8-2016

Wide-Area Measurement-Based Applications for Power System Monitoring and Dynamic Modeling

Jidong Chai
University of Tennessee, Knoxville, jchai@vols.utk.edu

Recommended Citation

Chai, Jidong, "Wide-Area Measurement-Based Applications for Power System Monitoring and Dynamic Modeling. " PhD diss., University of Tennessee, 2016.
https://trace.tennessee.edu/utk_graddiss/3900

To the Graduate Council:

I am submitting herewith a dissertation written by Jidong Chai entitled "Wide-Area Measurement-Based Applications for Power System Monitoring and Dynamic Modeling." I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Electrical Engineering.

Yilu Liu, Major Professor

We have read this dissertation and recommend its acceptance:

Fangxing Li, Husheng Li, Chien-fei Chen

Accepted for the Council:
Dixie L. Thompson

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

Wide-Area Measurement-Based Applications for Power System Monitoring and Dynamic Modeling

A Dissertation Presented for the

Doctor of Philosophy

Degree

The University of Tennessee, Knoxville

Jidong Chai

August 2016

To my beloved parents

ACKNOWLEDGEMENTS

I would like to express my deepest appreciation to my advisor Dr. Yilu Liu for her continuous support and patient guidance throughout my PhD studies. She has not only been an excellent mentor, but also a great friend. In addition, I want to thank the members of my committee, Dr. Fangxing Li, Dr. Husheng Li, and Dr. Chien-fei Chen for their insightful comments, continuous support and patience.

My sincere thanks go to my wonderful colleagues in the Power Information Technology Lab: Jiahui Guo, Yong Liu, Dao Zhou, Changgang Li, Yuming Liu, Zhiyong Yuan, Lingwei Zhan, Yin Lei, Ye Zhang, Gefei Kou, Hesen Liu, Zhuohong Pan, Penn Markham, Lang Chen, Yanzhu Ye, and Micah Till. This work would not be possible without their generous help. I would also like to thank the visiting scholars and students in the lab for the help and friendship they have given to me.

Last but not least, I want to express my appreciation to my family for their selfless love and support.

ABSTRACT

Due to the increasingly complex behavior exhibited by large-scale power systems with more uncertain renewables introduced to the grid, wide-area measurement system (WAMS) has been utilized to complement the traditional supervisory control and data acquisition (SCADA) system to improve operators' situational awareness. By providing wide-area GPS-time-synchronized measurements of grid status at high time-resolution, it is able to reveal power system dynamics which cannot be captured before and has become an essential tool to deal with current and future power grid challenges. According to the time requirements of different power system applications, the applications can be roughly divided into online applications (e.g., data visualization, fast disturbance and oscillation detection, and system response prediction and reduction) and offline applications (e.g., measurement-driven dynamic modeling and validation, post-event analysis, and statistical analysis of historical data).

In this dissertation, various wide-area measurement-based applications are presented. Firstly a pioneering WAMS deployed at the distribution level, the frequency monitoring network (FNET/GridEye) is introduced. For conventional large-scale power grid dynamic simulation, two major challenges are 1) accuracy of detailed dynamic models, and 2) computation burden for online dynamic assessment. To overcome the restrictions of the traditional approach, a measurement-based system response prediction tool using a Multivariate AutoRegressive (MAR) model is developed. It is followed by a measurement-based power system dynamic reduction tool using an autoregressive model to represent the external system. In addition, phasor measurement unit (PMU) data are employed to perform the generator dynamic model validation study. It utilizes both simulation data and measurement data to explore the potentials and limitations of the proposed approach. As an innovative application of using wide-area power system measurement, digital recordings could be authenticated by comparing the extracted frequency and phase angle from recordings with power system measurement database. It includes four research studies, i.e., oscillator error removal, ENF phenomenology, tampering detection, and frequency localization. Finally, several preliminary data analytics studies including inertia estimation and analysis, fault-induced delayed voltage recovery (FIDVR) detection, and statistical analysis of oscillation database, are presented.

TABLE OF CONTENTS

Chapter 1 Introduction to Frequency Monitoring Network (FNET/GridEye) 1

1.1 Frequency Monitoring Network 1

1.2 FNET/GridEye System Architecture 3

Chapter 2 Measurement-Based System Response Prediction 6

2.1 Introduction.. 6

2.2 Prediction Model with Autoregressive Model.. 7

2.2.1 Basics of System Identification 9

2.2.2 Multi-input Multi-output MAR Model 11

2.2.3 Model Training 13

2.2.4 Design Prediction Procedure.. 16

2.2.5 MAR Model Reduction Technique.. 18

2.3 Prediction Using Simulation Data 21

2.4 Prediction Using Real Measurement Data... 23

2.5 Prediction Time Delay Tests.. 24

2.6 MAR Model Reduction Tests.. 26

2.7 Conclusions.. 27

Chapter 3 Measurement-Based System Reduction 28

3.1 Introduction.. 28

3.2 The ARX Model 30

3.3 Model Accuracy Test with the Full Power System Model ... 31
3.3.1 23-bus System Model ... 31
3.3.2 The NPCC Model ... 33
3.3.3 The 25000-bus EI Model ... 33
3.4 Integrated Simulation in Power System Dynamic Simulation Software ... 39
3.4.1 23-bus System Model: Three Loads ... 41
3.4.2 23-bus System Model: One Load ... 41
3.4.3 The NPCC Model ... 44
3.4.4 The 25000-bus EI Model ... 46
3.5 Conclusions ... 46
Chapter 4 Measurement-Based Generator Dynamic Model Validation ... 49
4.1 Introduction ... 49
4.2 Validation Procedure and Simulation Method ... 50
4.3 Validation Using Simulation Data ... 52
4.3.1 Effects of PMU Reporting Rates and Event Selection ... 52
4.3.2 Parameter Analysis ... 54
4.3.3 Masking Effect ... 57
4.4 Validation Using Real Measurement Data ... 59
4.5 Conclusions ... 62
Chapter 5 Application of Power System Measurement for Digital Authentication ... 65

5.1 Introduction 65
5.2 Removing the Effects of Oscillator Error from ENF Comparisons 67
    5.2.1 Overview 67
    5.2.2 Methods 68
    5.2.3 Results 71
5.3 Phenomenology of Electric Network Frequency 75
    5.3.1 Overview 75
    5.3.2 Methods 76
    5.3.3 Results 81
5.4 Tampering Detection in Digital Recordings 91
    5.4.1 Overview 91
    5.4.2 Methods 92
    5.4.3 Results 98
5.5 Frequency Localization Using FDR Data 106
    5.5.1 Overview 106
    5.5.2 Methods 107
    5.5.3 Results 109
5.6 Conclusions 111
Chapter 6 Data Analytics Using Historical FNET/GridEye Database 113
6.1 EI Inertia Estimation and Analysis 113

6.1.1 System Inertia Estimation 114

6.1.2 Inertia Analysis Using SPSS 117

6.2 FIDVR Detection in FNET/GridEye Historical Database 117

6.3 Statistical Analysis of FNET/GridEye Oscillation Database 123

Chapter 7 Conclusions and Future Work 127

List of References 129

Vita 140

LIST OF TABLES

Table 3-1 Results of 23-bus system with different input signals 44

Table 3-2 Results of NPCC system with different input signals 46

Table 3-3 Results of EI system with different input signals 48

Table 5-1 Soundproof performance of the material 87

Table 5-2 Matching ratios for frequency localization study 110

LIST OF FIGURES

Figure 1-1 FDR deployment in North America 2

Figure 1-2 FDR worldwide deployment 2

Figure 1-3 FNET/GridEye system architecture 4

Figure 1-4 Second-generation FDR 4

Figure 2-1 Flowchart for prediction 17

Figure 2-2 One-line diagram of 23-bus model 22

Figure 2-3 Dynamics prediction of 23-bus system 22

Figure 2-4 Frequency and angle difference dynamics prediction of ISO-NE 24

Figure 2-5 Dynamics prediction with different reporting rates 25

Figure 2-6 Dynamics prediction of 23-bus model with reduced model 26

Figure 3-1 Model reduction using ARX 31

Figure 3-2 Model accuracy test on the 23-bus system 32

Figure 3-3 Accuracy test results of the 23-bus system 32

Figure 3-4 Model accuracy test on the NPCC system 34

Figure 3-5 Accuracy test results of the NPCC system 34

Figure 3-6 Model accuracy test on the EI system for case1 35

Figure 3-7 Accuracy test results of the EI system for case1 35

Figure 3-8 Model accuracy test on the EI system for case2 37

Figure 3-9 Accuracy test results of the EI system for case2 37

Figure 3-10 Model accuracy test on the EI system for case3 38

Figure 3-11 Accuracy test results of the EI system for case3.. 38

Figure 3-12 Model accuracy test on the EI system for case4 40

Figure 3-13 Accuracy test results of the EI system for case4.. 40

Figure 3-14 Integrated simulation in PSS/E 41

Figure 3-15 Integrated simulation of the 23-bus system 42

Figure 3-16 Integrated simulation results of the 23-bus system...................................... 42

Figure 3-17 Integrated simulation of the 23-bus system with input signals selection...... 43

Figure 3-18 Integrated simulation results of the 23-bus system with input signals (*v*1, *v*2) 43

Figure 3-19 Integrated simulation of the NPCC system with input signals selection 45

Figure 3-20 Integrated simulation results of the NPCC system with input signals (*v*1, *f*2) 45

Figure 3-21 Integrated simulation of the EI system with input signals selection............ 47

Figure 3-22 Integrated simulation results of the EI system with input signals (*v*1, *v*2) ... 47

Figure 4-1 Generator model validation using hybrid simulation..................................... 51

Figure 4-2 Simulation output with different PMU reporting rates for a line trip event.... 53

Figure 4-3 Simulation output with different PMU reporting rates for a nearby generation trip event 55

Figure 4-4 Parameter analysis results 56

Figure 4-5 Masking effect in PSS model (PSS2B).. 58

Figure 4-6 Validation of Unit#1 using generation trip#1 60

Figure 4-7 Validation of Unit#1 using generation trip#2 61

Figure 4-8 Validation of Unit#2 using generation trip#1 63

Figure 4-9 Validation of Unit#2 using generation trip#2 64

Figure 5-1 A plot of the ENF sequences computed using the same ENF extraction algorithm applied to two recordings made from the same grid sinusoidal voltage input signal 72

Figure 5-2 A plot of the two ENF sequences after frequency correction was performed by the Oscillator Error Correction/$d\sigma$ Metric algorithm at the end of its first iteration 73

Figure 5-3 A plot of the two ENF sequences after the time correction has been applied by the Oscillator Error Correction/$d\sigma$ Metric algorithm at the end of its first iteration 74

Figure 5-4 Cylinder made of the soundproofing material for the experiments 80

Figure 5-5 Frequency spectrum of an audio recording made by an iPhone 4 81

Figure 5-6 A comparison of ENF extracted from audio recording and voltage signal versus ENF measured by a Frequency Disturbance Recorder (FDR) 82

Figure 5-7 Test setup for effect of electric and magnetic fields on ENF coupling into battery-powered digital recordings 83

Figure 5-8 Frequency spectrum of recordings made by different recorders..................... 84

Figure 5-9 A comparison of ENF from different indoor recordings versus ENF measured by a Frequency Disturbance Recorder (FDR) 85

Figure 5-10 A comparison of ENF from different outdoor recordings versus ENF measured by a Frequency Disturbance Recorder (FDR) 85

Figure 5-11 Recording made when the fan is off 86

Figure 5-12 Recording made when the fan is on. 87

Figure 5-13 A comparison of ENF from audio recordings when fan is on versus ENF measured by a Frequency Disturbance Recorder (FDR) 87

Figure 5-14 A comparison of amplitude of the 120 Hz component from audio recordings made inside/outside soundproof materials 89

Figure 5-15 A comparison of ENF from a recording made near a computer versus ENF measured by a Frequency Disturbance Recorder (FDR) 89

Figure 5-16 An FFT analysis of a recording made with computer on............................. 90

Figure 5-17 An FFT analysis of a recording made with computer off 90

Figure 5-18 A comparison of phase angle estimated from a recording versus the phase angle of the grid as measured by a Frequency Disturbance Recorder (FDR) 93

Figure 5-19 Frequency change when 30s data of original recording is deleted with different window sizes 99

Figure 5-20 Phase change when 30s data of original recording is deleted with different window sizes.. 99

Figure 5-21 Sudden phase angle change due to a disturbance in the power grid 100

Figure 5-22 Frequency change when 30s data of original recording is deleted 102

Figure 5-23 Frequency change when 30s data of original recording is replaced 102

Figure 5-24 Frequency change with different deletion lengths of the original recording 103

Figure 5-25 Phase change with different deletion lengths of the original recording...... 103

Figure 5-26 Deletion length estimation using phase angle with the reference data 104

Figure 5-27 Frequency change with different lengths of replacement of the original recording 105

Figure 5-28 Phase change with different replacement lengths of the original recording 105

Figure 5-29 Approach of characteristics extraction.. 108

Figure 5-30 Approach of pattern recognition based on neural network 108

Figure 5-31 Location map of five states in EI 110

Figure 6-1 Inertia estimation results 116

Figure 6-2 Correlation coefficient of variables 118

Figure 6-3 Variables entered/removed 119

Figure 6-4 Model summary for Beta analysis 119

Figure 6-5 Model summary for inertia analysis 120

Figure 6-6 A typical FIDVR following a 230-kV transmission fault in Southern California 121

Figure 6-7 Flowchart of trigger design 122

Figure 6-8 Sample results of detected FIDVR in FNET/GridEye database 123

Figure 6-9 Probability distribution of important factors 124

Figure 6-10 Generation trip and oscillation data 125

Figure 6-11 Oscillation level distribution in EI 126

Chapter 1 Introduction to Frequency Monitoring Network (FNET/GridEye)

1.1 Frequency Monitoring Network

As a pioneering WAMS deployed at the distribution level, the frequency monitoring network FNET/GridEye has been continuously monitoring the grids for over ten years and various data visualization and analytics applications have been developed [1-5]. Unlike phasor measurement units (PMU) which require high manufacturing and installation costs [6-8], FNET/GridEye measures from normal single-phase electrical outlets with a simple procedure at a lower outlay. As a complete wide-area monitoring system, all the phasor measurements collected by frequency disturbance recorders (FDRs) are transmitted to the FNET/GridEye server hosted at the University of Tennessee, Knoxville (UTK), and Oak Ridge National Laboratory (ORNL) for cutting-edge research and development (R&D).

After over ten years of development and a number of improvements, the FNET/GridEye system is widely welcomed by the academia, industry as well as governments and has proved to be very stable and reliable. As of 2016, more than 200 FDR units have been deployed across the four North American Interconnections: Eastern Interconnections (EI), Western Electricity Coordinating Council system (WECC), Electric Reliability Council of Texas system (ERCOT), and Hydro Quebec area. Over 50 FDRs have been deployed worldwide, e.g., Europe, China, Egypt, etc. Figure **1-1** shows the current distribution of FDRs across the North American power grid and Figure **1-2** shows the

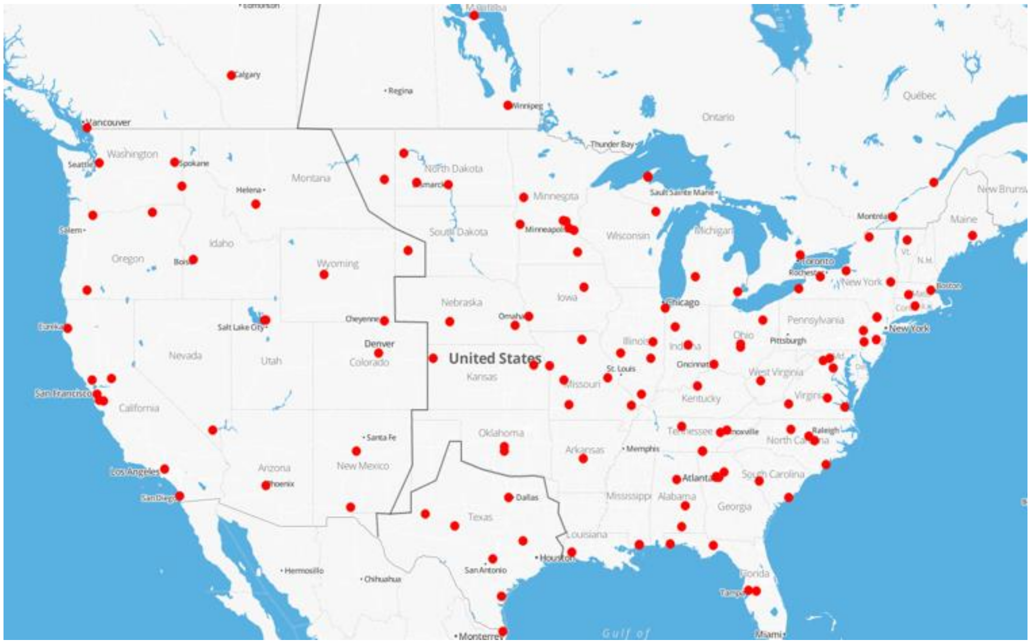


Figure 1-1 FDR deployment in North America

Figure 1-2 FDR worldwide deployment

world-wide FDR deployment map. Both global and local characteristics of frequency and phase angle variation can be monitored and analyzed based on these large volumes of data.

1.2 FNET/GridEye System Architecture

FNET/GridEye system consists of two major parts: sensors that are deployed across the power grids and data servers hosted by UTK and ORNL as shown in Figure **1-3**. The sensor (FDR) is an embedded microprocessor system with GPS time synchronization and Ethernet communications capability [2-3]. So far, three generations of FDRs have been developed to consistently pursue for higher measurement accuracy and better data quality. The current most-deployed FDR is Generation-II as shown in Figure **1-4**. Some of the new features of Generation-III include: 1) added power quality analysis function which can estimate harmonics composition and detect voltage sag and swell [9]; 2) higher steady-state phase angle and frequency measurement accuracy [10]. The error is less than 0.005° and 0.00006 Hz, respectively, compared with 0.01° and 0.0005 Hz for Generation-II; 3) improved dynamic-state measurement accuracy [11]; 4) use of atomic clock as the GPS timing backup. FDRs' holdover capability can be up to a day without losing accuracy.

The other part of FNET/GridEye system is the data center, where the measurements provided by FDRs are systematically managed, technically processed, and safely archived. The data center is a multi-layer data management system which is composed of the data server, application server, web server, and backup server, etc. Since power

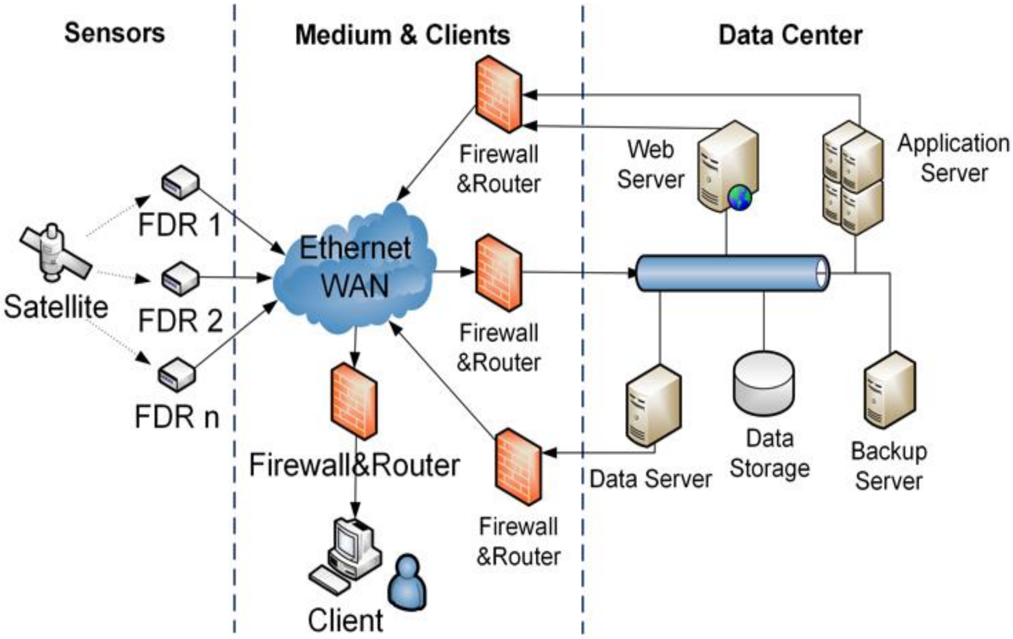


Figure 1-3 FNET/GridEye system architecture

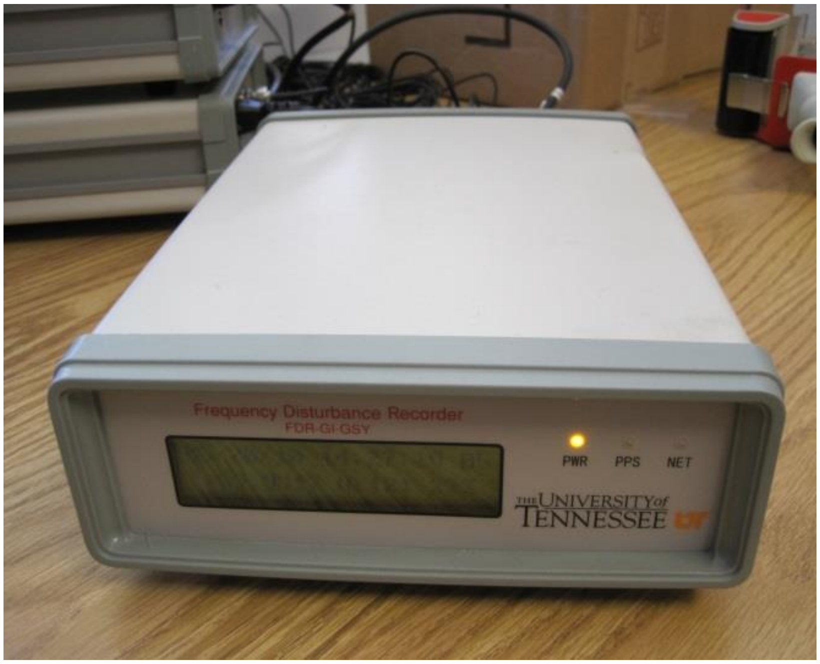


Figure 1-4 Second-generation FDR

system applications require various time requirements, the applications can be roughly divided into online applications (e.g., data visualization, fast disturbance detection and localization, oscillation and islanding detection) [12-17] and offline applications (e.g., measurement-driven dynamic modeling and validation, post-event analysis, various data analysis and forensic research) [18-23]. FNET/GridEye system and synchrophasor techniques will continue to help deal with future challenges in the grid due to larger penetration of renewables [24-28].

Chapter 2 Measurement-Based System Response Prediction

2.1 Introduction

This study is to develop a measurement-based system response-prediction tool using the transfer function approach in order to increase the overall speed of dynamic contingency screening. The advantage of this technique is faster results than that which can be obtained through the use of simulations for large systems. Another advantage is a potentially higher accuracy than existing model-based tools. This is because traditional circuit models are unable to include all possible details, especially as they relate to loads and may not be practical to update in real time. The assumption for this method to work effectively is the system under study is linear. This condition is met in the majority of situations where disturbances are small. For example, a trip of 2000 MW in the Eastern Interconnection (EI) is considered a small event except at the buses next to the trip.

There are two primary approaches for the study of power system dynamics: the time domain simulation approach and the measurement-based approach. Time-domain simulation [30], based on models of system equipment, is a sophisticated method for dynamics analysis, and there are commercial packages available, e.g., PSS/E, PSLF, DSATools, etc. It is usually used to perform "what-if" simulations to check the system behavior and control strategies. A reliable time-domain simulation should be performed with a very detailed model of the studied system. However, for dynamics prediction, an accurate model cannot be easily obtained. Thus, the power system model is not accurate enough, especially with respect to the loads. With constantly changing power system topology and operation status, it is not yet feasible to update power system models in real time; therefore, the result of model-based simulation is not reliable [31]. Meanwhile, it is time-consuming to simulate very large and complex systems. Even with an accurate model, model-based simulation can hardly be used for online dynamics prediction.

Benefiting from the accurate timing technology of the Global Positioning System (GPS), Phasor Measurement Units (PMUs) [32] and Frequency Disturbance Recorders (FDRs) [1] are developed to monitor system dynamics with high accuracy. Since PMU data are calculated directly from the measured voltage and current waveform, PMU measurements could reflect actual system conditions more authentically. Without knowledge of system parameters, some papers propose dynamics response prediction methods based on polynomial models, trigonometric functions, and reduced system models. The AutoRegressive (AR) model [33] and artificial intelligence based methods are also used for dynamics analysis and prediction [34].

Here we use a Multivariate AutoRegressive (MAR) model to predict dynamics from multiple measurement signals and formulate a systematical approach that can be used for online applications.

2.2 Prediction Model with Autoregressive Model

System Identification (SI) is an important technology that can be used to study a system with limited knowledge of its dynamic characteristics [35]. Basically, SI extracts system behaviors with measurement data from finite probes. For power systems, the probes can be PMUs or Frequency Disturbance Recorders (FDRs).

Since the disturbance information (the exact disturbance location, type, and severity) cannot be readily measured, a Multi-Input Multi-Output Multivariate AutoRegressive (MIMO MAR) model was used in this part to develop the measurement-based dynamics prediction tool.

In order to make the MAR model practical for a system with a large number of PMUs, a model reduction technique was used to reduce the model complexity while keeping the prediction accuracy.

The response-prediction tool developed in this research was achieved by following three steps, which are outlined below and are further explained in following sections.

1) Define prediction model structure. This is the basis of the prediction model and is done by choosing the Multivariate AutoRegressive (MAR) model as the basic model.

2) Construct the prediction model with measurement data. This is critical for developing the prediction model. By training the MAR model, a prediction model can be extracted to mimic power system dynamics.

3) Design prediction procedures with the developed prediction model. With the trained MAR model, proper prediction procedures can be designed to predict power system dynamics.

2.2.1 Basics of System Identification

A general system identification transfer function structure is shown:

$$A(q)y(t)=\frac{B(q)}{F(q)}u(t)+\frac{C(q)}{D(q)}e(t), \tag{2-1}$$

where

$$\begin{array}{l} A(q)=1+a_1q^{-1}+\ldots+a_{n_a}q^{-n_a}, B(q)=b_1q^{-1}+\ldots+b_{n_b}q^{-n_b}, C(q)=1+c_1q^{-1}+\ldots+c_{n_c}q^{-n_c} \\ D(q)=1+d_1q^{-1}+\ldots+d_{n_d}q^{-n_d}, F(q)=1+f_1q^{-1}+\ldots+f_{n_f}q^{-n_f} \end{array}.$$

Here, $y(t)$ is the output signal which is the observable (measurable) system response of interest, $u(t)$ is input signal which is a disturbance signal of the system or a stimuli manipulated by the observer, and $e(t)$ is a sequence of independent random variables. n_a, n_b, n_c, n_d, and n_f are the orders of each part. q^{-1} is a backward shift operator and $q^{-1}y(t)=y(t-h)$.

The interactions between input signals and the output signal is characterized by $A(q)$, $B(q)$, $C(q)$, $D(q)$ and $F(q)$. With different combinations of signals, different observation models can be developed. For example, if $B(q)=1$, $F(q)=1$, $C(q)=1$ and $D(q)=1$, (2-1) is a univariate AutoRegressive (AR) model. If $C(q)=1$, $D(q)=1$ and $F(q)=1$, (2-1) becomes an

AutoRegressive model with eXogenous inputs (ARX) which was used in the 2012 EPRI report [33].

For a physical system, disturbance signals are such physical quantities as voltage magnitude of a specific disturbance. They fall into two categories: those are directly measured, and those are unmeasured but observable through output signals. Measured disturbance signals are mathematically the same as stimuli signals expressed as $u(t)$. Other disturbance signals are usually modeled with random sequences $e(t)$. It should be noted here that the "disturbance signal" used for SI is different from power system "disturbances". To avoid confusion about the "disturbance signal" and power system "disturbance", "event" instead of "disturbance" is used in the following parts to describe the changing power system, e.g., load increase, and generation trip.

For power systems, dynamics response of frequency, voltage magnitude and phase angle is of key interest and thus can be chosen as output signals of the observation model (2-1). Though they can be directly measured with PMUs or FDRs, the disturbance signals of the event, i.e., input signals, are hard to measure due to the diversity of events. It is impractical to deploy measurement units at all possible locations. With only output signals, the univariate AR model is used in some research to model angle dynamics. However, the univariate AR model describes only the characteristics of each output signal individually. It is desirable to develop an observation model in order to describe the characteristics and interaction of all measured signals systematically.

2.2.2 *Multi-input Multi-output MAR Model*

In the 2012 EPRI report, the ARX model used for dynamics estimation is a multi-input single-output model. However, for dynamics prediction, a prediction model should be built to reflect the interaction between different measurement signals. For a synchronized power system, all physical quantities, e.g., voltage magnitude of each bus, bus frequency, generator outputs, etc, interact with each other via physical laws. In other words, the dynamics of a given measurement signal affects and is affected by other measurement signals. Intuitively, an observation model can be developed to model the interaction between those measurement signals by choosing one signal as the output signal and treating others as input signals.

With p interactional measurement signals $y_1(t)$, …, $y_p(t)$, if $y_i(t)$ is chosen as an output signal and all other signals are treated as input signals, a Multi-Input Single-Output (MISO) MAR model with $C(q)$=1, $D(q)$=1 and $F(q)$=1 can be developed as

$$A_i(q)\,y_i(t) = \sum_{j=1,\ j\neq i}^{p} B_{ij}(q)\,y_j(t) + e_i(t). \tag{2-2}$$

The MAR model is similar to ARX model except the input signals are, in fact, measured responses of other signals. Expanding (2-2) yields

$$y_i(t) + \sum_{k=1}^{n_{ai}} a_{ik}\,y_i(t-kh) = \sum_{j=1,\ j\neq i}^{p} \sum_{k=1}^{n_{bji}} b_{ijk}\,y_j(t-kh) + e_i(t), \tag{2-3}$$

where n_{ai} is the order of signal $y_i(t)$ and n_{bji} is the order of signal $y_j(t)$ when $y_i(t)$ is the output signal and $y_j(t)$ is the input signal. k is the index of time delay.

In the MAR model, the orders of each signal can be different. For simplicity, a uniform order n can be chosen for all signals. With this manipulation, (2-3) can be rewritten as

$$y_i(t) = -\sum_{k=1}^{n} a_{ik} y_i(t-kh) + \sum_{j=1,\ j\neq i}^{p} \sum_{k=1}^{n} b_{ijk} y_j(t-kh) + e_i(t). \tag{2-4}$$

or

$$y_i(t) = \sum_{j=1}^{p} \sum_{k=1}^{n} b_{ijk} y_j(t-kh) + e_i(t). \tag{2-5}$$

where $b_{iik} = -a_{ik}$.

Equation (2-5) can be further expressed in the vector form as

$$y_i(t) = \sum_{j=1}^{p} \mathbf{b}_{ij} \mathbf{y}_j(t) + e_i(t) \tag{2-6}$$

or

$$y_i(t) = \sum_{j=1}^{p} \mathbf{y}_j(t)^T \mathbf{b}_{ij}{}^T + e_i(t), \tag{2-7}$$

where

$$\mathbf{b}_{ij} = \left[b_{ij1}, \ldots, b_{ijn}\right], \mathbf{y}_j(t) = \left[y_j(t-h), \ldots, y_j(t-nh)\right]^T.$$

and superscript ^{T} indicates transpose.

In the MAR model, there is no difference between the output signal $y_i(t)$ and other signals in respect to mathematical status. So, MAR models with other signals as the output signal can be developed as follows

$$\begin{cases} y_1(t) = \sum_{j=1}^{p} \mathbf{b}_{1j}\mathbf{y}_j(t) + e_1(t) \\ \quad\quad \cdots \\ y_p(t) = \sum_{j=1}^{p} \mathbf{b}_{pj}\mathbf{y}_j(t) + e_p(t) \end{cases} . \tag{2-8}$$

Let $\mathbf{y}(t) = \left[y_1(t),...,y_p(t)\right]^T$, a Multi-Input Multi-Output (MIMO) MAR model can be written as

$$\mathbf{y}(t) = \begin{bmatrix} \mathbf{b}_{11} & \cdots & \mathbf{b}_{1p} \\ \cdots & \cdots & \cdots \\ \mathbf{b}_{p1} & \cdots & \mathbf{b}_{pp} \end{bmatrix} \begin{bmatrix} \mathbf{y}_1(t) \\ \cdots \\ \mathbf{y}_p(t) \end{bmatrix} + \begin{bmatrix} e_1(t) \\ \cdots \\ e_p(t) \end{bmatrix} . \tag{2-9}$$

Since the MISO MAR model (2-7) is part of the MIMO MAR model (2-9), (2-7) can be treated as a sub-model of the MIMO MAR model (2-9) and is denoted as "sub-model *i*".

2.2.3 Model Training

The MIMO MAR model (2-9) is comprised of p sub-models, and can be developed by training each sub-model separately. For sub-model *i* of (2-7), m-n equations can be written as (2-10) with an event of m measurement data points,

$$\mathbf{y}(t) = \begin{bmatrix} \mathbf{b}_{11} & \cdots & \mathbf{b}_{1p} \\ \cdots & \cdots & \cdots \\ \mathbf{b}_{p1} & \cdots & \mathbf{b}_{pp} \end{bmatrix} \begin{bmatrix} \mathbf{y}_1(t) \\ \cdots \\ \mathbf{y}_p(t) \end{bmatrix} \tag{2-10}$$

It can be re-written as

$$\mathbf{Y}_i = \mathbf{A}_i\mathbf{B}_i + \mathbf{E}_i , \tag{2-11}$$

where

$$\mathbf{A}_i = \begin{bmatrix} \mathbf{y}_1\left(\left(n+1\right)h\right)^T & \ldots & \mathbf{y}_p\left(\left(n+1\right)h\right)^T \\ \ldots & \ldots & \ldots \\ \mathbf{y}_1\left(mh\right)^T & \ldots & \mathbf{y}_p\left(mh\right)^T \end{bmatrix}, \quad \mathbf{Y}_i = \left[y_i\left(\left(n+1\right)h\right) \ \ \ldots \ \ y_i\left(mh\right)\right]^T .$$

$$\mathbf{B}_i = \left[\mathbf{b}_{i1} \ \ \ldots \ \ \mathbf{b}_{ip} \right]^T, \quad \mathbf{E}_i = \left[e_i\left(\left(n+1\right)h\right) \ \ \ldots \ \ e_i\left(mh\right)\right]^T$$

To get the best MAR model using (2-11), the error part E_i should be minimized. Parameter B_i can be estimated with the least-squares error estimator to minimize the error part E_i,

$$\mathbf{B}_i = \left(\mathbf{A}_i^T \mathbf{A}_i\right)^{-1} \mathbf{A}_i^T \mathbf{Y}_i . \tag{2-12}$$

It should be reminded that since the MAR model used is a linear model, B_i can capture the dynamic behavior of the system that is excited in the event for the model construction.

When the system reaches a new steady state for $t \rightarrow \infty$, the MISO MAR model (2-9) can be rewritten as

$$\mathbf{y}\left(\infty\right) = \mathbf{C}\mathbf{y}\left(\infty\right) \tag{2-13}$$

or

$$\left(\mathbf{I} - \mathbf{C}\right)\mathbf{y}\left(\infty\right) = 0 , \tag{2-14}$$

where $c_{ij} = \sum_{k=1}^{n} b_{ijk}$, and **I** is the $p \times p$ identity matrix.

It is clear that, the new steady state must be $y_i(\infty)=0$ for $i=1,\ldots,p$. However, it is seldom for directly measured signals to approach 0 when a new steady state is reached. The steady state value is mostly determined by the trend part of a time series which is a slow, gradual change in some properties of the series over the whole time window. To fulfill the condition at $t \rightarrow \infty$, the trend part of measured signals must be removed.

There are different detrending methods such as first differencing, curve fitting, and digital fitting. The first differencing method is a kind of high pass filter, and is good enough in most cases. In this study the first differencing method is used to remove signal trends and is defined as

$$y_i(t) = x_i(t) - x_i(t-h), \tag{2-15}$$

where $x_i(t)$ is the measurement signal, and $y_i(t)$ is the detrended signal.

For dynamics study, we focus on the dynamics of the original signal $x_i(t)$ instead of the detrended signal $y_i(t)$. To recover the original signal from the detrended signal, an inverse form of first differencing can be derived as

$$x_i(t) = y_i(t) + x_i(t-h). \tag{2-16}$$

It is clear that the recovery equation (2-16) depends on the current detrended signal $y_i(t)$ and the historical data $x_i(t-h)$ which is already known at time t. So the detrending method used can be easily implemented for field applications.

2.2.4 Design Prediction Procedure

In (2-3), the time delay index *k* starting from 1 and $y_i(t)$ is totally determined by historical data, i.e., data at a time prior to t. So, equations (2-3)-(2-9) give the one-step prediction of the studied dynamic system.

In equation (2-9), the random part $e(t)$ corresponds to the unmeasured disturbance signals and is hard to model. For simplicity, the $e_i(t)$ part is neglected and the MAR model becomes

$$\mathbf{y}(t)=\begin{bmatrix}\mathbf{b}_{11} & \cdots & \mathbf{b}_{1p}\\ \cdots & \cdots & \cdots\\ \mathbf{b}_{p1} & \cdots & \mathbf{b}_{pp}\end{bmatrix}\begin{bmatrix}\mathbf{y}_1(t)\\ \cdots\\ \mathbf{y}_p(t)\end{bmatrix} \qquad (2\text{-}17)$$

or

$$\mathbf{y}(t)=\mathbf{BY}(t), \qquad (2\text{-}18)$$

where

$$\mathbf{B}=\begin{bmatrix}\mathbf{b}_{11} & \cdots & \mathbf{b}_{1p}\\ \cdots & \cdots & \cdots\\ \mathbf{b}_{p1} & \cdots & \mathbf{b}_{pp}\end{bmatrix}$$

and

$$\mathbf{Y}(t)=\left[\mathbf{y}_1(t)^T \quad \cdots \quad \mathbf{y}_p(t)^T\right]^T .$$

Once the MAR model is constructed, dynamics can be predicted recursively as in the flowchart in Figure **2-1**, where t_{max} is the length of study time window. Since the $e_i(t)$ part is neglected, the prediction error may accumulate as t advances.

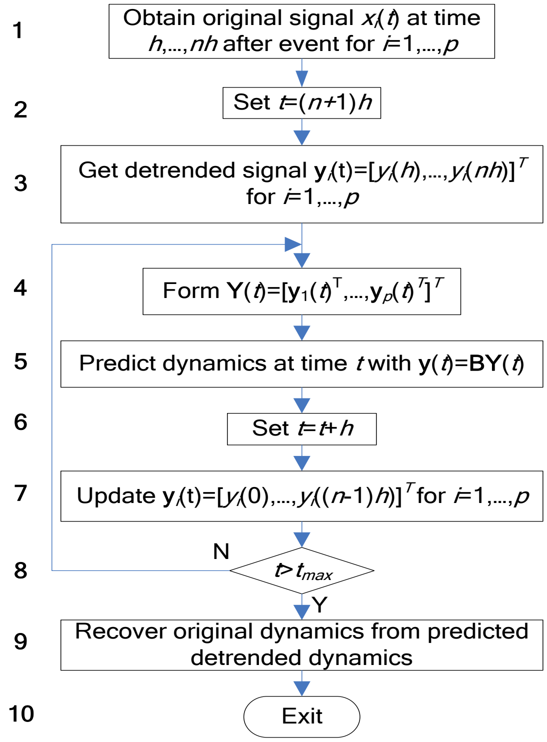


Figure 2-1 Flowchart for prediction

It can be seen from Figure **2-1** that the MAR model requires n data points for prediction. The first n points after an event (as in step 1) contain important information of the event, and trigger the following dynamics. There should be only one event or operation during the n points. Events with more than one operation, e.g., line faults which require fault operation and clearing operation, cannot be modeled with the MAR model and should be addressed by using other methods.

It should be noted that the MAR model used in this study is a linear model. It can be used for dynamics study of events that are relatively small compared with the system capacity, e.g., load increase or decrease, and generation trip. Line faults that will cause the topology change are out of the scope of the linear model.

2.2.5 MAR Model Reduction Technique

Unlike the univariate AR model, the MIMO MAR model describes the interaction between different measurements. Theoretically, all quantities of power system are affected by other quantities. More measurement signals would improve the prediction accuracy while increasing the model complexity.

The MISO MAR model developed with (2-4) includes n parameters for each input signal y_j. It is called full MAR model with np unknowns. With more and more PMUs deployed, the number of unknowns increases linearly with measurement number p. Though the MAR model is linear, too many unknowns may lead to a huge computation burden and model complexity, and thus prevent the full MAR model from being implemented for online applications.

The increase of an unknown number leads to a possible problem of under-determinacy. There are $m-n$ equations (constraints) but np unknowns in (2-4) for a MISO MAR model. If $m-n<np$ (or $m-n\geq np$ and A_i is singular), (2-4) is underdetermined and mathematically unsolvable. To overcome this difficulty, either the data point number should be increased or the unknown number should be reduced. However, m cannot be increased arbitrarily. In most cases, disturbance is damped in 20s or shorter. The time window length for

model training is usually between 10s and 20s. Time windows that are too long may include data points of a new steady state and make A_i singular. Under a specific PMU reporting rate, data length m is limited. So m cannot be increased unless the PMU is reporting at a higher rate. It is more practical to reduce the number of unknowns than to increase the measurement data length.

To summarize, the number of unknowns of the full MAR model should be reduced to make the MAR model less complex and easy to train.

Since the number of unknowns is *np*, there are two direct ways to reduce model complexity of a MISO MAR model. One way is to reduce model order n. However, n is a key parameter of the MAR model, and dynamics prediction accuracy is greatly affected by model order. In general, higher order usually yields better training accuracy and prediction accuracy. To keep high prediction accuracy, model order cannot be dramatically reduced.

The other way is to reduce the input signal number. The relationships between the output signal y_i and different input signals y_j are different. Those input signals more related to the output signal contribute more to the output signal. The Zero-delay Correlation Coefficient (CC) is a good choice to measure the contribution of each input signal. It is defined as (2-19).

$$r_{ij} = \left| \frac{\sum_{l=1}^{m} y_i(lh) y_j(lh)}{\sqrt{\sum_{l=1}^{m} y_i(lh)^2} \sqrt{\sum_{l=1}^{m} y_j(lh)^2}} \right| . \tag{2-19}$$

Obviously, those input signals with high zero-delay CCs can be selected as effective input signals and other signals are completely removed from (2-4). However, the main problem of this method is that it does not differentiate the input signals and specific input terms of each input signal. For each input signal y_j, there are n input terms of output signal y_i in (2-4), i.e., $b_{ij1}y_j(t-h)$, $b_{ij2}y_j(t-2h)$, …, and $b_{ijn}y_j(t-nh)$. Those input terms are the delayed time series of original signal y_j. If the zero-delay CC r_{ij} is high, it is not necessary that all the n delayed input terms of y_j contribute significantly to output signal y_i, and vice versa. It is preferable to study the contribution of each input term specifically, instead of the overall contribution of each input signal.

To overcome the drawbacks of the zero-delay CC, the delayed CC of each input term is used in this study to find a better combination of input terms. Take a k-delay input term of signal y_j, i.e., $y_j(t-kh)$ of (2-7), for example, the k-delayed CC between the input signal y_j and output signal y_i is

$$r_{ijk} = \left| \frac{\sum_{l=k+1}^{m} y_i(lh)\, y_j\left((l-k)h\right)}{\sqrt{\sum_{l=k+1}^{m} y_i(lh)^2}\sqrt{\sum_{l=1}^{m} y_j\left((l-k)h\right)^2}} \right| . \tag{2-20}$$

The delayed CC examines the correlation of each delayed input term with the output signal. Comparing (2-19) with (2-20), it can be found that the zero-delay CC is symmetric, i.e., $r_{ij} = r_{ji}$, however, the delayed CC is asymmetric, i.e., $r_{ijk} \neq r_{jik}$. This leads to the main difference between the input selections with the two CCs. With zero-delay CC, if signal y_j is highly correlated with signal y_i, signal y_j is the input signal of signal y_i, and vice versa. For delayed CC, if *k*-delay input term of signal y_j is highly correlated with signal y_i, $y_j(t-kh)$ can be selected as the input term of signal y_i. However, if the *k*-delay input term of signal y_i is not highly correlated with signal y_j, it does not have to be selected as an input term of y_j. So, **B** of (2-18) does not necessarily have symmetric structure when delayed CC is used for input term selection.

With the delayed CC, it is not necessary to find the optimal model order *n*. With a higher model *n*, all input terms can be examined with the delayed CC and only those with the delayed CC higher than a user-defined threshold are selected for developing the sub-model of y_i. The delayed CC lies in the range of [0, 1] and a threshold higher than 0.9 is usually suggested for input term selection.

2.3 Prediction Using Simulation Data

The 23-bus system was chosen as the test system for the validation of the prediction model. The one-line diagram of the 23-bus model is shown in Figure **2-2**. The simulation time step was set as a half cycle (1/120s). Frequency, voltage and phase angle of all 17 PQ buses were monitored. In this section, a full MAR model of those 51 signals was developed with 15s-long data with an under 1% load increase event and a prediction was made for a 5% load increase event. The order of each signal was 13. The predicted voltage dynamics of bus 151 and 203 predicted frequency dynamics of bus 151 and 203, and the phase angle difference between bus 151 and 203 are shown in Figure **2-3**. The diamonds indicate where the prediction starts from. Data before the diamond are historical data and data after the diamond are predicted.

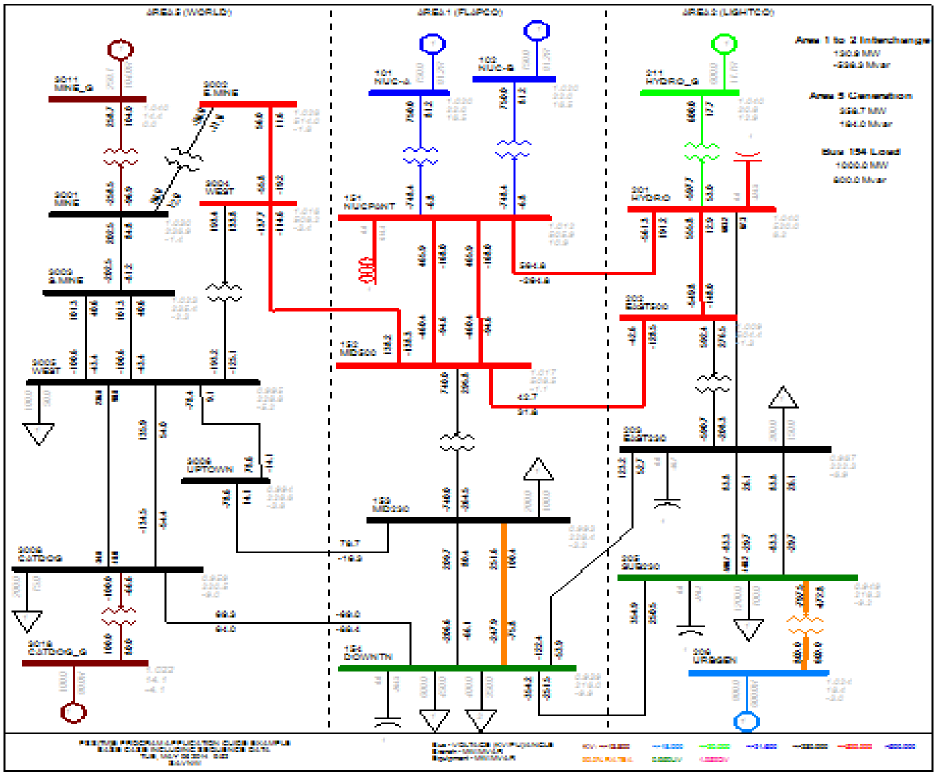
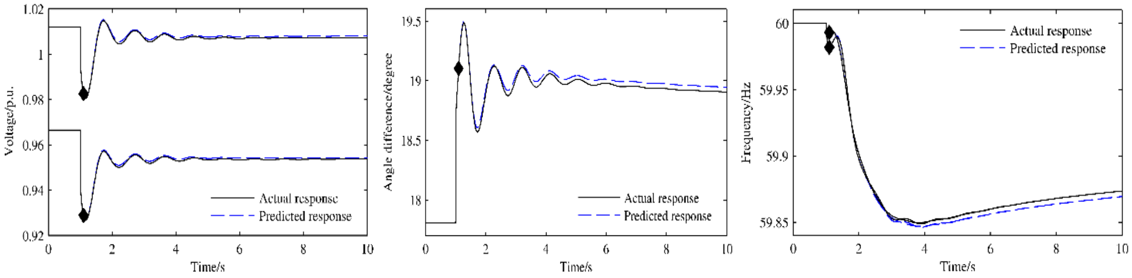


Figure 2-2 One-line diagram of 23-bus model



(a) Voltage (b) Angle (c) Frequency

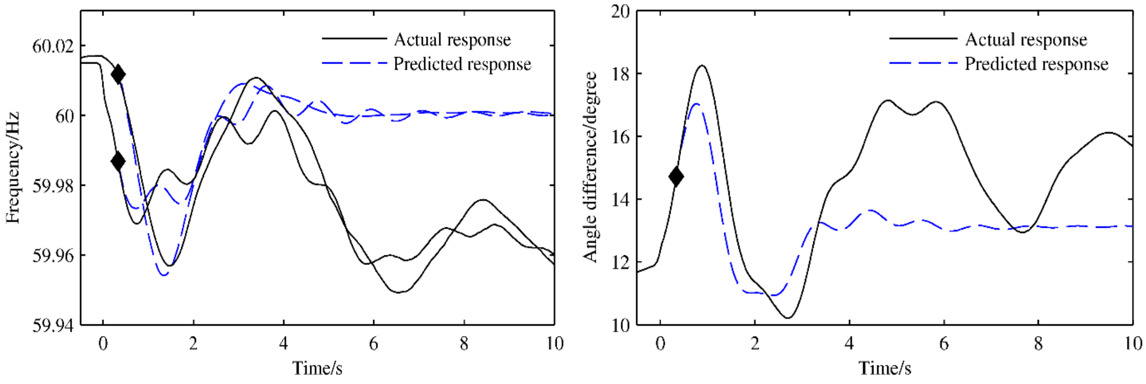
Figure 2-3 Dynamics prediction of 23-bus system

From Figure **2-3**(a) and Figure **2-3**(b), it can be seen that the predicted voltage and angle response are very accurate compared with actual responses. The frequency prediction in Figure **2-3**(c) is not as good since the frequency dynamics is monotonous. It can also be found from Figure **2-3** that predicted dynamics is more accurate in the first few tens of cycles than in latter cycles, which is mainly caused by the neglect of $e_i(t)$.

2.4 Prediction Using Real Measurement Data

With a generation trip at about 22:52:23 on July 1, 2012 near ISO New England of US, frequency, voltage and angle dynamics of 14 PMUs were measured with the reporting rate of 30 Hz. Since voltage dynamics is local and quite noisy, a MAR model was built with frequency and angle dynamics with 20s-long dynamic data after the event happened. The order of each signal was eight and all signals were filtered with a low pass filter (3T). MAR models were tested with another generation trip event at about 10:25:07 on March 21, 2013 near ISO New England. The predicted frequency of two stations is shown in Figure **2-4**(a). The angle difference between the two stations is shown in Figure **2-4**(b). The abscissa of the two figures is relative time with event time.

From Figure **2-4**, it can be found that the prediction result with field measurement data is not as good as model-based simulations. Only a prediction of the first 3-4 seconds is acceptable and there are mainly three reasons for the prediction inaccuracy. The first reason is that we only have dynamic data from 14 PMUs which are not enough for the model training of a large scale power system. It is desirable to have more PMUs deployed evenly in the power grid. The second reason is that the testing event occurred at a different location from the training event which makes the trained MAR model hard to predict the dynamics of the testing event. With more events recorded by PMUs, it would be possible to develop a library of MAR models for more accurate predictions. The third reason is that raw PMU data are not as clean as the simulation data. Though the noisy PMU data were filtered with a low pass filter in this work, it was still not ideal for dynamics prediction. The noisy PMU data and neglect of $e_i(t)$ affected the performance of the prediction method. More work needs to be done with field measurement data including the data filtering technology.



(a) Frequency dynamics

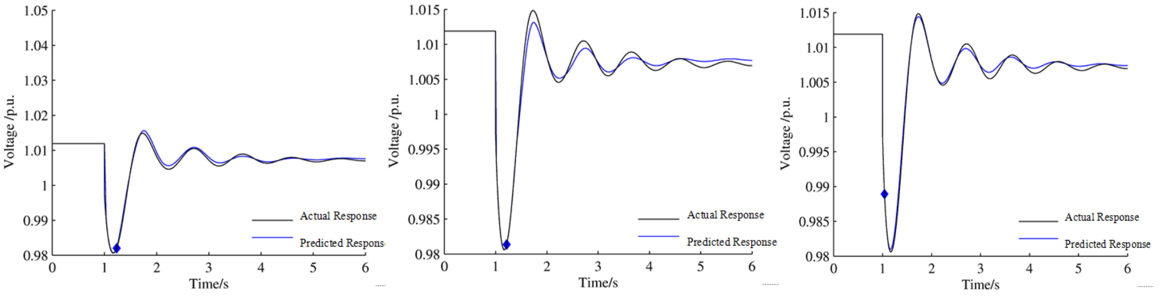
(b) Angle dynamics

Figure 2-4 Frequency and angle difference dynamics prediction of ISO-NE

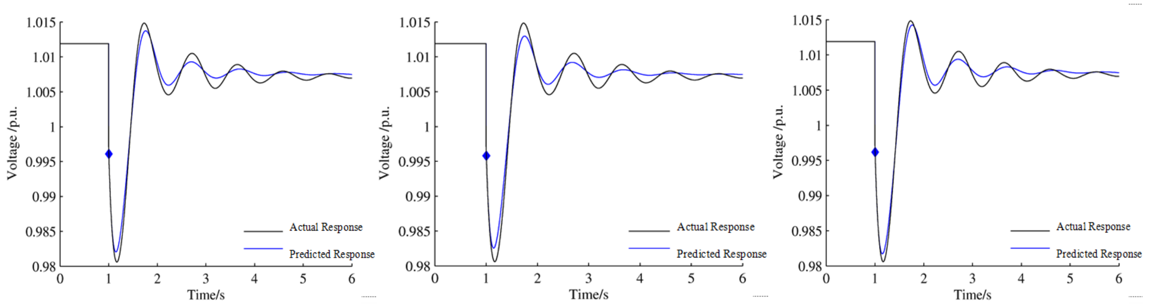
2.5 Prediction Time Delay Tests

As discussed before, the measurement-based system response prediction tool needs n (model order) data points to initiate the prediction procedure. In other words, there is a fixed time delay (equivalent to n data points) for prediction. With different measurement data reporting rates and different model orders, the prediction time delays will change.

In this part of the study, the prediction time delay was tested with simulation data of the 23-bus model. Simulation data were down sampled to mimic different PMUs. Figure **2-5** shows the influence of data reporting rates on the prediction time delay.



(a) 30 Hz (b) 60 Hz (c) 600 Hz



(d) 3k Hz (e) 6 kHz (f) 12 kHz

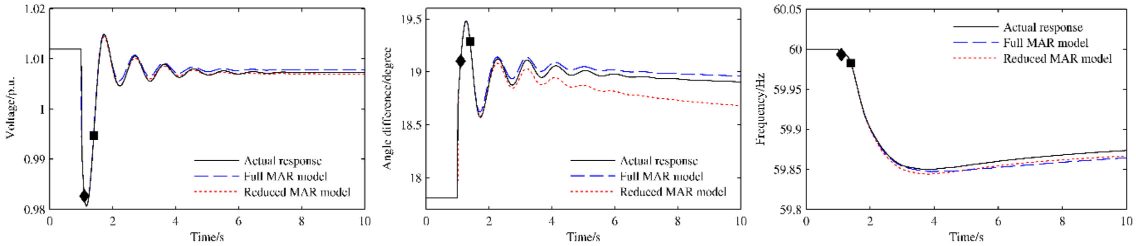
Figure 2-5 Dynamics prediction with different reporting rates

For the 30 Hz case in (a), with the model order of seven, the prediction delay was 14 cycles. With the model order of 13, the prediction delay of the 60 Hz case in (b) was 13 cycles, almost the same as the 30 Hz case. The two cases show that adequate measurement requires data with prediction delay of about 13 to 14 cycles to predict dynamics accurately if the data reporting rates are low.

When the data reporting rate increased to 600 Hz as shown in (c), the prediction delay reduced dramatically to 2.4 cycles with the model order of 24. This represents a turning point of the prediction delay. Cases with data reporting rates of 3 kHz, 6 kHz, and 12 kHz had similar prediction time delays at 1/4 to 1/3 cycles. In other words, further increasing the data reporting rate will not reduce the prediction delay.

2.6 MAR Model Reduction Tests

Suppose the data reporting rate of the 23-bus model is 30 Hz, which can be achieved by down sampling the 120 Hz simulation data. With 15s-long dynamic data, there are 450 data points. If a uniform order of 13 is selected, there would be 663 (13x51) unknowns and it would be impossible to train a full MAR model. By selecting input terms with delayed CC higher than 0.9, a reduced MAR model can be built. Prediction of frequency and voltage of bus 151 and the angle difference between bus 151 and bus 203 are shown in Figure **2-6**. The diamond shows where the prediction of the full MAR model starts when the data rate is 120 Hz, and the square shows where the prediction of the reduced MAR model starts when the data rate is 30 Hz.



(a) Voltage (b) Angle (c) Frequency

Figure 2-6 Dynamics prediction of 23-bus model with reduced model

In the reduced MAR model, there are only 44 unknowns for the sub-model of signal 1 (frequency of bus 151). The sub-model with the most unknowns is signal 15 (phase angle of bus 201), which was constituted by 190 unknowns. The number of unknowns in the reduced MAR model was 5030 (about 15% of full MAR model).

The reduced model still provided good prediction results. When the data reporting rate was 60 Hz, there were enough data points for training the full MAR model with uniform order of 13.

2.7 Conclusions

With increasing numbers of PMUs deployed, massive measurement data give an alternative approach to study power system dynamics without the traditional circuit models. The system response-prediction tool developed in this study was purely measurement-based and the MAR model can be easily updated online with typical events. Examples show that the accuracy of prediction is high for simulated cases. A real measurement case is not as accurate as the simulation case, but it still provides good prediction results for the first few seconds. To make the MAR model practical for a system with a large number of PMUs, the delayed correlation coefficient was defined to measure the contribution of each input term with delays. This study provided a good guideline for the input selection and model reduction. Simulation results show that the reduced model can keep the prediction accuracy while reducing the model complexity.

Chapter 3 Measurement-Based System Reduction

3.1 Introduction

With the infinite details especially in loads and constantly changing power system topology and operation status, it is not yet feasible to develop highly accurate large system models and update full power system dynamic models in real time. In addition, usually only the dynamics of a local system need to be studied while external systems can be reduced. Thus, dynamic equivalencing is one solution for the increasing size and complexity of power grids.

A number of methods for power system dynamic reduction have been developed. The coherency-based methods [36-37] identify coherent generators and aggregate them into single or multiple equivalent generators. However, it is very difficult to find the exact coherent groups and the accuracy of aggregation and network reduction depends on the detailed model parameters. In addition, it is time-consuming and thus not appropriate for online applications. The selective modal analysis-based methods [38-39] simplify the system by using dominant modes. But they require the computation of eigenvalues of the system matrix which is time-consuming and it is also very difficult to determine truly dominant modes of the system. To overcome the restrictions of traditional approaches, measurement-based system identification methods [40-41] were used to reduce the external systems. Most of system identification-based approaches use nonlinear models to represent the external systems considering the power system is nonlinear in general. But the nonlinear models may either be too complex or suffer convergence issues.

Based on some of our previous studies [42-44], this study is to treat the subsystem which needs to be reduced as a black-box and derive the equivalent model based on measured data using a linear autoregressive model. The derived black-box equivalents were integrated with the study system to perform the dynamic simulation. This approach may improve the modeling accuracy of the reduced system compared with the traditional approach and increase the speed of dynamic simulation through system reduction.

Results were achieved through the three methodological devices or techniques. They will be further explained in this chapter.

1) System identification with measurement data. The autoregressive with exogenous input (ARX) model was chosen and an equivalent model was derived by training the ARX model.
2) The model accuracy test with the full power system model. The derived ARX model was used to predict the dynamics of other events similar with the training event.
3) Integrated simulation in power system dynamic simulation software. The developed equivalent model was integrated into the study system in PSS/E for dynamic simulation.

3.2 The ARX Model

The mathematical structure expression of ARX model is given

$$A_k(z)\hat{y}_k(t) = \sum_{j=1}^{n_j} B_{jk}(z)u_j(t) + e(t) \tag{3-1}$$

where t is the sampled data number, $e(t)$ is the system noise, u_j and $\hat{y}_k$ are the model's j-th input and k-th output, respectively. $A_k(z)$ and $B_{jk}(z)$ are the ARX nominator and denominator polynomials respectively, which are given by:

$$A_k(z) = 1 + a_{k1}z^{-1} + \cdots + a_{kn_{a_k}} z^{-n_{a_k}} \tag{3-2}$$

$$B_{jk}(z) = b_{jk0} + b_{jk1}z^{-1} + \cdots + b_{jk(n_{b_{jk}}-1)} z^{-(n_{b_{jk}}-1)} \tag{3-3}$$

where n_{ak} and n_{bjk} are the orders of the model, and j and k are the numbers of inputs and outputs, respectively. The model parameters of a multi-variable ARX model can be estimated by a linear least square technique. The least-squares estimation problem is solved by using QR factorization to optimize the ARX model parameters. The least-squares loss function is defined as:

$$V_{LS} = \sum_{t=n_s+1}^{N} \varepsilon_{ARX}(t)^2 \tag{3-4}$$

where the equation error criterion $\varepsilon_{ARX}(t)$ is described by:

$$\varepsilon_{ARX}(t) = A_k(z)y_k(t) - \sum_{j=1}^{n_j} B_{jk}(z)u_j(t) \tag{3-5}$$

3.3 Model Accuracy Test with the Full Power System Model

The general idea is shown in Figure **3-1**. Area 1 is the study system and Area 2 represents the chosen area to be reduced. A tie line is between Area 1 and Area 2. For example, the voltage (*v*) and frequency (*f*) of the border bus are used as inputs; active power (*P*) and reactive power (*Q*) of the tie line are outputs. After training the ARX model with those inputs and outputs, a function $(P, Q) = g\ (v, f)$ was obtained. The derived model was then used to predict results of other test events and its accuracy was evaluated by actual full model responses.

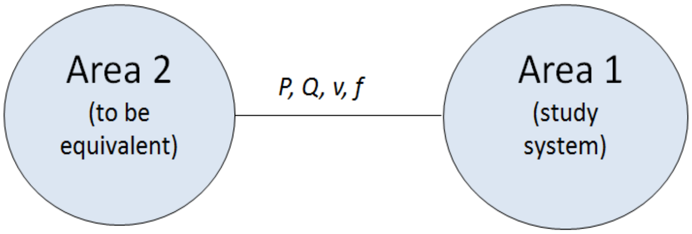


Figure 3-1 Model reduction using ARX

3.3.1 23-bus System Model

The study system and external system were connected by three tie lines as shown in Figure **3-2**. A generation trip at bus 206 was used for training, and generator trip at bus 211 was used for testing. The inputs were voltage and frequency of the border buses in the study system. The ARX model order is six. The active power deviation of each tie line is shown in Figure **3-3**. It shows the predicted response is very close to the actual response.

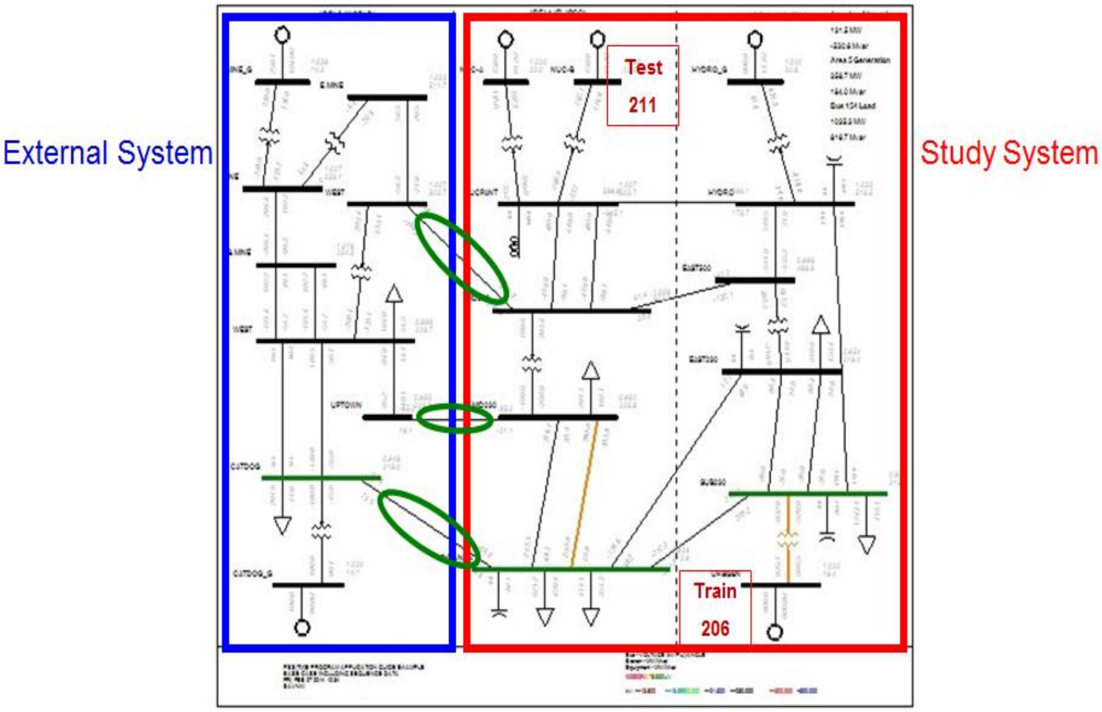
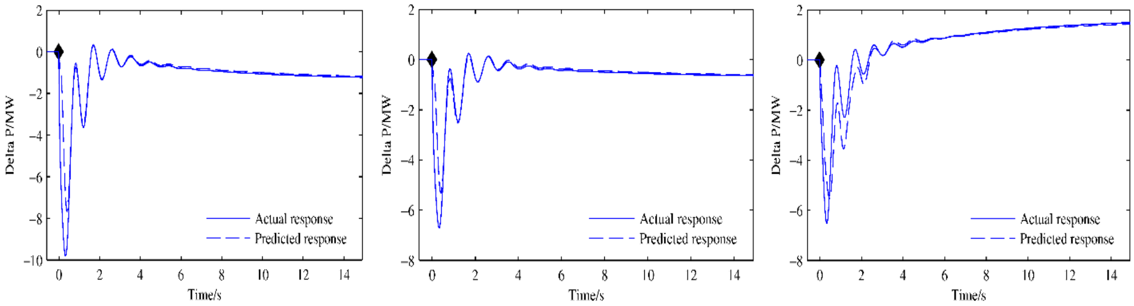


Figure 3-2 Model accuracy test on the 23-bus system



(a) Tie Line1

(b) Tie Line2

(c) Tie Line3

Figure 3-3 Accuracy test results of the 23-bus system

3.3.2 The NPCC Model

The study system and the external system are connected by two tie lines as shown in Figure 3-4. A generation trip at bus 22 is used for training and generation trips at other buses are used for testing. The inputs are voltage and frequency of border buses in the study system. The ARX model order is zero (static model). The active power and reactive power of two tie lines are shown in Figure 3-5. Both predicted active power and reactive power are very accurate when compared with the actual response.

3.3.3 The 25000-bus EI Model

Case1: Reduced Florida Area

The Florida area was reduced while other areas were retained as shown in Figure 3-6. A generation trip in the PJM area was used for training and generation trips at other locations used for testing. The inputs were voltage and frequency of border buses in the study system. The model order was 0 (static model), while the active power and the reactive power of one selected tie line are shown in Figure 3-7. Simulation results from DYNRED [45] (a commercial network reduction software) are also included in Figure 3-7. It can be seen that the ARX estimated response was very close to the actual response and is more accurate than DYNRED in this case. Note that results from DYNRED depended on how to reduce the system. In each of the following cases, a reasonable and straightforward method was employed to reduce the system using DYNRED and the results are shown together with the ARX model. Several tests were performed, but only

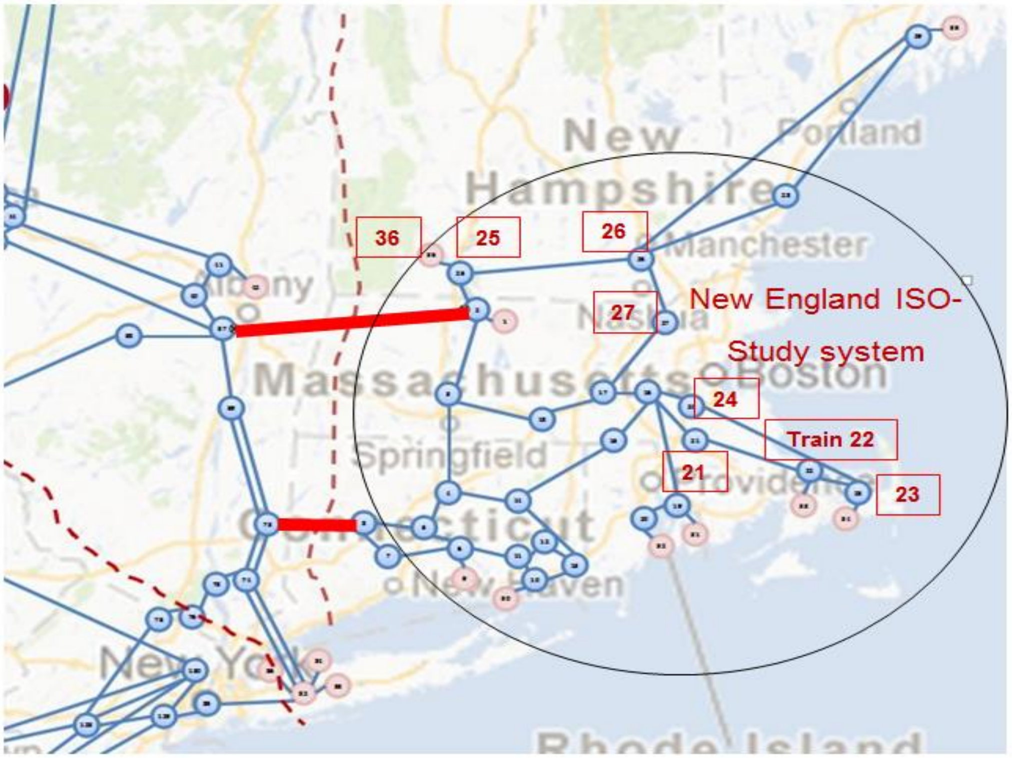
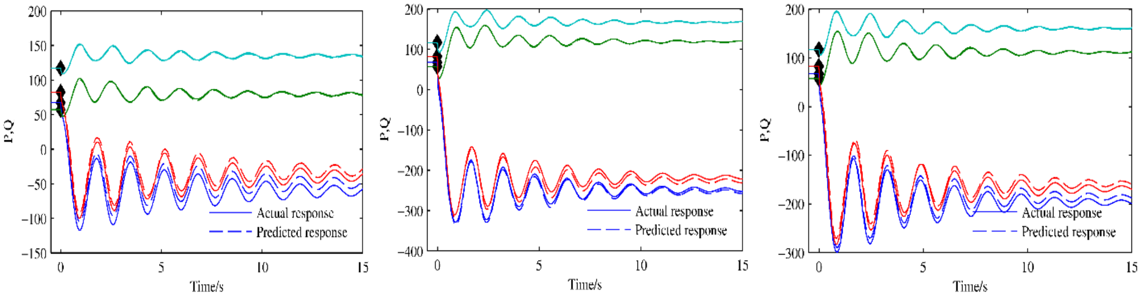


Figure 3-4 Model accuracy test on the NPCC system



(a) Gen. trip@23 (b) Gen. trip@24 (c) Gen. trip@25

Figure 3-5 Accuracy test results of the NPCC system

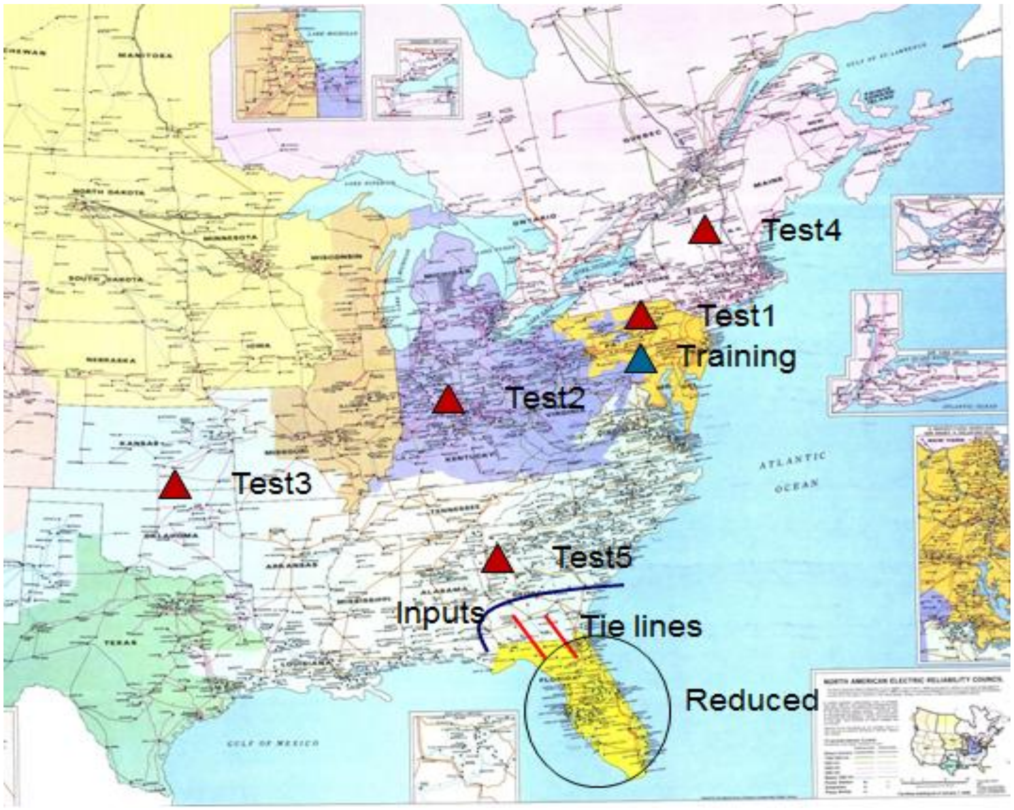
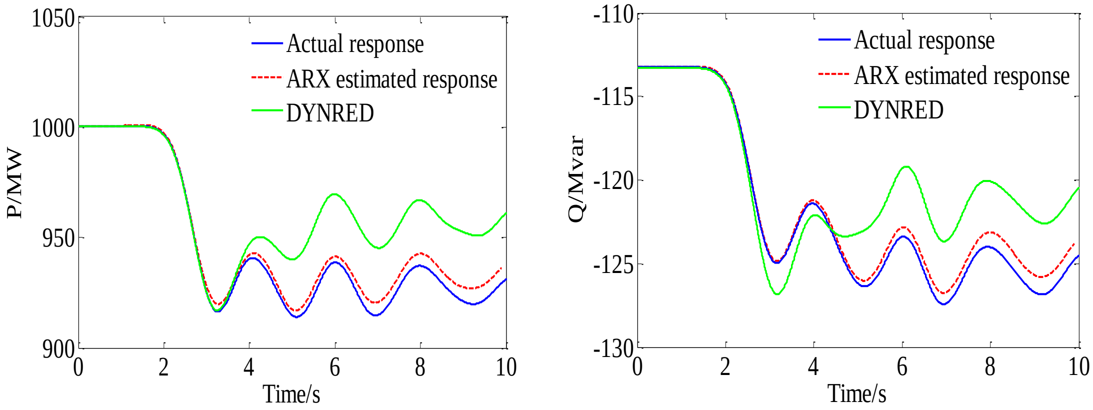


Figure 3-6 Model accuracy test on the EI system for case1



(a) Active power of tie line (b) Reactive power of tie line

Figure 3-7 Accuracy test results of the EI system for case1

one test (close to the training location) result is shown in the following figures because results of other tests were very similar.

Case2: Retained Florida Area

The Florida area was retained as a whole and other areas were reduced as shown in Figure 3-8. A generation trip in the Florida area was used for training and generation trips at other locations in the Florida area have been used for testing. The inputs are voltage and frequency of border buses in the study system. The model order is 0 (static model) and the active and the reactive power of one selected tie line are shown in Figure 3-9.

Both the ARX estimated active power and reactive power of the tie line are very accurate and closer to the actual response than DYNRED.

Case3: Reduced NPCC

The NPCC area is to be reduced and other areas are retained as shown in Figure 3-10. A generation trip in the TVA area is used for training and generation trips at other locations in the TVA area are used for testing. The inputs are voltage and frequency of the border buses in the study system. The model order is 0 (static model). Figure 3-11 shows the active power and reactive power of one selected tie line and indicates very good accuracy of the ARX estimated response.

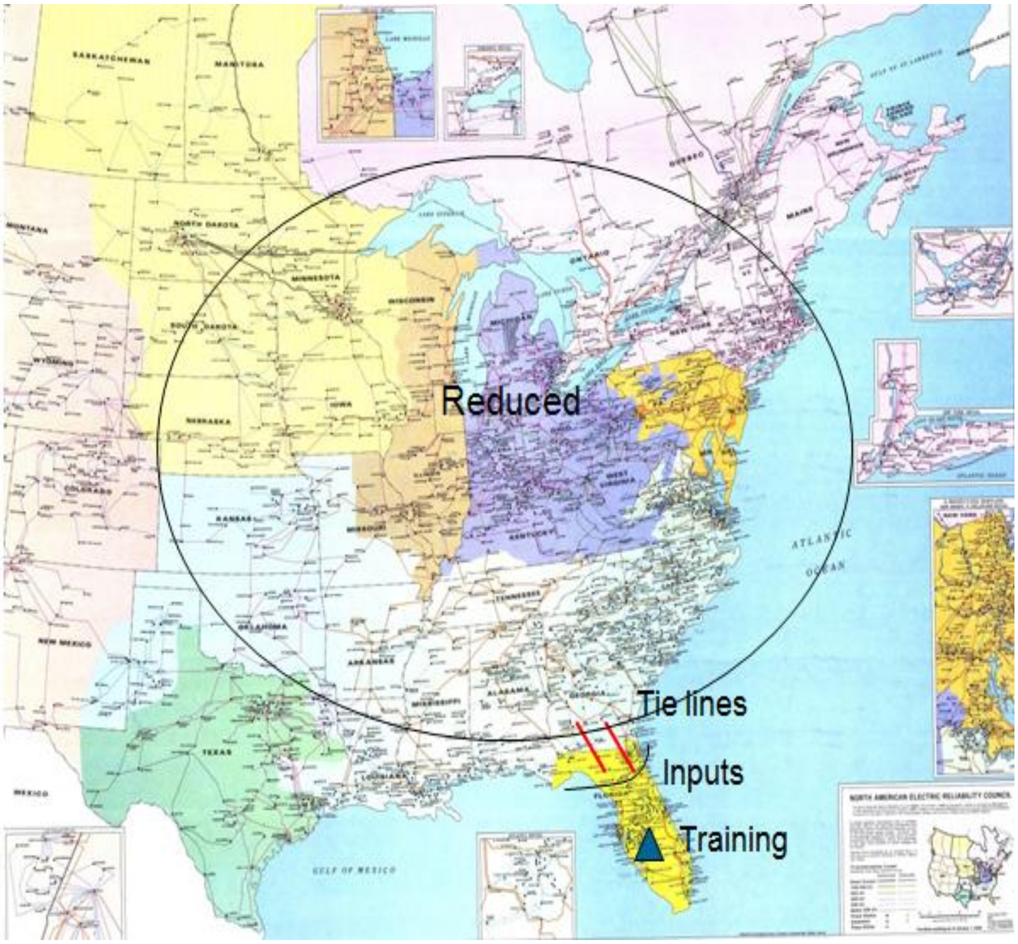
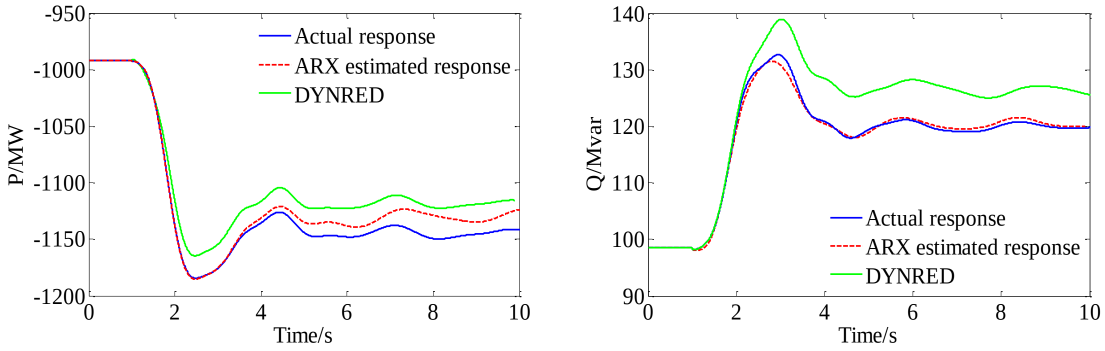


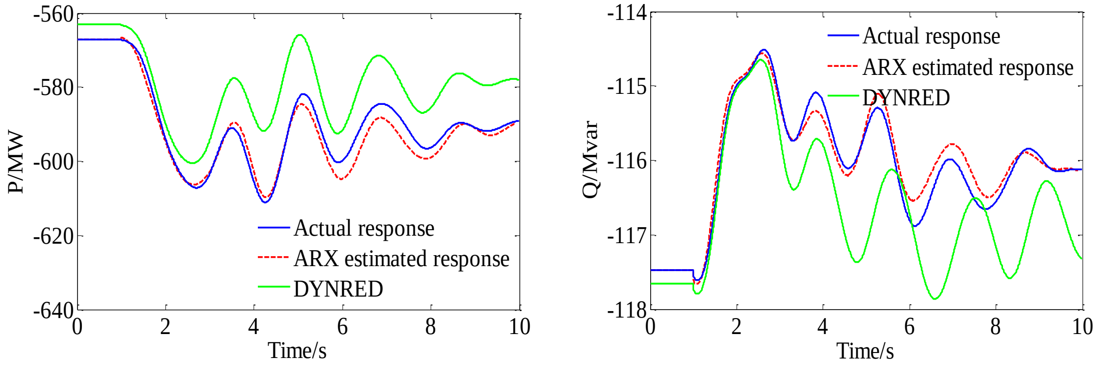
Figure 3-8 Model accuracy test on the EI system for case2



(a) Active power of tie line (b) Reactive power of tie line

Figure 3-9 Accuracy test results of the EI system for case2

Figure 3-10 Model accuracy test on the EI system for case3



(a) Active power of tie line

(b) Reactive power of tie line

Figure 3-11 Accuracy test results of the EI system for case3

Case4: Retained NPCC

The NPCC area was retained and other areas were reduced as shown in Figure 3-12. A generation trip in the NPCC area was used for training and generation trips at other locations in the NPCC area were used for testing. The inputs were voltage and frequency of the border buses in the study system. The model order is zero (static model). Active power and reactive power of one selected tie line are shown in Figure 3-13. It can be seen that the ARX estimated response is very accurate and it has a much better accuracy than DYNRED for the reactive power.

3.4 Integrated Simulation in Power System Dynamic Simulation Software

In order to implement the ARX model in PSS/E, a user-defined load-related model was chosen because it can do the calculation of state variables and the network solution can calculate current injections which are dependent on the bus voltage. The procedure is shown in Figure 3-14. The results of the reduced model and full model are compared.

A PSS/E user-defined load-related model was developed for the integrated simulation. Each tie line between the study system and the external system was replaced by connecting the load-related model to the border bus in the study system. In order to simplify the test case of input signal selection, some models were modified to retain only one tie line for the base case. The disturbance was a generation trip in the study system, and the ARX model order was zero (static model). The frequency and voltage at the disturbance location and interface location are shown. Note that the results in the

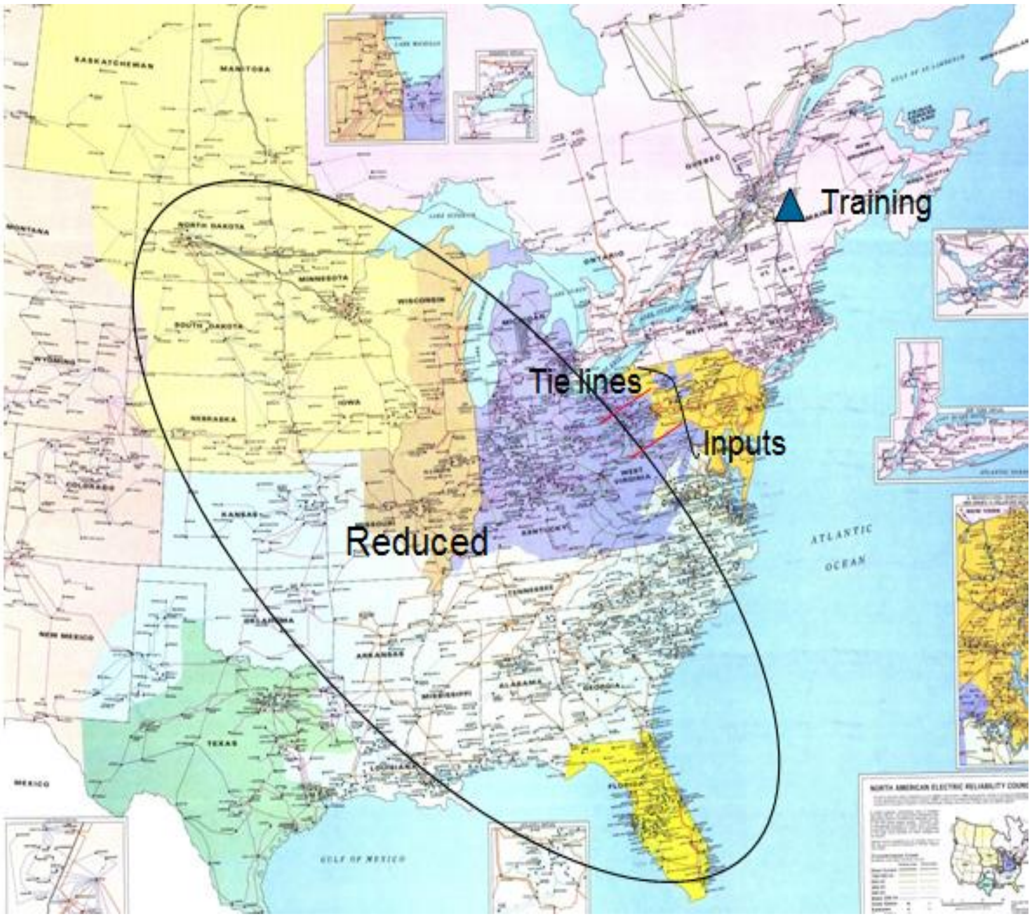
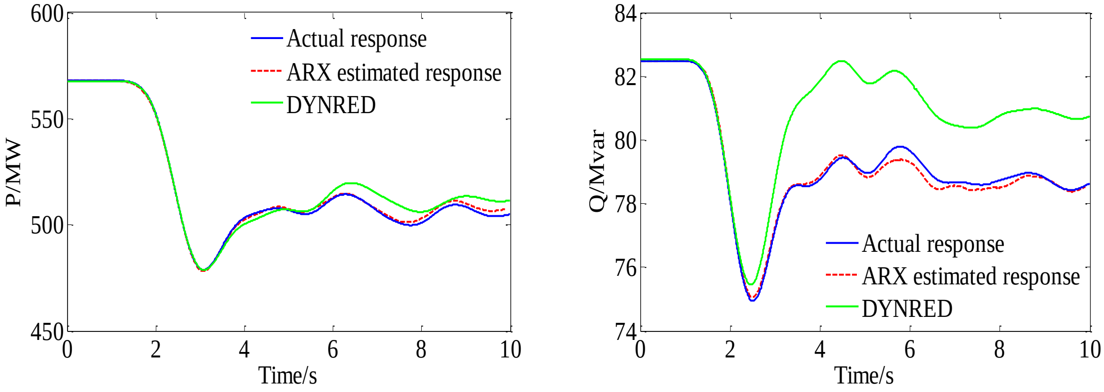


Figure 3-12 Model accuracy test on the EI system for case4



(a) Active power of tie line (b) Reactive power of tie line

Figure 3-13 Accuracy test results of the EI system for case4

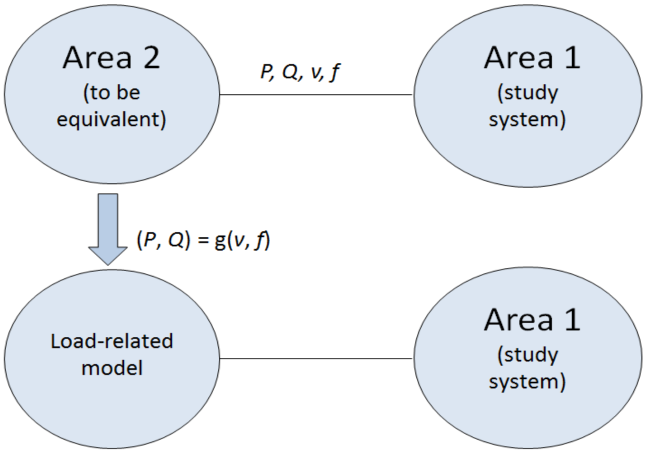


Figure 3-14 Integrated simulation in PSS/E

interface location were the worst case and a buffer area must be defined in real applications.

3.4.1 *23-bus System Model: Three Loads*

The study system and external system were connected by three tie lines as shown in Figure 3-15. The input of each ARX model is the voltage of the border bus of each tie line. The frequency and voltage at disturbance location and interface location are shown in Figure 3-16.

3.4.2 *23-bus System Model: One Load*

The study system and external system were connected by one tie line as shown in Figure 3-17. The inputs of the ARX model were either voltage ($v1/v2$) or frequency ($f1/f2$) from 2 buses near the interface. The input signals were varied to allow us to see their effect on the results. The frequency and voltage at disturbance location and interface location with input signals ($v1$, $v2$) are shown in Figure 3-18.

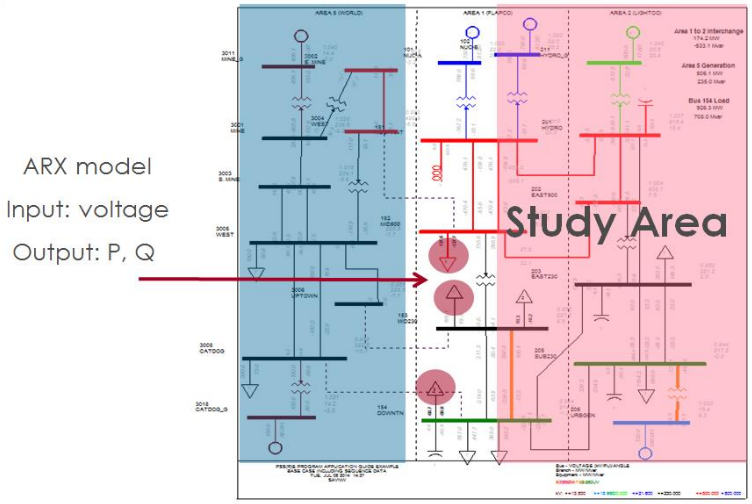
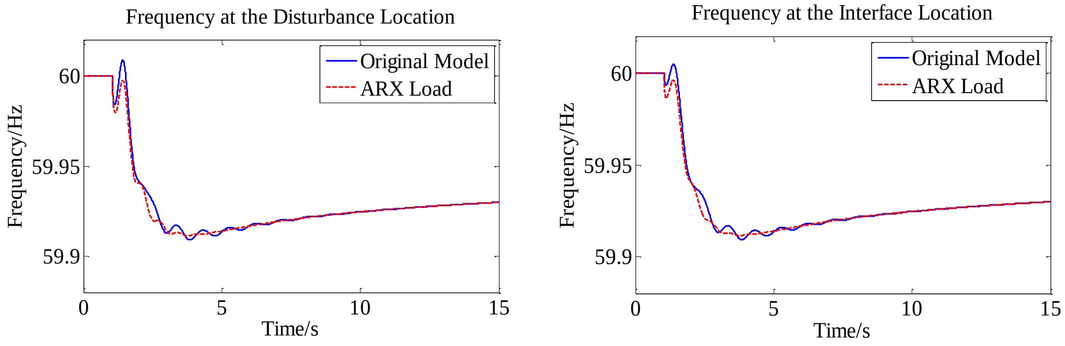
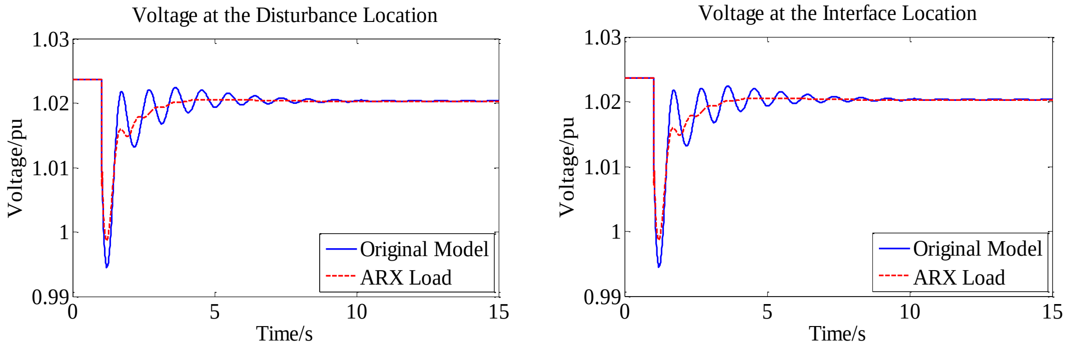


Figure 3-15 Integrated simulation of the 23-bus system



(a) Frequency at the Disturbance Location (b) Frequency at the Interface Location



(c) Voltage at the Disturbance Location (d) Voltage at the Interface Location

Figure 3-16 Integrated simulation results of the 23-bus system

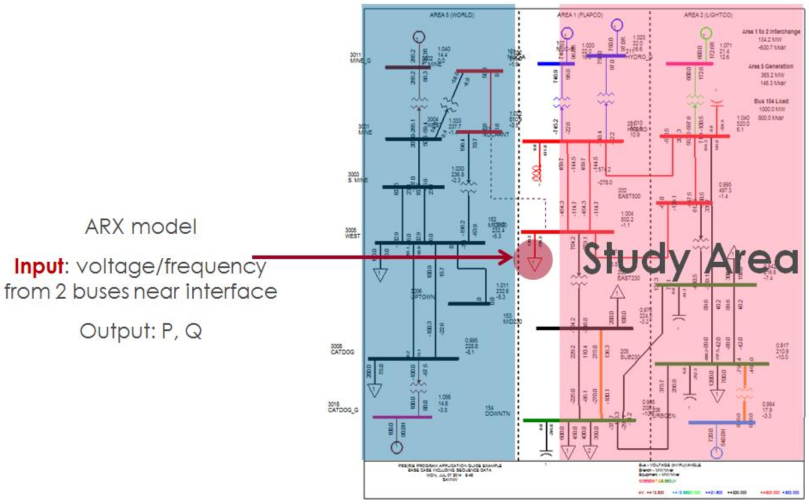
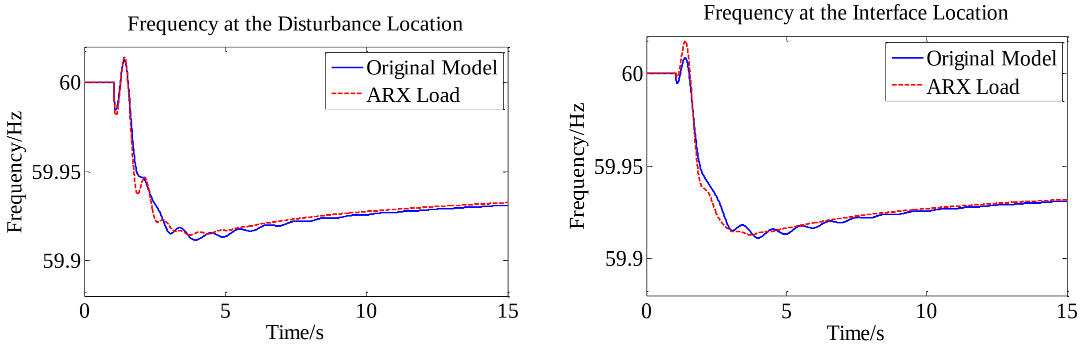
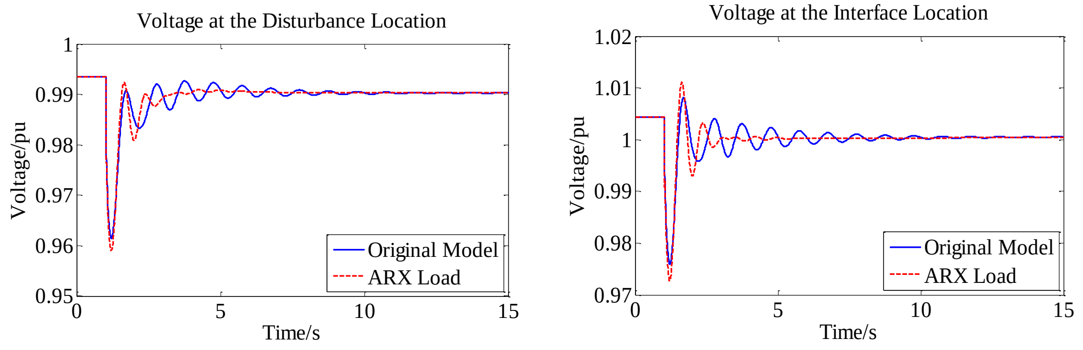


Figure 3-17 Integrated simulation of the 23-bus system with input signals selection



(a) Frequency at the Disturbance Location (b) Frequency at the Interface Location



(c) Voltage at the Disturbance Location (d) Voltage at the Interface Location

Figure 3-18 Integrated simulation results of the 23-bus system with input signals (*v*1, *v*2)

The results of other input signals are similar with (*v*1, *v*2), and their correlation coefficients with the original response are listed in Table 3-1.

Table 3-1 Results of 23-bus system with different input signals

Input Signals	f_CC_Dist	f_CC_Interface	v_CC_Dist	v_CC_Interface
(*v*1, *v*2)	0.9955	0.9955	0.9564	0.9084
(*v*1, *f*2)	0.9951	0.9951	0.9714	0.9390
(*f*1, *v*2)	0.9959	0.9958	0.9706	0.9366
(*f*1, *f*2)	0.9943	0.9946	0.9770	0.9539

f_CC_Dist — the correlation coefficient of frequency at the disturbance location;

f_CC_Interface — the correlation coefficient of frequency at the interface location;

v_CC_Dist — the correlation coefficient of voltage at the disturbance location;

v_CC_Interface — the correlation coefficient of voltage at the interface location.

3.4.3 The NPCC Model

The study system and external system were connected by two tie lines as shown in Figure 3-19. The inputs of the ARX model are either voltage (*v*1/*v*2) or frequency (*f*1/*f*2) from 2 buses near the interface. The input signals are varied to see their effects on the results. The frequency and voltage at disturbance location and interface location with input signals (*v*1, *f*2) are shown in Figure 3-20.

The results of other input signals are similar with (*v*1, *f*2) and their correlation coefficients with the original responses are listed in Table 3-2.

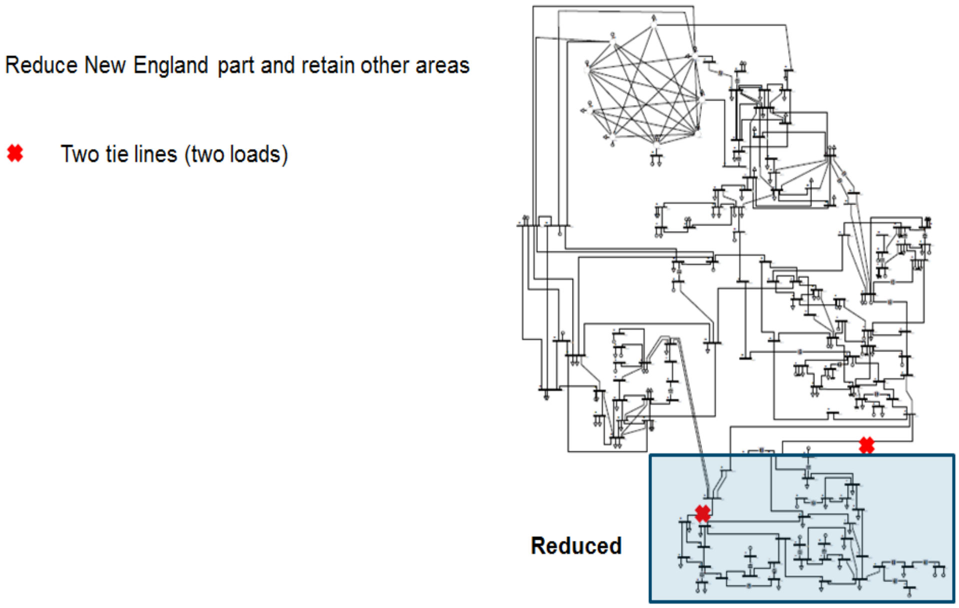
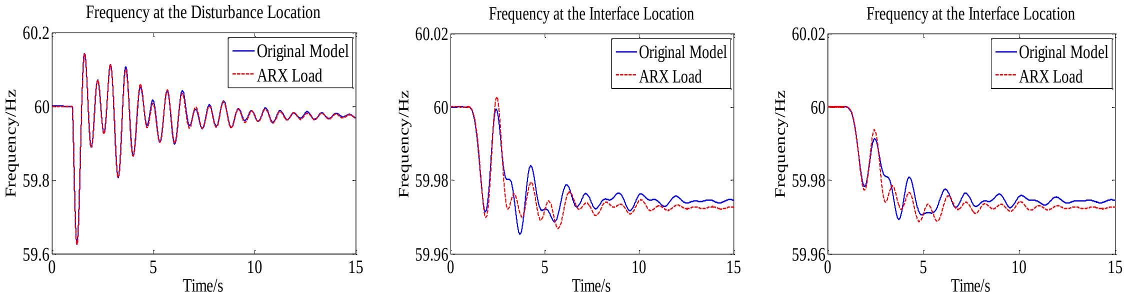
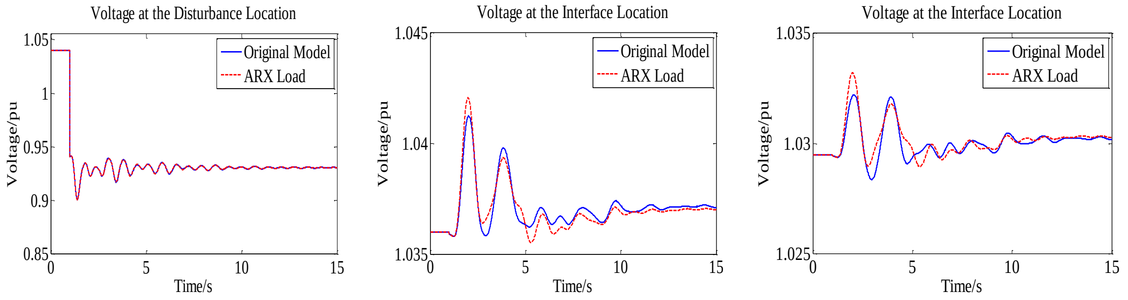


Figure 3-19 Integrated simulation of the NPCC system with input signals selection



(a) Frequency at Disturbance (b) Frequency at Interface1 (c) Frequency at Interface2



(d) Voltage at Disturbance (e) Voltage at Interface1 (f) Voltage at Interface2

Figure 3-20 Integrated simulation results of the NPCC system with input signals ($v1$, $f2$)

Table 3-2 Results of NPCC system with different input signals

Input Signals	f_CC_ Dist	f_CC_ Interface_1	f_CC_ Interface_2	v_CC_Dist	v_CC_ Interface_1	v_CC_ Interface_2
(*v*1, *v*2)	0.9943	0.8561	0.9224	0.9999	0.8708	0.7664
(*v*1, *f*2)	0.9986	0.9728	0.9792	1.0000	0.9612	0.8929

f_CC_Interface_1 — the correlation coefficient of frequency at the interface location of one tie line;

f_CC_Interface_2 — the correlation coefficient of frequency at the interface location of the other tie line.

Note: the results of $(v2, f1)$ and $(f1, f2)$ diverge in this particular case, and, thus, are not shown in the table.

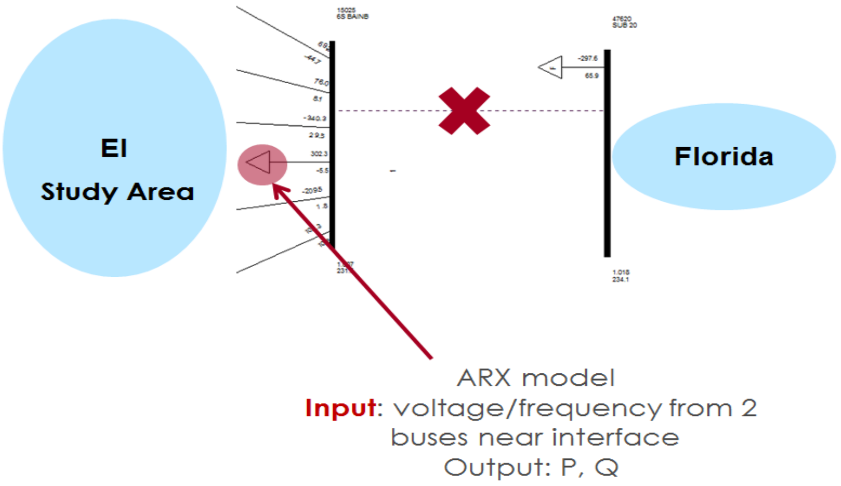
3.4.4 The 25000-bus EI Model

The study system and external system are connected by one tie line as shown in Figure 3-21. The inputs of the ARX model were either voltage (*v*1/*v*2) or frequency (*f*1/*f*2) from two buses near the interface. The input signals are varied to see their effects on the results. The frequency and voltage at disturbance location and interface location with input signals (*v*1, *v*2) are shown in Figure 3-22.

The results of other input signals were similar with (*v*1, *v*2) and their correlation coefficients with the original responses listed in Table 3-3.

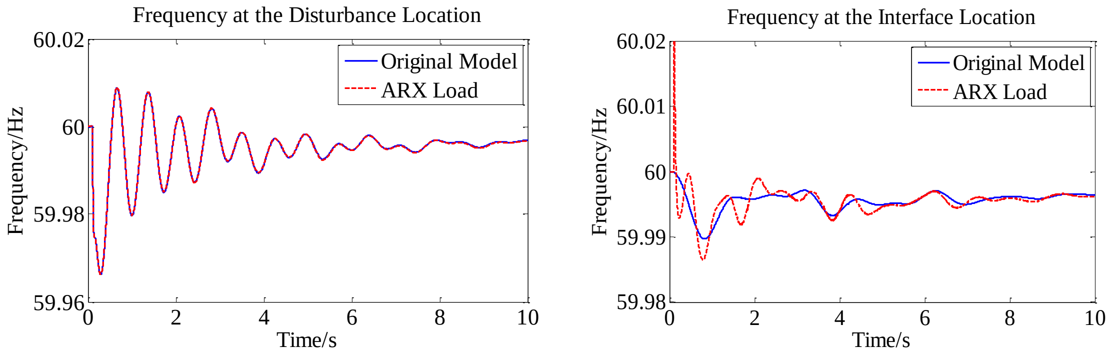
3.5 Conclusions

Our study has shown that measurement-based system reduction could be an attractive approach to overcome some disadvantages of the traditional model-based technique and increases the accuracy and speed of power system dynamic simulation. The ARX model was employed to reduce external systems, and the model accuracy was tested on three

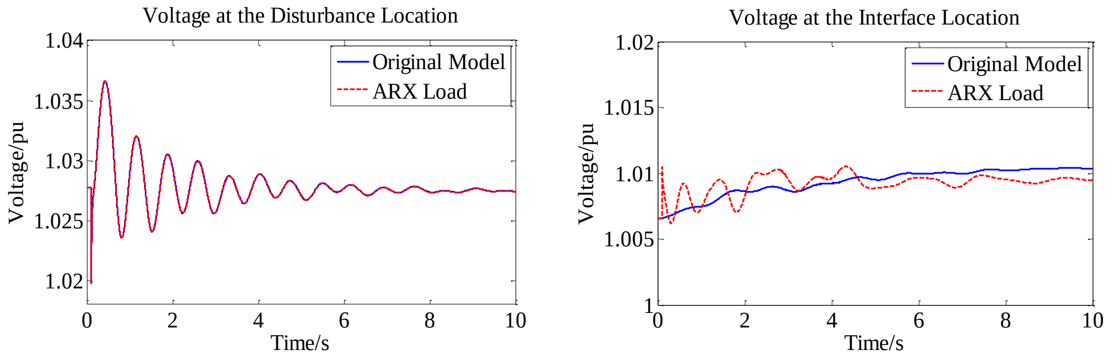


ARX model
Input: voltage/frequency from 2 buses near interface
Output: P, Q

Figure 3-21 Integrated simulation of the EI system with input signals selection



(a) Frequency at the Disturbance Location (b) Frequency at the Interface Location



(c) Voltage at the Disturbance Location (d) Voltage at the Interface Location

Figure 3-22 Integrated simulation results of the EI system with input signals ($v1$, $v2$)

Table 3-3 Results of EI system with different input signals

Input Signals	f_CC_Dist	f_CC_Interface	v_CC_Dist	v_CC_Interface
(v1, v2)	0.9999	0.6708	1.0000	0.6197
(v1, f2)	0.9995	0.8327	0.9999	-0.2280
(f1, v2)	0.9993	0.8027	0.9999	-0.1531
(f1, f2)	0.9989	0.5696	0.9998	0.3325

levels of systems, i.e., the 23-bus system, the NPCC system and the EI system. Results show that the ARX model was very accurate in prediction of similar disturbance events. A PSS/E user-defined load-related model was developed to integrate the ARX model with the study system and was implemented on those three system levels. One tie line or multiple tie lines between study systems and external systems were replaced by the load-related model. Frequency and voltage in both the disturbance location and interface location were shown and a preliminary study of input signals selection was conducted.

Future work includes using a higher-order ARX model, defining the buffer region and a more complete input signals selection study.

Chapter 4 Measurement-Based Generator Dynamic Model Validation

4.1 Introduction

The electric power industry relies heavily on power system dynamic models to perform extensive stability studies and to make critical decisions such as determining operating limit settings and corrective plans [46-48]. The accuracy of power system models is essential to safe power system operation and planning. However, several recent studies revealed that dynamic models fail to capture key characteristics of real systems [49-53]. The inconsistency between simulation and reality can lead to incorrect engineering judgement that eventually causes poor system dynamic performance, over or under utilization of resources, and even system instability.

Traditionally, as one of the most important components in system dynamics, generator model validation is conducted through staged testing, which require the tested equipment to be taken offline. Due to testing costs and economic opportunity loss, staged tests are usually performed every five to ten years. In contrast, generator operation modes could change within a day. Therefore offline testing methods cannot be guaranteed to maintain an up-to-date and accurate model for dynamic representation. This limitation led to recent studies focusing on online, PMU-based validation approaches [54-56]. In PMU-based approaches, a PMU is installed at the point of connection. Its event data are replicated by injecting them into a dynamic model through the use of hybrid simulation tools. Simulated generator responses are then compared to actual responses, validation, and model parameter adjustment, calibration, is done to best match the simulated and actual responses. However, these approaches focus on the calibration aspect while ignoring many key aspects of the validation portion. This paper presents a more detailed analysis on model validation and hybrid simulation to explore the potentials and limitations of such a method.

4.2 Validation Procedure and Simulation Method

Hybrid simulation takes real world measurements and incorporates them into the model. The theory behind hybrid simulation for generator model validation is that generator with a certain voltage magnitude and angle at its terminals will produce a set amount of real and reactive power. Thus, if voltage and angle measurements are injected to the generator terminals, the generator model real and reactive power curves should follow the real and reactive measurements.

The validation procedure is as follows:

1) Selecting events and preparing associated data (PMU, generator and GSU);

2) Adopting an appropriate simulation method, the simulated response of the generator model against the disturbance recorded by PMUs;

3) If noticeable mismatch is observed, specifying influential parameters;

4) Use multiple events to confirm;

5) Contact generator owners for further analysis.

The overall methodology of simulation is shown in Figure 4-1. The PMU records the voltage and current phasors, frequency, active and reactive power at the point of connection (POC). The voltage phasor is injected in the simulation through the use of a controllable voltage source at the POC. The controllable voltage source was created through an infinite machine and a lossless, ideal transformer with a tap changer and phase shifter. This method is known as the phase shifter method [57].

Active power and reactive power from the hybrid simulation are compared with the measurement values of the PMU to determine if the generator model is valid. A significant mismatch in the responses indicates the need for further analysis.

The tool developed for system validation is a Python-based PSS/E tool that provides the ability to easily validate any generator with a number of events

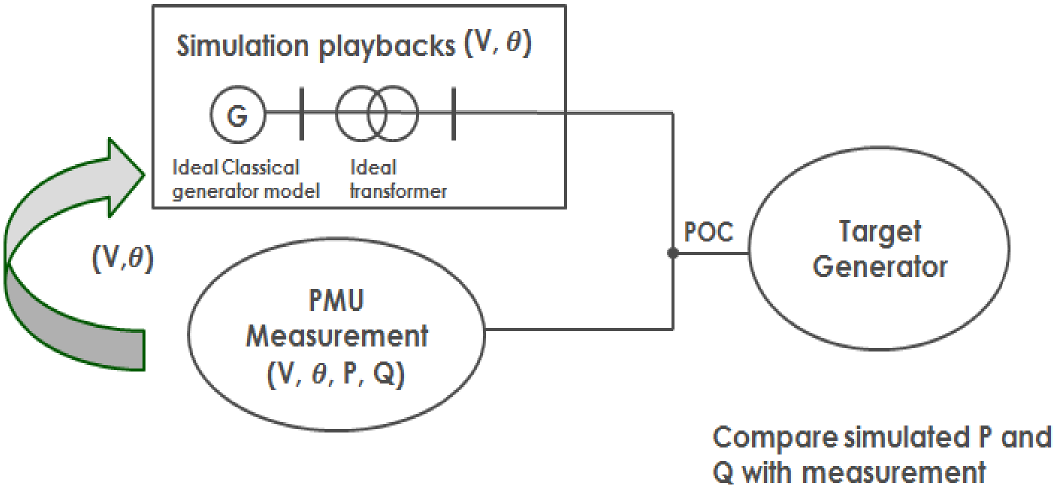


Figure 4-1 Generator model validation using hybrid simulation

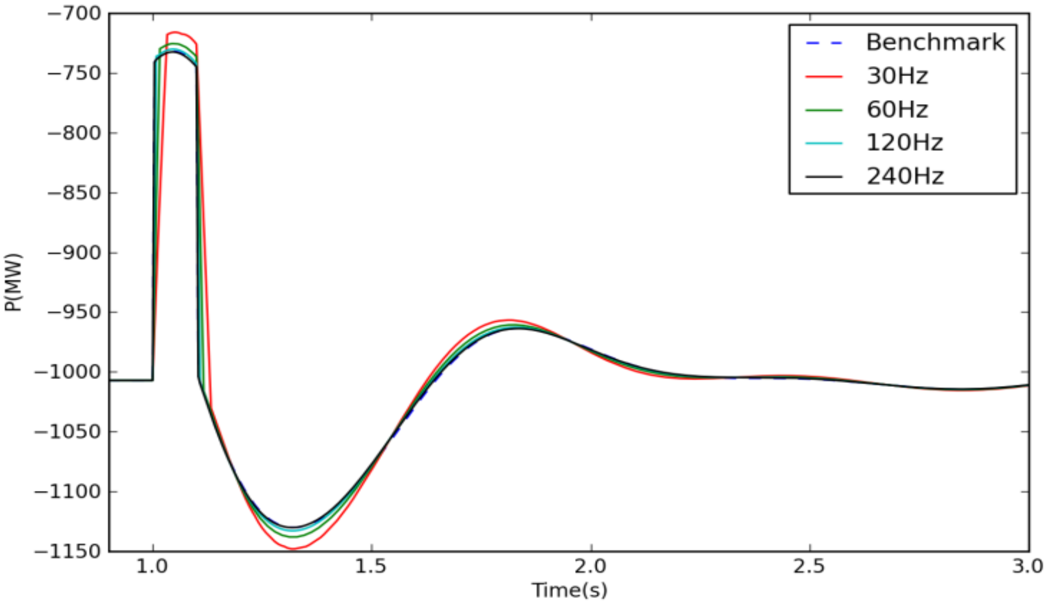
4.3 Validation Using Simulation Data

4.3.1 Effects of PMU Reporting Rates and Event Selection

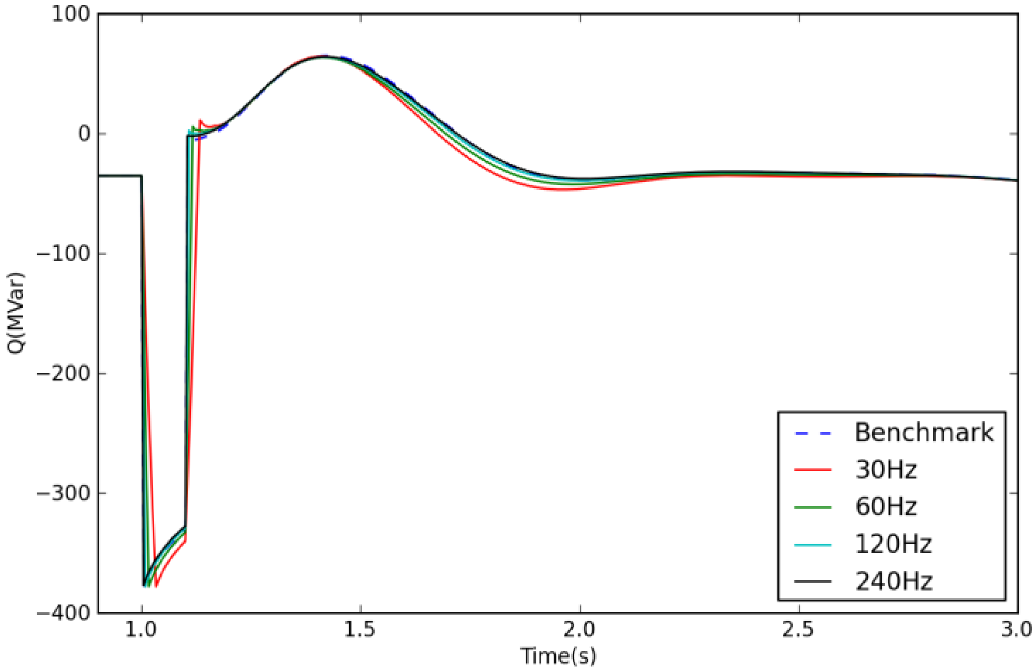
In order to study the potentials and limitations of the method, simulation data are used first. Certain events, such as line faults, generation trips, etc., were created in Eastern Interconnection (EI) model and the corresponding measurements were recorded to be used for the validation portion. Note that since the measurement data are recorded from the simulation, the parameters of the target generator are "perfect".

Two frequencies are involved in the study. One is the PMU reporting rate, typically it is 30/60Hz; the other is the time step used in the simulation software. It is typically larger than 120Hz and needs to be faster than the system dynamics, otherwise numerical issues may arise. Injecting the PMU measurements at its reporting rate would result in the generator seeing large steps at 30Hz intervals which would cause large inaccuracies. The PMU data should thus be interpolated to match the simulation frequency to help alleviate these steps. Linear interpolation is used here.

In this study, the EI and hybrid model simulation frequency is set to 240Hz. The PMU reporting rate is an important factor and lower reporting rates would introduce larger errors due to loss of information. The EI simulation output was down-sampled to various frequencies in Figure 4-2 to study the effect of reporting rates of PMUs on model validation. The event shown by Figure 4-2 is a 230kv 6-cycle line fault nine buses away from the POC of the generator.



(a) Active power



(b) Reactive power

Figure 4-2 Simulation output with different PMU reporting rates for a line trip event

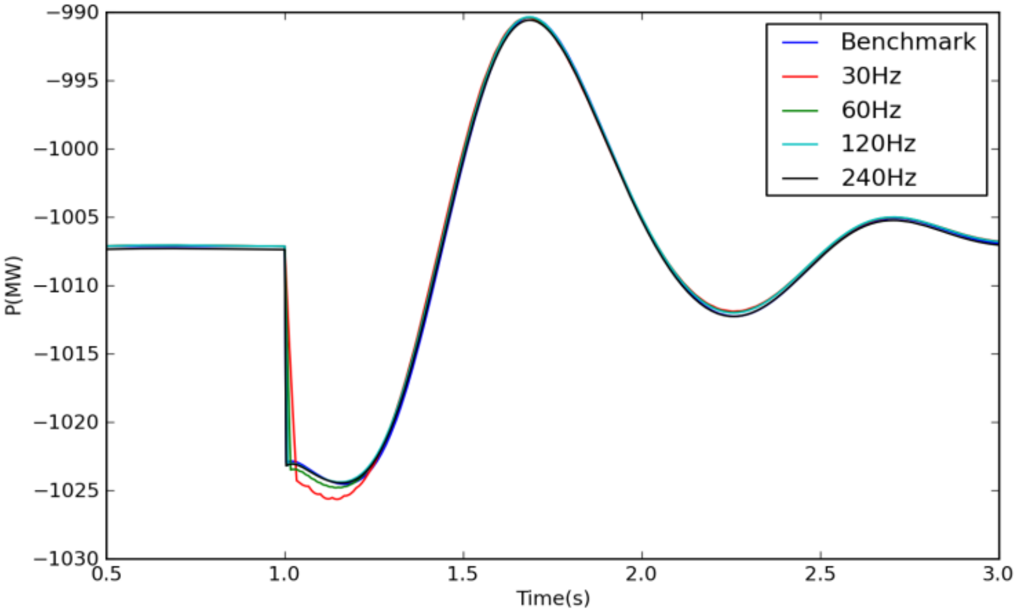
It can be seen that the 240Hz and 120Hz results match the benchmark well while the 60Hz and 30Hz reporting rates show a substantial mismatches. This mismatch may trick calibration techniques to calibrate parameters which they incorrectly think do not match the actual generator parameters. Thus, 60Hz and 30Hz PMU data of this event should not be used for validation. In addition, those mismatches caused by the reporting rate are dependent on the events. Figure 4-3 shows the result for a nearby generation trip using same procedure. In this event, even the 30Hz data mismatch is small after the first half swing and could therefore be used for validation only considering dynamics afterwards.

In general, the more severe the event is and the fewer points captured during the transient period, the larger the error introduced by PMU measurements becomes. Thus special attention needs to be paid when the event is chosen

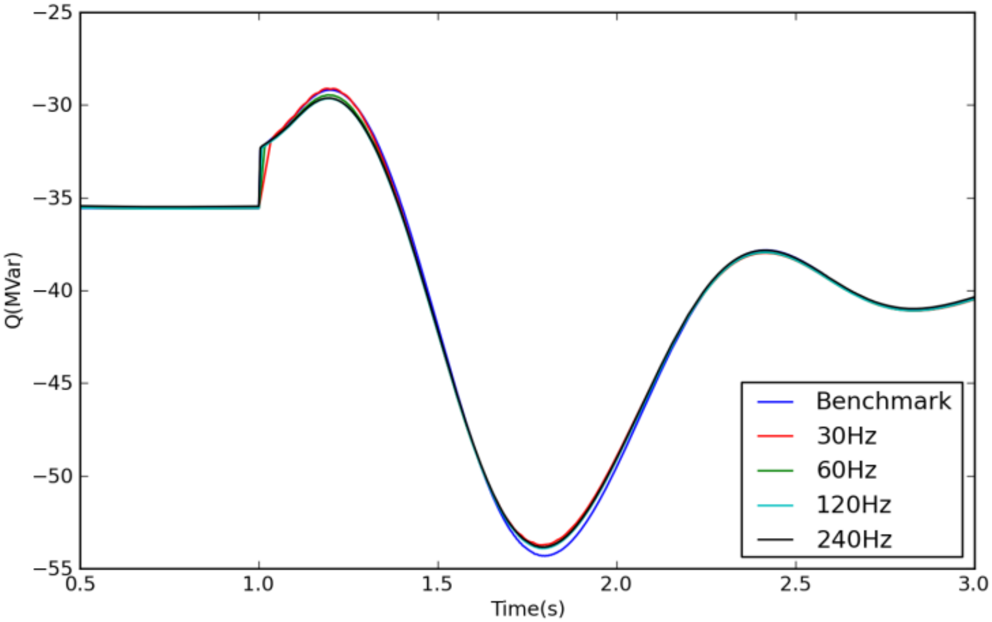
4.3.2 Parameter Analysis

Parameter analysis was initially conducted to see if changes in parameters were visible on the output of the hybrid simulation. Each parameter of each generator component was slowly changed by a set amount until a detectable mismatch was observed (based on engineer judgement). Simultaneous parameters were then changed together to study the masking effect.

Figure 4-4 include results for both active and reactive power after applying varied parameter changes ranging from a 5% to 40%. These tables specifically study a machine with a GENROU machine model, AC8B exciter model, and a PSS2B stabilizer model.

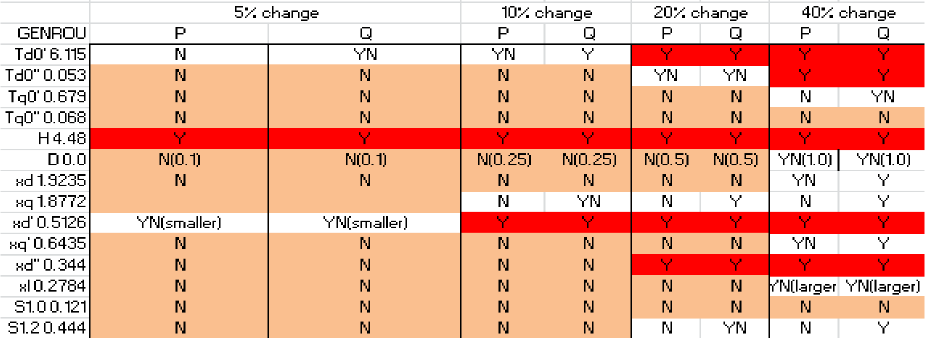


(a) Active power



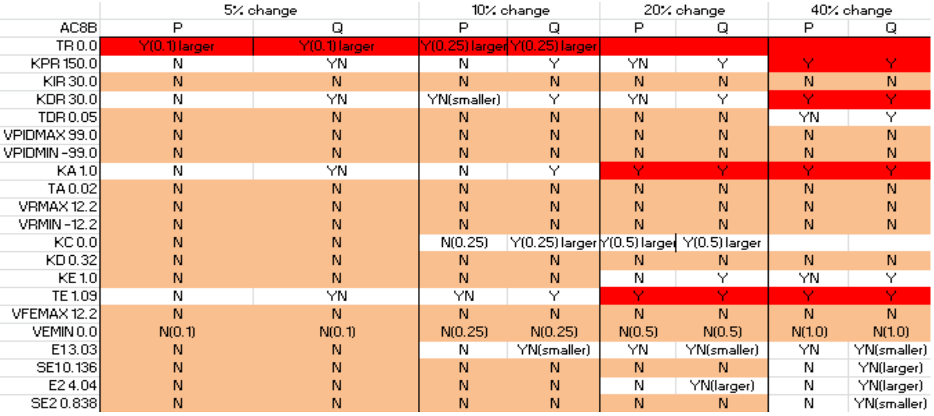
(b) Reactive power

Figure 4-3 Simulation output with different PMU reporting rates for a nearby generation trip event



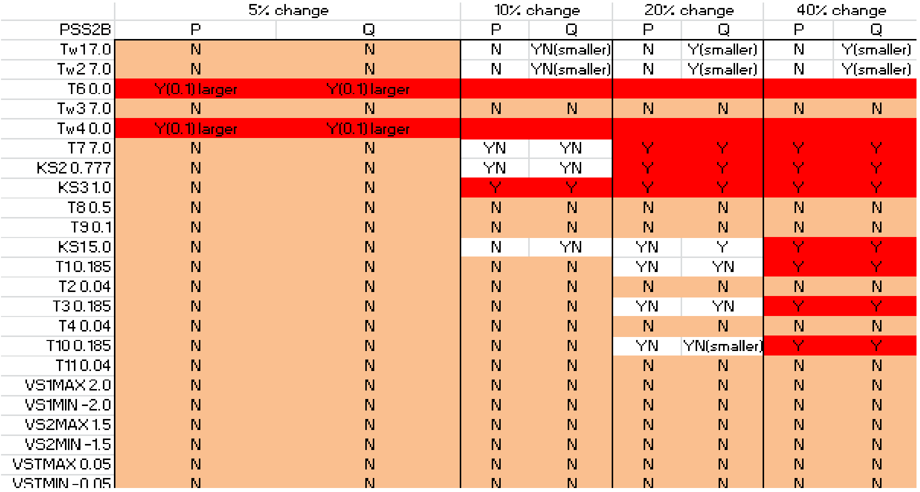
	5% change		10% change		20% change		40% change	
GENROU	P	Q	P	Q	P	Q	P	Q
Td0' 6.115	N	YN	YN	Y	Y	Y	Y	Y
Td0'' 0.053	N	N	N	N	YN	YN	Y	Y
Tq0' 0.679	N	N	N	N	N	N	N	YN
Tq0'' 0.068	N	N	N	N	N	N	N	N
H 4.48	Y	Y	Y	Y	Y	Y	Y	Y
D 0.0	N(0.1)	N(0.1)	N(0.25)	N(0.25)	N(0.5)	N(0.5)	YN(1.0)	YN(1.0)
xd 1.9235	N	N	N	N	N	N	YN	Y
xq 1.8772			N	YN	N	Y	N	Y
xd' 0.5126	YN(smaller)	YN(smaller)	Y	Y	Y	Y	Y	Y
xq' 0.6435	N	N	N	N	N	N	YN	Y
xd'' 0.344	N	N	N	N	Y	Y	Y	Y
xl 0.2784	N	N	N	N	N	N	YN(larger	YN(larger)
S1.0 0.121	N	N	N	N	N	N	N	N
S1.2 0.444	N	N	N	N	N	YN	N	Y

(a) Machine model (GENROU)



	5% change		10% change		20% change		40% change	
AC8B	P	Q	P	Q	P	Q	P	Q
TR 0.0	Y(0.1) larger	Y(0.1) larger	Y(0.25) larger	Y(0.25) larger				
KPR 150.0	N	YN	N	Y	YN	Y	Y	Y
KIR 30.0	N	N	N	N	N	N	N	N
KDR 30.0	N	YN	YN(smaller)	Y	YN	Y	Y	Y
TDR 0.05	N	N	N	N	N	N	YN	Y
VPIDMAX 99.0	N	N	N	N	N	N	N	N
VPIDMIN -99.0	N	N	N	N	N	N	N	N
KA 1.0	N	YN	N	Y	Y	Y	Y	Y
TA 0.02	N	N	N	N	N	N	N	N
VRMAX 12.2	N	N	N	N	N	N	N	N
VRMIN -12.2	N	N	N	N	N	N	N	N
KC 0.0	N	N	N(0.25)	Y(0.25) larger	Y(0.5) larger	Y(0.5) larger		
KD 0.32	N	N	N	N	N	N	N	N
KE 1.0	N	N	N	N	N	Y	YN	Y
TE 1.09	N	YN	YN	Y	Y	Y	Y	Y
VFEMAX 12.2	N	N	N	N	N	N	N	N
VEMIN 0.0	N(0.1)	N(0.1)	N(0.25)	N(0.25)	N(0.5)	N(0.5)	N(1.0)	N(1.0)
E1 3.03	N	N	N	YN(smaller)	YN	YN(smaller)	YN	YN(smaller)
SE1 0.136	N	N	N	N	N	N	N	YN(larger)
E2 4.04	N	N	N	N	N	YN(larger)	N	YN(larger)
SE2 0.838	N	N	N	N	N	N	N	YN(smaller)

(b) Exciter model (AC8B)



	5% change		10% change		20% change		40% change	
PSS2B	P	Q	P	Q	P	Q	P	Q
Tw1 7.0	N	N	N	YN(smaller)	N	Y(smaller)	N	Y(smaller)
Tw2 7.0	N	N	N	YN(smaller)	N	Y(smaller)	N	Y(smaller)
T6 0.0	Y(0.1) larger	Y(0.1) larger						
Tw3 7.0	N	N	N	N	N	N	N	N
Tw4 0.0	Y(0.1) larger	Y(0.1) larger						
T7 7.0	N	N	YN	YN	Y	Y	Y	Y
KS2 0.777	N	N	YN	YN	Y	Y	Y	Y
KS3 1.0	N	N	Y	Y	Y	Y	Y	Y
T8 0.5	N	N	N	N	N	N	N	N
T9 0.1	N	N	N	N	N	N	N	N
KS1 5.0	N	N	N	YN	YN	Y	Y	Y
T1 0.185	N	N	N	N	YN	YN	Y	Y
T2 0.04	N	N	N	N	N	N	N	N
T3 0.185	N	N	N	N	YN	YN	Y	Y
T4 0.04	N	N	N	N	N	N	N	N
T10 0.185	N	N	N	N	YN	YN(smaller)	Y	Y
T11 0.04	N	N	N	N	N	N	N	N
VS1MAX 2.0	N	N	N	N	N	N	N	N
VS1MIN -2.0	N	N	N	N	N	N	N	N
VS2MAX 1.5	N	N	N	N	N	N	N	N
VS2MIN -1.5	N	N	N	N	N	N	N	N
VSTMAX 0.05	N	N	N	N	N	N	N	N
VSTMIN -0.05	N	N	N	N	N	N	N	N

(c) PSS model (PSS2B)

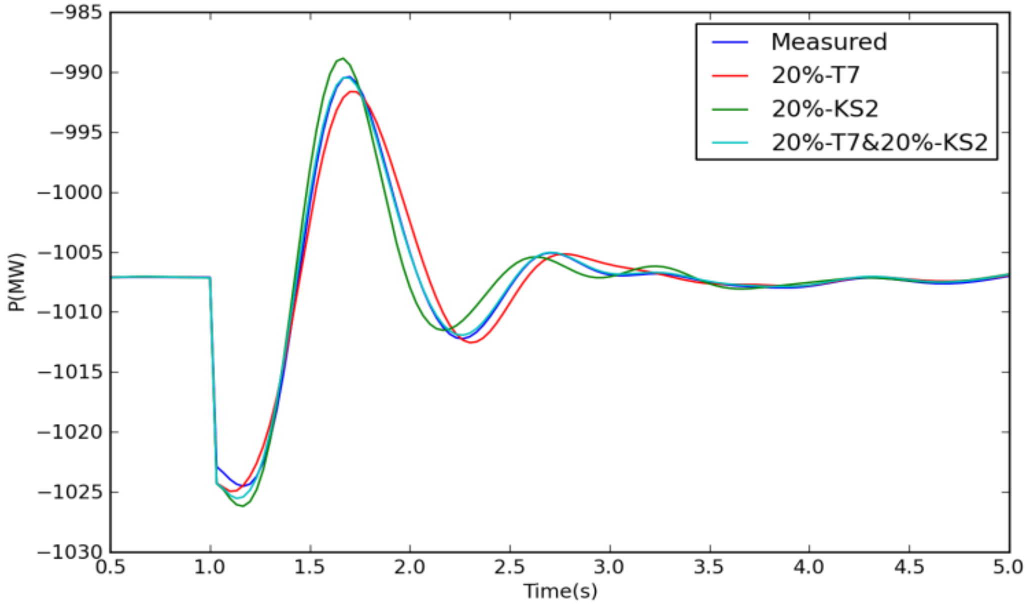
Figure 4-4 Parameter analysis results

Red indicates the mismatch is visible (Y) while brown is not visible (N). (YN) is between (Y) and (N). Note that if a smaller change is visible, a larger change will also be visible. Though the results vary for different generators and models, these tables and others like them can be used as a guideline for the validation procedure and to tell what level of accuracy to expect in each parameter. It can be seen that not all parameter inaccuracies may be detectable. Thus a PMU-based method may not be able to completely guarantee that all parameters are accurate.

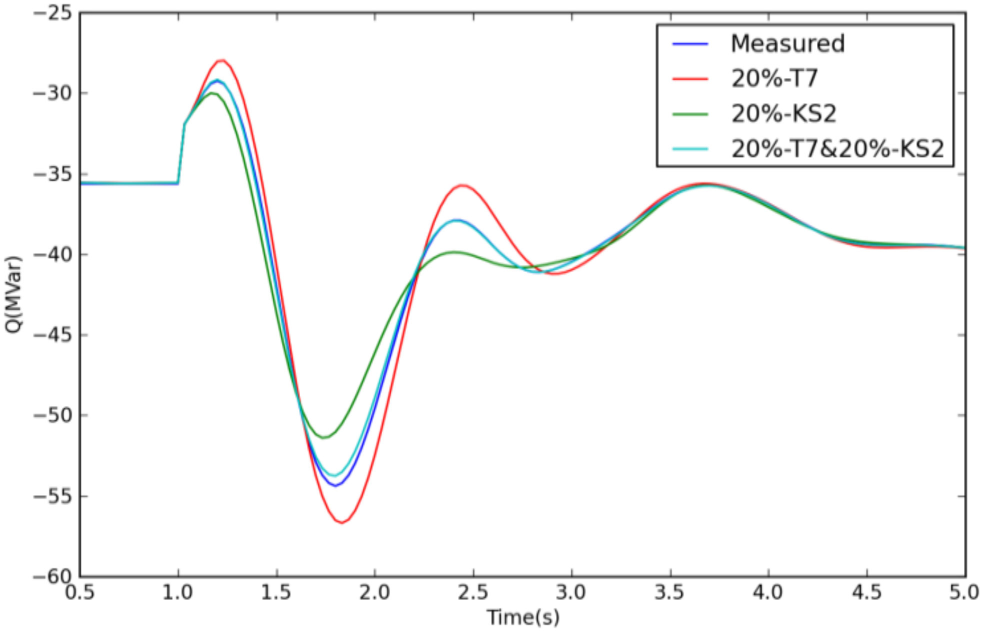
4.3.3 Masking Effect

Since there are so many parameters in a generator that are all coupled together in various ways, it is highly possible that the error introduced by one parameter could be masked by the error introduced in another parameter so that the mismatch in the output is not noticeable. This phenomenon is known as the masking effect.

For example, the effects of T7 and KS2 in the PSS2B model were found to have similar results in the parameter analysis. Thus, creating opposite errors in the two would result in opposite behaviors which may cause a masking effect. The results of varying these two parameters both individually and simultaneously are shown in Figure 4-5. It is observed that they create a masking effect, i.e. the mismatch is observable by a change of one parameter but it is not observable by a change of both. This indicates another limitation of this and any online validation method.



(a) Active power



(b) Reactive power

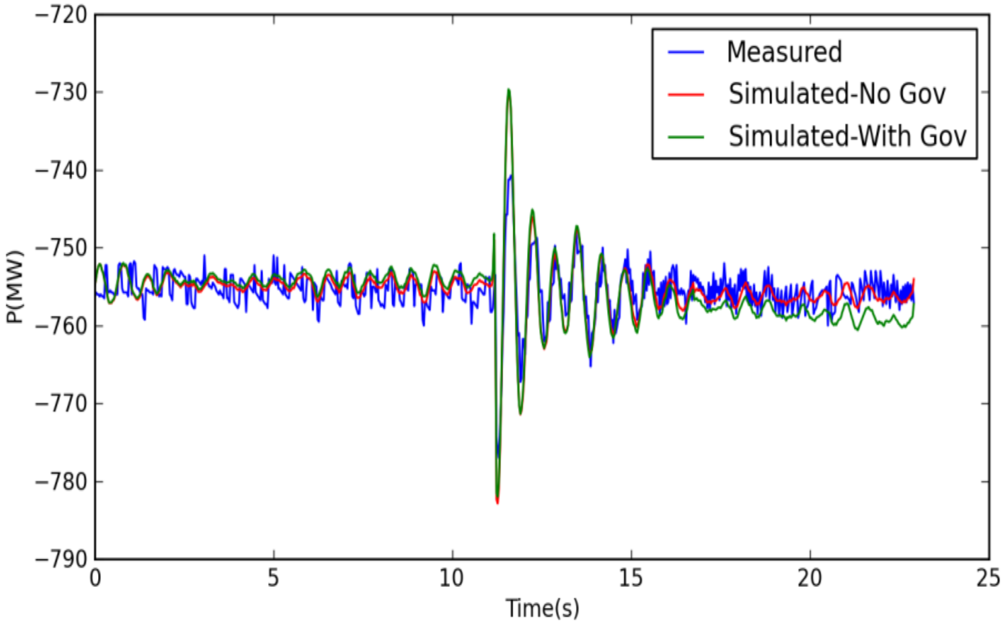
Figure 4-5 Masking effect in PSS model (PSS2B)

4.4 Validation Using Real Measurement Data

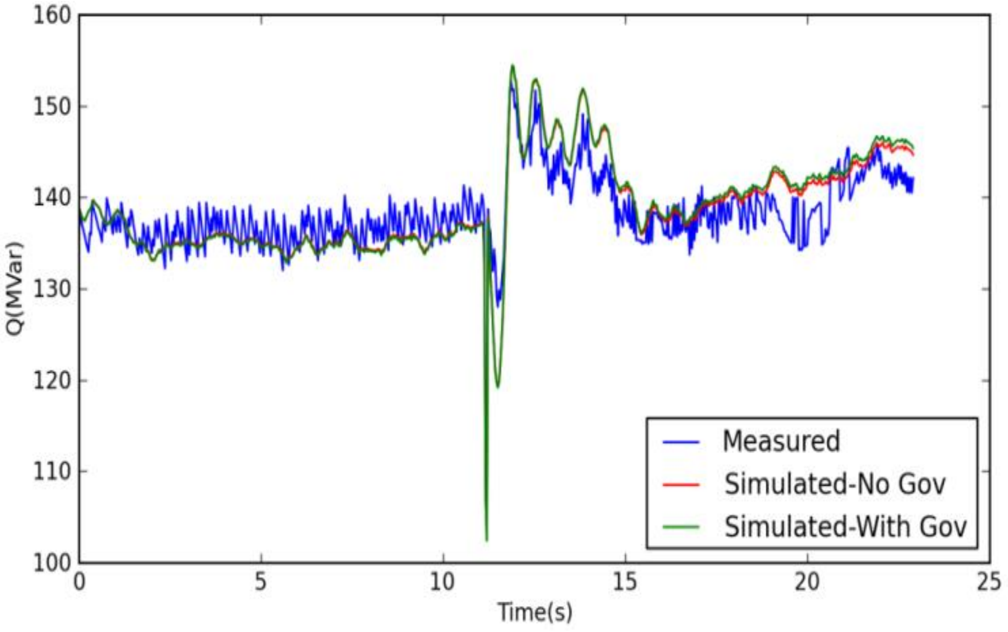
This section presents validation results using real PMU data. Initial condition tuning has to be done prior to inserting the actual measurement data to make sure the validation model's dynamic initial conditions match that of the generator at the beginning of the measurement window. Two separate units, Unit 1 and Unit 2, are each tested using two different events.

Unit 1 includes a machine model (GENROU), an exciter model (ESST1A), a governor model (IEEEG1), and is validated using two events. Both events are generation trips that occurred on the system while the PMU was measuring at the POC of the generator. Fig. 5 shows the first event, generation trip 1. It can be seen that the responses match quite well. It was discovered with this event that either the governor dead-band for this particular machine is modeled incorrectly or the governor was off as can be seen by the constant power drift in Figure 4-6. Such a phenomenon is also reported in [19]. Thus the governor in the model is turned off making the simulation response match the real response much better.

The second case, generation trip 2, tests the response of the same generator for a different event and is shown in Figure 4-7. In this case, the real power response mismatch is small, but the reactive power response gains an offset after the dynamics. It was not determined which parameter caused this response difference or operation of additional reactive power devices in the plant during the event, but it is most likely something in either the stabilizer or the exciter that needs to be further investigated if no additional devices were

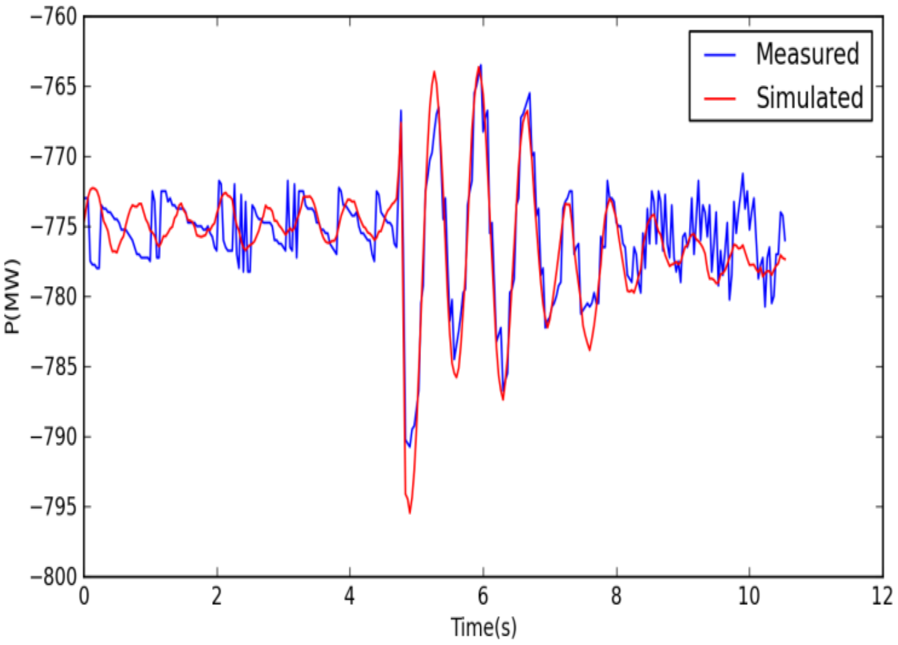


(a) Active power

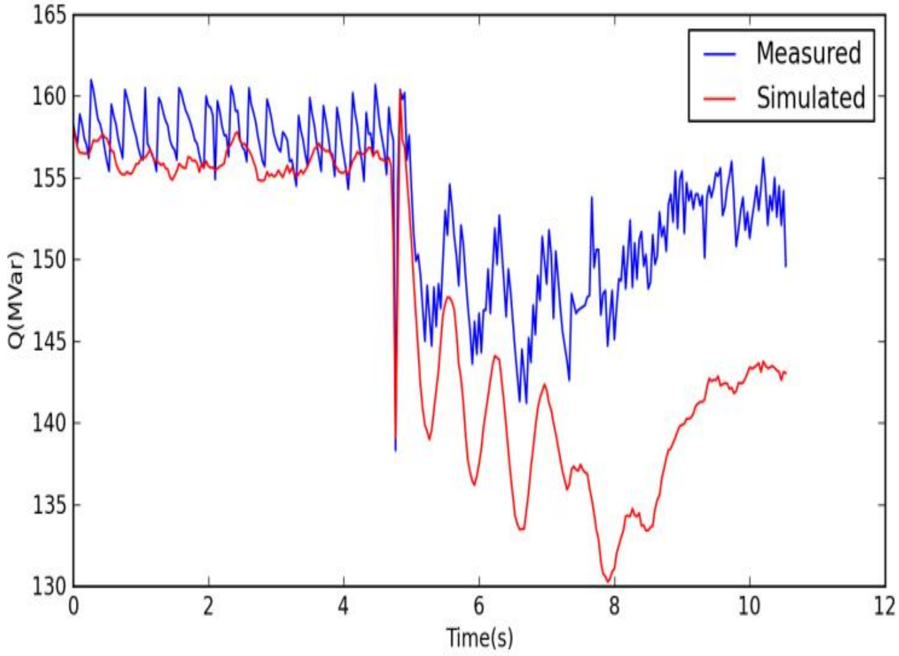


(b) Reactive power

Figure 4-6 Validation of Unit#1 using generation trip#1



(a) Active power



(b) Reactive power

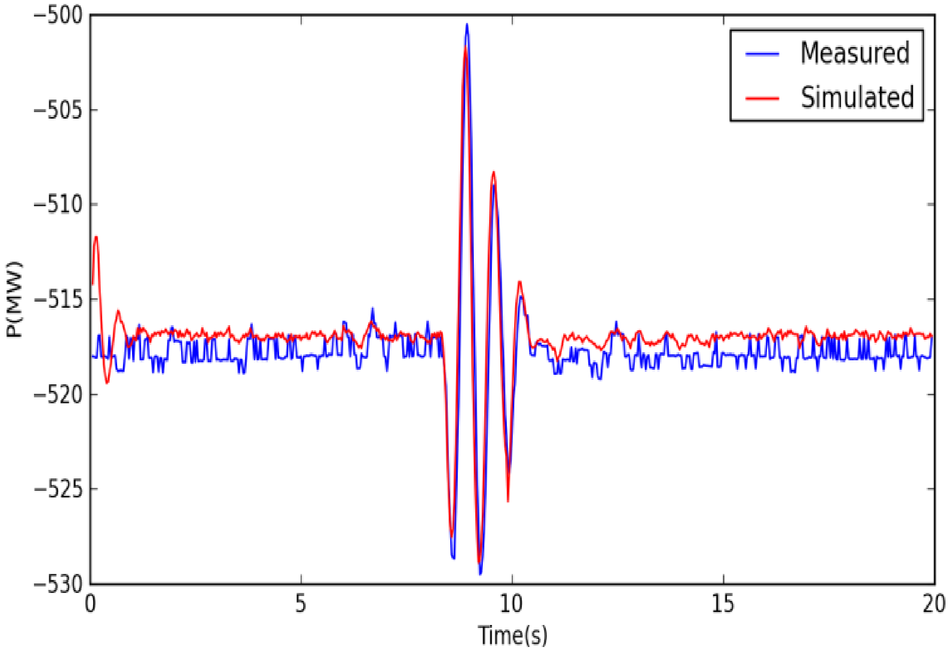
Figure 4-7 Validation of Unit#1 using generation trip#2

used. The point of note is that for one machine, one event gave a good validation after the governor was turned off while another event showed that some other component had inaccuracies that the first event did not capture. As such, careful consideration on the number of events used needs to be taken if validation or calibration were to be performed.

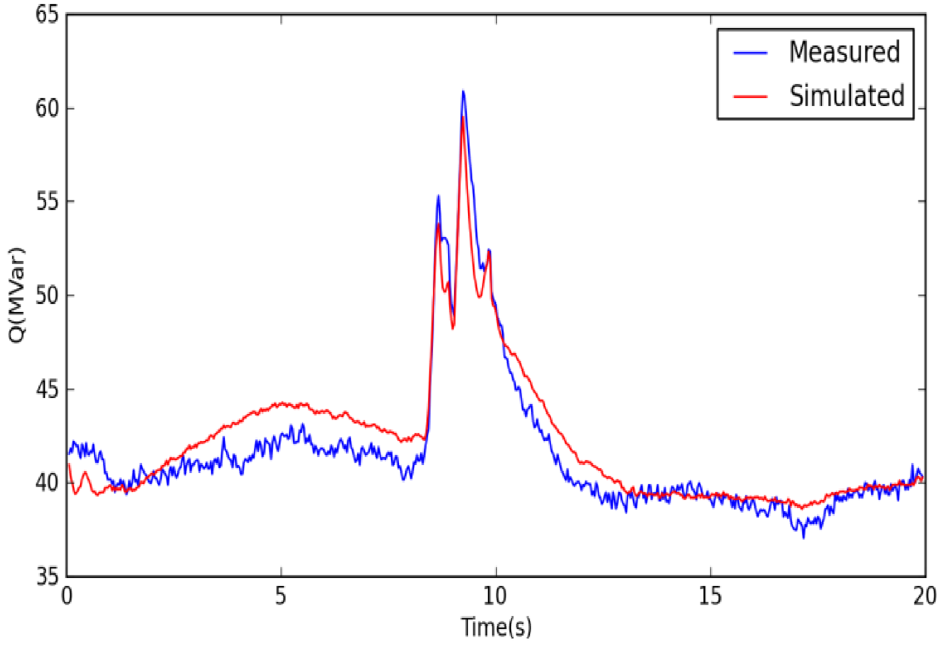
Unit 2 includes a machine model (GENROU), an exciter model (AC7B), a governor model (TGOV1), and a PSS model (PSS2A). As with the previous unit, the governor is turned off. Figure 4-8 and 4-9 show the validation results using two generation trip events. In this case, the generator's model matches well in both events for both real and reactive responses.

4.5 Conclusions

Measurement-based generator dynamic model validation is a promising tool which could be used for more frequent model validation. This tool is developed using PSS/E and Python so that it simplifies the procedure and eventually could perform validation automatically once proper measurement data are available. This chapter presents a detailed analysis of the validation procedure, including both potentials and limitations of this method. By using the simulation data, the effect of different PMU reporting rates and event selection, parameter analysis and masking effect are discussed. Finally, the validation utilizes real PMU data. It shows the necessity of using multiple events and the caution of using automatic calibration without thought to the validation.

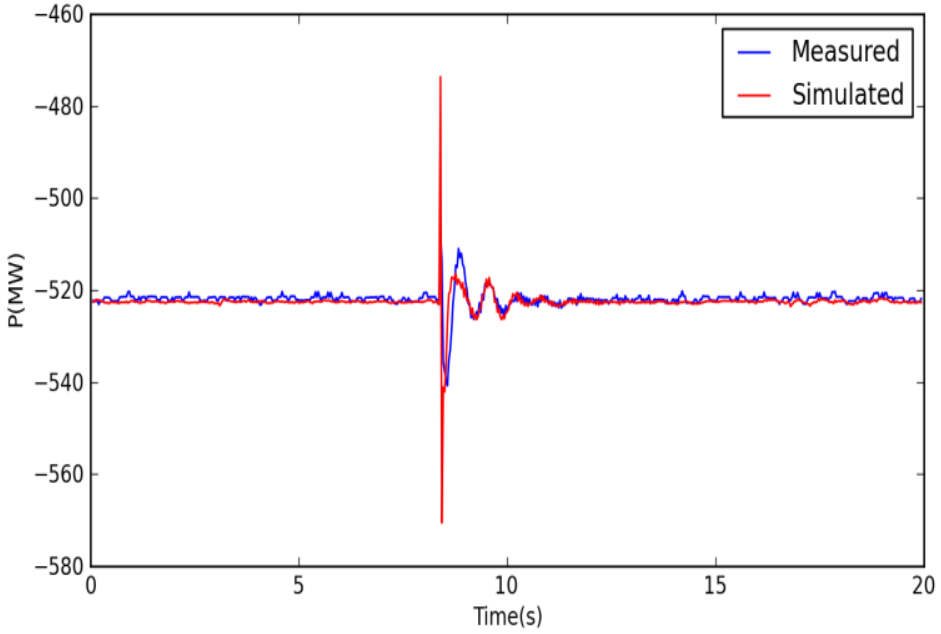


(a) Active power

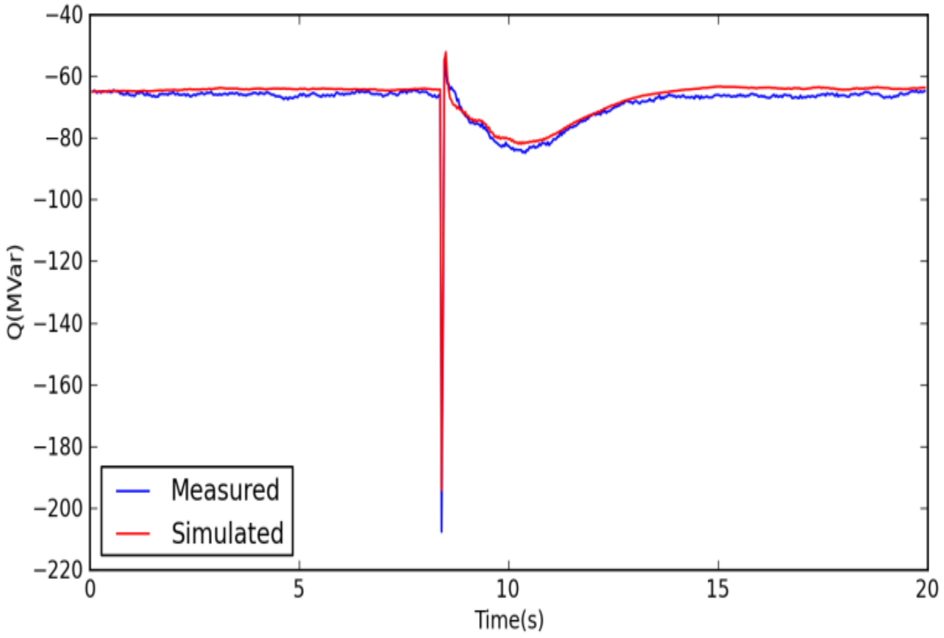


(b) Reactive power

Figure 4-8 Validation of Unit#2 using generation trip#1



(a) Active power



(b) Reactive power

Figure 4-9 Validation of Unit#2 using generation trip#2

Chapter 5 Application of Power System Measurement for Digital Authentication

5.1 Introduction

The so-called Electric Network Frequency (ENF) Criterion is based on an observation made by Grigoras [58] in 2003. This observation concerns a "hum" that is frequently found in digital recordings. In his 2003 paper Grigoras showed that this hum is related to the sinusoidal voltage signal that runs through electric power grids. This sinusoidal voltage signal is approximately constant across the whole grid. The importance of this observation comes from the fact that the frequency of this grid sinusoidal voltage signal, at any instant of time, is dependent on the amount of power being generated and the amount of power being consumed across the grid. Since the amount of power that is consumed at any instant of time is random, a sequence of the frequencies of this hum will be a random process. Consequently, given the "uniqueness" of patterns that occur in random processes, it ought to be possible to extract a frequency sequence from a recording, the target recording, and compare this frequency sequence against a database of grid reference frequencies to determine the time when a recording was made.

The two step process proposed by Grigoras makes up what will be called the ENF analysis process. The first step, where frequencies are estimated from the target recording, will be called the ENF extraction step. The output of this step is the target sequence. In the second step, this target sequence is compared to subsequences from a grid reference database. This step usually involves some type of distance measure. The subsequence of the grid reference database which has the minimum distance (maximum distance) to the target sequence is considered to be the "matching" sequence in the grid reference database. The time that the matching sequence occurred is considered to be the time that the target recording was made. This whole process of taking the target sequence generated by the ENF extraction step and creating a "match" with a subsequence in the grid reference database will be called the ENF comparison set.

Since the ENF Criterion was first proposed, numerous papers [59-62] have been written on using this methodology for doing forensic authentication of digital recordings. Since there are relatively few tools for doing the forensic examination of digital recordings, the ENF Criterion seemingly represents a valuable new tool for doing this authentication.

The goals of this research are to improve the understanding of the whole ENF analysis process using the synchronized, widely deployed, and high resolution FNET/GridEye system. To this end four research studies have been conducted. These studies have resulted in the following:

- A new ENF comparison algorithm that removes the effects of oscillator error [63].
- A better understanding of the phenomenology of how grid sinusoidal voltage signals get imbedded into digital recordings [64].
- An ENF based algorithm for detecting tampering in digital recordings [65].

- A frequency localization study that may determine the location when the digital recording was made.

5.2 Removing the Effects of Oscillator Error from ENF Comparisons

5.2.1 Overview

Oscillators power all digital recorders. They are the fundamental element in the clocks that is used to time when analog-to-digital conversions are made. Most oscillators do not run exactly at their specified frequencies. These timing errors affect both the values of the frequency estimates as well as the times between frequency estimates. In fact, a number of investigators have observed offsets between the frequencies making up a target sequence and the frequencies making up grid subsequences. While the investigators seemingly did not know it at the time, these offsets were caused by oscillator errors. To address these offsets, one of these investigators suggested using the correlation coefficient as the comparison distance measure. The investigator did this because the correlation coefficient is seemingly insensitive to oscillator error. However, it has other problems associated with its use, e.g., it is not really a metric under the standard mathematical definition of this term. Consequently, it was decided that a different approach was needed to remove these errors. This led to the creation of a new comparison algorithm, one that can remove the effects of oscillator errors from comparisons. The algorithm developed is an iterative nonlinear least squares method that attempts to overlay a target sequence on a grid reference subsequence. It is called the

Oscillator Error Correction/d_σ¯ Metric algorithm. This algorithm was implemented and tested. It always converged when using it in millions of comparisons that were made.

5.2.2 Methods

The Oscillator Error Correction $/d_{\overline{\sigma}}$ Metric algorithm has four inputs. The first input is denoted by G_r where $G_r = \{(t_{r_i}, f_{r_i}), i = 1, 2, \ldots, N\}$ is the subsequence of grid reference data and where, for each ordered pair in G_r, $t_{r_i}$is the time f_{r_i} was measured. The second input is the target sequence, $T = \left\{\left(t_{T_k}^{(0)}, f_{T_k}^{(0)}\right), k = 1, 2, \ldots, K\right\}$ where, for each ordered pair in T, t_{T_k} is the time that the frequency f_{T_k} was measured. Note that both the time, $t_{T_k}\mathrm{t}_{\mathrm{T_j}}$, and the frequency, $f_{T_k}\mathrm{f}_{\mathrm{T_j}}$, values will change from iteration to iteration so the superscript (0) indicates their starting value. The third input is the mean of the target frequency values, $\overline{M}_T^{(0)}$. It is computed using

$$\overline{M}_T^{(0)} = \frac{1}{K} \sum_{k=1}^{K} f_{T_k}^{(0)} \, . \qquad (5\text{-}1)$$

Again, since this mean value will change for each iteration, the superscript (0) indicates the starting, i.e., iteration 0, value. The fourth and final input value is $maxiterations$ which gives the maximum number of iterations.

Step 1. Set $deltafrequency = 1$, and $l = 1$.

Step 2. If $l = maxiterations$ exit with failure.

Step 3. Find $\left\{\left(t_{T_k}^{(l-1)}, f_{r_k}^{(l-1)}\right), k = 1, 2, \ldots, K\right\}$, i.e., the sequence of grid reference frequencies that occur at exactly the times given in the target sequence, using

$$t_2 = t_{r_i} \tag{5-2}$$

$$f_2 = f_{r_i} \tag{5-3}$$

$$t_1 = t_{r_i - 1} \tag{5-4}$$

$$f_1 = f_{r_i - 1} \tag{5-5}$$

$$t = t_{T_j}^{(l-1)} \tag{5-6}$$

$$f_{r_j}^{(l-1)} = \frac{f_2 - f_1}{t_2 - t_1}\,(t - t_1) + f_1 \tag{5-7}$$

where, $t_{r_i - 1} < t_{T_j}^{(l-1)} \leq t_{r_i}$.

Step 4. If $deltafrequency \leq 1.0 \times 10^{-13}$ then go to Step 10.

Step 5. Compute the mean of $\left\{\left(t_{T_k}^{(l-1)},\ f_{r_k}^{(l-1)}\right),\ k = 1, 2, \ldots, K\right\}$ using

$$\bar{M}_G^{(l-1)} = \frac{1}{K} \sum_{k=1}^{K} f_{r_k}^{(l-1)} \tag{5-8}$$

Step 6. Let $factor = \frac{\bar{M}_G^{(l-1)}}{\bar{M}_T^{(l-1)}}$ and let

$$\left\{\left(t_{T_k}^{(l)}, f_{T_k}^{(l)}\right), k = 1, 2, \dots, K\right\} = \left\{\left(t_{T_k}^{(l-1)}/factor,\ factor \times f_{T_k}^{(l-1)}\right), k = 1, 2, \dots, K\right\} \quad (5\text{-}9)$$

Step 7. Let $\bar{M}_T^{(l)} = factor \times \bar{M}_T^{(l-1)}$

Step 8. Let $deltafrequency = abs\left(\bar{M}_T^{(l)} - \bar{M}_T^{(l-1)}\right)$ and $l = l + 1$.

Step 9. Go to Step 2.

Step 10. Compute the distance between the target and grid subsequence using the $d_{\bar{\rho}}$ metric

$$d_{\bar{\rho}}\left(T^{(l-1)}, G_r^{(l-1)}\right) = \mathrm{sqrt}\left(\frac{1}{K}\sum_{k=1}^{K}\left(f_{r_k}^{(l-1)} - f_{T_k}^{(l-1)}\right)^2\right) \quad (5\text{-}10)$$

where both $f_{r_k}^{(l-1)}\mathrm{f}_{\mathrm{r_m}}^{(\mathrm{l}\text{-}1)}$ and $f_{T_k}^{(l-1)}\mathrm{f}_{\mathrm{T_m}}^{(\mathrm{l}\text{-}1)}$ occur at time $t_{T_k}^{(l-1)}\mathrm{t}_{\mathrm{T_m}}^{(\mathrm{l}\text{-}1)}$ for $k = 1, 2, \dots, K$.

5.2.3 Results

Consider the following experiment. The output of an analog device, i.e., the Electric Network Signal Conversion (ENSC) Box, that reduces the voltage level of a signal coming out of a standard wall power outlet and inputs it into two different sound cards. One of these is an external sound card. This external device is connected to the analysis computer by a USB-2 interface that transfers converted digital data to the computer in real-time. Note, the operation of this external device is based solely on its own internal clock. The second sound card receiving this analog signal is an internal sound card contained in a laptop computer. This sound card also has its own clock and collects data based on the output of its clock. Both cards were setup to collect data at 44,100 samples per second. The times that the two recordings were started was recorded. Both recordings were approximately 2,850 seconds long. After these two data sets were collected, one set from each sound card, a short time DFT based ENF extraction method similar to the one described by Cooper [66], i.e., the DFTVV algorithm, was applied to each recording. The window size used was 2 seconds and the hop size was 0.4 seconds. The important point here is that both recordings were processed in exactly the same manner to obtain each recording's ENF sequence.

Plots of the ENF estimates obtained from these two recordings are shown in Figure 5-1. As will be observed, the sequences of frequency estimates do not perfectly overlay onto one another as one might expect that they should. Rather they are separated by what appears to be a constant frequency offset, the same type of offset that was reported in the

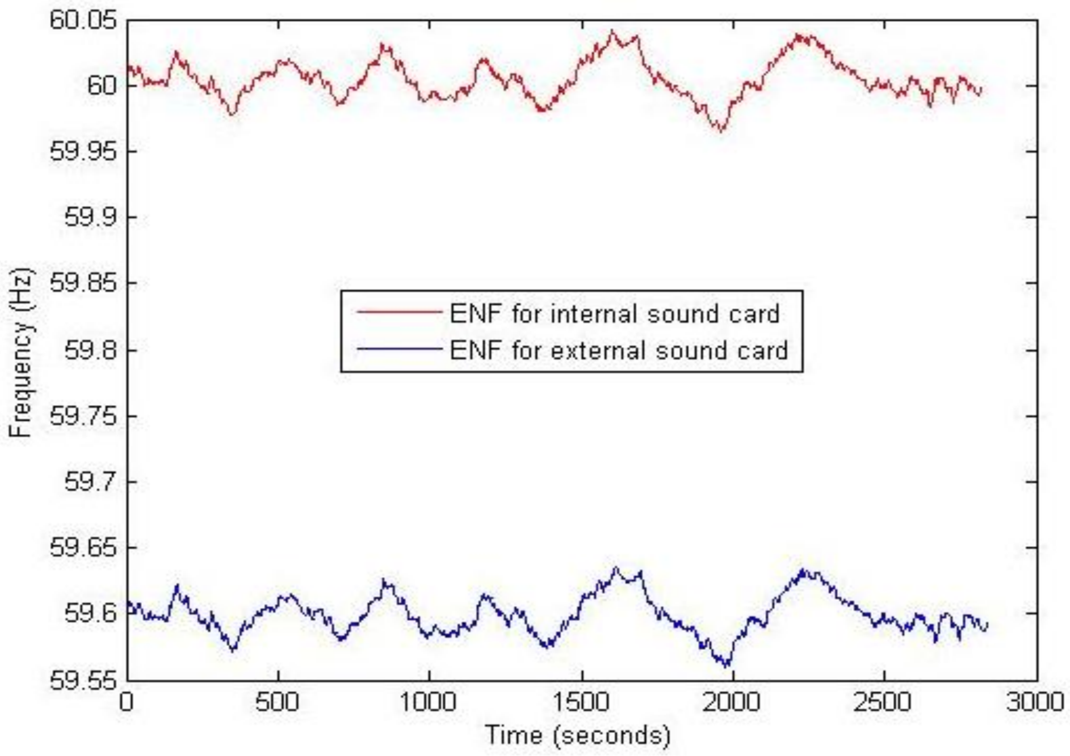


Figure 5-1 A plot of the ENF sequences computed using the same ENF extraction algorithm applied to two recordings made from the same grid sinusoidal voltage input signal

studies reported in [66-69]. In this controlled experiment the only possible cause for the observed offset is that the two sound card clocks and, hence, the two sound card oscillators are running at slightly different speeds.

These extracted ENF sequences were then put into the Oscillator Error Correction/d_σ¯ Metric algorithm. The sequence computed from the recording made from the data collected by the internal sound card was input as being the grid reference sequence which, of course, means that the ENF sequence computed from the recording made from the external sound card was considered to be the target sequence. As will be remembered, the algorithm attempts to make the target sequence, as perfectly as possible, overlay onto the reference ENF sequence.

The results of the oscillator error correction algorithm processing are shown in Figures 5-2 and 5-3. Figure 5-2 illustrates the frequency correction that was made during the first iteration. Note that the vertical frequency alignment looks correct. However, as will be noted, particularly when one examines the end elements of both ENF sequences, there is a noticeable timing misalignment that still needs correcting. The result of timing

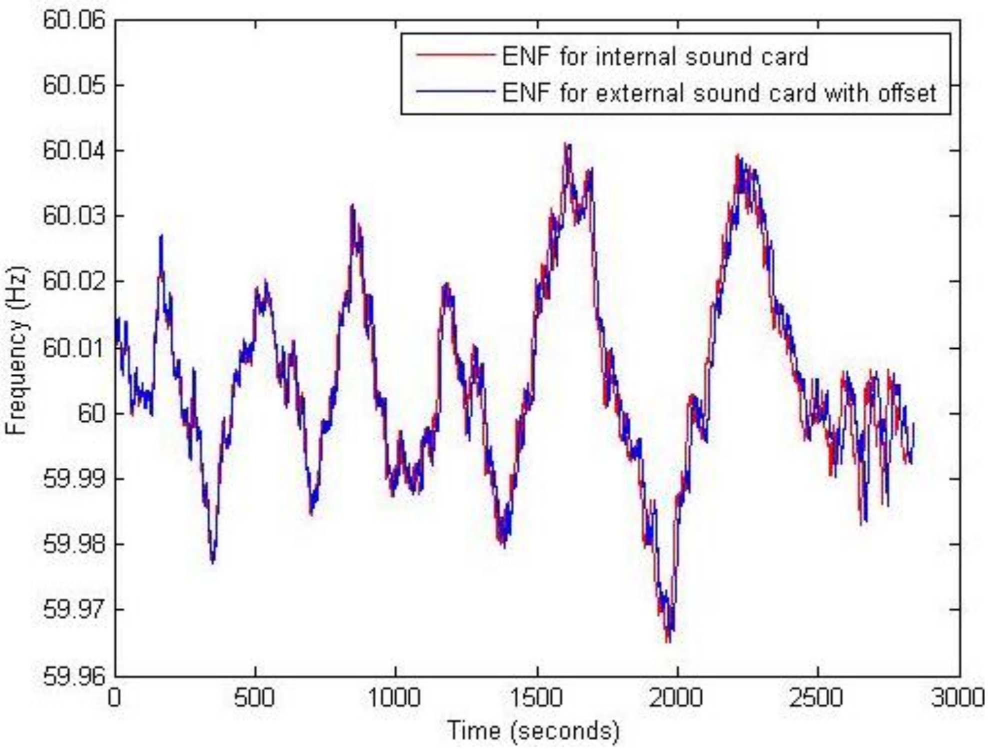


Figure 5-2 A plot of the two ENF sequences after frequency correction was performed by the Oscillator Error Correction/$d_{\bar{\sigma}}$ Metric algorithm at the end of its first iteration

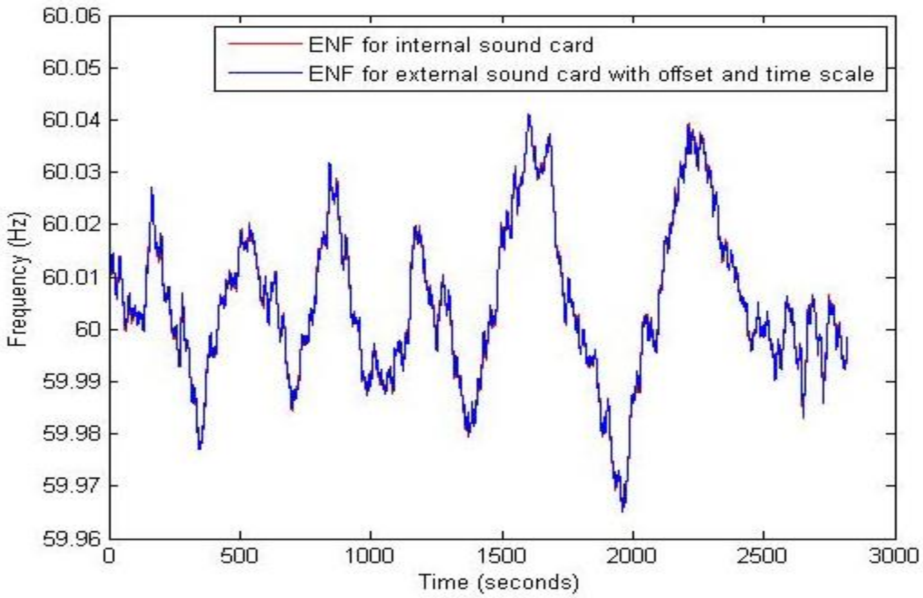


Figure 5-3 A plot of the two ENF sequences after the time correction has been applied by the Oscillator Error Correction/$d_{\bar{\sigma}}$ Metric algorithm at the end of its first iteration

correction that was made to the target sequence during the first iteration is shown in Figure 5-3. As will be noted in this figure the alignment between the two ENF sequences appears to be very good. The oscillator error correction algorithm took a total of three iterations to reach convergence with most of the corrections to both target frequencies and target times being made during the first iteration.

Admittedly, this one processing example does not conclusively prove that the Oscillator Error Correction/d_σ¯ Metric algorithm is robust since it is an iterative algorithm and one must be concerned about its convergence properties. However, it should be noted that this algorithm is used extensively in the study to be reported in the next section. This use involved the comparisons of several million target and grid reference sequences. On these comparisons the algorithm always converged. This should help allay any concerns about this algorithm convergence on real problems.

5.3 Phenomenology of Electric Network Frequency

5.3.1 Overview

The phenomenology by which sinusoidal grid voltage signals find their way onto recordings made using battery powered digital recorders is not well understood. A better understanding of this phenomenology might prove useful in creating yet other ways to forensically verify the authenticity of digital recordings. Consequently, a study was conducted that used varying strengths of electric fields, varying strengths of magnetic fields, and sound shielding to determine which had effects on the strength or signal-to-noise ratio (SNR) of the grid sinusoidal voltage signal that appear in recordings. Varying the sound level involved using sound silencing material to build a "soundproof" enclosure. Tests were conducted outside the sound enclosure and then inside the enclosures. Varying the electric field involved changing the charge across two parallel electrical plates. Varying the magnetic field involved changing the current that flows through a coil. Measuring instruments were used to gauge the intensity of the electric field and the magnetic field. A number of different battery powered recording devices were used in this study. To assure that the signal being observed was actually a grid sinusoidal voltage signal, frequencies were estimated from the recording, and the frequencies were compared with Frequency Disturbance Recorder (FDR) data. Interestingly, these studies suggest that the primary source for ENF in recordings appears to be an audible hum that seemingly comes from grid powered devices.

5.3.2 Methods

One continuing problem in the science of ENF analysis is that how sinusoidal grid signals find their way onto digital recordings is not well understood. Currently, there are very few experiments and papers on studying the source of ENF signals existing in digital recordings. In addition, it is unclear whether or not there are detectable ENF traces in battery-powered digital audio recordings. This study investigates the source of the ENF in battery-powered digital recordings, and reveals that the ENF in these recordings may not be caused by low-frequency electromagnetic field induction, but rather by low-frequency audible hum. A number of experiments were performed to explore the possible sources of ENF in battery-powered digital recordings. In these experiments, the electric and magnetic field strengths in different locations were measured, and the results of corresponding ENF extraction were analyzed. Understanding this underlying phenomenon is critical in verifying the validity of ENF-based techniques.

Recent papers [66, 70-71] on the ENF Criterion all mention that the ENF exists not only in digital recordings produced by devices powered by the grid, but also in recordings made using battery-powered devices located in proximity to other mains-powered equipment or transmission cables. However, while it is straightforward to understand how ENF exists in recordings made by mains-powered recorders, very little research has been done to study how ENF signals couple into battery-powered recordings. Brixen [72] performed several experiments using five different recording devices under different strengths of magnetic field trying to quantify the effect of magnetic field on ENF. His conclusion was that four out of the five recorders with electret microphones did not capture ENF in the recordings. Only one recorder with a dynamic microphone captured the ENF, however this was not dependent on the magnetic field strength. He also found out that under certain situations, the recordings will contain strong harmonics of the power grid frequency [72]. This leads to the question, is the magnetic field the dominant factor in embedding ENF in recordings?

In the experiments studying the ENF source for battery-powered recordings, the procedure of confirming whether a recording has ENF or not is as follows:

1) Digital recordings are made in a non-laboratory environment, i.e., one without measures in place to attenuate electromagnetic radiation or sound waves.
2) The frequency spectrum of recordings is analyzed using an FFT (a 10-second moving window is used, then the results are summed and averaged). If a strong 60 Hz peak or a peak at a harmonic of this frequency exists, then the recordings may have ENF.
3) Extract the suspicious 60 Hz or its harmonic components using a short-time Fourier Transform (STFT) and compare the result with reference measurement data to confirm.

The experiments were carried out as follows. First, a battery-powered iPhone 4 was used to make audio recordings, which were then examined for the presence of ENF. Since ENF was observed in those recordings, the possible sources for the ENF coupling mechanism were needed to be explored. Previously, it was thought that low-frequency electromagnetic field induction was the most dominant factor in such coupling. In order to study the effect of the magnetic field on ENF in recordings, an experiment was designed using an AC voltage source (Agilent 6811B), an external resistor and a coil of wire. The coil that carries an AC current can generate a magnetic field whose frequency is the same as the AC voltage. Then the strength of the magnetic field was measured by a Gauss Meter (range 1 nT - 1999 nT), while the electric field was held constant (about 1 V/m) during the experiment. Recordings were then made near the coil. By varying the frequency of the AC voltage, the effect of the frequency components on the ENF coupling could be examined.

In addition, the effect of the electric field on ENF in digital recordings was analyzed. In these experiments, a parallel-plate device that was powered by a DC voltage source was used to generate specific electric fields. The strength of the electric field can also be measured by the meter and audio recordings were made in the middle of the parallel plates. By varying the voltage amplitude, the effect of the electric field could be examined.

Through analysis of frequency spectrums of a number of recordings shows that the 120 Hz component is the strongest, while 60 Hz and 180 Hz components are very weak which are usually buried with other frequency components in the frequency spectrum. However, in the power grid it is well known that the 60 Hz component is the strongest, and the third harmonic (180 Hz) is larger than the usually very small second harmonic (120 Hz). Thus, it is highly possible that the 120 Hz harmonic in recordings is the audible hum, which is typically generated from the mechanical vibration of the transformer core from the electromagnetic force produced by the current in transformer windings. The frequency of such audio hum is twice that of the 60 Hz fundamental frequency; here in the U.S. it is 120 Hz.

The focus is then turned to study the effect of audible hum on ENF coupling into battery-powered recordings. Three recording devices (OLYMPUS WS-600S, OLYMPUS WS-300M, and an iPhone 4) were used in the following experiments. A number of recordings were made indoors and outdoors where audible hum was accessible while the magnetic field strength was low to find out whether ENF exists. In addition, several recordings were made in an apartment when the air conditioner fan was on or off to show the effect of audible hum from the fan.

In order to confirm the ENF is caused by audible hum, a soundproofing material was used to reduce the transmission of sound. The rubber-based material is often used for soundproofing multi-family residences, studios, and theatres. For the experiments, the soundproof cylinder was made of the 3.2-mm thick material. The cylinder has four layers with length of two feet and diameter of four inches as shown in Figure 5-4. The sound intensity was controlled by having a recorder either outside of the cylinder or in the cylinder. By putting the recorder inside of the cylinder while the cell phone was outside of the cylinder the ENF components in both recordings could be compared.

During the experiments, although there were sometimes strong 120 Hz components in the recordings, they were not actual ENF as compared with the reference data. After analyzing the experiments, it was found that the digital recorder and the soundproof cylinder were placed on the same table near a computer. Since the computer was always vibrating and producing a noise, further analysis was needed.

In order to figure out whether the 120 Hz component was from the vibration of the computer or not, the digital recorder was put on the table near the computer and two recordings were made. The computer was running during one recording and turned off during the other. The amplitude of the 120 Hz component from these two recordings was then compared.

Figure 5-4 Cylinder made of the soundproofing material for the experiments

5.3.3 Results

An iPhone 4 was used as recording device. During the ENF extraction work, the harmonics (i.e. 120 Hz, 180 Hz, 240 Hz, etc.) of 60 Hz (electric network frequency) were observed to exist in the audio recordings of battery-powered cell phones, which were made in an environment of very low electric and magnetic field strengths (as shown in Figure 5-5). The 120 Hz component was extracted from a typical audio recording (the extracted frequency was then divided by two) and compared with extraction of the 60 Hz voltage signal collected directly from electrical outlets and FDR measurements for the same time period (as shown in Figure 5-6). From the figure it can be seen that the three match well. Similar matching levels for 240 Hz and 360 Hz components were also confirmed. Since the iPhone 4 was used as the recording device, and the frequency

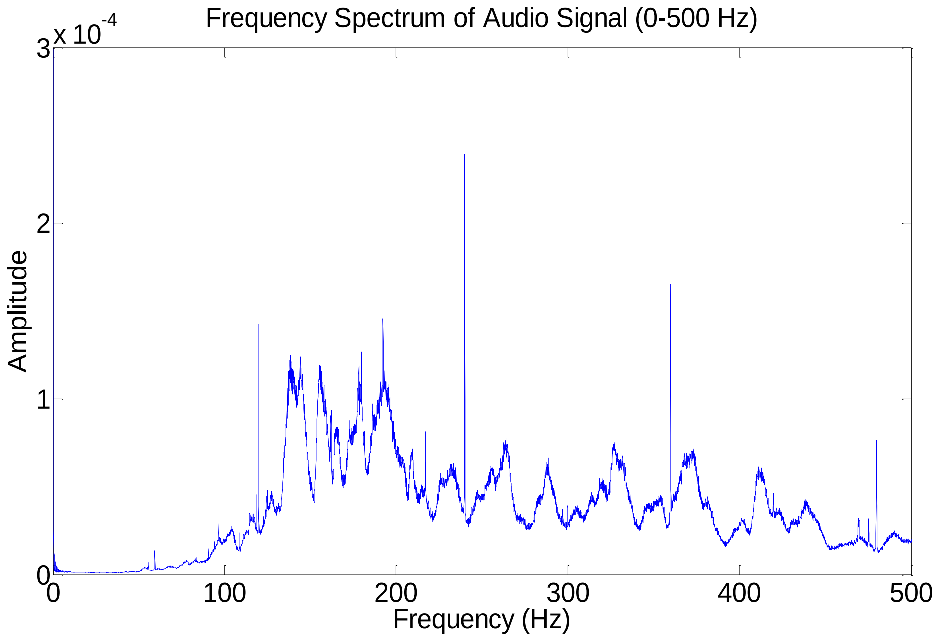


Figure 5-5 Frequency spectrum of an audio recording made by an iPhone 4

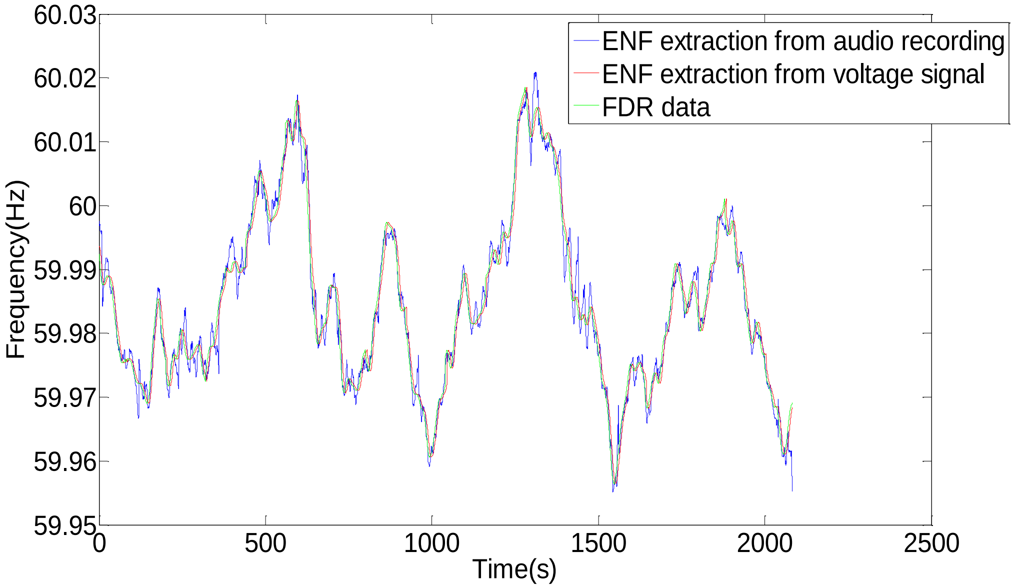


Figure 5-6 A comparison of ENF extracted from audio recording and voltage signal versus ENF measured by a Frequency Disturbance Recorder (FDR)

response of the built-in microphone is in the range of 100 Hz-20 kHz, only power grid harmonics higher than 100 Hz could be captured.

The test setup for the study of magnetic and electric fields is shown in Figure 5-7. The results show that the generated magnetic fields (such as 125 Hz, 150 Hz, 200 Hz) did not leave any trace in the audio recordings though the magnetic field was considered very strong (more than 1900 nT). Considering that ENF can be detected in audio recordings made in environment whose magnetic field is very weak (as low as 1 nT), the test shows that the magnetic field may not be a dominant source of ENF in recordings. During experiments for the study of electric fields, various voltage amplitudes (from 1 V/m to 50

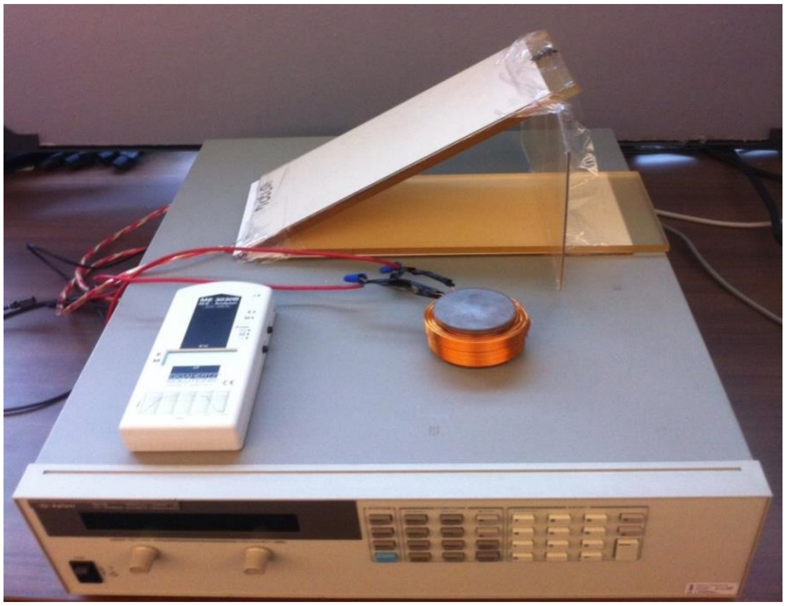
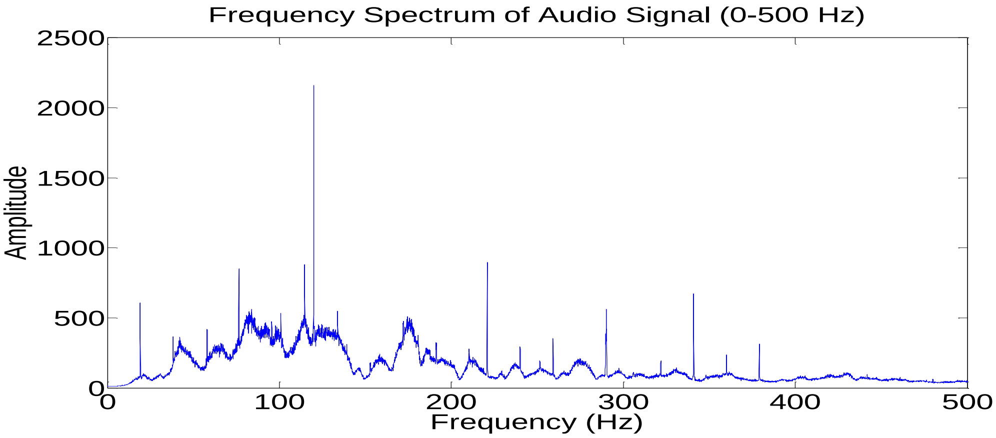


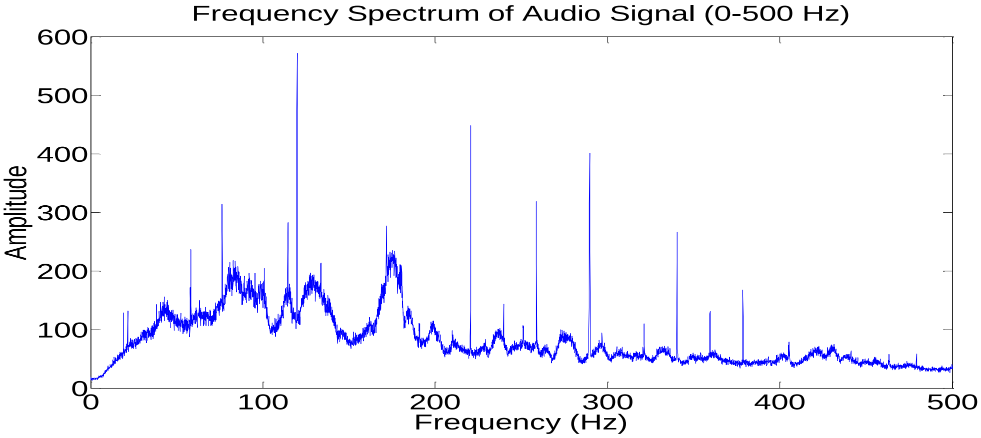
Figure 5-7 Test setup for effect of electric and magnetic fields on ENF coupling into battery-powered digital recordings

V/m) were chosen and corresponding recordings were made. The tests reveal that the electric field also has no influence on the existence of ENF in audio recordings.

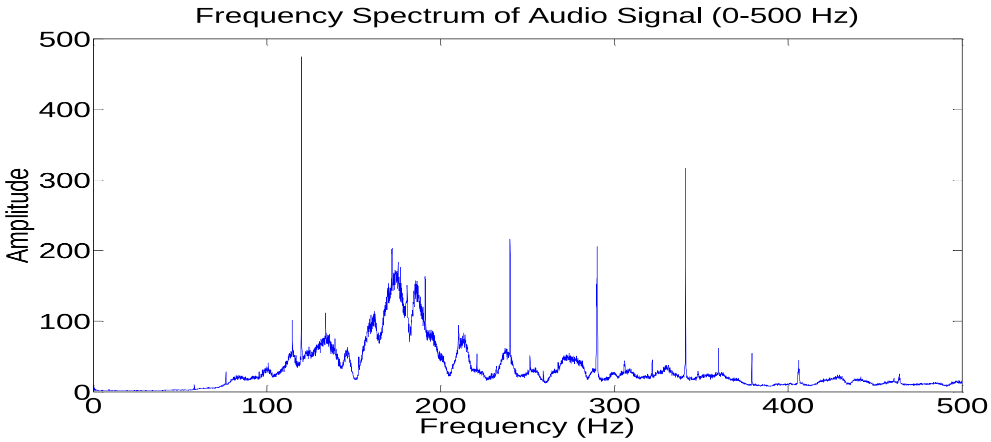
The electrical hum is difficult to escape indoors because most buildings contain many electrical devices with transformers and/or motors. The frequency spectrum of indoor recordings is shown in Figure 5-8 and the extracted frequency with reference data is shown in Figure 5-9. In addition, the extracted ENF from outdoor recordings where audible hum is accessible with measurement data is shown in Figure 5-10. Analysis of a number of these indoor/outdoor recordings revealed that most of them captured actual ENF.



(a) Frequency spectrum of a recording made by OLYMPUS WS-600S



(b) Frequency spectrum of a recording made by OLYMPUS WS-300M



(c) Frequency spectrum of a recording made by iPhone 4

Figure 5-8 Frequency spectrum of recordings made by different recorders

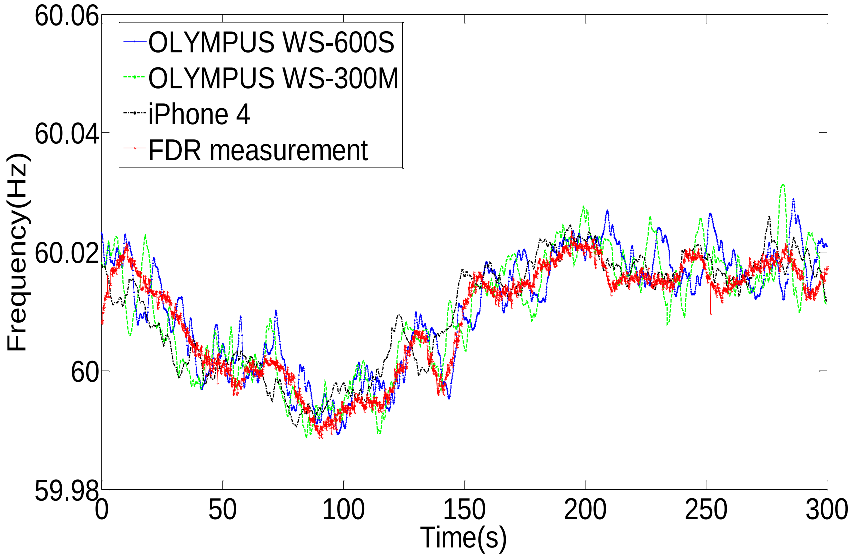


Figure 5-9 A comparison of ENF from different indoor recordings versus ENF measured by a Frequency Disturbance Recorder (FDR)

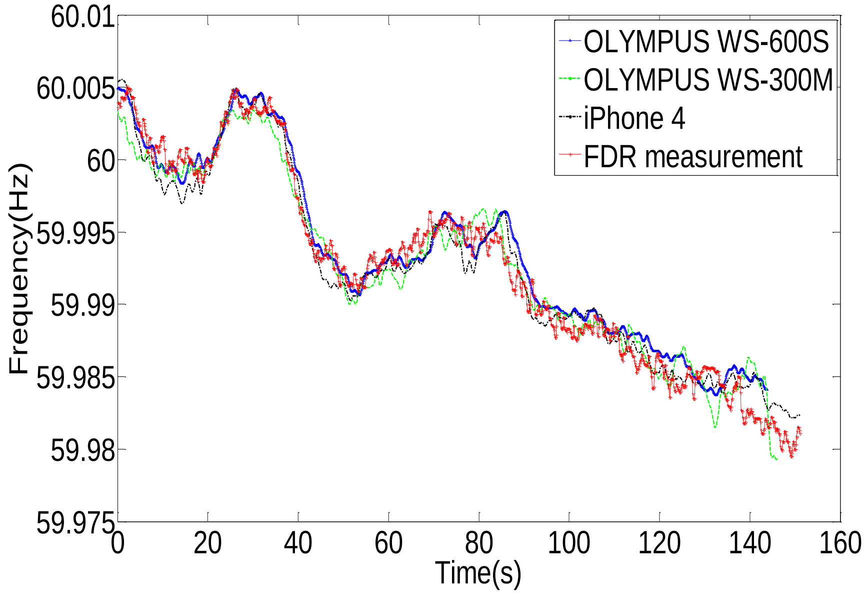


Figure 5-10 A comparison of ENF from different outdoor recordings versus ENF measured by a Frequency Disturbance Recorder (FDR)

For the experiments made in the apartment when the air conditioner fan was on or off, the Gauss Meter was used to make sure that the electric and magnetic fields didn't change during the experiments. When the fan was off, there were almost no ENF traces in recordings, as shown in Figure 5-11; when the fan was on, a strong 120 Hz component and its harmonics existed, which are shown in Figure 5-12. Figure 5-13 shows the comparison between extracted ENF from recordings and reference data from FDR measurements. From the figure, it can be seen that the two match well, which indicates the audible hum generated by the fan leaves ENF in the recordings. Note that after a while, there is a small drift between the two in Figure 5-13. This is probably due to the oscillation error of the recording device; the overall shape is still very close to the FDR measurements.

The attenuation effectiveness of the soundproofing material is shown in Table 5-1. Since the sensitivity of the two recorders is not the same, the results need to be calibrated by multiplying by a scalar, which can be calculated through recording under the same

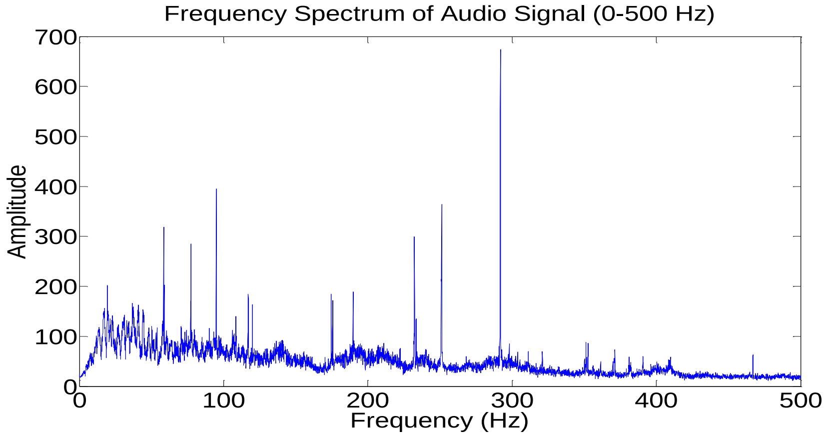


Figure 5-11 Recording made when the fan is off

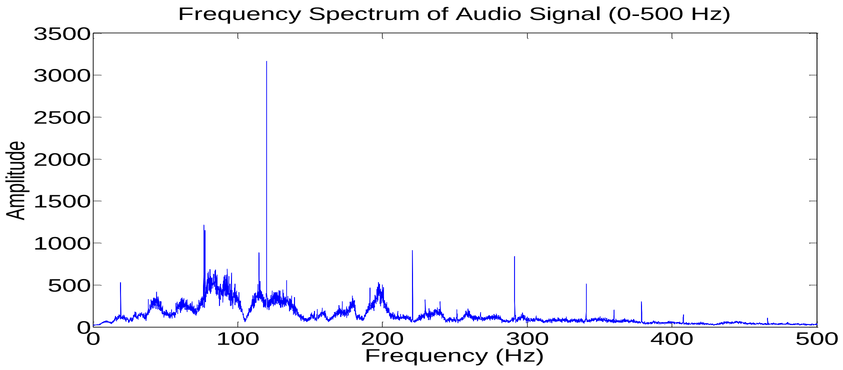


Figure 5-12 Recording made when the fan is on.

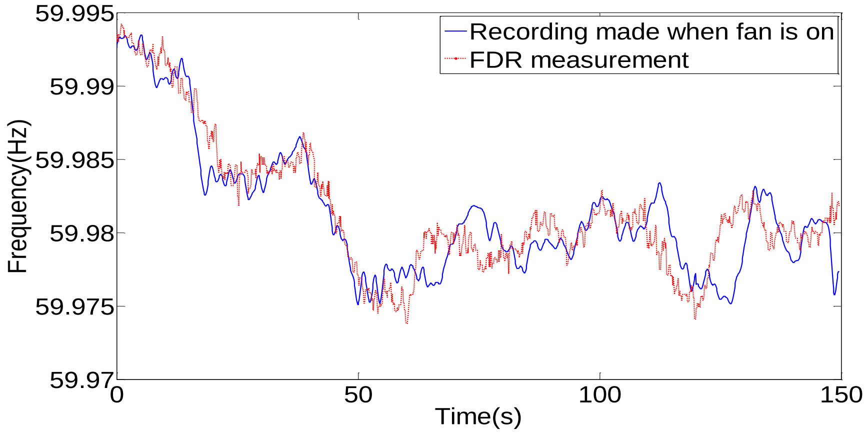


Figure 5-13 A comparison of ENF from audio recordings when fan is on versus ENF measured by a Frequency Disturbance Recorder (FDR)

Table 5-1 Soundproof performance of the material

2mm (Single Layer)	Freq. Hz	125	250	500	1000	2000	4000
	(dB)	10	8	10	11	12	14
2mm (Double Layer)	Freq. Hz	125	250	500	1000	2000	4000
	(dB)	11	12	14	17	21	23
3.2mm (Single Layer)	Freq. Hz	125	250	500	1000	2000	4000
	(dB)	13	13	16	19	21	23
6.4mm (Single Layer)	Freq. Hz	125	250	500	1000	2000	4000
	(dB)	7	5	3	4	5	5

conditions. The results of ENF in recordings made inside/outside the soundproofing material are shown in Figure 5-14. It shows that the 120 Hz component decreased significantly when the recorder was inside the cylinder and it confirms that the 120 Hz component is mainly audible hum.

During the experiments, it was observed that occasionally, strong 120 Hz components existed in the recordings that were not from the power grid. Figure 5-15 shows a comparison between the extracted 120 Hz component (which was divided by 2) and the corresponding FDR measurements. From the figure, it is obvious that this 120 Hz component is not an actual ENF signal.

In order to find the source of these 120 Hz components, two separate recordings were made when the computer was turned on and off. Each of the recordings was three minutes long. The FFT analysis results of the two are shown in Figure 5-16 and 5-17.

In Figure 5-16，there is a strong 120 Hz component when the computer was turned on. However, the 120 Hz component is much smaller and merged with other noises in Figure 5-17 when the computer was turned off. Therefore, the 120Hz component is mainly from the vibration of the computer, as its hard drive rotates at 7200 RPM which equates to 120 Hz. Attention needs to be paid to this phenomenon because an apparent 60 Hz signal and its harmonics are not necessarily due to the power grid frequency, but instead result from other sources.

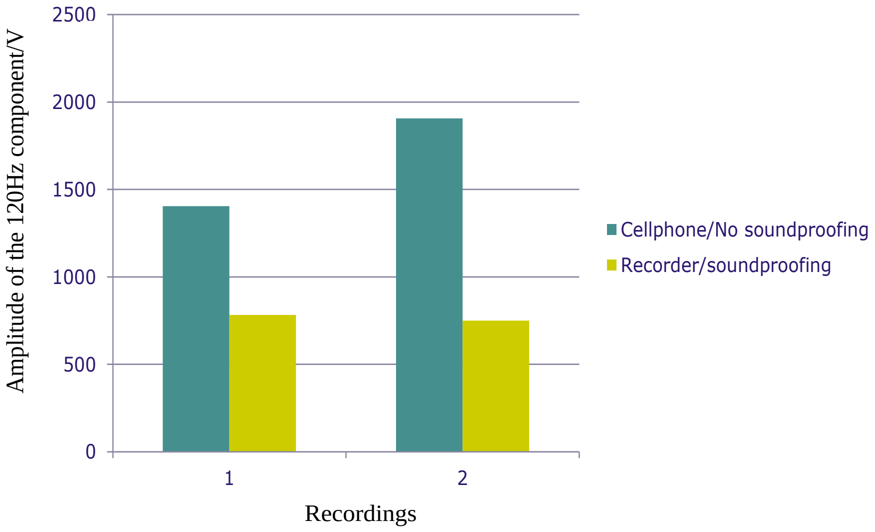


Figure 5-14 A comparison of amplitude of the 120 Hz component from audio recordings made inside/outside soundproof materials

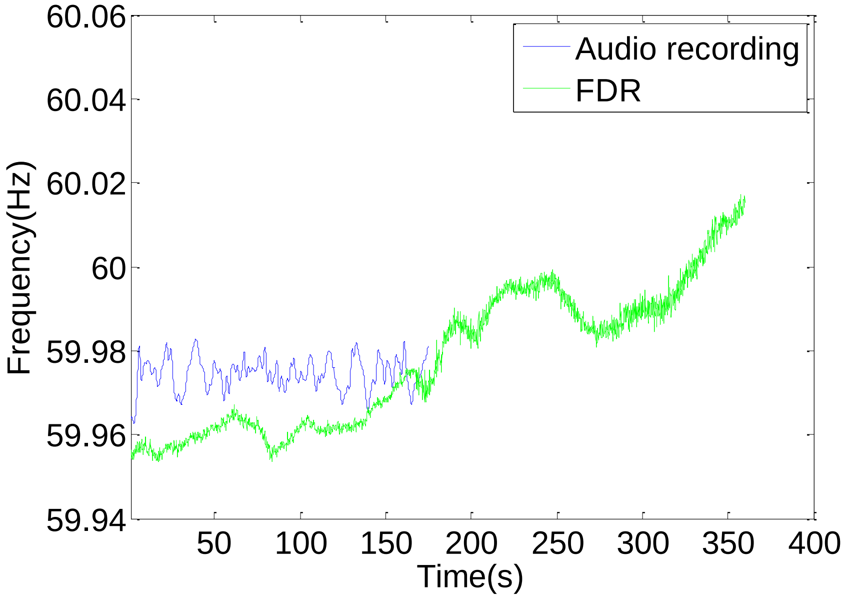


Figure 5-15 A comparison of ENF from a recording made near a computer versus ENF measured by a Frequency Disturbance Recorder (FDR)

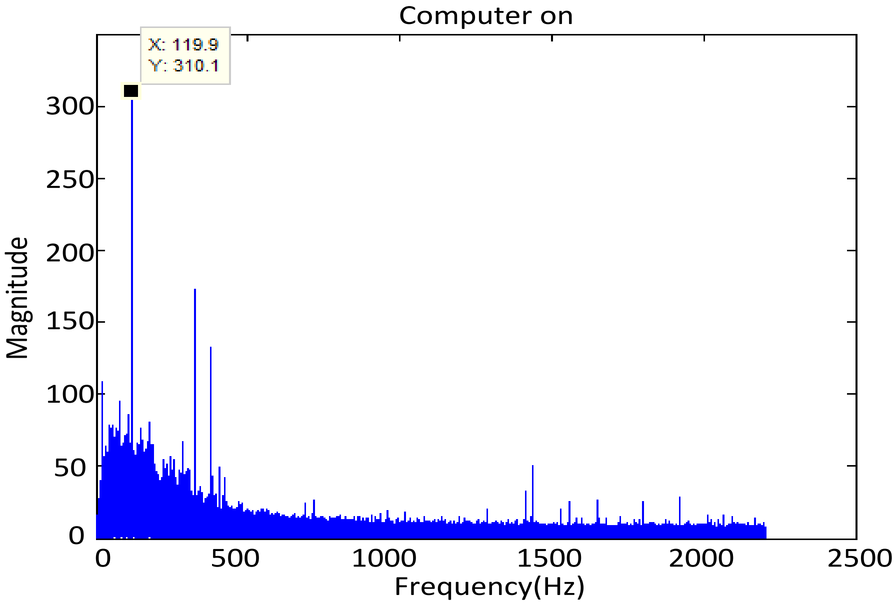


Figure 5-16 An FFT analysis of a recording made with computer on

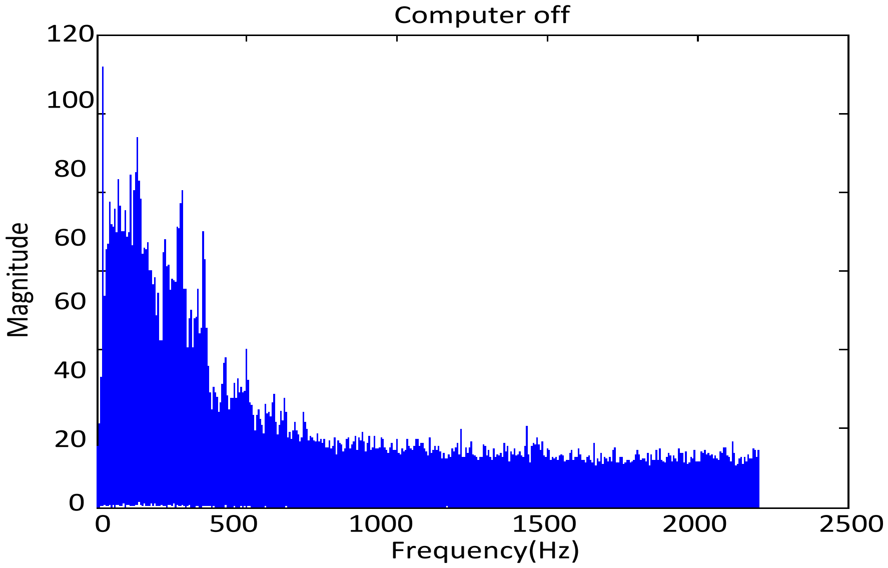


Figure 5-17 An FFT analysis of a recording made with computer off

5.4 Tampering Detection in Digital Recordings

5.4.1 Overview

Since this is a rather new area of research with few papers on the topic, a good starting point for this study seemed to be to determine which parameters to use, i.e., parameters that vary when insertions or deletions occur. The results of this study seemingly indicated that both frequency and phase angle usually change values at deletion points and at the start and stop points of an insertion. This then led to the creation of an improved algorithm for doing this type of analysis. This Discrete Fourier Transform (DFT) based algorithm uses both frequency and phase angle to determine whether tampering has occurred in a recording. It is believed that the use of both frequency magnitude and phase angle increases the "detectability" of both insertions and deletions.

The tampering detection algorithm created has a frequency and phase angle extraction step and a target sequence/grid subsequence comparison step. However, since both frequency and phase angle information are used, the grid reference database must contain not only frequency data but phase angle data as well. This means that the comparison methods used to compare target data to grid subsequence data must involve both frequency and phase angle. This resulting algorithm underwent rather extensive testing that included different lengths of insertions and deletions and with variations in both the frequency and phase angle as the insertion and deletion points. It is hoped that this work, over time, will lead to a greater theoretical understanding of the tampering detection problem. It certainly proved that both frequency and phase angle must be included in any detectability study.

5.4.2 Methods

An important task for forensic authentication is to determine whether someone has tampered with a recording, i.e., inserted sections into the recording or removed sections that were in the original recording. Previous work on tampering detection has involved looking for discontinuities in either the frequency or phase angle [62] that has been extracted from the digital recordings. However, using only frequency or phase angle to detect tampering may not be sufficient. As for using only frequency, frequency resolution depends on the extraction algorithm being used and its input parameters.

As for looking for discontinuities in the phase angle, it is well known that there are occasional, sudden phase shifts that occur in the grid's sinusoidal voltage signal. These shifts are caused by a variety of different disturbances, such as generation trip and transmission line trip. Such phase angle shifts are very similar to those created by either insertions or deletions. In order to distinguish between a sudden shift caused by a disturbance from a shift that is caused by tampering, a grid phase angle reference database is needed. Having this phase angle reference database should eliminate the many false detections. Consequently, the idea behind the new algorithm is to extract phase from audio recordings using DFT-based methods and attempt to match this phase data against a reference phase sequence. The starting time for this reference phase sequence is given by matching the target frequency sequence to a frequency subsequence of the grid reference database.

Figure 5-18 shows a plot of the phase that has been extracted from a recording and the phase of the grid as measured by a Frequency Disturbance Recorder (FDR). This plot clearly shows how well phase angles extracted from a recording match the phase angles in the phase reference sequence. The quality of the match seems very good, although a small drift is observed. Thus in the new algorithm, both frequency and phase angle data estimates are made from the recording and the result is compared to a grid frequency and phase angle reference database. It is believed that this provides a more reliable detection.

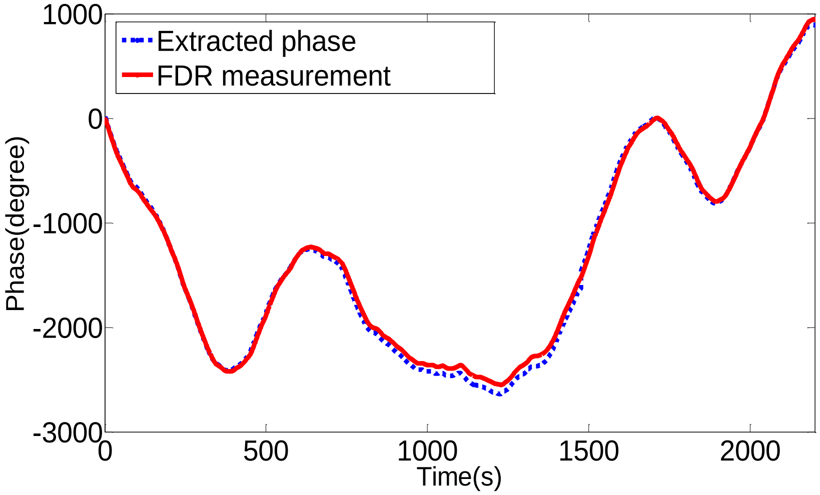


Figure 5-18 A comparison of phase angle estimated from a recording versus the phase angle of the grid as measured by a Frequency Disturbance Recorder (FDR)

As for the possible ways that the ENF can be extracted from recordings, there are currently three general methods: 1) time/frequency domain analysis using spectrograms, 2) frequency domain analysis based on selecting the maximal magnitude of a series of power spectra calculated from consecutive time segments of data, and 3) time-domain analysis based on zero-crossing measurements of a band-pass filtered signal [66]. A brief comparison of these three methods can be found in [59].

The method used in the new algorithm is a short-time Fourier transform (STFT)-based approach. A short-time Fourier transform (STFT) is a windowed Fourier transform where a signal is truncated by windowing function. This windowing function is moved through the signal in a manner that is specified by the hop size. Given a signal $x(n)$, $n = 1, 2, \ldots, N$, the STFT is calculated by an M-point discrete Fourier transform (DFT) [29]

$$\mathrm{STFT}_x(m,k) = \sum_{n=1}^{M} x(n)w(n-mP)e^{-\frac{j2\pi}{Mnk}} \tag{5-11}$$

where $m = 1, 2, \ldots, (N-M)/P$, $k = 1, 2, \ldots, M$, w is the window function and P is the size in each step, the so-called hop size.

Since the grid sinusoidal voltage signal always occurs within a certain frequency range, to reduce the computation burden, k in Equation (5-11) is constrained to some limited number of frequency bins according to a pre-set frequency range of interest, for example $[f_1, f_2]$. Only the frequency spectrum corresponding to $[f_1, f_2]$ is calculated. This adjusted STFT is now presented as

$$\mathrm{STFT}_x(m,k) = \sum_{n=1}^{M} x(n)w(n-mP)e^{-\frac{j2\pi}{Mnk}}$$

$$k \in [K_1 \;\; K_2] \tag{5-12}$$

$$K_i = f_i \frac{M}{f_s}, i = 1, 2. \tag{5-13}$$

where f_s is the sampling frequency.

When both time and frequency are discretized, (5-11) becomes a standard Discrete Fourier Transform (DFT). Window size and hop size are two important parameters that determine the length and shift of the selected window function.

Given a digital audio recording, the proposed new tampering detection procedure operates as follows:

1) *Signal pre-processing*: This step includes low-pass filtering, followed by signal decimation or downsampling. Then, a band pass filter is employed to select the frequency components that lie in the frequency range [f_1, f_2] from the decimated signal.

2) *Signal segmentation*: Break down the band pass-filtered signal into a series of overlapping frames according to the length and hop size of the moving window.

3) *Two-step frequency and phase estimation*: A coarse frequency estimation is obtained using the STFT. Then a polynomial interpolation procedure is applied. This procedure increases the frequency and phase resolution of the frequency and phase estimates. The procedure is then repeated for each frame until all the windows have been processed.

4) *Frequency reference database matching*: After the ENF sequence is extracted from the audio file, it is matched against the FNET/GridEye frequency database. Mean square error (MSE) is used to compare the ENF sequence to sub-sequences

of the grid reference frequency database [60]. The MSE is computed using

$$\varepsilon = \log\left(\frac{1}{M}\sum_{i=1}^{M}(ENF(i) - ref(i))^2\right) \tag{5-14}$$

where M is the length of the extracted ENF and *ref* stands for the reference frequency sequence. These distances are stored, and the subsequence of the grid reference frequency database that produces the smallest distance value is considered to be the "matching" sequence.

After matching, both frequency and phase angle are compared with the corresponding frequency and phase reference databases to determine whether any tampering has occurred in the recording.

Generally, the result in the frequency domain of an N-point DFT of a sinusoidal signal $x(n)$ is a series of discrete points $X(k)$ [74], $k = 0, \dots, N-1$.

$$\begin{aligned} X(k) &= \sum_{n=0}^{N-1} x(n) e^{-j\frac{2\pi}{N}nk} \\ &= \frac{A}{2} e^{j\theta} e^{j\frac{\pi(N-1)}{N}(l-k)} \frac{\sin \pi(l-k)}{\sin \pi(l-k)/N} \\ &+ \frac{A}{2} e^{-j\theta} e^{-j\frac{\pi(N-1)}{N}(l+k)} \frac{\sin \pi(l+k)}{\sin \pi(l+k)/N} \end{aligned} \tag{5-15}$$

where A denotes the amplitude, θ is the initial phase, and f_s is the sampling frequency. The real frequency of the signal corresponds to lf_s/N. In (5-15), the first term represents the positive frequency component, while the second term is the negative frequency component. The frequency spectrum can be obtained using the STFT and k_{peak} can be found by locating the largest magnitude sample. Then a fractional term δ ($|\delta| \leq 0.5$) is calculated based on three DFT samples around the peak to refine k_{peak} [74]. The real frequency corresponds to $l = k_{\text{peak}} + \delta$ and in the frequency domain it is always between the largest two bins. Replacing (5-15) with k_{peak} and letting $\alpha = \pi(N\text{-}1)/N$, the three bins around the peak value are obtained

$$\begin{aligned}
X(k_{peak}-1) &= \frac{A}{2}e^{j\theta}e^{j\alpha(\delta+1)}\frac{sin\,\pi(\delta+1)}{sin\,\pi(\delta+1)/N} \\
&\quad + \frac{A}{2}e^{-j\theta}e^{-j\alpha(2k_{peak}+\delta-1)}\frac{sin\,\pi(2k_{peak}+\delta-1)}{sin\,\pi(2k_{peak}+\delta-1)/N} \\
X(k_{peak}) &= \frac{A}{2}e^{j\theta}e^{j\alpha\delta}\frac{sin\,\pi\delta}{sin\,\pi\delta/N} \\
&\quad + \frac{A}{2}e^{-j\theta}e^{-j\alpha(2k_{peak}+\delta)}\frac{sin\,\pi(2k_{peak}+\delta)}{sin\,\pi(2k_{peak}+\delta)/N} \qquad (5\text{-}16) \\
X(k_{peak}+1) &= \frac{A}{2}e^{j\theta}e^{j\alpha(\delta-1)}\frac{sin\,\pi(\delta-1)}{sin\,\pi(\delta-1)/N} \\
&\quad + \frac{A}{2}e^{-j\theta}e^{-j\alpha(2k_{peak}+\delta+1)}\frac{sin\,\pi(2k_{peak}+\delta+1)}{sin\,\pi(2k_{peak}+\delta+1)/N}.
\end{aligned}$$

Since the amplitude of the positive frequency component is usually much larger than that of the negative frequency component [74], the negative frequency component can be neglected. Then the amplitude and phase angle of the signal can be estimated according to the expression of $X(k_{\text{peak}})$ [73].

$$A = \frac{2\pi\delta}{Nsin(\pi\delta)}\left|X(k_{peak})\right|$$

$$\theta = angle\left(X(k_{peak})\right) - \alpha\delta \qquad (6\text{-}17)$$

$$\delta = \mathrm{Re}\left[\frac{X(k_{peak}-1)-X(k_{peak}+1)}{2X(k_{peak})-X(k_{peak}-1)-X(k_{peak}+1)}\right]$$

The coarse frequency is then refined by

$$f = \frac{k_{peak}+\delta}{N} f_s. \tag{6-18}$$

5.4.3 Results

The effectiveness of detecting tampering may be different for frequency and phase angle for different extraction methods and parameter settings. Take the STFT method as an example. If different window sizes (*hop size* = 0.1s, *deletion length* = 30s) are chosen, the results are different. From Figure 5-19 it can be seen that the frequency change is less obvious as window size increases (sometimes a longer window time is needed to achieve acceptable results). Figure 5-20 shows the corresponding phase change. Since in different locations, the initial phase is different, the reference phase should be vertically shifted to match the starting phase if an exact match is needed. From Figure 5-20, unlike the frequency change, the phase change is more obvious as window size increases. Hence, besides using frequency, it is better to utilize the phase angle to detect tampering as well.

Figure 5-21 is the phase angle recorded by one FDR located in Florida when a line trip occurred nearby on 2/26/2008. From the figure, it can be seen that the sudden small phase change caused by disturbances is very similar to that due to tampering of recordings. In these cases, only looking for discontinuity of the phase angle without a phase reference

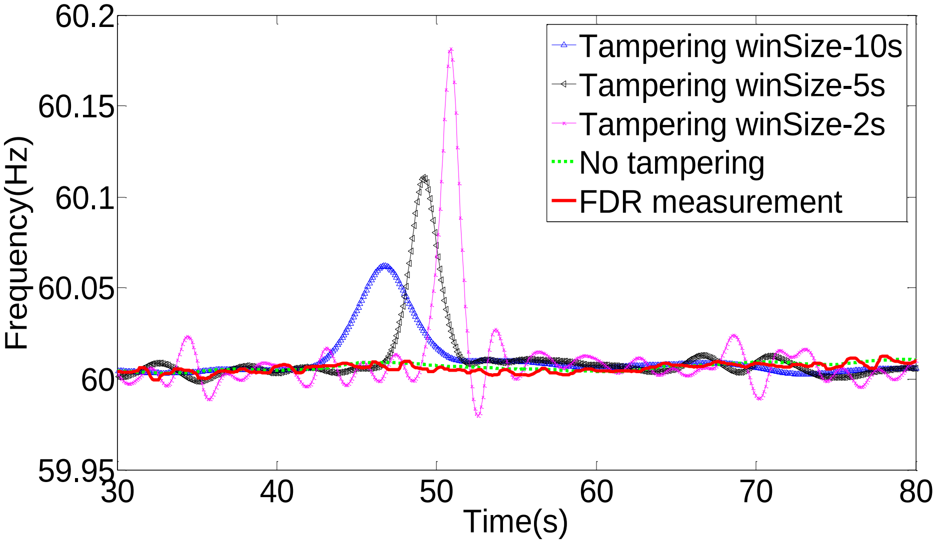


Figure 5-19 Frequency change when 30s data of original recording is deleted with different window sizes

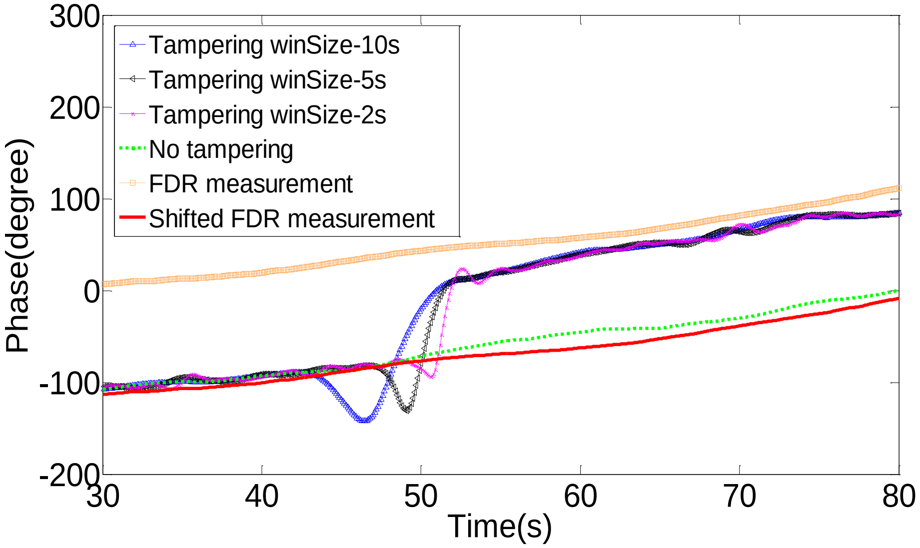


Figure 5-20 Phase change when 30s data of original recording is deleted with different window sizes

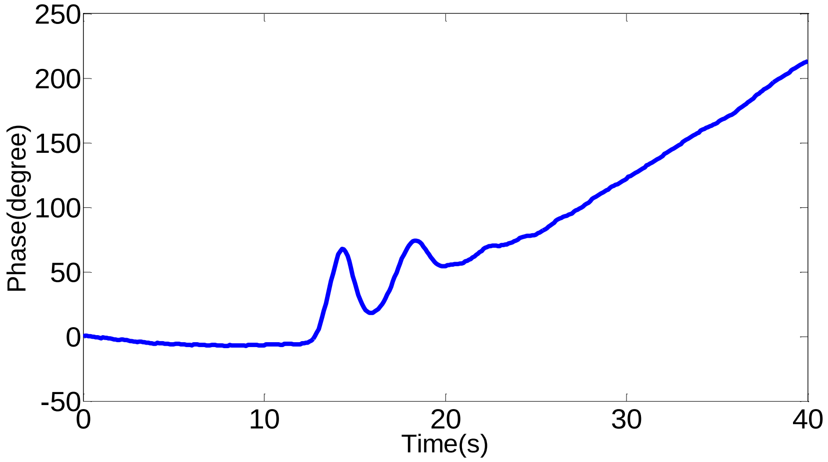


Figure 5-21 Sudden phase angle change due to a disturbance in the power grid

database is highly likely to cause a false detection. Thus, establishing a phase reference database is very helpful for tampering detection.

In this work, two typical types of tampering, deletion and replacement, are considered. The following audio recordings were collected using either an internal or external soundcard. The recordings were then saved as PCM WAV files. A thin band pass filter was designed and the Blackman Window was applied to each frame of the recording. Length of the window is determined by the parameter window size. The original sampling rate is 44,100 Hz and the downsampling rate is 35. Data analysis was performed in MATLAB.

Tampering detection using frequency is illustrated as follows. If a target sequence can match against frequency reference database correctly using MSE or correlation coefficient method [67], tampering can usually be easily detected by comparing the ENF with the reference frequency sequence. If there is a deletion in the recording, one spike corresponding to the deletion point in the ENF could be detected, as shown in Figure 5-22; if it is a replacement, there are two spikes which correspond to the beginning and ending point of the replacement in the ENF, as shown in Figure 5-23.

In practice, people may ask if the recording is tampered, i.e., there are frequency spikes in the extracted ENF, matching methods based on MSE or correlation coefficient may fail. How should the frequency matching be performed in these cases? It is suggested that once a frequency spike has been observed, an extracted ENF section that has no frequency spikes should be used to do the matching, assuming the recording is long enough.

A number of recordings have been analyzed to detect tampering using both frequency and phase, which result in more reliable detection. As an example, here are two cases of tampering detection with different lengths of deletion and replacement employing both frequency and phase.

A. Case 1- Deletion

A portion of the recording is deleted, then its frequency and phase are extracted as illustrated in previous Section.

Figure 5-24 shows the frequency change with different lengths of deletion, and Figure 5-25 shows the corresponding phase change. Besides detecting tampering using the phase, with a phase reference database, it is easy to roughly estimate the length of deletion.

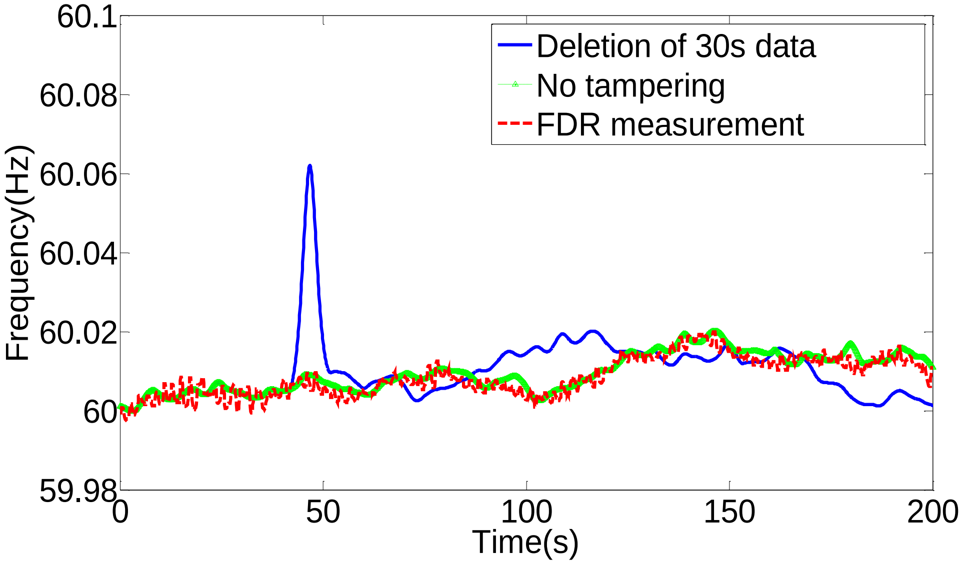


Figure 5-22 Frequency change when 30s data of original recording is deleted

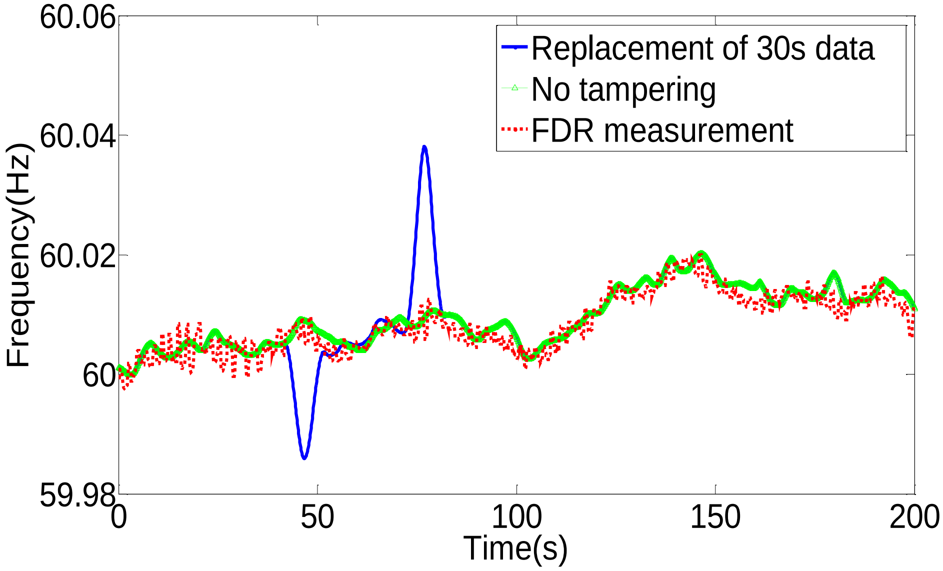


Figure 5-23 Frequency change when 30s data of original recording is replaced

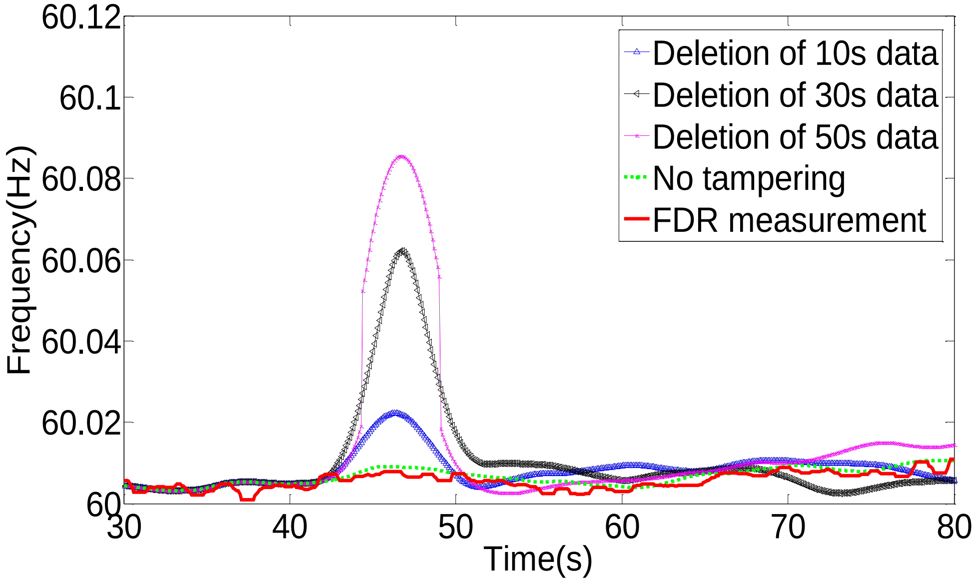


Figure 5-24 Frequency change with different deletion lengths of the original recording

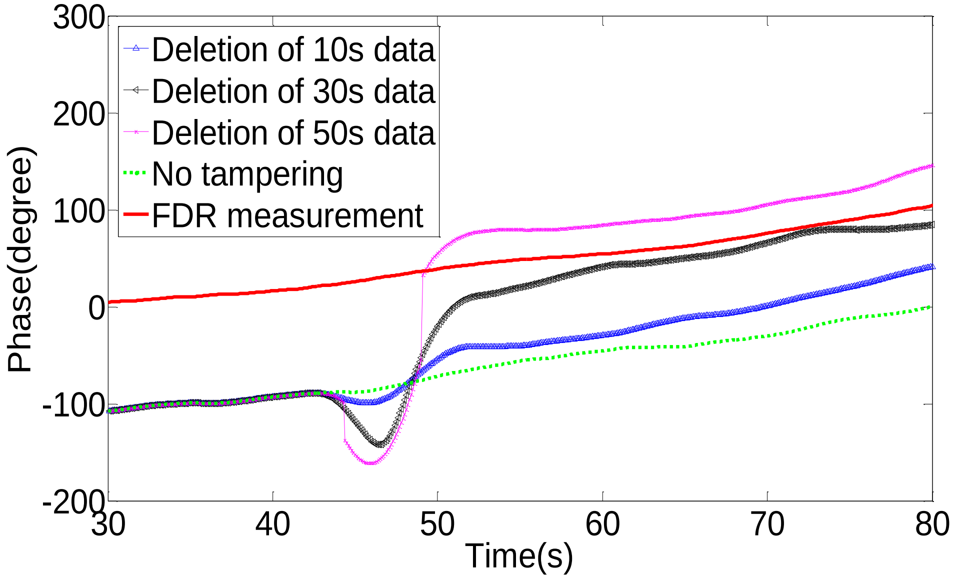


Figure 5-25 Phase change with different deletion lengths of the original recording

As shown in Figure 5-26, a point (52.7 s) right after the abrupt phase change and the point (102.1 s) in reference phase (here "No tampering" phase is used) which has the same phase value are chosen. Considering the phase value of the tampered recording and the reference should be the same after the tampering part, the deletion length can be estimated by measuring the time difference between those two points. It is also possible to estimate the deletion length using frequency with a similar procedure, but it is much less straightforward.

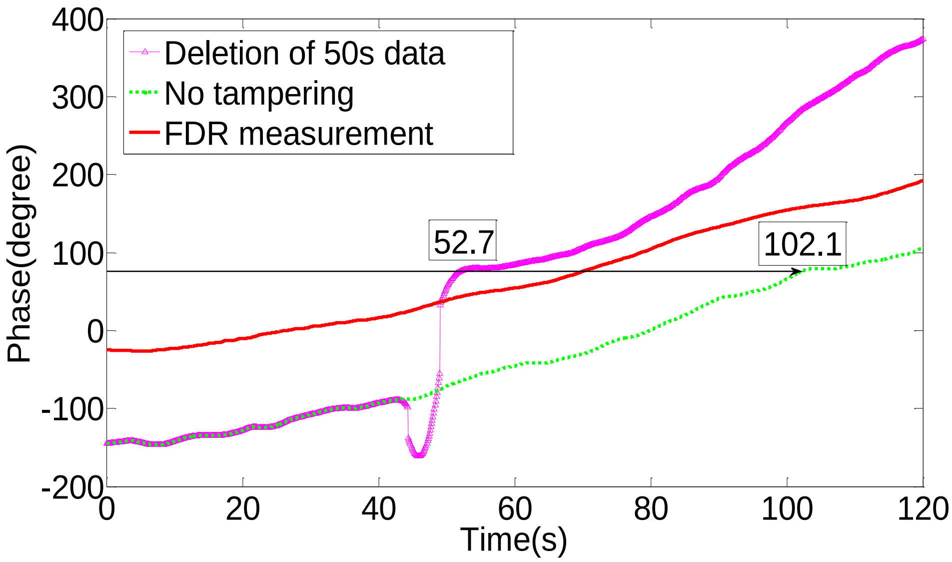


Figure 5-26 Deletion length estimation using phase angle with the reference data

B. Case 2 - Replacement

A section of the recording is replaced with another section of data. Figure 5-27 is the frequency change with different replacement lengths and Figure 5-28 shows the phase change. Note that these starting sections of tampered data are overlapped. As expected,

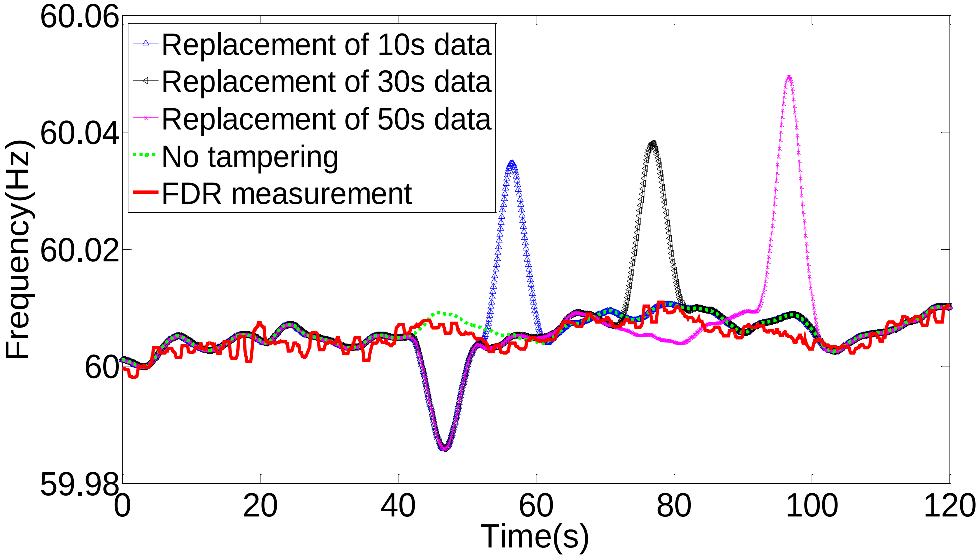


Figure 5-27 Frequency change with different lengths of replacement of the original recording

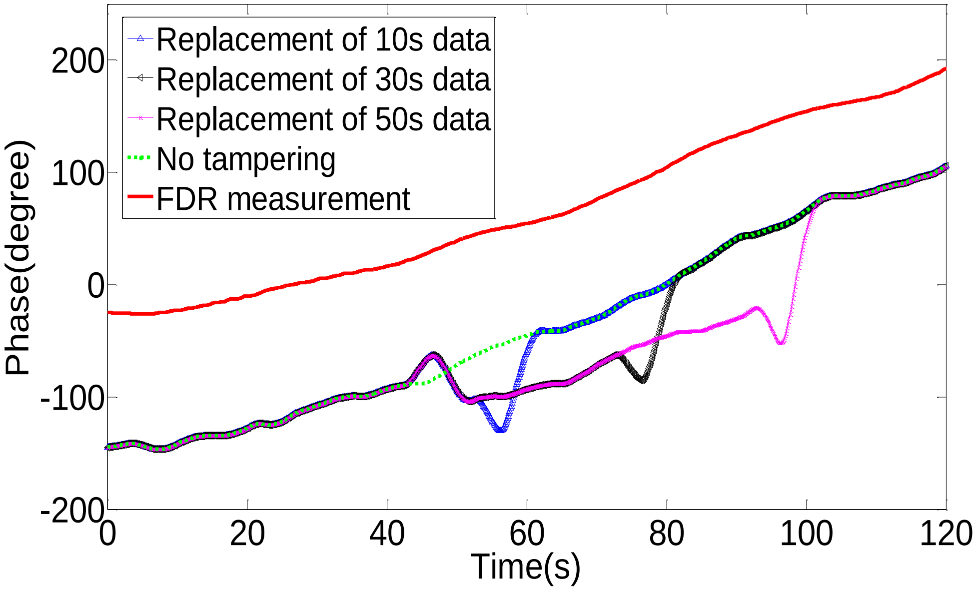


Figure 5-28 Phase change with different replacement lengths of the original recording

two frequency and phase spikes can be observed in the tampered cases. Furthermore, it is straightforward to estimate the length of replacement using either frequency or phase through measuring the time difference between the two spikes.

5.5 Frequency Localization Using FDR Data

5.5.1 Overview

The capability of previous research to determine the recording location is limited to the size of one interconnected grid. It is desirable to have a better spatial resolution. With the widely deployed FNET/GridEye system (over 200 units in U.S.), it provides a great opportunity to look into this topic. The FDR deployment map in the North America is shown in Figure 1-1.

Variations among ENF signals within the same grid are due to the local load characteristics. These background noises at different places are recorded by the local frequency measurement devices such as FDRs. Thus by using the noise characteristics, the location of a target frequency signal can be identified by comparing it with historical data from different places. Wide area power grid frequency measurement system, such as FNET/GridEye, can provide the historical data with high geography density and long time period, which can be used as a complete dictionary to identify the location of target frequency.

Three levels of spatial resolution, i.e., different states in one interconnection, different cities in one state, and multiple places in one city, are utilized to study the possibility of frequency localization.

5.5.2 Methods

To determine the location of the target signal, it basically needs two steps: 1) extracting the characteristics of the noise from target frequency as shown in Figure 5-29, and 2) compare the extracted characteristics with historical data from different places as shown in Figure 5-30. Therefore, a method needs to be developed to extract the characteristics of noises while filtering out the common part. In addition, a pattern recognition method needs to be used to identify the location of the target signal.

Here frequency domain analysis is employed to get the characteristics of signals. Due to the location dependent feature, the background noise shows different statistical characteristics in the frequency domain. To extract these characteristics, the "noise" is obtained by removing the common part, and then DFT is performed.

Neural network, as a novel and intelligent data analysis tool, can be used for pattern recognition. The frequency spectrum from historical data will be used to train the neural network, and the frequency spectrum of the target signal will be input into the trained network to identify its location.

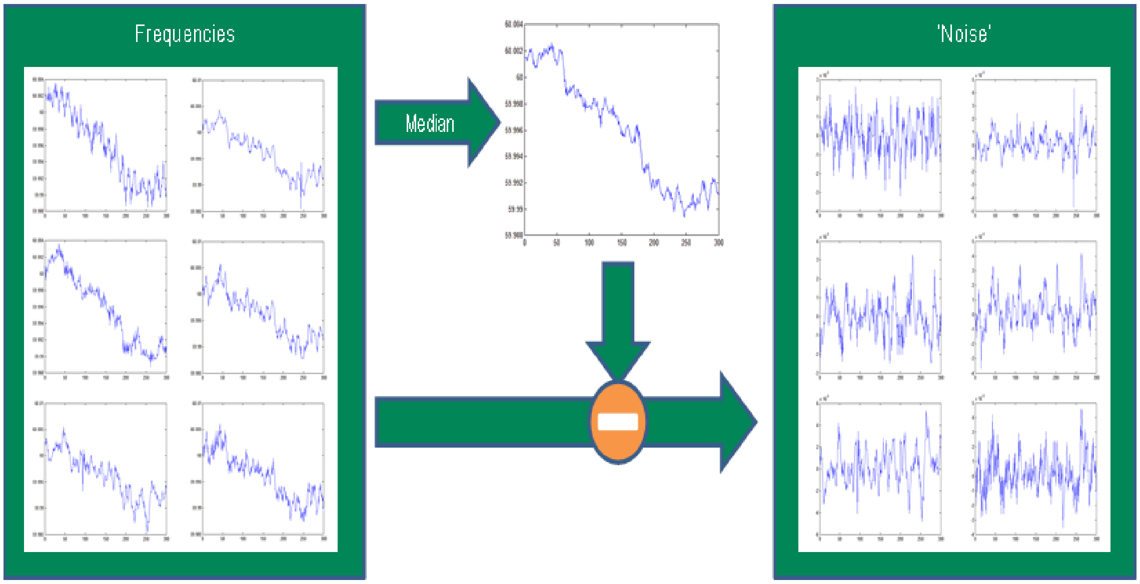


Figure 5-29 Approach of characteristics extraction

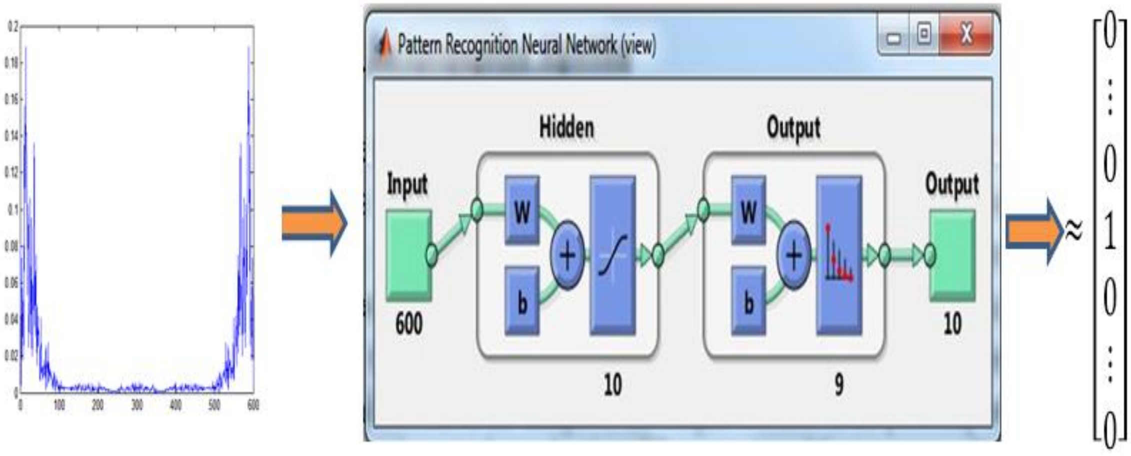


Figure 5-30 Approach of pattern recognition based on neural network

5.5.3 Results

In the current study, the FDR data are used and three levels of spatial resolution are studied.

As the first step, the location identification ability of signals from different states is tested. Five cities of different states are selected in Eastern Interconnection (EI) and shown in Figure 5-31. Frequency data recorded there are used as the historical data to train the neural network. A target signal from one of them measured at a different time is used to test if the developed technology can match it to the correct location of the five cities. 5-hour historical data are used to train the neural network. The target data length is 1 minute. Many samples of target data are tested to obtain the average matching ratio. Then, the localization area is reduced to a state (five cities of Missouri are selected) and same procedure is applied to obtain the results. Finally, five locations in the same city (Knoxville, TN) are tested.

For each of the three situations, different time intervals between target data and training data are considered, from 1 day to 2 years. This tested the effectiveness of the technology in different time scales. The results are listed in Table 5-2.

It can be seen from the table that for different states scenario, the location of the target signal can be identified with a high matching ratio even if the time interval between the target and historical data is up to 2 years. However, the matching ratio decreases as the time interval increases. Thus recent data are preferred for the frequency localization study

Figure 5-31 Location map of five states in EI

Table 5-2 Matching ratios for frequency localization study

Area	Time Interval	Matching Ratio
Different states	1 month	99.3%
	1 year	80.0%
	2 years	72.2%
Different cities in the same state	1 month	94.9%
	1 year	24.7%
Different locations in a city	1 day	39.3%

when they are available. Regarding different cities in the same state, the matching ratio of 1 month is lower, and the matching ratio of 1 year is too low to be useful. For the area as small as a city, the matching ratio even in one day is very low.

The reason is that in the same city, the background noise propagates to every outlet without too much attenuation. As a result, the frequency measurements show a very small difference between each other. That is, the "noise" extracted from the frequency is smaller than the resolution of frequency measurement devices which is 0.1 mHz [75]. Therefore, the characteristics of different locations are submerged by the measurement uncertainty and thus cannot be distinguished. To identify the location within a city, higher measurement accuracy is required [76].

Though the location cannot be distinguished within a city, the results actually indicate that it is the local frequency characteristics of power grid rather than measurement unit difference that enable the distinguishability. Therefore, this technology is useful regardless of what measurement recording devices are used as long as they are accurate enough.

5.6 Conclusions

This chapter presents an innovative application of wide area power system measurement for digital authentication using FNET/GridEye system. Four research studies, i.e., oscillator error removal, ENF phenomenology, tampering detection using both frequency and phase angle, and frequency localization, are performed. This provides a comprehensive guideline of using power system measurement for digital authentication.

Future work includes using different lengths of frequency data and extracted ENF for the localization study, and utilizing the local characteristics of frequency for cyber security related study.

Chapter 6 Data Analytics Using Historical FNET/GridEye Database

The large amounts of data collected by FNET/GridEye provide unique opportunities for data analytics related studies. This chapter explores several preliminary studies using historical FNET/GridEye data. Further investigation may be performed by expanding the preliminary studies.

6.1 EI Inertia Estimation and Analysis

The frequency drops when the power supply of the system is less than the demand due to a loss of generation. This study uses historical data to estimate the system inertia constant of EI. The various parameters in the estimation method were examined and determined in order to have an accurate estimation. In addition, analysis and interpretation of the results will be performed to help with system operators to evaluate the system condition.

The frequency of a power system changes when there is an imbalance between supply and demand. The frequency drops when the power supply of the system becomes insufficient due to a loss of major generation. If the amount of frequency drop is large, protection systems for low frequency may be activated in the power plants and the consequent shutdown of the plants may lead to the black-out of the power system. Thus it is extremely important to grasp the frequency response of the power system to the loss of generation. The rate of frequency change due to imbalance depends on the inertia of the system. System inertia is directly proportional to synchronously rotating mass in the system which includes synchronous generation and motor load. Knowing the system inertia is very helpful for system operations.

Estimation of system inertia is an important task for every power system [30, 77-78]. T. Inoue et al. [79] conducted research in this subject to estimate the inertia constant of Japanese system. A polynomial approximation with respect to time was fit to the waveform of the transients in estimating the inertia constant. D. P. Chassin et al. [80] developed computer models to compute the inertial constant of Western Electricity Coordination Council (WECC) system. S. Sharma et al. developed an on-line Inertial Frequency Response Estimation Tool for the ERCOT Interconnection [81].

This study focuses on analyzing the inertia of Eastern Interconnection (EI) using historical FNET/GridEye data. There are three major goals for this project:

1. Develop a method to estimate the system inertia using historical measurement data correctly.
2. Evaluate and improve the developed system inertia estimation method by refining the parameters.
3. Investigate the effects of different factors, such as system load and seasonal variation, on the system inertia.

6.1.1 System Inertia Estimation

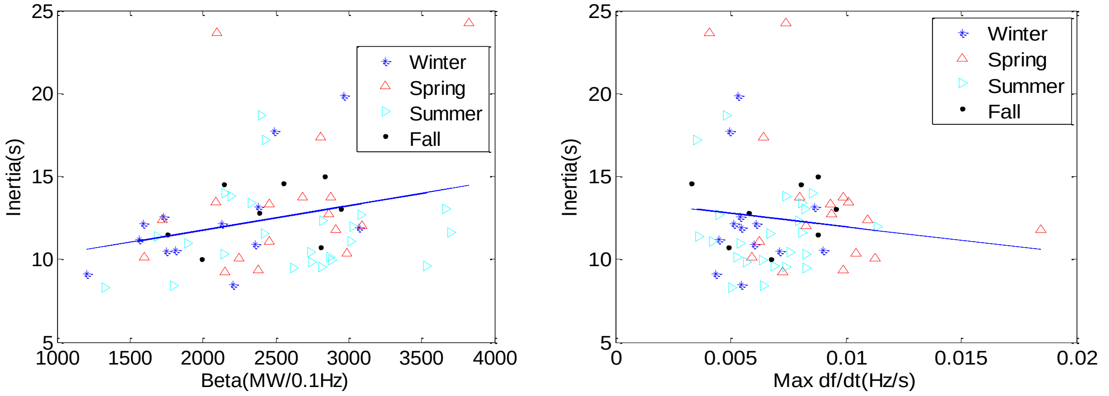
The general equation for calculating system inertia using rate of frequency change is illustrated below:

$$H = \frac{\Delta P}{df/_{dt}} \times \frac{f_0}{2} \tag{6-1}$$

where H is system inertia constant on system base (seconds), ΔP is power change (per unit in system load base), df/dt is rate of change of frequency (Hz/s) and f_0 is frequency at the time of disturbance (Hz).

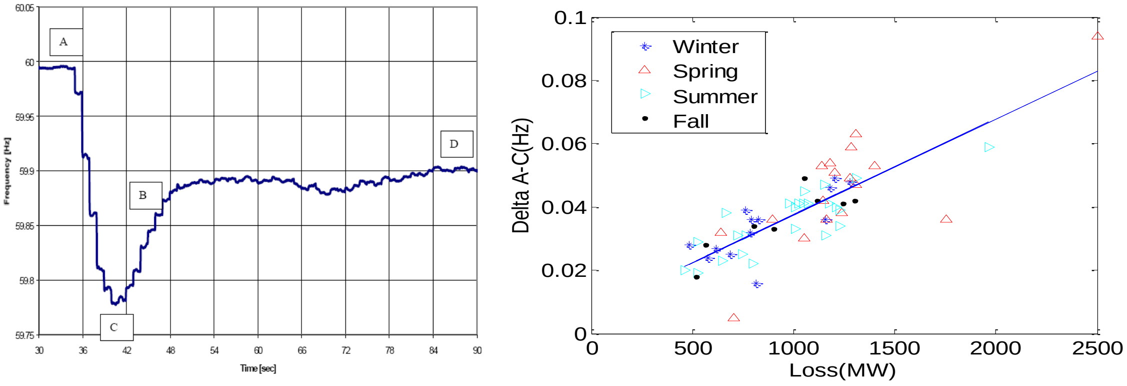
Confirmed large generation trip events in Year 2011 are analyzed. The estimated system inertia constant of EI versus system beta value in four different seasons (spring, summer, fall, winter) is shown in Figure 6-1(a) and the estimated inertia constant versus maximum rate of frequency change is shown in Figure 6-1(b). The blue line in the figures is the corresponding linear regression of the model. It could be seen that the inertia has a positive linear relationship with beta value while has a negative linear relationship with the rate of change of frequency.

NERC's frequency response standard is shown in Figure 6-1(c), where Point A is the frequency immediately before disturbance, Point B is the frequency at the point immediately after the frequency stabilizes due to governor action but before contingent control area takes action and Point C is the point of maximum frequency excursion due to the loss of rotating kinetic energy from the interconnection. The relationship between inertial frequency response and generation loss is high correlated as shown in Figure 6-1(d). This relationship may be used to estimate the generation loss based on the frequency deviation.



(a) System inertia vs. beta value

(b) System inertia vs. maximum df/dt



(c) NERC defined A, C, and B points

(d) Inertial frequency response vs. MW loss

Figure 6-1 Inertia estimation results

6.1.2 Inertia Analysis Using SPSS

A statistical software, SPSS, is used to perform the inertia analysis. Firstly, the correlation coefficient of the variables is examined, which is shown in Figure 6-2.

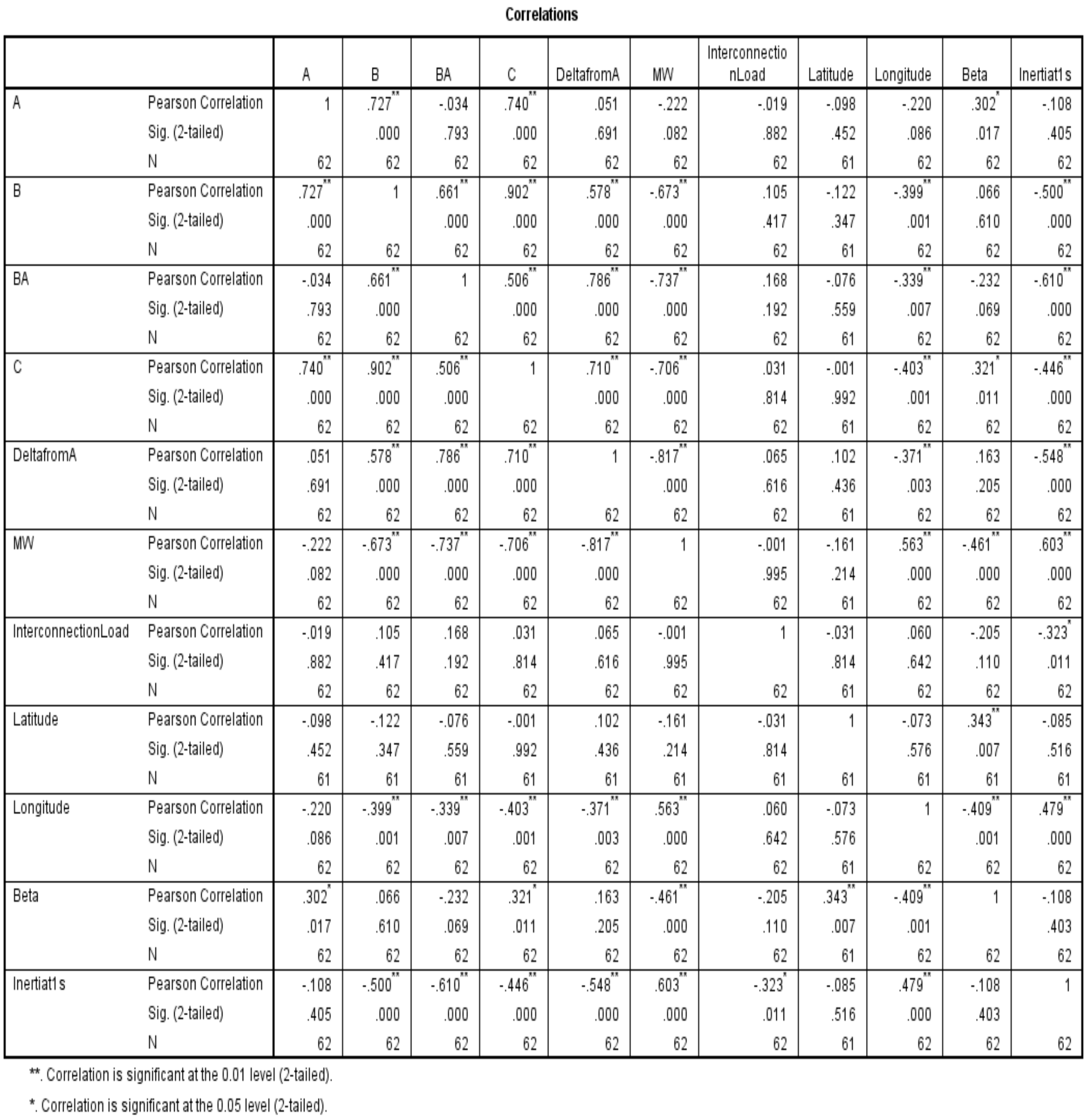
A number of factors, such as event location, load amount at event time, and frequency change during the event, are considered. In the analysis, "Beta" or "Inertia" is the dependent and other factors could be added as independents. Stepwise linear regression is chosen which adds only one variable at a time to the model as the model is "slowly" built.

For Year 2011, the data also include interconnection load for each event. The final model for the Beta analysis excluding BA frequency difference and loss amount is shown in Figure 6-3 and Figure 6-4. It can be seen from the figures that the final model is built in five steps. The combined variables explain approximately 60% of the variance of inertia.

Regarding the inertia, the final model is shown in Figure 6-5. It can be seen from the figure that the final model is built in three steps. The combined variables explain approximately 50% of the variance of inertia.

6.2 FIDVR Detection in FNET/GridEye Historical Database

A Delayed Voltage Recovery event, or more popularly today, a Fault Induced Delayed Voltage Recovery (FIDVR), is the phenomenon whereby system voltage remains at significantly reduced levels for several seconds after a transmission, sub-transmission, or distribution fault has been cleared [82]. FIDVR is caused by highly concentrated



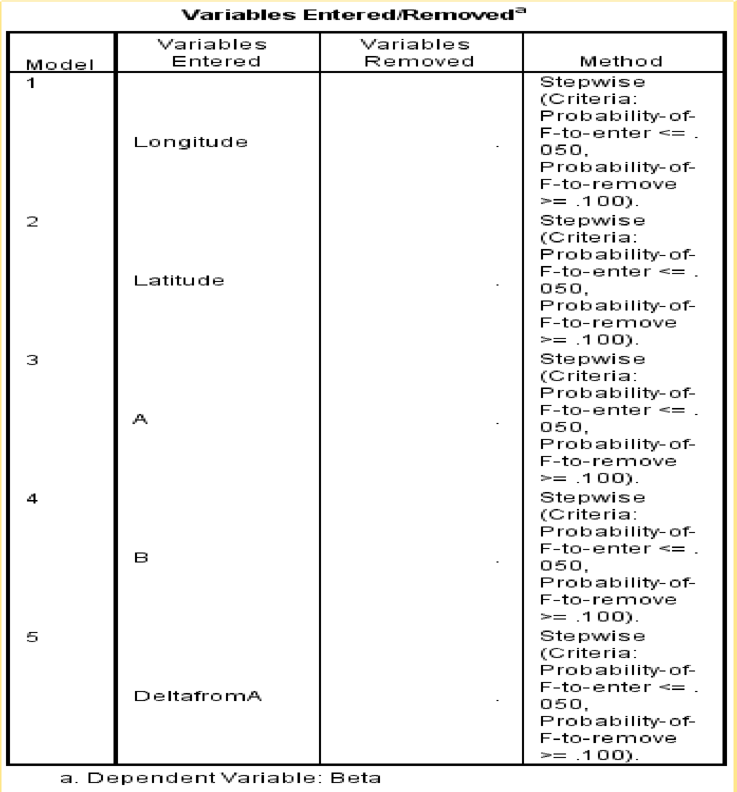
Correlations

		A	B	BA	C	DeltafromA	MW	InterconnectionLoad	Latitude	Longitude	Beta	Inertiat1s
A	Pearson Correlation	1	.727**	-.034	.740**	.051	-.222	-.019	-.098	-.220	.302*	-.108
	Sig. (2-tailed)		.000	.793	.000	.691	.082	.882	.452	.086	.017	.405
	N	62	62	62	62	62	62	62	61	62	62	62
B	Pearson Correlation	.727**	1	.661**	.902**	.578**	-.673**	.105	-.122	-.399**	.066	-.500**
	Sig. (2-tailed)	.000		.000	.000	.000	.000	.417	.347	.001	.610	.000
	N	62	62	62	62	62	62	62	61	62	62	62
BA	Pearson Correlation	-.034	.661**	1	.506**	.786**	-.737**	.168	-.076	-.339**	-.232	-.610**
	Sig. (2-tailed)	.793	.000		.000	.000	.000	.192	.559	.007	.069	.000
	N	62	62	62	62	62	62	62	61	62	62	62
C	Pearson Correlation	.740**	.902**	.506**	1	.710**	-.706**	.031	-.001	-.403**	.321*	-.446**
	Sig. (2-tailed)	.000	.000	.000		.000	.000	.814	.992	.001	.011	.000
	N	62	62	62	62	62	62	62	61	62	62	62
DeltafromA	Pearson Correlation	.051	.578**	.786**	.710**	1	-.817**	.065	.102	-.371**	.163	-.548**
	Sig. (2-tailed)	.691	.000	.000	.000		.000	.616	.436	.003	.205	.000
	N	62	62	62	62	62	62	62	61	62	62	62
MW	Pearson Correlation	-.222	-.673**	-.737**	-.706**	-.817**	1	-.001	-.161	.563**	-.461**	.603**
	Sig. (2-tailed)	.082	.000	.000	.000	.000		.995	.214	.000	.000	.000
	N	62	62	62	62	62	62	62	61	62	62	62
InterconnectionLoad	Pearson Correlation	-.019	.105	.168	.031	.065	-.001	1	-.031	.060	-.205	-.323*
	Sig. (2-tailed)	.882	.417	.192	.814	.616	.995		.814	.642	.110	.011
	N	62	62	62	62	62	62	62	61	62	62	62
Latitude	Pearson Correlation	-.098	-.122	-.076	-.001	.102	-.161	-.031	1	-.073	.343**	-.085
	Sig. (2-tailed)	.452	.347	.559	.992	.436	.214	.814		.576	.007	.516
	N	61	61	61	61	61	61	61	61	61	61	61
Longitude	Pearson Correlation	-.220	-.399**	-.339**	-.403**	-.371**	.563**	.060	-.073	1	-.409**	.479**
	Sig. (2-tailed)	.086	.001	.007	.001	.003	.000	.642	.576		.001	.000
	N	62	62	62	62	62	62	62	61	62	62	62
Beta	Pearson Correlation	.302*	.066	-.232	.321*	.163	-.461**	-.205	.343**	-.409**	1	-.108
	Sig. (2-tailed)	.017	.610	.069	.011	.205	.000	.110	.007	.001		.403
	N	62	62	62	62	62	62	62	61	62	62	62
Inertiat1s	Pearson Correlation	-.108	-.500**	-.610**	-.446**	-.548**	.603**	-.323*	-.085	.479**	-.108	1
	Sig. (2-tailed)	.405	.000	.000	.000	.000	.000	.011	.516	.000	.403	
	N	62	62	62	62	62	62	62	61	62	62	62

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Figure 6-2 Correlation coefficient of variables

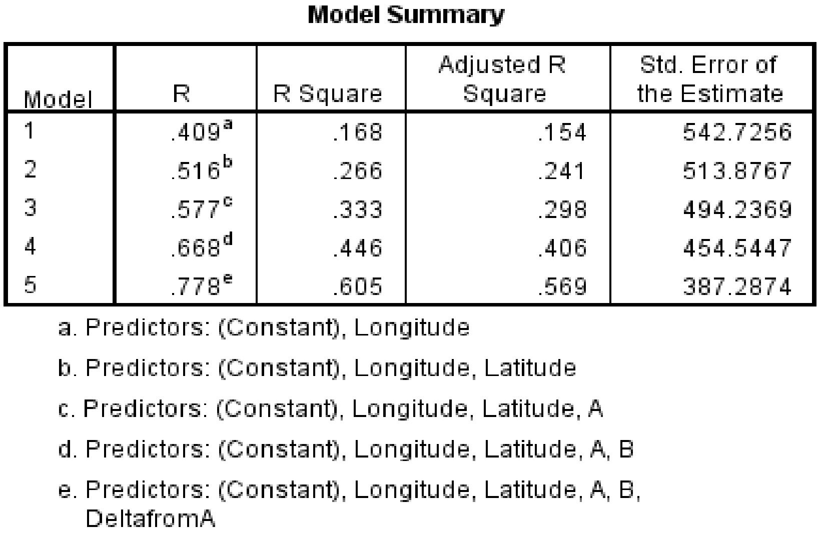


Variables Entered/Removed[a]

Model	Variables Entered	Variables Removed	Method
1	Longitude	.	Stepwise (Criteria: Probability-of-F-to-enter <= .050, Probability-of-F-to-remove >= .100).
2	Latitude	.	Stepwise (Criteria: Probability-of-F-to-enter <= .050, Probability-of-F-to-remove >= .100).
3	A	.	Stepwise (Criteria: Probability-of-F-to-enter <= .050, Probability-of-F-to-remove >= .100).
4	B	.	Stepwise (Criteria: Probability-of-F-to-enter <= .050, Probability-of-F-to-remove >= .100).
5	DeltafromA	.	Stepwise (Criteria: Probability-of-F-to-enter <= .050, Probability-of-F-to-remove >= .100).

a. Dependent Variable: Beta

Figure 6-3 Variables entered/removed



Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.409[a]	.168	.154	542.7256
2	.516[b]	.266	.241	513.8767
3	.577[c]	.333	.298	494.2369
4	.668[d]	.446	.406	454.5447
5	.778[e]	.605	.569	387.2874

a. Predictors: (Constant), Longitude

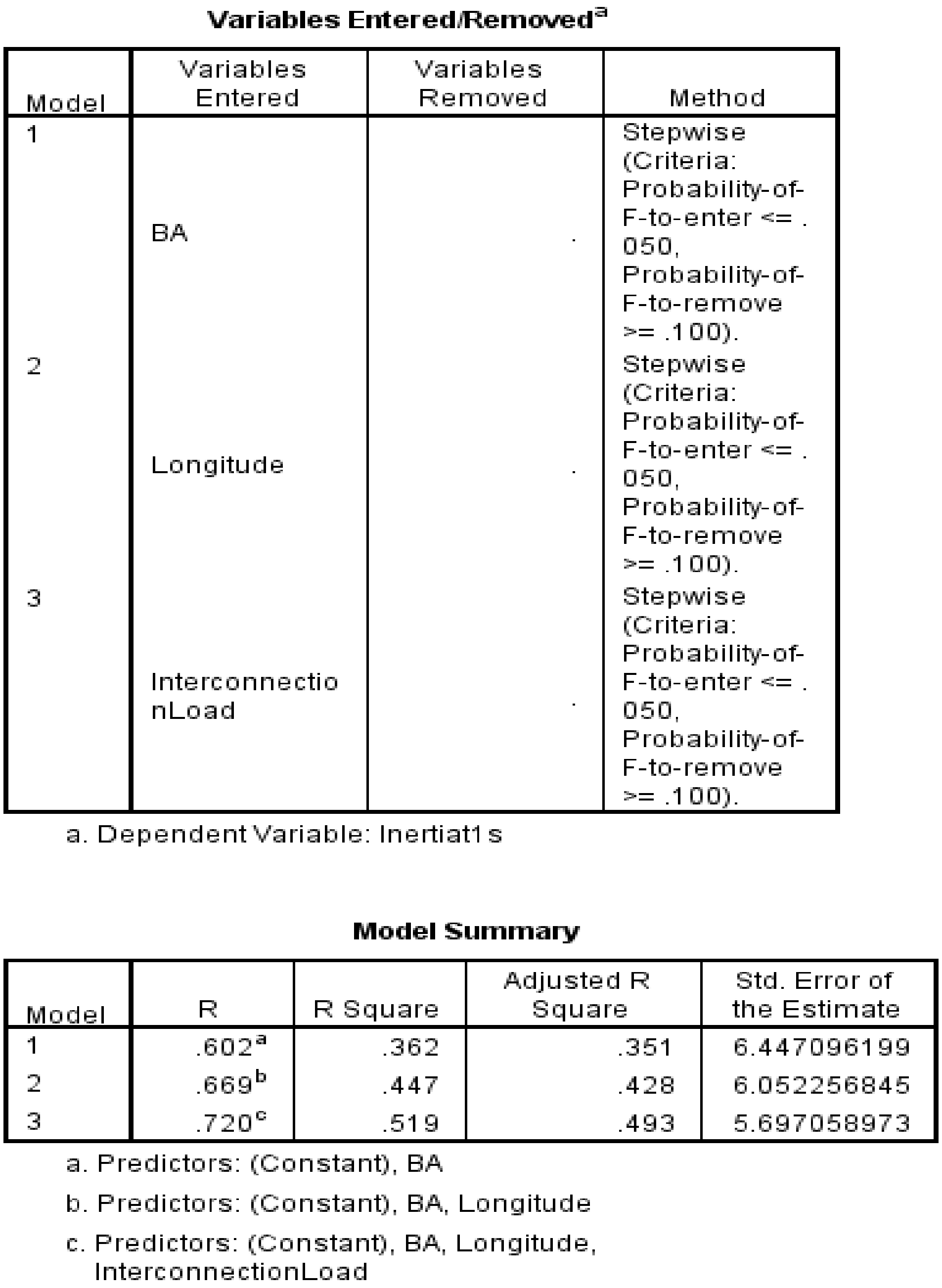
b. Predictors: (Constant), Longitude, Latitude

c. Predictors: (Constant), Longitude, Latitude, A

d. Predictors: (Constant), Longitude, Latitude, A, B

e. Predictors: (Constant), Longitude, Latitude, A, B, DeltafromA

Figure 6-4 Model summary for Beta analysis



Variables Entered/Removed[a]

Model	Variables Entered	Variables Removed	Method
1	BA	.	Stepwise (Criteria: Probability-of-F-to-enter <= .050, Probability-of-F-to-remove >= .100).
2	Longitude	.	Stepwise (Criteria: Probability-of-F-to-enter <= .050, Probability-of-F-to-remove >= .100).
3	InterconnectionLoad	.	Stepwise (Criteria: Probability-of-F-to-enter <= .050, Probability-of-F-to-remove >= .100).

a. Dependent Variable: Inertiat1s

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.602[a]	.362	.351	6.447096199
2	.669[b]	.447	.428	6.052256845
3	.720[c]	.519	.493	5.697058973

a. Predictors: (Constant), BA

b. Predictors: (Constant), BA, Longitude

c. Predictors: (Constant), BA, Longitude, InterconnectionLoad

Figure 6-5 Model summary for inertia analysis induction motor loads with constant torque that stall in response to low voltages associated with system faults and draw excessive reactive power from the grid. The study is to show if the phenomena is observable in FNET/GridEye database. Figure 6-6 shows a typical pattern for voltage data during a FIDVR event [82].

Figure 6-6 A typical FIDVR following a 230-kV transmission fault in Southern California

According to the characteristics of the FIDVR event as shown in the figure, a trigger is designed to detect such phenomena in the large FNET/GridEye historical database. The flowchart is shown in Figure 6-7. Several important parameters for designing the trigger are the voltage drop and time duration of going back to pre-trigger value. These parameters can be varied for the trigger. By using the designed trigger, some preliminary results have been generated. For example, a number of "FIDVR like" cases were detected from summer 2011 FNET/GridEye data at various locations as shown in Figure 6-8. The overall shape is very similar to the typical FIDVR pattern, but the voltage drop in some of them is not severe.

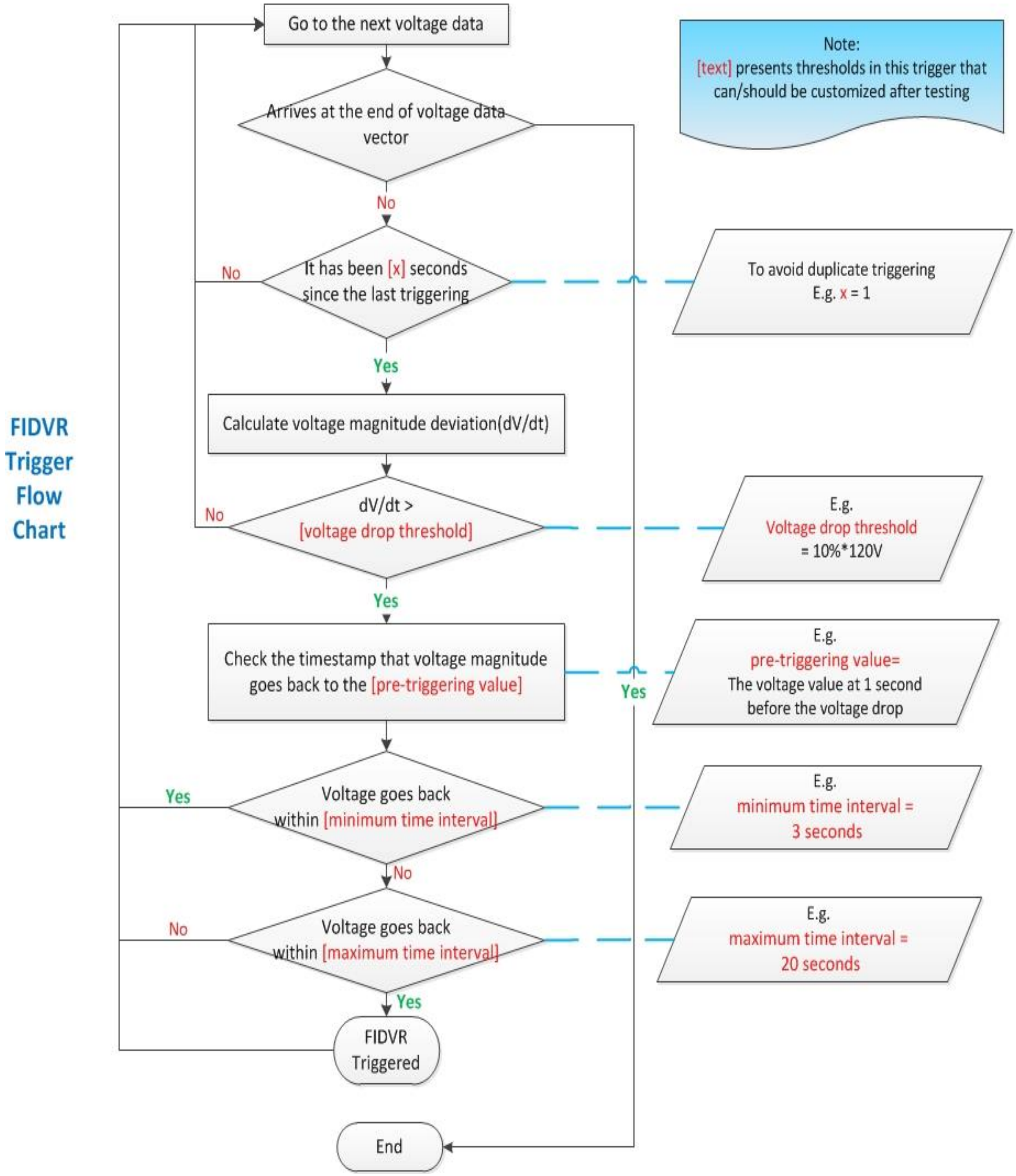


Figure 6-7 Flowchart of trigger design

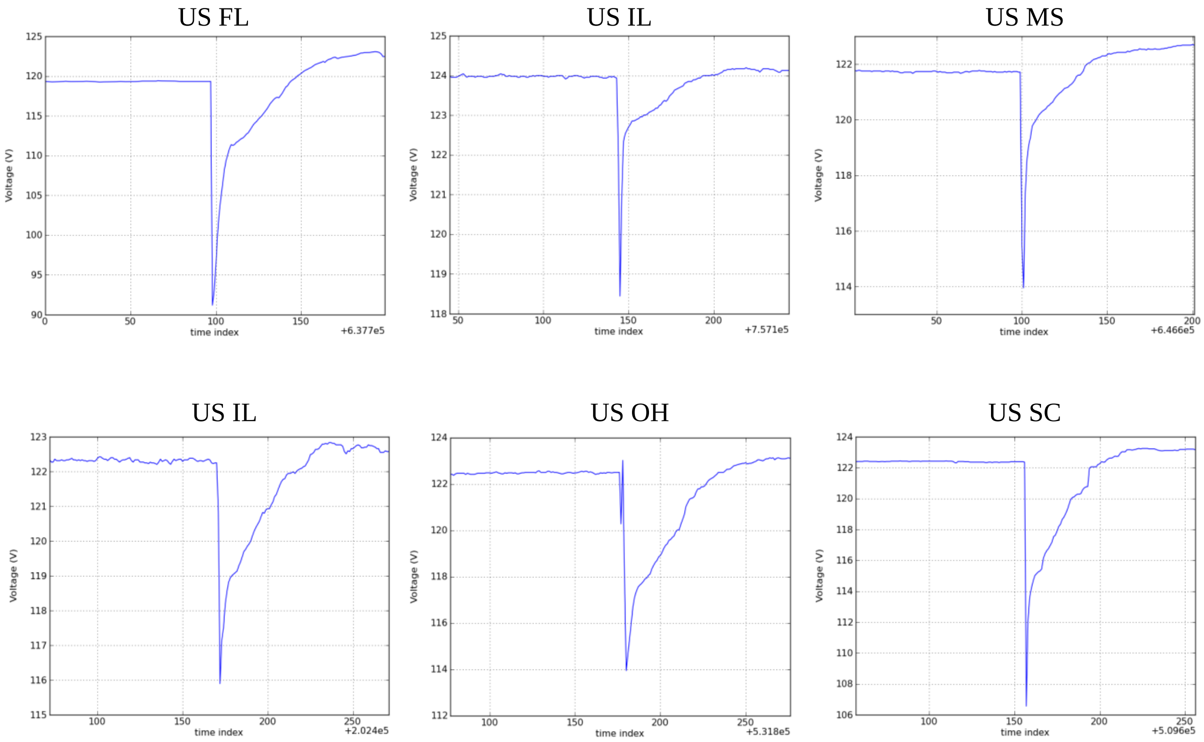


Figure 6-8 Sample results of detected FIDVR in FNET/GridEye database

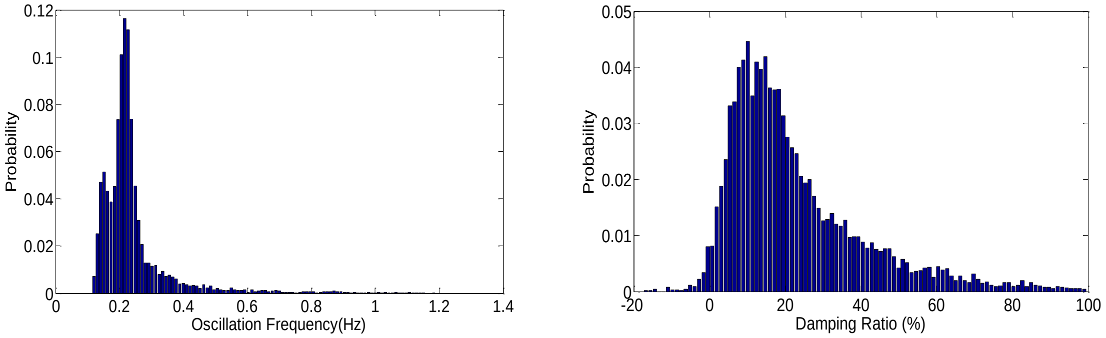
Future work includes a more complete analysis of the database, event occurrence mapping and trend analysis, correlation of FIDVR occurrence with recorded temperatures, and comparison with PMU data.

6.3 Statistical Analysis of FNET/GridEye Oscillation Database

The FNET/GridEye oscillation event database includes a great deal of useful information, such as frequency, damping ratio, maximum amplitude and time information [83]. Thanks to the wide deployment of FDRs, high-resolution geolocation information can also be found in the database.

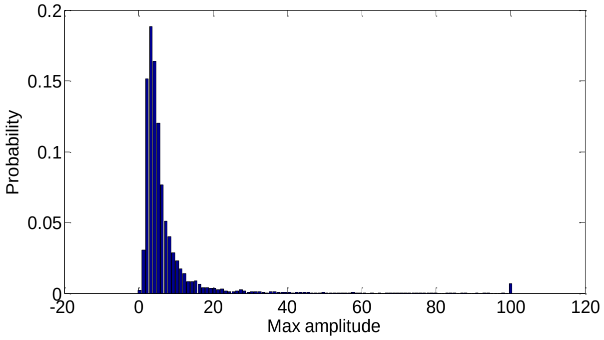
Generally speaking, the causes of oscillation can be generation trip, load shedding, line trip and non-obvious event. The probability distributions of frequency, damping ratio and maximum amplitude in EI are analyzed here.

The probability distributions of frequency, damping ratio and maximum amplitude for several years of data in EI are shown in Figure 6-9. It could be seen from Figure 6-9(a) that oscillation frequencies of most events are around 0.2 Hz, which is a typical inter-area oscillation frequency in EI. Figure 6-9(b) indicates that the damping ratios are mostly below 40%. Figure 6-9(c) shows that the maximum amplitudes are mostly below 20 degrees.



(a) Probability distribution of frequency

(b) Probability distribution of damping ratio



(c) Probability distribution of maximum amplitude

Figure 6-9 Probability distribution of important factors

Next, generation trip-caused oscillation events with location are specifically investigated. Take year 2014 EI data as an example, the generation trip and oscillation information is shown in Figure 6-10. Generation trip data include time, location, and size of the trip. Oscillation data include time, oscillation level, maximum negative and positive magnitude, dominant frequency, and damping ratio.

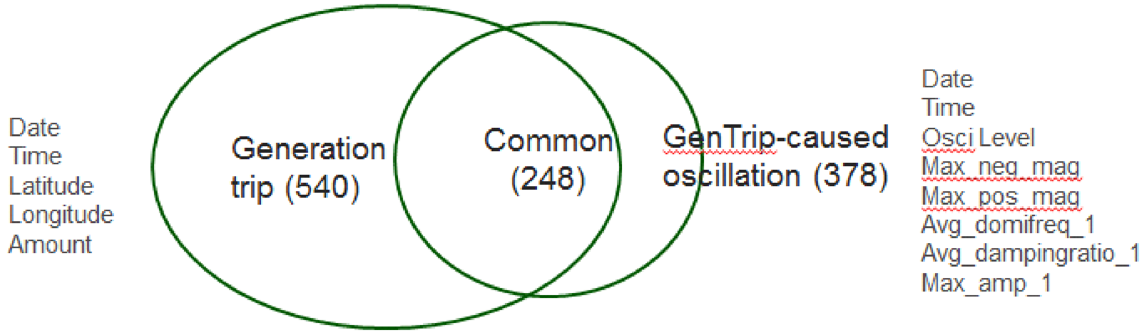


Figure 6-10 Generation trip and oscillation data

Figure 6-11 shows the oscillation level distribution in EI. Red indicates oscillation level is between 0 and 1 (highest oscillation magnitude), blue indicates oscillation level is between 2 and 3 (lowest oscillation magnitude), and green indicates oscillation level is between 1 and 2.

It can be seen from the figure that oscillation magnitudes are larger near the boundary and are not highly related with the generation trip amount.

Future work includes validation of the observation in simulation.

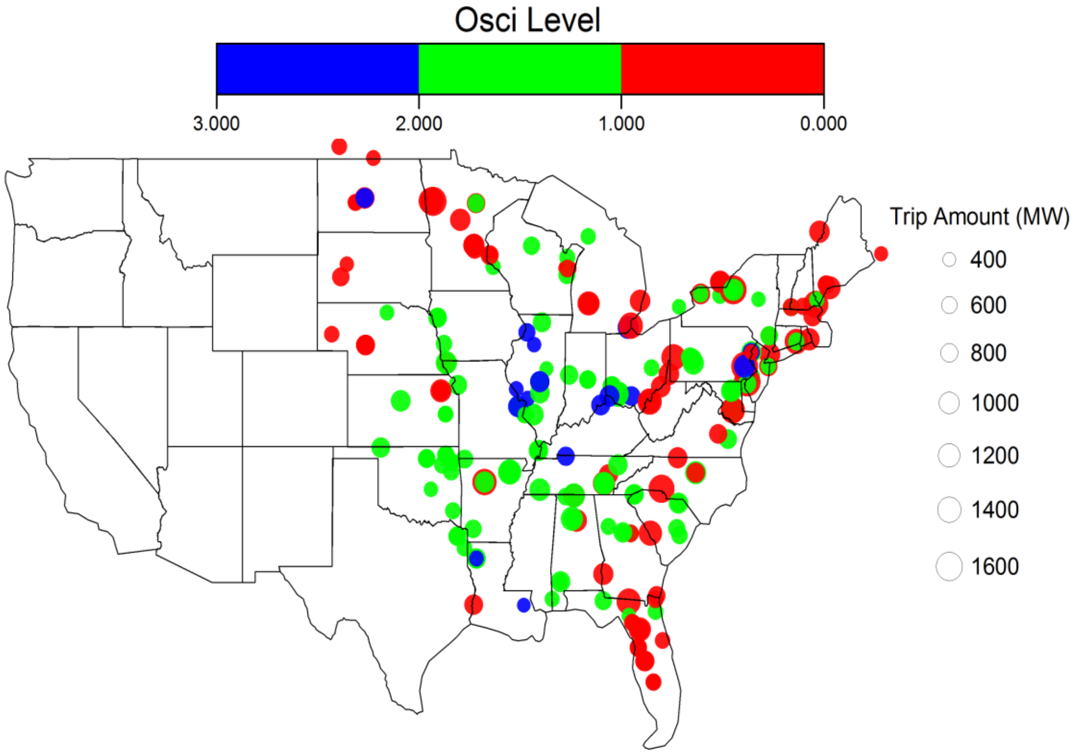


Figure 6-11 Oscillation level distribution in EI

Chapter 7 Conclusions and Future Work

Driven by both technological and economic considerations, electric power grids have become more interconnected across regions, countries, and even continents. However, planning and control of interconnection-level power grids is a demanding task that must carefully deal with interconnected system dynamics. With the bottlenecks of traditional circuit-based simulation, better utilization and integration of more and more power system measurement is highly needed. This dissertation presents several applications using either PMU or FDR which contribute to the development of power system monitoring and dynamic modeling.

For measurement-based system response prediction, the MAR model is developed for power system study of linear conditions with pure measurement. It can be easily updated online with typical events. To improve its feasibility with a large number of measurement inputs, a model reduction technique with delayed correlation coefficient is utilized for input selections. Future work includes a comprehensive test of the approach in a real system.

For measurement-based system reduction, the external system is reduced using the ARX model based on certain input/output data. This technique may increase the speed of power system dynamic simulation and improve the accuracy of equivalent systems.

Future work includes implementation of the approach in other more flexible tools and improvement of the ARX model for corresponding operating condition.

For measurement-based generator dynamic model validation, a detailed analysis using both simulation data and PMU data is performed. By utilizing the simulation data, several important factors, such as the effect of PMU reporting rates, event selection, and masking effect are discussed. This technique only requires a simple procedure and could allow more frequent validation.

For the application of using power system measurement for digital authentication, four research studies, i.e., oscillator error removal, ENF phenomenology, tampering detection using both frequency and phase angle, and frequency localization, are performed. It improves the understanding of the ENF criterion and provides a comprehensive guideline. With the discovered local characteristics of frequency, future work includes using it for power grid cyber security related research.

Finally, for the data analytics studies using historical data, preliminary results including inertia estimation and analysis, FIDVR detection, and statistical analysis of oscillation database are presented. These show the great potential of data analytics research to better understand the grid. Future work includes more thorough investigation and corresponding validation in the simulation.

List of References

[1] Z. Zhong, C. Xu, B. J. Billian, L. Zhang, S.-J.S. Tsai, R. W. Conners, V. A. Centeno, A. G. Phadke, and Y. Liu, "Power system frequency monitoring network (FNET) implementation," *IEEE Trans. Power Syst.*, vol. 20, no. 4, pp. 1914-1921, Nov. 2005.

[2] Y. Zhang, P. N. Markham, T. Xia, L. Chen, Y. Ye, Z. Wu, Z. Yuan, L. Wang, J. Bank, J. Burgett, R.W. Conners, and Y. Liu, "Wide-Area frequency monitoring network (FNET) architecture and applications," *IEEE Trans. Smart Grid*, vol. 1, no. 2, pp. 159-167, Sept. 2010.

[3] Y. Liu, et al., "Wide-Area Measurement System Development at the Distribution Level: an FNET/GridEye Example," *IEEE Trans. Power Del., vol.31, no.2, pp. 721-731, 2016.*

[4] D. Zhou, J. Guo, Y. Zhang, J. Chai, H. Liu, C. Huang, and Y. Liu, "Distributed data analytics platform for wide-area synchrophasor measurement systems," *IEEE Trans. Smart Grid*, in press.

[5] J. Chai, et al., "Wide-area measurement data analytics using FNET/GridEye: a review," in *Proc. 19th Power Systems Computation Conference (PSCC)*, 2016.

[6] A. G. Phadke, J. S. Thorp, and M. G. Adamiak, "A new measurement technique for tracking voltage phasors, local system frequency, and rate of change of frequency," *IEEE Trans. Power App. Syst.*, vol. PAS-102, no. 5, pp. 1025-1038, May 1983.

[7] A. G. Phadke and R. M. de Moraes, "The wide world of wide-area measurement," *IEEE Power Energy Mag.*, vol. 6, no. 5, pp. 52-65, Sept./Oct. 2008.

[8] D. Karlsson, M. Hemmingsson, and S. Lindahl, "Wide area system monitoring and control-Terminology, phenomena, and solution implementation strategies," *IEEE Power Energy Mag.*, vol. 2, no. 5, pp. 68-76, Sept.-Oct. 2004.

[9] L. Zhan, J. Zhao, S. Gao, J. Cullis, Y. Liu, and Y. Liu, "Universal grid analyzer design and development," in *Proc. 2015 IEEE Power Eng. Soc. Gen. Meet*, pp. 1-5, 2015.

[10] L. Zhan, Y. Liu, J. Culliss, J. Zhao, and Y. Liu, "Dynamic single-phase synchronized phase and frequency estimation at the distribution level," *IEEE Trans. Smart Grid*, vol. 6, no. 4, pp. 2013-2022, July 2015.

[11] L. Zhan and Y. Liu, "Improved WLS-TF algorithm for dynamic synchronized angle and frequency estimation," in *Proc. 2014 IEEE Power Eng. Soc. Gen. Meet,* pp. 1-5, 2014.

[12] R. M. Gardner, J. K. Wang, and Y. Liu, "Power system event location analysis using wide-area measurements," in *Proc. 2006 Power Eng. Soc. Gen. Meet*., pp. 1-7, 2006.

[13] S. You, J. Guo, Y. Liu and Y. Liu, "Oscillation Mode Identification Based on Wide-Area Ambient Measurements Using Multivariate Empirical Mode Decomposition," *Electric Power Systems Research*, vol. 134, pp. 158-166, May 2016.

[14] J. Guo, H. Liu, D. Zhou, J. Chai, Y. Zhang, and Y. Liu, "Real-time power system electromechanical mode estimation implementation and visualization utilizing synchrophasor data", in *Proc. 2016 IEEE T&D Conference and Exposition*, 2016.

[15] Z. Lin, T. Xia, Y. Ye, Y. Zhang, L. Chen, Y. Liu, K. Tomsovic, T. Bilke, and F. Wen, "Application of wide area measurement systems to islanding detection of bulk power systems," *IEEE Trans. Power Syst.*, vol. 28, no. 2, pp. 2006-2015, May 2013.

[16] J. Guo, Y. Zhang, M. A. Young, M. J. Till, A. Dimitrovski, Y. Liu, P. Williging, and Y. Liu, "Design and implementation of a real-time off-grid operation detection tool from a wide-area measurements perspective," *IEEE Trans. Smart Grid*, vol. 6, no. 4, pp. 2080-2087, July 2015.

[17] Y. Zhang, Y. Liu, L. Chen, J. Guo, and Y. Liu, "Visualization of distribution level voltage magnitude pattern trend in EI system using FNET data," in *Proc. 2014 IEEE PES T&D Conference and Exposition*, April 2014.

[18]J. Chai, Y. Liu, Y. Liu, N. Bhatt, A. D. Rosso, and E. Farantatos, "Measurement-based system reduction using autoregressive model," in *Proc. 2016 IEEE PES T&D Conference and Exposition*, 2016.

[19]G. Kou, P. N. Markham, S. Hadley, T. J. King, and Y. Liu, "Impact of governor deadband on frequency response of the U.S. eastern interconnection," *IEEE Trans. Smart Grid*, vol. 7, no. 3, pp. 1368-1377, June 2015.

[20]P. N. Markham, Y. Zhang, J. Guo, Y. Liu, T. Bilke, and D. Bertagnolli, "Analysis of frequency extrema in the Eastern and Western Interconnections, 2010-2011," in *Proc. 2012 IEEE Power and Energy Society General Meeting,* July 2012.

[21]J. Chai, J. Zhao, W. Yao, J. Guo and Y. Liu, "Application of wide area power system measurement for digital authentication," in *Proc. 2016 IEEE PES T&D Conference and Exposition*, 2016.

[22]Y. Lei, Y. Zhang, J. Guo, D. Zhou, and Y. Liu, "The impact of synchronized human activities on power system frequency," in *Proc. 2014 IEEE Power Eng. Soc. Gen. Meet.*, pp. 1-5, 2014.

[23]L. Liu, J. Chai, H. Qi and Y. Liu, "Power grid disturbance analysis using frequency information at the distribution level," in *Proc. 2014 IEEE International Conference on Smart Grid Communications,* Nov. 2014.

[24]X. Fang, Q. Hu, F. Li, B. Wang, and Y. Li, "Coupon-Based Demand Response Considering Wind Power Uncertainty: A Strategic Bidding Model for Load Serving Entities," *IEEE Trans. Power Syst.*, vol. 31, no. 2, pp. 1025-1037, 2016.

[25]J. Xiao, L. Bai, Z. Lu, and K. Wang, "Method, implementation and application of energy storage system designing," *Int. Trans. Elect. Energy Syst.*, vol. 24, pp. 378-394, 2014.

[26]X. Fang, F. Li, Y. Wei, R. Azim, and Y. Xu, "Reactive Power Planning with High Penetration of Wind Energy using Benders Decomposition," *IET Gener. Transm. Distrib.*, vol. 9, no. 14, pp. 1835-1844, 2015.

[27]L. Bai, T. Ding, Q. Hu, F. Li, and H. Sun, "Robust Mean-Variance Optimization Model for Grid-Connected Microgrids," in *Proc. 2015 IEEE PES General Meeting*, Denver, CO, Jul. 26-30, 2015.

[28]C. Huang, F. Li, J. Guo, and Y. Liu, "A bounded model of the communication delay for system integrity protection systems (SIPS)," *IEEE Trans. Power Delivery*, in press.

[29]H. Yuan and F. Li, "A Comparative Study of Measurement-Based Thevenin Equivalents Identification Methods," 46th North American Power Symposium, Pullman, WA, Sept. 7-9, 2014.

[30]Prabha Kundur, *Power System Stability and Control*, New York: McGraw-Hill, Inc. 1994.

[31]US-Canada Power System Outage Task Force. Final Report, August 14, 2003 Blackout in the United States and Canada: Causes and Recommendations. Online available : http://energy.gov/sites/prod/files/oeprod/DocumentsandMedia/BlackoutFinal-Web.pdf.

[32]J. De La Ree, V. Centeno, J. S. Thorp, and A. G. Phadke. "Synchronized phasor measurement applications in power systems," *IEEE Trans. Smart Grid*, vol. 1, no. 1, pp. 20-27, 2010.

[33]EPRI technical report, "Comprehensive Stability Analysis Using Synchrophasors: Development of Algorithms and Analytical Tools to Predict Rotor Angle Stability and Voltage Stability from Synchrophasor Data," 2012.

[34]A. G. Bahbah, and A. A. Girgis, "New Method for Generators' Angles and Angular Velocities Prediction for Transient Stability Assessment of Multimachine Power Systems Using Recurrent Artificial Neural Network," *IEEE Trans. Power Syst.*, vol. 19, no. 2, pp. 1015-1022, 2004.

[35]Lennart Ljung, "System Identification: Theory for the User", Englewood Cliffs: Prentice-Hall, Inc. 1987.

[36]P. M. Van Qirsouw, "A dynamic equivalent using modal coherency and frequency response," *IEEE Trans. Power Syst.*, vol. 5, no. 1, pp. 289-295, 1990.

[37]G. Troullinos and J. Dorsey, "Coherency and Model Reduction: State Space Point of View," *IEEE Trans. Power Syst.*, vol. 4, no. 3, pp. 988 - 995, Aug. 1989.

[38]Z. H. Wang, Y. P. Wang, and Y. D. Han, "A new external equivalent model for decentralized control design in multi-machine power system," IEEE International Conference on Advances in Power System Control, Operation and Management, Nov. 1991.

[39]A. N. Martins and P. E. M. Quintao, "Computing Dominant Poles of Power System Multivariable Transfer Functions," *IEEE Trans. Power Syst.*, vol. 18, no. 1, pp. 152-159, Feb. 2003.

[40]A. M. Stankovic and A. T. Saric, "Transient power system analysis with measurement-based gray box and hybrid dynamic equivalents," *IEEE Trans. Power Syst.*, vol. 19, no. 1, pp. 455-462, 2004.

[41] J. V. Milanovic and S. Mat Zali, "Validation of equivalent dynamic model of active distribution network cell,*" IEEE Trans. Power Syst*., vol. 28, no. 3, pp. 2101-2110, 2013.

[42]Y. Liu, K. Sun, and Y. Liu, "A Measurement-based power system model for dynamic response estimation and instability warning," *Electric Power Systems Research*, vol. 124, pp. 1-9, July 2015.

[43]C. Li, J. Chai, Y. Liu, N. Bhatt, A. D. Rosso, and E. Farantatos, "Power system dynamics prediction with measurement-based autoregressive model," in *Proc. 2014 CIGRE US Conference*, Houston, TX, US, 2014.

[44]H. Liu, et al., "ARMAX-Based Transfer Function Model Identification Using Wide-Area Measurement for Adaptive and Coordinated Damping Control," *IEEE Trans. Smart Grid*, in press.

[45]EPRI and Powertech, DYNRED User Manual, 2010.

[46] H. Yuan and F. Li, "Hybrid Voltage Stability Assessment (VSA) for N-1 Contingency," *Electric Power Systems Research*, vol. 122, pp. 65-75, May 2015.

[47] T. Jiang, L. Bai, H. Jia H. Yuan, and F. Li, "Identification of voltage stability critical injection region in bulk power systems based on the relative gain of voltage coupling," *IET Gener. Transm. Distrib.*, in press.

[48] H. Yuan, T. Jiang, H. Jia, F. Li, Y. Mishra, H. Chen, and G. Li, "RealTime Wide-area Loading Margin Sensitivity (WALMS) in Power Systems," in *Proc. 2015 IEEE PES General Meeting*, pp. 1-5, 2015.

[49] "Power System Model Validation - A White Paper by the NERC Model Validation Task Force of the Transmission Issues Subcommittee," North American Electric Reliability Corporation, Princeton, NJ, USA, Dec. 2010.

[50] N. W. Miller et al., "Eastern frequency response study," National Renewable Energy Laboratory (NREL), Golden, CO, USA, May 2013.

[51] "Frequency response initiative report - the reliability role of frequency response," North American Electric Reliability Corporation (NERC), Atlanta, GA, USA, Oct. 2012.

[52] D. Kosterev and D. Davies, "System model validation studies in WECC," in *Proc. 2010 IEEE Power & Energy Society General Meeting*, pp. 1-4, 2010.

[53] J. Undrill, "Power and frequency control as it relates to wind-powered generation," John Undrill LCC, Dec. 2010.

[54] Z. Huang, P. Du, D. Kosterev, and S. Yang, "Generator dynamic model validation and parameter calibration using phasor measurements at the point of connection," *IEEE Trans. Power Syst.*, vol. 28, no. 2, pp. 1939-1949, 2013.

[55] A. Hajnoroozi et al., "Generating Unit Model Validation and Calibration Through Synchrophasor Measurements," *IEEE Trans. Smart Grid*, vol. 6, no. 1, pp. 441-449, 2015.

[56] N. Zhou, D. Meng, Z. Huang, and G. Welch, "Dynamic State Estimation of a Synchronous Machine Using PMU Data: A Comparative Study," *IEEE Trans. Smart Grid*, vol. 6, no. 1, pp. 450-460, 2015.

[57] Z. Huang, D. Kosterev, R. T. Guttromson, and T. B. Nguyen, "Model validation with hybrid dynamic simulation," in *Proc. IEEE Power Eng Soc. Gen. Meeting*, Montreal, QC, Canada, Jun. 18-22, 2005.

[58] C. Grigoras, "Forensic Analysis of Digital Recordings - The Electric Network Frequency Criterion," *Forensic Science International*, vol. 136, (supp. 1), pp. 368-369, 2003.

[59] C. Grigoras, "Applications of ENF Analysis in Forensic Authentication of Digital Audio and Video Recordings," *Journal of the Audio Engineering Society*, vol. 57, no. 9, pp. 643-661, 2009.

[60] A. J. Cooper, "An automated approach to the electric network frequency (ENF) criterion: theory and practice," *Int. J. of Speech Language and the Law*, vol. 16, no. 2, pp. 193-218, 2009.

[61] Y. Liu, Z. Yuan, P. N. Markham, R. W. Conners, and Y. Liu, "Application of power system frequency for digital audio authentication," *IEEE Trans. Power Delivery*, vol. 27, no. 4, pp. 1820-1828, Oct. 2012.

[62] D. P. N. Rodriguez, J. A. Apolinario, and L. W. P. Biscainho, "Audio authenticity: detecting ENF discontinuity with high precision phase analysis," *IEEE Trans. Inf. Forensics and Security*, vol. 5, no. 3, pp. 534-543, Sept. 2010.

[63] Z. Yuan, Y. Liu, R. W. Conners and Y. Liu, "Effects of oscillator errors on electric network frequency (ENF) analysis," in *Proc. The Audio Engineering Society 46th Conference*, Denver, USA, 2012.

[64] J. Chai, F. Liu, Z. Yuan, R. W. Conners and Y. Liu. "Source of ENF in battery-powered digital recordings," in *Proc. The Audio Engineering Society 135th International Convention*, New York City, USA, 2013.

[65] J. Chai, Y. Liu, Z. Yuan, R. W. Conners and Y. Liu, "Tampering detection of digital recordings using electric network frequency and phase angle," in *Proc. The Audio Engineering Society 135th International Convention*, New York City, USA, 2013.

[66] A. Cooper, "The Electric Network Frequency (ENF) as an Aid to Authenticating Forensic Digital Audio Recordings - An Automated Approach," *Audio Engineering Society 33rd International Conference*, Denver, CO, June 5-7, pp. 1-10, 2008.

[67] M. Huijbregtse and Z. Geradts, "Using the ENF Criterion for Determining the Time of Recordings of Short Digital Audio Recordings," *International Workshop on Computational Forensics 2009, Lecture Notes in Computer Science 5718*, Springer-Verlag, Berlin Heidelberg, pp. 116-124, 2009.

[68] M. Kajstura, A. Trawinska, and J. Hebenstreit, "Application of the Electrical Network Frequency (ENF) Criterion: A Case of a Digital Recording," *Forensic Science International*, vol. 155, pp. 165-171, 2005.

[69] M. Michalek, "The Application of Power Line Hum in Digital Recording Authenticity Analysis," *Problems of Forensic Sciences*, vol. LXXX, pp. 355- 364, 2009.

[70] C. Grigoras, "Digital Audio Recording Analysis: The Electric Network Frequency (ENF) Criterion," *International Journal of Speech Language and the Law*, vol. 12, no. 1, pp. 63-76, 2005.

[71] C. Grigoras, "Applications of ENG Criterion in Forensic Audio, Video, Computer and Telecommunications Analysis," *Forensic Science International*, vol. 167, pp. 136 - 145, 2007.

[72]E. B. Brixen, "ENF; Quantification of the Magnetic Field," *Audio Engineering Society 33rd International Conference*, Denver, CO, June 5-7, pp. 1-4, 2008.

[73]E. Jacobsen and P. Kootsookos, "Fast, Accurate Frequency Estimator," *IEEE Signal Process. Mag.*, vol. 24, no. 3, pp. 123-125, 2007.

[74]J. Chen, "Accurate Frequency Estimation with Phasor Angles," M.S. thesis, Dept. Elect. Comput. Eng., Virginia Polytec. Inst. State Univ., Blacksburg, VA, 1994.

[75]T. Xia, "Frequency monitoring network (FNET) algorithm improvements and application development," Ph.D. dissertation, Dept. Electrical and Computer Eng., Virginia Polytechnic Institute and State University, 2009.

[76]J. Zhao, L. Zhan, Y. Liu, H. Qi, J. R. Gracia, and P. D. Ewing, "Measurement accuracy limitation analysis on synchrophasors," in *Proc. IEEE PES General Meeting*, Denver, USA, 2015.

[77]Frequency Task Force of the NERC Resources Subcommittee, "Frequency Response Standard Whitepaper," April, 2004.

[78]FERC’s Order, "Study for Reliable Integration of Wind and Other Renewable Resources using Frequency Response," May 13, 2009.

[79]T. Inoue, H. Taniguchi, Y. Ikeguchi, and K. Yoshida, "Estimation of power system inertia constant and capacity of spinning-reserve support generators using measured frequency transients," *IEEE Trans. Power Syst.*, vol. 12, no. 1, pp. 136-143, Feb. 1997.

[80]D. P. Chassin, Z. Huang, M.K. Donnelly, C. Hassler, E. Ramirez, and C. Ray, "Estimation of WECC System Inertia Using Observed Frequency Transients," *IEEE Trans. Power Syst.*, vol. 20, no. 2, pp. 1190-1192, May 2005.

[81] S. Sharma, S. Huang, and N. D. R. Sarma, "System inertial frequency response estimation and impact of renewable resources in ERCOT interconnection," in *Proc. IEEE Power and Energy Society General Meeting*, pp. 1-6, July 2011.

[82] NERC Transmission Issues Subcommittee and System Protection and Control Subcommittee - "Fault-Induced Delayed Voltage Recovery," Technical reference paper, version 1.2, Princeton, NJ, June 2009.

[83] Z. Yuan, T. Xia, Y. Zhang, et al., "Inter-area oscillation analysis using wide area voltage angle measurements from FNET," in *Proc. IEEE Power Engineering Society General Meeting*, pp. 1-7, 2010.

Vita

Jidong Chai received his B.S. degree from Tianjin University, China, in 2011. He started his Ph.D. study at The University of Tennessee, Knoxville, in August 2011 and his current research interests include wide-area power system monitoring and control, and power system dynamic simulation.