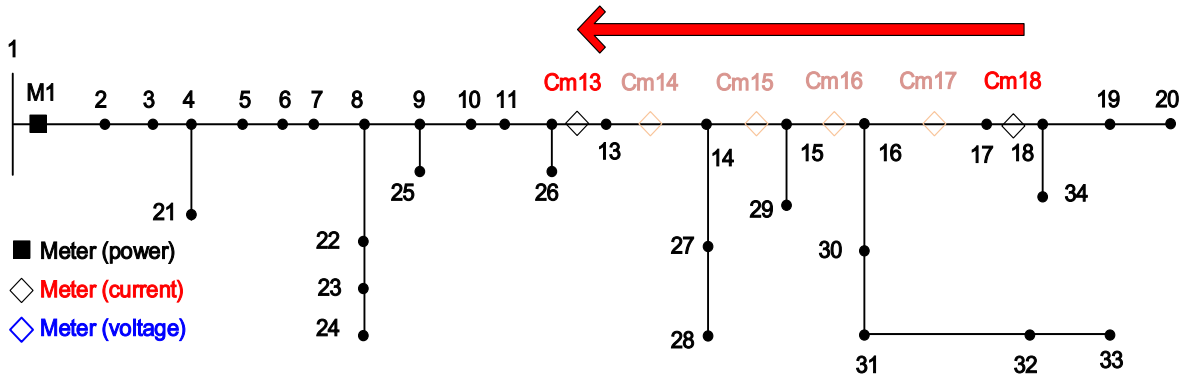


Figure (5-13): Moving one CM along the feeder from branch 18 (end of the feeder) to branch 13 (middle of the feeder).

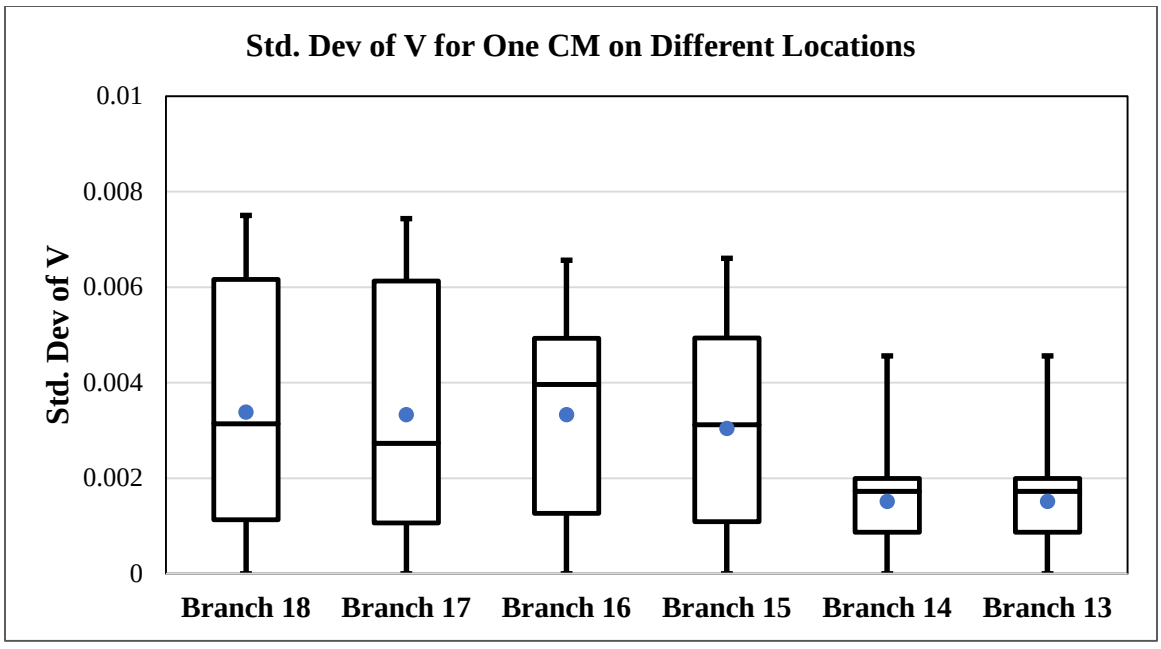


Figure (5-14): Boxplots of voltage standard deviation for different locations of one CM.

5-6-1-2 Study of VM Location

Two different setups were considered to investigate the impact of the location of the voltage sensor in the network. The first setup is Case 3 which has power measurements at the substation and four voltage measurements along the network at nodes 18, 33, 28, and 14. The second setup has power measurements at the substation and four VMs along the network at nodes 18, 33, 28, and 8. In other word, Vm14 has been moved to Vm8 to examine the impact of the voltage sensor location in the network. This location changes in voltage measurement set is shown in the Figure (5-15).

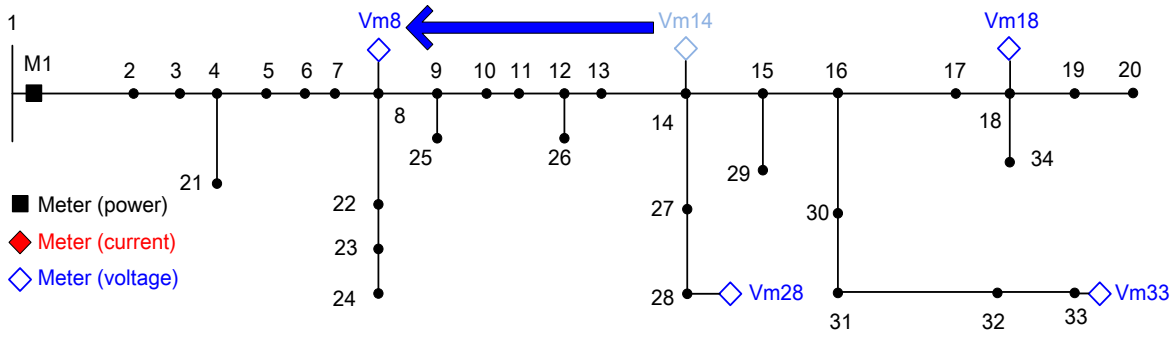


Figure (5-15): Moving one VM along the feeder from Vm14 to Vm8.

In the study at hand, the level of the load estimation and real-time measurement errors have been considered a constant at 30% and 1% for power and 0.3% for voltage measurements; respectively. Standard deviation of voltages and system states along the feeder are shown in the Figures.

It is clear from Figure (5-16), the location of the voltage sensor has a significant impact on the standard deviation of the estimated voltage magnitudes at different nodes, where smaller standard deviation yields to a better estimation. In this case, there are four voltage sensors in the network and for setups where only the small movement of the sensor is the one considered. In addition, the impact of the VM locations on the voltage estimation quality has been analytically investigated on a sample feeder with same loading conditions for each branches.

In addition, all branches have the same the line characteristics. This study is provided in Appendix 3.

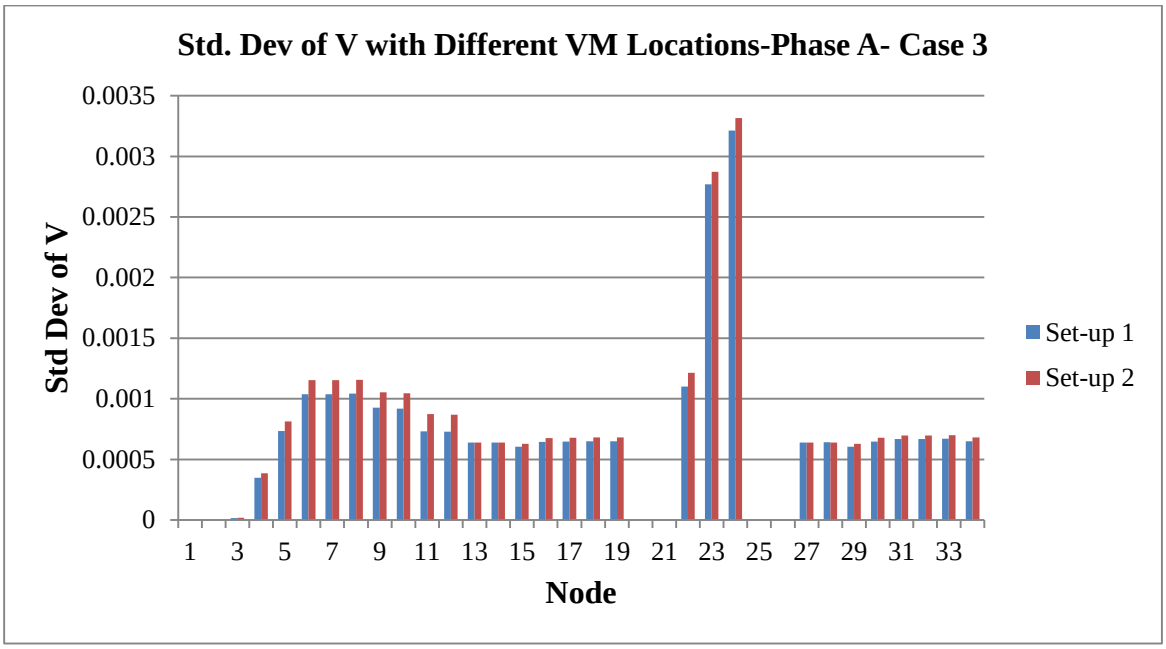


Figure (5-16): Standard deviation of voltage magnitude with different VM locations (Model 1) for phase A.

From the simulation results, by replacing voltage sensor from node 14 to node 8 in the network, standard deviation of estimated voltage at node 33 for phases A, B, and C became 0.007, 0.0073, and 0.0068; respectively from 0.0067, 0.0068, and 0.0063; respectively. This experiment shows that the impact of the voltage sensor on the voltage estimation depends on the sensor's location. Hence, the state estimator will provide the better output if the voltage sensor is placed close to the end of the feeder or at the end of the laterals. For further studies and more investigation to observe the pattern of voltage standard deviations, Model 2 was considered to be studied as well. In this study, the load error is 50 kW with a metering accuracy class of 1.2% was considered for power and voltage measurements. Standard deviation of voltage magnitudes along the feeder is shown in Figure (5-16). Clearly from Figure (5-17), the

location of the voltage sensor has a significant impact on the estimation of voltage profile. In this study, there are four voltage sensors considered at the network, for the two aforementioned setups and only small movement of only one sensor was considered. From the simulation results, standard deviation of estimated voltage at node 18 for phases A, B, and C became 0.0028, 0.0027, and 0.0027; respectively from 0.0022, 0.0023, and 0.0024; respectively by replacing the voltage sensor from node 14 to node 8 in the network.

Hence, it can then be concluded that the impact of the voltage sensor on the voltage estimation depends on the location. If voltage sensor is placed near the end of the feeder or at the end of the laterals, it could be expected that the SE can provide the better performance.

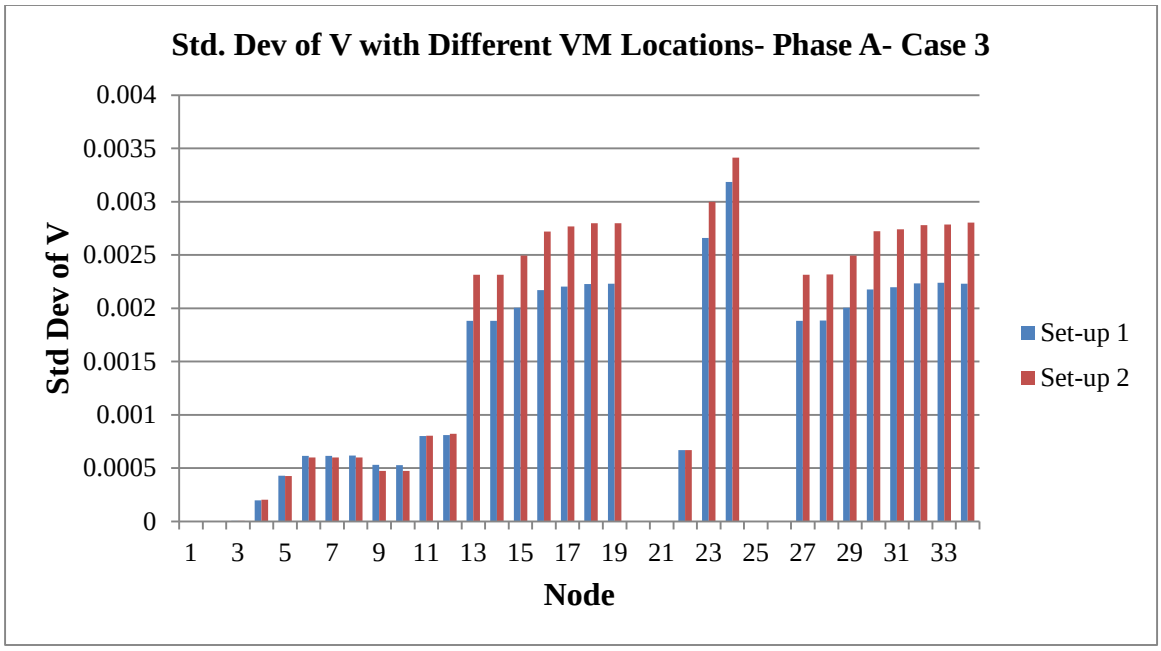


Figure (5-17): Standard deviation of voltage magnitude with different VM locations (Model 2) for phase A.

To further investigate this matter in regard of the VM locations on the feeder, it is considered that there is one VM with power and voltage measurements at the substation, i.e. M1. This VM was moved from different nodes consecutively to determine the impact of its location on the

quality of the voltage estimation. In this particular case, the VM has placed node 33 first. Following, this measurement was moved to node 31, node 16, node 14, and node 12, as shown in Figure (5-18). Nodes 33 and 31 were on one three phase lateral where it was connected to the main feeder at node 16. Moreover, nodes 16, 14, and 12 were placed on the main feeder. For each of these different locations, the profile of the voltage standard deviation was calculated. The profiles of the standard deviations are depicted in Figure (5-19).

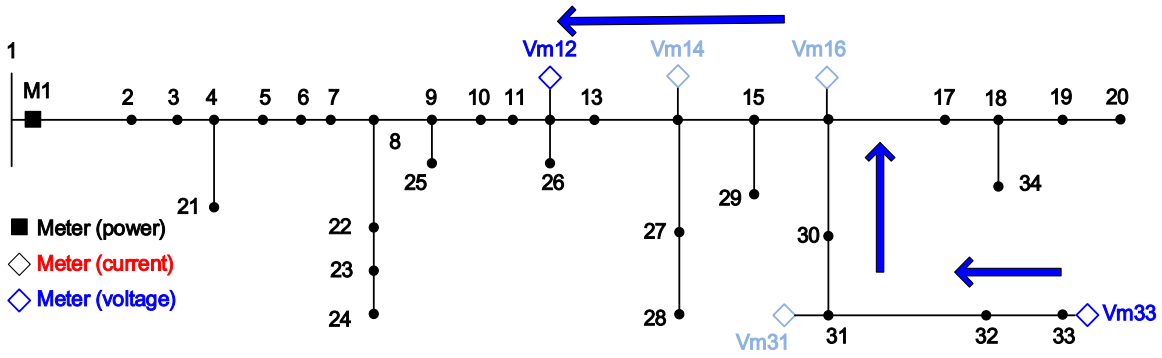


Figure (5-18): Moving one VM along the feeder from the end of the lateral, i.e. Node 33 to the middle of the feeder, i.e. Node 12.

As Figure (5-19) shows, the maximum of the voltage standard deviation increased when the VM moved from the end of the feeder to the middle. The performance of the SE to voltage estimation dropped significantly by moving from the end of the lateral to the middle of the main feeder. In other words, putting the VM in the beginning of the latter instead of the end deteriorates the estimation of the voltage. In addition, these studies show that VMs have better performance when they are placed in the end of the feeder of the laterals.

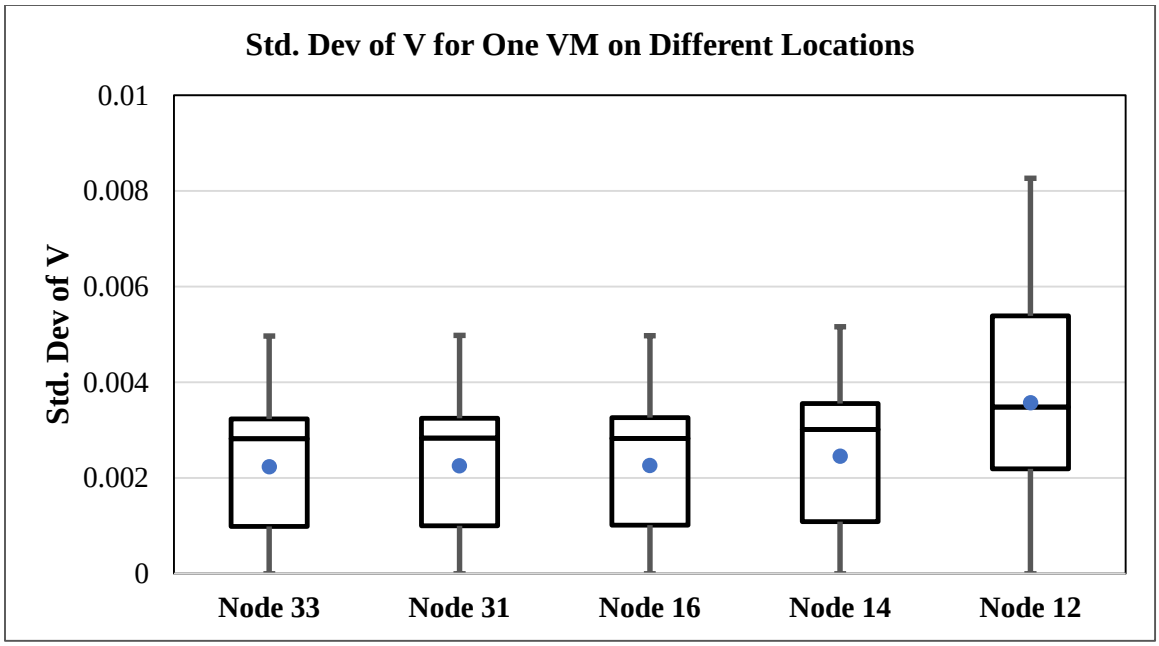


Figure (5-19): Boxplots of voltage standard deviation for different locations of one VM.

5-6-2 Impact of Voltage Measurements versus Current Measurements (VM vs CM)

In this part, two different experiments have been adopted to figure the impact the type of the measurement device has on the performance of the BCSE with a desired measurement set. The first experiment was designed to consider adding one new voltage or current measurement. As for the second experiment, it was planned to explore the impact of both CM and VM on the SE performance in Cases 2 and 3.

First Experiment

The goal of this analysis is to find the type of the measurement, voltage measurement or current measurement, when it is desired to put a new measurement to the measurement set. In this case, VMs are placed at the end of the main feeder or end of the laterals and the CMs are placed close to the load centers as well as at the beginning of the laterals. For the purpose of this study, Case 1 was considered with one power measurement (M1) at the substation, measurement accuracy class for real-time measurements was 1.2% and load estimation error was set at 50%.

One measurement was located on the feeder, one VM at node 18 and one CM at branch 14 has been employed for current measuring. Figure (5-20) summarizes the maximum and average of voltage standard deviation by adding one new measurement, either CM or VM, to the measurement set.

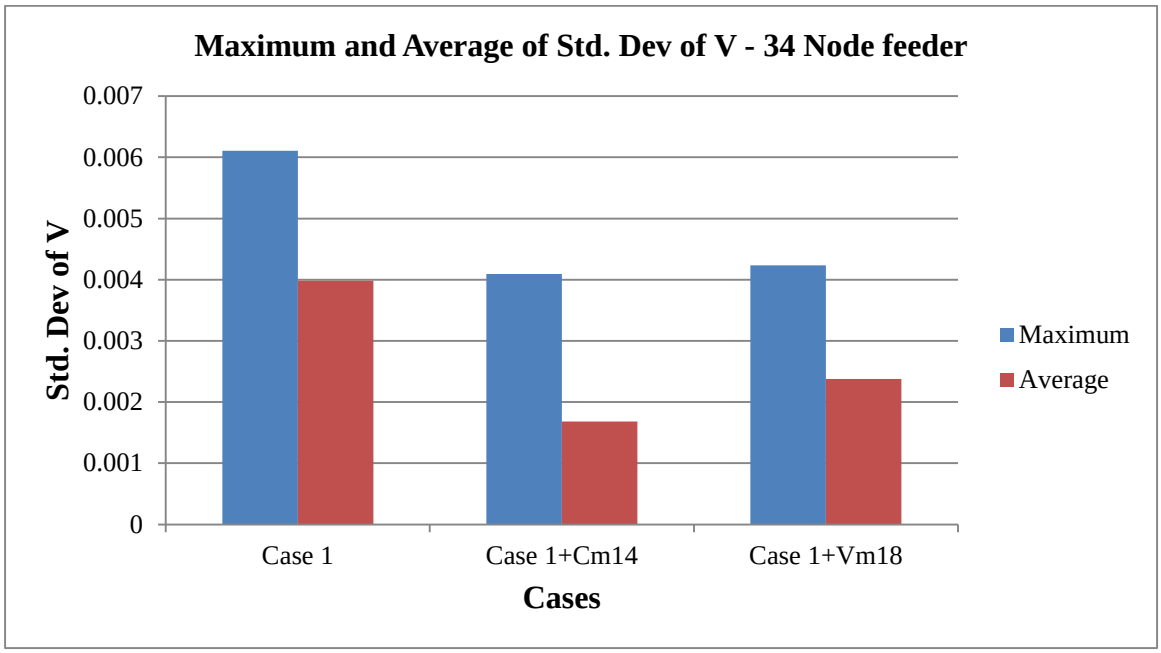


Figure (5-20): Comparison of maximum and average of voltage standard deviations by adding one new CM or VM.

By detecting the maximum standard deviation of voltages, it was found that the maximum standard deviation in both cases has dropped in the same manner. By adding one VM, the $\max\{\sigma_{Vi}\}$ has dropped from 0.0061 to 0.0042, similarly by adding one CM, the $\max\{\sigma_{Vi}\}$ has dropped from 0.0061 to 0.0041. Therefore, it is safe to declare that both CM and VM have approximately the same impact on the $\max\{\sigma_{Vi}\}$. However, CM has a better performance on the average of voltage standard deviations, $\text{avg}\{\sigma_{Vi}\}$, in comparison to VM. For instance, the $\text{avg}\{\sigma_{Vi}\}$ was found to be 0.0017 with one CM in comparison to a 0.0024 with one VM. In

summary, the type of the measurement does not have a remarkable impact on the maximum of the voltage standard deviation of whole system (all phases and nodes).

Second Experiment

In this experiment, the number of real-time measurements showed an increase in value, for example Cases 2 and 3 would have 3 CMs and 4 VMs; respectively. Voltage standard deviations for these cases are shown in Figure (5-21), in the case of 1.2% accuracy measurement class and 50% error on load estimation. As it is depicted in Figure (3-24), $\max\{\sigma_{Vi}\}$ for both cases have similar values 0.0025 and 0.0026 for Case 2 (3 CMs) and Case 3 (4VMs); respectively. Otherwise, the average voltage standard deviation in Case 2, 0.0015, is lower than that of Case 3, 0.0018. It is worth mentioning that in this specific experiment, a pattern is similar to that of the previous experiment was observed. By considering these two experiments, it can be noted that there is no significant difference between CM and VM performance when they are put in their appropriate places at the network, considering the maximum of the voltage standard deviations.

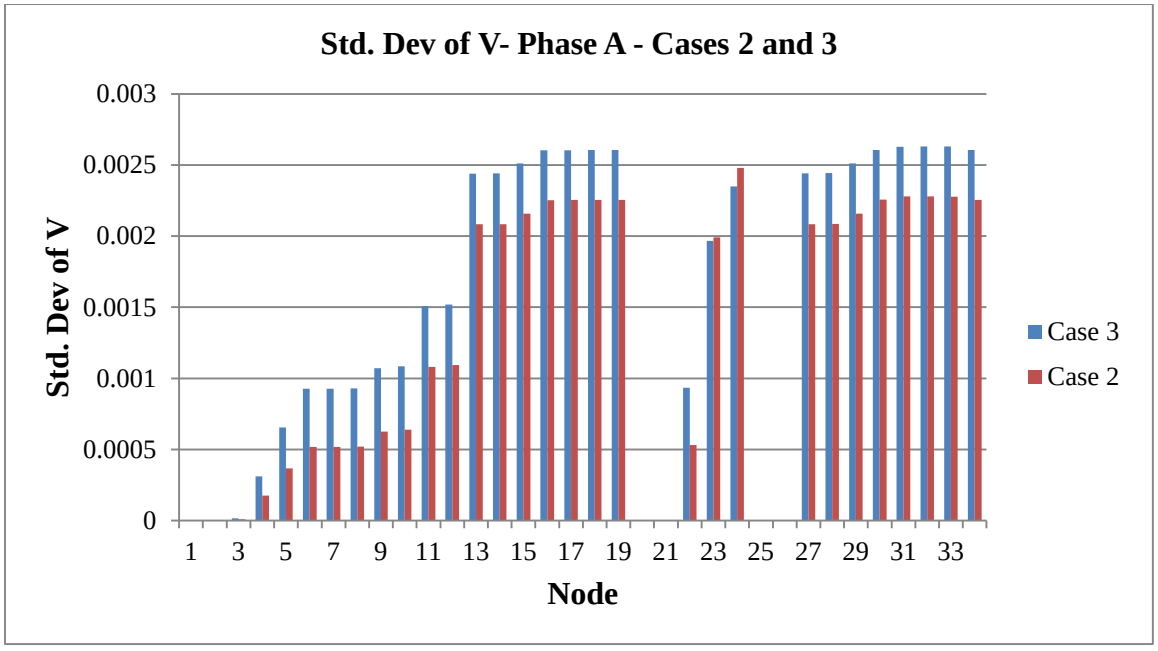


Figure (5-21): Comparison of voltage standard deviations for Cases 2 and 3.

To recall this fact, one of the important factors on standard deviation calculation is the way to model the load estimation uncertainty in addition to the level of error, factors that should be considered to analysis.

5-7 Observations

Based on the different case studies (Cases 1, 2 and 3) and setups for sensitivity analysis as well as other investigations, the following observations were obtained:

Observation 1: Based on impact study of load estimation uncertainty within different models (Models 1-3), it is found out that the load estimation accuracy has a direct impact on the voltage estimation accuracy. As long as the accuracy of the load estimation declines, the accuracy of the voltage profile estimation decreases as well. Therefore, in order to achieve the best possible voltage estimation, maintaining the most accurate load estimation is necessary.

Observation 2: It is clear from Cases 2 and 3 by having four and five measurements on the feeder; the accuracy class of the measurement devices which provides the real-time measurement for state estimator would have a direct impact on accuracy of the estimated interesting quantities (i.e. voltage profile). For example in Case 3, VM accuracy class has been changed from 1.2% to 0.3%, which in turn has affected the accuracy of the voltage estimation by becoming in the acceptable range (from 0.0029 to 0.0019 at node 24 for phase A). To summarize, adding more accurate measurement devices (sensors) provides more precise estimation for voltage profile along the feeder.

Observation 3: In regard to the CM allocation at the feeder, our study has shown that when CM moves from the middle of the feeder (near to the big loads) at branch 14 towards the end of the feeder (less load) at branch18 (i.e. setups 1 and 2 in Case 2 and moving one CM along the feeder from branch 18 to branch 13), the accuracy of the voltage profile estimation has dropped. In addition, the standard deviation of estimated voltages has become greater than the maximum acceptable standard deviation for voltages. For example, at node 19, standard deviation has increased from 0.0021 to 0.0029. As a matter of fact, the meter has become closer to node 19 at the end of the feeder, but the accuracy of the estimation has decreased. This

pattern has been observed in different load models. Therefore, CMs have better effects on the accuracy of the state estimation when they are added before the load centers and at the beginning of laterals provided that power measurements are available at the substation. It should be mentioned, this observation has been brought out in other references [32, 33].

Observation 4: Regarding the VM allocation at the feeder, the study at hand has shown that the standard deviation of the voltage profile estimation has increased when the VM moves from middle of the feeder (near to the big loads) at node 14 to the beginning of the feeder at node 8 (i.e. set-ups 1 and 2 in Case 3 and moving one VM along the feeder from node 33 to node 12), standard deviation, in addition the accuracy of estimated voltages has decreased. For example, at node 26-phase B, standard deviation has increased from 0.0019 to 0.0039. Actually the VM in this study has become farther to node 26, which has led to a decrease in the accuracy of the estimation. This situation has been observed in both load error models. Therefore, it is implied that VMs have a better impact on the accuracy of the state estimation when they are added at the end of the main feeder and the end of other laterals provided that the power measurements are available at the substation.

Observation 5: By adopting the two different experiments to investigate the impact of different types of measurements on SE performance, it was found that the impact of CM or VM on VVC objective ($\max\{\sigma_{Vi}\}$) was roughly the same when these types of measurements were located in their best spots. However, adding a CM provided a higher reduction on the average voltage standard deviation along the feeder in comparison to VM.

Observation 6: Based on the market research online in addition to the interview, it is found that the total cost was related to the number of measurement points on the circuits covering the cost of communication as well as other costs. All in all, the cost of the VM was roughly two thirds of the CM cost. Therefore, installing the VM at the feeder was approximately cheaper. On the contrary, the preferred location for VM would be at the end of the feeder or laterals. In case of any failures or disconnection take place at the network, the operator does not obtain any picture from these changes. In this condition CM could provide valuable information for monitoring the feeder.

5-8 The Proposed (Mixed) Meter Placement Scheme

Based on the aforementioned observations from extensive experiments and studies, a set of rules was developed for meter (sensor) placement to estimate the voltage profile along the feeder. In this approach all kinds of available measurements, i.e. VM and CM, were considered to for placement on the system. Following to that step, it is important to reach a compromise between cost and accuracy, therefore one algorithm for meter elimination will be proposed to remove the measurements with low contribution on voltage profile estimation. In other words, this proposed (mixed) scheme has two steps which are:

5-8-1 Step1: Basic Rules for Meter Placement Scheme

Here, it is assumed that power measurements are available at the beginning of the feeder. Then, other CMs and VMs will be placed on the distribution feeder based on the following set of rules which merges above observations.

Rule 1: 'Determine the load zones along the feeder.' Each load zone has similar loading conditions in three phase areas. For one phase / two phase loading area, same zoning will be considered like Zones in Figure (5-26). These zones can be found by analyzing the switch locations along the feeder.

Rule 2: 'Put CMs at the beginning of each load zones (three, two, and one phase load areas).' In addition, for very big load spots along the feeder, one other measurement could be placed. These measurements will provide data especially for feeder voltage profile monitoring.

Rule 3: 'Put VMs at the end of the main feeder and laterals with more than or equal to two nodes'. Voltage measurements at these places are desirable for VVC application. These additional measurements improve the accuracy of the voltage profile estimation.

It may be hard to justify that placing all the meters in the above scheme would be economical. In addition, by placing more meters, the voltage standard deviation profile will be estimated more accurate than the target value (maximum acceptable voltage standard deviation, $\overline{\sigma}_V$). Therefore, some of the meters are to be eliminated from the basic metering scheme. This procedure will be explained extensively in the next part.

5-8-2 Step2: Meter Elimination Procedure

From step 1, candidate measurements were selected to be placed on the feeder. However and as a second step, it is required to eliminate any extra measurements. For this purpose, let the types and places of nc candidate measurements, initially considered for installment, be given and the sets ncm, nvm contain the current and the voltage measurements; respectively. The general meter placement problem could be written as an optimization problem as follows:

$$\min\ f_O = \sum_{i=1}^{ncm} d_{c,i}.C_{CM,i} + \sum_{j=1}^{nvm} d_{v,j}.C_{VM,j} + C_0$$

$$s.t. \quad \mathbf{M}_1(\hat{\mathbf{x}}, \mathbf{z}) = 0$$

$$\hat{\boldsymbol{\sigma}}_V = \mathbf{M}_2(\hat{\mathbf{x}}, \mathbf{z})$$

$$\max\{\hat{\boldsymbol{\sigma}}_V\} \le \bar{\sigma}_V$$

$$d_{c,i}, d_{v,j} \in \{0,1\}$$

where:

C_0 is the fixed cost for power measurements and other requirements at substation,

$d_{c,i}$ corresponds to the decision that current measurement, i, is placed,

$d_{v,i}$ corresponds to the decision that voltage measurement, j, is placed,

$C_{CM,i}$ is the cost of the current measurement, i, placed on the designated branch,

$C_{VM,j}$ is the cost of the voltage measurement, j, placed on the designated node,

$\hat{\mathbf{x}}$ is the vector of the estimated system state from BCSE,

$\mathbf{z}$ is the vector of measurement for BCSE,

$\hat{\boldsymbol{\sigma}}_V$ is the calculated standard deviation of voltage at system node from BCSE,

$\mathbf{M}_1$ is the non-linear function of SE to calculate $\hat{\mathbf{x}}$ from ,

$\mathbf{M}_2$ is the non-linear function for $\hat{\boldsymbol{\sigma}}_V$ calculation from $\hat{\mathbf{x}}$ and $\mathbf{z}$, and

$\bar{\sigma}_V$ is the target value of voltage standard deviation for VVC application.

This is a non-linear, integer programming problem. Decision variables, $d_{c,i}$ for current measurement and $d_{v,i}$ for voltage measurement, are to be used to show which measurements are places along the feeder. Cost of the different current measurements is considered to be the same for all locations, i.e. $C_{CM,i} = C_{CM,j}$ for all i and j. In addition, costs of voltage measurements for all locations are assumed to stay constant as well. For VVC application, voltages must be estimated within certain accuracy. Here, the metric for the accuracy is sought to be the standard deviation, where a higher standard deviation means lower accuracy. Therefore, the maximum of standard deviation of the estimated voltages, $\max\{\hat{\boldsymbol{\sigma}}_V\}$, must be less than the target value. Vector of voltage standard deviations is calculated from the output of the state estimation (i.e. BCSE). Functions $\mathbf{M}_1$ and $\mathbf{M}_2$, which they are non-linear, present the relationship between the input measurement vector, $\mathbf{z}$, and estimated system state, $\hat{\mathbf{x}}$. Target value, $\bar{\sigma}_V$, is the maximum acceptable standard deviation for voltage estimation which is calculated from the VVC goal (voltage bandwidth). At the end, C_0 shows the cost of system requirements for providing the real-time measurement for the central control and power measurements at the substation which is assumed as a constant cost.

We follow a general approach in solving the master problem, which is an integer programming problem. Firstly, a decision graph is constructed, then an efficient search scheme to place the meters (i.e., to determine their numbers and places) is to be developed. Let there be nc initially given candidate meters. Ultimately, all possible decisions about choosing the meters for placement among the candidate measurements can be arranged as a decision graph assuming one decision is made at a time. Such a graph is shown in Figure (5-22) for $nc = 3$.

In Figure (5-22), each node represents a particular decision, $\mathbf{d} = [d_1\ d_2\ d_3]^T;\ d_i \in \{0,1\}$, where, $d_i = 0$ means that the meter is not selected, and $d_i = 1$ means it is still a candidate. A branch from a node to another show the fashion in which the transition could be conducted: taking out the candidate meter whose number is shown on the branch. Such a relationship is indicated in graph theory by calling a node and all the nodes coming from it, a *parent* and its *children*,

correspondingly. To find the local optimal set of the measurement, it is essential to run the Brute-force/ exhaustive search with 2^{nc} search order.

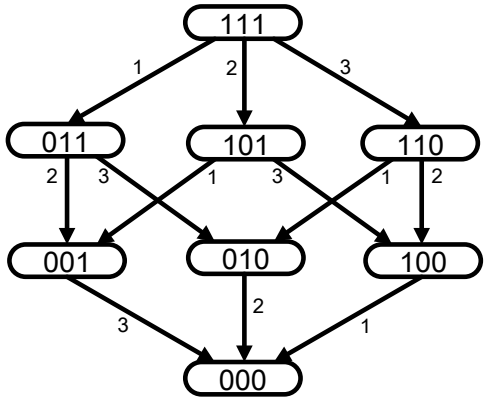


Figure (5-22): Decision graph for three variables [108].

Here, rather than employing the general search techniques directly, the children of a given node (parent) are sorted first according to their contribution to the objective or interested parameter. We use a "*sorting procedure*" to find the final measurement set which works like a "steepest descent" method in discrete case which was proposed in [108]. The procedure is as follows:

- Given a meter which is identified by its decision, $\mathbf{d}$ and the solution of the corresponding Problem, (i.e., the $\hat{\sigma}_{Vi,\max} = \max\{\hat{\boldsymbol{\sigma}}_V\}$ and the objective f_O)
- for all existing meters,. $k = 1,...,nc \;\; s.t. \;\; d_k \neq 0$
  - Construct $\mathbf{M}_1$ and $\mathbf{M}_2$ by removing the measurement k and keeping the rest measurements.
  - Calculate the new maximum voltage standard deviations, $\hat{\sigma}^k_{Vi,\max}$, and the new objective $f_O{}^k$ (this requires BCSE solutions: $\mathbf{M}_1$ and $\mathbf{M}_2$)
  - Let the contribution of the meter k to the interesting parameter and the objective be

$$\Delta f_O{}^k = f_O{}^k - f_O \text{ and } \Delta\hat{\sigma}^k_{Vi,\max} = \hat{\sigma}^k_{Vi,\max} - \hat{\sigma}_{Vi,\max}$$

- Sort and rank the existing meters by using $\Delta\hat{\sigma}^{k}_{Vi,\max}$ (when the cost of the different types of the measurement are the same the $\Delta f_O{}^{k}$ becomes the same, therefore the most important metric, $\hat{\sigma}^{k}_{Vi,\max}$, is used to rank the meters)

The $\Delta\hat{\sigma}^{k}_{Vi,\max}$ can used to pick the child to branch out (i.e., meter to take out) then the child with the smallest $\Delta\hat{\sigma}^{k}_{Vi,\max}$ is more likely the one who contributes to the problem the least.

Figure (5-23) illustrates this procedure, first the impact of each measurement in the current candidate set is assessed by removing the measurement and then calculating the resulting change in the accuracy of the voltage estimates, i.e. $\hat{\boldsymbol{\sigma}}_V$.

This search is of order nc^2 and requires at most $nc(nc+1)/2$ problem solutions in compare with an exhaustive search with 2^{nc} searches. Therefore this sorting procedure is computationally attractive. However, the search can be made much faster by reaching to the target value of $\max\{\hat{\boldsymbol{\sigma}}_V\}$.

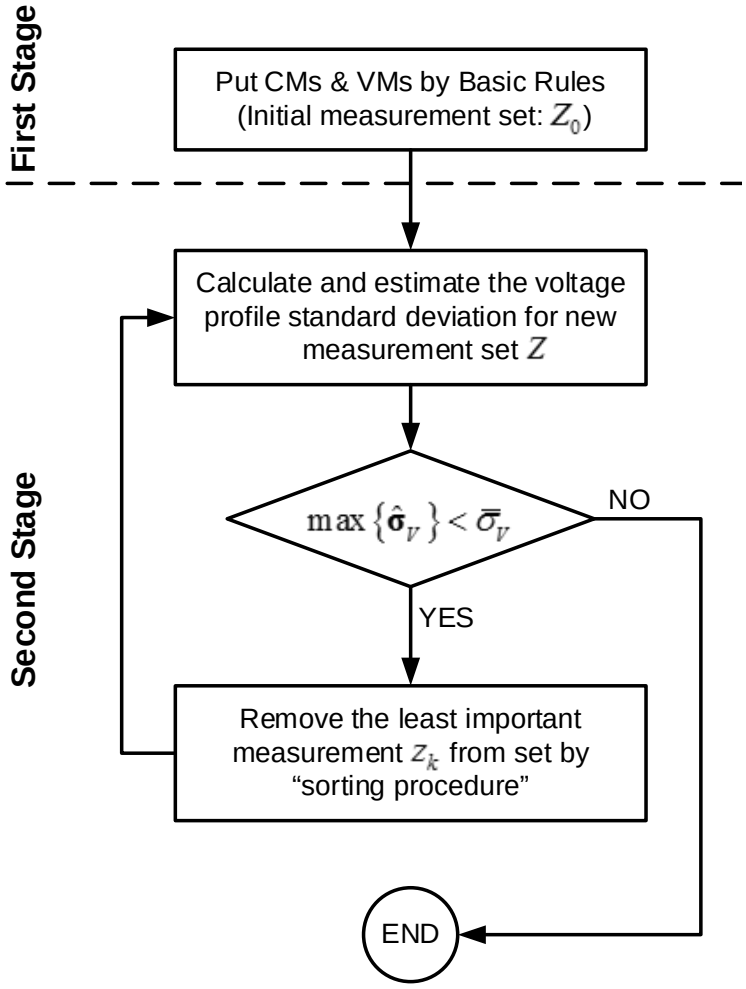


Figure (5-23): Algorithm of the mixed meter placement scheme

5-9 The Low Cost Meter Placement Scheme (VM based)

Based on the market search and observation 6, it was found voltage measurements are cheaper than the current measurements. Therefore, only the voltage measurements are placed on the feeder beside of power measurements in this scheme and we called it "Low Cost Meter Placement Scheme." This scheme has two steps:

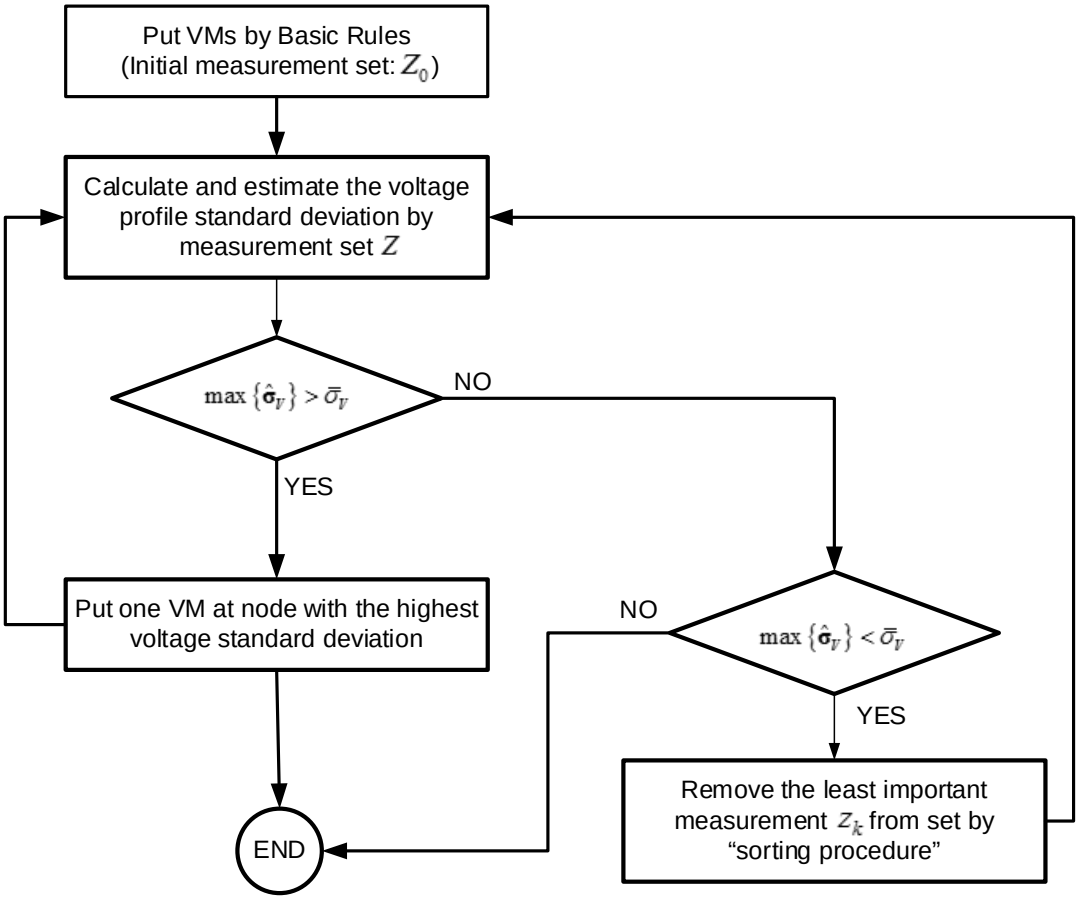


Figure (5-24): Algorithm of the low cost meter placement scheme.

Step 1: The voltage measurements are placed by this rule, 'Put VMs at the end of the main feeder and laterals with more or equal than two nodes'. Voltage measurements at these places are desirable for voltage profile estimation.

Step 2: In this step, we want to check whether that the goal of the meter placement problem has been fulfilled or not. In case of the maximum of the voltage standard estimation is greater than the target value, $\hat{\sigma}_{Vi,\max} > \bar{\sigma}_V$, then there is need to put extra voltage measurements on other nodes with this approach: putting extra voltages until the objective will be fulfilled. For case of $\hat{\sigma}_{Vi,\max} = \bar{\sigma}_V$, there is no need to add/eliminate any measurements. When the maximum

of the voltage standard estimation is smaller than the target value, $\hat{\sigma}_{Vi,\max} < \bar{\sigma}_V$, then there is a venue to run the "sorting procedure." The sorting procedure is explained with the details in the proposed meter placement scheme part. This meter placement approach together with voltage measurements are explained in this part and the main stages to run this scheme is shown in Figure (5-24). This scheme was adopted to use cheaper measurement devices in order to reduce the cost of the measurement setup for the given distribution feeder.

5-10 The Robust Meter Placement Scheme

Those previous methods are focus on finding the minimal number of needed measurement to estimate the voltage profile within certain accuracy. When one of the measurements, such as z_i, from measurement set, Z, is not available because of communication loss or malfunctioning of the meter. In this case, SE cannot provide accurate enough voltage profile for VVC application. Therefore, there is a need to have a robust scheme to loss of one measurement, i.e. N-1 redundancy.

The proposed solution for this issue is based on incorporating this N-1 redundancy check in the search method to find the significant meters from the initial set of measurements, Z_0. In this N-1 redundancy stage, every available measurement should be removed from the measurement set. Then, checking the $\max\{\hat{\boldsymbol{\sigma}}_V\}$ with the target value to find the quality of the voltage estimation with other N-1 measurements. Figure (5-25) shows this algorithm to find the robust set of the measurement.

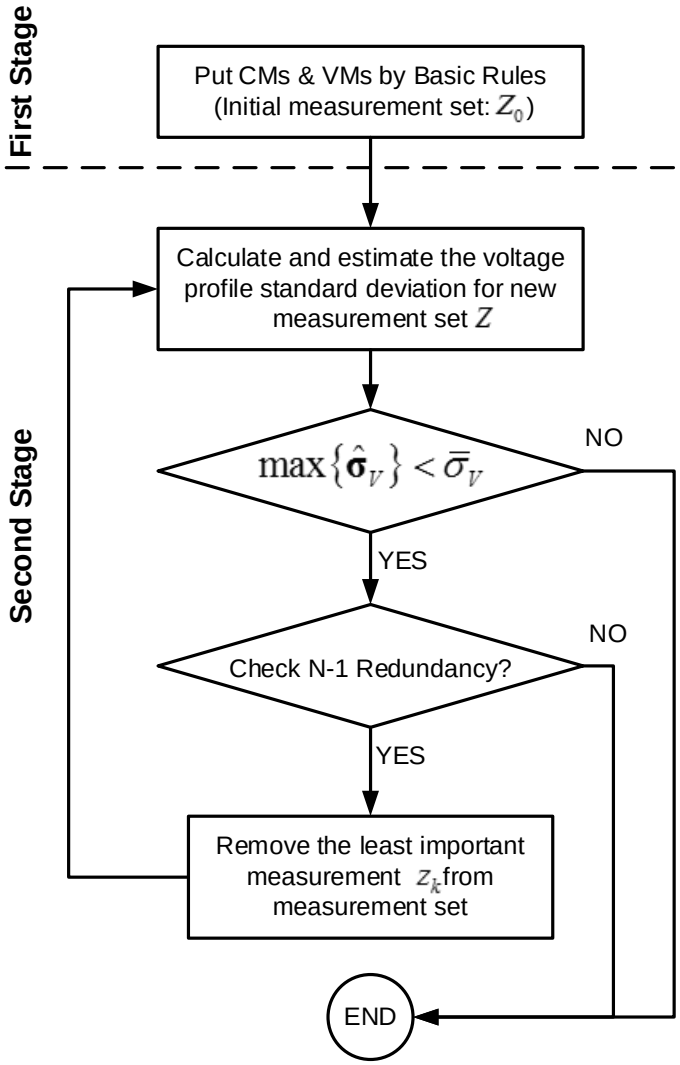


Figure (5-25): Algorithm of the robust meter placement scheme.

5-11 Test Results and Discussions

In this part, the low cost meter placement scheme and the proposed rule based meter placement method are implemented on a test feeder (IEEE 34 node feeder [87]) to estimate the voltage profile with certain accuracy. The results of these two methods will be given and discussed here.

5-11-1 Calculation of $\bar{\sigma}_V$ for VVC Application

To regulate the voltage along the feeder requires the good estimate of the voltage with certain accuracy. Output of state estimator will be used in Volt/VAR control; therefore, voltages must be estimated in the applicable range for VVC application. The common practice for VVC with VR, the considered bandwidth to the voltage variation is 2 V in 120 V scale [8, 109]. By applying the Empirical rule (three-sigma rule [110, 111]) "states that for a normal distribution, nearly all values lie within 3 standard deviations of the mean;" voltages must be estimated within this standard deviation:

$$\bar{\sigma}_V = \frac{BW}{6} = \frac{2/120}{6} \cong 0.00278 \ p.u.$$

This result gives the maximum acceptable standard deviation for voltage profile estimation along the feeder for VVC application.

5-11-2 Results for the Proposed Meter Placement Method

In this section, the results of the proposed meter placement scheme after adopting the scheme are presented. The first step is to put the measurements on the prototype system (IEEE 34 node feeder [87]) which is shown in Figure (5-26) using basic rules.

Figure (5-26-a) shows the trend used to apply rules 1 and 2 for the given system. Zones 1-4 are three phase load zones and Zone 5 is one phase load zone (phase A). Then, Figure (5-26-b) shows the sample feeder with VMs at the designated nodes where nodes 18, 33, and 28 are three-phase nodes and node 24 is a one-phase node (phase A) by following the rule 3. Eight measurements are placed along the feeder for monitoring the voltage profile.

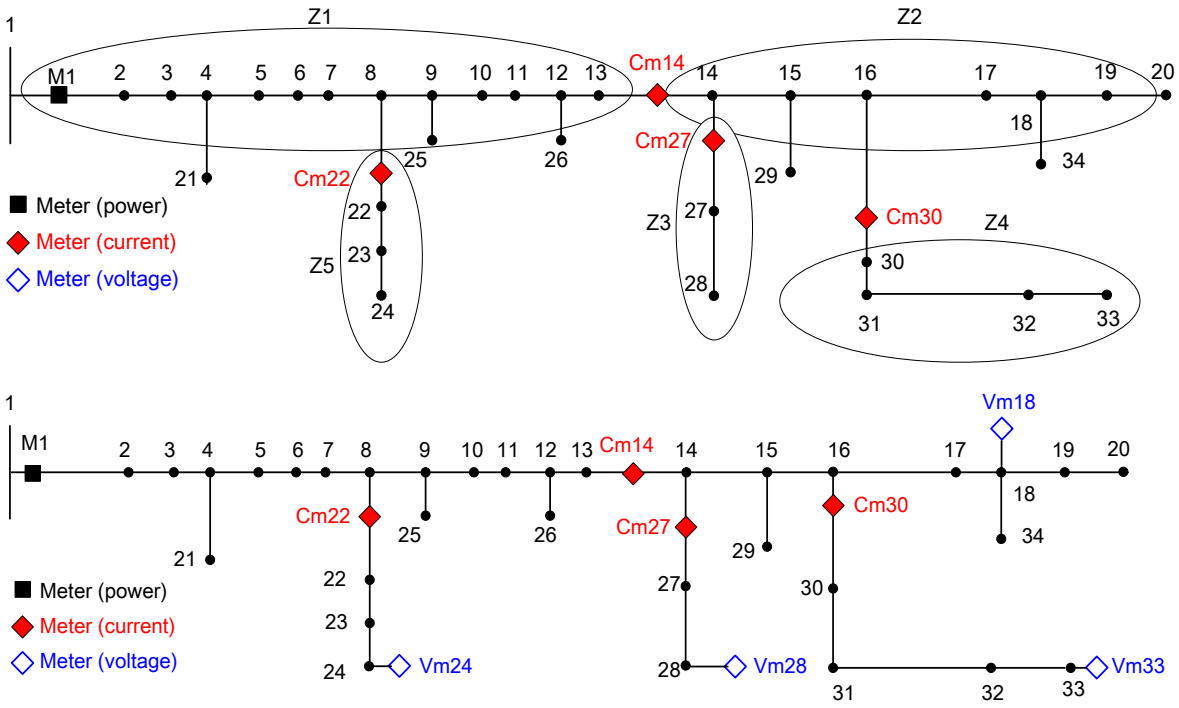


Figure (5-26): Load zones (a) and Initial set of measurements, Z_o, four CMs and four VMs (b).

By following the basic rules of meter placement, eight meters were placed along the prototype feeder. It may be hard to justify placing all the meters the above scheme will be economically feasible. In addition, with these eight meters, the voltage profile was estimated more accurate than the target value (maximum acceptable voltage standard deviation, $\overline{\sigma}_V$). Therefore, some of the meters will be eliminated from the basic metering scheme. In this process, meters are ranked based on their contribution on maximum voltage standard deviation along the feeder. Initial measurement set, Z_0, is constructed by adding the CMs and VMs through the basic rules. Then, meter elimination algorithm will be applied to find out the significant meters. Meters are ranked based on this fact that how much $\max\{\hat{\boldsymbol{\sigma}}_V\}$ changes by removing them. For example, first rank meter is the meter with the most impact on the objective, in case of the prototype feeder Cm14 has the first rank which means by removing this meter from

measurement set the maximum of the voltage standard deviation is increased much. With eight measurements at the system, the $\max\{\hat{\boldsymbol{\sigma}}_V\}$ is around 0.001, so there is a possibility to eliminate some sensors to reach the target value, 0.0027.

By running the first step for sorting procedure, the least important meter is Vm28. So this meter was removed from the measurement set and maximum voltage standard deviation has reached to 0.0013. After removing the least important meter, the accuracy of the voltage estimation was checked to see whether the $\max\{\hat{\boldsymbol{\sigma}}_V\}$ increased from the target value. When removing of the least important measurement from measurement set causes $\max\{\hat{\boldsymbol{\sigma}}_V\}$ increases from target, then this measurement will be backed to the measurement set. This procedure will be followed by removing the meters with low performance until $\max\{\hat{\boldsymbol{\sigma}}_V\}$ will reach the target value.

Table (5-3): Steps of meter elimination procedure for IEEE 34 node feeder with 8 measurements.

Step 1		Step 2		Step 3		Step 4		Step 5		Step 6	
8 Meters	Rank	7 Meters	Rank	6 Meters	Rank	5 Meters	Rank	4 Meters	Rank	3 Meters	Rank
Cm27	5	Cm27	6	Cm27	5	Cm27	5	Cm27		Cm27	
Cm14	1	Cm14	1	Cm14	1	Cm14	1	Cm14	1	Cm14	1
Cm30	6	Cm30	7	Cm30		Cm30		Cm30		Cm30	
Cm22	2	Cm22	2	Cm22	2	Cm22	2	Cm22	2	Cm22	2
Vm18	3	Vm18	3	Vm18	3	Vm18	3	Vm18	3	Vm18	2
Vm33	7	Vm33	4	Vm33	4	Vm33	4	Vm33	4	Vm33	
Vm28	8	Vm28		Vm28		Vm28		Vm28		Vm28	
Vm24	4	Vm24	5	Vm24	6	Vm24		Vm24		Vm24	

This procedure has been run for six times to remove the least significant measurements from the initial measurement set, Z_0, with eight measurement (4 CMs and 4 VMs). At the end, three measurements have remained in measurement set which is shown by steps in Table (5-3). This final measurement has two CMs and one VM. VM has located at the end of the main feeder (node 18) and CMs have located at Cm14 (beginning of the main load area at the feeder) and Cm22 at the beginning of the one phase lateral. With this measurement set, voltages along the feeder can be estimated with desired accuracy and the $\max\{\hat{\boldsymbol{\sigma}}_V\}$ is 0.002.

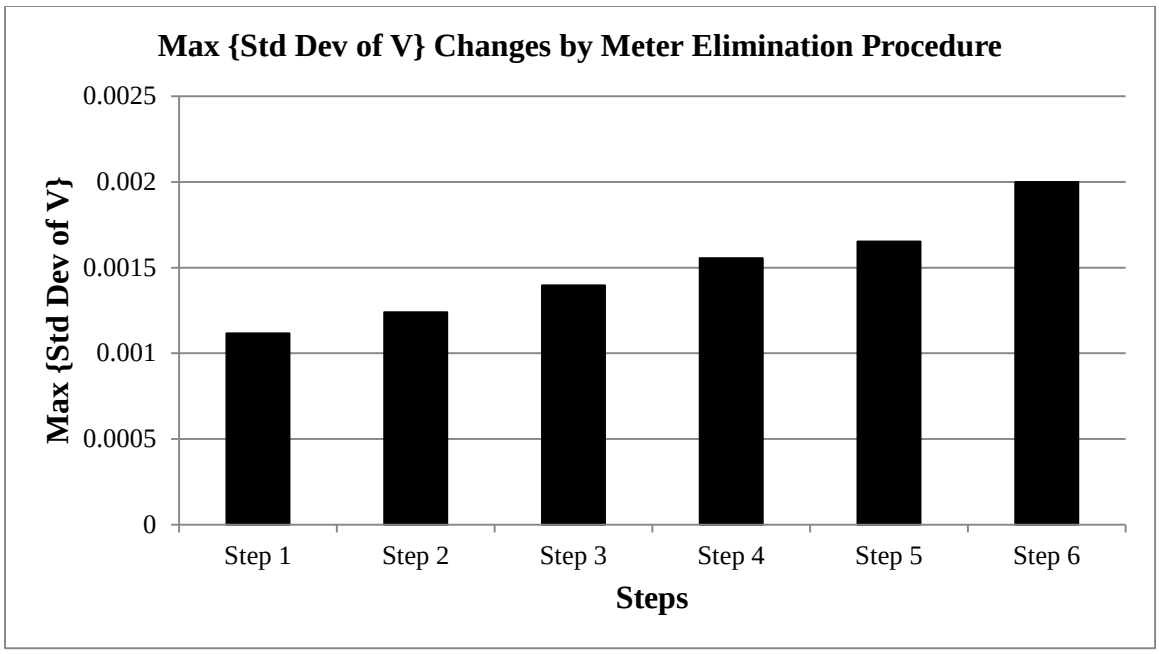


Figure (5-27): $\max\{\hat{\boldsymbol{\sigma}}_V\}$ changes by Meter Elimination Procedure.

Figure (5-27) shows how the $\max\{\hat{\boldsymbol{\sigma}}_V\}$ increased by eliminating one of the measurements at each step. In step 6, if any measurement was eliminated from the measurement set, the maximum voltage standard deviation will reach more than the target value, 0.0027. In the first step, one VM was removed and then one CM, this pattern has been seen the same up to sixth step. From VM set, Vm33, Vm28, and Vm24 was removed, and Cm30 and Cm27 was

eliminated from initial current measurement set. Final measurement set for voltage monitoring is shown in Figure (5-28).

From the cost analysis point of view, the cost of one VM and one CM for each phase are 2 (cost unit) and 3 (cost unit); respectively from market research and the sixth observation. In the final stage, these three sensors are placed along the test feeder with 18 (cost unit) as total cost.

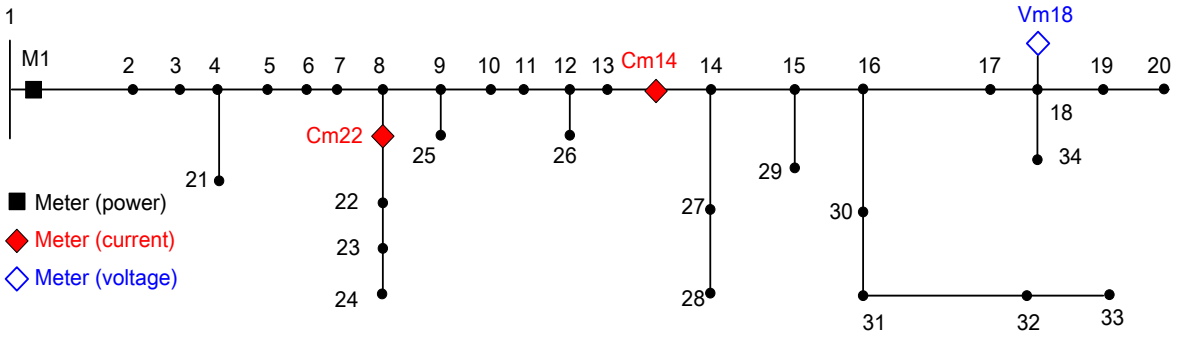


Figure (5-28): Final measurement set for prototype feeder (2 CMs and 1 VM).

5-11-3 Results for the Low Cost Meter Placement Method

This section previews the results of the low cost meter placement scheme after adopting the scheme have been presented. The first step is to put the voltage measurements on the prototype system (IEEE 34 node feeder [87]) which is shown in Figure (5-29) by one rule (VM placement).

In fact, after following the second step of the low cost meter placement method, it was found that there is no need to add any new voltage measurements to the system first, and none of the assigned meters in the first stage can be eliminated from the initial measurement set. Therefore, this scheme has four voltage measurements where three of them are three phase voltage measurements and one is one phase VM. The $\max\{\hat{\boldsymbol{\sigma}}_V\}$ is 0.0027 for node 24 at phase A. In addition, the total cost of this scheme is 20 (cost unit).

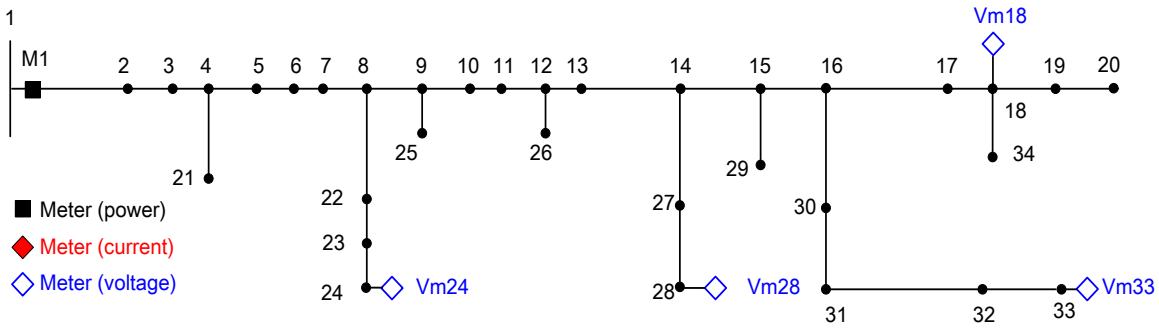


Figure (5-29): Initial (final) set of measurements, Z_o, for low cost meter placement approach.

5-11-4 Results for the Robust Meter Placement Method

This section previews the results of the robust meter placement scheme after adopting the aforementioned scheme. Here, like the mixed meter placement scheme, VMs and CMs are placed on the prototype feeder based on the basic rules of the meter placement. So, initial measurement set, Z_0, is constructed by four CMs and four VMs. By basic rules, initial meters have been employed along the feeder in the first stage. In the second stage, the meters with less impact on the voltage measurement estimation will be identified and removed from the initial measurement set. N-1 redundancy stage first checks that removing one measurement can cause any problem in the quality of the voltage estimation through BCSE. So, in the first step, one measurement can be removed from the measurement set and voltage still can be estimated by other measurement within the desired accuracy. Moreover, these eight meters are ranked based on the sorting procedure, measurement Cm14 is the first rank which means this measurement has the most impact on the voltage estimation. On the other hand, Vm28 has the least impact on the voltage estimation in comparison with other available measurements. Therefore, Vm28 can be removed from the measurement set. At the second step, there are seven measurements: three VMs and four CMs. First, N-1 redundancy stage checks the feasibility of measurement reduction. In other words, this stage wants to know that any measurement can be omitted from the measurement set and the performance of the SE cannot

be affected. Based on this checking it is possible to remove one measurement from measurement set and any measurement loss from measurement set cannot effect of the SE performance. Now, the least important measurement is Cm30 can be removed from the measurement set without having significant impact on the SE performance for voltage estimation.

At the third step, there are six measurements: three VMs and four CMs. First, N-1 redundancy checks the possibility of omitting one measurement from the measurement set. By checking this redundancy, it is detected that there is no chance to remove one measurement and still have a chance to loss of any measurement, then the voltage profile can be estimated within the desired accuracy. When sorting procedure runs, it found that the least impact measurement is Vm24. In case of Vm24 omitting from the measurement set, we should still have a robust measurement set with loss of any measurement that voltage can be estimated with the desired accuracy. The problem occurs with loss of Cm22 in any condition, which cause the uncovering the fifth zone, Z5. By loss of this measurement, we cannot estimate the voltage with confidence of losing any measurement from the available measurements. Table (5-4) depicts these steps clearly. Therefore, at this step, the search algorithm to find the smaller measurement set has been stopped and final measurement becomes Cm14, Cm22, Cm27, Vm18, Vm24, and Vm33. Final measurement set for voltage monitoring has been shown at Figure (5-30).

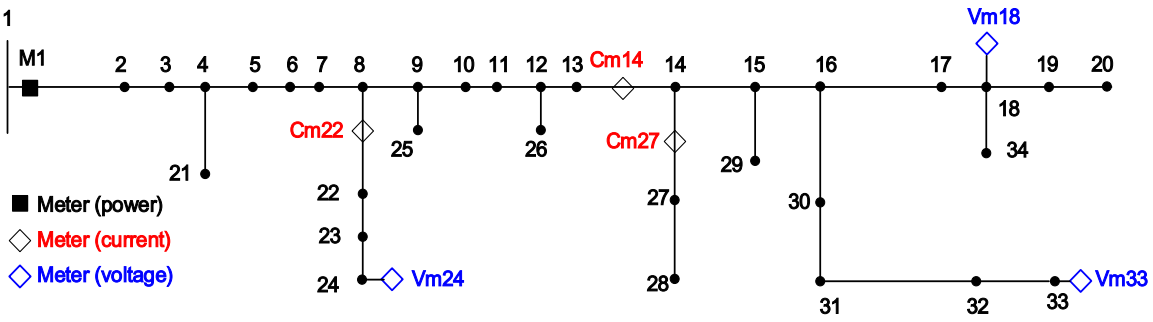


Figure (5-30): Final set of measurements for the robust meter placement algorithm.

In other words, Cm30 and Vm28 are omitted from the initial measurement set. It is clear that the measurement set is larger for this scheme in comparison with the proposed mixed meter placement scheme.

Table (5-4): Steps of meter elimination procedure for the robust meter placement scheme.

Step 1		Step 2		Step 3	
8 Meters	Rank	7 Meters	Rank	6 Meters	Rank
Cm27	5	Cm27	6	Cm27	5
Cm14	1	Cm14	1	Cm14	1
Cm30	6	Cm30	7	Cm30	
Cm22	2	Cm22	2	Cm22	2
Vm18	3	Vm18	3	Vm18	3
Vm33	7	Vm33	4	Vm33	4
Vm28	8	Vm28		Vm28	
Vm24	4	Vm24	5	Vm24	6

In the first step, one VM was removed and then one CM, this pattern has been seen the same in the mixed meter placement scheme. This method has been stopped at step 3 in comparison with the sixth step reduction of the first meter placement scheme. From VM set, Vm33 just has been removed, and Cm30 just has been eliminated from initial current measurement set. In addition, the total cost of this scheme is 35 (cost unit). This method provides the most expensive approach to place the meters on the feeder, because this method guarantees the quality of the voltage estimation with loss of any measurement from the real-time measurement set.

5-11-5 Comparison of the Proposed Schemes and Discussions

All the proposed meter placement methods are implemented on the IEEE 34 node feeder. Here, we want to compare these three meter placement schemes. As it is assumed that the

measurement set at distribution level includes power measurements (usually at the substation end), current measurements on certain line sections, and voltage measurements at nodes.

In “low cost” scheme, all measurements are VMs and power measurements are available from the substation. On the other hand, all kinds of available measurements (CMs and VMs) have been considered for placement in the “proposed” scheme. In addition, the mixed meter placement scheme has been strengthened in the “robust” method to loss of any measurement, N-1 redundancy plan. These three methods are implemented in the same prototype feeder.

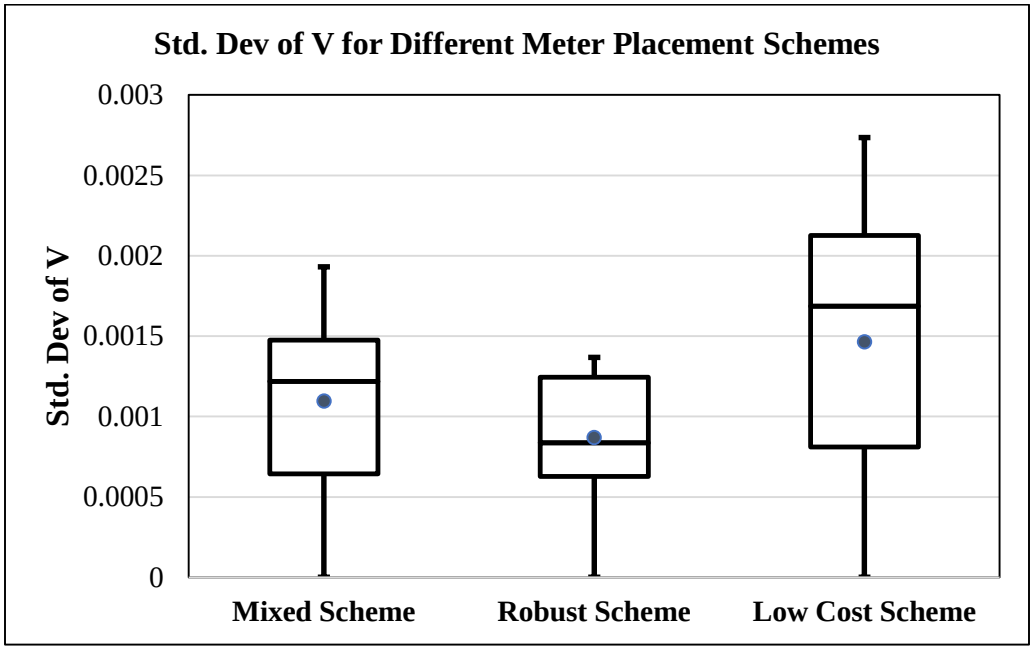


Figure (5-31): Boxplots of voltage standard deviations for different proposed meter placement schemes.

Figure (5-31) shows the voltage standard deviation profile for all phases, i.e. A, B, and C. Averagely, the proposed scheme provides a better performance in comparison to the low cost scheme where voltage profile of the system has been estimated by a better accuracy. In phase C, all of the voltage standard deviations for the proposed schemes are smaller than the low cost scheme. We can say that VMs have local impact on voltage profile estimation; otherwise CMs

impact on voltage estimation is more effective. Therefore, implementing a mixed measurement set, similar to the one in the proposed scheme, provides a better measurement set for voltage feeder monitoring with fewer number of measurements (10 measurement points for low cost scheme and 7 measurement points for the proposed one in phase-based calculation) in comparison to all the measurements considered only VMs, such as low cost scheme. When one measurement has been lost from the mixed measurement, there is not a guarantee to estimate the voltage in the desired accuracy. Therefore, there is a need to implement the robust scheme to consider the impact of the measurement lost. In this robust scheme, there are 14 measurement points which is higher than the other methods.

In addition, $\max\{\hat{\boldsymbol{\sigma}}_V\}$ of the mixed method, the robust method, the low cost method are 0.0019, 0.0014, and 0.0027 respectively as shown in Figure (5-31) which shows a better performance for the robust scheme which this maximum voltage standard deviation is far from the target value, 0.0014 vs 0.0027.

On average, as it is shown in blue points on the boxplots in Figure (5-31), average of the voltage standard deviation for the mixed method, the robust method, the low cost method are 0.0011, 0.0009, and 0.0015 respectively. This shows a better performance for the robust scheme which has more real-time measurements in the measurement set, as it is expected.

From the cost analysis point of view, the cost of one VM and one CM for each phase are 2 (cost unit) and 3 (cost unit) respectively from market research and sixth observation from previous sections. The total cost for the low cost approach, the mixed scheme, and the robust scheme are 20 (cost unit), 18 (cost unit), and 35 (cost unit); correspondingly. Hence, from an economical viewpoint, the proposed scheme offers a better scheme in comparison to the low cost scheme and the robust scheme. In addition, the most expensive measurement set provides from the robust method with 35 (cost units) which is expected to provide the enough real-time measurements to estimate the voltage within the desired level with the possibility of any measurement loss. We should mention that all these methods are heuristic methods and developed from the extensive searched and studies. Hereby, these approaches provide feasible solutions for the meter placement problem to find the minimal set of the measurement set from

available metering options on the distribution systems although they cannot guarantee the optimality of the proposed solutions.

5-12 Impact of the Load Variations on the Proposed Meter Placement Schemes

All the proposed meter placement methods are implemented on the full load condition of the feeder. One question can be asked that these final measurement sets from these proposed algorithms provide the enough measurement to estimate voltage for VVC. Consequently, the final measurement sets are should be tested in regard of the load variations. In this thesis, it has been shown that DT loads vary during the days. Two common condition is that the half load condition, i.e. 50% of full loads, and one quarter load, i.e. 25% of full loads. In addition to these cases, load growth can happen. So, impact study of load growth is necessary. Here, two cases have been considered first with 10% growth and second with 25% growth.

From the proposed meter placement schemes, there are two different schemes: the mixed meter placement scheme and the robust meter placement scheme. Final measurement sets for these different methods consist three measurements and six measurements, respectively. The mixed measurement set consists Cm14, Cm22, and Vm18. Furthermore, the robust measurement set consists Cm14, Cm22, Cm27, Vm18, Vm24, and Vm33.

The four aforementioned cases are tested for these two measurement sets. Figures (5-32) and (5-33) show the range of voltage standard deviation for these cases with fixed load error level. As the boxplots of the voltage standard deviation for these cases show that the $\max\{\hat{\boldsymbol{\sigma}}_V\}$ change by the load variation, but they are still in the required standard deviation, $\bar{\sigma}_V$. Figures (5-32) and (5-33) demonstrate that when the load level is less than the peak load, voltages are estimated with desired accuracy for both measurement sets. In case of load growth, for these common cases: 10% and 25% load growths, voltages can be estimated within the required accuracy. These studies show that the proposed meter placement outputs are robust to the load variations and load for the prototype feeder.

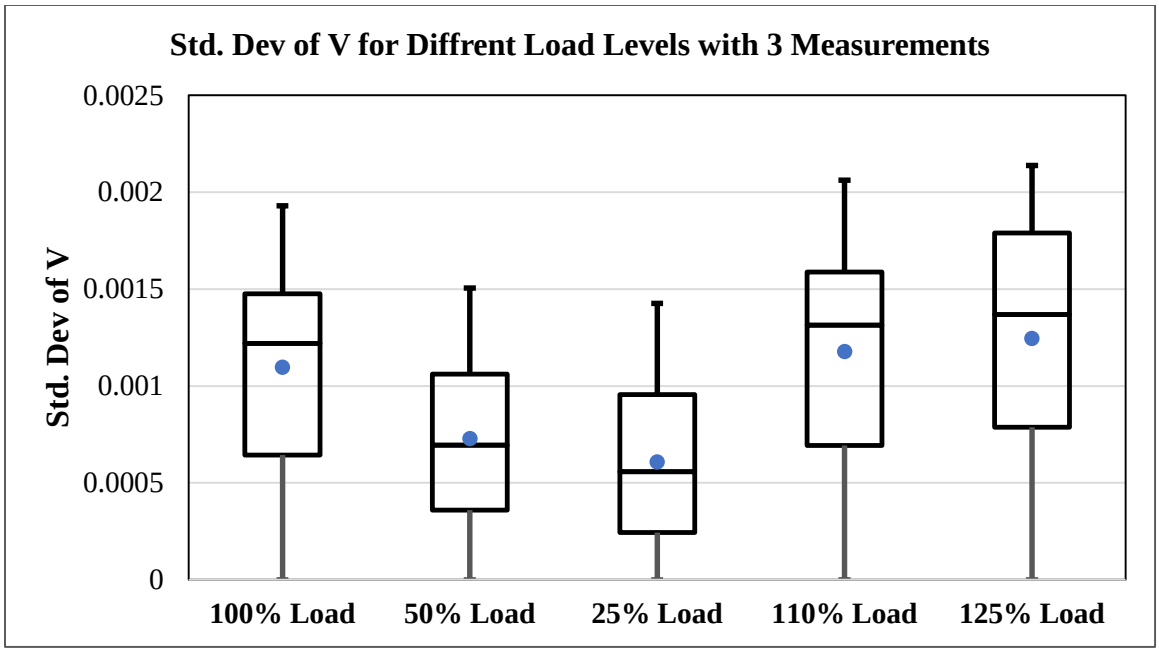


Figure (5-32): Boxplot of the voltage standard deviation to load variations with the mixed measurement set (3 real-time measurements), blue point: mean.

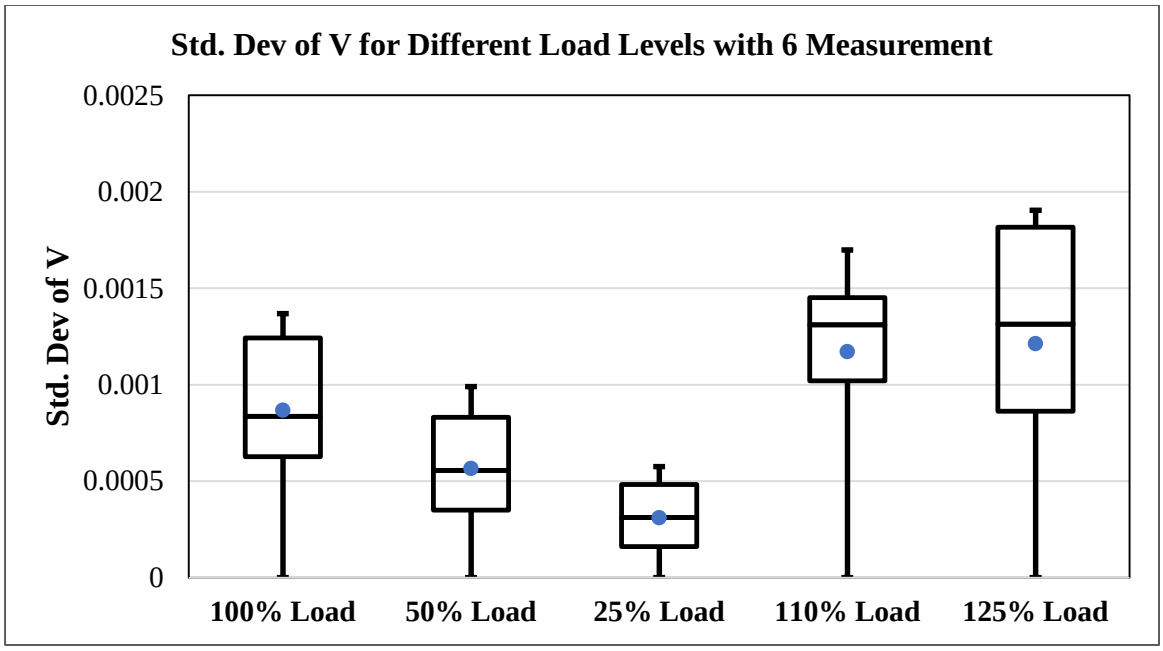


Figure (5-33): Boxplot of the voltage standard deviation to load variations with the robust measurement set (6 real-time measurements), blue point: mean.

5-13 Impact of the PV Generation on the Proposed Meter Placement Schemes

All the proposed meter placement methods are implemented on the full load condition of the feeder without any generation from PV generation at customer sides. One question can be asked that these final measurement sets from these proposed algorithms provide the enough measurement to estimate voltage for VVC with PV generation from customer's side. Consequently, the final measurement sets are should be tested in regard of the load variations. Here, the proposed approach in [129] is implemented [130]. First, it is aimed to find the normal cases for the feeder operation in the presence of the PV generation at every load point of the prototype feeder. In order to decide upon the normal operation conditions that may happen, actual load and PV daily profile is obtained. Average residential load data of November 1999 is provided by Progress Energy, and real power output of the roof top PVs in FREEDM System Center on a typical sunny day of November 2010 is also recorded [120, 129, 130]. The resolution of load profile is 30 minutes and the resolution of PV profile is 5 minutes. The raw data is analyzed and represented per unit. As shown in Figure (5-34), residential load usually has two peaks, one is in the morning and the heaviest load happens after work. The energy requirement pattern mainly depends on customer characteristics. PV starts to produce real power from 7:45 am until 6:15 pm. During noon time when customers needs are quite flat and solar radiation is sufficient, PV real power generation exceeds the load consumption from 11:00 am to 4:00 pm, which will result in reverse power flow and potentially high voltage.
Based on these observations, the following possible system operating conditions are simulated by scaling the PV and loads [129]:

- Case 1: Full load and no PV generation,
- Case 2: Full load and medium PV generation, i.e. 50%,
- Case 3: Medium load, i.e. 50% load, and medium PV generation,
- Case 4: Medium load, high PV generation, i.e. 100%, and
- Case 5: Medium load, no PV generation.

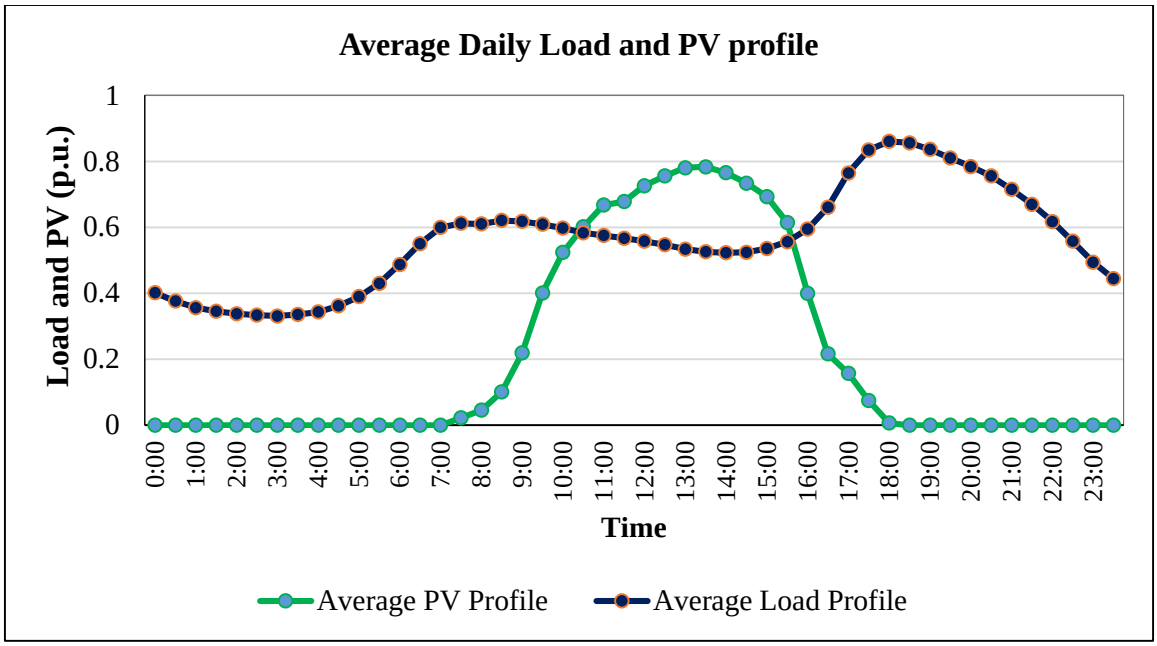


Figure (5-34): Average daily load profile and PV generation profile in NC [120, 129].

Actual loads and PVs output are expressed as a percentage of the peak load or its installation capacity. Full load is equivalent to a 100% of the nominal load at each node. At 100% conditions, loads are the heaviest and PVs are giving out its maximum real power. Case 3 presents the condition that the most PV generations have consumed on the load side when the load level is around 50%. Case 1 represents the peak load condition without any PV generations around 7 pm. Case 2 represents heavy load condition with certain PV real power output, like 5:30 pm. Case 3 is the crossing point when customer need equals to the PV generation, such as 11:00 am and 4:00 pm. Case 4 is the most severe situation that may happen in a conventional distribution feeder when the difference between PV generation and load consumption is maximum, such as 2:00 pm. Case 5 represents most of the day when load is moderate and no PV generation, from 12:00 am to 6:00 am. Out of these five cases, case 1 and case 4 are two extreme operating conditions during normal operation. Case 1 may result in severe voltage drop and low voltage profile at the end of the feeder, while case 4 may cause high voltage. In this study, these assumptions have been made:

1) Each load bus has solar PV connected to it with installation capability the same as the peak real power of its corresponding load
2) Load and PV generation can be estimated with the same accuracy at DT level with assumption of the availability of the AMI data.
3) VR and capacitor banks are kept fixed in this prototype distribution system.

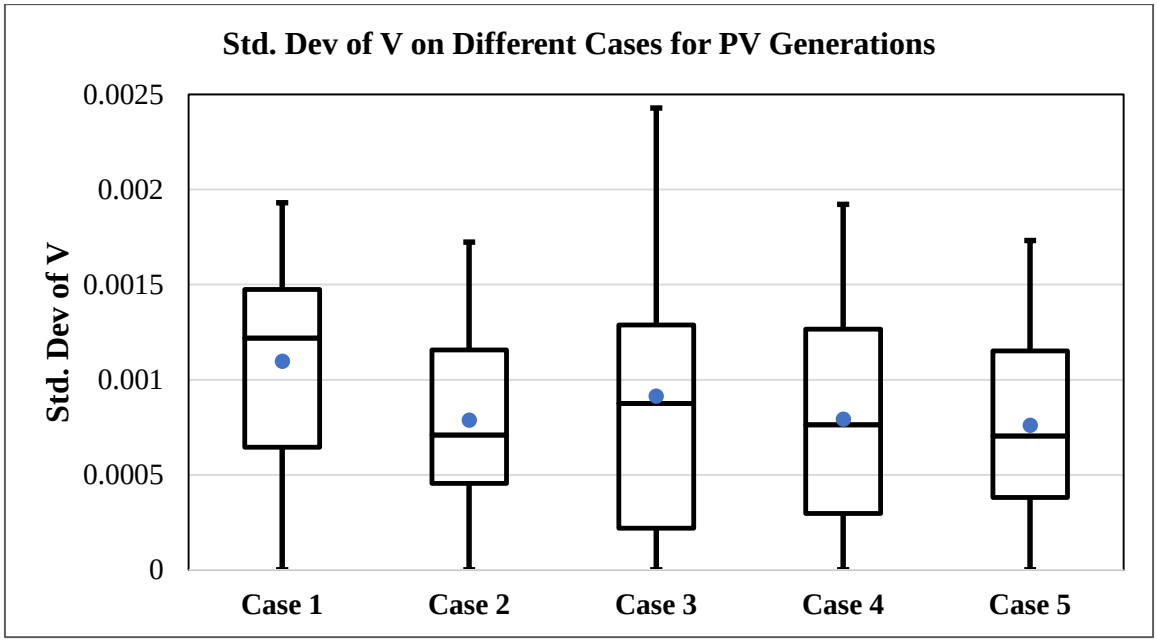


Figure (5-35): Boxplot of the voltage standard deviation for different PV generation cases with the mixed measurement set (3 real-time measurements), blue point: mean.

For these cases, the proposed meter placement algorithms are tested with this assumption that the net power of each DT with connected load and PV panel can be estimated with the same accuracy. Figure (5-35) shows the range of the voltage standard deviation with measurement set from the output of the mixed meter placement scheme. This measurement set has three measurements: Cm14, Cm22, and Vm18. As Figure (5-35) illustrates, voltage profile can be estimated within the desired accuracy according to the designed cases for considering the

different conditions of the PV generations on the feeder. Also, this figure presents these results too:

- The result of the Case 3 shows more variation in voltage estimation when active part of load has been covered by generation of PV, i.e. zero injection condition. In this case, accuracy of voltage estimation is less than other cases, because of this reason that small variation on the load or generation can change this zero injection and cause the remarkable voltage variation on the prototype feeder.
- The result of the Case 5 looks like the result of the Case 2.
- In all cases, all meter placement schemes provide enough data for voltage estimation within desired accuracy.

5-14 Summary and Conclusions

In this chapter, the meter placement problem on distribution feeders for different applications, such as load estimation and state estimation, was reviewed. For VVC application, this problem was formulated as a non-linear integer programming problem which requires running a Brute-force search to find the best location, type, and number of measurements to estimate the voltage profile within a certain accuracy. Due to the massive number of needed searches, one heuristic meter placement with two stages was proposed based on the extensive studies and searches to determine the better places to place the different types of measurements (VMs and CMs) and reduce the number of measurements by sorting procedure which is explained in details in previous parts.

IEEE 34 node feeder was used to test the proposed method and compare it with another scheme, i.e. low cost scheme. Voltage standard deviations were calculated using the output of the BCSE and Monte Carlo simulations. Three different schemes are proposed: low cost scheme, mixed measurement scheme, and robust scheme. Each of these proposed schemes has its own cost and benefits. It is obvious that the proposed scheme offers the practical and

feasible solution to the meter placement problem for VVC application. This chapter has these contributions in regard of the meter placement problem on the distribution systems:

- Performing a comprehensive sensitivity analyses to extract the initial rules of the meter placement. The following sensitivity studies were conducted:
- Adopting the branch current based SE (BCSE), the unbalanced nature of distribution systems in topology and loading was addressed.
- Estimating the voltage profile with an accuracy of +/- 1 V (with respect to 120 V base) using these proposed schemes with few real-time measurement along the feeder.
- Enhancing the meter placement schemes for loss of one real-time measurement from the measurement set.
- Investigating the impact of load variation on the proposed schemes.
- Investigating the impact of PV generation at customer sides.
- Considering different metering options to be placed on the feeder.

Chapter 6: Topology Processing for Distribution Systems by BCSE

This chapter of the dissertation focuses on the topology error detection and its identification on distribution systems using the branch current-based state estimation to offer the robust state estimation for monitoring and control of the distribution systems effectively. Firstly, the concept of the robust state estimation and bad data will be reviewed. Secondly, topology processing at transmission level of electrical power systems is surveyed. Finally and for the rest of the chapter, two common topology changes on the distribution system are considered to propose the algorithm to enhance the BCSE for detection and identification of these topology changes.

6-1 Overview

SE relies on the basic assumption that we know the exact network model so that we can write the measurement functions without any doubt. However, in the practical world, the status of switching devices is unknown or, for some reason, the current values in the database are under suspicion. Hence, the risk of assuming the wrong status for the switching device cannot be completely avoided, leading to potential topology errors. These errors usually cause the state estimate to be significantly biased. As a result, bad data detection & its identification routine may erroneously eliminate several analog measurements which appear as interacting bad data, yielding to an unacceptable state. Therefore, there is a need to develop effective mechanisms intended to detect and identify this kind of gross error [54].

"Keep changing topology" is one of the distribution system characteristics which have been listed in the first chapter. In transmission systems, the Topology Processor (TP) identifies the network configuration based on network connectivity model and dynamic switch status with highly redundant measurements. Distribution systems require a particular TP due to unavailability of switch status information. A few specific methodologies for distribution network are available at present. Therefore, with progressing of distribution automation, it is necessary to explore the existing transmission system TP techniques or develop new

methodologies for topology identification in order to carry out a meaningful state estimation. In this section, we will look at this problem in details and provide a proposal for topology processing in the distribution feeders with a limited real-time measurements and high uncertain load estimation [16, 17, 21, 62, 65-68].

6-2 Robust State Estimation on Power Systems

One of the essential functions of a state estimator is to detect measurement errors, and to identify and eliminate them if possible. Measurements may contain errors due to various reasons. Random errors usually exist in measurements due to the finite accuracy of the meters and the telecommunication medium [122-125, 15]. Large measurement errors can also occur when the meters have biases, drifts or wrong connections. Telecommunication system failures or noise caused by unexpected interference can also lead to large deviations in recorded measurements [15, 21]. Apart from these, a state estimator may be deceived by incorrect topology information which will be subsequently interpreted as bad data by the state estimator. Such situations deal with topology errors [15, 21].

6-2-1 What is “Bad Data”?

This term, “bad data”, mostly presents in technical literature (application point of view) and was mentioned by Schweppe in [3]. It was defined as: “In actual operation, some of the real-time meters or the corresponding communication links will fail at various times. If this failure is known to the computer, these are called lost data points. If the failure is not known, it is a bad data point”. In addition, [15] defines “bad data” as data that is much more inaccurate than is assumed by the mathematical models that may cause very poor estimates. Consequently, Koglin *et al.* prepared more detailed classifications [121] where he classified measurements into:

- Telemetered measurements,

- Virtual measurements which are this kind of information that does not require metering, e.g. highly accurate zero injection in passive (transit) nodes [15], and
- Pseudo-measurements with relatively low accuracy (e.g. injections/load of neighboring systems, forecast values, manually entered values etc.).

Normally, only telemetered measurements are subject to bad data analysis. “Bad data” can be caused by:

- Parameter errors such as false tap changer position, incorrect transformer models, incorrect branch impedances,
- Incorrect network topology, possibly due to manual updated switching positions or incorrect topology description, and/or
- Erroneous measuring devices or telecommunication errors.

“Bad data analysis” normally consists of three steps:

1) Detection procedure to determine whether bad data are present,
2) Identification procedure to determine which measurements/parameters/switch states are bad,
3) Elimination procedure to eliminate the influence of bad data on the state estimate.

If the estimated state remains insensitive to bad data, then the corresponding estimator will be considered (statistically) robust. Unfortunately, robustness is commonly achieved at the expense of computational complexity. The topic of robust estimation is quite broad and covered in several books and papers in literature. We will here focus on adopting the state estimation for detecting and identifying the topology changes [122-125].

Some bad data are obvious and can be detected and eliminated a-priori state estimation by simple plausibility checks [15, 21]. Negative voltage magnitudes, measurements with several orders of magnitude larger or smaller than expected values, or large differences between incoming and leaving currents at a connection node within a substation are some examples of such bad data. Unfortunately, not all types of bad data are easily detectable by such means.

Hence, state estimators have to be equipped with more advanced features that will facilitate the detection and identification of any type of bad data.

6-3 Topology Error Detection and Identification at Transmission Power Systems Level

At the transmission level of power systems, due to the available redundant measurement, topology error processing is one of the well-defined and well-developed state estimation functions [15].

The conventional network topology processing (NTP) function monitors the status of switching and switchable devices, and determines the model input to the state estimator [15, 21]. Circuit breaker statuses, isolator switch statuses, and transformer tap positions are examples of real-time inputs used by the network topology processor. A conventional NTP defines the connectivity of the electrical network, taken as input a complete model of the system, comprising nodes and switching devices. The NTP cuts the node-switching- device model to a “bus-branch” model, where the concept of bus defines a maximal sub-network interconnecting nodes and closed switching devices only. The goal of the conventional NTP is to eliminate all switching devices from the network model, by instantiating their “open” or “closed” statuses [65]. Hence, reliable and prompt detection of the switching device statuses is crucial for accurate state estimation. The output of the state estimator is a critical input to nearly all other network analysis, security, control and stability assessment applications. This problem has been formulated as both least absolute value (LAV) and weighted least squares (WLS) state estimation problem [15, 21].

Techniques for topology error analysis can also be categorized based on the moment at which the analysis is performed:

1. A priori processing: Much in the same way as analog measurements are pre-filtered before entering the SE, the assumed status of CBs can be validated in advance by means

of local consistency checks, rule-based techniques combined with recorded information, etc.

2. A posteriori processing: The bad data analysis stage can be made more sophisticated so as to consider the possibility of topology errors being responsible for biased estimates. If the topology error is not clearly identified, i.e., there are several candidates, then the second SE run mentioned above, containing detailed models for the suspected areas, can be resorted to.

A priori processing methods by nature are based on fast, approximate techniques, usually applied at the substation level, and hence they may fail to properly identify the topological problem. On the other hand, post-processing methods rely on the results of a converged SE, which may not always exist in the presence of certain topology errors.

A rule-based system, taking into account the temporal evolution of measurements and switch positions, is proposed in [71]. The basic idea is that, while normal evolution of load gives rise to smooth or incremental changes in the measurement values, switching operations are usually accompanied by more abrupt deviations. Trying to emulate the way an engineer would locally infer the coherency of present and past information around a given switch, including alarm and event lists, a knowledge base containing over 120 rules is developed. Practical experience gained in the integration of this expert system into a Swiss control center is reported in [71]. The application of artificial neural networks (ANN) to this problem is explored in [112].

In one of the earlier papers on topology error detection, Irving and Sterling [126] have formulated the problem as a linear program. The solution to this problem automatically rejects the erroneous power measurements and switch indications by taking into account redundant information and natural circuit laws. The location and degree of inaccuracy of the erroneous data are also obtained automatically.

In [127], Singh and Alvarado have formulated topology processing as a least absolute value (LAV) state estimation problem by introducing the status variables in the nonlinear norm problem. Normalized residual based approaches for identification of topology errors have been reported in [15] and [64]. Both papers have introduced the detectability and non-detectability

of topology error conditions with the help of a linearized residual sensitivity matrix. The detectability of topology errors has been proven to be associated with the criticality of the elements.

6-3-1 Branch status errors

Branch status errors can be identified and corrected by means of either normalized residuals or state vector augmentation. Both solutions will be discussed separately.

1- Residual Analysis: This approach makes use of the results of a converged state estimation in order to detect branch errors. The effect of these types of errors on the measurement equations can be modeled in the following manner [15]:

$$H = H_e + E$$

where, H is the true Jacobian, H_e is the incorrect Jacobian due to topology errors, and E is the Jacobian error matrix. Substitution using the linearized measurement model, $z = Hx + e$, yields:

$$z = H_e x + Ex + e$$

The statistical characteristics of the new residual vector can also be derived as below:

$$r = z - H_e\hat{x} = (I - K_e)(Ex + e)$$
$$E[r] = (I - K_e)Ex$$
$$Cov[r] = (I - K_e)R$$

where: $K_e = H_e(H_e^T R^{-1} H_e)^{-1} H_e^T R^{-1}$. Using the linearized measurement model, each entry of the bias vector, Ex, can be written as a linear combination of errors in network branch flows. Let f be a vector of branch flow errors, and M be the measurement-to-branch incidence matrix. Then, the measurement bias , Ex, can be written as:

$$Ex = Mf$$

and the residual vector will be given by: $r = (I - K_e)Mf$

Now, the expected value of the normalized residuals can be rewritten in terms of the branch flow errors:

$$E[r^N] = \Omega^{-\frac{1}{2}}(I - K_e)Mf = Sf$$

where: $\Omega = diag\{Cov[r]\}$, and $S = \Omega^{-\frac{1}{2}}(I - K_e)M$ is the sensitivity matrix for r^N with respect to branch flow errors, f.

Hence, topology error detection can proceed based on the normalized residual test, assuming that bad data in measurements have already been identified and eliminated.

2- State vector augmentation: The status of a non-zero impedance branch can be represented by means of a single integer variable, k, that multiplies all branch admittances. When this variable is considered, the π model is composed of the following admittances:

$$series: \quad (g_{ij} + jb_{ij})k$$
$$parallel: jb_{ij}^{p}k$$

Clearly, $k = 0$ represents a disconnected branch whereas $k = 1$ should be used if the branch is in service. The idea is to include the variable k into the state vector for any suspected branch.

As explained in [15, 66, 67], when the WLS estimator is used, it makes no sense to enforce simultaneously the two contradictory constraints. If they are considered as very accurate pseudo-measurements and are given the same weight, then $k \approx 0.5$ will be obtained, which is useless. On the other hand, if both constraints are ignored, the estimated value of k will be dictated by the analog measurements. A value of k close to 1 is an indication of the branch being in service, while a value approaching 0 implies a disconnected element. In practice, however, owing to measurement errors, the estimated value of k may significantly differ from

0 or 1. Even worse, in the presence of nearby bad data, the estimated value of k may approach 0.5, in which case the status of the element is ambiguous. Additionally, the estimated values of the remaining state vector components would be less accurate than if k had the right integer value. In order to elude these potential problems, the quadratic constraint:

$$k(1-k)=0$$

is used to enforce the estimator to converge to either of the two feasible statuses [15]. This idea can be applied to any of the WLS state estimation algorithms. Although it is ideally suited to those methods in which equality constraints are explicitly modeled, this constraint can also be handled as a very accurate measurement, provided orthogonal factorization of the Jacobian is employed in order to prevent the intrinsic ill-conditioning of the normal equations when large weighting factors are adopted.

All solutions mentioned in literature are based on the high redundancy of measurements to state variable, i.e. at least in the range of 1.7 to 2.2 [65], where they can add more state variables to the SE procedure and estimate the status of those variables. Therefore, the topology error detection and identification methods at transmission level cannot be applied to distribution networks with few measurements.

6-4 Topology Errors on Distribution Power Systems

“Keep changing topology” is one of the distribution system characteristics which has been listed in the first chapter. Proper operation of the distribution systems needs taking quick actions to restore continuity of electric power supply after forced outages and faults. Because of these faults and outages, as mentioned earlier, the distribution power systems are constructed by the feeders. In general, switching devices at distribution feeders are circuit breakers (CB), reclosers, sectionalizers, tie switches, and fuses as shown in Figure (6-1). Normally, fuses are not monitored by SCADA, in other words blown fuses are detected by sending the crew in the field based on the customers calls. On the other hand, CBs are monitored at the substation by

SCADA and it is possible that other switches like reclosers and sectionalizers are monitored too [16, 17, 21, 62, 65-68].

Figure (6-1) presents the different types of faults on the distribution systems. Some faults happen on the main feeder or laterals, like Fault 1 and Fault 2. These faults occur on the primary network of the distribution network. Here, these faults can be cleared through the feeder sectionalizer and the recloser functioning. If permanent faults happen some parts of the feeder become de-energized. Otherwise, faults can happen at the secondaries and they cause the outages for one set of the customers where they are energized from the specific distribution transformer. For example, several customers are connected through Fuse 201 to one lateral. When the fault happens as shown in Figure (6-1), all loads are disconnected from the electricity supplied. Some faults can happen on the primary part of the feeder and cause the change of the feeder topology. Other faults can occur on the secondary side of the feeder and cause the customer (load) outages. For many distribution systems, however, there are few real-time measurements on the feeders that the first and often only indications of an outage are telephone calls from customers reporting the power outages. In the typical radial distribution networks, opening of a normally-closed switching device generally results in some electric power outages. The analysis performed by collecting and mapping multiple customer calls into a suspected network part and then devices, such as a fuse, is a case of network topology estimation. Most of the existing outage management systems (OMS) are still based on the process of call collection [1, 54, 65-67], which can take from tens of minutes to hours (if it happens at night for example) to identify the culprit device.

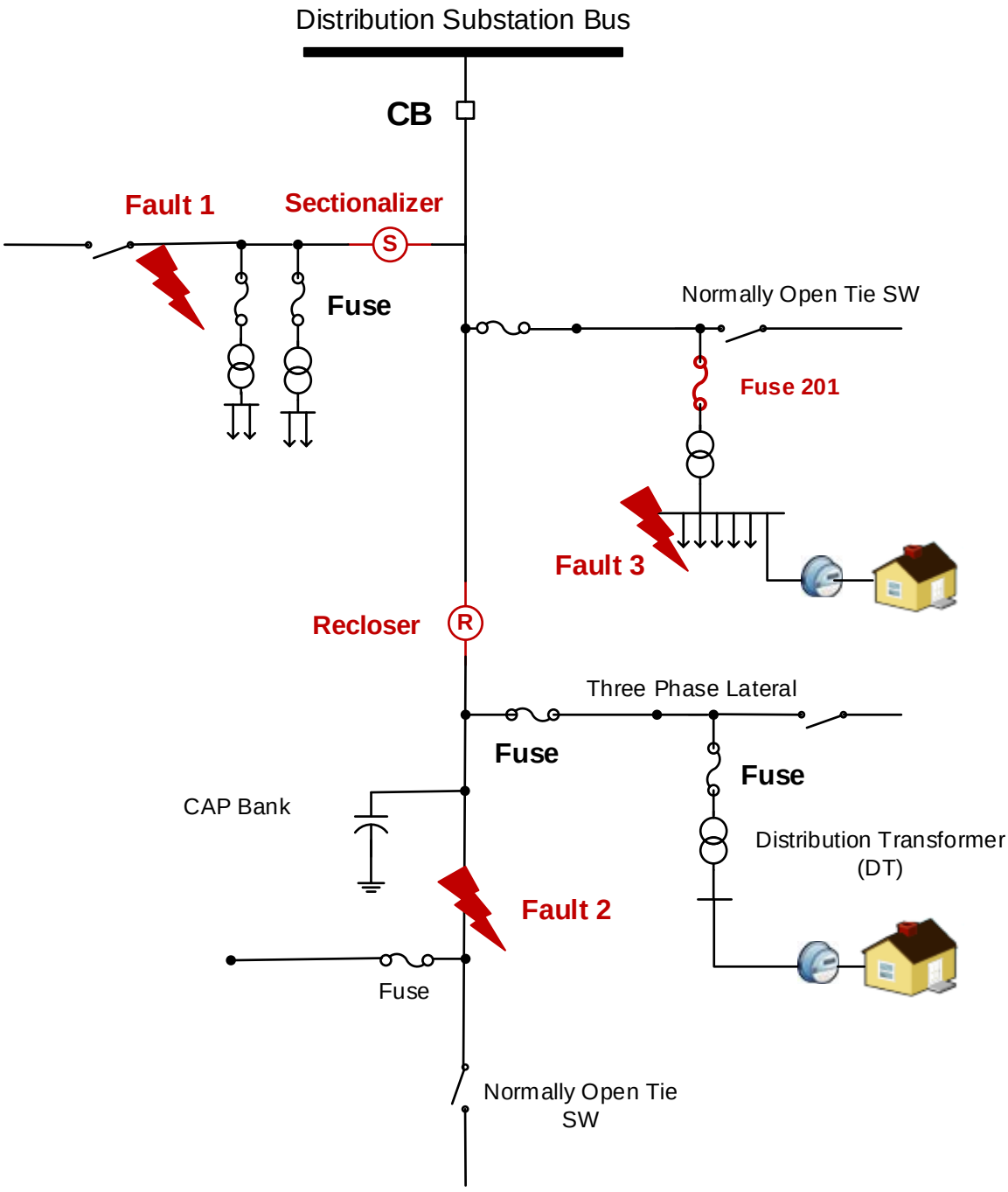


Figure (6-1): A sample distribution feeder with different faults at primary and secondary side of the network.

When a permanent fault, e.g. fuse blown, occurs on the distribution network, many customers connected to non-faulted ("healthy") parts of the feeder may experience a lengthy outage until repairs and restoration are made or until a switching takes place. This is shown in the time line depicted in Figure (6-2). Based on the EPRI white paper, 45-75 minutes are needed for power restoration of customers at a healthy part of the feeder through Outage Management System (OMS) [1]. One of the DMS functions as mentioned in first chapter is to restore power to customers served by healthy sections in less than one minute (before field crews arrive on to the scene) [1, 2]. In other words, monitoring of active distribution networks is essential to provide this service to the customers.

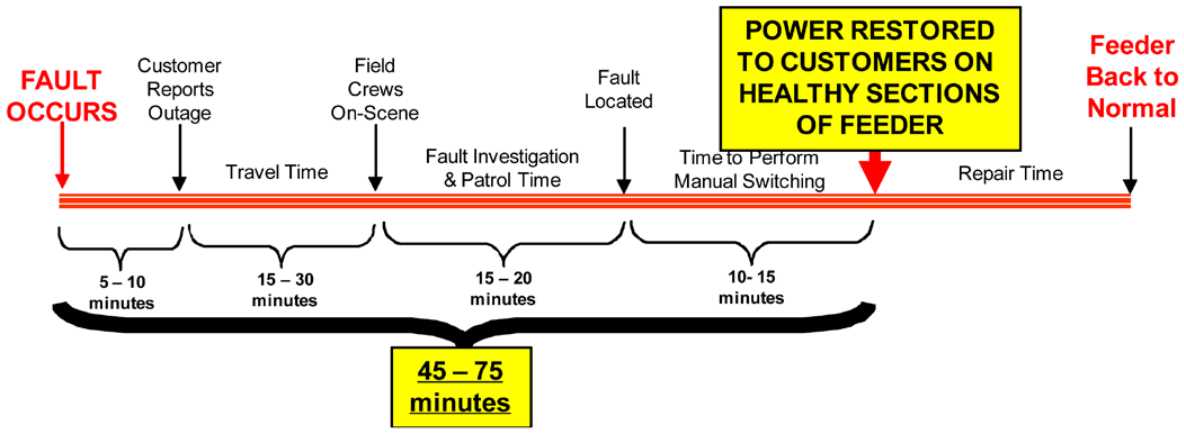


Figure (6-2): Time line for service restoration on distribution systems [1].

This monitoring framework provides electric distribution utilities with necessary information in order to locate faults more quickly and accurately, and automatically restore healthy portions of the distribution grid. For the sections that do require manual intervention, the above metering and grid information should help guide the crew to the fault's approximate location instead of having them manually locate the fault. Not only does this aid in reducing fuel costs and carbon monoxide emissions from crew trucks, this automated service also cuts customer

outage times by up to 30% [113], thus upgrading reliability indices (such as SAIDI, CAIDI) and improving customer satisfaction.

It is worth mentioning that some topology changes happen in order to balance the loads on the three phases. They can also happen for other reasons where they do not necessarily happen due the faults on the distribution systems.

One solution to detect and identify these topology changes on the distribution networks is to enhance the SE. Typically, this process happens in the SE at transmission system level. Here, the goal is to address these problems by adopting the BCSE to detect and identify the network changes and load outages.

6-4-1 Literature Review

In what follows, available literature on the approaches adopted to detect and identify the topology changes though state estimation processing on the distribution networks will be discussed. Previous literature shows that most of the proposed solutions for SE in distribution systems considered that the topology is accurate and fixed [18-47]. A few papers tried to address this challenge in the distribution networks by adopting the SE, such as M. Baran ***et al.*** [54] who presented the algorithm for topology error identification using the branch current state estimation, then, R. Singh ***et al.*** [66] proposed a recursive Bayesian approach for identification of the network configuration changes in distribution networks.

It should be mentioned that this problem has been approached by other methods in literature. For instance, Y. Sharon *et al.* presented a statistical method to detect and classify the changes in the status of switching devices on distribution networks by Support Vector Machine (SVM) method in 2012 [65]. Moreover, this problem has been approached though the linear programming based formulation in 1995 [128].

In this research, we have focused on the solutions based on the state estimation. The proposed method based on the recursive Bayesian method in [66] can be summarized. In this approach, different models representing various network configurations are stored in the form of a model

bank. The WLS estimators of all the models are run in parallel. All models use the same set of real measurements as common input. In addition to this, each model uses its own pseudo and virtual measurement sets which are prepared in accordance with the configuration of that model. The output of each estimator is compared with the common inputs in order to compute the error associated with each model. At a given point of time, only one model represents the correct configuration and later the conditional probability of this model being correct (given the errors in the output of all the models) attains its maximum value. The computation of the conditional probabilities for each model from the errors in estimates have been done by Recursive Bayesian Probability modeling. These drawbacks have been realized and listed:

1. In the model bank, a limited number of possible topologies has been considered to reduce the number of the different topologies which must be checked to find the current system topology. This can limit the operator to find the correct configuration of the system. But, it is suitable to reduce the computation burden of the algorithm.
2. For actual feeder with many switching devices at the feeder, constructing the bank of the network topology becomes time consuming procedures with exponential order for topology possibilities. In other words, it is needed to construct 2^{n_s} sample models, n_s number of switches, which is not a small number for actual feeders.
3. To find the most probable configuration, one hundred Monte Carlo simulations are used. In some cases, this number of simulations is not sufficient and further simulations are required which is not suitable for on-line applications and restricts the practicality of the proposed algorithm.
4. The model assumes balanced distribution system for performance test of the proposed approach which is not a practical assumption for the real distribution system which is a three-phase unbalanced system.

6-4-2 Topology Error Detection & Identification Using BCSE for Distribution Systems

Previous studies [82] have shown that checking only measurement consistency among the available measurements is not sufficient for detection and identification of the topology

changes. One of the proposed solutions is to enhance the SE for this purpose. The correct statuses of all CBs, reclosers, sectionalizers, tie switches, and fuses circuit breakers (CB) in the system are known all the time. However, in some rare cases, the assumed status of certain switches may be wrong, especially the switches without SCADA and fuses. When wrong statuses of the fuses or switches happen, the bus/branch model generated by the topology process is locally incorrect, leading to a topological error [54]. Unlike the parameter errors, topology errors usually cause the state estimate to be significantly biased [15, 82]. Accordingly, the bad data detection & identification routine may incorrectly reject several analog measurements which appear as interacting bad data, finally yielding an unacceptable state. It is also possible for the SE process to diverge, or have serious convergence problems, in the presence of topology errors. Therefore, there is a need to develop effective mechanisms intended to detect and identify these kinds of gross errors. For instance, Figure (6-3) shows that SW1 has been tripped and the zone of this switch has been de-energized, the gray area of the feeder. If the previous topology has been considered as the input of the SE, the result of the SE becomes erroneous and leads to wrong operation for the given system. Here, the zone of each switching device has been defined in Z_SW#, for example Z_F1 for fuse 1 zone.

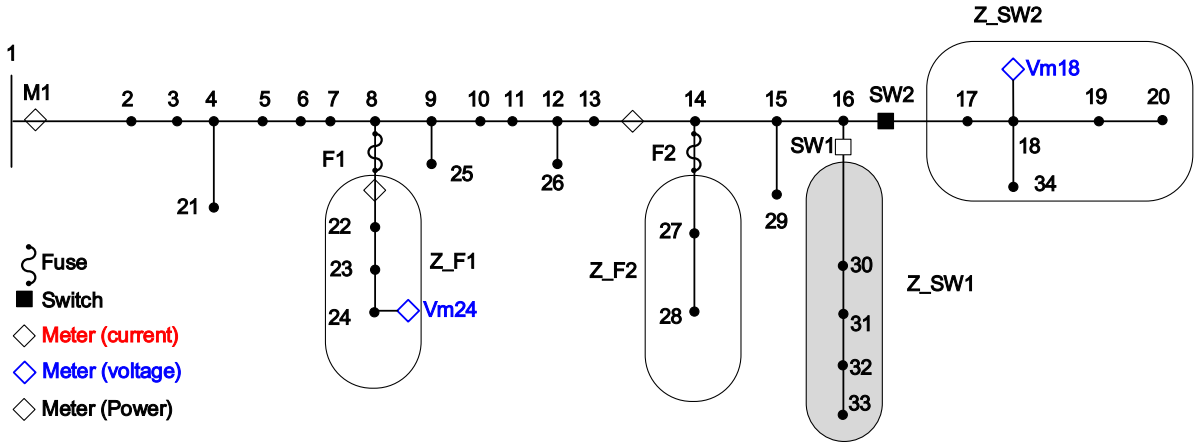


Figure (6-3): A sample distribution feeder with a fault at Z_SW1 and operated SW1.

The aim of this section is to present a solution to deal with topological errors and related matters using bad data detection & identification routine in the context of the Branch Current State Estimation (BCSE). The proposed method will be explained in the following section.

6-5 The Proposed Method for Topology Error Identification in Distribution Systems

The proposed method finds the topology changes by checking the residual of the estimation. When the system model is correct and the measurements are in acceptable ranges, the estimated states are acceptable from the WLS approach. On the other hand, if there is a bad data in the measurement set, it should be detected and then identified so that it can be removed from the estimator calculations to estimate the system state accurately. This is termed as bad data detection & identification in state estimation procedure [15, 21]. The bad data detection & identification procedure is presented as follows:

Step 1: Calculate the residuals $\hat{r}_j = z_j - \hat{z}_j$, where $\hat{r}_j$ is the residual of the measurement z_j which is estimated by $\hat{z}_j$, after running the BCSE program.

Step 2: Evaluate the weighted sum of residual squares $\hat{f} = \sum_{j=1}^{N_m} \hat{r}_j^2 / \sigma_j^2$, where σ_j^2 is the assumed error variance for the measurement j, and N_m is the number of the measurements.

Step 3: For the appropriate number of degrees of freedom $k = N_m - N_s$, where N_s is the number of the system states, and a specified probability α, i.e. p-value, determine whether or not the value of $\hat{f}$ is less than the critical value corresponding to α. In practice, this means we check that the inequality $\hat{f} < \chi^2_{k,\alpha}$, where $\chi^2_{k,\alpha}$ is the critical value of chi-squared distribution with k degrees of freedom (d.f.) at a certain

significance level, i.e. 0.05, is satisfied. If it is, then the measured raw data and the state estimates are accepted as being accurate.

Step 4: When the requirement of the inequality is not met, calculate $\hat{f}$ for each switching device zone (for example Z_F2 for fuse 2 or Z_SW2 for switch 2). Rank the $\hat{f}$ for each zone that does not satisfy the chi-square test of inequality for all switching zones. For example $\hat{f}_{F1}$ is for sum of the weighted squared residuals for Z_F1.

Here, ranking of residuals for each zones leads to finding the topology error. First, the area with the highest sum of the weighted residual will be removed. Then, it will be checked whether the new topology of the system can be fitted to the measurement set or not. Here is the rule to change the given feeder topology:

The Rule: Remove the zone with the highest rank from the topology.

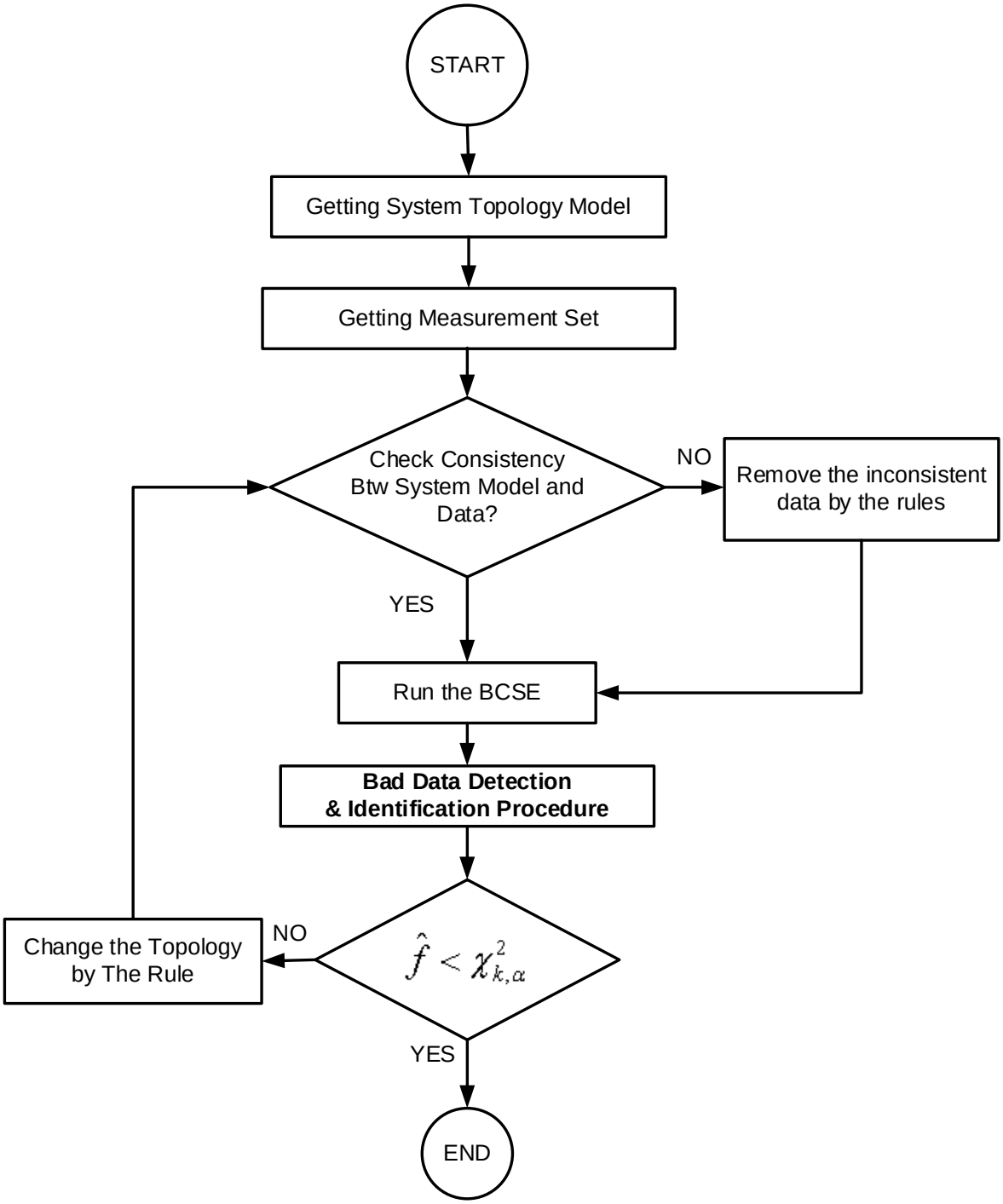


Figure (6-4): The proposed algorithm for topology error detection and identification by BCSE.

After applying the Rule, measurement set must be modified. Before starting the bad data detection and identification procedure, we should check the given measurement with very simple rules, such as rejecting the negative voltage measurements, negative loads with consumptions, or out of range actual measurements from the field.

The proposed algorithm for topology error detection and identification by BCSE has been illustrated in Figure (6-4). First, this algorithm takes the measurement set and the topology information to construct the bus/branch model of the feeder. Then, it checks the consistency in the data with the topology with simple rules. Afterwards, the BCSE will be executed and system states will be calculated from the SE. In next step, the bad data detection and identification procedure will be run to find the inconsistency among the measurements and the possible topology configuration. In other words, the weighted squared residual for each zone of the feeder will be calculated and tested by the proper statistical test to find the wrong status of the switches. When the measurement set and the processed topology matches with each other, this algorithm terminates and the final output of the BCSE will thus be reliable to operate the system.

6-6 Test Results

The IEEE 34 node test feeder [87] is used for test cases. The test feeder is assumed to have two line switches, two fuses and two tie switches as shown in Figure (6-5). As it is assumed voltage and power measurements are available at the substation. Other VMs and CMs are placed on the network based on two cases. In the first case, the measurement set has been constructed based on the output of the Mixed meter placement scheme. In the second case, measurement set has been constructed based on the output of the Robust meter placement scheme. First case has three real-time measurements and Second case has six real-time measurements, respectively.

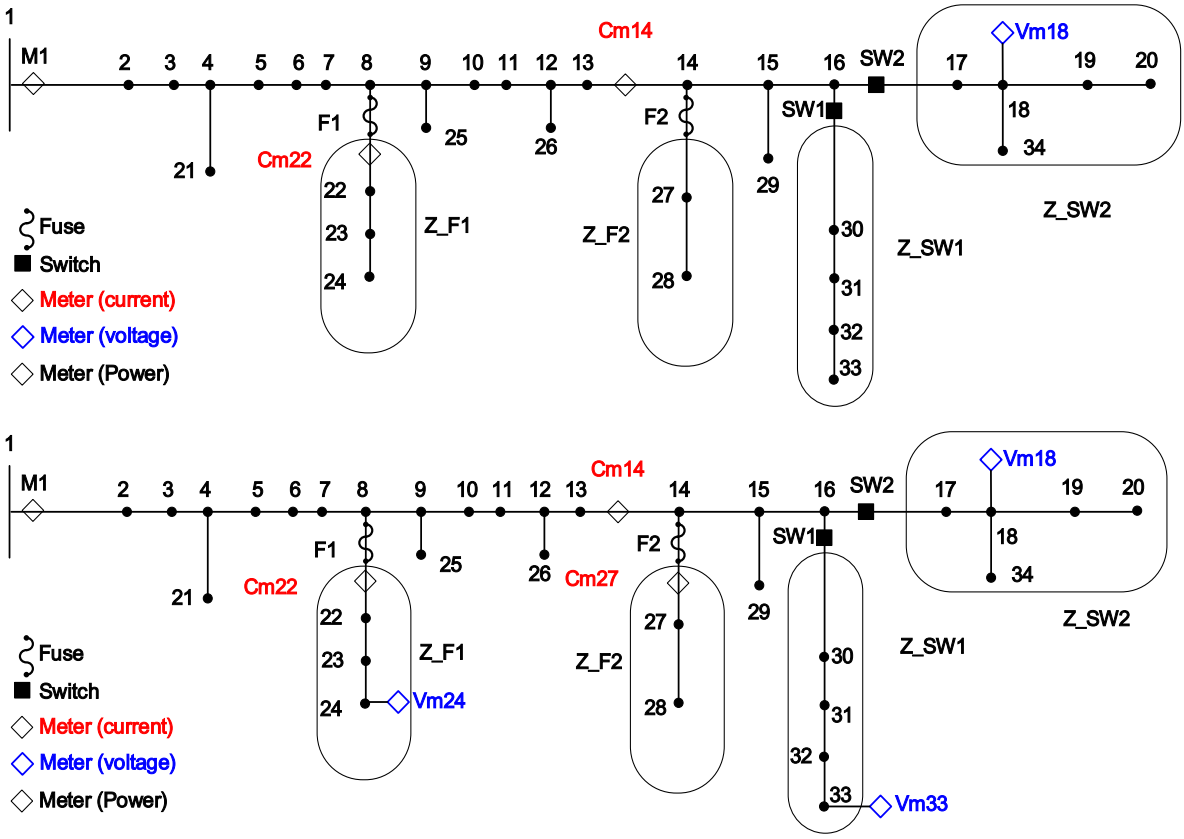


Figure (6-5): One-line diagram of the IEEE 34 node test feeder for case 1 with three measurements and case 2 with six measurements; respectively.

For testing topology error detection and identification algorithm, these different case are considered:

- *Case 1:* Power and voltage measurement (M1) with Cm14, Cm22, and Vm18.
- *Case 2:* Power and voltage measurement (M1) with Cm14, Cm22, Cm27, Vm33, Vm24, and Vm18.

Each case is simulated for different possible topologies. By switch and fuse location, these different cases are tested. To compare the results of those cases, a load flow program was used to provide the basic measurements data. The load flow solution can be considered as a true system state. If a case can obtain a solution closer to the LF solution, it is obviously a better

case and has more accurate solution. Then, each case is compared with this exact solution as a base case. For instance, in this prototype feeder, there are two fuses and two switches. So there are four feeder configurations which are tested for different cases [82].

To generate measurements data for testing purposes, measurement error was added to the actual measurements.

$$Z = Z^a \pm e_Z$$

where: Z^a is actual data and e_Z is generated by each data accuracy.

The accuracy level of the estimated load data is 50%. Monte Carlo cases are created by perturbing the actual load data by adding error of 50%. Here, for our studies, 300 MC cases have been constructed to test the performance of the proposed algorithm for topology error detection and identification. The power and current magnitude measurement errors are selected from Normal distribution with 1.2% accuracy level of their measured values. In addition, the voltage measurement data are generated by adding measurement error with 1.2% measurement error.

For bad data detecting & identification, the threshold value for inequality is chosen as $\alpha = 0.05$. Case 1 has 179 measurements (N_m) and 166 state variables (N_s) which has 9 degrees of freedom (k) so the threshold value is set at $\chi^2_{13,0.05} = 22.362$ for case 1. For case 2, it has 7 more measurements than case 1 so the threshold value is set at $\chi^2_{20,0.05} = 31.410$ for case 2.

6-6-1 Test Results for Case 1

In these test results, it is assumed that one switching device has been operated, i.e. just one fuse or one switch has been tripped. It is to be noted that considering the multiple topology changes is out of this thesis scope. Case 1 has three phase measurement for voltage as well as real and reactive power at the substation, also it has a three phase CM at 14th branch, a three

phase VM at 18th node, and one phase current measurement at 22th branch. There are four possibilities for changing the topology of the feeder, i.e. two fuses and two switches. A summary of the results is given in Table (6-1). $\hat{f}_{F1}$ presents the sum of the weighted residuals of the load estimation at zone of the fuse1, i.e. Z_F1. $\hat{f}_{F2}$ indicates the sum of the weighted residuals of the load estimation for Z_F2. $\hat{f}_{SW1}$ presents the sum of the weighted residuals of the load estimation at Z_SW1. $\hat{f}_{SW2}$ indicates the sum of the weighted residuals of the load estimation at zone of the fuse1, i.e. Z_F1. Here, $\hat{r}^{N}_{Load}$ is the highest normalized error for load estimation. In this case, the operation of the fuse on the circuit has been detected and identified completely. For instance, the sum of the partial weighted residuals is higher than the critical values of the chi-square distribution, $\chi^2_{13,0.05} = 22.362$. Therefore, topology error has been detected. Therefore, it is needed to identify the faulty area of the network. The proposed method has identified the topology error by having the highest $\hat{f}_{F1}$ in comparison with other zonal sum of the weighted residuals.

Table (6-1): Output for topology error detection & identification of Case 1.

	$\hat{f}_{F1}$	$\hat{f}_{F2}$	$\hat{f}_{SW1}$	$\hat{f}_{SW2}$	$\hat{r}^{N}_{Load}$
Fuse 1	113.9	1.13	1.75	0.21	L24
Fuse 2	0.055	50.38	56.94	4.76	L31
SW 1	0.452	79.17	101.8	0.22	L31
SW 2	0.118	92.88	411.6	210.6	L31

In case of fuse 2 operation, topology error has been detected by comparing the sum of weighted residuals with threshold chi-square value. Otherwise, the topology error has not been identified, because there is not any real-time measurement. The highest normalized error for load estimation is for load at 31st node where it is not in the zone of the fuse 2.

When SW 1 has been operated and zone of this switch has been de-energized by this operation, the sum of the weighted squared residuals has been calculated. This sum is larger than the threshold, so topology error has been detected. Now it is time to identify whether the area of the feeder has been changed. The highest value for zonal sum of the residual is $\hat{f}_{SW1}$, also the highest normalized load estimation error is for load at node 31. Therefore, the topology error has been identified by both approaches.

When SW 2 has been operated and Z_SW2 has been de-energized by this operation. The sum of the weighted squared residuals has been calculated. This value is bigger than the threshold, so topology error has been detected. Now it is time to identify the area of the feeder has been changed. The highest value for zonal sum of the residual is $\hat{f}_{SW1}$, also the highest normalized load estimation error is for load at node 31. Both of these measures shows the topology error in Z_SW1. On the other hand, there is real-time measurement at 18th node which is zero at this condition. The residual for this measurement is higher than the other measurement. Then, the topology error can been identified in stage of the consistency check of the topology with the measurement set. Table (6-2) summarizes these processes for topology error detection and identification.

Table (6-2): Summary for topology error detection & identification of Case 1.

	Detection	Identification
Fuse 1	Yes	Yes
Fuse 2	Yes	No
SW 1	Yes	Yes
SW 2	Yes	Yes

6-6-2 Test Results for Case 2

Case 2 has three phase measurements for voltage as well as real and reactive power at the substation, also it has two three phase CMs at 14th and 27th branch, two three phase VMs at

18th and 33th node, one phase current measurement at 22th branch, and one phase voltage measurement at 24th node. There are four possibilities for changing the topology of the feeder, because of the presence of two fuses and two switches on the feeder. A summary of the results is given in Table (6-3). Here, $\hat{r}^N_{Load}$ is the highest normalized error for load estimation. In this case, the operation of the fuse on the circuit has been detected and identified completely. For instance, the sum of the partial weighted residuals is higher than the critical values of the chi-square distribution, $\chi^2_{20,0.05} = 31.41$. Therefore, topology error has been detected. Hence it is required to identify the faulty area of the network. The proposed method has identified the topology error by having the highest $\hat{f}_{F1}$ in comparison with other zonal sum of the weighted residuals.

In case of fuse 2 operation, topology error has been detected by comparing the sum of weighted residuals with threshold chi-square value. In addition, the topology error has been identified. The highest normalized error for load estimation is for load at 28st node where it is in the zone of the fuse 2. The highest value for zonal sum of the residual is $\hat{f}_{F2}$ which shows the topology changing in the Z_F2. Hence, the topology error has been identified by both approaches.

Table (6-3): Output for topology error detection & identification of Case 2.

	$\hat{f}_{F1}$	$\hat{f}_{F2}$	$\hat{f}_{SW1}$	$\hat{f}_{SW2}$	$\hat{r}^N_{Load}$
Fuse 1	122.3	0.04	3.68	0.26	L24
Fuse 2	4.38	294.4	24.12	1.80	L28
SW 1	1.09	0.005	279.7	16.24	L31
SW 2	1.54	0.03	226.9	12.62	L31

When SW 1 was operated and the zone of this switch was de-energized by this operation, the sum of the weighted squared residuals was then calculated. This sum is higher than the threshold, so topology error has been detected. Now it is time to identify whether the area of

the feeder has been changed. The highest value for zonal sum of the residual is $\hat{f}_{SW1}$, also the highest normalized load estimation error is for load at node 31. Therefore, the topology error has been identified by both approaches.

In case of SW 2 operation, the Z_SW2 has been de-energized by this operation. The sum of the weighted squared residuals has been calculated. This value is bigger than the threshold, so topology error has been detected. Now it is time to identify the area of the feeder that was changed. The highest value for zonal sum of the residual is $\hat{f}_{SW1}$, 226.9, in addition the highest normalized load estimation error is for load at node 31. On the other hand, there is real-time measurement at the 18th node which is zero in this condition. The residual for this measurement is higher than that for other measurements. Consequently, the topology error can be identified in stage of the consistency check of the topology with the measurement set. Table (6-4) summaries these processes for topology error detection and identification.

Table (6-4): Summary for topology error detection & identification of Case 2.

	Detection	Identification
Fuse 1	Yes	Yes
Fuse 2	Yes	Yes
SW 1	Yes	Yes
SW 2	Yes	Yes

In summary, all topology changes were detected and identified in Case 2. On the other, measurement set of Case 1 cannot provide enough data to detect and identify of the topology changes. More investigations show that Case 1 does not have enough real-time measurements for each fuse and switch zones. For instance, the zone of the fuse 2 and the zone of the SW1 are not monitored by real-time measurements. However, each zone of the switches and fuses has at least one real-time measurement. Z_F1 has two real-time measurements, in addition Z_F2, Z_SW1, and Z_SW2 have one real-time measurement each. This study shows that more

real-time measurements at critical locations on the feeder help to detect and identify the topology changes properly by analysis of the residuals of the SE. It has been found that one real-time measurement for each switching is needed to detect and identify the topology changes and errors.

6-7 Impact of Topology Changes on BCSE Performance for VVC

As mentioned in previous sections, distribution systems are naturally exposed to topology changes. One of the possibilities that take place by transferring one part of the faulty feeder, from the neighbor feeder, to the main feeder without the fault. Usually, there are some normally open switches at the end of the feeder or laterals. For instance, one normally open switch is illustrated in Figure (6- 89), i.e. SW 12. This switch can add one lateral of each feeder to the other feeder. When the faults happen, these switches can be closed and provide another path to supply the assigned customers of this section by another main feeder. These topology changes can take place on normal operation of the feeder for load balancing among the feeders as well as sections of the feeders.

In this case, one part of the feeder will be added to the other feeder, meaning that the main circuit which has been considered for monitoring and control has been changed. By observing the state estimation procedures, the measurement functions are constructed based on the given feeder. Since the topology of the circuit has been changed, therefore, the measurement functions and the measurement matrix must change according to the new circuit.

To design this kind of topology changes in distribution systems, two IEEE 34 node test feeders, i.e. feeder 101 and feeder 201, are considered. After renumbering the feeders, these feeders are operated in radial when SW 12, i.e. normally open switch, is open. Feeder 101 has SW 101 which is the switch that is normally closed. Moreover, feeder 201 has a normally closed switch, i.e. SW 201. For load balancing, it is assumed that SW12 was closed and SW201 was opened.

This lateral from feeder 201 with nodes: 230, 231, 232, and 233 has been transferred to feeder 101.

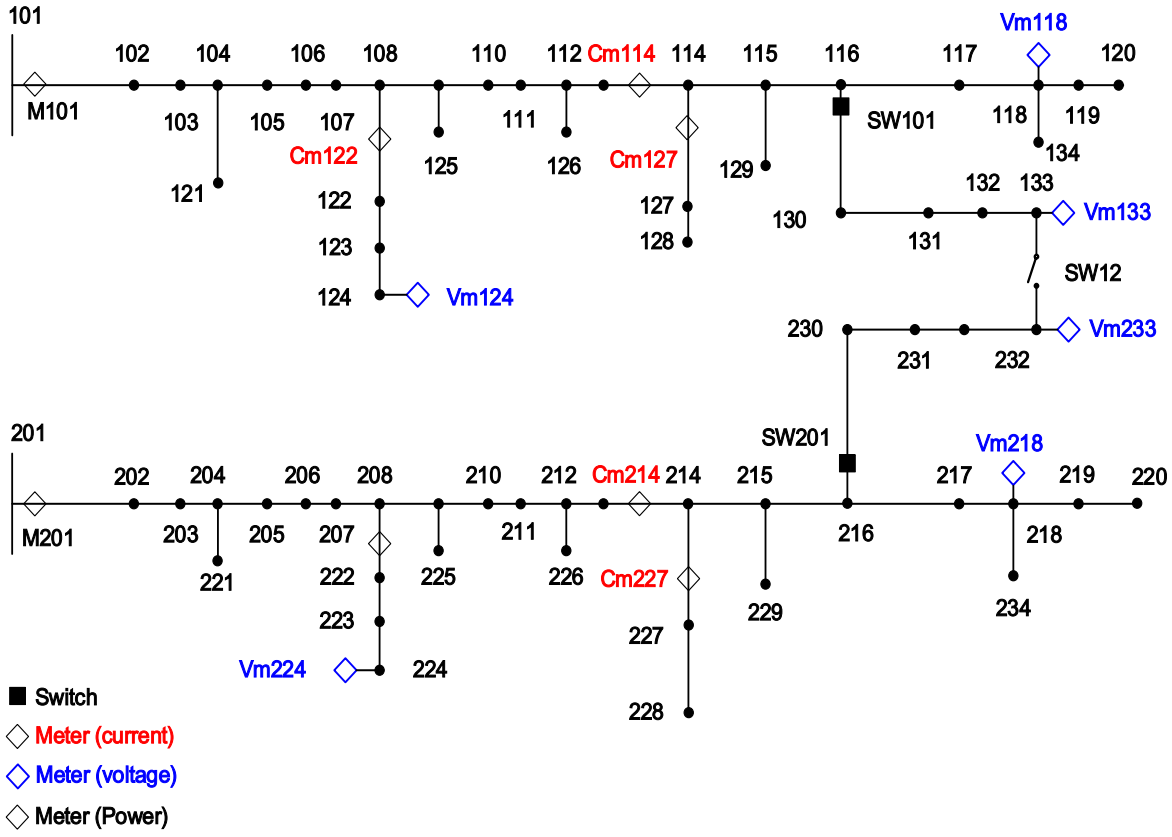


Figure (6-6): One-line diagram of two IEEE 34 node test feeders with normally opened switch, SW 12.

To investigate the topology changes, the following scenarios are considered.

Scenario 1: No SCADA for Switching with Case 1 measurement set

In this scenario, the control center has not been informed to these switching changes. Consequently, the SE was run with wrong topology information. Loads at nodes 230, 231, 232, 233 from the feeder 201 are shifted to the feeder 101. The measurement set for this scenario is

power and voltage measurement at the substation as well as Cm114, Cm122, and Vm118. This measurement set is the output of the mixed meter placement scheme. When the BCSE run, the sum of the weighted squared residuals becomes 52863.26, which is much higher than the acceptable value, i.e. $\chi^2_{13,0.05} = 22.362$. Based on the SE output, the normalized residual for these load estimations: L28, L31, and L33 are bigger than the threshold values. Therefore, the SE did not converge in this condition. In other words, the operator of the system cannot rely on the output of the SE. In addition, this result indicates that the measurement set of the mixed meter placement scheme is sensitive to the given topology.

Scenario 2: No SCADA for Switching with Case 2 measurement set

In this scenario, the control center has not been informed about these switching changes. Consequently, the SE was run with wrong topology information. Similarly, this scenario has the same load condition of scenario 1. The measurement set for this scenario is power and voltage measurement at the substation as well as Cm114, Cm127, Cm122, and Vm133, Vm118, and Vm124. This measurement set is the output of the robust meter placement scheme. In this case, the SE did not converge. In other words, the operator of the system can rely on the output of the SE for further control actions. The next concern for this output is whether that the voltages are estimated are within the desired accuracy for VVC application or not. The voltage standard deviation profile for this scenario is shown in Figure (6-7). The maximum voltage standard deviation in this case is 0.0055, which is higher than the target value, $\bar{\sigma}_V = 0.00278$.

In addition, this result indicates that the measurement set of the robust meter placement scheme can provide a better voltage estimate than the mixed meter placement scheme. On the other hand, the voltage cannot be estimated within the desired accuracy.

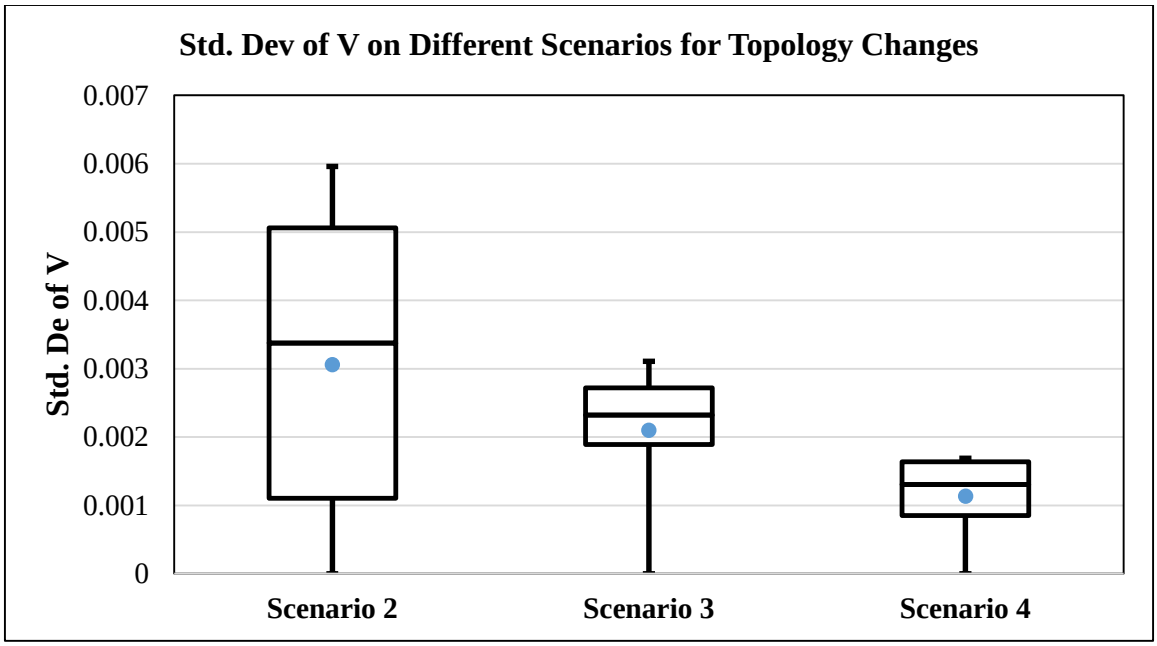


Figure (6-7): Range of the voltage standard deviation profiles for scenarios 2, 3, and 4.

Scenario 3: SCADA for Switching with Case 1 measurement set

In this scenario, the control center has been informed about the topology changes. Consequently, the SE was run with the new topology information. Similarly, this scenario has a similar load condition to that of same scenario 1. The measurement set for this scenario is the same as scenario 1. The BCSE in this case did converge. In other words, the operator of the system can rely on the output of the SE for further control actions. The next concern for this output is whether the voltages are estimated within the desired accuracy for VVC application or not. The voltage standard deviation profile for this scenario is shown in Figure (6-7). The maximum voltage standard deviation in this case is 0.0031 which is slightly larger than the target value, $\bar{\sigma}_V = 0.00278$. Therefore, the voltage profile cannot be estimated within the desired accuracy using the measurement set of the mixed meter placement scheme.

Scenario 4: SCADA for Switching with Case 2 measurement set

In this scenario, the control center was informed about these switching changes. Consequently, the SE was run with new topology information. Similarly, this scenario has the same load condition scenario as scenario 1.

The measurement set for this scenario is the same as scenario 2. The BCSE in this particular case has converged. In other words, the operator of the system can rely on the output of the SE for further control actions. The next concern for this output is, again and similar to the above, whether the voltages are estimated within the desired accuracy for VVC application or not. The voltage standard deviation profile for this scenario are shown in Figure (6-7). The maximum voltage standard deviation in this case is 0.0017 which is smaller than the target value, $\bar{\sigma}_V = 0.00278$. Therefore, with measurement set of the robust meter placement scheme the voltage profile can be estimated within the desired accuracy.

In summary, having the exact information about the topology of the distribution network is needed to run the SE to achieve accurate voltage profile estimation within the desired accuracy. In these scenarios, we can conclude that the availability of the real-time measurement from the new section of the feeder as well as the correct topology of the network are required to estimate the voltage along the feeder accurately.

6-8 Summary and Conclusion

In this chapter, the impact of the topology changes of the distribution systems on the SE was investigated. At the transmission level, TP provides accurate network configuration for the state estimator due to the availability of the redundant measurement set to offer the robust state estimation. However, due to the lack of the real-time measurement at distribution systems and un-monitored devices, such as fuses on the distribution feeder, providing accurate network configuration for real-time monitoring of the system is a challenging task. Topology changes occur by permanent faults on the main feeder and the laterals as well as for network re-configuration to balance the loads or for other reasons. Here, the proposed algorithm addresses these challenges by checking the consistency of the real-time measurement set and the network

and analysis of the residual at different zones of the switches and fuses. The first part of the study focused on cutting off some laterals or a section of the main feeder. Once this was performed, the proposed algorithm can then properly detect and identify these changes along the feeder by analyzing real-time measurements. Having the accurate topology of the system as well as few real-time measurements from critical locations of the network give enough input data for the BCSE to estimate the voltage along the feeder within the desired accuracy. This condition has been investigated in the last part of this chapter.

Chapter 7: Conclusion and Future Work

7-1 Conclusion

To manage and control the power distribution systems in an efficient and a reliable manner, implementing real-time monitoring structure is essential. The role of the SE in this structure is to determine the operating point of the system. First, the characteristics of the distribution systems which are constructed by the feeders were investigated using tools in the literature survey of Chapter 2. In addition, the main requirement for the operation of a distribution system, which is the voltage regulation, was explained and elaborated. After presenting the available SE methods for distribution systems, BCSE was selected for further statistical studies and future work. As clarified in this chapter 3, the performance of the BCSE relies mainly on the accuracy of the real-time measurements as well as the accuracy of the load estimation. By deploying the AMI infrastructure at distribution systems, more data about customer loads have become available. Hence, the load estimation method has been implemented to utilize these real-time measurements from AMI for improving the load estimation accuracy. The main contributions of this chapter are as follows:

1) Using the clustering analysis to determine the homogenous groups of the customers as well as to identify the different electrical load patterns,
2) Improving the load estimation method based on historical data by adding new real-time measurements from a few clusters of the customers, and
3) Implementing and testing statistical models to estimate the distribution transformer loads using the two weeks load data and real-time measurements from actual AMI data.
4) Testing the performance of the BCSE under load estimation errors

As mentioned before, providing the voltages within the acceptable range from the standard is the main goal for utilities to operate the system. Furthermore, the recent trend of utilities to implement the CVR programs to save more energy by adopting the advanced VVC schemes, accurate voltage estimate along the feeder is needed. It is shown in chapter 4, that load

estimation has an inherent huge error because of the electricity load nature. Therefore, estimating the voltage at the desired level requires real-time measurement from the feeder. This meter placement problem for VVC is defined and addressed in this chapter. Following are the main contributions of Chapter 5:

1) Performing a comprehensive sensitivity analyses to extract the initial rules of the meter placement. The following sensitivity studies were conducted:
    - Impact of load estimation on node voltage estimates,
    - Sensitivity of current measurements (CMs) on node voltage estimates, and
    - Sensitivity of Voltage measurements (VMs) on node voltages.
2) Adopting the branch current based SE (BCSE), the unbalanced nature of distribution systems in topology and loading was addressed.
3) Estimating the voltage profile with an accuracy of +/- 1 V (with respect to 120 V base) using these proposed schemes with few real-time measurement along the feeder.
4) Enhancing the meter placement schemes for loss of one real-time measurement from the measurement set.
5) Investigating the impact of load variation on the proposed schemes.
6) Investigating the impact of PV generation at customer sides.
7) Considering different types of measurements, those available at distribution level, to be placed on the feeder.
8) Estimating the node voltages with a desired accuracy using the search scheme that is computationally very efficient and test results showed that it is quite effective in identifying the minimal set of meters needed.

In daily operation of distribution power systems, the topology of the network may be changed. The last chapter of the work at hand has focused on the detection and identification of topology changes on distribution feeders. Chapter 7 has investigated the concept of the robust state estimation. Two common topology errors happening at the distribution feeder were explained. The first error was concerned about the disconnection of one lateral or section of the feeder.

As for the second error, it dealt with transferring one section of the feeder to the neighbor feeder by closing the normally open switch and opening the normally closed switch of the primary feeder. These issues are addressed in Chapter 7 by:

1) Proposing the detection and identification algorithm to find the incorrect part of the network.
2) Employing a proposed algorithm based on the analysis of the sum of the weighted residual for each zone of the switches and fuses.
3) Investigating the impact of adding the new section of another feeder, i.e. neighbor feeder, to the designated feeder for SE convergence as well as VVC application requirements to estimate the voltage profile along the new feeder within the desired accuracy.

7-2 Future Work

After the completion of this work, potential areas for further studies are hereby identified. These potential research topics are as follows:

- The implemented load estimation method needs more improvement to estimate the load at distribution transformers more accurately.
- Integrating the phase of load estimation stage with available real-time measurement to improve the quality of the SE output. This study can be done by measurements available from the DT levels as well as meters on the primary feeders.
- In the dissertation, statistical tests showed that population mean vector is not equal to the actual value. Thus, in the future, this biasness of the SE can be fixed by applying a proper statistical method to correct the average bias of the biased components.
- Here, it is assumed that the load and the PV generation are estimated within the same accuracy. For further study, the accuracy of the PV generation estimation must be investigated with actual PV data.

- The proposed algorithm for detection and identification of the topology errors is considered to run for one topology error. For future studies, multiple topology changes could be considered.
- Real-time measurements may be inaccurate or erroneous, due to several reasons as mentioned in the literature, and these bad data can deteriorate the performance of the BCSE. Therefore, it is crucial to propose the bad data detection and identification procedures to detect, identify, and eliminate these bad data from the input of the SE.
- At the transmission level of power systems, the synchronized phasor measurements, i.e. PMUs, are becoming more popular for monitoring applications. Thus, the BCSE method can include these PMU measurements in the future.
- Integration of the PMU measurements for distribution system applications, i.e. VVC, can be considered for the meter placement problem discussed in Chapter 5.
- AMI provides more information than the load profiles, such as customer outages. These information can be added in the process of the SE to more accurately estimate the operating condition of the system.
- Implementing the BCSE for the actual feeder, to verify the proposed algorithms in load estimation and state estimation.

REFERENCES

[1] B. Seal, B. Uluski, “Integrating Smart Distributed Energy Resources with Distribution Management Systems,” EPRI white paper, Sept. 2012.

[2] http://www.netl.doe.gov/smartgrid/refshelf.html#White Papers

[3] F. C. Schweppe and J. Wildes, “Power system static state estimation,: exact model, approximate model, and implementation ”, parts I-III, IEEE Trans. on Power Appar. and Syst., Vol. PAS-89, pp. 120-135, 1970.

[4] S. Grenard, O. Devaux, O. Carré, and O. Huet, “Power Steering: Advanced Distribution Applications for Operations by EDF R&D,” IEEE PES Magazine Sep/Oct2011.

[5] R.C. Dugan, R. Arritt, T. McDermott, S. Brahma, K. Schneider, “Distribution System Analysis to support the Smart Grid,” Power and Energy Society General Meeting, 2010 IEEE , vol., no., pp.1,8, 25-29 July 2010.

[6] ANSI C84.1, Electric Power Systems And Equipment – Voltage Ratings (60 Hertz), April 2011.

[7] B. Uluski, “Approaches to Volt-VAR Control and Optimization”, Transporting You into the 21st Century Distribution System, September 16, 2010.

[8] T. Gonen, Electric Power Distribution System Engineering, McGraw-Hill, 1986.

[9] E. Liu, J. Bebic, “Distribution System Voltage Performance Analysis for High-Penetration Photovoltaics” GE Global Research Niskayuna, New York, Subcontract Report, NREL/SR-581-42298, February 2008.

[10] http://www.dsireusa.org/incentives/index.cfm?SearchType=RPS&&EE=0&RE=1, DSIRE Database of State Incentives for Renewables & Efficiency

[11] F. Katiraei, J. R. Aguero, “Solar PV Integration Challenges,” IEEE power & energy magazine, May/June 2011

[12] M. Begovic, I. kim, D. Novosel, J. R. Agüero, A. Rohatgi,"Integration of Photovoltaic Distributed Generation in the Power Distribution Grid", 45th Hawaii International Conference on System Sciences 2012

[13] M. Thomson and D.G. Infield, "Impact of widespread photovoltaics generation on distribution systems", IET Renew. Power Generator, 2007, 1, (1), pp. 33-40

[14] Wenbo Zhang, "Fast Volt - VAR Control on PV Dominated Distribution Systems. (Under the direction of Mesut E Baran), " M. Sc. Thesis, North Carolina State University, 2012.

[15] A. Abur, A. G. Exposito, Power System State Estimation: Theory and Implementation, Marcel Decker, NY 2004.

[16] M. E. Baran, "Challenges of state estimation on distribution systems," 2001 PES Summer Meeting, page(s): 429-433. 2001.

[17] I. Roytelman, "Real-time distribution power flow – lessons from practical implementations", PSCE, page(s): 505-509. 2006.

[18] V. Thornley, N. Jenkins, and S. White, "State estimation applied to active distribution networks with minimal measurements," Power Systems Computation Conference, pp. 1–7, 2005.

[19] I. Roytelman and S. M. Shahidehpour, "State estimation for electric power distribution systems in quasi real-time conditions," IEEE Transactions on Power Delivery, vol. 8, no. 4, pp. 2009–2015, Oct. 1993.

[20] M. K. Celik, and W. H. E. Liu, "A practical distribution state calculation algorithm," in Proc. 1999 IEEE Power Eng. Soc. Winter Meeting, New York, Jan. 31-Feb. 3 1999.

[21] A. Monticelli, State Estimation Electric Power Systems, A Generalized Approach, Kluwer Academics Publishers, MA 1999.

[22] A. K. Ghosh, D. L. Lubkeman, M. J. Downey, and R. H. Jones, "Distribution circuit state estimation using a probabilistic approach," IEEE Trans. Power Syst., vol. 12, pp. 45–51, Feb. 1997.

[23] A. K. Ghosh, D. L. Lubkeman, and R. H. Jones, “Load modeling for distribution circuit state estimation,” IEEE Trans. Power Delivery, vol. 12, pp. 999–1005, Apr. 1997.

[24] D. L. Lubkeman, J. Zhang, A. K. Ghosh, and R. H. Jones, “Field results for a distribution circuit state estimator implementation,” IEEE Trans. Power Delivery, vol. 15, no.1, pp. 399–406, Jan. 2000.

[25] M. E. Baran and A. W. Kelley, “State estimation for real-time monitoring of distribution systems,” IEEE Trans. Power Syst., vol. 9, pp. 1601–1609, Aug. 1994.

[26] C. N. Lu, J. H. Teng, and W. H. E. Liu, “Distribution system state estimation,” IEEE Trans. Power Syst., vol. 10, pp. 229–240, Feb. 1995.

[27] W. M. Lin and J. H. Teng, “Distribution fast decoupled state estimation by measurement pairing,” Proc. Inst. Elect. Eng., Gen. Transm. Dist., vol. 143, no. 1, pp. 43–48, Jan. 1996.

[28] W. M. Lin and J. H. Teng, “State estimation for distribution systems with zero-injection constraints,” IEEE Trans. Power Syst., vol. 11, pp. 518–524, Feb. 1996.

[29] K. Li, “State estimation for power distribution system and measurement impacts,” IEEE Trans. Power Syst., vol. 11, pp. 911–916, May 1996.

[30] M. E. Baran and A.W. Kelley, “A branch-current-based state estimation method for distribution systems,” IEEE Trans. Power Syst., vol. 10, pp. 483–491, Feb. 1995.

[31] W. M. Lin, J. H. Teng, and S. J. Chen, “A highly efficient algorithm in treating current measurements for the branch-current-based distribution state estimation,” IEEE Trans. Power Delivery, vol. 16, pp. 433–439, July 2001.

[32] H. Wang, and N. N. Schulz, “A revised branch current-based distribution system state estimation algorithm and meter placement Impact,” IEEE Trans. Power Syst., vol. 19, no. 1, pp. 207–213, Feb. 2004.

[33] M. E. Baran, J. Zhu, and A. W. Kelley, “Meter placement for real-time monitoring of distribution feeders,” IEEE Trans. Power Syst., vol. 11, no. 1, pp. 332-337, Feb. 1996

[34] H. Wang and N. N. Schulz, “A load modeling algorithm for distribution system state estimation,” in Proc. 2001 IEEE/Power Eng. Soc. Transm. Dist. Conf. Expo., pp. 102–106.

[35] H. Wang and N. N. Schulz, “Using AMR data for load estimation for distribution system analysis,” J. Elect. Power Syst. Res., vol.76, pp.336-342, 2006.

[36] E. J. Beroset, “Advanced Metering Infrastructure (Elster Inc.),” presented in Smart Distribution System course at Dept. ECE., North Carolina State Univ., Raleigh, NC, USA, April 2011.

[37] A. P. S. Meliopoulos, and F. Zhang, “Multiphase power flow and state estimation for power distribution systems,” IEEE Trans. Power Syst., vol. 11, no. 2, pp. 939–946, May 1996.

[38] M. E. Baran, J. Jiang, and T. McDermott, “Including voltage measurements in branch current state estimation for distribution systems,” in 2009 PES General meeting, pp. 1-5, 2009.

[39] M. Baran and T. McDermott, “Distribution system state estimation using AMI data,” in Power Systems Conference and Exposition, 2009. PSCE '09. IEEE/PES, pp. 1-3, 2009.

[40] I. Dzafic, S. Henselmeyer, and H. T. Neisius, “High performance state estimation for smart grid distribution network operation,” in 2011 PES Innovative Smart Grid Technologies (ISGT), pp. 1-6, 2011.

[41] K. Samarakoon, J. Wu, J. Ekanayake, and N. Jenkins, “Use of delayed smart meter measurements for distribution state estimation,” in 2011 PES General meeting, pp. 1-6, 2011.

[42] R. Singh, B. C. Pal, and R. A. Jabr, “Choice of estimator for distribution system state estimation,” IET Gen., Transm., Distrib., vol. 3, no. 7, pp. 666-678, Aug. 2009.

[43] C. Gouveia, H. Leite, and I. Ferreira, “Voltage and current sensor for state estimation in distribution network with generation,” International Conference on Renewable Energies and Power Quality (ICREPQ’10), Granada, Spain, pp. 1-6, March 2010.

[44] O. Chilard, S. Grenard, O. Devaux, and L. A. Garcia, “Distribution state estimation based on voltage state variables: assessment of results and limitations,” 20th International Conference on Electricity Distribution (CIRED 2009), Prague, paper no. 0524, June 2009.

[45] A. Shafiu, V. Thornley, N. Jenkins, G. Strbac, and A. Maloyd, “Control of active networks,” 18th International Conference on Electricity Distribution (CIRED 2005), Turin, paper no. 0524, June 2009.

[46] G. A. Taylor, M. R. Irving, N. Nusrat, R. Liao, and S. Panchadcharam, “Smart distribution network operation: emerging techniques and standards,” in 2011 PES General meeting, pp. 1-6, 2011.

[47] R. Singh, B. C. Pal, and R. B. Vinter, “Measurement placement in distribution system state estimation,” IEEE Trans. Power Syst., vol. 24, no. 2, pp. 668-675, 2009.

[48] S. Kotz, T. J. Kozubowski, and K. Podgorski, The Laplace distribution and generalizations, Birkhauser, Boston, 2001.

[49] Selected Statistical Methods For Analysis of Load Research Data, EPRI Report EA-3467, May 1984.

[50] R. Herman and J.J. Kritzinger, “The Statistical Description of Grouped Domestic Electrical Load Currents,” Electric Power Systems Research Journal, vol. 27, pp. 43-48, 1993.

[51] A. Papoulis, Probability, Random Variables, and Stochastic Processes, 3rd Ed., McGraw-Hill Inc., New York, 1991.

[52] R.Y. Rubinstein, Simulation and the Monte Carlo Method, John Wiley & Sons, Inc., New York, 1981.

[53] I. Cobelo, A. Shafiu, N. Jenkins, and G. Strbac, “State estimation of networks with distributed generation,” Euro. Trans. Electr. Power, vol. 17, pp. 21-36, 2007.

[54] M. E. Baran, J. Jiang, and T. McDermott, “Topology error identification using branch current state estimation for distribution systems,” 2009 Asia and Pacific Transmission & Distribution Conference & Exposition, pp. 1-4, 2009.

[55] M.Z. Kamh, and R. Iravani, “Unbalanced model and power-flow analysis of microgrids and active distribution systems”, IEEE Transaction on Power Delivery, vol. 25, no.4, Oct. 2010.

[56] R. Hoffman, "Practical State Estimation for Electric Distribution Networks," in Proc. 2006 IEEE PSCE Power Syst. Conf. and Expo., pp. 510-517.

[57] “Assessment of Demand Response and Advanced Metering, Staff Report,” Federal Energy Regulatory Commission, USA, Aug-2006.

[58] “Assessment of Demand Response and Advanced Metering, Staff Report,” Federal Energy Regulatory Commission, USA, Aug-2012.

[59] K. Sridharan, and N. N. Schulz, “Outage management through AMR systems using an intelligent data filter,” IEEE Trans. Power Delivery, vol. 16, no. 4, pp. 669–675, Oct. 2001.

[60] E. Boardman, “The role of integrated distribution management systems in smart grid implementations,” IEEE PES General Meeting 2010, pp. 1-6, July 2010.

[61] A. Borghetti, M. Bosetti, S. Grillo, S. Massucco, C. A. Nucci, M. Paolone, and F. Silvestro, “Short-term scheduling and control of active distribution systems with high penetration of renewable resources”, IEEE Systems Journal, vol. 4, no. 3, Sep. 2010.

[62] I. Dzafic, S. Henselmeyer, and H.-T. Neisius, “Real-Time Distribution System State Estimation”, IPEC 2010, Suntec, Singapore, paper ID 389, Oct. 2010.

[63] I. Dzafic, S. Henselmeyer, and H.-T. Neisius, “Real Time Distribution System State Estimation Based on Interior Point Method”, 17th PSCC 2011, Stockholm, Sweden, Aug. 2011.

[64] A. Monticelli, “Electric power system state estimation,” Proc. IEEE, vol. 88, no. 2, pp. 262–282, 2000.

[65] Y. Sharon, A. Annaswamy, A. L. Motto, and A. Chakraborty, “Topology identification in distribution network with limited measurements,” in Proc. 2012 IEEE Innovative Smart Grid Technologies Conference (ISGT’12), Washington DC, USA, 2012.

[66] R. Singh, E. Manitsas, B. C. Pal, and G. Strbac, “A recursive Bayesian approach for identification of network configuration changes in distribution system state estimation,” IEEE Trans. Power Syst., vol. 25, no. 3, pp. 1329-1336, Aug. 2010.

[67] J. C. Pereira, J. T. Saraiva, V. Miranda, A. S. Costa, E. Lourengo, and K. Clements, “Comparison of approaches to identify topology errors in the scope of state estimation studies,” in 2001 IEEE Power Tech Proc., vol. 3, Porto, Portugal, 2001.

[68] Z. J. Simendic, V. C. Strezoski, and G. S. Svenda, “In-field verification of the real-time distribution state estimation,” 18th International Conference on Electricity Distribution (CIRED 2005), Turin, paper no. 0524, June 2005.

[69] V. Thornley, N. Jenkins, and S. White, “State estimation applied to active distribution networks with minimal measurements,” PSCC 2005, Liege, Belgium, 2005.

[70] A. S. Costa,M. C. dos Santos, "Real-Time Monitoring of Distributed Generation Based on State Estimation and Hypothesis Testing," IEEE Power Tech 2007, Lausanne , pp.538-543, 1-5 July 2007.

[71] N. Singh, and F. Oesch, “Practical experience with rule-based on-line topology error detection,” IEEE Trans. Power Syst., vol. 9, no. 2, pp. 841-847, May 1994.

[72] G. T. Heydt, “The next generation of power distribution systems,” IEEE Trans. Smart Grid., vol. 1, no. 3, pp. 225–235, Dec. 2010.

[73] A. P. Sakis Meliopoulos et al , “Smart grid technologies for autonomous operation and control” IEEE Trans. Smart Grid, vol. 2, no. 1, pp. 1-10, March 2011.

[74] S. Shao, M. Pipattanasomporn, and S. Rahman, “Challenges of PHEV Penetration to the Residential Distribution Network,” IEEE PES General Meeting 2009, pp. 1-8, July 2009.

[75] S. Babaei, D. Steen, L. Tuan, O. Carlson, L. Bertling, “Effects of Plug-in Electric Vehicles on distribution systems: A real case of Gothenburg,” IEEE ISGT Europe 2010, pp. 1-8, Oct. 2010.

[76] S. Bhattacharya, T. Zhao, G. Wang, S. Dutta, S. Baek, Y. Du, B. Parkhideh, X. Zhou, and A. Q. Huang, “Design and development of generation-I silicon based solid state transformer,” in Proc. 25th Annu. IEEE Appl. Power Electron. Conf., 2010, pp. 1666–1673.

[77] https://aep.com/about/IssuesAndPositions/Distribution/SmartGrid/CES.aspx

[78] D N. Ewart, “Whys and wherefores of power system blackouts,” IEEE Spectrum, vol. 15, pp. 36–41, Apr. 1978.

[79] A. De, W. Zhang, S. Bhattacharya, M. E. Baran, “DVC Modelling and Volt-Var Control for Distribution Feeders,” Varentec Quaterly Report, 30 July 2012.

[80] J. H. Teng, “ Using voltage measurements to improve the results of branch-current-based state estimators for distribution systems,” Proc. Inst. Elect. Eng., Gen. Transm. Dist., vol. 149, no. 6, pp. 667–672, Nov. 2002.

[81] E. Li, D. Boos, and M. Gumpertz, Simulation Study in Statistics, Department of Statistics, North Carolina State University, Oct. 2, 2001.

[82] J-Sung JUNG, “Branch Current State Estimation Method for Power Distribution Systems. (Under the direction of. Mesut E Baran), ” M. Sc. Thesis, North Carolina State University, 2009.

[83] D. D. Boos, L. A. Stefanski, Essential Statistical Inference : Theory and Methods. Springer, 2008.

[84] R. Singh, B. C. Pal, Rabih and A. Jabr, “Modelling of Pseudo Measurements in Distribution System State Estimation,” SmartGrids for Distribution, 2008. IET-CIRED. CIRED Seminar, vol.1, no. 4, pp 23-24, June 2008.

[85] H. Oja, and R. H. Randles, Multivariate Nonparametric Tests, Statistical Science, pp. 598-605, 2004.

[86] I. Lerche, and B. S. Mudford, “How Many Monte Carlo Simulations Does One Need to do?’, Energy Exploration & Exploitation, pp. 405-427, 2005.

[87] IEEE 34 Node Test Feeder, Distribution Test Feeders [online] available: http://ewh.ieee.org/soc/pes/dsacom/testfeeders/index.html.

[88] Multivariate Statistics, Statistics Toolbox of MATLAB R2011b: http://www.mathworks.com/products/statistics/

[89] N.T. Thomopoulos, Essentials of Monte Carlo Simulation: Statistical Methods for Building Simulation Models, Springer Science+Business Media, New York 2013.

[90] Haijun Liu; Yu, D.; Hsiao-Dong Chiang, "A heuristic meter placement method for load estimation," Power Systems, IEEE Transactions on , vol.17, no.3, pp.913,917, Aug 2002.

[91] Leou, R.C.; Lu, C.N., "Improving feeder voltage calculation results with telemetered data," Power Delivery, IEEE Transactions on , vol.11, no.4, pp.1914,1920, Oct 1996.

[92] Shafiu, A.; Jenkins, N.; Strbac, G., "Measurement location for state estimation of distribution networks with generation," Generation, Transmission and Distribution, IEE Proceedings- , vol.152, no.2, pp.240,246, 4 March 2005.

[93] Singh, R.; Pal, B.C.; Jabr, R.A.; Vinter, R.B., "Meter Placement for Distribution System State Estimation: An Ordinal Optimization Approach," Power Systems, IEEE Transactions on , vol.26, no.4, pp.2328-2335, Nov. 2011.

[94] Heuristic Methods: http://en.wikipedia.org/wiki/Heuristic_(computer_science)

[95] Thomas Weise: "Global Optimization Algorithms – Theory and Application," published by it-weise.de (self-published), Germany 2009.

[96] N. Kokash, “http://disi.unitn.it/~kokash/documents/Heuristic_algorithms.pdf., “An Introduction to Heuristic Algorithms.”

[97] A. Parkes, “Artificial Intelligence Methods- Course Notes and Lectures,” http://www.cs.nott.ac.uk/~ajp/.

[98] R. Billinton and D. Huang, "Effects of Load Forecast Uncertainty on Bulk Electric System Reliability Evaluation," Power Systems, IEEE Transactions on , vol.23, no.2, pp.418,425, May 2008.

[99] A. Seppala, "Statistical distribution of customer load profiles," Energy Management and Power Delivery, 1995. Proceedings of EMPD95, 1995 International Conference on , vol.2, pp.696-701, Nov 1995.

[100] Modeling Load Forecast Uncertainty (LFU) in Power System Reliability Studies, PJM working report, 2012.

[101] IEEE and ANSI Instrument Transformers STANDARDS (IEEE C57.13 Series), revision 2009 and 2005.

[102] GE Digital Energy: https://www.gedigitalenergy.com/multilin/catalog/fmc.htm#fmc4.

[103] Cooper Power Systems: http://www.landisgyr.com/products/electric-meters/.

[104] GridSense Products: http://www.gridsense.com/solutions-products/transmission-and-distribution-line-monitoring/

[105] SATEC Powerful Solutions: http://www.satec-global.com/eng/products.aspx?product=87

[106] Landis+Gyr Products: http://www.landisgyr.com/products/electric-meters/

[107] Electro Industries/GaugeTech: http://www.electroind.com/shark200-data-logging-power-meter.html

[108] M. E. Baran and F. F. Wu, "Optimal capacitor placement on radial distribution systems," Power Delivery, IEEE Transactions on , vol.4, no.1, pp.725,734, Jan 1989.

[109] W. H. Kersting, "Distribution Feeder Voltage Regulation Control," Industry Applications, IEEE Transactions on , vol.46, no.2, pp.620,626, March-april 2010.

[110] A. J. Wood and B.F. Wollenberg, “Power Generation, Operation & Control,” John Wiley and Sons, New York, 1984.

[111] Empirical Rule in Statistics: http://en.wikipedia.org/wiki/68-95-99.7_rule

[112] D. M. V. Kumar, S. C. Srivastava, S. Shah, and S. Mathur, “Topology processing and static state estimation using artificial neural networks,” Proc. Inst. Elect. Eng., Gen., Transm., Distrib., vol. 143, no. 1, pp. 99–105, Jan. 1996.

[113] ALSTOM Grid Smart Distribution Solution for Smart Utilities, ALSTOM-Grid white paper, http://www.alstom.com/Global/Grid/, 2013.

[114] Tao Huang, Zhongdong Yin and Renzhong Shan, “A New Dynamic Voltage-Var Compensator Based On SVPWM”, international Conference on Energy and Environment Technology, 2009

[115] Zhenyuan Wang, Jing Xu, K. Hatua, S. Madhusoodhanan, and S. Bhattacharya, "Solid state transformer specification via feeder modeling and simulation," Power and Energy Society General Meeting, 2012 IEEE , vol., no., pp.1,5, 22-26 July 2012.

[116] B. W. Kennedy and R. H. Fletcher, "Conservation voltage reduction (CVR) at Snohomish County PUD," IEEE Transactions on Power Systems, vol. 6, No. 3, August 1991.

[117] D. Kirshner, "Implementation of conservation voltage reduction at Commonwealth Edison," IEEE Transactions on Power Systems, vol. 5, No. 4, November 1990.

[118] C.A. McCarthy and J. Josken, “Applying Capacitors to Maximize Benefits of Conservation Voltage Reduction”, Rural Electric Power Conference, C4-1 – C4-5, 2003.

[119] R. Singh, F. Tuffner, J. Fuller, and K. Schneider, "Effects of Distributed Energy Resources on Conservation Voltage Reduction (CVR)," 2011 IEEE PES General Meeting, July 24-29, Detroit, Michigan, 2011.

[120] Zhan Shen, “Centralized and Decentralized Volt/Var Schemes for FREEDM Systems. (Under the direction of Mesut E Baran), ”Ph.D. Dissertation, North Carolina State University, 2013.

[121] H. J. Koglin, G. Beibler, and K. D. Schmitt, “Bad data detection and identification”, International Journal of Electrical Power and Energy Systems, Vol. 12, No. 2, pp. 94–103, April 1990.

[122] P. J. Rousseeuw and A. M. Leroy, “Robust Regression and Outlier Detection”, John Wiley & Sons 2003.

[123] F. Hampel, “Robust statistics: A brief introduction and overview”, 2001.

[124] P. J. Huber, and E. M. Ronchetti, “Robust Statistics”, 2nd Edition, Wiley series in Probability and Statistics, John Wiley & Sons Inc. 2009.

[125] F. R. Hampel, “The influence curve and its role in robust estimation”, J. of the American Stat. Assoc., vol. 69, pp. 383-393, 1974.

[126] M. Irving, R. Owen, and M. J. H. Sterling, “Power system state estimation using linear programming,” Proc. IEE, vol. 125, pp. 879–885, 1978.

[127] H. Singh, and F. L. Alvarado, “Network Topology Determination using Least Absolute Value State Estimation,” IEEE Transaction on Power Systems, Vol. 10, No. 3 , pp. 1159-1165, Aug. 1995.

[128] A. Abur, D. Shirmohammadi, and C. S. Cheng, “Estimation of switch statuses for radial power distribution systems,” proceeding of ISCAS '95, IEEE International Symposium on Circuits and Systems, vol.2, pp.1127-1130, May 1995.

[129] M.E. Baran, H. Hooshyar, Zhan Shen, A. Huang, "Accommodating High PV Penetration on Distribution Feeders," Smart Grid, IEEE Transactions on , vol.3, no.2, pp.1039,1046, June 2012.

[130] G. Karady, A. Q. Huang, M. Baran, "FREEDM system: An electronic smart distribution grid for the Future," Transmission and Distribution Conference and Exposition (T&D), 2012 IEEE PES , vol., no., pp.1,6, 7-10 May 2012.

[131] M. E. Baran, “Load estimation for distribution system analysis,” PES Summer Meeting, 2001. IEEE, vol.2, pp. 999-1001, 2001.

[132] U. Singh, and M. E. Baran, “Load estimation for distribution feeder monitoring and management,” PES General Meeting, 2010. IEEE, pp. 1-6, Jul. 2010.

[133] Smart Grid for Distribution Systems: The Benefits and Challenges of Distribution Automation (DA), NIST White Paper, Darft Version 2, 2012.

[134] K. Samarakoon, J. Wu, J. Ekanayake, N. Jenkins, "Use of delayed smart meter measurements for distribution state estimation," PES General Meeting, 2011. IEEE , pp. 24-29, Jul. 2011.

[135] J. H. Chow, F. F. Wu, and J. A. Momoh, Applied mathematics for restructured electric power systems, Optimization, Control, and Computational Intelligence. Springer, 2005.

[136] H. Wang and N. N. Schulz, "A load modeling algorithm for distribution system state estimation," Transmission and Distribution Conference and Exposition, 2001 IEEE/PES, vol.1, pp. 102-105, 2001.

[137] H. Wang and N. N. Schulz, "Using AMR data for load estimation for distribution system analysis," J. Elect. Power Syst. Res., vol.76, pp. 336-342, 2006.

[138] Y. Chen, P. B. Luh, C. Guan, Y. Zhao, L. D. Michel, M. A. Coolbeth, P. B. Friedland, and S. J. Rourke, "Short-term load forecasting: similar day-based wavelet neural networks," IEEE Trans. Power Syst., vol.25, no.1, pp. 322-330, Feb. 2010.

[139] T. Hong, M. Gui, M. E. Baran, and H. L. Willis, "Modeling and forecasting hourly electric load by multiple linear regression with interactions," PES General Meeting, 2010. IEEE, pp. 25-29, Jul. 2010.

[140] W. Zalewski, "Application of fuzzy inference to electric load clustering," Power India Conf. 2006 IEEE, pp.5, 2006.

[141] A. K. Ghosh, D. L. Lubkeman, and R. H. Jones, "Load modeling for distribution circuit state estimation," IEEE Trans. Power Delivery, vol.12, no.2, pp. 999-1005, Apr 1997.

[142] W. Charytoniuk, M. S. Chen, P. Kotas, P. Van Olinda, "Demand forecasting in power distribution system using non-parametric probability density estimation," IEEE Trans. Power Syst., vol.14, no.4, pp. 1200-1206, Nov. 1999.

[143] S. Kuzmanovic, G. Svenda, Z. Ovcin, "Practical statistical methods in distribution load estimation," CIRED 2009, vol. 1, no. 4, pp. 8-11, June 2009.

[144] A. Sargent, R. P. Broadwater, J. C. Thompson, and J. Nazarko, “Estimation of diversity an kWHR-to-peak-kW factors from load research data,” IEEE Trans. Power Syst., vol.9, no.3, pp. 1450-1456, Aug. 1994.

[145] C. S. Chen, J. C. Hwang, Y. M. Tzeng, C. W. Huang, and M. Y. Cho, “Determination of customer load characteristics by load survey System at Taipower,” IEEE Trans. Power Delivery, vol.11, no.3, pp. 1430-1436, Jul. 1996.

[146] A. K. Ghosh, D. L. Lubkeman, M. J. Downey, and R. H. Jones, “Distribution circuit state estimation using a probabilistic approach,” IEEE Trans. Power Syst., vol.12, no.1, pp. 45-51, Feb. 1997.

[147] V. P. Borozan and N. L. J. Rajakovic, “Load estimation for distribution systems with minimum information,” Euro. Trans. Electr. Power, vol. 14, pp. 331–345, Nov. and Dec. 2004.

[148] Westinghouse Distribution System, Electric Utility Engineering Reference Book. First Edition, vol. 3, Fifth Printing, 1965.

[149] Rural Electrification Administration, REA Bulletin 45-2 Demand Tables. June 1963.

[150] Y. C. Lee and M. Etezadi-Amoli, “An improved modeling technique for distribution feeders with incomplete information,” IEEE Trans Power Delivery, vol.8, no.4, pp. 1966-1972, Oct 1993.

[151] V. N. Mesa, Distribution Feeder Calculations: a Load, Voltage and Line Loss Analysis of Distribution Circuits. Pacific Gas & Electric (PG&E) Company.

[152] N. Lu, P. Du, X. Guo, and F. L. Greitzer, "," Transmission and Distribution Conference and Exposition (T&D), 2012 IEEE PES , vol. 7, no. 10, pp.1-6, May 2012.

[153] U. Singh, “Load estimation for distribution feeder monitoring & management,” M.Sc. thesis, Dept. ECE., North Carolina State Univ., Raleigh, NC, USA, 2010.

[154] W. Mendenhall and T. Sincich, A second course in statistics: regression analysis. Pearson Education, 2003.

[155] R. J. Freund and R. C. Littell, SAS System for Regression, Third Edition. SAS Institute, Cary, NC, 2000.

[156] S. Sharma, Applied multivariate techniques. John Wiley, 1996.

[157] SAS online manual [online] availabe: http://support.sas.com/onlinedoc/913/docMainpage.jsp

[158] G. Fernandez, Data mining using SAS applications. CRC Press, 2003.

[159] Notes on Cluster Analysis, University of Illinoise at Chicago, Department of Information & Decision Sciences, by S. L. Sclove [online] available: http://www.uic.edu/classes/idsc/ids472/clustering.htm.

[160] J. Wang, X. Li, and C. Li, "Load characteristics clustering based on improved FCM method," Indust. Tech. ICIT 2008. IEEE Int. Conf. on, pp. 1-4, Apr. 2008.

[161] R. O. Duda and P. E. Hart, Pattern classification and scene analysis, New York, JohnWiley, 1973.

[162] R. A. Johnson, Applied Multivariate Statistical Analysis. Prentice-Hall Inc. 1992.

[163] R. A. Fischer, "A polling scheme using an automated meter reading system for identifying system conditions," M.Sc. thesis, Dept. EE., Michigan Technological Univ., Houghton, MI, USA, 2000.

[164] M. E. Baran, L.-A. Freeman, F. Hanson, and V. Ayers, "Load estimation for load monitoring at distribution substations," IEEE Trans Power Syst., vol. 20, no.1, pp. 164- 170, Feb. 2005.

[165] T. Gonen, Electric Power Distribution System Engineering. Mcgraw-Hill Book Company,1986.

APPENDICES

APPENDIX 1 – State Estimation Test Results

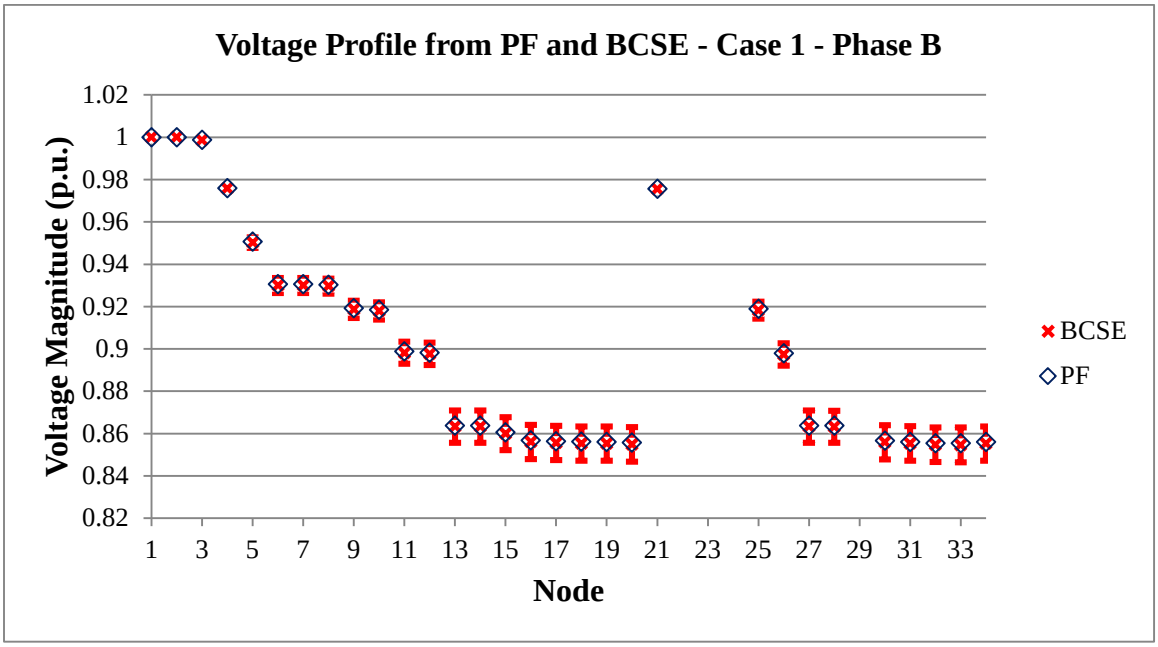


Figure (A1-1): Voltage profile for prototype feeder from PF and BCSE – Case 1 and Phase B.

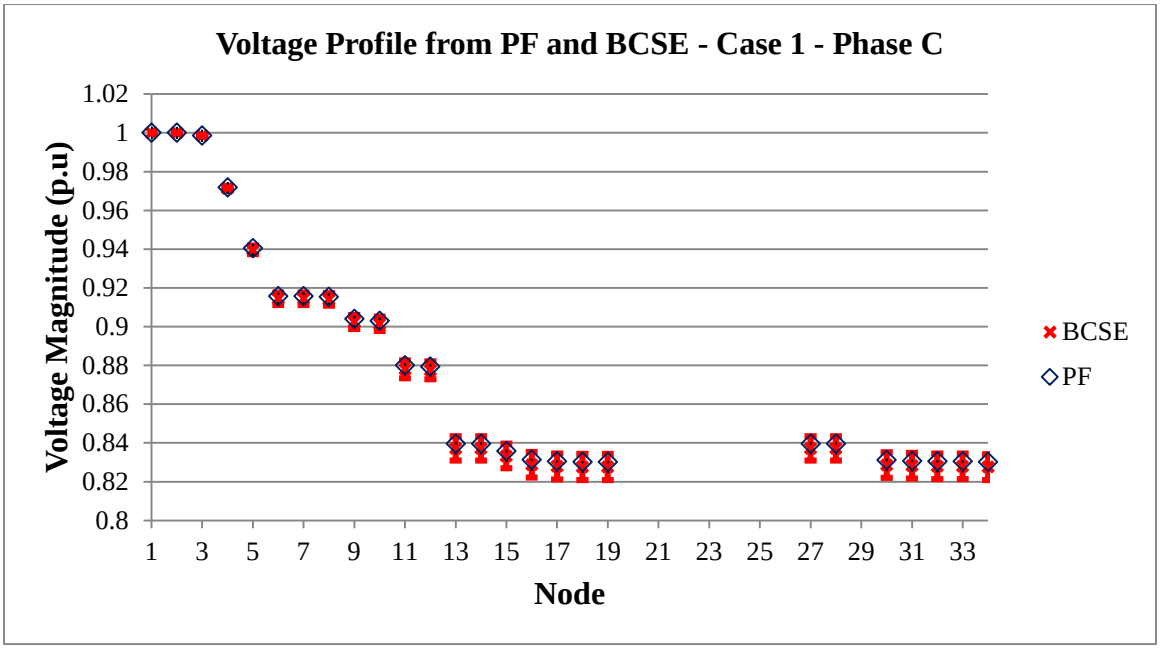


Figure (A1-2): Voltage profile for prototype feeder from PF and BCSE – Case 1 and Phase C.

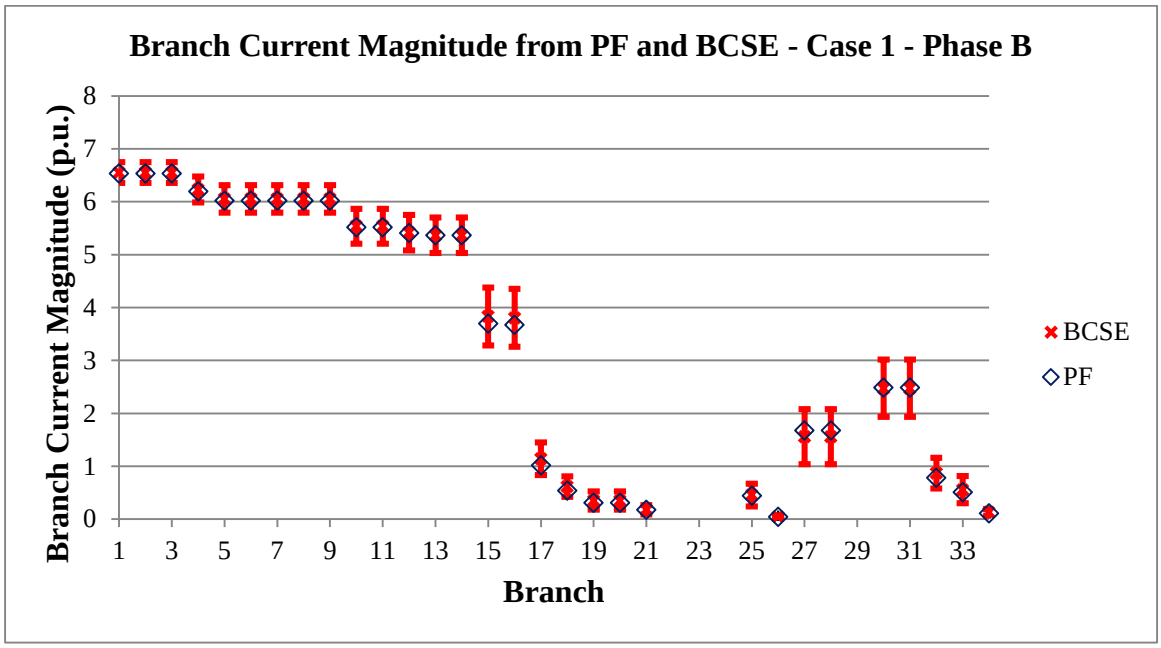


Figure (A1-3): Branch current magnitude for prototype feeder from PF and BCSE – Case 1 and Phase B.

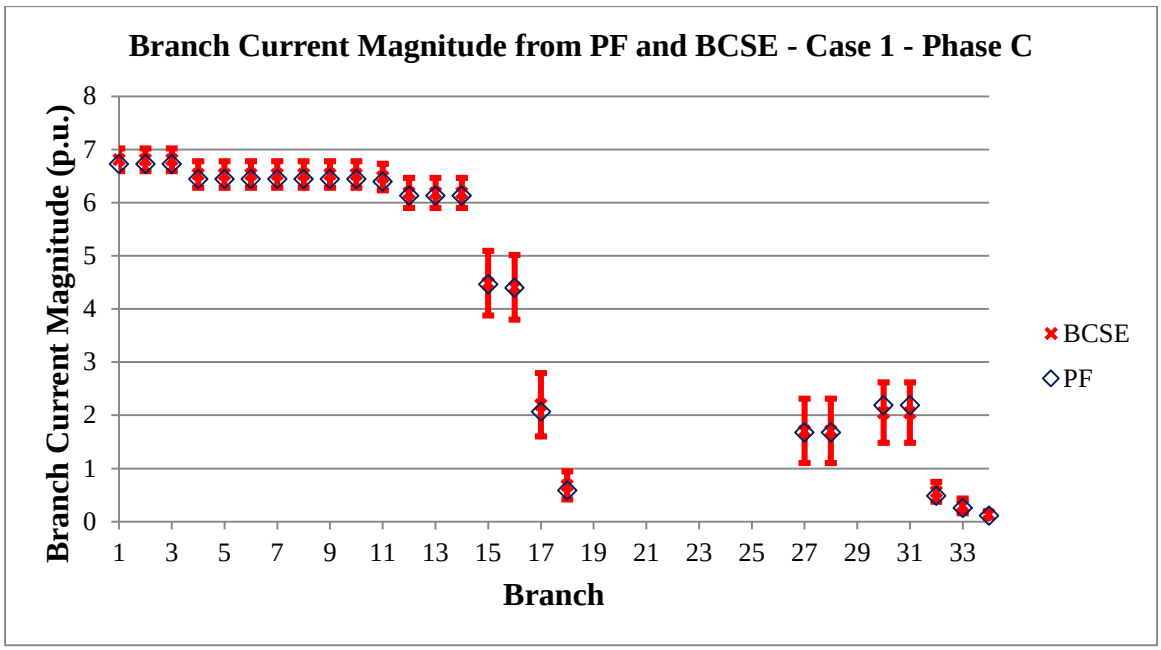


Figure (A1-4): Branch current magnitude for prototype feeder from PF and BCSE – Case 1 and Phase C.

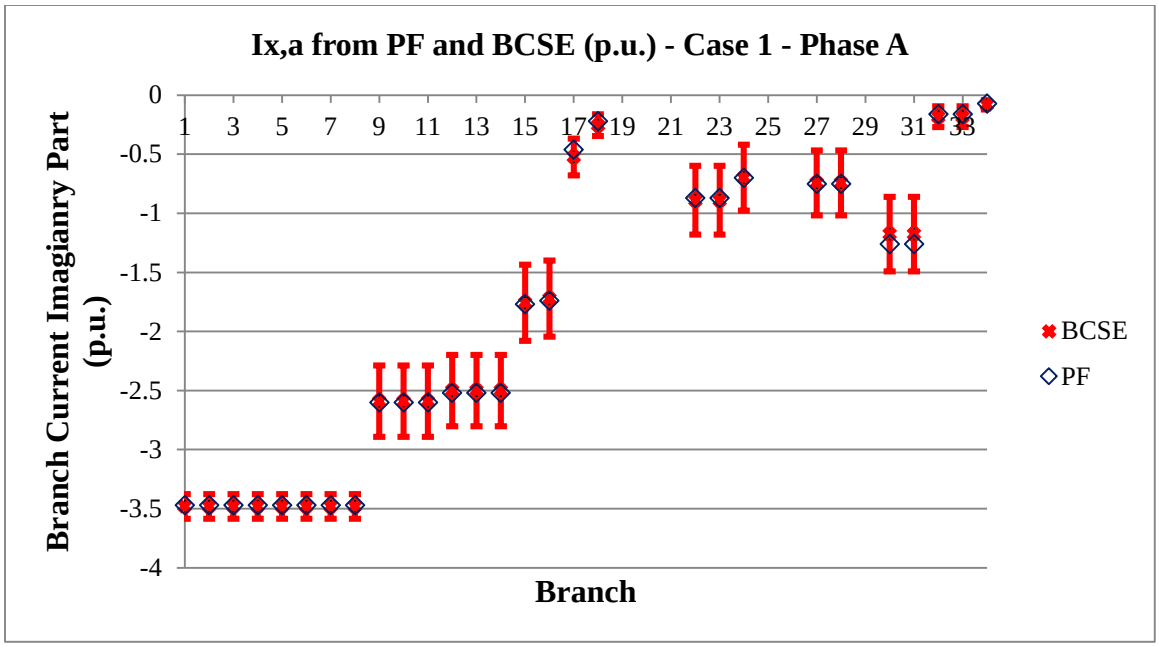


Figure (A1-5): Estimated system states and true values, $x = [I_r, I_x]$, from BCSE and PF; respectively for prototype feeder, Imaginary part – Case 1 and Phase A.

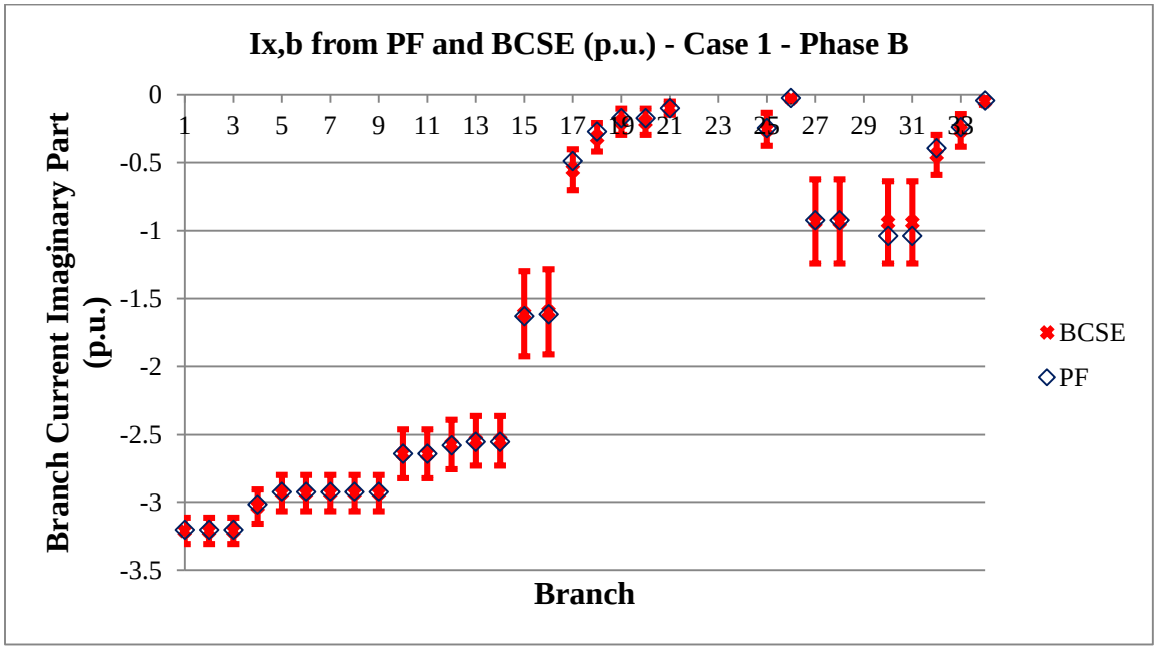


Figure (A1-6): Estimated system states and true values, I_x , from BCSE and PF; respectively for phase B– Case 1.

APPENDIX 2 - MC Simulations for Std Dev Calculation

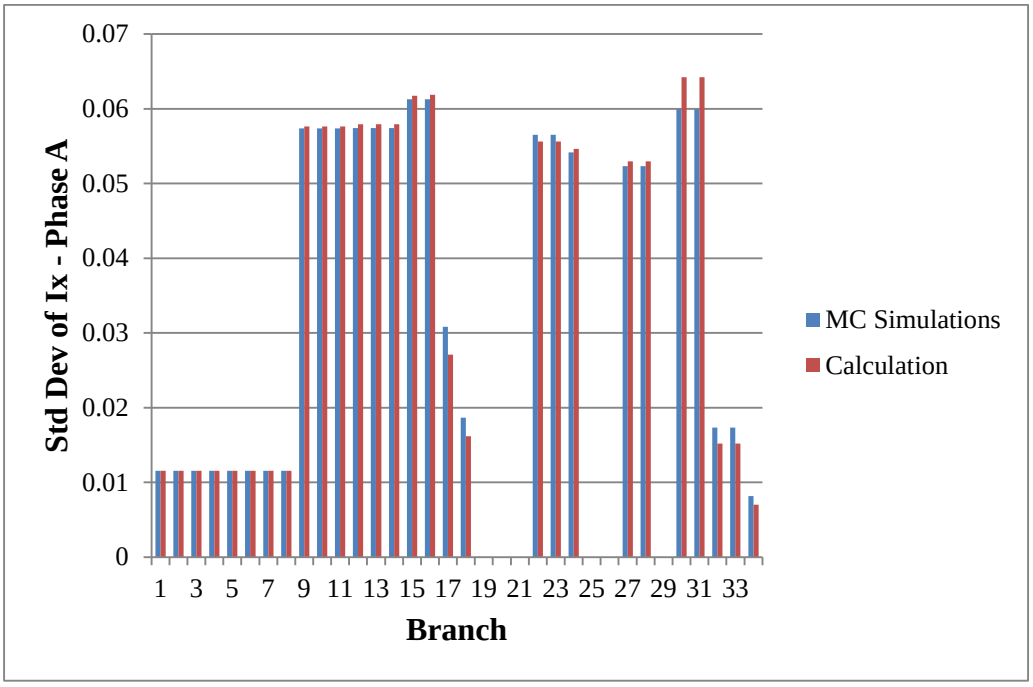


Figure (A2-1): MC simulation vs Calculation for Ix at Phase A.

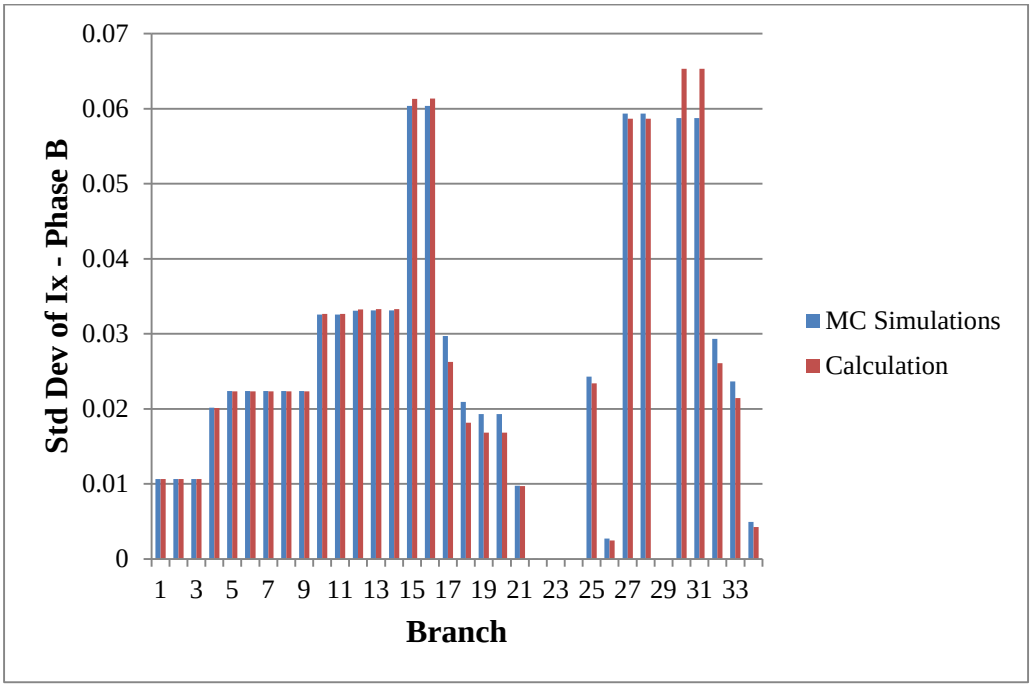


Figure (A2-2): MC simulation vs Calculation for Ix at Phase B.

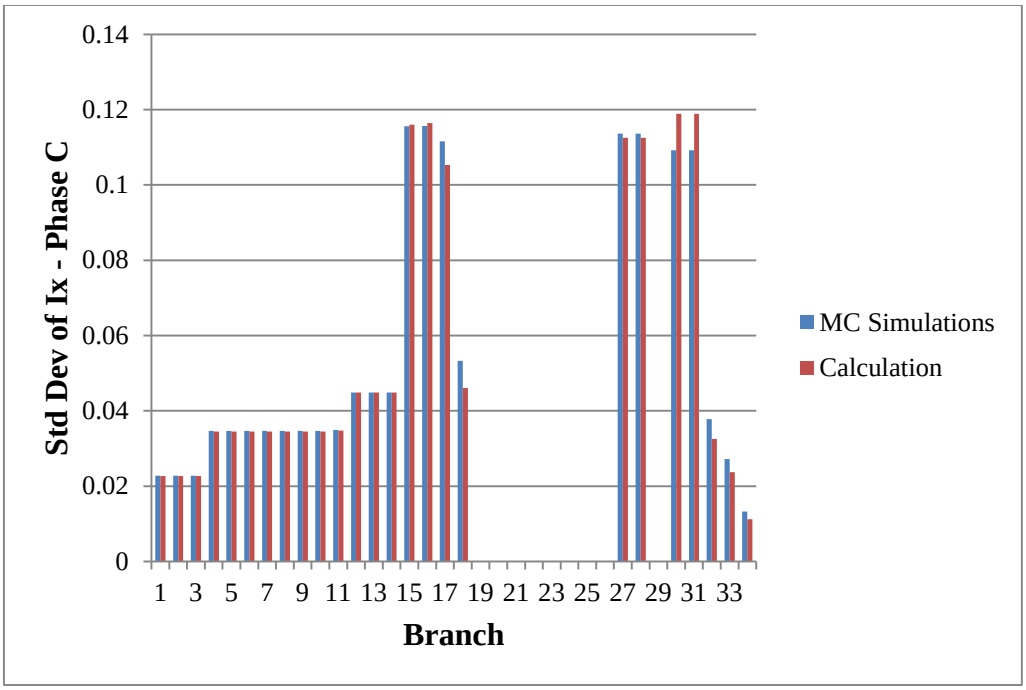


Figure (A2-3): MC simulation vs Calculation for Ix at Phase C.

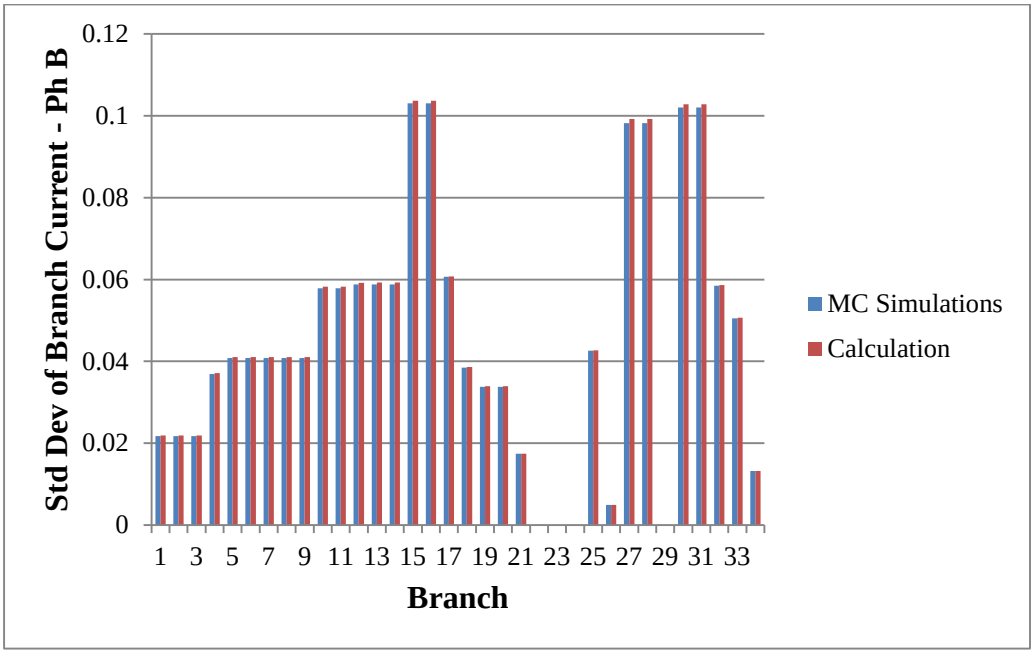


Figure (A2-4): MC simulation vs Calculation for standard deviation of branch current at Phase B.

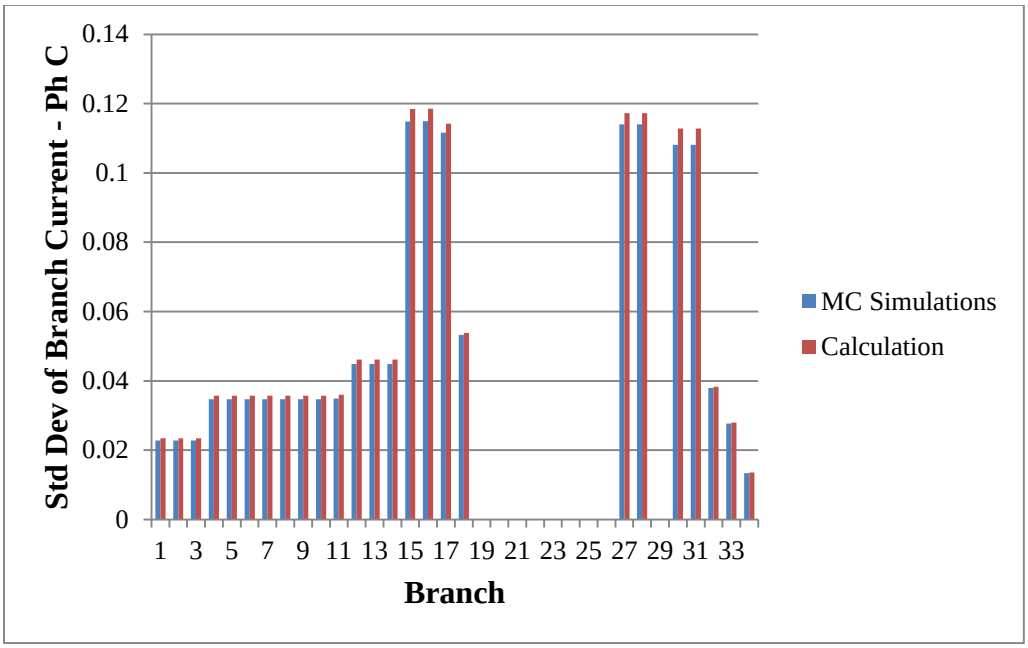


Figure (A2-5): MC simulation vs Calculation for standard deviation of branch current at Phase C.

APPENDIX 3 - Voltage Std Dev for Meter Placement

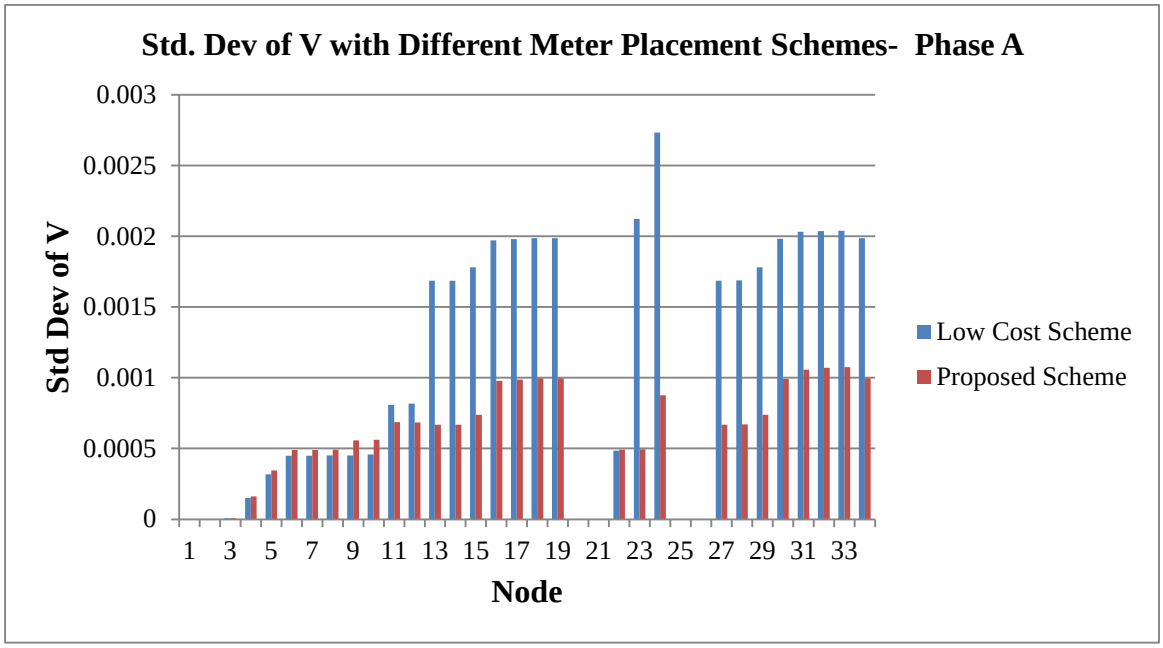


Figure (A3-1): Standard deviation of voltage magnitude profile with different meter placement schemes for phase A.

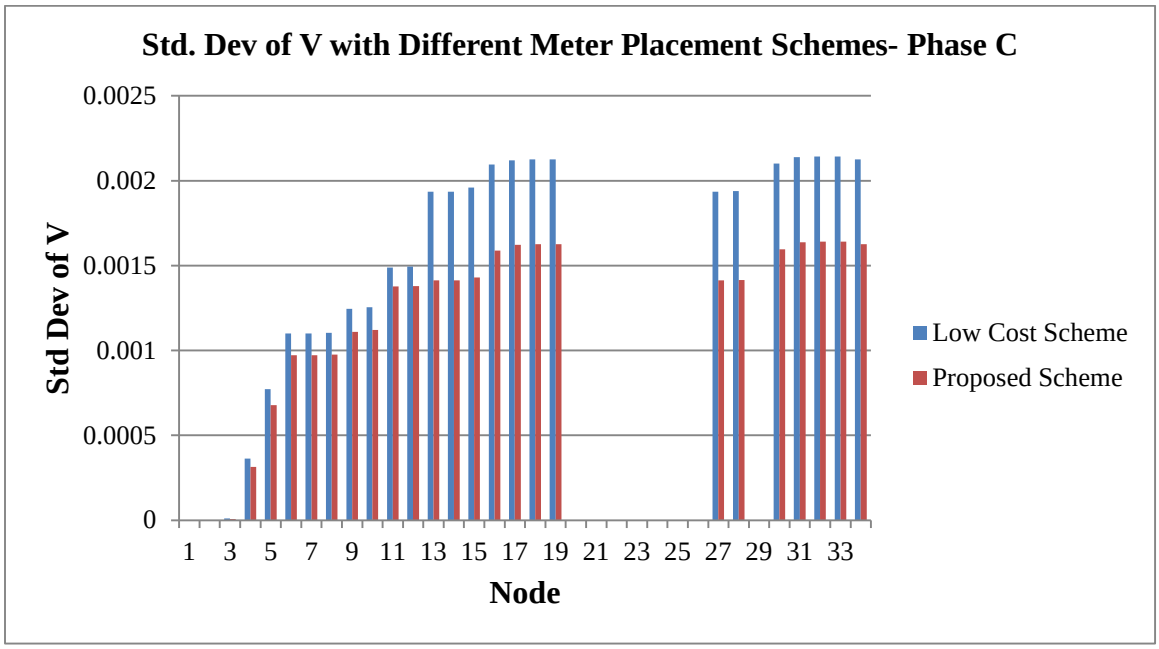


Figure (A3-2): Standard deviation of voltage magnitude profile with different meter placement schemes for phase C.

APPENDIX 4 – Analytical Sensitivity Analysis for a VM and CM Location on a Simplified Feeder

In the "Meter Placement on Distribution Feeders for Volt/VAR Control" chapter, different case studies have been designed for one prototype distribution feeder, i.e. IEEE 34 node feeder. After those case studies, experiments and simulations, the best locations to place one CM and one VM were found. Because of the complexity and non-linear nature of the SE procedure as well as complex nature of the distribution system especially in unbalanced network and loading condition. Here, in this appendix, the analytical approaches have been considered to find the best location of different measurement types: VM and CM, on the special feeder. This special feeder is a total three phase feeder with these characteristics: similar lines for all branches and similar loads for each section. This special feeder is shown in Figure (A4-1).

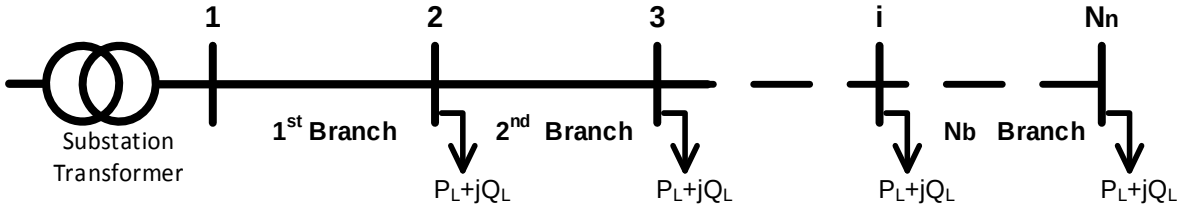


Figure (A4-1): The simplified feeder with same branches and same loads.

Based on the BCSE model, the Jacobian matrix for the measurement set is composed of bus injection, line flow, and voltage measurements, i.e. PMs, CMs, and VMs:

$$H = \begin{bmatrix} H_{PM,CM} \\ \cdots \\ H_{VM} \end{bmatrix} = \begin{bmatrix} H_{PM} \\ \cdots \\ H_{CM} \\ \cdots \\ H_{VM} \end{bmatrix} \quad \text{(A4-1)}$$

where: H_{PM} is the sub-Jacobian matrix for power measurements, H_{CM} is the sub-Jacobian matrix for current measurements and H_{VM} is the sub-Jacobian matrix for voltage measurements.

A4 - 1 : Adding One CM on the Feeder

Here, all the nodes have the same value with the same load estimation error. All the nodes have loads with the same weight: w_L, then $W = w_L.I$. By considering, N_n as the number of nodes and N_b is the number of branches. The sub-Jacobian matrix for power measurements becomes:

$$H_{PM} = \begin{bmatrix} 1 & -1 & \dots & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix}_{N_b \times N_b} \qquad \text{(A4-2)}$$

The Gain matrix then becomes:

$$G_{PM} = (H_{PM})^T W_r H_{PM} = w_L (H_{PM})^T H_{PM} = w_L \begin{bmatrix} 1 & -1 & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 \\ 0 & -1 & 2 & 0 & \vdots \\ 0 & 0 & 0 & \ddots & -1 \\ 0 & 0 & \cdots & -1 & 2 \end{bmatrix} \qquad \text{(A4-3)}$$

Covariance matrix of system states becomes:

$$P_x = (G_{PM})^{-1} = w_L^{-1} \begin{bmatrix} N_b & N_b - 1 & N_b - 2 & \cdots & 1 \\ N_{b-1} & N_b - 1 & N_b - 2 & \cdots & 1 \\ N_{b-2} & N_b - 2 & N_b - 2 & \cdots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 1 & \cdots & 1 \end{bmatrix} \qquad \text{(A4-4)}$$

Therefore, the variance of the system states: first branch: $Var(I_{r,1}) = \frac{N_b}{w_L}$, second branch: $Var(I_{r,i}) = \frac{N_b - i + 1}{w_L}$, …, and N_b^{th} branch: $Var(I_{r,\mathrm{N_b}}) = \frac{1}{w_L}$.

Now, one CM has been placed on the first branch, then the sub-Jacobian matrix for current measurement becomes:

$$H_{CM} = \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix}_{1\times N_b} = e_1^T \tag{A4-5}$$

Consecutively, the Jacobian matrix of the measurement forms by:

$$H = \begin{bmatrix} H_{PM} \\ H_{CM} \end{bmatrix} = \begin{bmatrix} H_{PM} \\ \cdots \\ e_1^T \end{bmatrix} \tag{A4-6}$$

Here, it is assumed: $w_C = w_L$. The gain matrix calculated by:

$$G = w_L H^T H = w_L \begin{bmatrix} H_{PM}^T & \vdots & e_1 \end{bmatrix} \begin{bmatrix} H_{PM} \\ \cdots \\ e_1^T \end{bmatrix} = w_L (e_1 e_1^T + H_{PM}^T H_{PM}) \tag{A4-7}$$

To derive the covariance matrix of the system states: $P_x' = G^{-1} = w_L^{-1}(e_1 e_1^T + H_{PM}^T H_{PM})^{-1}$, by Lemma 1:

$$\begin{aligned} P_x' = G^{-1} &= w_L^{-1}(H_{PM}^T H_{PM} + e_1 e_1^T)^{-1} \\ &= w_L^{-1}[(H_{PM}^T H_{PM})^{-1} - \frac{1}{1+G_{PM,(1,1)}^{-1}}(H_{PM}^T H_{PM})^{-1} e_1 e_1^T (H_{PM}^T H_{PM})^{-1}] \end{aligned}$$

$$\begin{aligned} P_x' &= P_x - \frac{1}{1+P_{x,(1,1)}} P_x e_1 e_1^T P_x \\ &= P_x - \frac{1}{1+P_{x,(1,1)}} \begin{bmatrix} P_{x,(1,1)}^2 & P_{x,(1,1)}P_{x,(1,2)} & P_{x,(1,1)}P_{x,(3,2)} & \cdots & P_{x,(1,1)}P_{x,(1,N_b)} \\ P_{x,(1,1)}P_{x,(2,1)} & P_{x,(2,1)}P_{x,(1,2)} & \cdots & \vdots & \vdots \\ P_{x,(1,1)}P_{x,(3,1)} & \cdots & P_{x,(3,1)}P_{x,(1,3)} & & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ P_{x,(1,1)}P_{x,(N_b,1)} & \cdots & \cdots & \cdots & P_{x,(N_b,1)}P_{x,(1,N_b)} \end{bmatrix} \end{aligned}$$

$$P'_x = \begin{bmatrix} P_{x,(1,1)} - \alpha_1 P^2_{x,(1,1)} & \cdots & \cdots & \cdot \\ \vdots & P_{x,(2,2)} - \alpha_1 P_{x,(2,1)} P_{x,(1,2)} & \cdots & \vdots \\ \vdots & & & \vdots \\ \vdots & \vdots & \vdots \ddots & \vdots \\ \cdot & \cdots & \cdots & P_{x,(N_b,N_b)} - \alpha_1 P_{x,(N_b,1)} P_{x,(1,N_b)} \end{bmatrix}$$

where: $\alpha_1 = \dfrac{1}{1 + P_{x,(1,1)}}$

New elements of covariance matrix become, if a CM has been placed on branch 1:

$$P'_{x,(i,i)} = P_{x,(i,i)} - \frac{1}{1 + P_{x,(1,1)}} P^2_{x,(1,i)} \tag{A4-8}$$

By adding one CM at the first branch, new variances become:

$$Var(I_{r,1}) = w_L^{-1}(N_b - \frac{1}{1+N_b} N_b) = w_L^{-1}(\frac{N_b}{1+N_b}) \tag{A4-9}$$

$$Var(I_{r,i}) = w_L^{-1}(N_b - i + 1 - \frac{(N_b - i + 1)^2}{1+N_b}) = w_L^{-1}(\frac{N_b i - i^2 + i}{1+N_b}) \tag{A4-10}$$

Now, place one CM on the last branch:

$$H_{CM} = \begin{bmatrix} 0 & 0 & \cdots & 1 \end{bmatrix}_{1 \times N_b} = e^T_{N_b} \tag{A4-11}$$

$$H = \begin{bmatrix} H_{PM} \\ H_{CM} \end{bmatrix} = \begin{bmatrix} H_{PM} \\ \cdots \\ e^T_{N_b} \end{bmatrix} \tag{A4-12}$$

In case of $w_C = w_L$, the gain matrix becomes:

$$G = w_L H^T H = w_L \begin{bmatrix} H_{PM}^T & \vdots & e_{N_b} \end{bmatrix} \begin{bmatrix} H_{PM} \\ \cdots \\ e_{N_b}^T \end{bmatrix} = w_L (e_{N_b} e_{N_b}^T + H_{PM}^T H_{PM}) \tag{A4-13}$$

To derive the covariance matrix of the system states: $P_x' = G^{-1} = w_L^{-1}(e_{N_b} e_{N_b}^T + H_{PM}^T H_{PM})^{-1}$, by Lemma 1 the covariance matrix of the system states become:

$$\begin{aligned} P_x' &= G^{-1} = w_L^{-1}(H_{PM}^T H_{PM} + e_{N_b} e_{N_b}^T)^{-1} \\ &= w_L^{-1}[(H_{PM}^T H_{PM})^{-1} - \frac{1}{1 + G_{PM,(N_b,N_b)}^{-1}} (H_{PM}^T H_{PM})^{-1} e_{N_b} e_{N_b}^T (H_{PM}^T H_{PM})^{-1}] \end{aligned} \tag{A4-14}$$

$$P_x' = \begin{bmatrix} P_{x,(1,1)} - \alpha_{N_b} P_{x,(N_b,1)}^2 & & & \cdots & & \\ \vdots & P_{x,(2,2)} - \alpha_{N_b} P_{x,(N_b,2)}^2 & \cdots & \vdots & & \vdots \\ \vdots & \cdots & & & & \vdots \\ \vdots & \vdots & \vdots & \ddots & & \vdots \\ \cdots & \cdots & \cdots & \cdots & P_{x,(N_b,N_b)} - \alpha_{N_b} P_{x,(N_b,N_b)}^2 & \end{bmatrix} \tag{A4-15}$$

where: $\alpha_{N_b} = \dfrac{1}{1 + P_{x,(N_b,N_b)}}$

If a CM has been placed on the last branch, new elements of covariance matrix become,:

$$P_{x,(i,i)}' = P_{x,(i,i)} - \frac{1}{1 + P_{x,(N_b,N_b)}} P_{x,(N_b,i)}^2 \tag{A4-16}$$

By adding one CM at the last branch, new variances become:

$$Var(I_{r,1}) = w_L^{-1}(N_b - \frac{1}{1+1}) = w_L^{-1}(N_b - 1/2) \tag{A4-17}$$

$$Var(I_{r,i}) = w_L^{-1}(N_b - i + 1 - \frac{1}{1+1}) = w_L^{-1}(N_b - i + 1/2) \tag{A4-18}$$

Generalization for placing one CM on the branch k^{th}:

The sub-Jacobian matrix for current measurement constructed through:

$$H_{CM} = \begin{bmatrix} 0 & \cdots & 1 & \cdots & 0 \end{bmatrix}_{1\times N_b} = e_k^T \tag{A4-19}$$

Thus, the Jacobian matrix of the measurement becomes:

$$H = \begin{bmatrix} H_{PM} \\ H_{CM} \end{bmatrix} = \begin{bmatrix} H_{PM} \\ \cdots \\ e_k^T \end{bmatrix} \tag{A4-20}$$

In case of $w_C = w_L$, the gain matrix is calculated by:

$$G = w_L H^T H = w_L \begin{bmatrix} H_{PM}^T & \vdots & e_k \end{bmatrix} \begin{bmatrix} H_{PM} \\ \cdots \\ e_k^T \end{bmatrix} = w_L (e_k e_k^T + H_{PM}^T H_{PM}) \tag{A4-21}$$

To derive the covariance matrix of system states: $P_x' = G^{-1} = w_L^{-1}(H_{PM}^T H_{PM} + e_k e_k^T)^{-1}$, Lemma 1 has been deployed and this matrix can be constructed by:

$$P_x' = P_x - \frac{1}{1+P_{x,(k,k)}} \begin{bmatrix} P_{x,(1,k)}^2 & P_{x,(1,k)}P_{x,(k,2)} & \cdot & \cdots & P_{x,(1,k)}P_{x,(k,N_b)} \\ P_{x,(2,k)}P_{x,(k,1)} & P_{x,(2,k)}P_{x,(k,2)} & \cdot & \cdot & \cdot \\ \vdots & \cdot & \ddots & & \vdots \\ \vdots & \vdots & \cdot & \ddots & \vdots \\ P_{x,(N_b,k)}P_{x,(k,1)} & P_{x,(N_b,k)}P_{x,(k,2)} & \cdots & \cdot & P_{x,(N_b,k)}P_{x,(k,N_b)} \end{bmatrix} \tag{A4-22}$$

$$P'_x = \begin{bmatrix} P_{x,(1,1)} - \alpha_k P^2_{x,(k,1)} & \cdot & \cdot & \cdot & \cdot \\ \vdots & P_{x,(2,2)} - \alpha_k P^2_{x,(k,2)} & & \cdot & \vdots \\ \vdots & \cdot & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ \cdot & \cdot & \cdot & \cdots & P_{x,(N_b,N_b)} - \alpha_k P^2_{x,(k,N_b)} \end{bmatrix} \quad \text{(A4-23)}$$

where: $\alpha_k = \dfrac{1}{1 + P_{x,(k,k)}}$

If a CM is placed on the branch k^{th}, new elements of covariance matrix become:

$$P'_{x,(i,i)} = P_{x,(i,i)} - \frac{1}{1 + P_{x,(k,k)}} P^2_{x,(k,i)} \quad \text{(A4-24)}$$

By adding one CM at the branch k^{th}, new variances will become:

$$Var(I_{r,1}) = w_L^{-1}(N_b - \frac{1}{1 + N_b - k + 1}(N_b - k + 1)^2) = w_L^{-1}[\frac{N_b k - (k-1)^2}{N_b - k + 2}] \quad \text{(A4-25)}$$

$$Var(I_{r,N_b}) = w_L^{-1}(1 - \frac{1}{N_b - k + 2}) = w_L^{-1}(\frac{N_b - k + 1}{N_b - k + 2}) \quad \text{(A4-26)}$$

These equations (A4-25) and (A4-26) show that the best place for one CM on the simple and consistent feeder is the first branch. If one CM is placed on the last branch, this measurement will have the least impact on the quality of the systems states.

A4 – 2 : Adding One VM on the Feeder

Here, one voltage measurement will be placed on the feeder. By considering all the nodes have the same loads with same load estimation error, i.e. weights of load estimation: w_L, then $W = w_L.I$ and from (A4-2). Then, the Gain matrix becomes:

$$G_{PM} = (H_{PM})^T W_r H_{PM} = w_L (H_{PM})^T H_{PM} = w_L \begin{bmatrix} 1 & -1 & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 \\ 0 & -1 & 2 & 0 & \vdots \\ 0 & 0 & 0 & \ddots & -1 \\ 0 & 0 & \cdots & -1 & 2 \end{bmatrix} \quad \text{(A4-27)}$$

Then, the Covariance matrix of system states can be calculated by:

$$P_x = (G_{PM})^{-1} = w_L^{-1} \begin{bmatrix} N_b & N_b - 1 & N_b - 2 & \cdots & 1 \\ N_{b-1} & N_b - 1 & N_b - 2 & \cdots & 1 \\ N_{b-2} & N_b - 2 & N_b - 2 & \cdots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 1 & \cdots & 1 \end{bmatrix} \quad \text{(A4-28)}$$

Variance of the system states are for the first branch $Var(I_{r,1}) = \frac{N_b}{w_L}$, for the second branch $Var(I_{r,i}) = \frac{N_b - i + 1}{w_L}$, and for the last branch $Var(I_{r,\mathrm{N}_b}) = \frac{1}{w_L}$.

Now, place one VM on the second node, construct Jacobian matrix through $H = \begin{bmatrix} H_{PM} \\ H_{VM} \end{bmatrix}$ and consider that all resistance and impedance for all branches are the same.

$$R_a = [r_1 \quad 0 \quad \cdots \quad 0]_{1 \times N_b}, X_a = [x_1 \quad 0 \quad \cdots \quad 0]_{1 \times N_b} \quad \text{(A4-29)}$$

Now, the gain matrix can be constructed by:

$$G = (H_{PM})^T W_L H_{PM} + R_a^T W_v R_a + X_a^T W_v X_a$$

$$= (H_{PM})^T W_L H_{PM} + w_v . r_1^2 \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} + w_v . x_1^2 \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} \quad \text{(A4-30)}$$

In case of equality of the weighting factors for load estimations and voltage measurements, i.e. $w_V = w_L$ and equal impedances for all the branches: $z_i^2 = r_i^2 + x_i^2$, the previous covariance matrix is $P_x = (G_{PM})^{-1}$, thus the new covariance matrix becomes:

$$P_x'' = G^{-1} = [(H_{PM})^T W_L H_{PM} + w_v (r_1^2 + x_1^2) \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}]^{-1}$$

$$= [(H_{PM})^T W_L H_{PM} + G_{VM}]^{-1} \qquad \text{(A4-31)}$$

By using Lemma 1, the new covariance matrix can be calculated through:

$$P_x'' = [(H_{PM})^T W_L H_{PM} + G_{VM}]^{-1} = (G_{PM} + G_{VM})^{-1}$$

$$= G_{PM}^{-1} - \frac{1}{1 + trace(G_{VM} G_{PM}^{-1})} G_{PM}^{-1} G_{VM} G_{PM}^{-1} \qquad \text{(A4-32)}$$

$$P_x'' = P_x - \frac{1}{1 + P_{x,(1,1)}} (w_v . z_1^2) \begin{bmatrix} P_{x,(1,1)}^2 & P_{x,(1,1)} P_{x,(1,2)} & \cdots & \cdots & P_{x,(1,1)} P_{x,(1,N_b)} \\ P_{x,(2,1)} P_{x,(1,1)} & P_{x,(2,1)} P_{x,(1,2)} & \cdots & \cdots & \vdots \\ \vdots & \vdots & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ P_{x,(N_b,1)} P_{x,(1,1)} & P_{x,(N_b,1)} P_{x,(1,2)} & \cdots & \cdots & P_{x,(N_b,1)} P_{x,(1,N_b)} \end{bmatrix} \qquad \text{(A4-33)}$$

$$P_x'' = \begin{bmatrix} P_{x,(1,1)} - \alpha_1 P_{x,(1,1)}^2 & & & \cdot & \cdot \\ \vdots & P_{x,(2,2)} - \alpha_1 P_{x,(2,1)} P_{x,(1,2)} & & \cdot & \vdots \\ \vdots & \cdot & \cdot & & \vdots \\ \vdots & \vdots & & & \\ \cdot & \cdot & \cdot & P_{x,(N_b,N_b)} - \alpha_1 P_{x,(N_b,1)} P_{x,(1,N_b)} & \end{bmatrix} \qquad \text{(A4-34)}$$

where: $\alpha_1 = \dfrac{z_1^2}{1 + P_{x,(1,1)}}$.

If a VM has been placed on 2nd node, new elements of covariance matrix become,:

$$P''_{x,(i,i)} = P_{x,(i,i)} - \frac{z_1^2}{1+P_{x,(1,1)}} P^2_{x,(1,i)} \tag{A4-35}$$

By adding one VM at 2nd node, new variances for system state will become:

$$\begin{aligned} Var(I_{r,1}) &= w_L^{-1}(N_b - \frac{z_1^2}{1+N_b} N_b^2) \\ &= w_L^{-1}(\frac{N_b^2 + N_b - z_1^2 N_b^2}{1+N_b}) \end{aligned} \tag{A4-36}$$

$$Var(I_{r,i}) = w_L^{-1}(N_b - i + 1 - \frac{z_1^2 (N_b - i + 1)^2}{1+N_b}) \tag{A4-37}$$

It is clear from the results that moving one VM on the feeder from the beginning of the feeder to the 2nd node is negligible. Now, place one VM on third node; VM has actually moved from the 2nd to the 3rd node, the gain matrix can be constructed through:

$$\begin{aligned} G' &= (H_{PM})^T W_L H_{PM} + w_v . \begin{bmatrix} r_1^2 & r_1 r_2 & 0 & 0 \\ r_2 r_1 & r_2^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} + w_v . \begin{bmatrix} x_1^2 & x_1 x_2 & 0 & 0 \\ x_2 x_1 & x_2^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} \\ &\approx G + B_1 + B_2 \end{aligned} \tag{A4-38}$$

where: $B_1 = w_v . \begin{bmatrix} r_1^2 + x_1^2 & 0 & 0 & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}$ and $B_2 = w_v . \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & r_2^2 + x_2^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}$ by removing the non-diagonal elements from the original matrix. Then, by applying Theorem 1:

$$\begin{aligned} P'''_x &= G^{-1} = (G_{PM} + B_1 + B_2)^{-1} \\ &= (G_{PM} + B_1)^{-1} - g_3 (G_{PM} + B_1)^{-1} B_2 (G_{PM} + B_1)^{-1} \end{aligned} \tag{A4-39}$$

Where: $g_3 = \dfrac{1}{1+trace((G_{PM}+B_1)^{-1}B_2)}$ and P''' can be calculated by:

$$\begin{aligned} P_x''' &= G^{-1} = (G_{PM}+B_1+B_2)^{-1} \\ &= P_x'' - g_3 P_x'' B_2 P_x'' \end{aligned} \tag{A4-40}$$

and

$$g_3 = \frac{1}{1+trace(P_x''B_2)} = \frac{1}{1+z_2^2(P_{x,(2,2)} - \alpha_1 P_{x,(2,1)}P_{x,(1,2)})} \tag{A4-41}$$

$$\begin{aligned} P_x''' &= P_x'' - g_3 P_x'' B_2 P_x'' \\ &= P_x'' - g_3 \begin{bmatrix} z_2^2 P''_{x,(1,2)}P''_{x,(2,1)} & z_2^2 P''_{x,(1,2)}P''_{x,(2,2)} & \cdots & \cdots & z_2^2 P''_{x,(1,2)}P''_{x,(2,N_b)} \\ z_2^2 P''_{x,(2,2)}P''_{x,(2,1)} & z_2^2 P''_{x,(2,2)}P''_{x,(2,2)} & \cdots & \cdots & \cdot \\ \vdots & \vdots & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ z_2^2 P''_{x,(N_b,2)}P''_{x,(2,1)} & z_2^2 P''_{x,(N_b,2)}P''_{x,(2,2)} & \cdots & \cdots & z_2^2 P''_{x,(N_b,2)}P''_{x,(2,N_b)} \end{bmatrix} \end{aligned} \tag{A4-42}$$

If a VM has been placed on node 3, new elements of covariance matrix become:

$$P'''_{x,(i,i)} = P''_{x,(i,i)} - g_3 z_2^2 P''_{x,(i,2)}P''_{x,(2,i)} = P''_{x,(i,i)} - g_3 z_2^2 (P''_{x,(i,2)})^2 \Rightarrow P'''_{x,(i,i)} < P''_{x,(i,i)}$$

These results show that the variances of the system states have been reduced by moving the VM from node 2 to node 3, like from node 1 to node 2. Therefore, placing VM at the end of the feeder improves the system states estimation process.

Lemma 1: If A and $A+B$ are invertible, and B has rank 1, then let $g = trace(BA^{-1})$. Then $g \neq -1$ and

$$(A+B)^{-1} = A^{-1} - \frac{1}{1+g}A^{-1}BA^{-1}$$

Theorem 1: From Lemma 1, we can take a general $A+B$ that is invertible and write it as $A+B = A+B_1+B_2+\cdots+B_r$, where B_i each have rank 1 such that each $A+B_1+B_2+\cdots+B_r$

is invertible (such a decomposition always exists if $A+B$ is invertible and $rank(B)=r$). Then we get:

Theorem: Let A and $A+B$ be non-singular matrices, and let B has a rank $r>0$. Let $B=B_1+B_2+\cdots+B_r$, where each B_i has rank 1, and each $C_{k+1}=A+B_1+B_2+\cdots+B_k$ is non-singular. Setting $C_1=A$, then

$$C_{k+1}^{-1}=C_k^{-1}-g_kC_k^{-1}B_kC_k^{-1}$$

where: $g_k=\dfrac{1}{1+trace(C_k^{-1}B_k)}$. In particular,

$$(A+B)^{-1}=C_r^{-1}-g_rC_r^{-1}B_rC_r^{-1}$$

If the rank B of is 0, then $B=0$, so $(A+B)^{-1}=A^{-1}$.

APPENDIX 5 – Impact of the System State Variances on Nodal Voltage Estimates in BCSE

In this appendix, the impact of the system state variances on voltage estimate variance has been investigated.

Now considering a branch k between node k and node $k+1$, we can write:

$$\tilde{V}_{k+1} = \tilde{V}_k - (r_k + jx_k)\tilde{I}_k \tag{A5-1}$$

Thus, we will have one real part and one imaginary part:

$$\begin{aligned} V_{k+1,r} &= V_{k,r} - r_k I_{k,r} + x_k I_{k,x} \\ V_{k+1,x} &= V_{k,r} - x_k I_{k,r} - r_k I_{k,x} \end{aligned} \tag{A5-2}$$

Now, partial derivation of $V_{k+1,r}$ and $V_{k+1,x}$ in regard of system states can be derived:

$$\frac{\partial V_{k+1,r}}{\partial I_{k,r}} = -r_k \,,\ \frac{\partial V_{k+1,r}}{\partial I_{k,x}} = x_k \tag{A5-3}$$

and

$$\frac{\partial V_{k+1,x}}{\partial I_{k,r}} = -x_k \,,\ \frac{\partial V_{k+1,x}}{\partial I_{k,x}} = -r_k \tag{A5-4}$$

By chain rules, i.e. $y = f(x)$, the voltage variances can be calculated. Then, the covariance matrix of interesting quantities can be calculated from [33, 42]:

$$R_y = F(H^T R^{-1} H)^{-1} F^T \tag{A5-5}$$

where: H is the Jacobian of the measurement functions, F is the Jacobian of the functions of interesting quantities, $f(.)$, and R is a diagonal covariance matrix containing the variances of the measurements, $W = R^{-1}$. Now, $(H^T R^{-1} H)^{-1} = G^{-1} = P_x$. Variances of real and imaginary voltages can thus be calculated as follows:

$$R_y = F(H^T R^{-1} H)^{-1} F^T = \begin{bmatrix} -r_k & x_k \\ -x_k & -r_k \end{bmatrix} \begin{bmatrix} \sigma^2_{I,r,k} & 0 \\ 0 & \sigma^2_{I,x,k} \end{bmatrix} \begin{bmatrix} -r_k & -x_k \\ x_k & -r_k \end{bmatrix}$$

$$R_y = \begin{bmatrix} r_k^2 \sigma^2_{I,r,k} + x_k^2 \sigma^2_{I,x,k} & r_k x_k \sigma^2_{I,r,k} - r_k x_k \sigma^2_{I,x,k} \\ r_k x_k \sigma^2_{I,r,k} - r_k x_k \sigma^2_{I,x,k} & x_k^2 \sigma^2_{I,r,k} + r_k^2 \sigma^2_{I,x,k} \end{bmatrix} \quad \text{(A5-6)}$$

In case of $\sigma^2_{I,r,k} = \sigma^2_{I,x,k}$, the variance of voltages can be derived using the following equation:

$$\sigma^2_{V,r,k+1} = (r_k^2 + x_k^2)\sigma^2_{I,r,k} \text{ and } \sigma^2_{V,x,k+1} = (r_k^2 + x_k^2)\sigma^2_{I,r,k} \quad \text{(A5-7)}$$

Based on the previous results, the variance of the nodal voltages depends on the variance of the real and imaginary parts of the branch currents, i.e. the system states in BCSE. In other words, the nodal voltages can be estimated more precisely by estimating the system states in a more accurate way.