

DEVELOPMENT OF NOVEL  
ELECTRICAL POWER DISTRIBUTION  
SYSTEM STATE ESTIMATION AND  
METER PLACEMENT ALGORITHMS  
SUITABLE FOR PARALLEL  
PROCESSING

A thesis submitted for the degree of Doctor of  
Philosophy  
by  
Nazia Nusrat

Brunel Institute of Power Systems (BIPS)  
Department of Electronic and Computer  
Engineering  
College of Engineering, Design and Physical  
Sciences  
Brunel University London

Supervisor:  
Prof Gareth A Taylor

8th May 2015

# ABSTRACT

The increasing penetration of distributed generation, responsive loads and emerging smart metering technologies will continue the transformation of distribution systems from passive to active network conditions. In such active networks, State Estimation (SE) tools will be essential in order to enable extensive monitoring and enhanced control technologies. In future distribution management systems, the novel electrical power distribution system SE requires development in a scalable manner in order to accommodate small to massive size networks, be operable with limited real time measurements and a restricted time frame. Furthermore, a significant phase of new sensor deployment is inevitable to enable distribution system SE, since present-day distribution networks lack the required level of measurement and instrumentation.

In the above context, the research presented in this thesis investigates five SE optimization solution methods with various case studies related to expected scenarios of future distribution networks to determine their suitability. Hachtel's Augmented Matrix method is proposed and developed as potential SE optimizer for distribution systems due to its potential performance characteristics with regard to accuracy and convergence. Differential Evolution Algorithm (DEA) and Overlapping Zone Approach (OZA) are investigated to achieve scalability of SE tools; followed by which the network division based OZA is proposed and developed. An OZA requiring additional measurements is also proposed to provide a feasible solution for voltage estimation at a reduced computation cost. Realising the requirement of additional measurements deployment to enable distribution system SE, the development of a novel meter placement algorithm that provides economical and feasible solutions is demonstrated. The algorithm is strongly focused on reducing the voltage estimation errors and is capable of reducing the error below desired threshold with limited measurements. The scalable SE solution and meter placement algorithm are applied on a multi-processor system in order to examine effective reduction of computation time. Significant improvement in computation time is observed in both cases by dividing the problem into smaller segments. However, it is important to note that enhanced network division reduces computation time further at the cost of accuracy of estimation. Different networks including both idealised (16, 77, 356 and 711 node UKGDS) and real (40 and 43 node EG) distribution network data are used as appropriate to the requirement of the applications throughout this thesis.

# CONTENTS

<b>ABSTRACT .....</b>	<b>II</b>
<b>CONTENTS.....</b>	<b>III</b>
<b>LIST OF FIGURES.....</b>	<b>VIII</b>
<b>LIST OF TABLES.....</b>	<b>XIII</b>
<b>ACKNOWLEDGEMENT .....</b>	<b>XIV</b>
<b>DECLARATION .....</b>	<b>XV</b>
<b>COPYRIGHT .....</b>	<b>XVI</b>
<b>ABBREVIATIONS.....</b>	<b>XVII</b>
<b>NOMENCLATURE .....</b>	<b>XVIII</b>
<b>CHAPTER 1</b>	
<b>INTRODUCTION.....</b>	<b>1</b>
<b>1.1 Research Background.....</b>	<b>1</b>
<b>1.2 Rationale and Objectives .....</b>	<b>3</b>
<b>1.3 Contribution of the Research .....</b>	<b>5</b>
<b>1.4 List of Publications .....</b>	<b>6</b>
<b>1.5 Thesis Outline .....</b>	<b>8</b>
<b>CHAPTER 2</b>	
<b>LITERATURE REVIEW .....</b>	<b>10</b>
<b>2.1. General Overview of Power Distribution System State Estimation .....</b>	<b>10</b>
2.1.1 Challenges of the Development of Distribution SE Tools.....	10
2.1.2 Existing Research on Distribution SE.....	12
2.1.3 Existing Research on Scalable and Distributed SE.....	16
<b>2.2. General Overview of Meter Placement Algorithms .....</b>	<b>19</b>
2.2.1 Motivation and Challenges .....	19
2.2.2 Existing Research.....	20
<b>2.3. Concluding Remarks and Discussion .....</b>	<b>23</b>
<b>CHAPTER 3</b>	
<b>POWER SYSTEM STATE ESTIMATION .....</b>	<b>25</b>

<b>3.1</b>	<b>Power System State Estimation Building Blocks</b> .....	<b>25</b>
3.1.1	Input and Output Data .....	26
3.1.2	Network Model and Components .....	28
<b>3.2</b>	<b>The Non-linear SE Model: Least Squares Problem</b> .....	<b>30</b>
<b>3.3</b>	<b>Measurement Functions and Measurement Jacobians</b> .....	<b>32</b>
<b>3.4</b>	<b>Assumptions and Hypothesis</b> .....	<b>36</b>
3.4.1	System Observability .....	36
3.4.2	Normal Error Distribution .....	36
3.4.3	Assumptions in Gauss-Newton Recursion .....	37
3.4.4	Static Topology and Balanced System .....	38
3.4.5	Transformer and Branches .....	38
3.4.6	Measurement Errors .....	38
<b>3.5</b>	<b>Conventional SE Solution Processes as Potential DSSE Tools</b> .....	<b>38</b>
3.5.1	Weighted Least Squares (WLS) .....	39
3.5.2	Orthogonal (QR) Decomposition.....	39
3.5.3	Constrained Weighted Least Squares (CWLS) .....	40
3.5.4	Hachtel's Augmented Matrix .....	41
<b>3.6</b>	<b>Novel Weighted Error Modulus (WEM) Method</b> .....	<b>43</b>
<b>3.7</b>	<b>Concluding Remarks and Discussion</b> .....	<b>46</b>
<b>CHAPTER 4</b>		
<b>SELECTION AND APPLICATION OF HACHTEL'S AUGMENTED MATRIX METHOD ...</b>		
.....		<b>47</b>
<b>4.1</b>	<b>Application of Candidate SE Solution Processes</b> .....	<b>47</b>
4.1.1	Network Descriptions .....	47
4.1.2	Measurement Data .....	48
4.1.3	General Overview of Assessment Processes .....	48
4.1.4	Computation Time .....	49
4.1.5	Assessment Scenarios .....	51
4.1.5.1	Assessment Scenario 1: Ideal Condition.....	53
4.1.5.2	Assessment Scenario 2: Errors in Branch Impedance .....	55
4.1.5.3	Assessment Scenario 3: Injection Sign Errors.....	57
4.1.5.4	Assessment Scenario 4: Combination of Injection Sign and Impedance Value Errors .....	59
4.1.5.5	Assessment Scenario 5: Very Short Branch with Low Impedance .....	61
<b>4.2</b>	<b>Selection of Hachtel's Augmented Matrix Method as Potential DSSE Tool</b> .....	<b>63</b>

<b>4.3</b>	<b>Application of Hachtel's Augmented Matrix Method on Real Data .....</b>	<b>66</b>
4.3.1	Description of Sensor Data.....	67
4.3.2	Sensor Data Processing .....	68
4.3.3	Off-line Field Data Test Applying Hachtel's Method .....	71
4.3.1.1	Feeder: Sencur.....	72
4.3.1.2	Feeder: Cerklje .....	77
4.3.4	Analysis of Test Results .....	82
<b>4.4</b>	<b>Concluding Remarks and Discussion .....</b>	<b>83</b>
<b>CHAPTER 5</b>		
<b>DEVELOPMENT OF SCALABLE DISTRIBUTION SYSTEM STATE ESTIMATION.....</b>		
<b>5.1</b>	<b>Choices of Scalable DSSE .....</b>	<b>86</b>
<b>5.2</b>	<b>Differential Evolution Algorithm (DEA) .....</b>	<b>88</b>
5.2.1	Basic Formulation.....	89
5.2.2	Variants of DEA .....	92
5.2.3	Defining Control Variables and Search Space .....	94
5.2.4	Application of DEA in DSSE.....	95
5.2.5	Performance Comparison of Fixed and Randomized $F$ Classical DEA.....	97
5.2.6	Performance Comparison of Three DEA Variants with Randomized $F$ .....	98
5.2.7	Observation of Computation Time and Accuracy with Generation .....	99
5.2.8	Limitation of Application of DEA for Scalable DSSE .....	100
<b>5.3</b>	<b>Overlapping Zone Approach.....</b>	<b>101</b>
5.3.1	Overlapping Zone .....	102
5.3.2	Local and Halo data.....	103
5.3.3	Network Splitting into Zones.....	104
5.3.4	The Complete Algorithm .....	104
<b>5.4</b>	<b>Application of Hachtel's Method in OZA .....</b>	<b>107</b>
5.4.1	OZA: Applying Measurements at GSP Only.....	107
5.4.2	OZA: Applying Measurements at GSP, Halo and Overlapping Nodes .....	119
<b>5.5</b>	<b>Application of DEA in OZA.....</b>	<b>127</b>
<b>5.6</b>	<b>Concluding Remarks and Discussion .....</b>	<b>130</b>
<b>CHAPTER 6</b>		
<b>METER PLACEMENT ALGORITHM .....</b>		
<b>6.1</b>	<b>Meter Placement Algorithm for MV Network .....</b>	<b>135</b>

<b>6.2</b>	<b>State Estimation as Two-Fold Error Reduction Process: WLS Optimization .....</b>	<b>136</b>
6.2.1	Reduction of True Residuals.....	136
6.2.2	Reduction of Measurement Residuals .....	136
<b>6.3</b>	<b>The Meter Placement Algorithm .....</b>	<b>137</b>
6.3.1	Step 1: Short-listing Potential Meter Position.....	137
6.3.2	Step 2: Selection of the Potential Meter Position .....	140
6.3.3	The Complete Algorithm with the Application of Parallel Processing .....	144
<b>6.4</b>	<b>Tests and Results .....</b>	<b>145</b>
6.4.1	Voltage Estimation Errors with Sequential Meter Placement .....	148
6.4.2	Effect of Unexpected Large Error Measurements from Faulty Sensors.....	150
6.4.3	Consistency of Singular Value Based and 100 Monte Carlo Study Based Algorithms.....	151
6.4.4	Application of the Algorithm Outcomes on Other SE Processes.....	153
<b>6.5</b>	<b>Discussion and Concluding Remarks .....</b>	<b>154</b>
<b>CHAPTER 7</b>		
<b>APPLICATION OF PARALLEL PROCESSING .....</b>		<b>157</b>
<b>7.1</b>	<b>Application of Parallel Processing to the DSSE Algorithm.....</b>	<b>157</b>
7.1.1	Principle of Operation .....	157
7.1.2	Performances of DSE on HPC Platform .....	159
<b>7.2</b>	<b>Application of Parallel Processing on Meter Placement Algorithm.....</b>	<b>164</b>
<b>7.3</b>	<b>Concluding Remarks and Discussion .....</b>	<b>167</b>
<b>CHAPTER 8</b>		
<b>CONCLUSION AND FUTURE WORK .....</b>		<b>169</b>
<b>8.1</b>	<b>Overall Discussion and Conclusion .....</b>	<b>169</b>
<b>8.2</b>	<b>Future Work.....</b>	<b>174</b>
<b>REFERENCE.....</b>		<b>176</b>
<b>APPENDIX.....</b>		<b>184</b>
1.	Calculation of Measurement Standard Deviation, $\sigma$ .....	184
2.	UKGDS 77 Node Network .....	185
3.	UKGDS 356 Node Network .....	185
4.	Location of Various Connection Points Discussed for the Assessment of 77 and 356 Node Networks (Chapter 4).....	186
5.	EG Network Feeders: Sencur, Cerklje .....	187
6.	Distributed Generators Connected to EG Feeders .....	189

7.	SCADA and QM Data Format Provided by EG .....	190
8.	Network Parameter Information .....	191
9.	Network Division.....	200
10.	Pseudo-Codes of OZA MATLAB Tool .....	204
11.	Voltage Estimation Errors for 2 Zone Division of the 356 Node Network with Measurements Only at GSP .....	205
12.	Location of Increased Measurements in OZA .....	206
13.	Effect of Expected Maximum Power Flow Measurement Errors in the Added Meter Data.....	208

# LIST OF FIGURES

Fig. 1. 1: DSSE requirements.....	4
Fig. 2. 1: Impact of data with leverage effect .....	11
Fig. 3. 1: DMS and DSSE .....	25
Fig. 3. 2: The equivalent $\pi$ model of transmission/distribution line .....	28
Fig. 3. 3: Transformer approximate equivalent model.....	29
Fig. 3. 4: Normal error distribution and various confidence intervals .....	37
Fig. 3. 5: WEM method flowchart .....	45
Fig. 4. 1: Computation time for 77 node network.....	49
Fig. 4. 2: Computation time for 356 node network.....	50
Fig. 4. 3: Required iteration for 77 node network.....	50
Fig. 4. 4: Required iteration for 356 node network.....	50
Fig. 4. 5:100 Monte Carlo studies for voltage, $ V $ estimation errors on 77 node network (scenario 1).....	53
Fig. 4. 6:100 Monte Carlo studies for phase angle, $ \theta $ estimation errors on 77 node network (scenario 1).....	53
Fig. 4. 7:100 Monte Carlo studies for voltage, $ V $ estimation errors on 356 node network (scenario 1).....	54
Fig. 4. 8:100 Monte Carlo studies for phase angle, $ \theta $ estimation errors on 356 node network (scenario 1).....	54
Fig. 4. 9:100 Monte Carlo studies for voltage, $ V $ estimation errors on 77 node network (scenario 2).....	55
Fig. 4. 10:100 Monte Carlo studies for phase angle, $ \theta $ estimation errors on 77 node network (scenario 2).....	56
Fig. 4. 11:100 Monte Carlo studies for voltage, $ V $ estimation errors on 356 node network (scenario 2).....	56
Fig. 4. 12:100 Monte Carlo studies for phase angle, $ \theta $ estimation errors on 356 node network (scenario 2).....	56
Fig. 4. 13:100 Monte Carlo studies for voltage, $ V $ estimation errors on 77 node network (scenario 3).....	58
Fig. 4. 14:100 Monte Carlo studies for phase angle, $ \theta $ estimation errors on 77 node network (scenario 3).....	58



Fig. 4. 15:100 Monte Carlo studies for voltage, $ V $ estimation errors on 356 node network (scenario 3).....	58
Fig. 4. 16:100 Monte Carlo studies for phase angle, $ \theta $ estimation errors on 356 node network (scenario 3).....	59
Fig. 4. 17:100 Monte Carlo studies for voltage, $ V $ estimation errors on 77 node network (scenario 4).....	60
Fig. 4. 18:100 Monte Carlo studies for phase angle, $ \theta $ estimation errors on 77 node network (scenario 4).....	60
Fig. 4. 19:100 Monte Carlo studies for voltage, $ V $ estimation errors on 356 node network (scenario 4).....	60
Fig. 4. 20:100 Monte Carlo studies for phase angle, $ \theta $ estimation errors on 356 node network (scenario 4).....	61
Fig. 4. 21:100 Monte Carlo studies for voltage, $ V $ estimation errors on 77 node network (scenario 5).....	62
Fig. 4. 22:100 Monte Carlo studies for phase angle, $ \theta $ estimation errors on 77 node network (scenario 5).....	62
Fig. 4. 23:100 Monte Carlo studies for voltage, $ V $ estimation errors on 356 node network (scenario 5).....	62
Fig. 4. 24:100 Monte Carlo studies for phase angle, $ \theta $ estimation errors on 356 node network (scenario 5).....	63
Fig. 4. 25: Voltage and power estimation of Sencur feeder with all SCADA and QM data .....	73
Fig. 4. 26: Voltage and power estimation errors of Sencur feeder with all SCADA and QM data .....	73
Fig. 4. 27: Voltage and power estimation of Sencur feeder with all SCADA and only power QM data.....	74
Fig. 4. 28: Voltage and power estimation errors of Sencur feeder with all SCADA and only power QM data.....	74
Fig. 4. 29: Voltage and power estimation of Sencur feeder assuming real measurement as constraints .....	75
Fig. 4. 30: Voltage and power estimation errors of Sencur feeder assuming real measurement as constraints.....	75
Fig. 4. 31: Voltage estimation ( $V_{est}$ ) and measurements ( $V_{mes}$ ) of Sencur for 10 sets of readings .....	76
Fig. 4. 32: % of deviation from measurement values (Sencur).....	77

Fig. 4. 33 : Voltage and power estimation of Cerklje feeder with all SCADA and QM data .....	78
Fig. 4. 34 : Voltage and power estimation errors of Cerklje feeder with all SCADA and QM data.....	78
Fig. 4. 35: Voltage and power of Cerklje feeder with all SCADA and only power QM data .....	79
Fig. 4. 36: Voltage and power estimation of Cerklje feeder with all SCADA and only power QM data.....	79
Fig. 4. 37: Voltage and power estimation of Cerklje feeder assuming real measurement as constraints .....	80
Fig. 4. 38: Voltage and power estimation errors of Cerklje feeder assuming real measurement as constraints.....	80
Fig. 4. 39: Voltage estimation ( $V_{est}$ ) and measurements ( $V_{mes}$ ) of Cerklje for 10 sets of readings .....	81
Fig. 4. 40: % of deviation from measurement values (Cerklje).....	82
Fig. 5. 1: Computation time with increased network sizes for various SE solutions.....	87
Fig. 5. 2: The DEA functional steps.....	89
Fig. 5. 3: Mutation mechanism in a two dimensional variable space .....	91
Fig. 5. 4: Basic DEA cross-over process .....	92
Fig. 5. 5: Exponential cross-over process .....	94
Fig. 5. 6: Performance of DEA for DSSE applying fixed and randomized mutation factor	97
Fig. 5. 7 : Classical DEA (DE/rand/1/bi) for 100 sets of measurements .....	98
Fig. 5. 8 : Randomized F - DEA ( DE/rand/2/bi) for 100 sets of measurements.....	99
Fig. 5. 9 : Randomized F - DEA ( DE/rand/2/exp) for 100 sets of measurements .....	99
Fig. 5. 10 : Parallel estimation process .....	102
Fig. 5. 11: Overlapping zone based network division.....	103
Fig. 5. 12: Overlapping zone approach algorithm flowchart .....	105
Fig. 5. 13: Studies on the 711 node network with 1 set of measurement taking a) only virtual measurement as constraints and b) both real and virtual power measurements as constraints .....	109
Fig. 5. 14: Studies on the 356 node network with 1 set of measurement taking a) only virtual measurement as constraints and b) both real and virtual power measurements as constraints .....	111
Fig. 5. 15: 2 zone division of 711 node network.....	113
Fig. 5. 16: 3 zone division of 711 node network.....	113

Fig. 5. 17: 4 zone division of 711 node network.....	114
Fig. 5. 18: 2 zone division of 356 node network.....	114
Fig. 5. 19: 3 zone division of 356 node network.....	114
Fig. 5. 20: 4 zone division of 356 node network.....	115
Fig. 5. 21: 2 zone division of 711 node network.....	116
Fig. 5. 22: 3 zone division of 711 node network.....	116
Fig. 5. 23: 4 zone division of 711 node network.....	116
Fig. 5. 24: 2 zone division of 356 node network.....	117
Fig. 5. 25: 3 zone division of 356 node network.....	117
Fig. 5. 26: 4 zone division of 356 node network.....	117
Fig. 5. 27: 2 zone division of 711 node network- a) Mismatches, b) Voltage estimation errors and -c) Phase angle estimation errors .....	121
Fig. 5. 28: 3 zone division of the 711 node network - a) Mismatches, b) Voltage estimation errors and -c) Phase angle estimation errors .....	122
Fig. 5. 29: 4 zone division of the 711 node network - a) Mismatches, b) Voltage estimation errors and -c) Phase angle estimation errors .....	123
Fig. 5. 30: 2 zone division of the 356 node network - a) Mismatches, b) Voltage estimation errors and -c) Phase angle estimation errors .....	124
Fig. 5. 31: 3 zone division of 356 node network- a) Mismatches, b) Voltage estimation errors and -c) Phase angle estimation errors .....	125
Fig. 5. 32: 4 zone division of 356 node network- a) Mismatches, b) Voltage estimation errors and -c) Phase angle estimation errors .....	126
Fig. 5. 33 : Randomized F - DEA (variant DE/rand/2/exp) in scalable DSSE tool .....	128
Fig. 5. 34 : Randomized F - DEA (variant DE/rand/2/bi) in scalable DSSE tool.....	128
Fig. 5. 35 : Voltage estimation errors applying centralized fundamental DEA variant and 5000 generation.....	129
Fig. 5. 36 : Voltage estimation errors applying combined OZA-fundamental DEA variant and 5000 generation .....	129
Fig. 5. 37 : Voltage estimation errors applying combined OZA-fundamental DEA variant and 1500 generation .....	130
Fig. 6. 1: Mapping effect and geometrical interpretation of SVD .....	138
Fig. 6. 2: The Monte Carlo assessment of $M_T$ .....	143
Fig. 6. 3 : Meter placement algorithm flowchart .....	145
Fig. 6. 4: Evaluating improvement ( $Qfac$ ) in PEPs for all $M_p$ to find 1st meter position.	147

Fig. 6. 5: Evaluating improvement ( $Qfac$ ) in PEPs for all $M_p$ to find 2nd meter position .....	147
Fig. 6. 6: Evaluating improvement ( $Qfac$ ) in PEPs for all $M_p$ to find 3rd meter position	148
Fig. 6. 7: Maximum Vdev applying classical WLS .....	149
Fig. 6. 8: Mean Vdev applying classical WLS .....	149
Fig. 6. 9: Effect on maximum Vdev for faulty sensor .....	151
Fig. 6. 10: Effect on mean Vdev for faulty sensor .....	151
Fig. 6. 11: Maximum Vdev applying WLS, QR and Hachtel's methods.....	153
Fig. 6. 12: Mean Vdev applying WLS, QR and Hachtel's methods .....	154
Fig.7. 1: Overlapping zone approach on MPI framework .....	158
Fig.7. 2 : Execution time and overlap mismatch of 356 node network for 500 times zonal interaction applying 2, 3 and 4 zone splitting .....	160
Fig.7. 3: (a) Execution time and (b) Overlap mismatch of 711 node network for 5, 30 and 300 times zonal interaction applying 2, 3 and 4 zone splitting.....	162
Fig.7. 4 : (a) Execution time and (b) Overlap mismatch of 411 node network for 3, 30 and 300 times zonal interaction applying 2, 3 and 4 zone splitting.....	163
Fig.7. 5: Work distribution of MATLAB workers/'labs'.....	165
Fig.7. 6: Application of SPMD .....	166
Fig.7. 7: Execution time applying varying number of 'labs' .....	167
Fig.7. 8: CPU usage applying varying number of 'labs'.....	167

# LIST OF TABLES

Table 4. 1: Average number of iteration for 100 Monte Carlo (scenario 1) .....	54
Table 4. 2: Average number of iteration for 100 Monte Carlo (scenario 2) .....	57
Table 4. 3: Average number of iteration for 100 Monte Carlo (scenario 3) .....	59
Table 4. 4: Average number of iteration for 100 Monte Carlo (scenario 4) .....	61
Table 4. 5: Average number of iteration for 100 Monte Carlo (scenario 5) .....	63
Table 5. 1: Calculation of slopes for various SE solutions .....	87
Table 5. 2: Computation time, voltage magnitude and phase angle estimation errors with various generation size .....	100
Table 5. 3 : Overlap mismatch values for 711 node network .....	118
Table 5. 4 : Overlap mismatch values for 356 node network .....	118
Table 6. 1: Sequential addition of voltage magnitude and power flow meters.....	146
Table 6. 2: Comparison of outcomes from two steps applying independently .....	152

# ACKNOWLEDGEMENT

I am grateful to my supervisor Prof. Gareth Taylor for considering me for HiPerDNO project. As my first supervisor, Prof. Taylor has been remarkably helpful and supportive with his valuable advice, suggestions and direction not only in research but also in many other aspects.

I would like to sincerely thank my second supervisor Prof. Malcolm Irving for his excellent support to guide me all thorough this research work with his valuable experience, expertise and knowledge. His appreciation and encouragement have always motivated me to keep focused, perform extensive research and achieve targets.

I am thankful to my husband and daughter for being my strength that helped me to accomplish my thesis. I must thank my parents and parent-in-laws for their blessings and continuous supports.

Last, but not the least all my colleagues and friends have been kind and helpful in many phases of my research by sharing their experience and advice; I am thankful and lucky for having them around.

This research was supported by 'High Performance Computing Technologies for Smart Distribution Network Operation' (HiPerDNO) project under Grant FP7 - 248135/2007-2013 (European Community's Seventh Framework Programme).

# DECLARATION

The work described in this thesis has not been previously submitted  
for a degree in this or any other university,  
and unless otherwise referenced it is the author's work

# COPYRIGHT

Attention is drawn to the fact that copyright of this thesis rests with the author. This copy of the thesis has been made available on the condition that anyone who refers to it is understood to recognize that the copyright rests with the author and that no information derived from it may be published without the prior written consent of the author. This thesis may also be made available for consultation within the university library and may be photocopied or loaned to other libraries for the purpose of dissemination, consultation and reference.



# ABBREVIATIONS

AMI	Advance Meter Infrastructure
CWLS	Constrained Weighted Least Squares
DEA	Differential Evolution Algorithm
DG	Distributed Generator
DMS	Distribution Management System
DNO	Distribution Network Operator
DSSE	Distribution System State Estimation
EG	Electro Gorenjska
HV	High Voltage
LV	Low Voltage
MPI	Message Passing Interface
MV	Medium Voltage
OZA	Overlapping Zone Approach
PEP	Performance Evaluation Parameter
QM	Quality Meter
SCADA	Supervisory Control and Data Acquisition
SE	State Estimation
SVD	Singular Value Decomposition
UKGDS	United Kingdom Generic Distribution Network
WEM	Weighted Error Modulus
WLS	Weighted Least Squares

# NOMENCLATURE

$B$	Imaginary Part of $Y$
$b_{km}$	Shunt Susceptance between Node $k$ and Node $m$
$c(x)$	Constraint Function
$e$	Error
$G$	Real Part of $Y$
$G(x)$	Gain Matrix as a Function of $x$
$g_{km}$	Shunt Conductance between Node $k$ and Node $m$
$H(x)$	Measurement Jacobian Matrix as a Function of $x$
$h(x)$	Measurement Equation as a Function of $x$
$I_{ij}$	Current Flow from Node $i$ to Node $j$
$J(x)$	Objective Function
$le$	Linearization Errors
$M_p$	Potential Meter Position
$M_C$	Candidate Meter Position
$M_T$	Trial/Test Meter Position
$M_{EX}$	Existing Meter Location
$M_B$	Most Beneficial Meter Position
$N$	Number of Nodes
$P_i$	Real Power Injection at Node $i$
$P_{ij}$	Real Power Flow from Node $i$ to Node $j$
$Q_i$	Reactive Power Injection at Node $i$
$Q_{ij}$	Reactive Power Flow from Node $i$ to Node $j$
$Q_{fac}$	Quality Factor
$r$	Measurement Residual
$r_{km}$	Series Resistance between Node $k$ and Node $m$
$r_t$	True Value Residual
$S_f$	Multiplying Factor
$z$	Measurement Value
$s_{est}$	Estimated Value of the State
$z_{km}$	Series Impedance between Node $k$ and Node $m$
$s_t$	True Value of the State

$V$	Voltage Magnitude
$V_i$	Voltage Magnitude of Node Index $i$
$V_{ij}$	Voltage Drop from Node $i$ to Node $j$
$V_{dev}$	% of Voltage Deviation from the True Value
$W$	Weighting Factor Matrix
$w$	Weighting Factor
$x$	Primary State
$x_{est}$	Estimated Value of Primary State
$x_{km}$	Series Reactance between Node $k$ and Node $m$
$Y$	Admittance Matrix
$y^{sh}$	Shunt Admittance
$y_{kk}^{sh}$	Shunt Admittance at Node $k$
$z$	Measurement
$\theta$	Phase Angle
$\theta_i$	Phase Angle of Node Index $i$
$\theta_{ij}$	Phase Angle Difference between Node $i$ and Node $j$
$\mathcal{L}$	Lagrange Multiplier
$\mu$	Mean/True Value
$\sigma$	Standard Deviation
$\sigma^2$	Measurement Variance
$\sigma_S$	Singular Value of a Matrix

# CHAPTER 1

## INTRODUCTION

### 1.1 Research Background

Modern power systems have been through considerable changes in the past few decades, with raising concerns of limited supply of conventional fuels and increased carbon footprints. The revolution in the power industry is based on three major key visions, which are: extensive utilization of renewable energy sources, efficient planning for energy consumption and implementation of sustainable techniques to reduce the dependency on fossil fuels and to reduce carbon dioxide production. The power transmission domain has already progressed sufficiently towards deploying smart and sustainable operating systems; however, most distribution networks are still being driven via traditional passive operating philosophies. The conventional distribution operating system is unsustainable and necessitates significant upgrades in order to comply with future requirements.

Several innovative schemes, such as: dynamic tariffing, household smart appliances, and smart meters have been introduced accordingly in modern power distribution management systems. Substantial numbers of consumers participating in demand side management will be expected in the near future. The consumer behaviour will impact significantly on the distribution load curve, with the increased use of electric vehicles, smart appliances and smart meters. However, the degree of impact is still subject to uncertainty. Furthermore, there will be significant changes in the conventional distribution voltage profiles due to increased penetration of small scale power resources known as Distributed Generators (DGs), connected directly to distribution feeders. In addition to contributing as local energy providers, DGs will also be participating as commercial suppliers in future distribution systems. As DGs trade their surplus energy, the corresponding distribution domain feeds the power to the main grid. The consequences are: possibilities of voltage rise at DG connection points and intermittent bi-directional power flow. In such circumstances, the conventional descending characteristics of voltage profile along feeders will no longer remain valid.

In essence, both the demand and voltage profiles of the Medium Voltage (MV) and Low Voltage (LV) networks would no longer respond in predictable manners in future distribution systems. The new generation of distribution networks are being called ‘active’, as there will be numerous DGs, active loads and henceforth bidirectional power flow through the circuits. Enhanced awareness of network status in real time is essential for the secure and reliable operation of active power networks and smart Distribution Management System (DMS). Efficient use of sensors and measurements data, backed up by appropriate communication architecture, are prerequisites for the required observability of the network. The development towards more sustainable and smart distribution systems will improve and maintain quality of service, reduce costs and increase the capacity of the grid to host DGs.

A State Estimation (SE) tool is the key enabler of network monitoring and control for both transmission and distribution levels. The Energy Management System (EMS) requires to be provided with network state information in real time to maintain secure and efficient operation. The network is provided with various measurement devices that scan the system status and send that information over to a control centre at regular intervals. Network operators receive a colossal amount of measured data transmitted through communication links. Despite the network being equipped with measurement devices for monitoring, the data obtained from these devices are not completely reliable as they are prone to metering and transducer errors, analogue to digital conversion errors, communication noise and erroneous measurements [1] [2] [3]. The presence of outliers and perturbed data in measurements, when fed into power management functionalities, can lead to problems for system stability and security. The data needs to go through some refinement processes using SE tools before they are fed into control functions and system management software. The SE tool takes the network measurement information, together with a parameterized network model, filters out noise through an optimization processes and calculates estimated values of network states. These estimated values are an approximation of the actual values of system states. State estimates are fed into control functions, asset management software, demand management tools and some other functionalities. Therefore, the Distribution System SE (DSSE) tool sits at the heart of a smart DMS.

## 1.2 Rationale and Objectives

DSSE is still an emerging area of development due to the fact that it was not a much required element for the operation of traditional distribution networks. Traditional distribution networks are supplied only from the transmission side, passively managed and the real time control problems are mostly resolved at the planning stage. The network equipment (transformers, switchgear, overhead lines, cables, etc.) are specified to operate under predefined conditions for voltage and power control without any requirement for real time network monitoring [4]. However as distribution networks are embracing significant changes necessitating enhanced network monitoring to enable smart grid functionalities, the DSSE is being considered as the core component of active distribution network management. The development of DSSE is therefore considered as a critical area of research and development in present power industry.

Existing estimation tools, suitable for transmission systems, are not directly applicable to distribution networks due to their structural differences. One important difference between transmission and distribution networks is their size. Transmission networks consist of a few hundred buses and distribution networks may have up to tens of thousands of nodes. Data originating from this high volume of nodes greatly increases the problem size for distribution networks. The higher ratio of resistance to inductance of distribution branches prevents the distribution state estimation from applying some reduced computation processes, such as decoupled SE, which are commonly used at transmission level. In such circumstances, the computation time of DSSE could be a major issue for real time state estimation.

It is clear that an SE tool can produce good quality estimated values when provided with sufficient real measurement data. Therefore, adequate instrumentation and communication support are important for reliable operation of the SE tool. Present distribution networks have very limited access to real time measurement data. Many of them provide incomplete measurement information that the active distribution network SE is unable to benefit from. For example, in some nodes there may exist power sensors that do not give signed readings. In the presence of a DG at a node connection point, it is absolutely important to get a signed reading indicating whether power is being consumed or supplied from that point. An unsigned power measurement data cannot be useful for DSSE in an active network. This underlines the requirement for implementation of more sensor devices in the distribution system to enable DSSE to perform accurately. Deployment of the required

number of meters on distribution networks is a pre-requisite for the satisfactory performance of DSSE.

A significant question is: how many meters should be installed to give a cost effective solution? It is not possible, with regard to economical and physical feasibility, to place meters at each node and along each branch for measurements of voltage and power. At the same time, the outcomes of the DSSE tool can be trusted more as more measurements are available to provide adequate system information. In general, the DSSE tool will be using pseudo-measurements mostly, along with several virtual measurements and real measurements that should be as few in number as possible due to financial constraints. The distribution system SE problem and its relation to the meter placement problem is depicted in Fig. 1. 1 as these are the major topics in this thesis.

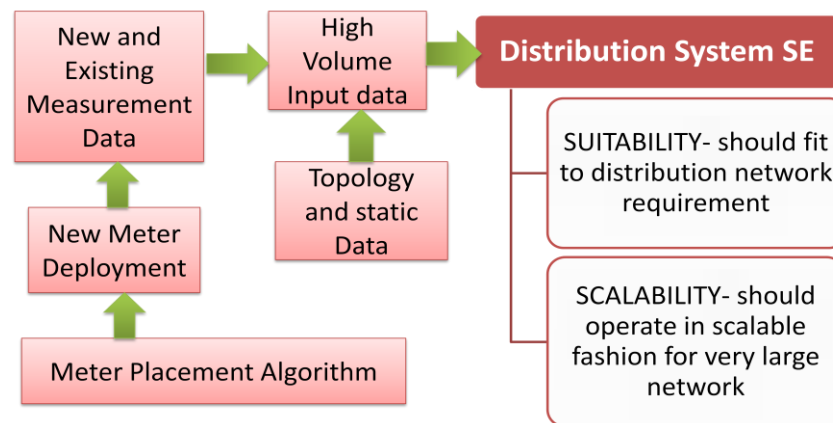


Fig. 1. 1: DSSE requirements

Achieving the desired estimation accuracy, to enable smart robust control over the network, with the requirement of a minimum number of meters is essentially the major challenge in the development of DSSE. The outcomes from the SE tool can be useful only if the required computation time constraints and accuracy thresholds are achieved. Reliable load modeling, additional sensor deployment, robust and efficient optimization techniques, parameter estimation and topology error tracing are a few of the closely related issues. Therefore, development of DSSE is not an isolated problem.

In this context, this thesis will present research on the development of DSSE along with addressing other issues such as the meter placement problem and execution time constraints. Here, various SE solution processes will be evaluated to find a SE tool suitable for distribution systems. The performance of various SE tools will be tested on

model networks and the selected DSSE tool will be applied on real datasets. Taking into consideration the substantial size of distribution networks, distributed DSSE methods for reduced computation time and enhanced scalability will be proposed and developed. As a pre-requisite of DSSE, a novel meter placement algorithm will also be proposed.

### **1.3 Contribution of the Research**

The main purpose of the thesis is to develop state estimation tools and meter placement algorithms suitable for MV distribution systems. The values of the research are explained below:

- As there has not been significant work on SE problem focusing strongly on typical distribution network infrastructure, this work takes into consideration the behaviour and challenges of distribution networks. In this thesis, typical characteristics that make distribution networks different from transmission networks are addressed and the impact of those on the quality of estimated data are assessed. In addition to classical Weighted Least Squares (WLS) solution processes, alternative WLS solution processes and novel Weighted Error Modulus (WEM) optimization techniques are also assessed as candidate DSSE tools. The assessments are performed on UKGDS 77 nodes and 356 nodes model networks.
- The effect of taking real measurements as constraints (in addition to virtual measurements) on estimation quality, when there are limited real measurements, is analysed in this thesis. Case studies are performed on real data sets provided by Elektro Gorenjska (EG); the Distribution Network Operator (DNO) in Slovenia. A DSSE that takes both real and virtual power measurements as constraints is believed to be a novel approach, which is proposed and studied in this thesis.
- There are several research studies on distributed state estimation methods that split a large network into smaller sized zones. The network splitting is executed based on tie-line connected areas or disjoint zones in most cases. In this research, an overlapping zone approach is proposed, in which the network division process is not dependent on the location or even the existence of tie-lines. The overlapping zone tool is applied on a distributed computer system in order to assess the scalability.



- Much of the existing research on distribution network meter placement algorithms does not address the high uncertainty of pseudo-measurements and the economical infeasibility to achieve full observability of network by deploying meters. This thesis proposes a novel meter placement algorithm taking into consideration these issues.
- The performance of the proposed meter placement algorithm is verified in different WLS-based solution processes to ensure its quality of performance. This is specifically important for power distribution systems, as currently DNOs are using various solution processes and it is not yet known which solution process would be accepted widely by DNOs.

## 1.4 List of Publications

A number of conference papers are accepted and presented. Two papers are submitted for consideration for journal publication. The details are listed below:

### Conference Papers (Presented and accepted)

1. Nusrat, N.; Irving, M.R.; Taylor, G.A., "Choice of state estimation solution process for medium voltage distribution systems," *IEEE Power and Energy Society General Meeting 2014 (IEEE PES GM 2014)*, pp.1-5, National Harbor, MD, US, 27-31 July 2014

The paper discusses and compares three optimization techniques as potential DSSE solutions. The proposed methods are: weighted least squares, constrained weighted least squares and weighted error modulus methods.

2. Nusrat, N.; Irving, M.; Taylor, G., "Novel meter placement algorithm for enhanced accuracy of distribution system state estimation," *IEEE Power and Energy Society General Meeting 2012 (IEEE PES GM 2012)*, pp.1-8, San Diego, CA, US, 22-26 July 2012

The paper presents the development of novel meter placement algorithms, based on singular value decomposition and Monte Carlo evaluation, suitable for distribution systems.

3. Nusrat, N.; Irving, M.; Taylor, G., "Stochastic meter placement algorithm for active distribution networks suitable for parallel processing," *47th International Universities*

*Power Engineering Conference 2012 (UPEC 2012)*, pp.1-6, London, UK, 4-7 Sept. 2012

The paper proposes a Monte Carlo evaluation based distribution system meter placement algorithm that is suitable for parallel processing.

4. Lopatka, P.; Salvini, S.; Nusrat, N.; De-Alvaro, L.; Wallom, D., "Performance study of distributed state estimation algorithms on the HiPerDNO HPC platform," *47th International Universities Power Engineering Conference 2012 (UPEC 2012)*, pp.1-6, London, UK, 4-7 Sept. 2012

The paper presents collaborative work on the application of DSSE algorithms developed by Brunel University and the French DNO (EDF R&D) on a high performance computation (HPC) platform provided and managed by Oxford University.

5. Nusrat, N.; Irving, M.; Taylor, G., "Development of distributed state estimation methods to enable smart distribution management systems," *IEEE International Symposium on Industrial Electronics 2011 (ISIE 2011)*, pp.1691-1696, Gdansk, Poland, 27-30 June 2011

Application of differential evolution algorithm as a distributed state estimation tool is investigated and presented in this paper.

6. Nusrat, N.; Irving, M.; Taylor, G., "Development of novel state estimation algorithms for active distribution networks," *46th International Universities Power Engineering Conference 2011 (UPEC 2011)*, pp.1-6, Soest, Germany, 5-8 Sept. 2011

A distributed stochastic state estimation method applying differential evolution algorithm is proposed in this paper.

7. Taylor, G.A.; Irving, M.R.; Nusrat, N.; Liao, R.; Panchadcharam, S., "Smart distribution network operation: Emerging techniques and standards," *IEEE Power and Energy Society General Meeting 2011(IEEE PES GM 2011)*, pp.1-6, Detroit, MI, USA, 24-29 July 2011

This paper focuses on the overall HiPerDNO project. The initial proposal for a distribution system SE tool is included in this work.

### **Journal Paper (Submitted)**

1. Nusrat, N.; Lopatka, P.; Irving, M.R.; Taylor, G.A.; Salvini, S.; Wallom, D.C.H., "An overlapping zone-based state estimation method for distribution systems," *IEEE Transactions on Smart Grid*, [available online: doi10.1109/TSG.2015.2393361]

The paper proposes further development and application of the network split based SE tool, initially proposed and presented in earlier conference papers [82,81].

2. N. Nusrat, M.R. Irving , G.A. Taylor, "A two-step meter placement algorithm for power distribution networks" (submitted to *IEEE Transactions of Measurement and Instrumentation* journal)

This paper presents details of the development and application of a novel two-step meter placement algorithm suitable for active distribution networks with a focus on reducing voltage estimation errors.

## **1.5 Thesis Outline**

This thesis is organized in eight chapters, as described as below:

**Chapter 1 Introduction** has introduced the overall background and explained the reasoning and aims of this research. It has highlighted the values of the work presented in the thesis. The list of published and submitted research work has been included in this chapter. Finally the structure and organisation of the thesis is discussed.

**Chapter 2 Literature review** discusses existing research work on the topics covered by this thesis. This chapter is divided into two main sections to review previous research on power system state estimation tools and meter placement algorithms.

**Chapter 3 Power System State Estimation** provides details of the power system state estimation model, relevant equations, formulae and definitions. Five methods have been discussed as candidate DSSE tools.

**Chapter 4 Selection and Application of Hachtel's Augmented Matrix Method** selects a potential DSSE tool based on experimental investigations on test case studies. Further application of the chosen DSSE tool on real data is described.

**Chapter 5 Development of Scalable Distribution System State Estimation** explains the motivation of this research topic and propose two methods suitable for distributed application of DSSE. Tests are performed on UKGDS test networks.

**Chapter 6 Meter Placement Algorithm** includes description of proposed algorithm with tests and results.

**Chapter 7 Application of Parallel Processing** describes how distributed DSSE is run on cluster computers. The application of parallel processing for a meter placement algorithm on a multi-core machine is also illustrated.

**Chapter 8 Conclusion And Future Work** ends the thesis with concluding remarks and proposals for future work.

# **CHAPTER 2**

## **LITERATURE REVIEW**

This chapter discusses about the existing research work on state estimation and meter placement algorithms. The challenges and limitations for the development of distribution system SE and meter placement methods are also discussed.

### **2.1. General Overview of Power Distribution System State Estimation**

The challenges of the development and existing work on DSSE are discussed in this section. Additionally, detailed discussion on previous research with regard to distributed and scalable state estimation are also included here.

#### **2.1.1 Challenges of the Development of Distribution SE Tools**

The majority of MV-LV networks are inadequately monitored and supported by weak communication facilities at present. However, an ideal measurement-communication infrastructure is not practical in the distribution domain due to the substantial size of the system and the associated costs. Severe limitations of the availability of real time measurements would be one of the key challenges while developing the DSSE tool. As a matter of fact, a power system state estimator is capable of producing acceptable outcomes when there is high redundancy of real measurements. Whereas, in a distribution system, not only is the redundancy very low, but also the various data origins such as measurements through field sensors and load estimation techniques lead to further challenges for DMS. Albeit, distribution network operators (DNOs) are considering deployment of new meters to enable smart DMS, the extensive number of nodes makes it economically infeasible to place real time measurement devices at all required connection points. The instrumentation of distribution networks will consist of a limited number of real measurements, accompanied by some virtual- and numerous pseudo-measurements to achieve the required observability. The degree of trust in the available information varies,

depending on the type of the data and the accuracy of the instruments, and is highly diverse.

These types of measurement characteristics can make the estimator more vulnerable to bad data, which includes outliers and leverage points. The effect of these bad data on estimation quality can be severely adverse; as depicted in Fig. 2. 1 for a set of measurement points with one incoherent datum. Supposedly, the system contains five measurements with the fifth acting as an outlier, having a strong negative leverage effect on the estimate. The solid line is the estimated value affected by the leverage point and the broken line indicates the expected estimate (in the case where no leverage effect exists) that reflects the true values. In Fig. 2. 1, the estimator tends to choose the solid line over the broken line as the sum of squared residuals is lower.

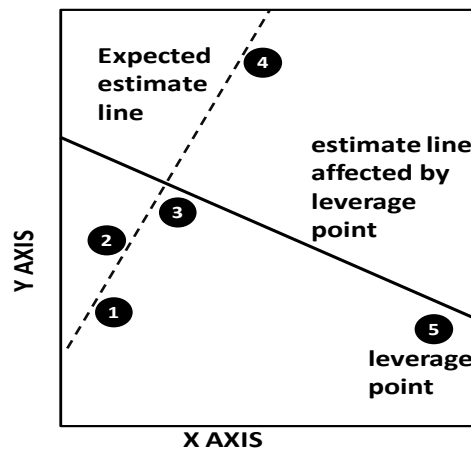


Fig. 2. 1: Impact of data with leverage effect

The outliers like this are normally detected in a pre-estimation process on the basis of incoherent residual and variance values [1] and removed from input data sets prior to feeding the estimator. With regard to power distribution system, detection of bad data points on the basis of variances and residual values is challenging, as existing pseudo-measurements already carry high error variances. Furthermore, it might not be very feasible to remove a measurement datum when the redundancy is already low and only a few real measurements are available to the estimator.

Furthermore, the measurement types and distribution network infrastructure itself can lead to ill-conditioning of Hessian and gain matrices in the normal equations based optimization problem. Due to high resistance to inductance ratios, the simplified algorithms adopting the

decoupling principle are not applicable at distribution system. Accordingly, the traditional estimation techniques that have been used at transmission level cannot be utilized effectively for distribution systems [5].

Another major issue in the development of DSSE is to complete the computation in limited time irrespective of the size of the network. The algorithms and procedures are required to be highly scalable and computationally inexpensive to achieve the required near to real-time state estimation. Decomposition of the calculation processes holds great potential to confront these challenges for the development of a scalable DSSE tool. The emerging technology of high performance parallel computation technology and high speed communication infrastructure can bring benefits to the application of DSSE tool in a distributed manner. Invention of efficient approaches towards the development of scalable DSSE, executable on advanced computation technology is critically challenging and a novel area of research.

In essence, the main challenges in power distribution system SE are:

- Producing high quality estimated data with the availability of a lesser number of real measurements and a much greater number of pseudo measurements
- Execution of the estimation process within a very limited time frame even for larger networks.

### **2.1.2 Existing Research on Distribution SE**

The power system state estimation problem was introduced by Schweppe [6] [7] [8] for transmission systems. However, research and development on SE for distribution systems is an emerging area. Most previous research considered distribution networks as unidirectional power flow passive networks.

Ghosh et al. applied a probabilistic approach for the distribution circuit state estimation based on forward and backward sweeps [9]. The algorithm takes into account the limitation of real measurements in distribution systems. The estimator performs effectively on a small system. Hoffman in [10] applied a similar load flow based estimation technique that is referred to as a ladder algorithm. The algorithm converts all measurements into current measurements. The current magnitude is considered as the primary state and the algorithm aims to match the feeder current magnitude measurements. Application of the proposed

method could be infeasible as a consequence of the following two assumptions that have been made in the presented research; 1) a number of the real measurements have zero variance, and 2) the magnitudes of the real power flow measurements on a radial feeder should always be monotonically decreasing the further it is located from the feeder source. It is important to note that both assumptions are not practical for active distribution systems. The current measurements data are extensively studied by Baran and Kelley [11]. These authors adopt a Weighted Least Squares (WLS) approach to develop a 3-phase DSSE tool. The DSSE tool considers the availability of only a few real measurements and large numbers of pseudo-measurements. The authors observe through a few case studies that power flow measurements are more effective in bad data identification than current measurements. The case studies demonstrate an important fact with regard to the power distribution system SE problem, which is, the improvement of the estimation quality largely depends on the accuracy of pseudo-measurements in presence of limited real measurements [11]. The same authors later have developed a branch-current based 3-phase DSSE tool in order to achieve more computational efficiencies and less sensitivity to line parameters than the conventional node voltage based tools [12]. Similar to reference [10], the authors convert all real power flow and pseudo load measurements into current measurements [12]. References [11][12] have successfully applied their proposed algorithms to obtain good quality of estimated values, however without considering the presence of distributed generators that may cause bidirectional power flows. Reference [13] also applies a 3-phase estimator that uses a current magnitude based formulation. Here the estimation problem is solved using WLS optimization criteria. Test cases imply the necessity of real measurement data for greater quality of estimated values. The proposed methodologies are successfully applied with limited real measurement data. A revised version of a branch current based estimation tool is developed by Wang and Schulz using current magnitude and phase angle as the primary states [14]. The algorithm undergoes additional computation to define the initial states but also decouples the three phases to improve computational efficiency. Significant improvement in estimation quality is observed by reducing real measurement errors from 5% to 3% and pseudo-measurement errors from 50% to 30% [14]. Reference [10]-[14] have applied branch-current magnitude as the key measurement element as well as the state variable to estimate. The outcomes are impressively correct in the presented case studies. However, there is high possibility of bidirectional power flow in future active networks; it will be then highly important to consider the direction of current flow along with its magnitude. The state estimation



algorithms based on current magnitudes may not work very effectively for future distribution systems.

New generation DMS issues like the impact of DG penetration, ill-conditioning problems resulting from normal equation based optimization, heavy computational burden arising from large distribution networks and the impact of smart grids have been addressed in some relatively recent works. Due to the presence of diverse confidence levels of multi-source measurements and various branch sizes, the normal-equation based state estimator is prone to matrix ill-conditioning. Many papers have considered the virtual measurements as equality constraints, which reduces this problem to some extent [15] [16] [17] [18] [19].

Xu et al. [20] developed a WLS optimization problem where the weight of the measurements is termed as 'quality tag'. All measurement data undergo novel bad data detection and processing and the quality tag for the measurement is calculated before they are fed into the WLS optimization tool. The field application shows promising results; however this method does not directly consider the impact of DGs at LV-MV levels [20]. Bignucolo et al. [18] develop a probabilistic voltage state estimation taking into consideration high penetration of DGs. This research proposes discrete step communication support to track down section-wise DG outputs and a load estimation method to reduce pseudo-measurement errors. Incorporation of partial knowledge of the DG in real time has significantly reduced the uncertainty of voltage magnitudes; while indicating suitable communication techniques to adopt and associated costs. The proposed methodology demonstrates its potential for practical applications [18]. Sing et al. [17] investigate compatibility of three different mathematical optimization algorithms (WLS, Weighted Least Absolute Values (WLAV) and Schweppe Huber Generalized M (SHGM) estimators) for DSSE with UK generic distribution networks in presence of DGs. Crucial studies are performed with various levels of measurement error probabilities and redundancy reflecting distribution system scenarios. It is concluded that the classical WLS method performs best when measurement errors are assumed to be Gaussian. The paper also indicates that the most suitable DSSE technique may be different if the measurement error distribution is modelled as other criteria than Gaussian distribution [17]. The authors in [21] have applied generalized three phase state estimation where three phase errors are considered to be correlated. The most significant contributions of the paper are performing three phase unbalanced DSSE and online test on a real distribution network. The algorithm succeeds to correctly identify the areas of the network which were in breach of regulatory

limits for voltage unbalance, however the paper does not explicitly discuss about the overall improvement of estimation quality applying the proposed method [21]. A two stage distribution substation SE solution to reduce computation burden is proposed in [22] for future smart grids. The paper represents one of few researches regarding network division based DSSE algorithm. The DSSE algorithm improves the convergence property over conventional method. Complete observability of the network is required to achieve by real measurements in this method, making its application to real distribution network limited.

In addition to conventional methods, new and extended concepts are being introduced into the distribution system SE problems. Heuristic methods like particle swarm optimization (PSO) is receiving attention to be applied as DSSE. In [23] [24] [25] [26], different types of PSOs are considered as DSSE solutions. Reference [23] [24] [25] apply a hybrid PSO (HPSO) base SE solution taking into account existence of DGs in the network and limited real measurements; with assumptions of contracted load and estimated power factor values at each node being known. The authors consider that there are impacts on the SE objective function due to the nonlinear characteristics of equipment. Such as the nonlinear characteristics of outputs of static VAR compensators (SVCs), induction generators (output equation expressed by constant impedance, constant current and constant power load) and the discrete tap control function of a transformer. Some improvement with respect to the estimation quality is observed when applying HPSO compare to the application of conventional PSO on a model network. The quality of estimation is compared with respect to the real measurement data; although the estimated values are usually compared with the true values to measure their accuracy for such model network. Chilard et al. [26] has adopted a generalized PSO approach as a meta-heuristic solution of DSSE and compare the performance with a constrained WLS approach. The paper discusses some important issues such as the accuracy of estimation and computation time with respect to the conventional normal equation based approach. The authors conclude that the constrained WLS approach outperforms generalized PSO to solve distribution SE. Formulation of the probabilistic estimation as a multi-objective combinatorial optimization problem is proposed by Hashimoto et al. [27]. The authors treat the objective functions by corresponding to an evaluation of occurrence probability and a proximity evaluation of calculated voltage parameters with values obtained by measurements. The meta-heuristic approach requires to set the occurrence probability to each interval, phase, active and reactive power. It is inferred that the proposed methodology may not need additional investments on

measurement equipment as the method is designed to consider different supervisory levels of systems developed in the utilities.

Most of the research into distribution state estimation has applied either power flow based algorithms or WLS minimization criteria. It is evident that most research in the past for distribution system estimation is strongly focused on passive distribution networks that considers unidirectional power flow. In recent decades, the development of smart grids is getting more attention and the profound requirements of the DSSE tool have been realized. Wide scale research, taking into account bidirectional load flow and the integration of DG, are being performed. Distribution networks that have been absolutely passive by nature in the past, are now required to undergo extensive infrastructural development to accommodate current and future requirements. Enhanced observability and robust automation control will be major functions of these active distribution networks enabling smart distribution management.

### **2.1.3 Existing Research on Scalable and Distributed SE**

Different approaches have been proposed for distributed state estimation, where most of the estimation problems are solved by Weighted Least Square (WLS) optimization and the system is assumed to have redundant measurements. Cutsem et al. in [28] [29] [30] have introduced a two level SE technique for large interconnected systems. The large network is sectioned by non-overlapping zones connected by tie lines. The zones perform local estimation with respect to their own slack buses in the first step which generates useful estimates for network operation. The second step involves estimation of tie-line power flow which may only be required for financial purposes rather than operational input. It is stated that the bad data processing is more convenient in each subsystem, however that becomes difficult to perform in the tie-line connected region. Therefore, it is necessary to connect more reliable measurements at tie-line connecting areas [28] [29]. The redundancy of the measurements are assumed to be more than two in the presented case studies in [30], which is quite a high number of measurements with respect to the distribution system measurement availability. Two stage hierarchical SE processes have been developed by many researchers in different ways as presented in [31] [32] [33], most of which are applicable to tie-line connected networks. Reference [31] adopts a multi-area state estimation where the zones are overlapping. The authors propose deployment of GPS provided phase measurements for estimation accuracy and bad data detection at tie-line

areas. The phasor measurements can contribute significantly to improve the quality of estimation and to provide better synchronization among subsystems, however their high cost can be an issue. The SE algorithm in [32] is applicable to large systems and is based on network geographical decomposition. The network division based on geographical differences can be effectively applied for tie-line connected transmission systems, however the application may not be completely feasible for the distribution network. The SE method in [33] does not depend on tie-line connected subsystems, two steps are executed by bi-factorizing the WLS SE problem into both a linear and a nonlinear component. The proposed method holds good potential to reduce computation time and relieve transmission bandwidth provided local computing power is available.

Mohagheghi et al. introduces a statistical estimation process termed ‘super calibrator concept’ that fits into the GPS synchronized measurements and takes advantage of the characteristic of Phasor Measurement Units (PMU). The subsystem can consist of one or more adjacent substations with their circuits, provided there will be at least one PMU or GPS-synchronized relay in each subsystem. In this case, a distributed estimator operates in a reduced system having very high redundancy of measurements, therefore fault data analysis is performed more effectively than centralized SE [34]. The algorithm provides flexibility for network splitting, however, the high cost of PMUs remains as an issue for such method. The border information of the overlapping zone is considered as equality constraints in a Lagrangian formulation in [35] [36] [37]. Reference [35] [36] propose zonal division of the area, where the number of overlapping buses are kept minimal in order to avoid communication bottlenecks. In their models, any injection in the overlapping area is replaced by a fictitious node connected through a zero impedance line and the measurement equations are divided into real and reactive parts based on the decoupling principle. The models benefit from applying asynchronous distributed approach, where the parallel processing can continue without waiting in absence of information from overlapping areas. The accuracy of the estimation and reduction in computation time are achieved at desired level requiring evenly distributed local redundancy ratio over the network. An asynchronous distributed state estimation has been proposed by Wu and Neyer using the load flow formulation exploiting the radial nature of distribution networks [37]. The estimation is based on the assumption that all power flow and injection measurements are available. Such availability of real measurements is more unlikely to occur in the real distribution system scenarios.

A parallel method for static state estimation suitable for MIMD (Multiple Instructions, Multiple Data) processing utilizing the gain matrix and its associated Factorization Graph Path (FPG) is proposed by Abur and Tapadiy [38]. The proposed algorithm has the advantages of low level interaction with other parallel estimations and simple scheduling of processors, however the computation efficiency has not achieved the desired level. Shahidehpour and Wang have described a more generalized parallel state estimation approach partitioning the measurement sets into subsystems equivalent to the number of available processors, so that each processor can be dedicated for one subsystem calculation [3]. Transmission of border node information that is electrically connected with internal nodes of other subsystems, is required in each Gauss-Newton iteration. The processor needs to wait for boundary information to be received before starting the next iteration. Well balanced computation tasks is crucial to assign to processors in this method, otherwise the communication overhead of border node information exchange would delay the computation progress. The Multi Agent System (MAS) based WLS algorithm adopting fast state estimation for large power system has been proposed by Nguyen et al [39]. Case studies have been performed where real time flow measurements in each zone and pseudo-measurement at each boundary node is assumed available. Due to completely decentralized nature, the approach shows promising improvement in scalability, while maintaining the quality of estimation at the desired level. However the presented case studies are performed for highly meshed networks with relatively high number of real measurements, whereas the distribution networks are mostly radial in structure having only limited real measurements available.

Nordman and Lehtonen introduce a novel agent based distribution state estimation involving both primary and secondary substations [40]. The secondary substation computer performs the core state estimation as well as topology, observability and bad data analysis based on a token that periodically traverse the secondary substations computer network, thus enabling interaction with neighbors. The primary substation is responsible for controlling the token handling and providing actions when any abnormal state arises. The proposed algorithm is also capable to analyze the topology, observability and existence of bad data. It is suggested to deploy a sensor at each node even for the small test network, which indicate the requirement of significant number of sensor deployment and associated cost for larger networks. Another multi-agent based SE approach which is suitable for DG connected distribution networks is proposed in [41]. The algorithm decomposes a centralized WLS optimization problem into smaller decentralized objective functions. Each

decentralized section is termed as a cell where cells perform local estimation. Coordination between cells occurs prior to the local estimation process. The algorithm tends to match the estimation of the common nodes between two cells in a WLS minimization problem. The proposed approach has the potential to be applicable to parallel processing, therefore a reduced computation time is expected. In addition to provide accurate estimation on the steady state, the algorithm is also capable to adapt itself to topology changes. The estimator is particularly suitable for a meshed configuration, which does not commonly present the typical distribution network structure.

As discussed above, the history of the distributed SE problem at transmission level is reasonably mature. The two stage hierarchical SE solution tools operate on tie-line based network decomposition and perform coordination of estimates in a second stage after completion of local estimation in the first stage. The decomposition and decoupling of the SE model itself is also proposed by some researchers. Development in distributed DSSE has occurred since the end of twentieth century introducing novel and innovative ideas. These algorithms require additional meter placement to achieve expected accuracy levels from DSSE.

## **2.2. General Overview of Meter Placement Algorithms**

In this section, the challenging issues for the development of meter placement algorithms are widely discussed. Existing research into distribution system meter placement algorithms is also reviewed.

### **2.2.1 Motivation and Challenges**

The quality of outcomes of a state estimation tool depends significantly on the accuracy of information provided to it. The state estimator performs by reducing the difference between the calculated and the measured quantities. In the presence of redundant measurements, the estimator tends to maximize the number of matches with the available measurements that provide information in coherence with the system model. It can effectively filter out the effects of non-coherent measurements which are more likely to carry deceptive information. As a consequence, most of the estimators can perform perfectly at transmission level, where there exists good redundancy of real measurement quantities. However, this is not the case in the distribution system. The substantial amount

of investment that will be required to achieve full observability of the distribution network, by placing meters at almost every node and line, is not practical.

Distribution networks can be equipped with a smaller number of real measurements than actually required for full observability, in order to remain within economical constraints. Further measurement information will be provided by pseudo and virtual measurements, in order to achieve observability. The resulting variable confidence level of measurements, along with the lack of sufficient knowledge of the networks, increases the possibility that the estimate may be biased by somewhat higher weighted measurements at the cost of compromising other (good) measurements. Thus the presence of gross errors, or a leverage point, may deceive the estimator towards infeasible solutions. As a matter of fact, placing a measurement at any point on the distribution network cannot ensure consistent improvement of the estimated information when real measurements are limited in number. A sensitive node can bring about considerable improvement of estimation quality if it produces highly accurate measurement data. On the other hand, it may result in an unacceptable quality of estimation, when generating some erroneous metered data. The placement of real measurements within distribution networks must be carefully determined.

### **2.2.2 Existing Research**

Research to develop both economically and computationally cost effective ways for placing measurement sensors over networks has always been subject to potential improvements for both transmission and distribution networks. The sensor positioning in active distribution network is a comparatively recent and emerging area of research. Most meter placement research solves the mathematical unobservability problems in the presence of redundancy of real measurements. However, the scenario is different at distribution level, which is predominantly radial in nature; every node is already assigned with a pseudo-measurement and the aim will be to replace some of these pseudo-measurements with real measurements.

The challenging issues and requirements of active distribution system meter placement have been well explained in [42] [43]. Shafiu et al. in [42] proposed a heuristic approach to identify the points on the network with higher voltage variations, through a series of load flows, as potential voltage sensor positions. The algorithm starts by assuming a potential

measuring position based on its topological location and loading quantity and then moves the position towards the node which is most prone to voltage deviation. This algorithm gives emphasis on improving network voltage magnitude estimation and places voltage sensors only. Test cases demonstrate significant improvement in voltage estimation quality, however the proposed algorithm does not ensure the optimum solution. It rather provides a methodological solution which can be considered to be better than relying on only engineering judgments. A meter placement algorithm for radial distribution networks with distributed generators is also developed using a stochastic approach in [44] [45]. In [44], the authors introduced a meter placement algorithm using a multi-stage decision making process known as dynamic programming. The process divides the problem into a number of decision stages that is equivalent to the number of sensors required in addition to the base case. Each component of a decision stage is linked with the components of the previous stage, to find the best choice at that stage. It is also a sequential meter placement process that continues to place meters until the desired estimation accuracy level is achieved. The method improves the overall voltage estimation quality. However the case studies are performed on a small network offering inadequate evidence to realize the usefulness of the algorithm. A more cost effective application of the proposed method is published in [45]. The authors propose to put more measurements at the same node where a measurement already exists and there is scope to measure other variables being placed at the same node, thereby reducing the installation cost and some other fixed costs. The dynamic programming is now biased towards those nodes where measurement already exists, by putting a premium value in the calculation of the objective function. Here the branch current flow is taken as the primary state variable, and only current sensors are placed by solving the dynamic programming problem. The improved approach is applied on larger networks and is able to provide desirable voltage estimation quality. However, the quality of estimates is affected by the presence DGs, requiring different measurement configurations for the both approaches in [44] [45]. The dynamic programming criterion is also applied in [85], where the authors take into account network parameter uncertainties and possible metrological degradations of the measurement devices. The meter placement problem is solved in two phases: first, by considering no gross measurement error and second, by considering large error from a measurement. Test cases show that by taking into account the parameter and measurement errors, the required number of measurement and the associated cost increases. However, the proposed method provides a robust meter placement solution with respect to error input data, which may justify the increased cost of measurement placement.



In a recent work, Singh et al. have proposed meter placement algorithms, exploiting the properties of the error covariance matrix, for active distribution networks. The algorithm reduces the size of error ellipsoids related to the voltage and phase angle of a node, in order to reduce the probability of state estimation error exceeding a certain threshold [46]. The process requires Monte Carlo simulations to find, not the optimal, but rather a feasible solution for the placement of meters. A further development of this work has been published in [43]; considering the meter placement as a stochastic optimization problem. They employ an ordinal optimization approach to short-list the number of possible solutions and thereafter perform an exhaustive evaluation process, only on the short-listed positions to find the best positions. In both approaches, the voltage and phase angle estimation errors remain within the expected error threshold values, while the approach in [43] requires reduced computation cost and fewer measurements making it more suitable for practical application. The proposed method in [43] however does not ensure to find the global optimum point, it rather ensures the solution would be at a near optimum point.

Liu et al. have proposed a robust meter placement algorithm incorporating phasor measurements and smart meter data in [86] [87]. The method applies genetic algorithm to search optimal meter placement for the objective function that is developed to minimize both the cost of placement and maximum deviation of estimates from true values. The method is also capable to take into account various network topologies, the loss of a measurement and the degradation of a measurement data. Although PMU and smart meter units can be expensive and may restrict their use at a very limited number, the method seems to provide robust solution for the meter placement for different network topologies, error measurement data and unmeasured DG connection points [86] [87]. Echternacht and Moser propose the deployment of new measurements based on the estimation deviation and present a method for the evaluation of the measurement placement's impact on state estimation [88][89]. The proposed method is suitable for both medium and low voltage networks and is able to reduce the voltage estimation errors significantly. However, the method does not follow any methodological principle to determine how many meters at what positions are required to be assessed. It rather randomly selects one or more measurements to evaluate their impact on the estimation deviation [88][89].

The Monte-Carlo technique has been heavily utilized in most of the algorithms [49] [45] [46][85]. The distribution network state estimation will be strongly influenced by the large

quantity of pseudo-measurements and be affected by any erroneous information in the static data. The Monte-Carlo study takes into account this uncertainty by adding different levels of random noise in each run. Thus it helps to evaluate the possible uncertainty of static and dynamic data through the state estimation process. The Monte Carlo and Genetic Algorithm based [86][87] approaches are usually computationally highly expensive. However that might not be any major concern since the meter placement problems are not a real-time problem, they are rather solved at the planning stage. The critical concerns of branch current based approaches [44] [45] are the availability of signed measurements and initialization of state variables to the unpredictable outputs from DGs. The power flow could be bi-directional in presence of DGs while most of the current measurement devices provide only current magnitude measurements at distribution systems. The application of PMUs and smart meters is being consider as a potential data source for robust measurement placement solutions, however their high cost is the major concern [86][87][89]. Many of the distribution meter placement problems are solved successfully by incremental measurement placement methods [42]-[46][85], along with that some recent work has provided optimum solution applying space exploration based methodologies [43] [86] [87]. It is yet difficult to state the advantages of incremental methods over space search methods and vice versa, since different algorithms apply different networks and choices of measurements.

### **2.3. Concluding Remarks and Discussion**

A general discussion on existing work relevant to the power system state estimation problem is presented in this chapter. Power system SE is an area of extensive research interest, however distribution system SE is, comparatively, a new era of study and development. The development of DSSE is more challenging because of the distribution system measurement and network structure. Existing research proposes classical optimization techniques as well as novel estimation processes as DSSE solutions.

One major challenge is to develop a scalable DSSE tool to fit small to very large networks. A scalable solution of DSSE is important, and this can be resolved by the development of distributed SE. This chapter has included discussion on various distributed SE methods in this context.

Current distribution networks lack sufficient real time measurements. To enable DSSE to achieve quality estimation, enhanced real time sensor deployment is essential. Similar to distribution SE, the existing measurement placement algorithms, which have been applied to transmission networks, cannot be directly used for distribution systems. The typical challenges and the relevant literature on the development of meter placement algorithms, for distribution systems, have been included in this chapter.

## CHAPTER 3

### POWER SYSTEM STATE ESTIMATION

The key elements that are essential to develop the power system SE problem is discussed in this chapter. The SE relevant functions, equations and definitions are also important subjects explained in this chapter. In addition to that, the conventional SE approaches as well as a novel algorithm is reported here in details as possible DSSE solutions.

#### 3.1 Power System State Estimation Building Blocks

This section describes the input-output variables, parameters and functions associated with a state estimator. Fig. 3. 1 shows how DSSE fits in DMS as a part of distribution Supervisory Control And Data Acquisition (SCADA) system [1] [2].

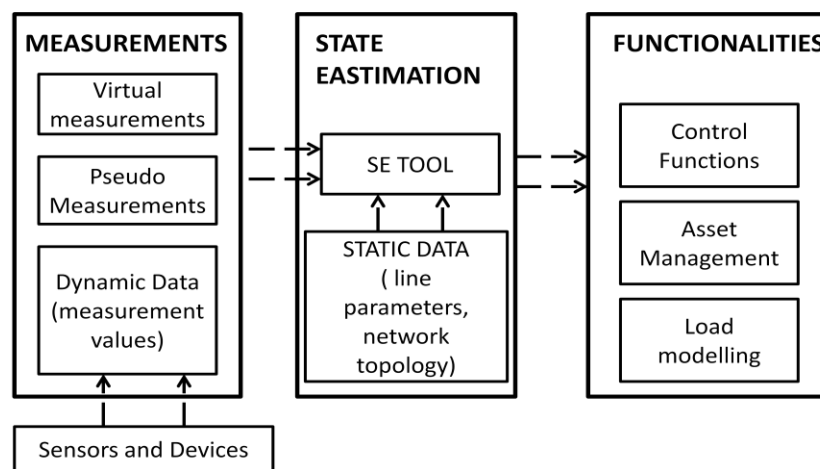


Fig. 3. 1: DMS and DSSE

SCADA is defined as a control system for real time monitoring of the network and setting the control functions required for DMS. The data acquisition is performed by instrumentation and communication facilities into the control room. The collected data from all over the network are passed through SE tools before they are monitored. Corrective and preventive actions are taken based on information obtained from SE tools, by various enabling control functions, e.g. Volt and VAr control, network reconfiguration,

economic dispatch and load control [26]. The data is also preserved for future use, e.g. load modelling and condition monitoring.

As can be seen in Fig. 3. 1, the SE tool requires to be fed with dynamic data transmitted by sensor devices, along with static data that comprises network parameters. The SE tool usually involves an optimization model, which solves non-linear power flow equations with the input data, to filter out the erroneous information. Detailed discussions of the input-output variables, parameters and network components, comprising the SE building blocks, are provided in the following sections.

### 3.1.1 Input and Output Data

*State Variables:* The network state variables are phasor voltages, current flows, real and reactive power injections and flows.

- The voltage magnitude ( $V$ ) and phase angle ( $\theta$ ) of connection points are referred to as 'primary states'. State values are often accompanied by a subscript to denote the associated node index.  $V_i$  and  $\theta_i$  refers to voltage and phase angle of node index  $i$ . And  $\theta_{ij}$   $V_{ij}$  present phase angle and voltage differences between node  $i$  and node  $j$ . All other state variables can be expressed as functions of the primary state values, provided network parameters are known.
- Other state variables (current, power injection and power flow) are termed 'secondary states'. These are dependent on and related to the primary variables by injection and flow balance equations associated with the observed network model. The current, real and reactive power injection and real and relative power flow are generally presented by  $I_{ij}$ ,  $P_i$ ,  $Q_i$ ,  $P_{ij}$ ,  $Q_{ij}$  respectively, where the subscripts present associated node indices. The order of the subscripts indicates the direction of flow.

*Measurement Data:* The measurements are classified into three types: The real, pseudo and the virtual measurements.

- Real measurements are actual telemeter data collected from sensors in real or near real time. The confidence in real measurement is calculated based on the sensor precision.
- The pseudo-measurements are load modelled data which are not real time measurements. As pseudo-measurements are assumption-based data, the degree of trust in these is very low.

- The virtual measurements are also not actually measured, but their values are known with greater confidence. These are zero injection zero load nodes, open circuit breaker flow measurements, etc.

Each measurement is associated with a weighting factor  $W$  that is calculated from the inverse of measurement variance,  $\sigma^2$ .  $\sigma$  is the standard deviation dependent on the precision of the corresponding measurement, which will be explained further in section 3.4.2. As measurement variances are smallest for virtual measurements and largest for pseudo-measurements, the weighting factor is highest for virtual and lowest for pseudo-measurements. The measurement values are also referred to as observed values in the statistical literature.

*Static and Dynamic Data:* The input data to the estimation tools are classified as static and dynamic data.

- Static data normally remains unchanged over time, e.g. network topology, transformer impedance, line parameters (resistance, reactance, admittance and susceptance), location and types of sensors and measurement weights.
- The dynamic data are measurement data that change over time. This includes voltage phasor values, current and power flows, power injections and transformer tap positions.

*True and Estimated Values:* Performances of an estimation tool can be verified by comparing true and estimated values.

- The actual value of the state variable is termed the true value. However, the real time true value is not generally known in practice. The simulated information obtained from load flow analysis for a known condition of the system is considered as the true values. True values are also termed as 'real values' and 'load flow values'.
- Estimated values are the information generated by the SE tool after rigorous data processing. These include both primary and secondary estimated state variables (as well as network parameters in case of generalized state estimation). In most cases, the SE tool executes an optimization process to estimate the primary state variables, and calculates the secondary state values using the most updated  $V$  and  $\theta$  estimates after achieving convergence within some criteria. The estimated values are afterwards provided to DMS functionalities.

### 3.1.2 Network Model and Components

*Nodes and Branches:* In distribution systems, connection points are generally termed as nodes and power distribution lines as branches. One of the nodes is considered as a reference node and is referred to as the 'slack' node. The equivalent  $\pi$  model of a transmission/distribution line is generally considered in power system state estimation problems. A  $\pi$  model of a branch connecting node  $k$  and  $m$  is shown in Fig. 3. 2 [1].

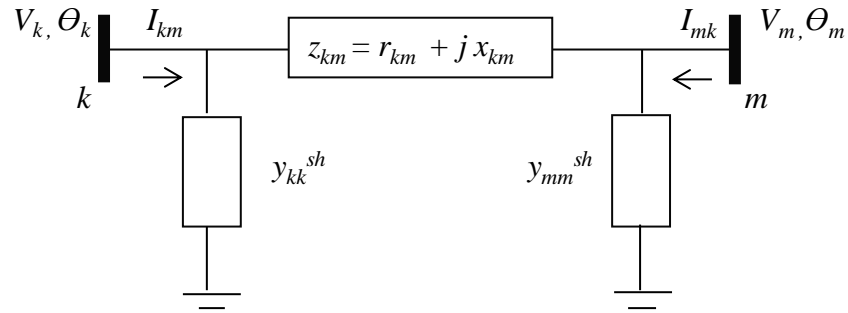


Fig. 3. 2: The equivalent  $\pi$  model of transmission/distribution line

The main components of the  $\pi$  model are: series impedance  $z_{km}$ , shunt admittance  $y_{kk}^{sh}$ ,  $y_{mm}^{sh}$ . The suffices indicate the bus/node index being connected by the branch. When series resistance and reactance are represented by  $r$  and  $x$  respectively, the following expressions are achieved for a branch connecting node  $k$  and  $m$

$$z_{km} = r_{km} + j x_{km}$$

$$y_{km} = z_{km}^{-1} = g_{km} + j b_{km}$$

The shunt conductance,  $g_{km}$  and shunt susceptance  $b_{km}$  are defined as follows

$$g_{km} = \frac{r_{km}}{r_{km}^2 + x_{km}^2}$$

$$b_{km} = - \frac{x_{km}}{r_{km}^2 + x_{km}^2}$$

*Transformers:* A transformer equivalent model is required for modelling network equations. The transformers are also treated as an equivalent branch/line while considering tap position. The equivalent model with turns ratio  $1:a$  and series impedance  $z^t$  is

presented in Fig. 3. 3. Where  $r^t$  represents transformer copper losses and  $x^t$  is leakage reactance [1].

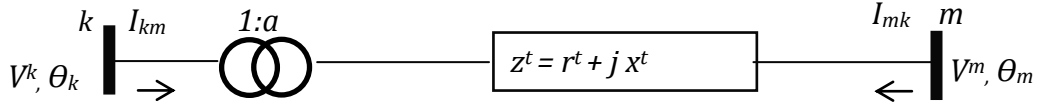


Fig. 3. 3: Transformer approximate equivalent model

*Nodal Equations:* The nodal formulation is obtained incorporating Kirchoff's current law that introduces the admittance matrix,  $Y$ . The real and imaginary parts of  $Y$  are presented by  $G$  and  $B$  therefore  $Y = G + jB$ .  $Y$  is defined as below for a  $N$  node network,

$$\mathbf{I} = \begin{bmatrix} I_1 \\ \vdots \\ I_N \end{bmatrix} = \begin{bmatrix} Y_{11} & \dots & Y_{1N} \\ \vdots & \dots & \vdots \\ Y_{N1} & \dots & Y_{NN} \end{bmatrix} \begin{bmatrix} V_1 \\ \vdots \\ V_N \end{bmatrix} = \mathbf{Y} \mathbf{V}$$

Where

$$Y_{kk} = G_{kk} + jB_{kk} = y_k^{sh} + \sum_{m \neq k}^N (y_{km} + y_{km}^{sh})$$

$$Y_{km} = G_{km} + jB_{km} = -y_{km}$$

It is assumed that  $y_{kk}^{sh} = y_{mm}^{sh} = y_k^{sh}$  and  $y_{km}^{sh} = y_{mk}^{sh}$ . For a tap changing transformer, the components of admittance matrix change as shown below:

$$Y_{kk} = Y_{kk} + \frac{y^t}{a^2}$$

$$Y_{mk} = Y_{km} - \frac{y^t}{a}$$

$$Y_{mk} = Y_{mk} + \frac{y^t}{a}$$

$$Y_{mm} = Y_{mm} + y^t$$

Where  $a$  is the transformation ratio from the receiving end side and  $y^t = 1/z^t$  as shown in Fig. 3. 3. For branches other than transformers, the value of  $a$  is set to one in the admittance matrix.

*Generators and Loads:* Load and power generation are considered as power consumption and injection respectively. Therefore, their presence does not have any effect on the



network model and they are treated in the load balance equations. The sign convention for power injection is  $+v$  and for power consumption is  $-v$ . In all cases, the aggregated power of a node is taken into consideration.

### 3.2 The Non-linear SE Model: Least Squares Problem

The data received at the control room does not necessarily reflect the actual values, as they are prone to low to high levels of communication noises, lack of synchronisation and erroneous sensor readings. If the true value of a system state is  $s_t$  for which the observed value is  $z$  with additive noise  $e$ , then

$$z = s_t + e \quad (3.1)$$

The observed data, or measurement information, is processed through regression analysis to fit equations which represent the network model. The regression model for power systems is non-linear by nature and this can be expressed below

$$z = s_{est} + e = h(x_{est}) + e \quad (3.2)$$

$h(x)$  are non-linear functions of independent state vector  $x$  i.e. primary states ( $[V_1, \theta_1, V_2, \theta_2 \dots V_N, \theta_N]$ ).  $s_{est}$  is the estimated value of the measured states and  $x_{est}$  is the estimated primary state vector. In the ideal case, (3.1) and (3.2) imply that

$$s_t = s_{est} \quad (3.3)$$

i.e. the estimated values should equal the true values, cancelling all noise from measurement information. However, in practice (3.3) cannot be achieved. The estimation tool actually distributes the measurement errors in such a way that some measure of the difference between the true and estimated values of the states is minimized. The measurement residual,  $r$  is defined by the difference between estimated and observed/measured value.

$$r = z - h(x_{est}) \quad (3.4)$$

The true value residuals,  $r_t$  are the differences between the estimated and true values

$$r_t = z - s_t \quad (3.5)$$

it is expected that  $|r|^{max} < |r_t|^{max}$  i.e. the estimator should reduce some measure of the true value residuals.

One of the most commonly applied optimization tools to filter out measurement noise is the Least Squares (LS) estimator. The LS optimizer minimizes the sum of the squares of measurement residuals to reach the optimum solution point. As all the measurements are not trusted equally, each residual is associated with a weighting factor depending on the accuracy of the measurement. LS with associated weighting factors is known as Weighted Least Squares (WLS), as expressed in (3.6). Here  $z_i$  is the measurement of state  $i$ ,  $h_i(x)$  is the corresponding measurement equation and  $M$  is the number of available measurements.

$$\text{Min} \sum_{i=1}^M W_i r_i^2 \quad (3.6)$$

$r_i$  is the residual of the  $i^{th}$  measurement and  $W_i$  is the associated weight [1] [2] [6] [47] [48]. Non-linear regression analysis performs the minimization iteratively to arrive at optimum estimates. It requires starting values to initiate the iterative process [48]. The objective function,  $J$  for a system of  $n$  nodes and  $m$  measurements, where  $m \geq n$  can be expressed as

$$\begin{aligned} J(x) &= \sum_{i=1}^m \left( \frac{z_i - h_i(x)}{\sigma_i} \right)^2 \\ &= [z - h(x)]^T W [z - h(x)] \end{aligned} \quad (3.7)$$

$\sigma_i$  is the standard deviation of the  $i^{th}$  measurement. The first order optimality condition is satisfied at the minimum point by making the gradient of  $J(x)$  equal zero

$$g(x) = \frac{\partial J(z)}{\partial x} = -H^T(x) W [z - h(x)] = 0 \quad (3.8)$$

where

$$\frac{\partial h(x)}{\partial x} = H(x)$$

$H(x)$  is the measurement Jacobian matrix which is a function of primary states  $x$ . By Taylor series expansion of  $g(x)$  around the state vector  $x$ , the gradient function can be approximated as

$$g(x + \Delta x) = g(x) + G(x)\Delta x \quad (3.9)$$

$G(x)$  is the Hessian matrix of  $J(x)$ .  $G(x)$  is also known as Gain matrix which is expressed as:

$$G(x) = \frac{\partial^2 J(x)}{\partial x^2} = H^T(x) W H(x) \quad (3.10)$$

The estimated state is obtained by the following iterative procedure as depicted in the equation below for the  $k^{th}$  iteration

$$H^T(x^k) W H(x^k) \Delta x^k = H^T(x^k) W^{-1} [z - h(x^k)] \quad (3.11)$$

$$x^{k+1} = x^k + \Delta x^k \quad (3.12)$$

Measurement residual,  $[z - h(x^k)]$  is also expressed by  $\Delta h$ . Equation (3.11) is the widely used normal equation for nonlinear regression analysis in this case. The iteration initiates with the start value of the primary state vectors and is carried out until convergence is achieved. The stopping criteria, for the convergence point, is achieved when there are very small changes in value of all the states between two consecutive iterations i.e.  $\Delta x^k \simeq 0$ . At the convergence point,  $x^k = x_{est}$  which is the final estimation output.

### 3.3 Measurement Functions and Measurement Jacobians

Measurement functions,  $h(x)$  are applied to calculate values of primary and secondary state variables for which real to near real time sensor data are provided. The measurement functions can be of various types e.g. power injections, power flows, voltage and phase angle magnitudes and current flows. The measurement functions for secondary state variables are expressed in terms of network static parameters and primary states. The estimator calculates  $(2N-1)$  primary states where  $N$  is the number of nodes of the system as

slack node phase angle is fixed at zero. The three-phase network is assumed to be a balanced system; that gives the advantage to be able to state all mathematical expressions in terms of a single phase system. While expressing a balanced three phase systems in term of a single phase system, the following assumptions are required.

- The mean of three phase values to be considered as single phase voltage and phase angle
- The sum of three phase values to be considered as total real and reactive power injections
- The sum of three phase flows to be considered as total power and current flows
- The mean of three phase drops to be the single phase voltage drops
- Branch impedances are the positive sequence equivalent values of all branches connecting two nodes

The measurement Jacobian,  $H$  is the matrix of first derivatives of measurement equations with respect to primary state variables. The structure of  $H$  is shown below for  $N$  nodes and  $M$  available measurements.

$$H(x) = \begin{vmatrix} \frac{\partial h_1(x)}{\partial \theta_1} & \frac{\partial h_1(x)}{\partial V_1} & \dots\dots & \frac{\partial h_1(x)}{\partial \theta_N} & \frac{\partial h_1(x)}{\partial V_N} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{\partial h_M(x)}{\partial \theta_1} & \frac{\partial h_M(x)}{\partial V_1} & \dots\dots & \frac{\partial h_M(x)}{\partial \theta_N} & \frac{\partial h_M(x)}{\partial V_N} \end{vmatrix} \quad (3.13)$$

The first column of the above matrix is normally excluded as  $\theta_1$  represents the slack node phase angle which is always fixed at zero. The Jacobian matrix is therefore an  $M \times (2N-1)$  matrix. In this thesis, primary state variables are expressed in polar form (voltage values in per unit and phase angles in radians) and the network branch  $\pi$  model is used. The voltage magnitude, power flow and power injection measurement equations are stated as follows. Current flow and phase angle measurement equations are not considered in this work, therefore they are not included here.

*Real and Reactive Power injection equations:*

- Injection measurement equations

$$P_i = V_i \sum_{j=1}^N V_j (G_{ij} \cos \theta_{ij} + B_{ij} \sin \theta_{ij}) \quad (3.14)$$

$$Q_i = V_i \sum_{j=1}^N V_j (G_{ij} \sin \theta_{ij} - B_{ij} \cos \theta_{ij}) \quad (3.15)$$

- Jacobian equation

$$V_i \frac{\partial P_i}{\partial V_i} = \sum_{j=1}^N V_i V_j (G_{ij} \cos \theta_{ij} + B_{ij} \sin \theta_{ij}) + V_i^2 G_{ii} \quad (3.16)$$

$$\frac{\partial P_i}{\partial \theta_i} = \sum_{j=1}^N V_i V_j (-G_{ij} \sin \theta_{ij} + B_{ij} \cos \theta_{ij}) - V_i^2 B_{ii} \quad (3.17)$$

$$V_j \frac{\partial P_i}{\partial V_j} = V_i V_j (G_{ij} \cos \theta_{ij} + B_{ij} \sin \theta_{ij}) \quad (3.18)$$

$$\frac{\partial P_i}{\partial \theta_j} = V_i V_j (G_{ij} \sin \theta_{ij} - B_{ij} \cos \theta_{ij}) \quad (3.19)$$

$$V_i \frac{\partial Q_i}{\partial V_i} = \sum_{j=1}^N V_i V_j (G_{ij} \cos \theta_{ij} - B_{ij} \sin \theta_{ij}) - V_i^2 B_{ii} \quad (3.20)$$

$$\frac{\partial Q_i}{\partial \theta_i} = \sum_{j=1}^N V_i V_j (G_{ij} \sin \theta_{ij} + B_{ij} \cos \theta_{ij}) - V_i^2 G_{ii} \quad (3.21)$$

$$V_j \frac{\partial Q_i}{\partial V_j} = V_i V_j (G_{ij} \sin \theta_{ij} - B_{ij} \cos \theta_{ij}) \quad (3.22)$$

$$\frac{\partial Q_i}{\partial \theta_j} = -V_i V_j (G_{ij} \cos \theta_{ij} + B_{ij} \sin \theta_{ij}) \quad (3.23)$$

*Real and Reactive Power flow equations:*

- Measurement equation

$$P_{ij} = V_i^2 (g_{si} + g_{ij}) - V_i V_j (g_{ij} \cos \theta_{ij} + b_{ij} \sin \theta_{ij}) \quad (3.24)$$

$$Q_{ij} = -V_i^2 (b_{si} + b_{ij}) - V_i V_j (g_{ij} \sin \theta_{ij} - b_{ij} \cos \theta_{ij}) \quad (3.25)$$

- Jacobian equation

$$V_i \frac{\partial P_{ij}}{\partial V_i} = 2V_i^2(g_{si} + g_{ij}) - V_i V_j(g_{ij} \cos \theta_{ij} + b_{ij} \sin \theta_{ij}) \quad (3.26)$$

$$\frac{\partial P_{ij}}{\partial \theta_i} = V_i V_j(g_{ij} \sin \theta_{ij} - b_{ij} \cos \theta_{ij}) \quad (3.27)$$

$$V_j \frac{\partial P_{ij}}{\partial V_j} = -V_i V_j(g_{ij} \cos \theta_{ij} + b_{ij} \sin \theta_{ij}) \quad (3.28)$$

$$\frac{\partial P_{ij}}{\partial \theta_j} = -V_i V_j(g_{ij} \sin \theta_{ij} - b_{ij} \cos \theta_{ij}) \quad (3.29)$$

$$V_i \frac{\partial Q_{ij}}{\partial V_i} = -2V_i^2(b_{si} + b_{ij}) - V_i V_j(g_{ij} \sin \theta_{ij} - b_{ij} \cos \theta_{ij}) \quad (3.30)$$

$$\frac{\partial Q_{ij}}{\partial \theta_i} = -V_i V_j(g_{ij} \cos \theta_{ij} + b_{ij} \sin \theta_{ij}) \quad (3.31)$$

$$V_j \frac{\partial Q_{ij}}{\partial V_j} = -V_i V_j(g_{ij} \sin \theta_{ij} - b_{ij} \cos \theta_{ij}) \quad (3.32)$$

$$\frac{\partial Q_{ij}}{\partial \theta_j} = V_i V_j(g_{ij} \cos \theta_{ij} + b_{ij} \sin \theta_{ij}) \quad (3.33)$$

*Voltage magnitude equation:*

- Jacobian equation

$$V_i \frac{\partial V_i}{\partial V_i} = V_i \quad (3.34)$$

$$\frac{\partial V_i}{\partial \theta_i} = 0 \quad (3.35)$$

$$V_j \frac{\partial V_i}{\partial V_j} = 0 \quad (3.36)$$

$$\frac{\partial V_i}{\partial \theta_j} = 0 \quad (3.37)$$

The common terms in (3.16), (3.18), (3.20), (3.22), (3.26), (3.28), (3.30), (3.32), (3.34), and (3.36) on right and left sides of these equations are not cancelled to express the voltage and the phase angle update equation in a generalized form. This makes the calculation process of state vectors simpler during the Gauss-Newton iteration.

## 3.4 Assumptions and Hypothesis

There are several assumptions that have been made while developing and applying DSSE in various case studies. These assumptions are discussed in detail in the following section.

### 3.4.1 System Observability

As the SE solution is based on a regression analytical model, the system is required to be over-determined or at least determined. This can be achieved by finding a network decision tree or computing the rank of the measurement matrix. The number of available sensor data must be more than or at least equal the number of unknown states of the system representing an over-determined or determined system respectively. For an  $N$  node network the number of unknown state variables is  $2N - 1$  as the slack node phase angle is always fixed at zero. The network is assumed to be capable of supplying  $M$  measurements (real, virtual and pseudo) forming  $M$  independent equations where  $M \geq (2N - 1)$ .

### 3.4.2 Normal Error Distribution

The measurement error is assumed to be an uncorrelated normal distribution function with known variance ( $\sigma^2$ ) and zero average value. The assumption of normal distribution of errors is highly compatible with any least squares based optimization problem. In probability theory, normal distribution function  $f(x)$  is defined as  $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ . Here, mean value  $\mu$  is defined as the expected value for the relevant observed value ( $m$ ), therefore  $\mu = E(m)$ . The parameter  $\sigma$  is the standard deviation of respected observed value, therefore its variance is  $\sigma^2$ . The positive and negative residuals with equal absolute values should clearly not be weighted the same unless this assumption is valid [47]. Fig. 3. 4 illustrates a normal distribution curve and varying confidence intervals for different margin of errors. The significant portion of the measured values are expected around the true values,  $\mu$  as shown. The normal distribution is almost zero as the value  $x$  lies more than a few standard deviation ( $\sigma$ ) away from the mean ( $\mu$ ). In normal distribution, about 68% of values are expected to be within  $1\sigma$  away from  $\mu$  and about 95.4% of the values lie within two standard deviations. Almost all values (about 99.7% ) are obtained within three standard deviations. The probability of finding almost all values within  $3\sigma$  tolerance interval is known as the *3-sigma rule*. In this thesis, a  $3\sigma$  error margin is considered to simulate the measurement data in later sections. That implies that the probability of measured values to be within the range of  $(\mu + \mu \times 3\sigma)$  and  $(\mu - \mu \times 3\sigma)$  is 99.73%.

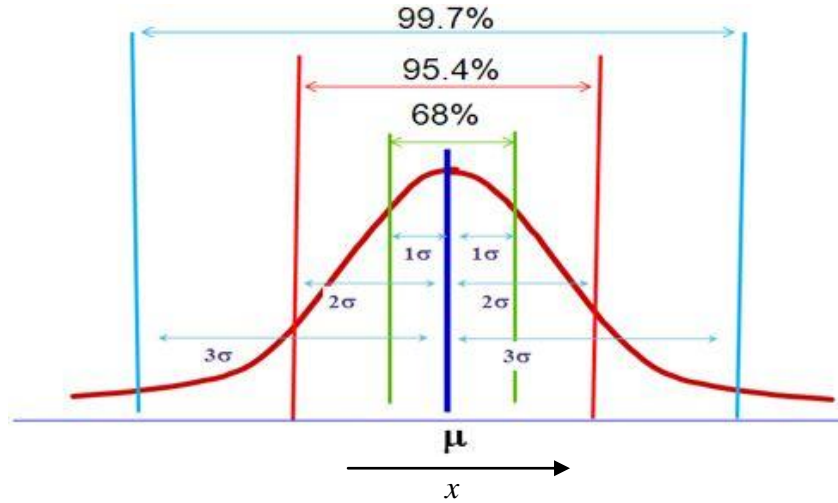


Fig. 3. 4: Normal error distribution and various confidence intervals

Equation (3.38) presents the formula used to calculate  $\sigma$ , the derivation is shown in appendix 1 and also in reference [46] applying a different procedure. The term *error* in (3.38) denotes percentage of expected maximum measurement errors with respect to corresponding  $\mu$ . The precision of measurement accuracy from which %*error* is calculated, is generally known for a SCADA sensor device.

$$\sigma = \frac{\mu \times \%error}{300} \quad (3.38)$$

It is assumed that any pair of errors are not correlated and each error is independent of others. The assumptions altogether could be defined as  $e \sim NID(0, \sigma^2)$  which implies that, in addition to normality any pair of errors  $e_i, e_j$  are uncorrelated and also independent [47].

### 3.4.3 Assumptions in Gauss-Newton Recursion

In the non-linear regression analysis, every primary element is required to be assigned with an initial value that should be reasonably close to the true value in order to initiate the recursion process. In power system problems, the primary and other states are usually expressed in per unit system. Therefore, starting with all voltage to be one p.u. and all phase angle to be zero radians are considered as a good initial guesses for a flat start of the SE tool.

Some equipments like static VAR compensators (SVCs) and DGs may show non-linear characteristics due to their saturation states and discrete outputs respectively. The effect on



the outcomes of the estimator [25] [24] due to such non-linear characteristics, are not considered in this work. In all cases, the objective function and measurement equations are assumed to be continuous and differentiable.

#### **3.4.4 Static Topology and Balanced System**

The topology of the network is assumed to remain unchanged. In other words, the SE tool considers a fixed topology of the network. The network parameters are supposed to be accurately known. The three phase characteristic is assumed to be symmetrical, balanced and fully transposed. This is justified as the DSSE is developed for MV network which is normally well balanced in the UK power system scenarios.

#### **3.4.5 Transformer and Branches**

Only tap changing transformers are taken into consideration in the presented case studies. The voltage drop is calculated for MV side when LV side measurement is available. In such cases, the core loss of the transformer is ignored when developing the equivalent circuit from short circuit test values. The static parameter values of all transformers and branches are assumed to be completely and precisely known. The tap positions of transformers are not considered as unknown states that need to estimate in this work.

#### **3.4.6 Measurement Errors**

The margin of error varies for various measurement types as discussed previously. It is assumed in this research that the pseudo-measurement errors are maximum  $\pm 50$  to  $\pm 70\%$  and real data errors are maximum  $\pm 1\%$  of corresponding true values. The virtual measurement error (when not considered as constraints) is assumed to be a very small number that is  $\pm 10^{-7}$  to  $\pm 10^{-9}$  times respective true values.

### **3.5 Conventional SE Solution Processes as Potential DSSE Tools**

Four conventional methods are discussed as potential methods for DSSE in this section. Weighted Least Squares (WLS) method is the most common and widely used state estimation (SE) method. Orthogonal (QR) Decomposition, Constrained Weighted Least

Square (CWLS), Hachtel's Augmented Matrix Method are alternative approaches to solve WLS optimization.

### 3.5.1 Weighted Least Squares (WLS)

The most commonly used state estimator is the normal equation based Weighted Least Squares (WLS) method, which provides an optimal solution when it is fed with known measurement variances and normally distributed measurement errors. The detail of this method has been discussed in section 3.2.

### 3.5.2 Orthogonal (QR) Decomposition

The objective function of Gauss-Newton iteration can be presented as follows, where we assume  $\mathcal{H} = W^{0.5}H$  and  $z = W^{0.5}z$

$$J(\Delta x) = \|z - \mathcal{H}\Delta x\|^2 \quad (3.39)$$

The orthogonal transformation method avoids squaring the gain matrix  $G = H^TWH$  by using the following decomposition technique:

$$\mathcal{H} = Q^T R = [Q_1^T \ Q_2^T] \begin{bmatrix} R_1 \\ 0 \end{bmatrix} \quad (3.40)$$

If  $Q$  is an orthogonal matrix then it yields  $Q^T Q = I$  and  $Q^T = Q^{-1}$  while satisfying

$$Q \mathcal{H} = R \quad (3.41)$$

Any  $m \times n$  matrix of full rank can be decomposed in the above way.  $R$  is an upper trapezoidal  $m \times n$  matrix.  $Q_1$  is  $m \times n$ ,  $Q_2$  is  $m \times (m - n)$  and  $R_1$  is a  $n \times n$  upper triangular matrix of  $R$ . Thereby,  $R_1$  preserves non-singularity as the system is assumed to be observable. This leads to recalculating the gain matrix as

$$G = H^T W H = \mathcal{H}^T \mathcal{H} = \mathcal{H}^T Q^T Q \mathcal{H} = (Q \mathcal{H})^T (Q \mathcal{H}) = R^T R \quad (3.42)$$

Furthermore, the right hand side of WLS normal equation in (3.11) can be recalculated as below

$$\begin{aligned}
H^T W [z - h(x)] &= (W^{0.5} H)^T W^{0.5} [z - h(x)] \\
&= \mathcal{H}^T W^{0.5} [z - h(x)] \\
&= R^T Q W^{0.5} [z - h(x)]
\end{aligned} \tag{3.43}$$

Therefore, replacing (3.11) by the gain matrix from (3.42) and the right hand side part from (3.43), the new Gauss-Newton state variable update equation is obtained as

$$R^T R \Delta x = R^T Q W^{0.5} [z - h(x)] \tag{3.44}$$

It is sufficient to build only the sub matrix  $Q1^T$  (3.40) rather than the full  $Q$ . Furthermore, as  $R$  is a regular matrix, (3.44) can be reduced as follows to solve the non-linear WLS [49] [2] [50] [1] [51]

$$R1 \Delta x = Q1 W^{0.5} [z - h(x)] \tag{3.45}$$

In the orthogonal factorization approach, (3.45) is the key equation that retains a smaller condition number while updating the state vector in each iteration, compare to classical WLS update equation in (3.11). As orthogonal transformations of vectors do not change their norm and the solution of basic WLS estimation problem can be considered as minimization of the Euclidean norm of the residual vector, the solutions from both algorithms are almost exact [1].

### 3.5.3 Constrained Weighted Least Squares (CWLS)

The CWLS method takes virtual measurement as constraints and hence improves the Gain matrix ill-conditioning problem. The CWLS method can be expressed as below [1] [2] [49] [50], when  $r = z - h(x)$  and virtual measurement equation  $c(x)$  is expressed as equality constraint.

$$\begin{aligned}
\text{Min } J(x) &= \frac{1}{2} [z - h(x)]^T W [z - h(x)] \\
\text{s. t. } c(x) &= 0
\end{aligned} \tag{3.46}$$

Solving using a Lagrange multiplier, the following is obtained:

$$\mathcal{L}(x, \lambda) = \frac{1}{2} (z - h(x))^T W (z - h(x)) - \lambda' c(x) \tag{3.47}$$

$$\frac{\partial \mathcal{L}}{\partial x} = H (z - h(x)) + C^T \lambda = 0 \quad (3.48)$$

$$\frac{\partial \mathcal{L}}{\partial \lambda} = c(x) = 0 \quad (3.49)$$

Here,

$$\frac{\partial h(x)}{\partial x} = H \quad \text{and} \quad \frac{\partial c(x)}{\partial x} = C.$$

$C$  is the Jacobian of  $c(x)$ . After deriving 1st order optimality conditions, the linearized Gauss-Newton update equation is obtained as

$$\begin{vmatrix} H^T W H & C^T \\ C & 0 \end{vmatrix} \begin{vmatrix} \Delta x \\ -\lambda \end{vmatrix} = \begin{vmatrix} H^T W r \\ \Delta c \end{vmatrix} \quad (3.50)$$

Here,  $\Delta x = \Delta x^{k+1} - \Delta x^k$  when  $k$  is the present iteration number.

The gain matrix ( $H^T W H$ ) in (3.50) excludes virtual measurements and therefore can avoid very high weighting values. This reduces the condition number of the gain matrix considerably. However, the coefficient matrix of (3.50) becomes indefinite and therefore requires more sophisticated numerical techniques [1] [49]. Advanced statistical problem solving software, such as MATLAB, can deal with such an indefinite matrix and solve the SE problem without requiring any additional coding.

### 3.5.4 Hachtel's Augmented Matrix

Hachtel's Augmented Matrix Method solves WLS optimization problems in such way that no  $H^T W H$  matrix is constructed in the solution process. The Hachtel method takes the residual equation as a constraint in addition to assuming virtual measurement constraints. Hachtel's method benefits from the properties of both orthogonal decomposition and CWLS methods. It considers the residual  $r$  as a parameter therefore calculation of  $H^T W H$  is avoided. The Hachtel's method [1] [49] [50] [52] can be expressed as shown below, when  $r = z - h(x)$

$$\begin{aligned} \text{Min } J(x) &= \frac{1}{2} r^T W r \\ \text{s. t. } &c(x) = 0 \end{aligned}$$

$$r - z + h(x) = 0 \quad (3.51)$$

Solving using Lagrange multipliers, the following is obtained

$$\mathcal{L}(x, r, \lambda) = \frac{1}{2} r^T W r - \lambda^T c(x) - \tau^T (r - z + h(x)) \quad (3.52)$$

Here, residual vector,  $r$  is treated as an unknown variable.  $\lambda, \tau$  are the Lagrange multipliers The Gauss-Newton linearization consists of three equations [50]

$$\frac{\partial \mathcal{L}}{\partial x} = H^T \tau + C^T \lambda = 0 \quad (3.53)$$

$$\frac{\partial \mathcal{L}}{\partial \lambda} = c(x) = 0 \quad (3.54)$$

$$\frac{\partial \mathcal{L}}{\partial \tau} = r - z + h(x) = 0 \quad (3.55)$$

$$\frac{\partial \mathcal{L}}{\partial r} = W r - \tau = 0 \quad , \quad (3.56)$$

Replacing  $r = W^{-1}\tau$  and linearising the remaining three equations, the Gauss-Newton iterative process generates the following:

$$\begin{vmatrix} 0 & 0 & C \\ 0 & \alpha W^{-1} & H \\ C^T & H^T & 0 \end{vmatrix} \begin{vmatrix} -\alpha^{-1} \lambda^{k+1} \\ \alpha^{-1} \tau^{k+1} \\ \Delta x \end{vmatrix} = \begin{vmatrix} \Delta c \\ \Delta z \\ 0 \end{vmatrix} \quad (3.57)$$

Here,

$$C \Delta x = \Delta c$$

$$\lambda^{k+1} = -\alpha^{-1} \lambda$$

$$\tau^{k+1} = \alpha^{-1} W \Delta r$$

$$\Delta r = \Delta z - H \Delta x$$

$$\Delta z = z - h \Delta x$$

$\alpha$  is a control parameter to control numerical stability of the problem.  $\alpha$  can be defined by fine tuning the solution, however in the SE problem  $\alpha = 1$  is frequently used. The coefficient matrix in (3.57) is called Hachtel's matrix or the Hachtel tableau. As this is a highly augmented formulation, hypothetically a lower condition number is expected. Due

to high sparsity, the method may not be computationally expensive when applied to a large system however it may require more efficient algorithms for row pivoting [1] [49]. However the implementation of the method in MATLAB does not require any additional coding as MATLAB acts as a black box to resolve the computation.

### 3.6 Novel Weighted Error Modulus (WEM) Method

Weighted Error Modulus (WEM) is a novel iteratively re-weighting approach introduced in this section as a potential DSSE tool. Although the WLS method can be applied successfully to detect and remove outliers, the method is not always efficient in detecting and overcoming the effect of gross measurement errors or bad data. In the presence of gross errors, an alternative optimization criteria known as Weighted Least Absolute Values (WLAV) is more effective. WLAV optimizes the absolute value of the residual vector instead of the quadratic value as in the WLS method [6]. A novel estimator, Weighted Error Modulus (WEM) method that exploits the benefits of both WLS and WLAV methods, is proposed as a candidate DSSE tool. In this approach, the weighting value associated with the measurement is modified iteratively within the WLS method.

In addition to the low level measurement errors, linearization errors from the Taylor series approximation of the optimization equation for  $J(x)$  and unexpected gross error can also be present in the measurement data. Assuming linearization and other unexpected errors are termed as  $le$ , the Gauss-Newton solution of WLS optimization function can be expressed as:

$$H \Delta x = r + le \quad (3.58)$$

Here,  $r = z - h(x)$  = vector of residuals. According to the Gauss-Newton principle,  $le$  is negligible provided that the initial guess of the state variable approximates the true value. The measurement is usually within a  $\pm 1\%$  error margin with respect to the true value and therefore the WLS method provides good estimation under normal conditions. This implies that the residual vector always has a considerably smaller value. However, this will not remain true if any gross error exists in the measurements. In that case, the assumption  $H \Delta x \approx r$  remains no longer valid since  $le$  in (3.58) would have a significant value. The proposed WEM method utilizes the characteristics of the variation in  $le$  depending on the accuracy of the measurement to re-weight the measurements. The principle of the algorithm is that

the measurement data corresponding to a larger residual eventually has reduced weight and the measurement with smaller residual has a gradually increased weighting factor. Hence,  $w_i$  is modified iteratively such that for  $k^{\text{th}}$  iteration

$$w_i^{k+1} = \frac{u_i}{|(H \Delta x)_i - r_i|^k} \quad (3.59)$$

Here  $u_i$  is the measurement re-weighting factor that corresponds to the actual weighting factor,  $W$  at the first iteration. As the recursion process approaches a convergence point,  $\Delta x$  becomes trivial and  $x^k \approx x^{k+1}$ . Consequently,  $(H \Delta x)_i$  becomes negligible,  $|z - h(x^k)| \approx |z - h(x^{k+1})|$  and therefore,  $|r_i|^k \approx |r_i|^{k+1}$  at that point. That also implies,

$$|(H \Delta x)_i - r_i|^k \approx |(H \Delta x)_i - r_i|^{k+1} \approx r_i^{k+1} \quad (3.60)$$

By replacing the value from (3.60) in (3.59),  $w_i^{k+1} = \frac{u_i}{|r_i|^{k+1}}$  is obtained. The WLS minimization problem can therefore be expressed as follows:

$$\text{Min} \sum_{i=1}^M w_i |r_i|^2 = \text{Min} \sum_{i=1}^M \frac{u_i |r_i|^2}{|r_i|} = \text{Min} \sum_{i=1}^M u_i |r_i| \quad (3.61)$$

Equation (3.61) represents the final objective function that resembles the WLAV minimization criteria. In such a way, the WEM commences the process with the WLS objective functions while the weighting factor is being recurrently updated. At near convergence point, (3.60) occurs and the minimization function is transformed to WLAV objective function. Essentially, the WEM method attempts to reduce low level measurement errors by the WLS method and gross errors by the WLAV method. Hence, WEM method combines the advantages of the objective functions of both the WLS and the WLAV methods [53]. The entire process is depicted in Fig. 3. 5.

There are two nested iteration loops in this algorithm: one outer and one inner iteration as shown in Fig. 3. 5. After accomplishment of each set of inner loop iterations, the elements of the weighting matrix are replaced by the most recent corresponding re-weighting factors. The outer loop controls Gauss-Newton recursion while the inner loop performs the re-weighting, gain matrix and mismatch calculations. The inner loop is assumed to be

converged when the mismatch vectors of two consecutive inner loop iterations are almost equivalent; while the outer loop's convergence criteria is set as mismatch values approximating zero. The inner loop iteration is capped at a smaller number of iteration (that is five in this thesis) and allowed to terminate even if full convergence is not achieved. The outer loop controls whether the algorithm is satisfying the Gauss-Newton convergence criteria and should terminate.

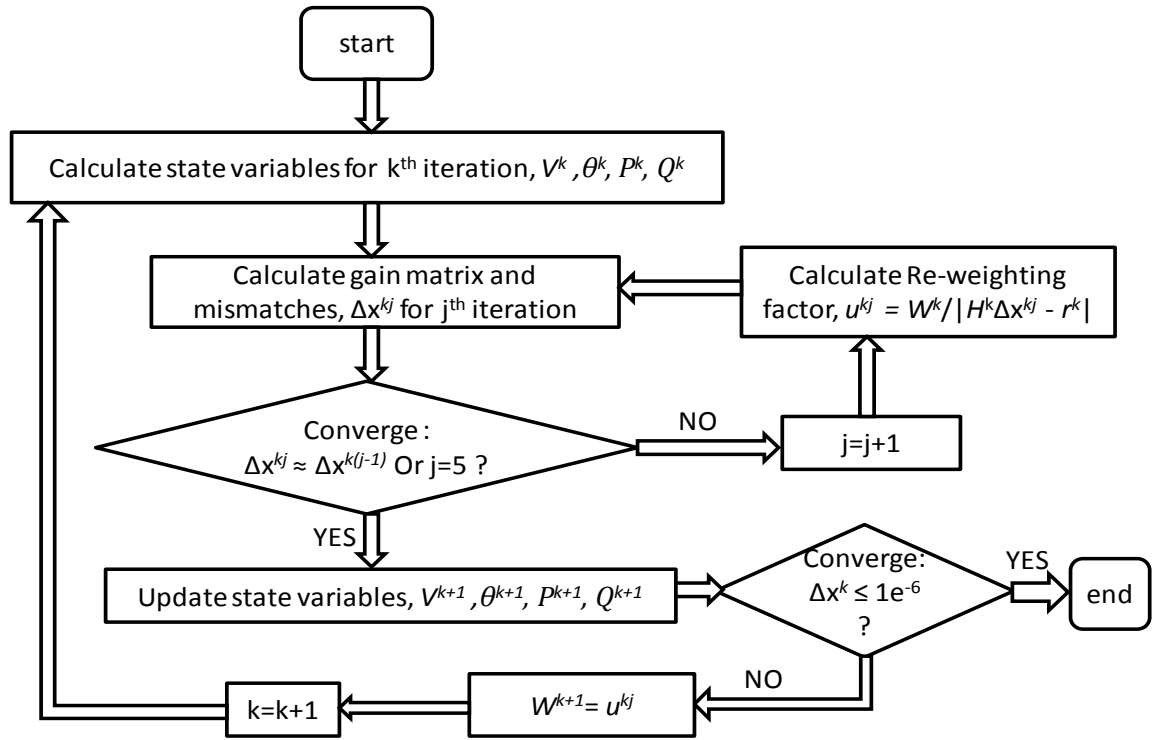


Fig. 3. 5: WEM method flowchart

The WEM method can be compared with the iteratively re-weighted method known as Schweppe-Huber Generalized-M (SHGM) estimation. The SHGM is developed as a robust estimator to suppress bad data and avoid the influence of the leverage point. It utilizes two different objective functions representing WLS and WLAV criteria. The value of residual remaining above or below certain threshold value determines whether the estimator will behave like a WLAV or WLS estimator respectively. The threshold value is set by the weighting factor and a tuning parameter [1] [17]. The proposed WEM method is similar to the SHGM estimator in a way, since both methods combine the benefits of WLAV and WLS estimators. However, unlike SHGM, WEM does not need to define two different objective functions. The process starts with the WLS objective function, which is gradually transformed to the WLAV equivalent optimization criteria inherently (through the re-weighting procedures). There is also no need to define any tuning parameter for WEM.



### **3.7 Concluding Remarks and Discussion**

The sensors distributed over the network to read and transmit the operating state information are prone to telemetry failure and erroneous information. Furthermore, it is not practically possible to achieve perfectly simultaneous measurement from all parts of the network. The state estimation tool uses a set measurements to calculate the system state. This estimated data can then be fed into DMS functions [1] [2]. An initial discussion of typical power system state estimation tools has been included in this chapter. The SE building blocks, network components and critical assumptions used for SE have been defined. Important mathematical formulations such as measurement equations, normal equation, and Jacobian matrix are identified and discussed in detail. These formulations are used while implementing DSSE in MATLAB. The chapter excludes current and phase angle measurement equations as the current and phase angle measurements are not considered in this work.

Furthermore, five different SE solution processes are discussed as candidate DSSE tools. In addition to normal equation based classical WLS, alternative solution processes are considered, which include orthogonal factorization, constrained WLS and Hachtel's Augmented Matrix Method. A novel WEM optimization method is introduced as a candidate solution in this chapter.

# **CHAPTER 4**

## **SELECTION AND APPLICATION OF HACHTEL'S AUGMENTED MATRIX METHOD**

The performance of five candidate DSSE solutions discussed previously in Chapter 3 is assessed through various case studies in order to select one of these as the most potentially useful DSSE tool in this chapter. The case studies are performed on simulated datasets and model networks that represent general features of UK distribution networks. A detail discussion is provided, leading to the selection of Hachtel's Augmented Matrix method, based on the outcomes of the case studies. One of the important contributions in this chapter is the demonstration of the performance of Hachtel's method on the real networks and authentic datasets in various off line case studies.

### **4.1 Application of Candidate SE Solution Processes**

The five candidate DSSE solution methods are applied on two different sizes of distribution network models to inspect their relative performances. The assessment process includes the comparison of computation time, convergence speed and the quality of estimated data under realistic scenarios.

#### **4.1.1 Network Descriptions**

The test networks are provided by the United Kingdom generic distribution system (UKGDS) datasets, that represents UK model distribution networks and is developed as a platform of benchmark tests for research purposes [54]. A 77 node and a 356 node radial network are selected for analysis and comparison of the performance of candidate solution processes. The 77 node radial network consists of 76 branches and the 356 node network contains 354 branches. The detail of the network data and line diagrams are provided in appendices 2, 3 and 8. The UKGDS default datasets contain network parameters, generation and load data; that enable complete load flow studies to be performed, in order to calculate true values of all states, i.e. phase angles, voltage magnitudes, power injections, power flows and current flows. It should be noted that in this thesis, the state

values obtained from the load flow studies are considered as the 'true values' which are generally never known in practice.

#### 4.1.2 Measurement Data

Taking into consideration the severe limitation of sensor data at the distribution level, the minimum availability of the real time measurement data is presumed for the assessment of five candidate methods. The minimum observability requirements for the case studies are, at least one set of real measurements should be available at the Grid Supply Point (GSP) and many pseudo-measurements for all other nodes. One of the nodes in both networks is a zero injection zero load point, that provides one pair of real and reactive power injection virtual measurements. The real measurement set includes one voltage magnitude, one pair of active and reactive power measurements; while the pseudo measurements are active and reactive power injection values on a node. A few more real measurements are assumed to be available in the network as per requirements in a couple of case studies. The measurement values are simulated data generated by adding Gaussian errors to the true values. The real and pseudo-measurement errors are of maximum of 1% and 50% respectively in  $3\sigma$  confidence interval [55] [56]. The measurement weights are greatly dependent on the expected errors in the measured and load-modelled values as explained in section 3.2 and 3.4. The virtual measurement weight is considered as  $10^{14}$ . MATLAB *normrnd* operator is used to generate the simulated measurement data that are normally distributed around the true values for the given standard deviation.

#### 4.1.3 General Overview of Assessment Processes

The relative performance is evaluated for equal network conditions and measurement availabilities. In addition to the computation time and required number of iterations, different scenarios reflecting possible real life situations are included in the evaluation process. The relative errors of primary state variables are observed as their accuracy strongly indicates the overall estimation correctness. The voltage and phase angle magnitudes are defined here as  $|V|$ ,  $|\theta|$  respectively and the relative errors are calculated by

$$\%error = 100 \times \frac{(true\ value - estimated\ value)}{true\ value} \quad (4.1)$$

The above equation is applicable to all primary state variables except the phase angle of slack node as it is fixed to zero value. In case studies incorporating extensive Monte Carlo

simulations, the estimation quality is quantified by taking the maximum and mean of absolute values of %*error* for  $|V|$  and  $|\theta|$  estimates.

It is understood that a single measurement configuration is not sufficient to appraise the performance in a case study in the presence of a significant number of highly erroneous pseudo-measurements. The range of the error in pseudo-measurements can be as high as 50%, which gives a very wide confidence interval leading to greater uncertainty in the outcomes from different estimation tools. Therefore, the Monte Carlo method, incorporating 100 runs, that creates different sets of measurement values within expected error distributions in each run, is applied in the assessment process. Scrutinizing such a high number of Monte Carlo studies offers more confidence to draw any conclusion with regard to the performance of each candidate solution.

#### 4.1.4 Computation Time

The mean computation times for a single run of the five candidate methods on both the networks are plotted in Fig. 4. 1 and Fig. 4. 2. The respective number of required iterations are plotted in Fig. 4. 3 and Fig. 4. 4. The network is assumed to be equipped with one set of real measurements, one virtual and pseudo-measurements for the remaining nodes. All methods are executed in the expected operating condition while no gross error exists. The time of execution is calculated for an Intel(R) Core(TM) i5-2450M 2.50GHz processor.

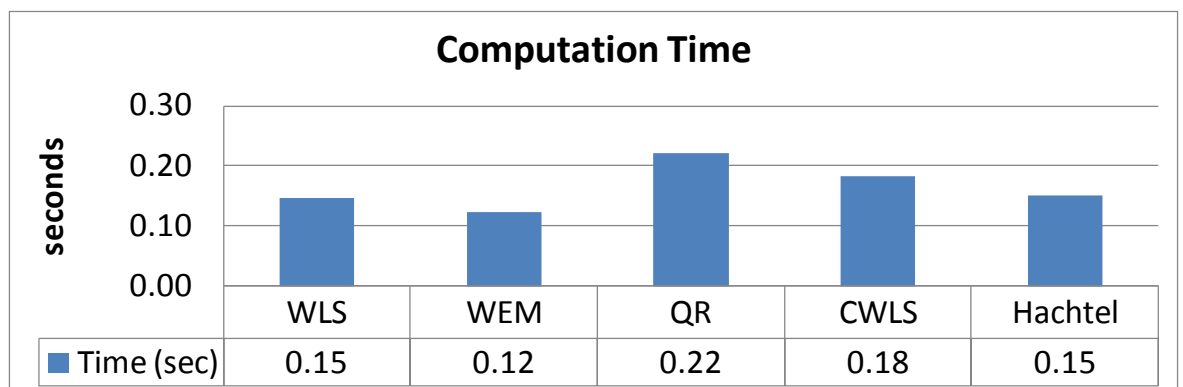


Fig. 4. 1: Computation time for 77 node network

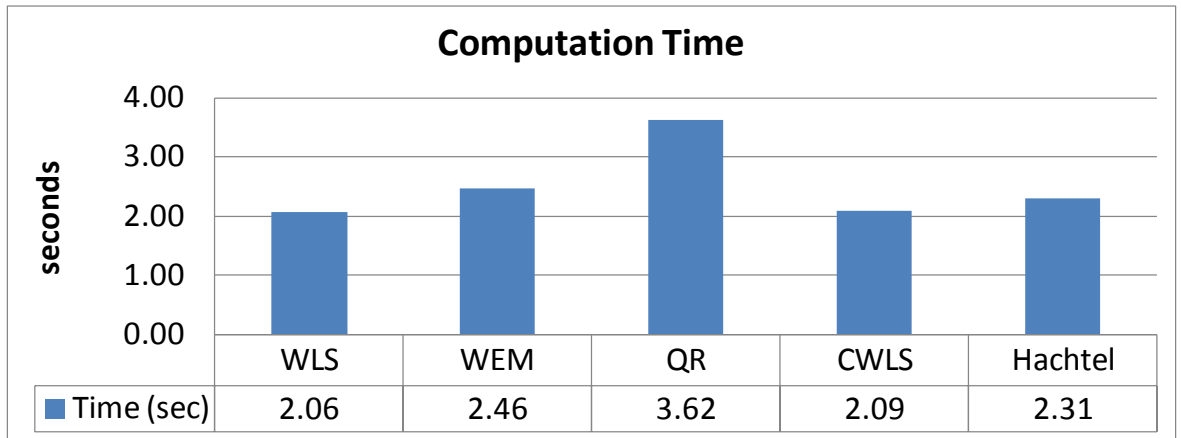


Fig. 4. 2: Computation time for 356 node network

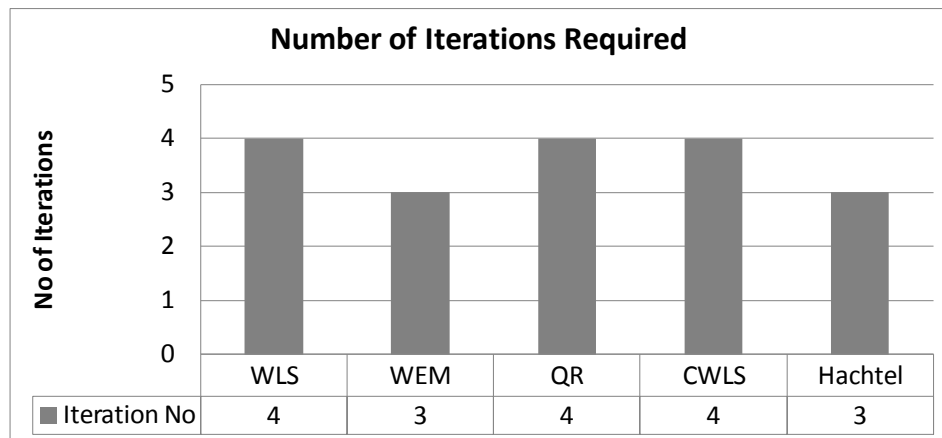


Fig. 4. 3: Required iteration for 77 node network

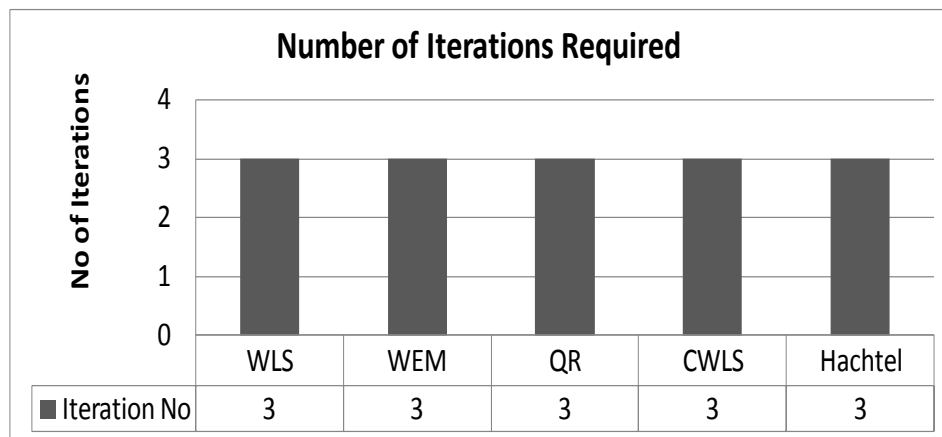


Fig. 4. 4: Required iteration for 356 node network

Although the required number of iterations for various DSSE tools are very similar, the execution times are different because of different computation methodologies. WLS is a normal equation based solution processes requiring the least computation steps, therefore the computation time it takes is expected to be the minimum for the same number of iterations as the others. On the other hand, WEM consists of one inner and one outer loop

which means the total iterations require are equal to the number of inner loop iteration times the number of outer loop iterations. Certainly, WEM is usually expected to take a longer time for SE calculation. The Orthogonal (QR) Decomposition method completes SE requiring a few additional steps for matrix decomposition and calculation, costing additional computation time. Since both CWLS and Hachtel's methods need further calculation and data processing compared to classical WLS, this may have some effect on the computation time. Accordingly, Fig. 4. 1 and Fig. 4. 2 show that the maximum time is required by QR methods for both networks, as expected. Since the WLS methods require four iterations (Fig. 4. 3) to complete SE for the 77 node networks, it requires somewhat larger execution time but is still the second minimum in Fig. 4. 1. Apparently, the WEM method requires the least time and only three iterations to solve SE for 77 node networks ( Fig. 4. 1 and Fig. 4. 3). The CWLS and Hachtel's Augmented Matrix methods require reasonable time in the cases of both the networks. As the number of required iterations are equal for the 356 node network, the computation time plot reflects perfectly the hypothetical interpretation- i.e. Orthogonal and WEM methods being computationally most expensive, with the WLS method requiring the shortest execution time.

#### 4.1.5 Assessment Scenarios

Five scenarios representing typical characteristics with regard to distribution networks and sensor data are simulated to observe the effect on the ( $|V|$ ,  $|\theta|$ ) estimation quality of the candidate SE processes. Each scenario is assessed for 100 Monte Carlo simulations in order to quantify the average effect of the presence of highly unpredictable pseudo-measurements. The absolute values of the mean and maximum voltage and phase angle estimation errors with respect to corresponding true values are plotted. The description of five scenarios is provided as follows.

*Scenario 1:* The performance for each method through extensive Monte Carlo based studies is observed under ideal conditions. In this case, ideal conditions are defined by the state of the system when all real and pseudo-measurement errors are within expected thresholds (real measurements around 1%, pseudo-measurements around 50% as maximum error margin with 99.7% confidence) and the topology parameters values are accurate. The topology parameters obtained from original UKGDS data sets and simulated measurement values without any gross errors represent the ideal conditions for the both networks.

*Scenario 2:* Given that all static parameter data are expected to be accurately known in the conventional SE problem, the effect of their being unpredictably inaccurate is measured under this scenario. As a matter of fact, the extensive volume of branch impedance data are vulnerable to human error. The effect of such erroneous branch impedance information on estimation quality is observed in this scenario.

*Scenario 3:* Here the effect of sign error in an injection measurement on the accuracy of different DSSE solution processes is observed. Traditionally passive feeders do not contain any power supplying node point; hence the power injection readings are always considered as negative representing consuming power irrespective of the signs the devices show. Under this circumstance, the sensor devices available at a node may not be installed in the correct direction introducing sign errors to the measurement data. This had not been any significant problem for the network characterized as passive, however the concern raises due to increased penetration of the DGs as that enhances the possibility of higher voltage and power supplying node points in a feeder. The DSSE tool is therefore likely to be fed with injection measurements prone to sign errors in an active network. In this case study scenario, additional simulated injection measurements are placed in the network and the input data is assumed to contain one sign error.

*Scenario 4:* The combination of scenario 2 and 3 is investigated i.e. how estimation quality is affected when a branch impedance error and a sign injection error are simultaneously present in the data fed into the estimator.

*Scenario 5:* One common problem at distribution system is the existence of very short and low impedance branches that can contribute to ill-conditioning problems and be accountable for bad results. Lack of real measurements on the low impedance branches may increase the variance of estimate values [2]. The performance of various DSSE solution processes is studied in the presence of such short branches. Since a very short branch is expected to have lower resistance and reactance values, they are normally located at the tail of the network tree. The end branch of a feeder is chosen to deliberately reduce its resistance and reactance values significantly to represent a very short branch. As network parameters are modified for this study, a load flow study prior to execution of SE is performed to obtain corresponding sets of true values of the state.

#### 4.1.5.1 Assessment Scenario 1: Ideal Condition

Voltage and phase angle estimation errors (4.1) with respect to load flow results for the state of the network have been plotted for 100 sets of simulated measurement applied to five processes. It is clearly visible that all five processes produce a similar quality of estimation (Fig. 4. 5 and Fig. 4. 7) except for small differences in maximum phase angle estimation errors by the WEM optimizer (Fig. 4. 6 and Fig. 4. 8). The mean voltage estimation errors are less than 0.4% in all cases however maximum voltage estimation errors are around 1.4%. The phase angle estimation errors are always significantly higher than the voltage estimation errors in both cases. This large values in errors can be attributed to the effect of highly erroneous pseudo-measurements and lack of real measurement data in required amount and lack of extensive real measurement data.

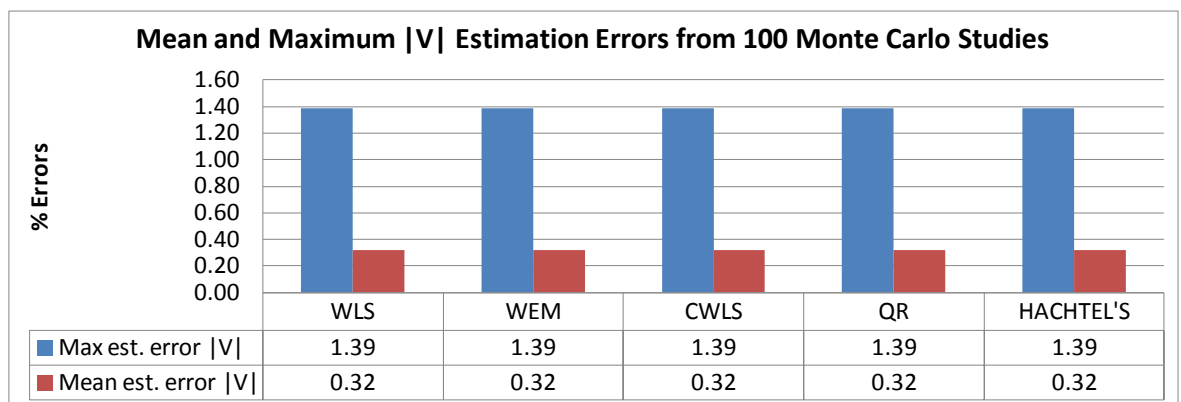


Fig. 4. 5:100 Monte Carlo studies for voltage,  $|V|$  estimation errors on 77 node network (scenario 1)

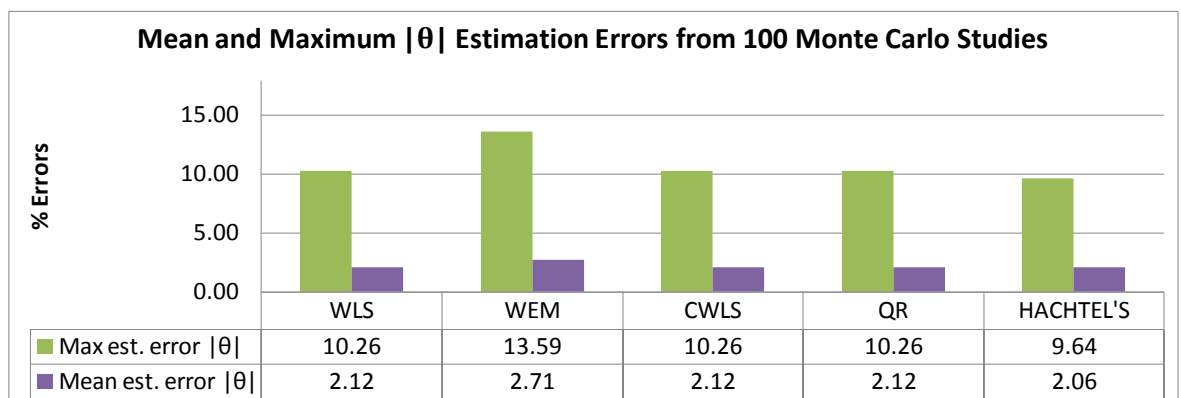


Fig. 4. 6:100 Monte Carlo studies for phase angle,  $|\theta|$  estimation errors on 77 node network (scenario 1)



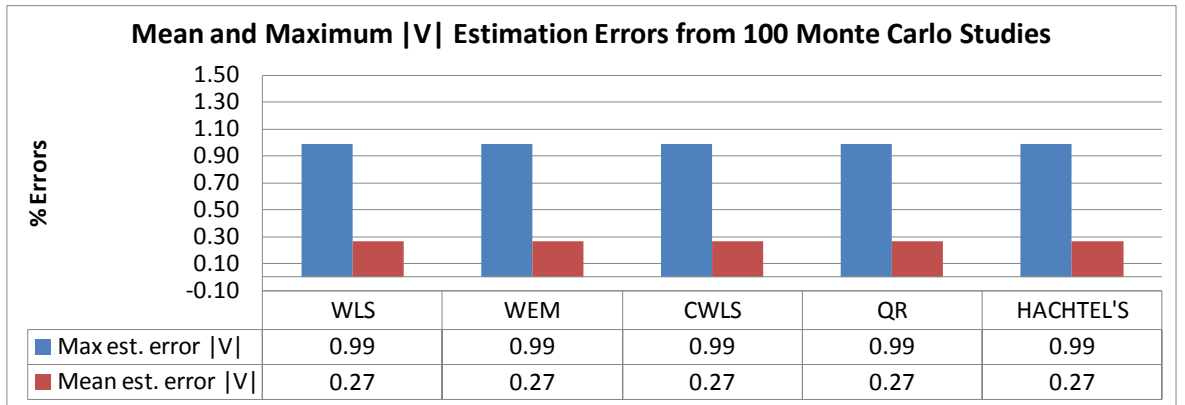


Fig. 4. 7:100 Monte Carlo studies for voltage,  $|V|$  estimation errors on 356 node network (scenario 1)

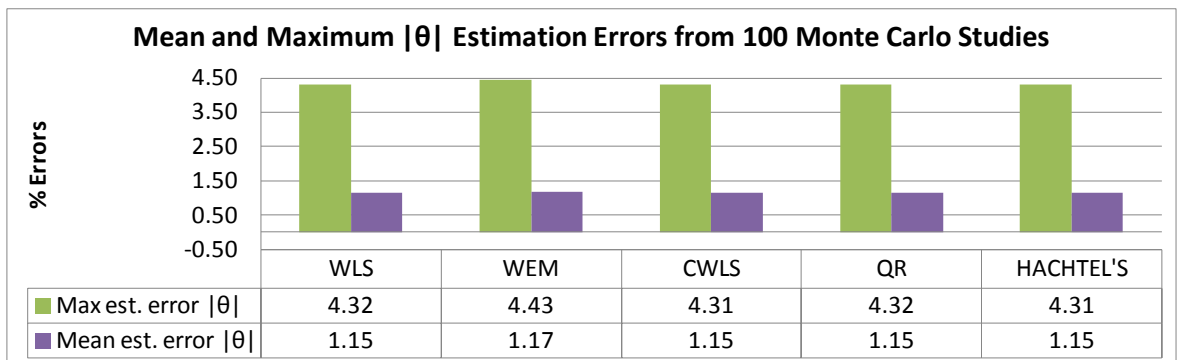


Fig. 4. 8:100 Monte Carlo studies for phase angle,  $|\theta|$  estimation errors on 356 node network (scenario 1)

Table 4. 1 presents the mean value of number of iterations required to converge for 100 Monte Carlo studies. The required number of iteration varies with different measurement sets for candidate DSSE tools in the case of the 77 node network. The WEM method requires the greatest number of iterations, on average; even after neglecting the nested inner loop iterations. Overall, Hachtel's method requires the least number of iterations, on average. The 356 node network requires a consistent number of iterations to converge in all cases.

<b>Scenario 1</b>					
<b>Networks</b>	<b>WLS</b>	<b>WEM</b>	<b>CWLS</b>	<b>QR</b>	<b>HACHTEL'S</b>
77 node	3.39	3.63	3.39	3.39	3.14
356 node	3	3	3	3	3

Table 4. 1: Average number of iteration for 100 Monte Carlo (scenario 1)

#### 4.1.5.2 Assessment Scenario 2: Errors in Branch Impedance

In this scenario, several branches are chosen at different locations of the network to significantly modify the impedance values, performing one change at a time. The actual impedances of these branches are reduced by a factor of  $10^{-5}$ . For example, the original impedance  $[R=0.061570248, X=0.047438017]$  pu for the branch indexed as 20 of the 77 node network, has been changed to  $[R=6.15702 \times 10^{-07}, X= 4.7438 \times 10^{-07}]$  pu to correspond to a human error in input parameter data. Fig. 4. 9 to Fig. 4. 12 show that the error in branch impedance value has the most adverse effect on the WEM estimates. The phase angle errors can increase quite significantly (Fig. 4. 12) in some instances due to deterioration of matrix conditioning for WEM. This kind of behaviour of WEM can be explained from the perspective of principles of the algorithms. As it tends to emphasize measurement that results in less residual values in every outer loop iteration, the error impedance value may deceive the estimator to produce smaller residual values for more erroneous measurements. As a result, WEM ends up trusting more erroneous observed values and deteriorating estimation quality. Other than that, the mean and maximum voltage estimation errors remain similar to those under ideal conditions.

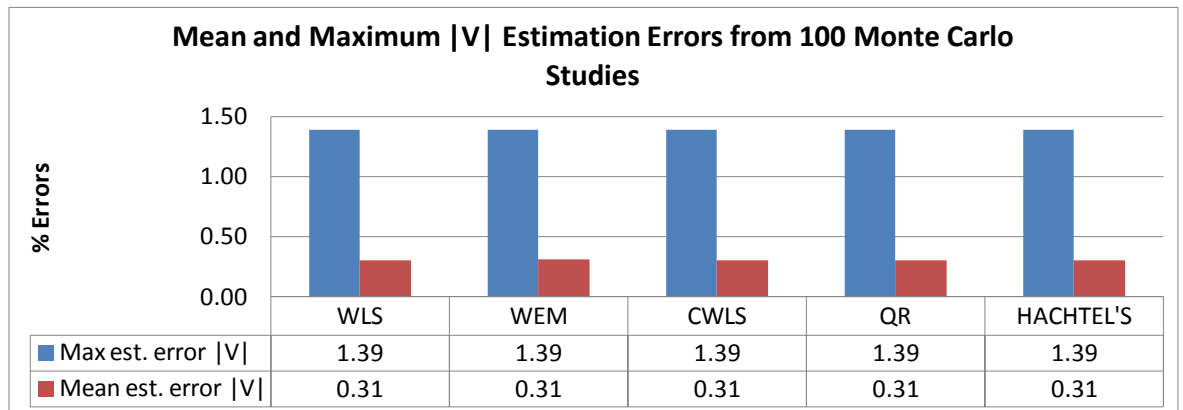


Fig. 4. 9:100 Monte Carlo studies for voltage, |V| estimation errors on 77 node network (scenario 2)

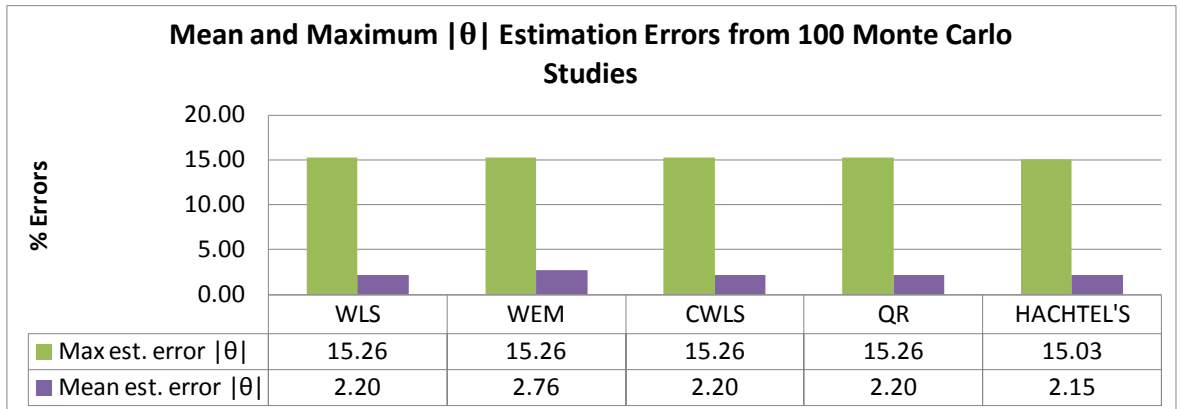


Fig. 4. 10:100 Monte Carlo studies for phase angle,  $|\theta|$  estimation errors on 77 node network (scenario 2)

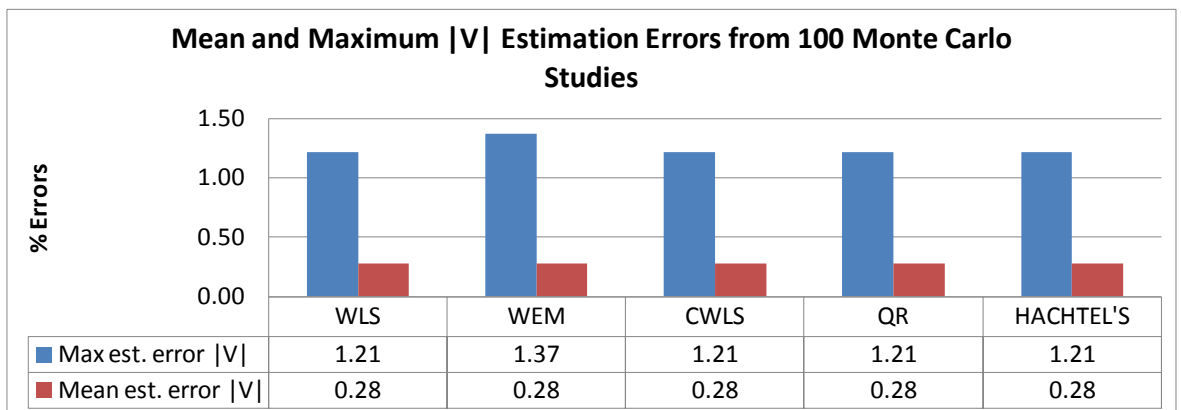


Fig. 4. 11:100 Monte Carlo studies for voltage,  $|V|$  estimation errors on 356 node network (scenario 2)

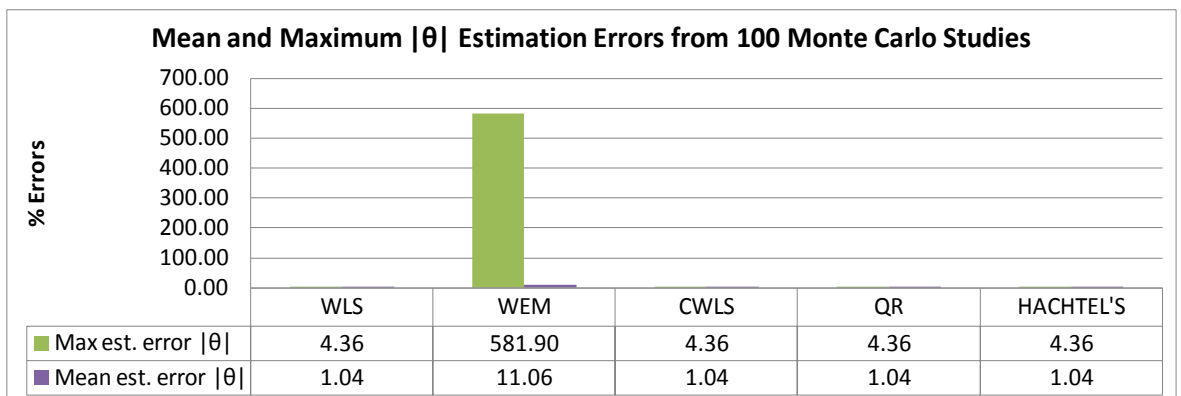


Fig. 4. 12:100 Monte Carlo studies for phase angle,  $|\theta|$  estimation errors on 356 node network (scenario 2)

Similar effects are achieved by selecting different locations of error impedance branch, as shown in appendix 4. The plot below shows voltage and phase angle estimation errors when branch 3 in 77 node network and branch 134 in 356 node network are assumed to be

erroneous. Table 4.2 depicts that the average iteration values for 100 Monte Carlo simulation is maximum with regard to WEM and classical WLS methods when applied on 77 and 356 node networks respectively. Hachtel's method requires the minimum average iterations for both the networks. The average iteration for CWLS is quite high for the 356 node network although the quality of estimation remains very similar to other estimators (except WEM).

<b>Scenario 2</b>					
<b>Networks</b>	<b>WLS</b>	<b>WEM</b>	<b>CWLS</b>	<b>QR</b>	<b>HACHTEL'S</b>
77 node	3.66	3.84	3.64	3.43	3.1
356 node	7.72	4.96	7.39	3	3

Table 4. 2: Average number of iteration for 100 Monte Carlo (scenario 2)

#### **4.1.5.3 Assessment Scenario 3: Injection Sign Errors**

The scenario based study is performed for both pseudo and real measurement sign errors. The injection error nodes are selected from three parts of the network representing, `

- beginning of feeders
- start node of sub-lateral branch
- end of the feeder

The location of these error injection measurement nodes are shown in appendix 4. One sign error at a time, as well as a couple of errors existing simultaneously, are studied for the assessment of relative performance by different estimators. In all the cases, the estimation quality is not affected much in the presence of injection sign errors, for both the test networks, as the mean and maximum errors are similar to those in section 4.1.5.1 in the ideal state. Such outcomes can be attributed to the insignificant effects of erroneous load values that are trivial with respect to the total loads. For instances, the load at node 53 of the 77 node network occupies around 1.39% and the load at node 144 of the 356 node network is only around 0.27% of collective load values.

The quality of estimates is similar for all SE solution processes, except the WEM results in slightly more erroneous estimates. Similar performance by different candidate DSSE tools are observed in cases such as different sign error locations, type of sign error measurements (real and pseudo) and number of sign errors (one and two). Only one case study is presented in Fig. 4. 13, Fig. 4. 14, Fig. 4. 15 and Fig. 4. 16. In the presented case

studies, sign injection real power errors on node 53 and node 144 respectively in the 77 and 356 node networks are assumed.

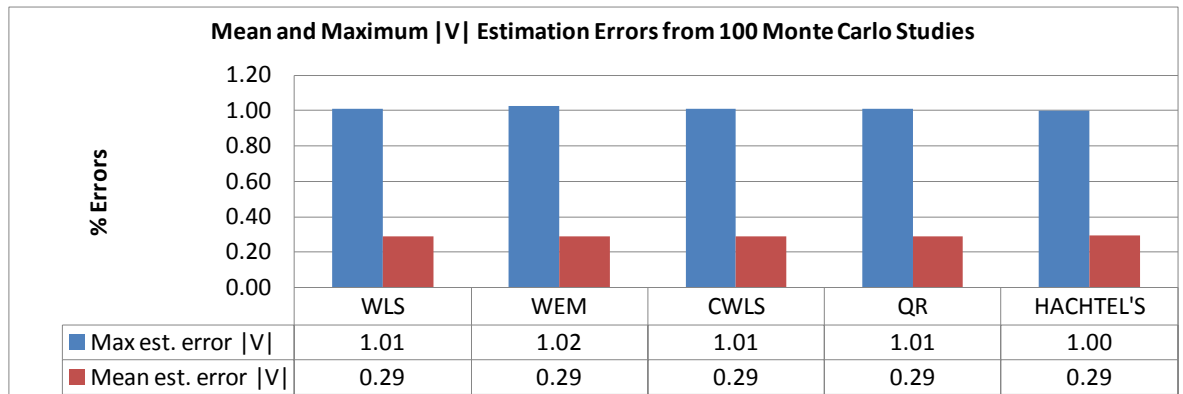


Fig. 4. 13:100 Monte Carlo studies for voltage,  $|V|$  estimation errors on 77 node network (scenario 3)

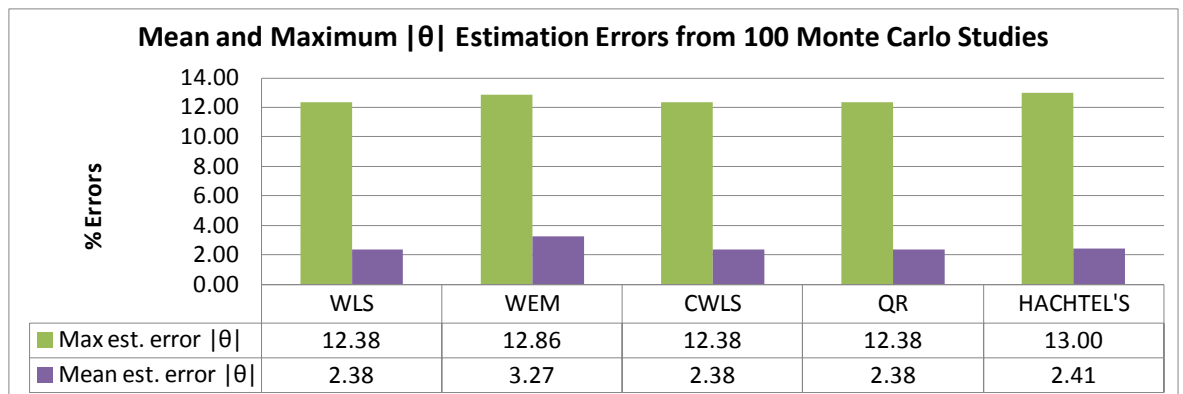


Fig. 4. 14:100 Monte Carlo studies for phase angle,  $|\theta|$  estimation errors on 77 node network (scenario 3)

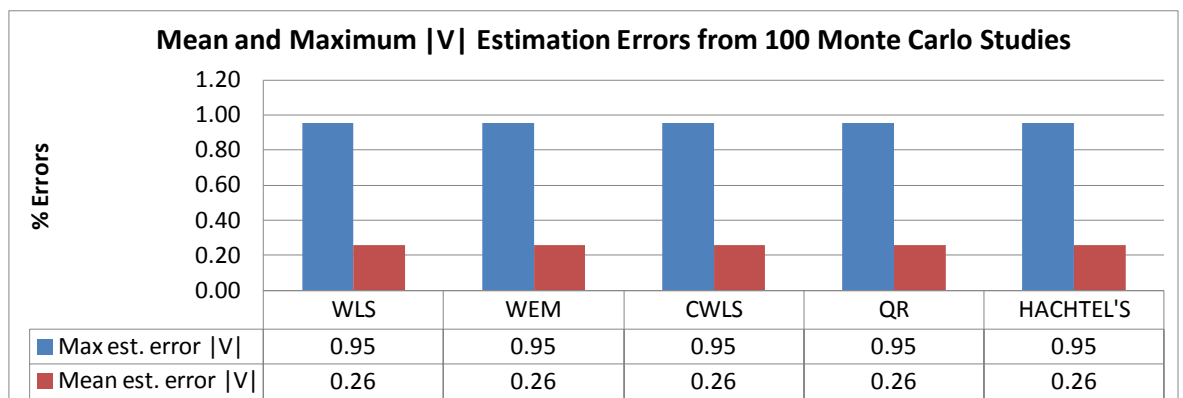


Fig. 4. 15:100 Monte Carlo studies for voltage,  $|V|$  estimation errors on 356 node network (scenario 3)

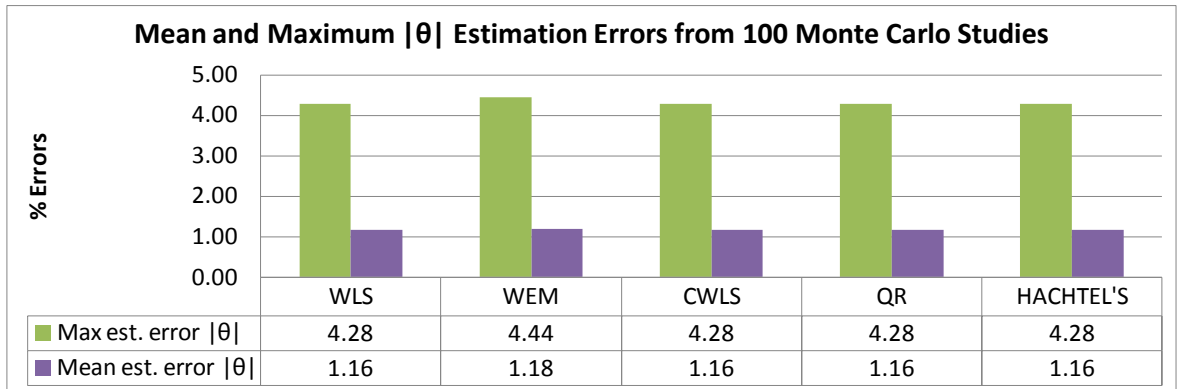


Fig. 4. 16:100 Monte Carlo studies for phase angle,  $|\theta|$  estimation errors on 356 node network (scenario 3)

The average number of iterations is the same for the 356 node network and varies slightly for the 77 node network, applying different processes. The Hachtel's method requires the least and WEM requires the most iterations to converge in the presence of injection sign errors for the 77 node network. The numbers in Table 4.3 also justify the insignificant effect of erroneous power measurement values. The erroneous load value is a relatively larger percentages of the total load in case of the 77 node network, when it is compared with the counterparts of the 356 node network. The SE for the 77 node network is observed to have some effect of the erroneous measurements as the convergence rate becomes slightly slow to filter out the errors (Table 4. 3). However, the 356 node network remains still unaffected with respect to the convergence rate as shown in Table 4. 3.

<b>Scenario3</b>					
<b>Networks</b>	<b>WLS</b>	<b>WEM</b>	<b>CWLS</b>	<b>QR</b>	<b>HACHTEL'S</b>
77 node	3.85	3.95	3.85	3.85	3.46
356 node	3	3	3	3	3

Table 4. 3: Average number of iteration for 100 Monte Carlo (scenario 3)

#### 4.1.5.4 Assessment Scenario 4: Combination of Injection Sign and Impedance Value Errors

The simultaneous existence of both erroneous branch impedance data and sign error injection real measurement is simulated to quantify their effect in Fig. 4. 17, Fig. 4. 18, Fig. 4. 19 and Fig. 4. 20. The WEM method generates significant phase angle estimation errors for 356 node network, which is similar to its performance in scenario 2. The

maximum voltage estimation error is also higher for the WEM method compared to other estimates. Since power injection sign errors are observed to have limited impact in the section 4.1.5.3, the error in impedance data should be liable for the deteriorated quality of estimation.

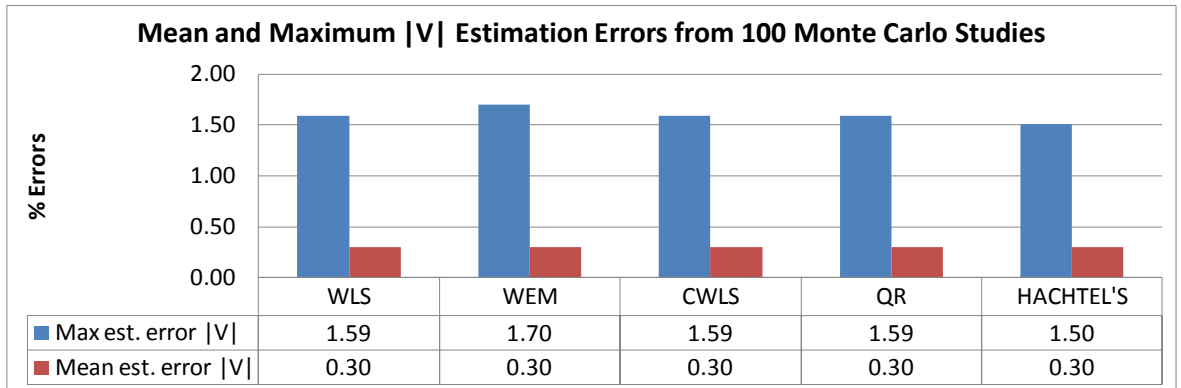


Fig. 4. 17:100 Monte Carlo studies for voltage, |V| estimation errors on 77 node network (scenario 4)

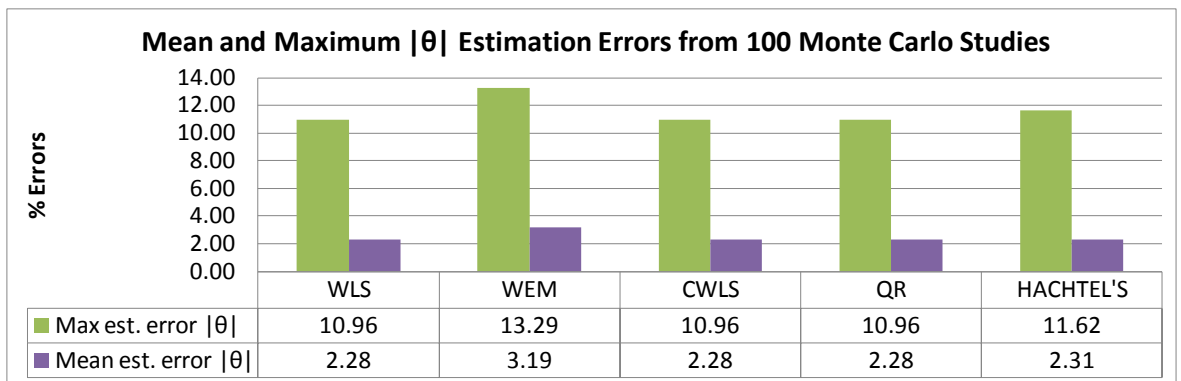


Fig. 4. 18:100 Monte Carlo studies for phase angle,  $|\theta|$  estimation errors on 77 node network (scenario 4)

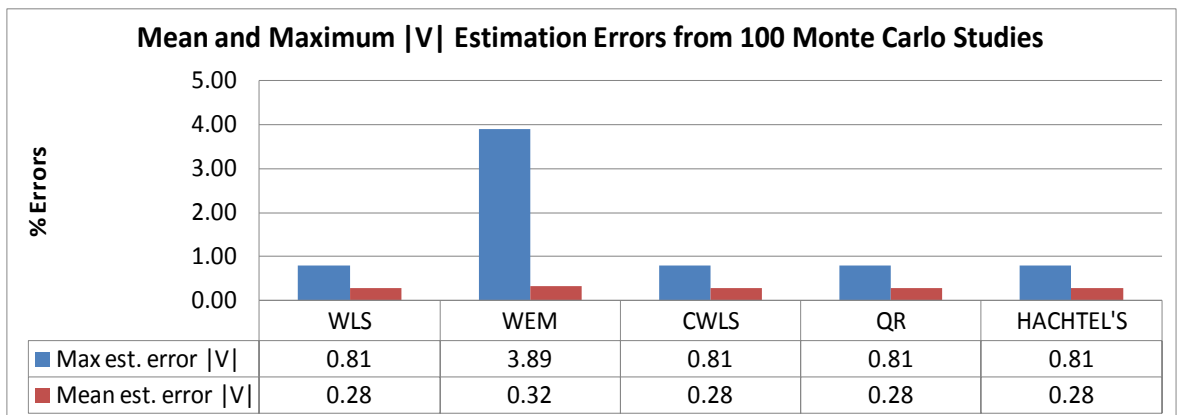


Fig. 4. 19:100 Monte Carlo studies for voltage, |V| estimation errors on 356 node network (scenario 4)

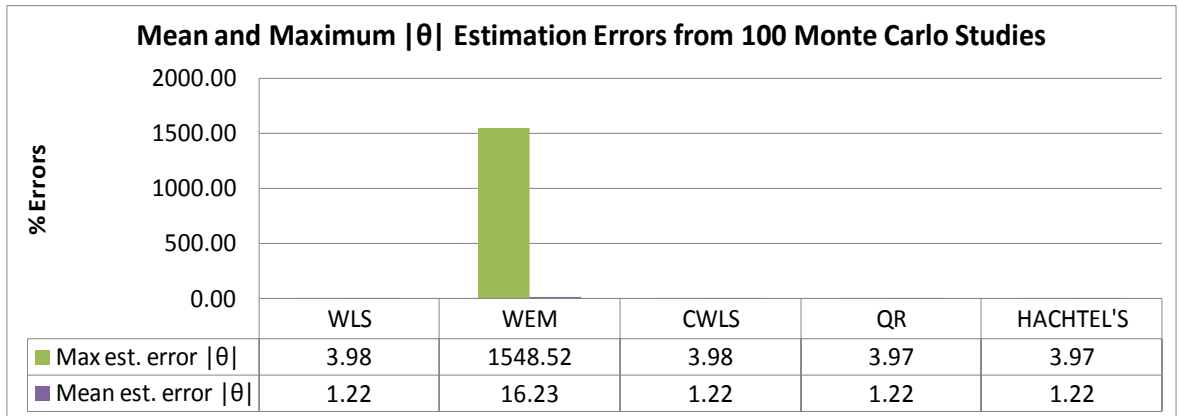


Fig. 4. 20:100 Monte Carlo studies for phase angle,  $|\theta|$  estimation errors on 356 node network (scenario 4)

CWLS and WLS require additional iterations to converge to estimate the 356 node network states compared to the requirement for the 77 node network as shown in Table 4.4. In contrast, Orthogonal (QR) Decomposition and Hachtel's method require fewer iterations when the network size becomes larger. The average iteration counts do not change considerably for WEM, QR and Hachtel's method when network size changes from 77 to 356 nodes.

Scenario 4					
Networks	WLS	WEM	CWLS	QR	HACHTEL'S
77 node	3.87	3.95	3.88	3.78	3.4
356 node	7.53	5.76	7.51	3	3

Table 4. 4: Average number of iteration for 100 Monte Carlo (scenario 4)

#### 4.1.5.5 Assessment Scenario 5: Very Short Branch with Low Impedance

The low impedance branches are generally located at the end of the feeder as they do not require to supply loads beyond their own nodes. The location of the selected branches, indexed as 3 and 134, for 77 and 356 node networks respectively, are treated as a very low impedance branch are shown in appendix 4. Fig. 4. 21, Fig. 4. 22,

Fig. 4. 23 and Fig. 4. 24 show the influence of the existence of very short branch in the network on the estimated values. Apparently, most of the candidate optimizers are not affected significantly due to the existence of one very short branch with the exception of voltage and phase angle estimation by WEM method.



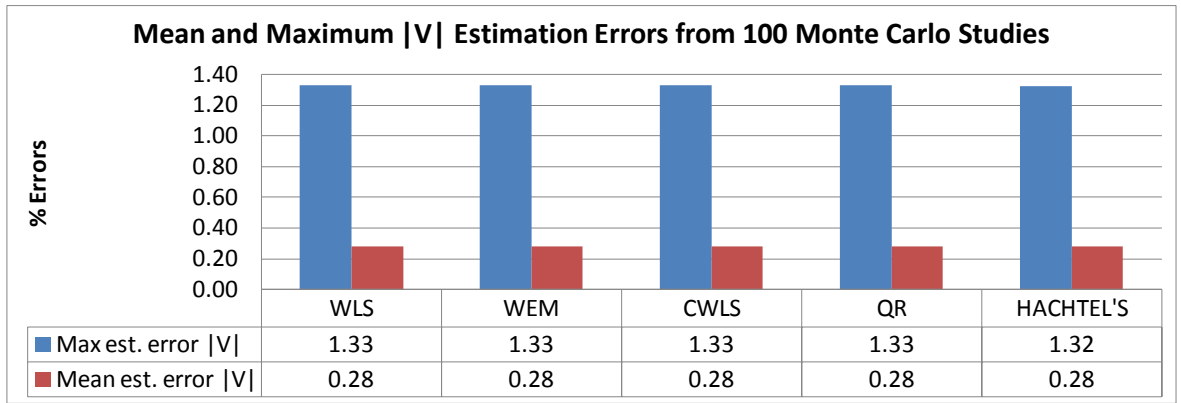


Fig. 4. 21:100 Monte Carlo studies for voltage, |V| estimation errors on 77 node network (scenario 5)

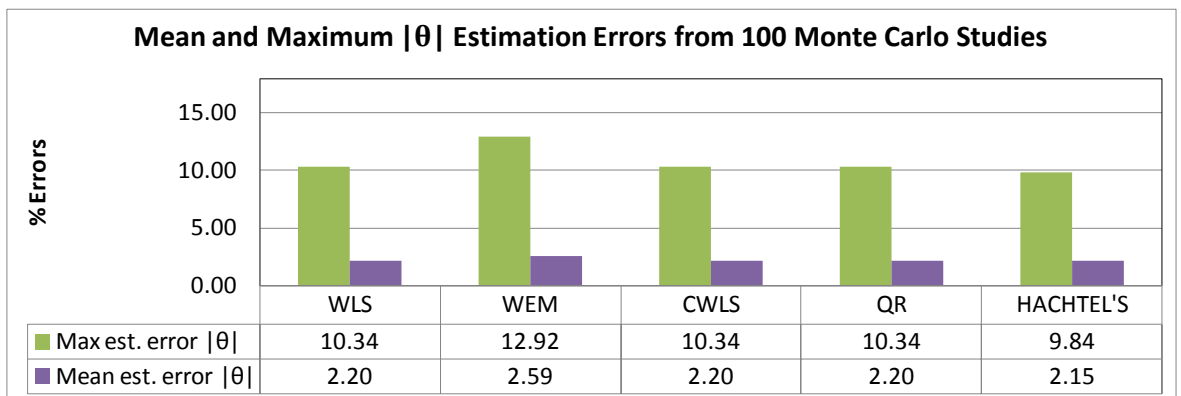


Fig. 4. 22:100 Monte Carlo studies for phase angle,  $|\theta|$  estimation errors on 77 node network (scenario 5)

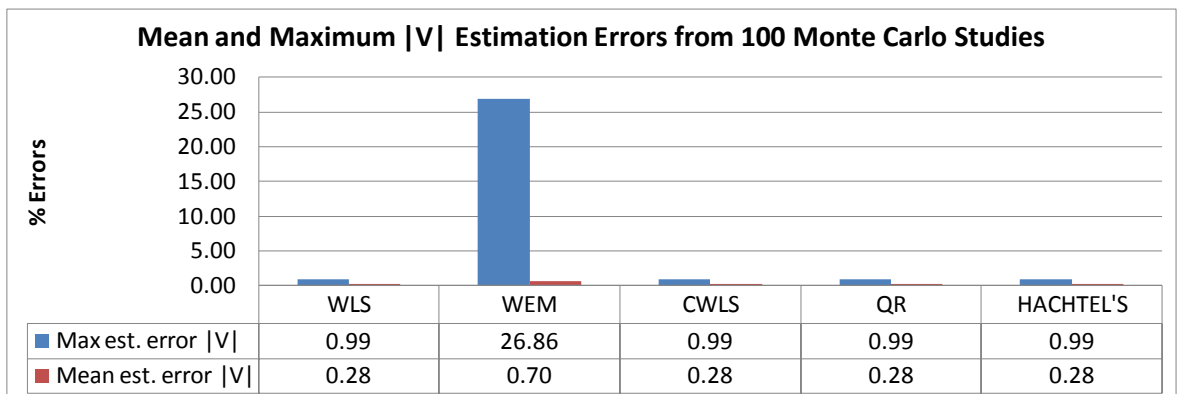


Fig. 4. 23:100 Monte Carlo studies for voltage, |V| estimation errors on 356 node network (scenario 5)

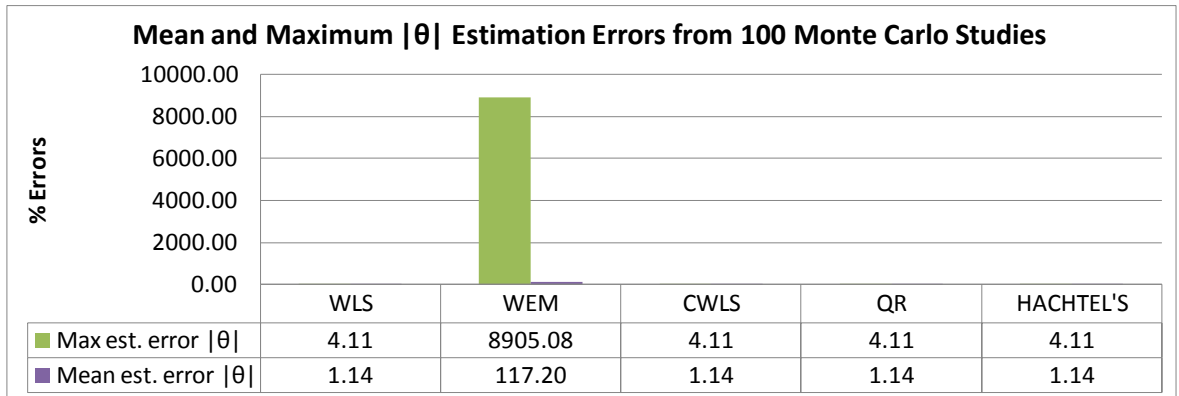


Fig. 4. 24:100 Monte Carlo studies for phase angle,  $|\theta|$  estimation errors on 356 node network (scenario 5)

The average iteration number increases by more than two as WLS, WEM and CWLS methods are applied to the larger network i.e. 356 node network in Table 4.5. On the other hand, the QR and Hachtel's methods converge slowly for the smaller (i.e. 77 nodes) network and the differences in average number of iterations are trivial.

<b>Scenario 5</b>					
<b>Networks</b>	<b>WLS</b>	<b>WEM</b>	<b>CWLS</b>	<b>QR</b>	<b>HACHTEL'S</b>
77 node	3.7	3.79	3.78	3.53	3.03
356 node	6.34	5.88	7.53	3	3

Table 4. 5: Average number of iteration for 100 Monte Carlo (scenario 5)

## 4.2 Selection of Hachtel's Augmented Matrix Method as Potential DSSE Tool

In section 3.5, 3.6 and 4.1, five SE solution processes which are classical WLS, CWLS, Orthogonal (QR) Decomposition, Hachtel's Augmented Matrix and WEM methods, are described as candidate DSSE tools. In this section their hypothetical values and performance on test cases are analysed to select one as the most potentially useful method to be used for distribution SE. The selection will be based on the computation time, convergence property, robustness of the process and most importantly, the quality of estimation they can provide.

Classical WLS is a popular and widely used optimizer in power system SE problems especially for transmission systems, due to its excellent performance in removing errors

generated from low level noises. However, the application of WLS is more challenging at distribution levels where significant numbers of pseudo and virtual measurements may cause deterioration of the gain matrix condition number. One of the major sources of matrix ill-conditioning is the high weighting factors assigned to virtual measurements in the normal equation based SE solution strategy. The matrix in CWLS and Hachtel's Augmented method do not contain such significant values, as virtual measurements are considered as equality constraints, is in a separate equation in these methods. Furthermore, the Hachtel's method does not form any normal equation or gain matrix at all and the equality equation is never squared in the solution equation for CWLS and Hachtel's methods. Hence the CWLS and Hachtel's methods acquire robustness with regard to the ill-conditioning problem due to measurement weights. However any possibility of the presence of bad data in the enforced constraint may leave catastrophic effects on convergence and estimation quality [49], but that is usually an unlikely event to occur.

The Orthogonal ( $QR$ ) Decomposition and the Hachtel's methods can be considered as the most robust methods, since they both do not form a gain matrix or normal equation. Studies under various scenarios give strong evidence in support of the robustness of both methods. In addition to consistency in producing good quality estimated data, the mean iteration values under different scenarios and networks appear not to be affected significantly. However the  $QR$  method benefits from being strongly well conditioned by default. The  $QR$  factorization is mathematically more robust than  $LU$  factorization as stated in [1]. A few drawbacks of this method include lower sparsity than the gain matrix, memory storage requirement [50] and computation time.

As Hachtel's method tactically avoids  $H^TWH$  calculation, it gives somewhat better conditioning than CWLS method. The Hachtel's method tends to be numerically more stable and is theoretically expected to generate less erroneous solutions than constrained normal equation e.g. CWLS [50]. While Orthogonal Decomposition has been proved to be the most robust but a computationally costly SE solution process, CWLS is more economic in computation time and preserves good conditioning. The Hachtel method can be considered as a compromise between Orthogonal factorization and CWLS, as it provides better robustness than CWLS and faster computation than the  $QR$  method [50] [52]. While avoiding formation of complete normal equations, the coefficient matrix of Hachtel's and CWLS remain no longer positive definite, therefore they require more sophisticated ordering and factorization. This can be treated as a trivial problem to be concerned about

since advanced software technology, such as MATLAB, can execute such complex computation efficiently.

Significant errors in estimated values are obtained from the WEM estimation tool in a few case studies, which indicates that the iteratively reweighted least squares based methods may not be efficiently applied to DSSE where limited real measurements are available. WEM tends to adjust the weight with magnitudes of relevant measurement residual values to attribute greater emphasis to coherent measurement data. The method, therefore, instead of treating all pseudo-measurements equally, prefers a few of them to gain more weight as the solution approaches convergence. The reweighting factors are expected to elect those pseudo-measurements which are closer representation of real states- which is not happening as observed in case studies. The iteratively weight assignment is possibly leading the WEM estimator to trust somewhat more erroneous measurements, deteriorating estimation quality. The error becomes quite large in the cases of impedance data errors and existence of very short branches. One possible reason of such unexpected performance can be attributed to the assumption of Gaussian error distribution. At near convergence point, the WEM method starts behaving more like the WLAV estimator, which is based on the maximum log-likelihood of the Laplace error density function. Therefore, the Gaussian error assumption no longer remains consistent with the objective function as the process approaches to the convergence [17]. Hence, the principle of the WEM method is not performing at the expected level in the case studies.

The quality of estimation is very much similar for the least squares error minimization based algorithms i.e. WLS, CLWS, QR and Hachtel's methods. The calculation time and iteration requirements vary, as the solution processes are different. The convergence characteristics based on required iteration numbers are always better in the case of Orthogonal Decomposition and Hachtel's methods, even in presence of various errors and specific conditions of the network. Their operation time however is longer for the same number of iterations when compared with that of WLS and CWLS. Hachtel's and Orthogonal methods can still be considered to have better convergence properties, as the convergence of WLS and CWLS is delayed at some operational states. The relative convergence characteristics are shown in Table 4. 1 to Table 4. 5.

Considering the pros and cons of five candidate solutions, WEM can be excluded to be considered as a potential DSSE tool, deeming it's lower estimation quality and higher

execution time. The Orthogonal method is very robust, however its quite high computation time may turn out infeasible to apply to large networks. Completely normal and semi-normal equation based WLS and CLWS methods respectively suffer from ill-conditioning problems and increased iteration requirements when applied to larger (356 node) networks with erroneous DSSE inputs. This may raise strongly the problem of scalability. Hachtel's Augmented Matrix method takes longer estimation time than WLS and CWLS and shorter than WEM or Orthogonal Decomposition methods for the same number of iterations. However, the average number of iterations it requires remain steadily within the range of 3 to 3.5 for various network states and sizes. The quality of estimation is consistent and very similar for WLS, CWLS, QR Decomposition and Hachtel's methods.

The critical analysis confirms that Hachtel's Augmented Matrix method can provide most robust and consistent outcome while maintaining acceptable estimation quality and convergence speed in various scenarios as well as network sizes. Therefore, it is proposed as the most useful potential tool for DSSE.

### **4.3 Application of Hachtel's Augmented Matrix Method on Real Data**

In this section, Hachtel's Augmented Matrix as the DSSE tool is applied to real networks and sensor datasets representing Slovenian distribution networks operated by the DNO, Elektro Gorenjska (EG), Slovenia. Two 20 kV feeders known as Cerklje and Sencur, are selected for case studies. The feeders are connected to a 110kV-20kV transformer known as Primiskovo. The proposed Hachtel's Augmented Matrix method is applied to solve the SE problem, utilizing the offline and time tagged real measurement sensor data. The sensor readings are available in the intervals of 5 to 15 minutes over a week. Both the static and dynamic data have been provided in MS Excel data format by EG. The network diagram and static data are presented in appendix 5. The case studies demonstrated in this section, include the comparison of relative voltage estimates, when feeding the estimator all voltage sensor data as well as only SCADA voltage sensor data, to observe the effect of reduced voltage measurements. One special scenario is demonstrated assuming real measurements as constraints. The offline field data is applied for ten sets of reading instances to observe how good the estimated values match the measured data.

### 4.3.1 Description of Sensor Data

Sencur and Cerklje feeders are equipped with a few real measurements and the remainder are unmeasured values. Real measurement data includes SCADA data, Quality Meter (QM) data and Advance Meter Infrastructure (AMI) data. The useful sensor data generally include voltage measurements, power injection measurements with signs indicating power consumption/generation and current magnitudes. The average active and reactive load consumption information of the unmeasured nodes are provided, albeit these are not obtained from concrete load modelling and not time tagged. The branch reactance and resistance are presented in the actual *ohm* values along with rated current and voltage information for each branch.

The SCADA data provides telemeter sensor data from HV-MV substations. The accuracy of SCADA measurement is within range of 1% of rated voltage and current and the data is provided in 15 minute regular intervals for an entire week. SCADA measurements include single phase voltage, three phase voltage, three phase current, total active-reactive power injection, temperature and transformer tap position data. Each sensor is associated with a unique number while recording the transmitted measurement data as shown in an example in appendix 7.

The QM data are measured from the LV side of MV-LV substation providing measurements on LV side for three phase voltages, three phase currents, overall and three phase active and reactive powers, harmonic distortion and energy consumption. The QM data are provided within the time interval of 5 to 10 minutes. All the QM measurements provided are not an instantaneous value at that time stamp, rather the average values measured from the prior readings [56]. An example QM data is shown in appendix 7.

In addition to QM and SCADA measurements, EG also has provided AMI measurements that is power consumption and generation data. However these data have not been used for DSSE, as AMI collects consumption data only from a subset of the customers connected to the LV substation. Hence, the AMI data only provides a fraction of the total power injection/generation values at a node.

Most of the measured data are used as DSSE inputs, except harmonic, energy and temperature information since these are not relevant information. The average load information are assumed as the pseudo measurement and their errors are assumed to be up

to 70% to calculate their weighting factors, to take into account greater uncertainty. It should be noted that the pseudo measurement errors are considered higher than those used in the previous section, since the pseudo measurement data are not load-modelled in this case. The current magnitude data do not incorporate the sign to indicate the direction of flow. It is therefore not directly used in the measurement equation; rather it is used for transformer voltage drop calculation.

The sensor data provided by EG are pre-processed before being fed into the Hachtel's DSSE tool. The network parameters and load data are converted to the per unit values exploiting the corresponding rated voltage and current values. The line voltage data are converted to phase voltage and the three phase power injection/consumption data are summed to calculate the total amount. The data accuracy range varies with the type of variables measured. For instance, SCADA voltage, current and real power measurements provides 1% accuracy, whereas the accuracy of QM counterpart data is 0.5%, as provided by EG datasets.

### 4.3.2 Sensor Data Processing

The sensor data require pre-processing before they are fed into the DSSE.  $V_L, V_{ph}, I_{ph}, P, Q, S$  symbols represents line voltage, phase voltage, current, real power, reactive power and apparent power respectively in the following discussion.

- As the SCADA measured power values are very small in some time steps, they appear as zero in the datasheet. In such cases, power values are calculated from voltage and current magnitude measurements. The general formula for this is as below.

$$P = \sqrt{3}V_L I_L \cos\theta,$$

$$Q = \sqrt{3}V_L I_L \sqrt{(1 - \cos^2 \theta)}$$

However, since per phase current data is available and power factor is assumed to be ~0.97 as suggested by the EG DNO, the actual equation applied for calculation looks as below

$$P = \sqrt{3}V_L \times (I_{ph1} + I_{ph1} + I_{ph1}) \times 0.97$$

$$Q = \sqrt{3}V_L \times (I_{ph1} + I_{ph1} + I_{ph1}) \times \sqrt{(1 - 0.97^2)}$$

- Several numbers of QM do not read reactive powers, in such cases the reactive power ( $Q$ ) is calculated from real ( $P$ ) and apparent power ( $S$ ) values.

$$P^2 + Q^2 = S^2$$

- The SCADA voltage measurements are obtained by converting line voltage data to phase voltages.

$$V_{ph} = \sqrt{3} \times V_L$$

- All the sensor values are converted to the per unit system before they are fed to the DSSE tool. In all cases, 20kV is the base voltage,  $V_{base}$  and 31.5 MVA is the base Volt-Amp,  $S_{base}$  in per unit calculations. The base impedance  $Z_{base}$ , is calculated as

$$Z_{base} = \frac{V_{base}^2}{S_{base}}$$

The per unit conversion is performed by dividing the actual value by the corresponding base value

$$per\ unit\ value = \frac{actual\ value\ (Z, V, I, P, Q)}{base\ value\ (Z_{base}, V_{base}, I_{base}, P_{base}, Q_{base})} PU$$

- The exact data format supplied by EG in MS Excel file is not directly used. The node names are numbered according to node index that has been used to calculate admittance matrix. EG has provided all SCADA data in 15 minutes interval collectively in one excel file and QM data in 5 to 10 minutes interval in separate files for different QM. In order to simplify input files uploading process and to reduce the time, feeder wise excel file is created where relevant static and dynamic data of the feeder are transferred. Different QM and SCADA meter data are deposited in separate sheets of the same excel file as DSSE input. The DSSE tool accesses the file from MATLAB to read and load required information.
- Since QM data are measured from LV side of the 21-0.4kV transformer, the following data processing are performed
  - The QM voltage measurements are calculated taking the mean of all available phase voltage measurements followed by their transformation to the MV side.



- The real power measured at LV side is considered as power consumption from MV side. Transformer losses are added to LV real power measurement values when considering from MV side. MV-LV transformer data provides the short circuit voltage and transformer rated voltage and powers, the transformer voltage drop is calculated using that information.
- The snapshot below presents a transformer data set provided by EG. The respective calculation of voltage drop and MV substation voltages are shown for the transformer coded as T0279.

		Specifications of transformers								
MV/LV substation	Quality meter type	$S_n$ [kVA]	$U_p$ [kV]	$U_s$ [kV]	connection	tap type	tap position	$u_k$ [%]	$P_{Fe}$ [kW]	$P_{Cu}$ [kW]
T0279 GRAD CERH	MI7150	250	21	0.42	Dyn5	A	3	4	0.425	3.25

Short circuit voltage ( $U_k$ ) is stated as 4% of rated voltage ( $U_p, U_s$ ). Considering secondary side

$$U_k = \frac{U_s \times 4}{100}$$

The rated current,

$$I_s^{rated} = \frac{S_n}{\sqrt{3} \times U_s}$$

Approximate equivalent impedance from LV side (ignoring core losses),

$$\begin{aligned} Z_s &= R_s + j X_s \\ &= \frac{u_k \times U_s^2}{100 \times S_n} \end{aligned}$$

Copper losses,

$$P_{cu} = 3 \times (I_s^{rated})^2 \times R_s$$

$$\text{Therefore, } R_s = \frac{P_{cu}}{3 \times (I_s^{rated})^2}$$

$$\text{and } X_s = \sqrt{(Z_s^2 - R_s^2)}$$

Phase angle,

$$\cos \theta = \left| \frac{P_{sensor}}{S_{sensor}} \right|, \quad \sin \theta = \left| \frac{Q_{sensor}}{S_{sensor}} \right|$$

The voltage drop,

$$\Delta V_{transformer} = (R_s + jX_s) \times I_{sensor(mean)} \angle \pm \theta \text{ (leading/lagging)}$$

Here,

$$I_{sensor(mean)} = \text{average value of three phase LV current sesnor readings}$$

$$\angle \theta = \cos \theta \pm j \sin \theta$$

The MV side transformer voltage,  $V_{MV}$  is calculated using the following formula

$$V_{MV} = p (V_{sensor(mean)} + |\Delta V_{transformer}|)$$

$$V_{sensor(mean)} = \text{average value of three phase LV voltage sesnor readings}$$

$$p = \text{transformation ration based on tap position}$$

$p = \frac{21}{0.42}$  in this case. The obtained values are then converted to the per unit convention by dividing by the base value.

### 4.3.3 Off-line Field Data Test Applying Hachtel's Method

The quality of phase angle estimation, as well as secondary state estimation values, may not remain within the DNO's required error threshold limit in the presence of a number of poorly defined pseudo-measurement inputs. Realizing the present situation, DNOs are more interested to obtain quality voltage estimation to apply to voltage control functionalities. Four case studies are performed on real data sets to analyse the comparative performances. The first three cases are studied to observe how well estimated values match the measurements under various scenarios for one random set of measurements. The fourth case presents off line application of state estimation using real SCADA and QM data provided for different time steps. The four case studies are:

- Voltage estimation assuming all SCADA and QM measurement data are available
- Voltage estimation assuming all power injection and only SCADA voltage measurement data are available
- Voltage estimation assuming real and reactive power measurements as constraints
- Voltage estimation for ten sets of measurement data

A voltage measurement can have a significant effect in overall voltage estimation quality; similarly an error in a voltage sensor can deteriorate the estimation quality severely. The effect can be even stronger in small networks such as Ceklje and Sencur. An individual

power measurement value on the other hand, has less effect on the overall voltage estimation quality. As power injection measurement values are quite small in per unit with very small error probabilities (0.5% to 1%), they can be treated as close approximation to true values. This justifies performing an experimental study by assuming real measurement as constraints in addition to virtual measurements. The Hachtel's Augmented Matrix objective function is thereby modified from (3.51) to:

$$\begin{aligned}
 \text{Min } J(x) &= \frac{1}{2} r^T W r \\
 \text{s.t. } z_c - c(x) &= 0 \\
 r - z + h(x) &= 0
 \end{aligned} \tag{4.2}$$

$z_c$  is the measured value of corresponding measurement equation for the constrained values.  $z_c = 0$  for virtual measurements.

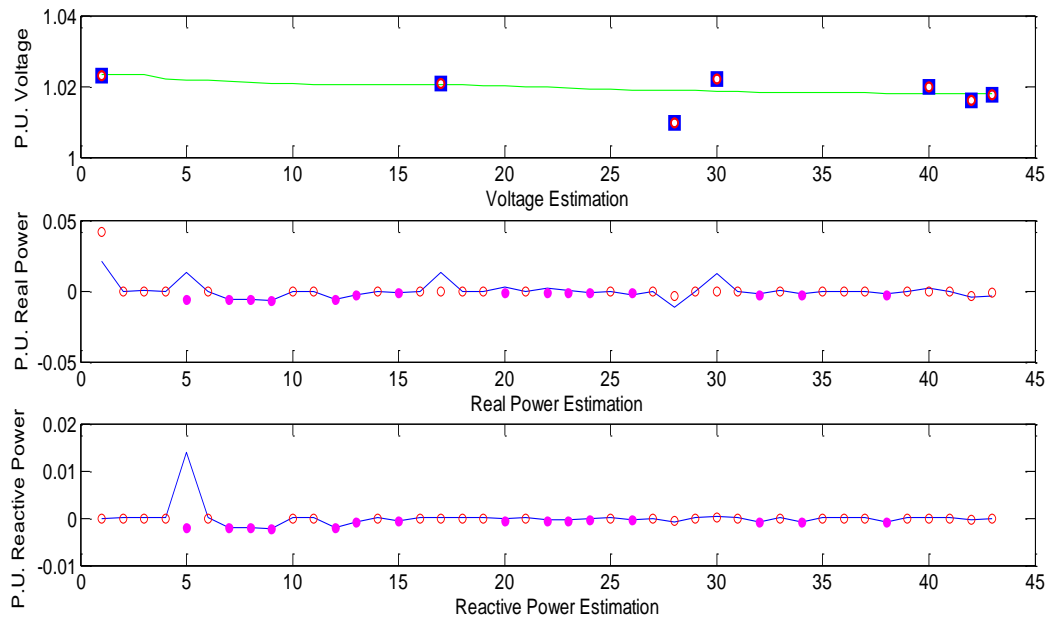
As the true values of the states are not known in practice, the quality of estimation is measured based on how accurately the estimated values match the measured values. In the plots illustrated in this section, a square shape (both blue and red) presents the voltage measurement as DSSE inputs, red circle shapes are all available real measurements, magenta dots are all available pseudo-measurements. In a few cases, all available voltage measurements are shown on the plot however they may not be considered as DSSE inputs e.g. in the case study when QM voltage sensor data are not fed into the estimator. In other cases, since all voltage sensor data are exploited by the estimator, therefore either the red square or the blue square overlapped by the red circle are visible in the graphical representations. All available real and reactive powers are fed into estimator in four case studies. Green lines represent voltage estimation, blue lines for real and reactive power estimation.

#### 4.3.1.1 Feeder: Sencur

This is a 43 nodes 42 branches network. The nodes and branches are numbered according to node index given in appendix 5. DSSE for the feeder is provided with voltage and power measurements from one SCADA and six QMs datasets. The feeder is connected with 11 solar DGs with power generation capacity upto 240 kVA (appendix 6).

**All measurements**

Fig. 4. 25 shows estimated values using all SCADA and QM measurements for one instance. Fig. 4. 26 shows percentage deviation of the voltage and power estimation with respect to real and pseudo-measurement values. The voltage estimation errors remains within satisfactory thresholds as the maximum error value is less than 1%; while the real and reactive power estimation values are of significantly lower qualities.



○ Data available from sensors    □ Input voltage measurement data    ● Input Pseudo-measurement data  
 — Voltage estimation    — active/reactive power estimation

Fig. 4. 25: Voltage and power estimation of Sencur feeder with all SCADA and QM data

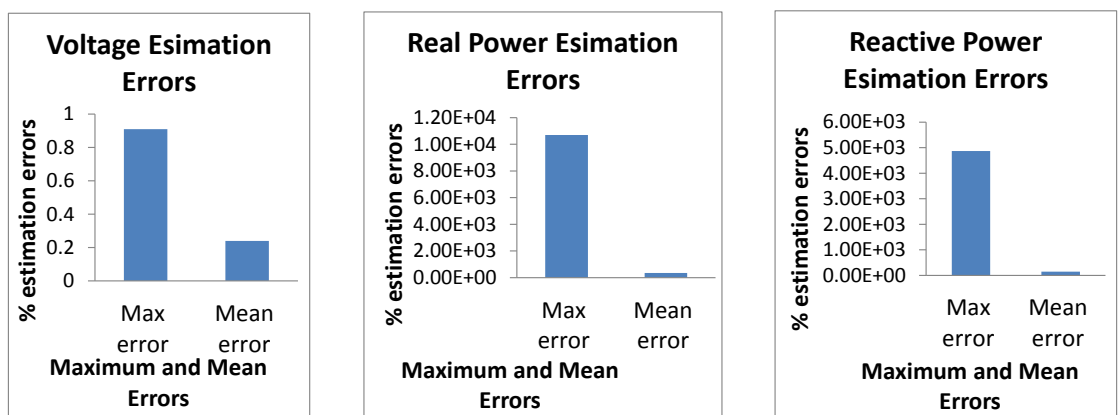


Fig. 4. 26: Voltage and power estimation errors of Sencur feeder with all SCADA and QM data

**SCADA V, SCADA and QM PQ**

Fig. 4. 27 shows estimated values using all SCADA and QM power measurements for the same time instance. Given that only SCADA voltage measurements are provided to the estimator, one voltage measurement value is in a square box overlapped by a red circle and the rest are represented by only red circles indicating the available yet unused QM measurements data in Fig. 4. 27. Fig. 4. 28 presents the voltage and power estimation deviation with respect to real and pseudo-measurement values. Similar to the previous case study, the voltage estimation retains good quality while significant errors are visible in real and reactive power estimation.

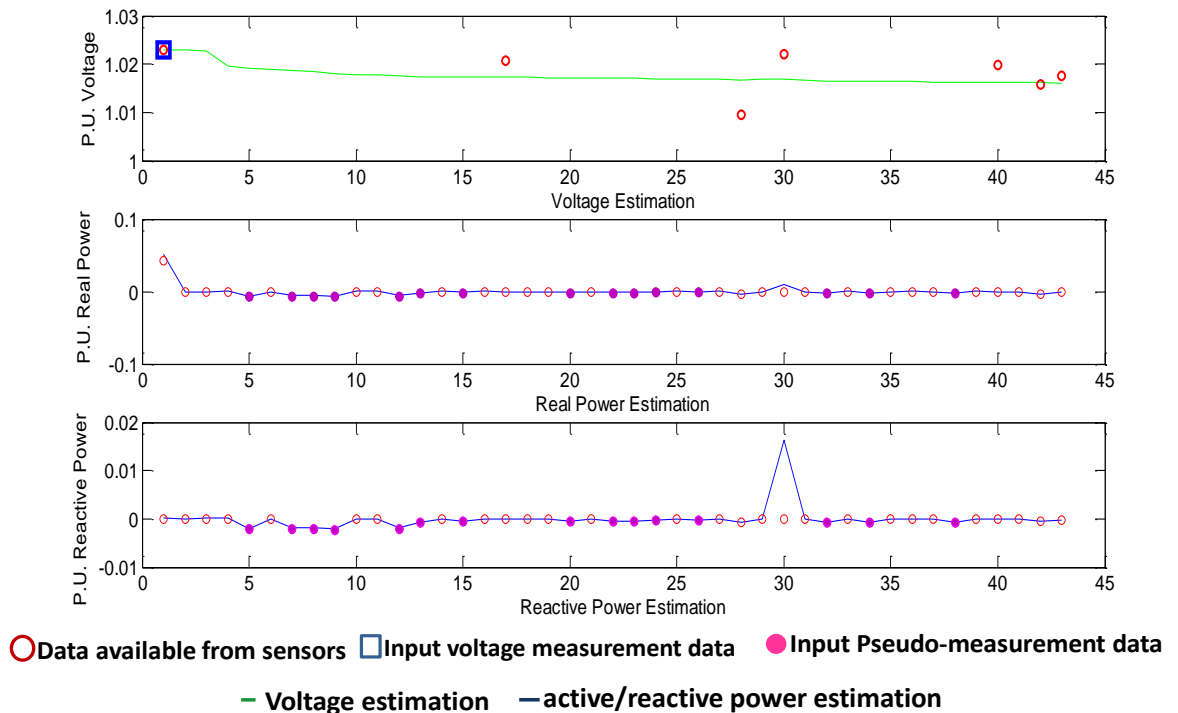


Fig. 4. 27: Voltage and power estimation of Sencur feeder with all SCADA and only power QM data

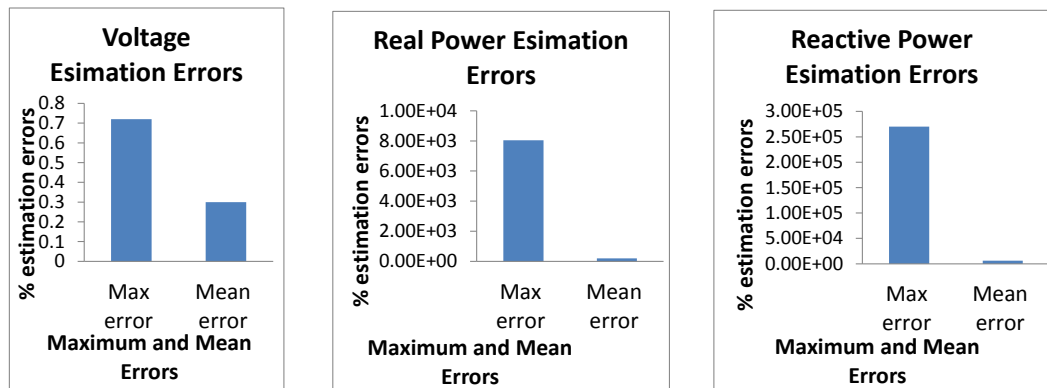
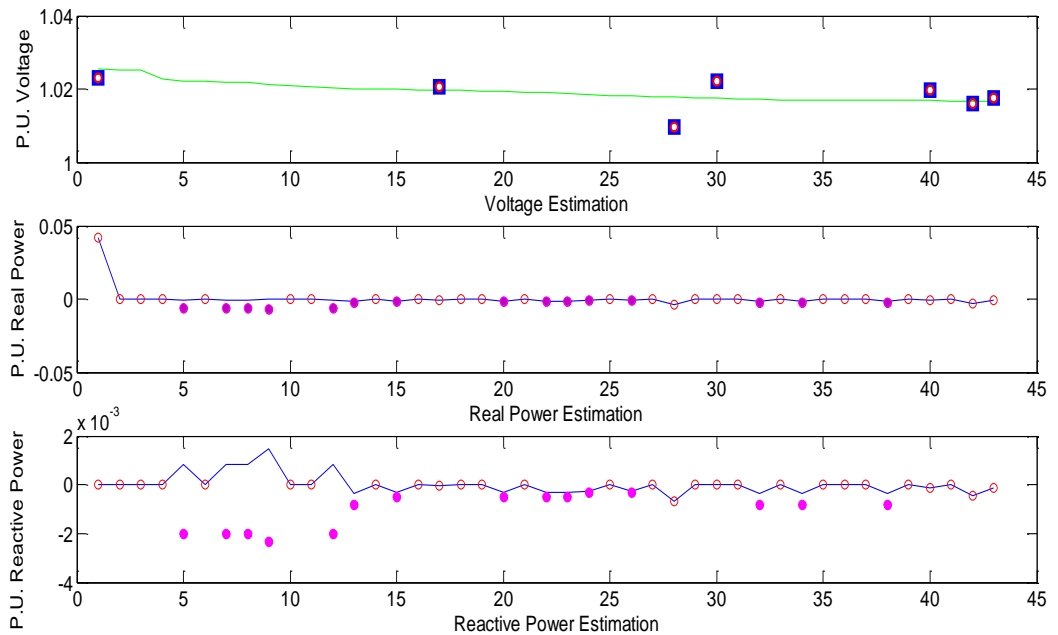


Fig. 4. 28: Voltage and power estimation errors of Sencur feeder with all SCADA and only power QM data

**Assuming real measurement as constraints**

All available measurements are provided to the estimation tool in this case for the same time instance. In addition to virtual measurements, the real and reactive power measurements (excluding pseudo-measurement data) are also considered as equality constraints. The estimated values and estimation deviation are plotted in Fig. 4. 29 and Fig. 4. 30 respectively. Comparing to Fig. 4. 25 and Fig.4. 27, the real and reactive power profiles are following more closely the measured values (Fig. 4. 29) leading to reduced maximum power estimation errors to some degree in Fig 4. 30.



○ Data available from sensors   □ Input voltage measurement data   ● Input Pseudo-measurement data  
 — Voltage estimation   — active/reactive power estimation

Fig. 4. 29: Voltage and power estimation of Sencur feeder assuming real measurement as constraints

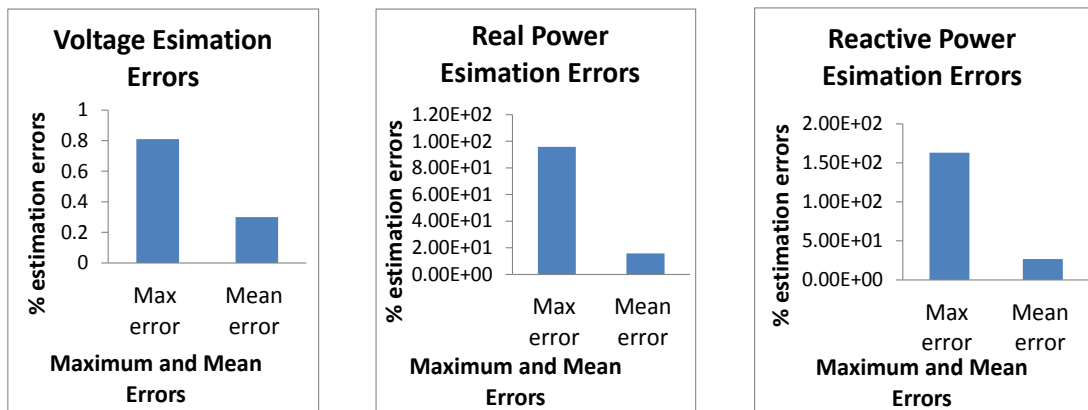


Fig. 4. 30: Voltage and power estimation errors of Sencur feeder assuming real measurement as constraints

### Voltage estimation for ten sets of real meter data

Fig. 4. 31 shows voltage estimation for 10 readings starting from 01/02/2012 00:00 hours to 01/02/2012 02:15 hours in 15 minutes interval. Transformer voltage drops are calculated and added to the measured values in each case as discussed in section 4.3.2. Fig. 4. 32 illustrates the percentage of voltage estimation deviation with respect to the measured values. The colour bars and the associated numbers represent the nodes having voltage sensor data in Fig. 4. 32.

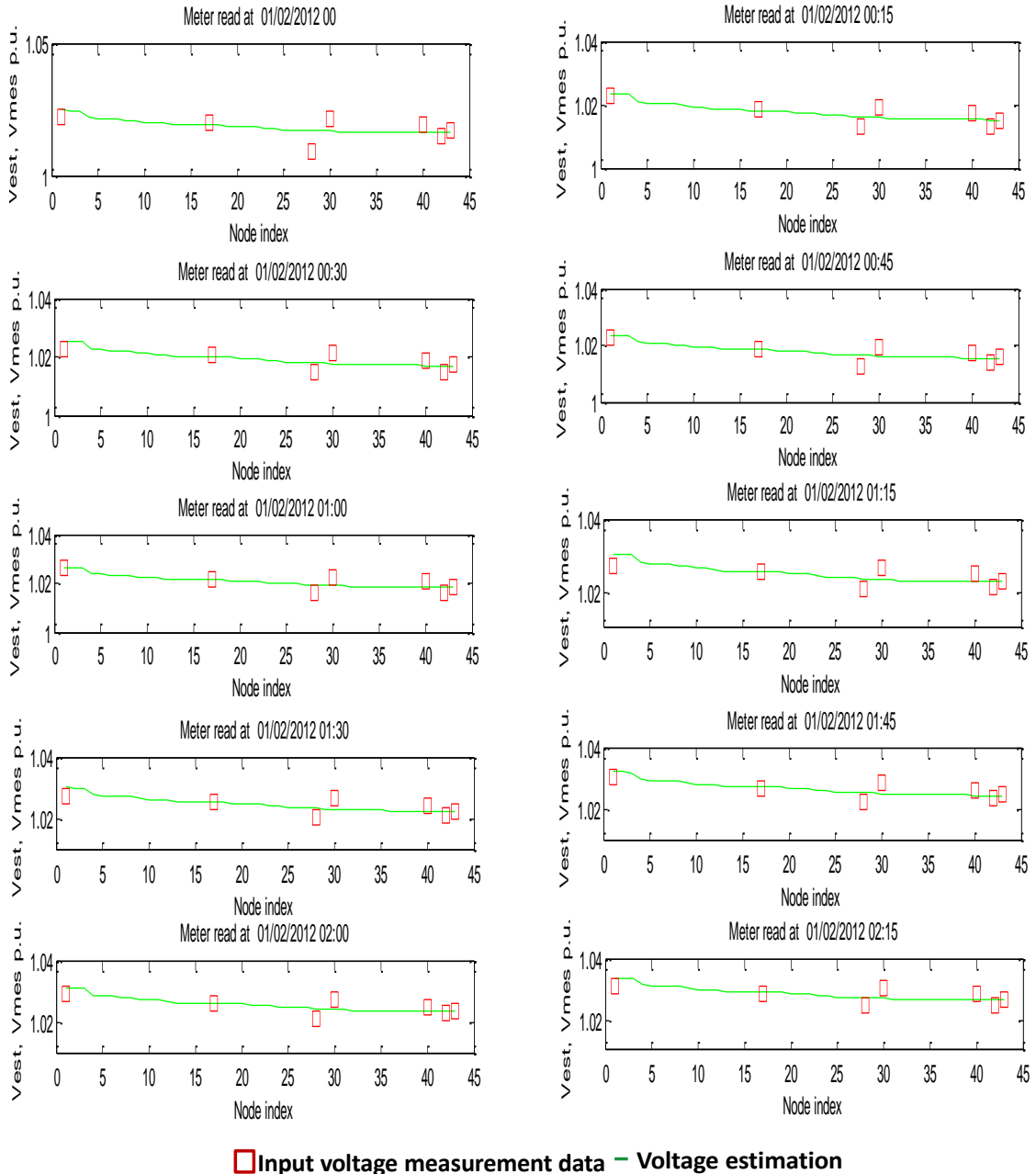


Fig. 4. 31: Voltage estimation ( $V_{est}$ ) and measurements ( $V_{mes}$ ) of Sencur for 10 sets of readings

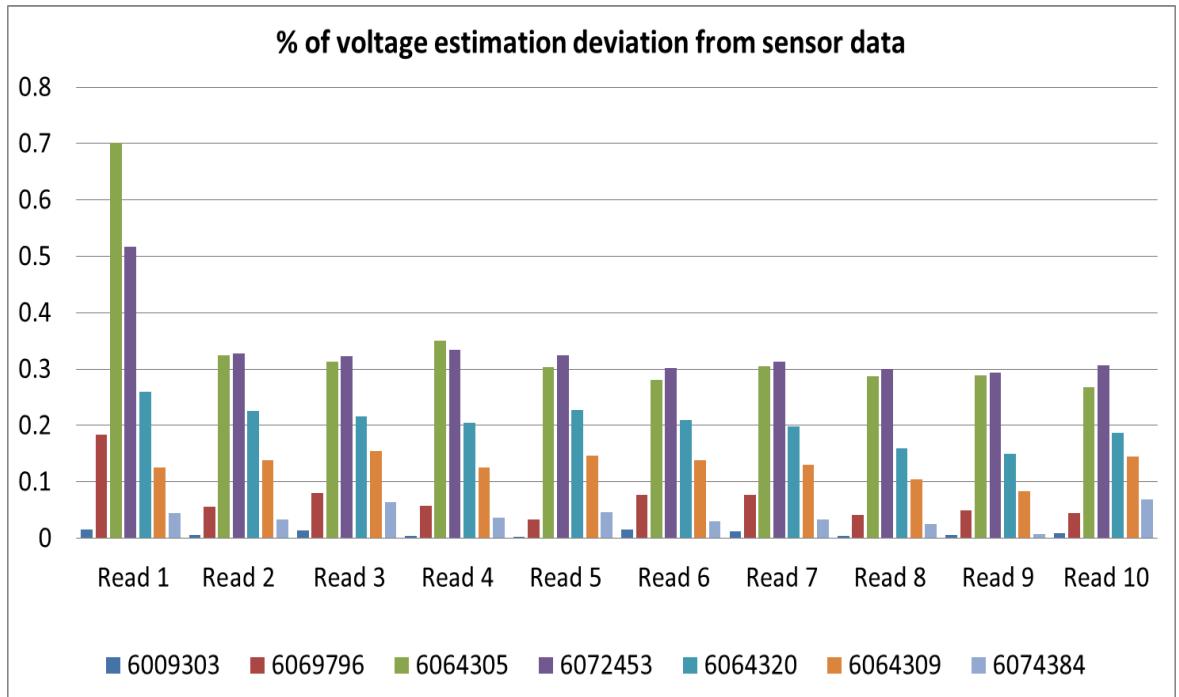


Fig. 4. 32: % of deviation from measurement values (Sencur)

#### 4.3.1.2 Feeder: Cerklje

This is a 40 nodes 39 branches network. The nodes and branches are renumbered according to node index given in appendix 5. DSSE for the feeder is provided with voltage and power measurements from one SCADA and three QMs data. The feeder is connected with 2 small DGs (hydro power plants of 40 kVA and 95 kVA) as shown in appendix 6.

#### *All measurements*

Fig. 4. 33 shows estimated voltage, real and reactive power values using all SCADA and QM measurement for one time instance. Fig. 4. 34 presents the voltage and power estimation deviation with respect to real and pseudo-measurement values. The maximum voltage estimation error is observed to be around 1%, however the maximum real and reactive power estimation errors have significant values.



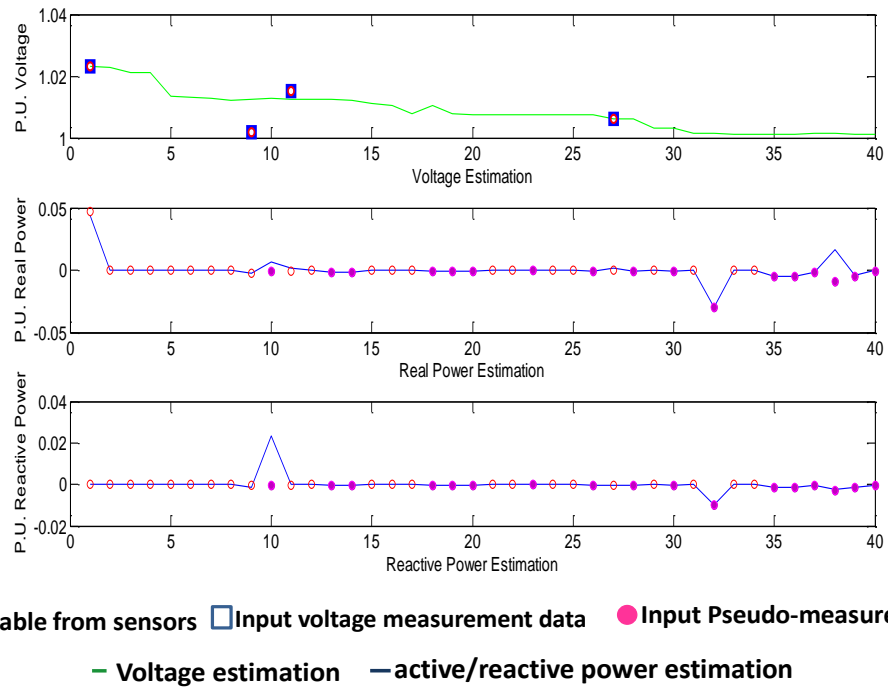


Fig. 4. 33 : Voltage and power estimation of Cerklje feeder with all SCADA and QM data

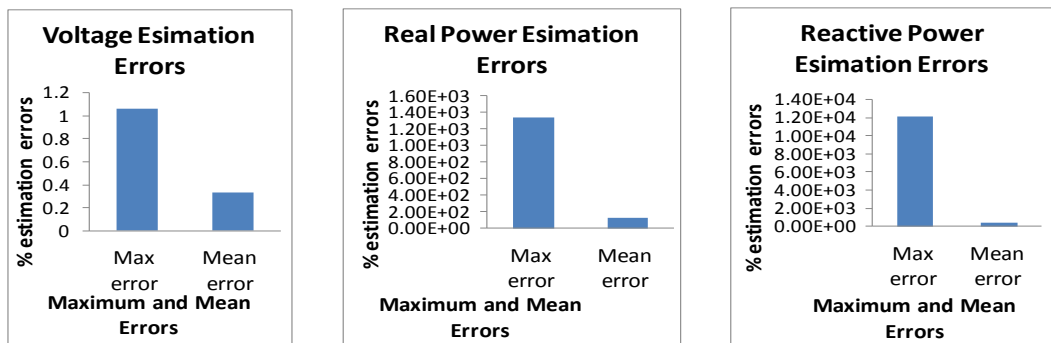


Fig. 4. 34 : Voltage and power estimation errors of Cerklje feeder with all SCADA and QM data

### SCADA V, SCADA and QM PQ

Fig. 4. 35 shows estimated values using all SCADA and QM power measurements for the same time instance. Similar to equivalent case studies for Sencur feeder, only SCADA voltage measurements value is in a square box overlapped by red circles while the rest of the voltage measurements are represented by only red circles in Fig.4.35. Fig. 4. 36 presents the voltage and power estimation deviation with respect to real and pseudo-measurement values. The voltage estimation values are quite satisfactory as they remain below 1% error margin.

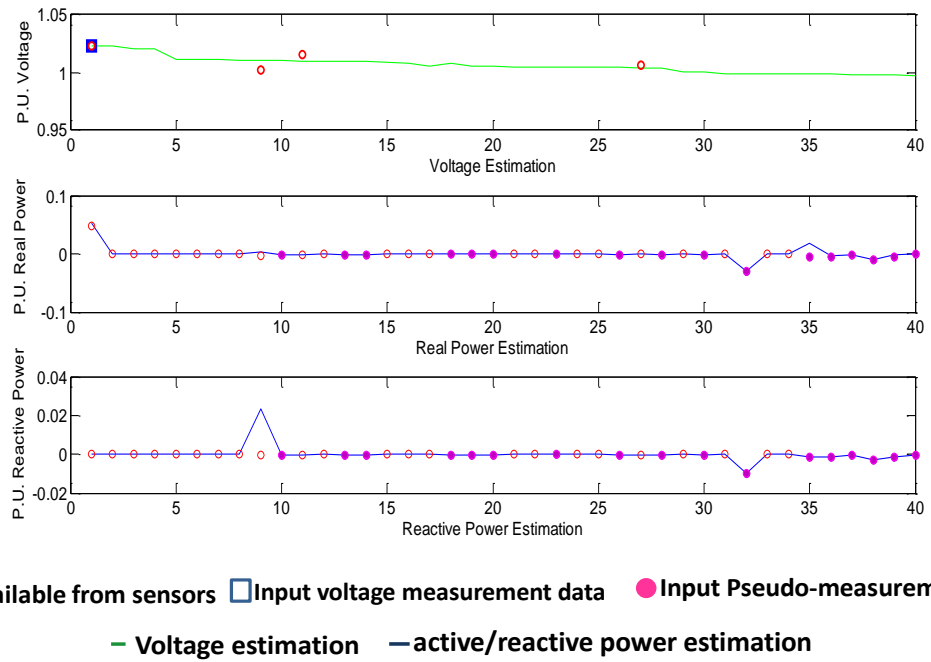


Fig. 4. 35: Voltage and power of Cerklje feeder with all SCADA and only power QM data

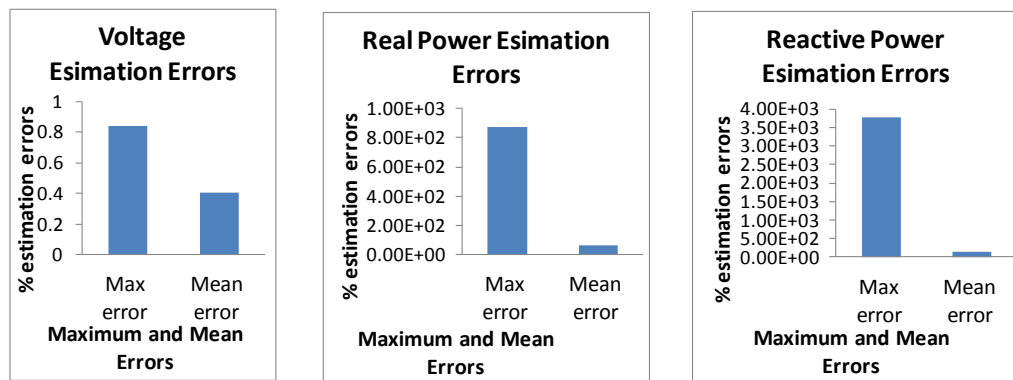


Fig. 4. 36: Voltage and power estimation of Cerklje feeder with all SCADA and only power QM data

### *Assuming real measurement as constraints*

All available measurements are fed into the estimation tool for the same time instance in this case study. In addition to virtual measurements, real and reactive power measurements provided by SCADA and QM devices are also considered as equality constraints. Graphical representation of estimated voltage and power profiles are provided in Fig. 4. 37; the corresponding deviation values are plotted in Fig. 4. 38. As seen in the counterparts of Sencur feeder case studies, taking all sensor provided power measurements as constraints apparently improves the real and reactive power estimation to some degree while retaining good voltage estimation.

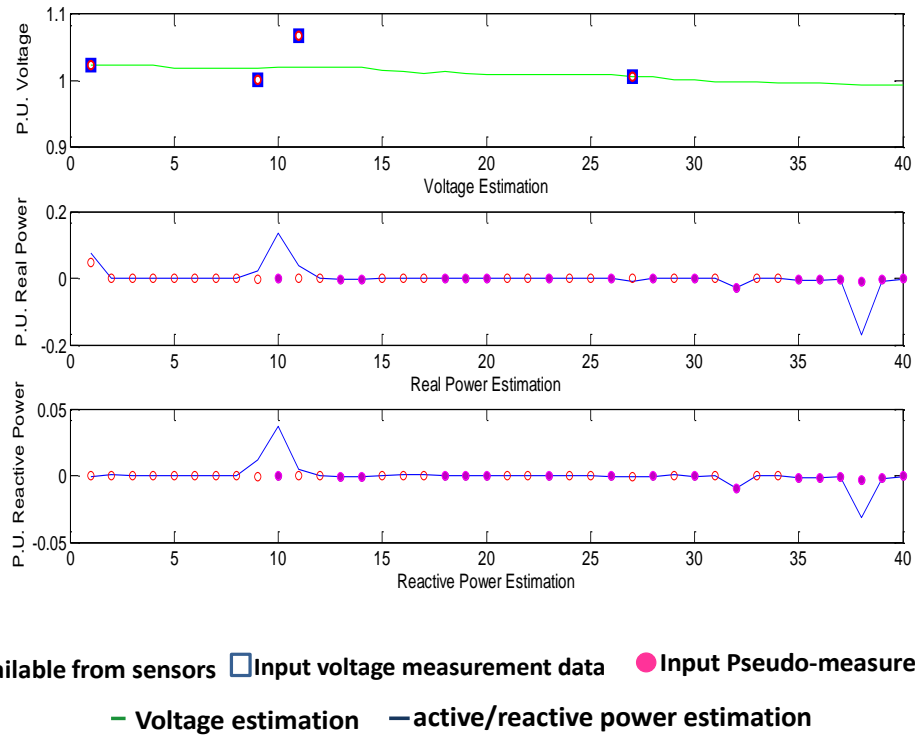


Fig. 4. 37: Voltage and power estimation of Cerklje feeder assuming real measurement as constraints

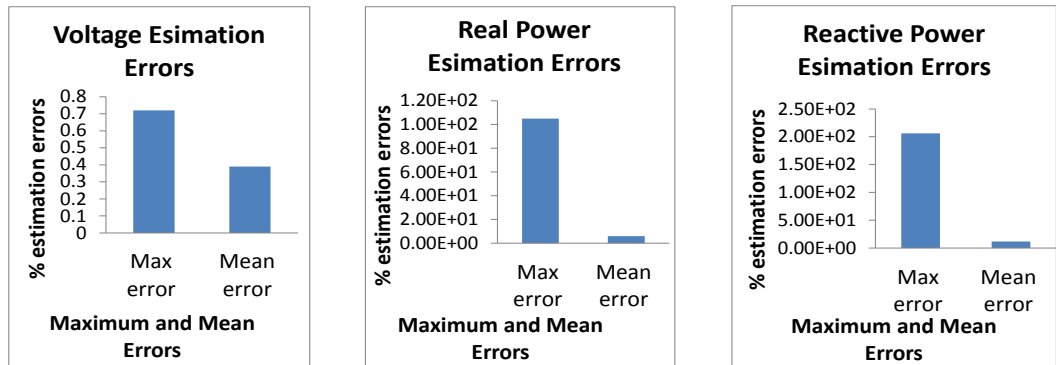


Fig. 4. 38: Voltage and power estimation errors of Cerklje feeder assuming real measurement as constraints

### ***Voltage estimation for ten sets of real meter data***

Fig. 4. 39 shows estimated voltage profiles for 10 sets of meter readings for 10 time instances starting from 01/02/2012 00:00 hours to 01/02/2012 02:15 hours in 15 minutes interval. The transformer voltage drops are added to LV transformer voltage measurements in all ten sets. Fig. 4. 40 illustrates the corresponding deviation of estimated values from the measured values where each colour and associated numbers represent the node connected to the voltage sensor.

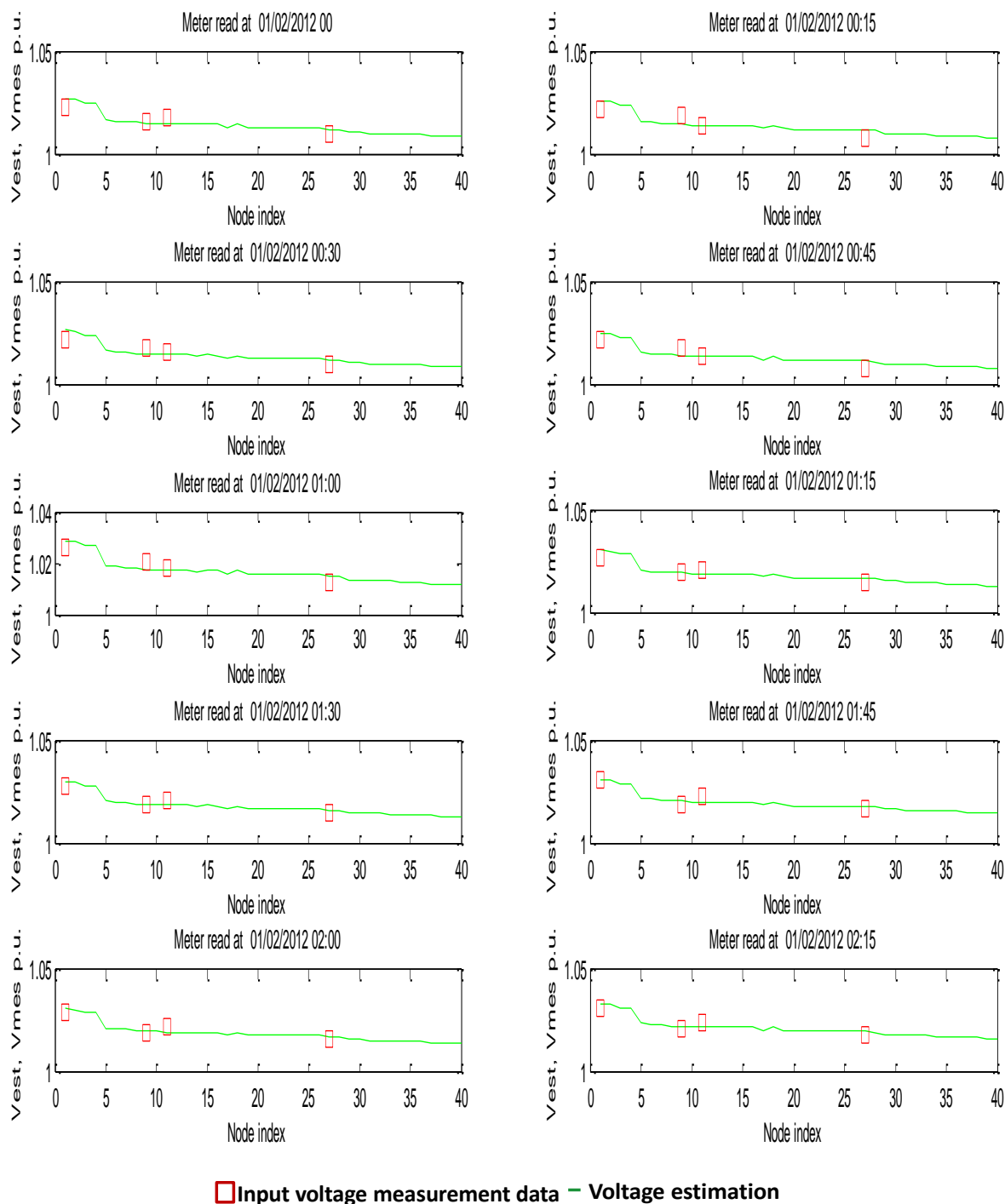


Fig. 4. 39: Voltage estimation ( $V_{est}$ ) and measurements ( $V_{mes}$ ) of Cerklje for 10 sets of readings

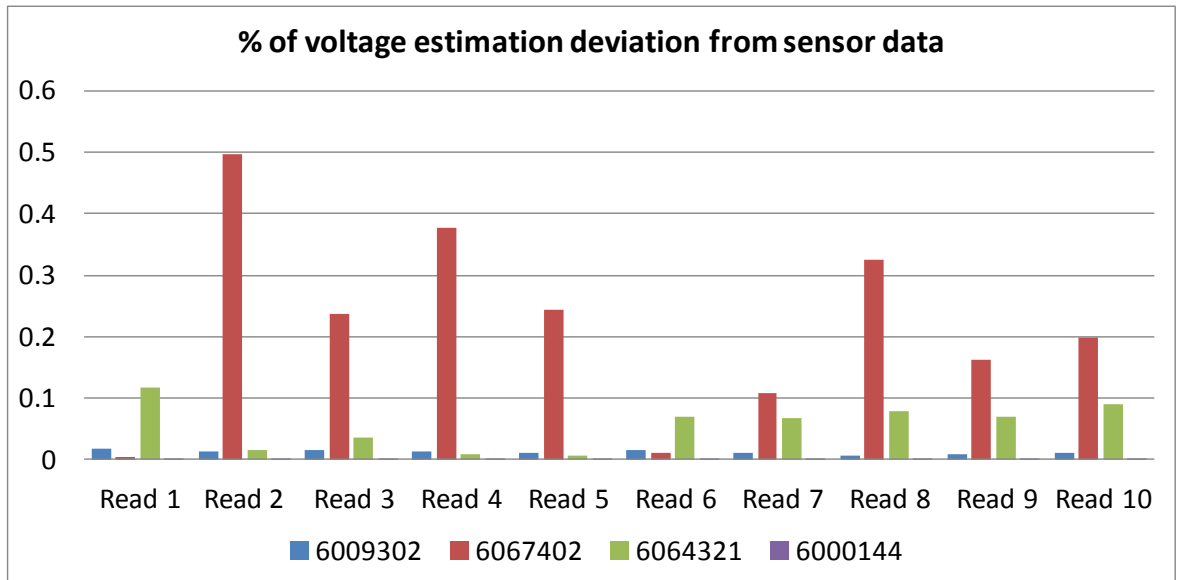


Fig. 4. 40: % of deviation from measurement values (Cerklje)

#### 4.3.4 Analysis of Test Results

In this section, application of Hachtel’s augmented matrix method on real network data is observed. Two EG feeders are selected for studies. In all cases, the estimator strongly matches SCADA voltage data and produces rather a compromise value for all other QM voltage data. Studies include how estimation quality is affected, considering all voltage and all power measurements, only SCADA voltage and all power measurements and all measurements but constrained real power measurements. One interesting observation is the voltage estimation quality worsens slightly in presence of many voltage measurements (Fig. 4. 26 and Fig. 4. 34) than availability of only SCADA voltage measurement (Fig. 4. 28 and Fig. 4. 36 ). As discussed earlier, voltage sensor data has quite a strong effect on overall voltage estimation qualities, in the presence of errors in multiple voltage measurements the estimation quality can be adversely affected. On the contrary, when there is only one voltage measurement accompanied by a number of quite accurate power measurements, the estimates are more coherent.

However, using real power measurements as constraints, an improvement in both voltage and power estimation quality is observed in Fig. 4. 30 and Fig. 4. 38. As for EG network case studies, the real power measurement errors are mostly within 0.5%, if SCADA measurements also provide good data, the incoherent data from pseudo-measurements can be detected as found in Fig. 4. 29 and Fig. 4. 37. In those cases, the power estimates closely match the real measurements while some pseudo-measurements fall completely out

of the network power injection profile. These injection powers are very much incoherent to the real power measurement values. This infers that when there are small numbers of real measurements available compare to the network size, enforcing them as constrained values may be beneficial provided the real measurement values are subject to greater accuracy.

Further studies are performed up to ten sets of meter reading data focusing on voltage estimation using only virtual power measurement as constraints. In most of the cases in Fig. 4. 32 and Fig. 4. 40, the gaps between estimated and measured values are less than 0.5% which can be considered as a good match. Observing ten voltage profiles considering ten sets of measurements in different time tags (Fig. 4. 31 and Fig. 4. 39), the overall voltage profile of the network looks consistent and coherent to the measured values. The observation of power injection values are excluded from the ten meter reading studies as considerable errors are obtained in other case studies.

Albeit study shows that using real measurement also as constrained can bring benefits to the overall estimation quality, this holds true only when the difference between measured and true power values are very small. As SCADA power measurements are still subject to greater errors (up to around 1%), this may lead to bad estimation when significant noise is added. The estimator is then forced to trust the bad data at the cost of ignoring good measurements in such cases. Considering this, for EG network scenarios, use of real power injection measurements as constrained is not recommended. It is also proposed to apply Hachtel's method with real and virtual injection measurement as constraints for overall estimation accuracy, when there is a limited supply of real measurements for which the expected noise levels are low.

#### **4.4 Concluding Remarks and Discussion**

Rigorous assessment of five SE solution processes as discussed in Chapter 3 applying simulated network data has been demonstrated in this chapter. One of these optimization methods is selected as the most potential DSSE tool based on the assessment outcomes. The selected DSSE tool is then applied offline on a real network using real datasets. The chapter performs extensive multi-dimensional studies, which can be divided into three major parts.

- Evaluation of five candidate DSSE tools
- Selection of Hachtel's Method as the most potentially useful DSSE

- Application of Hachtel's method on real data

The major challenges of application of real data to DSSE is the high possibility of incomplete or erroneous data that impedes the estimator in producing quality outcomes. The proposed DSSE candidate tools, as discussed in chapter 3, are therefore examined in five scenario studies taking into account the ideal conditions as well as existence of erroneous inputs. The scenarios are simulated to replicate possible network states in order to observe how the primary states and convergence are affected. Rigorous Monte Carlo study is performed to quantify the average effect of the presence of significantly erroneous pseudo data on two UKGDS model networks. The advantage of using a model network to perform such benchmark tests is that the true values are known (these cannot be known in practical application) therefore performance verification can be executed more confidently, by inspecting relative errors in the estimated data. The investigation outcomes along with the hypothetical characteristics of the algorithms are taken into account to select one out of five methods as the potential DSSE tool. Both WLS and WEM requires the least time for execution in an ideal condition however the number of iterations needed increases quite significantly once the network sizes or states change. The QR method, in spite of providing robust performance, has not been considered as a strong DSSE candidate due to the required operation time. The WEM method has proved to be the weakest candidate of all since it generates deteriorated quality of estimation in a many cases. Hachtel's Augmented Matrix method is eventually chosen as the most useful potential DSSE tool, since it has shown overall acceptable performance with regard to estimation quality, robustness and execution time consistently in the presented case studies.

The selection of Hachtel's method is followed by its application to off line field datasets of Slovenian distribution networks provided by DNO EG. Two small feeders, Sencur and Cerklje are selected for case studies, both of which are equipped with several DGs. The major drawbacks confronted in order to apply EG data is the deficiency of well defined pseudo-measurement data for the unmeasured nodes. The estimated voltage values still provide good matches with the respective measurements in all cases. The quality of the real and reactive power estimation is consistent with the quantity and quality of the input data, however, at this stage they are infeasible as inputs to the control functions. Improvement in input data is an essential requirement to realize the potential of DSSE in the future. The graphical representations of deviation of estimated values from the real values for 10 meter reading instances provides consistently acceptable voltage estimation by Hachtel's method.

In addition to virtual measurements, SCADA and QM provided real and reactive power measurements are also exploited as equality constraints in a few case studies to observe the effect of imposing enhanced emphasis on real power data. Somewhat better quality estimations are achieved by applying constrained real measurements. However, the idea of constrained real measurements is subject to further investigation and therefore, has not been applied further due to the likelihood of significant risk in the presence of bad data.



# **CHAPTER 5**

## **DEVELOPMENT OF SCALABLE DISTRIBUTION SYSTEM STATE ESTIMATION**

This chapter presents research on the development of DSSE tools executable in scalable fashion. Two approaches are investigated in order to verify their feasibility as scalable tools. These are Differential Evolution Algorithm (DEA) and Overlapping Zone Approach (OZA). In addition to that, a particular case study is performed by combining both DEA and OZA in a single process. The algorithms are initially applied to a small network and subsequently on larger networks, where satisfactory performance was achieved in the initial investigations. The development of scalable tools at this stage is strongly focused on the improvement of voltage estimation quality.

### **5.1 Choices of Scalable DSSE**

The computation time of DSSE becomes a crucial issue when the network creates a considerable amount of SE input data and the estimation process is required to be completed in a limited time window. A study of computation time needed for three DSSE candidate solution processes (WLS, WEM, Hachtel's) generates the following outcomes in Fig. 5.1. The WLS method is the classical approach, Hachtel's method applies constrained virtual measurements and the WEM method proposes a novel iteratively re-weighting approach. These three methods represent diverse SE solution processes and their 'execution time versus network size' relation can be analysed.

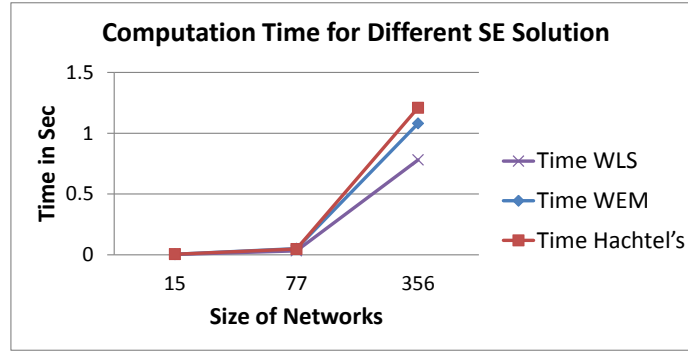


Fig. 5. 1: Computation time with increased network sizes for various SE solutions

The relation between the SE computation time ( $T$ ) to the size of the network can be expressed in the following equation where  $N$  is the number of nodes,  $c$  is a constant number and  $n$  is the degree of order to which network size relates computation time.

$$T = N^n + c$$

Taking  $\log$  on the both side of above equation

$$\log_{10}(T) = \log_{10}(N^n) + \log_{10}(c)$$

i.e.

$$\log_{10}(T) = n \log_{10}(N) + \log_{10}(c) \quad (5.1)$$

The above equation can be compared with the linear equation which is conventionally expressed as  $y = mx + c$ . Comparing (5.1) with the conventional linear equation,  $n$  can be considered equivalent to the slope of the linear problem consisting of two variables:  $\log_{10}(T)$  and  $\log_{10}(N)$ . Applying this notion, the following chart shows the calculated value of slopes that determines to which degree of order the computation time relates the size of the network. Table 5. 1 refers the values for the three SE processes illustrated in Fig. 5. 1.

Calculation between networks sizes of	Slope WLS	Slope WEM	Slope HACHTEL's
15 Node and 77 Node	1.40765872	1.55955269	1.478356333
356 Node and 77 Node	2.08575875	2.00682028	2.124998671
15 Node and 356 Node	1.73550693	1.77579789	1.790995361
<b>Average Slope Values</b>	1.7429748	1.780723619	1.798116788

Table 5. 1: Calculation of slopes for various SE solutions

As shown in Table 5. 1, the computation time of most of the observed SE optimization method increases in a quadratic manner with the size of the network, demonstrating a weak scalable property. This becomes an extreme challenge when a centralized SE is executed

for a very large network. Alternative approaches are required for the scalable solution of the DSSE problem while considering input data generated by a distribution system comprising a substantial number of nodes. In this research, two options are chosen to investigate their potential as scalable DSSE solution processes, these are:

- **Differential Evolution Algorithm (DEA):** The DEA is a derivative-free heuristic optimization process inspired by genetic algorithms (GA). Similar to GA, DEA performs optimization based on generation and evolution of a set of population members. DEA therefore inherently retains the property of applicability to parallel processing in a multi-core/multi-processor system.
- **Overlapping Zone Approach (OZA):** A novel network division based DSSE approach applicable in scalable fashion is introduced. The network is split into a number of zones which have common nodes with other zones. The algorithm is therefore referred to as the ‘Overlapping Zone Approach’ (OZA).

## **5.2 Differential Evolution Algorithm (DEA)**

‘Differential Evolution Algorithm (DEA)’ is a meta-heuristic based stochastic optimization method first proposed by Storn and Price in 1995 [58] [59] [60]. DEA is a derivative-free direct search algorithm based on ‘generate-and-test’ and possesses inherent ‘scalability’. The algorithm is inspired by biological motivations and is capable of generating optimal solution on rough, discontinuous and multi-modal surfaces requiring no gradient information of the function to be optimized. The DEA uses only primitive mathematical operators and is simple to implement, yet has been proven as an effective optimizer [58] [60]. The application of this emerging state-of-art optimization tool has been investigated in a wide range of real life problems. With regard to some other common heuristic methods, DEA outperforms evolutionary algorithms such as classical Evolutionary Algorithms (EA), Simulated Annealing (SA) and Particle Swarm Optimization (PSO), in solving optimization problems [58] [61] [62] [64]. DEA is robust, able to deliver consistent results and has few parameters to set. Therefore, DEA is considered as one of the greatest potential optimization tools among all kinds of evolutionary algorithms. Hence, DEA is selected for investigation as a potential scalable solution for the DSSE tool.

One added advantage of choosing DEA, in application to DSSE, is that it can still operate even if there are some measurements missing due to maintenance or malfunction of the instruments, giving rise to unobservability. Since the DEA is a derivative-free process, it does not require to generate the gain matrix, which is prone to ill-conditioning. It also can adopt different optimization objective functions such as WLS, WEM, Weighted Least Absolute Values (WLAV), without requiring any major software modifications. The major drawbacks could be the communication bottleneck and computation expenses, which are expected to be resolved by the promising development of emerging scalable High Performance Computation Platforms (HPC) and novel Information and Communications Technology (ICT) solutions. References [64][65] [66] [67] [68] reveal promising outcomes in parallelization modelling of DEA in terms of efficient reduction of computation time with enhanced population search spaces for improved solutions. Adopting the state-of-art technology is expected to enable successful application of DEA as a potential scalable tool.

### 5.2.1 Basic Formulation

The underlying operational steps of DEA are very similar to many other evolutionary and genetic algorithms. The algorithm evolves a fixed population containing  $P$  sets of vectors. Each set of  $P$  is an  $N$  dimensional decision parameter vector,  $D$  which can be referred to as the state vector in the power system state estimation problem. The initial set of population is the first generation member. This population evolves into the next generation through mutation and cross-over processes to gradually move towards a fitter candidate solution [58] [60] [58] [69] [70]. The fitness of candidate vectors is evaluated by objective function,  $f$ , often referred to as the cost function that is to be minimized. The overall functional steps of DEA [71] [72] are shown in Fig. 5. 2 and explained as follows.

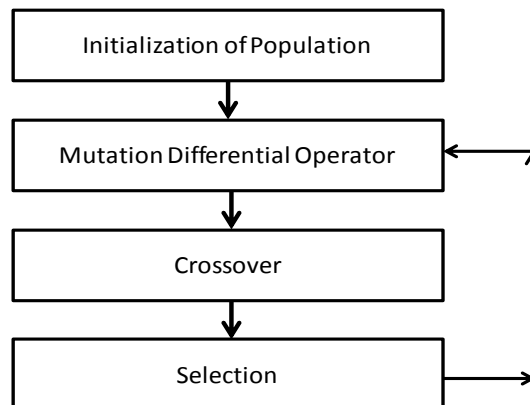


Fig. 5. 2: The DEA functional steps

*Population:* Each generation,  $G$  contains a  $P \times N$  dimensional population matrix. The decision vector is indexed as  $D_i$  for  $i^{th}$  element of the current population.  $D_i$  consists of  $N$  number of members that constitute cost functions,  $f: D \subset \mathbb{R}^N \rightarrow \mathbb{R}$  [75] [75]. If vector members are expressed as  $D_{i,j}$  where  $j$  is the vector member index ( $j = 1, 2, \dots, N$ ) and  $i$  is the population index ( $i = 1, 2, \dots, P$ ), therefore  $D_i = [D_{i,1}, D_{i,2}, \dots, D_{i,N}]$ . A complete set of population,  $P_G$  for a generation  $G$  forms the following matrix

$$P_G = \begin{bmatrix} D_{1,1} & D_{1,2} & \dots & D_{1,N} \\ D_{2,1} & D_{2,2} & \vdots & D_{2,N} \\ \vdots & \vdots & \vdots & \vdots \\ D_{P,1} & D_{P,2} & \dots & D_{P,N} \end{bmatrix} \quad (5.2)$$

DEA generates a randomly distributed initial population,  $P_{G=0}$  within the decision space which is generally defined within parameter bounds. The initial assumption for each member of  $P_{G=0}$  is calculated from the following formula.

$$D_{i,j} = D_{i,j}^{lo} + rand[0,1].(D_{i,j}^{hi} - D_{i,j}^{lo}) \quad (5.3)$$

The term  $rand[0,1]$  is used to state uniformly distributed random variables ranging from zero to one. The subscript  $hi$  and  $lo$  represent the upper and lower limit of parameter bounds respectively [75].

*Mutation:* The mutation process is controlled by mutation parameter or step size  $F$ . The mutant vector set  $D_{M,j}$  is obtained by taking the weighted differences of equivalent members of other vectors in the population as stated in the following the formula.

$$D_{M,j} = D_{X,j} + F \times (D_{Y,j} - D_{Z,j}) \quad (5.4)$$

Here index  $X, Y, Z$  are three different sets of array of  $P$  number of population in (5.2) and their relation must satisfies  $X \neq Y \neq Z$ . The subscript  $M$  refers mutant vector index. The mutation process corresponding to (5.4) for a two dimensional cost function, showing its contour lines, is illustrated as follows.

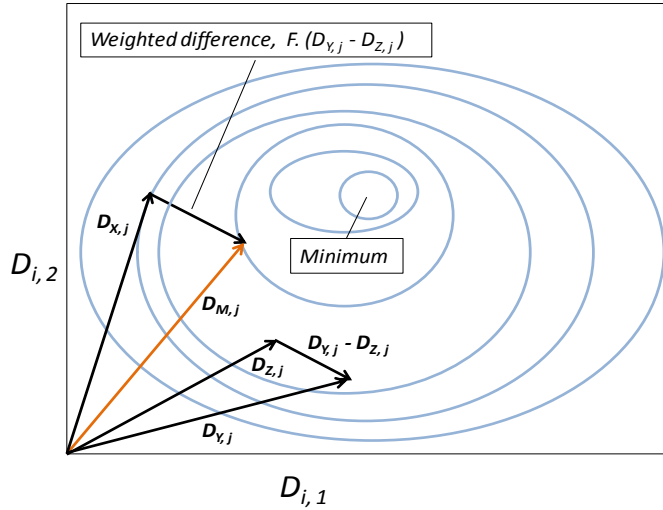


Fig. 5. 3: Mutation mechanism in a two dimensional variable space

The difference vector and the resulting perturbation become smaller when the difference between vectors of the parent population reduces. Thus, as the population gets closer to the optimum point, the step length also automatically decreases enabling a more fine-tuned search. This makes DEA a self-adaptive evolutionary approach.

*Cross-over*: The cross-over operation is performed after the mutation and it is controlled by setting a value for the cross-over parameter,  $CR$ . It is completed by mixing the mutant vector  $D_M$  with another vector set from the population of the generation. The new vector set is called the target vector,  $D_T$  and the product after the cross-over operation is called the trial vector,  $D_C$ . The target vector  $D_T$  is such that,  $T \neq X \neq Y \neq Z$ . The cross-over formulation follows a rule as specified in (5.6).

$$\begin{aligned}
 D_{C,j} &= D_{M,j}; & \text{if } RND_{i,j} < CR \\
 &= D_{T,j}; & \text{otherwise}
 \end{aligned} \tag{5.6}$$

The vector  $RND$  in (5.6) consists of random numbers generally developed in such a way that it ensures the cross-over operation occurs at least once in each vector. The subscript  $C$  refers to trial vector notation. The cross-over process for basic DEA variant is shown as below

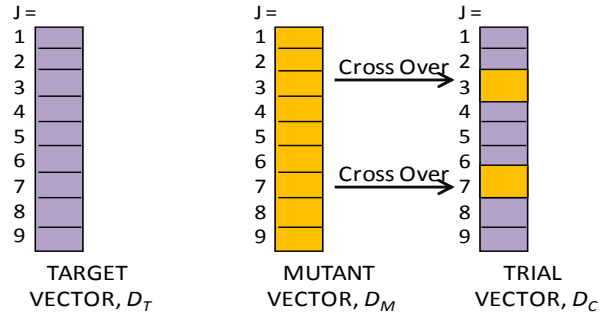


Fig. 5. 4: Basic DEA cross-over process

The cross-over is also referred to as a recombination process. The control parameter  $CR$  essentially acts as mutation probability as it controls the number of components that would be inherited from the mutant vector [75].

*Selection:* Finally, the parameter vector for the next generation ( $G + 1$ ) is chosen in a greedy selection criteria, based on the fitness values of the trial vector  $D_C = [D_{C,1}, D_{C,2} \dots \dots D_{C,N}]$  and the target vector  $D_T = [D_{T,1}, D_{T,2} \dots \dots D_{T,N}]$ . For an optimization problem where the aim is to achieve the minimum value of the cost function,  $f$ , the selection criteria is as follows:

$$\begin{aligned}
 D_i^{G+1} &= D_C; \quad \text{if } f(D_C) < f(D_T) \\
 &= D_T; \quad \text{otherwise}
 \end{aligned} \tag{5.7}$$

Here  $D_i^{G+1}$  is the new vector set for the population,  $P_{G+1}$  of the next generation ( $G + 1$ ). The trial vector is selected for the next generation if the cost value is smaller than the target vector, otherwise the target vector again serves as the parent vector in the next generation. To summarize, each vector set of a population has to go through the same process of mutation, cross-over and finally selection based on cost function evaluation. At the end of a number of population equivalent cost function evaluations, the next generation population is obtained. The evolution operation continues until the target value or the maximum number of iterations is reached. Finally the vector set that results in the lowest cost value is selected as the best and the final solution,  $D_{i_{best}}^{G_{max}}$ .

### 5.2.2 Variants of DEA

The basic DEA variant, expressed as  $DE\backslashrand\backslash1\backslashbin$ , performs mutation incorporating the difference between one pair of vectors as stated in (5.4) involving three vectors for

mutation [58] [76]. The general notation of the variants is  $DE\backslash x\backslash y\backslash z$ . There are a few other variants of DEA using different strategies. They are referenced according to the way they perform mutation and cross over operations.

$x$  specifies the vector to be mutated which can be a randomly chosen population vector (*rand*) or the vector of lowest cost from the current population (*best*).  $D_{x,j}$  in (5.4) is a randomly selected member for ( $x: x = rand$ ) and the corresponding element from the best set of vector for ( $x: x = best$ ). A third strategy for  $x$  is *rand – to – best*, which involves two differences from two pairs of vectors in the mutation process. One of the pairs is the corresponding best member ( $D_{i_{best}}^G$ ) from the current population in this strategy.

$y$  represents the number of difference vectors used for mutation. Equation (5.4) uses one difference involving three members of the population, therefore  $y = 1$ . In some optimization problems, involving more members for mutation can be beneficial. For those cases the mutation equation involves five members using two differences. The mutation equation for these cases when  $y = 2$  is as below

$$D_{M,j} = D_{X,j} + F \times (D_{Y,j} - D_{Z,j}) + F \times (D_{V,j} - D_{W,j}) \quad (5.8)$$

Where  $V \neq W \neq X \neq Y \neq Z$ . For strategy  $x = rand - to - best$ ,  $D_{Y,j}$  in (5.8) should be the member of corresponding element of the best vector,  $D_{i_{best}}^G$  in current population i.e.

$$D_{Y,j} = D_{i_{best},j}^G.$$

$z$  denotes the cross-over criteria whether it is binary (*bin*) or exponential (*exp*) characteristics. Following this notation, the basic DE strategy is written as:  $DE\backslash rand\backslash 1\backslash bin$ . To define binomial and exponential cross-over

- Binomial cross-over ( $z = bi$ ): similar to uniform cross-over. It selects individual member of a vector in a random manner. The basic DEA strategy follows binomial cross-over in Fig. 5. 4.
- Exponential cross-over ( $z = exp$ ): similar to one and two point cross-over variant. This process also chooses members of a vector in a random fashion, but it chooses a sequence of consecutive members as shown in Fig. 5. 5.



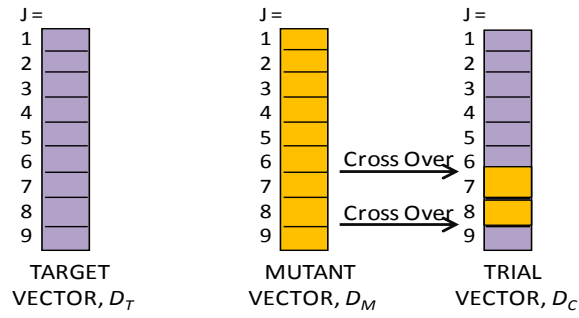


Fig. 5. 5: Exponential cross-over process

The DSSE variants using different strategies are given below

- Strategy 1  $DE\backslash best\backslash 1\backslash exp$ :
- Strategy 2  $DE\backslash rand\backslash 1\backslash exp$
- Strategy 3  $DE\backslash rand - to - best\backslash 1\backslash exp$
- Strategy 4  $DE\backslash best\backslash 2\backslash exp$
- Strategy 5  $DE\backslash rand\backslash 2\backslash exp$
- Strategy 6  $DE\backslash best\backslash 1\backslash bin$
- Strategy 7  $DE\backslash rand\backslash 1\backslash bin$
- Strategy 8  $DE\backslash rand - to - best\backslash 1\backslash bin$
- Strategy 9  $DE\backslash best\backslash 2\backslash bin$
- Strategy 10  $DE\backslash rand\backslash 2\backslash bin$

### 5.2.3 Defining Control Variables and Search Space

The two control parameters  $F$  and  $CR$  influence the behaviour of DEA and the scale of variation of decision parameters from parent to child generations. The step-size or mutation constant  $F$  is normally taken between 0 and 2, indicating the degree of perturbation,  $F \in [0,2]$ . Cross-over variable  $CR$  is chosen from 0 to 1, to define the degree of recombination,  $CR \in [0,1]$ . The values of control parameters for a specific problem are generally obtained by parameter tuning. A smaller value of  $CR$  indicates that fewer members of the mutant vector would be transferred to the trial vector. Too small value of  $CR$  may cause stagnation of decision parameters in a local minima. A larger value of  $CR$  enhances population diversity and faster convergence; however it should not exceed the threshold limit, to avoid premature convergence. The smaller the value of  $F$  causes limited perturbation, which also increases premature convergence due to insignificant variation between two consecutive generations.  $F$  should have sufficient magnitude to counteract

selection pressure; however too large a value is also not recommended, as that requires additional function evaluations, delaying the convergence considerably. In the ideal case, as DEA proceeds to the optimum point, the diversity of population declines and therefore the differences between two vectors reduces. This consequently reduces the step size and variation due to parameter mixing, resulting in a finer grained search criteria. Hence DEA provides self adaptive parameter control strategies [58] [75] [77] [79].

The size of population is also a crucial factor to determine how well the process can achieve global optima. The population size should not be chosen too small as that may lead to premature convergence limiting the search space, also too large a population can make the process sluggish as more function evaluation is required. The size of population largely depends on the characteristics of the problem to which DEA is being applied. Some literature suggests that a population size of  $4D$  to  $8D$  is sufficient for solving an optimization problem [58] [75] [77] [79].

#### 5.2.4 Application of DEA in DSSE

In the application of DEA for DSSE optimization, the primary state vectors, as defined in chapter 3, are selected as the decision vectors. If the number of nodes of a system is  $N_{nd}$ , the number of unknown elements of the vector is  $2 \times N_{nd} - 1$ , that excludes the voltage angle state of the slack node. Here WLS optimization criteria is selected as optimization criteria, the cost function  $f$  is equivalent to WLS objective function,  $J(x)$  in (3.7) to minimize. Here state vector  $x$  is equivalent to decision vector  $D$ .

Preliminary investigation has shown that adopting strategies that involve five vectors of a population and two vector differences (5.8) can improve the estimation quality considerably compare to applying other DEA variants. A randomized mutation parameter  $F$  has been applied to increase perturbation and reduce the possibility to be stagnated in local minima. Therefore, a 'randomized  $F$  - DEA' method that incorporates an extended mutation function and randomized mutation parameter, have been applied to perform DSSE. The randomized scaling factor  $F$  has been implemented for feasibility studies of three DE variants to observe the effects of binary and exponential cross-over schemes. The variants are referred as below:

*DE\rand\1\bin*

*DE\rand\2\bin*

*DE\rand\2\exp*

A good estimator should be able to reduce the estimation error and to recognize the voltage profile of the network through evaluation of the measurement equation based cost function. In the following section, the effectiveness of the three DEA schemes is assessed, based on the quality of voltage estimation achieved, for similar sizes of population and generation. Realizing possible risks regarding computation cost, the initial analysis is performed on a small UKGDS [54] network of 16 nodes. Investigations have been performed using an error scenario of up to around 1% real measurement and 50% pseudo-measurement errors. It is assumed that one voltage, one real and reactive power injection measurements are located in the network at GSP. The true values are obtained from conventional load flow results for a known state of the network. The measurements are simulated value taking true values as the mean in a Gaussian distribution curve (section 3.4.2).

The generation size is kept consistent at 2500 iterations for comparison studies. As it is a 16 node network, the population size is  $8 \times (2 \times 16) = 256$ . All the values are presented in per unit (p.u.).  $F$  is varied between 0.3 to 0.5 in each generation in a random fashion for randomized  $F$  cases and is kept fixed to 0.5 for basic DEA. In all cases  $CR$  is 0.98. The higher and lower boundary are set for voltage magnitude,  $|V| = [1.03, 0.97]$  p.u. and for phase angle,  $|\theta| = [0.02, -0.02]$  radians. If an element of the vector tends to exceed the boundary value, it is forced to remain within the boundary. In such case, the boundary exceeding value is replaced by generating a random value that is still larger than the corresponding element from a parent vector but still within the boundary. Each trial vector is checked to see whether there is any violation of boundary values and correction is executed, if required, before initiating the selection process. The formula for corrections of boundary violation are as below

- For violation of upper boundary,  $D_{C,j} = D_{i,j}^{hi} - rand[0,1].(D_{i,j}^{hi} - D_{i,j})$
- For violation of lower boundary,  $D_{C,j} = D_{i,j}^{lo} + rand[0,1].(D_{i,j} - D_{i,j}^{lo})$

$D_{C,j}$ ,  $D_{i,j}^{hi}$  and  $D_{i,j}^{lo}$  are defined in (5.6) and (5.3),  $D_{i,j}$  is an element of a parent vector participating in the mutation process.

It is intended to observe the feasibility of the algorithm in a small network by scrutinizing the effect on voltage estimation quality, when applying the DEA based DSSE approach

through case studies presented in this section. The parallel segments of the algorithms are actually run serially one after another; therefore the parallelization is performed hypothetically rather than practically in the following test cases. The core DEA tool in MATLAB M files is available online<sup>1</sup> [55] [80].

### 5.2.5 Performance Comparison of Fixed and Randomized $F$ Classical DEA

As observed in background research, classical DEA exerting a randomized scaling factor ( $F$ ) can attain more success in identifying the network voltage profile than that with fixed  $F$  for DSSE problems. Five sets of measurement scenarios have been simulated to observe the impact of the fixed and randomized  $F$  on the quality of voltage estimation in Fig. 5.6. The basic variant  $DE\backslash rand\backslash 1\backslash bin$  is exploited in both cases. Two sub-plots in Fig. 5.6 represent the percentage of voltage estimation errors relative to the true values and the simultaneous true and estimated voltage profile plot. As the plots are created in continuous fashion for five measurement sets, the number of data points for five sets of measurements of the 16 node network is  $16 \times 5 = 80$ . The top sub-plot illustrates the mean error values in addition to the absolute values of voltage estimation error percentage.

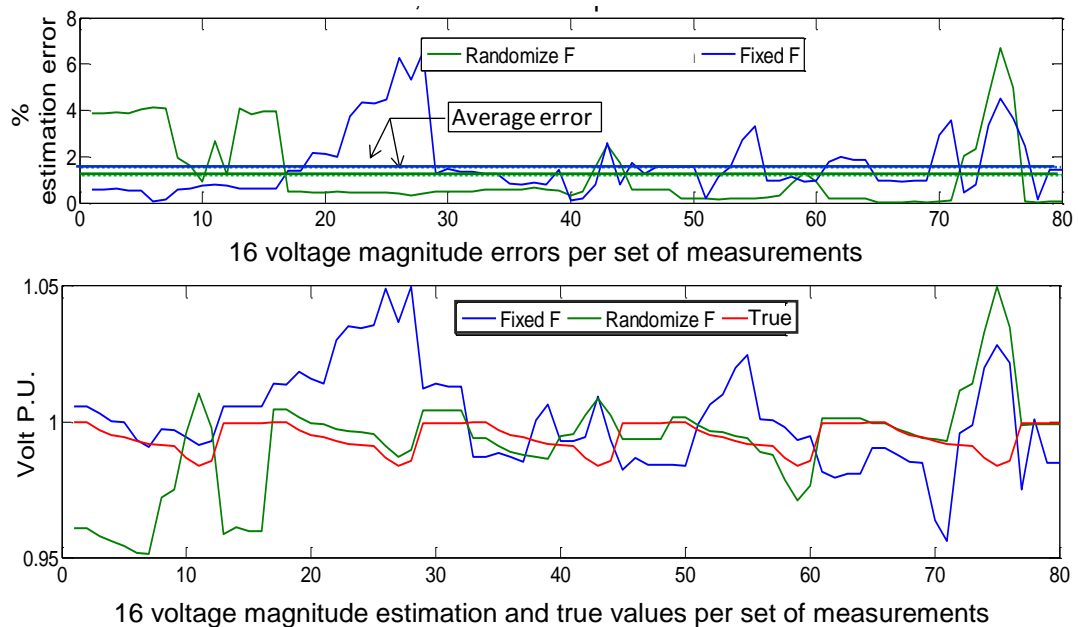


Fig. 5. 6: Performance of DEA for DSSE applying fixed and randomized mutation factor

<sup>1</sup> <http://www1.icsi.berkeley.edu/~storn/code.html>

It can be seen that DEA using randomized  $F$  (green continuous line) reduces the average voltage estimation errors more effectively. The bottom sub-plot shows that the voltage profile (green line) using randomized  $F$  matches the true voltage profile (red line) more closely than for fixed  $F$  DEA counterparts (blue line), for all five sets of case studies. This clearly indicates that the randomized  $F$  DEA outperforms the fixed  $F$  DEA consistently, to some degree. As a consequence, the randomized  $F$  DEA was chosen for further analysis and assessment of performance of the aforementioned three DEA variants.

### 5.2.6 Performance Comparison of Three DEA Variants with Randomized $F$

In this section, the efficacy of three DEA variants (conventionally expressed as  $DE\backslashrand\backslash 1\backslash bin$ ,  $DE\backslashrand\backslash 2\backslash bin$  and  $DE\backslashrand\backslash 2\backslash exp$ ) incorporating randomized  $F$ , are compared with regard to filtering out measurement noise, as DSSE optimizers. The DEA variants are executed for one hundred sets of simulated measurements with randomized Gaussian errors under similar system states, measurement configurations, generation and population sizes. The precision of estimated voltage profiles with respect to corresponding true values is observed in Fig. 5. 7, Fig. 5. 8 and Fig. 5. 9. The true and estimated voltage profiles are provided for 100 Monte Carlo studies plotted in continuous fashion, hence generating  $16 \times 100 = 1600$  data points for the 16 node network.

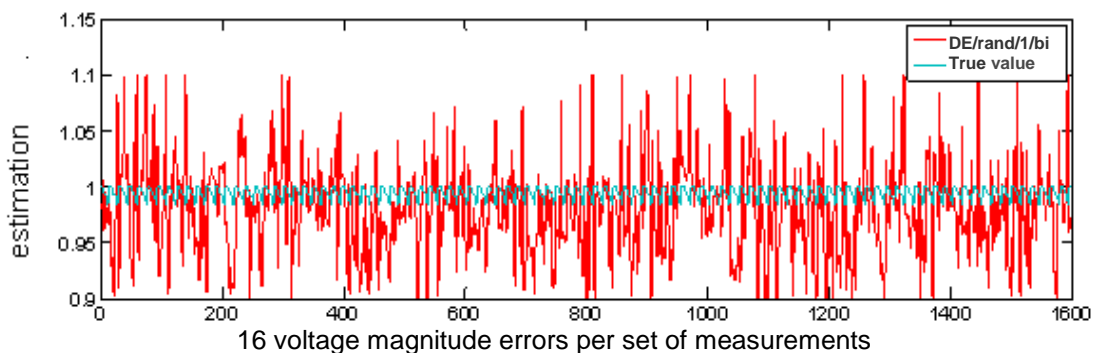


Fig. 5. 7 : Classical DEA ( $DE\backslashrand\backslash 1\backslash bi$ ) for 100 sets of measurements

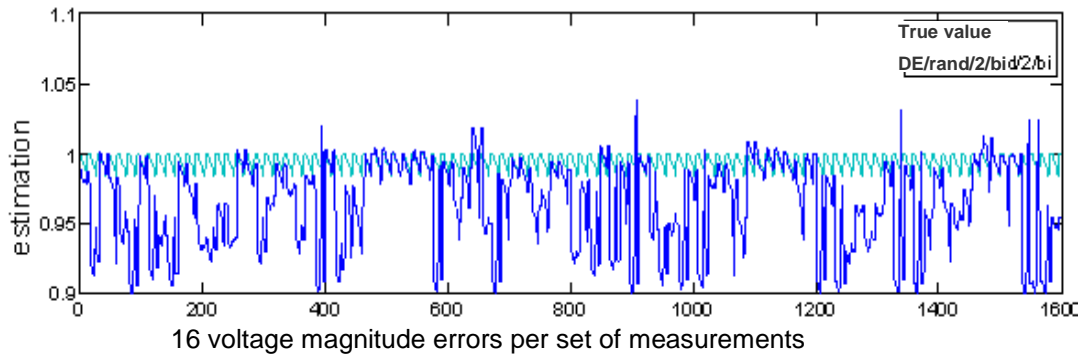


Fig. 5. 8 : Randomized F - DEA ( DE/rand/2/bi) for 100 sets of measurements

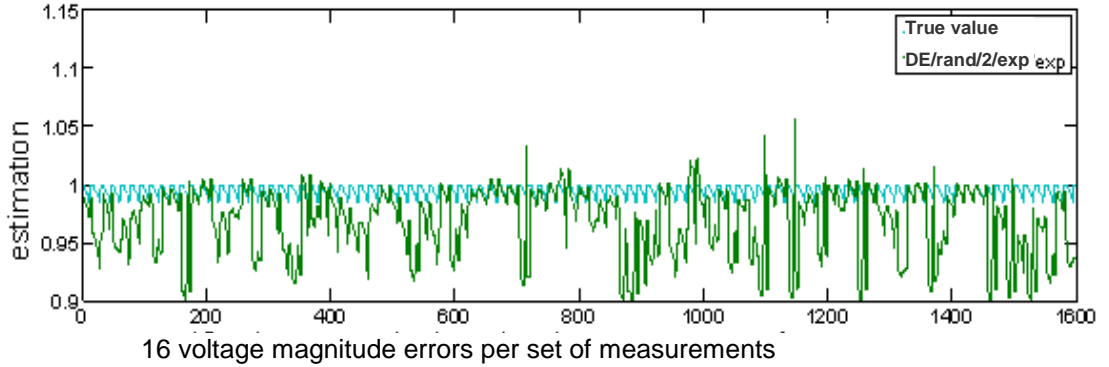


Fig. 5. 9 : Randomized F - DEA ( DE/rand/2/exp) for 100 sets of measurements

The classical DEA scheme in Fig. 5. 7 has apparently been entirely ineffective in attempting to follow the true voltage curves and generates significant levels of noise in the estimated voltage values. Both randomized  $F$  - DEA with binary cross-over strategy and with exponential cross-over strategy in Fig. 5. 8 and Fig. 5. 9 respectively, have been successful in some measure, compared to the fundamental DEA, in estimating the correct voltage profiles. The quality of performance does not exhibit any significant differences when applying two different cross-over schemes with the randomized  $F$  in Fig. 5. 8 and Fig. 5. 9. Since more accurate voltage estimation is expected from the DSSE tool, further improvement is essential for practical applications. In this context, increase in generation size and thereby enhancement of the search space is proposed to achieve acceptable quality in voltage estimation.

### 5.2.7 Observation of Computation Time and Accuracy with Generation

In previous case studies, the quality of voltage estimation improves somewhat when incorporating randomized  $F$  and taking two sets of vector differences for mutation. Since the quality is not yet as good as required for application at operational levels, an increased population-generation size is proposed, due to the fact that greater size of population and generation enhances the possibility of convergence to the global minimum point. However,

the computation time also increases due to the additional number of functions required to evaluate the augmented population and/or generation size. In this context, a comparison study is performed for a fixed set of simulated measurements and varying the generation size applying randomized  $F$  DEA with two vector differences. Table 5. 2 enlists the mean and maximum voltage magnitude and phase angle estimation errors (with respect to the true values) and computation time with various generation sizes.

<b>Generation Size</b>	<b>Time (sec)</b>	<b>%Maximum Error,  V </b>	<b>%Mean Error,  V </b>	<b>%Maximum Error, <math>\theta</math></b>	<b>%Mean Error, <math>\theta</math></b>
200	46	4.20	1.26	2751	378.20
500	116	2.80	1.22	1060	230.20
1000	234	1.54	0.83	184.50	96.25
2000	466	0.47	0.40	7.30	3.14
5000	1134	0.68	0.65	4.33	2.15

Table 5. 2: Computation time, voltage magnitude and phase angle estimation errors with various generation size

As observed in Table 5. 2, quite high values of generation ( $G > 1000$ ) are required to obtain reasonably small voltage and phase angle estimation errors. For  $G \leq 1000$ , the  $|V|$  estimation errors are even larger than the expected maximum sensor data errors (i.e. 1%) and the  $|\theta|$  estimation errors are significantly high. At  $G = 2000$ ,  $|V|$  estimation errors are at acceptable level (much less than the maximum expected sensor data errors i.e. 1%) and phase angle estimation errors are comparatively small as well. As generation size is increased further ( $G = 5000$ ), the phase angle estimation errors continue improving, however not much improvement is noticed in the voltage estimation errors. The overall outcomes are now more acceptable. However, the 1134 seconds of computation time required for this is significantly high and not suitable for the practical application. In essence, the augmented generation is certainly improving estimation quality, but with a significant cost in computation time.

### 5.2.8 Limitation of Application of DEA for Scalable DSSE

Understanding the potentiality of DEA as a scalable DSSE tool, various DEA schemes are tested to assess their effect on estimation quality, particularly on the voltage estimation. It is observed that enhanced population diversity is required for DSSE to generate good estimates. Randomized mutant vector and taking two vector differences in the mutation

process are adopted to increase population diversity and reduce the possibility of premature convergence. As maintaining the quality of estimation, while completing the calculation within a narrow time window, are essential for DSSE, the computation time for centralized DSSE for a 16 node network is observed. In practice, distribution networks can consist of thousands of nodes and DSSE must be completed within a few minutes. The DEA population for a large network can be distributed over a number of processors, so that each processor would not exceed the time constraints imposed on DSSE.

Considering time of execution in Table 5. 2, the average computation time per vector evaluation for a  $G$  generation and  $P$  population is  $\frac{\text{Total Execution Time}}{P \times G}$ . Here, five cases are for different size generation and a population of 256, the approximate average calculation time per vector evaluation is

$$T_{vector} \cong \frac{\frac{36}{256 \times 200} + \frac{116}{256 \times 500} + \frac{234}{256 \times 1000} + \frac{466}{256 \times 2000} + \frac{1134}{256 \times 5000}}{5} \cong 0.000903 \text{ sec}$$

If the population evaluation is distributed equally to  $K$  number of processors, each processor requires approximately  $0.000903 \times \frac{P}{K} \text{ sec}$  for one generation and  $0.000903 \times \frac{P}{K} \times G \text{ sec}$  for the total number of generations, ignoring communication overheads. Assuming four processors are deployed for the 16 node network cases, the computation time for one generation approximates  $\left(0.000903 \times \frac{256}{4}\right) = 0.0578 \text{ sec}$ . The approximate evaluation time for different generation sizes (Table 5. 2) will be 11 sec, 29 sec, 58 sec, 115 sec and 289 sec, for 200, 500, 1000, 2000 and 5000 generations successively. The 16 node network requires around 2000 generations to achieve the  $|V|$  estimation quality of acceptable level. The computation time after distributing computation among four processor would be at least 115 seconds. Whereas, classical WLS based methods take only fractions of a second to perform centralized DSSE for the same network. The very expensive computation time therefore is the major drawback in application of DEA as a scalable distributed DSSE.

### 5.3 Overlapping Zone Approach

One much practised approach in the development of scalable state estimation tools is to divide a large network into smaller size zones and perform estimation locally in the zones.



The local estimation is executed in parallel, while required communication facilities between zones need to be provided to enable exchange of specific local information. Allowing each zone to perform their own estimation, the substantial amount of data computation burden to centralized SE process can be distributed over the local estimators. The idea is presented in Fig. 5. 10 where circles represent zones and the arrow signs show the direction of information flows indicating the interaction with neighbouring zones.

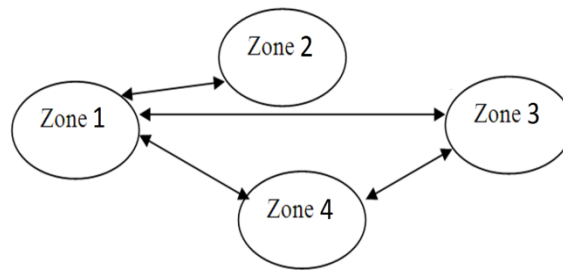


Fig. 5. 10 : Parallel estimation process

A zone may require communicating with more than one zone and each zone should exchange information with at least one other zone. As seen in Fig. 5. 10, zone 2 is linked with only one zone, zone 3 and 4 with two zones and zone 1 with three zones. Hypothetically, there is no restriction on the number of zones a network can be split into and the number of other zones that a zone can be connected with. The communication between zones, to exchange local information, is essential to provide co-ordination of information and parameter estimation for the overall process. The emerging technology of high performance parallel computation technology and high speed communication support are expected to bring benefits to the parallel application of a DSSE tool in zones. In this work the overlapping zone based estimation method is proposed as a scalable DSSE solution and is referred to as Overlapping Zone Approach (OZA).

### 5.3.1 Overlapping Zone

In OZA, zones are formed in such way that each zone has at least one common node with at least one other zone. The zone performs local estimation that leads to global estimation through information exchange, coordination and communication among them. In the proposed method, nodes are overlapped and are the common elements between two zones. Therefore, there will be more than one set of voltage, phase angle and power injection parameter estimation for overlapping nodes.

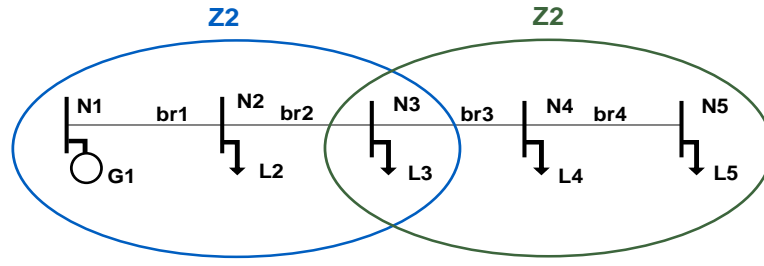


Fig. 5. 11: Overlapping zone based network division

The network line diagram of one example case presented in Fig. 5. 11 explains the idea of overlapping zone division. The small network is supposed to consist of five nodes (N1, N2, N3, N4 and N5), four branches (br1, br2, br3 and br4), one generator (G1) and four loads (L2, L3, L4 and L5). The network is divided into two overlapping zones, Z1 and Z2 where each zone contains three nodes. N3 is common to each zone and so are the parameters associated to N3.

### 5.3.2 Local and Halo data

Zones are required to be provided with two types of data to enable local estimation for OZA based DSSE. These are local( $N_l$ ) and halo data( $N_h$ ).

- Local Data: All static and dynamic data of nodes and branches inside a zone. In Fig. 5. 11, local data for Z1 is information associated to N1, N2, N3, br1, br2 and local data for Z2 is information involving N3, N4, N5, br3, br4.
- Halo Data: This includes static and dynamic data associating nodes and branches outside a zone, however the local estimation cannot be accurately accomplished without this information. The halo data generally comprises of information from outside nodes and branches which are directly connected to the overlapping nodes. For example, N3 of Z1 is connected to N4 through br3. N3 is the overlapping node which is also inside Z2, however br3 and N4 are components of Z2 only; hence N4 is a halo node, br3 is a halo branch. The branch parameter and flow information of br3, voltage, phase angle and power injection information of Z2 can be considered as halo data in this case. Similarly, halo node and branch for Z2 is N2 and br2 respectively and halo data is all information related to N2 and br2 from Z1.

The requirement of halo data information can be realized if the local estimation process of a zone is considered. For example, when Z1 calculates total power injection of the overlapping node N3, it requires power injection/consumption data of all generator/load as

well as power flow information from all branches connected to N3. The local estimator of Z1 therefore needs to be provided with either, power flow information from Z2 of br3 or, voltage and phase angle of N4 and line parameters of br3, so Z1 can calculate the power flow value. The power injection calculation of the overlapping node cannot be completed without the halo data information in this example.

### **5.3.3 Network Splitting into Zones**

Zones of sizes as similar as possible should be chosen to balance the local estimation computation times. This method allows great flexibility in the definition of zones and their overlaps for consistency, size and location of measurements. Any part of the network can form a zone provided each node at least has one real or pseudo or virtual measurement. However, investigations have shown that the estimation quality improves when zones include electrically close lateral and sub-lateral branches. The overlap nodes can be common to more than two zones. It was proposed initially that zones should be heavily overlapping in terms of number of nodes to facilitate extensive data co-ordination among zones. Nevertheless, the heavy overlap is observed to affect the local estimation quality adversely when the estimated data from neighbouring zones lacks coherency with the local estimates. It also contributes to increased communication overheads as a consequence of additional transferable information. Further investigation concludes that minimizing overlaps (in terms of nodes) leads to better estimates and faster convergence. Local estimation for lightly overlapping zones uses local data at maximum level while being affected reasonably by estimates from neighbouring zones. In case of incoherency in the data from neighbouring zones, the local estimator is not heavily affected, but rather tends to adjust to the neighbouring zone estimates steadily.

### **5.3.4 The Complete Algorithm**

As discussed previously, the OZA splits a large network into smaller size overlapping zones that perform local estimation recurrently. Completing all local estimation in a repetitive manner is termed as 'zonal iteration'. Local estimation is executed in parallel, while communication facilities between zones needs to be provided to enable exchange of specific local information. The entire process is depicted in Fig.5. 12.

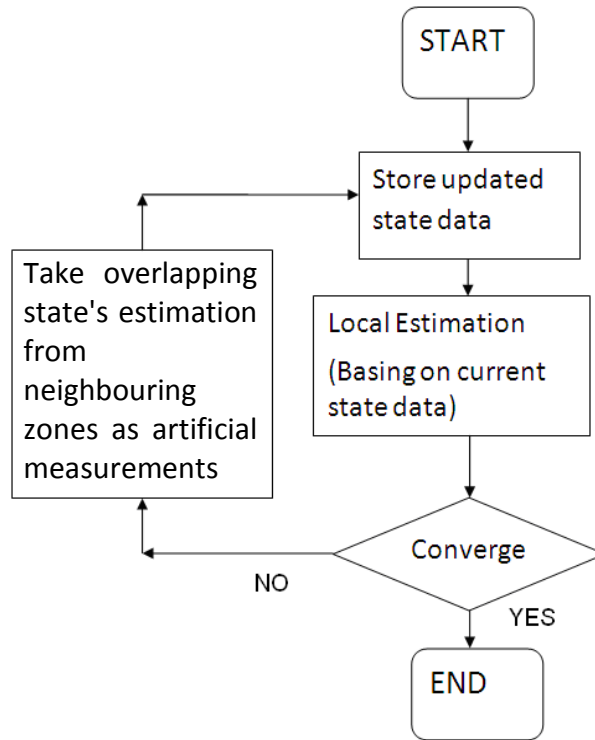


Fig. 5. 12: Overlapping zone approach algorithm flowchart

The local estimator performs 'flat-' or 'semi-hot' estimation during the first zonal iteration. In a 'semi-hot' start, the voltage values  $|V|$ , are set to one if measurements are not available at that node, otherwise they are set to the measurement values. The phase angle values  $|\theta|$ , are always set to zero in the first zonal iteration. Each local estimator is updated with the new local  $|V|$  and  $|\theta|$  on completion of every zonal iteration. The most updated  $|V|$ ,  $|\theta|$  are used as initial values in the next round of local estimation to give a 'hot-start'. Application of this combination of hot and semi-hot start leads the estimator towards faster convergence and more accurate estimation.

Hachtel's Augmented Matrix method is integrated as the optimization method for local estimation. Only one zone can have the slack node, as assumed generally in a centralized process. As there is continuous data flow among zones, the other zones can eventually adjust and estimate with respect to the slack node values. It should be noted that, two types of convergences and iterations are associated with OZA: local and zonal:

- Local or inner loop variables: Convergence and iterations are related to the conventional state estimation Gauss-Newton recursive process which is performed within the local SE optimization process, which is Hachtel's method in this case.
- Zonal or outer loop variables: Convergence and iterations are not related to Gauss-Newton process. The zonal iteration or interaction refers to the exchange of

information between zones in the form of artificial measurements. The zonal or overall convergence is controlled by overlap node mismatch (differences between voltage and phase estimates of overlapping nodes from associated zones) values or maximum allowable zonal interactions. This means, if the mismatch values are small enough (ideally,  $10^{-6}$ ), or the number of zonal data exchanges reaches the maximum allowable number, the entire algorithm is considered to have converged and results are retrieved.

In Fig. 5.12, the data exchange and coordination is executed when local estimation is at or near the optimum point. The exchange of information between zones occurs in the form of ‘artificial measurements ( $M_a$ )’. Estimates of  $|V|$ ,  $|\theta|$  of shared nodes, provided from a neighbouring zone, are referred to as ‘artificial measurements’. The local estimator considers the artificial measurements as additional measurements with relatively low weights. In addition to local node ( $N_l$ ) data, the local estimation also needs primary state ( $|V|$ ,  $|\theta|$ ) information of halo nodes ( $N_h$ ) from the adjacent zones to calculate overlapping nodes’ power values. The local Jacobian matrix is a  $N_l \times (M_l + M_a)$  matrix where  $M_l$  represents locally available real and pseudo-measurements. The local estimation is allowed to run the optimization process until it achieves a local convergence point and after that the data exchange between zones is executed. The overall algorithm is deemed to have converged when the voltage and phase angle estimates of overlapping nodes match to within a given tolerance. When the mismatch of overlap nodes is negligible, the estimates of OZA achieve equivalent values to that of the centralized process. The mismatch is defined as the  $L_1$  norm of the difference between common node estimates by a pair of overlapping zones. In Fig. 5. 11, node 3 is common for zone 1 and zone 2, the mismatch will be a two elements vector,  $\left[ \left| V_{z1(N3)} - V_{z2(N3)} \right|_1, \left| \theta_{z1(N3)} - \theta_{z2(N3)} \right|_1 \right]$ . The algorithm must stop if reaches the maximum allowable zonal iteration count. Several zonal iterations are executed until global convergence is attained.

The main control variables of the algorithm are zonal convergence, interactions and mismatches, which are observed in the test cases in the following sections. The local estimator and the Gauss-Newton elements are rather considered as inside a black box. In addition to Hachtel’s Augmented Matrix method as WLS solution process, DEA is also applied for local WLS estimation in OZA in the following assessment studies. However, the OZA approach also provides flexibility to choose different optimization processes as the local estimator.

## 5.4 Application of Hachtel's Method in OZA

Since typical Gauss-Newton based optimization solvers require only fractions of a second in general to complete SE for a small network, networks having larger size are selected to perform the OZA based DSSE (that has not been done for DEA DSSE related studies). Several comparative studies are performed on a 356 and a 711 node UKGDS network. The 711 node network is developed by concatenating two 356 node network to provide a larger size problem to apply OZA. Load flow studies are performed for both networks before creating simulated measurement sets and scenario based data. The performance is observed in two case studies. In the first study, the network is assumed to have measurements only at a GSP. In the second case, more measurements are assumed to be available at overlapping and halo nodes. The local estimation is set to stop when the Gauss-Newton update vector values,  $\Delta x$  are less than or equal to  $10^{-5}$ , with the maximum allowed local Gauss-Newton iterations being 5. To elaborate further, the local estimation is allowed to run the optimization process unless it achieves a local near convergence point ( $\Delta x \leq 10^{-5}$ , or Gauss-Newton iteration equals 5). The global or zonal convergence is achieved when the maximum overlap mismatch value is equal to or less than  $10^{-6}$ . The maximum allowed zonal iteration varies from 10 to 1000 depending on the case study scenarios. The purpose of the case studies presented here is to observe the effect of voltage and phase angle estimation quality using the proposed OZA.

The algorithms are actually run serially, applying the parallel processing logic. In the practical MATLAB tool, a loop is controlled in such a way that it at first accomplishes the local estimation serially in each zone without exchanging information. Once all local estimation is achieved, the data transfer in zones is executed sequentially. Following that, the local information is updated and the mismatches in zones are checked. If mismatches in all zones are not below the thresholds i.e. global convergence is not achieved, the next zonal iteration initiates. The pseudo-code including syntax regarding the practical MATLAB generated tool is given in appendix 10.

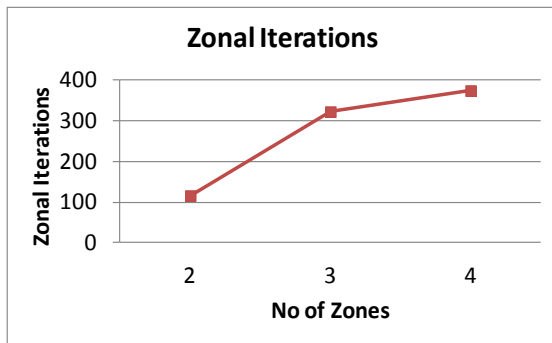
### 5.4.1 OZA: Applying Measurements at GSP Only

In this case study, the network is assumed to have only one set of voltage and power (P, Q) injection measurements at GSP. Background surveys substantiate that the common node state estimates, from the overlapping zones, match well when the number of zonal iterations is high, in this case. The consistency of the algorithm depends on the pattern of

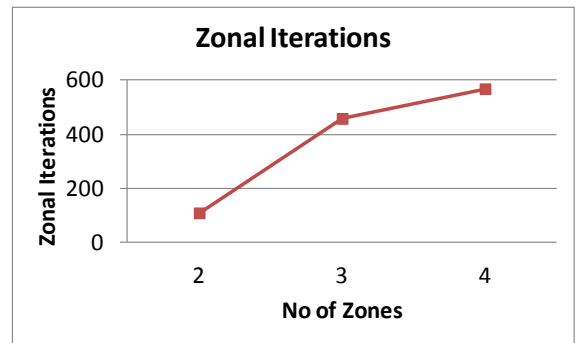
overlap mismatch values. In principle, more zonal interactions should lead to smaller overlap mismatches, that would approximate 'zero' at the convergence point. The algorithm is set to perform as many zonal iterations as 1000 and the maximum of all absolute overlap mismatch values is observed in each iteration.

*Voltage and Phase Angle Deviation Relative to True Value*

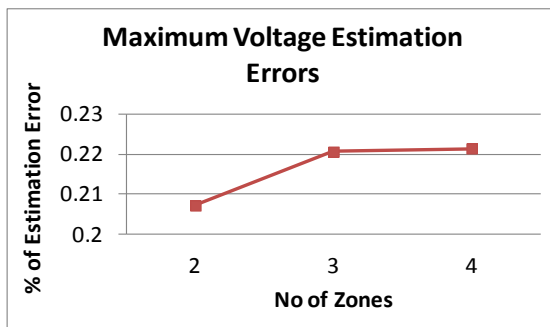
The voltage and phase angle estimation errors with respect to the true values are plotted for the 711 node network in Fig. 5. 13a. Another set of outcomes is observed under the same conditions and measurement scenarios, however taking real P, Q measurements as equality constraints in Fig. 5. 13b. The number of zonal iterations required to achieve the expected overlap mismatch values are plotted in Fig. 5. 13a1 and Fig. 5. 13b1. The maximum and mean voltage estimation errors, the maximum and mean phase angle estimation errors of 711 nodes are plotted in Fig. 5. 13(a2, b2), Fig. 5. 13(a3, b3), Fig. 5. 13(a4, b4) and Fig. 5. 13(a5, b5) successively.



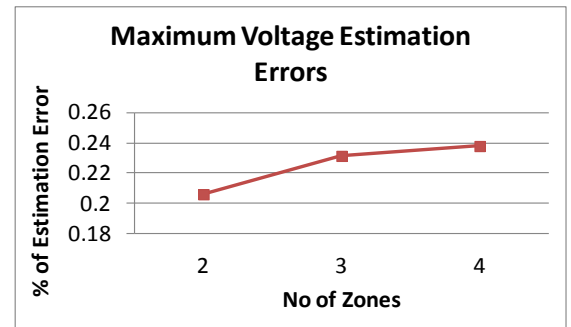
(a1)



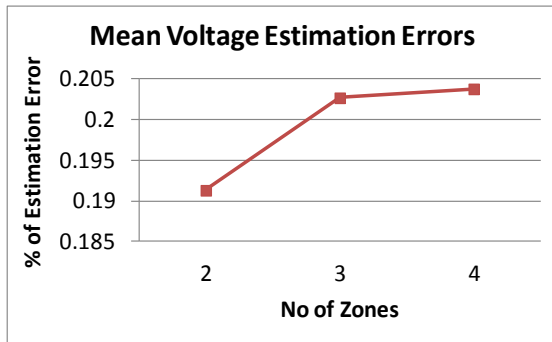
(b1)



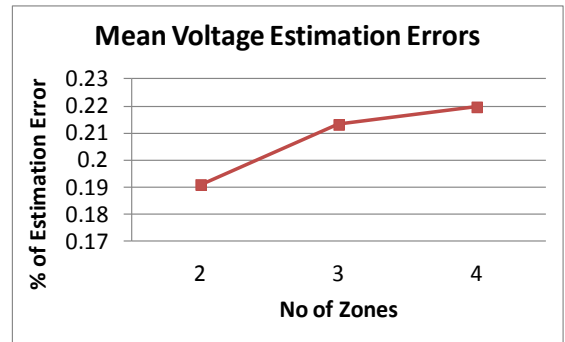
(a2)



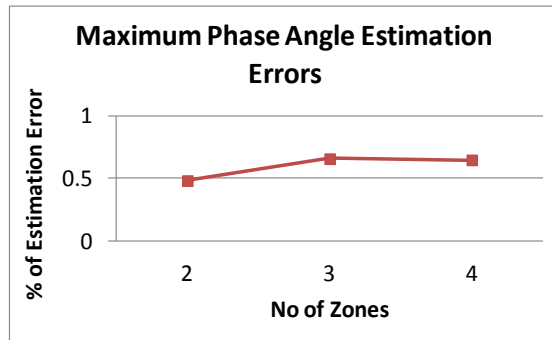
(b2)



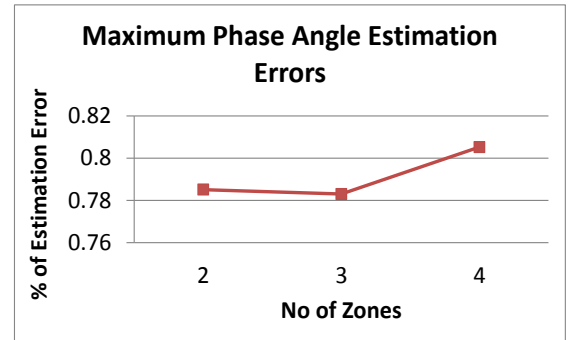
(a3)



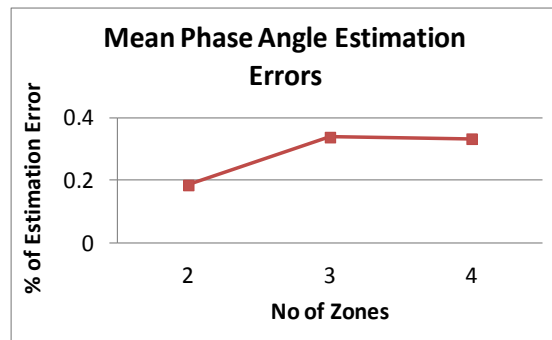
(b3)



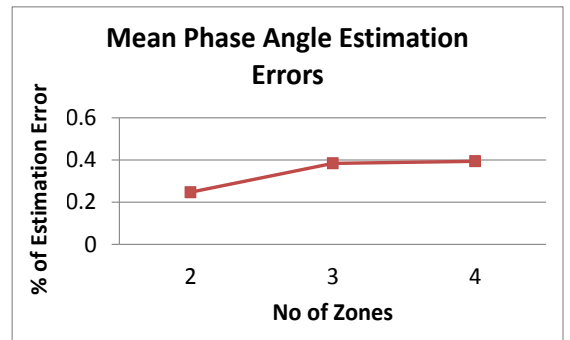
(a4)



(b4)



(a5)



(b5)

Fig. 5. 13: Studies on the 711 node network with 1 set of measurement taking a) only virtual measurement as constraints and b) both real and virtual power measurements as constraints

a1,b1: Number of zonal iteration required to attain convergence i.e. the maximum absolute zonal mismatch value less than or equal to  $10^{-6}$

a2,b2: Maximum voltage estimation errors with respect to the true values at convergence

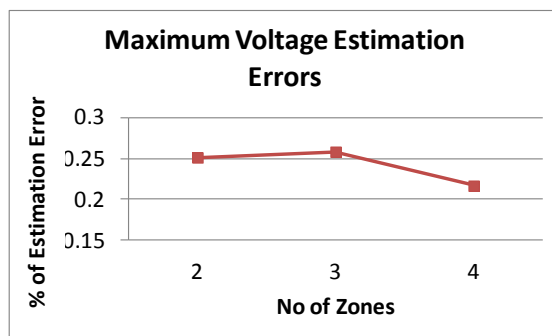
a3,b3: Mean voltage estimation errors with respect to the true values at convergence

a4,b4: Maximum phase angle estimation errors with respect to the true values at convergence

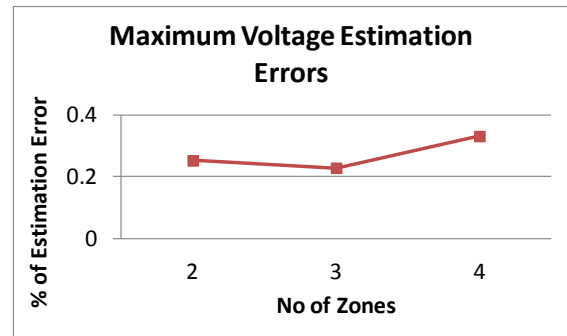
a5,b5: Mean phase angle estimation errors with respect to the true values at convergence



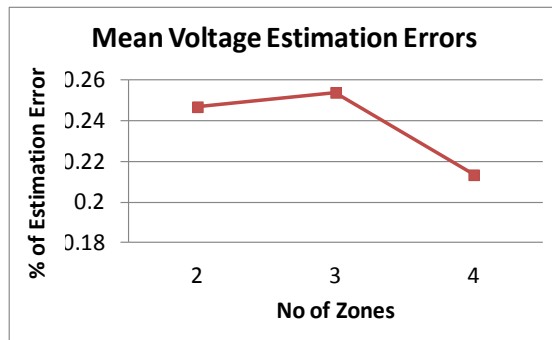
A similar study is performed for the 356 node UKGDS network, for which the expected overlap mismatch value is never obtained in a few cases associated with three and four zone divisions. The estimator therefore operates until the maximum allowed number of zonal iterations (1000) is reached. In order to justify the comparison in equivalent settings, the entire case studies are allowed to run for 1000 zonal interactions and the resultant voltage and phase angle estimation errors plots are illustrated in Fig. 5. 14. Since the zonal iteration is fixed at 1000 times, the corresponding 'iteration versus zones' (as in Fig. 5. 13a1 and Fig. 5. 13b1) plots are excluded in this case. The maximum and mean voltage estimation errors, the maximum and mean phase angle estimation errors of 356 nodes are plotted in Fig. 5. 13(a1, b1), Fig. 5. 13(a2, b2), Fig. 5. 13(a3, b3) and Fig. 5. 13(a4, b4) successively.



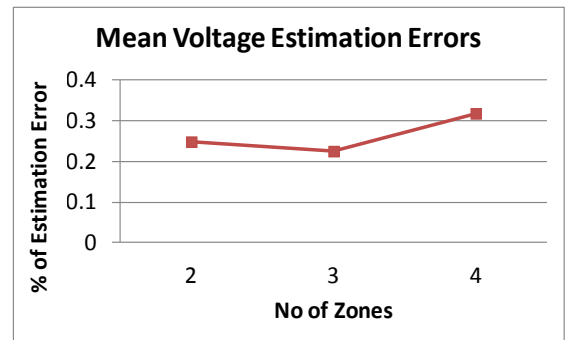
(a1)



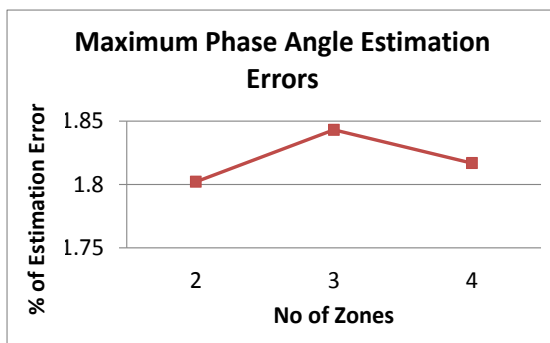
(b1)



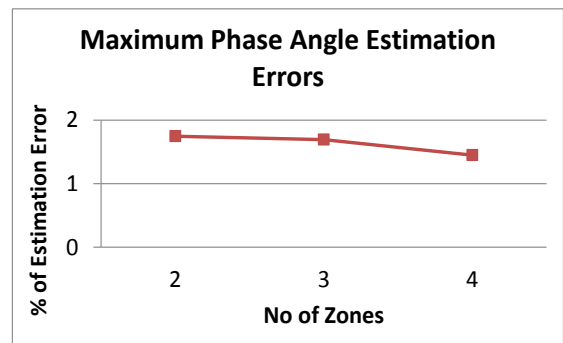
(a2)



(b2)



(a3)



(b3)

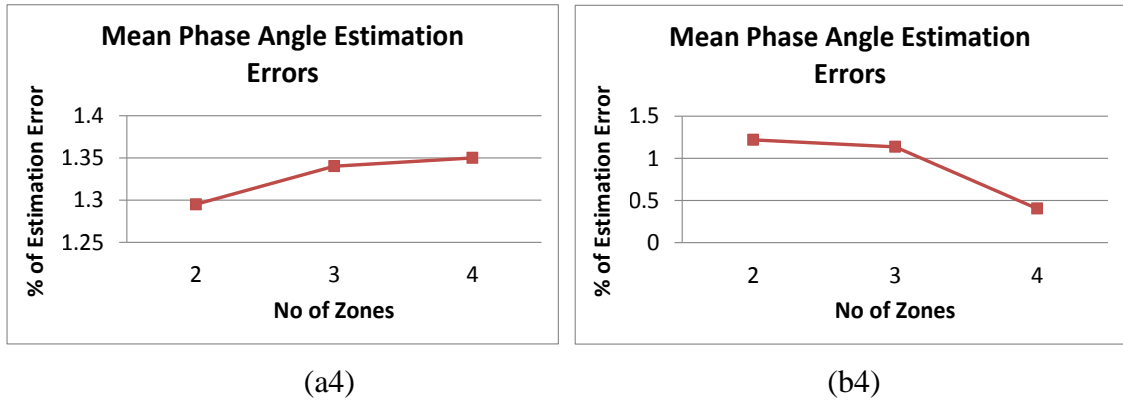


Fig. 5. 14: Studies on the 356 node network with 1 set of measurement taking a) only virtual measurement as constraints and b) both real and virtual power measurements as constraints

a1,b1: Maximum voltage estimation errors with respect to the true values at 1000 iterations

a2,b3: Mean voltage estimation errors with respect to the true values at 1000 iterations

a3,b3: Maximum phase angle estimation errors with respect to the true values at 1000 iterations

a4,b4: Mean phase angle estimation errors with respect to the true values at 1000 iterations

In Fig. 5. 13(a1, b1), the 711 node network requires quite a significant number of zonal iterations to achieve the desired overlap mismatch to convergence. The 356 node network does not achieve convergence point even after 1000 zonal iteration in several cases; therefore the relevant plot illustrating the convergence speed with various network splitting cannot be provided. However the maximum overlap mismatch value for 1000<sup>th</sup> zonal iteration is observed to be  $3.30 \times 10^{-09}$ ,  $3.41 \times 10^{-07}$  and  $1.30 \times 10^{-05}$  assuming only virtual measurement as constraints and  $4.93 \times 10^{-09}$ ,  $4.65 \times 10^{-06}$  and  $1.70 \times 10^{-04}$  assuming both real and virtual measurements as constraints when the 356 node network is divided into two, three and four zones respectively. The convergence therefore becomes slow when real time power measurements are assumed to be constrained. Similar impacts of constrained P,Q measurements are observed on the 711 nodes network in Fig. 5. 13(a1,b1) as the OZA requires more zonal iterations to converge (i.e. the maximum absolute mismatch value is equal to or less than  $10^{-6}$ ). These certainly indicate that assumption of real measurements as equality constraints can slow down the OZA in achieving convergence with the desired overlap mismatch values.

The voltage estimation errors somewhat increases when taking real measurements as constraints for both 711 and 356 network case studies. On the other hand, some degree of improvements are observed in the maximum phase angle estimation errors for the 711

node network and the mean phase angle estimation errors for the 356 node network in Fig. 5. 13(a4, b4) and Fig. 5. 14 (a4, b4) respectively. The increase in zone division has a variable effect on different sizes networks. The case studies associated with the 711 node network illustrate that more zone divisions increases the overall estimation errors. However, the outcomes for the 356 node network depict an increase in the voltage estimation errors from two to three zone divisions case studies, and then again a decrease of that for three to four zone divisions when using only virtual measurement as constraints. The estimation errors are the least for three zone divisions using both real and virtual measurements as constraints for the 356 node network. The variations in improvement are absolutely close in many cases in Fig. 5. 13 and Fig. 5. 14, therefore minor increases or decreases in estimation quality could be ignored and they could be assumed to have similar levels of estimation accuracy.

The centralized SE using Hachtel's Augmented Matrix method, with the exact measurement sets used in the above cases, reduces voltage estimation errors down to a maximum of 0.474% and a mean of 0.469% for the 711 node network. And, the maximum and mean phase angle estimation errors are obtained as 1.68% and 0.772% respectively. The 356 node network results in maximum voltage estimation error as 0.245% and mean voltage estimation error as 0.241% when a centralized SE is applied. The maximum and mean phase angle estimation errors are 1.88% and 1.412% respectively in this case. Comparing with voltage and phase angle estimation errors presented in Fig. 5. 13 and Fig. 5. 14 for the 711 and 356 node network respectively, it is evident that the OZA has been more effective in reducing the overall voltage and phase angle estimation errors, compared to centralized SE methods, when applying exact real and pseudo-measurement sets and an optimization process (Hachtel's method) in this study.

#### *Estimated and True Voltage and Phase Angle Profiles*

The small percentage errors in voltage and phase angle estimation in the previous section clearly indicate that the quality of primary variable estimation should be satisfactory. Nevertheless, simultaneous plots of the true and estimated values are presented to obtain a more comprehensible illustration. The voltage and phase angle profiles, with respect to the nodes, are plotted at convergence if the required overlap mismatch is achieved (in the 711 node network case), otherwise plots are drawn at the 1000<sup>th</sup> iteration (in the case of the 356 node network) for two, three and four zone divisions. Fig. 5. 15, Fig. 5. 16 and Fig. 5. 17 show the voltage and phase angle profiles of the estimated and true values for the 711 node

network. Fig. 5. 18, Fig. 5. 19 and Fig. 5. 20 illustrate similar plots for the 356 node network. Only virtual measurements are used as constraints for the plots in Fig. 5. 15 to Fig. 5. 20. The  $|V|, |\theta|$  graphical representations in this section are provided for one measurement instance.

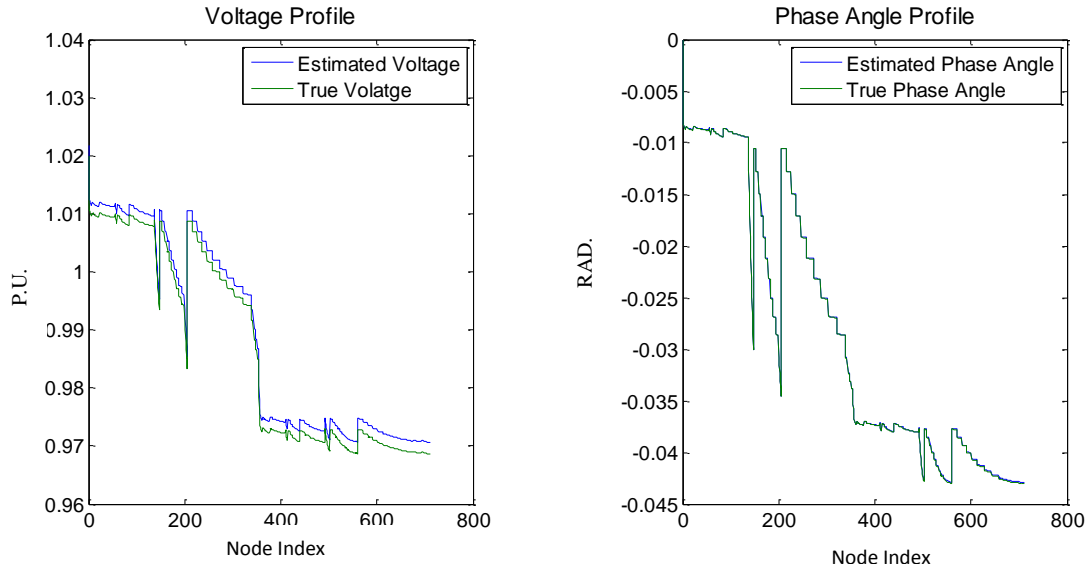


Fig. 5. 15: 2 zone division of 711 node network

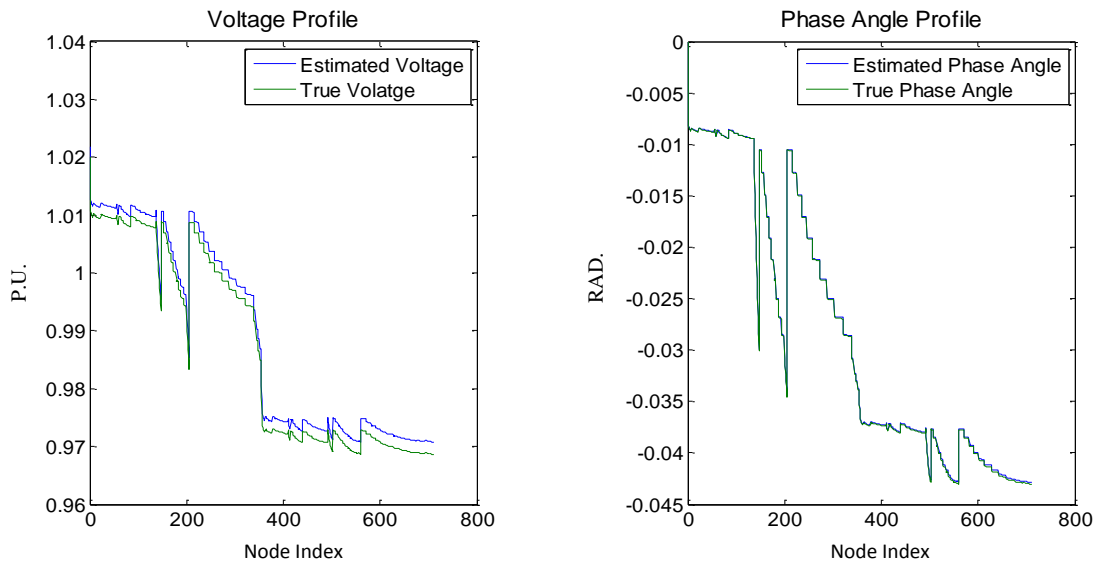


Fig. 5. 16: 3 zone division of 711 node network

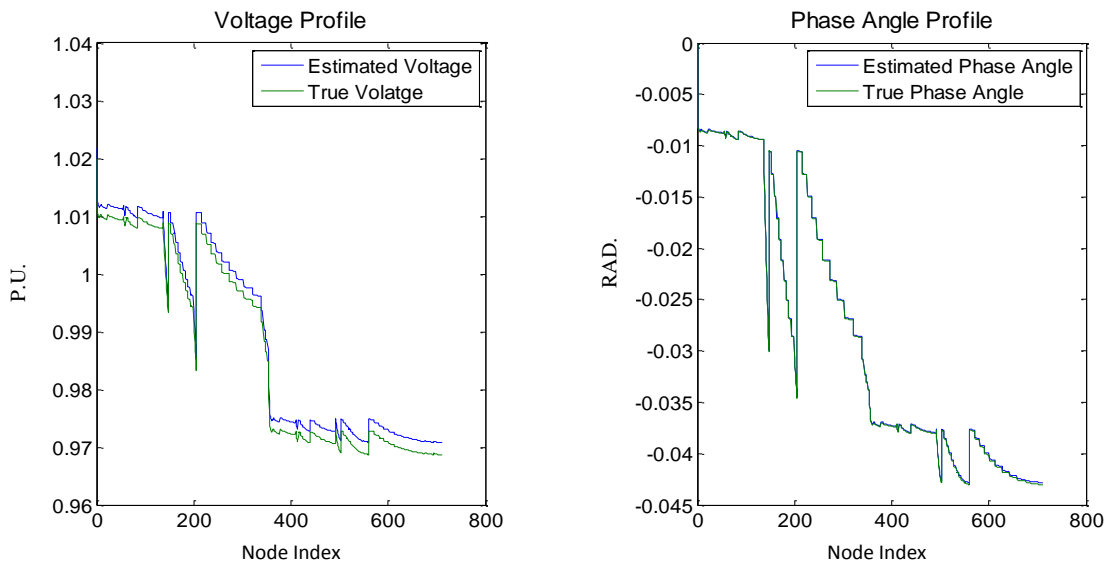


Fig. 5. 17: 4 zone division of 711 node network

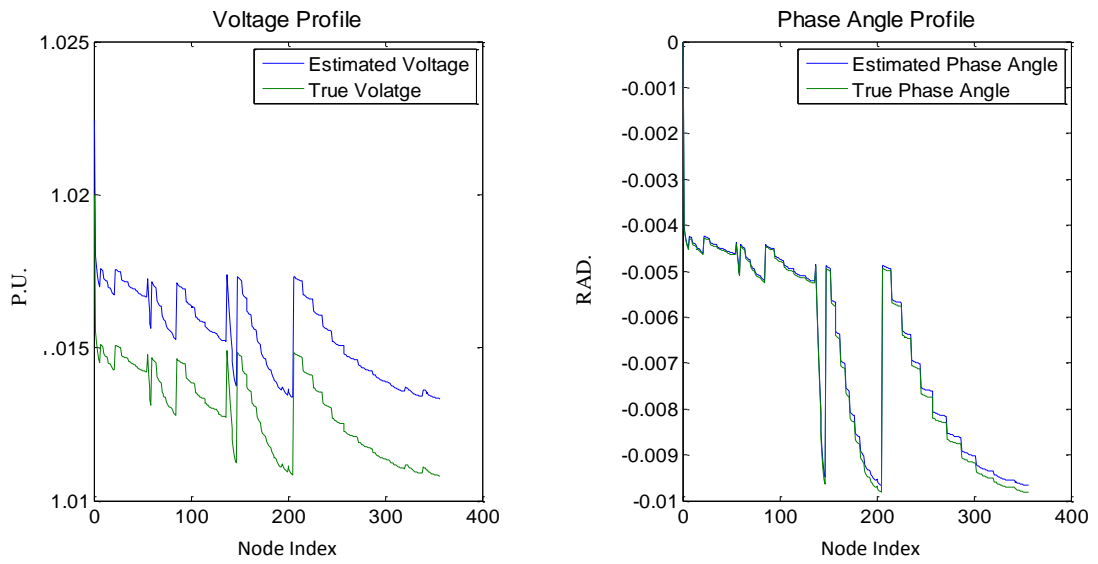


Fig. 5. 18: 2 zone division of 356 node network

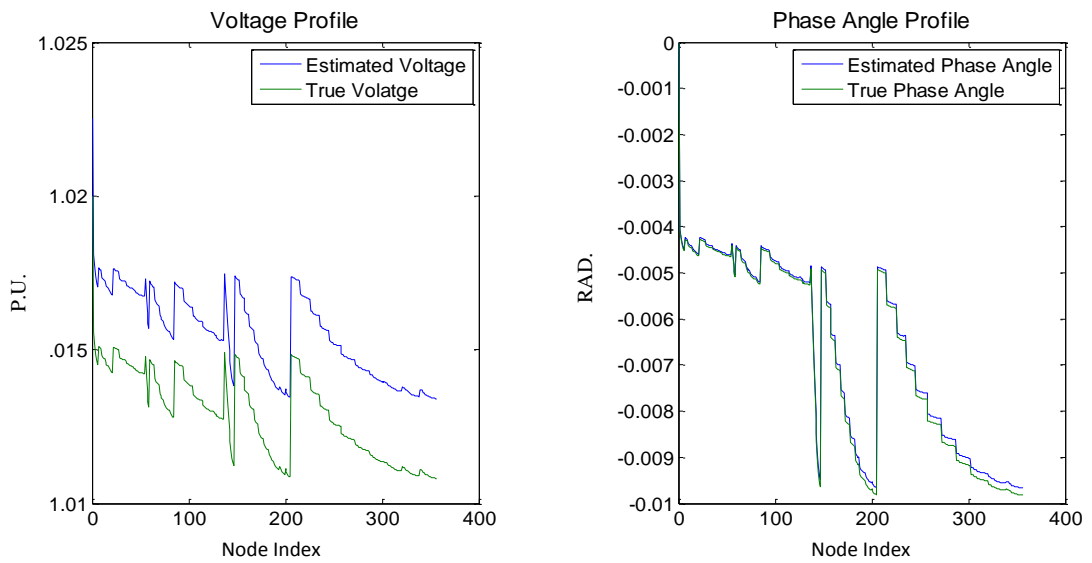


Fig. 5. 19: 3 zone division of 356 node network

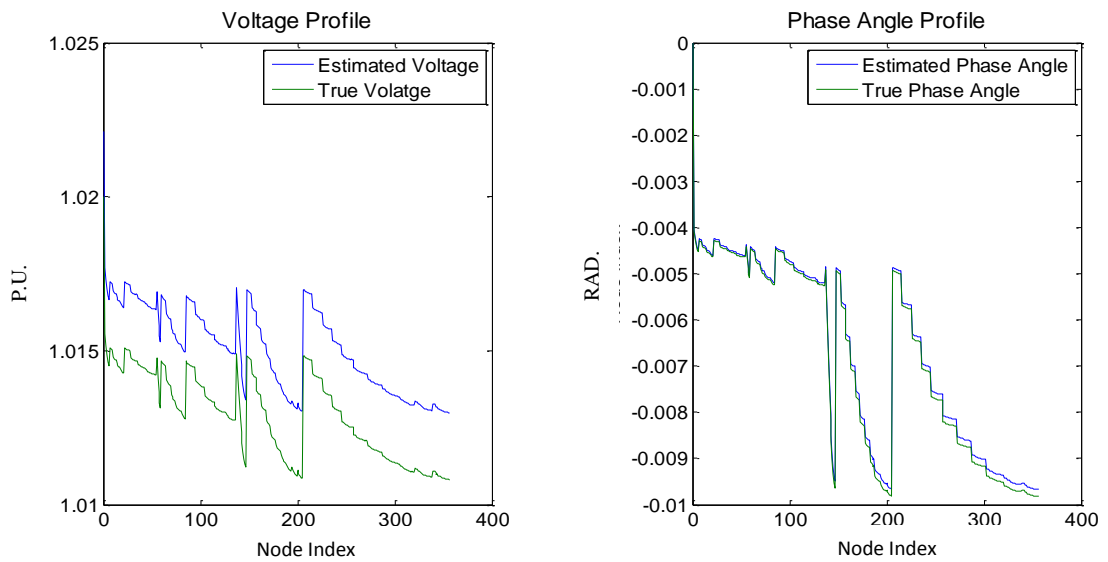
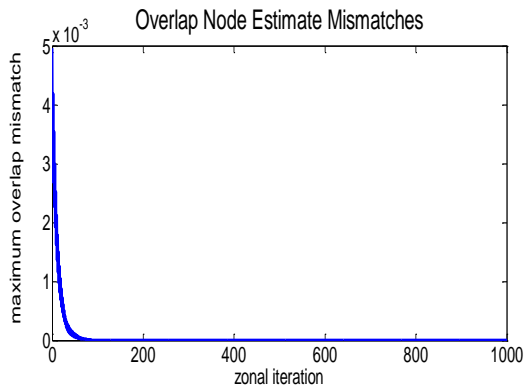


Fig. 5. 20: 4 zone division of 356 node network

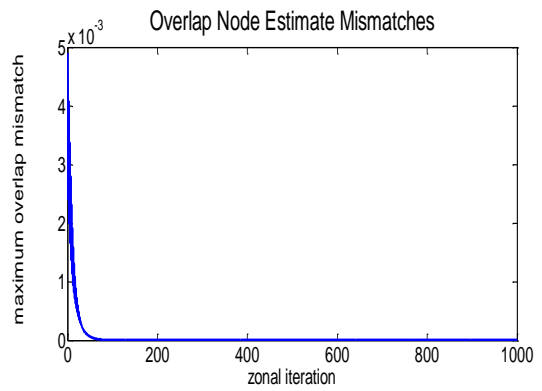
It is observed in all cases in Fig. 5. 15 to Fig. 5. 20 that the estimated voltage magnitudes and phase angle values do not exactly match the true values as there always exist small to large errors in estimated values; such small deviations are expected. An important observation is that in all cases the estimated voltage and phase angle values closely follow the pattern of overall voltage and phase angle profiles of the networks. As no incoherent value is observed in estimated quantities, the successful application of OZA is sufficiently justified through the above studies.

#### *Verification of Consistency*

The consistency of the algorithm depends on the pattern of overlap mismatch values. Ideally, more zonal interaction should lead to smaller overlap mismatch. If the overlap mismatch values increases with the increase of zonal interactions, the OZA may not be considered as a robust and stable method. In this respect, the algorithm is set to perform as high a number of zonal iterations as 1000 times and the maximum absolute overlap mismatch value is observed after each iteration for both cases of using only virtual measurements as constraints and also using both virtual and real measurements as constraints. Fig. 5. 21, Fig. 5. 22 and Fig. 5. 23 show the mismatch curves for two, three and four zone division where index (a) presents cases using only virtual measurement as constraints and index (b) represents cases where both virtual and real measurements are used as constraints. Fig. 5. 24, Fig. 5. 25 and Fig. 5. 26 illustrates similar cases for the 356 node network.

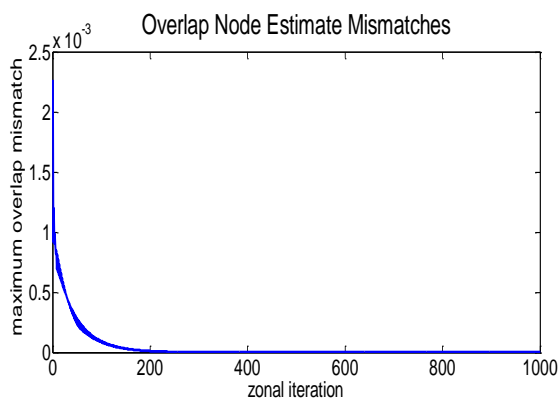


(a)

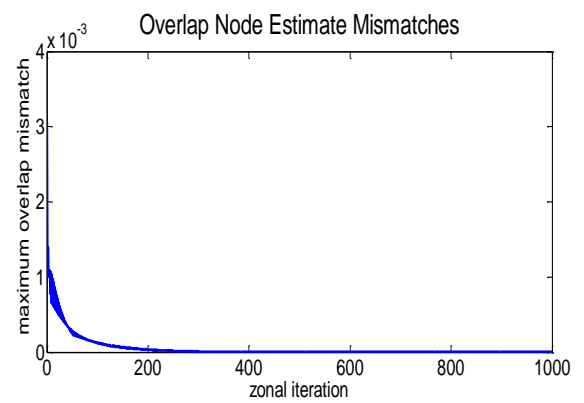


(b)

Fig. 5. 21: 2 zone division of 711 node network

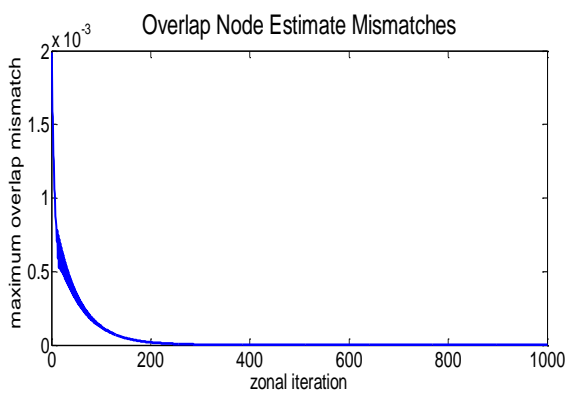


(a)

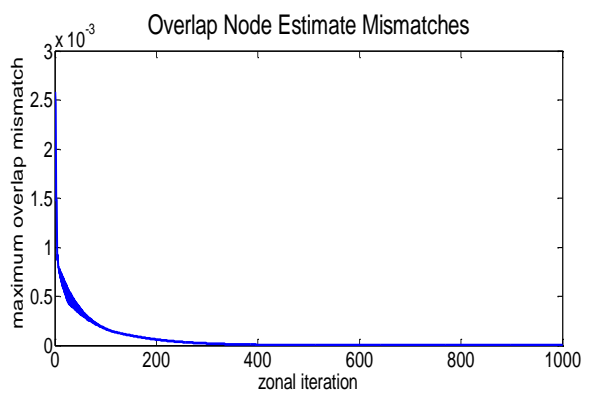


(b)

Fig. 5. 22: 3 zone division of 711 node network



(a)



(b)

Fig. 5. 23: 4 zone division of 711 node network

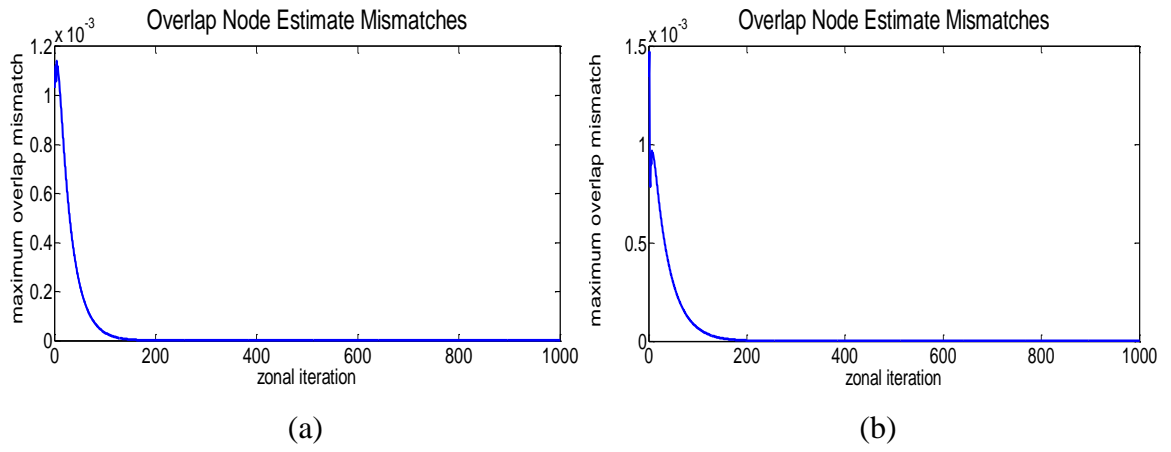


Fig. 5. 24: 2 zone division of 356 node network

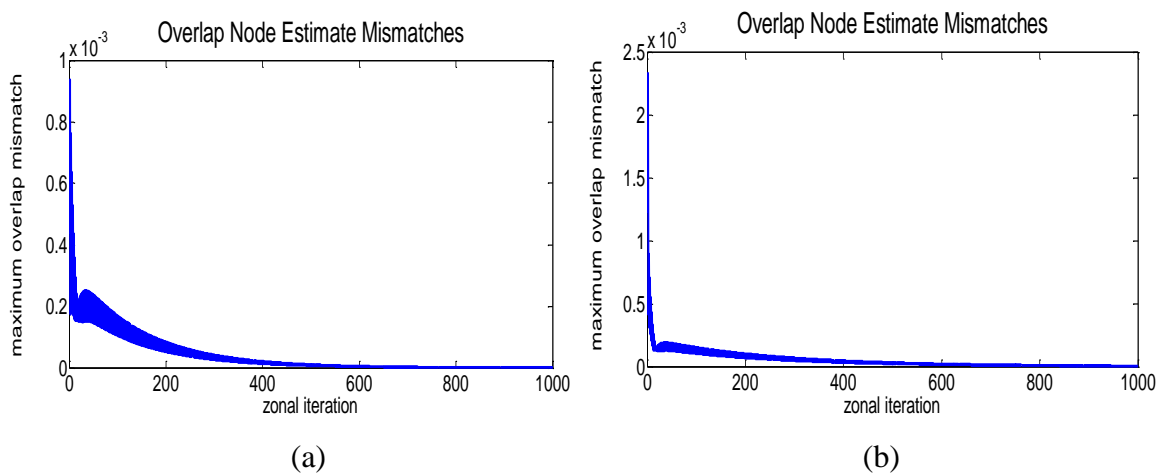


Fig. 5. 25: 3 zone division of 356 node network

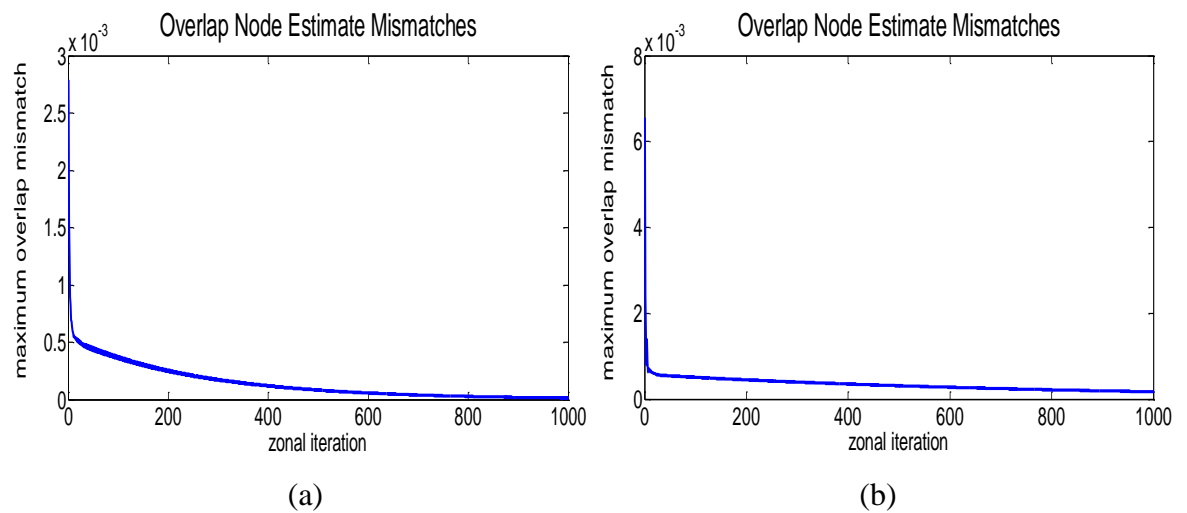


Fig. 5. 26: 4 zone division of 356 node network

The overlap mismatch values show gradually descending trends, although there are in some cases fluctuations from one value to the next. The fluctuating behaviour results from the effect of the new input values, from halo and overlapping nodes, after every zonal



iteration. Although fluctuating, the consistently decreasing pattern of overlap mismatch indicates the solution is approaching the optimum point. The maximum absolute overlap mismatch values after the 1st and 1000th zonal iterations are included in Table 5. 3 and Table 5. 4.

<b>Zone Divisions</b>	<b>Mismatch for Virtual Measurements as Constraints, 1st Iteration (p.u./rad.)</b>	<b>Mismatch for Virtual Measurements as Constraints, 1000th Iteration (p.u./rad.)</b>	<b>Mismatch for Real and Virtual Measurements as Constraints, 1st Iteration (p.u./rad.)</b>	<b>Mismatch for Real and Virtual Measurements as Constraints, 1000th Iteration. Mismatch (p.u./rad.)</b>
2 zones	0.00496	6.06e-9	0.00491	9.53e-9
3 zones	0.00227	1.39e-9	0.00306	5.47e-9
4 zones	0.00199	3.05e-9	0.00258	8.065e-9

Table 5. 3 : Overlap mismatch values for 711 node network

<b>Zone Divisions</b>	<b>Mismatch for Virtual Measurements as Constraints, 1st Iteration (p.u./rad.)</b>	<b>Mismatch for Virtual Measurements as Constraints, 1000th Iteration (p.u./rad.)</b>	<b>Mismatch for Real and Virtual Measurements as Constraints, 1st Iteration (p.u./rad.)</b>	<b>Mismatch for Real and Virtual Measurements as Constraints, 1000th Iteration (p.u./rad.)</b>
2 zones	0.00103	3.3e-9	0.001342	4.93e-9
3 zones	0.00044	3.41e-7	0.00121	4.6e-6
4 zones	0.00279	1.3e-5	0.00656	1.7e-4

Table 5. 4 : Overlap mismatch values for 356 node network

In the cases with the 711 node network, in Table 5. 3, any particular relation between convergence and network divisions is not apparent. In every case, the mismatch values approximate zero, indicating very good match of overlapping variables. However, Table 5. 4 for the 356 node network shows that the convergence becomes slow, and consequently mismatch values become larger, as network division increases. Assuming real and virtual

measurements as constraints shows an increase in overlap mismatch values to some degree for both networks, as shown in Table 5. 3 and Table 5. 4.

In OZA, the estimation quality is best achieved when a zone includes a complete lateral or sub-lateral branch or part of a feeder. A zone consisting of part of two or more feeders may account for slow convergence. Therefore, network division should be in proportion to the size of the network. The 711 node network is large enough to divide into up to four zones. However, the 356 node network is too small to be divided into more than two zones. The three and four zone divisions of the 356 node network lead to the creation of a zone containing part of two or three feeders; which is avoided to some extent in the case of the 711 node network. A large network can give more flexibility in network division.

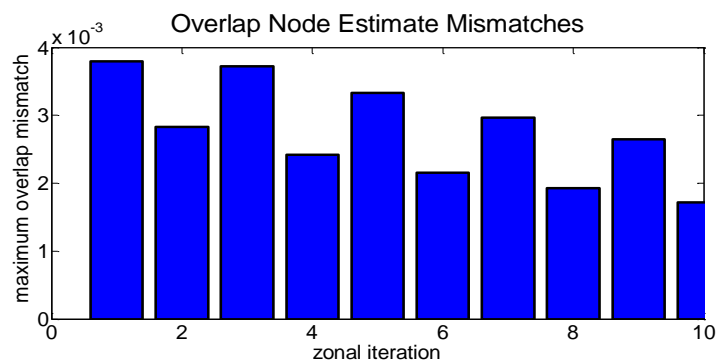
As the common node mismatch values remain near zero after 1000 iterations in both cases, the  $|V|$  and  $|\theta|$  estimation errors remain fairly similar for different zone divisions as observed in the previous section. This can be considered as a good indication that different zone splits can deliver similar quality results, provided heavy zonal interaction occurs. In this test case, only real measurements are available at a GSP that is not common to other zones, nevertheless the  $|V|$  and  $|\theta|$  estimation errors remain very small with respect to the true values and similar to those of the centralized DSSE. This justifies the fundamental idea of sharing and coordinating locally available information through the overlapping nodes to achieve the ultimate estimates in the OZA.

#### **5.4.2 OZA: Applying Measurements at GSP, Halo and Overlapping Nodes**

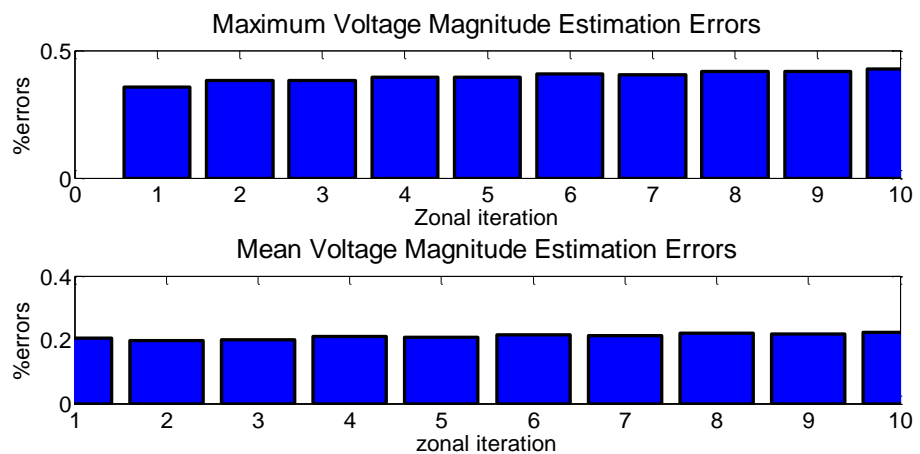
As observed in section 5.4.1, OZA can estimate voltage and phase angle providing even greater quality estimates than that from centralized DSSE; however the number of zonal iterations required for that is certainly a significant number. This raises the possibility of increased communications overhead, when the algorithm is actually applied on a parallel computation platform. Where computation time is a vital factor for DSSE, a reduced zonal interaction has an adverse impact on estimation quality. The trade-off between computation time and estimation accuracy brings challenges in the practical application of OZA. In this section, OZA applications requiring reduced zonal interactions but more sensor deployment are proposed. As the reduced level of information exchange (to reduce computation time) costs the coherency of the estimates, more sensor deployment is required to compensate for the loss of coherency and to achieve a feasible solution.

Additional  $|V|$  and P,Q injection sensors are proposed to be placed on all halo and common nodes. The added  $|V|, P, Q$  sensors required on the halo and overlapping nodes determines the number of new sensors. The 711 node network is assumed be equipped with nine sets of sensors when divided into two zones and nineteen sets of sensors when divided into three and four zones. The number of sensors assumed to be available for the 356 node network is five, seven and ten when network is divided into two, three and four zones respectively. The voltage and phase angle estimation errors with respect to the true values are observed for up to ten zonal iterations.

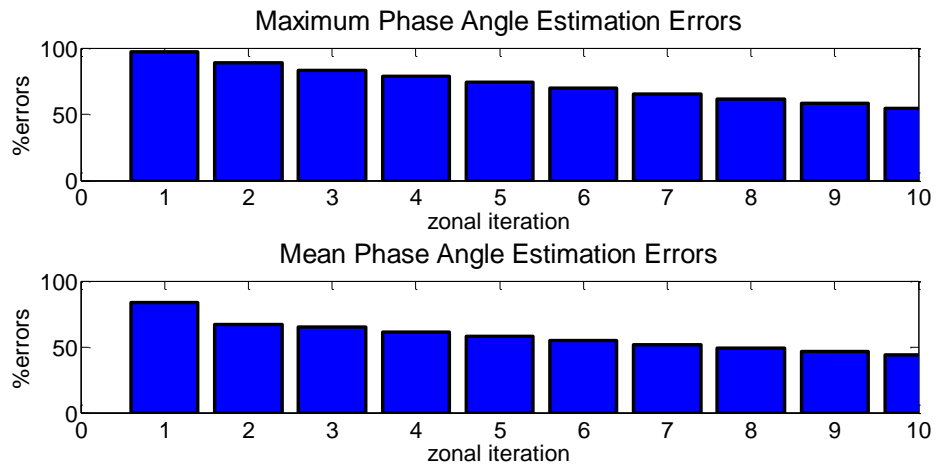
The match of estimated values with the true values is observed with every zonal iteration/ overlap exchange, to obtain a view of improvement of estimates with increased zonal interactions. Fig. 5. 27, Fig. 5. 28 and Fig. 5. 29 demonstrate the overlap mismatch values, maximum and mean voltage and phase angle estimation errors for the 711 node network being divided into two, three and four zones respectively.



(a)

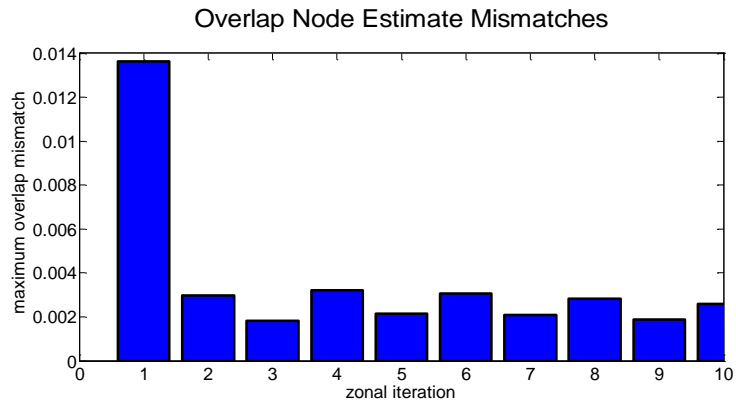


(b)

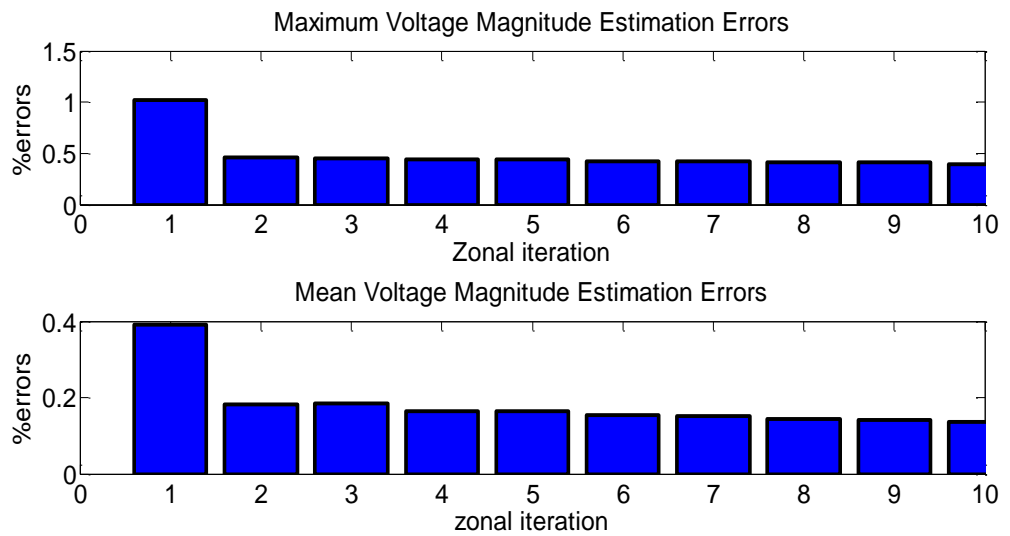


(c)

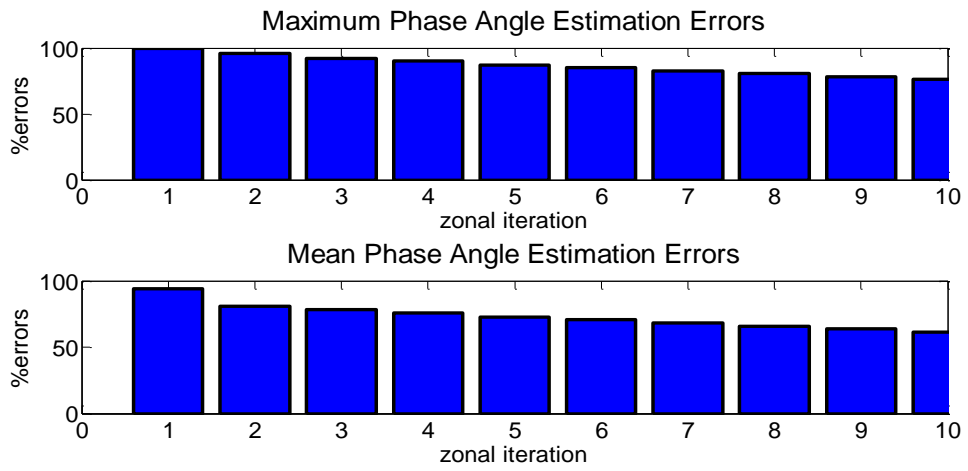
Fig. 5. 27: 2 zone division of 711 node network- a) Mismatches, b) Voltage estimation errors and -c) Phase angle estimation errors



(a)

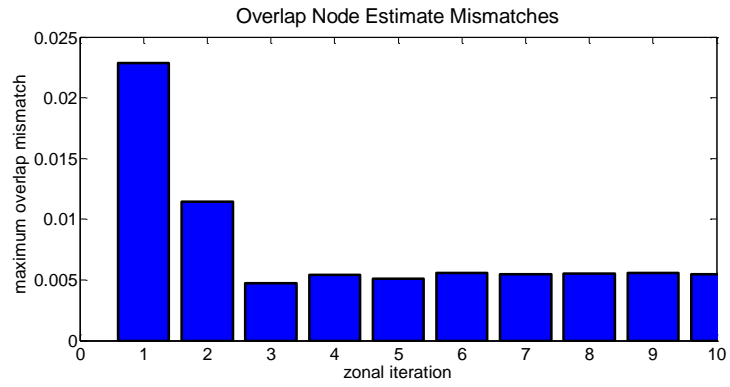


(b)

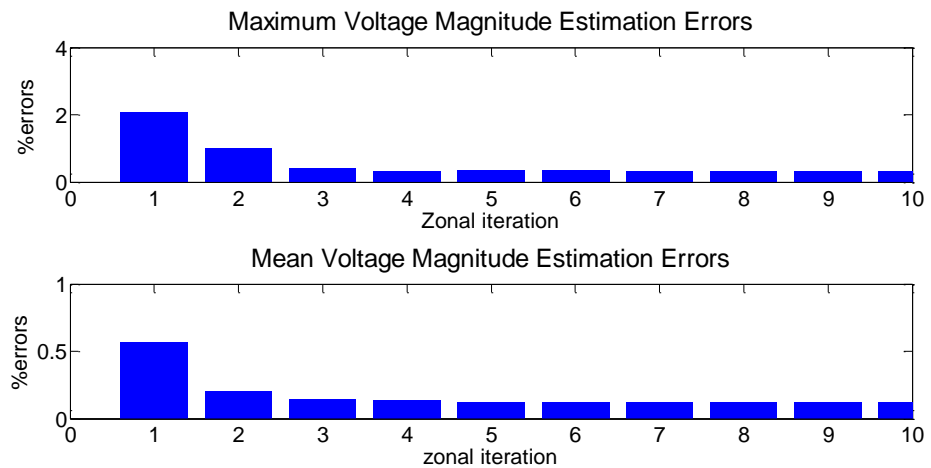


(c)

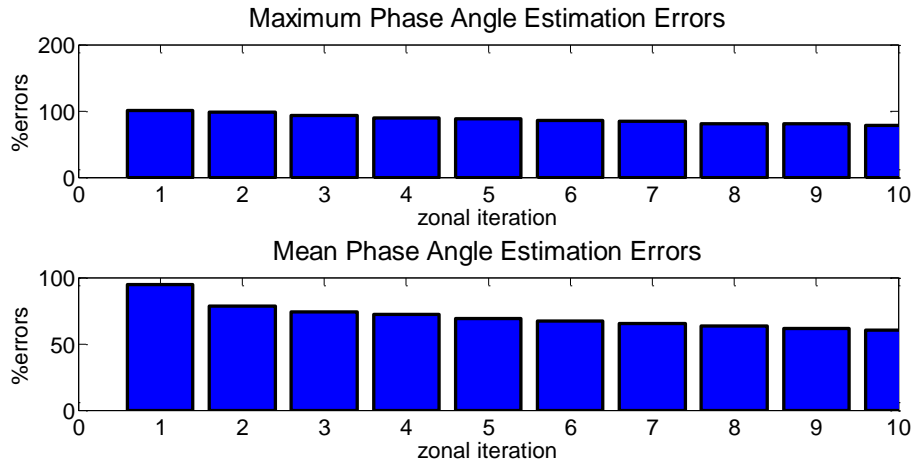
Fig. 5. 28: 3 zone division of the 711 node network - a) Mismatches, b) Voltage estimation errors and -c) Phase angle estimation errors



(a)



(b)



(c)

Fig. 5. 29: 4 zone division of the 711 node network - a) Mismatches, b) Voltage estimation errors and -c) Phase angle estimation errors

The overlap mismatch values show fluctuating, but gradually descending, trends for two, three and four zone divisions in Fig. 5. 27(a), Fig. 5. 28(a) and Fig. 5. 29(a). Fig. 5. 27(b, c) show a gradual decrease in phase angle estimation errors; on the contrary the voltage estimation errors increases slightly. However the voltage estimation errors constantly remain below 0.5% and the mean errors are around 0.2%. As mentioned earlier in section 5.4.1., the two and three zone division cases show gradual decrease in both voltage and phase angle estimation errors in Fig. 5. 28(b, c) and Fig. 5. 29(b, c). The decrease in voltage estimation errors is quite sharp at the end of the second zonal iteration and the change in phase angle estimations in these cases is quite slow. The maximum voltage estimation errors approaches below 0.5% after second and third zonal iterations for three and four zone divisions cases respectively; while the mean voltage estimation errors achieve values less than 0.2%. As the centralized SE for the 711 node network reduces voltage estimation errors down to a maximum of 0.474% and a mean of 0.469%, the voltage estimation quality achieved with additional meter placement can be considered as quality outcomes with reduced zonal iterations. The phase angle estimation errors in all cases of Fig. 5. 27(c), Fig. 5. 28(c) and Fig. 5. 29(c) are quite significant compare to those obtained in centralized SE solutions which are 1.68% and 0.772% for maximum and mean phase angle estimation errors respectively.

Fig. 5. 30, Fig. 5. 31 and Fig. 5. 32 show the overlap mismatch values, maximum and mean voltage and phase angle estimation errors for the 356 node network dividing into two, three and four zones respectively

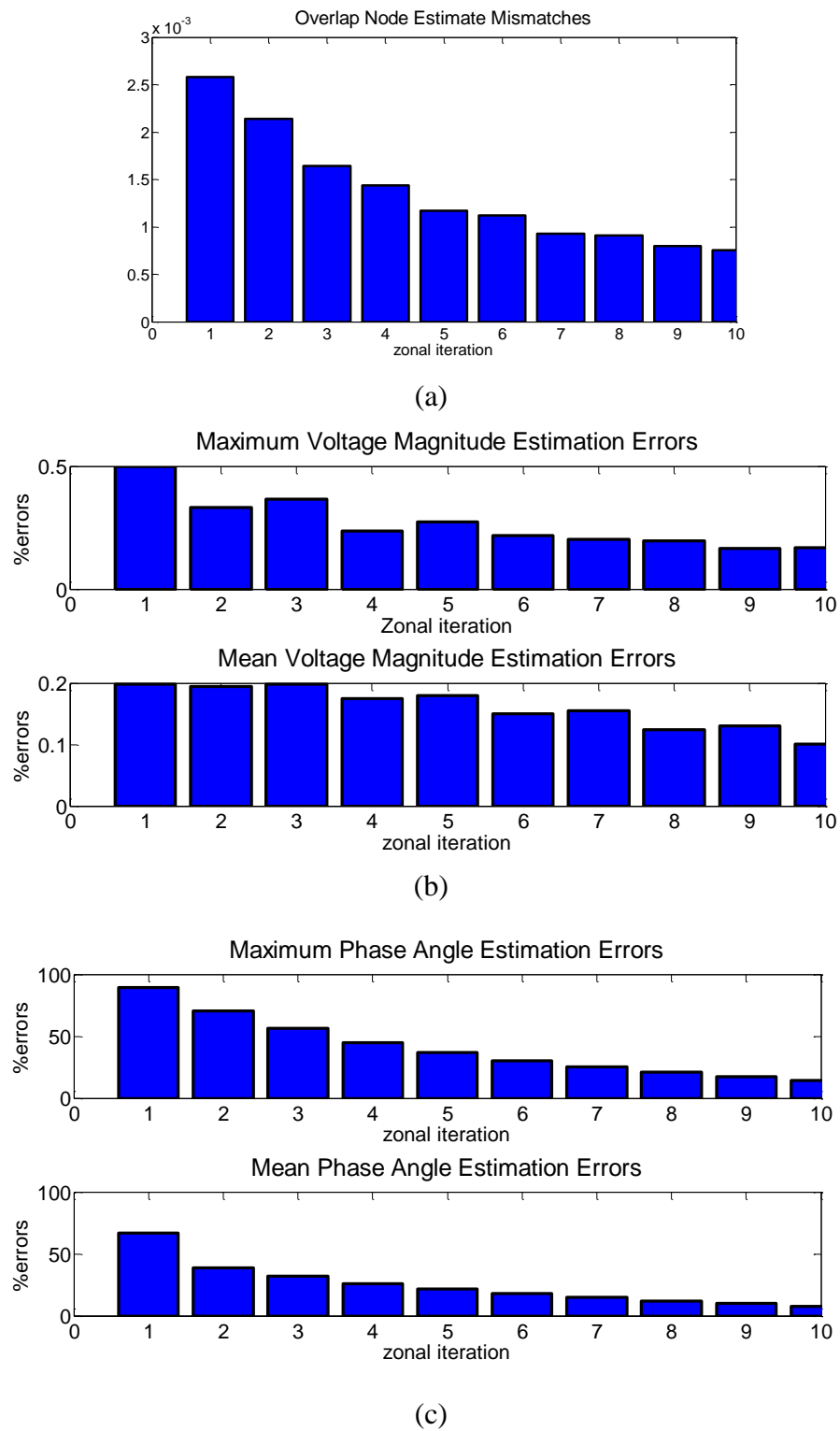
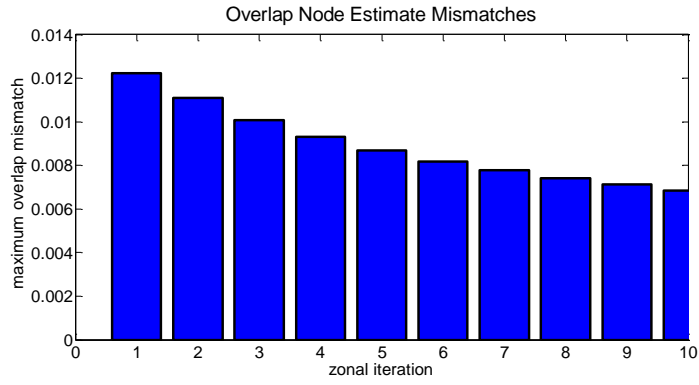
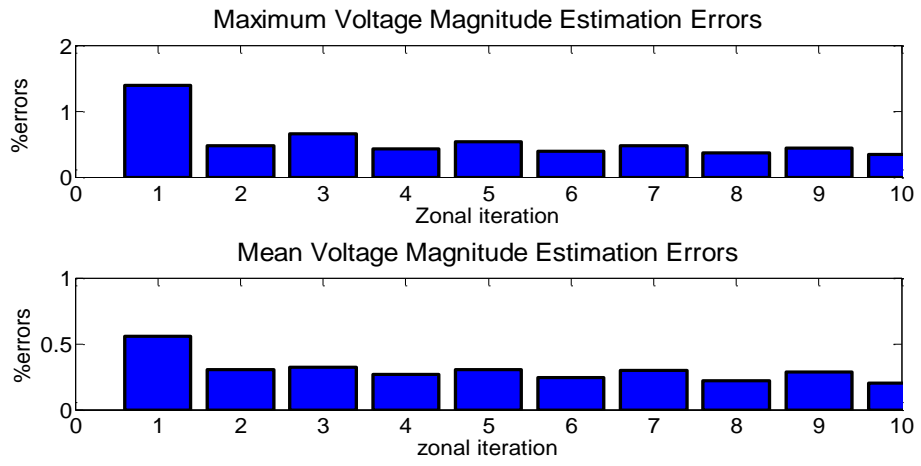


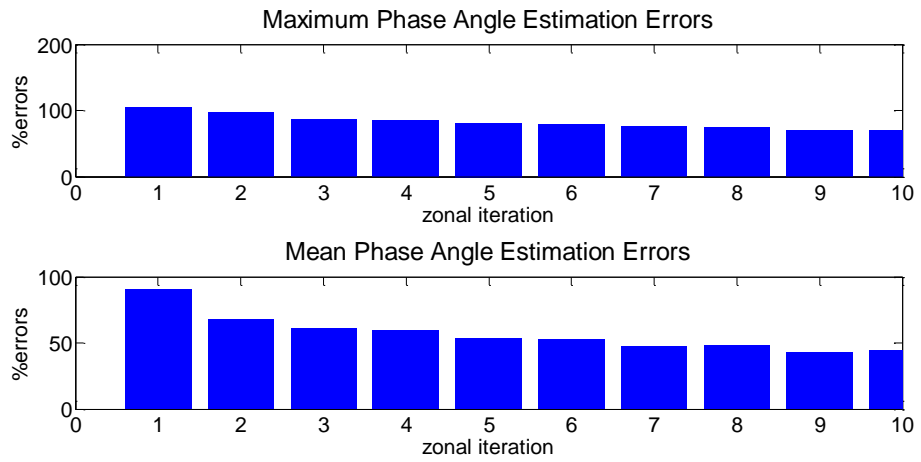
Fig. 5. 30: 2 zone division of the 356 node network - a) Mismatches, b) Voltage estimation errors and -c) Phase angle estimation errors



(a)



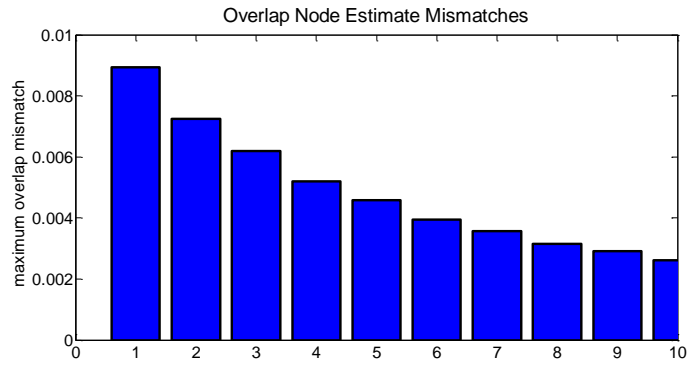
(b)



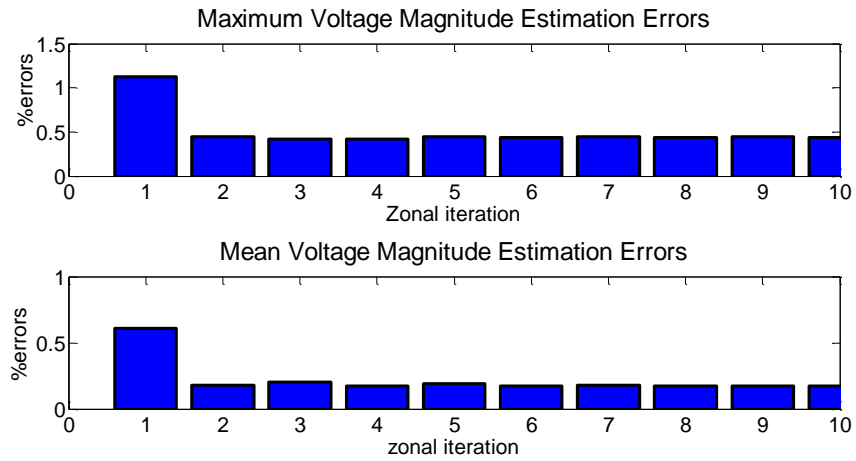
(c)

Fig. 5. 31: 3 zone division of 356 node network- a) Mismatches, b) Voltage estimation errors and -c) Phase angle estimation errors

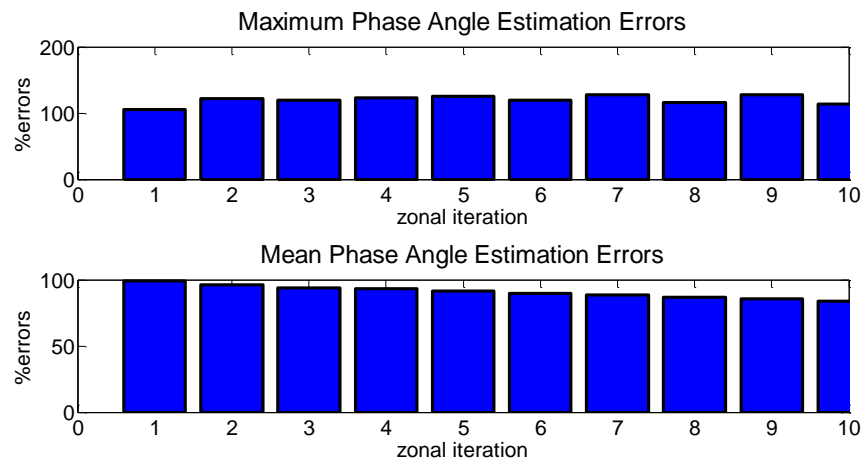




(a)



(b)



(c)

Fig. 5. 32: 4 zone division of 356 node network- a) Mismatches, b) Voltage estimation errors and -c) Phase angle estimation errors

Unlike the 711 node network, the overlap mismatch trends with zonal iterations show consistent gradual decrease in Fig. 5. 30(a), Fig. 5. 31(a) and Fig. 5. 32(a) for two, three and four zone divisions respectively. The maximum and mean voltage estimation errors in

all three cases reduce below 0.5% and 0.2% respectively after the second zonal iteration. Here again, the phase angle estimation errors remain significantly large in 10 zonal interactions. Since the 356 node network results in maximum voltage estimation error of 0.245% and mean voltage estimation error of 0.241% when a centralized SE is applied, some deterioration of maximum voltage estimation errors are observed when applying OZA in this case. The maximum and mean phase angle estimation error is 1.88% and 1.412% respectively for centralized SE which is much less than what was obtained in Fig. 5. 30(c), Fig. 5. 31(c) and Fig. 5. 32(c). In essence, there is much potential for applying OZA with increased sensors and reduced computation time to obtain the desired accuracy in voltage estimations.

## 5.5 Application of DEA in OZA

In this section, a special scenario is presented, combining the application of DEA and OZA methods. It was observed in the previous section that by applying Hachtel's Augmented Matrix method as local estimator in OZA, the quality of estimation can be improved over centralized DSSE in some cases. On the contrary, the DEA based DSSE shows comparatively reduced performance with regard to achieving the desired estimation quality. In both cases, there is a high possibility of the computation time becoming the major drawback. The special scenario is developed to observe whether the combination of both can bring any further improvements over the centralized DSSE, in terms of both estimation quality and computation cost [55] [81]. The randomized  $F$  - DEA based DSSE is applied by dividing the 16 node UKGDS network into four zones (appendix 9). The network splitting and zonal interactions are performed according to the proposed OZA while WLS DEA is applied as the local DSSE solver. In short, the combined method replaces only the Hachtel's Augmented Matrix method of the previously discussed OZA by the DEA as the local SE optimizer. Another significant difference from the underlying OZA method is that the interaction among zones is well controlled and executed after accomplishment of pre-defined varying numbers of generations. The outcomes are presented in Fig. 5. 33 and Fig. 5. 34 applying exponential and binary cross-over processes respectively in fifty Monte Carlo simulations.

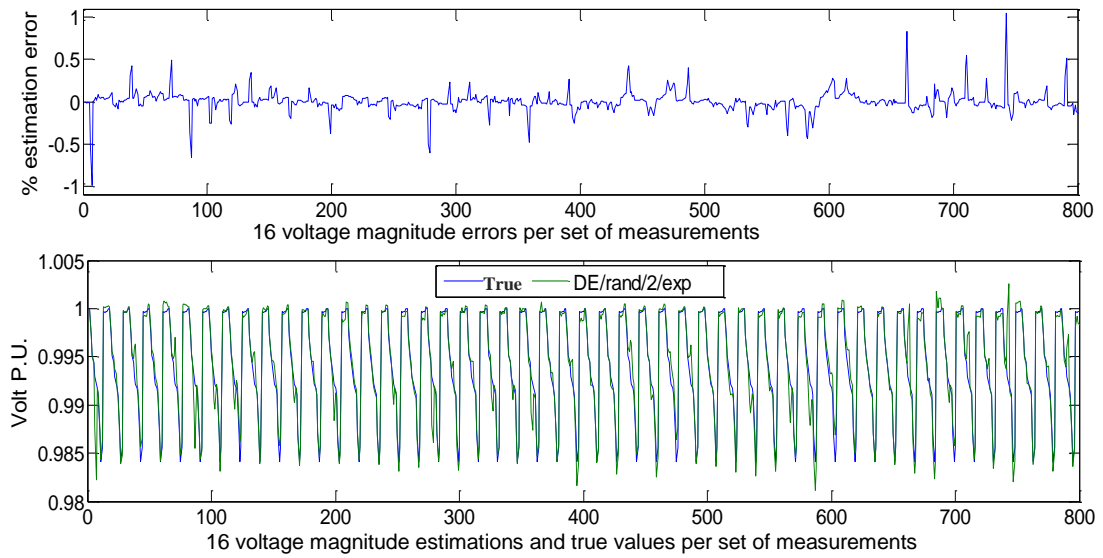


Fig. 5. 33 : Randomized  $F$  - DEA (variant DE/rand/2/exp) in scalable DSSE tool

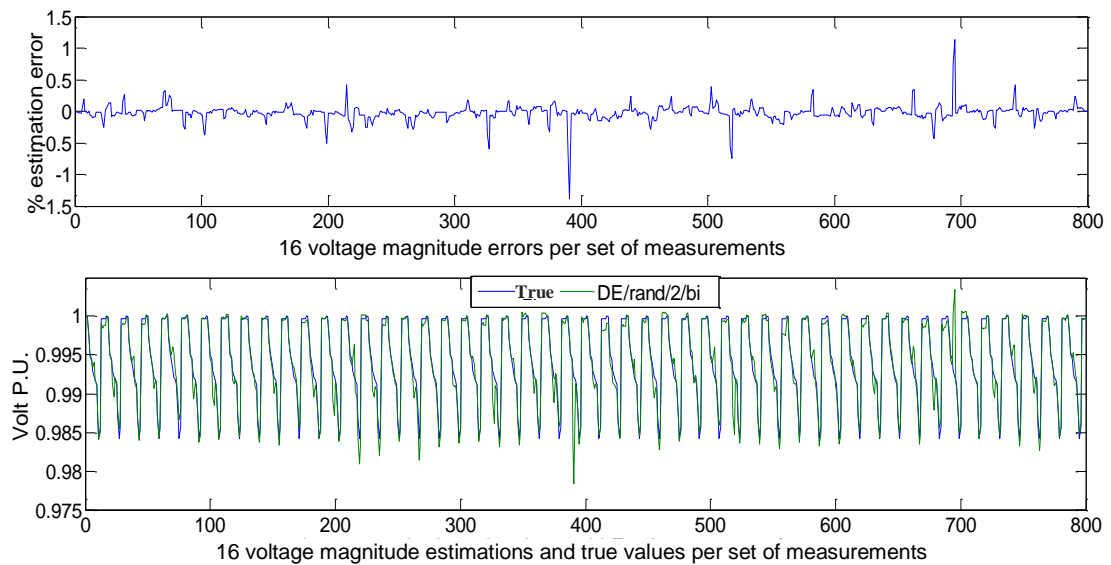


Fig. 5. 34 : Randomized  $F$  - DEA (variant DE/rand/2/bi) in scalable DSSE tool

Here, the percentage of voltage estimation errors relative to the true values, as well as the true and estimated voltage profiles, are plotted. The frequency of zonal interaction increases as the generation count approaches its maximum value. The first information exchange occurs after 820 generations, then after 640, 460, 280 and 100 further generations, leading to a total of 2300 generations. Such varying frequency of data exchange allows the DEA to achieve tentative outcomes before data exchange is executed, and enhances the data exchange rate as the all local estimator approaches convergence. Fig. 5. 33 and Fig. 5. 34 depict how close voltage estimations match the true voltage profiles for fifty sets of Monte Carlo studies plotted in continuous form. The number of

data points in these plots are therefore  $16 \times 50 = 800$  in a row. Similar to the results obtained for the centralized DSSE cases in section 5.2.6, the use of exponential and binary *CR* does not cause significant difference in the voltage estimation quality in Fig. 5. 33 and Fig. 5. 34. In both figures, the voltage estimation errors remain within 1% threshold values. The most impressive outcome is the observation that DEA incorporating OZA zone division and data exchange schemes can outperform the centralized DEA (section 5.2.6) with regard to the improvements in voltage estimation. The improvement in voltage estimation quality can be observed even by applying the fundamental DEA. In Fig. 5.35 to Fig. 5.37, the voltage estimation errors for fifty Monte Carlo simulations, applying fundamental DEA variant in centralized and combined manners are observed. The zonal interactions are provided 25 times for the generation size of 5000 and 5 times for the generation size of 1500.

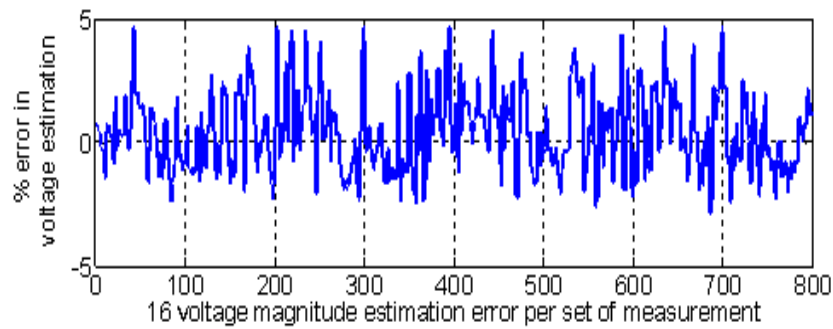


Fig. 5. 35 : Voltage estimation errors applying centralized fundamental DEA variant and 5000 generation

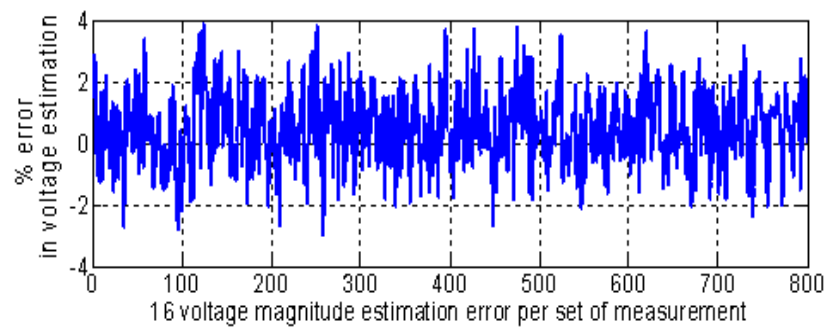


Fig. 5. 36 : Voltage estimation errors applying combined OZA-fundamental DEA variant and 5000 generation

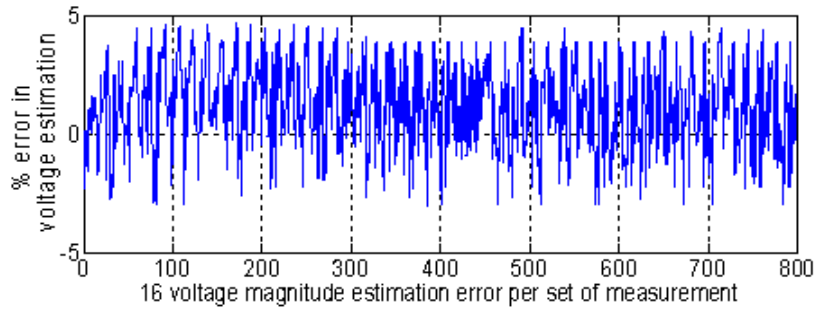


Fig. 5. 37 : Voltage estimation errors applying combined OZA-fundamental DEA variant and 1500 generation

The voltage estimation improves applying combined DEA for the exact generation size compare to the centralized DEA as illustrated in Fig.5.35 and Fig.5.36 respectively. As generation size and interactions are reduced to 1500 and 5 times respectively in Fig. 5. 37, the maximum voltage estimation errors by the combined approach are quite similar to those by the centralized DEA. It is therefore believed that the combined approach brings considerable and multidimensional benefits over the centralized DEA approach with respect to the computation cost, communication overhead and voltage estimation improvement.

## 5.6 Concluding Remarks and Discussion

Parallelization of the estimator is a potential option to achieve a linear 'execution time versus network size' relation, while the computation time increases in a quadratic manner with the size of the network in the case of most of the conventional SE optimizers. Hence, the DSSE tool is required to be operable on the parallel platform and thereby the computation time is expected to be reduced for useful practical applications. The motivation and reasoning behind the research on a scalable state estimation tool has been discussed previously in Chapter 2. In that context, this chapter proposed two methods, DEA and OZA, that are suitable for parallel processing. The objectives included: obtaining satisfactory quality of voltage estimation, reduction of computational cost and improvement in scalability of the DSSE tool. Although parallelization is one of the objectives of the proposed algorithms, this chapter essentially focused on the quality of voltage state estimation and only predicts the potentiality of the algorithms to benefit from parallel processing for computation time reduction. The parallel code is executed in a serial manner since the fundamental focus is on the performance of the algorithms with regard to maintaining required voltage estimation quality. The relevant pseudo-code of parallel

functioning logic executed in a serial manner is provided in appendix 10. The actual scalability in terms of execution time, applying parallel computation, will be observed in chapter 7, only for those cases (i.e. Hachtel's OZA) in which a greater quality of voltage estimation is achieved while the estimated operation time is expected to be reasonable.

Initially, DEA incorporating the WLS solver has been applied in centralized fashion, since DEA inherently possesses a parallel application property. A 16 node network was selected for assessments, revealing that the DEA optimizer requires a significant number of generations and high population size to generate the minimum satisfactory quality of voltage estimations. Only a successful application of DEA with regard to estimation quality and computation time on such a small network can justify its further application on a larger network and on parallel computation platforms. Whereas, the voltage estimation quality in this case has not been achieved to the level that can be obtained from the classical Gauss-Newton recursion method. The minimum expected level of accuracy was achieved once the generation size is increased to 5000, but that takes more than 1000 minutes to execute in a centralized manner. Section 5.2.8 provides further evidence with regard to the heavy computational cost of performing DEA based DSSE for the 16 node network. Hence the possibility of the expected reduced computation time is low for such a large number of generations/population size, even if applied on a parallel computation platform. As a consequence, it is considered not to be worth continuing with further application of DEA DSSE on a larger network and parallel platforms.

The novel OZA is based on estimation of overlapping zones and has proven to be a more feasible method as a scalable DSSE solution. The main advantage of the OZA is that the network division is not dependent on the location of circuit breakers or tie lines, providing great flexibility of zone division for any large network. At this stage, the network splitting process requires manual inputs and judgements, and the network diagram is required to be known to select electrically connected close lateral and sub lateral branch nodes. The zone splitting depends on the size of the network. Splitting a network into very small zones is not expected to obtain much benefit from this algorithm. The algorithm is better suited to very large networks, where the network can be split into medium to small size zones. The number of zones can depend on the number of processors available, as each processor can handle one zone. The increasing zone splitting can effectively reduce computation time however the convergence deteriorates to some extent. It is however difficult to predict how

the presence of sensors, network splitting and zonal interaction would simultaneously affect the zonal convergence and this is subject to next stage of the research. As a rule of thumb, for a smaller network (356 nodes) the reduced zonal interactions is set to 3 and for larger network (711 nodes), it is set to a larger number, i.e. 5. However it is recommended to set the minimum zonal interaction to 5 in order to ensure adequate co-ordination among zones.

The OZA integrating Hachtel's method was studied by applying: 1) only virtual measurement as constraints and 2) using both virtual and real measurements as constraints. Similar studies with the same input data using and not using real measurements as equality constraints did not show considerably different outcomes. However, better outcomes were observed in convergence using unconstrained real measurements in most cases. The case studies also imply that if the network split is increased, the algorithm requires more zonal interactions to achieve similar values of overlap mismatch as when there are fewer zone splits. The stability of the algorithm was established by observing the persistently decreasing trends of the mismatch curves with the number of zonal iterations. As a local estimation can preserve co-ordination and coherency with other zones by matching overlapping node estimation, only measurements at GSP can lead to optimum solutions and the resulting overall voltage and phase angle estimates are almost an exact representation of the centralized estimates. However this may require longer execution time with a slower convergence rate.

Further studies were performed applying more measurements in addition to one at the GSP, since it was observed that the proposed OZA with additional sensor data can provide voltage estimation up to the expected quality in less computer time. Placing real measurements, preferably on zone start, overlapping and halo nodes, can lead to better estimates and economic computation time. The stop criteria/convergence was set to a fixed number of zonal iteration to restrict the overall computation time. An example background study provided in appendix 11 demonstrates the voltage estimation errors (for the first 10 zonal interactions) when meters are available only at the GSP in the 356 node network (split into two zones). The plots in appendix 11 clearly depict that the maximum errors never remain below 1%, whereas the equivalent study with additional measurements provides voltage estimates with maximum errors less than 0.5% as illustrated in Fig.5. 30. Greater accuracy in sensor data from slack and common nodes would significantly benefit

this algorithm, since they have strong influence on the overall quality of estimation. The voltage estimation quality may deteriorate slightly compared to the centralized SE approach (only a fraction of a percentage) with reduced zonal iteration and additional sensor data. However the percentage of errors relative to true values still remains well below 1% (when maximum expected sensor data errors are 1%). The splitting into zones and reduced interactions are not applicable in order to obtain the most accurate phase angle estimates. Therefore the solution of voltage estimates with reduced zonal interaction is referred to as a 'feasible' solution (for reduced computation time) as opposed to an 'optimum' solution.

DEA was also integrated with OZA leading to the observation of important improvements, however the required number of DEA generations was still high in the special scenario. An impressive development in voltage estimation quality was observed compared to centralized DEA methods, irrespective of the types of the DEA variants. Since the computation time was still significant, requiring a generation size as large as 1500-5000, the combined method was applied only on the 16 node network. On the other hand, OZA applying Hachtel's method showed evidence of bringing benefits to both voltage and phase angle quality in a relatively shorter time span. Therefore, Hachtel's Augmented Matrix based OZA was selected to perform further studies on the parallel platform.

It is worth mentioning here that the assessments in section 5.4, applying OZA, were performed for only one measurement configuration, unlike other case studies involving extensive Monte Carlo studies. As the algorithm actually operates the process in a serial manner, the execution time is high and so is the required computer data space. That makes it infeasible to perform heavy Monte Carlo based study for OZA DSSE with the available resources. However, a rigorous Monte Carlo based study is not essential in this case, since the objective was to explore the required algorithm design to achieve similar or better quality estimation relative to the centralized estimation. That has been achieved with the measurement configuration as discussed in section 5.4. Furthermore, the outcomes are observed for various measurement sets by repetitive execution of the tool to achieve confidence over the assessment outcomes. The expected error threshold is assumed to be much lower when studies are performed for a single measurement configuration. A maximum  $|V|$  estimation error below 1% is deemed as within an expected threshold in the



cases involving rigorous Monte Carlo studies; whereas the maximum of 0.5%  $|V|$  estimation errors are considered as the desired threshold for a single measurement set.

## CHAPTER 6

### METER PLACEMENT ALGORITHM

A novel meter placement algorithm compatible with distribution networks is proposed in this chapter. The algorithm finds meter positions that offer the greatest potential benefit to the DSSE tool. The proposed algorithm also benefits from the solution procedure being applicable on a parallel platform, providing reduced computation time. The feasibility of the algorithm with regard to consistency and voltage estimation improvement is validated using case studies.

#### 6.1 Meter Placement Algorithm for MV Network

MV and LV distribution networks that are usually characterized as radial in structure, theoretically require measurements at every node or branch to provide complete observability. Since the network consists of significant numbers of nodes and branches, it is not economically and technically viable to provide such extensive sensors and communication support in distribution systems. The real measurements deployed, in practice, are expected to be comparatively small in number, and therefore the measurement locations should be selected cautiously. The remaining measurements would be provided by virtual and pseudo measurements obtained from historic load data. Under these circumstances, there arises the requirement of the development of a tool to find the most suitable locations for the deployment of new sensors. The additional sensors would assist the DSSE tool to generate high quality estimated data by limiting maximum estimation errors. In this context, a novel meter placement algorithm for the distribution system is proposed in this chapter. The proposed algorithm focuses exclusively on reducing the voltage estimation errors as far as possible, with minimum instrumentation over the network.

## 6.2 State Estimation as Two-Fold Error Reduction Process: WLS Optimization

Power system SE can be considered as a two-fold error reduction problem: reduction of measurement residuals and reduction of true residuals. Measurement residuals are defined as the differences between measurements and estimated values. True residuals are the differences between true values and the estimated values of the respective states. The true values of all the variables at a known state of the system are calculated by performing load flow studies.

### 6.2.1 Reduction of True Residuals

The main purpose of SE tools is to filter out measurement noises, consequently the gap between the true values and the estimated values are reduced. For reliable operation and control of the system, the estimated values should satisfy certain accuracy threshold values,  $th$ , that can be defined as

$$|x_{est_{ss}} - \mu_{ss}| < |\pm th_{ss}|. \quad (6.1)$$

Here,  $x_{est}$  is estimated state value,  $\mu$  is true/real value and  $ss$  represents system state index.  $|\pm th_{ss}|$  in (6.1) should be small enough so network operators can trust the estimated value ( $x_{est_{ss}}$ ) as much as the corresponding true value ( $\mu_{ss}$ ). Index  $ss$  may represent all the states of the system including voltage magnitude, phase angle, real and reactive power, current etc. Since both power and current variables can be represented in terms of voltage magnitudes and phase angles, scrutinizing true residuals of only voltage magnitudes and phase angles can provide adequate information regarding the quality of the estimated data.

### 6.2.2 Reduction of Measurement Residuals

In practice, the true values in the system are never known. Therefore, SE is modeled to reduce the measurement residuals in order to achieve (6.1). In the WLS SE minimization problem (defined in Chapter 3), the Gauss-Newton recursion process reduces measurement residuals over several iterations by updating the state vector by  $\Delta x$  (3.11) in each iteration where

$$\Delta x = (H^T W H)^{-1} H^T W [z - h(x)]. \quad (6.2)$$

$H$  is the Jacobian matrix defined as the gradient of  $h(x)$ . The state estimator can redistribute the measurement errors to keep all estimation errors (even for higher erroneous

measurements) within acceptable levels. It provides a compromise solution between the more accurate and less accurate measurements and is expected to reduce the gap between estimated and true values as expressed in (6.3).

$$\text{Max } |z - h(x_{est})|_i < \text{Max } |z - \mu|_i, \quad i = 1..M. \quad (6.3)$$

### 6.3 The Meter Placement Algorithm

The meter placement algorithm presented in this chapter aims to assist the distribution SE to improve voltage estimation quality for reliable voltage control operation. The algorithm generates outcomes in two steps: first, by short-listing potential positions to deploy meters and second, by performing rigorous assessments only on the short-listed position to find the best potential placement. The basic and most commonly used classical WLS SE model is applied in the presented approach.

#### 6.3.1 Step 1: Short-listing Potential Meter Position

A singular value of a matrix is an indication of its scaling property. Performing Singular Value Decomposition (SVD) of a  $m \times n$  matrix  $A$ , the following equation is obtained,

$$A = U \cdot \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix} \cdot V^T \quad (6.4)$$

$U$  and  $V$  are orthogonal or unitary matrices. The singular value matrix is expressed by  $\Sigma$ , which is a diagonal  $r \times r$  matrix consisted of non-negative real numbers (where,  $r = \min |m, n|$ ). The diagonal elements of  $\Sigma$  are called singular values  $[\sigma_1, \sigma_2, \dots, \sigma_r]$  conventionally arranged in descending order. Therefore the first element  $\sigma_1$  is the largest singular value of matrix  $A$ . The largest singular value is a measure of energy preserved by the matrix; it also represents the spectral norm of  $A$ ,  $\|A\|_2 = \sigma_1$ . The role of singular value is quite significant when a matrix represents a transformation to another vector space for an under-determined or over-determined problem. The linear mapping imposed by the matrix  $A$  on a vector  $x$  to vector  $y$  where  $y = A \cdot x$  can be decomposed as in Fig. 6. 1(a).  $V^T$  and  $U$  represent the gain at input and output directions respectively. The singular value matrix  $\Sigma$  scales the magnitude. Further illustration of a geometrical interpretation of the scaling property of singular value is given in Fig. 6. 1(b). Here  $m = n = 2$  for matrix  $A$ , that transforms a unit sphere to an ellipsoid. Fig. 6. 1(b) shows that the axes  $v_1$  and  $v_2$  are

rotated by  $V^T$ , then stretched by  $\Sigma$  forming an ellipse and finally  $U$  rotates the ellipse to its final direction. These considerations imply that the maximum singular value of a matrix is a measure of its ability to expand or contract a vector it is mapping. The larger/smaller is the value of maximum singular value, the higher is the scale of expansion/contraction.

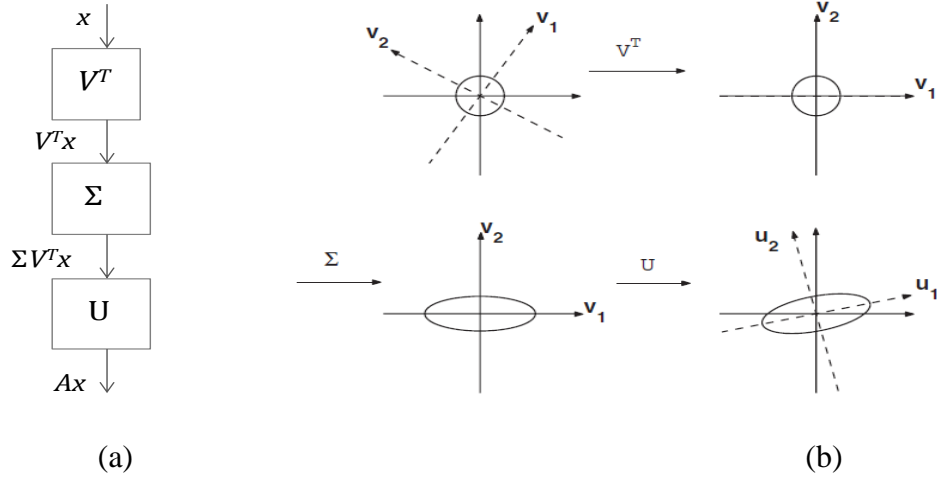


Fig. 6. 1: Mapping effect and geometrical interpretation of SVD

The scaling impact of the maximum singular value of a matrix has been exploited in the SE problem to search for potential locations to deploy additional sensors. In the SE optimization, the Gauss-Newton linearization problem enables the SE to reduce the sum of measurement residuals  $[z - h(x)]$  in every iteration of (3.11). At the end of recursion process,  $[z - h(x)]$ , as well as  $\Delta x$ , should be as small as possible. A multiplying factor,  $S_f$  is introduced here, where  $S_f = (H^T W H)^{-1} H^T W$ . Replacing  $S_f$  in (6.2), the following (6.5) is obtained.

$$\begin{bmatrix} \Delta x_1 \\ \Delta x_2 \\ \vdots \\ \Delta x_N \end{bmatrix} = S_f \begin{bmatrix} z_1 - h_1(x) \\ z_2 - h_2(x) \\ \dots \\ z_M - h_M(x) \end{bmatrix} \quad (6.5)$$

Applying SVD (6.4) to  $S_f$  in (6.5) for a network that has  $M$  measurements and  $K$  states and  $M \geq K$ , the following expressions in (6.6) and (6.7) can be written.

$$S_f = [U_1 \ U_2 \ \dots \ U_K] \begin{bmatrix} \sigma_{S1} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \sigma_{SK} \end{bmatrix} [V_1 \ V_2 \ \dots \ V_K]^T \quad (6.6)$$

$$\Delta x = (U_1 \sigma_{S1} V_1^T) \begin{bmatrix} z_1 - h_1(x) \\ z_2 - h_2(x) \\ \dots \\ z_M - h_M(x) \end{bmatrix} + \dots + (U_K \sigma_{SK} V_K^T) \begin{bmatrix} z_1 - h_1(x) \\ z_2 - h_2(x) \\ \dots \\ z_M - h_M(x) \end{bmatrix} \quad (6.7)$$

Here,  $\sigma_{S1} > \sigma_{S2} > \dots > \sigma_{SK}$ . The term  $[z - h(x)]$  consists of various types of measurements having different scales of magnitudes. The voltage measurements are in kV and close to unity in the per unit system; whereas the power measurements are in MW/MVAr and they may vary down to small fractional numbers in per unit values. As  $[z - h(x)]$  is not normalized but rather calculates algebraic differences between the measurements and estimated values, it does not provide complete information about the degree of reduction of measurement residuals corresponding to each measurement. When alternative measurement configurations are to be considered and it has been checked which configuration is most effective to reduce the measurement residuals; instead of examining  $[z - h(x)]$  directly, it would be more useful to observe the transformation effect of  $S_f$  on  $[z - h(x)]$  for various measurement configurations. Since the Gauss-Newton WLS method approaches convergences by reducing measurement residuals, the mapping effect on  $[z - h(x)]$  is expected to be contracting in the last iteration. The maximum singular value  $\sigma_{S1}$  of  $S_f$  can be considered as a measure of the mapping effect on  $[z - h(x)]$  as explained in Fig. 6. 1. If  $\sigma_{S1}$  has a small value, the contraction effect on residual vector will be greater.

In addition to the scaling property, the  $\sigma_{S1}$  is also useful as an indication of the sensitivity of a measurement configuration. In (6.6) and (6.7), the perturbation of  $\Delta x$  is mostly affected by the largest component,  $U_1 \sigma_{S1} V_1^T$  from the SVD analysis. The greatest changes in  $\Delta x$  will occur to the direction of  $V_1^T$  and in proportion to  $\sigma_{S1} U_1$ . As both  $V_1^T$  and  $U_1$  set the direction of changes,  $\sigma_{S1}$  is the magnitude of sensitivity of estimation to  $[z - h(x)]$ . The estimated values become less sensitive to the residuals by achieving smaller values of  $\sigma_{S1}$ . The sensitivity of a measurement configuration is important for the distribution system SE as there will be only few real measurements and placing a sensor on a highly sensitive node may generate erroneous estimation when the sensor provides inaccurate information. Although selection of a more sensitive node to achieve greater effect from the more accurate data expected from real measurements apparently seems rational, that may be true only when there exists an adequate amount of real measurement data. Having an erroneous input can be compensated by many other real measurement values even though

the flawed data is introduced at a sensitive node. This would not be the case with regard to distribution system scenarios. In the presence of fewer real measurements compared to the amount of pseudo-measurement data, the DSSE optimizer will be greatly influenced by those greater weighted real time data. The probability of having a strong negative effect from erroneous real measurement values is much higher when that is generated at a sensitive node. Therefore, selecting a node having less sensitivity to the SE outcomes is more reasonable for measurement placement when real time measurements data are limited. In essence, selecting a measurement configuration that gives the minimum value of the maximum singular value assists the estimator to achieve reduced estimation errors in two ways: (a) by reducing measurement residuals to achieve good convergence and (b) by selecting less sensitive measurement setups to minimize the adverse effect of erroneous data. The summary of the procedure of short-listing potential meter positions is as follows.

Suppose, there are  $C$  available candidate meter positions ( $M_C$ ) to introduce sensors. The algorithm adds sensors in one place at a time and then performs SE; followed by recording  $\sigma_1$  in the last Gauss-Newton iteration step. Addition of sensors at each candidate position thus corresponds with a singular value and a vector of  $(\sigma_1)_i$  with  $i = 1$  to  $C$  is generated after performing SE for all  $M_C$ . After surveying the magnitude of each element of  $(\sigma_1)_i$ , the algorithm detects those having smaller values.  $P$  number of locations are suggested out of  $M_C$  as potential meter positions ( $M_P$ ), since the measurement setup including real time data from those positions, attributes smaller values of  $\sigma_1$ . Background surveys imply that consideration of the 20-30% of all  $M_C$  that generate comparatively smaller values of maximum singular value, is sufficient for the most beneficial meter location ( $M_B$ ) to occur in the short-listed  $M_P$  vector.

### **6.3.2 Step 2: Selection of the Potential Meter Position**

Placing one meter,  $M_T$  (when  $M_T \in M_P$ ) at a time at each of  $P$  positions, the estimation quality before and after placement can be compared for a known system. The subscript  $T$  denotes the meter position under observation. In distribution system SE, there are a large number of pseudo-measurements which are typically power injections having high error variance. Extensive Monte Carlo based studies are very effective to quantify the performance of an SE process, being fed by uncertain measurements. Simulated measurement sets with various noise distributions can be created in the Monte Carlo

studies to evaluate the overall quality of estimation. In this context, 100 Monte Carlo studies are performed before and after placing  $M_T$  to quantify the effect. The evaluation is performed only for  $M_P$  which have already been assessed as contributing towards the SE tool achieving reduced measurement residuals. Each Monte Carlo study is performed with the same input data except the additional information of the new measurement,  $M_T$ . As the other measurement data remains unaltered, the comparison studies attribute the effects on estimation quality in a similar operation state. In these studies, measurements are simulated from the true values and relevant error variance calculated from the expected deviation from the true values. On completion of the comparison studies for all  $M_P$ , the measurement placement having the most positive impact on the quality of estimation is finally selected. The quality of estimation is evaluated in terms of accuracy of voltage estimation after placing the new meter(s). It is desirable to achieve minimum estimated voltage deviation  $V_{dev}$ , with respect to the true value.

$$V_{dev} = 100 \times \frac{|x_{est_{ss}} - \mu_{ss}|}{\mu_{ss}} \quad (6.8)$$

$V_{dev}$  can be calculated for a known system where the true values of all the states are calculated from the load flow analysis. In this work, meter placements are found such that the voltage estimation errors are reduced to the greatest possible extent.

Performance of every  $M_T$  placement is evaluated by three criteria; which are referred to as ‘Performance Evaluation Parameters’ (PEPs). These are calculated in (6.9), (6.10) and (6.11), where Monte Carlo simulations are performed  $MC$  times ( $MC=100$  in this algorithm) and  $V_{dev}$  is calculated from (6.8). The PEPs are classified as primary and secondary PEPs.

### ***Primary PEPs***

*PEP1*: The  $V_{dev}$  for each node of the network is calculated and the maximum voltage estimation errors are traced before and after including  $M_T$ . The same procedure is executed for all Monte-Carlo simulations, therefore there will be a set of maximum  $V_{dev}$  values comprising 100 elements before adding  $M_T$  and an equivalent set of maximum  $V_{dev}$  after the placement of  $M_T$ . The mean of each set of maximum 100  $V_{dev}$ , expressed as  $(V_{dev})^{mean\ of\ max}$ , is calculated. Hence, there will be a pair of  $(V_{dev})^{mean\ of\ max}$  that indicate the quality of voltage estimation before and after considering  $M_T$  measurements.



The entire process performed for all elements of  $M_P$ .  $(V_{dev})^{mean\ of\ max}$  is termed as PEP1 and expressed as below.

$$(V_{dev})^{mean\ of\ max} = \frac{\sum_{j=1}^{MC} [Max\{(V_{dev})_i\}]_j}{MC}, \quad i = 1..N \quad (6.9)$$

*PEP2*: Similar to the previous case, the  $V_{dev}$  is calculated for each node of the network for 100 Monte Carlo simulations before and after adding  $M_T$ . The mean value of all  $V_{dev}$ , expressed as  $(V_{dev})^{mean}$  are obtained in this case. Hence, there will be a pair of  $(V_{dev})^{mean}$  before and after inclusion of  $M_T$ . The similar procedure is executed for all elements  $M_P$ . The  $(V_{dev})^{mean}$  values indicate the effect of placement of  $M_T$  on voltage estimation quality and are termed as PEP2.

$$(V_{dev})^{mean} = \frac{\sum_{j=1}^{MC} \sum_{i=1}^N \{(V_{dev})_i\}_j}{N \times MC}, \quad \begin{array}{l} i = 1..N \\ j = 1..MC \end{array} \quad (6.10)$$

### **Secondary PEPs**

*PEP3*: After calculation of  $V_{dev}$  for each node of the network, the maximum voltage estimation error is traced for every Monte Carlo simulation. At the end of 100 Monte Carlo studies, two sets of maximum  $V_{dev}$  vectors each containing 100 elements will be obtained before and after the placement of  $M_T$ . The largest of 100 maximum  $V_{dev}$ , referred to as  $(V_{dev})^{max\ of\ max}$  represents PEP3 for each set. Similar to other PEP values, a pair of  $(V_{dev})^{max\ of\ max}$  are obtained and the procedure is continued for all  $M_P$ .

$$(V_{dev})^{max\ of\ max} = Max[Max\{(V_{dev})_i\}]_j, \quad \begin{array}{l} i = 1..N \\ j = 1..MC \end{array} \quad (6.11)$$

The Monte Carlo based study is illustrated in Fig. 6. 2, where it is shown that each potential placement actually involves 200 Monte Carlo simulations altogether to assess the effect of an added meter. The algorithm creates a set of data array, performs SE, updates the existing measurement setup with the data associated with  $M_T$  and executes SE again. These steps are executed for 100 times followed by storing a set of corresponding PEPs

values. The entire process is repeated for other elements of  $M_P$  and the associated PEPs are stored for further assessments.

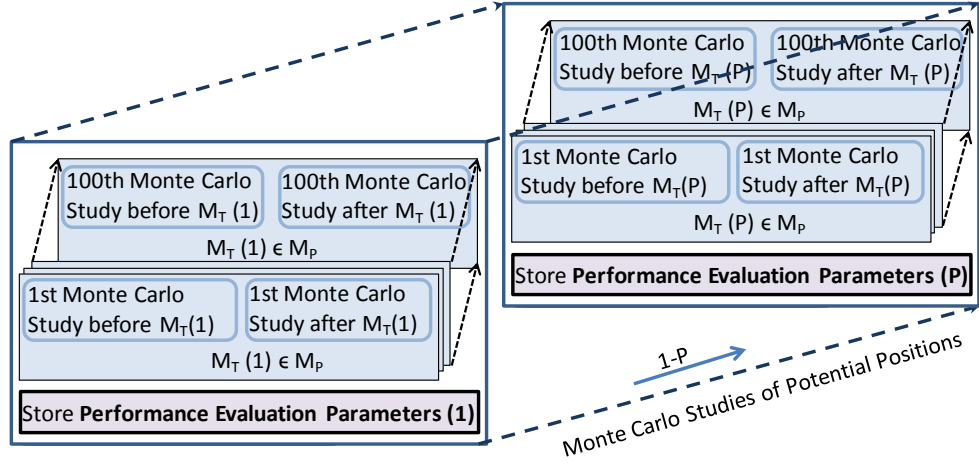


Fig. 6. 2: The Monte Carlo assessment of  $M_T$

Since, the three PEP values are calculated for both the cases of before and after placing  $M_T$ , there will be  $M_P$  pairs of PEP1, PEP2 and PEP3 on completion of Monte-Carlo simulation studies. The effect of addition of  $M_T$  on the voltage estimation quality can be explicitly understood by calculating improvement factor,  $Q_{fac}$  (6.12) for every  $M_T$  placement. The  $Q_{fac}$  quantifies the percentage of reduction of voltage estimation errors after placing a new meter.

$$Q_{fac} = \frac{100 \times (M_{err} - (M+1)_{err})}{M_{err}} \quad (6.12)$$

$M_{err}$  = Performance evaluation parameters [  $(V_{dev})^{mean\ of\ max}$ ,  $(V_{dev})^{max\ of\ max}$ ,  $(V_{dev})^{mean}$  before placing  $M_T$

$(M+1)_{err}$  = Performance evaluation parameters [  $(V_{dev})^{mean\ of\ max}$ ,  $(V_{dev})^{max\ of\ max}$ ,  $(V_{dev})^{mean}$  after placing  $M_T$

Consequently, a positive  $Q_{fac}$  value in (6.12) indicates an improvement and the negative counterpart indicates deterioration of voltage estimation quality after placing the meter. A smaller value of  $(M+1)_{err}$  than  $M_{err}$  is always desirable in the meter placement problem. The location for which the three PEPs collectively achieve the maximum positive  $Q_{fac}$ , is considered as the most beneficial position ( $M_B$ ) to place a new meter. However, values of  $Q_{fac}$  for primary PEPs (PEP1, PEP2) are prioritized over the secondary PEPs (PEP3) in selection criteria for a few specific scenarios described as follows.

- If the cumulative  $Q_{fac}$  values of three PEPs are very close for more than one  $M_P$ , in such case the one provides the larger  $Q_{fac}$  for the primary PEPs is selected as  $M_B$ .
- When the largest  $Q_{fac}$  values of primary PEPs are significantly greater than the secondary PEPs, while the secondary PEPs having one of their best values, the corresponding position is selected as  $M_B$ .

Excluding the above cases, the most beneficial position  $M_B$  is chosen that has the best possible improvement in both primary and secondary PEPs.

### 6.3.3 The Complete Algorithm with the Application of Parallel Processing

The complete meter placement algorithm is depicted in Fig. 6. 3. Here,  $M_{EX}$  stands for the already existing metered locations and  $M_C$  represents the candidate placement nodes/branches where no real measurement already exists, therefore  $M_C : M_{EX} \cap M_C = \emptyset$ . In the background studies, the true values of all the variables at a known state of the system is calculated by performing load flow studies. Feeding SE real or simulated measurements of such known cases, the true residuals can be calculated and hence the accuracy of estimated values can be verified. In Fig. 6. 3,  $M_C$  are the virtual- and pseudo-measurement nodes and their connecting branches where applicable. As discussed, the proposed meter placement algorithm selects a set of  $M_P$  locations ( $M_P \subset M_C$ ) that contributes to reduced measurement residuals in the first step and then  $M_B$  ( $M_B \in M_P$ ) that reduces the true residuals most in the second step. After obtaining  $M_B$ , the meter is added at the newly found position, the existing meter vector ( $M_{EX}$ ) is increased by one and the algorithm searches for the next position. The loop continues while economical feasibility persists or the desired level of accuracy is achieved, as defined in (6.1), which are set as convergence criteria.

Execution of 100 times P number of Monte Carlo assessments twice (before and after the addition of a new meter) results in 200P times of execution of the SE process, which is considerably computationally expensive. Each set of 100 Monte-Carlo assessments for a  $M_T$ , performs the same task with different input data and is independent of assessments of other elements of  $M_P$ . It is therefore possible to distribute two or more assessment slots over a multi-processor system (Fig. 6. 3). Application of parallel processing will be

particularly beneficial in order to reduce computation burden for large networks and multiple measurement options involving increased computation.

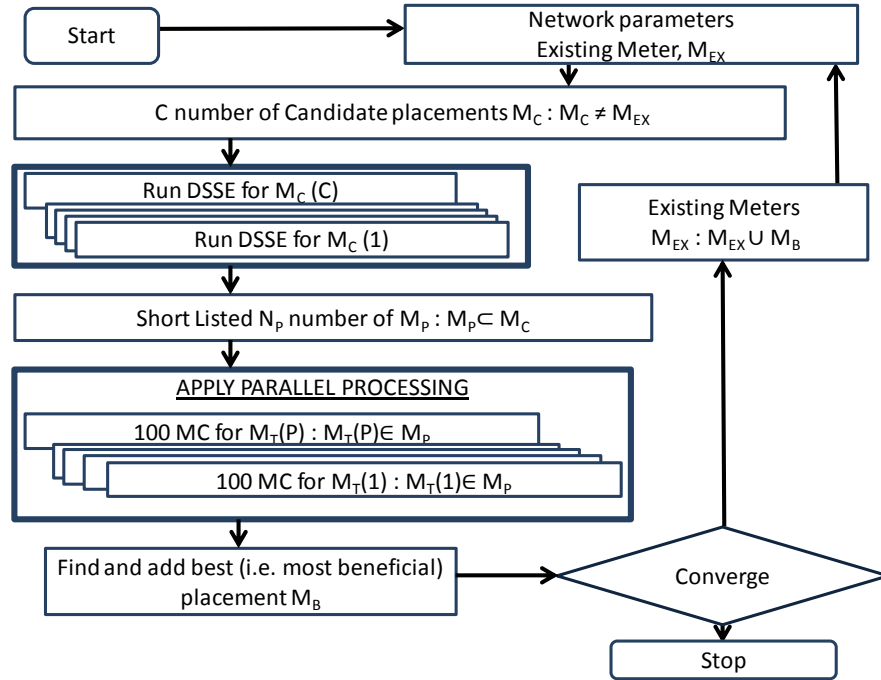


Fig. 6. 3 : Meter placement algorithm flowchart

## 6.4 Tests and Results

The algorithm is applied on a 77 node UK distribution system model network (appendix 2). The expected measurement errors are assumed to have 3-sigma confidence limits of 1% for real sensors and 50% for pseudo measurements. Power flow sensors on line 76 and voltage magnitude sensor on node 1 are assumed to be available initially. The meters are placed sequentially as stated in Table 6. 1. Here, a set of meters include one voltage magnitude and one pair of real and reactive power flow meters. The voltage measurement is assumed to be located at either the sending end of the flow measurement branch or the receiving end if a voltage meter already exists at the sending node. The voltage measurement being accompanied by the power measurements are observed to provide robust improvement in voltage estimation quality.

Sets of Meters	Power Flow Meter at Branch	Voltage Magnitude Meter at Node
1	76	1

<b>2</b>	52,76	53,1
<b>3</b>	51, 52, 76	2,53,1
<b>4</b>	1, 51, 52, 76	3, 2,53,1
<b>5</b>	4, 1, 51, 52, 76	6, 3, 2,53,1
<b>6</b>	37, 4, 1, 51, 52, 76	39, 6, 3, 2,53,1
<b>7</b>	53, 37, 4, 1, 51, 52, 76	54, 39, 6, 3, 2,53,1
<b>8</b>	7, 53, 37, 4, 1, 51, 52, 76	9, 54, 39, 6, 3, 2,53,1
<b>9</b>	54, 7, 53, 37, 4, 1, 51, 52, 76	55, 9, 54, 39, 6, 3, 2,53,1
<b>10</b>	10, 54, 7, 53, 37, 4, 1, 51, 52, 76	12, 55, 9, 54, 39, 6, 3, 2,53,1
<b>11</b>	11, 10, 54, 7, 53, 37, 4, 1, 51, 52, 76	13, 12, 55, 9, 54, 39, 6, 3, 2,53,1
<b>12</b>	67, 11, 10, 54, 7, 53, 37, 4, 1, 51, 52, 76	69,13, 12, 55, 9, 54, 39, 6, 3, 2,53,1
<b>13</b>	15,67, 11, 10, 54, 7, 53, 37, 4, 1, 51, 52, 76	17, 69,13, 12, 55, 9, 54, 39, 6, 3, 2,53,1
<b>14</b>	55, 15,67, 11, 10, 54, 7, 53, 37, 4, 1, 51, 52, 76	56, 17, 69,13, 12, 55, 9, 54, 39, 6, 3, 2,53,1
<b>15</b>	69, 55, 15,67, 11, 10, 54, 7, 53, 37, 4, 1, 51, 52, 76	57, 56, 17, 69,13, 12, 55, 9, 54, 39, 6, 3, 2,53,1
<b>16</b>	68, 69, 55, 15,67, 11, 10, 54, 7, 53, 37, 4, 1, 51, 52, 76	70, 57, 56, 17, 69,13, 12, 55, 9, 54, 39, 6, 3, 2,53,1
<b>17</b>	26, 68, 69, 55, 15,67, 11, 10, 54, 7, 53, 37, 4, 1, 51, 52, 76	28, 70, 57, 56, 17, 69,13, 12, 55, 9, 54, 39, 6, 3, 2,53,1
<b>18</b>	56, 26, 68, 69, 55, 15,67, 11, 10, 54, 7, 53, 37, 4, 1, 51, 52, 76	58, 28, 70, 57, 56, 17, 69,13, 12, 55, 9, 54, 39, 6, 3, 2,53,1

Table 6. 1: Sequential addition of voltage magnitude and power flow meters

Meter positions are selected first by short-listing  $M_P$  and then observing which one provides the best value of  $Q_{fac}$  for both primary and also in secondary PEPs. Fig. 6. 4, Fig. 6. 5 and Fig. 6. 6 present an illustration of how the algorithm locates  $M_B$  based on the values of  $Q_{fac}$ . The example in Fig. 6. 4 shows values of  $Q_{fac}$  for all  $M_P$  to find the 1<sup>st</sup> placement. The cumulative bar plots apply different colours to represent three PEPs. As a set of power flow measurement on line 52 (along with corresponding voltage measurement) achieves the best  $Q_{fac}$  in all three PEPs, it is selected as 1st meter position. The cumulative plot of three PEPs for that position achieves the maximum height.

Similarly, Fig. 6. 5 and Fig. 6. 6 show the best  $Q_{fac}$  is obtained when meters are placed at branch 51 and 1 (and connecting nodes).

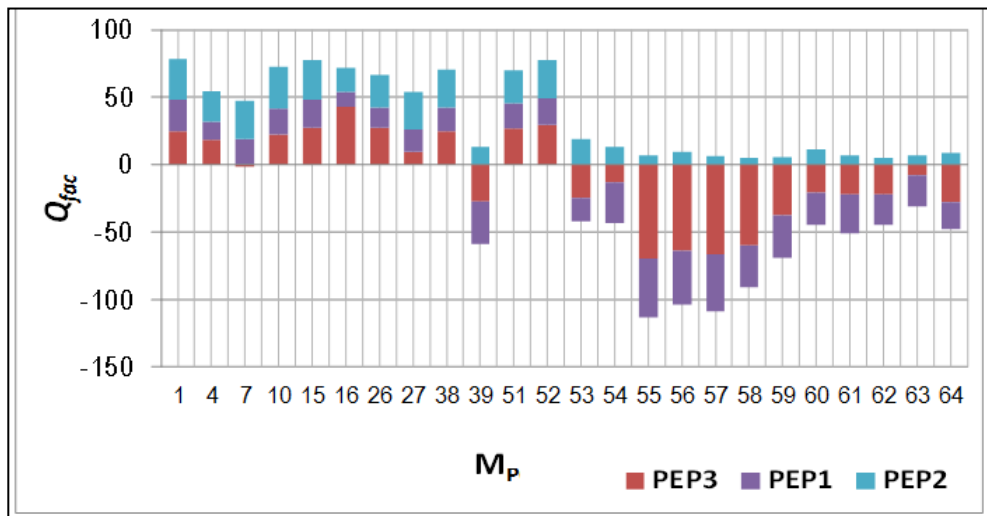


Fig. 6. 4: Evaluating improvement ( $Q_{fac}$ ) in PEPs for all  $M_p$  to find 1st meter position

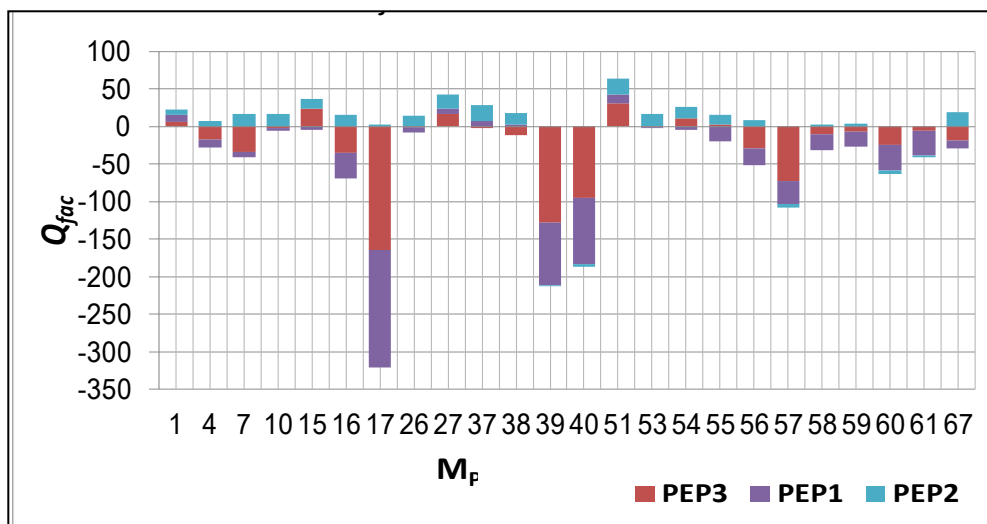


Fig. 6. 5: Evaluating improvement ( $Q_{fac}$ ) in PEPs for all  $M_p$  to find 2nd meter position

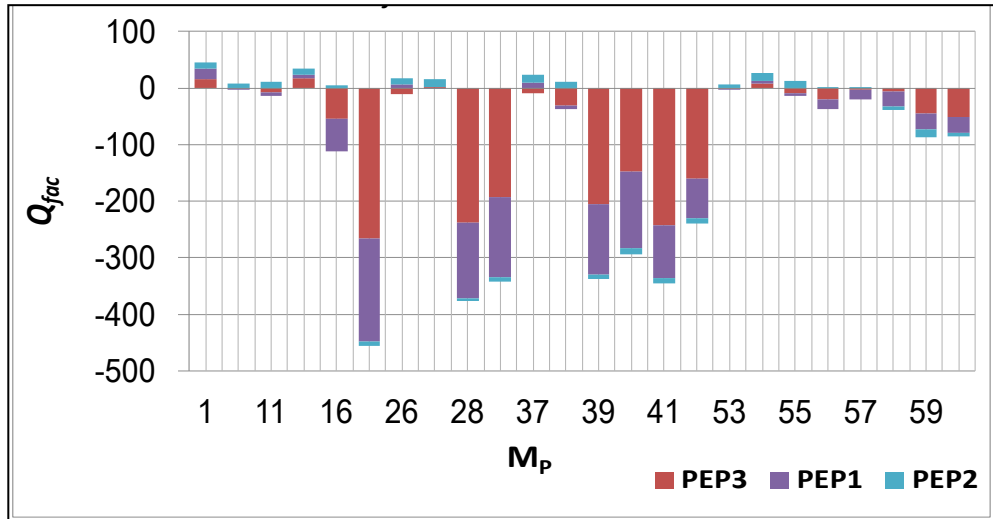


Fig. 6. 6: Evaluating improvement ( $Q_{fac}$ ) in PEPs for all  $M_p$  to find 3rd meter position  
 In the following sections, four features are observed to demonstrate the performance of the algorithm and the quality of results.

1. Voltage estimation errors with sequential meter placement
2. Effect of unexpected large error measurements from faulty sensors
3. Consistency of *singular value based* and *100 Monte Carlo stochastic* study based algorithm
4. Application of the algorithm outcomes on other SE process

#### 6.4.1 Voltage Estimation Errors with Sequential Meter Placement

The placement of meters is continued up to 17<sup>th</sup> placement to observe how the  $V_{dev}$  reduction trend changes. As observed, it is not financially worthwhile to spend more on installation of meters that are not providing sufficient consequentially benefit due to slow improvement of voltage estimation quality as shown in Fig. 6. 7 and Fig. 6. 8. Therefore, the algorithm is allowed to stop adding meters at that point. In fact, the overall voltage estimation error does not reduce much after placement of the 9<sup>th</sup> meter (when total number of meters in the network is 10). Fig. 6. 7 and Fig. 6. 8 show the mean and maximum  $V_{dev}$  from extensive Monte Carlo studies with sequential implementation of meters as shown in Table 6. 1. The reason for performing such a rigorous Monte Carlo assessment is to be confident that placement of meters are providing improvement as expected with measurements affected with possible levels of noise. Studies depict that the large reduction of the maximum  $V_{dev}$  occurs in the first 8 sets of meter placements; from 9<sup>th</sup> to 17<sup>th</sup> it does

not change much in Fig. 6. 7. The mean  $V_{dev}$  reduces constantly with addition of more meters in Fig. 6. 8, the rate of reduction becomes slower from the 9th placement.

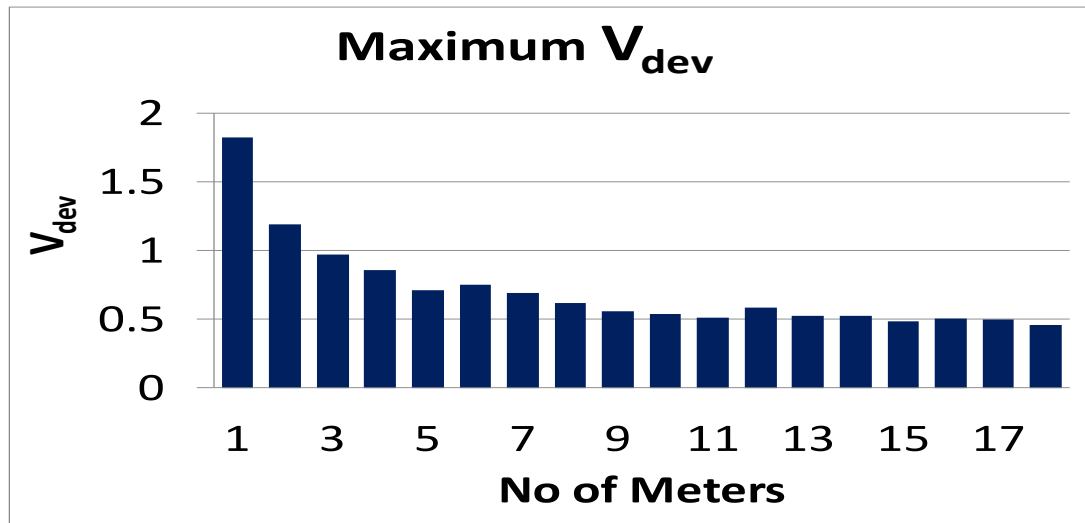


Fig. 6. 7: Maximum  $V_{dev}$  applying classical WLS

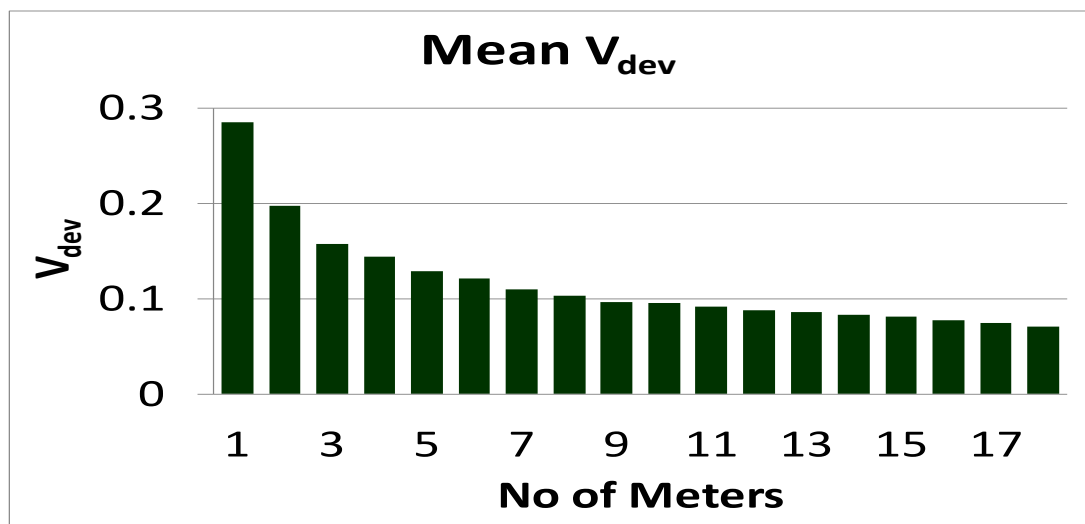


Fig. 6. 8: Mean  $V_{dev}$  applying classical WLS

In the presented case studies, normal random noise is added to the measurements from the new meter using MATLAB operator *normrnd*, while a maximum of 1% measurement errors are expected with 99.7% confidence. As part of background surveys, it is also observed the effect of the expected highest possible erroneous data originated at the added meter. The measurements of the new meter is forced to generate data fixed as  $(1 \pm 0.10) \times true\ value$  for all 100 Monte Carlo studies. The  $Q_{fac}$  values are then compared with the equivalent parameter when randomized error data is obtained from the added meters. Since the algorithm tends to select less sensitive nodes, the quality of voltage estimation is



observed not to be affected much in spite of the highest expected erroneous readings originating constantly from the newly added meter. Appendix 13 depicts the relevant graphical representation of  $Q_{fac}$  when the power flow measurements for  $M_P$  are fixed at  $(1 \pm 0.10) \times \text{true power value}$ . The outcomes of the background studies suggest that, the newly added meter is expected to provide the same quality of performance in the presence of maximum expected errors measurements, which might not be the case if the most sensitive node was chosen to deploy the new meters.

#### **6.4.2 Effect of Unexpected Large Error Measurements from Faulty Sensors**

The effect of unexpected large error measurements from faulty sensors is assessed in Fig. 6. 9 and Fig. 6. 10. Maximum and mean  $V_{dev}$  are plotted for 100 Monte Carlo studies assuming one faulty sensor at a time of eighteen power sensors (Table 6. 1). The blue and red bars represent  $V_{dev}$  before and after adding a large error to the sensor respectively. Each Monte Carlo study simulates higher error (up to 50%) to a power measurement considered as real measurement. The horizontal axis denotes the branch number with faulty power sensors. According to the principle of the algorithm, the less sensitive nodes are selected for placing meters to minimize the effect of erroneous measurements on the voltage estimation quality. In Fig. 6. 9 and Fig. 6. 10, the measurement data fed into the estimator before and after adding large errors are the same except for the erroneous ones; therefore the differences in  $V_{dev}$  can be attributed exclusively to the unexpected bad data. As seen in Fig. 6. 9 and Fig. 6. 10, the level of blue and red bars are similar in most cases. That provides sufficient evidence that the presence of gross power errors in the new measurement data has a minor effect on voltage estimation in most cases.

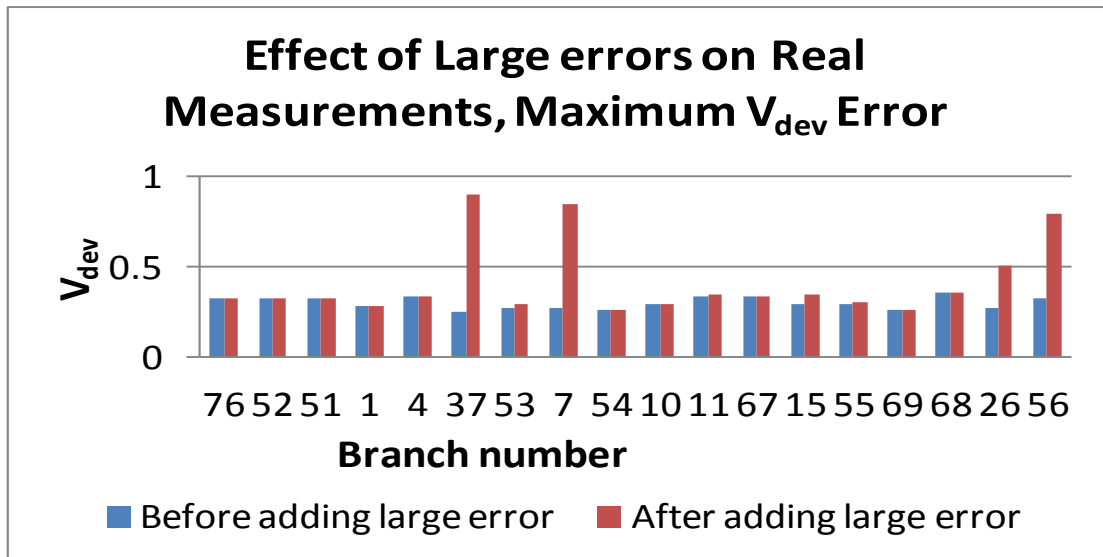


Fig. 6. 9: Effect on maximum  $V_{dev}$  for faulty sensor

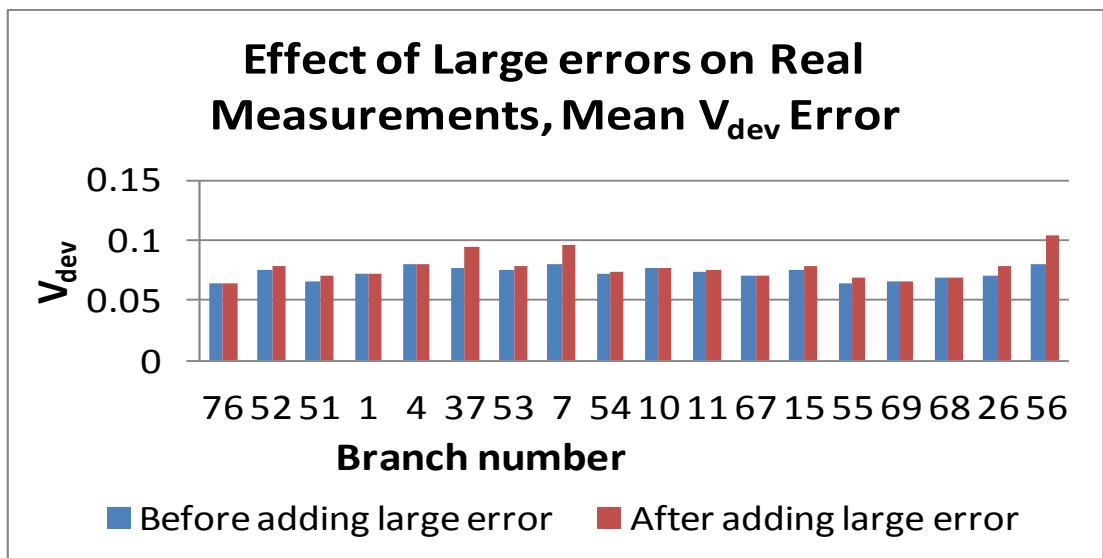


Fig. 6. 10: Effect on mean  $V_{dev}$  for faulty sensor

### 6.4.3 Consistency of Singular Value Based and 100 Monte Carlo Study Based Algorithms

The complete algorithm combines two different approaches to find the most beneficial meter position.

- Find  $M_P$  out of  $M_C$  applying the *Singular value based* approach. Here  $M_P \subset M_C$  and  $M_C \neq M_{EX}$ .
- Perform *100 Monte Carlo stochastic evaluation* approach only for  $M_P$ .

Tests show an excellent consistency of selecting the most beneficial meter position following two approaches independently. Here the ten most beneficial meter positions are listed in Table 6. 2 to perform a comparison study following:

- *Singular value based approach*, involving one random SE study. This step is performed actually in the original algorithm to find  $M_p$ .
- *100 Monte Carlo stochastic evaluation approach*, performing 100 SE studies. In the actual algorithm this step is performed only for  $M_p$ , however here it is being calculated for all  $M_C$  to compare and order ten most beneficial meter placements.

Table 6. 2 shows the outcomes for the 6th, 12th and 18th placements. The large bold numbers represent branches that are common findings in both approaches. This indicates that the position has been suggested as the possible best placement by the *100 Monte Carlo stochastic studies* approach, can also be found in the top ten positions by the *singular value based* approach. This establishes an excellent coherency of the outcomes applying two different approaches within the algorithm. Such outcomes are expected as the reduction of measurement residuals acts very effectively to reduce true value residuals. However, the impact of noise on the algorithm is not very predictable, causing the slight discrepancies. The results, applying the two approaches, are shown in Table 6. 2.

<b>18th Meter</b>	<b>18th Meter</b>	<b>12th Meter</b>	<b>12th Meter</b>	<b>6th Meter</b>	<b>6th Meter</b>
1 MC $\sigma_1$ based	100 MC Stochastic	1 MC $\sigma_1$ based	100 MC Stochastic	1 MC $\sigma_1$ based	100 MC Stochastic
57	<b>56</b>	56	27	26	15
<b>56</b>	27	57	26	1	<b>53</b>
58	5	55	<b>15</b>	52	10
59	16	58	55	37	54
5	38	38	39	4	2
41	8	<b>15</b>	12	10	38
33	13	2	38	54	7
29	69	1	8	<b>53</b>	67
28	4	27	68	15	8
39	67	67	2	2	26

Table 6. 2: Comparison of outcomes from two steps applying independently

#### 6.4.4 Application of the Algorithm Outcomes on Other SE Processes

The meter placement algorithm is entirely based on the classical WLS SE criteria due to the fact that most SE algorithms work on the principle of reduction of the weighted measurement residuals. In this section two different WLS based approaches are applied to validate that the proposed algorithm performs equally well using other WLS based SE techniques. The SE techniques used here are- Orthogonal (QR) Decomposition and Hachtel's Augmented Matrix approaches, in addition to the classical WLS method for comparison. Both QR and Hachtel's methods solve the WLS optimization criteria by applying different solution processes discussed detail in Chapter 3. Hachtel's Augmented Matrix method takes the virtual measurements as constraints, improving matrix ill-conditioning. The QR method provides higher robustness to solve the WLS problem. Both the algorithms avoid forming the 'normal equation'. Therefore multiplying factor,  $S_f$  (6.2) is not explicit in their solution processes. The deployed meters affect estimation quality for various solution processes in similar fashion with each placement, as can be seen in Fig. 6. 11 and Fig. 6. 12. Here, meters are implemented sequentially, as stated in Table 6. 1, in the 77 node network applying WLS, QR and Hachtel's approaches. The maximum and mean  $V_{dev}$  from 100 Monte Carlo simulations are illustrated below.

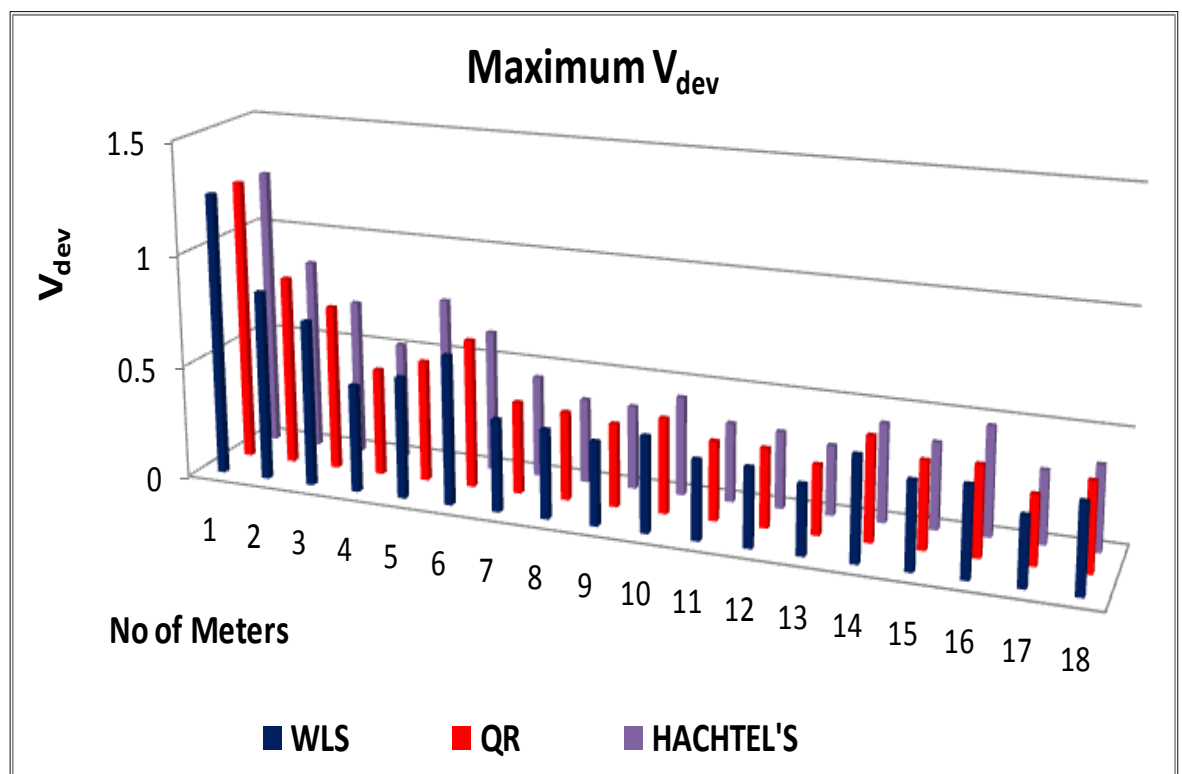


Fig. 6. 11: Maximum  $V_{dev}$  applying WLS, QR and Hachtel's methods

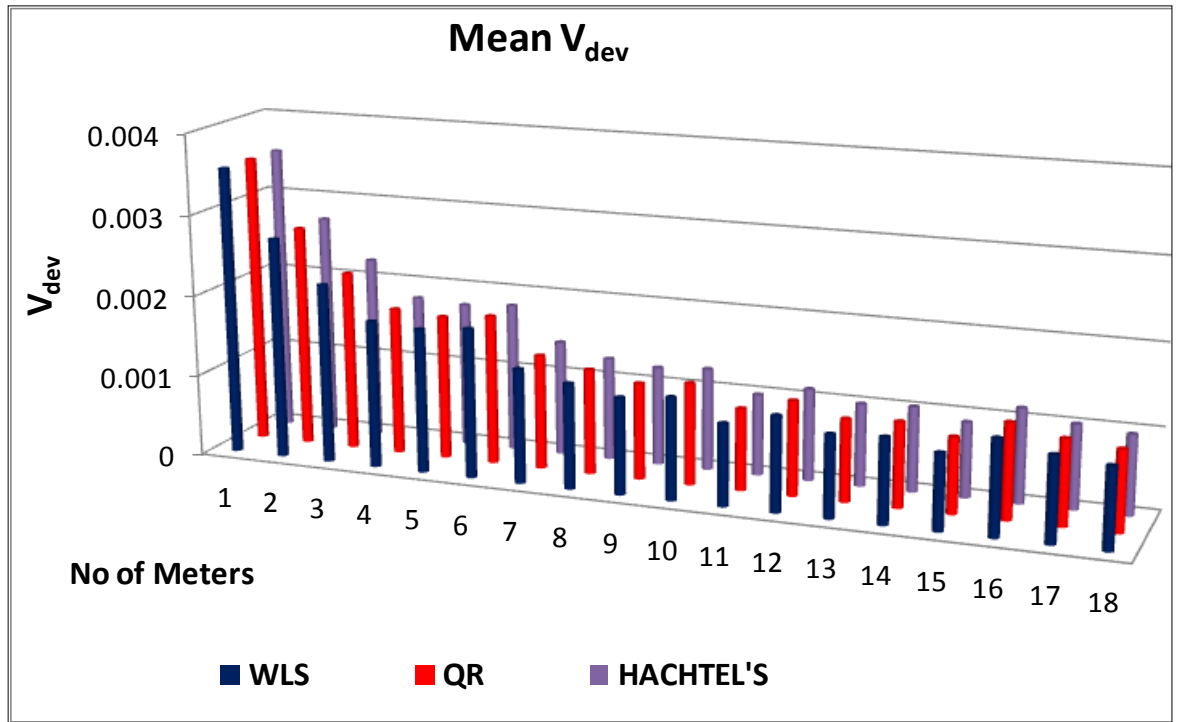


Fig. 6. 12: Mean  $V_{dev}$  applying WLS, QR and Hachtel's methods

## 6.5 Discussion and Concluding Remarks

In the chapter a semi-stochastic meter placement algorithm suitable for parallel processing is established and applied to a power distribution network model. The key objective is to reduce the voltage estimation error deploying limited numbers of meters, and that has been achieved successfully applying the proposed algorithm. Meters are placed sequentially, unless the improvement in maximum voltage estimation errors becomes negligible, as at that point further introduction of meters would not be economically feasible. The algorithm is executed in two steps: first, by short-listing potential locations and then second, selecting the best i.e. the most beneficial locations out of the short-listed positions. The short-listed locations are obtained by setting a criterion of reduced measurement residuals, achieved by inspecting the singular values for various measurement configuration. The suggested meter placements are further assessed by a very high number of Monte Carlo studies for simulated measurement sets to calculate maximum and mean voltage estimation errors. Integrating such extensive Monte Carlo studies provides more confidence that the estimation quality improves as expected and is not being affected by any abnormalities due to presence of pseudo measurements. The algorithm provides the information of how estimation accuracy is improving with new meter placement. In this way network operators have the flexibility to decide when to stop placing meters to achieve a good accuracy as well as to stay within budget.

The voltage estimation errors are observed to decrease consistently with the addition of new meters in most cases. Although the randomly distributed errors are expected to be  $\sim 0$  to  $\sim 1\%$ , it is also observed how estimation quality is affected if the new meter provides data containing maximum expected errors (i.e. 1%) all the time. Furthermore, the effect of faulty sensors that generate errors up to 50%, nevertheless being weighted as high as the real measurements, is examined. The maximum and mean voltage estimation errors remain negligibly affected in most cases of both expected and unexpected significantly erroneous measurements from the new sensors. It is believed that the insignificant effect of larger error data from the added meter can be attributed to the low sensitivity of the meter positions. It is further shown that the most potentially beneficial meter positions obtained from the algorithm improve voltage estimation quality at a similar level, when applying different types of WLS based approaches e.g. classical WLS, Orthogonal (QR) decomposition and Hatchel's augmented matrix method. The condition number is observed to remain consistently within the range of  $10^{13}$  to  $10^{15}$ , that indicates the robustness of SE problems for any measurement configuration. Application of parallel processing is discussed and applied in Chapter 7.

The meter placement method applying parallel processing can be also important to select an already installed meter to connect to the SCADA system. It is not usually required to send the information to the control room from every smart meter connected to each node when the system is heavily equipped with smart meters. By applying the parallel processing enabled meter placement method, the critical node point from a number of closely located meters can be selected to dispatch measurement data.

In this work, each of the real measurement sets consist of a pair of power flow and a voltage magnitude sensors, since the combination of a voltage sensor along with the power measurement data can provide more robust outcomes. The proposed algorithm is however capable of incorporating any types of measurements including power injection, current flow and phasor measurement units. The newly introduced meters improve the accuracy of voltage magnitude estimation exclusively in the presented case studies. As a matter of fact, the principle of the algorithm can also be applied to improve the accuracy of other state variables, e.g. phase angles, power flow etc. by defining and observing the related quality factors. For example, if the algorithm is required to improve both voltage and phase angle

estimation data, it will calculate three pairs of PEPs for voltage magnitude and another three pairs for phase angle related assessments. Hence, there will be six quality factor values to signify the improvement of estimation quality. The algorithm requires to perform additional calculations for the relevant PEPs and  $Q_{fac}$  in step 2, provided the improvement of other states are required. The first step remains unaltered in all cases. It is also possible to assess multiple measurement options for each potential location by observing corresponding  $Q_{fac}$  values. Furthermore, the general nature of the algorithm enables it to be applied to any network topology, including radial and non-radial systems.

# CHAPTER 7

## APPLICATION OF PARALLEL PROCESSING

This chapter presents the practical application of parallel processing of the presented scalable distributed DSSE tool and meter placement algorithm introduced in chapter 5 and chapter 6 respectively. The DSSE tool is applied on a cluster computer system and the parallelization of the meter placement algorithm is tested on a multi-core machine.

### 7.1 Application of Parallel Processing to the DSSE Algorithm

The scalable DSSE tool introduced as the Overlapping Zone Approach (OZA) is applied on a cluster computer system and the scalability with respect to zone division is observed. The Hachtel's Augmented Matrix method is used as the local estimator. Tests are performed on one real and two model distribution networks utilizing real and simulated data respectively. In this case, a co-ordination of MATLAB and C++ settings is applied, while the MATLAB part performs the core DSSE algorithm and the C++ section executes data co-ordination and transmission for parallelization. Although the C++ coding segment relevant to the parallelization process was not explicitly contributed by the author, the author was required to develop the MATLAB DSSE tool providing compatibility and effective co-ordination with the cluster computer system. The core MATLAB coding developed for the assessment in chapter 5 was modified significantly for that. The input data is generated into .MAT files providing a generalized format for all types of networks to feed into the parallelized code. The entire MATLAB coded section is developed by the author and the C++ coded section integrating with the cluster computer system was developed by Oxford University<sup>2</sup> as one of the key contributions in HiPerDNO project.

#### 7.1.1 Principle of Operation

Each processor can deal with one zone, therefore the number of processors required is equal to the number of zones the network is split into. A parallel Message Passing Interface (MPI) frame is used to enable exchanges of overlap and halo data with neighbouring zones at the end of local estimation. Application of MPI allows execution of independent



estimation by zones as well as implementation of the data collection-transmission procedure to process the matching overlapping data [82]. Each overlap and halo dataset and zone/processor are tagged with a specific number to determine the destination processor(s) of the dataset, i.e. 'which data should go to which processor'. Since all local estimators will not necessarily achieve convergence concurrently, the local stop/convergence parameter is defined for each zone when the overlap mismatch values of two consecutive zonal iterations remains within a predefined threshold value. A global stop/convergence is considered to be achieved when all processors attain local convergence simultaneously. This occurs when the local stop/convergence parameters all indicate their corresponding zones have achieved convergence.

In this process, there are two MPI synchronization-communication steps:

1. The first requires exchange of overlap and halo data between overlapping zones. This step starts when all processors accomplish their local estimation.
2. The second evaluates the global stop criterion based on all overlap matches or the maximum allowed number of zonal iterations having been attained.

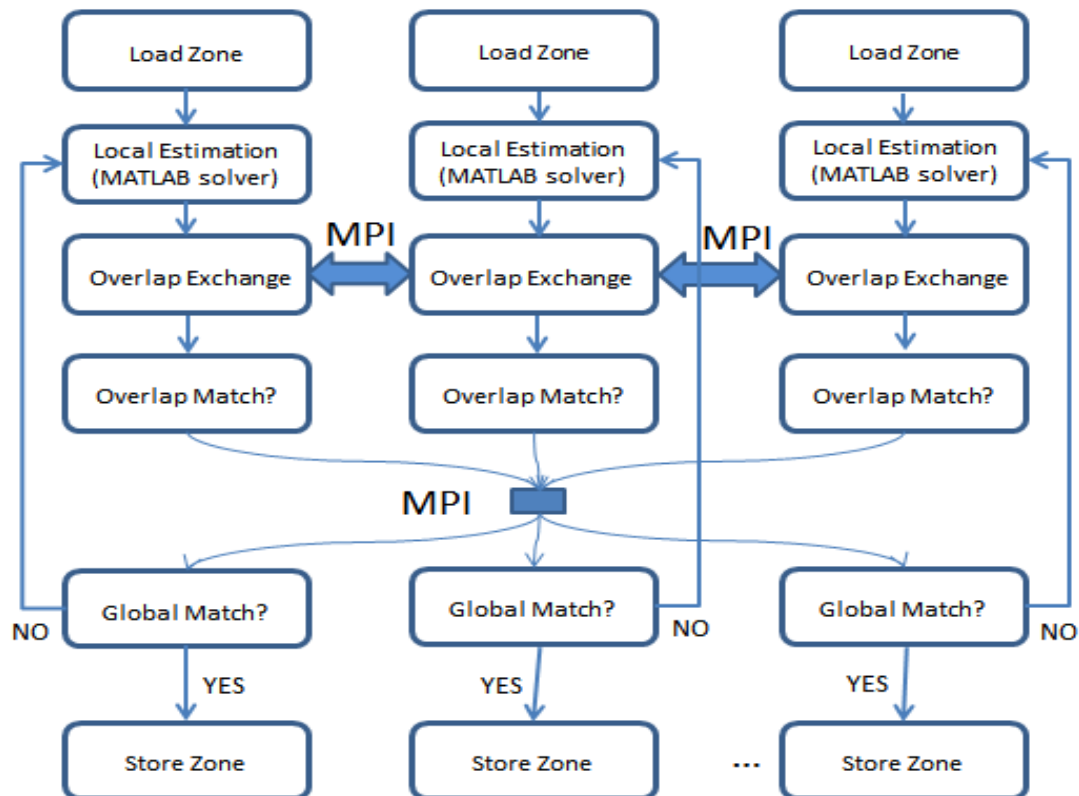


Fig.7. 1: Overlapping zone approach on MPI framework<sup>2</sup>

<sup>2</sup> This work was supported by the Oxford e-Research Centre, Oxford University as part of the HiperDNO project.

When a global stop is achieved, the final estimation is stored in the form of .MAT file. Fig.7. 1 demonstrates how the MPI skeleton provides data communication and storage support in the flow chart of the algorithm [82].

### **7.1.2 Performances of DSE on HPC Platform**

This section consists of three parts-

1. Initial application on UKGDS 356 Node Networks
2. Advance application on UKGDS 711 Node Networks
3. Application on Real Data of an EG Network

The first two case studies are performed on UKGDS networks which have been developed as UK model distribution networks for research purposes. The 711 network is a simulated network, concatenating two 356 UKGDS networks. The network parameters are obtained through load flow studies. The 3<sup>rd</sup> case is part of the EG distribution network (in Slovenia) connected to Primiskovo HV-MV transformers. The network parameters and measurement data were provided as part of HiPerDNO project. The centralized DSSE for 356, 711 and 411 node networks takes 2, 11 and 9 seconds respectively. However, after increasing the number of voltage measurements, the 411 node network takes about 11 seconds, as more iterations are required to achieve convergence. Measurement values are simulated for the 356 and 711 node networks. The EG network provides some real measurement data otherwise real and pseudo-measurements in all networks are simulated with an error margin up to around 50%. The network information, diagram and sensor positioning are stated in appendices 3, 8, 9 and 12.

A load flow study has been performed beforehand with available information for the three networks. The primary and secondary state values obtained from the load flow simulation, are assumed as the true values of a particular state of the system. The available real measurement data provided by the EG DNO could offer limited utilization as those were not located on the overlapping or halo nodes in many cases. The pseudo-measurement and the required real measurement data are simulated based on the true values obtained from the load flow studies.

The initial case studies for 356 node networks reflect the preliminary expectations regarding OZA. Initially, heavy zonal overlap was allowed, to ensure strong co-ordination among zones. However, investigation shows that better estimation can be achieved by

making the zones lightly overlapped as confirmed in chapter 5. Consequently, the number of overlapping nodes among zones are reduced in further studies for 711 and 411 node networks.

### 1. Case study 1: Initial Test on 356 Node Networks

The objective of this study is to observe the mismatch when zones are heavily overlapped (up to 100 overlapping nodes between zones, as shown in appendix 9) and the network is split into two, three and four zones. The 356 nodes network contains one  $|V|$  and a pair of PQ injection measurements at grid supply point, a pair of virtual measurements and 354 pairs of pseudo-measurements. Except for the overlapping and halo node measurement data, identical measurement configurations and parameter data are used for the case studies pertaining to various zone splits. The convergence characteristics as well as the computation time are analysed when the maximum allowed exchange of overlapping data is set to 500 times. The mismatch values and computation time [82] for different zone splits are plotted in Fig.7. 2. The network diagram showing zone splitting is included in appendix 9 and parameter details are provided in appendix 8.

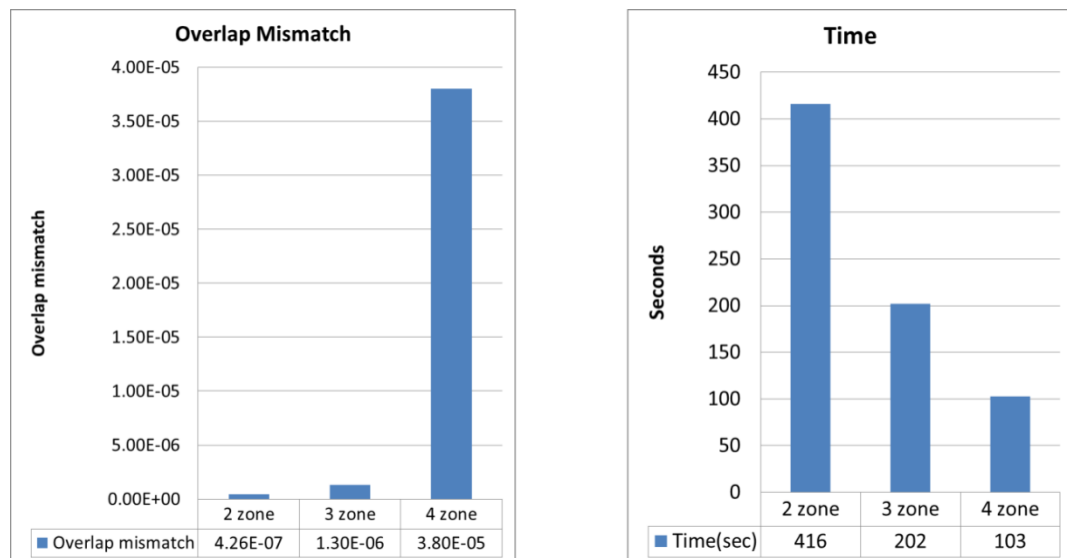
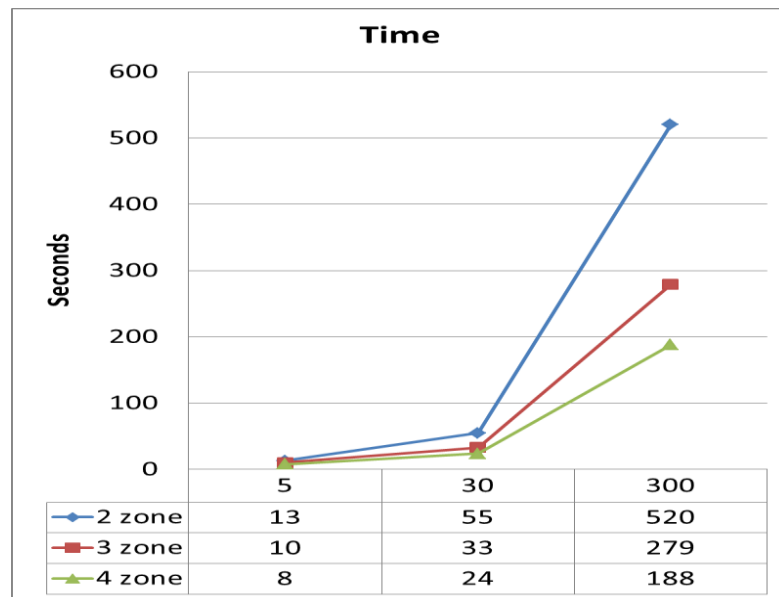


Fig.7. 2 : Execution time and overlap mismatch of 356 node network for 500 times zonal interaction applying 2, 3 and 4 zone splitting

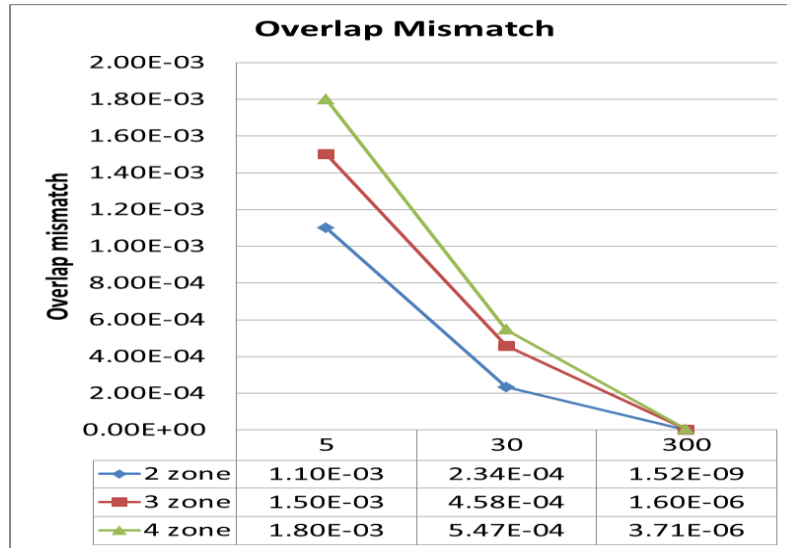
The overlap mismatch values increase with the number of zones the network is divided into. The mismatch values with for 500 zonal information interchanges are significantly small, which gives an indication of strongly coordinated estimates by each zone. The execution time reduces to almost one fourth ( $=416/103$ ) for two to four zone splitting, however the overlap mismatch values deteriorate.

## 2. Case study 2: 711 Node Networks

Extensive studies are performed on 711 node simulated UKGDS networks (appendix 8 and 9) where zones are lightly overlapped (2-3 overlapping nodes between zones, as shown in appendix 9). The network is split into 2, 3 and 4 zones (appendix 9). The zonal information exchange is allowed to execute 5, 30 and 300 times. In cases of greater number of zonal interactions ( i.e. 30 and 300 times), only one set of  $|V|$ , PQ sensors at the beginning of the feeder is considered. The sensors are assumed to be more numerous when the zones are allowed to exchange information only for a few times (i.e. 5). There are 15  $|V|$  and 15 PQ injection measurements assumed to be available, two of which are virtual power measurements. The sensor locations are shown in appendix 12. As reduced level of information exchange costs the coherency of the estimates, more sensors deployment at common nodes and halo nodes (nodes connected to common nodes) are required to compensate the loss of coherency and to achieve a feasible solution. The execution time and overlap mismatch values are shown in Fig.7. 3.



(a)



(b)

Fig.7. 3: (a) Execution time and (b) Overlap mismatch of 711 node network for 5, 30 and 300 times zonal interaction applying 2, 3 and 4 zone splitting

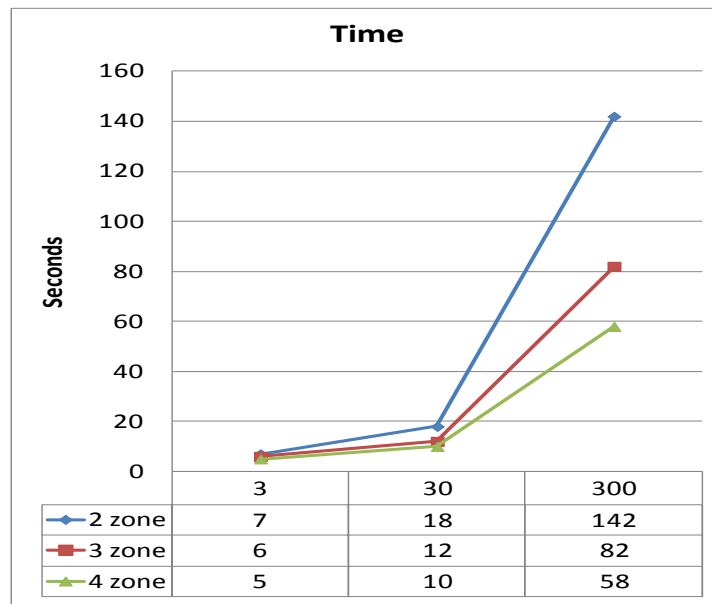
As seen in the 356 node network case study, the execution time reduces considerably with increased zone splitting particularly when interactions/information exchanges among zones are higher. In case of 711 node networks in Fig.7. 3(a), the execution time is greatly reduced by increasing the network split from 2 to 4 zones, especially when there are a large number of interactions (reduces to 1.6, 2.3 and 3.7 times for 5, 30 and 300 times zonal interactions respectively). Fig.7. 3(b) shows overlap mismatch is reduced to significantly small values when information exchange is allowed up to 300 times. However the computation time increases significantly as in Fig.7. 3(a) at the expenses of better convergence (smaller overlap mismatch) in Fig.7. 3(b).

### 3. Case study 3: EG Feeder Network

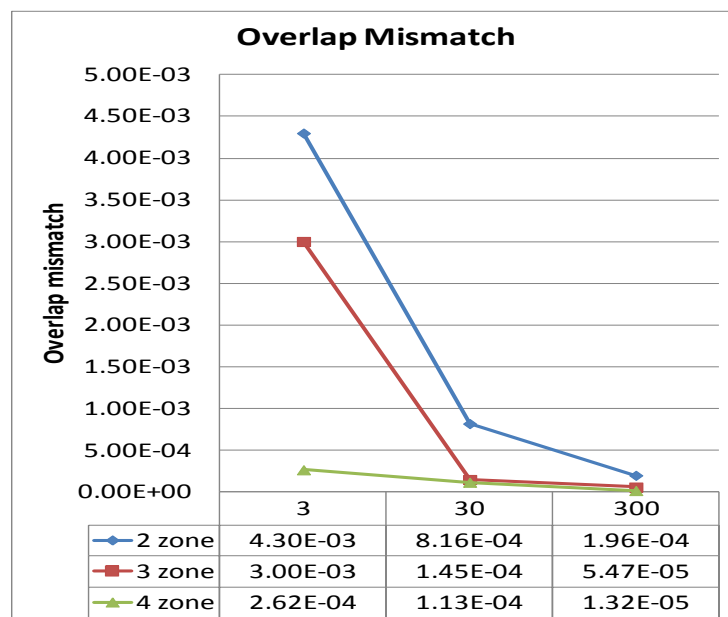
The overlapping zone approach is applied on part of the EG 20 kV network which is connected to the 110 kV HV side through two Primiskovo 110/20 kV transformers. The network contains 411 nodes and 412 branches. The Primiskovo feeder information is retrieved from EG network data provided by the HiPerDNO project. The node information, network connection etc. are described in appendix 8.

As addition of more measurements improves the convergence and execution time, the Primiskovo feeder is assumed to be equipped with twelve voltage and six power injection measurements that include zone start nodes and a few halo nodes (appendix 12). The

measurements are simulated values obtained except for node indexed 1, 2 and 3 as the associated MV sensor data are provided by EG.



(a)



(b)

Fig.7. 4 : (a) Execution time and (b) Overlap mismatch of 411 node network for 3, 30 and 300 times zonal interaction applying 2, 3 and 4 zone splitting

Fig.7. 4 shows the improvement of overlap mismatch and computation time with zone splitting and zonal iterations/interactions. Fig.7. 4 (a) shows that with more zone splitting, the computation time reduces significantly for higher number of zonal iterations. The two zone splitting takes around 142 seconds for 300 time zonal interactions (iterations) whereas four zone splitting takes less than half of that (~58 seconds) for the same number of zonal

iterations. Fig.7. 4 (b) shows that the convergence or overlap mismatch decreases rapidly for 3 to 30 zonal interactions and the decrease gets slower from 30 to 300 interactions for two and three zone splitting. The mismatch values do not change significantly in four zone splitting. Although overlap mismatch values are a minimum when zonal interactions are 300 times, the execution times are significantly high compare to the centralized DSSE. The outcomes infer that the increase of network splitting may improve convergence when additional sensors are deployed in some cases.

## **7.2 Application of Parallel Processing on Meter Placement Algorithm**

The Monte Carlo based evaluation process in searching for potential meter positions, is the most computationally expensive part of the proposed meter placement algorithm in chapter 6. Taking into consideration the extensive computation cost that the algorithm may confront due to considerable data sizes from a large distribution network, the application of parallel processing is proposed as in Fig. 6. 3. In this chapter the parallel application is tested on a four core machine applying MATLAB specific parallelizing operators.

The computationally and data-intensive Monte-Carlo based assessment process is parallelized applying MATLAB parallel computation support using multi-core processors. Applying Parallel For-loops (PARFOR) and Single Programme-Multiple Data (SPMD) operators, the process is parallelized without need of Compute Unified Device Architecture (CUDA) or MPI programming model. MATLAB parallel computation toolbox provides a number of MATLAB engines known as workers/'labs' to execute applications locally on a multi-core desktop as depicted in Fig.7. 5. The algorithm is developed in MATLAB (2010a) on an Intel core i5- quad-core machine. The four core machine is able to assign task to up to four "labs" where each "lab" represents a processor.

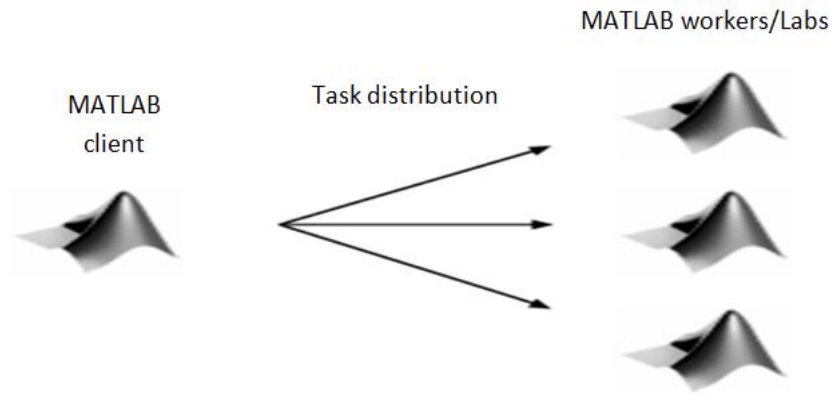


Fig.7. 5: Work distribution of MATLAB workers/'labs'

There are some functional differences in the application of PARFOR and SPMD. PARFOR provides less control over parallel operation whereas SPMD has more control over data distribution and operation. As PARFOR performs task distribution automatically, it is easy to implement in parallel processing. SPMD requires data distribution and management defined by users therefore implementation is somewhat complex. PARFOR distributes iterations of a for loop over available workers provided each iteration is independent of others. In case of meter placement Monte-Carlo assessment part, the parallel processing enabled MATLAB automatically distributes assessment computation for different elements of  $M_P$  to the 'labs' as depicted in Fig.7. 5. On the other hand, SPMD requires assigning which elements of  $M_P$  should go to which 'lab' for the assessment as stated in Fig.7. 6. Each 'lab' is assigned similar number of  $M_C$  for assessment when applying SPMD. Therefore, each 'lab' has its own local best. Three local bests are collected from each 'lab' and sent them to the client or one of the chosen 'lab' for the final evaluation and selection.



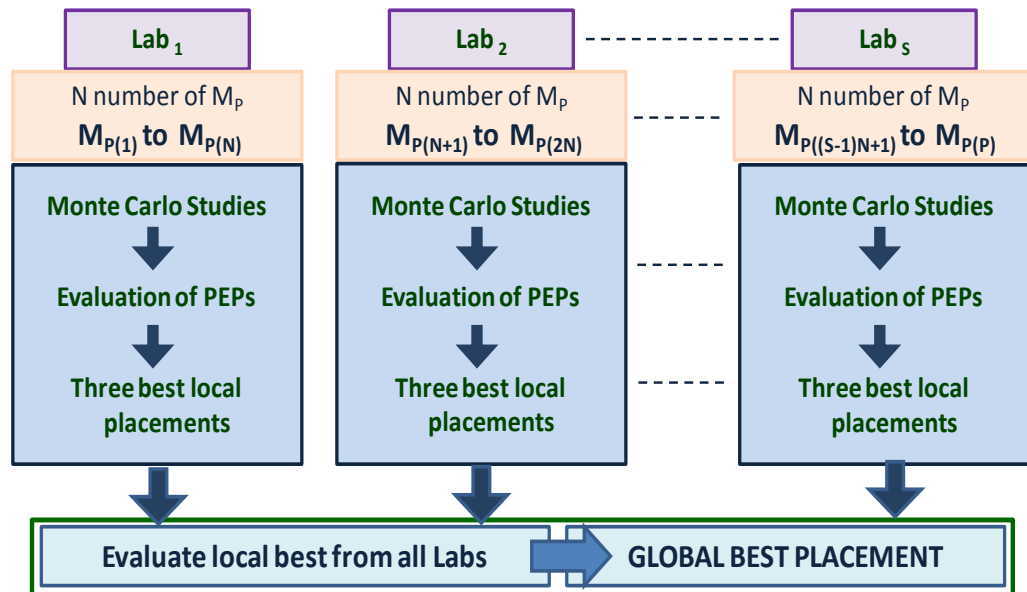


Fig.7. 6: Application of SPMD

Engaging more workers does not necessarily guarantee reduced computation time. It may run slower on multiple workers than it does on one desktop computer when the time it takes to transfer data and initialize a task is far greater than the actual time it takes for the worker to evaluate the task function. In the case study, the number of elements in  $M_P$  is approximately twenty four. The 100 Monte-Carlo assessments for  $M_P$  can be distributed to four 'labs'. The computation time and CPU utilization engaging different number of 'labs' are plotted in Fig.7. 7 and Fig.7. 8 respectively. The heading 'lab' 'None' represents when no 'lab' is created i.e. without applying parallel process (apart from MATLAB's inherent multi-threading in a multi-core machine). Both operators, in this case, perform at similar efficiency level. The computation time is reduced most deploying two 'labs' although CPU utilization is maximum applying four 'labs' with respect to the non-parallel process. The code runs slower on three and four 'labs' than it does on two. This is occurring because the task becomes too fine-grained. As more workers get involved, the task size reduces to each worker and the task start up and stop time become significant relative to the task run time. For this particular case, engaging two workers gives the best computation time.

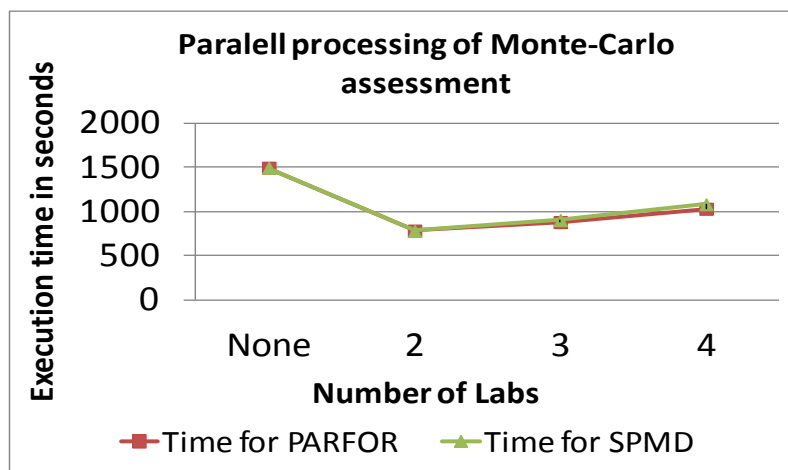


Fig.7. 7: Execution time applying varying number of 'labs'

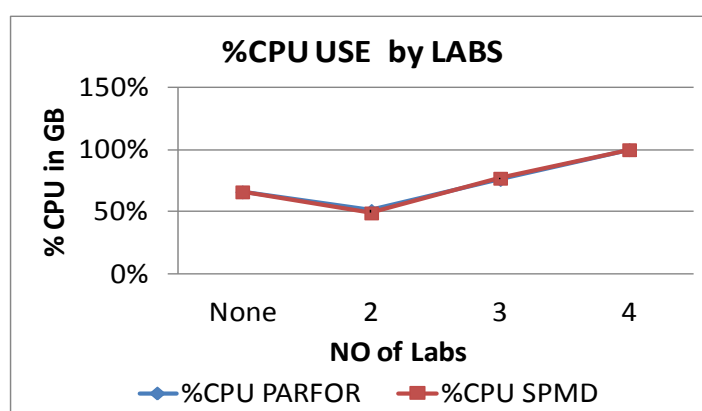


Fig.7. 8: CPU usage applying varying number of 'labs'

### 7.3 Concluding Remarks and Discussion

The core OZA is developed on the MATLAB solver, suitable to be integrated inside a parallel framework which is developed and coded with the support of the Oxford University research group<sup>2</sup>. Application of parallel processing on DSSE Hachtel's Augmented Matrix method based OZA shows that, increasing zone splitting leads to significantly reduced computation time. In the case of 711 and 411 node networks, allowing reduced zonal interaction and network split into four zones, the computation time required is 8 and 5 seconds respectively. Comparing to centralized DSSE for these networks which takes about 11 and 9 seconds respectively, the reduction in computation time is 1.4 and 2 times that of centralized DSSE for feasible voltage estimate solutions. That signifies the improvement of scalability to some degree.

As observed, increase in zone split may deteriorate overlap mismatch values i.e. convergence to some degree. The test cases imply that if network split is increased, the algorithm requires more zonal interactions to achieve similar values of overlap mismatch, than for fewer zone splits in most cases. As the reduction of computation time is crucial, further instrumentation on overlapping and connected (halo) nodes is required in the case of reduced zonal interaction and increased zone splits; which is observed in the cases of 5 zonal iterations for 711 node network and the EG network. However, it is difficult to predict how the presence of sensors, network splitting and zonal interaction simultaneously would affect the zonal estimation. As a rule of thumb, for the smaller network (411 nodes) the number of zonal interactions is set to 3 and for the larger network (711 nodes), it is set to a larger number, i.e. 5 times. However it is recommended to set the minimum zonal interaction to 5 times in general to ensure adequate co-ordination among zones.

In the Monte Carlo based assessment for meter placement algorithm, the parallel processing has been tested on a multi-core machine applying two different parallelization operators. PARFOR requires insignificant amendment to non-parallel code apart from code slicing, however the SPMD operator requires distribution of  $M_P$  to different workers/'labs' providing more control over parallel distribution. The two functions provide similar results with regard to computation time and CPU usages as shown in Fig.7. 7 and Fig.7. 8. The computation time is reduced to almost half by applying two 'labs', however more 'labs' increases overheads resulted in increased processing time in both cases. The presented parallel processing application can also be applied to a computer cluster or a grid computing service for very large networks.

## CHAPTER 8

# CONCLUSION AND FUTURE WORK

### 8.1 Overall Discussion and Conclusion

The SE tool can estimate the actual system state utilizing metered data which are prone to small or large errors, provided that the SE optimizer is fed with sufficient real measurement data. The development of distribution system SE tools is receiving increasing attention as a consequence of the requirement for enhanced network observability, in order to provide secure operation and control of future active distribution networks. As passive distribution system status has in general been predictable and consistent, real time network monitoring was not considered as an essential requirement for the purpose of distribution system operation and control. Existing SE algorithms are mostly suitable for transmission networks that are usually characterized by meshed systems, and benefiting from redundant real measurements and reduced computation strategies such as decoupled methods. None of these are applicable to DSSE. The development of the DSSE tool becomes challenging since the distribution estimator is vulnerable to ill-conditioning problems caused by distribution network characteristics and poor estimation due to lack of sufficiently accurate measurements data. Present power distribution systems do not usually provide adequate real time measurement data, due to their limited instrumentation and communication infrastructure.

Furthermore, the probability of MV-LV distribution networks being substantial in size, may result in the execution of the DSSE becoming computationally costly. A parallel approach allows the estimation for each zone to be performed simultaneously, and a substantial amount of the data computation burden can be distributed over the local estimators. The existing parallel computation based SE algorithms are mostly deemed suitable for transmission systems, many of which take advantages of inter-connected power systems defining each tie-line connected area as a zone. Almost all the existing distributed SE tools achieve the minimum global observability of the network by the real time measurement data. Since the distribution system measurement configuration is expected to consist of limited real measurement data due to economical constraints, the

distributed algorithms proposed in the literature, are not directly applicable to distribution systems. A solution for DSSE to be computed in a restricted time window could be obtained by introducing 'parallel but not completely independent' estimation; which can also enable the DSSE to achieve some degree of scalability.

Aforementioned, the transition from largely passive to more active distribution systems will require distribution system real time monitoring at a much larger scale. To enable future distribution system state estimation, a significant phase of new sensor deployment is an essential prerequisite. The distribution system observability can be achieved by limited real time measurement, some virtual and many pseudo measurement data. There remains a fundamental risk of low quality estimation due to gross measurement errors, particularly when the real measurement data assigned with greater weight is limited. The meter placement procedure for the distribution system should be able to achieve the maximum improvement in estimation quality taking into consideration such measurement configurations and risk factors.

In these contexts, this research aimed to propose solutions for MV network DSSE related issues in three aspects.

- Development of a DSSE method
- Development of a scalable DSSE solution
- Development of a meter placement algorithm

Five DSSE optimization solutions are case studied extensively as candidate DSSE tools. The candidate tools include the normal equation based classical WLS method as well as alternative WLS solution processes such as CWLS, Orthogonal Decomposition and Hachtel's Augmented Matrix methods. A novel algorithm known as WEM is also introduced and assessed as a potential candidate tool. The candidate solution processes are preliminarily selected based on their potential aspects in theory to be applied as DSSE solutions. Investigation is performed under various scenarios representing typical distribution network characteristics to assess their practical viability. Five scenarios are studied on a 77 and a 356 node network, applying the candidate solutions to appraise the effect of very short lines, erroneous parameter values and error measurements on the candidate solutions. Extensive Monte Carlo SE studies that generate a variety of measurement data are performed for this assessment. The quality of estimates is quantified by comparing true values and the estimated values of primary states. Taking into

consideration the computation time, convergence property, consistency of overall performance and the accuracy of estimated data in various scenarios, the Hachtel's Augmented Matrix method is proposed as the most useful DSSE solver.

Further studies are consequently performed with real sensor data and network parameters applying Hachtel's Augmented Matrix method on two feeders of Slovenian distribution networks operated by EG. The estimated data produced by the DSSE tool matches the measurement data in most cases with a few exceptions, which indicates the possibility of the presence of bad data and inaccurate parameter information. As true values of primary states are unknown for the EG feeders, the quality of estimation is determined by how good the estimated value matches the respective the measured data. In addition, the use of real measurements as equality constraints is introduced as a novel concept to be investigated in this research. The quality of outcomes with regard to voltage and power estimation have been evaluated by comparing the accuracy of estimates applying with and without constrained real measurements. The investigation leads to the conclusion that the constrained real measurements can be beneficial when they are limited and very accurate compare to other available measurements. Otherwise, there are potential risks of catastrophic degradation in estimation quality in cases where the constrained real measurement contains gross errors.

The DSSE tool should be operable within a very short time window. However due to the quadratic relationship between size of the DSSE problem and the computation time, the execution time could be a major concern in the practical application for large networks. Therefore, parallelization of DSSE tools is suggested as one of the promising methods to rectify this issue. Two methods are proposed for initial investigation: DEA and OZA. The proposed methods are applicable to any large network having limited real measurements, and the network subdivision can be taken advantage of but is not completely dependent on the MV feeder configurations or the location of tie-lines. Since the key priority of the distribution network operator will be observing the voltage profiles at MV levels, the objectives of the proposed algorithms are to achieve satisfactory quality of voltage estimations, reduced overall computation time and improved scalability of the DSSE. After comparison of computation time and estimation quality, the OZA using Hachtel's method has outperformed DEA in all respects. Hence, the OZA is proposed as a scalable as well as

a distributed DSSE solution. The OZA is extensively applied on simulated larger networks consisting of 711 nodes and 356 nodes.

As a local estimation can preserve co-ordination and coherency with other zones by matching overlap node estimation, only measurement at GSP can lead to optimum solutions and the resulted zonal voltage and phase angle estimates are almost exact representation of the centralized estimates. However this may take a longer time and have very slow convergence. It is also recommended to deploy additional sensors at least on halo nodes, overlapping nodes and zone start nodes to achieve good quality voltage estimation in a faster way. The studies infer that the overlapping zone approach has great potentiality to preserve the property of scalability while increased network splitting and additional sensor deployment are provided. The reduced zonal interactions with increased instrumentation can provide feasible solution of voltage estimates. In addition, a special case study is developed in which the DEA is integrated in OZA as the local estimator. The combined approach results in better quality estimated data compared to that of only DEA based DSSE method. The network division in OZA is performed manually and mostly based on some general rules in this thesis.

The OZA method shows evidence of improved scalability, computation cost and its capability of estimating voltage magnitude requiring only a few real measurements, encouraging its consideration for realistic application at distribution systems in the present and near-future scenarios. Nevertheless, there is scope for further improvement of the novel algorithms introduced as the scalable DSSE in terms of sensor placements, accuracy and methods of network divisions.

At present, power distribution networks usually do not provide adequate real time measurement data, due to their limited instrumentation and communication infrastructure. However, fast growing DG and responsive loads are transforming distribution networks into active systems, necessitating enhanced network monitoring and automated voltage control operation. Present distribution networks will therefore need large scale meter deployment at MV to LV substation levels. As the pre-requisite of DSSE, a meter placement algorithm has been proposed in this work. The semi-stochastic meter placement algorithm is focused on reduction of voltage estimation error, which is very important for

reliable operation of active distribution networks. The proposed algorithm requires two major steps:

- By restricting measurement residual exploiting property of singular value of a matrix
- By reducing true residuals applying rigorous Monte Carlo based studies

In each placement, a voltage magnitude and a pair of real-reactive power flow sensors are deployed. It is suggested to accompany a voltage sensor along with a pair of power/current sensors to obtain greater improvement in estimation quality. The nodes having lower sensitivity are preferred to add new meters, since that makes the estimated data to be less affected by erroneous sensor data. The principle of adding sensors to the less sensitive nodes is particularly feasible for the distribution system, where real measurements will be limited compared to the availability of much larger numbers of pseudo-measurement data. The algorithm is applied on the 77 node network in various case studies. It is observed that the effect of error measurements from power injection meters on the voltage estimation quality is negligible in most cases, which again justifies the principle of adding sensors to less sensitive nodes. One important test shows that the quality of voltage estimates is similar, when meters are placed as obtained from the algorithm, applying different DSSE solution processes which are WLS, Orthogonal Decomposition (QR) and Hachtel's Augmented Matrix methods. Studies are performed for high number of Monte Carlo studies ranging from 100 to 7500, which gives added confidence to conclude the investigation outcomes. Meters are placed sequentially and the proposed meter position by the algorithm is referred to as a feasible solution.

The last part of this thesis includes the practical application of DSSE tool and the meter placement algorithm on parallel platforms. For DSSE, the cluster computer system provided by Oxford University was used. Part of the work is contributed by the Oxford research group. The computation time and convergence with zone split are studied and reported in this research, also incorporating real network and real data sets. It is suggested that, for good quality voltage estimation in a reduced time frame, additional measurements should be added at halo, overlapping and zone start nodes and the zonal interaction is required to be restricted. At this stage, each processor can handle one zone only. The zones, as well as the overlapping and halo data, are required to be tagged with individual numbers, therefore the transferable data arrays can identify their destinations based on the tagged numbers. Networks are split into two, three and four zones. Overall, some



improvement in scalability is observed. The expected reduction in computation time with respect to the time required for the centralized DSSE is obtained when the network is split into four zones.

The most computationally expensive part of the meter placement algorithm is applied on a four core machine applying MATLAB parallel operators: PARFOR and SPMD. The performance of both functions with regard to computation time and CPU usage is very similar, however SPMD practises more control over the algorithm than PERFOR. The outcomes show that involving more workers/'labs' does not always bring reduction of computation time due to increasing communication overhead.

In addition to introduction of several novel approaches, an important contribution of the research is to use the real data set and real networks in parallel to simulated networks. The DSSE tool applying Hachtel's method is applied to two MV feeders of Slovenia distribution network using real sensor data. The line parameters and sensor data are provided by Slovenian DNO, EG. The distributed scalable DSSE is applied on a larger EG network consisting of 411 nodes. In most cases the real sensor data are used, however for parallel DSSE application, simulated measurements are used in absence of real measurements. Application of model networks also carries significant value for assessment purposes. The major advantage of using a model network is that, the true values of system states can be calculated. Once true values are known, the quality of estimated data can be assessed ideally and confidently with respect to the true values. The model networks (UKGDS networks in this research) can provide an excellent platform to evaluate the performance of novel algorithms and comparison of various algorithms.

## **8.2 Future Work**

In this research several novel algorithms are introduced which have potential scope for further research. The DEA is applied as distribution system DSSE, which has not performed as expected unfortunately. However, there are several advanced and improved types of DEA e.g. Hybrid Differential Evolution (HDE) [83], Self Adaptive Differential Evolution (SADE) [69] methods which have proven to be more efficient and speedier than classical DEA. Their application on DSSE is recommended as a future research direction.

A successful application of enhanced DEA based DSSE should be followed by parallel application of DEA [64] [65] [66] [67] [68].

The novel OZA approach shows promising outcomes in voltage estimation applying Hachtel's Augmented Matrix Method. As the zone splitting is a manual process in this work, development of a zone splitting tool will be beneficial in application of the OZA.

The Meter placement algorithm is introduced and applied for the improvement of voltage estimation. Application of the algorithm for the perfection of other states has further scope for research. Presented parallel processing application on MTALAB can also be applied to a computer cluster or a grid computing platform, which is the next step in application research. The proposed algorithm is suitable to utilize for the improvement of other state estimates, e.g. phase angle, power etc. and for radial or non-radial networks in similar fashion as applied for voltage estimation.

As mentioned earlier, this research is strongly focused on the development of MV network scalable state estimation and meter placement algorithms, the LV network SE is considered an extension of this research for the future. It will be useful to investigate the outcomes of proposed DSSE and meter placement algorithms, applying to LV networks. The LV SE can utilize the smart meter data as pseudo measurements, however the major work for LV SE would be transforming the algorithm to an unbalanced three phase SE formulation.

It is also proposed to perform DSSE for MV radial distribution network taking the phase angle difference as the primary state. Conventionally, the phase angle itself is considered as primary states, therefore the number of state variables remains same as the number of nodes of the systems. The radial distribution networks benefit from the advantage of number of branches being one less than the number of nodes in general. Furthermore, the phase angle differences between nodes are closer to the initial guess 'zero' values which can give the estimator a better start. This can assist the DEA estimator to overcome the issue of poor phase angle estimation. The author therefore believes that taking phase differences between nodes can help the estimator to produce further improved outcomes.

## REFERENCE

- [1] A. Abur and A. Exposito, *Power System State Estimation Theory and Implementation*, New York: Marcel Dekker Inc, 2004.
- [2] A. Monticelli, "Electric Power System State Estimation," *Proceedings of the IEEE*, vol. 88, no. 2, pp. 262 - 282, Feb 2000.
- [3] M. Shahidehpour and Y. Wang, "Parallel and Distributed State Estimation," in *Communication and Control in Electric Power Systems: Applications of Parallel and Distributed Processing*, pp. 135-164, Wiley-IEEE Press, 2003.
- [4] G. Strbac, C. Ramsay and D. Pudjianto, "Integration of Distributed Generation into the UK Power System," Mar 2007. [Online]. Available: <http://www.sedg.ac.uk/>. [Accessed Dec 2012].
- [5] E. Manitsas, R. Singh, B. Pal and G. Strbac, "Distribution System State Estimation Using an Artificial Neural Network Approach for Pseudo Measurement Modeling," *IEEE Transactions on Power Systems*, vol. 27, no. 4, pp. 1888-1896, Nov 2012.
- [6] F. Schweppe and J. Wildes, "Power System Static-State Estimation, Part I: Exact Model," *IEEE Transactions on Power Apparatus and Systems*, Vols. PAS-89, no. 1, pp. 120-125, Jan 1970.
- [7] F. Schweppe and D. Rom, "Power System Static-State Estimation, Part II: Approximate Model," *IEEE Transactions on Power Apparatus and Systems*, Vols. PAS-89, no. 1, pp. 125-130, Jan 1970.
- [8] F. Schweppe, "Power System Static-State Estimation, Part III: Implementation," *IEEE Transactions on Power Apparatus and Systems*, Vols. PAS-89, no. 1, pp. 130-135, Jan 1970.
- [9] A. Ghosh, D. Lubkeman, M. Downey and R. Jones, "Distribution Circuit State Estimation Using A Probabilistic Approach," *IEEE Transactions on Power Systems*, vol. 12, no. 1, pp. 45-51, Feb 1997.
- [10] R. Hoffman, "Practical State Estimation for Electric Distribution Networks," in *IEEE PES Power Systems Conference and Exposition (PSCE '06)*, Oct-Nov 2006.
- [11] M. Baran and A. Kelley, "State Estimation for Real-Time Monitoring of Distribution Systems," *IEEE Transactions on Power Systems*, vol. 3, no. 9, pp. 1601-1609, Aug

1994.

- [12] M. Baran and A. Kelley, "A Branch-Current-Based State Estimation Method for Distribution Systems," *IEEE Transactions on Power Systems*, vol. 10, no. 1, pp. 483-491, Feb 1995.
- [13] C. Lu, J. Teng and W. Liu, "Distribution System State Estimation," *IEEE Transactions on Power Systems*, vol. 10, no. 1, pp. 229-240, Feb 1995.
- [14] H. Wang and N. Schulz, "A Revised Branch Current-Based Distribution System State Estimation Algorithm and Meter Placement Impact," *IEEE Transactions on Power Systems*, vol. 19, no. 1, pp. 207 - 213, Feb 2004.
- [15] V. Thornley, N. Jenkins and S. White, "State Estimation Applied to Active Distribution Networks with Minimal Measurement," in *15th Power System Computation Conference*, 2007.
- [16] O. Chillard, S. Grenard, O. Devaux and L. Garcia, "Distribution State Estimation Based on Voltage State Variables: Assessment of Results and Limitations," in *CIREN 20th International Conference on Electricity Distribution*, Jun 2009.
- [17] R. Sing, B. Pal and R. Jabr, "Choice of Estimator for Distribution System State Estimation," *IET Generation, Transmission and Distribution*, vol. 7, no. 3, pp. 666-678, Jul 2009.
- [18] F. Bignucolo, R. Caldon and M. Valente, "Probabilistic Voltage Estimation for Active Control of Distribution Networks," in *CIREN 19th International Conference on Electricity Distribution*, May 2007.
- [19] M. Biserica, Y. Besanger, R. Caire, O. Chillard and P. Deschamps, "Neural Networks to Improve Distribution State Estimation—Volt Var Control Performances," *IEEE Transactions on Smart Grid*, vol. 3, no. 3, pp. 1137-1144, Sep 2012.
- [20] W. Xu, D. Liu, R. Lu, J. Qiu, Pan, H. and Z. Wei, "A Novel State Estimation Method Based on Quality Tag for Distribution Networks," in *CIREN 20th International Conference and Exhibition on Electricity Distribution - Part 1*, Jun 2009.
- [21] N. Woolley and J. Milanovic, "Statistical Estimation of the Source and Level of Voltage Unbalance in Distribution Networks," *IEEE Transactions on Power Delivery*, vol. 27, no. 3, pp. 1450-1460, Jul 2012.
- [22] C. Gomez-Quiles, A. Gomez-Exposito and A. de la Villa Jaen, "State Estimation for Smart Distribution Substations," *IEEE Transactions on Smart Grid*, vol. 3, no. 2, pp. 986-995, Jun 2012.

- [23] S. Naka, T. Genji, T. Yura, Y. Fukuyama and N. Hayashi, "Distribution State Estimation Considering Nonlinear Characteristics of Practical Equipment Using Hybrid Particle Swarm Optimization," in *International Conference on Power System Technology 2000*, 2000.
- [24] S. Naka, T. Genji, T. Yura and Y. Fukuyama, "Practical Distribution State Estimation Using Hybrid Particle Swarm Optimization," in *IEEE Power Engineering Society Winter Meeting 2001*, 2001.
- [25] S. Naka, T. Genji, T. Yura and Y. Fukuyama, "A Hybrid Particle Swarm Optimization for Distribution State Estimation," *IEEE Transactions on Power Systems*, vol. 118, no. 1, pp. 60- 68, Feb 2003.
- [26] O. Chilard, S. Grenard and O. Devaux, "Comparison of the Performances of Distribution State Estimation Algorithms: Classical Newton Approach and PSO Approach," in *IEEE Power and Energy Society General Meeting 2012*, Jul 2012.
- [27] K. Hashimoto, N. Kagan and S. Santoso, "A Metaheuristic Evolutionary Method for Electrical Performance Estimation of Distribution Networks," in *IEEE Power Engineering Society General Meeting 2005*, Jun 2005.
- [28] T. Van Cutsem, J. Horward and M. Ribbens-Pavella, "Hierarchical State Estimator: A New Way for Large Power Power Systems Estimation," in *Mathematics and Computers in Simulation XXII*, 1980.
- [29] T. Van Cutsem, J. Horward and M. Ribbens-Pavella, "Hierarchical State Estimation," *International Journal of Electrical Power & Energy Systems 2*, vol. 2, no. 2, pp. 70-80, Apr 1980.
- [30] T. Van Cutsem, J. Horward and M. Ribbens-Pavella, "A Two-Level Static State Estimator for Electric Power Systems," *IEEE Transactions on Power Apparatus and Systems*, Vols. PAS-100, no. 8, pp. 3722-3732, Aug 1981.
- [31] M. Zhao and A. Abur, "Multi Area State Estimation Using Synchronized Phasor Measurements," *IEEE Transaction on Power Systems*, vol. 20, no. 2, pp. 611-617, May 2005.
- [32] S. Iwamoto, M. Kusano and V. Quintana, "Hierarchical State Estimation Using a Fast Rectangular-Coordinate Method [Power System Analysis Computing]," *IEEE Transactions on Power Systems*, vol. 4, no. 3, pp. 870-880, Aug 1989.
- [33] A. Gomez-Exposito and A. de la Villa Jaen, "Two-Level State Estimation with Local Measurement Pre-Processing," *IEEE Transactions on Power Systems*, vol. 24, no. 2,

pp. 676-684, May 2009.

- [34] S. Mohagheghi, R. Alaileh, G. Cokkinides and A. Meliopoulos, "Distributed State Estimation Based on The Supercalibrator Concept - Laboratory Implementation," in *2007 iREP Symposium- Bulk Power System Dynamics and Control- VII, Revitalizing Operational Reliability*, Aug 2007.
- [35] D. Falcao, F. Wu and L. Murphy, "Parallel and Distributed State Estimation," *IEEE Transactions on Power Systems*, vol. 10, no. 2, pp. 724-730, May 1995.
- [36] J. Carvalho and F. Barbosa, "Parallel and Distributed Processing in State Estimation of Power System Energy," *9th Mediterranean Electrotechnical Conference*, vol. 2, no. 18-20, pp. 969-973, May 1998.
- [37] F. Wu and A. Neyer, "Asynchronous Distributed State Estimation for Power Distribution System," in *Proc. of The 10th Power Systems Computation Conference*, Aug 1990.
- [38] A. Abur and P. Tapadiy, "Parallel State Estimation Using Multiprocessors," *Electric Power Systems Research*, vol. 18, no. 1, pp. 67-73, Jan 1990.
- [39] P. Nguyen, W. Kling and J. Myrzik, "Completely Decentralized State Estimation for Active Distribution Network," in *Proceedings of the EuroPES conference*, Sep 2009.
- [40] M. Nordman and M. Lehtonen, "Distributed Agent-Based State Estimation for Electrical Distribution Networks," *IEEE Transactions on Power Systems*, vol. 20, no. 2, pp. 652- 658, May 2005.
- [41] P. Nguyen and W. Kling, "Distributed State Estimation for Multi-Agent based Active Distribution Networks," in *IEEE Power and Energy Society General Meeting 2010*, Jul 2010.
- [42] A. Shafiu, N. Jenkins and G. Strbac, "Measurement Location for State Estimation of Distribution Networks with Generation," *IEE Proceedings Generation, Transmission and Distribution*, vol. 152, no. 2, pp. 240 - 246, Mar 2005.
- [43] R. Singh, B. Pal, R. Jabr and R. Vinter, "Meter Placement for Distribution System State Estimation: An Ordinal Optimization Approach," *IEEE Transactions on Power Systems*, vol. 26, no. 4, pp. 2328- 2335 , Nov 2011.
- [44] C. Muscas, F. Pilo, G. Pisanor and S. Sulis, "Optimal Placement of Measurement Devices in Electric Distribution Systems," in *Proceedings of the IEEE Instrumentation and Measurement Technology Conference 2006*, Apr 2006.

- [45] C. Muscas, F. Pilo, G. Pisanor and S. Sulis, "Optimal Allocation of Multichannel Measurement Devices for Distribution State Estimation," *IEEE Transactions on Instrumentation and Measurement*, vol. 58, no. 6, pp. 1929- 1937, Jun 2009.
- [46] R. Singh, B. Pal and R. Vinter, "Measurement Placement in Distribution System State Estimation," *IEEE Transactions on Power Systems*, vol. 24, no. 2, pp. 668- 675, May 2004.
- [47] P. Rousseeuw and A. Leroy, *Robust Regression and Outlier Detection*, John Willey and Sons, inc., 2005.
- [48] T. Ryan, *Modern Regression Methods*, John Willey and Sons, 2008.
- [49] A. Monticelli, *State Estimation in Electric Power Systems A Generalized Approach*, USA: Kluwer Academic Publisher, 1999.
- [50] L. Holten, A. Gjelsvik, S. Aam, F. Wu and W. Liu, "Comparison of Different Methods for State Estimation," *IEEE Transactions on Power Systems*, vol. 3, no. 4, pp. 1798-1804, Nov 1988.
- [51] A. Chatterjee, "An Introduction to the Proper Orthogonal Decomposition," *Current Science*, vol. 78, no. 7, p. 808–817, Apr 2000.
- [52] M. Gilles and R. Nucera, "A Blocked Sparse Matrix Formulation for the Solution of Equality-Constrained State Estimation," *IEEE Transaction on Power Systems*, vol. 6, no. 1, pp. 214 - 221, Feb 1991.
- [53] G. Taylor, M. Irving, N. Nusrat, R. Liao and S. Panchandaram, "Smart Distribution Network Operation: Emerging Techniques and Standards," in *IEEE Power and Energy Society General Meeting 2011*, Jul 2011.
- [54] "United Kingdom Generic Distribution System," [Online]. Available: <http://monaco.eee.strath.ac.uk/ukgds/>. [Accessed Sept. 2011].
- [55] N. Nusrat, M. Irving and G. Taylor, "Development of Novel State Estimation Algorithms for Active Distribution Networks," in *Proceedings of 46th International Universities' Power Engineering Conference 2011*, Sep 2011.
- [56] KORONA, EG, "HPC Platform Testing Specification," HiPerDNO Project, 2012.
- [57] R. Storn and K. Price, "Differential Evolution—A Simple and Efficient Heuristic for Global Optimization over Continuous Spaces," *Journal of Global Optimization* 11, vol. 11, no. 4, p. 341–359, Dec 1997.
- [58] R. Storn and K. Price, "Minimizing the Real Functions of the ICEC'96 Contest by

- Differential Evolution,” in *Proceedings of IEEE International Conference on Evolutionary Computation 1996*, May 1996.
- [59] R. Storn, “On the Usage of Differential Evolution for Function Optimization,” in *Biennial Conference of the North American Fuzzy Information Processing Society 1996*, Jun 1996.
- [60] S. Das, A. Abraham and A. Konar, “Particle Swarm Optimization and Differential Evolution Algorithms: Technical Analysis, Applications and Hybridization Perspectives,” in *Advances of Computational Intelligence in Industrial System*, 2008.
- [61] J. Vesterstrom and R. Thomsen, “A Comparative Study of Differential Evolution, Particle Swarm Optimization, and Evolutionary Algorithms on Numerical Benchmark Problems,” in *Congress on Evolutionary Computation 2004* , Jun 2004.
- [62] J. Brest, S. Greiner, B. Boskovic, M. Mernik and V. Zumer, “Self-Adapting Control Parameters in Differential Evolution: A Comparative Study on Numerical Benchmark Problems,” *IEEE Transactions on Evolutionary Computation*, vol. 10, no. 6, pp. 646 - 657, 2006.
- [63] T. Robič and B. Filipič, “DEMO: Differential Evolution for Multiobjective Optimization,” in *Proc. 3rd Int. Conf. Evol. Multi-Criterion Optimization, LNCS 3410*, 2005.
- [64] D. Tasoulis, N. Pavlidis, V. Plagianakos and M. Vrahatis, “Parallel Differential Evolution,” in *Congress on Evolutionary Computation 2004*, 2004.
- [65] M. Weber, F. Neri and V. Tirronen, “Parallel Random Injection Differential Evolution,” in *Applications of Evolutionary Computation*, 2010.
- [66] W. Kwedlo and K. Bandurski, “A Parallel Differential Evolution Algorithm for Neural Network Training,” in *Proceedings of the International Symposium on Parallel Computing in Electrical Engineering 2006*, Sep 2006.
- [67] P. Bujok and J. Tvrdik, “Parallel Migration Models Applied to Competitive Differential Evolution,” in *13th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing*, Sep 2011.
- [68] M. Weber, F. Neri and V. Tirronen, “Distributed Differential Evolution with Explorative–Exploitative Population Families,” *Genetic Programming and Evolvable Machines*, vol. 10, no. 4, pp. 343-371, 2009.
- [69] V. Huang, A. Qin and P. Suganthan, “Self-adaptive Differential Evolution Algorithm for Constrained Real-Parameter Optimization,” in *IEEE Congress on Evolutionary*



*Computation 2006*, 2006.

- [70] R. Gämperle, S. D. Müller and P. Koumoutsakos, "Parameter Study for Differential Evolution," *Advances in Intelligent Systems, Fuzzy Systems, Evolutionary Computation*, no. 10, pp. 293-298, 2002.
- [71] A. Abou El Ela, M. Abido and S. Spea, "Optimal Power Flow Using Differential Evolution Algorithm," *Electric Power Systems Research*, vol. 80, no. 7, pp. 878-885, Jul 2010.
- [72] S. Das and P. Suganthan, "Differential Evolution: A Survey of the State-of-the-Art," *IEEE Transactions on Evolutionary Computation*, vol. 15, no. 1, pp. 4-31, Feb 2011.
- [73] D. Zaharie, "Critical Values for the Control Parameters of Differential Evolution Algorithms," in *Proceedings of MENDEL 2002, 8th International Conference on Soft Computing*, Jun 2002.
- [74] D. Zaharie, "Influence of Crossover on the Behavior of Differential Evolution Algorithms," *Applied Soft Computing*, vol. 9, no. 3, pp. 1126-1138, 2009.
- [75] D. Corne, M. Dorigo and F. Glover, "An Introduction to Differential Evolution," in *New Idea in Optimization*, pp. 79-160, McGraw-Hill, 1999.
- [76] E. Mezura-Montes and C. Monterrosa-López, "Global and Local Selection in Differential Evolution for Constrained Numerical Optimization," *Journal of Computer Science & Technology*, vol. 9, no. 2, pp. 43-52, Oct 2009.
- [77] A. Cuello-Reyna and J. Cedeno-Maldonado, "Differential Evolution-Based Weighted Least Squares State Estimation with Phasor Measurement Units," in *49th IEEE International Midwest Symposium on Circuits and Systems 2006*, Aug 2006.
- [78] N. Figueroa and J. Cedeño, "A Differential Evolution Solution Approach for Power System State Estimation," in *Proceedings of Seventh IASTED International Conference, Power and Energy Systems*, 2004.
- [79] K. Price, R. Storn and J. Lampinen., *Differential Evolution: A Practical Approach to Global Optimization(Ch.1-5)*, Springer-Verlag, 2005.
- [80] "Differential Evolution (DE) for Continuous Function Optimization (an algorithm by Kenneth Price and Rainer Storn)," International Computer Science Institute (ICSI), [Online]. Available: <http://www1.icsi.berkeley.edu/~storn/code.html>.. [Accessed Oct 2010].
- [81] N. Nusrat, M. Irving and G. Taylor, "Development of Distributed State Estimation

- Methods to Enable Smart Distribution Management Systems,” in *IEEE International Symposium on Industrial Electronics 2011*, Jun 2011.
- [82] P. Lopatka, S. Salvini, N. Nusrat, L. De-Alvaro and D. Wallom, “Performance Study of Distributed State Estimation Algorithms on the HiPerDNO HPC Platform,” in *47th International Universities Power Engineering Conference 2012*, Sep 2012.
- [83] J. Chiou, “A Variable Scaling Hybrid Differential Evolution for Solving Large-Scale Power Dispatch Problems,” *IET Generation, Transmission & Distribution*, vol. 3, no. 2, pp. 154-163, Feb 2009.
- [84] S. Trek, “Margin of Error,” [Online]. Available: <http://stattrek.com/estimation/margin-of-error.aspx>. [Accessed Aug 2014].
- [85] Pegoraro, P.A.; Sulis, S., "Robustness-Oriented Meter Placement for Distribution System State Estimation in Presence of Network Parameter Uncertainty," *IEEE Transactions on Instrumentation and Measurement*, vol.62, no.5, pp.954-962, May 2013
- [86] Liu, J; Tang, J; Ponci, F.; Monti, A.; Muscas, C.; Pegoraro, P.A., "Trade-Offs in PMU Deployment for State Estimation in Active Distribution Grids," *IEEE Transactions on Smart Grid*, vol.3, no.2, pp.915-924, Jun. 2012
- [87] Liu, J; Ponci, F.; Monti, A.; Muscas, C.; Pegoraro, P.A.; Sulis, S., "Optimal Meter Placement for Robust Measurement Systems in Active Distribution Grids," *IEEE Transactions on Instrumentation and Measurement*, vol.63, no.5, pp.1096,1105, May 2014
- [88] D. Echternacht, C. Linnemann and A. Moser, "Optimized Positioning of Measurements in Distribution Grids," in *3rd IEEE PES International Conference and Exhibition on Innovative Smart Grid Technologies 2012*, pp. 1-7, Oct. 2012.
- [89] D. Echternacht and A. Moser, "Cost Optimal Meter Placement in Low and Medium Voltage Grids Considering Stochastic Dependencies," in *4th IEEE PES Innovative Smart Grid Technologies Europe 2013*, pp.1-5, Oct. 2013.

# APPENDIX

## 1. Calculation of Measurement Standard Deviation, $\sigma$

The relation between margin of error and standard deviation,  $\sigma$  for a normally distributed function is,

$$\text{margin of error} = z_{\alpha/2}^* \times \sigma$$

Here  $z_{\alpha/2}^*$  is the critical value. Here, the measurement standard deviation,  $\sigma$  is calculated for  $3\sigma$  distribution i.e. 99.73% confidence interval. In case of  $3\sigma$  confidence interval, the z-score value would be such that the area between  $-z$  and  $+z$  is 0.9973. Therefore

$$\frac{\alpha}{2} = \frac{1 - 0.9973}{2} = 0.00135$$

The critical probability,

$$p^* = 1 - \frac{\alpha}{2} = 1 - 0.00135 = 0.99865$$

Looking at the standard normal distribution chart for the above critical probability equivalent z-score value, it is obtained that  $z_{\alpha/2}^* = 3.00$ . [84]

As the margin of error is defined by the maximum expected difference between the true value and observed value, for a sensor it is equivalent to the maximum expected error from the device. Therefore, for a true or mean value,  $\mu$

$$\text{margin of error} = \frac{\mu \times \%error}{100}$$

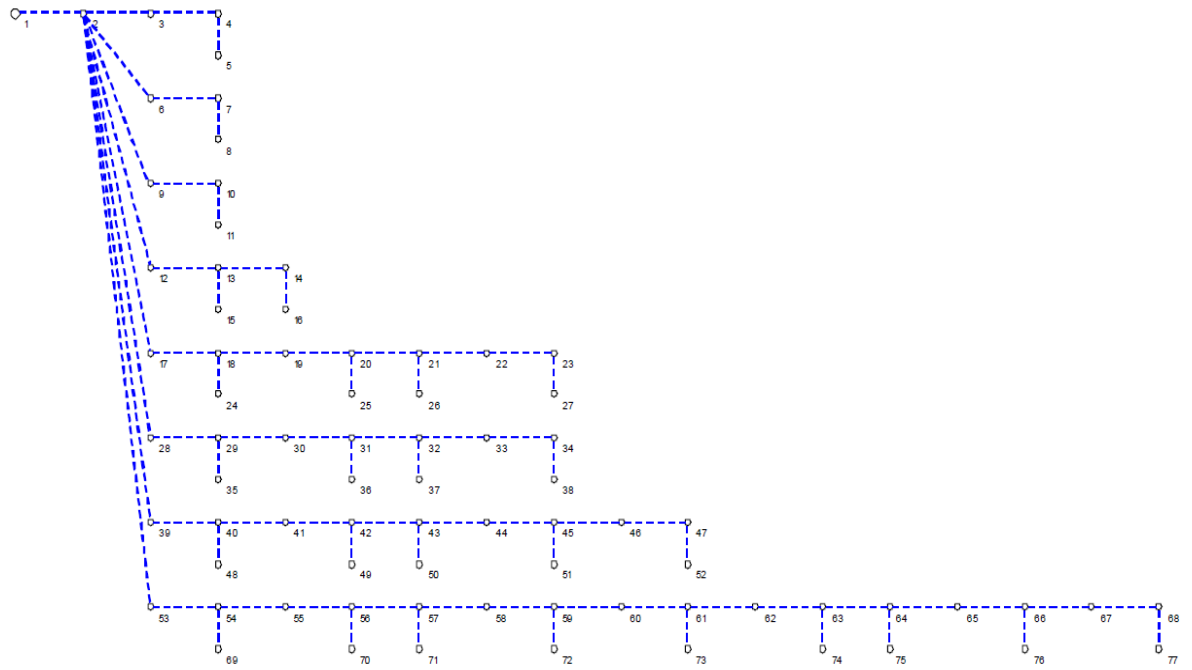
Placing the values of *margin of error* and  $z_{\alpha/2}^*$ , the equation looks like as below

$$\frac{\mu \times \%error}{100} = 3 \times \sigma$$

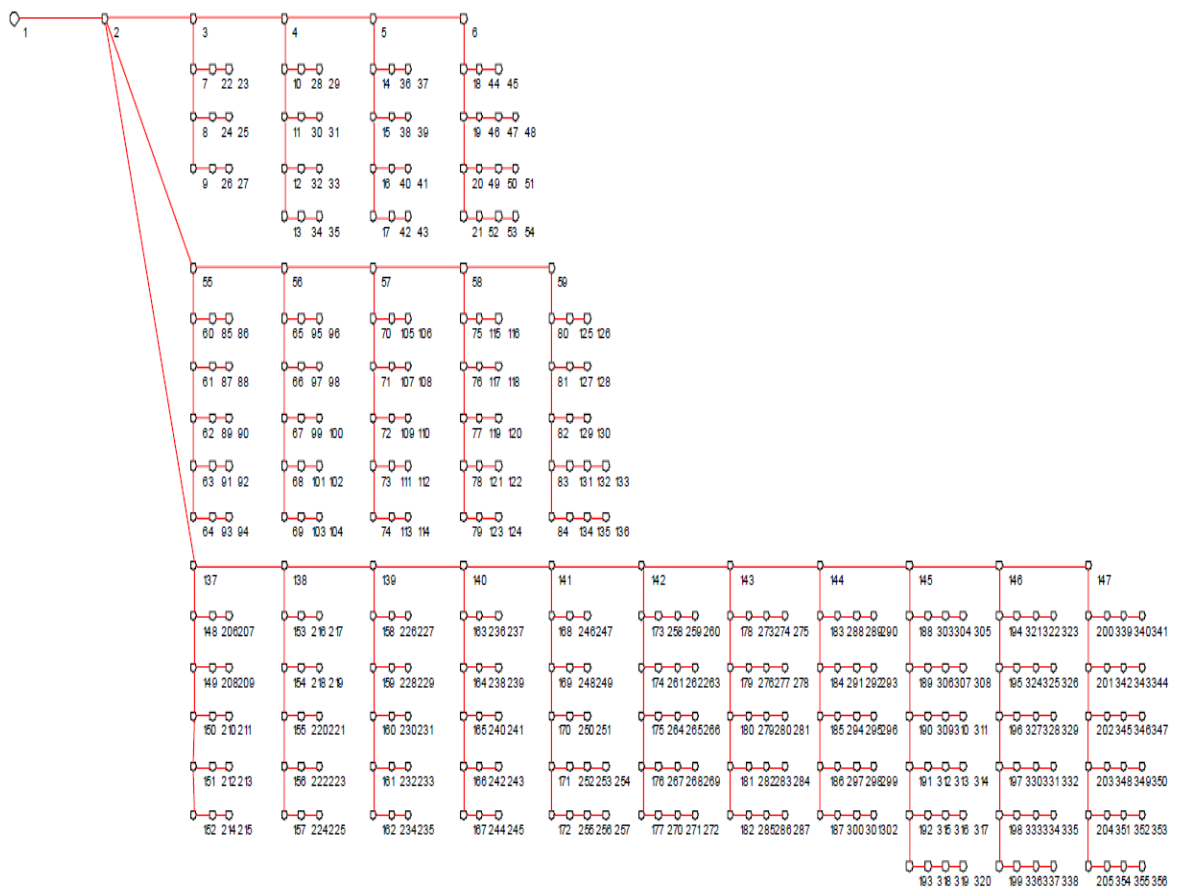
i.e.

$$\text{Standard Deviation, } \sigma = \frac{\mu \times \%error}{3 \times 100} = \frac{\mu \times \%error}{300}$$

## 2. UKGDS 77 Node Network

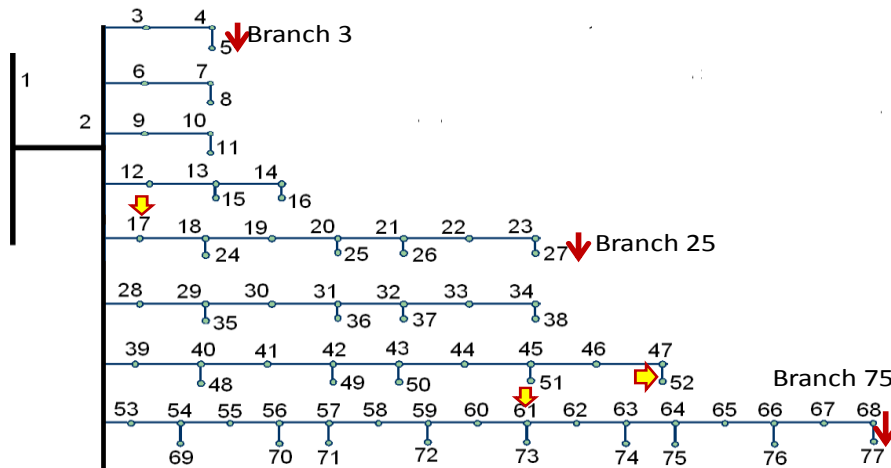


## 3. UKGDS 356 Node Network

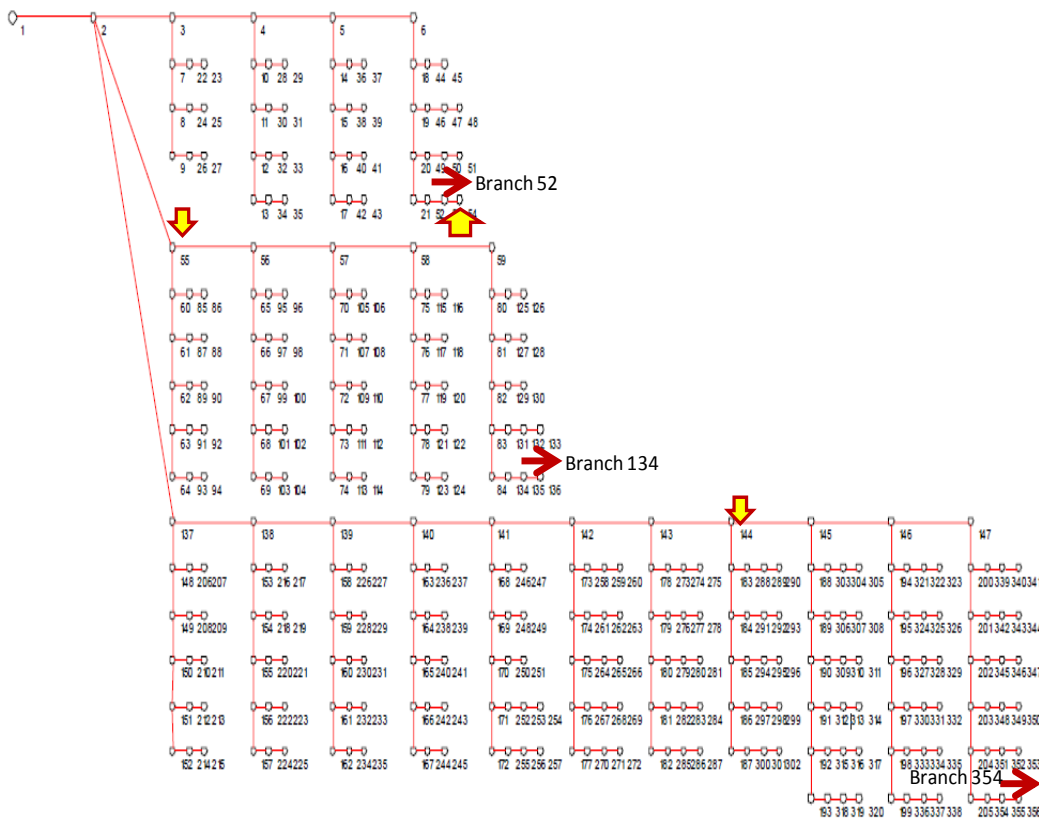


#### 4. Location of Various Connection Points Discussed for the Assessment of 77 and 356 Node Networks (Chapter 4)

Figures below show the location of nodes and branches as discussed in assessment studies in section 4.1. All node index are shown in the following figures. The location of the assumed error injection nodes are indicated by yellow arrows. The location of assumed very short branches are shown by red arrows along with the corresponding branch index information.



77 node networks: Location of error injection and short branches



356 node networks: Location of error injection and short branches

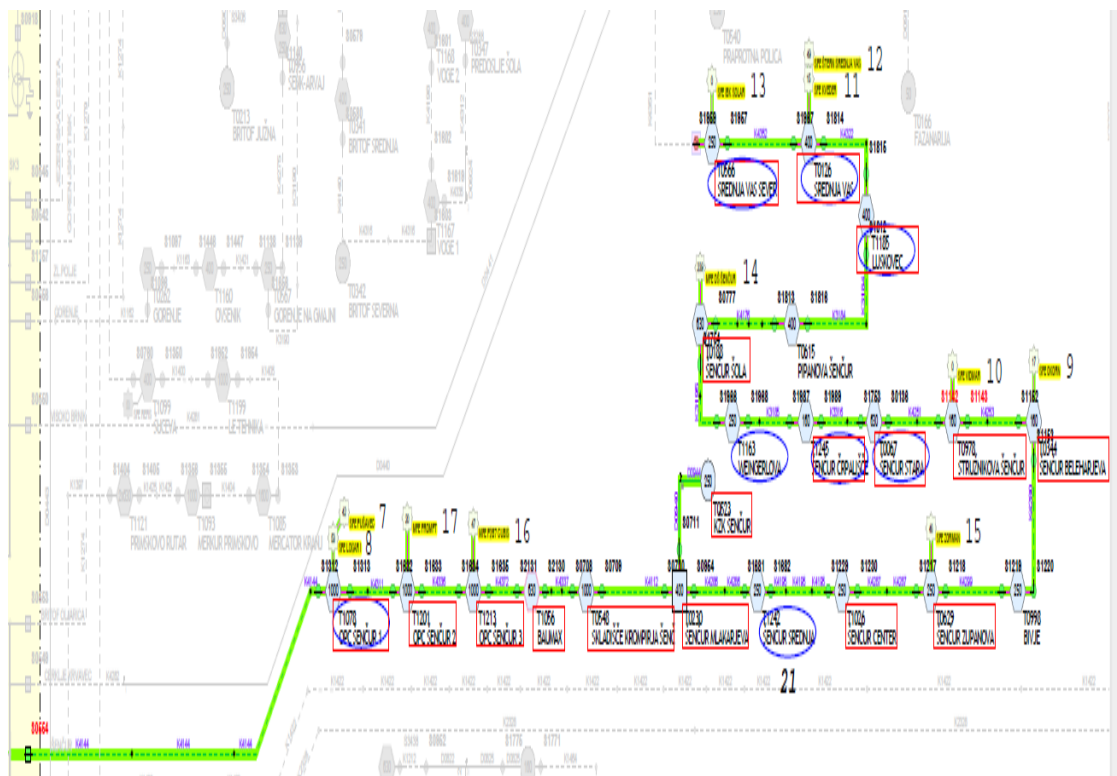
## 5. EG Network Feeders: Sencur, Cerklje

### Sencur Network Information

Node indexing and actual node codes:

Node ID	Node_description	Node index	Node ID	Node_description	Node index	Node ID	Node_description	Node index
6009303	J34 ŠENČUR	1	6000498	T0523 KŽK ŠENČUR	15	6072453	T1245 ŠENČUR ČRPALIŠČE	30
6068510	SPOJKA K4144 SP'C	2	6076434	SPOJKA K4286 SP'A"	16	6107544	SPOJKA K3185 SP'A"	31
6068508	SPOJKA K4144 SP'B"	3	6069796	T1242 ŠENČUR SREDNJA	17	6063250	T1163 WEINGERLOVA	32
6065748	SPOJKA K4144 SP'A"	4	6076992	SPOJKA K4195 SP'B"	18	6081704	SPOJKA K3195 SPOJKE 1	33
6044757	T1078 OPC ŠENČUR	5	6069768	SPOJKA K4195 SP'C"	19	6067503	T0188 ŠENČUR ŠOLA	34
6068701	SPOJKA K4311 SP'D"	6	6027682	T1026 ŠENČUR CENTER	20	6068026	SPOJKA K4176 SP'A"	35
6064411	T1201 OPC ŠENČUR 2	7	6069767	SPOJKA K4287 SP'B"	21	6067602	SPOJKA K4176 SP'B"	36
6067216	T1213 OPC ŠENČUR 3	8	6000591	T0629 ŠENČUR ZUPANOVA	22	6067601	SPOJKA K4176 SP'C"	37
6077343	T1056 BAUMAX	9	6010793	T0998 BIVJE	23	6000578	T0615 PIPANOVA ŠENČUR	38
6129673	SPOJKA K4337 "F" SKL - BAUMAX	10	6000337	T0344 ŠENČUR BELEHARJEVA	24	6084407	SPOJKA K3184 SP'A"	39
6068700	SPOJKA K4337 SP'E"	11	6051779	SPOJKA K4252-K4253	25	6064320	T1185 LUSKOVEC	40
6000518	T0548 SKLADIŠČE KROMPIRJA ŠENČ	12	6000796	T0978 STRUŽNIKOVA ŠENČUR	26	6084410	SPOJKE A K4322	41
6000225	T0230 ŠENČUR MLAKARJEVA	13	6051778	SPOJKA K4250-K4251	27	6064309	T0126 SREDNJA VAS	42
6030394	D0940-036	14	6064305	T0067 ŠENČUR STARA	28	6074384	T0566 SREDNJA VAS SEVER	43
			6107521	SPOJKA K3316 SP'A"	29			

The network diagram:

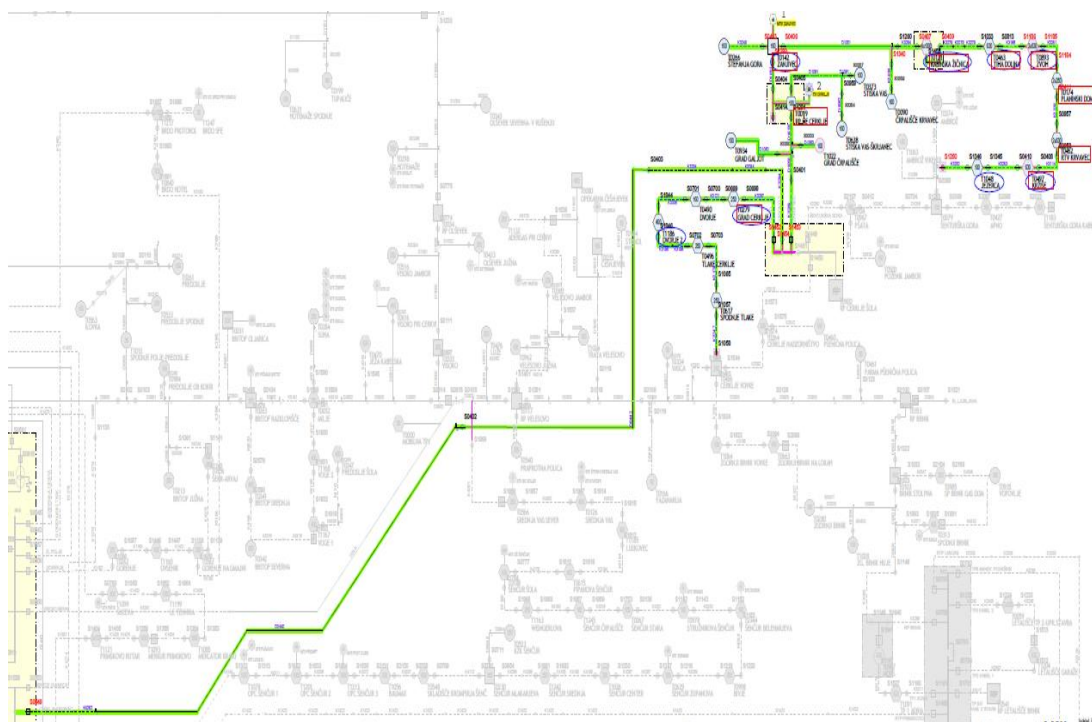


## Cerklje Network Information

Node indexing and actual node code:

Node ID	Node_description	Node index	Node ID	Node_description	Node index
6009302	J38 CERKLJE-KRVAVEC	1	6030686	D1061-013	21
6030152	D0440-001	2	6059507	S3357 T0373 STIŠKA VAS-OPOR	22
6030162	D0440-014	3	6000361	T0373 STIŠKA VAS	23
6041360	S0402 VISOKO-GRAD	4	6041909	S0959 ODCEP TP STIŠKA VAS ŠKRJANC	24
6041361	S0403 GRAD-POVEZ.DV BRNIK-KRVAVEC	5	6059504	S3354 T0638 STIŠKA VAS-ŠKRJANEC- OPO	25
6104124	SPOJKA K3294 SP"B"	6	6000599	T0638 STIŠKA VAS-ŠKRJANEC	26
6084762	SPOJKA K3294 SP"A"	7	6000144	T0142 ZANJIVEC	27
6063480	S1453-D RP CERKLJE	8	6000261	T0266 ŠTEFANJA GORA	28
6067402	T0279 GRAD CERKLJE	9	6030543	D1051-091	29
6000467	T0490 DVORJE	10	6000097	T0090 ČRPALIŠČE KRVAVEC	30
6064321	T1186 DVORJE 2	11	6042204	S1260 KRVAVEC-PRED TP KABIN.ŽIČNICA	31
6100091	SPOJKA K3196 SP"A"	12	6000125	T0121 KABINSKA ŽIČNICA	32
6000473	T0496 TLAKE CERKLJE	13	6082540	SPOJKA K3279 A	33
6000580	T0617 SPODNJE TLAKE	14	6082541	SPOJKA K3279 B	34
6041359	S0401 GRAD-ZA ODCEPOM ŠENT.GORA	15	6000441	T0463 TIHA DOLINA	35
6030473	D1050-022	16	6000718	T0893 ZVOH	36
6030492	D1050-041	17	6000174	T0174 PLANINSKI DOM	37
6000755	T0934 GRAD GALJOT	18	6000440	T0462 RTV KRVAVEC	38
6027654	T1022 GRAD ČRPALIŠČE	19	6000474	T0497 KRŽIŠE	39
6000035	T0019 RP HE CERKLJE	20	6027699	T1048 JEZERCA	40

The network diagram:



## 6. Distributed Generators Connected to EG Feeders

Name	Type	Power	Connected in MV/LVsubstation
SFE Pušavec	Solar	42 kVA	T1078 OPC ŠENČUR
SFE Logar 1	Solar	53 kVA	T1078 OPC ŠENČUR
SFE Okorn	Solar	17 kVA	T0344 ŠENČUR BELEHARJEVA
SFE Vidmar	Solar	20 kVA	T0978 ŠENČUR STRUŽNIKOVA
SFE Kveder	Solar	15 kVA	T0126 SREDNJA VAS
SFE Štern Sr. Vas	Solar	49 kVA	T0126 SREDNJA VAS
SFE IBK solar	Solar	34,6 kVA	T0566 SREDNJA VAS SEVER
SFE Šenčur Šola	Solar	239 kVA	T0188 ŠENČUR ŠOLA
SFE Zorman	Solar	46 kVA	T0639 SENCUR
SFE P2B7 Cubis	Solar	47 kVA	T1213 OPC SENCUR 3
SFE Prompt	Solar	20 kVA	T1201 OPC SENCUR 2

List of DGS connected to Sencur feeder

Name	Type	Power	Connected in MV/LVsubstation
MHE Zanjivec	Hydro	45 kVA	T0142 TP ZANJIVEC
HE Cerklje	Hydro	90 kVA	T0019 RP HE CERKLJE

List of DGS connected to Cerklje feeder



## 7. SCADA and QM Data Format Provided by EG

Description:	KN Kranj																
Location:	TP T0293 Elektro - TR 1																
Date	U1 [V]	U2 [V]	U3 [V]	THD U1 [%]	THD U2 [%]	THD U3 [%]	I1 [A]	I2 [A]	I3 [A]	P [W]	Q [var]	S [VA]	E1 [Wh]	E2 [varh]	E3 [Wh]	E4 [varh]	
01/02/2012	235.52	234.81	235.43	1.75	1.46	0.58	145.3	118.7	126.3	89360	20290	91630	4946080000	1.85E+09	0	0	
01/02/2012 00:05	235.31	234.74	235.4	1.82	1.58	0.6	147.7	104.7	120.4	84870	21250	87490	4946087000	1.85E+09	0	0	
01/02/2012 00:10	235.29	234.65	235.15	1.83	1.58	0.6	141.5	106.1	139.1	87980	21740	90670	4946094000	1.85E+09	0	0	
01/02/2012 00:15	234.83	234.09	234.67	1.99	1.74	0.74	155.3	114.7	133.1	91770	21670	94390	4946102000	1.85E+09	0	0	
01/02/2012 00:20	235.13	234.37	234.9	1.94	1.71	0.69	150	114.6	124.9	88460	22490	91290	4946109000	1.85E+09	0	0	
01/02/2012 00:25	235.01	234.16	234.6	1.87	1.53	0.59	147.4	115.7	137.6	91840	23810	94940	4946117000	1.85E+09	0	0	

### QM ID number and data

LOCATION	HV/MV substation RTP <u>Prmskovo</u>		
device	transformer 1		
voltage level	110 kV		
Quantity	SCADA point	SCADA account	Meas. ID
L1 voltage	PRIMSK_VN.6009273_NAR0	PRIMSK_VN.6009273_NAR0< 15MAVG	394
L2 voltage	PRIMSK_VN.6009273_NAS0	PRIMSK_VN.6009273_NAS0< 15MAVG	395
L2 voltage	PRIMSK_VN.6009273_NAT0	PRIMSK_VN.6009273_NAT0< 15MAVG	396
L1-L2 voltage	PRIMSK_VN.6009273_NARS	PRIMSK_VN.6009273_NARS< 15MAVG	393
L1 current	PRIMSK_VN.6009273_TOKR	PRIMSK_VN.6009273_TOKR< 15MAVG	397
L2 current	PRIMSK_VN.6009273_TOKS	PRIMSK_VN.6009273_TOKS< 15MAVG	398
L2 current	PRIMSK_VN.6009273_TOKT	PRIMSK_VN.6009273_TOKT< 15MAVG	399
active power	PRIMSK_VN.6009273_PDEL	PRIMSK_VN.6009273_PDEL< 15MAVG	391
reactive power	PRIMSK_VN.6009273_PJAL	PRIMSK_VN.6009273_PJAL< 15MAVG	392
temperature	PRIMSK_VN.6009273_TEMP	PRIMSK_VN.6009273_TEMP< 15MAVG	332
tap level	PRIMSK_VN.6009273_REGS	PRIMSK_VN.6009273_REGS< 15MINS	409
LOCATION	HV/MV substation RTP <u>Prmskovo</u>		
device	transformer 1		

### SCADA sensor ID number

	A	B	C	D	E	F	G	H	I	J	K	L
1	Date	329	330	331	332	333	334	335	336	337	338	339
2	1/2/2012 0:00:00	9.321384429	44.8968	32.6444054	14.4368	23.45279	320.3249	-59.4092	20494.55	11812.73	11834.15	11832.39
3	1/2/2012 0:15:00	8.594027519	42.59529	32.0821152	14.4368	23.45279	278.1751	-61.403	20489.61	11810.89	11831.68	11836.37
4	1/2/2012 0:30:00	8.252724647	42.47889	32.0065689	14.4368	23.45279	273.8307	-61.4866	20486.15	11810.81	11832.23	11835.91
5	1/2/2012 0:45:00	7.614887237	41.54235	32.0541115	14.4368	23.45279	242.4907	-71.3438	20487.78	11808.69	11826.8	11833.86
6	1/2/2012 1:00:00	7.430253982	41.14957	31.5673771	14.4368	23.45279	237.1809	-70.8498	20568.76	11855.65	11876.08	11881.87
7	1/2/2012 1:15:00	7.318350791	41.49138	31.835907	14.4368	23.45279	236.9174	-66.5659	20591.67	11869.64	11893.61	11897.95
8	1/2/2012 1:30:00	7.4078722	41.16083	31.7464008	14.4368	23.45279	239.9295	-65.1397	20602.25	11872.94	11897.33	11903.37
9	1/2/2012 1:45:00	6.982646942	40.68179	31.6233349	14.4368	23.45279	227.7341	-65.7198	20671.93	11914.21	11939.39	11939.75
10	1/2/2012 2:00:00	6.837171554	40.87374	31.9086533	14.4368	23.45279	226.5891	-66.3835	20654.22	11903.89	11925.87	11927.75
11	1/2/2012 2:15:00	6.69169569	40.17897	31.4554769	14.4368	23.45279	216.6593	-67.6189	20679.77	11915.62	11941.89	11945.61

### SCADA sensor data

## 8. Network Parameter Information

### 77 Node Network Data

Node nam	Node No	Line No	From	To	Line R	Line X	Line B	Line G	Node nam	Node No	Line No	From	To	Line R	Line X	Line B	Line G
301	1	1	2	3	0.16843	0.087273	0	0	1137	39	39	40	41	0.075785	0.058347	0	0
1100	2	2	3	4	0.16843	0.087273	0	0	1138	40	40	41	42	0.075785	0.058347	0	0
1101	3	3	4	5	0.05157	0.01405	0	0	1139	41	41	42	43	0.075785	0.058347	0	0
1102	4	4	2	6	0.16843	0.087273	0	0	1140	42	42	43	44	0.075785	0.058347	0	0
1103	5	5	6	7	0.16843	0.087273	0	0	1141	43	43	44	45	0.075785	0.058347	0	0
1104	6	6	7	8	0.05157	0.01405	0	0	1142	44	44	45	46	0.075785	0.058347	0	0
1105	7	7	2	9	0.16843	0.087273	0	0	1143	45	45	46	47	0.075785	0.058347	0	0
1106	8	8	9	10	0.16843	0.087273	0	0	1144	46	46	40	48	0.04719	0.01281	0	0
1107	9	9	10	11	0.05157	0.01405	0	0	1145	47	47	42	49	0.04719	0.01281	0	0
1108	10	10	2	12	0.219835	0.113884	0	0	1146	48	48	43	50	0.04719	0.01281	0	0
1109	11	11	12	13	0.219835	0.113884	0	0	1147	49	49	45	51	0.04719	0.01281	0	0
1110	12	12	13	14	0.219835	0.113884	0	0	1148	50	50	47	52	0.04719	0.01281	0	0
1111	13	13	13	15	0.054793	0.014876	0	0	1149	51	51	2	53	0.054959	0.042314	0	0
1112	14	14	14	16	0.054793	0.014876	0	0	1150	52	52	53	54	0.054959	0.042314	0	0
1113	15	15	2	17	0.06157	0.047438	0	0	1151	53	53	54	55	0.054959	0.042314	0	0
1114	16	16	17	18	0.06157	0.047438	0	0	1152	54	54	55	56	0.054959	0.042314	0	0
1115	17	17	18	19	0.06157	0.047438	0	0	1153	55	55	56	57	0.054959	0.042314	0	0
1116	18	18	19	20	0.06157	0.047438	0	0	1154	56	56	57	58	0.054959	0.042314	0	0
1117	19	19	20	21	0.06157	0.047438	0	0	1155	57	57	58	59	0.054959	0.042314	0	0
1118	20	20	21	22	0.06157	0.047438	0	0	1156	58	58	59	60	0.054959	0.042314	0	0
1119	21	21	22	23	0.06157	0.047438	0	0	1157	59	59	60	61	0.054959	0.042314	0	0
1120	22	22	18	24	0.044793	0.012149	0	0	1158	60	60	61	62	0.054959	0.042314	0	0
1121	23	23	20	25	0.044793	0.012149	0	0	1159	61	61	62	63	0.054959	0.042314	0	0
1122	24	24	21	26	0.044793	0.012149	0	0	1160	62	62	63	64	0.054959	0.042314	0	0
1123	25	25	23	27	0.044793	0.012149	0	0	1161	63	63	64	65	0.054959	0.042314	0	0
1124	26	26	2	28	0.06157	0.047438	0	0	1162	64	64	65	66	0.054959	0.042314	0	0
1125	27	27	28	29	0.06157	0.047438	0	0	1163	65	65	66	67	0.054959	0.042314	0	0
1126	28	28	29	30	0.06157	0.047438	0	0	1164	66	66	67	68	0.054959	0.042314	0	0
1127	29	29	30	31	0.06157	0.047438	0	0	1165	67	67	54	69	0.060248	0.016364	0	0
1128	30	30	31	32	0.06157	0.047438	0	0	1166	68	68	56	70	0.060248	0.016364	0	0
1129	31	31	32	33	0.06157	0.047438	0	0	1167	69	69	57	71	0.060248	0.016364	0	0
1130	32	32	33	34	0.06157	0.047438	0	0	1168	70	70	59	72	0.060248	0.016364	0	0
1131	33	33	29	35	0.044793	0.012149	0	0	1169	71	71	61	73	0.060248	0.016364	0	0
1132	34	34	31	36	0.044793	0.012149	0	0	1170	72	72	63	74	0.060248	0.016364	0	0
1133	35	35	32	37	0.044793	0.012149	0	0	1171	73	73	64	75	0.060248	0.016364	0	0
1134	36	36	34	38	0.044793	0.012149	0	0	1172	74	74	66	76	0.060248	0.016364	0	0
1135	37	37	2	39	0.075785	0.058347	0	0	1173	75	75	68	77	0.060248	0.016364	0	0
1136	38	38	39	40	0.075785	0.058347	0	0	1174	76	76	1	2	0.002305	0.032032	0	0

### 356 Node Network Data

Node name	Node No	Line No	From	To	Line R	Line X	Line B	Line G	Node name	Node No	Line No	From	To	Line R	Line X	Line B	Line G
301	1	1	2	3	0.02303	0.015592	0	0	1188	90	90	91	92	0.036364	0.013609	0	0
1100	2	2	3	4	0.02303	0.015592	0	0	1189	91	91	64	93	0.036364	0.013609	0	0
1101	3	3	4	5	0.02303	0.015592	0	0	1190	92	92	93	94	0.036364	0.013609	0	0
1102	4	4	5	6	0.02303	0.015592	0	0	1191	93	93	65	95	0.036364	0.013609	0	0
1103	5	5	3	7	0.022773	0.016345	0	0	1192	94	94	95	96	0.036364	0.013609	0	0
1104	6	6	7	8	0.022773	0.016345	0	0	1193	95	95	66	97	0.036364	0.013609	0	0
1105	7	7	8	9	0.022773	0.016345	0	0	1194	96	96	97	98	0.036364	0.013609	0	0
1106	8	8	4	10	0.01708	0.012259	0	0	1195	97	97	67	99	0.036364	0.013609	0	0
1107	9	9	10	11	0.01708	0.012259	0	0	1196	98	98	99	100	0.036364	0.013609	0	0
1108	10	10	11	12	0.01708	0.012259	0	0	1197	99	99	68	101	0.036364	0.013609	0	0
1109	11	11	12	13	0.01708	0.012259	0	0	1198	100	100	101	102	0.036364	0.013609	0	0
1110	12	12	5	14	0.01708	0.012259	0	0	1199	101	101	69	103	0.036364	0.013609	0	0
1111	13	13	14	15	0.01708	0.012259	0	0	1200	102	102	103	104	0.036364	0.013609	0	0
1112	14	14	15	16	0.01708	0.012259	0	0	1201	103	103	70	105	0.036364	0.013609	0	0
1113	15	15	16	17	0.01708	0.012259	0	0	1202	104	104	105	106	0.036364	0.013609	0	0
1114	16	16	6	18	0.01708	0.012259	0	0	1203	105	105	71	107	0.036364	0.013609	0	0
1115	17	17	18	19	0.01708	0.012259	0	0	1204	106	106	107	108	0.036364	0.013609	0	0
1116	18	18	19	20	0.01708	0.012259	0	0	1205	107	107	72	109	0.036364	0.013609	0	0
1117	19	19	20	21	0.01708	0.012259	0	0	1206	108	108	109	110	0.036364	0.013609	0	0
1118	20	20	7	22	0.032654	0.012222	0	0	1207	109	109	73	111	0.036364	0.013609	0	0
1119	21	21	22	23	0.032654	0.012222	0	0	1208	110	110	111	112	0.036364	0.013609	0	0
1120	22	22	8	24	0.032654	0.012222	0	0	1209	111	111	74	113	0.036364	0.013609	0	0
1121	23	23	24	25	0.032654	0.012222	0	0	1210	112	112	113	114	0.036364	0.013609	0	0
1122	24	24	9	26	0.032654	0.012222	0	0	1211	113	113	75	115	0.036364	0.013609	0	0
1123	25	25	26	27	0.032654	0.012222	0	0	1212	114	114	115	116	0.036364	0.013609	0	0
1124	26	26	10	28	0.032654	0.012222	0	0	1213	115	115	76	117	0.036364	0.013609	0	0
1125	27	27	28	29	0.032654	0.012222	0	0	1214	116	116	117	118	0.036364	0.013609	0	0
1126	28	28	11	30	0.032654	0.012222	0	0	1215	117	117	77	119	0.036364	0.013609	0	0
1127	29	29	30	31	0.032654	0.012222	0	0	1216	118	118	119	120	0.036364	0.013609	0	0
1128	30	30	12	32	0.032654	0.012222	0	0	1217	119	119	78	121	0.036364	0.013609	0	0
1129	31	31	32	33	0.032654	0.012222	0	0	1218	120	120	121	122	0.036364	0.013609	0	0
1130	32	32	13	34	0.032654	0.012222	0	0	1219	121	121	79	123	0.036364	0.013609	0	0
1131	33	33	34	35	0.032654	0.012222	0	0	1220	122	122	123	124	0.036364	0.013609	0	0
1132	34	34	14	36	0.032654	0.012222	0	0	1221	123	123	80	125	0.036364	0.013609	0	0
1133	35	35	36	37	0.032654	0.012222	0	0	1222	124	124	125	126	0.036364	0.013609	0	0
1134	36	36	15	38	0.032654	0.012222	0	0	1223	125	125	81	127	0.036364	0.013609	0	0
1135	37	37	38	39	0.032654	0.012222	0	0	1224	126	126	127	128	0.036364	0.013609	0	0
1136	38	38	16	40	0.032654	0.012222	0	0	1225	127	127	82	129	0.036364	0.013609	0	0
1137	39	39	40	41	0.032654	0.012222	0	0	1226	128	128	129	130	0.036364	0.013609	0	0
1138	40	40	17	42	0.032654	0.012222	0	0	1227	129	129	83	131	0.024242	0.009073	0	0
1139	41	41	42	43	0.032654	0.012222	0	0	1228	130	130	131	132	0.024242	0.009073	0	0
1140	42	42	18	44	0.032654	0.012222	0	0	1229	131	131	132	133	0.024242	0.009073	0	0
1141	43	43	44	45	0.032654	0.012222	0	0	1230	132	132	84	134	0.024242	0.009073	0	0
1142	44	44	19	46	0.021772	0.008145	0	0	1231	133	133	134	135	0.024242	0.009073	0	0
1143	45	45	46	47	0.021772	0.008145	0	0	1232	134	134	135	136	0.024242	0.009073	0	0
1144	46	46	47	48	0.021772	0.008145	0	0	1233	135	135	2	137	0.008623	0.017612	0	0
1145	47	47	20	49	0.021772	0.008145	0	0	1234	136	136	137	138	0.008623	0.017612	0	0
1146	48	48	49	50	0.021772	0.008145	0	0	1235	137	137	138	139	0.008623	0.017612	0	0
1147	49	49	50	51	0.021772	0.008145	0	0	1236	138	138	139	140	0.008623	0.017612	0	0
1148	50	50	21	52	0.021772	0.008145	0	0	1237	139	139	140	141	0.008623	0.017612	0	0
1149	51	51	52	53	0.021772	0.008145	0	0	1238	140	140	141	142	0.008623	0.017612	0	0
1150	52	52	53	54	0.021772	0.008145	0	0	1239	141	141	142	143	0.008623	0.017612	0	0
1151	53	53	2	55	0.029247	0.019798	0	0	1240	142	142	143	144	0.008623	0.017612	0	0
1152	54	54	55	56	0.029247	0.019798	0	0	1241	143	143	144	145	0.008623	0.017612	0	0
1153	55	55	56	57	0.029247	0.019798	0	0	1242	144	144	145	146	0.008623	0.017612	0	0
1154	56	56	57	58	0.029247	0.019798	0	0	1243	145	145	146	147	0.008623	0.017612	0	0
1155	57	57	58	59	0.029247	0.019798	0	0	1244	146	146	147	148	0.01786	0.012819	0	0
1156	58	58	55	60	0.017374	0.01247	0	0	1245	147	147	148	149	0.01786	0.012819	0	0
1157	59	59	60	61	0.017374	0.01247	0	0	1246	148	148	149	150	0.01786	0.012819	0	0
1158	60	60	61	62	0.017374	0.01247	0	0	1247	149	149	150	151	0.01786	0.012819	0	0
1159	61	61	62	63	0.017374	0.01247	0	0	1248	150	150	151	152	0.01786	0.012819	0	0
1160	62	62	63	64	0.017374	0.01247	0	0	1249	151	151	138	153	0.01786	0.012819	0	0
1161	63	63	56	65	0.017374	0.01247	0	0	1250	152	152	153	154	0.01786	0.012819	0	0
1162	64	64	65	66	0.017374	0.01247	0	0	1251	153	153	154	155	0.01786	0.012819	0	0
1163	65	65	66	67	0.017374	0.01247	0	0	1252	154	154	155	156	0.01786	0.012819	0	0
1164	66	66	67	68	0.017374	0.01247	0	0	1253	155	155	156	157	0.01786	0.012819	0	0
1165	67	67	68	69	0.017374	0.01247	0	0	1254	156	156	139	158	0.01786	0.012819	0	0
1166	68	68	57	70	0.017374	0.01247	0	0	1255	157	157	158	159	0.01786	0.012819	0	0
1167	69	69	70	71	0.017374	0.01247	0	0	1256	158	158	159	160	0.01786	0.012819	0	0
1168	70	70	71	72	0.017374	0.01247	0	0	1257	159	159	160	161	0.01786	0.012819	0	0
1169	71	71	72	73	0.017374	0.01247	0	0	1258	160	160	161	162	0.01786	0.012819	0	0
1170	72	72	73	74	0.017374	0.01247	0	0	1259	161	161	140	163	0.01786	0.012819	0	0
1171	73	73	58	75	0.017374	0.01247	0	0	1260	162	162	163	164	0.01786	0.012819	0	0
1172	74	74	75	76	0.017374	0.01247	0	0	1261	163	163	164	165	0.01786	0.012819	0	0
1173	75	75	76	77	0.017374	0.01247	0	0	1262	164	164	165	166	0.01786	0.012819	0	0
1174	76	76	77	78	0.017374	0.01247	0	0	1263	165	165	166	167	0.01786	0.012819	0	0
1175	77	77	78	79	0.017374	0.01247	0	0	1264	166	166	141	168	0.01786	0.012819	0	0
1176	78	78	59	80	0.017374	0.01											

Node name	Node No	Line No	From	To	Line R	Line X	Line B	Line G	Node name	Node No	Line No	From	To	Line R	Line X	Line B	Line G
1277	179	179	180	181	0.01786	0.012819	0	0	1366	268	268	177	270	0.023912	0.008944	0	0
1278	180	180	181	182	0.01786	0.012819	0	0	1367	269	269	270	271	0.023912	0.008944	0	0
1279	181	181	144	183	0.01786	0.012819	0	0	1368	270	270	271	272	0.023912	0.008944	0	0
1280	182	182	183	184	0.01786	0.012819	0	0	1369	271	271	178	273	0.023912	0.008944	0	0
1281	183	183	184	185	0.01786	0.012819	0	0	1370	272	272	273	274	0.023912	0.008944	0	0
1282	184	184	185	186	0.01786	0.012819	0	0	1371	273	273	274	275	0.023912	0.008944	0	0
1283	185	185	186	187	0.01786	0.012819	0	0	1372	274	274	179	276	0.023912	0.008944	0	0
1284	186	186	145	188	0.014894	0.010689	0	0	1373	275	275	276	277	0.023912	0.008944	0	0
1285	187	187	188	189	0.014894	0.010689	0	0	1374	276	276	277	278	0.023912	0.008944	0	0
1286	188	188	189	190	0.014894	0.010689	0	0	1375	277	277	180	279	0.023912	0.008944	0	0
1287	189	189	190	191	0.014894	0.010689	0	0	1376	278	278	279	280	0.023912	0.008944	0	0
1288	190	190	191	192	0.014894	0.010689	0	0	1377	279	279	280	281	0.023912	0.008944	0	0
1289	191	191	192	193	0.014894	0.010689	0	0	1378	280	280	181	282	0.023912	0.008944	0	0
1290	192	192	146	194	0.014894	0.010689	0	0	1379	281	281	282	283	0.023912	0.008944	0	0
1291	193	193	194	195	0.014894	0.010689	0	0	1380	282	282	283	284	0.023912	0.008944	0	0
1292	194	194	195	196	0.014894	0.010689	0	0	1381	283	283	182	285	0.023912	0.008944	0	0
1293	195	195	196	197	0.014894	0.010689	0	0	1382	284	284	285	286	0.023912	0.008944	0	0
1294	196	196	197	198	0.014894	0.010689	0	0	1383	285	285	286	287	0.023912	0.008944	0	0
1295	197	197	198	199	0.014894	0.010689	0	0	1384	286	286	183	288	0.023912	0.008944	0	0
1296	198	198	147	200	0.014894	0.010689	0	0	1385	287	287	288	289	0.023912	0.008944	0	0
1297	199	199	200	201	0.014894	0.010689	0	0	1386	288	288	289	290	0.023912	0.008944	0	0
1298	200	200	201	202	0.014894	0.010689	0	0	1387	289	289	184	291	0.023912	0.008944	0	0
1299	201	201	202	203	0.014894	0.010689	0	0	1388	290	290	291	292	0.023912	0.008944	0	0
1300	202	202	203	204	0.014894	0.010689	0	0	1389	291	291	292	293	0.023912	0.008944	0	0
1301	203	203	204	205	0.014894	0.010689	0	0	1390	292	292	185	294	0.023912	0.008944	0	0
1302	204	204	148	206	0.035868	0.013425	0	0	1391	293	293	294	295	0.023912	0.008944	0	0
1303	205	205	206	207	0.035868	0.013425	0	0	1392	294	294	295	296	0.023912	0.008944	0	0
1304	206	206	149	208	0.035868	0.013425	0	0	1393	295	295	186	297	0.023912	0.008944	0	0
1305	207	207	208	209	0.035868	0.013425	0	0	1394	296	296	297	298	0.023912	0.008944	0	0
1306	208	208	150	210	0.035868	0.013425	0	0	1395	297	297	298	299	0.023912	0.008944	0	0
1307	209	209	210	211	0.035868	0.013425	0	0	1396	298	298	187	300	0.023912	0.008944	0	0
1308	210	210	151	212	0.035868	0.013425	0	0	1397	299	299	300	301	0.023912	0.008944	0	0
1309	211	211	212	213	0.035868	0.013425	0	0	1398	300	300	301	302	0.023912	0.008944	0	0
1310	212	212	152	214	0.035868	0.013425	0	0	1399	301	301	188	303	0.023912	0.008944	0	0
1311	213	213	214	215	0.035868	0.013425	0	0	1400	302	302	303	304	0.023912	0.008944	0	0
1312	214	214	153	216	0.035868	0.013425	0	0	1401	303	303	304	305	0.023912	0.008944	0	0
1313	215	215	216	217	0.035868	0.013425	0	0	1402	304	304	189	306	0.023912	0.008944	0	0
1314	216	216	154	218	0.035868	0.013425	0	0	1403	305	305	306	307	0.023912	0.008944	0	0
1315	217	217	218	219	0.035868	0.013425	0	0	1404	306	306	307	308	0.023912	0.008944	0	0
1316	218	218	155	220	0.035868	0.013425	0	0	1405	307	307	190	309	0.023912	0.008944	0	0
1317	219	219	220	221	0.035868	0.013425	0	0	1406	308	308	309	310	0.023912	0.008944	0	0
1318	220	220	156	222	0.035868	0.013425	0	0	1407	309	309	310	311	0.023912	0.008944	0	0
1319	221	221	222	223	0.035868	0.013425	0	0	1408	310	310	191	312	0.023912	0.008944	0	0
1320	222	222	157	224	0.035868	0.013425	0	0	1409	311	311	312	313	0.023912	0.008944	0	0
1321	223	223	224	225	0.035868	0.013425	0	0	1410	312	312	313	314	0.023912	0.008944	0	0
1322	224	224	158	226	0.035868	0.013425	0	0	1411	313	313	192	315	0.023912	0.008944	0	0
1323	225	225	226	227	0.035868	0.013425	0	0	1412	314	314	315	316	0.023912	0.008944	0	0
1324	226	226	159	228	0.035868	0.013425	0	0	1413	315	315	316	317	0.023912	0.008944	0	0
1325	227	227	228	229	0.035868	0.013425	0	0	1414	316	316	193	318	0.023912	0.008944	0	0
1326	228	228	160	230	0.035868	0.013425	0	0	1415	317	317	318	319	0.023912	0.008944	0	0
1327	229	229	230	231	0.035868	0.013425	0	0	1416	318	318	319	320	0.023912	0.008944	0	0
1328	230	230	161	232	0.035868	0.013425	0	0	1417	319	319	194	321	0.023912	0.008944	0	0
1329	231	231	232	233	0.035868	0.013425	0	0	1418	320	320	321	322	0.023912	0.008944	0	0
1330	232	232	162	234	0.035868	0.013425	0	0	1419	321	321	322	323	0.023912	0.008944	0	0
1331	233	233	234	235	0.035868	0.013425	0	0	1420	322	322	195	324	0.023912	0.008944	0	0
1332	234	234	163	236	0.035868	0.013425	0	0	1421	323	323	324	325	0.023912	0.008944	0	0
1333	235	235	236	237	0.035868	0.013425	0	0	1422	324	324	325	326	0.023912	0.008944	0	0
1334	236	236	164	238	0.035868	0.013425	0	0	1423	325	325	196	327	0.023912	0.008944	0	0
1335	237	237	238	239	0.035868	0.013425	0	0	1424	326	326	327	328	0.023912	0.008944	0	0
1336	238	238	165	240	0.035868	0.013425	0	0	1425	327	327	328	329	0.023912	0.008944	0	0
1337	239	239	240	241	0.035868	0.013425	0	0	1426	328	328	197	330	0.023912	0.008944	0	0
1338	240	240	166	242	0.035868	0.013425	0	0	1427	329	329	330	331	0.023912	0.008944	0	0
1339	241	241	242	243	0.035868	0.013425	0	0	1428	330	330	331	332	0.023912	0.008944	0	0
1340	242	242	167	244	0.035868	0.013425	0	0	1429	331	331	198	333	0.023912	0.008944	0	0
1341	243	243	244	245	0.035868	0.013425	0	0	1430	332	332	333	334	0.023912	0.008944	0	0
1342	244	244	168	246	0.035868	0.013425	0	0	1431	333	333	334	335	0.023912	0.008944	0	0
1343	245	245	246	247	0.035868	0.013425	0	0	1432	334	334	199	336	0.023912	0.008944	0	0
1344	246	246	169	248	0.035868	0.013425	0	0	1433	335	335	336	337	0.023912	0.008944	0	0
1345	247	247	248	249	0.035868	0.013425	0	0	1434	336	336	337	338	0.023912	0.008944	0	0
1346	248	248	170	250	0.035868	0.013425	0	0	1435	337	337	200	339	0.023912	0.008944	0	0
1347	249	249	250	251	0.035868	0.013425	0	0	1436	338	338	339	340	0.023912	0.008944	0	0
1348	250	250	171	252	0.023912	0.008944	0	0	1437	339	339	340	341	0.023912	0.008944	0	0
1349	251	251	252	253	0.023912	0.008944	0	0	1438	340	340	201	342	0.023912	0.008944	0	0
1350	252	252	253	254	0.023912	0.008944	0	0	1439	341	341	342	343	0.023912	0.008944	0	0
1351	253	253	172	255	0.023912	0.008944	0	0	1440	342	342	343	344	0.023912	0		

## 711 Node Network Data

Line No	From	To	Line R	Line X	Line B	Line G	Node name	Node No	Line No	From	To	Line R	Line X	Line B	Line G
1	2	3	0.023030303	0.015592287	0	0	1189	91	91	64	93	0.036363636	0.013608815	0	0
2	3	4	0.023030303	0.015592287	0	0	1190	92	92	93	94	0.036363636	0.013608815	0	0
3	4	5	0.023030303	0.015592287	0	0	1191	93	93	95	95	0.036363636	0.013608815	0	0
4	5	6	0.023030303	0.015592287	0	0	1192	94	94	95	96	0.036363636	0.013608815	0	0
5	3	7	0.022773186	0.016345271	0	0	1193	95	95	66	97	0.036363636	0.013608815	0	0
6	7	8	0.022773186	0.016345271	0	0	1194	96	96	97	97	0.036363636	0.013608815	0	0
7	8	9	0.022773186	0.016345271	0	0	1195	97	97	67	99	0.036363636	0.013608815	0	0
8	4	10	0.01707989	0.012258953	0	0	1196	98	98	99	100	0.036363636	0.013608815	0	0
9	10	11	0.01707989	0.012258953	0	0	1197	99	99	68	101	0.036363636	0.013608815	0	0
10	11	12	0.01707989	0.012258953	0	0	1198	100	100	101	102	0.036363636	0.013608815	0	0
11	12	13	0.01707989	0.012258953	0	0	1199	101	101	69	103	0.036363636	0.013608815	0	0
12	5	14	0.01707989	0.012258953	0	0	1200	102	102	103	104	0.036363636	0.013608815	0	0
13	14	15	0.01707989	0.012258953	0	0	1201	103	103	70	105	0.036363636	0.013608815	0	0
14	15	16	0.01707989	0.012258953	0	0	1202	104	104	105	106	0.036363636	0.013608815	0	0
15	16	17	0.01707989	0.012258953	0	0	1203	105	105	71	107	0.036363636	0.013608815	0	0
16	6	18	0.01707989	0.012258953	0	0	1204	106	106	107	108	0.036363636	0.013608815	0	0
17	18	19	0.01707989	0.012258953	0	0	1205	107	107	72	109	0.036363636	0.013608815	0	0
18	19	20	0.01707989	0.012258953	0	0	1206	108	108	109	110	0.036363636	0.013608815	0	0
19	20	21	0.01707989	0.012258953	0	0	1207	109	109	73	111	0.036363636	0.013608815	0	0
20	7	22	0.032653811	0.012222222	0	0	1208	110	110	111	112	0.036363636	0.013608815	0	0
21	22	23	0.032653811	0.012222222	0	0	1209	111	111	74	113	0.036363636	0.013608815	0	0
22	8	24	0.032653811	0.012222222	0	0	1210	112	112	113	114	0.036363636	0.013608815	0	0
23	24	25	0.032653811	0.012222222	0	0	1211	113	113	75	115	0.036363636	0.013608815	0	0
24	9	26	0.032653811	0.012222222	0	0	1212	114	114	115	116	0.036363636	0.013608815	0	0
25	26	27	0.032653811	0.012222222	0	0	1213	115	115	76	117	0.036363636	0.013608815	0	0
26	10	28	0.032653811	0.012222222	0	0	1214	116	116	117	118	0.036363636	0.013608815	0	0
27	28	29	0.032653811	0.012222222	0	0	1215	117	117	77	119	0.036363636	0.013608815	0	0
28	11	30	0.032653811	0.012222222	0	0	1216	118	118	119	120	0.036363636	0.013608815	0	0
29	30	31	0.032653811	0.012222222	0	0	1217	119	119	78	121	0.036363636	0.013608815	0	0
30	12	32	0.032653811	0.012222222	0	0	1218	120	120	121	122	0.036363636	0.013608815	0	0
31	32	33	0.032653811	0.012222222	0	0	1219	121	121	79	123	0.036363636	0.013608815	0	0
32	13	34	0.032653811	0.012222222	0	0	1220	122	122	123	124	0.036363636	0.013608815	0	0
33	34	35	0.032653811	0.012222222	0	0	1221	123	123	80	125	0.036363636	0.013608815	0	0
34	14	36	0.032653811	0.012222222	0	0	1222	124	124	125	126	0.036363636	0.013608815	0	0
35	36	37	0.032653811	0.012222222	0	0	1223	125	125	81	127	0.036363636	0.013608815	0	0
36	15	38	0.032653811	0.012222222	0	0	1224	126	126	127	128	0.036363636	0.013608815	0	0
37	38	39	0.032653811	0.012222222	0	0	1225	127	127	82	129	0.036363636	0.013608815	0	0
38	16	40	0.032653811	0.012222222	0	0	1226	128	128	129	130	0.036363636	0.013608815	0	0
39	40	41	0.032653811	0.012222222	0	0	1227	129	129	83	131	0.024242424	0.009072544	0	0
40	17	42	0.032653811	0.012222222	0	0	1228	130	130	131	132	0.024242424	0.009072544	0	0
41	42	43	0.032653811	0.012222222	0	0	1229	131	131	132	133	0.024242424	0.009072544	0	0
42	18	44	0.032653811	0.012222222	0	0	1230	132	132	84	134	0.024242424	0.009072544	0	0
43	44	45	0.032653811	0.012222222	0	0	1231	133	133	134	135	0.024242424	0.009072544	0	0
44	19	46	0.021772268	0.008145087	0	0	1232	134	134	135	136	0.024242424	0.009072544	0	0
45	46	47	0.021772268	0.008145087	0	0	1233	135	135	2	137	0.00862259	0.017612489	0	0
46	47	48	0.021772268	0.008145087	0	0	1234	136	136	137	138	0.00862259	0.017612489	0	0
47	20	49	0.021772268	0.008145087	0	0	1235	137	137	138	139	0.00862259	0.017612489	0	0
48	49	50	0.021772268	0.008145087	0	0	1236	138	138	139	140	0.00862259	0.017612489	0	0
49	50	51	0.021772268	0.008145087	0	0	1237	139	139	140	141	0.00862259	0.017612489	0	0
50	21	52	0.021772268	0.008145087	0	0	1238	140	140	141	142	0.00862259	0.017612489	0	0
51	52	53	0.021772268	0.008145087	0	0	1239	141	141	142	143	0.00862259	0.017612489	0	0
52	53	54	0.021772268	0.008145087	0	0	1240	142	142	143	144	0.00862259	0.017612489	0	0
53	2	55	0.029247016	0.01979798	0	0	1241	143	143	144	145	0.00862259	0.017612489	0	0
54	55	56	0.029247016	0.01979798	0	0	1242	144	144	145	146	0.00862259	0.017612489	0	0
55	56	57	0.029247016	0.01979798	0	0	1243	145	145	146	147	0.00862259	0.017612489	0	0
56	57	58	0.029247016	0.01979798	0	0	1244	146	146	137	148	0.017860422	0.0128191	0	0
57	58	59	0.029247016	0.01979798	0	0	1245	147	147	148	149	0.017860422	0.0128191	0	0
58	55	60	0.017373737	0.012470156	0	0	1246	148	148	149	150	0.017860422	0.0128191	0	0
59	60	61	0.017373737	0.012470156	0	0	1247	149	149	150	151	0.017860422	0.0128191	0	0
60	61	62	0.017373737	0.012470156	0	0	1248	150	150	151	152	0.017860422	0.0128191	0	0
61	62	63	0.017373737	0.012470156	0	0	1249	151	151	138	153	0.017860422	0.0128191	0	0
62	63	64	0.017373737	0.012470156	0	0	1250	152	152	154	155	0.017860422	0.0128191	0	0
63	56	65	0.017373737	0.012470156	0	0	1251	153	153	154	155	0.017860422	0.0128191	0	0
64	65	66	0.017373737	0.012470156	0	0	1252	154	154	155	156	0.017860422	0.0128191	0	0
65	66	67	0.017373737	0.012470156	0	0	1253	155	155	156	157	0.017860422	0.0128191	0	0
66	67	68	0.017373737	0.012470156	0	0	1254	156	156	139	158	0.017860422	0.0128191	0	0
67	68	69	0.017373737	0.012470156	0	0	1255	157	157	158	159	0.017860422	0.0128191	0	0
68	57	70	0.017373737	0.012470156	0	0	1256	158	158	159	160	0.017860422	0.0128191	0	0
69	70	71	0.017373737	0.012470156	0	0	1257	159	159	160	161	0.017860422	0.0128191	0	0
70	71	72	0.017373737	0.012470156	0	0	1258	160	160	161	162	0.017860422	0.0128191	0	0
71	72	73	0.017373737	0.012470156	0	0	1259	161	161	140	163	0.017860422	0.0128191	0	0
72	73	74	0.017373737	0.012470156	0	0	1260	162	162	163	164	0.017860422	0.0128191	0	0
73	58	75	0.017373737	0.012470156	0	0	1261	163	163	164	165	0.017860422	0.0128191	0	0
74	75	76	0.017373737	0.012470156	0	0	1262	164	164	165	166	0.017860422	0.0128191	0	0
75	76	77	0.017373737	0.012470156	0	0	1263	165	165	166	167	0.017860422	0.0128191	0	0
76	77	78	0.017373737	0.012470156	0	0	1264	166	166	141	168	0.017860422	0.0128191	0	0
77	78	79	0.017373737	0.012470156	0	0	1265	167	167	168	169	0.017860422	0.0128191	0	0
78	59	80	0.017373737	0.012470156	0	0	1266	168	168	169	170	0.017860422	0.0128191	0	0
79	80	81	0.017373737	0.012470156	0	0	1267	169	169	170	171	0.017860422	0.0128191	0	0
80	81	82	0.017373737	0.012470156	0	0	1268								

Line No	From	To	Line R	Line X	Line B	Line G	Node name	Node No	Line No	From	To	Line R	Line X	Line B	Line G
181	144	183	0.017860422	0.0128191	0	0	1369	271	271	178	273	0.023911846	0.008943985	0	0
182	183	184	0.017860422	0.0128191	0	0	1370	272	272	274	274	0.023911846	0.008943985	0	0
183	184	185	0.017860422	0.0128191	0	0	1371	273	273	274	275	0.023911846	0.008943985	0	0
184	185	186	0.017860422	0.0128191	0	0	1372	274	274	179	276	0.023911846	0.008943985	0	0
185	186	187	0.017860422	0.0128191	0	0	1373	275	275	276	277	0.023911846	0.008943985	0	0
186	145	188	0.014894399	0.010688705	0	0	1374	276	276	277	278	0.023911846	0.008943985	0	0
187	188	189	0.014894399	0.010688705	0	0	1375	277	277	180	279	0.023911846	0.008943985	0	0
188	189	190	0.014894399	0.010688705	0	0	1376	278	278	279	280	0.023911846	0.008943985	0	0
189	190	191	0.014894399	0.010688705	0	0	1377	279	279	280	281	0.023911846	0.008943985	0	0
190	191	192	0.014894399	0.010688705	0	0	1378	280	280	181	282	0.023911846	0.008943985	0	0
191	192	193	0.014894399	0.010688705	0	0	1379	281	281	282	283	0.023911846	0.008943985	0	0
192	146	194	0.014894399	0.010688705	0	0	1380	282	282	283	284	0.023911846	0.008943985	0	0
193	194	195	0.014894399	0.010688705	0	0	1381	283	283	182	285	0.023911846	0.008943985	0	0
194	195	196	0.014894399	0.010688705	0	0	1382	284	284	285	286	0.023911846	0.008943985	0	0
195	196	197	0.014894399	0.010688705	0	0	1383	285	285	286	287	0.023911846	0.008943985	0	0
196	197	198	0.014894399	0.010688705	0	0	1384	286	286	183	288	0.023911846	0.008943985	0	0
197	198	199	0.014894399	0.010688705	0	0	1385	287	287	288	289	0.023911846	0.008943985	0	0
198	147	200	0.014894399	0.010688705	0	0	1386	288	288	289	290	0.023911846	0.008943985	0	0
199	200	201	0.014894399	0.010688705	0	0	1387	289	289	184	291	0.023911846	0.008943985	0	0
200	201	202	0.014894399	0.010688705	0	0	1388	290	290	291	292	0.023911846	0.008943985	0	0
201	202	203	0.014894399	0.010688705	0	0	1389	291	291	292	293	0.023911846	0.008943985	0	0
202	203	204	0.014894399	0.010688705	0	0	1390	292	292	185	294	0.023911846	0.008943985	0	0
203	204	205	0.014894399	0.010688705	0	0	1391	293	293	294	295	0.023911846	0.008943985	0	0
204	148	206	0.035867769	0.013425161	0	0	1392	294	294	295	296	0.023911846	0.008943985	0	0
205	206	207	0.035867769	0.013425161	0	0	1393	295	295	186	297	0.023911846	0.008943985	0	0
206	149	208	0.035867769	0.013425161	0	0	1394	296	296	297	298	0.023911846	0.008943985	0	0
207	208	209	0.035867769	0.013425161	0	0	1395	297	297	298	299	0.023911846	0.008943985	0	0
208	150	210	0.035867769	0.013425161	0	0	1396	298	298	187	300	0.023911846	0.008943985	0	0
209	210	211	0.035867769	0.013425161	0	0	1397	299	299	300	301	0.023911846	0.008943985	0	0
210	151	212	0.035867769	0.013425161	0	0	1398	300	300	301	302	0.023911846	0.008943985	0	0
211	212	213	0.035867769	0.013425161	0	0	1399	301	301	188	303	0.023911846	0.008943985	0	0
212	152	214	0.035867769	0.013425161	0	0	1400	302	302	304	304	0.023911846	0.008943985	0	0
213	214	215	0.035867769	0.013425161	0	0	1401	303	303	304	305	0.023911846	0.008943985	0	0
214	153	216	0.035867769	0.013425161	0	0	1402	304	304	189	306	0.023911846	0.008943985	0	0
215	216	217	0.035867769	0.013425161	0	0	1403	305	305	306	307	0.023911846	0.008943985	0	0
216	154	218	0.035867769	0.013425161	0	0	1404	306	306	307	308	0.023911846	0.008943985	0	0
217	218	219	0.035867769	0.013425161	0	0	1405	307	307	190	309	0.023911846	0.008943985	0	0
218	155	220	0.035867769	0.013425161	0	0	1406	308	308	309	310	0.023911846	0.008943985	0	0
219	220	221	0.035867769	0.013425161	0	0	1407	309	309	310	311	0.023911846	0.008943985	0	0
220	156	222	0.035867769	0.013425161	0	0	1408	310	310	191	312	0.023911846	0.008943985	0	0
221	222	223	0.035867769	0.013425161	0	0	1409	311	311	312	313	0.023911846	0.008943985	0	0
222	157	224	0.035867769	0.013425161	0	0	1410	312	312	313	314	0.023911846	0.008943985	0	0
223	224	225	0.035867769	0.013425161	0	0	1411	313	313	192	315	0.023911846	0.008943985	0	0
224	158	226	0.035867769	0.013425161	0	0	1412	314	314	315	316	0.023911846	0.008943985	0	0
225	226	227	0.035867769	0.013425161	0	0	1413	315	315	316	317	0.023911846	0.008943985	0	0
226	159	228	0.035867769	0.013425161	0	0	1414	316	316	193	318	0.023911846	0.008943985	0	0
227	228	229	0.035867769	0.013425161	0	0	1415	317	317	318	319	0.023911846	0.008943985	0	0
228	160	230	0.035867769	0.013425161	0	0	1416	318	318	319	320	0.023911846	0.008943985	0	0
229	230	231	0.035867769	0.013425161	0	0	1417	319	319	194	321	0.023911846	0.008943985	0	0
230	161	232	0.035867769	0.013425161	0	0	1418	320	320	321	322	0.023911846	0.008943985	0	0
231	232	233	0.035867769	0.013425161	0	0	1419	321	321	322	323	0.023911846	0.008943985	0	0
232	162	234	0.035867769	0.013425161	0	0	1420	322	322	195	324	0.023911846	0.008943985	0	0
233	234	235	0.035867769	0.013425161	0	0	1421	323	323	324	325	0.023911846	0.008943985	0	0
234	163	236	0.035867769	0.013425161	0	0	1422	324	324	325	326	0.023911846	0.008943985	0	0
235	236	237	0.035867769	0.013425161	0	0	1423	325	325	196	327	0.023911846	0.008943985	0	0
236	164	238	0.035867769	0.013425161	0	0	1424	326	326	327	328	0.023911846	0.008943985	0	0
237	238	239	0.035867769	0.013425161	0	0	1425	327	327	328	329	0.023911846	0.008943985	0	0
238	165	240	0.035867769	0.013425161	0	0	1426	328	328	197	330	0.023911846	0.008943985	0	0
239	240	241	0.035867769	0.013425161	0	0	1427	329	329	330	331	0.023911846	0.008943985	0	0
240	166	242	0.035867769	0.013425161	0	0	1428	330	330	331	332	0.023911846	0.008943985	0	0
241	242	243	0.035867769	0.013425161	0	0	1429	331	331	198	333	0.023911846	0.008943985	0	0
242	167	244	0.035867769	0.013425161	0	0	1430	332	332	333	334	0.023911846	0.008943985	0	0
243	244	245	0.035867769	0.013425161	0	0	1431	333	333	334	335	0.023911846	0.008943985	0	0
244	168	246	0.035867769	0.013425161	0	0	1432	334	334	199	336	0.023911846	0.008943985	0	0
245	246	247	0.035867769	0.013425161	0	0	1433	335	335	336	337	0.023911846	0.008943985	0	0
246	169	248	0.035867769	0.013425161	0	0	1434	336	336	337	338	0.023911846	0.008943985	0	0
247	248	249	0.035867769	0.013425161	0	0	1435	337	337	200	339	0.023911846	0.008943985	0	0
248	170	250	0.035867769	0.013425161	0	0	1436	338	338	339	340	0.023911846	0.008943985	0	0
249	250	251	0.035867769	0.013425161	0	0	1437	339	339	340	341	0.023911846	0.008943985	0	0
250	171	252	0.023911846	0.008943985	0	0	1438	340	340	201	342	0.023911846	0.008943985	0	0
251	252	253	0.023911846	0.008943985	0	0	1439	341	341	342	343	0.023911846	0.008943985	0	0
252	253	254	0.023911846	0.008943985	0	0	1440	342	342	343	344	0.023911846	0.008943985	0	0
253	172	255	0.023911846	0.008943985	0	0	1441	343	343	202	345	0.023911846	0.008943985	0	0
254	255	256	0.023911846	0.008943985	0	0	1442	344	344	345	346	0.023911846	0.008943985	0	0
255	256	257	0.023911846	0.008943985	0	0	1443	345	345	346	347	0.023911846	0.008943985	0	0
256	173	258	0.023911846	0.008943985	0	0	1444	346	346	203	348	0.023911846	0.008943985	0	0
257	258	259	0.023911846	0.008943985	0	0	1445	347	347	348	349	0.023911846	0.008943985	0	0
258	259	260	0.023911846	0.0089439											

Line No	From	To	Line R	Line X	Line B	Line G	Node name	Node No	Line No	From	To	Line R	Line X	Line B	Line G
361	362	363	0.022773186	0.016345271	0	0	10194	451	451	452	453	0.036363636	0.013608815	0	0
362	363	364	0.022773186	0.016345271	0	0	10195	452	452	422	454	0.036363636	0.013608815	0	0
363	359	365	0.01707989	0.012258953	0	0	10196	453	453	454	455	0.036363636	0.013608815	0	0
364	365	366	0.01707989	0.012258953	0	0	10197	454	454	423	456	0.036363636	0.013608815	0	0
365	366	367	0.01707989	0.012258953	0	0	10198	455	455	456	457	0.036363636	0.013608815	0	0
366	367	368	0.01707989	0.012258953	0	0	10199	456	456	424	458	0.036363636	0.013608815	0	0
367	360	369	0.01707989	0.012258953	0	0	10200	457	457	458	459	0.036363636	0.013608815	0	0
368	369	370	0.01707989	0.012258953	0	0	10201	458	458	425	460	0.036363636	0.013608815	0	0
369	370	371	0.01707989	0.012258953	0	0	10202	459	459	460	461	0.036363636	0.013608815	0	0
370	371	372	0.01707989	0.012258953	0	0	10203	460	460	426	462	0.036363636	0.013608815	0	0
371	361	373	0.01707989	0.012258953	0	0	10204	461	461	462	463	0.036363636	0.013608815	0	0
372	373	374	0.01707989	0.012258953	0	0	10205	462	462	427	464	0.036363636	0.013608815	0	0
373	374	375	0.01707989	0.012258953	0	0	10206	463	463	464	465	0.036363636	0.013608815	0	0
374	375	376	0.01707989	0.012258953	0	0	10207	464	464	428	466	0.036363636	0.013608815	0	0
375	362	377	0.032653811	0.012222222	0	0	10208	465	465	466	467	0.036363636	0.013608815	0	0
376	377	378	0.032653811	0.012222222	0	0	10209	466	466	429	468	0.036363636	0.013608815	0	0
377	363	379	0.032653811	0.012222222	0	0	10210	467	467	468	469	0.036363636	0.013608815	0	0
378	379	380	0.032653811	0.012222222	0	0	10211	468	468	430	470	0.036363636	0.013608815	0	0
379	364	381	0.032653811	0.012222222	0	0	10212	469	469	470	471	0.036363636	0.013608815	0	0
380	381	382	0.032653811	0.012222222	0	0	10213	470	470	431	472	0.036363636	0.013608815	0	0
381	365	383	0.032653811	0.012222222	0	0	10214	471	471	472	473	0.036363636	0.013608815	0	0
382	383	384	0.032653811	0.012222222	0	0	10215	472	472	432	474	0.036363636	0.013608815	0	0
383	366	385	0.032653811	0.012222222	0	0	10216	473	473	474	475	0.036363636	0.013608815	0	0
384	385	386	0.032653811	0.012222222	0	0	10217	474	474	433	476	0.036363636	0.013608815	0	0
385	367	387	0.032653811	0.012222222	0	0	10218	475	475	476	477	0.036363636	0.013608815	0	0
386	387	388	0.032653811	0.012222222	0	0	10219	476	476	434	478	0.036363636	0.013608815	0	0
387	368	389	0.032653811	0.012222222	0	0	10220	477	477	478	479	0.036363636	0.013608815	0	0
388	389	390	0.032653811	0.012222222	0	0	10221	478	478	435	480	0.036363636	0.013608815	0	0
389	369	391	0.032653811	0.012222222	0	0	10222	479	479	480	481	0.036363636	0.013608815	0	0
390	391	392	0.032653811	0.012222222	0	0	10223	480	480	436	482	0.036363636	0.013608815	0	0
391	370	393	0.032653811	0.012222222	0	0	10224	481	481	482	483	0.036363636	0.013608815	0	0
392	393	394	0.032653811	0.012222222	0	0	10225	482	482	437	484	0.036363636	0.013608815	0	0
393	371	395	0.032653811	0.012222222	0	0	10226	483	483	484	485	0.036363636	0.013608815	0	0
394	395	396	0.032653811	0.012222222	0	0	10227	484	484	438	486	0.024242424	0.009072544	0	0
395	372	397	0.032653811	0.012222222	0	0	10228	485	485	486	487	0.024242424	0.009072544	0	0
396	397	398	0.032653811	0.012222222	0	0	10229	486	486	487	488	0.024242424	0.009072544	0	0
397	373	399	0.032653811	0.012222222	0	0	10230	487	487	489	489	0.024242424	0.009072544	0	0
398	399	400	0.032653811	0.012222222	0	0	10231	488	488	489	490	0.024242424	0.009072544	0	0
399	374	401	0.021772268	0.008145087	0	0	10232	489	489	490	491	0.024242424	0.009072544	0	0
400	401	402	0.021772268	0.008145087	0	0	10233	490	490	357	492	0.00862259	0.017612489	0	0
401	402	403	0.021772268	0.008145087	0	0	10234	491	491	492	493	0.00862259	0.017612489	0	0
402	375	404	0.021772268	0.008145087	0	0	10235	492	492	493	494	0.00862259	0.017612489	0	0
403	404	405	0.021772268	0.008145087	0	0	10236	493	493	494	495	0.00862259	0.017612489	0	0
404	405	406	0.021772268	0.008145087	0	0	10237	494	494	495	496	0.00862259	0.017612489	0	0
405	376	407	0.021772268	0.008145087	0	0	10238	495	495	496	497	0.00862259	0.017612489	0	0
406	407	408	0.021772268	0.008145087	0	0	10239	496	496	497	498	0.00862259	0.017612489	0	0
407	408	409	0.021772268	0.008145087	0	0	10240	497	497	498	499	0.00862259	0.017612489	0	0
408	357	410	0.029247016	0.01979798	0	0	10241	498	498	500	500	0.00862259	0.017612489	0	0
409	410	411	0.029247016	0.01979798	0	0	10242	499	499	500	501	0.00862259	0.017612489	0	0
410	411	412	0.029247016	0.01979798	0	0	10243	500	500	501	502	0.00862259	0.017612489	0	0
411	412	413	0.029247016	0.01979798	0	0	10244	501	501	492	503	0.017860422	0.0128191	0	0
412	413	414	0.029247016	0.01979798	0	0	10245	502	502	503	504	0.017860422	0.0128191	0	0
413	410	415	0.017373737	0.012470156	0	0	10246	503	503	504	505	0.017860422	0.0128191	0	0
414	415	416	0.017373737	0.012470156	0	0	10247	504	504	505	506	0.017860422	0.0128191	0	0
415	416	417	0.017373737	0.012470156	0	0	10248	505	505	506	507	0.017860422	0.0128191	0	0
416	417	418	0.017373737	0.012470156	0	0	10249	506	506	493	508	0.017860422	0.0128191	0	0
417	418	419	0.017373737	0.012470156	0	0	10250	507	507	508	509	0.017860422	0.0128191	0	0
418	411	420	0.017373737	0.012470156	0	0	10251	508	508	509	510	0.017860422	0.0128191	0	0
419	420	421	0.017373737	0.012470156	0	0	10252	509	509	510	511	0.017860422	0.0128191	0	0
420	421	422	0.017373737	0.012470156	0	0	10253	510	510	511	512	0.017860422	0.0128191	0	0
421	422	423	0.017373737	0.012470156	0	0	10254	511	511	494	513	0.017860422	0.0128191	0	0
422	423	424	0.017373737	0.012470156	0	0	10255	512	512	513	514	0.017860422	0.0128191	0	0
423	412	425	0.017373737	0.012470156	0	0	10256	513	513	514	515	0.017860422	0.0128191	0	0
424	425	426	0.017373737	0.012470156	0	0	10257	514	514	515	516	0.017860422	0.0128191	0	0
425	426	427	0.017373737	0.012470156	0	0	10258	515	515	516	517	0.017860422	0.0128191	0	0
426	427	428	0.017373737	0.012470156	0	0	10259	516	516	495	518	0.017860422	0.0128191	0	0
427	428	429	0.017373737	0.012470156	0	0	10260	517	517	518	519	0.017860422	0.0128191	0	0
428	413	430	0.017373737	0.012470156	0	0	10261	518	518	519	520	0.017860422	0.0128191	0	0
429	430	431	0.017373737	0.012470156	0	0	10262	519	519	520	521	0.017860422	0.0128191	0	0
430	431	432	0.017373737	0.012470156	0	0	10263	520	520	521	522	0.017860422	0.0128191	0	0
431	432	433	0.017373737	0.012470156	0	0	10264	521	521	496	523	0.017860422	0.0128191	0	0
432	433	434	0.017373737	0.012470156	0	0	10265	522	522	524	524	0.017860422	0.0128191	0	0
433	414	435	0.017373737	0.012470156	0	0	10266	523	523	524	525	0.017860422	0.0128191	0	0
434	435	436	0.017373737	0.012470156	0	0	10267	524	524	525	526	0.017860422	0.0128191	0	0
435	436	437	0.017373737	0.012470156	0	0	10268	525	525	526	527	0.017860422	0.0128191	0	0
436	437	438	0.017373737	0.012470156	0	0	10269	526	526	497	528	0.017860422	0.0128191	0	0
437	438	439	0.017373737	0.012470156	0	0	10270	527	527	528	529	0.017860422	0.0128191	0	0
438	415	440	0.036363636	0.01											





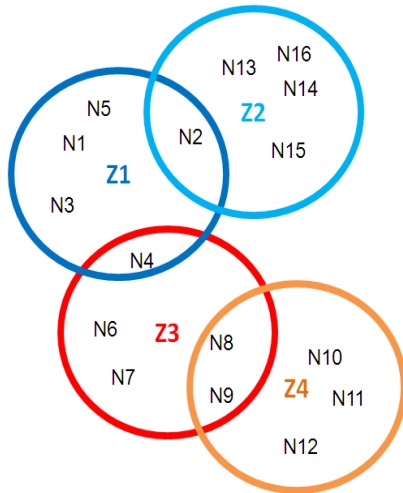
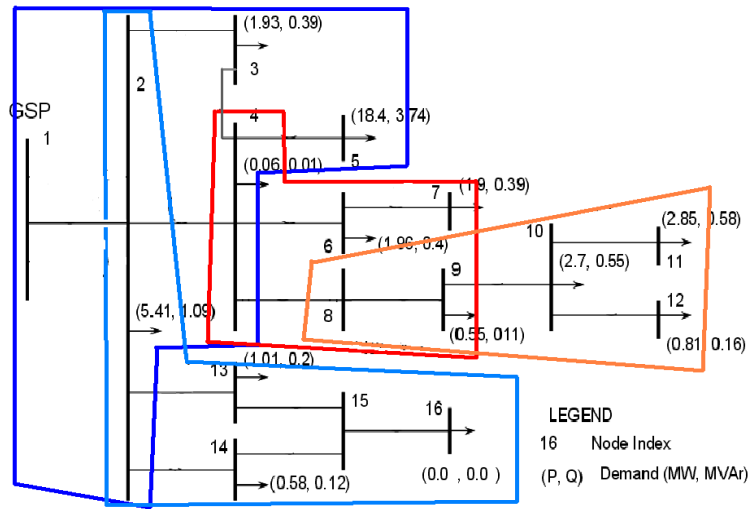
*411 Node EG Network Data (with node renumbering and actual node code)*

Node ID	Node Index	Node ID	Node Index	Node ID	Node Index	Node ID	Node Index
8000065	1	6000640	104	6027676	207	6030609	310
8000335	2	6000753	105	6030411	208	6000111	311
8000334	3	6000208	106	6027688	209	6029258	312
6000619	4	6000256	107	6000650	210	6051779	313
6000616	5	6000122	108	6000664	211	6000174	314
6000688	6	6000454	109	6000225	212	6030713	315
6000514	7	6063260	110	6000580	213	6000760	316
6068191	8	6000244	111	6000144	214	6029850	317
6000433	9	8000653	112	6030686	215	6029875	318
6132298	10	6000332	113	6068146	216	6000756	319
6105761	11	6028973	114	6000335	217	6000086	320
6000563	12	8000656	115	6031708	218	6000477	321
6000415	13	6000716	116	6030079	219	6029259	322
6059841	14	6000127	117	6000759	220	6000796	323
6000257	15	6051933	118	6029695	221	6000440	324
6031369	16	6067216	119	6029785	222	6030724	325
6030250	17	8000310	120	6062641	223	6000491	326
6000694	18	6029713	121	6029188	224	6027681	327
6000359	19	6027665	122	6000622	225	6000275	328
6078455	20	6029586	123	6000067	226	6029879	329
6054071	21	6051534	124	6000673	227	6062644	330
6051759	22	6000154	125	6030394	228	6030664	331
6030152	23	6000507	126	6069796	229	6082698	332
6122767	24	6141255	127	6030543	230	6000198	333
6029998	25	8000066	128	6000261	231	6051778	334
6000617	26	8000067	129	6000361	232	6000474	335
6000615	27	6000511	130	6000599	233	6009918	336
6000163	28	6055805	131	6064304	234	6000414	337
6000722	29	6029957	132	6031722	235	6042097	338
6064946	30	6000750	133	6000489	236	6043145	339
6052693	31	6028977	134	6029929	237	6030671	340
8000061	32	8000654	135	6000057	238	6122172	341
6000418	33	6064947	136	6029699	239	6064305	342
6077344	34	6077343	137	6000444	240	6027699	343
6000288	35	6067402	138	6029351	241	6030726	344
61057611	36	6030467	139	6029191	242	8000634	345
6000431	37	6029714	140	6000135	243	6000362	346
6000564	38	6000798	141	6000498	244	6064314	347
6000416	39	6031678	142	6051762	245	6072453	348
6053169	40	6029572	143	6042204	246	6030728	349
6063027	41	6000055	144	6000097	247	6000292	350
6054066	42	6000227	145	6068233	248	6062718	351
6029755	43	6000698	146	6029738	249	6063250	352
6000705	44	6029663	147	6063446	250	6030745	353
6000360	45	6030345	148	6031627	251	6000336	354
6102042	46	6000276	149	6000612	252	6067503	355
6044757	47	6029978	150	6029935	253	6030748	356
6030162	48	6000350	151	6000264	254	6000578	357
6042081	49	6061932	152	6029702	255	6030764	358
6030016	50	6000644	153	6064319	256	6000346	359
6000618	51	6066481	154	6043146	257	6064320	360
6000566	52	6129673	155	6029345	258	6030776	361
6000801	53	6000467	156	6000303	259	6000188	362
6065747	54	6030473	157	6029238	260	6064309	363
6000721	55	6029720	158	6000752	261	6030777	364
6000513	56	6030296	159	6027682	262	6074384	365
6068192	57	6031683	160	6000125	263	6030812	366
6000213	58	6000505	161	6000279	264	6000459	367
6000540	59	6030051	162	6029746	265	6030842	368
6000565	60	6000548	163	6063445	266	6000092	369
6000417	61	6067689	164	6000241	267	6030870	370
6082953	62	6029672	165	6000242	268	6000426	371
6029623	63	6030442	166	6000056	269	6027657	372
6000045	64	6030350	167	6059652	270	6000079	373

6000559	65	6000780	168	6063249	271	6000598	374
8000657	66	6000411	169	6009920	272	6030880	375
6000214	67	6004768	170	6000439	273	6030900	376
6000026	68	6062646	171	8000311	274	6000160	377
6063336	69	6064393	172	6062284	275	6000483	378
6028388	70	6068700	173	6000621	276	6000562	379
8000329	71	6064321	174	6000302	277	6030991	380
6087627	72	6030492	175	6029239	278	6031009	381
6068701	73	6000755	176	6000633	279	6000452	382
6041361	74	6000258	177	6000751	280	6031030	383
6029600	75	6000047	178	6000591	281	6031218	384
6051533	76	6031693	179	6065751	282	6031031	385
6000545	77	6000530	180	6029756	283	6000100	386
6000069	78	6030060	181	6062715	284	6000538	387
6000512	79	6078032	182	6030357	285	6000726	388
6077345	80	6076533	183	6029967	286	6064262	389
6000535	81	6029682	184	6064391	287	6031262	390
6000308	82	6000166	185	6029854	288	6031053	391
6000330	83	6000050	186	6030573	289	6000684	392
6000075	84	6000184	187	6000072	290	6000725	393
6000774	85	6030422	188	6010793	291	6031058	394
6029629	86	6059518	189	6000441	292	6000625	395
6000291	87	6000352	190	6029758	293	6031059	396
6000410	88	6004784	191	6000579	294	6000294	397
8000655	89	6053507	192	6041527	295	6031062	398
6000051	90	8000103	193	6029976	296	6031067	399
6028969	91	6000518	194	6064392	297	6000689	400
8000658	92	6000473	195	6000109	298	6031068	401
6000701	93	6000035	196	6000785	299	6000099	402
6000656	94	6027654	197	6029254	300	6031086	403
6028380	95	6029728	198	6000337	301	6000791	404
6100086	96	6000334	199	6000718	302	6031092	405
6064411	97	6000046	200	6000049	303	6000098	406
6084762	98	6000044	201	6062725	304	6027671	407
6029709	99	6030074	202	6041902	305	6031100	408
6029596	100	6000365	203	6000553	306	6000453	409
6044250	101	6029684	204	6029870	307	6000209	410
6000290	102	6000327	205	6029836	308	6141261	411
6000191	103	6000087	206	6000799	309		

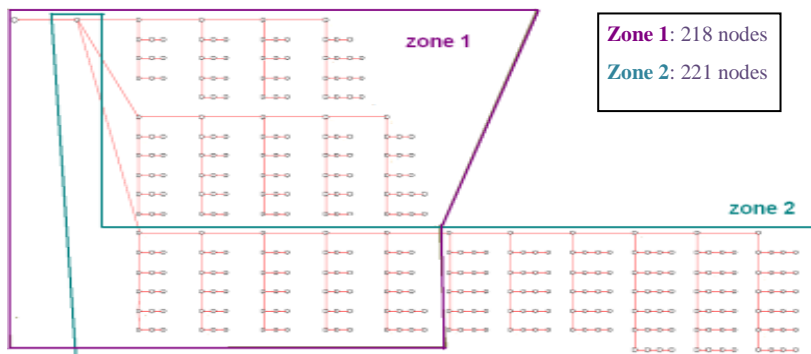
## 9. Network Division

### 16 Node Network

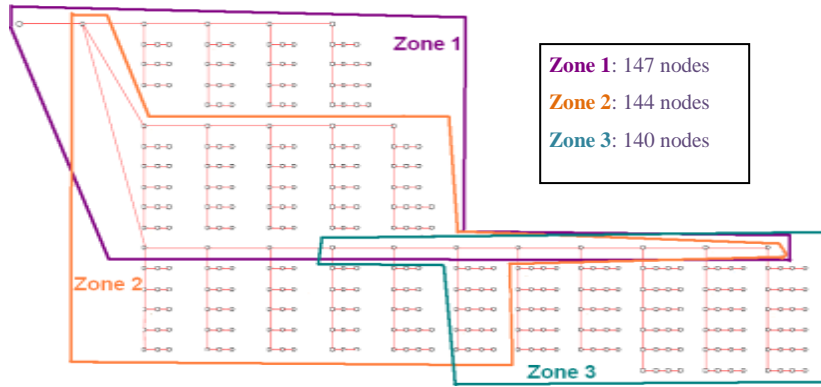


### 356 Node Network

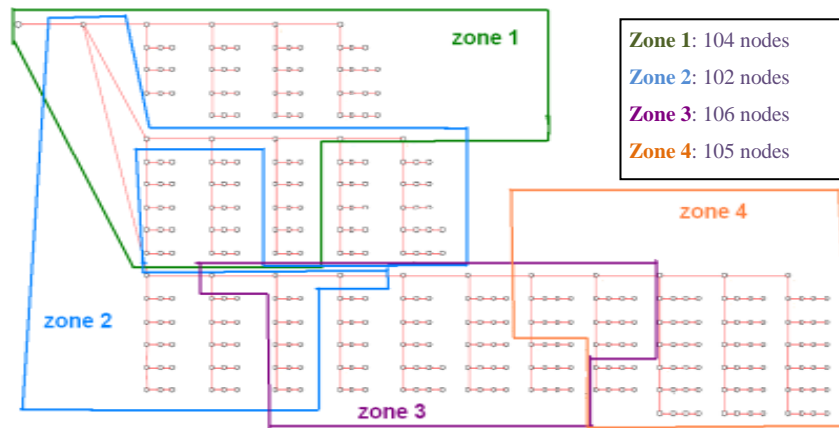
Figures below show the network division using different colour codes and size of each zone in terms of number of nodes. The size of overlapping nodes varies from 16-100 depending on the zone numbers and structure.



### 2 zone splitting



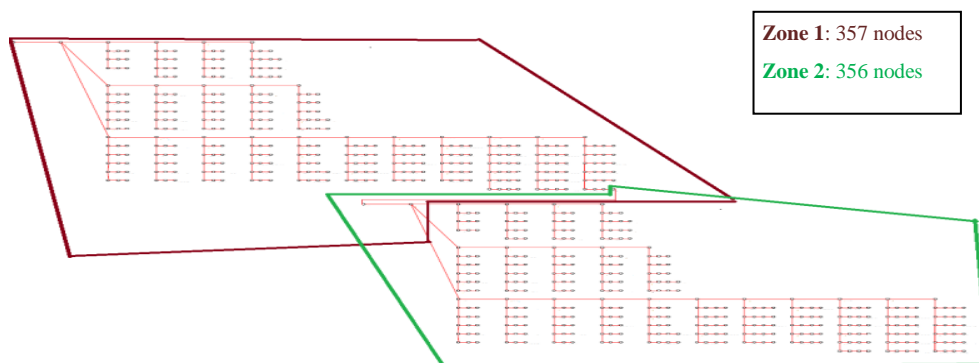
### 3 zone splitting



### 4 zone splitting

#### **711 Node Network**

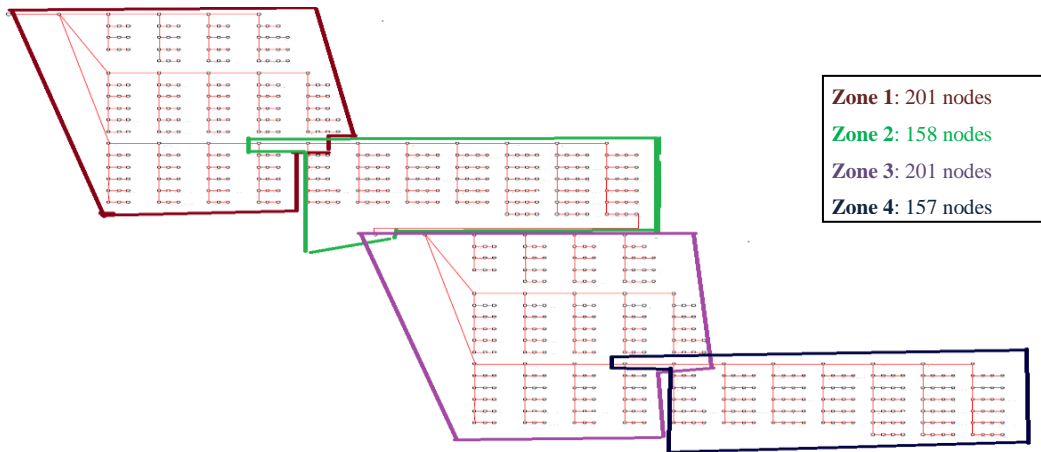
Figures below show the network division using different colour codes and size of each zone in terms of number of nodes. The size of overlapping nodes varies from 2-3 depending on the zone numbers and structure.



### 2 zone splitting



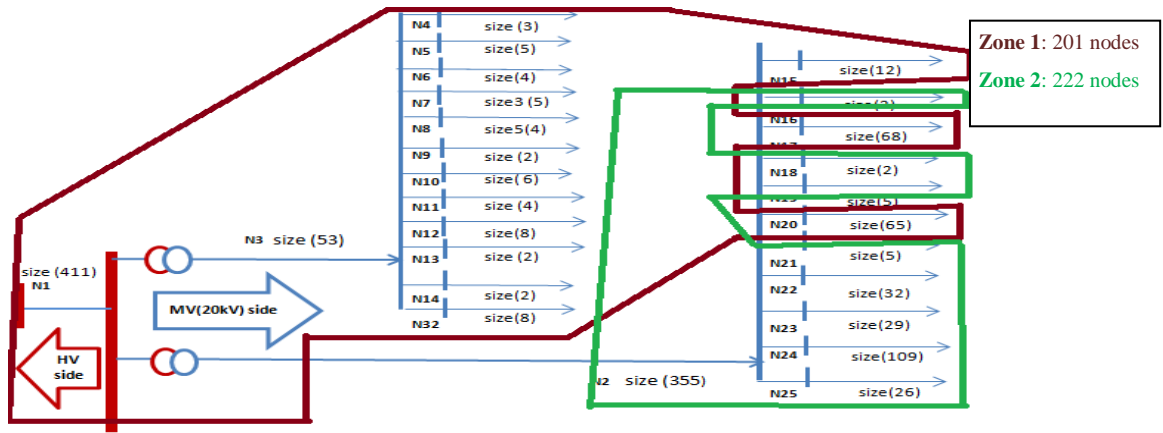
3 zone splitting



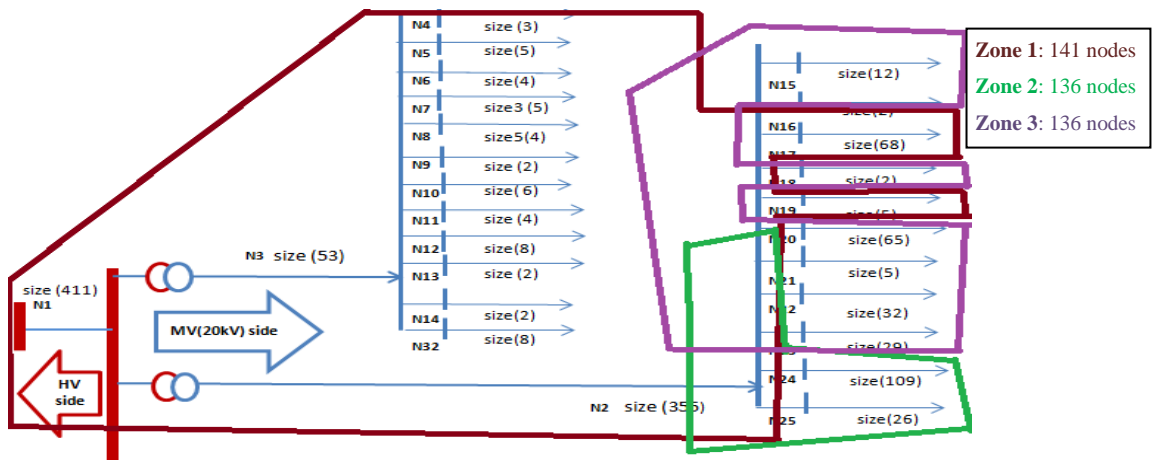
4 zone splitting

***411 Node Network***

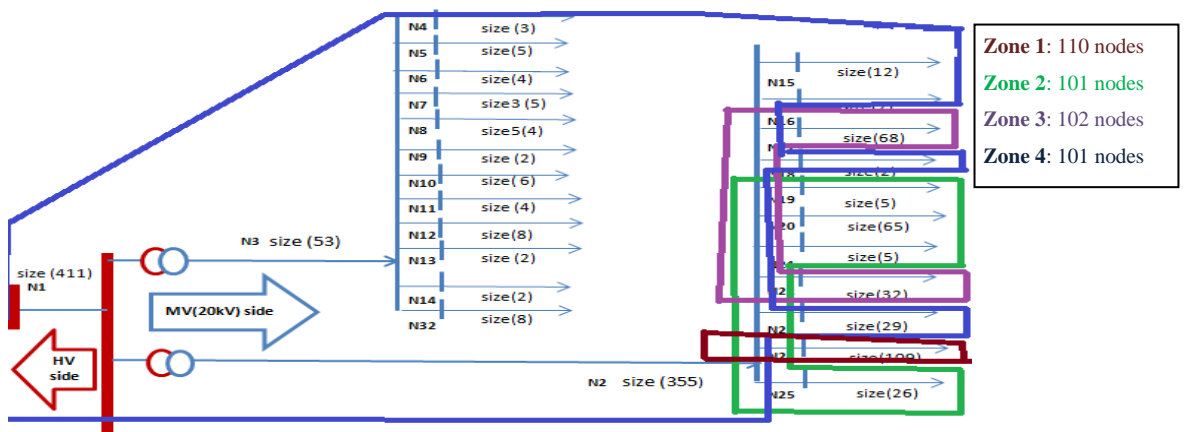
Figures below show the network division using different colours and size of each zone in terms of number of nodes. The size of overlapping nodes varies from 1-3 depending on the zone numbers and structure. In the picture below, only the MV feeder start locations are shown. Each zone contains one or more complete feeders in this case. The size of the feeder information i.e. number of nodes in each feeder is given next to each feeder.



2 zone splitting



3 zone splitting



4 zone splitting

## 10.Pseudo-Codes of OZA MATLAB Tool

```
local_converge=zeros(1:number of zones);
i=1; j=1; k=1; m=1;

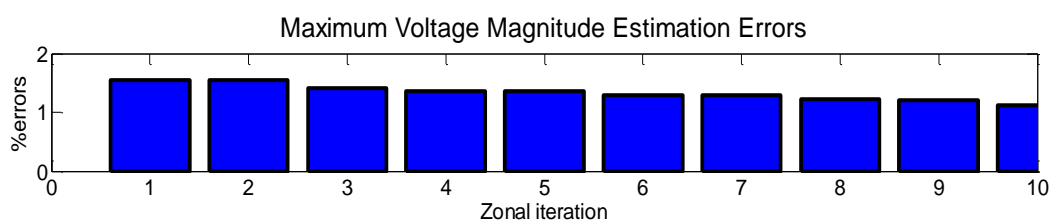
for ( i:1:number of zones)
    Local estimation (i);
end for

for ( j:1:number of zones)
    for (k:1:number of zones)
        if zone(j)  $\cap$  zone (k)  $\neq$  0
            data transfer (j,k);
        end if
    end for
end for

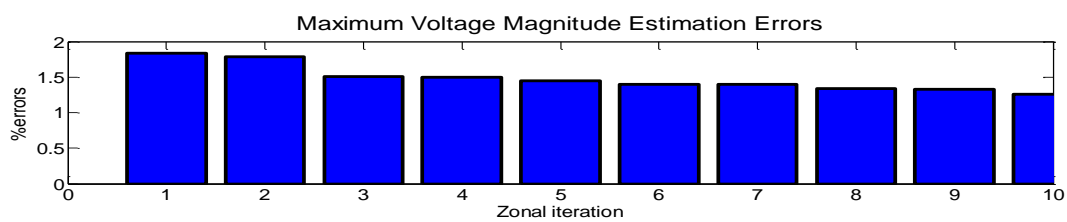
for ( m:1:number of zones)
    data update (m);
    local_converge(m)= max(abs(local mismatch));
end for

if max( local_converge(1:number of zones))<=1e-6 || maximum zonal iteration
    save results
    stop
else
    continue
end if
```

## 11. Voltage Estimation Errors for 2 Zone Division of the 356 Node Network with Measurements Only at GSP



Local maximum voltage estimation errors for zone 1



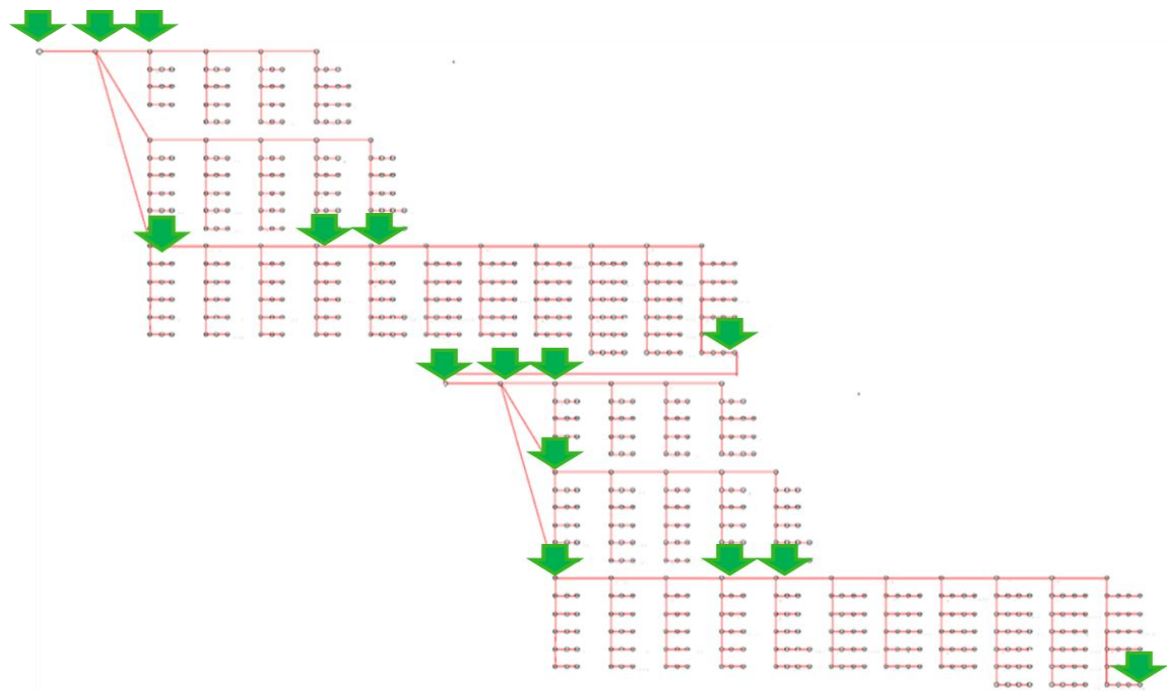
Local maximum voltage estimation errors for zone 2



## 12. Location of Increased Measurements in OZA

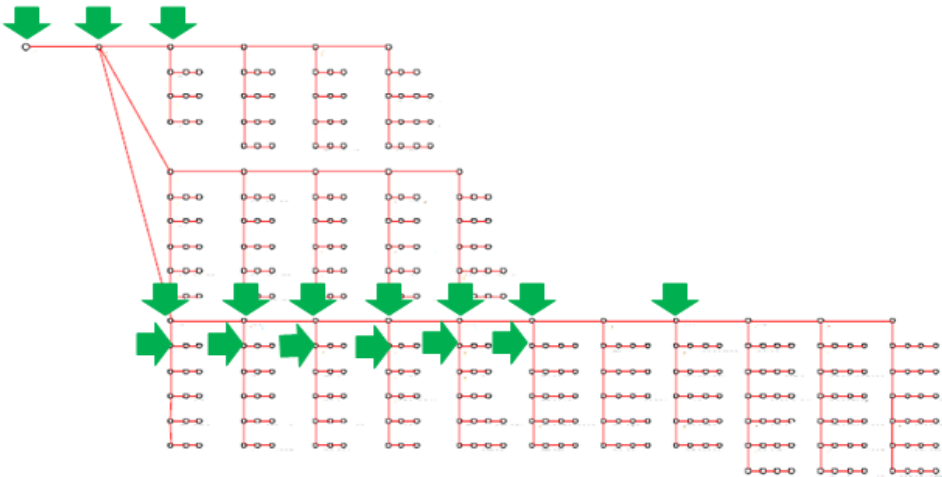
The green arrows in the networks below represent the location of voltage and power injection measurements. The measurement locations are shown on the network diagram. For the case studies of the 356 and 711 node networks, the location of meters for various zone divisions is stated in the tables. The location of meters along with the node index are visible for the case studies (in chapter 7) on the 411 node network, therefore the relevant table is not included. The 411 node network retains this measurement configuration for all relevant case studies. However, the 711 node network applies this sensor configuration only for the case of 5 zonal interactions in chapter 7.

### 711 Node Network Sensor Locations



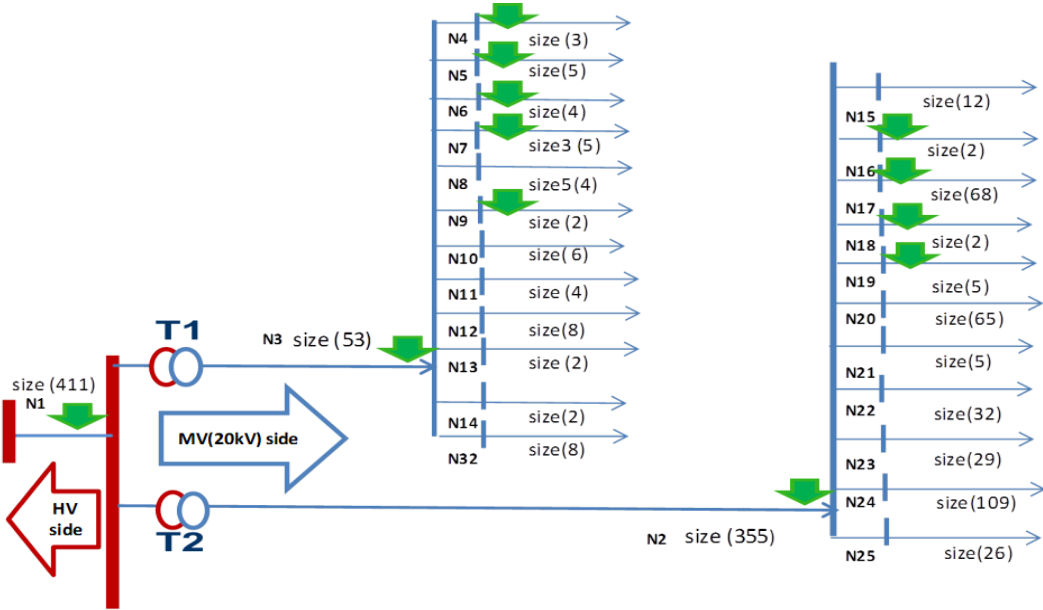
Zone division	No of node positions	Node positions
2 zone	9	1, 2, 355, 356, 357, 138, 358, 410, 492
3 zone	19	1, 2, 137, 140, 141, 142, 143, 168, 173, 356, 357, 491, 358, 359, 362, 410, 411, 417, 492
4 zone	19	1, 2, 139, 140, 141, 142, 163, 168, 354, 355, 356, 357, 494, 495, 496, 497,

**356 Node Network Sensor Locations**



Zone division	No of node positions	Node positions
2 zone	5	1, 138, 141, 158,163
3 zone	7	1, 2, 140, 143, 148, 168, 173
4 zone	10	1, 2, 3, 55, 137, 139, 141, 144, 153, 173

**411 Node Network Sensor Locations**



### 13. Effect of Expected Maximum Power Flow Measurement Errors in the Added Meter Data

In figures below, 'random errors' present the values when normal randomized errors are added. 'maximum expected error at new node (1) and (2)' represent the cases when the measurement is fixed at  $(1.10 \times \text{true value})$  and  $(0.99 \times \text{true value})$  respectively.

