STATIC STATE ESTIMATION AND LOAD FORECASTING VIA ADVANCED COMPUTATIONAL TECHNIQUES

by

Hossam Mosbah

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy

at

Dalhousie University
Halifax, Nova Scotia
June 2019

This thesis is dedicated to my parents. For their endless love, support, and encouragement

Contents

List of Tables . . . . . **viii**

List of Figures . . . . . **xi**

Abstract . . . . . **xvi**

Acknowledgements . . . . . **xvii**

Chapter 1 Introduction **1**

1.1 Motivation 1

1.2 Thesis Objective 2

1.3 Thesis Contribution 3

1.4 Thesis Outline 3

Chapter 2 Literature Review **6**

2.1 Introduction 6

2.2 State Estimation 6

2.3 The operating condition of a power system 10

2.4 Power System Security Analysis 12

Chapter 3 Hourly Electricity Price Forecasting For The next Month Using Multilayer Neural Network **14**

3.1 Introduction 14

3.2 System Description 17

3.3 Simulation Results and Discussion 19

3.4 Conclusion 32

Chapter 4 Optimization of Neural Network Parameters by Stochastic Fractal Search for Dynamic State Estimation under Communication Failure 34

4.1 Introduction 34

4.2 Stochastic Fractal Search Technique (SFS) 38

4.3 Overview of DSE under Communication Failure 38

4.4 Theoretical background 39
4.4.1 Backpropagation Neural Network (BPN) 39
4.4.2 Overview of BPN related works 39

4.5 Proposed method and problem description of MLP via SFS under DSE 40
4.5.1 Short Term Load Forecasting in Time Series using Support Vector Regression (STLF-SVR) 40
4.5.2 NYISO Load Normalization 42
4.5.3 Optimal power flow (OPF) 42
4.5.4 White Gaussian Noise (WGN) 43

4.6 Implementation of MLP based on SFS 43
4.6.1 Initialization 44
4.6.2 Encoding Strategy 46
4.6.3 The MLP as a fitness function of SFS 47
4.6.4 Diffusion Process 49
4.6.5 Ranking 49
4.6.6 Updating Process 49
4.6.7 Training Termination of the MLP Based SFS process 50

4.7 Results and discussion 50
4.7.1 No Communication Failure 56
4.7.2 Sustained communication failure 61
4.7.3 Communication failure Duration 66
4.7.4 Multi-Step ahead forecasting 67
4.7.5 Multi-station failure 69
4.7.6 Failure at different time of the day 71

4.8 Conclusion 73

Chapter 5 Optimized Neural Network Parameters Using Stochastic Fractal Technique to Compensate Kalman Filter for Power System-Tracking-State Estimation 75

5.1 Introduction 75
5.2 Tracking State Estimation 77
5.3 Kalman Filtering Compensated Via Optimized Neural Network Parameters based Stochastic Fractal Technique 78
5.4 Case study system 83
5.5 Results and discussion 85
5.5.1 Normal operating condition 87
5.5.2 Presence of bad data measurements 88
5.5.3 Presence of Sudden Load change conditions 90
5.5.4 During sudden loss of transmission line 91
5.5.5 During sudden loss of Generations 95
5.6 Conclusion 96

Chapter 6 Conclusion 98

6.1 Conclusion 98
6.2 Scope of Future Work 100

Appendix A Multilayer Artificial Neural Networks for Real Time Power System State Estimation 101

A.1 Introduction 101
A.2 State Estimation 103
A.3 Proposed Neural Network Architecture 105
A.4 Tested System and Results 106
A.5 Conclusion 118

Appendix B Power System Tracking State Estimation Based on Stochastic Fractal Search Technique under Bad Measurements Conditions 119

B.1 Introduction 119
B.2 Tracking State Estimation 120
B.3 Solution Methodology of SFS 121
B.4 Simulation Study 122

B.4.1 WLAV . . . 123
B.4.2 Constraints . . . 123
B.4.3 Bad Measurements Condition . . . 123
B.4.4 Test Systems . . . 123
B.4.5 Performance Evaluations . . . 124

B.5 Results and Discussion . . . 124

B.6 Conclusion . . . 130

Appendix C Power System Tracking State Estimation Based on Stochastic Fractal Search Technique under Sudden Load Changing Conditions . . . 131

C.1 Introduction . . . 131

C.2 Sudden Load Changes . . . 132

C.3 Solution Methodology of SFS algorithm . . . 133

C.4 Simulation Study . . . 134
C.4.1 Sudden Load Change . . . 134
C.4.2 Test System . . . 136
C.4.3 Performance Indices . . . 136

C.5 Results and Discussion . . . 136

C.6 Conclusion . . . 142

Appendix D Power System Static State Estimation using Modified Stochastic Fractal Search Technique . . . 143

D.1 Introduction . . . 143

D.2 Stochastic Fractal Search Technique . . . 144

D.3 Modified Stochastic Fractal Search Technique . . . 146
D.3.1 Modified α Parameter . . . 146
D.3.2 Implementation of Chaotic Maps . . . 146

D.4 Simulation Study . . . 148

D.5 Results and discussions . . . 148

D.6 Conclusion . . . 151

Appendix E Evaluating the Impact of Phasor Measurement Units on the Accuracy of State Estimation 152

E.1 Introduction 152

E.2 Hybrid state estimator formulation 153

E.3 Hybrid techniques 155

E.4 Simulation Study 156

E.5 Results and discussions 158

E.6 Conclusion 161

Appendix F A Distributed Multiarea State Estimation 162

F.1 Introduction 162

F.2 Multiarea State Estimation 163

F.2.1 Network Partitioning 163

F.2.2 Local SE Formulation 164

F.2.3 System-Wide SE Formulation 164

F.3 Proposed and implemented Simulation System 165

F.3.1 Stochastic Fractal Search (SFS) 165

F.3.2 Simulated Annealing (SA) 166

F.3.3 Implemented Simulation Systems 166

F.4 Case Study System 167

F.5 Results and discussions 168

F.6 Conclusion 172

List of Tables

3.1 Weights For Three Interconnected Nets 27

3.2 Final Comparision In Cascade and Parallel Hybrid Connections in Different Performance 28

3.3 Comparison Of Performance in Multilayer Topologies For Actual and Predicted Electricity Price For January 2006 29

3.4 Ranking Training Algorithms Of Feed Forward Neural Network With 10 Hidden Layers in Hybrid Parallel Topology in Case Of Accuracy . 30

3.5 Ranking Training Algorithms Of Feed Forward Neural Network With 10 Hidden Layers in Hybrid Cascade Topology in Case Of Accuracy . 31

4.1 Control Parameters used in Optimized MLP based SFS for IEEE 14-bus system combined with NYISO load data 52

4.2 Control Parameters used in Optimized MLP based SFS for IEEE 118-bus system combined with NYISO load data 53

4.3 Selecting the best control parameters used in Optimized MLP based SFS for IEEE 14 Bus System combined with NYISO load data in case of accuracy and CPU Time 54

4.4 Selecting the best control parameters used in Optimized MLP based SFS for IEEE 118 Bus System combined with NYISO load data in case of accuracy and CPU Time 54

4.5 Non-optimized MLP state estimation error for IEEE 14-bus system at different measurement error under no communication failure 58

4.6 Optimized MLP-SFS state estimation error for IEEE 14-bus system at different measurement error under no communication failure 59

4.7 Optimal Power flow for IEEE 14-bus system combined with NYISO load data under no communication failure in six hours ahead forecasting 61

4.8 Computational time comparison at measurement error k=5 under no communication failure 61

4.9 Mean Absolute Percentage error for IEEE 14-bus system combined with NYISO loads in sustained communication failure in six hours ahead forecasting 64

4.10 Mean phase and magnitude voltage percentage error at failure moment in bus 14, West region for IEEE 14-bus system combined with NYISO load data 72

5.1 Ranking Training Algorithm of Feed Forward Neural Network Using compensated KF-MLP Based SFS for IEEE 57 bus system in case of accuracy and computational time 86

5.2 Selecting the best number of neurons Using compensated KF-MLP Based SFS in case of accuracy and computational time 86

5.3 Control Parameters used in Compensated KF-MLP based SFS for IEEE 57-bus system 87

5.4 Comparison of proposed KF-MLP Based SFS technique with two other conventional approaches in Filter Effect for IEEE 57 bus system 97

A.3 Comparison Parallel and Cascade Hybrid Connection in Regression, Time, and Neurons 113

A.1 Comparison Parallel and Cascade Hybrid Connection in Voltage Angle Vector 114

A.2 Comparison Parallel and Cascade Hybrid Connection in Voltage Magnitude Vector 114

A.4 Weights For different Topologies 115

A.5 Ranking Training Algorithms Of Feed Forward Neural Network With 10 Hidden Layers In Hybrid Cascade and Parallel Topologies In IEEE 118 Bus System 116

B.1 Control Parameters settings 125

B.2 Computational time(s) taken by proposed SFS technique 125

B.3 Comparison Of State Vector, Estimation, and True Values On Bus 3 In IEEE 5-Bus System For 24 Instant Of Time 129

C.1 Control Parameters settings 136

C.2 Computational time(s) taken by proposed SFS technique 137

C.3 Comparison Of State Vector, Estimation, and True Values On Bus 3 In IEEE 5-Bus System For 24 Instant Of Time 138

C.4 Statistical Parameter Values Under Sudden Load Condition For Voltage Angle . . . 139
C.5 Statistical Parameter Values Under Sudden Load Condition For Voltage Magnitude . . . 139

D.1 The Presented Replacement Functions For α . . . 147
D.2 The Mathematical Description Of Chaotic Maps μ . . . 147
D.3 Results of modified SFS by α parameter . . . 149
D.4 Results of modified SFS by Chaotic maps . . . 149
D.5 MAPE of Voltage angle for modified SFS . . . 149
D.6 MAPE of Voltage magnitude for modified SFS . . . 150
D.7 Computational time for modified SFS . . . 150

E.1 Control Parameters used in hybrid MLP-SFS . . . 158
E.2 Six Different Cases by adding Phasor Measurement units (PMUs) for IEEE 14-bus system . . . 159
E.3 Six Different Cases by adding Phasor Measurement units (PMUs) for IEEE 30-bus system . . . 159
E.4 Six Different Cases by adding Phasor Measurement units (PMUs) for IEEE 57-bus system . . . 159
E.5 MAPE of Voltage magnitude for only PMUs . . . 160
E.6 MAPE of Voltage angle for only PMUs . . . 161

F.1 Bus Partitioning for the IEEE 118-bus system . . . 169
F.2 Measurement Configuration for each area . . . 169
F.3 Measurement Configuration for the overall system . . . 170
F.4 Speed comparison between different techniques . . . 170

List of Figures

2.1 Power System Operating States 11

2.2 Function diagram of EMS 13

3.1 Cascade Topology 17

3.2 Parallel Topology 18

3.3 Parallel- Cascade in Cascade Connection 19

3.4 Cascade – Parallel in cascade Connection 19

3.5 Cascade-Parallel in Parallel Connections 20

3.6 Parallel-Cascade in Parallel Connections 20

3.7 Hybrid Parallel Topology 21

3.8 Hybrid Cascade Topology 22

3.9 Hybrid topology in Parallel connection, actual and predicted hourly electricity price for Jan. 2006 23

3.10 Hybrid topology in Cascade connection, actual and predicted hourly electricity price for Jan. 2006 23

3.11 Hybrid Parallel topology, Absolute Percentage error for hourly electricity price for Jan. 2006 24

3.12 Hybrid Cascade topology, Absolute Percentage error for hourly electricity price for Jan. 2006 24

3.13 . Hybrid Parallel topology, Residual between the actual and the predicted hourly price forecasting for Jan.2006 25

3.14 Hybrid Cascade topology, Residual between the actual and the predicted hourly price forecasting for Jan.2006 25

3.15 Hybrid Cascade topology, regression model for the actual and the predicted hourly price forecasting for Jan.2006 26

3.16 Hybrid Parallel topology, regression model for the actual and the predicted hourly price forecasting for Jan.2006 27

4.1 Multilayer Perceptron Neural network model 48

4.2 Flowchart of the hybrid MLP based SFS technique 51

4.3 Ranking Training Algorithm of Feed Forward Neural Network Using MLP-SFS SFS for IEEE 14-bus system in case of accuracy and Computational time (Secs) 55

4.4 Ranking Training Algorithm of Feed Forward Neural Network Using MLP-SFS for IEEE 118-bus system in case of accuracy and Computational time (Secs) 56

4.5 Average Phase Percentage error comparison at different measurement error with no communication for IEEE 14-bus system 57

4.6 Average Voltage Percentage error comparison at different measurement error with no communication for IEEE 14-bus system 58

4.7 Average Phase Percentage error comparison at different measurement error with no communication for IEEE 118-bus system 59

4.8 Average Voltage Percentage error comparison at different measurement error with no communication for IEEE 118-bus system 60

4.9 Average Phase Percentage error comparison at different measurement error with single station communication failure for IEEE 14-bus system. X-axis: measurements error. Y-axis: bus with communication failure. Z-axis mean phase Percentage error comparison 62

4.10 Mean Voltage Percentage error comparison at different measurement error with single station communication failure for IEEE 14-bus system. X-axis: measurements error. Y-axis: bus with communication failure. Z-axis mean voltage percentage error comparison 63

4.11 Mean Phase Percentage error comparison at different measurement error with single station communication failure on bus 59 for IEEE 118-bus system 64

4.12 Mean Voltage Percentage error comparison at different measurement error with single station communication failure on bus 59 for IEEE 118-bus system 65

4.13 Mean Phase Percentage error comparison at different failure duration with single station communication failure for IEEE 14-bus system. X-axis communication failure duration. Y-axis Real Power Injection. Y-axis Mean Phase Percentage error 66

4.14 Mean Voltage Percentage error comparison at different failure duration with single station communication failure for IEEE 14-bus system. X-axis communication failure duration. Y-axis Reactive Power Injection. Y-axis Mean Voltage Percentage error 67

4.15 Mean Phase Percentage error comparison between SVM and RBF when load data from Bus2, Central region of the IEEE 14-bus system is not available for different failure duration 68

4.16 Mean Voltage Percentage error comparison between SVM and RBF when load data from Bus2, Central region of the IEEE 14-bus system is not available for different failure duration 69

4.17 Mean Phase Percentage error comparison when both bus 5 and 6 fail at different measurement error for the IEEE 14-bus system 70

4.18 Mean Voltage Percentage error comparison when both bus 5 and 6 fails at different measurement error for the IEEE 14-bus system . . . 70

5.1 Tracking State Estimation 78

5.2 Structure of KF compensatd by optimized MLP based SFS (KF-MLP based SFS) 79

5.3 Flowchart of Kalman filtering compensated by optimized MLP based SFS 84

5.4 Performance of voltage magnitude at buses 3, 19, 30, and 56 for IEEE 57-bus system under normal operating condition 88

5.5 Mean Absolute Percentage error (MAPE) for estimated states in IEEE 57-bus system under normal operation condition 89

5.6 Performance of estimated states at buses 2 and 49 for IEEE 57-bus system under bad data condition 90

5.7 Mean Absolute Percentage error (MAPE) for estimated states in IEEE 57-bus system under bad data condition 91

5.8 Performance of voltage magnitude at buses 10, 19, and 51 for IEEE 57-bus system under sudden load changing condition 92

5.9 Mean Absolute Percentage error (MAPE) for estimated states in IEEE 57-bus system under sudden load changing condition 93

5.10 Performance of voltage angles at buses 22, 23, 36, and 40 for IEEE 57-bus system under sudden loss of transmission lines condition . . . 93

5.11 Mean Absolute Percentage error (MAPE) for estimated states in IEEE 57-bus system under transmission line loss condition 94

5.12 Performance of voltage magnitude at buses 8 and 12, Generator 5 and 7 for IEEE 57-bus system under sudden loss of Generations condition 95

5.13 Mean Absolute Percentage error (MAPE) for estimated states in IEEE 57-bus system under Generators loss condition 96

A.1 Architecture of Multilayer Perceptron 105

A.2 Hybrid Parallel and Cascade topologies, Estimated and Actual voltage angle in IEEE 14 bus system 107

A.3 Hybrid Parallel and Cascade topologies, Estimated and Actual magnitude voltage in IEEE 14 bus system 108

A.4 Hybrid Parallel and Cascade topologies, voltage angle residual in IEEE 14 bus system 108

A.5 Hybrid Parallel and Cascade topologies, magnitude voltage residual in IEEE 14 bus system 109

A.6 Hybrid Parallel and Cascade topologies, Relative voltage angle Error in IEEE 14 bus system 109

A.7 Hybrid Parallel and Cascade topologies, Relative magnitude voltage Error in IEEE 14 bus system 110

A.8 Hybrid Parallel and Cascade topologies, Estimated and Actual voltage angle in IEEE 118 bus system 110

A.9 Hybrid Parallel and Cascade topologies, Estimated and Actual magnitude voltage in IEEE 118 bus system 111

A.10 Hybrid Parallel and Cascade topologies, voltage angle residual in IEEE 118 bus system 111

A.11 Hybrid Parallel and Cascade topologies, Magnitude voltage Residual in IEEE 118 bus system 112

A.12 Hybrid Parallel and Cascade topologies, Relative voltage angle Error in IEEE 118 bus system 112

A.13 Hybrid Parallel and Cascade topologies, Relative magnitude voltage Error in IEEE 118 bus system 113

B.1 Statistical parameter values under bad data condition for voltage angle 126

B.2 Statistical parameter values under bad data condition for magnitude voltage 126

B.3 Performance Index J (k) in the presence of normal and bad data condition 127

B.4 Normalized Measurements vector for bad data condition during 11th time instant 128

C.1 Performance Index J(k)in the presence of normal and sudden load condition 141

C.2 Normalized Measurements vector for sudden load change condition during 11th time instant 141

D.1 Mean absolute Percentage error of Voltage angle 150

D.2 Mean absolute Percentage error of Voltage Magnitude 151

F.1 Multiarea Partitioning of the IEEE 118-bus system 168

F.2 Sum of absolute percentage error (objective function) for cases 1 & 2 in local areas and case 1 in coordinator level 171

F.3 Sum of absolute percentage error (objective function) for cases 2 & 3 in coordinator level and case 3 in local areas 171

Abstract

The system operator faced high level of uncertainties due to the evolving complexity of the electric power systems. Therefore, efficient solutions for power system prediction, monitoring and state estimation can be found in new methods. These methods improve the secure operations of the network.

This thesis covers three different state estimations incorporating static, tracking, and dynamic state estimation in order to estimate all possible operating conditions of power systems.

This research also introduces new meta heuristic methods such as stochastic and fractal search technique. The SFS technique is implemented in real time nonlinear power system applications under various scenarios. New methodology of multilayer neural networks exhibited in composite typologies are proposed in this thesis to improve the estimation performance. Optimized Neural Network by Stochastic Fractal Search technique is used and applied to both tracking and dynamic state estimation. The proposed methods are validated utilizing diverse benchmark optimization methods.

The combination of conventional and synchronized measurement is also studied in this thesis. This is used to increase the reliability of electric power systems in real-time.

Additionally, the research is extended to evaluate the benefits of multiarea state estimators and how it uses to reduce the computational time.

Finally, all formulations proposed in this work were validated in different IEEE test systems.

Acknowledgements

I take this opportunity to express my thankfulness to those who made this thesis a success.

Special acknowledgement to my family for their support and patience during my study.

I would like to express my gratitude to my advisor Dr. M. E. El-Hawary for his guidance and assistance throughout the thesis work

I also wish to express a sincere gratefulness to the Ministry of Education in my country, Libya, the staff of the Libyan Embassy in Canada and the Canadian Bureau for International Education, CBIE, for their mentor and financial support.

Chapter 1

Introduction

1.1 Motivation

Reliable operation of today's large electric power systems heavily depend upon the state estimation techniques for different applications. These estimators give important measurement used for different application functions of control center. Therefore, a computationally efficient and robust state estimator is an important tool to improve the performance, security and safety of the power delivery system.

State estimation (SE) is becoming one of the key functions in the distribution control centers in the deregulated and competitive environment [1]. The state estimation of electric power system is an imperative part of modern energy management system (EMS).

The major goal of the research will be proposing new robust techniques include meta-heuristic, artificial intelligent, and decoupling techniques in several SE types such as static, tracking, and dynamic state estimation to prevent the security risks in the modern electric grids. And also to provide us with a reliable power system operation exhibited in less computational time (a few seconds), detection and identification of bad measurements, and robustness.

The state estimation of electric power system utilizes the redundancy of data brought by SCADA to enhance the precision of data, automatically exclude error messages brought on by arbitrary interference, and estimate the running condition of the system. Engineers in the control center utilize the yields of the state estimation process as the base for incident analysis. In the incident analysis, they decide potential operational issues, and provide operational direction to maintain a strategic distance from these issues, and in addition, conceivable reactions brought on by these activities.

State estimation expected to perform in under a second to allow operators the ability to detect, identify and react to any abnormal status.

A general state estimation process used in power delivery systems involves the

following functions: [2]

• Network topology analysis: This is used to determine the real time network structure based on digital measurements.

• Observability analysis: This is utilized to find whether the power system is observable based on the available measurements. If the system is not-observable then pseudo-measurements will be added to make the system observable. Only observable parts of the system can be estimated. If the rank of the measurement Jacobian matrix is equal to the number of state variables, then network is observable [3].

• State estimation computation: This is used to estimate the state of the system according to the usable measurements. The state estimation computation encompasses the bad data detection and identification process.

• Bad data analysis: This is used to detect and identify bad data in the measurement data sets based on the analysis of measurement residuals. When accurate estimate is failed to be determined, it is either due to the bad data, or modeling error, or both. If any bad data are detected, it will be removed from measurement set and state estimation will be repeated.

1.2 Thesis Objective

For the achievement of the objective, the research method involves the following:

1. Combined studies of both state estimation includes static, tracking, and dynamic estimation and three different techniques include neural network, meta-heuristic, and decoupling techniques.

2. Detect and identify bad measurements in distribution power system to eliminate these measurements to increase the reliability of SE.

3. Contingency analysis is one of a challenge part in electrical networks, so we can use the SE to evaluate the type of emergency.

4. Reduce the security risk in electric grid by enhancing PSSE exhibited in robust performance and less computational time.

5. Comparison between these techniques to determine the capabilities.

6. Design a Matlab program to solve the Power System State Estimation problem with different techniques.

7. Modify the existence test system to be used in testing the Matlab program.

8. Test the PSSSE program using different test systems which are used to test the PSSE solution.

1.3 Thesis Contribution

The contributions of this thesis can be stated as follows:
1. Propose and test a new mathematical optimization techniques include meta heuristic and artificial intelligent.
2. Study different state estimations including static, tracking, and dynamic state estimations in order to estimate all possible operating conditions of power system.
3. Propose new methodologies of multilayer neural typologies to improve the estimation performance.
4. Propose different hybrid optimization technique such as optimized neural network parameters by stochastic fractal search technique.Then, apply these techniques to tracking and dynamic state estimations.
5. Asses the advantage of multi area state estimation using new proposed methods.
6. Study the benefit of gradually combing PMU measurement with conventional measurement.
7. Develop the proposed methods to improve the estimation performance.
8. Compare the proposed methods to various benchmark optimization techniques.
9. Validate the proposed methods using different IEEE bus systems.

1.4 Thesis Outline

This thesis is organized to six chapters and six appendices as follows:

Chapter 1 This chapter includes the motivation, the objectives of the thesis, the major contributions and the thesis outline.

Chapter 2 This chapter contains literature review for the four power system state estimation problems under consideration include the state estimation, supervisory control and data acquisition, energy control center, and the contingency analysis

problems.

Chapter 3 This chapter will cover hourly electricity price forecasting using composite multilayer neural network. This includes hybrid parallel and hybrid cascade topologies .

Chapter 4 This chapter solves dynamic state estimation problem under communication failure using hybrid optimization technique. Stochastic Fractal Search is used to optimized multilayer perceptron neural network parameters.

Chapter 5 This chapter is the extension of previous chapter. This chapter applied to the improved hybrid optimization technique to solve the tracking state estimation problem under several contingency conditions.

Chapter 6 This chapter covers the conclusion and future work.

Appendix A This appendix applies the main contribution of chapter 3, composite multilayer neural network to solve static state estimation problem.

Appendix B This appendix solves the tracking state estimation problem under bad measurement condition using Stochastic fractal search technique.

Appendix C This appendix solves the tracking state estimation problem under sudden load changing condition using Stochastic fractal search technique.

Appendix D This appendix solves the static estimation problem using modified stochastic fractal search technique.This appendix improves the original stochastic fractal search technique to obtain accuracy along with speed calculation.

Appendix E This appendix evaluates the impact of phasor measurement units on the accuracy of state estimation. This appendix uses hybrid back propagation neural network with stochastic fractal search technique.

Appendix F This appendix presents hybrid stochastic fractal search technique with simulated annealing algorithm to solve multi area state estimation problem.

Chapter 2

Literature Review

2.1 Introduction

This chapter provides an overview of power system operation and the function of state estimators in the energy management system. This includes state estimation, the operating condition of a power system, and power system security analysis.

2.2 State Estimation

In dependable operation of the power system, a robustness, less computational time, and accurate state estimation is a fundamental for most energy administration system implementations in control centers [4].

In general, a state estimation utilizes for characterizing a reliable assessment of the power state vectors incorporate bus voltage magnitudes and angles from system parameters, basic information, and an arrangement of real time available data including net active and reactive power injections and flows. Bad measurements are recognized and cleansing with a specific end goal to state estimator gives a dependable information base for power system operation and control. The estimation framework chooses how well the state estimator performs this limit[5]. An iterative technique for computing the state estimate and concepts of detecting and identifying the errors of the model is discussed in [6]. The key thoughts of essential a static-state estimator is presented and arranged into three sections, part I, clarifying the scientific issues and general calculations for state estimation, In part II, examining several approximate scientific models and it is resulted simplification in estimation in case of detection and identification, and in part III, diverse execution issues associated with dimensionality, PC rate and limit, and the time changing nature of real power systems are discussed.

Linear programming (LP) is presented in [7] to minimize the weighted sum of the absolute values of the measurement residuals. The sequence of linear programing

strategy first recognizes the bad measurements using the measurement residuals of those data rejected by the LP estimator. At that point, the bad data is identified and eliminated by assessing the data errors of the zero residual measurements. The execution of the proposed procedure is attempted and the results are shown, using AEP's 14, 30, 57 and 118 bus power systems.

Another computational strategy is proposed in [8] for solving equality constraint in PSSE without the Lagrange multipliers. It is an imperative to process a complete fractional factorization (L, U) of grid which improve numerical modeling. Further examination is required for the treatment of basic estimations. Critical measurements have zero residuals and, in this case, these measurements are proportional to equality constraints. The technique was shown to perform viably on a couple of viable measured examples. Standardized residuals can be prepared proficiently with the new technique.

A new methodology of artificial neural networks topology based on topology processing and static state estimation is proposed in [9]. Two ANN based on models include the counter-propagation network (CPN) and functional link network (FLN) compared in order to determine topology transforming and static state estimation on several IEEE test systems include 14, 19, 57 useful Indian system. Moreover, the Hopfield neural network and the conventional fast decoupled state estimator (FDSE) have been determined and compared. It is concluded that the ANN based on models has the least CPU time and much faster, and function accurately notwithstanding for non- Gaussian noise.

Ali Abur and Antonio Goomez Expoosito [10] proposed a technique in 1997 which investigates the problem of bad data identification for power systems containing line current magnitude measurements. Two commonly used state estimation methods (Least Squares and Least Absolute Value) and the associated bad data identification techniques are studied. The definitions for the critical and redundant measurements are modified in order to account for the non-uniquely observable systems where line current measurements play a critical role in the network observability determination. A post estimation procedure that checks the unique redundancy of the candidate measurements to be eliminated is proposed. Numerical examples are included to illustrate the proposed procedure.

In 1998, Hua Wei H. et al [11] proposed, a new Interior point algorithm to solve the power system weighted non-linear state estimation problem. The effectiveness of the proposed algorithm has been verified by the extensive simulating computations.

In [12], the authors introduced another system focuses on the neural network prepared by prop algorithm, in which topological perceptibility is considered. The sufficiency of preparing of the system is tried by feeding some novel information designs that were excluded in the preparation set and the yield results acquired were accepted by looking at the results from alternate techniques.

In [13], Hopfield neural network (HNN) and Parallel Genetic algorithm (PGA) applied in neural network in term of the best robust static state estimation method on 5 bus test system. It is concluded that Hopfield method has a long training time and Parallel Genetic method provides the optimal converge which is not rely on initial values.

In 2004, Mai-Hoa Vuong et al [14] proposed a method for determining redundant and pseudo-measurement locations to enhance reliability of detection and identification of topology errors. Its effectiveness is the low number of redundant measurements needed and the systematic positioning of pseudo- and redundant measurements. Directions for future work include noisy data containing bad values, and system parameter estimation. Also, the results obtained were sufficiently encouraging to incorporate the method in the modeling of real-time external equivalents network for static security analysis.

In [15], the authors presented another system focused around the artificial neural network method in order for deciding the recognizability of the power systems proposed a strategy on prop algorithm. Notwithstanding the standard back-propagation algorithm additionally presented.

In 2006, Bei Gou [16] designed numerical algorithms to determine observable islands and to decide pseudo-measurements to merge all islands of a measured network. It is shown that transforming the measurement Jacobian matrix by Gaussian elimination is possible to obtain all the necessary information for observability analysis. This constitutes the main contribution of this paper. Results of test examples are provided to illustrate the proposed algorithms. Implementation of the algorithms is easy since Gaussian elimination is the only computation required.

In 2009, Antonio Gómez-Expósito [17] proposed, a two-level hierarchical estimation methodology, arising from a factorization of the conventional WLS nonlinear model. In the first step, a majority of raw measurements are processed by a linear State Estimation. The results of this step comprising both state variables and associated covariance are then passed to a conventional nonlinear State Estimation.

In 2009, M. R. Irving [18] proposed a robust generalized state estimator method. The use of binary decision variables throughout allows all three types of gross errors to be judged on a comparable basis. Simultaneous gross errors in measurements, parameters, and topology can be identified and corrected. The computational effort of solving the required mathematical program probably limits the practical application of the method to relatively small networks, or sections of the network.

In 2010, Eduardo Car et al [19] proposed a procedure to identify the on/off statuses of breakers at substations throughout a power network. It is intended as a structural data preprocessor prior to running a state estimator. The procedure relies on a dc model of power lines and bus-bars, and requires the solution of a well-behaved mixed integer quadratic programming problem.

George N. Korres [20] proposed a robust method in the year 2010 for the solution of the power system state estimation with equality constraints. In this method the formulated coefficient matrix has a unique triangular factorization which can be accomplished symbolically using only the sparsity criterion.

In 2011, Tao Yang et al [21] proposed a decentralized two-level linear state estimator based on the phasor measurements. The main contribution in this paper is that this two-level processing removes the bad data and topology errors, which are major problems today, at the substation level. In Part I of the paper, the layered architecture of databases are discussed, communications, and the application programs that are required to support this two level linear state estimator. In Part II, the mathematical algorithms that are different from those in the existing literature are explained.

In 2011, Farrokh Aminifar et al, [22] presented a model for the PMU placement that considers observability requirements according to a set of probabilistic criteria. The PMU placement incorporates the stochastic nature of components and their outage probabilities. The nonlinear expression associated with the probability of

observability index is converted to a linear representation by an effective linearization technique. The proposed mathematical model is therefore compatible with the MIP solution tools. The approach considers the PMU placement in a staged time span.

In 2011, H. T. Kung et al [23] proposed a new method called separation based method, for detecting and also identifying bad measurements in compressive sensing. If errors introduced by bad measurements are small, then the separation-based method requires a smaller number of good measurements than the previous method based on joint l1- minimization. In this method it has been told that if there are sufficient good measurements, the joint l1- minimization works independently of error magnitudes. That is, if the joint method fails due to an insufficient amount of good measurements, then they proposed the separation-based method, as it can survive with a smaller number of good measurements.

Gavrilas [24] proposed a multilayer perceptron for the static state estimator and employed on IEEE 14 bus test system. It is proven that ANN estimator can provide less CPU time compared to conventional estimation algorithms in case of bus voltage varies dues to load conditions and there is no change on operating system. Also the results of IEEE 14 bus system have a precise estimation.

2.3 The operating condition of a power system

As the operating states change, the system undergoes into one of three potential states include normal, emergency and restorative states. A power system is considered to operate in a normal condition if both loads and operating constraints are not violated. In other words, the demands of all customers are met by the operation of both generation and transmission system. Operational constraints include the limits on bus voltage magnitudes, standard frequency and the limits on the transmission line flows. When small relative changes in loads occurred, the system is assumed to operate in a quasi-steady-state condition. This leads the system to move from one normal condition to another. Moreover, the security of normal state is based on the ability of the system to remain in a normal condition even after the occurrence of a severe disturbance [25], [26]. Three common emergencies are studied. The steady-state emergency, the system remains stable with secure operating constraints. However, the system remains vulnerable with respect to some of the considered disturbances.

This type of emergency condition is endured for a short period of time. Proper corrective control action should be taken to prevent its move into emergency state. The emergency state which occurs when some of the operating constraints violated, while the power system continually supplies power to all the loads. Operators should take corrective action to bring the system back to a normal state. The extreme emergency state occurs when both the operating and the load constraints are not satisfied. This causes the power system to become unstable. An immediate corrective control action must be taken to prevent the system collapse such as disconnecting various equipment include loads , lines, and transformers. This will limit the violations and recover the system stability. Then, the load-generation balance restored to start supplying power to all the loads [27]. The state diagram in Figure 2.1 shows the possible transitions between the various operating conditions defined above [28].

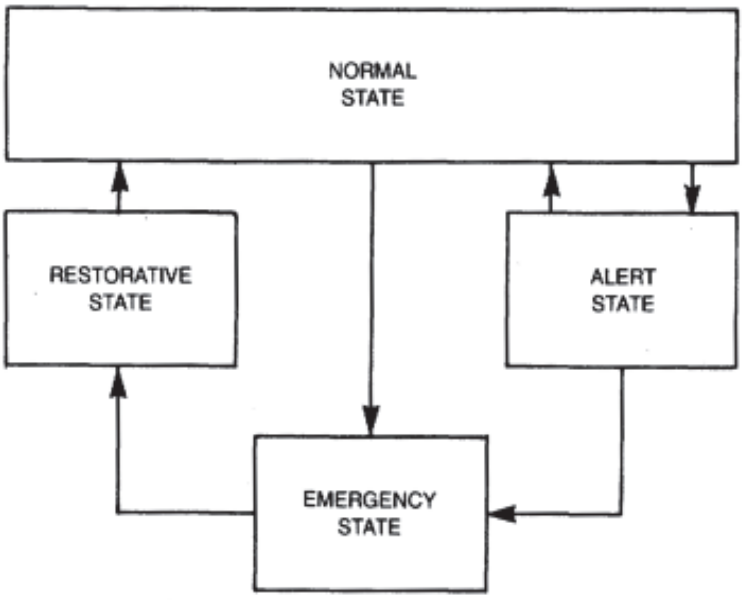

Figure 2.1: Power System Operating States

2.4 Power System Security Analysis

Due to the change in operating conditions during daily operation, the normal secure state is considered as the main goal of the system operator. This can be achieved by continuous monitoring of the system states, identifying of the operating condition, and determining proper preventive control actions to maintain the system in normal state. The initial step of security analysis is to monitor the present condition of the system. This includes acquisition of measurements from all components of the system to determine the state of the system [27].

The data is presented in the form of analog and digital quantities which is computed at the electric utility substations Sensors. These sensors transmit the data to the remote terminal units (RTUs). The data received from the RTUs is processed by the system operators in control center. The system operator estimates these data in order to determine the current condition of the electric power grid. In addition, the system operator monitors the system by switching on or off remote circuit breakers based on the available data [29].the evolving of communication and computer technologies has changed the design of the control centers. The way of gathering data has transformed since the digital computers involved in electric network [30].

Real time measurements include phase, voltage, real and reactive power, isolators, bus lines, circuit breakers, switches, transformers, and generators. These raw data and measurements are collected by remote terminal units (RTUs). State estimators process these data to eliminate the measurement noise and detect the bad errors. Optimal estimate of the system state is provided by state estimators based on the raw measurements. The output of state estimators will be used in energy management system (EMS) applications functions such as automatic generation control (AGC), economic dispatch (ED), optimal power flow, and contingency analysis [31].

Electric power system has Initially been monitored and controlled by Supervisory control and data acquisition. The supervisory control and data acquisition were first established by augmenting supervisory control systems by real time wide data acquisition. This allows the control center to collect all different measurements and circuit breakers status data from the power system. The real time operating states of the system should be determined to perform different application functions such as corrective real and reactive power dispatch and contingency analysis. Due to the

measurements error and instrument failure, and communication noise, the SCADA system may not be reliable. Moreover, The corresponding A.C. operating condition of the system may not be directly extracted from the available set of measurements due to the absence of all possible measurements [27].The function diagram of energy management system is shown in figure 2.2 [32].

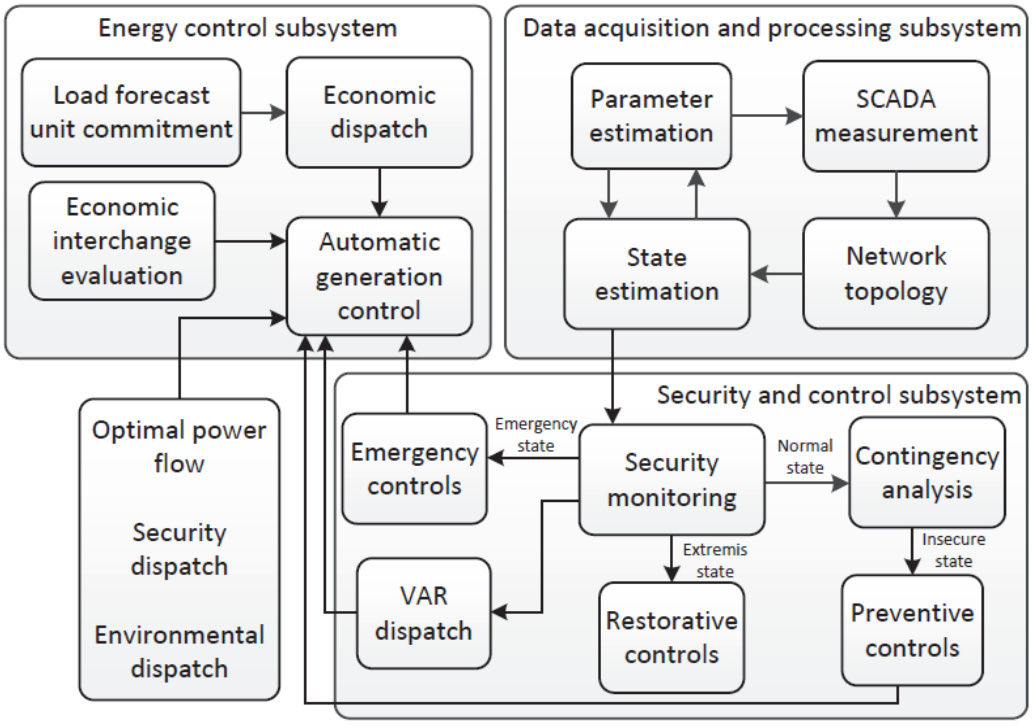

Figure 2.2: Function diagram of EMS

Chapter 3

Hourly Electricity Price Forecasting For The next Month Using Multilayer Neural Network

3.1 Introduction

In deregulated electricity markets electricity prices will fluctuate as a result of competition among power suppliers. Profit maximization has become a major motivation in electric markets. The imbalance between supply and demand results in volatile electricity prices. In a regulated market, load forecasting was a main focus of the electric power industry. Subsequently, the more complex electricity price forecasting has become more important in the deregulated electric power industry [33]–[37].

Electricity price forecasting is a critical factor in decision making in power systems. Its principal target is to lower the cost of electricity through rivalry, and augment efficient generation and consumption of electricity. In view of the non-storable nature of electricity, all generated electricity must be consumed. Thusly, both producers and consumers need accurate price forecasting so as to build their own particular methodologies for profit or utility augmentation [38].

Electricity price forecasting depends on input variables such as available historical price and load data, system operating conditions, weather conditions and temperature values, fuel prices, time indices including hours, weekdays, and seasons, demand, bidding strategies, operating reserves, imports, temperature effects, predicted power shortfalls, and generation outages [39], [40].

Bunn [41] reviewed several major innovative methodologies and techniques focused on daily loads and prices forecasting in challenging electricity markets such as variable segmentation, multiple modeling, combinations, and neural networks. He focused on forecasting strategy in both supply and demand side and discussed the benefit side of considering more influential factors such as weather. He suggested that further research and integration are needed in order to provide more accurate forecasts, so he

concluded that "the forecasting of loads and prices are mutually intertwined activities and the economic perspective alone cannot be an accurate basis for daily".

Skantze, and Ilic, classified power value models and existing significant distributions into six categories, and discussed their goals, attributes, and drawbacks [42]. Shahidehpour, et al. discussed the basics of electricity demand estimation including price, unpredictability, and exogenous variables and proposed a value determining module focused on neural networks [43].

Koritarov, promoted using Agent Based Models to uncover and clarify the perplexing and total framework practices that arise out of the interaction of the heterogeneous individual elements [44]. Ventosa, et al. discussed the diverse methodologies to study electricity markets, including the Nash-Cournot framework and the supply work balance approach [45].

In IEEE Power and Energy Magazine talk article, Amjady and Hemmati clarify the requirement for short-term price forecasts, survey issues identified with Electric Price forecasting EPF, and set forward suggestions for such expectations. They contend that time series techniques including autoregressive (AR), autoregressive integrated moving average model (ARIMA,) and Generalized Autoregressive Conditional Hetero-skedasticity (GARCH) are by and large just as effective in the regions where the frequency of the information is low, for example, week after week designs. Moreover, they advocate the utilization of computational insights and cross methodologies (neural systems, fuzzy regression, fuzzy neural networks, cascaded architecture of neural networks, and committee machines, which are fit for following the hard nonlinear behaviors of hourly load and particularly value signals [46].

Weron, provided an outline of modelling methodologies, then focused on pragmatic applications of statistical strategies for day-ahead forecasting including autoregressive integrated moving average model (ARMA), autoregressive moving average with exogenous variable model (ARMAX), and Generalized Autoregressive Conditional Hetero-skedasticity (GARCH) models [47]. Moreover he discussed interval forecasts and proceeded to review quantitative stochastic models for subordinates estimating "jump-diffusion models and Markov regime-switching".

Zareipour, reviewed time series of linear models such as autoregressive with exogenous input (ARX), ARIMA, and ARMAX and nonlinear models including regression

splines and multilayer neural networks, and then used them to forecast hourly costs in the Ontario force market [48].

Amjady briefly reviewed electricity price forecasting (EPF) techniques, then concentrated on artificial intelligence-based methods, and specifically emphasized choice of procedures and hybrid forecast engines [49].

Garcia-Martos and Conejo investigated short- and medium-term EPF, with a concentrate on time series models. Particularly, they consider ARIMA and regular ARIMA models aligned to hourly costs for day-ahead forecasts, and vector ARIMA [50].

Hong discussed spatial load forecasting, short-term load forecasting, EPF, and two smart grid era research areas including demand-response and renewable-generation forecasting. He categorized EPF models into three classes including simulation systems of the power markets, load forecasts, and blackout data, and offers from business members, measurable techniques, and artificial intelligence strategies (AI) [51], [52].

In the latest study of structural models, distributed as a chapter in the book Quantitative Energy Finance by Carmona and Coulon in 2014 presented a point by point investigation of the structural methodology for power demonstrating, underlining its benefits in respect to customary diminished structure models. Expanding on a few late articles, they advocate a wide and adaptable structural system at spot costs, consolidating demand, limit and fuel costs in a few ways, while computing closed structure forward costs all through. The aforementioned articles, book chapters and Ph.D. these are supplemented by a couple of overview meeting papers of changing quality [53].

In this chapter, composite multilayer neural network topologies including hybrid parallel and hybrid cascade topologies are applied to enhance hourly electricity price forecasting for the next month in Australian electricity markets. Both the hourly temperature, hourly electricity load, hourly natural gas data, and other hourly historical data have been considered in forecasting. The neural network models are trained on hourly historical data for the year of 2005 from Australian electricity markets published by the Australian Energy Market Operator (AEMO) to predict the hourly electricity price forecasting for January 2006. The simulation results obtained have shown that hybrid parallel topology is more accurate and less in computational time.

3.2 System Description

This chapter presents an attempt to design interconnections of neural networks with different connection topologies to improve the overall forecasting performance accuracy. A review of a number of successful implementations of multiple connection topologies is given first. Neural Networks may be connected in many configurations such as:

- Cascade or Serial Topology.
- Parallel Topology.
- Cascade Parallel Topology in cascade and in parallel.
- Parallel-cascade Topology in cascade and in parallel.
- Hybrid Parallel Topology
- Hybrid Cascade Topology

Figure 3.1 shows the cascade topology interconnected in Feed forward connection where the output of the first network is used as input for the second network. The neural network design is divided into stages including training stage and simulation stage. Each network is trained to optimize its inputs with respect to the target values.

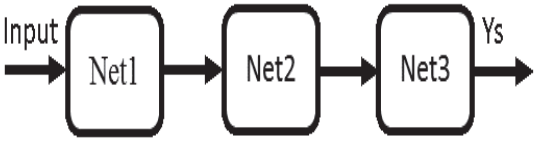

Figure 3.1: Cascade Topology

As a result, Net 1 will be trained using 'x' input value and 't' target value while Net 2 will be trained with the output of Net 1 "y1" as an input to it with respect to "t" target value, so connecting neural networks in cascade topology and finding the best cascade sequence. For example, speed and accuracy are evaluated for the cascade topology. The sequence (or order) is very important in this connection topology. It can be concluded that accurate Network should be placed first to minimize the overall fitting error and to enhance the overall performance. In the parallel topology shown in figure 3.2, one starts using the same input applied to all networks and then the output of all networks is aggregated into to single output neuron. Different weights

are assigned to each network output where the best network performing is assigned higher weight.

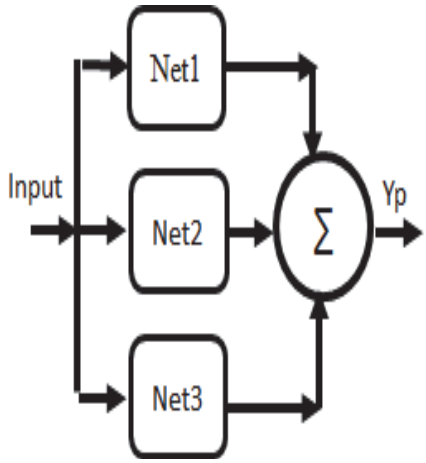

Figure 3.2: Parallel Topology

In connecting neural networks in parallel, there is no difference in connection sequence and the final output is improved compared with the output of each network resulting from the averaging mechanism. Enhanced results can be obtained by assigning a higher weight to the network with best performance and lower or even zero weight for the weak network. There are four connection combinations for Cascade Parallel such as cascade -parallel in Cascade, Cascade -parallel in parallel, parallel-cascade in series and parallel- cascade in parallel. From figure 3.3 to figure 3.6 show the difference between Cascade - parallel in cascade with respect to Parallel-Cascade in cascade and also sequence between Cascade-parallel in Parallel with respect to Parallel-Cascade in Parallel.

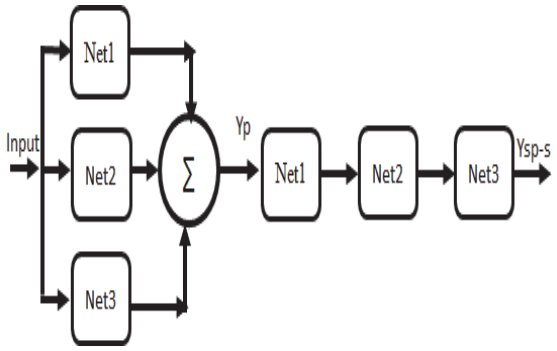

Figure 3.3: Parallel- Cascade in Cascade Connection

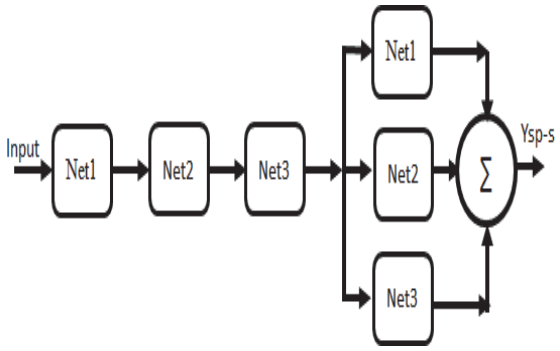

Figure 3.4: Cascade – Parallel in cascade Connection

From figure 3.7 to figure 3.8 demonstrate hybrid parallel topology and hybrid cascade topology to determine the most efficient neural network interconnection topology in terms of accuracy and time management. The output and input of each hyper topology could be connected in layers and the connection terminal of each one is connected to the other layers in whether cascade or parallel. For evaluating each network topology with three different Feed forward networks is trained with the same input-output data and each network has 6 hidden layers for network 1, 7 hidden layers for network 2, and 8 hidden layers for network 3 which will result to a good fit and then different hybrid connection topology is evaluated.

3.3 Simulation Results and Discussion

Our computational experiment involves the hourly historical data for the year of 2005 as a training data to predict the hourly electricity price for January 2006. The data

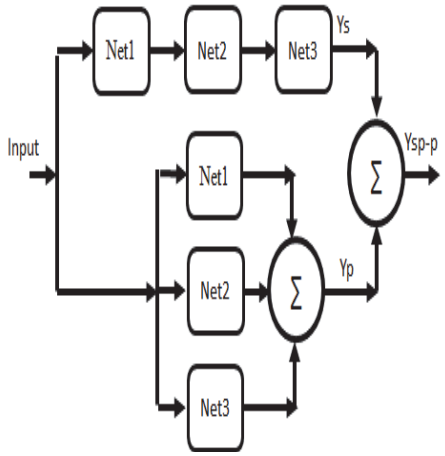

Figure 3.5: Cascade-Parallel in Parallel Connections

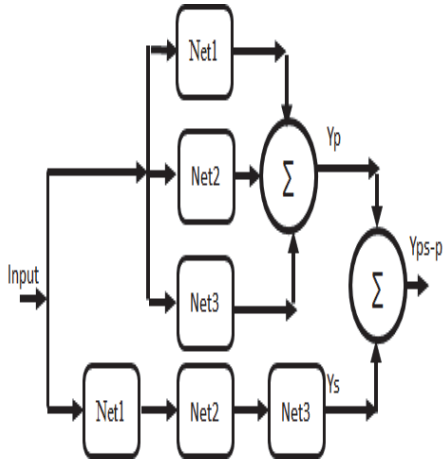

Figure 3.6: Parallel-Cascade in Parallel Connections

used is published by the Australian Energy Market Operator (AEMO) in electricity load forecasting for the Australian market [54].

There are many impact factors that can be used to forecast the electric price such as:

• Time in hours including previous day same hour load, previous week same hour load, and previous 24 hour average load.

• Weather conditions including dry bulb and dew point.

• Electricity price including previous day same hour price, previous week same hour price, and previous 24 hour average price.

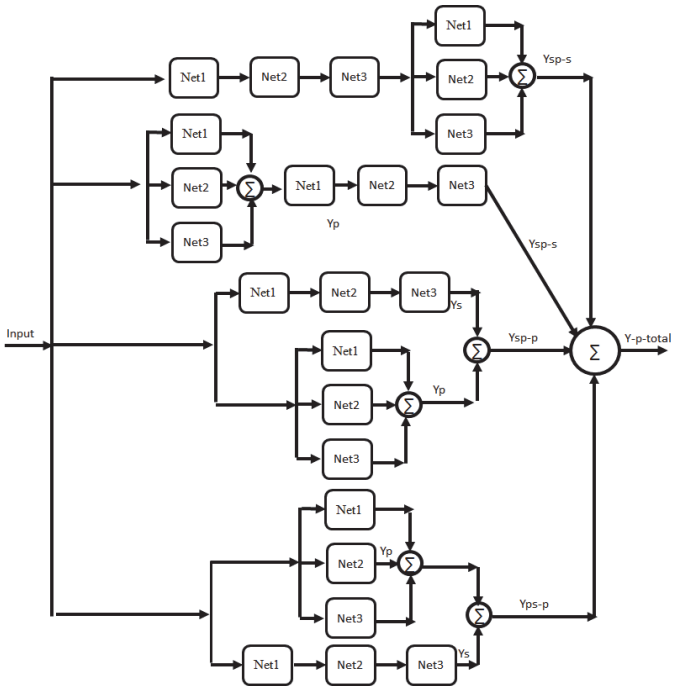

Figure 3.7: Hybrid Parallel Topology

• Dates including days of the week.

• Natural gas price including previous day same hour price, previous week same hour price.

• System loads including previous load forecasting.

• None business days working including holidays.

Where inputs include hourly load data, hourly weather conditions, hourly natural gas price and hourly other factors for the year of 2005 and target includes the electricity price for year of 2005. We use the training data includes inputs and targets to predict the hourly electricity forecasting for January 2006.

$$MSE = \frac{1}{N_i}\sum_{i=1}^{n}(y_i - \hat{y}) \tag{3.1}$$

$$MAPE = \frac{1}{N_i}\sum_{i=1}^{n}\frac{y_i - \hat{y}}{y_i} \times 100 \tag{3.2}$$

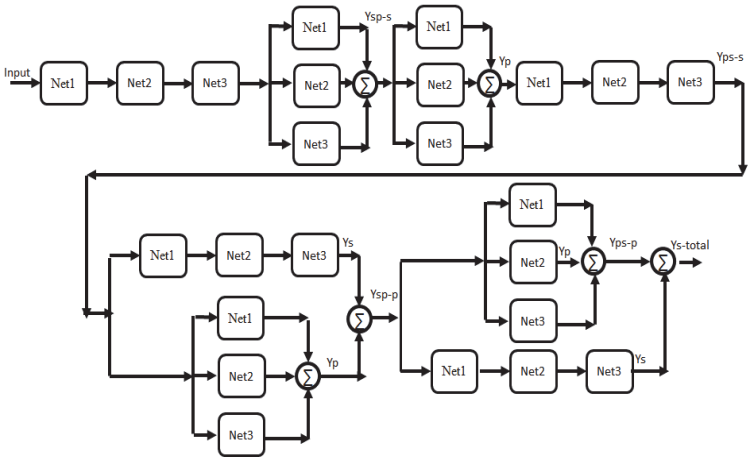

Figure 3.8: Hybrid Cascade Topology

$$R^2 = \frac{\sum_{i=1}^{n} \hat{y} - \bar{y}}{\sum_{i=1}^{n} y_i - \hat{y}} \tag{3.3}$$

Where y_i and $\hat{y}$ are the actual and predicted price respectively for January 2006 respectively, and N_i is number of predicted data. R^2 is coefficient of determination, $\bar{y}$ is the average of the data [38].

A few comments on the selections made in our experiment:

• We arbitrarily selected three different layers to compare performance and accuracy behavior between proposed topologies. More layers require more computation, but their use might result in the network solving complex problems more efficiently.

• We chose the number of neurons in each network based on the best performance, so we put 6 neurons for net1, 7 neurons for net2, and 8 neurons for net3. More neurons require more computation, and they have a tendency to overfit the data when the number is set too high, but they allow the network to solve more complicated problems.

• We selected the initial weights of a neural network from the range (1/d, 1/d), where d is the number of inputs to a given neuron. It is assumed, that the sets are normalized - mean 0, variance 1.

• We run the polynomial regression, different model selections include forward selection and backward elimination, to improve the full regression model by whether eliminating bad predictors (input factors) or selecting good predictors (input factors). It is concluded that natural gas price has the high impact on the electricity price.

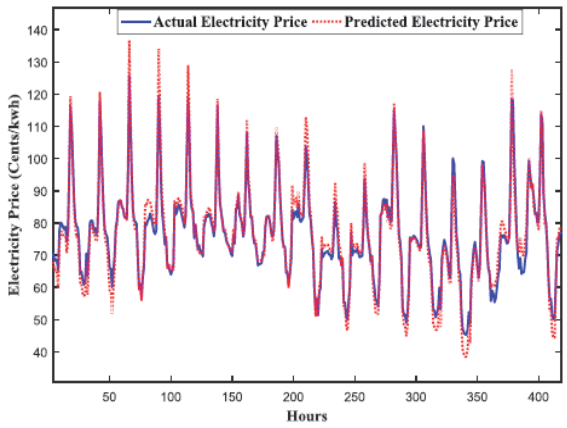

Figure 3.9: Hybrid topology in Parallel connection, actual and predicted hourly electricity price for Jan. 2006

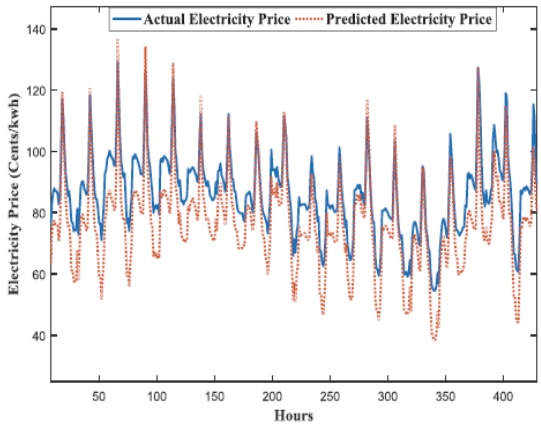

Figure 3.10: Hybrid topology in Cascade connection, actual and predicted hourly electricity price for Jan. 2006

Figures 3.9 and 3.10 show the predicted and actual electricity prices for January 2006. The total hours are 720 hours for a month, however we just take 400 hours

to see the variation between actual and predicted electricity price. It can be seen that the hybrid parallel topology forecasting is less variable than the hybrid cascade topology.

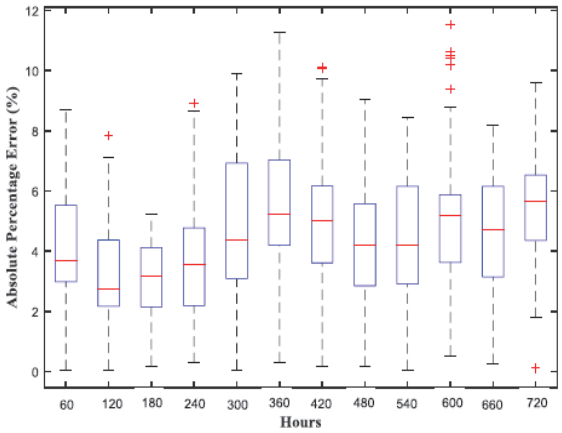

Figure 3.11: Hybrid Parallel topology, Absolute Percentage error for hourly electricity price for Jan. 2006

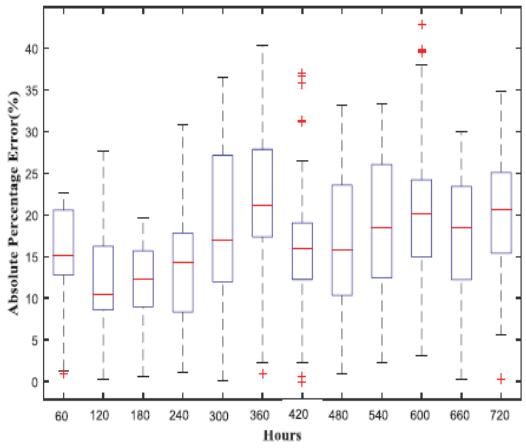

Figure 3.12: Hybrid Cascade topology, Absolute Percentage error for hourly electricity price for Jan. 2006

Figures 3.11 and 3.12 display the absolute percentage error of the hybrid parallel and cascade topologies for actual and predicted electricity price for January 2006.

We have divided the total number of hours (720 hours a month/60 hours=12 boxes). Each box consists of six hours of APE data and also each box represents different range includes maximum, average, and minimum for APE. The range of APE of the hybrid parallel topology is less than that of the hybrid cascade topology.

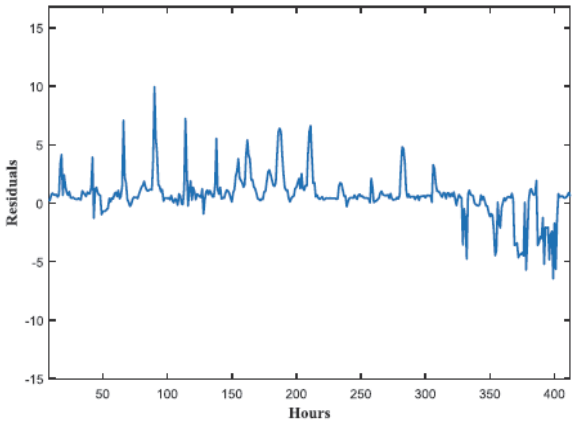

Figure 3.13: . Hybrid Parallel topology, Residual between the actual and the predicted hourly price forecasting for Jan.2006

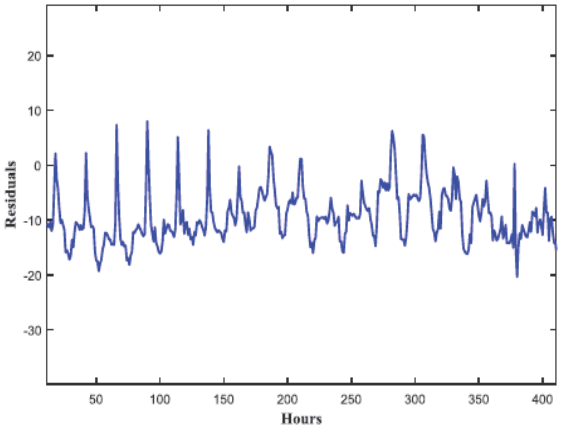

Figure 3.14: Hybrid Cascade topology, Residual between the actual and the predicted hourly price forecasting for Jan.2006

Figures 3.13 and 3.14 show results of the hybrid parallel and cascade topologies' residuals between actual and predicted hourly electricity price for January 2006. It

is evident that the residual range of hybrid parallel topologies is less variant than hybrid cascade topology.

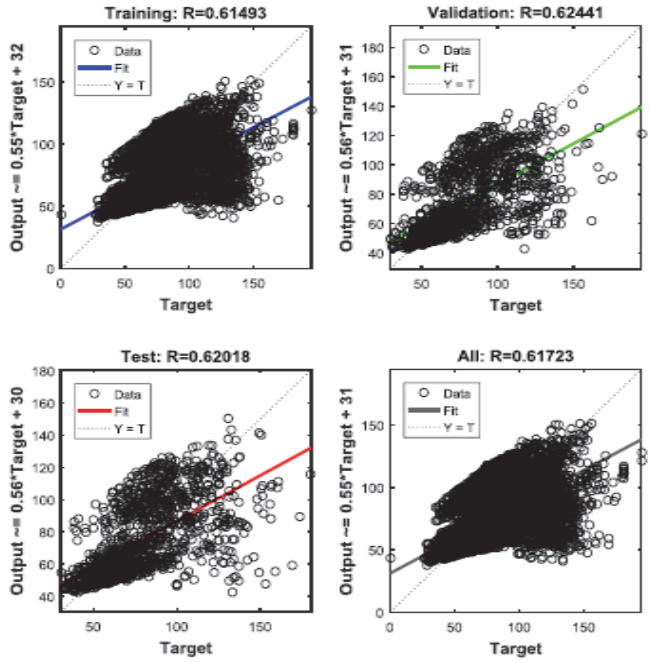

Figure 3.15: Hybrid Cascade topology, regression model for the actual and the predicted hourly price forecasting for Jan.2006

Figures 3.15 and 3.16 show regression plots displaying the output of both hybrid parallel and cascade topologies respect to targets for hourly data (Jan.2006), validation, and test sets. The data should fall along a 45 degree line to satisfy the equality of both output and target. For this case, the regression value in each and total case is 0.97118 in hybrid parallel topology compared to hybrid cascade topology is 0.61723. To improve the regression value, we may retrain the network many times to update the initial weights and biases of the network. According to the value of the regression in hybrid cascade topology, R=0.61732, It can be seen that the hybrid cascade topology has a bad fitting model. We can also increase the number of hidden layers or number of training vectors or using other algorithms for training.

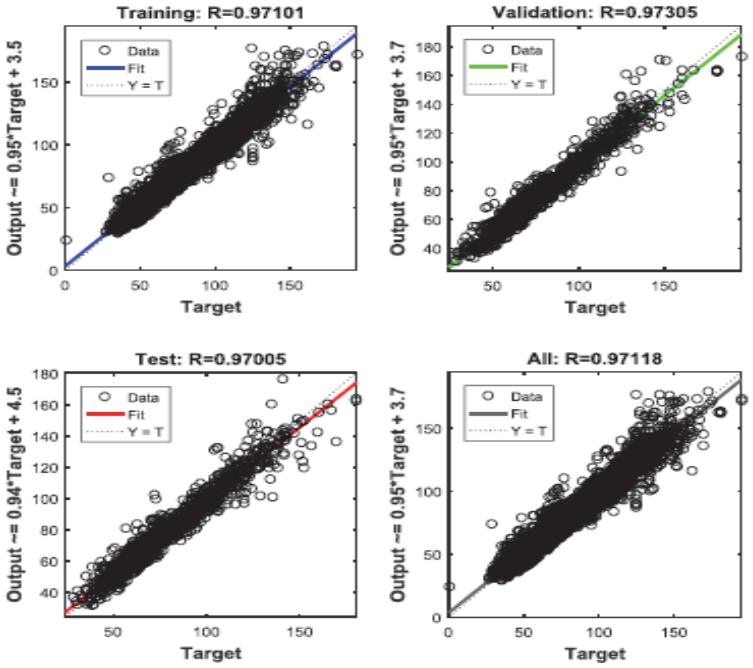

Figure 3.16: Hybrid Parallel topology, regression model for the actual and the predicted hourly price forecasting for Jan.2006

The weights in Table 3.1 are selected to improve the overall hybrid topology by giving high weights for the topology with best contributed performance and low weights for the topologies with weak performance. The actual values of the weights were selected through trial and error.The data is categorized to 75% used for training, 15% used for both of validation and testing each network.The non-linear relationships between electricity price output and the sixteen input variables including load data, natural gas price, crude oil price, temperature, humidity and other factors are important considerations.

Table 3.1: Weights For Three Interconnected Nets

Topologies	Cascade-Parallel in Cascade W1	Parallel-Cascade in Parallel W2	Cascade-Parallel in Parallel W3	Cascade-Parallel in Parallel W4
Hybrid Parallel	0.5	0.95	0.7	1

Table 3.2: Final Comparision In Cascade and Parallel Hybrid Connections in Different Performance

Hybrid Connections	Mean Square Error	Mean Absolute Percentage Error	Coefficient of determination	Time (Secs)	
	MSE	MAPE	R^2	Training	Simulation
Hybrid Topology in Parallel	9.11	3.25	0.9815	228.21	0.086
Hybrid Topology in Cascade	87.14	12.27	0.942	912.82	0.065

Table 3.1 shows the weights assigned to multilayer neural networks in parallel cascade. There is no difference in interconnection sequence because all networks are in parallel and the final output is improved compared with the output of each network due to the averaging mechanism in the parallel analysis topology. A better outcome may be obtained by introducing a higher weight to the best performing network such as w2 = 0.95 and w4=1.0. The higher weight is because these topologies contributed to improve hybrid parallel topology. Both of w1= 0.5 and w3= 0.7, the lowest weight is given to the network that makes the total, averaging parallel topology inefficient.

Table 3.2 offers a comparison of the hybrid cascade topology and hybrid parallel topology. In terms of training and simulating times, we conclude that hybrid parallel topology has the lowest time in seconds. For instance, hybrid parallel topology takes 228.21 seconds in training time and in simulating time takes 86 milliseconds. However the hybrid Cascade topology takes 912.82 seconds in training time and 65 milliseconds in simulation. In terms of accuracy, the hybrid Cascade topology shows MSE of 87.4 and the MSE error of the hybrid parallel topology 9.11. For the MAPE, the hybrid parallel topology error is 3.25 compared to 12.27 for the hybrid cascade topology. We conclude that hybrid parallel topology is efficient in comparison with hybrid cascade topology.

Table 3.3 gives a better picture of which interconnection topology is the most efficient. If we take a training and a simulating time, we can conclude that parallel topologies have the lowest time in seconds to produce the most efficient solution.

Table 3.3: Comparison Of Performance in Multilayer Topologies For Actual and Predicted Electricity Price For January 2006

Cascade	Parallel	Cascade-Parallel in Cascade	Cascade-Parallel in Parallel	Parallel-Cascade in Cascade	Parallel-Cascade in Parallel	Performance Indices
38.98	4.11	4.95	45.44	11.85	11.85	Mean Square Error (MSE)
6.93	1.42	1.54	7.66	3.76	3.76	Mean Absolute Percentage Error (MAPE %)
10 Iterations at the 6 validation checks	1000 Iteration at the 0 validation checks	181 Iteration at the 0 validation checks	1000 Iteration at the 0 validation checks	11 Iteration at the 6 validation checks	1000 Iteration at the 0 validation checks	Maximum number of training Iteration (Epoch) at the validation checks
0.9439	0.9912	0.9975	0.9172	0.9745	0.9898	Coefficient of determination
0.31	0.084	0.25	0.084	0.022	0.086	Simulation Time (s)
102.91	50.32	480.25	127.65	100.48	204.43	Training Time (s)

Table 3.4: Ranking Training Algorithms Of Feed Forward Neural Network With 10 Hidden Layers in Hybrid Parallel Topology in Case Of Accuracy

#	Algorithm	Training Time (secs)	Operation Time (secs)	Mean Square Error	Coefficient of determination	Mean Absolute Percentage Error
1	LM	276.03	0.086	9.11	0.9815	3.25
2	BFG	237.06	0.083	14.95	0.9596	4.74
3	OSS	21.27	0.085	15.21	0.9554	4.93
4	RP	21.29	0.093	15.65	0.9536	5.07
5	SCG	18.32	0.092	22.3	0.9355	7.12
6	GDX	21.09	0.088	30.05	0.9032	10.39
7	GDM	13.15	0.091	50.21	0.7853	18.75

For instance, the Cascade–Parallel-in-Cascade interconnected topology takes 480.26 seconds and in simulating time takes 250 millisecond to produce the optimum solution. However in cascade -parallel in parallel interconnected topology takes just 100.48 seconds in training time and in simulating time takes 84 milliseconds to do that. In terms of accuracy, it can be seen that the MAPE for the Cascade -Parallel in Cascade topology has 1.54%. However, the Cascade -Parallel in Parallel topology shows 7.66%, also the Parallel-Cascade in Cascade and Parallel-Cascade in Parallel topologies show 3.76%. Starting and ending with Cascade topology can enhance the total performance. Also starting with Parallel and ending with Parallel and Cascade will not improve the performance according to MAPE=3.76% in both of Parallel-Cascade in Cascade and Parallel-Cascade in Parallel. It is affirmed that Parallel topologies have the least MAPE. Moreover, starting and ending with parallel topology can minimize the training time. For instance, Parallel-Cascade in Cascade topology requires less training time, 127.65 seconds, than parallel-Cascade in parallel topology, which required 204.43 seconds.

Table 3.4 shows an evaluation of each network topology in hybrid topology based on the actual and predicted electricity price for January 2006, three different Feed forward networks is trained. Each network has 6, 7, and 8 respectively hidden layers. In terms of accuracy, the Levenberg-Marquardt algorithm (LM) yields the most accurate performance exhibited by MSE=9.11 and MAPE =3.25 compared to the variable Learning Rate Back propagation (GDX), showing MSE = 10.39 and MAPE= 30.05%.

Table 3.5: Ranking Training Algorithms Of Feed Forward Neural Network With 10 Hidden Layers in Hybrid Cascade Topology in Case Of Accuracy

#	Algorithm	Training Time (secs)	Operation Time (secs)	Mean Square Error	Coefficient of determination	Mean Absolute Percentage Error
1	LM	509.74	0.086	87.14	0.942	12.27
2	SCG	27.72	0.092	90.71	0.9154	16.15
3	OSS	33.51	0.085	100.86	0.8754	17.43
4	BFG	505.41	0.083	112.89	0.8356	25.12
5	RP	31.77	0.093	120.15	0.7785	27.05
6	GDX	28.38	0.088	127.34	0.7037	40.32
7	GDM	15.17	0.091	222.04	0.23	105.75

In the case of operating time, LM takes 276.03 seconds to compute the optimum solution compared to the variable Learning Rate Back propagation (GDX) which has the minimum operating time 21.09 seconds. The gradient descent with momentum back propagation (GDM) is the least efficient algorithm in terms of both accuracy and operating time.

Table 3.5 shows an evaluation of each network topology in hybrid topology based on the actual and predicted electricity price for January 2006, three different Feed forward networks is trained. Each network has 6, 7, and 8 respectively hidden layers. In terms of accuracy, the Levenberg-Marquardt algorithm (LM) scores the most accurate performance exhibited by MSE= 87.14 and MAPE= 12.27% compared to the Resilient Backpropagation (RP), MSE = 120.15 and MAPE= 27.05%. The variable Learning Rate Back propagation (GDX), shows high error. In terms of execution time, LM takes 509.74 seconds in training to compute the optimum solution compared to Scaled conjugate gradient backpropagation (SCG) which has the lowest operating time of 18.32 seconds. Gradient descent with momentum back propagation (GDM) is the most inefficient algorithm in both accuracy and operating time.

The cascade-forward-net structure is less efficient and therefore, is not discussed further. We decided to set fit net and feed forward to use the different three multilayer nets and it seems to be working efficiently. But we are looking for more enhancement of neural network, so we set the feed forward for the three layer which gives the most efficient results with small improvement in minimizing the errors with less training

and simulating time. We also attempted to increase the number of neurons, but this did not improve the performance of the multilayer neural network, so we have tried to change the number of hidden layers to 2 for each network and 10 for each network, however it achieved slightly higher error. We implemented several algorithms with back propagation neural network and the best one is the Bayesian regulation back-propagation algorithm, but it takes a long time to complete training. We set feed forward net in the first and second layers net, and fit net in the third layer, however we obtained less training time with some outliers of errors which is makes our option is not efficient as we expected. We tried to set the first two layers as a fit net and the third one as a feed forward net, but mean square error is high with good results in parallel connection in regression and mean absolute percentage performance and also less time training.

After testing all aforementioned options, we come up with the best option, we decided to set Levenberg-Marquardt algorithm for both Cascade and Parallel training functions to enhance the performance. Also we set Bayesian regulation back-propagation algorithm for the first net and Levenberg-Marquardt algorithm for the second net2 and third net3 to reduce the accumulated error. Moreover, we set 6 neurons for network 1, 7 neurons for network 2, and 8 neurons for network 3. We finally conclude that these options give the most accurate performance especially in hybrid parallel topology.

3.4 Conclusion

This paper has demonstrated a new computational approach to applying composite back-propagation multilayer neural networks to forecast hourly electricity price for the next month based on hourly important factors such as previous hourly load, hourly natural gas, and hourly weather conditions for electricity load forecasting for the Australian market in January 2006. We interconnected three separate nets in a cascade and a parallel topology to acquire the best performance based on simulation results. A comparison of the Cascade and Parallel alone topology results in performance superior to Parallel topology. A comparison of the Cascade-Parallel in Cascade topology with Cascade-Parallel in Parallel, reveals that starting topology by a Cascade and ending with a Cascade connection results in performance superior

to starting topology by a parallel and ending with a Parallel connection. Therefore, the Cascade-Parallel in Cascade topology is superior to Cascade-Parallel in Parallel. A comparison of the performance of the Cascade-Parallel in parallel with that of Parallel- Cascade in Parallel, we concluded that starting the topology with Parallel yields superior performance to starting with the Cascade connection. Likewise, we found that there is no different in accuracy between a Parallel-Cascade in Cascade and Parallel-Cascade in Parallel connection. From above experiment, it could be reasoned that the performance of the Cascade– Parallel in Cascade topology is superior to all different aforementioned topologies.

Following the analysis of performance of each topology, it is concluded that the parallel topology which averages the output errors improves the performance in the systems. On the other hand, the Cascade topology accumulates the error from each training stage which mostly improves the overall performance in the systems with high data distribution.

Chapter 4

Optimization of Neural Network Parameters by Stochastic Fractal Search for Dynamic State Estimation under Communication Failure

4.1 Introduction

Power flow pattern has been less anticipated due to the expansion in electric power grid throughout the recent ten years. The expansion is exhibited in the entrance of renewable energy sources and deregulation of the power system. The power system grid must be observed and controlled proficiently by control centers to be operated in a secure and reliable manner. Modern energy management (EMSs) and control systems could assist in maximizing energy savings and minimizing the cost of energy. However, they are intricately connected to accomplish the desired objective. Different proposals to EMSs with distinctive optimization techniques and different grid structures have already been presented in literature [55]–[64]. SE is the main part of the EMS system, which utilizes a redundant set of data to determine precise and dependable states of the power system. The system states (voltage phasors) are further used in real-time applications including security investigation, economic operation dispatch and load flow [65].

Conventional WLS state estimation has been widely used due to its high precision and easier implementation. The accuracy of WLS estimator is based on the availability of instantaneous measurement to assess the system state. For instance, the system state can be assessed at time, step t when the measurement at only step t is utilized. At the point when the system encounters some sort of communication failures, such as, delayed or lost measurement data, the execution of the WLS estimator breaks down. Therefore, specific parts of the electric power grid will be unobservable and control centers will have no information about these parts. Another drawback of the WLS state estimation is that, it cannot measure the system state for the next time,

step t+1[66]. Moreover, several DC power flow SE studies have been represented to overcome the AC state estimation problems [67]. In the aforementioned referenced study, the author built a DC power flow model (DCPF) to improve the accuracy of power flow estimation by introducing the measurement noise detection approach. Conventional SCADA system suffers from low updating rate and inaccuracy. These limitations can restrict the reliability of SE. PMU devices play an important role in improving the precision of power system.

PMUs can increase the sampling rate of measurements from several seconds per measurement (conventional SCADA) to over 30 measurements per second. This is vital for improving monitoring and analysis of the dynamic conduct of power systems. PMUs give synchronous measurements (via global synchronous time stamps), which can synchronize measurements from far off areas to provide a continuous picture of the entire power system. Both the voltage and current phasors of a given bus can be measured directly from where it is installed on power network [68]. However, it is difficult to introduce a PMU on each node of the power grid due to the expensive cost of PMU devices in addition to the constrained communication channel bandwidth. Therefore, past reviews state estimations that concentrate on PMUs and optimal meters placement. There are various earlier techniques that deal with PMUs and meter placement in the distribution system. Ref. [69], the authors proposed circuit representation model of SE error to demonstrate the relationship between estimations and SE errors. A logarithmic articulation of the circuit representation model is implemented where the optimal meter placement problem is effectively changed into an optimal network expansion problem. The model has proven superiority for enhancing SE in distribution network. Ref. [70] provides literature review on Phasor Measurement Units and Optimal Meters Placement for State Estimation Studies.

Generally, Kalman Filter (KF) is utilized to find solutions for the DSE problem because of its easier implementation, ability to forecast the system states and efficiency. Therefore, KF has been utilized extensively in many DSE applications. In Ref. [71], two-level DSE is presented using EKF that combines both SCADA and PMU data. Parallel DSE is improved on a graphic processing unit (GPU), which is particularly intended for preparing large data sets. The approach is proposed to improve the computational time in the DSE for large-scale grids. A two-stage Kalman

filtering technique is introduced in Ref.[72] to assess the static conditions of voltage phasors, the dynamic conditions of rotor angles and speeds of the generator. In the first stage, it takes the raw PMU data into the Adaptive Kalman Filter with Inflatable Noise Variances. The output is then taken into the second stage, which utilizes an Extended Kalman filter to estimate the true dynamic states. It concluded that EKF can provide an accurate solution to the DSE problem under normal operating conditions but under sudden load changing conditions and drastic generation variation, there can be a challenge due to a higher degree of non-linearity in the measurement function [73]. Approach [74] has implemented optimal KF instead of EKF to obtain the unknown complex bus voltages in addition to finding the linear relationship between complex element voltage and complex bus voltage. KF gives a more precise and quicker estimation. Hybrid technique combines unscented transformation with KF, which has been proposed in Ref. [75] to enhance the accuracy and to conquer the drawbacks of EKF technique. Decentralized DSE of power systems using unscented Kalman filter based on consensus algorithm has been discussed in Ref.[76].

The objective function turns out to be exceptionally non-linear, discontinuous and undifferentiable due to the presence of non-linear devices, such as, var compensators, distributed generators and transformers with on-load tap changers in the power system. This further reduces the solution of conventional optimization methods [77]. To overcome the limitations of classical optimization methods, artificial intelligence techniques, evolutionary algorithms, fuzzy and neural networks have been used as a prospect to add new scopes to the area of DSE [78]. Recently, a method based on fuzzy control theory improved with a sliding surface in Ref. [79] to explore the dynamic state estimation where the sudden load variation happened. This method can cleverly direct the solution to a close ideal trajectory to reduce the computational time. The technique combined the error and the rate of error as an incorporated input variable. Another approach based on ANN has been discussed in Ref. [80]. In Ref. [80], ANN was based on bus load prediction (DLP) for dynamic state estimation. The method classifies the process into two stages: short-term load forecasting and rectangular coordinate formulation for filtering. It concluded that ANN based on DLP gave a precise prediction, an accurate estimation and less computational time.

The principle objective of evolutionary techniques is to determine the optimum

solution in the favorable region already identified. However, for dynamic optimization environment, this sort of approach may encounter a convergence issue due to exist dynamic landscape changes in one region and if there are no particles of the algorithm in this region, the algorithm will not be able to react to the change proficiently, and might not succeed in tracking the changing global optimum [81]. Recently, metaheuristic techniques have been employed to find the solution of difficult problems, which require investigating a larger space. Normally, metaheuristics depend on four primary purposes: investigating the space effectively with being insensitive to the magnitude of the search space, reducing the computational time, solving large and complex problems, and obtaining robustness. Additionally, there is simplicity of design and implementation [82]. Natural phenomena behaviors are the inspiration of metaheuristic techniques. Genetic Algorithm (GA) presented in Ref. [83] mimics the behavior of natural evolution processes. Particle Swarm Optimization (PSO) is proposed in Ref. [84] to simulate the behavior of flocks of birds in searching their food. Both Artificial Bee Colony (ABC) and Ant Colony (AC) are represented in Ref. [85], [86] to simulate the foraging behavior of bee swarm and ant colonies respectively. These algorithms are utilized to tackle complex computational optimization problems, however, fast convergence along with accuracy is not ensured. Therefore, another powerful metaheuristic technique is proposed that uses a mathematical concept called fractal to overcome the above disadvantages. The proposed method shows superiority compared to other metaheuristic algorithms in case of finding the global minima, avoiding being stuck in local minima and for reducing the computational time.

This chapter presents a newly improved hybrid optimization technique for multilayer perceptron neural network (MLP) parameters (weights and thresholds) based on Stochastic Fractal Search technique (MLP-SFS) to solve the dynamic state estimation under communication failure problem. Communication failure is a type of cyber attack that affects the secure and reliable operation of power grid. The effectiveness of MLP (random weights and thresholds) and optimized MLP based SFS technique (optimized weights and thresholds) have been examined on IEEE 14-and 118-bus test systems using realistic load patterns from the New York Independent System Operator (NYISO) under different communication failure and measurement error. The

hybrid technique (MLP-SFS) compared to other techniques such as nonoptimized MLP, optimized MLP based Particle Swarm Optimization (PSO), and Genetic algorithm (GA). The obtained results clearly pointed out the superiority of the hybrid technique (MLP-SFS) exhibited in good accuracy and least computational time. The estimation problems treated are at high voltage transmission level, and therefore the generation and distribution level models are not involved in the analysis.

4.2 Stochastic Fractal Search Technique (SFS)

The process of SFS algorithm is categorized into diffusing and updating process. Diffusing process, every particle diffuses around its present position to fulfill the exploitation property. This procedure expands the possibility of determining the global minima, furthermore avoids being stuck in the local minima. Static diffusing process is only considered in this paper. We just consider the best generated particles and discard the others. Updating process, the SFS algorithm evaluates how a perspective in the sets updates its position dependent upon the different positions of the points in the group. Two statistical methods are considered in creating particles from diffusing process include Levy flight walk and Gaussian walk. Both originated from Markovian stochastic form [82]. Reference [87] explains in details the levy flight walks. The Gaussian random walk (GRW) is random walk utilized in Diffusion Restricted Aggregation(DLA). The advantage of using Gaussian walk is fast to converge and capable of finding the global minima [88]. Reference [89] provide complete study and development of the Diffusion Restricted Aggregation(DLA).

4.3 Overview of DSE under Communication Failure

The primary objective of high economical and reliable operation is supply continuity at high power quality to the user. This can be achieved by developing the interconnected networks to monitor, control, and protect power grids [36]. Today, power grids are more likely to be exposed to cyber-attacks which can block the transmission of real time measurements from RTUs to control centers. References [90]–[92] studied the likelihoods of cyber-attacks in details in power grid. Cyber-attacks can be sorted

into three classifiers; the data accessibility, data integrity, and data security. Supervisory control and data acquisition (SCADA) at control centers receive the real-time data transmit by remote terminal units (RTUs) [93]. Reference [94] provides one of the most common attack in data availability, Denial of service (DoS) attacks. This attack leads to significant block of the internet between RTUs and control center. Subsequently it causes disruptions in the power grid operation. After communication failure occurred the restoration time of the devices will be delayed [95]. The sequential Monte Carlo simulation technique is used to prove that load shedding will be increased due to communication failure [96].

4.4 Theoretical background

4.4.1 Backpropagation Neural Network (BPN)

Feed forward NN includes input, hidden, and output layers. The input layer is fully linked to output layer by one or more hidden layers. The output of the hidden layer is summed up of weighted input and biased, then activation function is applied to smooth the output which are computed and provided on to the next hidden layer. The process continues straightforwardly to the output layer. Connection weights and biases are set backwardly using perceptron (neurons)convergence. The training procedure through the generalization rule is an iterative procedure. Every training iterations incorporates two process, forward propagation, updating weights from input to output layers and backward propagation, error signals are determined at the output layer and minimized by re-adjusting the weights of forward propagation process at each iteration. The learning of BPN relies on upon the quantity of nodes in the hidden layer(s) and the example of associations between nodes in adjacent layers [97].

4.4.2 Overview of BPN related works

Optimization of neural network parameters (weights and thresholds) have recently been known as one of the most effective techniques in increasing the rates of convergence and finding the global minima [98]. The author in [99] presented specific curvature matrix vector to iteratively approximate second order gradient methods. Stochastic meta-descent (SMD) implemented to accelerate the presented technique.

It concluded that the results showed improved performance stability. In [100], the authors used partial least squares (PLS) to optimize weight initialization and size of hidden layer in BPN. The comparison between PLS and BPN individually with the hybrid PLS-BPN showed that hybrid PLS-BPN provided minimum root average square error and increased rates of convergence. In [101], the authors introduced polynomial bases using elliptical partial differential equation to BPN weight initialization. It pointed out that the technique converged 90% faster than random weight initialization. Genetic algorithm (GA) has widely been used to determine the optimal BPN parameters (weights, thresholds, and size of neurons) due to its capability of finding global minima [102], [103]. Many algorithms have recently been implemented to optimize BP neural network such as particle swarm optimization (PSO-BPN) [104], [105], Anti Colony (AC-BPN) [106], Simulating Annealing (SA-BPN) [107], and hybrid of genetic algorithm and particle swarm in BP neural network (HGAPSO-BPN)[108].

4.5 Proposed method and problem description of MLP via SFS under DSE

The following steps describe the process of optimized neural network via stochastic fractal search technique under dynamic state estimation.

4.5.1 Short Term Load Forecasting in Time Series using Support Vector Regression (STLF-SVR)

Support Vector Machine Regression (SVR) is used to forecast the unavailable load data due to communication failure for six hours ahead. SVR was initially presented by Vapnik in 1995. He extended the idea of SVM to solve the nonlinear regression function using -insensitive loss function. The fundamental concept of SVR is to map the input vector of input space into high dimensional component space by nonlinear mapping [109]. The SVR is formulated in optimization problem as following,

$$f(x, w) = w^T \phi(x) + b \tag{4.1}$$

$$Minimize\frac{1}{2}||w||^2 + c\sum_{i=1}^{l}(\zeta_i + \zeta_i^*) \tag{4.2}$$

Subject to the constraints:

$$\begin{cases} y_i - (w, x_i) - b \leq \varepsilon + \xi_i \\ (w, x_i) + b - y_i \leq \varepsilon + \xi_i{}^* \\ \zeta_i, \xi_i{}^* \quad \geq 0 \end{cases} \tag{4.3}$$

Where x_i represents the input vector which will be mapped into m-dimensional space, y_i denotes the target or the desired output, (x) denotes the mapping function, w represents weight vector, b is a bias form, C is the penalty coefficient, ε denotes Vapnik's insensitive loss width. Non-negative slack variables ζ_i,ζ_i^* are added to measure the deviation amongst input and target samples out of the -insensitive zone. The ideal hyperplane coefficients $f(x, w)$ can be computed by utilizing Lagrangian multipliers. Kernel function is used to transform input data from low dimensional space into high dimensional feature space. The execution of SVR relies on the best selected coefficients incorporate the error cost C, the loss function width , and the type of kernel function. Weka software is used for SVR modeling. Sequential Minimal Optimization (SMO) algorithm is utilized to train SVR coefficients. The control parameters for SVR, $C = 10$, $\varepsilon = 1e-3$, kernel type: polynomial kernel. The technique for the load forecasting is done by using the hourly historical data of 45 days from the day before the forecast day at the same year, and past 45 days before and after the forecast day in the previous year. The hourly historical data, March 22nd to Jun 19th 2015 and March 22nd to May 5th 2016 will be used to forecast the NYISO load for six hours ahead On Thursday May 5th 2016 from 18:00 PM to 23:00 PM during the communication failure. The New York Independent System Operator (NYISO) regions are combined with IEEE 14-bus system because both consist of 11 load regions with five minutes' interval [110]. Each load bus of IEEE 14-bus system is linked to one region of NYISO; the first row represents the IEEE 14-bus system and the second row of the index matrix represents the NYISO load regions,

$$\begin{pmatrix} 2 & 3 & 4 & 5 & 6 & 9 & 10 & 11 & 12 & 13 & 14 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \end{pmatrix}$$

Due to the high correlation between 11 load regions, time series forecasting is utilized to overcome the inaccuracy. The mean absolute percentage error (MAPE) for forecasted load in six hours ahead calculated based on Eq. (4.4),

$$MAPE = \frac{1}{N_i} \sum_{i=1}^{n} \left| \frac{L_i - \hat{L}}{L_i} \right| \times 100 \tag{4.4}$$

Where L_i and L denote the actual and predicted load respectively, and N_i is number of predicted data.

4.5.2 NYISO Load Normalization

Due to the unavailable data of reactive power load, we assume that power factor=0.8 according to National Electric Manufacturing Association (NEMA) website [111], so the reactive power load will be computed by multiplying the power factor to the NYISO forecasted active power load. NYISO forecasted loads will be normalized to operate near to initial system states of IEEE 14-and 118-bus system load.

4.5.3 Optimal power flow (OPF)

Conventional optimal power flow using interior point method (OPF-IPMs) is used to minimize the generation cost by setting all the control vectors within their constraints [112]. The input of the OPF is the forecasted load by SVR. The output of OPF consists of real time measurement, active and reactive power injections and flows and system states, voltage angles and magnitude voltages. The following equation is formulated to minimize the cost function,

$$Minimize \quad \sum_{i=1}^{n} F_i(P_{Gi}) \tag{4.5}$$

Subjects to Constraints:

$$\begin{cases} Q_k = 0 \\ P_k = 0 \\ P_{Gi\,\min} \leq P_{Gi} \leq P_{Gi\,\max} \\ V_{i\,\min} \leq V_i \leq V_{i\,\max} \end{cases} \tag{4.6}$$

Where F_i denotes the cost function, P_Gi denotes the generation at bus i, P_k and Q_k denotes the nonlinear power flow equations, V_i denotes the voltages at bus load i.

4.5.4 White Gaussian Noise (WGN)

Due to the lack of reasonable field measurements, white Gaussian noise with zero mean and a standard deviation is added to the output of OPF, real-time measurement, active and reactive power injections and flows. In order to make the simulated data resemble field data.White Gaussian noise is added to the output of OPF. The standard deviation for the active and reactive power injections σ(injections)=0.05, and the standard deviation for the active and reactive power flow σ(flows)=0.001. We create five sets of measurements error. The measurement error will be increased by the standard deviation, k=0,1,2,3,4,5, where k=0 implies there is no measurement error. The standard deviation will be increased by zero mean and the following standard deviation, σ(injections)=$k\times$ 0.05, and σ(flows)=$k\times$ 0.01. White Gaussian noise is added using the following formula,

$$Z_s(t) = Z(t) + rand \times k \times \sigma + \mu \tag{4.7}$$

Where $Z(t)$ and $Z_s(t)$ denote the measurements before and after adding white Gaussian noise respectively at time step t, k denotes the measurement error, σ denotes the standard deviation, and μ denotes the mean which equals to zero.

4.6 Implementation of MLP based on SFS

The flow chart of the proposed hybrid MLP based on SFS approach is shown in figure 4.2.

4.6.1 Initialization

This includes SFS and MLP parameters initialization. Different computations of parameters initialization have been made include population size, upper and lower boundaries, and size of hidden layer are shown in Tables 4.3 and 4.4. The best initial parameters will be used to start the experiment. The best initial parameters are based on the least mean square error (MSE).

Algorithm 1 Hybrid MLP based SFS

Set the input, real-time available measurement $z(t)$ and the target, true system states $x(t)$ (load flow) at time step t.
Initialize a random population size N_p , No. of hidden layers N_h, maximum No. of generations h, the maximum diffusion walk k, the side walk SW.
Set the upper and the lower constraints of MLP NN parameters (weights and thresholds) $P_j = LB + \varepsilon \times (UB - LB)$
while $h \leq maximumNo.ofIteration$ **do**
  Encode the population size (particles) based on Eqs.(4.8-4.10)
  Calculate the fitness function (MLP) as shown in Eq.(4.17)
  for every point P_i in the population size **do**
    Call diffusing process
    for $j = 1 : k$ **do**
      If (for simple problem, first Gaussian walk is applied) **do**
      A new point will be created:
      $GW_1 = Gaussian(\mu_{BP}, \sigma) + (\varepsilon \times BP - \varepsilon^{'} \times P_i)$
    end for
    Else (for complex problem, second Gaussian walk is applied) **do**
    A new point will be created:
    $GW_2 = Gaussian(\mu_P, \sigma)$
    end If
    end Call
  end for
end while
Call updating process
First updating process
First, all points of MLP parameters (weights and thresholds)are ranked : $P_{ai} = \frac{rank(P_i}{N}$
for every point P_i **do**
  for each component$j$ in P_i **do**
    If $rand[0, 1] \leq P_{ai}$
    Update the the MLP parameters

28: $P_i^{'}(j) = P_r(j) - \varepsilon \times (P_t(j) - P_i(j))$
29: **end** If
30: **Else** do nothing
31: **end** Else
32: **end for**
33: **Second updating process**
34: all points determined by the first updating process are re-ranked
35: **for** every new point $P_i^{'}$ of MLP parameters **do**
36: **If** $rand[0,1] \leq P_{ai}^{'}$
37: **Update** the position
38: $P_i^{"} = P_i^{'} - \varepsilon^{'} \times X(P_t^{'} - BP) > 0.5$
$P_i^{"} = P_i^{'} - \varepsilon^{'} \times X(P_t^{'} - P_r^{'}) < 0.5$
39: **end** If
40: **Else** do nothing
41: **end** Else
42: **end for**
43: **end for**
44: **Train** the optimized MLP parameters (weights and thresholds) based SFS by simple Multilayer MLP NN Using LM algorithm
45: **Calculate** MAPE for system states based on Eqs.(4.18-4.19)
46: **When** $z(t) \rightarrow z(t+1)$ & $x(t) \rightarrow x(t+1)$
47: **Go** step 1

4.6.2 Encoding Strategy

4.6.2.1 Vector encoding strategy

In this encoding strategy, every particle is encoded for a vector. For Multilayer Perceptron, each particle represents all weights and thresholds of a MLP's structure. MLP structure for IEEE 14-bus system consists of 1–25–2, real time measurement includes real and reactive power injections and flows as one input vector, number of hidden neurons equals 25, and two output vectors include phase and magnitude

voltage.

$$gen(i) = [w_{1,1}...w_{1,25}\ b_1...b_{25}\ v_{1,1...}v_{1,25}\ d_{1...}d_{25}]$$
$$population \quad = \quad [gen\,(\,1\,); .....; gen\,(\,M\,)] \tag{4.8}$$

Where M is the number of the total population size, i=1,2,.....M

4.6.2.2 Matrix encoding strategy

In this encoding strategy, every particle is encoded for a matrix. We also take the MLP structure for IEEE 14-bus system, 1–25–2 for an example, the encoding strategy can be written as:

$$W = (w_{1,1}\ w_{1,2.....}w_{1,25})\,,\ b = \begin{pmatrix} b_1 \\ b_2 \\ : \\ b_{25} \end{pmatrix}, v = \begin{pmatrix} v_{1,1}\ v_{1,2.....}v_{1,25} \\ \\ v_{2,1}\ v_{2,2.....}v_{2,25} \end{pmatrix},\ d = \begin{pmatrix} d_1 \\ d_2 \end{pmatrix} \tag{4.9}$$

Where W denotes the input weight matrix, v denotes the output weight matrix, while b denotes the bias matrix in hidden layer, and d denotes the output bias matrix. The total number of points created to establish MLP for IEEE 14-bus system equals 102, so these points are encoded into weights and biases matrices represented by rows $\times$ columns, w=[1×25], b=[25×1], v=[2×25], d=[2×1]. Moreover, MLP structure for IEEE 118-bus system consists of 1-60-2. The total number of points created to establish MLP for IEEE 118-bus system equals 242, so these points are encoded into weights and biases matrices, w=[1×60], b=[60×1], v=[2×60], d=[2×1].

4.6.3 The MLP as a fitness function of SFS

After encoding all the initial parameters into neural network parameters (weights and thresholds), the input layer i has n nodes, the hidden layer j has H hidden nodes, and the output layer k has O output nodes. The fitness function is given by assuming the hidden transfer function is Hyperbolic tangent sigmoid function, and the output transfer function is linear based on the best selected control parameters shown in Table (4.3) for IEEE 14-bus system and Table (4.4) for IEEE 118-bus system.

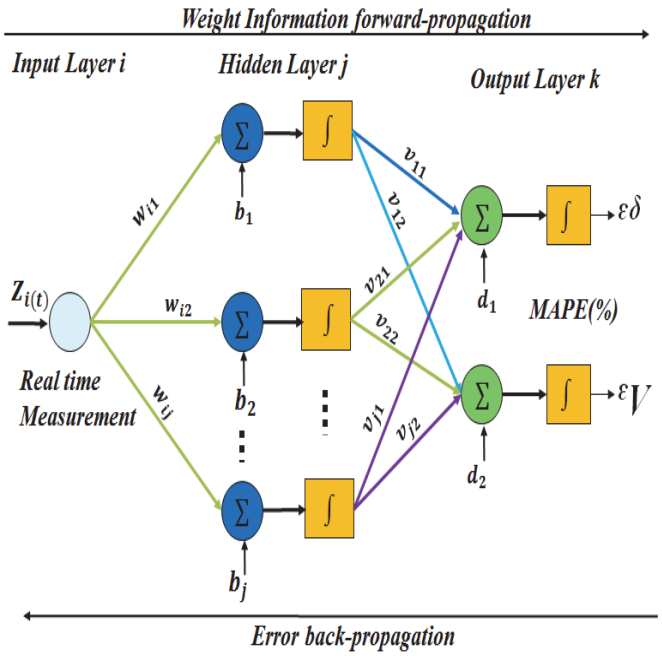

Figure 4.1: Multilayer Perceptron Neural network model

The input layer

$$z_i = [z_1,\ z_2\ , ... \ z_n] \tag{4.10}$$

The hidden layer input S_j ,

$$s_{j=} \sum_{i=1}^{n} w_{ij} \times z_i - b_j \tag{4.11}$$

The hidden layer output y_j,

$$y_i = f(s_j) \tag{4.12}$$

$$y_{i=} f(\sum_{i=1}^{n} w_{ij} \times z_i - b_j) \tag{4.13}$$

$$y_i = \frac{1}{1 + exp(-(\sum_{i=1}^{n} w_{ij} \times z_i - b_j))}, j = 1, 2, ...H \tag{4.14}$$

The computation of the output node to the input,

$$\begin{aligned} &u_{k=}f(\textstyle\sum_{j=1}^{H} v_{kj} \times y_j - d_k) \\ &\text{,k=1,2,...o} \end{aligned} \tag{4.15}$$

The sum of square error (MSE) of the output node is calculated,

$$MSE(e_k) = \frac{1}{n}\sum_{k=1}^{n}(u_k - t)^2 \tag{4.16}$$

Where z denotes real time available measurements include real power injections and flows, $MSE(e_k)$ denotes the mean square error (MSE) of the output node, t denotes the desired output (target), system states of the optimal power flow and u_k denotes the output vector of the kth layer. The least mean square of the output error will provide us with the optimal weights and thresholds.

4.6.4 Diffusion Process

While the number of generations are less than the maximum, the variables of the fitness function, Multilayer Perceptron (weights and thresholds) will be diffused around its current position to obtain the best points of weights and thresholds (BP) among all points (weights and thresholds) of the search space.

4.6.5 Ranking

All weights and thresholds obtained from diffusing process are ranked. This rank will be based on the output error of the fitness function in Eq. (4.16). The ranking will be based on the probability, so the best weights and thresholds among others will have high probability. This technique will increase the possibility of computing optimal weights and thresholds in the next iteration.

4.6.6 Updating Process

After ranking all weights and thresholds. These weights and thresholds will be updated and ranked again to the output error of the fitness function (MLP). If the

condition, the ranking value $\leq$ MSE(e_k) is satisfied, then second update is used to improve the quality of the search space and satisfy the diversification property. Otherwise, no update will be occurred, and second update process will not be represented. The reference [82] provides in detail the SFS technique.

4.6.7 Training Termination of the MLP Based SFS process

The training process terminates if the maximum number of generations is reached and the least mean square error of the fitness function (MLP-SFS) is determined, otherwise it returns to step 5.5.2. The optimal MLP neural network parameters (weights and thresholds) will be used as the initial weights and thresholds to the simple MLP program. The backpropagation algorithms shown in figures 4.3 and 4.4 will train the optimized and non-optimized neural network parameters (weights and thresholds), then the comparison in terms of accuracy and least computational time will be computed. The following equations are used to compute the comparison of non-optimized MLP and optimized MLP based on SFS,

$$\varepsilon_{\theta i}(k) = \frac{1}{n_b - 1} \sum_{i=2}^{n_b} \mid \frac{\theta_i(k) - \theta_i^T(k)}{\theta_i^T(k)} \mid \times 100 \tag{4.17}$$

$$\varepsilon_{vi}(k) = \frac{1}{n_b} \sum_{i=1}^{n_b} \mid \frac{v_i(k) - v_i^T(k)}{v_i^T(k)} \mid \times 100 \tag{4.18}$$

Where $\varepsilon_{\theta i}(k)$ and $\varepsilon_{vi}(k)$ denote the mean absolute percentage error for voltage angles and magnitude voltages at bus i and time step k respectively. $\theta_i(k)$ and $\theta_i^T(k)$ denote the estimated and actual voltage angles at bus i respectively. $V_i(k)$ and $V_i^T(k)$ denote the estimated and actual magnitude voltage at bus i respectively. n_b denotes the number of buses.

4.7 Results and discussion

The experiment has been applied on IEEE 14-and 118-bus systems which was published by the official website of University of Washington [113]. New York Independent System Operator (NYISO) has been combined with IEEE 14-bus system. In IEEE

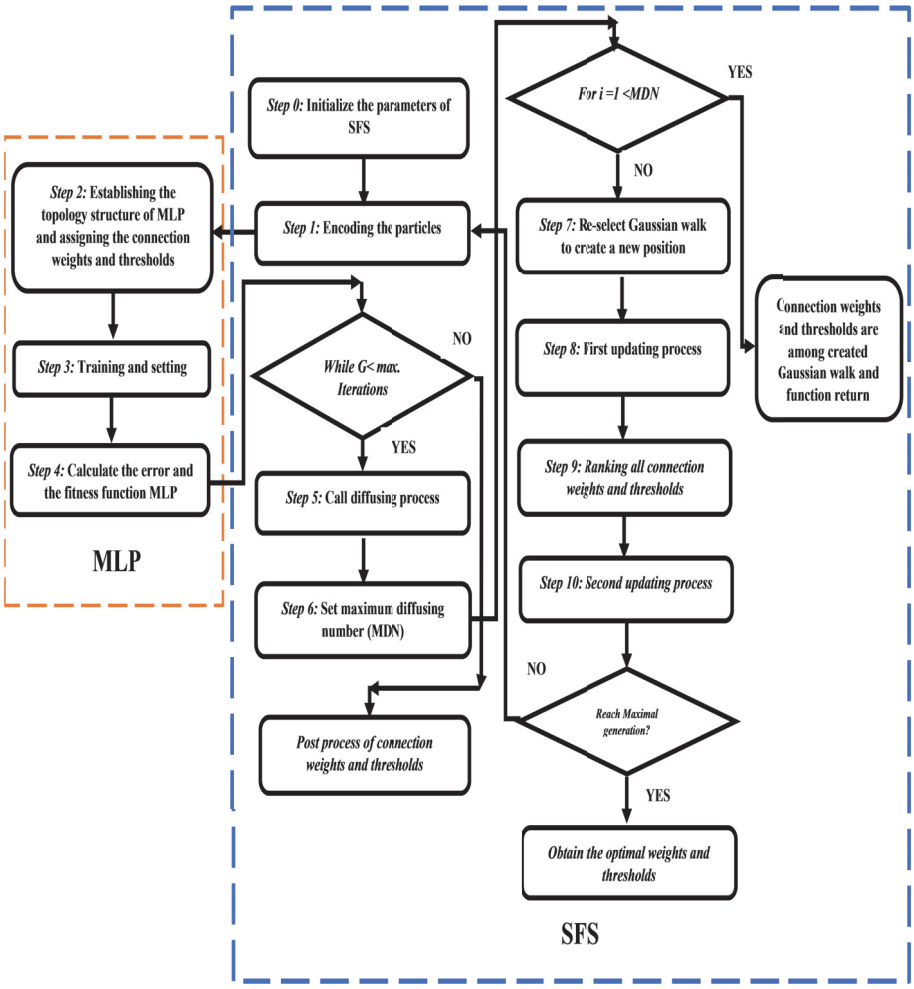

Figure 4.2: Flowchart of the hybrid MLP based SFS technique

118-bus system, we take the 11 regions of NYISO data and distribute these regions arbitrarily, then we compute the proportional rate by dividing the total normalized load capacity of IEEE 118-bus system with NYISO data over the total load capacity of IEEE 118-bus system without NYISO data. Then the proportional rate will be applied to each bus load to be updated after each hour load forecasting. Due to the lack of the hourly historical data of reactive power load, we assume that power factor=0.8 according to the National Electric Manufacturing Association (NEMA), so the reactive power load will be computed by applying active power load to the power factor. Support vector regression (SVR) in time series is used to forecast the hourly load for six hours ahead using hourly historical data of NYISO. Optimal power flow is utilized to serve the forecasted load at the least generation cost and to convert

the loads and generations into real time measurements include active and reactive power injections and flows and system state (phase and magnitude voltage). White Gaussian noise (WGN) is added with zero mean and standard deviation to the real-time measurement (real and reactive power injections and flows). Then, stochastic fractal search technique is used to find the optimal MLP neural network parameters (weights and thresholds) as shown in figure 4.1. The final parameters of hybrid MLP-SFS will be trained by simple backpropagation neural network (BPN) to obtain the mean absolute percentage error (MAPE) for system state. The hybrid technique (MLP-SFS) results compared with other techniques in case of accuracy and the computational time. Ten different computations have been made to select the best control parameters of the MLP neural network based on stochastic fractal search technique (SFS) as shown in Tables 4.3 and 4.4. In each computation, we adjust the lower and upper boundaries of neural network parameters (weights and thresholds), then we set the transfer function of the hidden layer and the activation function of the output layer. Then, we vary both of population size from 10 to 200 and number

Table 4.1: Control Parameters used in Optimized MLP based SFS for IEEE 14-bus system combined with NYISO load data

Architecture	Multilayer perceptron
Learning algorithm	Feed-forward
Population Size	50
Number of Neurons	25
Input Transfer Function	Tang-Sigmoid
Output Activation Function	Linear
Upper and Lower boundaries	[1, -1]
Maximum Diffusion Number (MDN)	1
Maximum Generation	150
Side Walk	0

of neurons from 1 to 50 to obtain the least mean square error (MSE) in the fitness function. Number of neurons should not exceed 50 in IEEE 14-bus system and should not exceed 90 in IEEE 118-bus system as well to prevent ending up with overfitting. Increasing the number of neurons and population size can affect the computational time because it takes the proposed technique (MLP-SFS) long time to terminate the process. On contrary, increasing number of neurons and population size can enhance

the accuracy.

Table 4.2: Control Parameters used in Optimized MLP based SFS for IEEE 118-bus system combined with NYISO load data

Architecture	Multilayer perceptron
Learning algorithm	Feed-forward
Population Size	70
Number of Neurons	60
Input Transfer Function	Tang-Sigmoid
Output Activation Function	Linear
Upper and Lower boundaries	[1, -1]
Maximum Diffusion Number (MDN)	10
Maximum Generation	550
Side Walk	1

In IEEE 118-bus system, we increase the maximum diffusion walk to 10 instead of 1 as it used for IEEE 14-bus system because the increment of maximum diffusion walk can reduce the number of generations and subsequently the computational time will be reduced, however, it will not have any impact on the accuracy. Tables 4.1 and Table 4.2 provide the best control parameters used in hybrid MLP-SFS to start the experiment. It is concluded that computation number eight has the best control parameters of MLP-SFS for both IEEE-14 and 118-bus systems. In IEEE 14-bus system, the presented MLP-SFS consists of 50 for population size, 25 for neurons, [-1,1] for the lower and the upper bound, and Hyperbolic tangent transfer function for the hidden layer and linear activation function for the output layer. These control parameters have provided the least mean square error (MSE)=1.82E-9 computed by Eq. (4.17) and the least computational time=1.05 seconds. Moreover, in IEEE 118-bus system, the presented MLP-SFS consists of 70 for population size, 60 for neurons, [-1,1] for the lower and the upper bound, and Hyperbolic tangent transfer function for the hidden layer and linear activation function for the output layer. These control parameters have provided the least mean square error (MSE)=1.39E-1 computed by Eq. (4.17) and the least computational time=16.89 seconds. The implementation and control parameters of both MLP-GA and MLP-PSO in details are provided in Ref. [102], [103] and Ref. [104], [105] respectively.

Table 4.3: Selecting the best control parameters used in Optimized MLP based SFS for IEEE 14 Bus System combined with NYISO load data in case of accuracy and CPU Time

No.	**Population Size**	**No. Of Neurons**	**Lower & Upper bound**	**Hidden layer**	**Output layer**	**MSE**	**Time(s)**
1	150	50	[-50,50]	Tan-sigmoid	Tan-sigmoid	$1.35\times10^{-}4$	5.91
2	100	25	[-40,40]	Linear	Linear	$6.08\times10^{-}1$	20.7
3	100	25	[-30,30]	Log-sigmoid	Log-sigmoid	$1.57\times10^{-}3$	2.6
4	90	25	[-25,25]	Tan-sigmoid	Log-sigmoid	$1.92\times10^{-}3$	1.39
5	90	25	[-20,20]	Log-sigmoid	Linear	$7.31\times10^{-}5$	6.6
6	70	25	[-10,10]	Log-sigmoid	Tan-sigmoid	$8.57\times10^{-}2$	3.35
7	50	25	[-0.5,0.5]	Tan-sigmoid	Log-sigmoid	$1.59\times10^{-}4$	10.6
8	50	25	[-1,1]	Tan-sigmoid	Linear	$1.82\times10^{-}9$	1.05
9	50	25	[-2,2]	Linear	Tan-sigmoid	$1.57\times10^{-}1$	16.5
10	50	25	[-5,5]	Linear	Log-sigmoid	$1.60\times10^{-}1$	20.1

Table 4.4: Selecting the best control parameters used in Optimized MLP based SFS for IEEE 118 Bus System combined with NYISO load data in case of accuracy and CPU Time

No.	**Population Size**	**No. Of Neurons**	**Lower & Upper bound**	**Hidden layer**	**Output layer**	**MSE**	**Time(s)**
1	70	35	[-50,50]	Tan-sigmoid	Tan-sigmoid	1.4	53.81
2	50	40	[-40,40]	Linear	Linear	1.54	28.89
3	50	40	[-30,30]	Log-sigmoid	Log-sigmoid	1.4	30.3
4	50	35	[-25,25]	Tan-sigmoid	Log-sigmoid	1.41	30.85
5	50	45	[-20,20]	Log-sigmoid	Linear	$3.89\times10^{-}1$	26
6	70	35	[-10,10]	Log-sigmoid	Tan-sigmoid	8.57×10^{2}	33.85
7	50	55	[-0.5,0.5]	Tan-sigmoid	Log-sigmoid	2	33.68
8	70	60	[-1,1]	Tan-sigmoid	Linear	$1.39\times10^{-}1$	16.89
9	50	30	[-2,2]	Linear	Tan-sigmoid	1.42	21.69
10	60	40	[-5,5]	Linear	Log-sigmoid	1.6	37.02

Figures 4.3 and 4.4 illustrate the ranking of Backpropagation neural network (BPN) algorithms in case of robustness. The robust performance is exhibited in the least mean absolute percentage error (MAPE) computed by Eqs. (4.18) and (4.19) and the least computational time. The robust algorithm will be utilized as a part of the simple MLP program to train the optimized neural network parameters (weights and thresholds). In IEEE 14-bus system, Levenberg–Marquardt algorithm (LM) provided MAPE (voltage angle) =2.08%, 0.17% for MAPE (Magnitude voltage), and 1.33 seconds compared to Gradient descent with momentum (GDM), MAPE (voltage angles) =7.91%, 2.95% for MAPE (magnitude voltages), and 10.15 seconds. In addition, in IEEE 118-bus system, Levenberg–Marquardt algorithm (LM) provided MAPE (voltage angle) =3.75%, 1.64% for MAPE (Magnitude voltage), and 4.75 seconds compared to Gradient descent (GD), MAPE (voltage angles) =17.11%, 9.87% for MAPE (magnitude voltages), and 12.45 seconds. Also, in IEEE 118 bus system, BFGS Quasi-Newton algorithm (BFG) has the longest computational time =111.2 seconds. It is confirmed from the above discussion that the LM algorithm shows reliable and robust performance. Both GDM and GD algorithms have ranked as the worst performance for IEEE 14-and 118-bus systems respectively.

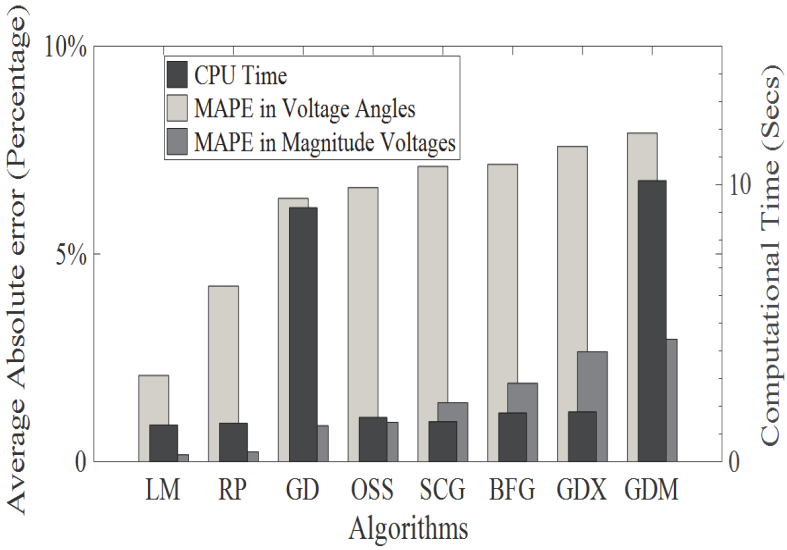

Figure 4.3: Ranking Training Algorithm of Feed Forward Neural Network Using MLP-SFS SFS for IEEE 14-bus system in case of accuracy and Computational time (Secs)

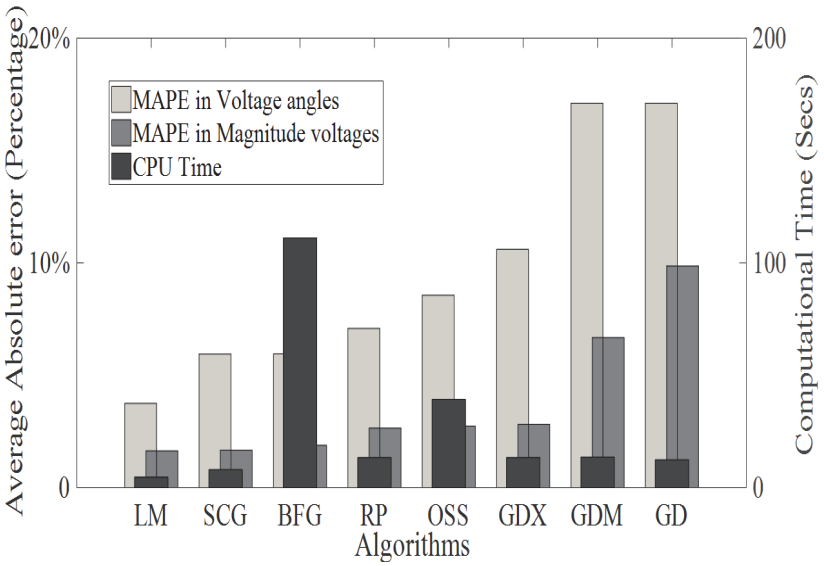

Figure 4.4: Ranking Training Algorithm of Feed Forward Neural Network Using MLP-SFS for IEEE 118-bus system in case of accuracy and Computational time (Secs)

4.7.1 No Communication Failure

State estimation SE is mostly performed as there is no communication failure where all different sensors distributed on the electric power grid work perfectly by transmitting real time measurements include active and reactive power injections and flows from RTUs to control centers with no delay. Then, system states are determined and electric grid status is evaluated. IEEE 14-bus system with no communication failure consists of measurement vector Z(t) ϵ R^{68x1} at step time t, 2×14 for active and reactive power injections and 2×20 for active and reactive power flows.

Figures 4.5 and 4.6 demonstrate the comparison in mean absolute phase and voltage percentage error (MAPE) respectively between non-optimized Multilayer Perceptron (MLP) parameters (random weights and thresholds) and optimized Multilayer Perceptron (MLP) parameters (optimal weights and thresholds) based on Stochastic Fractal Search technique (SFS), Genetic algorithm (GA), and Particle Swarm Optimization (PSO) for IEEE 14-bus system. As measurement error increases from

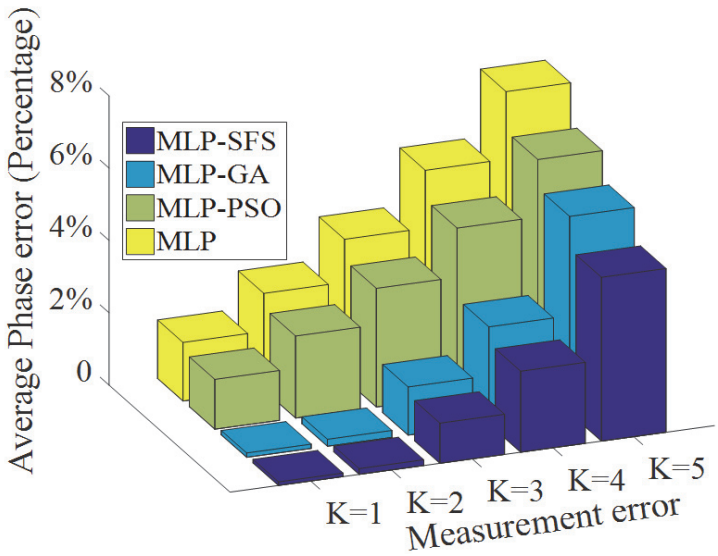

Figure 4.5: Average Phase Percentage error comparison at different measurement error with no communication for IEEE 14-bus system

k=1 to k=5, the mean absolute percentage error is incremented. For example, at k=1, MAPE (voltage angles) = 1.62% for MLP compared to 1.38% for MLP-PSO, 0.15% for MLP-GA, and 0.10% for MLP-SFS. Also, MAPE (magnitude voltages) = 0.92% for MLP compared to 0.66% for MLP-PSO, 0.12% for MLP-GA, and 0.06% for MLP-SFS. At k=5, MAPE (voltage angles) = 7.33% for MLP compared to 6.23% for MLP-PSO, 5.44% for MLP-GA, and 3.53% for MLP-SFS. Also, MAPE (magnitude voltages) = 4.59% for MLP compared to 3.31% for MLP-PSO, 3.04% for MLP-GA, and 2.60% for MLP-SFS. As can be seen that optimized MLP based SFS has the least variation. It is concluded from the results that DSE in Optimized MLP parameters (weights and thresholds) based on SFS technique improve accuracy and convergence rate of system state even with full communication.

Further, a complete comparison between optimized and non-optimized Multilayer Perceptron NN parameters by SFS technique is tabulated in tables 4.5 and 4.6. Mean absolute percentage error in system states (magnitude voltages and angles) increases due to the variation in measurement error from k=1 to k=5. As can be observed from the tables at the maximum measurement error k=5, the optimized MLP based SFS

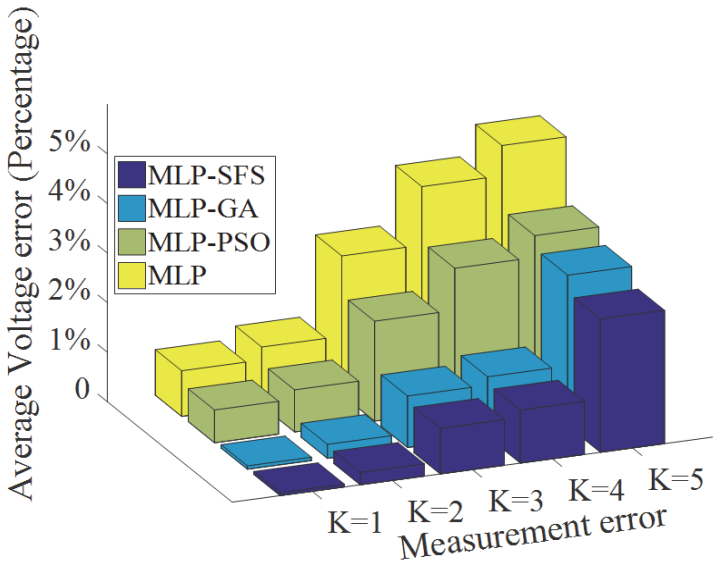

Figure 4.6: Average Voltage Percentage error comparison at different measurement error with no communication for IEEE 14-bus system

Table 4.5: Non-optimized MLP state estimation error for IEEE 14-bus system at different measurement error under no communication failure

	ε_θ			ε_V			**CPU Time**
k	Mean	Max.	Std.	Mean	Max.	Std.	(seconds)
1	1.62	3.05	0.24	0.92	1.93	0.15	1.38
2	2.67	7.66	0.69	1.18	5.87	0.18	1.41
3	3.86	11.5	1.13	2.8	7.81	0.32	1.68
4	5.46	16.4	1.58	3.98	10.1	0.68	1.72
5	7.33	21.8	2.02	4.59	14	0.87	3.05

provides mean absolute percentage error in voltage angles=4.53% compared to non-optimized MLP, mean absolute percentage error in voltage angles=7.33%. Also, maximum absolute percentage error in voltage angles=9.78% for Optimized MLP based SFS compared to 21.08% for non-optimized MLP. Moreover, the optimized MLP based SFS gives mean absolute percentage error in magnitude voltage =2.69%, compared to 4.59% for non-optimized MLP. In addition to that, maximum absolute percentage error in magnitude voltage=6.64% compared to non-optimized MLP=14.0%. It is concluded that the optimized MLP based SFS has the least variation. Computational time is also taken and tabulated in tables 4.5 and 4.6. For example, at k=5, the CPU time of the proposed technique (MLP-SFS) =1.98 seconds compared

Table 4.6: Optimized MLP-SFS state estimation error for IEEE 14-bus system at different measurement error under no communication failure

	ε_θ			ε_V			**CPU Time**
k	Mean	Max.	Std.	Mean	Max.	Std.	(seconds)
1	0.1	0.74	0.1	0.06	0.17	0.08	1.32
2	0.16	2.93	0.38	0.25	1.16	0.17	1.35
3	1.11	4.9	0.66	0.92	2.5	0.27	1.44
4	2.24	7.54	1.03	1.07	3.4	0.35	1.63
5	3.53	9.78	1.22	2.6	6.64	0.43	1.98

to CPU time of the non-optimized MLP =3.05 seconds. It is confirmed from the tables that the proposed technique (MLP-SFS) has enhanced the multilayer Perceptron in both precision and least computational time. Figures 4.7 and 4.8 show a

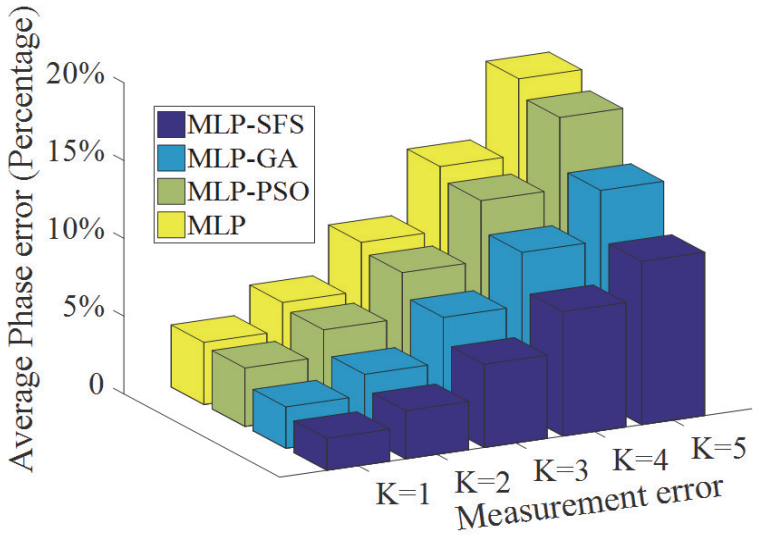

Figure 4.7: Average Phase Percentage error comparison at different measurement error with no communication for IEEE 118-bus system

comparison between non-optimized and optimized Multilayer Perceptron parameters (weights and thresholds) based on SFS, GA, and PSO techniques for IEEE 118-bus system. For example, at k=1, MAPE (voltage angles) = 4% for MLP compared to 3.76% for MLP-PSO, 2.67% for MLP-GA, and 2.05% for MLP-SFS. Also, MAPE (magnitude voltages) = 2.93% for MLP compared to 2.64% for MLP-PSO, 1.28% for MLP-GA, and 1.05% for MLP-SFS. At k=5, MAPE (voltage angles) = 18.01% for

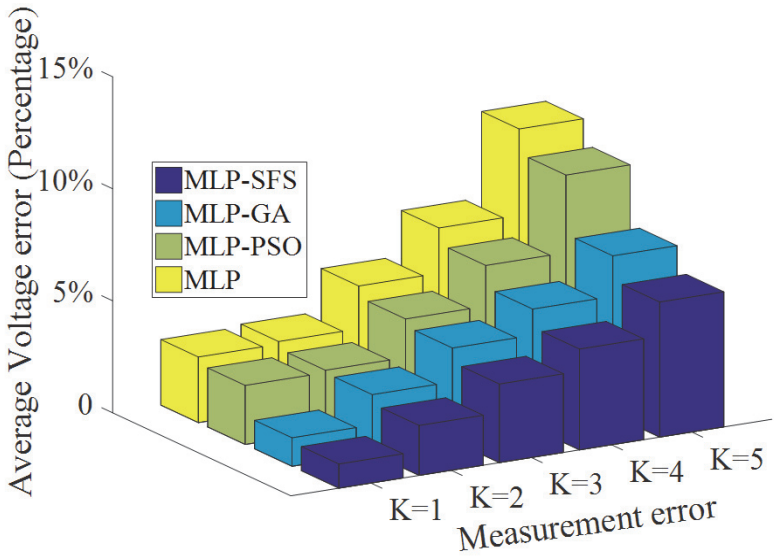

Figure 4.8: Average Voltage Percentage error comparison at different measurement error with no communication for IEEE 118-bus system

MLP compared to 16.93% for MLP-PSO, 13.65% for MLP-GA, and 1.05% for MLP-SFS. Also, MAPE (magnitude voltages) = 10.83% for MLP compared to 9.75% for MLP-PSO, 7.12% for MLP-GA, and 5.03% for MLP-SFS.It is concluded that DSE in Optimized MLP parameters based on SFS technique has the least variation. Also, it can be observed from figures that MLP based PSO is sensitive to the noise variation from k1 to k5. In contrast with GA and SFS, PSO is ranked as the worst technique used to optimize MLP parameters. Table 4.7 shows the optimal power flow (OPF) for six hours ahead forecasting in full communication failure. After six hours ahead forecasting using support vector machine (SVM), optimal power flow will be determined to serve the forecasted load for each hour ahead forecasting at the minimum cost. It can be observed that there is a small increment in the generators during six hours ahead forecasting; for example, the main generator P1 increases from 216.08MW to 216.25MW leads to increase the generator cost from 9457.83\$/h to 9778.54\$/h. This increment in the generator leads to have small increment in power losses from 21.88MW to 22.93MW. There is no specific trend in computational time. Table 4.8 shows computational time comparison at measurement error k=5 under no communication failure. In IEEE 14 bus system, the computational time for the proposed

Table 4.7: Optimal Power flow for IEEE 14-bus system combined with NYISO load data under no communication failure in six hours ahead forecasting

	One-hour-ahead	Two-hour-ahead	Three-hour-ahead	Four-hour-ahead	Five-hour-ahead	Six-hour-ahead
P1 (MW)	216.08	216.11	216.16	216.19	216.25	216.25
P2 (MW)	39.45	39.45	39.46	39.47	39.48	39.48
P3 (MW)	18.98	19.3	19.73	19.89	20.2	20.4
P6 (MW)	17.7	17.84	18.33	18.34	18.81	18.92
P8 (MW)	11.84	11.8	13.48	13.59	15.32	16.97
Cost($/hr.)	9457.83	9476.14	9583.27	9595.58	9699.43	9778.54
Losses (MW)	21.88	21.91	21.97	22.08	22.91	22.93
CPU(sec.)	1.26	1.33	1.27	1.28	1.26	1.27

Table 4.8: Computational time comparison at measurement error k=5 under no communication failure

Bus Systems	MLP	MLP-PSO	MLP-GA	MLP-SFS
IEEE-14	3.05	2.97	2.36	1.98
IEEE-118	18.67	18.11	16.73	9.67

technique (MLP based SFS), 1.98 seconds compared to non-optimized MLP, 3.05 seconds, MLP based PSO, 2.97 seconds, and 2.36 seconds for MLP-GA. In IEEE 118 bus system, the computational time for the proposed technique (MLP based SFS), 9.67 seconds compared to non-optimized MLP, 18.67 seconds, MLP based PSO, 18.11 seconds, and 16.73 seconds for MLP-GA. It concluded that MLP based SFS has the least computational time.

4.7.2 Sustained communication failure

SCADA system should receive data at time interval from 2 to 10 seconds, so one time step interval of five minutes delay which is somehow is not very practical in increasingly interconnected power grid. Hence a more intensive investigation of communication delay/failure is required. In this case study, we begin with an examination of sustained single bus communication failure. When sustained communication, failure takes a place at a single station, we assume that the real-time measurements

include active and reactive power injections and magnitude voltage at this single bus are lost. However, measurements include active and reactive power flows can easily be measured at the other end of transmission line. Sustained communication failure of six hours ahead is examined. The unavailable data of maintained single bus communication failure for six hours ahead is not very practical because the system operator will respond to this failure within one to two hours after the communication failure occurred. In any case, the motivation behind this contextual analysis is to test the ability of the proposed approach. During communication failure, the information from the failure bus will not be transmitted to the control center, so six hours ahead load forecasting using support vector machine (SVM) will be utilized to estimate the variation of this bus load until the communication failure fixed. Table 4.9 shows

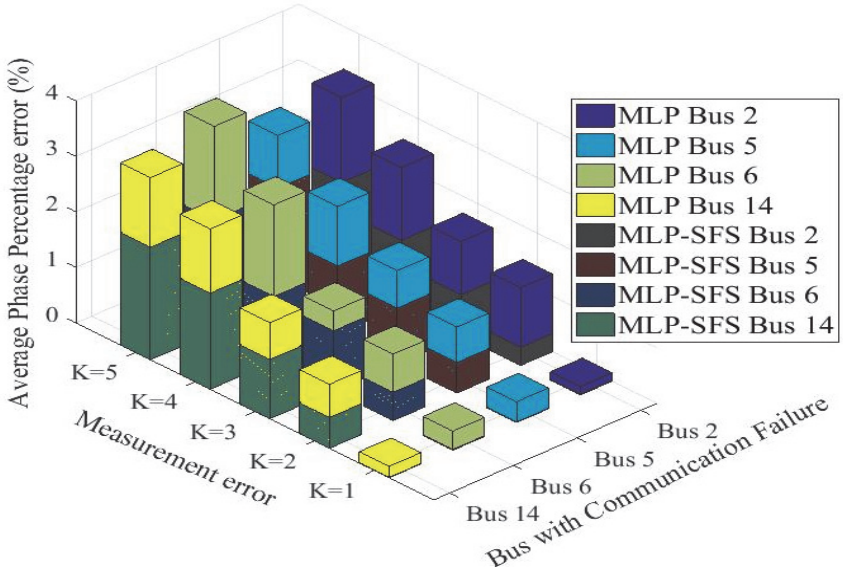

Figure 4.9: Average Phase Percentage error comparison at different measurement error with single station communication failure for IEEE 14-bus system. X-axis: measurements error. Y-axis: bus with communication failure. Z-axis mean phase Percentage error comparison

load forecasting for six hours ahead in several failure buses or regions. Mean absolute percentage error is used to observe the behavior of load pattern during the six hours ahead. In each hour load forecasting, we can see that mean absolute percentage error is increased due to the increment in the error in each hour ahead forecasting, then optimal power flow will be applied to serve the forecasted load at the least generation

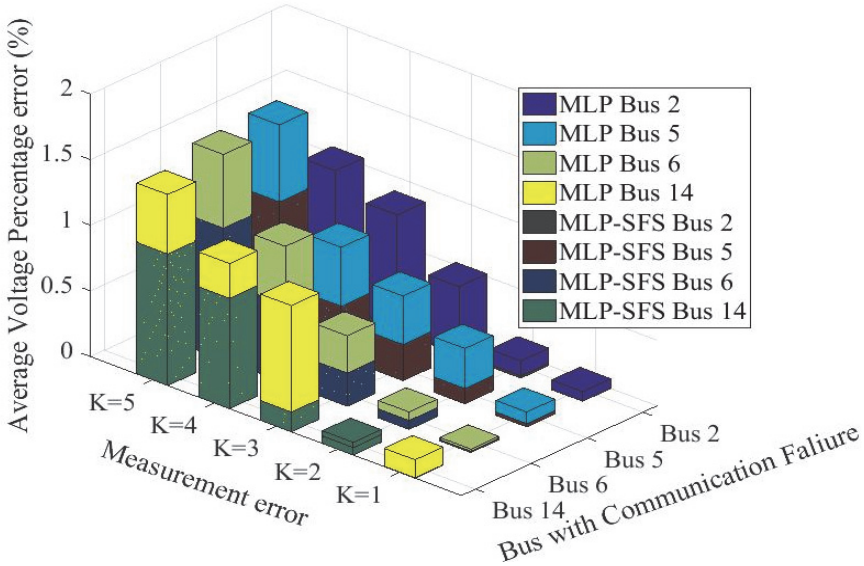

Figure 4.10: Mean Voltage Percentage error comparison at different measurement error with single station communication failure for IEEE 14-bus system. X-axis: measurements error. Y-axis: bus with communication failure. Z-axis mean voltage percentage error comparison

cost. Two 3-D plots that analyze and compare the performance of the non-optimized MLP and optimized MLP based SFS for the IEEE 14-bus system combined with NY-ISO load data are represented in Figures 4.9 and 4.10. Overall, it is obvious that the optimized MLP parameters based SFS performs superior than non-optimized MLP when there is a maintained single bus communication failure.

Figures 4.11 and 4.12 show a comparison in MAPE for system states between optimized MLP based SFS and non-optimized MLP on single bus communication failure, bus 59 in IEEE 118-bus system. Bus 59 is selected due to the highest active load=277MW and reactive load=113 MVAR. As measurement error varies from $k = 1$ to $k = 5$, the MAPE for system states is increased. It is concluded that optimized MLP based SFS is more robust.

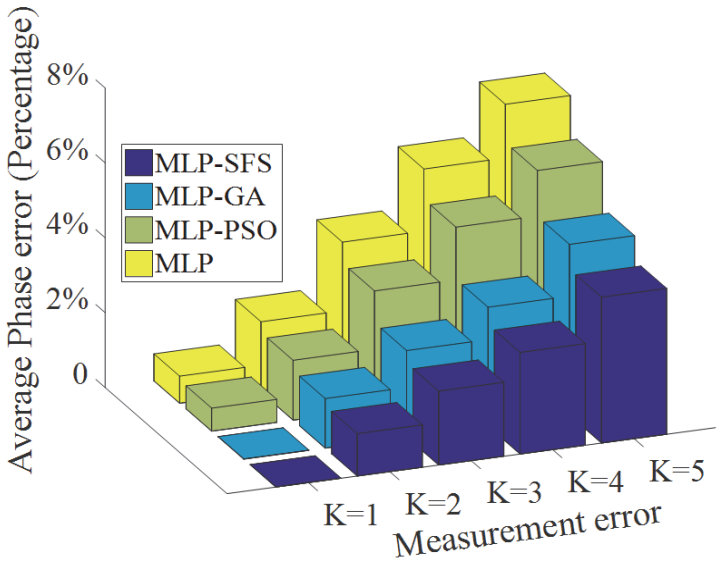

Figure 4.11: Mean Phase Percentage error comparison at different measurement error with single station communication failure on bus 59 for IEEE 118-bus system

Table 4.9: Mean Absolute Percentage error for IEEE 14-bus system combined with NYISO loads in sustained communication failure in six hours ahead forecasting

Bus No.	Regions	One-hour-ahead	Two-hour-ahead	Three-hour-ahead	Four-hour-ahead	Five-hour-ahead	Six-hour-ahead
Bus 2	Central	2.3×10^{-1}	3.2×10^{-1}	1.26	2.89	3.79	3.89
Bus 5	Hud VI	1.85	1.47	3.49	4.45	4.49	4.43
Bus 6	Longil	7.9×10^{-1}	7.5×10^{-1}	2.8	3.97	4.35	4.59
Bus 14	West	9.5×10^{-1}	1.01	1.17	1.58	1.83	1.99

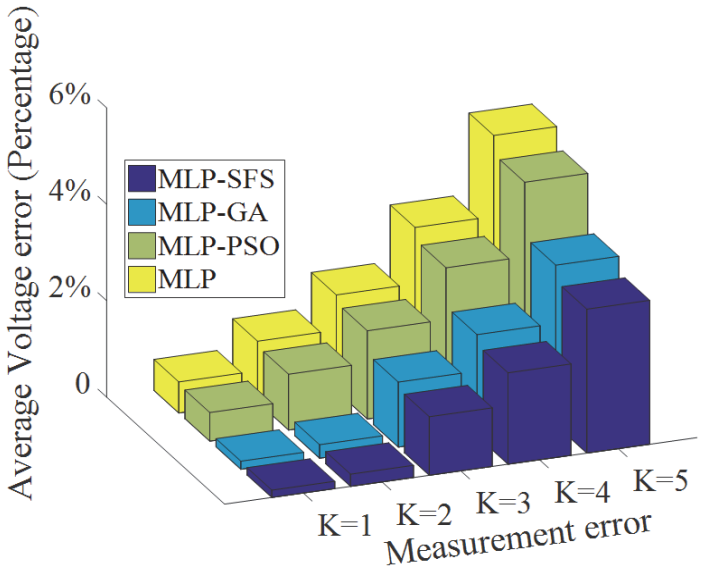

Figure 4.12: Mean Voltage Percentage error comparison at different measurement error with single station communication failure on bus 59 for IEEE 118-bus system

4.7.3 Communication failure Duration

Communication failure duration is addressed the effectiveness of the restoration time on the electric power grid. The case study is applied to bus 2 in IEEE 14-bus system, Central region in NYISO load data with several failure duration at measurements error k=1.

Figures 4.13 and 4.14 show the mean percentage estimation error in both phase and magnitude voltage when the delay duration varies from 5 min to six hours for IEEE 14-bus system combined with NYISO load data. Real and reactive power injection variation during six hours ahead forecasting is plotted in the same figures of 4.13 and 4.14 respectively to observe the trend relation between mean percentage phase and voltage error in both non-optimized MLP and optimized MLP based SFS. The injected real and reactive power calculated as $Z_{injections}(t) - Z_{injections}(t-d)$, where d is the time delay and $Z_{injections}(t)$ denotes the real and reactive power injections at time step t.

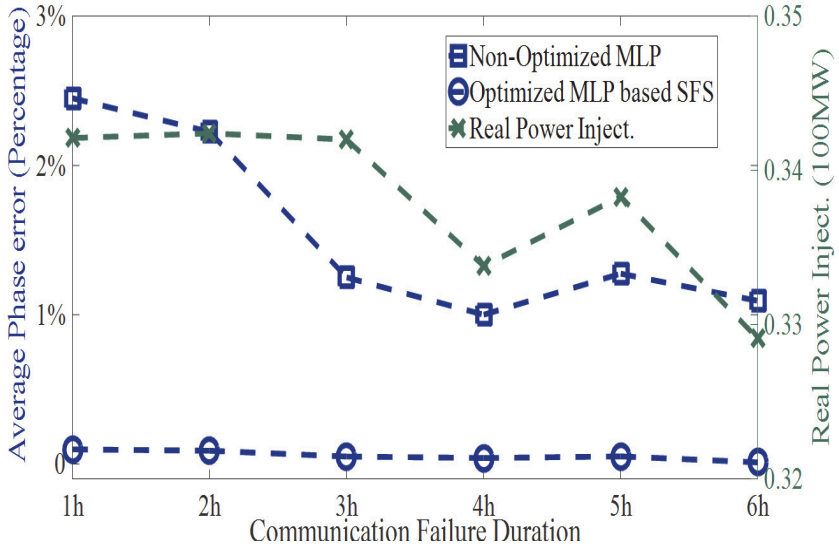

Figure 4.13: Mean Phase Percentage error comparison at different failure duration with single station communication failure for IEEE 14-bus system. X-axis communication failure duration. Y-axis Real Power Injection. Y-axis Mean Phase Percentage error

As can be seen from figures 4.13 and 4.14 that non-optimized MLP, the trend of MAPE is similar to the hourly trend of injected real and reactive power. However,

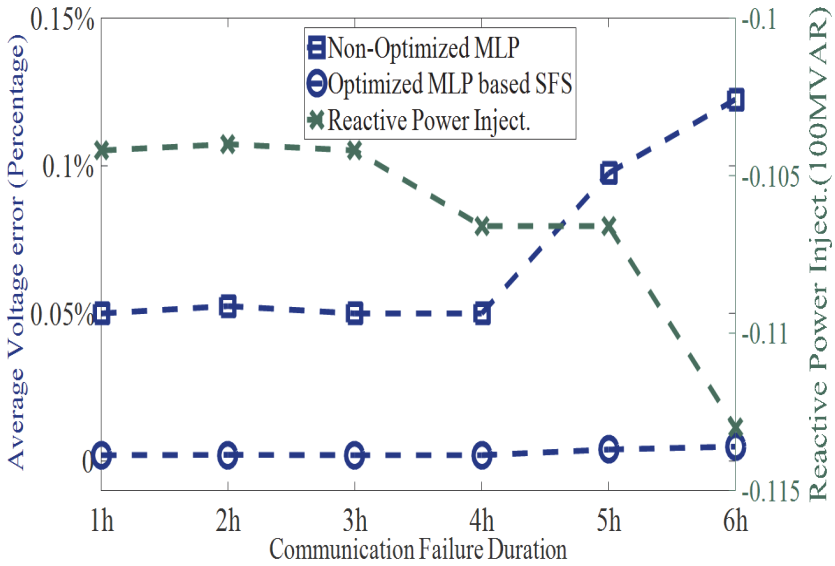

Figure 4.14: Mean Voltage Percentage error comparison at different failure duration with single station communication failure for IEEE 14-bus system. X-axis communication failure duration. Y-axis Reactive Power Injection. Y-axis Mean Voltage Percentage error

the optimized MLP based SFS has no specific hourly trend in both mean voltage phasors percentage estimation error due to the un-utilized delayed data in MLP-SFS. It is concluded that the presented approach MLP based SFS has the least variation error.

4.7.4 Multi-Step ahead forecasting

Radial basis function (RBF) and Support Vector Regression (SVR) are used in multi-step ahead load forecasting to forecast the NYISO load for IEEE 14-bus system. RBF is utilized to perform hourly multi-step ahead load forecasting due to good nonlinear approximate, faster, understandable, and easier to be implemented. RBF network is used to train a regression model by minimizing the squared error of the cost function using Broyden Fletcher Goldfarb Shanno (BFGS). The load forecasting is done by Weka software. The parameters of RBF are set as following, number of basis function=2, size of threads pool=1, the ridge penalty factors=0.01, and the tolerance parameter=1e-6.

The MAPE in system states comparison between SVM and RBF when load data

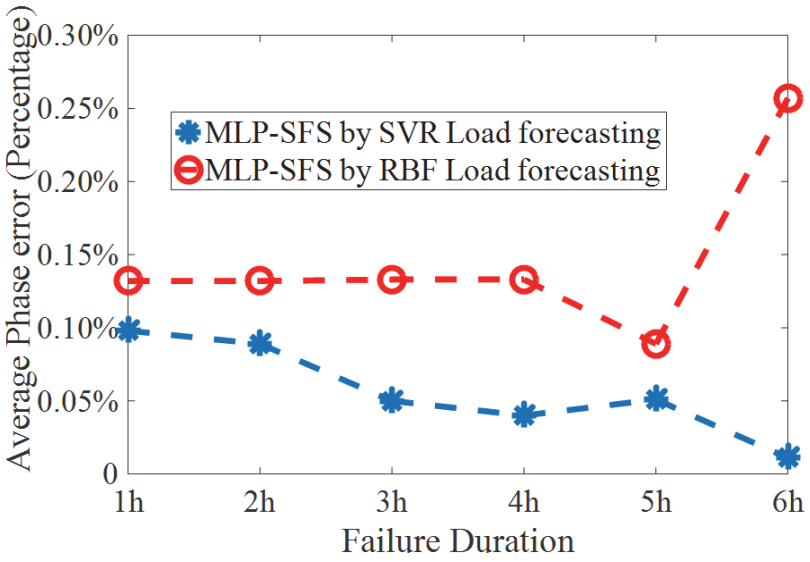

Figure 4.15: Mean Phase Percentage error comparison between SVM and RBF when load data from Bus2, Central region of the IEEE 14-bus system is not available for different failure duration

from Bus 2, Central region of the IEEE 14-bus system combined with NYISO load data is not available from 5 min to six hours ahead failure duration. The load forecasting by both SVM and RBF will be used to estimate the system states in optimized MLP based SFS. It can be seen from figures 4.15 and 4.16 that when the communication failure duration is short, the performance of RBF and SVM based on the presented approach MLP-SFS are almost the same. However, when the hourly communication failure duration is increased, this will cause the MAPE in system states to drop in SVM and to increase in RBF. It is confirmed from the results that SVR is superior compared to RBF.

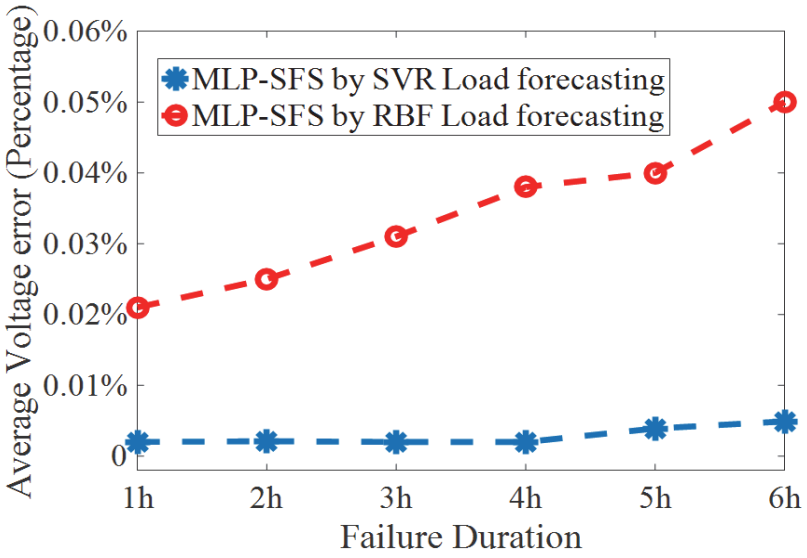

Figure 4.16: Mean Voltage Percentage error comparison between SVM and RBF when load data from Bus2, Central region of the IEEE 14-bus system is not available for different failure duration

4.7.5 Multi-station failure

The possibility of multi-station communication failure is studied. Power grids are more likely to be exposed to cyber-attacks such as Denial of service (DoS) which leads to block the transmission of real time measurements from RTUs to control centers, then measurements from multi-station will be unavailable.

Figures 4.17 and 4.18 represent a case study of communication failure at both of bus 5 and 6 together for IEEE 14-bus system combined with NYISO load data, from the figures, the proposed approach, MLP based SFS is superior compared to other techniques.

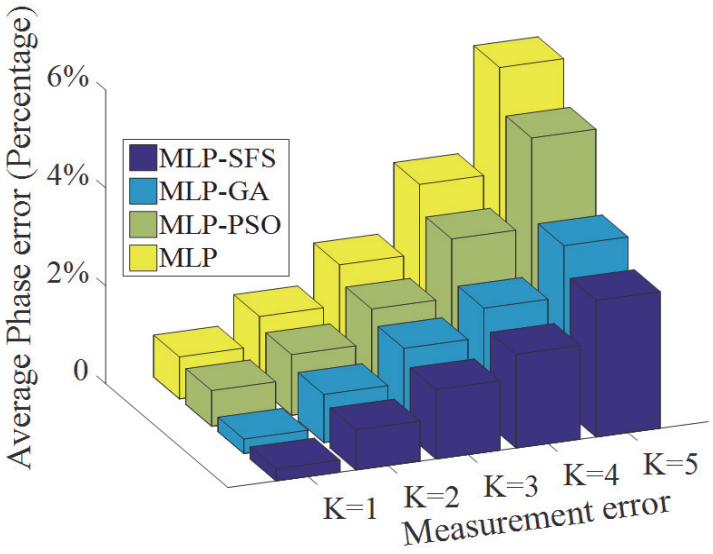

Figure 4.17: Mean Phase Percentage error comparison when both bus 5 and 6 fail at different measurement error for the IEEE 14-bus system

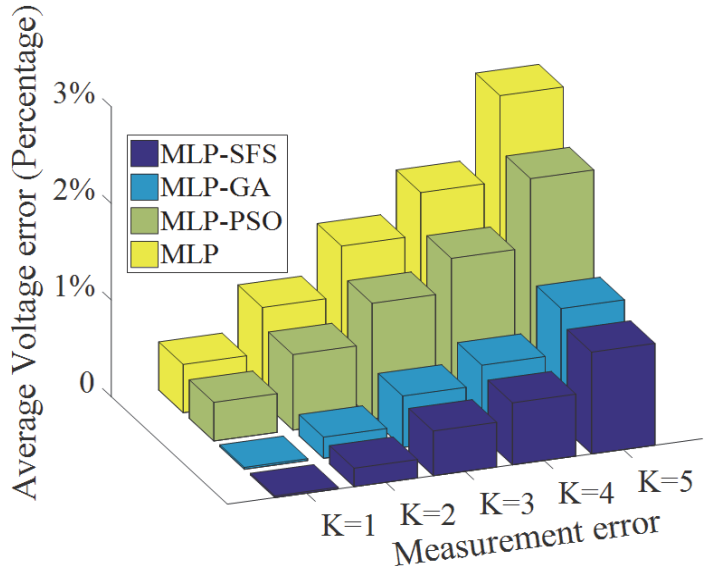

Figure 4.18: Mean Voltage Percentage error comparison when both bus 5 and 6 fails at different measurement error for the IEEE 14-bus system

4.7.6 Failure at different time of the day

The examining period of this case study is from May 1st to May 7th 2016, from Sunday to Saturday. As can be seen from table 4.9 the MAPE for system state of the non-optimized MLP and optimized MLP based PSO, GA, and SFS has the same trend of the active and reactive power load; for instance, when the active power load drops down from 32.365MW to 29.046MW during May 6th to May 7th 2016, subsequently it drops the MAPE for system state (voltage angles and magnitude voltages) from 2.502% to 2.302% and from 1.691% to 1.494% for MLP respectively. Also, MAPE for MLP-PSO. However, the proposed approach MLP based SFS is de-trended which has no hourly, daily, or weekly trend. The hybrid MLP based SFS technique provides very accurate state estimation whenever the communication failure occurs.

Table 4.10: Mean phase and magnitude voltage percentage error at failure moment in bus 14, West region for IEEE 14-bus system combined with NYISO load data

Date	Day of the Week	Time of the Day		Pload	Qload	MLP		MLP-PSO		MLP-GA		MLP-SFS	
		From	To	(MW)	(MVAR)	ε_θ	ε_V	ε_θ	ε_V	ε_θ	ε_V	ε_θ	ε_V
May 1st 2016	Sunday	0:00AM	5:00AM	28.475	2.406	2.22	1.461	1.687	1.146	0.107	0.037	0.098	0.033
May 2nd 2016	Monday	6:00AM	11:00AM	32.025	3.302	2.475	1.68	2.004	1.326	0.274	0.037	0.254	0.033
May 3th 2016	Tuesday	12:00PM	17:00PM	31.92	3.276	2.435	1.676	2.169	1.233	0.306	0.039	0.264	0.034
May 4th 2016	Wednesday	18:00PM	23:00PM	31.607	3.196	2.414	1.65	1.905	1.184	0.253	0.033	0.228	0.029
May 5th 2016	Thursday	4:00AM	9:00AM	27.707	2.212	2.17	1.42	1.844	1.056	0.068	0.039	0.057	0.034
May 6th 2016	Friday	10:00AM	15:00PM	32.365	3.388	2.502	1.691	2.126	1.344	0.321	0.039	0.268	0.034
May 7th 2016	Saturday	16:00PM	21:00PM	29.046	2.55	2.302	1.494	1.996	1.09	0.114	0.032	0.095	0.028

4.8 Conclusion

Stochastic Fractal Search technique (SFS) is applied to optimize Multilayer Perceptron NN parameters (weights and thresholds). The purpose of using SFS involves providing an enhancement in case of accuracy, convergence rate and easier operation. The SFS is used to obtain a set of appropriate initial weights and thresholds.

The hybrid MLP based SFS is presented to improve the current MLP Neural Network in terms of the accuracy and the convergence time. Simple Backpropagation Neural Network (BPN) is used to adjust the final parameters as illustrated in Figures 4.4 and 4.5. The presented approach has been compared to non-optimized MLP parameters (random weights and thresholds) and optimized MLP based on PSO and GA individually. The hybrid (MLP-SFS) technique has been tested on IEEE 14- and 118-bus systems combined with the New York Independent System Operator (NYISO) under various scenarios of communication failure and measurement error. Three different stages discussed include short-term load forecasting using support vector machine (SVM) to forecast the unavailable load due to communication failure, optimal power flow (OPF) to serve the forecasted load at the minimum generation cost and filtering using the presented hybrid (MLP-SFS) technique to estimate the system states.

The status of the electric power grid is evaluated based on the behavior of the system state (voltage angles and magnitude voltages) during the six hours ahead of communication failure. Experimental results prove that the hybrid (MLP-SFS) provides superiority compared to non-optimized MLP and optimized MLP based PSO and GA individually. In IEEE 14 bus system, the maximum mean absolute percentage error (MMAPE) of the system state for the proposed technique (MLP based SFS) is within 4% compared to 8% for non-optimized MLP, 7% for MLP based PSO and 6% for MLP-GA. In IEEE 118 bus system, the maximum mean absolute percentage error (MMAPE) of the system state for the proposed technique (MLP based SFS) is within 8% compared to 18% for non-optimized MLP, 17% for MLP based PSO and 14% for MLP-GA.

In IEEE 14 bus system, the computational time for the proposed technique (MLP based SFS) constitutes 1.98 s compared to 3.05 s for non-optimized MLP, 2.97 s for MLP based PSO and 2.36 s for MLP-GA. In IEEE 118 bus system, the computational

time for the proposed technique (MLP based SFS) is 9.67 s compared to 18.67 s for non-optimized MLP, 18.11 s for MLP based PSO and 16.73 s for MLP-GA. Optimized (MLP-SFS) NN has better behavior and tolerance under different level of noise. It has determined the precision of MLP increments to be around 20%–50% after the hybrid proposed technique presented in contrast with existing MLP optimization techniques. Also, computational time of the hybrid approach was reduced by about 20%–50% compared to other MLP optimization techniques.

The presented hybrid (MLP-SFS) approach can be further extended to detect false injected data, topology error, Micro and Smart grid applications. MATLAB has been used to model and simulate the hybrid (MLP-SFS) approach. MATLAB was used successfully in the following example applications [114]–[117].

Chapter 5

Optimized Neural Network Parameters Using Stochastic Fractal Technique to Compensate Kalman Filter for Power System-Tracking-State Estimation

5.1 Introduction

The electric power industry is experiencing changes due to deregulation. Currently, the competitive power markets provide lower retail costs along with efficient power generation. In general, SE is classified into three categories: Static SE (SSE), Tracking SE (TSE), and Dynamic SE or Forecasted-Aided SE (DSE). In TSE, the power system states are evaluated based on a single set of data, which include real and reactive power injections and flows [65]. Oftentimes, measured data are contaminated by methodical errors, instrument errors, measurement-data scaling, wiring errors, and so forth. These measurements are described as bad measurement data [118].

Kalman filter has been extensively utilized for ongoing SE applications. KF is the minimum-variance estimator to estimate the unknown state vector for the case of white Gaussian noise. However, due to heavy dependence of KF on prior statistical knowledge of system noise, it is hard to compute the optimal solution when it is not identified [119]. Neural Networks have been successfully used in several areas of electric power systems engineering [120], [121]. Considerable attempts have been made recently to compensate for the shortcomings of the conventional Kalman filter by neural networks to improve the estimation error.

Xiong and Zhou [122] implemented a feedforward neural network to replace the Kalman filter gain. The filter has a robust capability to estimate the states of the plant in a stochastic domain without prior knowledge of noise statistics. Neural network filtering methodologies of colored noise considering the Kalman filter structure are used to reestablish the cephalometric images of stomatology.

The authors of this paper [123], [124] proposed a hybrid neural network approach

to correct the Kalman filter errors. The technique is applied to solve the multitarget tracking problem. The approach combines the capabilities of both tracking of the Kalman filter and learning of neural network.

Gao et al. [125] presented an application of Kalman filter compensated by radial basis function NN (RBFNN,) where both KF gain and measurement noise errors are replaced by RBFNN to suppress the filter divergence and improve the robustness of the results. The approach deals with seam tracking during high-power fiber laser butt-joint welding. The paper concluded that the Kalman filter compensated by RBFNN is efficient in tracking accuracy.

Gao et al. [126] employed Elman neural network embedded into adaptive Kalman filter (AKF). The approach is used to improve the stability of adaptive Kalman filter. Elman neural network is applied as an error estimator to compensate for the estimated error of AKF. The eigenvector is used to establish the state and measurement equations for weld seam position which is derived from the weld seam position variable. The results demonstrated that hybrid technique of AKF and Elman neural network improve the performance of weld seam tracking.

The authors of the paper [127] applied an adaptive neural network to power angle stability analysis. The work intended to anticipate the stability status for each generator under a contingency. The rotor angle stability is only selected as the parameter impacting system stability.

Parlos et al. [128] presented an effective non-adaptive and adaptive state filtering for nonlinear dynamic systems utilizing feedforward and recurrent neural networks. Both off-line and on-line learning stages are presented with state filtering algorithms. The algorithms include a prediction and an update steps. The neural networks are used to approximated state equations. Moreover, extended Kalman filters (EKFs) are improved and contrasted to the proposed filter algorithms.

Chin [129] employed a Hopfield neural network to aid the Kalman filter to decrease the estimation error. The hybrid approach amalgamates the capabilities of tracking and learning for Kalman filter and neural network respectively. The state estimate of Kalman filter is compensated by the output of the trained neural network. His approach aims to achieve the multitarget target tracking task without introducing

additional computing resources. The advantage of using the hybrid technique (KF-MLP based SFS) is more effective in enhancing tracking accuracy and reducing the disturbance influences. Another benefit is that it can guarantee robustness which is exhibited in a fast convergence rate along with high accuracy.

This chapter presents KF compensated by a newly improved hybrid optimization technique for MLP NN parameters (weights and thresholds) based on stochastic fractal search technique (MLP-SFS). The hybrid technique is used to solve the tracking-state estimation problem under normal operation conditions, bad data conditions, and sudden loss of loads, generators, and transmission lines. The proposed technique(KF-MLP-based SFS) has been examined on the IEEE 57-bus test systems. The technique (KF-MLP-based SFS) was compared to other techniques, such as NRLF, KF, and KF compensated with radial basis (RBF)-NN (KF-RBF-NN). The results obtained clearly pointed out the superiority of the presented technique (KF-MLP-based SFS) exhibiting high accuracy and less computational time.

5.2 Tracking State Estimation

Two techniques are utilized to process the measured data transmitted through SCADA system and received in the control center. If the measured data are presented into SE program as one vector, then this technique is called snapshot processing of measured data. If the measured data are processed as and when they arrive into the control center and update the estimate, then this technique is called sequential processing of measured data [130].

Debs and Larson presented a basic model for the timely conduct of the power system. This model combined with the real-time measured data. This straightforward procedure is called sequential processing of real time measured data. This can be used to compute the minimum variance estimate [131].

Schweppes and Masiello proposed digital feedback loop to track the changes in static estimate during the daily load cycle. The technique required new measured data to determine the new estimate. The old estimate is rectified through feedback error signal adjusted by a gain matrix [132]. The authors [133] introduced a TSE technique for coordinated AC/DC systems. This technique combined AC and DC measurement equations in rectangular coordinates.

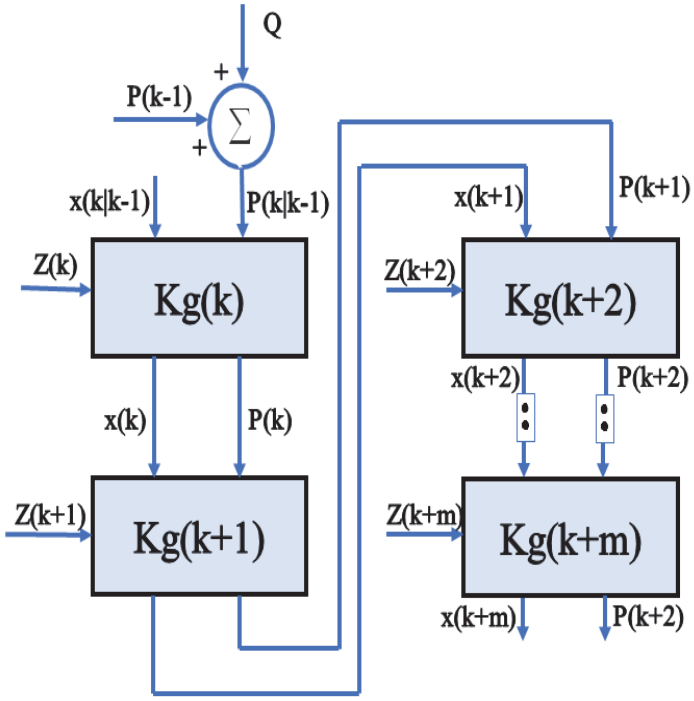

Figure 5.1: Tracking State Estimation

Tracking SE using KF is presented in figure 5.1. Both covariance and state vector are initialized at k-1, then the measurements noise are computed. The measurement noise will be used to compute Kalman gain at k, then update the posterior estimate and covariance at k. When new measured data are represented at k+1, the new estimate should be determined. Both covariance and state vector at k are used as the initial parameters to determine the new Kalman gain, estimate and covariance at k+1. The sequential process continues to the maximum number of time instant k+m.

5.3 Kalman Filtering Compensated Via Optimized Neural Network Parameters based Stochastic Fractal Technique

Optimized neural network parameters (weights and thresholds) based stochastic fractal technique is described in details in [134]. Two important errors are included in the output of linear KF, mismodeling and measurement noise error. Optimized MLP

based SFS is utilized to improve the KF algorithm performance and precision. This is achieved by embedding the optimized MLP based SFS into KF to compensate the aforementioned errors as shown in figure 5.2. A complete flowchart of Kalman filtering compensated by optimized MLP based SFS is provided in figure 5.3.

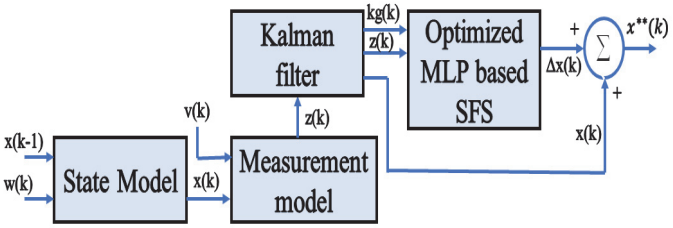

Figure 5.2: Structure of KF compensatd by optimized MLP based SFS (KF-MLP based SFS)

Step 0) At k-1, the initial estimate $X(0|0)$ and covariance $P(0|0)$ were provided, and the statistical parameters Q and R of noises are assumed.

Step 1) Calculate the measurement noise $\bar{y}$ based on Eq. (5.1),

$$\bar{y}(k) = z(k) - H(k)x(k|k-1) \tag{5.1}$$

Where $z(k)$ denotes real time available measurement received from various substations include real and reactive power injections and flows. H(k) denotes jacobian matrix for nonlinear load flow equations. Nonlinear load flow equations are explained in details in [135]. Also $x(k$ |k-1) denotes initial estimate of system state from $k-1$ to k. In this paper, IEEE 57 bus system is implemented which consists of $R^{147\times 1}$ at step k, 2×25 for active and reactive power injections, 2×48 for active and reactive power flows, and 1×1 for voltage magnitude (slack bus).

Step 2) Update the residual covariance S(k) based on Eq. (5.2),

$$S(k) = H(k)P(k|k-1)H^T(k) + R(k) \tag{5.2}$$

Step 3) Compute KF gain $K_g(k)$ based on Eq. (5.3),

$$Kg(k) = P(k|k-1)H^T(k)S^{-1}(k) \tag{5.3}$$

Step 4) Update Posterior states based on Eq. (5.4),

$$x^*(k|k) = x(k|k-1) + Kg(k)\bar{y}(k) \tag{5.4}$$

Step 5) Update Posterior covariance based on Eq. (5.5),

$$P(k|k) = (I - Kg(k)H(k))P(k|k-1) \tag{5.5}$$

Step 6) After computing both measurement noise $\bar{y}$ and KF gain K_g, both quantities will be used as inputs to optimized MLP based SFS technique as shown in Eq. (5.6),

$$G(k) = \begin{bmatrix} \bar{y}(1) & \bar{y}(2) & \bar{y}(n) \\ K_g(1) & K_g(2) & K_g(n) \end{bmatrix} \tag{5.6}$$

Step 7) Initialization,

This includes population size, upper and lower constraints, number of neurons in the hidden layer, side walk, and maximum diffusing number (MDN) as shown in table 5.3.

Step 8) Matrix encoding strategy,

Each particle is encoded for a matrix. We take the MLP structure for IEEE 57-bus system, 2-40-1 for instance, the encoding strategy can be described as:

$$w = \begin{bmatrix} w_{1,1} & w_{1,2}.....w_{1,40} \\ w_{2,1} & w_{2,2}.....w_{2,40} \end{bmatrix}, b = \begin{bmatrix} b_1 \\ b_2 \\ . \\ . \\ b_{40} \end{bmatrix}$$

$$v = \begin{bmatrix} v_{1,1} & v_{2,1}.....v_{40,1} \end{bmatrix}, d = d_1$$

where W denotes the input-hidden layer weight matrix, v denotes the hidden-output layer weight matrix, while b denotes the hidden layer bias matrix, and d denotes the output bias matrix. The overall number of particles generated to create

MLP for IEEE 57-bus system equals 161, so these particles are encoded into weights and biases matrices described by rows × columns, $w = [2 \times 40]$, $b = [40 \times 1]$, $v = [1 \times 40]$, $d = [1 \times 1]$.

Step 9) The MLP as a fitness function of SFS,

For IEEE 57 bus system, it is assumed that the input layer i consists of 2 nodes shown in Eq. (5.6), the hidden layer (size of neurons) j has 40 nodes, and the output layer h has 1 output node. Hyperbolic tangent sigmoid is set as an input transfer function in the fitness function, and linear as an output activation function based on the best chosen control parameters according to table 5.3.

The hidden layer input L_j,

$$L_j = \sum_{i=1}^{n} w_{ij} G(k) - b_j \tag{5.8}$$

The hidden layer output C_j,

$$C_j = f(L_j) \tag{5.9}$$

$$C_j = f(\sum_{i=1}^{n} w_{ij} G(k) - b_j) \tag{5.10}$$

$$C_j = \frac{1}{1 + exp(-(\sum_{i=1}^{n} w_{ij} G(k) - b_j))} \tag{5.11}$$

The computation of the output node $\Delta(x)$to the input,

$$\Delta x(k) = g(\sum_{j=1}^{H} v_{hj} C_j - d_h), h = 1, 2, ...O \tag{5.12}$$

$$\Delta x(k) = \Delta x_{mod}(k) + \Delta x_{nos}(k) \tag{5.13}$$

$$\Delta\hat{x}(k) = \hat{x}(k) - x^*(k) \tag{5.14}$$

Where f denotes hyperbolic tangent sigmoid function, Δx denotes the output of the optimized MLP based SFS or it can also be denoted as mismodeling and measurement noise errors respectively, $\Delta\hat{x}$denotes as true (target) errors which computes by subtracting true states $\hat{x}(k)$ to estimated states $x^*(k)$ as shown in Eq. (5.14).

The mean square error (MSE) of the output node is calculated based on Eq.(5.15),

$$MSE(e_h) = \frac{1}{n}\sum(\Delta x(k) - \Delta\hat{x}(k)) \tag{5.15}$$

Step 10) Diffusion process,

While the number of iterations are not reached, the parameters (weights and thresholds) of the objective function , Multilayer Perceptron (MLP) will be diffused around its current position to compute the optimal parameters among all created points of the search space.

Step 11) Ranking,

All parameters (weights and thresholds) are ranked based on the value of the objective function which is shown in Eq. (5.16). Every parameter in the group will be given a probability value, then the parameters with the highest probability will only be considered in the search space. This technique will enhance the possibility of finding the optimal points in the next generation.

$$P_{ai} = \frac{rank(P_i)}{N} \tag{5.16}$$

Where P_i denotes all parameters,P_{ai} denotes the parameters with high probability, and N denotes the number of all parameters.

Step 12) Updating process,

The optimal parameters (weights and thresholds) will be updated and re-ranked. Second update is required if and only if the ranking values of the objective function is lesser than ε, random number [0,1] to enhance the quality of the search space and fulfill the diversification property. Ref. [82] describes in details the SFS algorithm.

Step 13) Training termination of the KF-(MLP based SFS) process,
If the MSE of objective function is computed or just small enhancement is made in certain number of generations, then the training process terminated, otherwise it goes back to step 7.

Step 14) Simple program of MLP is utilized to train the obtained optimal parameters (weights and threshold) using backpropagation NN algorithm. Moreover, the results will be compared in case of accuracy and least computational time (secs) to true states obtained by NRLF, and estimated states obtained by KF and KF-RBF NN. The equation below are used to determine the final compensated estimated states (voltage phasors),

$$x^{**}(k) = x^{*}(k) + \Delta x(k) \tag{5.17}$$

Where $x^{**}(k)$ denotes the final compensated estimated states (phasors), $x^{*}(k)$denotes the estimated states computed by Eq. (5.4), and $\Delta(x)$ denotes the output of the optimized MLP based SFS as shown in Eq. (5.12). $\Delta(x)$ can also be denoted as mismodeling and measurement noise errors respectively which is shown in Eq. (5.13).

5.4 Case study system

The main object of the case study system is to investigate the presented technique (KF-MLP based SFS) by applying it to TSE. In this study, we only consider over-determined SE when the number of available measurement is larger than the number of state vector $N_s < N_m$. IEEE 57 bus system has been utilized to approve the presented hybrid technique [113]. IEEE 57 bus test system consists of $R^{147\times1}$ at step k, 2×25 for active and reactive power injections, 2×48 for active and reactive power flows, and 1×1 for voltage magnitude (slack bus). The measurements error e_i assumed to have Gaussian normal distribution. The simulated experiment undergoes

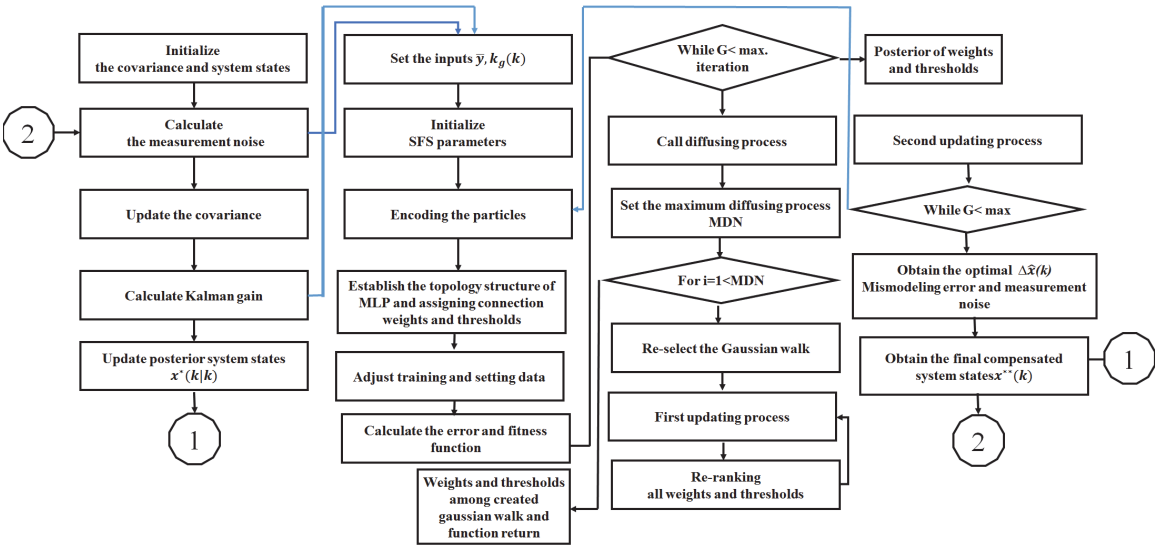

Figure 5.3: Flowchart of Kalman filtering compensated by optimized MLP based SFS

different operating conditions such as load fluctuated linearly with random variations of 4% of the trend component, the presence of bad measurement, sudden load changes, loss of transmission lines, and loss of generations. The efficacy of the presented approach is illustrated by comparing the experimental results of the presented hybrid technique (KF-MLP based SFS) to load flow (true values), conventional Kalman filter (KF), and compensated KF by radial basis neural network (KF-RBF) [125]. Power factor is assumed to be constant so that reactive power follows the active power. Due to the absence of practical field data, Matpower, Newton Raphson Load flow method is presented for each time instant to make the simulated data resemble to the practical field data with ±5% Gaussian noise which is arbitrarily generated according to ref. [136]. The final measurement vector thus obtained are used as an input to the hybrid technique (KF-MLP based SFS). This study is carried out over a period of 16 time samples. In order to approve the execution of the presented hybrid technique, the acquired results for IEEE 57-bus test system is carried out based on the following execution indices:

The mean absolute percentage error is used to evaluate the states (voltage phasors),

$$Mape(x) = \frac{1}{n_b - 1} \sum_{i=2}^{n_b} \mid \frac{x^{**}(k) - \hat{x}(k)}{\hat{x}(k)} \mid \tag{5.18}$$

where $x^{**}(k)$ denotes the final compensated estimated states using presented technique (KF-MLP based SFS) and $\hat{x}(k)$ denotes the true states (NRLF)

The filter effect FE is calculated to evaluate the overall estimation achievement,

$$FE = \frac{\sum_{i=1}^{n} \mid \hat{z}_i(k) - z_i^t(k) \mid}{\sum_{i=1}^{n} \mid z_i(k) - z_i^t(k) \mid} \tag{5.19}$$

Where $\hat{z}_i(k)$, $z_i^t(k)$, and $z_i(k)$ final estimate, true, and measured measurements values, so the performance index FE should equal to less than one for accurate estimation.

5.5 Results and discussion

The execution of the proposed hybrid technique(KF-MLP based SFS) is examined on IEEE 57-bus system. In this paper, we only consider $N-1$ contingency analysis. Various contingency analysis are included in this experiment such as normal operating condition, a presence of bad data, sudden load changing condition, sudden loss of transmission line and generations.

The best control parameters are tabulated in tables 5.1 and 5.2. This includes the best performance of an algorithm and the optimal size of neurons. Levenberg Marquardt (LM) has the least MAPE which is 1.2017% for voltage angles and 0.8290% for magnitude voltage. This can be compared to Gradient descent with momentum (GDM) which has the highest MAPE, 15.3335% for voltage angles and 13.5157% for magnitude voltages. Moreover, Levenberg Marquardt (LM) algorithm has the least computational time which is 19 seconds compared to Bayesian regularization (BR) which has the longest computational time, 5.32 minutes. In case of the optimal size

Table 5.1: Ranking Training Algorithm of Feed Forward Neural Network Using compensated KF-MLP Based SFS for IEEE 57 bus system in case of accuracy and computational time

#	Algorithms	MAPE(δ)	MAPE($\|V\|$)	Training Time (Secs)
1	LM	1.2017	0.8290	00:00:19
2	BFG	5.8059	4.8508	00:02:35
3	SCG	10.0878	8.4583	00:00:30
4	RP	10.1466	8.6433	00:00:49
5	OSS	10.7535	9.6272	00:00:53
6	GDX	14.0550	12.4443	00:00:49
7	BR	14.4926	14.6085	00:05:32
8	GD	15.3335	13.5155	00:00:41
9	GDM	15.3335	13.5157	00:00:41

Table 5.2: Selecting the best number of neurons Using compensated KF-MLP Based SFS in case of accuracy and computational time

#	No. Of Neurons	MAPE(δ)	MAPE($\|V\|$)	Training Time (Secs)
1	5	9.9363	8.0660	00:00:09
2	15	7.7155	6.6711	00:00:14
3	30	4.9906	3.8677	00:00:25
4	40	1.2017	0.8290	00:00:30
5	50	2.4557	1.3031	00:01:04
6	60	2.3009	1.5205	00:01:52
7	70	7.8×10^{-4}	8.0×10^{-4}	00:03:37

of neurons, 40 neurons in the hidden layer provide us with the least MAPE which is 1.2017% for voltage angles and 0.8290% for magnitude voltages. Also, it gives a reasonable computational time which is 30 seconds. Due to overfitting, computation NO. 7, 70 neurons in the hidden layer are neglected. It concluded that the LM algorithm and 40 neurons have been chosen to be the best control parameters for the proposed hybrid technique.

Table 5.3 shows a summary of different computations. This table provides the best control parameters used in the proposed hybrid KF-(MLP based SFS). These parameters include input transfer function, output activation function, the maximum number of generations, and others. Moreover, the measurements are divided into 90% for training and 10% for testing. These measurements include real and reactive power injections and flows.

Table 5.3: Control Parameters used in Compensated KF-MLP based SFS for IEEE 57-bus system

R (power injections)	10
R (power flows)	5
Q (voltage angles)	0.001
Q (voltage magnitude)	0.1
Training Algorithm	LM
No. Of Neurons	40
Setup Division of Input Data	90% for training, 10% for testing
Population Size	125
Input Transfer Function	Tang-Sigmoid
Output Activation Function	Linear
Upper and Lower boundaries	[1, -1]
Maximum Diffusion Number (MDN)	1
Maximum No. Of Generations	250
Gaussian Side Walk (GSW)	1

5.5.1 Normal operating condition

Loads fluctuate linearly at each bus from 70% to 130% with a trend of 4% . Measurements are considered to be noisy, however, these measurements are not assumed to be as bad measurements.

Figure 5.4 describes the performance of the proposed KF-(MLP based SFS), KF-RBF, linear KF, and NRLF techniques of IEEE 57- bus system. The figure represents the time behavior of the true and the estimated voltage magnitudes at buses 3, 19, 30, and 56. These buses are arbitrarily selected. As can be observed, the voltage magnitudes drop as the loads increase linearly. For instance, the voltage magnitude on bus 30 drops from 1.1pu to 0.95pu, however, it remains constant on bus 3 due to closeness to the main generator. Moreover, the proposed approach follows the same behavior of the true technique NRLF, however, both KF-RBF and KF failed to provide us with the true change of the voltage magnitudes during the 16-time samples. This can make the proposed technique more reliable compared to others.

Figure 5.5 shows the comparison of the proposed KF-(MLP based SFS), KF, and KF-RBF techniques in terms of MAPE. As can be seen, the proposed technique (KF-MLP based SFS) has the lowest MAPE, 5% compared to 20% for KF-RBF and 26% for KF. Further, there is no specific trend of the MAPE for all presented techniques.

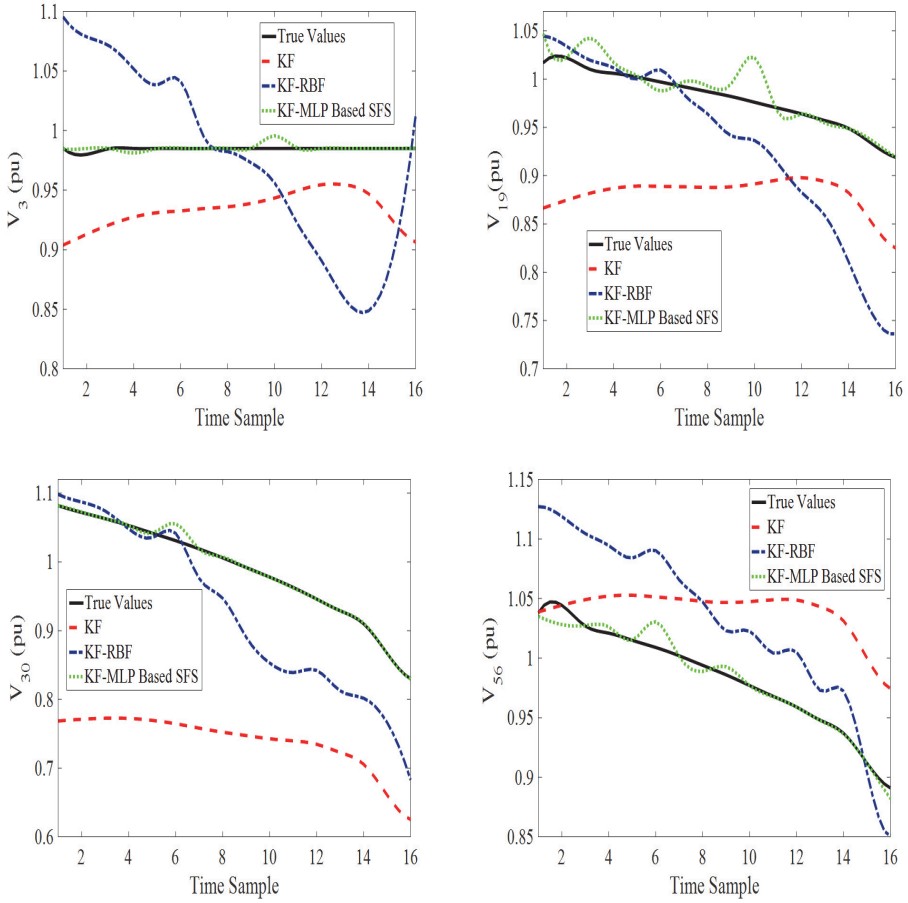

Figure 5.4: Performance of voltage magnitude at buses 3, 19, 30, and 56 for IEEE 57-bus system under normal operating condition

It concluded that the proposed technique is superior compared to others.

5.5.2 Presence of bad data measurements

i) One bad measurement is chosen randomly with invert polarity from 2nd to 4th time instants.

ii) Two bad measurements chosen randomly, one with invert polarity and the other is set to zero from 6th to 8th time instants.

iii) Three bad measurements are chosen randomly, one with invert polarity, second measurement is set to zero, and the third measurement is presented with gross data error with 12σ from 10th to 12th time instants.

iv) Four bad measurements are chosen randomly, one with invert polarity, the

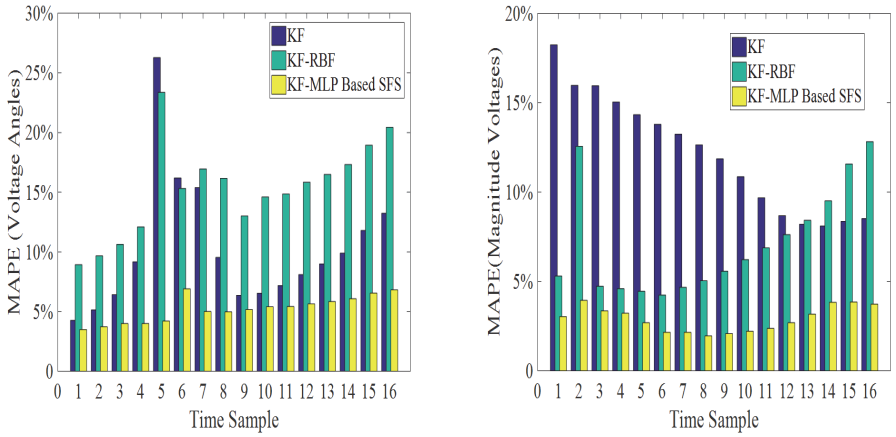

Figure 5.5: Mean Absolute Percentage error (MAPE) for estimated states in IEEE 57-bus system under normal operation condition

second is set to zero, and the third and the fourth one is presented with gross data error with 15σ and 25σ from 14th to 15th time instants.

Figure 5.6 shows the performance of the techniques for IEEE 57-bus system under bad data condition. The figure represents the time behavior of the true and the estimated states at buses 2 and 49. It can be seen that KF is more sensitive to the bad data. For instance, From 6th to 8th, the voltage magnitude on buses 2 and 49 increase suddenly to 2 pu which exceeds the nominal voltage. Further, from 10th to 12th, we introduce more bad data which lead to the voltage magnitude to increase more than 2 pu. KF-RBF is less sensitive to the bad data, however, the proposed KF-(MLP based SFS) rejects the bad data and follow the behavior of true states.

Figure 5.7 illustrates the comparison of the proposed KF-(MLP based SFS), KF, and KF-RBF techniques in terms of MAPE. The figure verifies the sensitivity of KF to bad data. For instance, from 6th to 8th, the MAPE is 70%, from 10th to 12th, the MAPE is 90%, and so forth. On contrary, the proposed KF-(MLP based SFS) has the lowest MAPE, less than 5% compared to KF-RBF, less than 25%. This confirms the efficacy of the proposed technique.

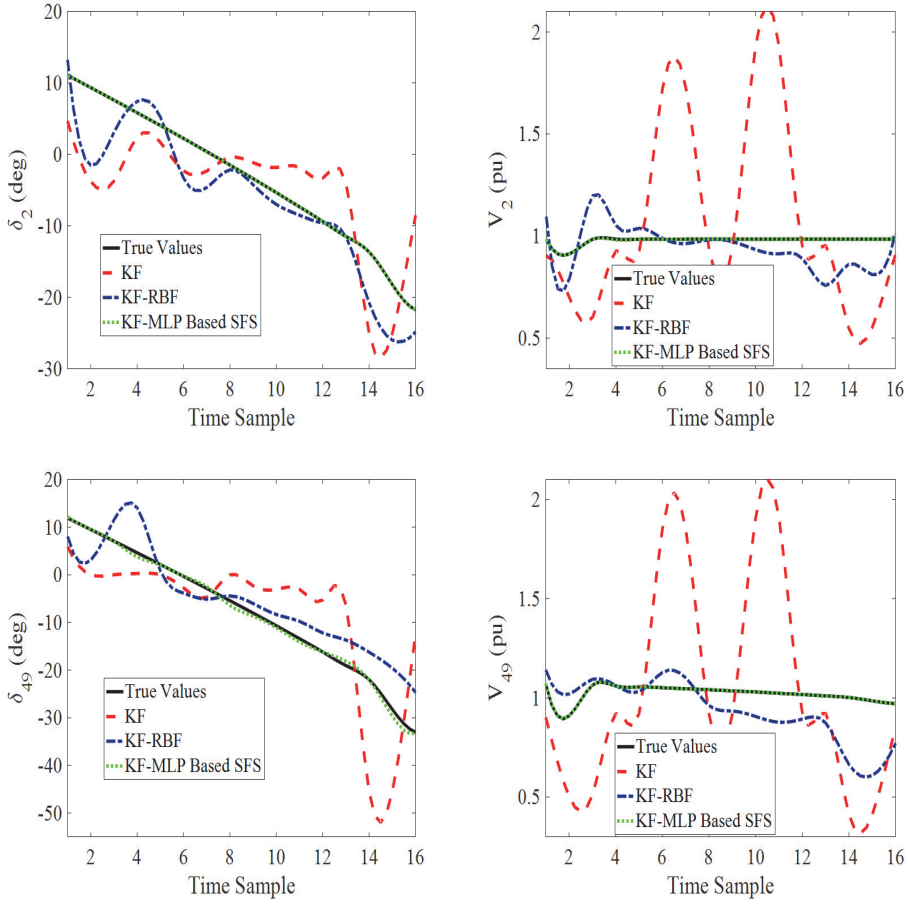

Figure 5.6: Performance of estimated states at buses 2 and 49 for IEEE 57-bus system under bad data condition

5.5.3 Presence of Sudden Load change conditions

i) 40% load increase from 2nd to 4th time instants on bus 10.

ii) Sudden loss of load from 5th to 7th time instants on bus 19.

iii) 50% load decrease from 8th to 10th time instants on bus 51.

iv) 20% load decrease from 11th to 13th time instants on all the load buses.

v) 20% load increase from 14th to 16th time instants on all the load buses.

Figure 5.8 shows the performance of the techniques for voltage magnitude at buses 10, 19, 37 and 51 of IEEE 57-bus system. As can be observed, 20% load reduction on all buses leads to increase the voltage magnitude on bus 19 from 0.990pu to 0.995pu,

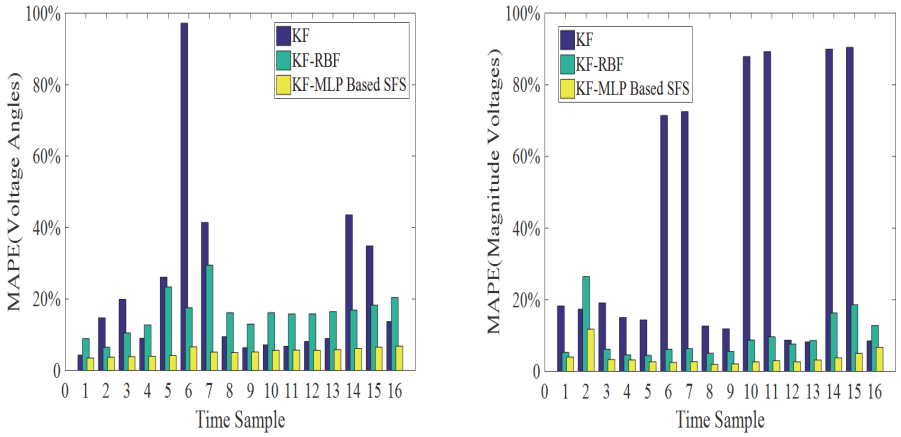

Figure 5.7: Mean Absolute Percentage error (MAPE) for estimated states in IEEE 57-bus system under bad data condition

from 1.01pu to 1.03pu on bus 19, from 1.060pu to 1.066pu on bus 51, and from 1.010pu to 1.015pu on bus 37. Further, 20%load increment on all buses leads to decrease the voltage magnitude to less than 0.915 on bus 19, less than 1.044pu on bus 51, and less than 0.97pu on bus 37. This load increment does not affect bus 10 due to the closeness to the main generator. The variation of the magnitudes of the proposed technique is lower than other techniques. This shows the reliability of the proposed technique in comparison to other.

Figure 5.9 depicts the comparison of the proposed KF-(MLP based SFS), KF, and KF-RBF techniques in case of MAPE. It can be seen that the proposed hybrid technique has the lowest MAPE, 5% for both voltage angle and magnitude compared to KF, 25% for voltage angles and 18% for voltage magnitude, and 15% for voltage angle and 10% for voltage magnitude for KF-RBF. This demonstrates the better filtering execution of the proposed technique in comparison to other.

5.5.4 During sudden loss of transmission line

i) A transmission line between buses 22 and 23 are lost from 3th to 6th time instants. ii) A transmission line between buses 36 and 40 are lost from 11th to 13th time instants.

Figure 5.10 shows the true and estimated voltage angles at buses 22, 23, 36, and 40 of IEEE 57-bus system. As can be observed, the voltage angle on bus 22 drops

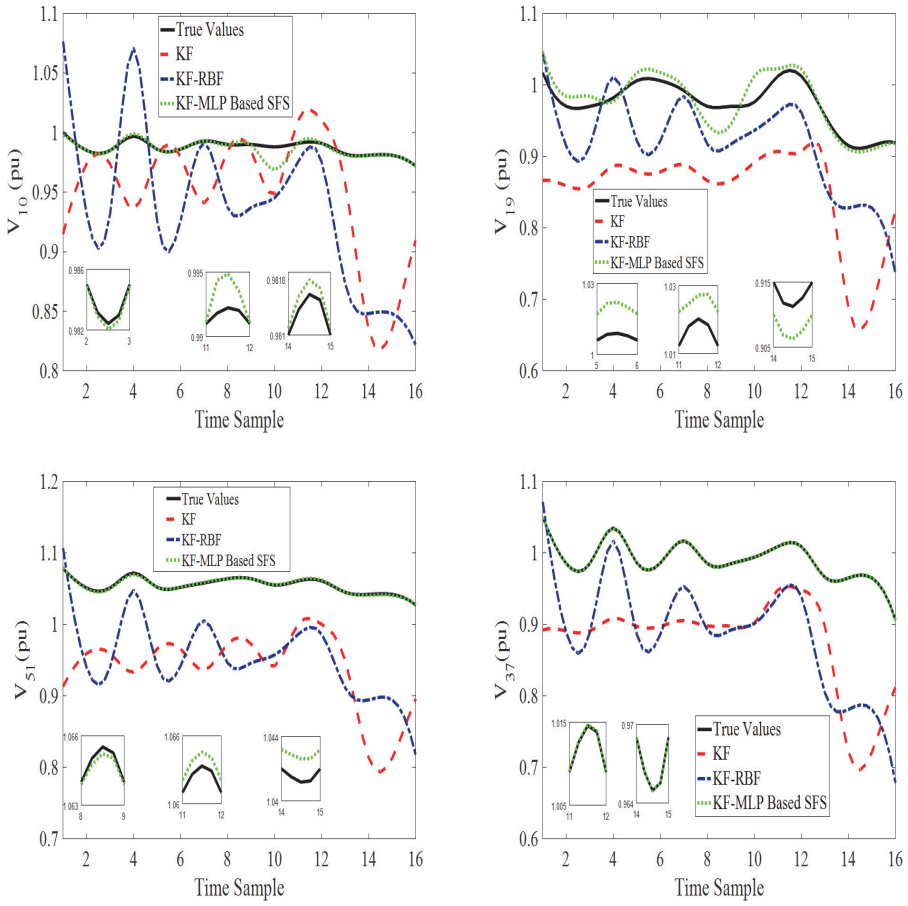

Figure 5.8: Performance of voltage magnitude at buses 10, 19, and 51 for IEEE 57-bus system under sudden load changing condition

down from -11 deg to -17 deg. This due to the removal of transmission line between buses 22 and 23. Moreover, bus 36 is affected by the line outage between buses 22 and 23. This due to the closeness to the line outage. Further, there is a small reduction in the voltage angle on bus 40 from -14 deg to -16 deg due to the loss of transmission line between buses 36 to 40. The proposed technique shows better performance in capturing precisely the change in the estimated states compared to others.

Figure 5.11 shows the comparison of the proposed KF-(MLP based SFS), KF, and KF-RBF techniques in case of MAPE. It can be seen, the MAPE of the proposed technique is smaller, 5% compared to KF, 20%, and 18% for KF-RBF. This shows the better filtering performance of the proposed method.

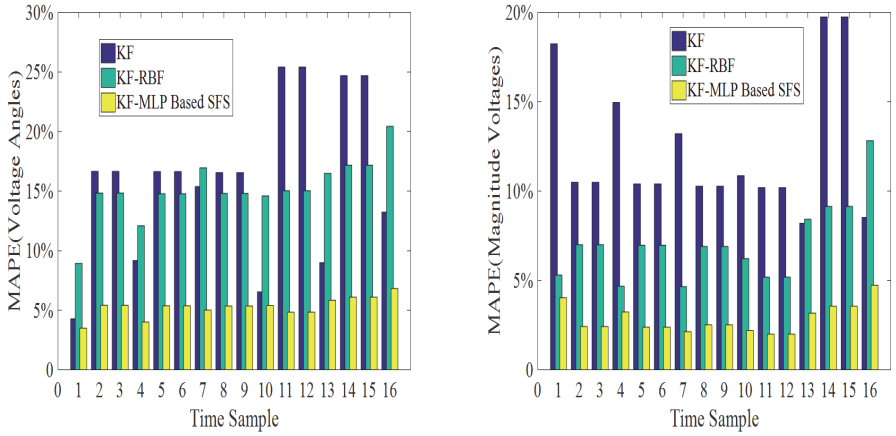

Figure 5.9: Mean Absolute Percentage error (MAPE) for estimated states in IEEE 57-bus system under sudden load changing condition

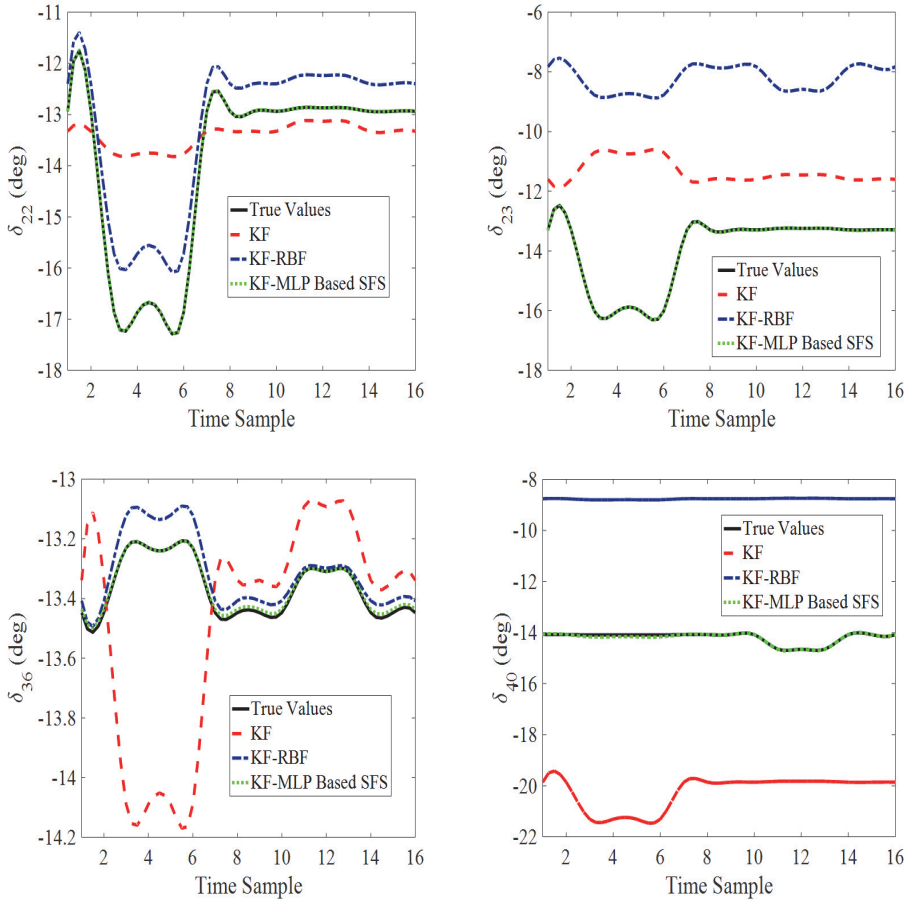

Figure 5.10: Performance of voltage angles at buses 22, 23, 36, and 40 for IEEE 57-bus system under sudden loss of transmission lines condition

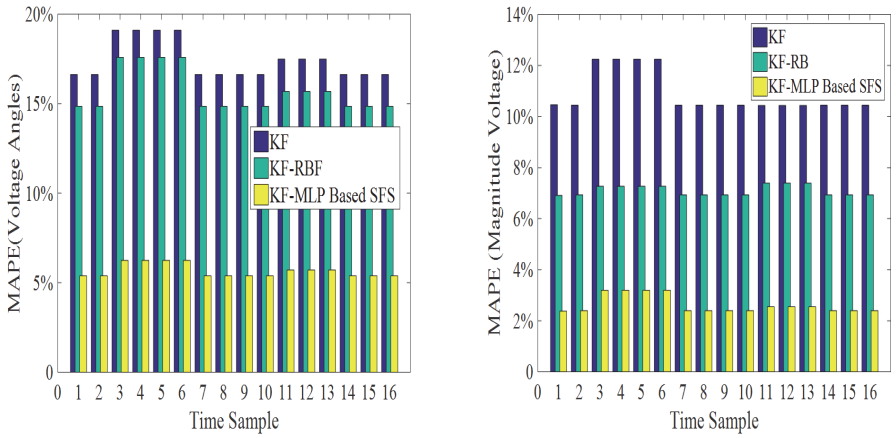

Figure 5.11: Mean Absolute Percentage error (MAPE) for estimated states in IEEE 57-bus system under transmission line loss condition

5.5.5 During sudden loss of Generations

i) Generator 5 on bus 8 is lost from 4th to 6th time instants.

ii) Generator 7 on bus 12 is lost from 11th to 12th time instants.

Figure 5.12 illustrates the true and estimated voltage magnitude at buses 8 and 12 of IEEE 57-bus system. As can be observed, the voltage magnitude on bus 8 drops down from 1pu to 0.95pu.

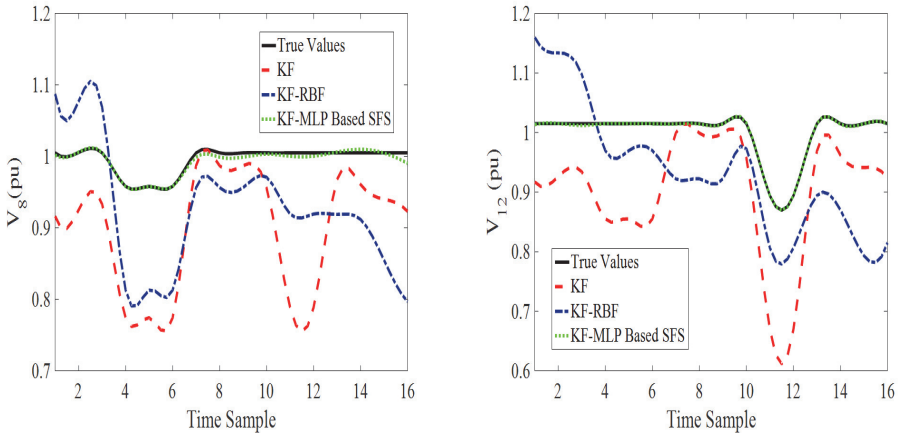

Figure 5.12: Performance of voltage magnitude at buses 8 and 12, Generator 5 and 7 for IEEE 57-bus system under sudden loss of Generations condition

This due to the sudden loss of generator 5. Moreover, the loss of generator 7 on bus 12 leads the voltage magnitude to drop down from 1pu to 0.9pu. Figure 5.13 shows the comparison of the proposed KF-(MLP based SFS), KF, and KF-RBF techniques in case of MAPE. As can be observed, the MAPE of the estimated states for the proposed KF-MLP based SFS technique is very lower, 5% compared to KF, 30%, and KF-RBF, 15%. This shows the superiority of the proposed KF-(MLP based SFS) technique.

Table 5.4 presents filter effect index (FE) obtained under different operating conditions. The filter effect (FE) should be less than one for accurate estimation. It can

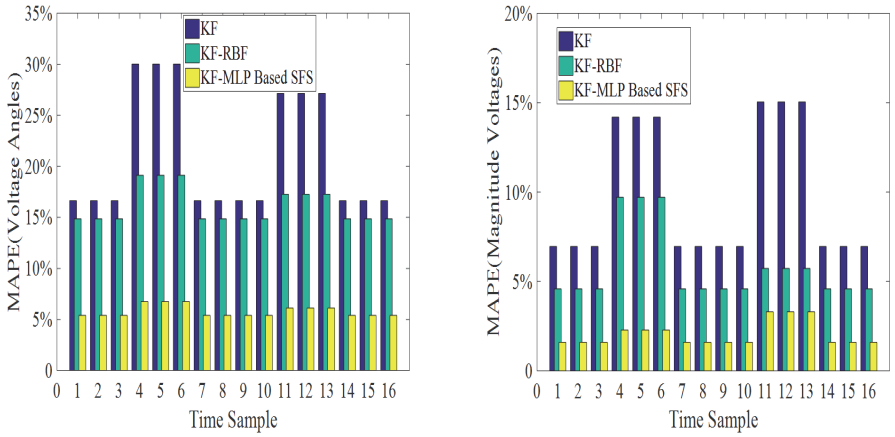

Figure 5.13: Mean Absolute Percentage error (MAPE) for estimated states in IEEE 57-bus system under Generators loss condition

be seen, under normal operating condition, the FE of the proposed technique is 0.6959 compared to 1.8924 for KF, and 0.9462 for KF-RBF. Moreover, under the presence of bad data, the FE of the proposed technique is 0.8340 compared to 1.8570 for KF, and 0.9026 for KF-RBF. In addition, sudden load change, the proposed technique has FE of 0.700 compared to 1.6938 for KF, and 0.8141 for KF-RBF. Sudden loss of transmission line, the proposed technique has 0.6500 of FE compared to 1.0548 for KF, and 0.8230 for KF-RBF. Finally, sudden loss of generations, the FE of the proposed technique is 0.7650 compared to 1.1280 for KF, and 0.9664 for KF-RBF. Both KF-(MLP based SFS) and KF-RBF techniques have achieved a better filtering property, however, KF-(MLP based SFS) has the least FE value. On the other hand, the FE of KF is higher than 1 which makes KF impractical for SE.

5.6 Conclusion

Optimized MLP parameter based stochastic fractals search technique (KF-MLP based SFS) was applied to improve KF performance and accuracy. Both KF gain (mismodeling error) and measurement noise are substituted by MLP based SFS technique. This could suppress the filter divergence and improve the accuracy in the estimated states. Different operating conditions were used to validate this experiment include normal operating condition, bad data condition and sudden loss of loads, generations,

Table 5.4: Comparison of proposed KF-MLP Based SFS technique with two other conventional approaches in Filter Effect for IEEE 57 bus system

Filter Effect Index	KF	KF-RBF	KF-MLP-SFS
Normal operating condition			
Min. FE	0.3414	0.1707	0.1018
Mean. FE	1.1461	0.5731	0.4432
Max. FE	1.8924	0.9462	0.6959
Presence of bad data			
Min. FE	0.0362	0.2055	0.2500
Mean. FE	0.7857	0.5052	0.4526
Max. FE	1.8570	0.9026	0.8340
Sudden Load change			
Min. FE	0.5000	0.1000	0.1137
Mean. FE	0.9308	0.5941	0.3313
Max. FE	1.6938	0.8141	0.7000
Sudden loss of transmission line			
Min. FE	0.8635	0.7577	0.3649
Mean. FE	0.9889	0.8037	0.5546
Max. FE	1.0548	0.8230	0.6500
Sudden loss of Generations			
Min. FE	0.8282	0.7346	0.3909
Mean. FE	0.9923	0.8506	0.6722
Max. FE	1.1280	0.9664	0.7650

and transmission lines. Further, IEEE 57 bus system was implemented to illustrate the results of the proposed (KF-MLP based SFS) and other techniques. The estimated states of the proposed approach were contrasted to the true states obtained by NRLF and estimated states obtained KF and KF-RBF. Experimental results show the superiority of the proposed technique (KF-MLP based SFS) compared to KF and KF-RBFNN. The superiority of the presented technique was exhibited in the least MAPE. The presented technique (KF-MLP based SFS) increased the precision to approximately 50%-70% in contrast with other techniques.

Chapter 6

Conclusion

6.1 Conclusion

In this research work, new robust techniques include metaheuristic and artificial intelligent is proposed to solve load forecasting and state estimation problem.

Hourly electricity price forecasting problem is discussed and solved using composite back-propagation multilayer neural networks. Three individual nets are interconnected in cascade and parallel topologies. A comparison of different topologies is made to investigate the superior performance based on accuracy and computational time. Electricity load forecasting for the Australian market is used to validate the experiment. It is concluded that cascade topology improves the overall execution in the system due to its ability in accumulating the error from each training stage. On contrary, parallel topology improves the execution in the systems due to its ability in averaging the errors.

Multilayer perceptron neural network parameters (weights and thresholds) is optimized using stochastic fractal search technique. The purpose of using optimized hybrid technique is to enhance the current multilayer neural network. The enhancement is exhibited in high accuracy, low convergence rate, and easier operation. The main goal of SFS algorithm is to provide optimal initial parameters to improve the performance of MLP NN. A combination of IEEE bus systems and New York Independent System Operator (NYISO) is made. Various scenarios of communication failure and measurement error is discussed to validate the presented technique. The presented hybrid technique provides superiority compared to other techniques.

The aforementioned technique is applied to tracking state estimation problem. The presented technique improves the Kalman filter performance by replacing both Kalman filter gain (mismodeling error) and measurement error. This suppresses the Kalman filter divergence and improve the performance. Various operating conditions are applied to validate the experiment. A comparison of the proposed and other

techniques is made to show the improvement. The results prove the superiority of the presented techniques among others.

Stochastic fractal search technique is applied to tracking state estimation under various operating conditions. Four IEEE test cases are implemented to verify the SFS technique. The presented SFS is carried out to be compared to other techniques. The results show that SFS technique provides accurate sate estimates compared to other techniques. Moreover, the SFS technique has been used in hybrid technique with back propagation neural network (MLP-SFS) and simulated annealing (SFS-SA). The MLP-SFS is applied to improve the accuracy as well as measurement redundancy by gradually increase the number of PMUs. The least mean absolute percentage error is provided by the proposed technique. The SFS-SA technique is used to solve distributed multi area state estimation problem. The SFS is used to perform the local SE and SA technique is used to perform the global SE. A significant reduction in computational time is achieved by the presented technique.

Stochastic fractal search technique is modified and applied to power system static sate estimation. Logarithmic functions in diffusing process is substituted with several benchmarks functions. In addition, the uniform distribution parameter in both diffusing and updating processes is replaced by chaotic maps. The M-SFS technique improves the performance of its original SFS technique.

In conclusion, The presented techniques are applied to different load forecasting and state estimation problems. Different modern electric network are used to verify the experiments. A comparison of the presented techniques to others is made to validate the outcomes. The proposed techniques addressed and achieved reliable power system operation by preventing the security risks in the modern electric grids.

6.2 Scope of Future Work

This thesis can be further extent:

• The contribution discussed in this thesis can be further extent to implement large size systems such as IEEE 300 bus systems.

• The presented techniques can be further extent to detect false injected data and topology error.

• This thesis can be further extent to study load change and lack of generation production challenges.

• Contingency analysis can be extent to include N-2 emergency state. This represents multi generations or transmission lines outages.

• The techniques introduced in this thesis can be employed in micro and smart grid area.

• New hybrid optimization techniques can be explored and implemented in this thesis.

• Renewable energy includes solar, tidal, and wind energies can be studied in this thesis.

• The impact of Phasor measurement units can be deeply studied in this thesis. This study can include the advantage and disadvantage of PMUs.

Appendix A

Multilayer Artificial Neural Networks for Real Time Power System State Estimation

A.1 Introduction

A state estimation uses for defining a reliable estimate of the power state vectors include bus voltage magnitudes and angles from system parameters, structural data, and a set of real time available measurements including net active and reactive power injections and flows. Bad data is identified and purging in order to state estimator gives a reliable data base for power system operation and control. The measurement system outline decides how well the state estimator performs this capacity [5].

The methods of state estimation are focused around the work of a lot of scientists. In [137], the authors applied an interactive type block-partitioning algorithm on a single instruction, single data (SISD) type computer in order to assess its ability of the speed up of estimation computation.It functions accurately for both of IEEE 118-node with numerous loops and IEEE 135 node.

Abur [7] gave a brief outline of the LP estimator and he proposed the bad data identification technique exhibited in a succession of identification and elimination cycles. He clarified a few contemplations with the methodology. Several numerical cases such as AEP's 14, 30, 57 and 118 test systems have been utilized to demonstrate the method. He conclude that "elimination of such bad data will yield an unbiased estimator and an efficient way of eliminating designated measurements through weight changing".

In [8], the authors proposed a new computational technique for solving equality constraint in PSSE by excluding the Lagrange multipliers. It is an important to compute a complete partial factorization (L, U) of matrix which enhance numerical conditioning. A second region for further examination is the treatment of critical measurements, for example, the pseudo-estimations needed for observability. Critical

measurements have zero residuals and, in this sense, they are proportional to equality constraints. The algorithm was indicated to perform effectively on a few practical measured samples. Standardized residuals can be processed in a proficient way with the new algorithm.

In [9], the authors introduced a new methodology of artificial neural networks topology based on topology processing and static state estimation. They have compared two ANN based on models include the counter-propagation network (CPN) and functional link network (FLN) in order to determine topology transforming and static state estimation on several IEEE test systems include 14, 19, 57 useful Indian system. Moreover, the Hopfield neural network and the conventional fast decoupled state estimator (FDSE) have been determined and compared. It is concluded that the ANN based on models has the least CPU time and much faster, and function accurately notwithstanding for non- Gaussian noise.

In [12], the authors introduced another system focuses on the neural network prepared by prop algorithm, in which topological perceptibility is considered. The sufficiency of preparing of the system is tried by feeding some novel information designs that were excluded in the preparation set and the yield results acquired were accepted by looking at the results from alternate techniques.

In [13], Hopfield neural network (HNN) and Parallel Genetic algorithm (PGA) applied in neural network in term of the best robust static state estimation method on 5 bus test system. It is concluded that Hopfield method has a long training time and Parallel Genetic method provides the optimal converge which is not rely on initial values.

In [15], the authors presented another system focused around the artificial neural network method in order for deciding the recognizability of the power systems proposed a strategy on prop algorithm. Notwithstanding the standard backpropagation algorithm additionally presented.

Ivanov and Gavrilas [24] proposed a multilayer perceptron for the static state estimator and employed on IEEE 14 bus test system. It is proven that ANN estimator can provide less CPU time compared to conventional estimation algorithms in case of bus voltage varies dues to load conditions and there is no change on operating system. Also the results of IEEE 14 bus system have a precise estimation.

In this chapter, composite multilayer neural network topologies exhibited in hybrid Parallel and Cascade topologies are applied to address the conduct of differing composite topologies to compare the best performance indices exhibited by the maximum relative error, mean absolute percentage error (MAPE), and mean square error (MSE). The state estimation performance of the proposed system is evaluated utilizing real time data for IEEE 14, 30, 57, 118, and 300 bus test systems in electric girds from the American Electric Power System in the Midwestern US which is published by the official website of University of Washington. The terminology of this work is minimizing the residual of equivalent node measurements and estimate the state vectors including magnitude voltages and angles which is useful for security assessments and system monitoring. Chapter 3 explains the neural networks structure tested.

A.2 State Estimation

The main objective of state estimation (SE) algorithms is to determine the unknown sate vectors on all buses in electric grid include angles and magnitude voltages. Where inputs are real time available measurements of net active and reactive power injections of generations and loads and initial state vectors. Here the output is the predicted state vectors. The following terminology is utilized by reformulating nodal nonlinear net power system injections' equations in (A.1) and (A.2) as a general mathematical model of the WLS algorithm represented in equation (A.3).

$$P_{Gi} - P_{Di} - \sum_{i=1}^{n} |V_i||V_j|(G_{ij}\cos(\theta_i - \theta_j) + (B_{ij}\sin(\theta_i - \theta_j)) = 0 \tag{A.1}$$

$$Q_{Gi} - Q_{Di} - \sum_{i=1}^{n} |V_i||V_j|(G_{ij}\sin(\theta_i - \theta_j) + (B_{ij}\cos(\theta_i - \theta_j)) = 0 \tag{A.2}$$

$$[z] = h([x]) - e[x] \tag{A.3}$$

Where z is a known vector of M matrix, real time measurements including active and reactive power generations and loads, $h(x)$ is represented as $h : R^n \times R^m (n < m)$

matrix of nonlinear power flow equations which is relating measurements to state vectors,e is M vector of the measurements white noise, x is unknown vector of M matrix, state variables (voltage magnitudes and angles).

State vectors will be corrected by linearization, in an iterative procedure, utilizing Newton's method.

$$[x]^{k+1} = [x]^k + [\Delta x]^k \tag{A.4}$$

The following terminology is minimized:

$$J([x]) = ([z] - h([x]))^T W([z] - h([x])) \tag{A.5}$$

Where W is M vector of diagonal matrix of measurements weights.

Newton's linearization method:

$$G(x)[\Delta x]^k = H^T[x]^k R^{-1}[\Delta z]^k \tag{A.6}$$

Where Δz is M vector of Gain matrix which represents the error between scheduled and computed real time available measurements. Each iteration step, Jacobian and Gain matrices will be computed by bus admittance matrix which is created by bus and line data of the grid elements [138].

In this chapter, composite multilayer neural network topologies exhibited in hybrid Parallel and Cascade topologies are applied to address the conduct of differing composite topologies to compare the best performance indices exhibited by the maximum relative error, mean absolute percentage error (MAPE), and mean square error (MSE). The state estimation performance of the proposed system is evaluated utilizing real time data for IEEE 14, 30, 57, 118, and 300 bus test systems in electric girds from the American Electric Power System in the Midwestern US which is published by the official website of University of Washington. The terminology of this work is minimizing the residual of equivalent node measurements and estimate the state vectors including magnitude voltages and angles which is useful for security assessments and system monitoring. Chapter 3 explains the neural networks structure tested.

A.3 Proposed Neural Network Architecture

The Multilayer Perceptron (MLP) are feedforward networks utilized for approximation and forecasting, trained with back-propagation algorithms which supervised the networks by involving the desired output.

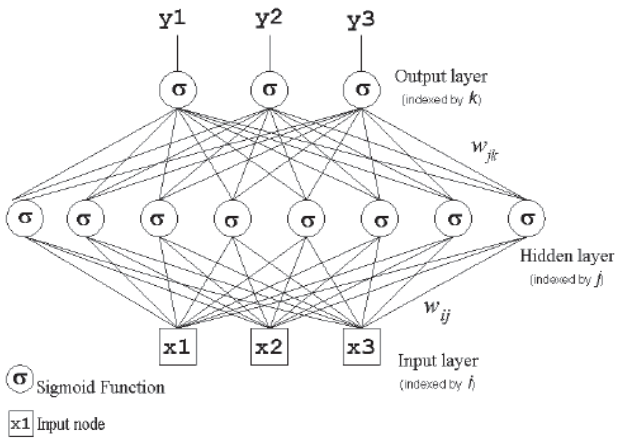

Figure A.1: Architecture of Multilayer Perceptron

They comprise of one input layer, one or more hidden layers and one yield layer. The sizes of input and yield layers are computed by the tackled problem. Every neuron has weighted associations with all the neurons from the adjoining layers, depending on the ANN structure.

Determining the solution with a MLP obliges two stages: training and generalization. In the training stage, the network is fabricated and the neurons' weights and inputs are arbitrarily instated and after that refined in an iterative forward-in reverse procedure, until the yields of the MLP coordinate nearly with the desired outputs, for all the models in the training information set [139].

$$E = \frac{1}{2} \sum_{m=1}^{M} ||d^m - o^m||^2 \tag{A.7}$$

Where E is the error between the $[d]$ desired output or target and $[o]$ is the computed output. After that weights will be updated at each iteration t, through

backpropagation with [24].

$$w_{ij}^{t+1} = w_{ij}^{t} - \eta \frac{\partial E}{\partial w_{ij}} \downarrow w_i j \tag{A.8}$$

There are some designed roles must be followed such as number of hidden layer size (neurons) should be among the number of input and output layer because large number of neurons in the hidden layer leads to obtain Over fitting happens when the neural networks has so much data transforming limit that the constrained measure of data contained in the training set is insufficient to train the majority of the neurons in the hidden layers. A second issue can happen notwithstanding when the training information is enough. An unnecessarily substantial number of neurons in the hidden layer can expand the time it takes to train the network. All the aforementioned issues can be used to adjust the weights and thresholds [139].

A.4 Tested System and Results

The comparison of both hybrid Cascade and Parallel topologies is carried out on five test systems. A brief description of these system follows:

• Real time Measurements of the part of the American Electric Power System in the Midwestern US include scheduled net active and reactive power generations and loads injections and net nonlinear power injections equations are subjected to minimize the nodal measurements residuals.

• We subjectively selected three distinct layers to compare performance and accuracy conduct between proposed topologies. More layers require more calculations, yet their utilization may bring about the system taking care of complex problems all the more proficiently.

• We picked the quantity of neurons in each net taking into account the best performance, more neurons require more calculations, and they tend to over-fit the data when the number is situated too high, yet they permit the network to tackle more problems.

• We chose the initial weights of a neural network from the reach $(-1/\sqrt{d}, 1/\sqrt{d})$, where d is the quantity of inputs to a given neuron. It is accepted, that the sets are standardized - mean 0, variance 1.

• We set Net1, feed-forward structure, n hidden layer size (neurons), Bayesian regularization backpropagation (BR) as a training algorithm.

• We set Net2, cascade-forward structure, n hidden layer size (neurons), Levenberg-Marquardt backpropagation (LM) as a training algorithm.

• We set Net3, fit-net structure, n hidden layer size (neurons), quasi-Newton backpropagation (BFGS) as a training algorithm.

• We set activation function to hyperbolic tangent sigmoid for input and output layers.

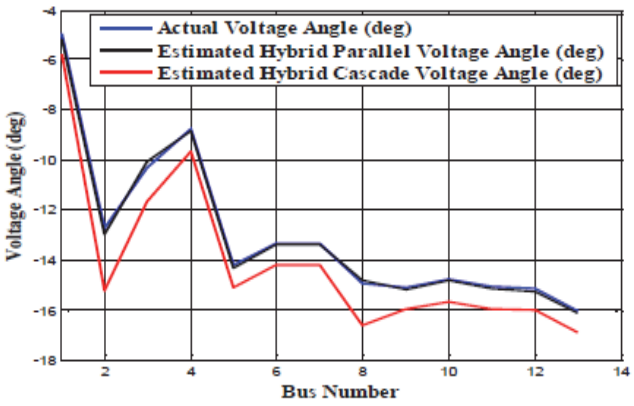

Figure A.2: Hybrid Parallel and Cascade topologies, Estimated and Actual voltage angle in IEEE 14 bus system

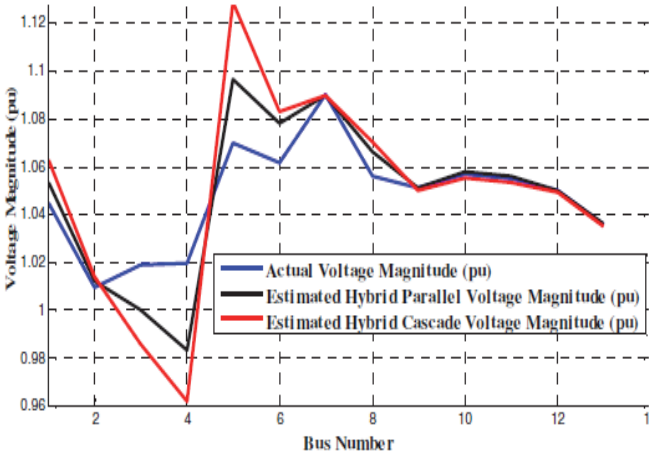

Figure A.3: Hybrid Parallel and Cascade topologies, Estimated and Actual magnitude voltage in IEEE 14 bus system

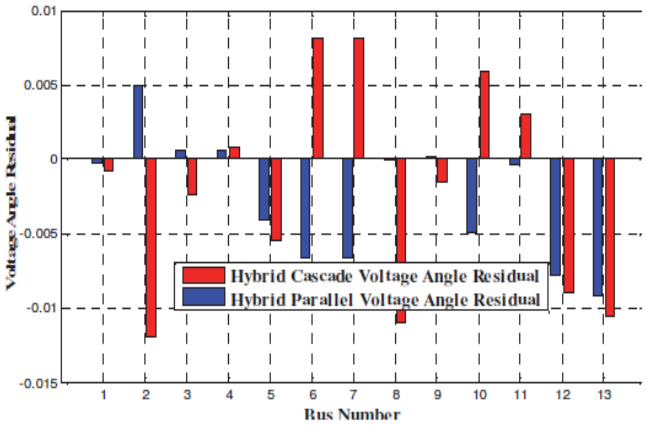

Figure A.4: Hybrid Parallel and Cascade topologies, voltage angle residual in IEEE 14 bus system

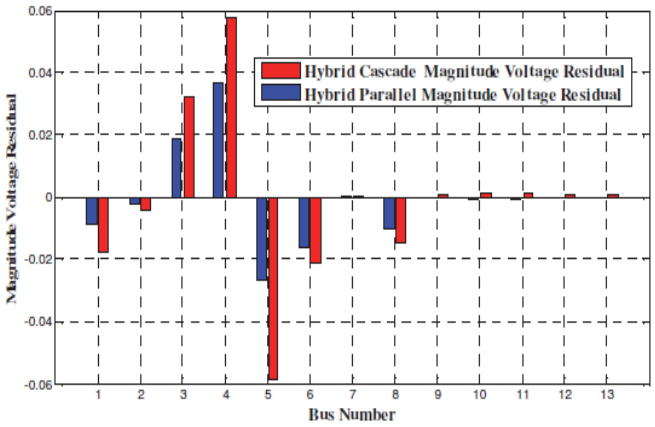

Figure A.5: Hybrid Parallel and Cascade topologies, magnitude voltage residual in IEEE 14 bus system

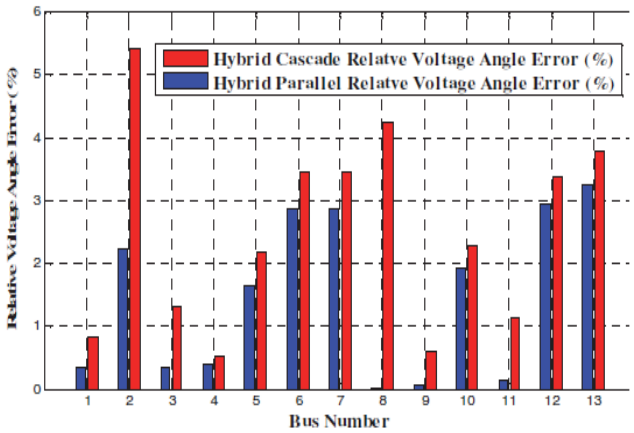

Figure A.6: Hybrid Parallel and Cascade topologies, Relative voltage angle Error in IEEE 14 bus system

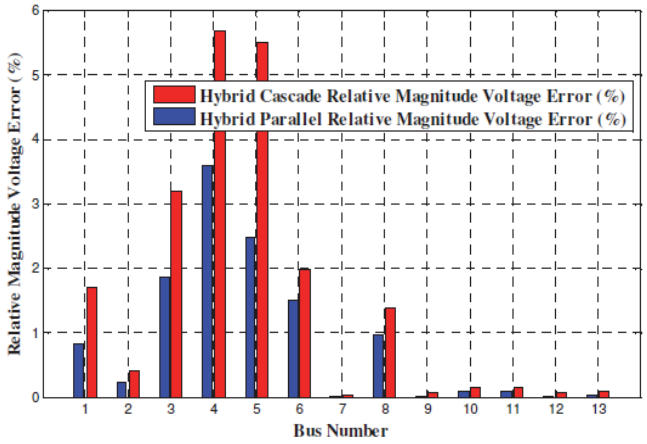

Figure A.7: Hybrid Parallel and Cascade topologies, Relative magnitude voltage Error in IEEE 14 bus system

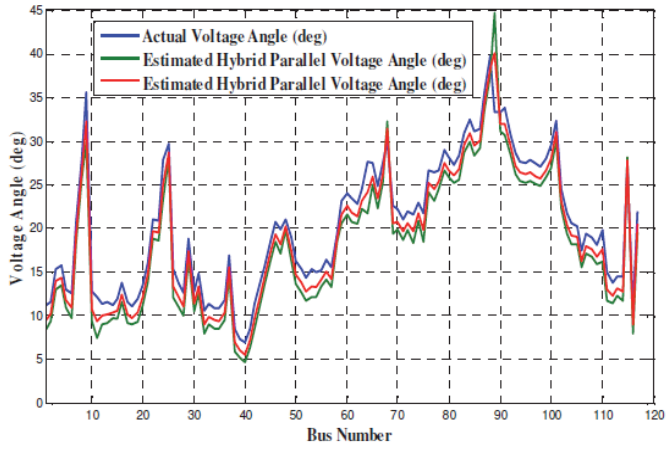

Figure A.8: Hybrid Parallel and Cascade topologies, Estimated and Actual voltage angle in IEEE 118 bus system

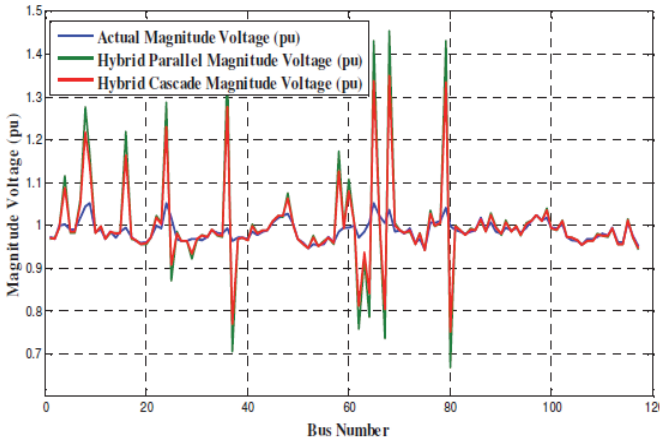

Figure A.9: Hybrid Parallel and Cascade topologies, Estimated and Actual magnitude voltage in IEEE 118 bus system

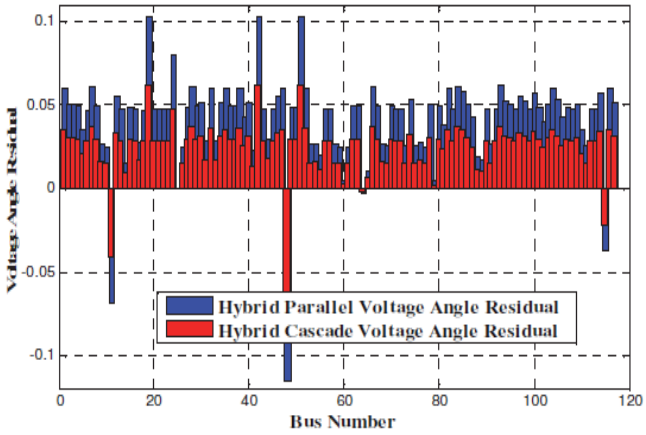

Figure A.10: Hybrid Parallel and Cascade topologies, voltage angle residual in IEEE 118 bus system

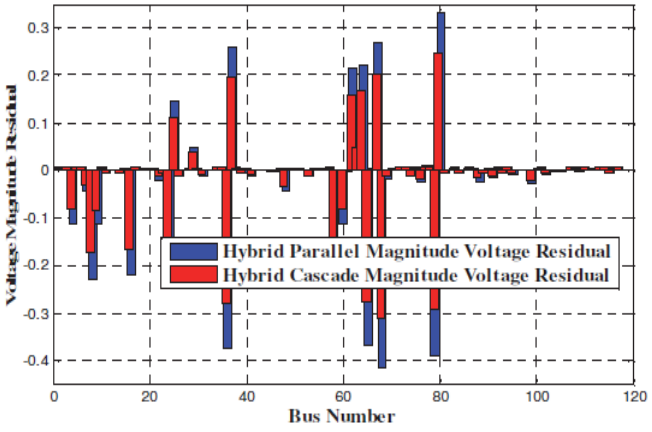

Figure A.11: Hybrid Parallel and Cascade topologies, Magnitude voltage Residual in IEEE 118 bus system

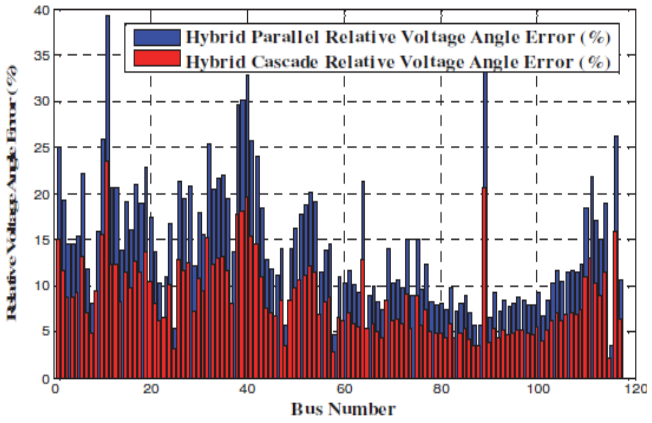

Figure A.12: Hybrid Parallel and Cascade topologies, Relative voltage angle Error in IEEE 118 bus system

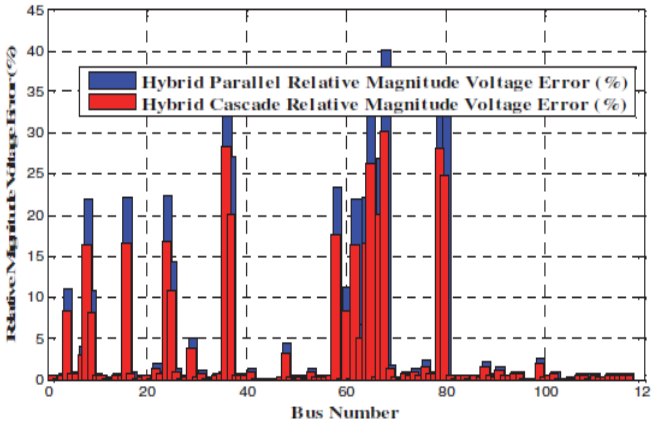

Figure A.13: Hybrid Parallel and Cascade topologies, Relative magnitude voltage Error in IEEE 118 bus system

Table A.3: Comparison Parallel and Cascade Hybrid Connection in Regression, Time, and Neurons

Bus	**Regression**		**Time(Secs)**				**No. Of Neurons**		
System									
	Cascade	**Parallel**	**Cascade**		**Parallel**		**Net1**	**Net2**	**Net3**
			Trai	Sim	Trai	Sim			
IEEE 5	0.997	0.9989	25.18	0.088	15.78	0.078	4	5	6
IEEE 14	0.9103	0.9367	53.45	0.092	27.52	0.083	9	10	11
IEEE 30	0.8507	0.9124	81.75	0.095	34.54	0.089	10	12	14
IEEE 57	0.976	0.9696	110.02	0.098	45.26	0.091	14	15	16
IEEE 118	0.7505	0.6095	150.75	0.098	60.32	0.094	19	20	21

Table A.1: Comparison Parallel and Cascade Hybrid Connection in Voltage Angle Vector

Bus System	**Max.APE**		**RMSE**		**MSE**		**MAPE**	
	Cascade	**Parallel**	**Cascade**	**Parallel**	**Cascade**	**Parallel**	**Cascade**	**Parallel**
IEEE5	3.03×10^{-1}	3.34×10^{-1}	1.24×10^{-1}	1.19×10^{-1}	2.27×10^{-4}	3.48×10^{-4}	3.38×10^{-3}	3.50×10^{-3}
IEEE14	5.41	3.245	7.76×10^{-2}	5.96×10^{-2}	5.14×10^{-5}	2.29×10^{-5}	4.63×10^{-4}	2.74×10^{-4}
IEEE30	8.55	7.69	7.41×10^{-2}	7.03×10^{-2}	5.28×10^{-5}	4.28×10^{-5}	1.89×10^{-4}	1.70×10^{-4}
IEEE57	12.09	18.99	5.61×10^{-2}	7.04×10^{-5}	3.76×10^{-5}	3.10×10^{-1}	5.63×10^{-3}	8.84×10^{-5}
IEEE118	23.56	39.27	1.64×10^{-1}	2.12×10^{-1}	8.44×10^{-4}	2.30×10^{-3}	2.20×10^{-4}	3.62×10^{-4}

Table A.2: Comparison Parallel and Cascade Hybrid Connection in Voltage Magnitude Vector

Bus System	**Max.APE**		**RMSE**		**MSE**		**MAPE**	
	Cascade	**Parallel**	**Cascade**	**Parallel**	**Cascade**	**Parallel**	**Cascade**	**Parallel**
IEEE5	6.47	1.66	1.51×10^{-1}	1.03×10^{-1}	1.20×10^{-3}	1.56×10^{-4}	5.70×10^{-3}	2.70×10^{-3}
IEEE14	5.69	3.60	1.28×10^{-1}	9.66×10^{-2}	6.83×10^{-4}	2.20×10^{-4}	1.30×10^{-3}	7.19×10^{-4}
IEEE30	10.48	6.89	1.22×10^{-1}	9.87×10^{-2}	9.94×10^{-4}	4.30×10^{-4}	5.11×10^{-4}	3.37×10^{-4}
IEEE57	10.07	14.55	1.26×10^{-1}	1.51×10^{-1}	7.30×10^{-4}	1.50×10^{-3}	2.81×10^{-4}	4.07×10^{-4}
IEEE118	30.29	40.39	1.78×10^{-1}	2.06×10^{-1}	5.90×10^{-3}	1.04×10^{-2}	2.71×10^{-4}	3.62×10^{-4}

Table A.4: Weights For different Topologies

Bus System	Cascade-Parallel in Cascade W1	Parallel-Cascade in Parallel W2	Cascade-Parallel in Parallel W3	Cascade-Parallel in Parallel W4
IEEE5	1	0.6	0.98	0.95
IEEE14	1	0	0.95	1
IEEE30	1	1	1	0
IEEE57	1	1	1	1
IEEE118	1	1	0	1

From the above figures, we concluded that in small scale electric grid, both of hybrid Cascade and Parallel topologies are efficient, however hybrid Parallel topology, approximated or averaging topology, is more accurate and smooth. Also in large scale electric grid, hybrid Cascade topology, accumulated error topology, is proficient compared to hybrid Parallel topology.

Tables A.1 and A.2 offer a comparison of the hybrid cascade and parallel topology in both of voltage angle and magnitude voltage vectors in several IEEE bus systems. We conclude that Parallel topology, approximated topology, is only efficient in small scale electric grid and also both topologies work accurately which is exhibited in relative error within 10%. Moreover, in large scale electric grid, accumulated error topology, hybrid Cascade topology can be very robust based on the efficient algorithms in each net. In IEEE 118 bus system, the relative error seems high for both of hybrid topologies and it is out of 10% because when we add more data, the percentage of corrupted measurements is increased to affect accuracy of the system. We have applied IEEE 300 bus system, but unfortunately it did not give a proper convergence because of unreliable data or multilayer neural network is not successful in large scale electric grid.

Table A.3 discusses training, simulating times, and number of neurons, we conclude that hybrid parallel topology in five IEEE cases bus systems have the lowest time in seconds. For instance in IEEE 5 bus system, hybrid parallel topology takes 15.78 seconds in training time and in simulating time takes 78 milliseconds. However

Table A.5: Ranking Training Algorithms Of Feed Forward Neural Network With 10 Hidden Layers In Hybrid Cascade and Parallel Topologies In IEEE 118 Bus System

NUMBERS	ALGORITHM	Training Time(sec)		Operation Time (sec)		MSE		Regression		MAPE	
		Hybrid Parallel Topology	Hybrid Cascade Topology	Hybrid Parallel Topology	Hybrid Cascade Topology	Hybrid Parallel Topology	Hybrid Cascade Topology	Hybrid Parallel Topology	Hybrid Cascade Topology	Hybrid Parallel Topology	Hybrid Cascade Topology
1	LM	60.32	61.55	0.094	0.094	4.40×10^{-3}	8.80×10^{-3}	0.9588	0.9589	5.95×10^{-2}	4.35×10^{-2}
2	BFG	26.17	34.89	0.077	0.077	5.80×10^{-3}	1.08×10^{-1}	0.9516	0.9516	6.33×10^{-2}	4.31×10^{-2}
3	RP	19.62	28.65	0.07	0.07	9.80×10^{-3}	7.53×10^{-3}	0.9494	0.9585	7.57×10^{-2}	4.50×10^{-2}
4	SCG	15.04	23.44	0.069	0.069	7.90×10^{-2}	6.86×10^{-2}	0.9421	0.9585	6.93×10^{-2}	4.42×10^{-2}
5	OSS	23.41	30.03	0.085	0.085	8.70×10^{-2}	8.11×10^{-2}	0.9314	0.9314	7.08×10^{-1}	4.31×10^{-2}
6	GDX	24.85	43.76	0.061	0.061	1.57×10^{-1}	7.99×10^{-2}	0.8607	0.9535	9.17×10^{-1}	4.63×10^{-2}
7	GDM	90.85	102.82	0.083	0.083	7.85×10^{-1}	5.61×10^{-2}	0.5388	0.0709	1.24	1.34

the hybrid Cascade topology takes 25.18 seconds in training time and 88 milliseconds in simulation. Also in IEEE 118 bus system, hybrid parallel topology takes 60.32 seconds in training time and in simulating time takes 94 milliseconds. However the hybrid Cascade topology takes 150.75 seconds in training time and 98 milliseconds in simulation. Neurons should be increased when number of real time available measurements are high to do more computations to obtain the optimal solution.

Table A.4 offers the weights for three different nets in five IEEE bus system cases. There is no distinction in interconnection sequence because all networks are in parallel and the final output is enhanced contrasted with the output of each network due to the averaging component in the parallel analysis topology. A superior outcome may be obtained by acquired a higher weight to the best performing network and also inferior topology set to zero weight. For instance, In IEEE 118 bus system, the superior topologies include w1 (Cascade-Parallel in Cascade topology weight) = 1.0, w3 (Cascade-Parallel in Parallel topology weight) =0.98, w4 (Parallel-Cascade in Parallel topology weight) =0.95. The higher weight is because these topologies contributed to improve hybrid parallel topology, w2 (Parallel-Cascade in Cascade topology weight) = 0.60, the lowest weight is given to the network that makes the total, averaging parallel topology inefficient.

Table A.5 shows an evaluation of each network in the hybrid Cascade and Parallel topologies for IEEE 118 bus system, three different Feed forward networks is trained with the same input- output data. Each network has 10 hidden layers. In terms of accuracy, the Levenberg-Marquardt algorithm (LM) scores the most accurate performance exhibited by MSE = 7.53E-3 and MAPE = 4.35E-2 compared to the Resilient Backpropagation (RP), MSE = 8.80E-3 and MAPE = 4.50E-2. The variable Learning Rate Back propagation (GDX), shows high error. In terms of execution time, LM takes 61.55 seconds in training to compute the optimum solution compared to Scaled conjugate gradient backpropagation (SCG) which has the lowest operating time of 23.44 seconds. Gradient descent with momentum back propagation (GDM) is the most inefficient algorithm in both accuracy and operating time.

We chose to set fit net and feed forward to utilize the diverse three multilayers nets and it is working efficiently. Yet we are searching for more improvement of neural

network, in large scale electric grid such as IEEE 118, we set the the Levenberg-Marquardt algorithm (LM) algorithm for the three layers and which gives the most proficient results with little change in minimizing the error with less training and simulating time. We tried to expand the quantity of neurons, yet this did not enhance the performance of the multilayers neural system. We applied a few algorithms with back propagation neural network and the best one is the Bayesian regulation regulation back-propagation algorithm, yet it takes quite a while to finish training. We set feed forward net for three different nets and LM algorithm for net1, RP for net2, and SCG for net3. Also we set the transfer or activation function to hyperbolic tangent sigmoid transfer function for two layers, input and output layers to obtain better results.

Finally at last we chose to set LM algorithm for Cascade and Parallel training functions to improve the performance. So finally we got the most precise solution in all connections. In addition to that, in large scale electric grid, we have attempted to change the quantity of hidden layers to be 25 neurons for net1, 30 neurons for net2, and 36 neurons for net3 because we want to do more computations to obtain good results, but, we end up with over-fitting, so we decided to not give high values of the neurons.

A.5 Conclusion

This paper has proposed a new computational technique in backpropagation multilayers neural network to estimate state vectors including angle and magnitude voltage based on real time available measurements including scheduled net active and reactive power injection and net active and reactive power nonlinear injections equations for diverse IEEE bus systems. We associated three different nets in a cascade and a parallel topology to obtain the best performance based on simulation results. Conducting analysis, performance of each topology, it is summed up that in small scale electric grid, the hybrid parallel topology which averages the output errors slightly enhances the performance in the systems. On the other hand, the Cascade topology collects the error from each training stage which mostly enhances the overall performance in the systems with large scale electric grid.

Appendix B

Power System Tracking State Estimation Based on Stochastic Fractal Search Technique under Bad Measurements Conditions

B.1 Introduction

Static state estimation evaluates the state vectors consisting of all bus voltage magnitudes and angles in an electric power system. Bad measurements need to be recognized and purged to enable the state estimator to provide dependable information base for power system operation and control [140], [141].

To allow continuing monitoring and controlling of the electric power grid, state estimation must be conducted at short intervals of time. Yet, as the electric grid grows, with expansion of power generations, service areas, and loads, the system turns out to be too vast for the state estimation to be done at short intervals of time. This is the motivation behind Debs and Larson's "Tracking State Estimation (TSE) [130], [131]. This estimator relies on modeling the dynamics of the system over short time intervals. The TSE is coupled to a basic anomaly identification plan, to prevent any significant error in the modeling of the system and the measurements.

Schweppes and Masiello extended this work by presenting a TSE as a digital feedback loop which utilizes new measurements to determine new estimates by correcting an old estimate by adding a feedback error signal modified through a gain matrix [132].

Reference [133] proposed a TSE strategy for coordinated AC/DC systems, using rectangular coordinates for both AC and DC measurement equations.

The present chapter proposes using Fractal Search algorithm (FS) to solve the TSE problem under normal operation where the load increases linearly from 20% to 130% with trend of 4% each time instant for 24 consecutive instants. Bad measurements are introduced as well. Performance indices are evaluated by maximum mean

square error (MMSE), average absolute error (AAE), maximum root mean square error (MRMSE), and maximum sum of the square error (MSSE). The effectiveness of the FS technique has been examined on four IEEE standard test systems (5, 14, 30, and 57 bus systems.) The results of the proposed technique are compared with those obtained using Genetic Algorithm (GA) and Particle Swarm Optimization (PSO).

B.2 Tracking State Estimation

Real time measurements of the power system are transferred to the control center through a remote terminal unit (RTU). These measurements may contain bad ones because of communication and systematic errors, erroneous wiring or lack of instrument adjustment.

Tracking state estimation (TSE), determines the state variable and updates for the next instant of time k+1 from a new set of measurement data. The power system is thought to be a semi static system. Henceforth the adjustments in the system happen gradually. It implies that the state may not change much in a short period of time. In some cases, it turns out to be important to monitor events in the system, such as through the picking up of load by a generator or through some contingency.

The SCADA system, which performs the continuous monitoring, and control of power systems receives field measurements through sensors and sends them to the control center at defined instants of time. To have a continuous monitoring system, state estimation must be executed as and when each new measurement set arrives. In any case, as state estimation is computationally intensive (particularly as the system size gets greater), many control centers will not have adequate processing assets to perform a fast and accurate state estimate at high frequency. If the length of time between two instants is too substantial, it results in a weak co-connection between the estimated states, rendering the detected bad measurements extremely erroneous [142].

Tracking state estimation is utilized to give up-to-date results [143]. In this paper, TSE is implemented according to the method described in [118]. Without access to reasonable field measurements, Newton Raphson load flow analysis is executed for each time instant on the test systems under study to produce simulated measurements. In order to make the simulated data resemble field data, an arbitrarily

generated data error of 5% is added according to the technique described in [136]. The last measurement vector thus acquired acts as an input to the proposed stochastic fractal search state estimator.

B.3 Solution Methodology of SFS

The SFS technique begins with initializing the population of size N_p that follows a uniform distribution which is considered as a start point with equal energy E_i. Two statistical methods should be considered to generate new particles from the diffusing procedures including Levy flight and Gaussian random walk [82]. In term of fast convergence, Levy flight is better in a few generations, however, Gaussian random walk is more guarantee in determining the global minima. SFS algorithm utilize Gaussian random walk distribution is the only arbitrarily search walk employed in the diffusion limited aggregation (DLA) growth procedure using the following equations,

$$GW_1 = Gaussian(\mu_{BP}, \sigma) + (\varepsilon \times BP - \varepsilon^{'} \times P_i) \tag{B.1}$$

$$GW_2 = Gaussian(\mu_P, \sigma) \tag{B.2}$$

Here ε and $\varepsilon^{'}$ are arbitrarily distributed and limited to [0, 1], BP and P_i are indicated as the best position point and i^th point in the group respectively. $\mu_B P$ is equal to $|BP|$, μ_p is equal to P_i and σ is the standard deviation. The fitness function of each point is determined to obtain the best point (BP) among all points. As indicated by the utilization of diffusion property, every particle is diffused around its current position to allow proper exploitation of the search space. Two statistical procedures intended to build the better space investigation are considered because of the exploitation property. The principal statistical method performs on every individual vector index, and the second statistical technique is then applied to all points [82].

Each point is initialized arbitrarily, so constraints should be selected to describe the minimum and the maximum bounds.

$$P_(j) = LB + \varepsilon \times X(UB - LB) \tag{B.3}$$

Where LB and UB are the lower and the upper limits of search space and ε is a uniformly distributed arbitrary which is limited to [0, 1]. For the principal statistical technique, each of the points are positioned based on the estimates of the fitness function. Each point i in the assembly is then given a likelihood probability which complies with a straightforward uniform distribution.

$$P_{ai} = \frac{rank(P_i)}{N} \tag{B.4}$$

Where the rank (P_i) is the rank of the P_i point among all different points in the group and N is the total number of all points in the group. Eq. (B.4) is used to maximize the probability of re-positioning the points when the optimal solution is hard to find [144]. The points' positions are updated according to Eq. (B.5).

$$P_i^{'}(j) = P_r(j) - \varepsilon \times (P_t(j) - P_i(j)) \tag{B.5}$$

Where $P_i^{'}$ is the new updated position of P_i and both of P_r and P_t are arbitrary chosen points in the group. The second statistical change is expected to change the position of a point considering the position of different points in the gathering. This property enhances the nature of investigation, and it fulfills the diversification property. All points acquired from the first statistical method are positioned taking into account Eq. (B.5). If the condition $P_{a_i} < e$ holds for a new point Pi, the present position of Pi is adjusted by Eqs.B.6 and B.7.

$$P_i^{"} = P_i^{'} - \varepsilon^{'} \times X(P_t^{'} - BP) > 0.5 \tag{B.6}$$

$$P_i^{"} = P_i^{'} - \varepsilon^{'} \times X(P_t^{'} - P_r^{'}) < 0.5 \tag{B.7}$$

Where $P_i^{"}$ and $P_r^{'}$ are arbitrary chosen point and $\varepsilon^{'}$ arbitrary number created by Gaussian normal distribution (GND).

B.4 Simulation Study

The main object of this simulation experiment is to analyze and to minimize the residual based on SFS technique. The validation of the performance has been carried

out with bad measurements data.

B.4.1 WLAV

The weighted Least Absolute Value is formulated to estimate unknown state vectors by minimizing the following objective,

$$J(x) = \sum_{i=1}^{m} w_i |z - h(x)| \tag{B.8}$$

The nonlinear formulae $f_i(x)$ are functions of the state vectors to be estimated. $f_i(x)$ include active and reactive power injections and flows [135].

B.4.2 Constraints

Both of the magnitude and angle of the voltages are constrained in order to meet the operating requirements as follows,

$$\begin{aligned} \mathrm{v}_i^{min} \leq v_i \leq v_i^{max} \\ \delta_i^{min} \leq \delta_i \leq \delta_i^{max} \end{aligned} \tag{B.9}$$

B.4.3 Bad Measurements Condition

Bad measurements are introduced at several time instants for all test systems as shown. For all test cases, the following bad measurements are observed.

1. Introducing one bad data measurement at 3^{rd} time instant.
2. Introducing two bad data measurements at 6^{th} time instant.
3. Introducing three bad data measurement at 10^{th} time instant.
4. Introducing four bad data measurements at 14^{th} instant.
5. Introducing five bad data measurements at 20^{th} instant.
6. Introducing six bad data measurements at 24^{th} instant.

B.4.4 Test Systems

The performance of the technique has been examined on four test systems, viz. IEEE 5-bus, IEEE 14-bus, IEEE 30-bus, and IEEE 57-bus systems [113].

B.4.5 Performance Evaluations

Several performance indices and statistical parameters have been used to evaluate the performance of the proposed algorithm and its comparison is carried out with Particle Swarm Optimization (PSO) and Genetic Algorithm (GA).

1. Filter performance is introduced to evaluate the overall estimation for all IEEE bus test systems as shown in Eq.(5.19).

2. The maximum of mean square error of the state estimate is calculated according to

$$MMSE = Max[\frac{1}{2(N_i-1)}\sum_{K=1}^{2(N_i-1)}[\hat{x}_i(k) - x_i^t(k)]^2] \tag{B.10}$$

Where N_i is number of time instants,$\hat{x}_i(k)$ and $x_i^t(k)$ are estimated and true state vectors at kth time instant.

3. The maximum of sum of square error of the state estimate is computed according to

$$MSSE = Max[\sum_{K=1}^{2(N_i-1)}[\hat{x}_i(k) - x_i^t(k)]^2] \tag{B.11}$$

4.The average of absolute error of the state estimate is calculated according to

$$AAE = \frac{1}{2(N_i-1)}[\sum_{K=1}^{2(N_i-1)}|\hat{x}_i(k) - x_i^t(k)|] \tag{B.12}$$

5. The maximum of the root mean square error of state estimate is computed according to

$$MMSE = Max\sqrt{\frac{1}{2(N_i-1)}[\sum_{K=1}^{2(N_i-1)}[\hat{x}_i(k) - x_i^t(k)]^2]} \tag{B.13}$$

B.5 Results and Discussion

The proposed technique is implemented on four IEEE test systems. The estimation performance of the proposed technique is assessed using time data from the American Electric Power Systems in the Midwestern of US which is published by the official

website of University of Washington [113]. Several scenarios have been applied under bad measurements.

Table B.1: Control Parameters settings

Algorithm	Control Parameters settings
SFS	Population Size= 50 Maximum No of generation =400 Maximum No of diffusing =9 Maximum No of diffusing Walk =9 No. of dimension= No. of state vectors
PSO	Swarm Size= 20 Maximum No of generation =50 maximum particle velocity=4 Learning factors, $c_1 = 2$, $c_2 = 2$ Inertia weight,$w_{initial} = 0.9$, $w_{final} = 0.2$
GA	Population Size= 20 Crossover rate=0.7 Crossover type: arithmetic Mutation rate=0.0175 Mutation type: non-uniform Maximum No of generation =50

Table B.2: Computational time(s) taken by proposed SFS technique

Test Systems	Total CPU Time(s)	
	Normal Condition	Bad data Condition
IEEE 5-bus	0.4991	0.5165
IEEE 14-bus	0.9112	1.2136
IEEE 30-bus	2.1168	3.0598
IEEE 57-bus	5.6807	7.7146

Table B.1 shows the selected control parameters for the proposed technique, stochastic fractal search with other two algorithms include Genetic and Particle Swarm algorithms to validate the proposed technique and to obtain the optimal solution. A detailed description about these two algorithms with their control parameters are described in ref [145], [146].

Table B.2 demonstrates the total computational time of the proposed SFS technique. Table B.2 comprises of three segments indicating the test systems under both

normal operating and bad data conditions. It is obvious that the primary challenge of the proposed SFS technique is a computational time which increase with an increment in the system size. This is due to an increment in the search space and diffusing process for computing the optimal solution. On the other hand, the proposed SFS technique gives more accurate estimates and profoundly robust against outliers contrasted to other methods. One of the suggestion to overcome the drawback of computational time is by decomposing the vast systems into multi-systems. This proposal can decrease the total computational time to be more reasonable.

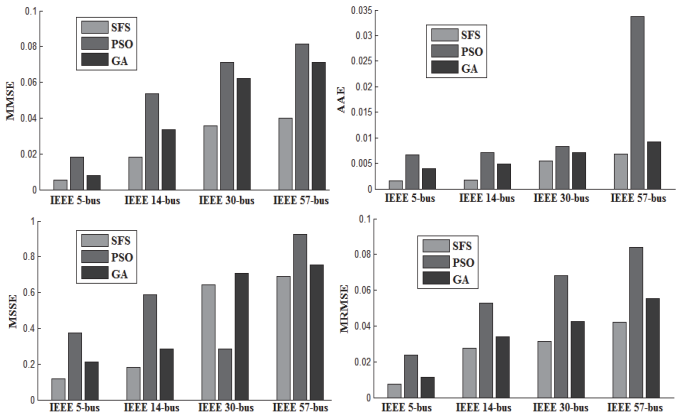

Figure B.1: Statistical parameter values under bad data condition for voltage angle

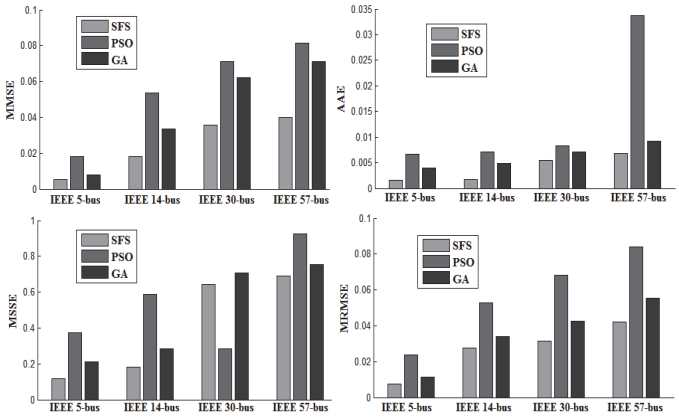

Figure B.2: Statistical parameter values under bad data condition for magnitude voltage

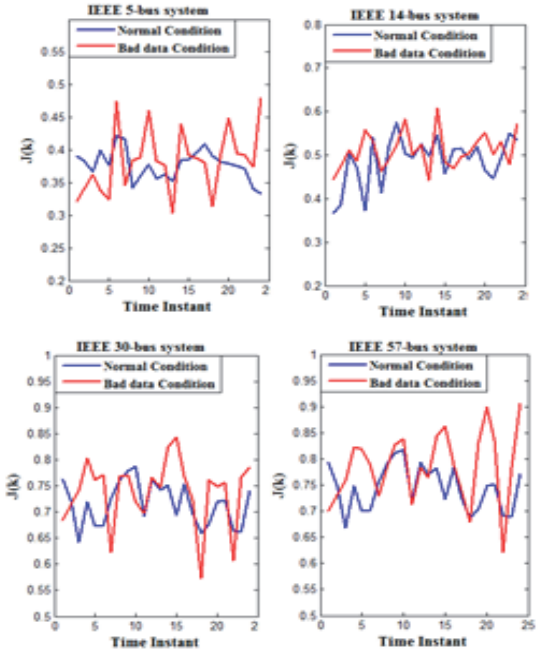

Figure B.3: Performance Index J (k) in the presence of normal and bad data condition

Table B.3 presents the measured and the true state vectors on bus 3 for IEEE 5-bus system for 24 time instants under normal operating conditions. It can be observed that voltage angles are more sensitive to the load change condition as compared to magnitude voltages. Relative error is presented to assess the accuracy in bus 3 during 24 time instant. Also performance index is used to evaluate the behavior of bus 3 during 24 time instants.

Figures B.1 and B.2 compare the statistical parameters values of the proposed SFS estimator under bad data condition. It can be concluded that SFS estimator gives better assessed states contrasted with other SE based strategies. Be that as it may, statistical parameter values of SFS, PSO, and GA based SE method increased with an increment in the system size. This is an evidence that FS estimator is more accurate compared with other SE methods.

Figure B.3 compares the performance index J (k) of the proposed SFS estimator under both normal operation and bad data conditions for several test systems. It can be observed that the performance indices are close for all systems and the proposed

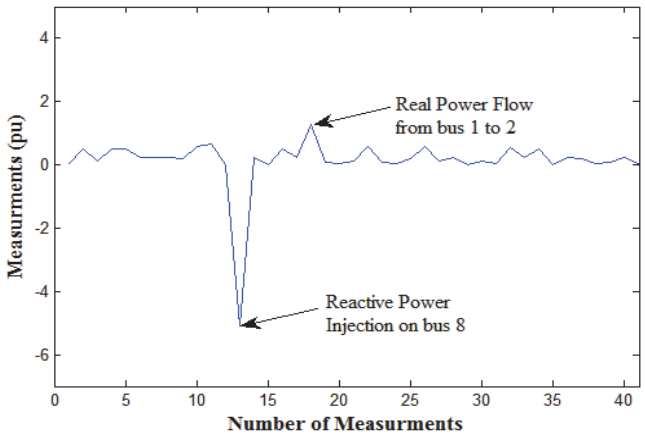

Figure B.4: Normalized Measurements vector for bad data condition during 11th time instant

algorithm gives a competitive technique to solve such hard problems by moving the bad data measurements toward the most fitted distribution.

Figure B.4 demonstrates the normalized measurements computed in per unit according to the primary generator, slack bus generator values include apparent power, 100 MVA and base voltage, 69 KV for IEEE 14-bus system corresponding to 11th time instant. It shows that normalized measurements for bad data condition with anomaly data identification such as reactive power injection on bus 8 and real power flow from bus 1 to 2. These anomalies are eliminated using backward elimination which selects one data which has high anomaly value each step in order to improve the entire estimation.

Table B.3: Comparison Of State Vector, Estimation, and True Values On Bus 3 In IEEE 5-Bus System For 24 Instant Of Time

k	True		Estimation		Error		J(k)	k	True		Estimation		Error		J(k)
	v_3	δ_3	v_3	δ_3	v_3	δ_3			v_3	δ_3	v_3	δ_3	v_3	δ_3	
1	0.97	0.208	0.9987	0.5523	1.89137	0.78467	0.45262	13	1.008	-0.1318	1.013	-0.1377	0.49358	0.58243	0.40739
2	0.97	0.208	0.982	0.2122	1.222	2.01923	0.44339	14	1.01	-0.1479	1.009	-0.1497	0.09911	0.17839	0.44515
3	0.964	0.2427	0.973	0.2595	0.92497	2.69094	0.423	15	1.01	-0.2261	0.9897	-0.216	2.05113	1.02051	0.44669
4	0.96	0.1371	0.971	0.141	1.13285	2.84464	0.46231	16	1.01	-0.3053	1.0044	-0.303	0.55755	0.22899	0.45908
5	0.96	0.1013	0.95	0.1019	1.05263	0.5923	0.43546	17	1.03	-0.1713	1.033	-0.176	0.29042	0.45499	0.47323
6	0.96	0.0653	0.9591	0.0667	0.09384	2.14395	0.488	18	1.03	-0.4104	0.991	-0.4288	3.93542	1.77598	0.45254
7	0.96	0.0291	0.9573	0.0282	0.28204	3.09278	0.48146	19	1.01	-0.4499	0.998	-0.4864	1.2024	3.65731	0.44192
8	0.96	-0.0074	0.9674	-0.0065	0.76494	1.5625	0.39562	20	0.99	-0.4898	0.9867	-0.39	0.33445	0.02027	0.43892
9	0.943	-0.044	0.947	-0.045	0.42239	2.27263	0.41585	21	0.93	-0.3902	0.927	-0.3909	0.32362	0.07551	0.434
10	0.99	-0.0897	0.98	-0.0884	1.02041	1.44928	0.43669	22	0.903	-0.5784	0.892	-0.568	1.23318	1.79806	0.43031
11	0.99	-0.1268	0.9986	-0.127	0.86121	0.15773	0.412	23	0.85	-0.2465	0.869	-0.2718	2.18642	2.91139	0.39385
12	1.0056	0.1771	0.9997	0.1757	0.59018	0.79051	0.41885	24	0.73	-0.0334	0.749	-0.0377	2.62072	0.56653	0.38477

B.6 Conclusion

In this chapter, the stochastic fractal search technique is proposed and implemented to solve tracking state estimation (TSE) problem. The SFS estimator has been applied for tracking of the time varying static state of the power system. The proposed SFS technique has been verified on four IEEE test cases including IEEE 5-bus, IEEE 14-bus, IEEE 30-bus, and IEEE 57-bus systems. Two different conditions including normal operating and bad data conditions have been used to assess the performance of the proposed SFS estimator. The proposed SFS performance is carried out to be compared to other estimators such as Particle Swarm Optimization (PSO) and Genetic Algorithm (GA) in several statistical parameters values. The numerical results demonstrate that the proposed SFS technique has provided us with an accurate state estimates compared with other proposed estimators. However, this technique suffers from the large computational time for large systems.

Appendix C

Power System Tracking State Estimation Based on Stochastic Fractal Search Technique under Sudden Load Changing Conditions

C.1 Introduction

The control center receives measurements transmitted through the SCADA system, which depend on the scanning rate. It utilizes one of two techniques: the snap shot operating, when the set of measurements at a given time is smaller than the base measure of time, and sequential operating if the set of measurements at a given time is larger than the base measure of time [147].

SE is computationally substantial, particularly as the system size gets larger, and many control centers will not have adequate assets to perform an exact and quick state estimates. Be that as it may, when the length of time between two estimates is large, will result in a weak co-connection between the accessed states [148]. Tracking state estimation is used to give better estimation results.

Debs and Larson [131] introduce a minimum variance optimal state estimate based on all the real time measurements up to time t. This procedure operates on all measurements sequentially (Batch Mode,) and is compared with static estimator. The tracking estimator can be analyzed by quadratic-deviation.

In [132] Schweppe and Masiello introduce a tracking static state estimator (TSE) as an algorithm which uses new measurements to create new estimates by amending old estimates by a term related to the feedback error signal.

In [143] the authors introduce three tracking estimators using pre-estimation bad data detection/ elimination scheme based on the exponential smoothing of past estimations and logical checks followed by an estimation stage. The algorithms tested are plain weighted least squares, quadratic square-root and a linear criterion.

Reference [133] authors present a tracking state estimation algorithm for integrated AC/DC systems based on rectangular coordinate formulation of the network equations. The first and second order derivatives of the equations are used to develop a fast tracking state estimator.

This chapter presents the application of a stochastic search technique named fractal search algorithm (SFS) to solve the tracking state estimation (TSE) problem under sudden load changes scenario. The performance of the method is evaluated based on maximum average square state error (MASSE), average absolute state error (AASE), maximum root average square state error (MRASSE), and maximum sum of the square state error (MSSSE). The procedure is tested on four IEEE standard test systems including 5, 14, 30, and 57 bus test systems. The results of the proposed strategy are compared to Genetic Algorithm (GA) and Particle Swarm Optimization (PSO). The solution of assessed state vectors are compared to the true state variables utilizing Newton Raphson Load Flow method (NR-LF). The estimation performance of the proposed technique is assessed utilizing real time data from the American Electric Power System in the Midwestern US which is hosted by the website maintained at the University of Washington [113].

C.2 Sudden Load Changes

Power system loads fluctuate as dictated by pattern cycles, creating deviations of relatively small impact on system dynamic performance. Sudden load changes are not common but may take place due to unscheduled events such as outages of customer loads as well as load curtailment, planned and unscheduled outages of equipment, unplanned climatic conditions, and so forth. System variables including voltages, power flows/injections, transformer taps are monitored and controlled to meet system operational constraints [78].

When a sudden change takes place, the set of nearest state variables could experience larger changes. Consequently, there could be appreciable differences between actual and estimated values of system states. Operators would then question the practical value of DSE [149].

The authors of [150] offered one of the earliest proposed techniques to detect sudden changes in operating condition. The approach compares the anticipated and

the actual estimates and decides whether or not a sudden change condition has taken place. To deal with this situation the significance of suspect data is de-emphasized by reducing their weight and labeling the expectation as untrustworthy.

Gross differences between estimates are considered as bad data within all the measured data. Two principal sources of gross errors include errors in the metering system because of broken switch, inaccurate electrical switch status data. If abnormalities are not identified accurately, then the output of the estimator is unreliable. The computational time taken to evaluate the condition of the power system states is important and there is a need to build up a proficient state estimator to assess the conditions of the power system.

This chapter presents the application of a stochastic search technique named fractal search algorithm (SFS) to solve the tracking state estimation (TSE) problem under sudden load changes scenario. The performance of the method is evaluated based on maximum average square state error (MASSE), average absolute state error (AASE), maximum root average square state error (MRASSE), and maximum sum of the square state error (MSSSE). The procedure is tested on four IEEE standard test systems including 5, 14, 30, and 57 bus test systems. The results of the proposed strategy are compared to Genetic Algorithm (GA) and Particle Swarm Optimization (PSO). The solution of assessed state vectors are compared to the true state variables utilizing Newton Raphson Load Flow method (NR-LF). The estimation performance of the proposed technique is assessed utilizing real time data from the American Electric Power System in the Midwestern US which is hosted by the website maintained at the University of Washington [113].

C.3 Solution Methodology of SFS algorithm

Fractals were proposed by Benoit Mandelbrot in 1975. He attempted to utilize the idea of fractal speculations to depict patterns in nature [144]. The Stochastic Fractal Search is based on this concept and maybe outlined as follows [82]:

SFS Procedure

1. Initializing the population size N_p that follows a uniform distribution which is considered as a starting point with equal energy E_i, maximum generation (iteration), maximum diffusing number (MDN), and the dimension size N_dim.
2. Set diffusing walk which is considered as Search walk=1 for first Gaussian (simple problems), and diffusing Search walk=2 for second Gaussian (hard problems). Any choice of second order Gaussian distribution should consider increasing number of generations (iterations). SFS algorithm utilize Gaussian distribution, search walk employed in the DLA growth procedure.
3. Set time instant k=1.
4. Minimizing the fitness function.
5. Constraints should be selected.
6. Generating random points in considered search space.
7. Computing the fitness value of the first created point.
8. Find the best point (BP) in the group.
9. First updating the points' position .
10. All points obtained by the first update process are ranked for each new point.
11. Second updating Process.
12. Steps from 2-9 are then repeated until the number of generations reaches the maximum.
13. Update the time instant k as k=k+1.
14. Stop.

C.4 Simulation Study

C.4.1 Sudden Load Change

The execution of the proposed technique has been examined under sudden burden change conditions. Measurement set for this condition has been resolved under the following situations,

For test case 1, i.e. IEEE 5-bus system, the following sudden load change conditions are observed,

1) 50% load reduction at 3^th time instant on bus 2.
2) Sudden loss of load at 11^th time instant on bus 3.
3) 50% load increase at 16^th time instant on bus 4.
4) 20% load increase at 19^th time sample on all the load buses.
5) Sudden loss of load at 22^th time instant on bus 5.
6) 20% load decrease at 24^th times instant on all the load buses.

For test case 2, i.e. IEEE 14-bus system, the following sudden load change conditions are observed:
1) 50% load reduction at 3^th time instant on bus 3.
2) Sudden loss of load at 11^th time instant on bus 8.
3) 50% load increase at 16^th time instant on bus 12.
4) 20% load increase at 19^th time sample on all the load buses.
5) Sudden loss of load at 22^th time instant on bus 14.
6) 20% load decrease at 24^th times instant on all the load buses.

For test case 3, i.e. IEEE 30-bus system, the following sudden load change conditions are observed:
1) 50% load reduction at 3^th time instant on bus 3.
2) Sudden loss of load at 11^th time instant on bus 8.
3) 50% load increase at 16^th time instant on bus 12.
4) 20% load increase at 19^th time sample on all the load buses.
5) Sudden loss of load at 22^th time instant on bus 14.
6) 20% load decrease at 24^th times instant on all the load buses.

For test case 4, i.e. IEEE 30-bus system, the following sudden load change conditions are observed:
1) 50% load reduction at 3^th time instant on bus 5.
2) Sudden loss of load at 11^th time instant on bus 10.
3) 50% load increase at 16^th time instant on bus 15.
4) 20% load increase at 19^th time sample on all the load buses.
5) Sudden loss of load at 22^th time instant on bus 26.

6) 20% load decrease at 24^th times instant on all the load buses.

For test case 5, i.e. IEEE 57-bus system, the following sudden load change conditions are observed:
1) 50% load reduction at 3^th time instant on bus 5.
2) Sudden loss of load at 11^th time instant on bus 13.
3) 50% load increase at 16^th time instant on bus 25.
4) 20% load increase at 19^th time sample on all the load buses.
5) Sudden loss of load at 22^th time instant on bus 45.
6) 20% load decrease at 24^th times instant on all the load buses.

C.4.2 Test System

The execution of this analysis has been applied on four test systems, viz. IEEE 5-bus, IEEE 14-bus, IEEE 30-bus, and IEEE 57-bus systems.

C.4.3 Performance Indices

A few execution indices and parameters have been used to evaluate the performance utilized to assess the execution of the proposed [151].

C.5 Results and Discussion

Table C.1: Control Parameters settings

Algorithm	Control Parameters settings
SFS	Population Size= 50 Maximum No of generation =400 Maximum No of diffusing =9 Maximum No of diffusing Walk =9 No. of dimension= No. of state vectors

Table C.2: Computational time(s) taken by proposed SFS technique

Test Systems	Sudden Load Change Condition	
	First CPU Time(s)	Total CPU Time(s)
IEEE 5-bus	0.0309	0.7479
IEEE 14-bus	0.0575	1.4573
IEEE 30-bus	1.0338	3.2231
IEEE 57-bus	1.6807	7.9832

Table C.1 illustrates the algorithm control parameters which are chosen to obtain the optimal solution.

Table C.2 demonstrates the first and total computational time of the proposed SFS technique under sudden load condition. Also this table consists of three segments showing the test systems under consideration, the entire computational time, and the time taken for the initial appearance of the solution with adequate precision respectively. It is seen from the table that the real problem of the proposed SFS procedure is the computational time which increments with an increment in the system size. This is because of an increment in the search space and diffusing process for determining the optimal solution. Although, the proposed SFS technique gives more accurate appraisals and profoundly vigorous against outliers contrasted with other techniques. One of the suggestion to overcome the drawbacks of computational time is by decomposing the vast systems into multi-systems. This proposal can decrease the total computational time to be more reasonable. Moreover, the heavy computational burden can be solved by compelling programming and with an increment of PC speeds as being seen each year. Under these conditions, the transformative methodology should be quick soon. Subsequently, in the present day situation the proposed SFS technique is a viable and effective apparatus for disconnected from the net studies (offline studies).

Table C.3 merely offers estimated and true values of V3, and 3 on bus 3 on IEEE 5 bus system for 24 time instants. It concluded that errors are within 10% which makes our estimation to be accurate. Also performance assessment J (k) for this case is represented in table C.3, measurements are not effected and the performance is far enough from one which makes our estimation to be reliable and robust. Four statistical parameters of the SFS technique based SE strategy is exhibited in Table C.4 and C.5 for voltage angle and magnitude voltages respectively. From the aforementioned

Table C.3: Comparison Of State Vector, Estimation, and True Values On Bus 3 In IEEE 5-Bus System For 24 Instant Of Time

k	True		Estimation		Error		J(k)	k	True		Estimation		Error		J(k)
	v_3	δ_3	v_3	δ_3	v_3	δ_3			v_3	δ_3	v_3	δ_3	v_3	δ_3	
1	0.9700	0.2432	1.0027	0.1951	3.3711	5.4276	0.4089	13	0.972	-0.3091	0.9508	-0.4272	2.1779	4.5293	0.5143
2	0.9700	0.1992	0.9557	0.1623	1.4742	2.0582	0.4182	14	0.9729	-0.4077	0.9532	-0.479	2.0249	4.7829	0.4658
3	0.9623	0.1549	0.9738	0.1106	1.1951	4.7773	0.3842	15	0.9732	-0.458	0.9212	-0.8267	5.347	4.5852	0.3068
4	0.9723	0.1103	0.943	0.0629	3.0135	0.2719	0.4235	16	0.9689	-0.8767	0.9252	-0.5344	4.5162	5.7032	0.5863
5	0.9700	0.0653	0.9472	0.0211	2.3505	3.7366	0.4268	17	0.9700	-0.5606	0.9473	-0.6516	2.3454	4.6736	0.3511
6	0.9700	0.02	0.9155	-0.0262	5.6186	5.35	0.4282	18	0.9700	-0.6131	0.9183	-0.8362	5.3299	6.2796	0.4429
7	0.9700	-0.0257	0.9304	-0.0715	4.0825	1.7509	0.4269	19	0.9654	-0.8895	0.9083	-0.7211	5.9176	5.9888	0.5845
8	0.9700	-0.0718	0.9791	-0.1159	0.9381	0.4178	0.4283	20	0.9721	-0.7206	0.9572	-0.7567	1.5328	0.0666	0.3074
9	0.963	-0.1183	1.0129	-0.1623	5.1817	2.0287	0.4389	21	0.9723	-0.7758	0.9782	-0.5261	0.6119	2.4619	0.5571
10	0.965	-0.1653	0.9831	0.6321	1.8757	1.833	0.4466	22	0.9700	-0.5377	0.9797	-0.8523	0.9979	2.1573	0.3591
11	1.0236	0.6108	0.9775	-0.2556	4.5037	3.4872	0.4201	23	0.97	-0.8895	0.9702	-0.7569	0.0237	4.1821	0.5832
12	0.9873	-0.2606	0.9503	-0.2951	3.7495	1.9187	0.4754	24	0.9769	-0.7206	0.9895	-0.4272	1.2876	5.0375	0.5466

Table C.4: Statistical Parameter Values Under Sudden Load Condition For Voltage Angle

Test Systems	MASSE			AASE			MSSSE		
IEEE	SFS	PSO	GA	SFS	PSO	GA	SFS	PSO	GA
5-bus	6.1×10^{-2}	9.9×10^{-2}	6.8×10^{-2}	6.1×10^{-3}	1.3×10^{-2}	8.6×10^{-3}	4.4×10^{-1}	9.8×10^{-1}	6.8×10^{-1}
14-bus	8.2×10^{-2}	1.3×10^{-1}	9.9×10^{-2}	1.3×10^{-2}	8.8×10^{-2}	4.5×10^{-2}	7.5×10^{-1}	1.01×10^{0}	8.7×10^{-1}
30-bus	2.4×10^{-1}	5.6×10^{-1}	3.8×10^{-1}	5.8×10^{-2}	1.8×10^{-1}	9.9×10^{-2}	9.2×10^{-1}	1.6×10^{0}	9.9×10^{-1}
57-bus	6.8×10^{-1}	9.9×10^{-1}	8.7×10^{-1}	1.4×10^{-1}	9.9×10^{-1}	6.7×10^{-1}	9.9×10^{-1}	1.35×10^{0}	1.16×10^{0}

Table C.5: Statistical Parameter Values Under Sudden Load Condition For Voltage Magnitude

Test Systems	MASSE			AASE			MSSSE		
IEEE	SFS	PSO	GA	SFS	PSO	GA	SFS	PSO	GA
5-bus	1.0×10^{-2}	7.9×10^{-2}	4.6×10^{-2}	1.6×10^{-3}	7.8×10^{-3}	3.5×10^{-3}	5.8×10^{-1}	9.9×10^{-1}	7.9×10^{-1}
14-bus	2.2×10^{-2}	9.9×10^{-2}	6.5×10^{-2}	1.1×10^{-2}	6.7×10^{-2}	3.5×10^{-2}	6.9×10^{-1}	1.09×10^{0}	9.9×10^{-1}
30-bus	8.5×10^{-2}	1.4×10^{-1}	9.9×10^{-1}	6.3×10^{-2}	9.6×10^{-2}	8.9×10^{-2}	8.1×10^{-1}	1.16×10^{0}	1.08×10^{-1}
57-bus	3.4×10^{-1}	9.9×10^{-1}	6.7×10^{-1}	1.3×10^{-1}	4.6×10^{-1}	2.0×10^{-1}	9.9×10^{-1}	1.46×10^{0}	1.13×10^{0}

tables, it can be seen that like ordinary operation, the statistical parameter parameters are little influenced by these measurements, however statistical parameters values for sudden load variety conditions are reasonable compared to other techniques. Subsequently it can be inferred that the proposed SFS estimator gives better estimates even in the exhibiting of sudden load changes. This shows the robustness of the proposed SFS strategy.

Figure C.1 demonstrates the comparison in the presence of the performance index J (k) of the proposed SFS estimator under sudden load condition for several test systems during 24 instant of time. It can be observed from figure C.1 that the distinction in performance variation index is little far for all case systems, subsequently, in IEEE 5-bus system, the variation is more obvious in the last ten instant of time where most of sudden load changes occurred from 16th to 24th time instant. For other cases, performance variation J (k) is distributed adequately for all time instants. Due to the robustness, the proposed SFS technique is a competitive technique to solve such hard problems by moving the influenced measurements, by different circumstances and conditions, toward the most fitted distribution. It is affirmed from figure C.1 that the estimated results are influenced after the presence of sudden load changes. The evaluation of robustness is taken as an assumption of performance index value J (k) being less than one. In other words, as long as the performance index of the proposed SFS technique less than one, the estimation is considered as reliable and robust.

Figure C.2 shows normalized measurements in per unit according to the primary generator, slack bus generator values include apparent power, 100 MVA and base voltage, 69 KV for IEEE 14-bus system corresponding to 11th time instant for sudden load change condition with anomaly data identification at 11th time instant including reactive power injection on bus 2, real power injection on bus 14, real power flow from bus 6 to 13, real power flow from bus 5 to 6, and reactive power flow from bus 4 to 7.These anomalies are eliminated using backward elimination which is selected one anomaly measurement every time and eliminated, then re-estimate the system at the same time instant in order to improve the quality of the estimation.

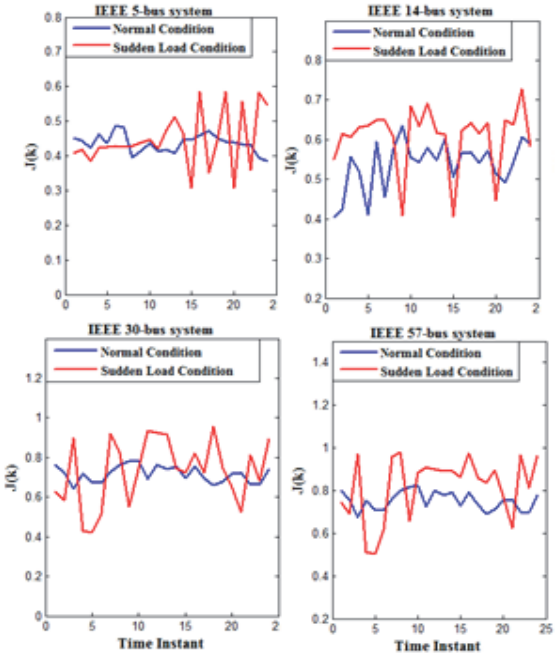

Figure C.1: Performance Index J(k)in the presence of normal and sudden load condition

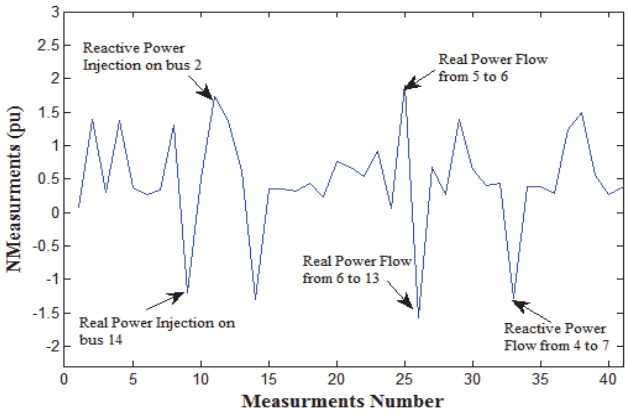

Figure C.2: Normalized Measurements vector for sudden load change condition during 11th time instant

C.6 Conclusion

In this chapter, a meta-heuristic procedure named a stochastic fractal search technique is proposed and applied to determine tracking state estimation (TSE). The SFS estimator has been implemented for following the time fluctuating static condition of the power system. The proposed SFS method has been checked on four IEEE test cases. Sudden load changes conditions has been used to assess the performance of the proposed SFS estimator. The proposed SFS execution is done to be compared with different estimators including Particle Swarm Optimization (PSO) and Genetic Algorithm (GA) in a few factual parameters. The numerical results show that the proposed SFS system has given us a precise state estimator differentiated to other proposed estimators. On the other hand, this procedure experiences the long computational time in large system. This strategy can be more successful in multi-systems by decoupling the large system into multi-small scale systems. The Matlab code used for the estimation is described in [82].

Appendix D

Power System Static State Estimation using Modified Stochastic Fractal Search Technique

D.1 Introduction

Reliable operation of today's large electric power systems heavily depend upon the state estimation techniques for different applications. These estimators give important measurement used for different application functions of control center. Therefore, a computationally efficient and robust state estimator is an important tool to improve the performance, security and safety of the power delivery system. State estimation (SE) is becoming one of the key functions in the distribution control centers in the deregulated and competitive environment [1]. The innovative idea of SE was first proposed by Schweppe in ref. [152]. He presented three sequential papers on SSE. In the first paper, he introduced mathematical modeling and general techniques. In the second paper, he tested several approximate mathematical models and it is resulted simplification in estimation in case of detection and identification. In the third paper, he discussed several implementation issues associated in with dimensionality, and the time changing nature of real power systems. Linear programming (LP) is presented in ref. [7] to minimize the weighted sum of the absolute values of the measurement residuals. The sequence of linear programing strategy first recognizes the bad measurements using the measurement residuals of those data rejected by the LP estimator. Another computational strategy is proposed in ref. [8] for solving equality constraint in PSSE without the Lagrange multipliers. It is an imperative to process a complete fractional factorization (L, U) of grid which improve numerical modeling. A new technique in ref. [10] investigated the problem of bad data identification under line current magnitude measurements. Two common SE methods include least squares and least absolute value estimation are studied. The least squares estimation is used as a post estimation problem and the least absolute value estimation is used as

an outlier rejection problem. Interior point algorithm is presented in ref. [11] to solve the power system weighted non-linear state estimation problem. The effectiveness of the proposed algorithm has been verified by the extensive simulating computations. Mosbah and Elhawary were first introduced SFS technique to PSSE in ref. [153], [154]. The authors implemented the SFS technique under a few scenarios incorporating normal operation where load fluctuates linearly, bad data , and sudden load changing conditions presented at various time samples during the study period. The technique is tested on different IEEE bus systems and the results are contrasted with other meta-heuristic techniques.

This chapter presents modified stochastic fractal search technique to solve static SE problem. The proposed M-SFS technique is utilized to improve the accuracy and computational time of the original SFS algorithm. Two modifications are considered in M-SFS. This includes various benchmark functions replaced the logarithmic function in the diffusing process and a few chaotic maps replaced the uniform distribution in diffusing and updating processes. The M-SFS is tested on IEEE 30, 57, and 118-bus systems. The results illustrated the superiority of M-SFS technique compared to its original SFS.

D.2 Stochastic Fractal Search Technique

SFS grouped into two main processes. The diffusing process, the exploitation property is satisfied by diffusing every particle around its present position. Therefore, the chance of preventing local minima and finding global minima will be increased. Updating process, The position of each point will be updated based on the other points in the group. Fitness function is used to rank all points in the group. The particles with high probability are only considered and others are discarded. This is called static diffusing process which is used in SFS algorithm. Gaussian walk is only considered in proposed SFS technique to produce new particles due to its ability in finding the global minima. DLA growth process of SFS employed Gaussian distribution as the only random walk [154].

Algorithm 2 Stochastic Fractal Search

Initialize the SFS parameters
Set the lower and the upper constraints
while $g \leq max.Iteration$ **do**
  for each point i in the system **do**
    Call diffusing process
    for $j = 1 : q(max.No.of diffusion)$ **do**
      If first Gaussian walk is applied **do**
      $GW_1 = Gaussian(\mu_{BP}, \sigma) + (\varepsilon \times BP - \varepsilon^{'} \times P_i)$
      Else second Gaussian walk is applied **do**
      $GW_2 = Gaussian(\mu_P, \sigma)$
    end for
  end for
  end If
  Call updating process
  for each point i in the system **do**
    for each component$j$ in P_i **do**
      If $rand[0, 1] \leq P_{ai}$ **do**
      Update the component **do**
      $P_i^{'}(j) = P_r(j) - \varepsilon \times (P_t(j) - P_i(j))$
      Else do nothing
      Rank all points obtained by the update process
    end for
  end for
  end If
end while

D.3 Modified Stochastic Fractal Search Technique

This section shows the presented modified SFS technique and its execution evaluations utilizing different standard benchmark functions. The logarithmic function in the diffusing process is replaced by well-known benchmarks functions and the uniform distribution parameter in both diffusing and updating process is replaced with a few chaotic maps.The execution ranks and significant contrasts among the modified and the original technique are assessed.

D.3.1 Modified α Parameter

The logarithmic function in standard deviation of the diffusing process will be replaced by parameter α. The α parameter represents the benchmark functions which are shown in table D.1. This will limit the number of Gaussian jumps as the size of iteration increments. This will lead to efficient exploration of search space.

$$GW_1 = Gaussian(\mu_{BP}, \sigma) + (\varepsilon \times BP - \varepsilon^{'} \times P_i) \tag{D.1}$$

$$GW_2 = Gaussian(\mu_P, \sigma) \tag{D.2}$$

$$\sigma = |\frac{log(g)}{g} \times (P_i - BP)| \tag{D.3}$$

$$\sigma = |\alpha \times (P_i - BP)| \tag{D.4}$$

Note that equations (D.1-D.3) represent the Gaussian walk of the diffusing process in SFS technique. However, equation (D.4) replaced the logarithmic function by α parameter. The α parameter denotes the benchmark functions show in table D.1.

D.3.2 Implementation of Chaotic Maps

The uniform distribution in both diffusing and updating processes will be replaced by μ parameter. The μ parameter represents the chaotic maps which are shown in table D.2. The randomization of every particle utilizing uniform distribution more often is the fundamental execution of the metaheuristic techniques. This choice carries

Table D.1: The Presented Replacement Functions For α

Algorithms	Functions	Names
SFS	$log(g)/g$	Logarithmic
Modified SFS	$(1/1 + exp(-g)) * g$	Sigmoid
	$\mid acos(g)/g) \mid$	Inverse cosine
	$exp(-g)/g$	Exponential
	$cosh(g)/g$	Hyperbolic cosine
	$csch(g)/g$	Hyperbolic cosecant

drawbacks when dealing with complex nonlinear and multimodal issues. Therefore, chaotic maps will improve accuracy and convergence speed.

$$GW_1 = Gaussian(\mu_{BP}, \sigma) + (\varepsilon \times BP - \mu \times P_i) \tag{D.5}$$

$$P_i^{'}(j) = P_r(j) - \mu \times (P_t(j) - P_i(j)) \tag{D.6}$$

Table D.2: The Mathematical Description Of Chaotic Maps μ

Map Name	Equation
Chebyshev map	$x_{i+1} = cos(i\cos^{-1}(x_i))$
Circle map	$x_{i+1} = mod(x_i + b - (\frac{a}{2\pi})\sin(2\pi x_i), 1)$
Gauss/mouse map	$x_{i+1} = \begin{cases} 1, for x_i = 0 \\ mod(\frac{1}{x_i}, 1) \end{cases}$
Iterative map	$x_{i+1} = sin(\frac{a\pi}{x_i})$
Logistic map	$x_{i+1} = ax_i(1 - x_i)$
Piecewise map	$x_{i+1} = \begin{cases} \frac{x_i}{p} for 0 \le x_i < p, \\ \frac{x_i - p}{0.5 - p} for p \le x_i < 0.5 \\ \frac{1 - p - x_i}{0.5 - p} for 0.5 \le x_i < 1 - p \\ \frac{1 - x_i}{p} for 1 - p \le x_i < 1 \end{cases}$
Sine map	$x_{i+1} = sin(\pi x_i)$
Sinusoidal map	$x_{i+1} = 2.3x_i^2 sin(\pi x_i)$
Tent map	$x_{i+1} = (\frac{10}{3})(1 - x_i)$

The above equations (D.5-D.6) represent the diffusing and updating processes

respectively in the SFS technique. The particles will be forced to move toward the optimal solution by chaotic maps. Chaotic maps were first proposed by both Saremi and Mitic in [9], [10]. Every chaotic map has an initial point=0.7 and normalized to the range of [0,1]. Here ε and $\varepsilon^{'}$ are uniform distributed to [0, 1], BP and P_i denote as the best position point and i^{th} point respectively. μ_{BP} denotes $|BP|$, μ_P is equal to P_i and σ is the standard deviation.

D.4 Simulation Study

The presented M-SFS technique is implemented in Matlab and illustrated on IEEE 30-bus, 57-bus and IEEE 118-bus systems [113]. The weighted Least squares is defined to evaluate the unknown states by minimizing the following objective,

$$J(x) = \sum_{i=1}^{m} w_i (z - h(x))^2 \tag{D.7}$$

Mean absolute percentage error has been utilized to assess the execution of the proposed M-SFS technique and its comparison is carried out with its original SFS.

$$Mape(x) = \frac{1}{n_b - 1} \sum_{i=2}^{n_b} \mid \frac{x^{est} - x^t}{x^t} \mid \tag{D.8}$$

Where x^{est} and x^t denote the estimated (M-SFS) and true states (load flow) respectively.

D.5 Results and discussions

The presented M-SFS technique is demonstrated on three IEEE bus systems. The estimation performance of the proposed technique is evaluated using real time data from the American Electric Power Systems in the Midwestern of US.

Table D.3 investigates the effect of various benchmark functions represented by α parameter. The α parameter is replaced the logarithmic function of the original SFS. Moreover, the weighted sum of square error is used to evaluate the accuracy of all benchmark functions under various IEEE bus systems. As can be seen that Modified SFS technique with inverse cosine and exponential functions outperformed the original SFS algorithm in most of the benchmark functions. However, inverse cosine function has the least sum of square error. Table D.4 shows the implementation of chaotic maps

Table D.3: Results of modified SFS by α parameter

Algorithms	Functions	IEEE bus systems		
		30-bus	57-bus	118-bus
SFS	Logarithmic	2.6021	29.8204	50.3070
M-SFS	Sigmoid	2.6985	38.6251	54.9006
	Inverse cosine	1.3922	18.2271	40.7387
	Exponential	2.4434	29.0701	53.1143
	Hyperbolic cosine	7.821962	63.4263	212.7553
	Hyperbolic cosecant	3.7998	26.6832	60.6161

represented by μ parameter. The μ parameter is replaced the uniform distribution of the original SFS by various chaotic maps. The replacement occurred in both diffusing and updating process. In addition, the accuracy of all chaotic maps is assessed by the weighted sum of square error under three different IEEE bus systems. Iterative map improved the accuracy of SFS algorithm far better than other chaotic maps. It concluded that chaotic fractal search algorithm (CFS) has demonstrated significant enhancement due to its capability in finding the best optimal solution compared to its original SFS algorithm.

Table D.4: Results of modified SFS by Chaotic maps

Map Name	IEEE 30-bus	IEEE 57-bus	IEEE 118-bus
Chebyshev map	3.3242	34.1794	35.1063
Circle map	3.9886	29.1277	39.3523
Gauss/mouse map	4.4310	25.8382	55.8319
Iterative map	2.4479	21.7495	34.1999
Logistic map	5.7923	35.4989	55.5910
Piecewise map	5.0481	36.3972	41.9784
Sine map	4.5570	32.0554	39.7002
Sinusoidal map	3.7310	25.1363	49.9311
Tent map	3.9293	38.6699	65.6037

Table D.5: MAPE of Voltage angle for modified SFS

Algorithms	IEEE 30-bus	IEEE 57-bus	IEEE 118-bus
SFS	4.8073	9.9565	16.0715
Modified SFS	1.7376	5.9327	8.8347

The MAPE of estimated states for various IEEE bus systems is tabulated in tables (D.5-D.6). Both modified SFS and its original SFS algorithm are compared in these

Table D.6: MAPE of Voltage magnitude for modified SFS

Algorithms	IEEE 30-bus	IEEE 57-bus	IEEE 118-bus
SFS	3.6929	6.6718	9.8929
Modified SFS	1.4152	4.1859	6.0970

tables. As can be seen that modified SFS has the least MAPE in all three IEEE test systems. For instance, in IEEE 30-bus system, the MAPE for voltage magnitude is 3.6929% and 4.8073%for voltage angles in SFS technique. This compared to 1.4152% for voltage magnitude and 1.7376% for voltage angles in modified SFS. Moreover, in IEEE 118-bus system, the MAPE for voltage magnitude is 9.8929% and 16.0715% for voltage angles in SFS technique. This compared to 6.0970% for voltage magnitude and 8.8347% for voltage angles in modified SFS. It concluded that modified SFS has increased the accuracy significantly compared to its original SFS algorithm.

Table D.7: Computational time for modified SFS

Algorithms	IEEE 30-bus	IEEE 57-bus	IEEE 118-bus
SFS	16.227668	28.484373	72.541932
Modified SFS	7.401716	12.521130	38.775233

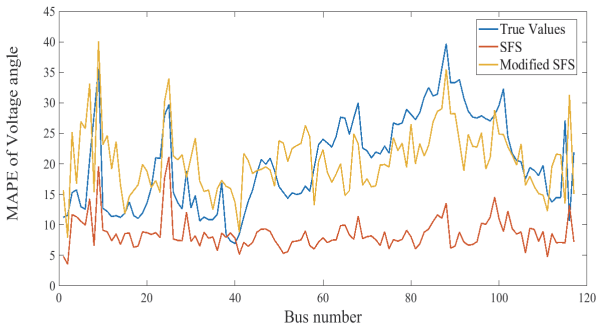

Figure D.1: Mean absolute Percentage error of Voltage angle

Table D.7 shows the computational time for three different IEEE bus systems. Both modified SFS and its original SFS algorithm are compared in this table. As can be observed that modified SFS has the least CPU time in all three different IEEE test systems. For example, in IEEE 30-bus system, the CPU time is 16.227668 seconds for SFS algorithm. This can be compared to 7.401716 seconds for modified SFS technique. In addition, in IEEE 118-bus system, the CPU time is 72.541932 seconds

for SFS algorithm. This can be compared to 38.775233 seconds for modified SFS technique. It affirmed that modified SFS has the least computational time compared to its original SFS algorithm.

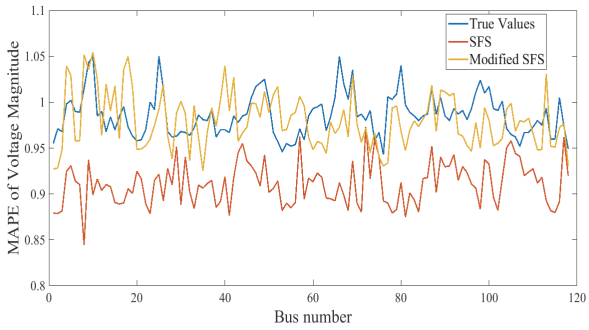

Figure D.2: Mean absolute Percentage error of Voltage Magnitude

Figures (D.1-D.2) present the state estimation behavior of system states for IEEE 118-bus system. Both modified SFS technique and its original SFS algorithm are compared with true states (load flow). It can be observed from these figures that modified SFS has a better performance than its original SFS. It concluded that modified SFS technique is more robust.

D.6 Conclusion

The chapter introduced modified stochastic fractal search technique to PSSE. Two important modifications are studied in this paper. Various benchmarks functions replaced logarithmic function in the diffusing process. Moreover, a few chaotic maps replaced the uniform distribution parameter in both diffusing and updating processes. These modifications will have a crucial effect on the algorithm execution. Therefore, the accuracy with the least computational time will be fulfilled. Several IEEE bus systems are used to validate the experiment. The obtained results demonstrated a significant improvement in both accuracy and the computational time in comparison to its original algorithm.

Appendix E

Evaluating the Impact of Phasor Measurement Units on the Accuracy of State Estimation

E.1 Introduction

PMUs can monitor, protect, and control real-time wide area of modern power systems [155]. State estimation is a function which is used to process a set of measurements to obtain the best estimated states. Weighted least square method is conventionally used to solve the SE problem with SCADA measurements include voltage magnitudes, real and reactive power injections and flows [27]. The advent of phasor measurement units (PMUs) eased the problem of measuring phase angles of the bus voltages in real time by synchronizing the voltage and current waveforms from far off areas with respect to a Global Positioning System (GPS) [156]. When a PMU installed at a bus, it can measure the voltage phasor of the installed bus as well as the current phasors in the lines associated with that bus. PMUs can help in enhancing the operation and real-time control of vast transmission systems, however, PMUs cannot make state estimation obsolete [136]. Linear estimator can be expressed as a complete observability of the system when only PMU measurements are used to estimate the system states. However, this can economically be impractical for a vast system because it requires a large number of PMUs. The present practice is to introduce PMUs in an incremental manner in conjunction with traditional measurements [157]. In ref. [158], state estimation problem is defined without utilizing any actual or virtual reference bus. The detection and identification of incorrect PMUs is the main benefit of using this formulation. In ref. [159], the state estimation problem is defined such that the current SCADA-based state estimator can be attached effectively to represent the PMU measurements. In ref. [160], the authors analyzed the impact of measurement synchronicity includes SCADA and additional PMUs measurements on the accuracy of the state estimation.A hybrid state

estimator combined SCADA and PMUs measurements in rectangular coordinates is proposed in [161]. In ref. [162], the authors presented a modified formulation utilizing rectangular coordinates. This formulation solves numerical difficulties of using phasor voltage and current measurements. Detection and identification of errors in these phasor measurements are included. In ref. [163], the authors presented a phasor assisted hybrid for non-linear state estimator. This approach combined SCADA and PMUs measurements to improve the precision of the traditional weighted least square state estimator.

This chapter presents hybrid multilayer perceptron neural network with stochastic fractal search to solve the hybrid state estimator problem. The performance of the hybrid technique is assessed based on mean absolute percentage error and redundancy indices. The experiment is tested on three standard IEEE bus systems including 14, 30, and 57 bus systems. Several cases are tested to demonstrate the effect of PMUs on the accuracy of state estimation. The results of the hybrid technique are contrasted with stochastic fractal search and multilayer perceptron neural network individually.

E.2 Hybrid state estimator formulation

The main idea of state estimation is to estimate the system states (voltage phasors) at all buses. These estimated states are computed based on available set of measurements. Eq.(E.1) provides a general mathematical formula of state estimation.

$$z = h(x) + e \tag{E.1}$$

where z denotes as a measurement vector of $R^{m\times 1}$ computed from both conventional and PMU measurements, x is $R^{n\times 1}$ estimated states ($n < m$), $h(x)$ denotes as a non-linear function of $R^{m\times 1}$ vector relating measurements to states, n is size of estimated states($n = 2 \times N$), m is size of measurements, N is number of buses, and e denotes as a measurement noise error vector $R^{m\times 1}$ which has a Gaussian distribution with zero

mean and standard deviation.

$$z = \begin{bmatrix} z_{SCADA} \\ \\ z_{PMUs} \end{bmatrix} \tag{E.2}$$

$$z_{SCADA} = \begin{bmatrix} P_i \\ Q_i \\ P_{ij} \\ Q_{ij} \end{bmatrix}, z_{PMUs} = \begin{bmatrix} v_{real} \\ v_{imag} \\ I_{real} \\ I_{imag} \end{bmatrix}$$

The above matrices describe the ways to combine the conventional measurements by SCADA along with the voltage and the current measurements by PMUs. Here, P_i and Q_i denote the real and reactive power injections respectively, P_{ij} and Q_{ij} represent the real and reactive power flows respectively,V_{real} and V_{imag} stand for real and imaginary part of bus voltage, I_{real} and I_{imag} represent the real and imaginary of line currents recorded by PMUs.

$$h(x) = \begin{bmatrix} h(x)_{Conv} \\ \\ h(x)_{PMUs} \end{bmatrix} \tag{E.4}$$

$$h(x)_{Conv} = \begin{bmatrix} V_i \sum_{i=1}^{m} V_j |Y_{ij}| \cos(\delta_i - \delta_j - \theta_{ij}) \\ \\ -V_i \sum_{i=1}^{m} V_j |Y_{ij}| \sin(\delta_i - \delta_j - \theta_{ij}) \\ \\ |V_i^2| G_{ii} + \sum_{i=1}^{m} |V_i||V_j|Y_{ij}| \\ \cos(\delta_i - \delta_j + \theta_{ij}) \\ \\ -|V_i^2| B_{ii} - \sum_{i=1}^{m} |V_i||V_j|Y_{ij}| \\ \sin(\delta_i - \delta_j + \theta_{ij}) \\ \\ \end{bmatrix} \tag{E.5}$$

$$h(x)_{pmus} = \begin{bmatrix} V_{real} \\ \\ V_{imag} \\ \\ V_i[(G_{ij} + G_{si})\cos\delta_i - (B_{ij} + B_{si}) \\ \sin\delta_i] - V_j[G_{ij}\cos\delta_j - B_{ij}\sin\delta_i] \\ \\ V_i[(B_{ij} + B_{si})\cos\delta_i + (G_{ij} + G_{si}) \\ \sin\delta_i] - V_j[B_{ij}\cos\delta_j + G_{ij}\sin\delta_i] \\ \\ \end{bmatrix} \tag{E.6}$$

Where $h(x)$ represents the nonlinear measurement function relates to both SCADA and PMUs measurements, both V_i and V_j are voltage magnitudes on bus i and j respectively, G_{ij} denotes as the series conductance of branch i-j, B_{ij} denotes as the series susceptance of brunch i-j, G_{si} denotes as the shunt conductance between bus i and ground, B_{si} denotes as the shunt susceptance between bus i and ground.

The mathematical formula of hybrid state estimator problem is constructed by the typical pi-model of a branch connecting the PMU bus with regular bus, This mathematical formula is described in details in ref. [157].
Generally, the objective function below is minimized using WLS method to compute the estimated states as given in Eq. (E.7) ,

$$J(x) = [z - h(x)]^T R^{-1}[z - h(x)] \tag{E.7}$$

where,R represents the measurement covariance matrix of $R^{m\times m}$.

E.3 Hybrid techniques

The hybrid MLP-SFS is classified into two phases. The first phase, multilayer perceptron NN is used to determine the initial system states. The dataset for all IEEE bus systems are randomly divided into three subsets, i.e. 70% for training and 15%

for both validation and testing. Input layer comprised of one node representing as the available measurements. An output layer with one node served as the initial system states. The desired output (target) is set as true system states (load flow). Levenberg-Marquardt back propagation (LMBP) is used as a training algorithm. Hidden neurons were equipped with non-linear transfer function, tan-sigmoid, while output layer has (linear) purlin-neurons. This system can determine the initial estimated states. Finally, the desired model is selected based on minimum mean square error (MSE). The second phase, stochastic fractal search is used to determine the final estimated states. The output of MLP NN is used an input into SFS. SFS begins with diffusing process, the exploitation property is fulfilled by the diffusing of each particle around its present position. Gaussian random walk is used to generate new particles from the diffusing process. The optimal point among all points is computed by the fitness function. SFS ends with updating process, the fitness function is used to rank all the points. A probability value is applied to every point i in the group. The position of each point is changed based on the position of other points in the group. Finally, algorithm stops by a termination criterion and determine the best fitness value.

E.4 Simulation Study

The proposed technique is implemented in Matlab and demonstrated with IEEE 14-bus, IEEE 30-bus, and 57-bus systems [113]. The available measurements incorporate both conventional and phasor measurements. Every PMU will measure the voltage phasor of the bus where it is installed and the current phasors in all lines connected to that bus. Bus 1 is thought to be the slack bus for conventional measurements. In this experiment, we set an error of 2% for the traditional measurements and 0.5% for the PMU measurements. Reference [164] is used to obtain the optimal locations of PMUs. The weighted Least Absolute Value is defined to assess the unknown system states by minimizing the following objective,

$$J(x) = \sum_{i=1}^{m} w_i |z - h(x)| \tag{E.8}$$

Mean absolute percentage error has been used to evaluate the performance of the proposed technique and its comparison is carried out with multilayer perceptron NN

Algorithm 3 Backpropagation NN

Initialize the connection weights randomly
Set error, threshold error, and max epoch
while $epoch \leq max.epoch$ **do**
    For each input node i in the input layer **do**
    $a_i \leftarrow x_i$
    end for
    For each node j in each hidden layer **do**
    $z \leftarrow \sum w_{ij} a_i$
    $a_j \leftarrow f(z)$
    end for
    For each node k in the output layer **do**
    $y \leftarrow \sum w_{jk} a_j$
    $a_k \leftarrow f(y)$
    end for
end while
Calculate error
Increment epoch

(MLP) and Stochastic fractal search individually.

$$Mape(x) = \frac{1}{n_b - 1} \sum_{i=2}^{n_b} | \frac{x^{est} - x^t}{x^t} | \tag{E.9}$$

where x^{est} denotes the estimated states using the proposed technique and x^t denotes the true system states (load flow).

E.5 Results and discussions

The presented hybrid technique is illustrated on three IEEE test systems. The estimation execution of the presented technique is assessed using real time data from the American Electric Power Systems in the Midwestern of US. Seven cases are examined to demonstrate the effectiveness of PMUs on the accuracy and computational time in electric grid.

The best control parameters are tabulated in table E.1. The population size varies from 50 to 90, the size of neurons varies from 20 to 35, and [-1,1] is selected for the lower and the upper bound for all IEEE bus systems. The hyperbolic tangent transfer function is selected as the hidden layer for IEEE 14 and 30-bus system. However, the Log-sigmoid transfer function is chosen as the hidden layer for IEEE 57-bus system. Moreover, linear activation function is set as an output layer for all IEEE bus systems. These control parameters have provided the optimal condition.

Table E.1: Control Parameters used in hybrid MLP-SFS

Control parameters	14-bus	30-bus	57-bus
Side Walk	0	1	1
Population Size	50	70	90
No. of Neurons	20	25	35
Max. Generation	150	200	350
Input Transfer Funct.	TangSigmoid	TangSigmoid	Logsigmoid
Output Activation Funct.	Linear	Linear	Linear
boundaries	[1, -1]	[1, -1]	[1, -1]
Diffusion No. (MDN)	1	10	15

Tables (E.2-E.4) investigate the impact of phasor measurement units (PMUs) on the accuracy of state estimation. The tables illustrate Six different cases by gradually

Table E.2: Six Different Cases by adding Phasor Measurement units (PMUs) for IEEE 14-bus system

Case	Measurement Configuration	No. of Meas. (m)	No. of States (n)	Redundancy(r=m/n)	MAPE(V)	MAPE(δ)
1	Only traditional Measurement	43	27	1.59	4.2190	8.4960
2	Conventional Meas. with PMUs of (10% of bus number)	45	27	1.66	3.8160	5.4091
3	Conventional Meas. with PMUs of (20% of bus number)	48	27	1.77	3.1871	4.6678
4	Conventional Meas. with PMUs of (30% of bus number)	50	27	1.85	2.8286	3.8101
5	Conventional Meas. with PMUs of (40% of bus number)	52	27	1.93	1.6892	3.5483
6	Conventional Meas. with PMUs of (50% of bus number)	53	27	1.96	1.2907	2.8595

Table E.3: Six Different Cases by adding Phasor Measurement units (PMUs) for IEEE 30-bus system

Case	Measurement Configuration	No. of Meas. (m)	No. of States (n)	Redundancy(r=m/n)	MAPE(V)	MAPE(δ)
1	Only traditional Measurement	93	59	1.58	5.0046	13.3776
2	Conventional Meas. with PMUs of (10% of bus number)	96	59	1.63	4.4130	8.9042
3	Conventional Meas. with PMUs of (20% of bus number)	98	59	1.66	3.2594	8.6339
4	Conventional Meas. with PMUs of (30% of bus number)	100	59	1.69	3.0168	7.3343
5	Conventional Meas. with PMUs of (40% of bus number)	105	59	1.78	2.8380	6.6932
6	Conventional Meas. with PMUs of (50% of bus number)	116	59	1.97	2.7494	5.2005

Table E.4: Six Different Cases by adding Phasor Measurement units (PMUs) for IEEE 57-bus system

Case	Measurement Configuration	No. of Meas. (m)	No. of States (n)	Redundancy(r=m/n)	MAPE(V)	MAPE(δ)
1	Only traditional Measurement	150	113	1.33	12.7819	21.4200
2	Conventional Meas. with PMUs of (10% of bus number)	189	113	1.67	8.6840	18.9688
3	Conventional Meas. with PMUs of (20% of bus number)	202	113	1.79	9.1460	18.5572
4	Conventional Meas. with PMUs of (30% of bus number)	208	113	1.84	8.1125	16.3300
5	Conventional Meas. with PMUs of (40% of bus number)	212	113	1.88	6.7805	17.6341
6	Conventional Meas. with PMUs of (50% of bus number)	220	113	1.95	6.7641	15.4042

incorporating PMUs measurements with conventional measurements for three different IEEE bus systems. From the tables can be seen that the increment in PMUs number leads to increase the redundancy and subsequently decrease the mean absolute percentage error of estimated states. In IEEE 14-bus system, in case of no PMUs, the mean absolute percentage error is 1.59% and 4.2190% for voltage magnitudes and angles respectively. However, when PMUs increases to 10% of bus number, the MAPE decreases to 3.8160% for voltage magnitudes and 5.4091% for voltage angles. In addition, the bus number of PMUs increases to 50%. This leads to increase the redundancy to 1.96 and decrease the MAPE to 1.2907% for voltage magnitudes and 2.8595% for voltage angles. In IEEE 30-bus system, from no PMUs to 50% PMUs, the MAPE is reduced from 5.0046% to 2.7494% for voltage magnitudes and from 13.3776% to 5.2005% for voltage angles. Also, the redundancy increases from 1.58 to 1.97. In IEEE 57-bus system, from no PMUs to 50% PMUs, the MAPE is reduced from 12.7819% to 6.7641% for voltage magnitudes and from 21.4200% to 15.4042% for voltage angles. Also, the redundancy increases from 1.33 to 1.95. It concluded that phasor measurement units (PMUs) can increase the observability of the power system. This results in improving the accuracy of estimated states.

Tables (E.5-E.6) show the mean absolute percentage error of estimated states for various IEEE bus systems. Linear state estimator is considered by the only PMUs. This case can examine the effect of the only PMUs on the accuracy in the electric grid. As can be observed that only PMUs case has the least MAPE compared to previous tables. For instance, with 50%PMUs in IEEE 57-bus system, the MAPE for voltage magnitude is 6.7641% and 15.4042% for voltage angles. This compared to 5.5562% for voltage magnitude and 13.6281% for voltage angles in only PMUs case. Moreover, the proposed hybrid technique has the least error compared to other techniques.

Table E.5: MAPE of Voltage magnitude for only PMUs

IEEE system	14-bus system	30-bus system	57-bus system
MLPNN-SFS	0.6511	2.0788	5.5562
SFS	2.2521	5.1193	9.0031
MLP	3.4223	5.9270	9.0203

Table E.6: MAPE of Voltage angle for only PMUs

IEEE system	14-bus system	30-bus system	57-bus system
MLPNN-SFS	1.8894	4.7660	13.6281
SFS	6.6470	8.2391	19.0417
MLP	8.3476	11.1304	21.7104

E.6 Conclusion

Hybrid MLP-SFS was applied to enhance the accuracy of state estimation. The measurement redundancy and accuracy were improved by gradually increasing the size of PMUs. Only PMU measured data is used to investigate the linear formulation of the state estimation. Seven cases were used to validate this experiment. Moreover, three different IEEE bus systems were implemented to illustrate the results of the proposed technique. The proposed technique was contrasted with other techniques include MLP and SFS individually. The presented technique provided the least MAPE.

Appendix F

A Distributed Multiarea State Estimation

F.1 Introduction

Robust SE is essential for wide-area monitoring, protection, and control (WAMPAC). Due to the large-scale power systems, the measurements and the system states turn out to be very large. This leads to transmit extremely large amounts of data from the measurement units to the centralized SE. Therefore, the required cost for the communication link expanded. Decentralized SE is an alternative technique to provide a reliable estimate with significantly reduced computational requirements for a large scale power system. This can be done by partitioning the large power system into multi-regions and running local state estimators in each area to compute the global state estimate [76]. A two-level technique is proposed in ref. [165]. A traditional state estimation is performed in parallel for all areas in the lower level. The coordination of these local estimations is acknowledged in the upper level. Dantzig-Wolfe Decomposition Principle is presented in ref. [166] to solve the linear programing SE problem in large scale electric system. The approach reduces the LP problems into smaller dimension of a coordinated sequence of LP. In ref. [167], a parallel algorithm is presented based on border measurements and overlapping subsystems. The Auxiliary Problem Principle is applied to improve a distributed state estimator. Multiarea decentralized state estimation approach is presented in [168]. The approach allows the boundary buses to exchange its information among area operators without considering the central coordinator level. A distributed state estimator is proposed in ref. [169] to solve multi area SE problem. The entire system is partitioned into a particular number of non-overlapping subsystems. Each subsystem independently perform their own state estimator based on its local measurements. Boundary measurements and states received by central coordinator to estimate the system-wide solution. Both Observability and bad data analysis are also included in this approach. A fully distributed state estimation technique is presented in ref. [170] for

wide-area monitoring in power systems. In this algorithm, the sub-areas exchange the information iteratively. In ref. [171], conventional algorithms connected with a coupling constraints optimization technique is proposed for parallel and distributed state estimation. where overlapping subareas are considered with a common zero injection boundary bus. A diakoptic-based distributed SE algorithm is presented in ref. [172] for large-scale power systems. the proposed algorithm is partitioned the large-scale power system into a particular number of subareas by removing tie line measurements. Each subarea performs their own state estimation based on its local computational resources. The tie line measurements sent to a central computer for completing the state estimation process by the intermediate subarea SE solutions.
This chapter presents hybrid stochastic fractal search with simulated annealing algorithm to solve multiarea SE problem. The proposed hybrid technique is decomposed the IEEE 118-bus system into four non-overlapping subareas. Each subarea performs its own SE based on their local measurements. The tie line measurements sent to a coordination center to compute the system-wide solution. Stochastic fractal search and simulated annealing techniques are used to perform the local and the coordinator SE respectively. Three different measurement configurations are included in this chapter. The results of the hybrid technique are contrasted with integrated SE. The obtained results demonstrated a significant reduction in the computational time.

F.2 Multiarea State Estimation

F.2.1 Network Partitioning

Large scale power system is partitioned into a particular number of non-overlapping subareas. These subareas are connected by tie-lines. Each local area performs their own state estimation based on the available local measurements in each area. The tie line available measurements (boundary measurements) transmit to a coordination center to compute the system-wide solution. The buses of each subarea are classified into three categories,

- Internal bus, A bus of a subarea i that is not associated to any external boundary bus.
- Boundary bus, A bus of a subarea i that is associated to at least one external

boundary bus.

- External bus, A bus that belongs to another subarea and associated to an internal boundary bus of the subarea i.

F.2.2 Local SE Formulation

The weighted least-squares (WLS) method is commonly used to solve the SE problem for each subarea. Each subarea in local SE problem is formulated as follows,

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & J_i(x) = r_i^T R_i^{-1} r_i \\ \text{subject to} \quad & z_i = h_i(x_i) + r_i \end{aligned} \tag{F.1}$$

$$\mathrm{x}_i = [x_i^{int} \quad x_i^b]^T$$

where

z_i measurement vector of $R^{m\times 1}$ in area i. This includes all injection and flow measurements;

r_i measurement residual vector $R^{m\times 1}$;

R_i measurement noise error variance matrix for area i;

$h_i(x_i)$ nonlinear function of $R^{m\times 1}$ vector relating measurements to states for area i;

x_i system states of $R^{n\times 1}$ for area i. This includes all internal and boundary buses;

It is supposed that enough redundancy should be available in each subarea to identify and eliminate bad data for all internal subarea measurements, so in this case, state estimate for each subarea are assumed to be unbiased.

F.2.3 System-Wide SE Formulation

The SE solutions from all subareas sent to the central coordinator to process the wide-area solution. The unbiased estimate of the overall system state is reached

by including raw measurements from subarea boundary buses. The SE problem is formulated as the following optimization problem,

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & J_c(x) = r_c^T R_c^{-1} r_c \\ \text{subject to} \quad & z_c = h_i(x_c) + r_c \end{aligned} \tag{F.2}$$

$$\mathrm{z}_c = [z_{bm}^T \quad x_b^{int^T} \quad x_b^{ext^T}]$$

where

z_c all available measurement data vector of $R^{m\times 1}$.

r_c measurement residual vector $R^{m\times 1}$ for z_c;

$h_c(x_c)$ nonlinear function of $R^{m\times 1}$ vector relating all available measurements to states;

x_c system states of coordination $R^{n_c\times 1}$;

The central coordinator receives the boundary measurements and its estimated states with covariance matrix. Moreover, two pseudo measurements may associate with state of a boundary bus. One provided by local SE in each subarea and another by the neighbors SE. Topology information around boundary nodes should be given to process the boundary injections [173].

F.3 Proposed and implemented Simulation System

F.3.1 Stochastic Fractal Search (SFS)

SFS is classified into diffusing and updating processes, each particle is diffused around its present position to fulfill the propriety of search space. New particles are produced by Gaussian random walk from the diffusing process. The optimal point among all points is computed by the fitness function. The fitness function is utilized to compute the optimal point among various points in the search space. In updating process, all the points are ranked by the fitness function. Each point in the group has a probability value. Each point in the group is re-positioned based on the position of other points. Finally, algorithm terminates based on the best fitness value [154].

F.3.2 Simulated Annealing (SA)

The SA technique begins by initializing a random solution with a large temperature. The initial temperature should be high enough to make transition probabilities. A few choices have a significant impact on the performance of the SA algorithm include probability transition function, the neighbor selection method, the state space, and the annealing schedule. As the temperature approaches to a cooling schedule, the likelihood in solution quality decays exponentially toward zero as indicated by the Boltzman distribution. Different transitions are made by the solution state variable in each temperature. Consequently, the last solution is close optimal when the temperature reaches zero [174].

Algorithm 4 Simulated Annealing

1: Initialize the SA parameters
2: $T \leftarrow T_{max}$
3: $x_{old} \leftarrow x_0$
4: **while** $T > T_{min}$ **do**
5: $\quad x_{new} \leftarrow x_{old} + \Delta(x)$
6: $\quad \Delta E \leftarrow E(x_{new}) - E(x_{old})$
7: $\quad$ **If** $\Delta E < 0$
8: $\quad x_{best} \leftarrow x_{new}$
9: $\quad$ **Else**
10: $\quad prob = exp(-\frac{\Delta E(x)}{T})$
11: $\quad (x_{new}, prob) \leftarrow x_{old}$
12: $\quad$ **end** If
13: $\quad T_{new} \leftarrow \alpha T_{old}$
14: **end while**
15: **return**

F.3.3 Implemented Simulation Systems

The steps of the hybrid SFS-SA technique are stated as follows:

Hybrid SFS-SA Procedure

1. At each area i run in parallel:
   - Reads the local initial states.
   - Reads the local measurements.
   - Runs the local SE based on SFS.
   - Identify and eliminate internal bad data.
   - Reruns the local SE based on SFS.
   - Obtain the local estimated states for each area i.
2. At the central coordinator area c run:
   - Reads the boundary measurements.
   - Reads the boundary estimated states in each area i.
   - Runs the central coordinator SE based on SA.
   - Identify and eliminate boundary bad data.
   - Reruns the central coordinator SE based on SA.
   - Obtain the coordinator estimated states for area c.

F.4 Case Study System

The presented hybrid SFS-SA technique is implemented in Matlab. IEEE 118-bus is used to illustrate the results [113]. The network consists of 19 generators, 177 lines, 9 transformers, and 91 loads. The network is partitioned into four areas as shown in Figure F.1. Three observable measurement configurations are tested include,
$Case1$. All internal and boundary power flow are measured.
$Case2$. All internal power flows and all boundary power injections are measured.
$Case3$. All power flows and injections are measured.
In this experiment, measurements error is generated by adding white Gaussian noise with zero mean and a standard deviation to load flow results selected as, $\sigma = 0.005$ for voltage magnitude, $\sigma = 0.01$ for power injections, and $\sigma = 0.004$ for power flows.

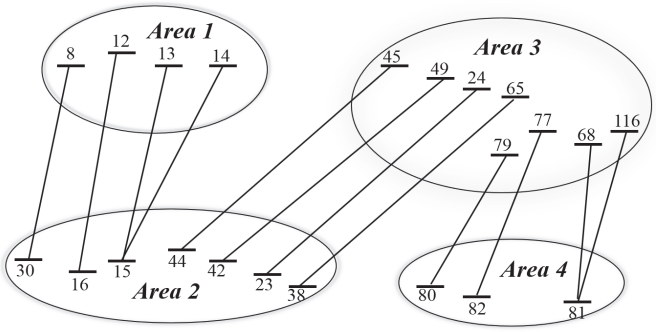

Figure F.1: Multiarea Partitioning of the IEEE 118-bus system

The following objective function is minimized,

$$J(x) = \sum_{i=1}^{m} w_i |z - h(x)| \tag{F.3}$$

F.5 Results and discussions

The proposed hybrid SFS-SA technique is demonstrated on IEEE 118-bus system. Two stat estimation levels are used including local and central coordinator SE. Three different measurement configurations are used. Computational time is assessed based on the results of the hybrid technique.

Bus partitioning for IEEE 118 bus system is tabulated in table F.1. The network is split into four areas. Local phase angle and one voltage magnitude measurement are selected in each local reference bus. Each area has its own local reference bus. For instance, in area 1, local reference bus is bus 10, bus 26 is selected in area 2, bus 69 for area 3, and bus 89 for area 4. The local reference bus in each area is selected based on the highest active and reactive power generation. Active and reactive pairs are used to specify the injection and flow measurement. The tie-line flow measurements are 8-30, 12-16, 13-15, 14-15, 44-45, 42-49, 23-24, 38-65,79-80, 77-82, 68-81, 116-81.

Tables F.2 and F.3 demonstrate the points of the different measurement configurations. These measurement configurations are sorted out as follows: internal and boundary flow measurements are measured, internal flow and only boundary injection measurements are measured, and lastly internal and boundary flow and injection measurements are measured. As can be seen that both cases 1 and 2 have identical

Table F.1: Bus Partitioning for the IEEE 118-bus system

		Type of Buses		No. of Branches
Area	Buses	Internal	Boundary	Internal
1	1-14,117	1-7,9-11,117	8,12-14	18
2	15-23,25-44, 113-115	17-23,25-29 31-37,39-41, 43,113-115	15-16,23,30, 38,42,44	43
3	24,46-79, 116,118	46-48,50-64, 66,67,69-76, 78,118	24,45,49, 65,68,77, 79,116	60
4	80-112	83-112	80-82	50

Table F.2: Measurement Configuration for each area

	Case 1				Case 2				Case 3			
Area	1	2	3	4	1	2	3	4	1	2	3	4
Local measurement $(m_i + 1)$	33	83	105	91	33	83	105	91	47	131	128	155
Local states $(2n_i)$	28	64	72	72	28	64	72	72	28	64	72	72
Local redundancy $R_i = \frac{(m_i+1)}{2n_i}$	1.18	1.30	1.46	1.26	1.18	1.30	1.46	1.26	1.68	2.05	1.78	2.16
Boundary measurement $m_c i$	9	14	16	8	17	28	30	14	17	28	30	14

internal measurements sets, but case 3 has more internal measurements sets than the previous cases. Therefore, the redundancy index is higher in case 3 than cases 1 and 2. This due to the inclusion of internal injection measurements in local areas in case 3. Moreover, Cases 2 and 3 have identical boundary measurement sets, but case 1 has fewer boundary measurements than cases 2 and 3. Consequently, the redundancy index is smaller in case 1 than cases 2 and 3. This due to the inclusion of boundary injection measurements in coordinator in cases 2 and 3. Table F.4 shows the comparison in CPU time between local areas, coordinator, and distributed algorithms. the overall CPU time of the distributed algorithm is the sum of the CPU times of the coordinator and the slowest area 4 estimator in seconds. As can be seen, cases 1

Table F.3: Measurement Configuration for the overall system

	Case 1	Case 2	Case 3
Total measurement (m)	359	401	550
Total states ($2n$)	236	236	236
Global redundancy ($r = \frac{m}{2n}$)	1.52	1.70	2.33
Boundary measurement ($m_c + r - 1$)	47	89	89

Table F.4: Speed comparison between different techniques

SE techniques	Case 1	Case 2	Case 3
Integrated	6.6472489 (100%)	6.6472489 (100%)	7.7036457 (100%)
Area 1	1.16724907	1.16724907	1.30637165
Area 2	1.26432514	1.26432514	1.52406487
Area 3	1.70786793	1.70786793	2.13189080
Area 4	2.12901911	2.12901911	2.94049257
Coordination	1.41732837	1.48411886	1.48411886
Distributed	3.54634748 (53.35%)	3.61313797 (54.36%)	4.42461143 (57.44%)

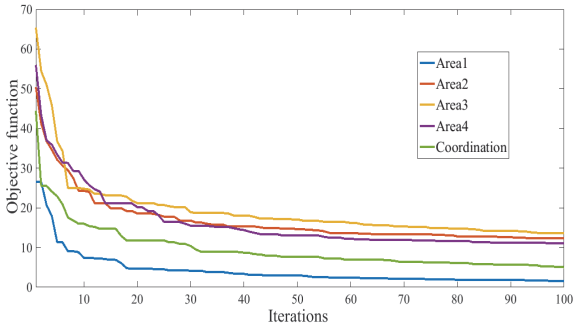

Figure F.2: Sum of absolute percentage error (objective function) for cases 1 & 2 in local areas and case 1 in coordinator level

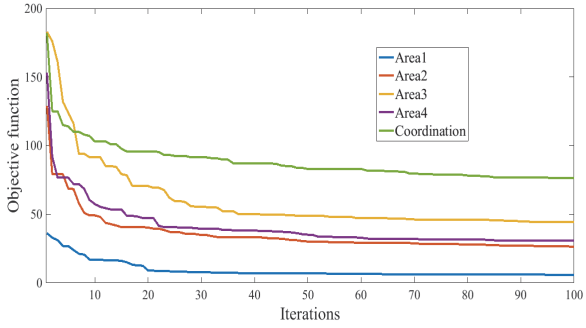

Figure F.3: Sum of absolute percentage error (objective function) for cases 2 & 3 in coordinator level and case 3 in local areas

and 2 have the same CPU time due to the identical internal measurements in local areas. Case 3 in local areas has slower CPU time than cases 1 and 2. The reason is that case 3 has more internal measurements. Moreover, cases 2 and 3 have the same CPU time due to the identical boundary measurements in the coordinator. Case 1 in coordinator has faster CPU time than other cases. The reason is that case 1 includes only boundary line flow measurements. In addition, CPU time of the distributed algorithm is slower in case 3 than cases 1 and 2 because case 3 includes all measurements such as internal and boundary line flows and injections.

Figures F.1 and F.2 present the sum of absolute error of the objective function for each local estimator and coordinator. As can be seen, Area 1 has the least sum of absolute error in all cases compared to other areas. The sum of absolute error for coordinator estimator in case 1 is lesser than other cases.

F.6 Conclusion

The chapter presented a distributed multiarea SE. The IEEE 118-bus system is split into four areas. Each area executes its own local SE and the only minimum of border information is shared to a coordination. Stochastic fractal search (SFS) technique is used to perform local SE. Moreover, simulated annealing technique(SA)technique is used to perform wide-area SE. Three different measurement configurations are implemented. The obtained results demonstrated a significant reduction in the computational time.

Bibliography

[1] A. Dubey and S. V. Sagar, "Optimization of static transfer switch operation for impedance matching in feeder reconfiguration", *ITSI Transactions on Electrical and Electronics Engineering*, vol. 1, no. 1, 2320 – 8945, 2013.

[2] J. Chen, "Power system state estimation using phasor measurement units", 2013.

[3] T. Mahmood and M. A. Choudhry, "Application of state estimation technique for impedance matching to optimize static transfer switch operation", *2007 International Conference on Emerging Technologies*, pp. 269–273, 2007.

[4] S. Wang, W. Gao, and A. P. S. Meliopoulos, "An alternative method for power system dynamic state estimation based on unscented transform", *IEEE Transactions on Power Systems*, vol. 27, no. 2, pp. 942–950, 2012.

[5] K. A. Clements, G. R. Krumpholz, and P. W. Davis, "Power system state estimation residual analysis: An algorithm using network topology", *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-100, no. 4, pp. 1779–1787, 1981.

[6] R. D. Schweppe F.C. Wildes J., "Power system static-state estimation, part i: Exact model, part ii: Approximate model, part iii: Implementation", *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-89, no. 1, pp. 120–125, 1970.

[7] A. Abur, "A bad data identification method for linear programming state estimation", *IEEE Transactions on Power Systems*, vol. 5, no. 3, pp. 894–901, 1990.

[8] K. A. Clements, G. W. Woodzell, and R. C. Burchett, "A new method for solving equality-constrained power system static-state estimation", *IEEE Transactions on Power Systems*, vol. 5, no. 4, pp. 1260–1266, 1990.

[9] D. M. V. Kumar, S. C. Srivastava, S. Shah, and S. Mathur, "Topology processing and static state estimation using artificial neural networks", *IEE Proceedings - Generation, Transmission and Distribution*, vol. 143, no. 1, pp. 99–105, 1996.

[10] A. Abur and A. G. Exposito, "Bad data identification when using ampere measurements", *IEEE Transactions on Power Systems*, vol. 12, no. 2, pp. 831–836, 1997.

[11] H. Wei, H. Sasaki, J. Kubokawa, and R. Yokoyama, "An interior point method for power system weighted nonlinear l/sub 1/ norm static state estimation", *IEEE Transactions on Power Systems*, vol. 13, no. 2, pp. 617–623, 1998.

[12] A. Jain, Y. Kawazoe, R. Balasubramanian, and S. C. Tripathy, "Network observability: A solution technique using neural networks", *TENCON 2003. Conference on Convergent Technologies for Asia-Pacific Region*, vol. 3, 1007–1011 Vol.3, 2003.

[13] T. Q. D. Khoa, P. T. T. Binh, and T. V. Khoa, "Hopfield network and parallel genetic algorithm for solving state estimate in power systems", *2004 International Conference on Power System Technology, 2004. PowerCon 2004.*, vol. 1, pp. 845–849, 2004.

[14] M.-H. Vuong, S. Lefebvre, and P. Lagace, "Redundancy in topology errors detection and identification", *2004 IEEE International Conference on Industrial Technology, 2004. IEEE ICIT '04.*, vol. 3, pp. 1494–1498, 2004.

[15] A. Jain, R. Balasubramanian, S. C. Tripathy, B. N. Singh, and Y. Kawazoe, "Power system topological observability analysis using artificial neural networks", *IEEE Power Engineering Society General Meeting, 2005*, vol. 1, pp. 497–502, 2005.

[16] B. Gou, "Jacobian matrix-based observability analysis for state estimation", *IEEE Transactions on Power Systems*, vol. 21, no. 1, pp. 348–356, 2006.

[17] 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, 2009.

[18] M. R. Irving, "Robust algorithm for generalized state estimation", *IEEE Transactions on Power Systems*, vol. 24, no. 4, pp. 1886–1887, 2009.

[19] E. Caro, A. J. Conejo, and A. Abur, "Breaker status identification", *IEEE Transactions on Power Systems*, vol. 25, no. 2, pp. 694–702, 2010.

[20] G. N. Korres, "A robust algorithm for power system state estimation with equality constraints", *IEEE Transactions on Power Systems*, vol. 25, no. 3, pp. 1531–1541, 2010.

[21] T. Yang, H. Sun, and A. Bose, "Transition to a two-level linear state estimator—part ii: Algorithm", *IEEE Transactions on Power Systems*, vol. 26, no. 1, pp. 54–62, 2011.

[22] F. Aminifar, M. Fotuhi-Firuzabad, M. Shahidehpour, and A. Khodaei, "Probabilistic multistage pmu placement in electric power systems", *IEEE Transactions on Power Delivery*, vol. 26, no. 2, pp. 841–849, 2011.

[23] H. T. Kung, T.-H. Lin, and D. Vlah, "Identifying bad measurements in compressive sensing", *2011 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, pp. 1054–1059, 2011.

[24] O. Ivanov and M. Garvrilaş, "State estimation for power systems with multilayer perceptron neural networks", *11th Symposium on Neural Network Applications in Electrical Engineering*, pp. 243–246, 2012.

[25] T. E. Dy Liacco, "Real-time computer control of power systems", *Proceedings of the IEEE*, vol. 62, no. 7, pp. 884–891, 1974, ISSN: 0018-9219.

[26] ——, "The adaptive reliability control system", *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-86, no. 5, pp. 517–531, 1967, ISSN: 0018-9510.

[27] A. Gomez-Exposito and A. Abur, *Power system state estimation: theory and implementation*. CRC press, 2004.

[28] R. P. Schulz and W. W. Price, "Classification and identification of power system emergencies", *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-103, no. 12, pp. 3470–3479, 1984, ISSN: 0018-9510.

[29] D. J. Gaushell and H. T. Darlington, "Supervisory control and data acquisition", *Proceedings of the IEEE*, vol. 75, no. 12, pp. 1645–1658, 1987.

[30] F. F. Wu, K. Moslehi, and A. Bose, "Power system control centers: Past, present, and future", *Proceedings of the IEEE*, vol. 93, no. 11, pp. 1890–1908, 2005.

[31] A. P. Meliopoulos, A. Feliachi, A. G. Bakirtzis, and G. Cokkinides, "Development of courses on power system energy control centers", *IEEE Transactions on Education*, vol. 27, no. 2, pp. 66–72, 1984.

[32] K. Emami, "Dynamic state estimation and control of power systems", PhD thesis, 2015.

[33] N. Amjady and A. Daraeepour, "Mixed price and load forecasting of electricity markets by a new iterative prediction method", *Electric power systems research*, vol. 79, no. 9, pp. 1329–1336, 2009.

[34] A. Karsaz, H. R. Mashhadi, and M. M. Mirsalehi, "Market clearing price and load forecasting using cooperative co-evolutionary approach", *International Journal of Electrical Power & Energy Systems*, vol. 32, no. 5, pp. 408–415, 2010.

[35] R. Weron, "Electricity price forecasting: A review of the state-of-the-art with a look into the future", *International journal of forecasting*, vol. 30, no. 4, pp. 1030–1081, 2014.

[36] J. Catalão, S. Mariano, V. Mendes, and L. Ferreira, "An artificial neural network approach for short-term electricity prices forecasting", *Intelligent Systems Applications to Power Systems, 2007. ISAP 2007. International Conference on*, pp. 1–6, 2007.

[37] V. Iyer, C. C. Fung, and T. Gedeon, "A fuzzy-neural approach to electricity load and spot-price forecasting in a deregulated electricity market", *TENCON 2003. Conference on Convergent Technologies for the Asia-Pacific Region*, vol. 4, pp. 1479–1482, 2003.

[38] K. B. Sahay and M. Tripathi, "Day ahead hourly load and price forecast in iso new england market using ann", *India Conference (INDICON), 2013 Annual IEEE*, pp. 1–6, 2013.

[39] G. Li, C.-C. Liu, J. Lawarree, M. Gallanti, and A. Venturini, "State-of-the-art of electricity price forecasting", *CIGRE/IEEE PES, 2005. International Symposium*, pp. 110–119, 2005.

[40] A. Karsaz, H. R. Mashhadi, and R. Eshraghnia, "Cooperative co-evolutionary approach to electricity load and price forecasting in deregulated electricity markets", *Power India Conference, 2006 IEEE*, 6–pp, 2006.

[41] D. W. Bunn, "Forecasting loads and prices in competitive power markets", *Proceedings of the IEEE*, vol. 88, no. 2, pp. 163–169, 2000.

[42] P. L. Skantze and M. D. Ilic, "Valuation, hedging and speculaton in competitive electricity markets: A fundamental approach", 2001.

[43] M Shahidehpour, H Yamin, and Z. Li, "Market operations in electric power systems: Forecasting, scheduling, and risk management", 2002.

[44] V. S. Koritarov, "Real-world market representation with agents", *IEEE Power and Energy Magazine*, vol. 2, no. 4, pp. 39–46, 2004.

[45] M. Ventosa, A. Baıllo, A. Ramos, and M. Rivier, "Electricity market modeling trends", *Energy policy*, vol. 33, no. 7, pp. 897–913, 2005.

[46] N. Amjady and M. Hemmati, "Energy price forecasting-problems and proposals for such predictions", *IEEE Power and Energy Magazine*, vol. 4, no. 2, pp. 20–29, 2006.

[47] R. Weron, "Modeling and forecasting electricity loads and prices: A statistical approach (the wiley finance series)", 2006.

[48] H. R. Zareipour, *Price-based Energy Management in Competitive Electricity Markets: Price Forecasting and Optimal Operation of Wholesale Customers*. VDM Publishing, 2008.

[49] N Amjady, "Short-term electricity price forecasting", *Electric power systems: advanced forecasting techniques and optimal generation scheduling*, pp. 1–58, 2012.

[50] C. García-Martos and A. J. Conejo, "Price forecasting techniques in power systems", *Wiley Encyclopedia of Electrical and Electronics Engineering*, pp. 1–23, 2013.

[51] T. Hong *et al.*, "Energy forecasting: Past, present, and future", *Foresight: The International Journal of Applied Forecasting*, no. 32, pp. 43–48, 2014.

[52] C. P. Rodriguez and G. J. Anders, "Energy price forecasting in the ontario competitive power system market", *IEEE transactions on power systems*, vol. 19, no. 1, pp. 366–374, 2004.

[53] R. Carmona and M. Coulon, "A survey of commodity markets and structural models for electricity prices", pp. 41–83, 2014.

[54] AMEO. (2015). The australian energy market operator@ONLINE.

[55] M. Marzband, S. S. Ghazimirsaeid, H. Uppal, and T. Fernando, "A real-time evaluation of energy management systems for smart hybrid home microgrids", *Electric Power Systems Research*, vol. 143, pp. 624–633, 2017.

[56] M. Marzband, M. M. Moghaddam, M. F. Akorede, and G. Khomeyrani, "Adaptive load shedding scheme for frequency stability enhancement in microgrids", *Electric Power Systems Research*, vol. 140, pp. 78–86, 2016.

[57] M. Marzband, M. Javadi, J. L. Domínguez-García, and M. M. Moghaddam, "Non-cooperative game theory based energy management systems for energy district in the retail market considering der uncertainties", *IET Generation, Transmission & Distribution*, vol. 10, no. 12, pp. 2999–3009, 2016.

[58] M. Marzband, N. Parhizi, M. Savaghebi, and J. M. Guerrero, "Distributed smart decision-making for a multimicrogrid system based on a hierarchical interactive architecture", *IEEE Transactions on Energy Conversion*, vol. 31, no. 2, pp. 637–648, 2016.

[59] M. Marzband, F. Azarinejadian, M. Savaghebi, and J. M. Guerrero, "An optimal energy management system for islanded microgrids based on multiperiod artificial bee colony combined with markov chain", *IEEE Systems Journal*, vol. 11, no. 3, pp. 1712–1722, 2017.

[60] M. Marzband, E. Yousefnejad, A. Sumper, and J. L. Domínguez-García, "Real time experimental implementation of optimum energy management system in

standalone microgrid by using multi-layer ant colony optimization", *International Journal of Electrical Power & Energy Systems*, vol. 75, pp. 265–274, 2016.

[61] M. Marzband, M. Ghadimi, A. Sumper, and J. L. Domínguez-García, "Experimental validation of a real-time energy management system using multi-period gravitational search algorithm for microgrids in islanded mode", *Applied energy*, vol. 128, pp. 164–174, 2014.

[62] M. Marzband, N. Parhizi, and J. Adabi, "Optimal energy management for stand-alone microgrids based on multi-period imperialist competition algorithm considering uncertainties: Experimental validation", *International transactions on electrical energy systems*, vol. 26, no. 6, pp. 1358–1372, 2016.

[63] M. Marzband, A. Sumper, A. Ruiz-Álvarez, J. L. Domínguez-García, and B. Tomoiagă, "Experimental evaluation of a real time energy management system for stand-alone microgrids in day-ahead markets", *Applied Energy*, vol. 106, pp. 365–376, 2013.

[64] M. Marzband, A. Sumper, J. L. Domínguez-García, and R. Gumara-Ferret, "Experimental validation of a real time energy management system for microgrids in islanded mode using a local day-ahead electricity market and minlp", *Energy Conversion and Management*, vol. 76, pp. 314–322, 2013.

[65] V. Basetti, A. K. Chandel, and R. Chandel, "Power system dynamic state estimation using prediction based evolutionary technique", *Energy*, vol. 107, pp. 29–47, 2016.

[66] C. Gu and P. Jirutitijaroen, "Dynamic state estimation under communication failure using kriging based bus load forecasting", *IEEE Transactions on Power Systems*, vol. 30, no. 6, pp. 2831–2840, 2015.

[67] M. H. Amini, M. Rahmani, K. G. Boroojeni, G. Atia, S. S. Iyengar, and O. Karabasoglu, "Sparsity-based error detection in dc power flow state estimation", *arXiv preprint arXiv:1605.04380*, 2016.

[68] X. Tai, D. Marelli, E. Rohr, and M. Fu, "Optimal pmu placement for power system state estimation with random component outages", *International Journal of Electrical Power & Energy Systems*, vol. 51, pp. 35–42, 2013.

[69] X. Chen, J. Lin, C. Wan, Y. Song, S. You, Y. Zong, W. Guo, and Y. Li, "Optimal meter placement for distribution network state estimation: A circuit representation based milp approach", *IEEE Transactions on Power Systems*, vol. 31, no. 6, pp. 4357–4370, 2016.

[70] H. G. Abood and V. Sreeram, "A review on phasor measurement units placement for state estimation studies", *Power Engineering Conference (AUPEC), 2014 Australasian Universities*, pp. 1–6, 2014.

[71] H. Karimipour and V. Dinavahi, "Extended kalman filter-based parallel dynamic state estimation", *IEEE transactions on smart grid*, vol. 6, no. 3, pp. 1539–1549, 2015.

[72] J. Zhang, G. Welch, G. Bishop, and Z. Huang, "A two-stage kalman filter approach for robust and real-time power system state estimation", *IEEE Transactions on Sustainable Energy*, vol. 5, no. 2, pp. 629–636, 2014.

[73] K.-R. Shih and S.-J. Huang, "Application of a robust algorithm for dynamic state estimation of a power system", *IEEE Transactions on Power Systems*, vol. 17, no. 1, pp. 141–147, 2002.

[74] G. D. Prasad and S. Thakur, "A new approach to dynamic state estimation of power systems", *Electric power systems research*, vol. 45, no. 3, pp. 173–180, 1998.

[75] G. Valverde and V. Terzija, "Unscented kalman filter for power system dynamic state estimation", *IET generation, transmission & distribution*, vol. 5, no. 1, pp. 29–37, 2011.

[76] X. Qing, H. R. Karimi, Y. Niu, and X. Wang, "Decentralized unscented kalman filter based on a consensus algorithm for multi-area dynamic state estimation in power systems", *International Journal of Electrical Power & Energy Systems*, vol. 65, pp. 26–33, 2015.

[77] Y. Del Valle, G. K. Venayagamoorthy, S. Mohagheghi, R. G. Harley, and J. Hernandez, "Particle swarm optimization: Basic concepts, variants and applications in power systems", 2008.

[78] M. B. Do Coutto Filho and J. C. S. de Souza, "Forecasting-aided state estimation—part i: Panorama", *IEEE Transactions on Power Systems*, vol. 24, no. 4, pp. 1667–1677, 2009.

[79] J.-M. Lin, S.-J. Huang, and K.-R. Shih, "Application of sliding surface-enhanced fuzzy control for dynamic state estimation of a power system", *IEEE Transactions on Power Systems*, vol. 18, no. 2, pp. 570–577, 2003.

[80] A. Sinha and J. Mondal, "Dynamic state estimator using ann based bus load prediction", *IEEE Transactions on Power Systems*, vol. 14, no. 4, pp. 1219–1225, 1999.

[81] T. T. Nguyen, S. Yang, and J. Branke, "Evolutionary dynamic optimization: A survey of the state of the art", *Swarm and Evolutionary Computation*, vol. 6, pp. 1–24, 2012.

[82] H. Salimi, "Stochastic fractal search: A powerful metaheuristic algorithm", *Knowledge-Based Systems*, vol. 75, pp. 1–18, 2015.

[83] M. Mitchell, *An introduction to genetic algorithms*. 1998.

[84] R. Eberhart and J. Kennedy, "A new optimizer using particle swarm theory", *Micro Machine and Human Science, 1995. MHS'95., Proceedings of the Sixth International Symposium on*, pp. 39–43, 1995.

[85] D. Karaboga, "An idea based on honey bee swarm for numerical optimization", Technical report-tr06, Erciyes university, engineering faculty, computer engineering department, Tech. Rep., 2005.

[86] M. Dorigo and G. Di Caro, "Ant colony optimization: A new meta-heuristic", *Evolutionary Computation, 1999. CEC 99. Proceedings of the 1999 congress on*, vol. 2, pp. 1470–1477, 1999.

[87] A. V. Chechkin, R. Metzler, J. Klafter, V. Y. Gonchar, *et al.*, "Introduction to the theory of lévy flights", *Anomalous transport: Foundations and applications*, vol. 49, no. 2, pp. 431–451, 2008.

[88] P. K. Rawlings, "Modes of a gaussian random walk", *Journal of statistical physics*, vol. 111, no. 3-4, pp. 769–788, 2003.

[89] T. A. Witten and L. M. Sander, "Diffusion-limited aggregation", *Physical Review B*, vol. 27, no. 9, p. 5686, 1983.

[90] Y. Cherdantseva, P. Burnap, A. Blyth, P. Eden, K. Jones, H. Soulsby, and K. Stoddart, "A review of cyber security risk assessment methods for scada systems", *Computers & security*, vol. 56, pp. 1–27, 2016.

[91] K. G. Boroojeni, M. H. Amini, and S. Iyengar, "Overview of the security and privacy issues in smart grids", *Smart grids: security and privacy issues*, pp. 1–16, 2017.

[92] K. G. Boroojeni, M. H. Amini, and S. S. Iyengar, *Smart grids: security and privacy issues*. Springer, 2017.

[93] C.-C. Liu, A. Stefanov, J. Hong, and P. Panciatici, "Intruders in the grid", *IEEE Power and Energy magazine*, vol. 10, no. 1, pp. 58–66, 2012.

[94] G. Chaojun, D. Yang, P. Jirutitijaroen, W. M. Walsh, and T. Reindl, "Spatial load forecasting with communication failure using time-forward kriging", *IEEE Transactions on Power Systems*, vol. 29, no. 6, pp. 2875–2882, 2014.

[95] M. Long, C.-H. Wu, and J. Y. Hung, "Denial of service attacks on network-based control systems: Impact and mitigation", *IEEE Transactions on Industrial Informatics*, vol. 1, no. 2, pp. 85–96, 2005.

[96] M. Panteli and D. S. Kirschen, "Assessing the effect of failures in the information and communication infrastructure on power system reliability", *Power Systems Conference and Exposition (PSCE), 2011 IEEE/PES*, pp. 1–7, 2011.

[97] X. Wu and J. D. Westervelt, "Engineering applications of neural computing: A state-of-the-art survey", Construction Engineering Research Lab (ARMY) Champian IL, Tech. Rep., 1991.

[98] Y. Liu, C.-f. Zhou, and Y.-w. Chen, "Weight initialization of feedforward neural networks by means of partial least squares", *Machine Learning and Cybernetics, 2006 International Conference on*, pp. 3119–3122, 2006.

[99] N. N. Schraudolph, "Fast curvature matrix-vector products for second-order gradient descent", *Neural computation*, vol. 14, no. 7, pp. 1723–1738, 2002.

[100] T.-C. R. Hsiao, C.-W. Lin, and H. K. Chiang, "Partial least-squares algorithm for weights initialization of backpropagation network", *Neurocomputing*, vol. 50, pp. 237–247, 2003.

[101] T. M. Varnava and A. J. MEADE JR, "An initialization method for feedforward artificial neural networks using polynomial bases", *Advances in Adaptive Data Analysis*, vol. 3, no. 03, pp. 385–400, 2011.

[102] D. D. Silalahi, C. E. Reaño, F. P. Lansigan, R. G. Panopio, and N. C. Bantayan, "Using genetic algorithm neural network on near infrared spectral data for ripeness grading of oil palm (elaeis guineensis jacq.) fresh fruit", *Information Processing in Agriculture*, vol. 3, no. 4, pp. 252–261, 2016.

[103] L. Hu, L. Qin, K. Mao, W. Chen, and X. Fu, "Optimization of neural network by genetic algorithm for flowrate determination in multipath ultrasonic gas flowmeter", *IEEE Sensors Journal*, vol. 16, no. 5, pp. 1158–1167, 2016.

[104] W. Sun and Y. Xu, "Using a back propagation neural network based on improved particle swarm optimization to study the influential factors of carbon dioxide emissions in hebei province, china", *Journal of Cleaner Production*, vol. 112, pp. 1282–1291, 2016.

[105] J.-R. Zhang, J. Zhang, T.-M. Lok, and M. R. Lyu, "A hybrid particle swarm optimization–back-propagation algorithm for feedforward neural network training", *Applied mathematics and computation*, vol. 185, no. 2, pp. 1026–1037, 2007.

[106] M. Dorigo, V. Maniezzo, and A. Colorni, "Ant system: Optimization by a colony of cooperating agents", *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 26, no. 1, pp. 29–41, 1996.

[107] C. S. Koh, O. A. Mohammed, and S.-Y. Hahn, "Detection of magnetic body using artificial neural network with modified simulated annealing", *IEEE transactions on magnetics*, vol. 30, no. 5, pp. 3644–3647, 1994.

[108] S. Du, W. Li, and K. Cao, "A learning algorithm of artificial neural network based on ga-pso", *Intelligent Control and Automation, 2006. WCICA 2006. The Sixth World Congress on*, vol. 1, pp. 3633–3637, 2006.

[109] A. J. Smola and B. Schölkopf, "A tutorial on support vector regression", *Statistics and computing*, vol. 14, no. 3, pp. 199–222, 2004.

[110] NYISO. (2016). New york independent system operator@ONLINE.

[111] NEMA. (2016). The association of electrical equipment and medical imaging manufacturers@ONLINE.

[112] S Dhivya and T Vigneswaran, "Primal dual interior point algorithm for constrained economic load dispatch and optimal power flow", *Power, Energy and Control (ICPEC), 2013 International Conference on*, pp. 360–365, 2013.

[113] UWEE. (2016). Power systems test case archive@ONLINE.

[114] M. Valipour, "Optimization of neural networks for precipitation analysis in a humid region to detect drought and wet year alarms", *Meteorological Applications*, vol. 23, no. 1, pp. 91–100, 2016.

[115] M. Valipour, M. E. Banihabib, and S. M. R. Behbahani, "Comparison of the arma, arima, and the autoregressive artificial neural network models in forecasting the monthly inflow of dez dam reservoir", *Journal of hydrology*, vol. 476, pp. 433–441, 2013.

[116] M. Valipour, "Use of average data of 181 synoptic stations for estimation of reference crop evapotranspiration by temperature-based methods", *Water resources management*, vol. 28, no. 12, pp. 4237–4255, 2014.

[117] M. Rezaei, M. Valipour, and M. Valipour, "Modelling evapotranspiration to increase the accuracy of the estimations based on the climatic parameters", *Water Conservation Science and Engineering*, vol. 1, no. 3, pp. 197–207, 2016.

[118] V. Basetti and A. K. Chandel, "A robust lws state estimation including anomaly detection and identification in power systems", *Neurocomputing*, vol. 166, pp. 122–132, 2015.

[119] X. Gao, Y. Chen, D. You, Z. Xiao, and X. Chen, "Detection of micro gap weld joint by using magneto-optical imaging and kalman filtering compensated with rbf neural network", *Mechanical Systems and Signal Processing*, vol. 84, pp. 570–583, 2017.

[120] L. Wu, Z. Feng, and W. X. Zheng, "Exponential stability analysis for delayed neural networks with switching parameters: Average dwell time approach", *IEEE Transactions on Neural Networks*, vol. 21, no. 9, pp. 1396–1407, 2010.

[121] P. Shi, F. Li, L. Wu, and C.-C. Lim, "Neural network-based passive filtering for delayed neutral-type semi-markovian jump systems", *IEEE Transactions on Neural Networks and Learning Systems*, vol. 28, no. 9, pp. 2101–2114, 2017.

[122] S.-S. Xiong and Z.-Y. Zhou, "Neural filtering of colored noise based on kalman filter structure", *IEEE Transactions on Instrumentation and Measurement*, vol. 52, no. 3, pp. 742–747, 2003.

[123] V Vaidehi, N Chitra, M Chokkalingam, and C. Krishnan, "Neural network aided kalman filtering for multitarget tracking applications", *Computers and Electrical Engineering*, vol. 2, no. 27, pp. 217–228, 2001.

[124] L Chin, "Applications of the neural network approaches to multiple target tracking", *IEEE Trans. Aerospace and Electronic Systems*, vol. 30, no. 1, pp. 1220–1223, 1994.

[125] X. Gao, X. Zhong, D. You, and S. Katayama, "Kalman filtering compensated by radial basis function neural network for seam tracking of laser welding.", *IEEE Trans. Contr. Sys. Techn.*, vol. 21, no. 5, pp. 1916–1923, 2013.

[126] X. Gao, D. You, and S. Katayama, "Seam tracking monitoring based on adaptive kalman filter embedded elman neural network during high-power fiber laser welding", *IEEE Transactions on Industrial Electronics*, vol. 59, no. 11, pp. 4315–4325, 2012.

[127] A. N. AL-Masri, M. Z. A. Ab Kadir, H. Hizam, and N. Mariun, "A novel implementation for generator rotor angle stability prediction using an adaptive artificial neural network application for dynamic security assessment", *IEEE Transactions on Power Systems*, vol. 28, no. 3, pp. 2516–2525, 2013.

[128] A. G. Parlos, S. K. Menon, and A Atiya, "An algorithmic approach to adaptive state filtering using recurrent neural networks", *IEEE Transactions on Neural Networks*, vol. 12, no. 6, pp. 1411–1432, 2001.

[129] L. Chin, "Application of neural networks in target tracking data fusion", *IEEE Transactions on Aerospace and Electronic Systems*, vol. 30, no. 1, pp. 281–287, 1994.

[130] A. Jain and N. Shivakumar, "Power system tracking and dynamic state estimation", *IEEE PES Power Systems Conference Exposition (PSCE)*, 2009.

[131] A. S. Debs and R. E. Larson, "A dynamic estimator for tracking the state of a power system", *IEEE Transactions on Power Apparatus and Systems*, no. 7, pp. 1670–1678, 1970.

[132] F. Schweppe and R. Masiello, "A tracking static state estimator", *IEEE Transactions on Power Apparatus and Systems*, no. 3, pp. 1025–1033, 1971.

[133] A. Sinha, L Roy, and H. Srivastava, "A new and fast tracking state estimator for multiterminal dc/ac power systems", *TENCON'89. Fourth IEEE Region 10 International Conference*, pp. 949–952, 1989.

[134] H. Mosbah and M. El-Hawary, "Optimization of neural network parameters by stochastic fractal search for dynamic state estimation under communication failure", *Electric Power Systems Research*, vol. 147, pp. 288–301, 2017.

[135] D. Tungadio, B. Numbi, M. Siti, and A. A. Jimoh, "Particle swarm optimization for power system state estimation", *Neurocomputing*, vol. 148, pp. 175–180, 2015.

[136] G. N. Korres and N. M. Manousakis, "State estimation and bad data processing for systems including pmu and scada measurements", *Electric Power Systems Research*, vol. 81, no. 7, pp. 1514–1524, 2011.

[137] H Sasaki, K Aoki, and R Yokoyama, "A parallel computation algorithm for static state estimation by means of matrix inversion lemma", *IEEE transactions on power systems*, vol. 2, no. 3, pp. 624–631, 1987.

[138] O. Ivanov and M. Gavrilaş, "State estimation with neural networks and pmu voltage measurements", *Electrical and Power Engineering (EPE), 2014 International Conference and Exposition on*, pp. 983–988, 2014.

[139] G. Panchal, A. Ganatra, Y. Kosta, and D. Panchal, "Behaviour analysis of multilayer perceptronswith multiple hidden neurons and hidden layers", *International Journal of Computer Theory and Engineering*, vol. 3, no. 2, p. 332, 2011.

[140] H. Mosbah and M. El-Hawary, "Multilayer artificial neural networks for real time power system state estimation", *Electrical Power and Energy Conference (EPEC), 2015 IEEE*, pp. 344–351, 2015.

[141] F. C. Schweppe, "Power system static state estimation", 1969.

[142] S. Chohan, "Static and tracking state estimation in power systems with bad data analysis", PhD thesis, 1993.

[143] D. Falcao, P. Cooke, and A Brameller, "Power system tracking state estimation and bad data processing", *IEEE Transactions on Power Apparatus and Systems*, no. 2, pp. 325–333, 1982.

[144] B. B. Mandelbrot, "The fractal geometry of nature/revised and enlarged edition", *New York, WH Freeman and Co., 1983, 495 p.*, 1983.

[145] B. Birge, "Psot-a particle swarm optimization toolbox for use with matlab.", *SIS*, vol. 3, pp. 973–990, 2003.

[146] A. Chipperfield and P. Fleming, "The matlab genetic algorithm toolbox", 1995.

[147] E. Handschin, "Real-time data processing using state estimation in electric power systems", *Real-Time Control of Electric Power Systems*, pp. 29–61, 1972.

[148] F. Shweppe, "Power system static-state estimation, part i-iii", *IEEE Trans. Power Apparatus System*, vol. 89, no. 1, pp. 120–135, 1970.

[149] F. Aminifar, M. Shahidehpour, M. Fotuhi-Firuzabad, and S. Kamalinia, "Power system dynamic state estimation with synchronized phasor measurements", *IEEE Transactions on Instrumentation and Measurement*, vol. 63, no. 2, pp. 352–363, 2014.

[150] A. L. Da Silva, M. Do Coutto Filho, and J. Cantera, "An efficient dynamic state estimation algorithm including bad data processing", *IEEE transactions on Power Systems*, vol. 2, no. 4, pp. 1050–1058, 1987.

[151] S. Mallick, S. Ghoshal, P Acharjee, and S. Thakur, "Optimal static state estimation using improved particle swarm optimization and gravitational search algorithm", *International Journal of Electrical Power & Energy Systems*, vol. 52, pp. 254–265, 2013.

[152] F. Shweppe, "Power system static-state estimation, part i-iii", *IEEE Trans. Power Apparatus System*, vol. 89, no. 1, pp. 120–135, 1970.

[153] H. Mosbah and M. El-Hawary, "Power system tracking state estimation based on stochastic fractal search technique under bad measurements conditions", *Electrical Power and Energy Conference (EPEC), 2016 IEEE*, pp. 1–6, 2016.

[154] H. Mosbah and M. El-Hawary, "Power system tracking state estimation based on stochastic fractal search technique under sudden load changing conditions", *Electrical and Computer Engineering (CCECE), 2016 IEEE Canadian Conference on*, pp. 1–6, 2016.

[155] S. Chakrabarti, E. Kyriakides, T. Bi, D. Cai, and V. Terzija, "Measurements get together", *IEEE Power and Energy Magazine*, vol. 7, no. 1, 2009.

[156] S. Chakrabarti and E. Kyriakides, "Optimal placement of phasor measurement units for power system observability", *IEEE Transactions on power systems*, vol. 23, no. 3, pp. 1433–1440, 2008.

[157] S Chakrabarti, E Kyriakides, G Ledwich, and A. Ghosh, "Inclusion of pmu current phasor measurements in a power system state estimator", *IET generation, transmission & distribution*, vol. 4, no. 10, pp. 1104–1115, 2010.

[158] J. Zhu and A. Abur, "Effect of phasor measurements on the choice of reference bus for state estimation", *Power Engineering Society General Meeting, 2007. IEEE*, pp. 1–5, 2007.

[159] S. Chakrabarti, E. Kyriakides, G. Valverde, and V. Terzija, "State estimation including synchronized measurements", *PowerTech, 2009 IEEE Bucharest*, pp. 1–5, 2009.

[160] M. Hurtgen and J.-C. Maun, "Advantages of power system state estimation using phasor measurement units", *16th Power Systems Computation Conference*, pp. 1–7, 2008.

[161] T. Bi, X. Qin, and Q. Yang, "A novel hybrid state estimator for including synchronized phasor measurements", *Electric Power Systems Research*, vol. 78, no. 8, pp. 1343–1352, 2008.

[162] J. Zhu and A. Abur, "Bad data identification when using phasor measurements", *Power Tech, 2007 IEEE Lausanne*, pp. 1676–1681, 2007.

[163] R. Sodhi, S. Srivastava, and S. Singh, "Phasor-assisted hybrid state estimator", *Electric Power Components and Systems*, vol. 38, no. 5, pp. 533–544, 2010.

[164] S. Chakrabarti, E. Kyriakides, and D. G. Eliades, "Placement of synchronized measurements for power system observability", *IEEE Transactions on Power Delivery*, vol. 24, no. 1, pp. 12–19, 2009.

[165] T. Van Cutsem, J.-L. Horward, and M Ribbens-Pavella, "A two-level static state estimator for electric power systems", *IEEE transactions on power apparatus and systems*, no. 8, pp. 3722–3732, 1981.

[166] A. El-Keib, J Nieplocha, H Singh, and D. Maratukulam, "A decomposed state estimation technique suitable for parallel processor implementation", *IEEE transactions on power systems*, vol. 7, no. 3, pp. 1088–1097, 1992.

[167] R. Ebrahimian and R. Baldick, "State estimation distributed processing [for power systems]", *IEEE Transactions on Power Systems*, vol. 15, no. 4, pp. 1240–1246, 2000.

[168] A. J. Conejo, S. de la Torre, and M. Canas, "An optimization approach to multiarea state estimation", *IEEE Transactions on Power Systems*, vol. 22, no. 1, pp. 213–221, 2007.

[169] G. N. Korres, "A distributed multiarea state estimation", *IEEE Transactions on Power Systems*, vol. 26, no. 1, pp. 73–84, 2011.

[170] L. Xie, D.-H. Choi, S. Kar, and H. V. Poor, "Fully distributed state estimation for wide-area monitoring systems", *IEEE Transactions on Smart Grid*, vol. 3, no. 3, pp. 1154–1169, 2012.

[171] D. M. Falcao, F. F. Wu, and L. Murphy, "Parallel and distributed state estimation", *IEEE Transactions on Power Systems*, vol. 10, no. 2, pp. 724–730, 1995.

[172] W. Jiang, V. Vittal, and G. T. Heydt, "Diakoptic state estimation using phasor measurement units", *IEEE Transactions on Power Systems*, vol. 23, no. 4, pp. 1580–1589, 2008.

[173] L. Zhao and A. Abur, "Multi area state estimation using synchronized phasor measurements", *IEEE Transactions on Power Systems*, vol. 20, no. 2, pp. 611–617, 2005.

[174] P. J. Van Laarhoven and E. H. Aarts, "Simulated annealing", *Simulated annealing: Theory and applications*, pp. 7–15, 1987.