MONITORING OF POWER SYSTEM DYNAMICS USING A HYBRID STATE ESTIMATOR

by
Gabriel Ortiz

Submitted in partial fulfillment of the requirements
for the degree
of
DOCTOR OF ENGINEERING

Faculty of Electrical Engineering and Information Technology
Technical University of Dortmund

Supervisor: Univ. -Prof. Dr. -Ing. habil. Christian Rehtanz
Co-supervisor: Univ. -Prof. Dr. -Ing. Delia Graciela Colomé

Dortmund, July 2021

Abstract

Modern power systems are undergoing a transformation process where distributed energy resources together with complex load technologies are increasingly integrated. This, in addition to a sustained growth in electricity consumption and a lack of significant investment in transmission infrastructure, leads power systems to face with new stochastic operating behavior and dynamics and to operate under stressed conditions. Under such operating conditions, the occurrence of a potential disturbance may cause a partial or a total collapse. Therefore, in order to minimize the risk of collapses and their impact, new monitoring tools must be adopted, capable of providing the right conditions for dynamic wide-area monitoring.

The thesis presents a hybrid state estimator, that is a monitoring tool that combines fast synchronized phasor measurements with traditional measurements into a single scheme. It has the ability to estimate at high speed power system dynamics associated to slow and fast transient phenomena considering a reduced amount of phasor measurement units (PMUs). The developed scheme consists of two phases depending on the power system operating regime. In phase one the system is in stationary regime and bus voltages (magnitude and angle) together with related variables like power flows, current through lines, etc. are estimated by a static estimator at a low speed, which is determined by the supervisory control and data acquisition (SCADA) system. When a physical disturbance happens and the system is in transient regime phase two comes into operation. This time, two estimators work in sequence at high speed. First, a static state estimator is used to estimate bus voltages as soon as the synchronized phasor measurement set arrives. Then, a dynamic estimator is in charge of estimating dynamic states of all generators and motors in the system, even if the unit is not observed by a PMU. Full observability is restored through a novel data-mining based methodology, which defines, first, a PMU topology that allows monitoring the post-contingency bus voltage dynamics of the entire power system and, second, generates a number of bus voltage pseudo-measurements to extend the observability to the whole system.

This page intentionally left blank

Acknowledgments

Finally, after a long time and so much effort this research comes to an end. It has been a long and arduous journey with plenty of doubts and uncertainties. However, today I can certainly say: it was worth taking. As a famous argentine song goes: "*...tarda en llegar y al final, al final hay recompensa...*". I could not have done it without the unconditional support of my family, friends and colleagues. The journey to the completion of this thesis can be summarized in two parts, both different from each other, but equally beautiful and enriching.

It all started in my hometown San Juan, Argentina, at the Institute of Electrical Engineering of the National University of San Juan – CONICET and with the help of my mentor during my first steps as an engineer, Prof. Dr. Graciela Colomé. My sincere thanks to her for all the support given during my work. Her teachings, advices and discipline played a key role in the completion of this thesis as well as in my professional development.

My commitment to my personal growth and the desire to share my work in an international environment led me to meet my mentor in Germany, Prof. Dr. Christian Rehtanz. Many thanks to him for accepting me as a member of his team and trusting me since the beginning. My time as a research assistant at the "ie3" of the Technical University of Dortmund has been a unique and pleasing experience, which gave me the possibility to grow and develop skills I did not even know I had. Many thanks to the "ie3" team, it has been a pleasure working with all of you guys.

Thanks to those friends I made during this journey in Argentina, Ecuador, Colombia, Paraguay, Bolivia and Germany. I have fond memories of each of you.

Thanks to my father, Mario Ortiz, for supporting me and giving me strength at times I thought it would not be possible to finish. Thanks to my mother, Silvana Robles. No matter how far life takes me away, in every step I took you were always by my side. Today, I feel you closer than ever. Thank you for everything mom.

A very special thanks to the two people I love the most. Thanks to my wife, Sol España. Without your unconditional support and patience this work would not be finished. Thanks for leaving so much behind so that this obstinate man can fulfill his dreams. Thanks to my beloved daughter Emma. In your eyes I find the strength to never give up and to believe any dream is possible.

For Sol, Emma and Silvana

Table of contents

1 Introduction

Integrated energy systems (IESs) based on renewable energy resources are considered to be the future of energy systems [1]. Sector coupling is of particular interest because it suggests supplying heat and mobility sectors with non-conventional energy. Heat pumps are increasing in the heat sector. Mobility will rely on alternative energy sources such as batteries. Simultaneously, photovoltaic is being deployed in great amounts [2]. This will lead to a radical change in the nature of the electrical grid behavior bringing more uncertainties to the operation of transmission systems [3]. The variable and uncontrollable nature of renewable energy sources represents a challenge to the secure operation of power systems [4]. In addition, there is a sustained growth in electricity demand and a lack of investment in transmission infrastructure that leads power systems to operate close to their limits. A contingency occurring in the system under such conditions can lead to more complex dynamics, which increases the dynamic insecurity risk and even blackout risk. Experts around the world agree that one of the primary causes of cascading outages that led to blackouts in the past "was due to a lack of information on system conditions and a lack of readiness to take action" [5].

In order to improve the power system stability and security during and after disturbances, new strategies for enhancing operator situational awareness and power grid global and local controllers must be developed. The enhancement of the system stability using new control schemes implies having access to information on system dynamics to perform online dynamic security assessment (DSA) at high speed [6]. However, the system monitoring tool of the Energy Management Systems (EMS) in charge of delivering such information, i.e. the state estimator, is based on a steady state system model, which can only capture very slow dynamics of power systems. This is mainly because the EMS depends on slow update rates of Remote Terminal Units (RTUs) from the Supervisory Control and Data Acquisition (SCADA) system [7]. The sampling rate is too low (2 to 5 seconds) to reveal the electromechanical dynamics in power systems. In addition, RTU-SCADA data are not synchronized [8]. That is why, only local states are used for controlling electromechanical phenomena. "Without global objectives and systematic coordination over wide areas, the influence of local controllers at the grid level may have adverse effects" [9, 10].

Phasor measurement units (PMUs) are the means to improve conventional estimators or even change the paradigm of the estimation process. These devices have been developed since the beginning of the 80s and now they are widely accepted as an additional source of information by most systems around the world [11].

Recent research into wide area monitoring, protection and control schemes has shown that PMUs could significantly increase the performance of state estimators [12] by improving their accuracy and reliability as well as the bad data identification [13, 14]. Their accuracy, reporting speed and synchronization capacity by means of the global positioning system (GPS) make them suitable for global monitoring of the power system dynamics and open-up the possibility of developing distributed state estimation approaches, which have been an active research topic for the last past few years [15].

If enough PMUs could be deployed in the system so that to achieve a PMU-only state estimator, it would be possible to remove the issues related to the slow update rate and time skew. Under such conditions the static assumption would be dismissed. Because voltage and current are linear functions of the state, the estimation problem becomes linear and can be solved in just one iteration. Since PMU data is reported at high speed (60 times a second), a truly dynamic estimate would be available. However, the deployment of such an approach for a large transmission system is still very expensive, where the average overall costs per PMU range between $40,000 and $180,000. Here, communication systems and upgrades are the most significant factor followed by cybersecurity requirements, installation and commission of the unit. The PMU hardware cost represents less than 5% of the total cost [16]. A more viable and economically alternative to exploit the benefits offered by PMUs is a hybrid state estimator that combines synchronized phasor measurements with the existing conventional measurements [12, 17, 18]. Unfortunately, state estimators with a few synchronized measurements usually have the same limitations as conventional estimator regarding dynamics monitoring. That is, they only provide information in the form of a sequence of steady states or quasi-steady states (defined as a set of bus voltage magnitudes and angles). PMU high speed is usually not considered, and thus no benefit is derived from it with respect to dynamics monitoring. Besides, the dynamic state is completely ignored by such schemes.

Therefore, it is required to develop a hybrid state estimator capable of capturing power system dynamics using measurements from RTU-SCADA as well as measurements from a limited number of PMUs installed in the power grid.

Additionally, some challenges that usually must be faced when combining PMU measurements with conventional RTU-SCADA measurements into a hybrid algorithm to obtain an optimal estimated state [17]:

- Numerical stability issues: Current measurements from PMUs tend to provoke ill-conditioning problems in the gain matrix.
- Increased computational burden and data-tsunami: The dimensions of vectors and matrices are increased due to the additional PMU data. In addition, the deployment of PMUs will lead to huge amount of data to be transmitted to control centers.
- Time skewness: Conventional and synchronized phasor measurements are reported at different rates and are not synchronized with each other.

1.1 Research questions

Previously, a number of issues that require further research in the field of hybrid state estimation applied to power system dynamics monitoring were presented. This thesis addresses these issues by answering the following research questions:

i. The number of PMUs to be installed in current power grids is one of the main obstacles when trying to get a hybrid state estimator capable of monitoring power system dynamics. A reduced number of units results in electric networks partially observed by such devices. Under such circumstances: Can full PMU observability be restored so that power system dynamics can be accurately estimated? If that is the case: How many units would be needed and where should they be located (PMU topology)?

ii. How should synchronized phasor measurements be incorporated into a hybrid estimator considering aspects like time skewness and numerical issues?

iii. Which algorithm/s is/are appropriate for solving the dynamic state estimation problem considering the ratio speed-accuracy? In case of needing more than one algorithm: How should they be integrated together in one methodology considering aspects like speed requirements, accuracy and convergence properties?

iv. Which variables should be estimated from a stability and security point of view?

v. In case of adopting a model-based estimator: how should the system be modelled to achieve an acceptable degree of accuracy of results?

vi. It is well known that additional data provided by PMUs together with their high reporting speed contribute to an increase in the computational burden and a saturation of the communication and data processing systems. What measures could be taken to counteract such effects?

1.2 Objectives

Main objective

The main objective of this thesis is to develop a new hybrid state estimator capable of accurately monitoring the power system dynamics associated to slow and fast transient events considering a limited amount of PMUs.

Specific objectives

- To study and analyze with the assistance of data-mining tools the time evolution of bus voltage magnitudes and angles associated to transient events in order to:

a) Identify groups of buses with similar voltage behavior (coherent areas) in magnitude and angle.

b) Define the number of PMUs and their location in the grid (topology) using the knowledge gained in a) so that post-contingency bus voltage dynamics of the entire system can be observed by such units.

c) Define a new methodology for generating voltage pseudo-measurements with the aim of extending the observability to the whole system.

- To study and analyze the estimation algorithms proposed by the scientific community considering accuracy, speed and convergence properties with the aim of defining a feasible option to be used in the proposed estimator.
- To define the state variables that must be estimated considering their potential use in future online dynamic security assessment functions.
- To define the mathematical models to be used by the estimator to represent the behavior of the power system components so that to achieve an acceptable degree of accuracy of estimated results.
- To design an architecture so that the estimator can handle measurements at different reporting rates.

- To merge the hybrid state estimator with the methodology for generating voltage pseudo-measurements into a single methodology.

1.3 Thesis outline

This thesis has been organized as follows:

Chapter 2 presents the theoretical framework of this thesis. A new monitoring tool that informs the power system dynamics will lead to an improvement of the stability and therefore the security as well. For this reason, these two attributes are addressed in first place. Next, the state estimation function and related concepts such as estimation methods and algorithms, observability, bad data processing and multi-area state estimation are defined. After that, an analysis of the state of the art on state estimation with PMUs is carried out. Finally, mathematical definitions and tools used in this thesis are presented.

Chapter 3 presents the approach for defining the number of PMUs and their location and the methodology for generating bus voltage pseudo-measurements. In first place, a data-mining based approach for identifying group of buses with similar post-contingency voltage behavior is presented. Next, the approach that determines the PMU topology in the grid using the knowledge gained before is described. Finally, a classifier capable of predicting coherent areas and the methodology for generating voltage pseudo-measurements are explained in detail.

Chapter 4 presents the hybrid state estimator developed in this thesis for estimating power system dynamics. First, an overview of it is given. Each component is described in detail, their functions and the variables that are estimated. A detailed description of the mathematical models used to represent the behavior of power grid elements such as lines, transformers, synchronous and asynchronous machines and loads is also provided.

In Chapter 5 the proposed hybrid state estimator, the approach that defines the PMU topology and the methodology for generating the pseudo-measurements are evaluated in the New England benchmark system. Several operating scenarios which entail different kind of contingencies are simulated. Results and analysis are given in detail.

Chapter 6 summarizes the conclusions of this thesis, highlights its main contributions and provides an outlook on potential follow-up research topics.

This page intentionally left blank

2 Theoretical framework

Current power systems are characterized by a growing uncertainty of their dynamic behavior in response to disturbances. Sector coupling, i.e. the use of electricity in other sectors aiming at eliminating the need for fossil fuels, is leading to a large-scale integration of renewable energy resources and to the use of more sophisticated electronics-based components. Because of that, power systems are becoming more complex. Additionally, electricity consumption is increasing rapidly. Under such circumstances, results delivered by conventional security assessment functions are not reliable enough. A new scheme to perform an online dynamic security analysis is needed and therefore a new monitoring tool that provide accurate information about system dynamics. Hence, this chapter addresses first power system security concepts. The traditional approach for assessing the power system security and the changes that should be carried out to make an improvement in this sense are presented. Power system stability, as an essential component of power system security, is defined and classified.

Improved analysis tools require new and better monitoring functions. In a second step, the state estimation function for monitoring the system state is addressed. The traditional state estimator is first introduced and its limitations are identified. After that, a review of alternative estimation methods and algorithms as well as related concepts such as observability, bad data processing and multi-area state estimation are performed. Next, the PMU technology and its potential benefits and current limitations for practical use are introduced. An analysis of the state of the art on state estimation using PMU measurements is carried out.

The proposed approaches for defining the PMU topology and the pseudo-measurements involve an analysis of post-contingency bus voltage behavior aiming to gain useful knowledge from it. In this thesis, data-mining techniques capable of performing such tasks are used. In the last part of this chapter such techniques are introduced and explained.

2.1 Security concepts

This subsection reviews the main concepts involved in power system security. The term dynamic security assessment is explained. Besides, the relationship between security and stability is discussed. Here, the different categories of power system instability together with the devices and the electrical quantities involved in of each of them are identified and described.

2.1.1 Security assessment

Power system security is related to the probability of supplying the energy demand without interruption. The more secure the system, the lower the probability of loss of load [19]. Security analysis aims at determining the robustness of the power system under imminent disturbances. According to [20] there are two main components in the security analysis: “For a power system subjected to changes (small or large), it is important that when the changes are completed, the system settles to new operating conditions such that no physical constraints are violated. This implies that, in addition to the next operating conditions being acceptable, the system must survive the transition to these conditions”. Dynamic security assessment will find out if the system can meet certain security specifications in dynamic and steady state and for all credible contingencies. A secure system complies with pre-defined operating specifications before and after the contingency. Because of this, an analysis involving all aspects of the system security is mandatory. Line loadings, bus voltages, frequency deviations and all kind of stability must be assessed. Security assessment involves high computational resources. That is why, it has been carried out in “an offline operation-planning environment” where the system performance (in static and dynamic condition) considering short-term forecasted scenarios is evaluated by means of power flow calculations and time-domain simulations [21].

Power systems were more secure in times of regulated and vertically integrated electricity markets. This was due to the fact that those systems were designed, built and operated only by one entity (monopoly). Integrated system planning allowed generation and transmission systems to keep up with the load increase, thereby avoiding overloading and equipment failures that could lead to system disturbances. System maintenance were also more strict. Forecasting system conditions was simpler because there were fewer market players involved in the generation and transmission sectors and they “were operating in a carefully planned and cooperative manner” [21]. As a result, systems were more robust when being subjected to disturbances and were more predictable in their behavior. Nevertheless, the evolution of electricity markets over the last decade has introduced a number of factors that increases the possible sources for system disturbances, reduced the system robustness and predictability [21].

Table 2.1 Factors that currently contribute to a reduction in the system security [21]

Factor	Potential consequence
Aging transmission infrastructure	▪ Increased probability of component failures and malfunction leading to system disturbance
Lack of new transmission facilities	▪ Overloading of transmission facilities leading to protection operation or contributing to phenomena such as voltage collapse ▪ Bottlenecks in key transmission corridors leading to congestion
Increased dependence on controls and special protection systems	▪ Increased probability of inadvertent/incorrect operation of protections ▪ Increased unpredictability of cascading events
Large number of small distributed generators	▪ Increased difficulty in adequate system design due to uncertainty generation plans ▪ Uncertainty in dispatch
New technologies such as advanced control systems, wind power, fuel cells, etc.	▪ Lack of operating experience with technologies which may have unique dynamics characteristics ▪ Unpredictable behaviors during disturbances
Trend toward interconnection	▪ Exposure to cascading disturbances brought on by events in neighboring systems

In today's competitive markets, the uncertainty about future operating conditions leads to the need for an online approach in order to evaluate the system security: the online dynamic security assessment. This new method analyzes the system stability for the current operating condition with enough speed so that automatic control actions can be performed if the contingency under analysis is found to be insecure. Because the evaluation is performed on the current system operating condition, the uncertainty of the offline analysis using forecasted scenarios is removed. Such an approach would provide early warning to operators so that remedial actions can be taken [21]. The quality of the online dynamic security assessment and system operation is directly affected by the speed, accuracy and reliability of the state estimator [22]. Thus, power system state estimation constitutes the core of the online dynamic security analysis function.

2.1.2 Power system stability

Stability analysis is an integral component of system security assessment. A Task Force, set up jointly by the CIGRE Study Committee 38 and the IEEE Power System Dynamic Performance Committee, has defined the power system stability as "the ability of an electric power system, for a given initial operating condition, to regain a state of operating equilibrium after being subjected to a physical disturbance, with most system variables bounded so that practically the

entire system remains intact". This definition does not take into consideration the different types of stability. The analysis can be simplified when power system stability is classified into appropriate categories [23].

Stability can be classified by focusing on only one variable (i.e. voltage, frequency or rotor angle). Such a classification is known as partial stability and is based on the following assumptions [20, 23]:

- The nature of the resulting mode of instability as indicated by the main variable in which instability can be detected.
- The size of the disturbance, which defines the method of calculation.
- The devices, processes, and the time span that must be considered in order to evaluate stability.

Rotor angle stability

Rotor angle stability is the ability of synchronous machines to remain in synchronism after being subjected to a disturbance. It is defined by the ability to conserve/restore the balance between electromagnetic and mechanical torque. The system becomes unstable when the angular swings of some generators increase until reaching loss of synchronism [20].

In stationary regime, the mechanical and electromagnetic torque of each generator are equal, and the speed remains constant. When the system is disturbed, this balance is upset, which leads to acceleration or deceleration of the machine shaft according to the laws of motion of a rotating body. The rotor angular position of the faster generators will advance the position of the slower machines. This angular difference causes part of the load to be transferred from the slow generator to the fast one, depending on the power-angle relationship (highly nonlinear), which reduces the speed difference and hence the angular distance. From a certain limit, an increase in the angular distance leads to a decrease of power transfer so that the angular separation rises even more. Instability will occur when the system is not capable of absorbing the kinetic energy related to the rotor speed differences. Thus, stability depends on whether or not deviations in angular positions provoke sufficient restoring torques [20].

According to [20], angle stability can be classified in the following two subcategories

- Small-disturbance angle stability is concerned with the ability of the system to remain in synchronism under small disturbances. The disturbances are considered to be small enough

to linearize the system of equations for analysis purposes. The time frame of interest is in the order of 10 to 20 seconds following the disturbance [20].

- Large-disturbance angle stability or transient stability is concerned with the ability of the system to remain in synchronism after the occurrence of a large system upset, such as a short circuit. “The system response involves large excursions of generator rotor angles and is influenced by the nonlinear power-angle relationship”. The time frame of interest goes from 3 to 5 seconds after the disturbance [20].

Voltage Stability

Voltage stability is defined as the power system ability to keep acceptable voltage levels under normal operating conditions and following a disturbance. Imbalance between load demand and load supply may lead to voltage instability. Voltage drops through the inductances in transmission grids contribute to voltage stability issues. Power flowing through inductive reactance limits the power transfer capacity of the network and the voltage support. Reactive power limits of generators and loads and compensation devices characteristics represent also a key factor that contributes to voltage instability [20, 23, 24]. Some measures that may help to restore the voltage level and to prevent the voltage collapse are load shedding, transformer tap changing, generator excitation forcing, etc. Voltage can be controlled by means of reactive power support. As stated by [23], “the voltage stability conditions are met if the reactive power injection increase in one bus will lead to voltage increase in the same bus. In other words, a system is voltage stable if the V-Q sensitivity is positive for every bus and voltage unstable if the V-Q sensitivity is negative for at least one bus”.

According to [20], voltage stability can be classified into the following subcategories:

- Large-disturbance voltage stability is defined by the system’s ability to maintain voltage levels within boundaries after large disturbances. Its analysis involves the evaluation of the system nonlinear response over a certain period of time so that the performance and interactions of devices such as motors can be detected. The time frame of interest goes from a few seconds to tens of minutes.
- Small-disturbance voltage stability is defined by the system’s ability to maintain voltage levels within limits when subjected to small upsets such as small changes in the system load. Under such circumstances, the system equations can be linearized for analysis purposes.

Voltage stability may be either a short-term or a long-term phenomenon [20]:

- Short-term voltage stability "involves dynamics of fast acting load components such as induction motors, electronically controlled loads, and HVDC converters". The time frame of interest is in the order of seconds and its assessment entails the solution of differential equations. The load behavior must be represented in detail using dynamics models.
- Long-term voltage stability "involves slower acting equipment such as tap-changing transformers and thermostatically controlled loads". The time frame of interest may cover several minutes. Its analysis entails long-term simulations. Static analysis can be employed for defining the stability margins.

Frequency stability

Frequency stability is defined by the system ability to maintain steady frequency after the occurrence of a large disturbance that leads to an important imbalance between generation and demand. Frequency stability can be ensured if the system is capable of keeping the balance between generation and load, with minimum amount of load being unintentional disconnected. When the system becomes unstable the frequency either suddenly drops or swings permanently. This phenomenon will lead eventually to generation outage and/or load curtailment [20]. As stated in [20], "frequency stability problems are generally associated with inadequacies in equipment responses, poor coordination of control and protection equipment, or insufficient generation reserve".

Times behind the processes and equipment involved in frequency stability goes from fraction of seconds (under frequency load shedding and generator controls and protections) to several minutes (response of equipment such as prime mover systems). That is why, frequency stability may be either a short-term or a long-term phenomenon [20].

2.2 Power system monitoring: state estimation

At the beginning, power systems were monitored only by supervisory control systems. They basically monitor and control circuit breakers, generator outputs and the system frequency. Then, these systems were improved with real-time wide-data acquisition capabilities, which allowed control centers to collect all measurements from the power system. This led to the creation of the first SCADA system. The main goal behind this development was the possibility of performing a security analysis. Nevertheless, errors in the measurement set and communication noises, to name a few examples, may result in unreliable information from the SCADA

system. In addition, the system state may not be accessible from the set of measurements. For example, voltage angles are not usually measured, and transmission line flows are only known in some lines. Metering all possible measurements is not economically feasible. Professor Schweppe was the one who identified all these limitations and proposed the state estimation function. The state estimation function augmented the capabilities of the SCADA system and led to the creation of the EMS, a system that would now be equipped with a number of specific applications, one of them is the state estimator [25].

State estimator allow identifying the current operating state by monitoring with efficiency quantities such voltage magnitudes and line loadings. The security assessment function uses this information in order to analyze contingencies and to determine any remedial actions [25]. State estimator usually include the following functions [25]:

- Topology processor: it collects data on the circuit breakers and defines the single line diagram of the system.
- Observability analysis: it defines if the estimation problem for the entire system can be solved using the available set of measurements.
- State estimation: it finds the system state, i.e. bus voltages in magnitude and angle for the entire power system, based on the network model and the available set of measurements. In addition, it finds the line flows, loads, transformer taps, and generator outputs.
- Bad data processing: it detects and identifies gross errors in the measurement set in order to remove them.

Limitations on current state estimators

Traditional state estimators currently in use do not meet the necessary requirements for the implementation of an online dynamic security assessment. This is mainly due to the following reasons [15]:

- Measurements used by traditional estimators are delivered by SCADA systems from RTUs at a low update rates (2 to 5 seconds) to capture the dynamics associated with transient events. Besides, “they are not accurately time-stamped, in a unified way, independent of the measurement location, and thus are not synchronized”. This can lead to “significant biases in the estimation, and may even be practically useless in a dynamic estimation framework”.

- "No dynamics are included in the models, since most of even the slower system dynamics have time scales of the order of milliseconds or at most seconds, and usually a very limited amount of system dynamic data is available by electric utilities".

Conventional approaches for estimating the power system state are not capable of capturing the system dynamics (with the exception of some slow system changes). They only give information "in the form of a sequence of steady states, or quasi-steady states" (bus voltage magnitudes and angles). That is why, wide area control actions are limited to slow steady state control that is manually executed by system operators. Fast automatic control in transient regime is provided only locally without considering the wide area system behavior [15].

From recent major power system outages emerged the need for more efficient methods of monitoring the state of power systems and in particular the stability of the system. Experts from around the world agree that one of the primary causes of cascading outages on recent blackouts "was due to a lack of information on system conditions and a lack of readiness to take action". This particular issue can be addressed with a better monitoring tool. The concept of a dynamic wide-area monitoring scheme will provide additional real-time information. Besides, it will help to understand the reason for the blackout and will enable a postmortem analysis. Having access to real-time information about the actual conditions of the electrical grid would allow detecting and recognizing more easily emergency conditions. Then, a better analysis could be carried out and remedial actions could be taken in a "quicker and more controlled fashion" [5].

2.2.1 Methods

Thinking about the term "state", one can imagine that it gives a "complete summary" about the system condition at a given time instant. By conventional estimators, it is supposed that the power system is working under normal condition, known as quasi-stationary regime. This means that the system is subjected to variations coming from small and smooth load changes and the corresponding adjustments in generation. Under these assumptions, the system operating condition is fully defined at a certain time instant by variables such as bus voltages and related variables like power flows, generator outputs, etc. This state is known as static-state and its time evolution (a sequence of stable steady-state conditions) is usually observed through measurements acquired by RTUs. "From the dynamic point of view, these devices have intrinsic limitations (no time stamps and low sampling rates)" [26]. Methods used to solve the problem in such a cases are known as static state estimation methods. "This mean that, whenever a

new measurement set is processed by the estimator, previously estimated states are not considered to obtain an estimate of the current system state. In other words, estimated states are extracted from a single scan of measurements viewed as a true snapshot of the system state" [26]. The following nonlinear measurement model is used:

$$\mathbf{z} = \mathbf{h(x)} + \mathbf{e} \tag{2.1}$$

where **z** is the measurement, **x** is the true state, **h** is a nonlinear vector function relating measurements to states, and **e** is the measurement error [19] whose components are commonly assumed to have Gaussian (Normal) distribution with zero mean and variance σ^2.

The dynamic state estimation method gets information not only from current measurements but also from past measurements. It is assumed that the system may experience a sudden disturbance. Such a condition is known as transient-regime and electromechanical transients are considered. Hence, the system is mathematically modelled by the following differential algebraic equations (DAEs) [27]:

$$\dot{\mathbf{x}}=\mathbf{f}\,(\mathbf{x}_{(t)}, \mathbf{y}_{(t)}, \mathbf{u}_{(t)}, \mathbf{p}) \tag{2.2}$$

$$\mathbf{0}=\mathbf{g}(\mathbf{x}_{(t)}, \mathbf{y}_{(t)}, \mathbf{u}_{(t)}, \mathbf{p}) \tag{2.3}$$

where **x** represents the dynamic-state, such as the internal states of a machine or a dynamic load, etc.; **y** represents the algebraic state, such as bus voltages and currents; **u** is the system input; **p** represent model parameters; and **f** and **g** are nonlinear functions. Through discretization of the continuous-time model and considering the equality constraints in (2.3) as pseudo-measurement and processing them with the incoming measurements, a "more general state space model for dynamic state estimation" is [27]:

$$\mathbf{x}_k=\mathbf{f}\,(\mathbf{x}_{k-1}, \mathbf{y}_{k-1},\mathbf{u}_k, \mathbf{p})+\mathbf{w}_k \tag{2.4}$$

$$\mathbf{z}_k=\mathbf{h}(\mathbf{x}_k, \mathbf{u}_k, \mathbf{p})+\mathbf{e}_k \tag{2.5}$$

where $\mathbf{w}_k$ is an error that includes discretization and model approximation errors, $\mathbf{z}_k$ is the measurement, including pseudo-measurements, measured algebraic variables, etc.; **h** is the nonlinear function of measurements; and $\mathbf{e}_k$ is the measurement error.

The forecasting-aided state estimation method is a particular application of the dynamic estimator concept to quasi-stationary state regime, "in which the state-transition model is driven only by slow enough stochastic changes in the power injections and the dynamics of $\mathbf{x}_k$ are

sufficiently small to be neglected" [27]. In the forecasting-aided method, the state-transition model in (2.4) is assumed linear, which leads to [27]:

$$\mathbf{y}_k = \mathbf{A}_k \mathbf{y}_{k-1} + \zeta_{k-1} + \mathbf{w}_k \tag{2.6}$$

$$\mathbf{z}_k = \mathbf{h}(\mathbf{y}_k, \mathbf{p}) + \mathbf{e}_k \tag{2.7}$$

where $\mathbf{y}_k$ "represents algebraic state variables that specifically refer to the bus voltage magnitudes and angles". The transition matrix $\mathbf{A}_k$ and the trend vector ζ_{k-1} are defined from historical time series data. When the input or trend vector changes smoothly, the forecasting-aided method delivers good results. Nevertheless, its performance can be degraded in presence of abrupt changes as "the state transitions coefficients take a while to adapt to the new situation" [27].

Assuming that the state transition matrix $\mathbf{A}_k$ is an identity matrix and the change in the state vector from one time instant to the next one is small enough, the forecasting-aided method simplifies to the tracking state estimation method. In this case the model used is [27]:

$$\mathbf{y_k} = \mathbf{y}_{k-1} + \mathbf{w}_k \tag{2.8}$$

$$\mathbf{z_k} = \mathbf{h}(\mathbf{y}_k, \mathbf{p}) + \mathbf{e}_k \tag{2.9}$$

The tracking approach "assumes a quasi-stationary state regime, where the system state remains unchanged other than an additive Gaussian noise $\mathbf{w}_k$". One main drawback of the tracking method is "the lack of an appropriate model to represent the system dynamics". This becomes more relevant in today's power system where the large deployment of renewable energy sources leads to unpredictable behavior of the power system. Hence, adopting tracking methods for practical applications is far from becoming real [27].

2.2.2 Algorithms

This section introduces and explains some of the most commonly used state estimation algorithms.

Weighted least squares

According to [25], "most conventional state estimators in practical use are formulated as overdetermined systems of nonlinear equations and solved as weighted least square (WLS) problems". Considering the nonlinear measurement model (2.1) and assuming m measurements and n state variables, where n<m, the WLS algorithm will minimize the following objective function [25]:

$$J(\mathbf{x})=\sum_{i=1}^{m}\frac{(z_i - h_i(\mathbf{x}))^2}{R_{ii}}=[\mathbf{z}-\mathbf{h(x)}]^T\,\mathbf{R}^{-1}\,[\mathbf{z}-\mathbf{h(x)}] \tag{2.10}$$

It is assumed that measurement errors are independent from each other and they follow a Gaussian distribution with zero mean and variance σ^2. Under such assumptions the error covariance is:

$$\mathrm{Cov}(\mathbf{e})=\mathbf{R}=\mathrm{diag}\left\{\sigma_1^2, \sigma_2^2, \ldots, \sigma_m^2\right\} \tag{2.11}$$

The standard deviation σ_i of each measurement error i is calculated to reflect the expected accuracy of the corresponding meter. At the minimum, the first order optimality must be satisfied. These can be expressed as follows:

$$\begin{gathered}\mathbf{g(x)}=\frac{\partial J(\mathbf{x})}{\partial \mathbf{x}}=-\mathbf{H}^T(\mathbf{x})\,\mathbf{R}^{-1}[\mathbf{z}-\mathbf{h}(\mathbf{x})]=\mathbf{0}\\ \mathbf{H(x)}=\frac{\partial \mathbf{h(x)}}{\partial \mathbf{x}}\end{gathered} \tag{2.12}$$

Expanding the nonlinear function $\mathbf{g}(\mathbf{x})$ into its Taylor series around the state vector $\mathbf{x}^k$ yields:

$$\mathbf{g(x)}=\mathbf{g}(\mathbf{x}^k)+\mathbf{G}(\mathbf{x}^k)(\mathbf{x}-\mathbf{x}^k)+\ldots=\mathbf{0} \tag{2.13}$$

Neglecting the higher order terms leads to an iterative solution scheme known as the Gauss-Newton method:

$$\begin{gathered}\mathbf{x}^{k+1}=\mathbf{x}^k-\left[\mathbf{G}(\mathbf{x}^k)\right]^{-1}\mathbf{g}(\mathbf{x}^k)\\ \mathbf{G}(\mathbf{x}^k)=\frac{\partial \mathbf{g}(\mathbf{x}^k)}{\partial \mathbf{x}}=\mathbf{H}^T(\mathbf{x}^k)\mathbf{R}^{-1}\mathbf{H}(\mathbf{x}^k)\\ \mathbf{g}(\mathbf{x}^k)=-\,\mathbf{H}^T(\mathbf{x}^k)\mathbf{R}^{-1}\left[\mathbf{z}-\mathbf{h}(\mathbf{x}^k)\right]\end{gathered} \tag{2.14}$$

$\mathbf{G}(\mathbf{x})$ is called the gain matrix. It is sparse, positive definite and symmetric if the system is fully observable. It is decomposed into its triangular factors and the following sparse linear set of equations are solved using forward/back substitutions at each iteration k:

$$\begin{gathered}\left[\mathbf{G}(\mathbf{x}^k)\right]\Delta\mathbf{x}^{k+1}=\mathbf{H}^T(\mathbf{x}^k)\mathbf{R}^{-1}[\mathbf{z}-\mathbf{h}(\mathbf{x}^k)]\\ \Delta\mathbf{x}^{k+1}=\mathbf{x}^{k+1}-\mathbf{x}^k\end{gathered} \tag{2.15}$$

The set of equations given by (2.15) is known as the normal equations. The WLS algorithm is used to estimate the static state of the system.

M-Estimators

The solution obtained with the WLS method will be affected under the presence of bad data (outliers). The objective function is a proof of this (residuals have a quadratic influence in the objective function) [19]. In this sense, the WLS is not robust and the accuracy will be degraded even in the presence of a single outlier. Hence, the estimation process is followed by a post-estimation bad data processor that will detect, identify, and correct any existing outlier [28].

M-Estimators involve a number of robust estimators. They can automatically detect measurements with growing residuals and remove their influence on the result. An M-Estimator is a maximum likelihood estimator that minimizes a non-quadratic objective function subjected to constraints given by the measurements equations [25]:

$$\begin{aligned} &\text{Minimize} \quad \sum_{i=1}^{m} \rho(r_i) \\ &\text{Subject to} \quad \mathbf{z}=\mathbf{h}(\mathbf{x})+\mathbf{r} \end{aligned} \tag{2.16}$$

where $\rho(r_i)$ is a function of the measurement residual, $\mathbf{z}$ is the measurement, $\mathbf{x}$ is the state and $\mathbf{h}(\mathbf{x})$ is the measurement function. As stated in [25], "the influence of bad data on the estimated state and methods of their suppression can be addressed by changing the estimation algorithm in such a way that the bad measurements will be screened during the iterative process". This concept produced several M-estimators. The Least Absolute Value (LAV) estimator represents an option among them.

Considering the following linear regression model [25]:

$$z_i=\mathbf{d}_i^T\mathbf{x}+e_i \tag{2.17}$$

where, a set of observations $\{z_i, i=1,\ldots,m\}$ are assumed to be linearly related to a set of vectors $\{\mathbf{d}_i \in R^n, i=1,\ldots,m\}$ and an unknown vector $\mathbf{x} \in R^n$, and each observation i contains some random error e_i. Then, the LAV estimate $\hat{\mathbf{x}}$ for the unknown vector $\mathbf{x}$ will be obtained by the solution of the following optimization problem [25]:

$$\begin{aligned} &\text{Minimize } \mathbf{c}^T|\mathbf{r}| \\ &\text{Subject to } \mathbf{z}\text{-}\mathbf{D}\mathbf{x}=\mathbf{r} \end{aligned} \tag{2.18}$$

where, $\mathbf{D}$ is a m x n matrix with $\mathbf{d}_i^T$ being its i-th row, $\mathbf{c} \in R^m$ is a vector with all of its entries equal to 1, and $\mathbf{r} \in R^m$ is the observation residual.

LAV estimators can be applied using linear programming solvers and they can automatically process large errors and remove them in the same estimation process. Despite these qualities, LAV estimators have not been implemented in the past for two reasons: in the first place, the estimation problem is nonlinear and computationally inefficient when using conventional measurements and, secondly, leverage points in measurements make the LAV estimator fail to automatically remove big errors [29]. A measurement is called a "leverage point if it lies far away from bulk of the rest of the measurements in the sample space". They are similar to critical measurements in the sense they "force the estimated state to closely satisfy their values and thus biasing the estimated state if they carry gross errors" [29]. As in the case of the WLS, LAV estimators are employed to estimate the static state.

Discrete-time Kalman filter

Kalman filters (KFs) are the right choice for systems in change. Their efficiency and high speed make them appropriate for solving real-time problems and embedded systems. KFs are recursive estimators, which means they only use the estimated state at the previous time and the current measurement to estimate the system state [30]. They are employed to estimate the dynamic state of a system.

The KF consists of two phases: "Predict" and "Update". The predict phase uses the estimated state from the previous time step to get an estimate at the current time step. This state is known as the "a priori estimated state" since the measurement set from the current time step is not considered. In the update phase, the a priori estimated state is combined with the current measurement set to "refine" the estimated state. This improved solution is known as the "a posteriori estimated state" [30]. "The KF operates by propagating the mean and covariance of the state through time" [31].

The standard KF assumes that the state x at time k evolves from time k-1 according to the following linear discrete-time equation [30]:

$$\mathbf{x}_k = \mathbf{F}_{k-1}\,\mathbf{x}_{k-1} + \mathbf{U}_{k-1}\,\mathbf{u}_{k-1} + \mathbf{w}_{k-1} \tag{2.19}$$

where $\mathbf{x}_k$ is the true state at time k, $\mathbf{F}_{k-1}$ and $\mathbf{U}_{k-1}$ are the transition and input matrices respectively at time k-1 (considering time-variant system), $\mathbf{u}_{k-1}$ is the input at time k-1 and $\mathbf{w}_{k-1}$ is a Gaussian zero-mean white noise with covariance $\mathbf{Q}_{k-1}$. At time k, a measurement z_k is obtained with the following equation:

$$\mathbf{z}_k = \mathbf{H}_k\,\mathbf{x}_k + \mathbf{e}_k \tag{2.20}$$

where $\mathbf{H}_k$ is the measurement matrix and $\mathbf{e}_k$ is a noise measurement, which follows a Gaussian zero-mean distribution with covariance $\mathbf{R}_k$. Both noises $\mathbf{w}_k$ and $\mathbf{e}_k$ are assumed uncorrelated.

The filter is initialized defining the initial estimated state $\mathbf{x}_0$. It is named $\hat{\mathbf{x}}_0^+$, where the "+" superscript denotes that the estimated state is "a posteriori", despite not having any available measurement yet, since the first one is taken at t = k. The covariance of the initial estimated state $\mathbf{P}_0^+$ it is also defined here [30, 31]:

$$\begin{aligned} \hat{\mathbf{x}}_0^+ &= \mathrm{E}(\mathbf{x}_0) \\ \mathbf{P}_0^+ &= \mathrm{E}\left[(\mathbf{x}_0 - \hat{\mathbf{x}}_0^+)\,(\mathbf{x}_0 - \hat{\mathbf{x}}_0^+)^{\mathrm{T}}\right] \end{aligned} \tag{2.21}$$

If the initial state is perfectly known, then $\mathbf{P}_0^+ = \mathbf{0}$, otherwise $\mathbf{P}_0^+ \gg \mathbf{I}$. When time k arrives, before processing the measurements, covariance and state are predicted using the following general time-update equations [30, 31]:

$$\begin{aligned} \hat{\mathbf{x}}_k^- &= \mathbf{F}_{k-1}\hat{\mathbf{x}}_{k-1}^+ + \mathbf{U}_{k-1}\,\mathbf{u}_{k-1} \\ \mathbf{P}_k^- &= \mathbf{F}_{k-1}\mathbf{P}_{k-1}^+\mathbf{F}_{k-1}^{\mathrm{T}} + \mathbf{Q}_{k-1} \end{aligned} \tag{2.22}$$

where the "-" superscript denotes that the estimate is a priori. Because in this step there is not any additional measurement, the state is defined using the knowledge from the system dynamics. Next, when the measurements at time k are finally available, the estimated state and its covariance are updated using the measurement-update equations, which are obtained from the recursive least squares estimator [30, 31]:

$$\begin{aligned} \mathbf{K}_k &= \mathbf{P}_k^-\mathbf{H}_k^{\mathrm{T}}(\mathbf{H}_k\mathbf{P}_k^-\mathbf{H}_k^{\mathrm{T}} + \mathbf{R}_k)^{-1} \\ \hat{\mathbf{x}}_k^+ &= \hat{\mathbf{x}}_k^- + \mathbf{K}_k(\mathbf{z}_k - \mathbf{H}_k\hat{\mathbf{x}}_k^-) \\ \mathbf{P}_k^+ &= (\mathrm{I} - \mathbf{K}_k\mathbf{H}_k)\mathbf{P}_k^- \end{aligned} \tag{2.23}$$

The matrix $\mathbf{K}_k$ in the above equations is called the Kalman filter gain.

The standard KF assumes that the system model is linear and well known, the process and measurement noises are unrelated, they follow a Gaussian distribution with zero mean and known covariance matrices. Nevertheless, in the reality these assumptions are seldom fulfilled [30]. That is why, some variants have been proposed in the literature. For instance, the extended Kalman filter (EKF) represents a suitable alternative to nonlinear problems based on the Taylor expansion. Unscented Kalman filter (UKF) and particle filter (PF) are appropriate for nonlinear and non-Gaussian systems [30].

Since the EKF is one of the algorithms used in this thesis, it will be briefly introduced and explained below. For further information regarding UKF and PF, readers can refer [30, 31].

Discrete-time extended Kalman filter

In case of nonlinear system models, the EKF can be used to estimate the dynamic state of the system. To achieve this, "the system is linearized around the nominal state trajectory, which is obtained from the KF estimate. This is a sort of a bootstrap method. The nonlinear system is linearized around the KF estimate, and the KF estimate is based on the linearized system" [31].

Assuming the following general and nonlinear discrete-time system model [31]:

$$\begin{gathered}\mathbf{x}_k=\mathbf{f}_{k-1}(\mathbf{x}_{k-1},\mathbf{u}_{k-1},\mathbf{w}_{k-1})\\ \mathbf{z}_k=\mathbf{h}_k(\mathbf{x}_k,\mathbf{e}_k)\\ \mathbf{w}_k\sim(0,\mathbf{Q}_k)\\ \mathbf{v}_k\sim(0,\mathbf{R}_k)\end{gathered} \tag{2.24}$$

The filter is initialized as follows [31]:

$$\begin{gathered}\hat{\mathbf{x}}_0^+=\mathrm{E}(\mathbf{x}_0)\\ \mathbf{P}_0^+=\mathrm{E}\left[(\mathbf{x}_0-\hat{\mathbf{x}}_0^+)\,(\mathbf{x}_0-\hat{\mathbf{x}}_0^+)^{\mathrm{T}}\right]\end{gathered} \tag{2.25}$$

Then, for k = 1, 2,…, the following operations are performed [31]:

1) Matrices from partial derivatives

$$\begin{gathered}\mathbf{F}_{k-1}=\left.\frac{\partial\mathbf{f}_{k-1}}{\partial\mathbf{x}}\right|_{\hat{\mathbf{x}}_{k-1}^+}\\ \mathbf{L}_{k-1}=\left.\frac{\partial\mathbf{f}_{k-1}}{\partial\mathbf{w}}\right|_{\hat{\mathbf{x}}_{k-1}^+}\end{gathered} \tag{2.26}$$

2) Time update of the state estimate and its covariance

$$\begin{gathered}\hat{\mathbf{x}}_k^-=\mathbf{f}_{k-1}(\hat{\mathbf{x}}_{k-1},\mathbf{u}_{k-1},0)\\ \mathbf{P}_k^-=\mathbf{F}_{k-1}\mathbf{P}_{k-1}^+\mathbf{F}_{k-1}^{\mathrm{T}}+\mathbf{L}_{k-1}\mathbf{Q}_{k-1}\mathbf{L}_{k-1}^{\mathrm{T}}\end{gathered} \tag{2.27}$$

3) Matrices from partial derivatives

$$\mathbf{H}_k = \frac{\partial \mathbf{h}_k}{\partial \mathbf{x}}\Big|_{\hat{\mathbf{x}}_k^-} \qquad \mathbf{M}_k = \frac{\partial \mathbf{h}_k}{\partial \mathbf{e}}\Big|_{\hat{\mathbf{x}}_k^-} \tag{2.28}$$

4) Measurement update of the state estimate and its covariance

$$\mathbf{K}_k = \mathbf{P}_k^- \mathbf{H}_k^T (\mathbf{H}_k \mathbf{P}_k^- \mathbf{H}_k^T + \mathbf{M}_k \mathbf{R}_k \mathbf{M}_k^T)^{-1}$$
$$\hat{\mathbf{x}}_k^+ = \hat{\mathbf{x}}_k^- + \mathbf{K}_k (\mathbf{z}_k - \mathbf{h}_k(\hat{\mathbf{x}}_k^-, 0)) \tag{2.29}$$
$$\mathbf{P}_k^+ = (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \mathbf{P}_k^-$$

2.2.3 Observability

As stated in [27], "observability is defined as the ability to determine uniquely the system state from available measurements". Depending on the system to be estimated and the estimation method adopted observability analysis will be different in nature.

Observability analysis for static state estimation

In relation to traditional estimators, a network observability analysis is usually performed prior to state estimation with the aim of determining if the number of measurements and their location make state estimation possible. When there are not enough measurements, observability analysis allows finding which parts of the network contain states that can still be estimated, as well as where pseudo-measurements can be added to improve observability.

The relationship between the number of measurements and the number of the states is known as measurement redundancy. In practice, its value usually is in the range 1.7 - 2.2. A critical measurement is a non-redundant measurement and will make the system unobservable if it is removed [32].

There are three main types of approaches to perform observability analysis for static state estimation problems: topological, numerical and hybrid [32]:

a) The topological observability method was thought for networks represented by bus/branch models and is based on the concept of maximal spanning tree. A maximal spanning tree contains every node in the graph representing a network. This method does not necessarily ensure that the estimation problem can be solved [32].

b) Numerical approaches can be based on fully coupled or decoupled models. They operate by factorizing the measurement Jacobian or the gain matrices and are formulated using either branch or nodal variables [25].

c) Hybrid approaches benefit from the best of both, topological and numerical methods. As stated in [32], "a basic topological algorithm with simple injection conversion to obtain one or more islands, which are as large as possible, and a numerical algorithm for application to the reduced system".

Observability analysis for dynamic state estimation

A dynamic system may be highly or poorly observable. Since the problem is nonlinear and time-dependent, observability will be also time-dependent. As regards linear time-invariant dynamic system, observability can be determined according to the rank of the observability matrix (a full rank matrix denotes an observed system). In relation to nonlinear system, the system response can be linearized around a given operating point and then observability can be locally evaluated. Despite such an approach may be computationally efficient, it will be inaccurate under "highly nonlinear operating conditions" [27]. The following methods do not perform any linearization:

a) Lie derivatives: a nonlinear dynamic system is observable at $\mathbf{x}_0$ if the observability matrix obtained by using Lie derivatives at $\mathbf{x} = \mathbf{x}_0$ has a full rank [27].

b) Empirical observability Gramian: the empirical observability Gramian allows analyzing the output behavior of a nonlinear system. It is defined over an operating region and gives the observability of the full nonlinear dynamics in the given domain [27].

2.2.4 Bad data processing

One main function of the state estimator is to detect errors and to identify and eliminate them if possible. Errors in measurements are generally due to the "finite accuracy" of the meters. If the measurement redundancy is enough, such errors will be filtered by the estimator. Big errors can occur because the meters have "biases, drifts or wrong connections" [25].

When using the WLS algorithm, detection and identification of errors are performed only after the estimation approach by processing the measurement residuals [25].

Bad data can be classified as [25]:

a) Single bad data: only one measurement is affected by a large error.

b) Multiple bad data: more than one measurement is affected by a large error.

Bad data processing involves two processes: detection and identification. The detection process defines whether or not the measurement set contains any error. The identification process finds which specific measurements actually have bad data. Bad data can be detected as long as their removal does not make the system unobservable. This means, bad data in critical measurements cannot be detected [25, 33].

Next, some of the most known methods to detect and identify bad data in conventional estimators will be introduced and briefly explained.

Detection

An approach used for detecting bad data is the Chi-squares test based on the properties of the χ^2 distribution. The method involves the following steps [25]:

1) Solve the WLS problem and compute the objective function $J(\hat{\boldsymbol{x}})$ in (2.10)
2) Obtain the value corresponding to a detection confidence with probability p (e.g. 95 %) and m-n degree of freedom from a table containing χ^2 cumulative distribution function values for different degree of freedom, where m is the number of measurements and n is the number of states. The value will be $\chi^2_{(m-n),\,p}$
3) Test if $J(\hat{\mathbf{x}}) \geq \chi^2_{(m-n),\,p}$. If yes, then bad data will be suspected. Else, measurements will be assumed free of bad data

Normalized residuals can also be used for detecting bad data. The normalized residual r for the measurement i can be obtained by dividing its absolute value by the corresponding diagonal entry in the residual covariance matrix [25]:

$$r_i^N = \frac{|r_i|}{\sqrt{\Omega_{ii}}} \qquad (2.30)$$

$$\boldsymbol{\Omega} = \mathbf{RS}$$

Where **S** is called sensitivity matrix and represents the sensitivity of the measurement residuals to the measurement errors [25]:

$$\mathbf{S} = \boldsymbol{I} - (\mathbf{H}\mathbf{G}^{-1}\mathbf{H}^{T}\mathbf{R}^{-1}) \qquad (2.31)$$

The largest element in $\mathbf{r}^N$ can be compared against a statistical threshold to decide on the existence of bad data [25].

Identification

Bad data can be identified in the measurement set by further processing of the residuals. The largest normalized residual test can be used for this purpose. The method consists on the following steps [25]:

1) Solve the WLS estimation and obtain the elements of the measurement residual vector:

$$r_i = z_i - h_i(\hat{\mathbf{x}}), \qquad i = 1, \ldots, m \tag{2.32}$$

2) Compute the normalized residuals in (2.30)

3) Find k such that r_k^N is the largest among all the r_i^N

4) If $r_k^N > c$, then the k-th measurement will be suspected as bad data. Else, stop, no bad data will be suspected. Here, c is a identification threshold

5) Eliminate the k-th measurement from the measurement set and go to step 1)

2.2.5 Multi-area state estimation

Deregulation of electricity markets can lead companies to monitor network across large geographical areas. As a result, a large number of meters should be available, and with it an important number of measurements, which prevents a centralized estimator to operate with efficiency. In this sense, multi-area state estimation (MASE) represents an alternative, where a power system is divided into smaller areas and local estimators are run in each area in order to obtain the global estimated state (distributed state estimator). Multi-area approach reduces the amount of data that each estimator processes (hence reduces complexity) and this allows improving the computing performance [17, 34].

Two processing techniques are used by multi-area state estimation methods [35]:

a) Parallel*:* it uses a number of closely coupled processors inside a physical reduced space. This technic speeds up the solution process [35].

b) Distributed*:* it uses a number geographically distributed computers. For a large power system, distributed processing can be a more flexible and reliable option for performing monitoring and controlling tasks. In addition, it can save large investment in communication networks [35].

Two computer architectures have been proposed [35]:

a) Hierarchical*:* a master computer assigns the work among slave computers performing local estimation and then it coordinates the local results [35].

b) Decentralized*:* there is no central coordinator computer. Each local computer "communicates only with processors of neighboring areas, exchanging border information" [35].

2.3 PMU and its application in state estimation: State of the Art

A PMU is a meter that provides synchronized voltage and current phasor measurements. PMU features were first implemented in stand-alone units (the first one was built in 1990 at Virginia Tech). Currently, Intelligent Electronic Devices (IEDs) have been upgraded to produce synchronized phasor measurements in addition to their own function. About two decades ago, the PMU deployment was scarce but after a number of successive blackouts it became more frequent all around the globe [36, 37].

The phasor estimation technique allows obtaining the phasor representation of a sinusoid quantity. Samples of the wave are taken over a data window which is usually one period of the fundamental frequency. The phasor is estimated using the Discrete Fourier Transform (DFT) [37].

PMUs use GPS to achieve a common reference for the synchronized phasor acquisition process. By synchronizing the sampling processes for signals that can be far away from each other, it is possible to assign their phasors to the same diagram. This allows the measurements to be compared directly and facilitates the analysis [37]. Measurements taken at one location are globally valid and can be used together with data collected at different locations [15]. Other estimates of interest are frequency and rate of change of frequency.

Phasor measurements are transmitted using digital communication networks to higher level applications at high speed (rate up to 60 samples per second in 60 Hz networks). Phasor data from different PMUs are then collected by a computer called Phasor Data Concentrator (PDC). It correlates the data using time-stamps with the aim of obtaining a system-wide measurement set [36].

Synchronized phasor measurements are usually more precise than conventional measurements. PMU measurements are time tagged with precision better than 1 microsecond. Their accuracy in magnitude is above 0.1% [36].

One of the applications that can be improved with the use of the PMU technology is the power system state estimator. The state estimator is one of the most critical EMS functions [38]. PMUs allow improving state estimation: both algorithmic and architectural [37]. If enough PMUs are deployed, the measured voltage and current phasors can make the network observable, thus making state estimation feasible. Such PMU-only state estimator is linear and can run at PMU speed. Thus the linear estimator can observe the dynamics associated to slow and fast phenomena of the entire power system. A linear PMU-only state estimator is possible, but is still expensive and far from becoming real in large scale applications. Besides, conventional measurements provide useful information that should not be discarded. Therefore, a more realistic and feasible option entails the use of both kind of measurements in a hybrid state estimator [39].

2.3.1 PMU based state estimators

Two kind of state estimators based on PMUs are found in the literature: PMU-only state estimators and hybrid state estimators. In this thesis both kinds are analyzed considering the following aspects: estimation method, reporting speed, variable/s to be estimated, algorithm used to find the solution and approach employed to detect and identify bad data.

PMU-only state estimators

A PMU-only state estimator may rely on static [28, 29, 40–42], forecasting-aided [43] or dynamic methods [6, 7, 10, 15, 44–54]. For either of the first two methods previously mentioned bus voltages (magnitude and angle) are chosen as state variables to be estimated (static state). This task can be performed at PMU speed [29, 40, 43], which allows observing fast phenomena, like transient events after a fault or switching, which is not possible to do with traditional approaches or at SCADA speed [28, 41, 42]. Algorithms commonly used by static estimators are WLS (linear and non-iterative) and LAV (solved with linear programming). In the first case a separate post-estimation bad data processing function based on the Chi-square test and normalized residuals is performed. In the second case, the LAV allows processing gross errors and remove them automatically as a part of the estimation process. In the forecasting-aided approach first the state is predicted through a linear technique and then a filtering process is performed using predictions and measurements. The problem is initially formulated based on the WLS technique and then adapted to be solved using Mixed-Integer Programming (MIP). Predictions are discarded when a sudden change takes place by using the normalized residual test concept and adding a binary variable to the prediction residual. Bad data detection and identification could be implemented following the same rule. Redundancy enhancement provided by

this method permits increasing the robustness of the estimator. With regard to dynamic state estimation, either only the dynamic state related to electromechanical dynamics (e.g. rotor angle and speed of generators) [6, 7, 10, 15, 44–49, 51, 52, 54, 55] or both type of states (static and dynamic in sequence) [50, 53] are estimated at PMU speed. Two kinds of approaches are identified. The first one is the model-based approach [6, 7, 10, 15, 45–53, 55] where the system to be estimated is mathematically modelled by DAEs and the problem is solved by recursive filters (e.g. least square, Kalman filters and Particle Filters). Approaches used to detect and identify large errors include in [7] the normalized innovation vector-based test, which detects observation outliers and is based on the ratio between the deviation of actual measurement from the predicted one and the expected standard deviation of the predicted measurement, the normalized residual test is employed in [15, 53] to find measurements with gross errors and in [55] a projection statistics-based multiple hypothesis approach is presented. It is capable of finding three kinds of outliers, namely the observation, innovation and structural outliers. Observation outliers refer to the received measurements providing unreliable metered values due to gross errors, innovation outliers are caused by system process noise, whereas structural outliers are induced by incorrect parameters of generators or its associated controllers. Model-based approaches work well under the validity of certain assumptions. First, the system process and observation noise are assumed to have at each time instant zero means and known covariance matrices; second, they are assumed to follow a Gaussian distribution and, finally, the system model is assumed to be known with good accuracy [27]. When the chosen models do not match the actual system behavior, and/or the PMU measurements contain significant errors, the estimated state may deviate from the true state rapidly and the estimator may not converge [27, 50]. The second approach is known as data-driven approach [44, 54]. In this case the non-linear behavior of the system is described without needing a mathematical model. The problem is solved using artificial intelligence (AI) methods, e.g. artificial neural networks (ANN). These are trainable dynamical systems capable of estimating input-output functions by learning from experience with numerical sample data. Model-free approaches have the advantages of fast parallel computation and fault tolerance characteristics. The research in this field demonstrated that these estimators can lead to improved performance, being robust to parameter variation, adaptive and easy to extend, and not computationally demanding. However, the performance greatly depends on the training set, making the selection crucial. The training data must represent all possible transient and steady state characteristics of the machine under consideration. The model used needs to be updated when major changes occur in the system, such a new transmission or generation equipment.

All the works analyzed assume that the power system is fully observed by PMUs, which represents a strong assumption. Although the number of installed PMUs in large interconnected systems is gradually increasing, there is still a long way to equip all buses of a large system with PMUs [48].

Hybrid state estimators

Hybrid state estimators focus on the static state of the network, i.e. bus voltages and related variables. They may be based on static, tracking or forecasting-aided methods. In the first case, most of the works analyzed [12–14, 35, 56–64] do not take into consideration the fact that PMUs update rate is remarkably higher than SCADA. Different ways for including phasor measurements and how they impact on accuracy, convergence properties and bad data detection capabilities is mostly studied. In distributed approaches [61, 63, 64], PMUs are mostly used to coordinate angles between subsystems thanks to their synchronization capacity. The optimal state is found by means of the conventional WLS technique and bad data are detected and identified by the largest normalized residual test. Unlike most of its counterparts, the paper in [65] proposes a hybrid estimator based on a static method that runs at PMU speed. The solution is found using a modified WLS formulation. Voltages at non-observed buses are calculated from phasor measurements and an interpolation matrix, which is obtained from network admittances as a function of voltages and power injections at non-observed buses and is updated whenever conventional measurements arrive or in contingencies that provoke a line outage or a significant load change. The author achieves good results regarding accuracy and computing time. Nevertheless, in the construction of the interpolation matrix loads are represented by static models when actually in real systems approximately 60-70 % of loads are dynamic. The normalized residual test is used to detect and identify large errors. In relation to forecasting-aided methods, it has been found a few interesting works which propose hybrid estimators running at PMU speed [66, 67]. In [66] an estimator that has been though to work in large systems using parallel computing by means of a graphic processing unit (GPU) is presented. Since the number of PMUs is not enough to observe the entire system, those buses without units are predicted using previous data. A data collation method extrapolates SCADA measurements to update them as fast as PMU measurements arrive. The estimation procedure is based on the EKF. Its performance is evaluated under both normal and contingency conditions. Despite the author claims that the approach can accurately perform under steady and dynamic conditions, results do not permit verifying the latter due to the chosen time scale. The state-transition model is linear. Therefore, the lack of an appropriate dynamic model to represent the behavior of the system in

the presence of sudden changes can represent a major source of error. Paper [67] proposes a multi-area hybrid estimator that operates at PMU speed. It estimates the state by using a cubature Kalman filter (CKF). The approach is run in different areas in parallel. A multi-agent based software is used to exchange information and to run the estimator for SCADA and PMU measurements separately and then integrate their results to get the final estimated state in each area. The method is evaluated regarding its accuracy and computing time. Results prove that the method can run fast in large systems using all the available measurements. It is capable of estimating with accuracy and quite fast the static state under steady or quasi-steady conditions. Bad data analysis is not performed in either of the two works previously mentioned. Papers [18, 68] propose hybrid estimators using tracking methods at PMU speed. In [18] a two-stage estimator capable of tracking bus voltages is presented. Both stages are WLS based. Second stage is solved in one iteration using a linear state estimator. PMU observability is ensured by adding current pseudo-measurements from previous estimated states. The method is prone to errors in the presence of sudden changes and presents limitations when the electrical distance between PMUs and non-observed buses is greater than two buses. Paper [68] introduces another hybrid scheme to track bus voltages at high sampling rate. Between two consecutive SCADA scans a limited number of phasor measurements are received. To maintain full observability a set of old SCADA measurements are used. A conventional WLS is employed at instances when both kind of measurements are refreshed simultaneously. When only PMU measurements are refreshed, a weighted least absolute value estimator (LAV) is utilized. The last algorithm allows gaining robustness against gross errors and therefore post-estimation bad data processing methods can be avoided. The proposed approach can monitor voltage collapses, sags and swells but is not designed to follow large disturbances like faults, generator outages, etc.

From the works presented above, it can been seen that the integration of a few synchronized phasor measurements into an existing nonlinear estimator results mostly in a hybrid estimator that has most of the limitations of the original conventional estimator when trying to estimate the dynamics of the power system, i.e. only steady state models are employed and the high PMU speed is not exploited. The most significant assumption that is made by traditional estimators, i.e. the system does not change during scan, remains valid. Only a few papers propose schemes at PMU speed with the aim of tracking the time evolution of the static states. Nevertheless, they may not perform well in the presence of sudden disturbances (transient-regime) like faults. Besides the power system dynamic is completely ignored. The proposed estimators

does not deliver information regarding those variables which have a great meaning in stability analysis, that is, the dynamic states, e.g. speed and rotor angle of generators.

2.4 Mathematical tools

2.4.1 Monte Carlo method

Security assessment is generally carried out by means of a deterministic approach. Power systems are designed and operated to cope with a number of contingencies whose probability of occurrence is significantly high [20].

The deterministic approach worked well in the past since it led to high security levels and minimum study effort. One main limitation, however, is the fact it assumes all security-limiting scenarios as having the same risk. In addition, it does not define how likely or unlikely various contingencies are [20] and it dismisses the stochastic nature of real systems. Under such conditions, some severe events that could lead the system to insecure conditions may be ignored [69]. In today's electricity markets with a number of participants and different business interests, the deterministic approach may not be suitable. The stochastic nature of the system should be taken into account [20]. Monte Carlo based simulation represents a possibility of obtaining more realistic results by considering the uncertainties that influence the system dynamic response [69].

The Monte Carlo method allows simulating stochastic phenomena, such those associated with power system operating conditions. It is a repetitive procedure that evaluates at each iteration the system response by means of an uncertainty function and using a set of input variables which are generated randomly from their probability distribution functions (PDFs) [69].

In this thesis, Monte Carlo method is used to obtain post-contingency dynamic data taking into account probable operating scenarios and contingencies. Some of the simulated scenarios result in unstable operating conditions. The obtained data will be later analyzed by means of data-mining techniques with the aim of defining a PMU topology and the methodology for generating pseudo-measurements.

2.4.2 Data-mining techniques

As stated in [70], "Data mining is the process of discovering meaningful new correlations, patterns and trends by sifting through large amounts of data stored in repositories, using pattern recognition technologies as well as statistical and mathematical techniques".

Other terms that are also generally used to describe this process are: analytics, predictive analytics, big data, machine learning, and knowledge discovery in databases. All of them share the

same main goal, to obtain knowledge from large data sets [70]. The knowledge discovery process is an “iterative sequence” that involves the following steps [71]:

1) Data cleaning: noise and inconsistent data are removed
2) Data integration: multiple data sources can be combined
3) Data selection: data that is relevant to the analysis are obtained from the database
4) Data transformation: data are transformed into forms appropriate for mining
5) Data mining: essential process where intelligent methods are applied to extract data patterns
6) Pattern evaluation: interesting patterns representing knowledge are identified
7) Knowledge presentation: visualization and knowledge representation techniques are used

Data objects and attributes

Data sets are composed of data objects. A data object represents an entity and is described by attributes. Data objects are also samples, examples, instances, data points, or objects. If data objects are located in a database, they are data tuples. The rows of a database are data objects and the columns are attributes. An attribute is a feature of the data object. The terms attribute, feature, and variable are usually synonyms [71].

Similarity and dissimilarity

In data mining applications, such as clustering and classification, there is the need to assess how alike or unalike objects are in relation to one another. Similarity and dissimilarity are measures of proximity. Two objects with similarity equal to 0 will be different. The higher the similarity vale, the more similar the objects are. The dissimilarity measure works in the opposite way. If the objects are the same the dissimilarity value will be 0. The higher the dissimilarity value, the more different the objects are [71].

Data preprocessing

Real-world databases are highly vulnerable to noise, missing and inconsistent data. This is mainly because their big size and their heterogeneous origin from multiple sources. Low-quality data will result in low-quality mining results [71].

High quality data satisfy the requirements of the intended use. Among the factors affecting the data quality are: accuracy, completeness, consistency, timeliness, believability, and interpretability [71].

Data quality can be improved using the following data preprocessing techniques [71]:

- Data cleaning:

Real data tend to be incomplete, noisy, and inconsistent. Data cleaning techniques are intended to correct these imperfections by filling in missing values, smoothing out noise, and removing data inconsistencies [71].

- Data integration

Data integration involves the merging of data from multiple data sources. Data integration represents a means of reducing and avoiding data redundancy and inconsistency. As a result, the accuracy and speed of the data mining process can be improved [71].

Redundancy is not a minor issue in data integration. An attribute may be redundant if it can be obtained from other attribute. Some redundancies can be detected using correlation analysis. Such analysis can measure the degree of relationship between attributes based on the available data. For nominal data, the χ^2 (chi-square) test is used. For numeric attributes, the correlation coefficient and covariance can be employed to this end [71].

- Data reduction

Data reduction techniques allows obtaining a reduced representation of the data set that is smaller in size but still keeps the integrity of the original data. Data mining methods applied on the reduced database should be more efficient and produce the same (or almost the same) results [71].

Data reduction approaches include dimensionality reduction. This approach reduces the number of variables or attributes. Principal Component Analysis (PCA) is a dimensionality reduction method employed in this thesis. Assuming that the data to be reduced consist of tuples described by n attributes or dimensions, principal components analysis searches for k n-dimensional orthogonal vectors that can be used to represent the data, where $k \leq n$. The original data are projected onto a smaller space, resulting in dimensionality reduction [71]. The basic procedure is the following [71]:

1) The input data are normalized so that each attribute is in the same range.
2) PCA defines k orthonormal vectors (principal components) that provide a basis for the normalized input data. These are unit vectors perpendicular to each other. The input data are a linear combination of the principal components.

3) The principal components are ranked in order of decreasing importance. The principal components provide information about the variance. The first axis shows the most variance among the data, the second axis shows the next highest variance, and so on.

4) The data size can be reduced by removing those components with less importance, that is, those with low variance. The strongest principal components should allow obtaining a good approximation of the original data.

- Data Transformation

In data transformation, the data are transformed into forms that are appropriate for mining. Next, some techniques for data transformation [71]:

a) Smoothing: it removes noise from the data. Some techniques include binning, regression, and clustering

b) Attribute construction: new attributes are built and added from the given set of attributes to improve the mining process

c) Aggregation: aggregation operations are applied to the data

d) Normalization: the attribute data are scaled

e) Discretization: the values of a numeric attribute are replaced by interval labels or conceptual labels

f) Concept hierarchy generation for nominal data: attributes can be classified to higher-level concepts

Data-mining tasks

In this thesis two data-mining tasks are used for defining the PMU topology and generating the pseudo-measurements that will make PMU full observability restoration possible. These two tasks are: classification and clustering.

- Classification

Classification is a data analysis approach that build models describing data classes. The models (called classifiers) predict categorical (discrete, unordered) class labels. Classification involves two steps: a learning step (where the classification model is built) and, a classification step (where the model is used to predict class labels) [71].

First, a classifier capable of describing a predetermined set of data classes is built. This is the learning step (or training phase), where a classification algorithm creates the classifier by analyzing ("learning from") a training set. A tuple, **t**, is an n-dimensional attribute vector **t**=(at_1, at_2,…, at_n). Each tuple, **t**, belongs to a predefined class according to the class label attribute. The class label attribute is discrete and unordered. It is categorical in the sense that each value serves as a category or class. The tuples that are part of the training set are randomly sampled from the database under analysis. Since the class label of each training tuple is known, this step is also named supervised learning [71].

The accuracy of the classifier is determined by means of a test set that consists of test tuples and their associated class labels. They are independent of the training tuples, i.e. they were not used to build the classifier. The accuracy of a classifier is defined as the percentage of test set tuples that are correctly classified [71].

In this thesis two classification methods are used: Random Forest and k-Nearest-Neighbor. The Random Forest is an ensemble method. An ensemble method is a composite model that consists of a combination of classifiers. The individual classifiers vote and the result will be obtained from the ensemble method based on the collection of votes. Ensemble methods tend to be more accurate than the individual classifiers [71].

In the case of the Random Forest method, it is formed by a number of Decision Tree classifiers. A decision tree is a chart or diagram with a shape similar to a tree that is used to make decisions. Internal nodes represent tests on attributes or variables and branches constitute the outputs for such tests. Classes will be included in the terminal nodes (leaf node). The node at the top in a tree is known as root node. Given a tuple, **t**, for which the class is unknown, the attributes of the tuple are evaluated against the decision tree so that "a path is traced from the root to a leaf node, which holds the class prediction for that tuple". Decision trees can be used to build classification rules. No domain knowledge is needed for building a decision tree. That is why, it is suitable for "exploratory knowledge discovery". Besides, they can work with multidimensional data. Decision trees are intuitive and easy to assimilate by humans and the process behind it is simple and fast [71].

Nearest-neighbor classifiers learn by comparing a certain test tuple with similar training tuples. Each tuple denotes a point in an n-dimensional space. Training tuples are stored in an n-"dimensional pattern space". The k-nearest-neighbor classifier finds in the pattern space those k

training tuples (k nearest neighbors) that are closest to the unknown tuple. The proximity between tuples is determined using a distance metric such as the Euclidean distance [71]. Considering the following points or tuples, $\mathbf{t}_1=(at_{11}, at_{12},\ldots, at_{1n})$ and $\mathbf{t}_2=(at_{21}, at_{22},\ldots, at_{2n})$, the Euclidean distance is defined as [71]:

$$\text{dist}(\mathbf{t}_1,\mathbf{t}_2)=\sqrt{\sum_{i=1}^{n}(at_{1i}-at_{2i})^2} \tag{2.33}$$

Usually, the values of each attribute are normalized before using (2.33). By doing this the same weight is assigned to each attribute regardless of their initial ranges. The k-nearest-neighbor assigns to the unknown tuple the most common class among its k-nearest neighbors. If k=1, the nearest training tuple defines the class of the unknown tuple [71].

- Clustering

Clustering is a process that splits a set of data objects into subsets. Each subset is a cluster. Objects that belong to the same cluster are similar to each other and dissimilar to objects in other clusters. Different clustering methods may lead to different results on the same data set. Clustering can be used to find groups within the data which were previously unknown [71].

Since a cluster is a group of objects that are similar to each other and dissimilar to objects in other clusters, a cluster can be seen as “an implicit class”. That is why, clustering is sometimes called “automatic classification”. “Unsupervised learning” is another term used for clustering since the class label is not known. “Clustering is a form of learning by observation, rather than learning by examples”. It can be used for detecting outliers [71].

Clustering methods can be classified into partitioning or hierarchical methods. A partitioning method splits the data into k groups so that each group contain at least one object. The clusters are built to optimize a dissimilarity function based on distance so that the objects in a cluster are similar to each other and dissimilar to objects in other clusters. A hierarchical clustering method groups data objects into a hierarchical way. Hierarchical representation of the data objects is useful for “data summarization and visualization”. However, there may be inconveniences related to the selection of merge or split points. This is a key aspect by hierarchical methods, because once a group is split, the process at the next step will operate on the new generated clusters. There is no way to undo what was done or to exchange objects between clusters. “Thus, merge or split decisions, if not well chosen, may lead to low-quality clusters” [71].

In this thesis, two partitioning methods are selected: the k-means and the k-medoids, respectively. The k-means algorithm is a iterative method that splits a certain data set into a prespecified number of clusters k. The algorithm works on a set of d-dimensional objects, $\mathbf{O}=\{\mathbf{o}_i,\ldots,\mathbf{o}_N\}$, where $\mathbf{o}_i \in \Re^d$. The algorithm is initialized by choosing k objects in $\Re^d$ as the initial k cluster representatives or "centroids". These initial seeds can be obtained by randomly sampling the dataset, by clustering a small subset of the data and using the solution for defining the seeds or by perturbing the global mean of the data k times [72]. Then, the algorithm iterates between the following two steps [72]:

1) Data Assignment. Each object is assigned to its closest centroid based on the Euclidean distance. "This results in a partitioning of the data".
2) Relocation of "means". Each centroid is "relocated to the center (mean) of all objects assigned to it".

The quality of cluster cl_j can be determined using the within sum of squares, which is the sum of squared error between all objects in cl_j and the centroid $\mathbf{ce}_j$, defined as [71]:

$$WSS=\sum_{j=1}^{k}\sum_{o_i \in cl_j} \text{dist}\,(\mathbf{o}_i,\mathbf{ce}_j)^2 \qquad (2.34)$$

where dist ($\mathbf{o}_i$, $\mathbf{o}_j$) is the Euclidean distance between two points $\mathbf{o}_i$ and $\mathbf{o}_j$. The k-means algorithm improves the within sum of squares. The algorithm runs until the assignment is stable, i.e. the clusters determined in the current iteration are the same as those from the previous one [71].

The k-means method often reaches a local optimum. The initial random selection of centroids can influence the "quality" of the results. Running the algorithm multiple times with different initial centroids is common practice when trying to get good results [71].

The k-medoid algorithm is similar to k-means with the exception that the centroids have to belong to the data set being clustered. This change makes the k-medoids more robust than k-means under the presence of noise and outliers because a medoid is less susceptible to outliers than a mean. However, a bigger complexity in each iteration makes the k-medoids computationally less efficient than its counterpart [71].

2.5 Summary

At the beginning of the preceding chapter power system security and power system stability are defined. Following, the state estimation function for power system monitoring is presented. An analysis of the current situation in this regard allows identifying the limitations of the current monitoring functions and the impact that such limitations have on power system security. A review and analysis of the state of the art on alternative PMU based monitoring functions was also performed. The solutions offered by each of them and the conditions and challenges to be met in their implementation were given in detail. In the last part, data-mining techniques used in this thesis are presented and explained.

The changes that power systems have experienced lately have a negative effect on stability and therefore on dynamic security as well. In this context, the occurrence of a disturbance may provoke a cascading failure and eventually a system outage. In order to improve power system stability, a new monitoring tool capable of informing system dynamics must first be developed. Thus, online dynamic security assessment will become a reality.

Traditional state estimators are not able to inform power system dynamics (except maybe for some very slow system changes). They are based on steady state models. The main reason behind this is the slow update rate of RTUs-SCADA. Besides, traditional measurements are not synchronized which means that they are not globally valid.

The arrival of PMUs represents an opportunity to improve the existing estimators. Their high speed, accuracy and synchronization capacity make them suitable for global monitoring of power system dynamics. An estimator based only on PMUs would be capable of solving the estimation problem at high speed in just one iteration. This would allow removing the static assumption and the dynamic state could be estimated. However, such an estimator would need a power system fully observed by such units, which is financially and technically unviable in the near future. That is why, the current trend is to incorporate PMUs gradually into the grid together with conventional RTUs-SCADA.

The state of the art on state estimators using PMUs is analyzed. Hybrid state estimators as well as PMU only state estimators are considered. The analysis focuses on aspects like algorithms, working speed, estimation methods and estimated variables.

PMU only state estimators are able to work at high speed and to deliver the time evolution of static (network bus voltage magnitudes and angles) and dynamic states (speed and rotor angle

of generators) when a dynamic method is employed. Here, two different approaches are proposed. The model based approach and the data-driven approach. The first one works well as long as certain assumptions are satisfied: process and measurement noises follow a Gaussian distribution with zero mean and the system model is well known. The data-driven approach depends to a great extent on how well the training data represents the system response regardless of the transient event that is happening. As mentioned before, all works analyzed here assume that there are PMUs in a number enough to achieve full observability.

Since only PMU based power system monitoring represents a remote reality this thesis focuses on a nearer future which involves the use of mixed measurement technologies such as RTUs-SCADA and PMUs applied together in hybrid state estimators. Nevertheless, hybrid estimators result mostly in approaches that have most of the limitations of the conventional estimators. Only steady state models are used and the PMU speed is not considered. This is mainly due to the reduced number of units. A few works propose schemes capable of operating at high speed and getting the time evolution of the static state but without taking into account those variables which have great meaning in stability analysis and therefore a significant influence in power system dynamic security, i.e. the dynamic state (for instance speed and rotor angle of generators).

That is why, this thesis will develop a hybrid state estimator capable of accurately monitoring power system dynamics associated to slow and fast transient events considering a limited amount of PMUs.

Since the lack of PMUs is the main obstacle when trying to get a hybrid estimator capable of monitoring the system dynamics, this issue will be addressed first. Initially, an analysis of the time evolution of post-contingency bus voltages will be performed in order to define the PMU topology, that is, number of units and their location in the grid so that bus voltage dynamics of the whole system can be observed by such units. Then, a data-mining based methodology for generating pseudo-measurements aiming at restoring the PMU full observability will be developed.

The development of the hybrid state estimator includes the study and analysis of the following aspects:

1. Incorporation of PMU measurements in traditional estimators: this is one of the main concerns regarding the design of a hybrid estimator. Their high accuracy and speed

provokes numerical stability problems in the first case and time skewness between both kinds of measurements in the second one.

2. Estimation algorithm: several algorithms have been proposed in the international literature depending on the estimation method adopted. In this thesis, an analysis of all of them is performed in order to find an adequate option for estimating the static and dynamic states considering aspects like accuracy, speed and convergence.
3. Estimated variables: the development of a new monitoring tool aims to improve the power system dynamic security. Therefore, power system stability will be studied and analyzed in order to find those variables of great importance in stability assessment.
4. Architecture of the estimator: aspects 1) and 2) will allow defining the structure of the hybrid estimator, their components and how they relate each other.

Finally, the hybrid state estimator will be coupled together with the methodology for generating pseudo-measurements.

3 Methodology for generating bus voltage pseudo-measurements

As mentioned earlier, most of the hybrid estimators proposed in the international literature estimate only the static state. Dynamic characteristics of power system are left aside and only models that represent the system behavior in stationary regime are considered. This is mainly because there are not enough PMUs to observe the entire electrical grid. Thus, certain characteristics of this equipment, such as its high reporting speed, which would allow operators to monitor the power system as a whole in transient regime, are not taken into consideration.

Aiming to remove this obstacle, this thesis proposes a data-mining based methodology that generates bus voltage pseudo-measurements, thereby restoring full PMU observability. First, the PMU topology, i.e. the number of units and their location, is defined with the help of data-mining techniques so that dynamics of post-contingency bus voltages are fully monitored. Then, with the help of a classifier, bus voltage pseudo-measurements are generated and observability is extended to the entire network.

This chapter presents and explains the methodology for generating bus voltage pseudo-measurements. First, an overview of the methodology is given. Next, the probabilistic approach that defines the PMU topology is given in detail. Then, the classifier in charge of predicting bus voltage coherency is explained. Finally, the procedure for generating bus voltage pseudo-measurements is presented and explained.

3.1 Overview

The methodology for generating bus voltage pseudo-measurements that is proposed in this thesis is based on the concept of bus voltage coherency. Buses that are coherent with each other show similar voltage behavior during a transient event and, therefore, the behavior in one of these buses represents approximately the behavior of the remaining buses. The proposed approach relies on this idea to define bus voltage pseudo-measurements by means of voltage magnitudes and angles at PMU observed buses, the pre-contingency system loading condition and information regarding the contingency (kind of fault, affected device and time of occurrence), which is assumed to be known.

Nevertheless, coherency depends mainly on the last two characteristics mentioned before [73]. That is why, a PMU topology, that is, number of units and their location, must be defined first so that all the coherent areas are observed with high probability. This is accomplished by means of an offline data-mining based knowledge discovery process which consists of three steps. In the first step, a large set of scenarios that are representative of normal, during fault and post-fault behavior in a broad range of conditions is obtained from numerical simulations using Monte Carlo simulation and afterwards bus voltage coherency is determined for each simulated scenario using clustering technique. In the second step, a data transformation process takes place in order to facilitate data analysis. Finally, coherency analysis is performed with the aim of defining the PMU topology which allows observing all bus voltage coherent areas (magnitude and angle) in a large number of simulated scenarios. This tool constitute a contribution of this work in the sense that it focuses on observing coherent areas in voltage magnitude and angle instead of observing the whole network, as is often the case with approaches proposed in the literature [74–77]. One main advantage is the reduction in the amount of equipment required.

Once PMU topology has been defined, a classifier based on the k-nearest neighbors algorithm that is designed offline will forecast coherency at high speed whenever a contingency happens and with that an assignment vector that shows the connection between observed and non-observed buses. The classifier is supported by a classification matrix that is generated offline from the pre-processed data during the definition of the PMU topology. Finally, bus voltage pseudo-measurements will be obtained using information about the estimated static state at the previous time, the assignment vector predicted by the classifier and PMU measurements at observed buses. The proposed approach for generating bus voltage pseudo-measurements represents another important contribution of this thesis. In the literature few are the papers working on this topic and getting promising results. In [65] voltages at non-observed buses are calculated from phasor measurements and an interpolation matrix, which is obtained from network admittances as a function of voltages and power injections at non-observed buses and is updated whenever conventional measurements arrive or in contingencies that provoke a line outage or a significant load change. A main drawback of this methodology is that in the construction of the interpolation matrix loads are represented by static models. In paper [18] pseudo-measurements are obtained from estimated currents in the previous time. A major drawback here is the loss of accuracy in transient regime.

The following flowchart gives an overview of the methodology explained before.

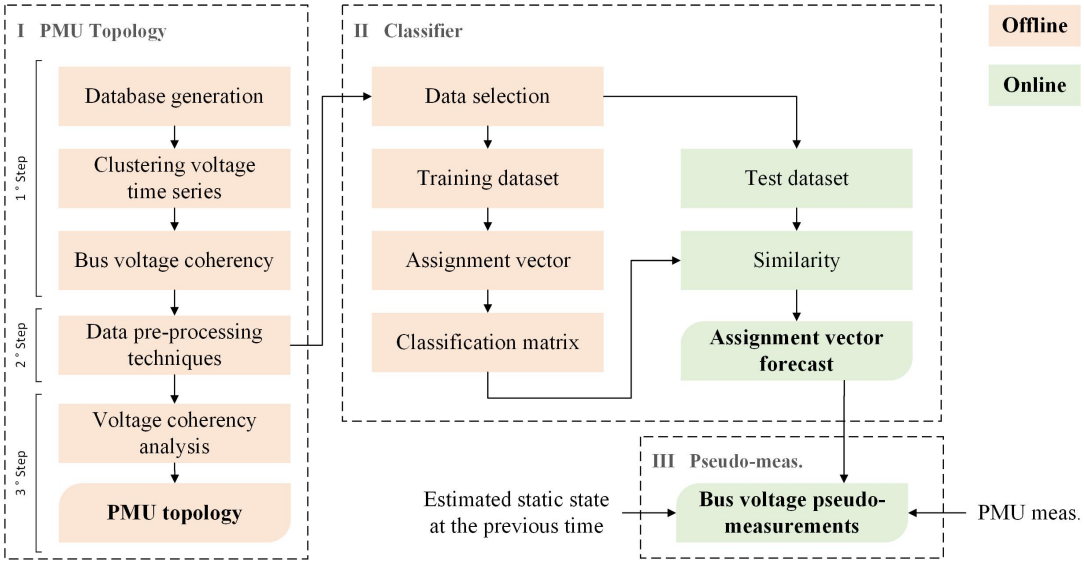

Fig. 3.1. Methodology for generating bus voltage pseudo-measurements.

3.2 Probabilistic approach for defining the PMU topology

The PMU topology is defined offline and the process behind it essentially consists of three steps as mentioned before: in the first one with the help of simulation software and the Monte Carlo method, several operating scenarios for evaluating the dynamics of post-contingency bus voltage and coherency are generated. The Monte Carlo method is an iterative procedure that evaluates the response of the system in each iteration, using input variables randomly generated from their probability density function [73].

The database is generated from the dynamic model considering a large number of operating conditions with uncertainties in: the type of contingency (following the N-1 criterion commonly used by the EMS in control centers for real-time security analysis [78]), the failed component and fault location (line, load or generator and percentage of line in which the fault occurs) and the load-generation scenario. The procedure for generating the database is as follows: a random load scenario is constructed using probability distribution functions and short-term demand forecasting; the optimal power flow is calculated, and the dispatch of generators is determined; N-1 contingency (three-phase short circuit in transmission line, load curtailment or generator outage) and the affected device (line, load or generator) are randomly selected; the time-domain simulation is performed and time series of voltage magnitudes and angles are stored. This is a procedure in which are simulated the different scenarios that a power system can face during the operation, see flowchart in Figure 3.2.

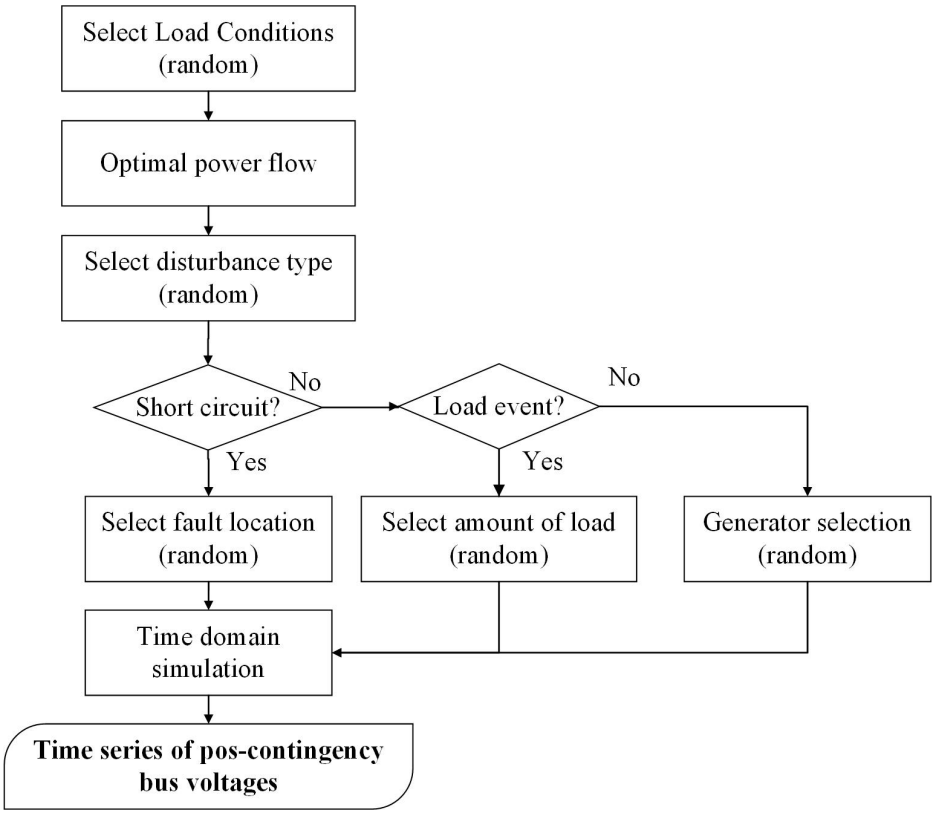

Fig. 3.2. First step: database generation.

The obtained dataset is composed of input and output variables. Inputs are the kind of contingency, the failed element, the fault location or amount of load disconnected and the pre-contingency system loading condition. Outputs are the time series of bus voltages in magnitude and angle. For each scenario, bus voltage coherency, i.e. the class to be predicted, is determined by means of an iterative clustering procedure that operates on time series of post-contingency bus voltages (magnitude and angle) in each scenario and is based on the partitioning algorithm k-medoids (see subsection 2.4.2, Data-mining techniques). The algorithm organizes a number of objects "n" from a data set "D" into "k" partitions where each partition represents a cluster. Conceptually, the medoid is an object that represents the center point of the cluster. The method is performed based on the principle of minimizing the sum of the dissimilarities between each object and its corresponding medoid. Bus voltage coherency is defined by means of an iterative process where the number of clusters evolves from a minimum of two to a maximum predefined value. For each "k", the partitioning algorithm runs a certain number of times and solutions are evaluated employing cluster validity indices (CVIs) [79]. This double-loop routine aims, first, to obtain the optimal cluster from a voting matrix (which reduces uncertainties associated with the initial solution that is randomly chosen) and, second, to find the optimal "k" through the

Elbow rule [80] by analyzing the evolution of the total intra-cluster variance while varying the number of clusters, see Figure 3.3.

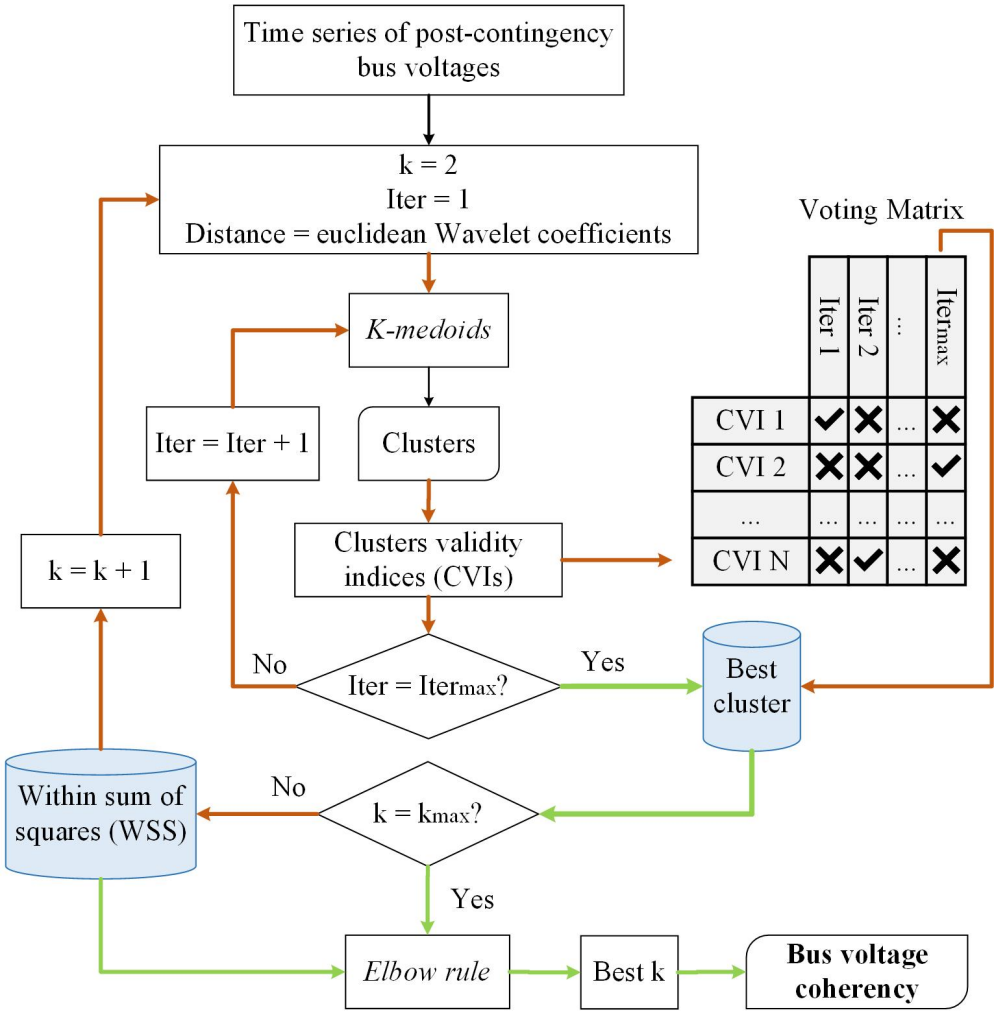

Fig. 3.3. First step: determination of bus voltage coherency.

In the second step, data pre-processing techniques are carried out. The main goal here is to make data analysis and comprehension easier and to obtain data that is suitable for creating good predictive models. These techniques work on input variables (features) and involve the following operations:

1) Variable normalization and dimensionality reduction without loss of information aiming to reduce the number of features under consideration through Principal Component Analysis (subsection 2.4.2, Data-mining techniques) of the pre-contingency system loading condition and data clustering according to principal components.

2) The distribution of coherent scenarios or classes and their boundaries are analyzed due to two problems that frequently arise in multi-class classification tasks, which are data imbalance and class overlapping.

   Traditional learning models assume that data classes are well distributed. Nevertheless, in many real-world data domains this does not happen. If the numbers of scenarios for each class are similar there is no problem. However, if at least one of the classes has many more examples than the other classes, the model tends to favor the class with more examples [81]. Moreover, objects from minority classes located in overlapping regions of the data attribute space makes things even more complex for learning models [82]. In reality, data from different classes may have similar attributes. In such cases, boundaries may not be clearly defined, being quite complicated to be correctly learned [83]. In order to overcome these problems a three-step filtering process takes place. This procedure alters the original data with the aim of reducing the impact of class imbalance and overlapping in the classifier performance and consists of three stages: a) in stage one, "boundary" scenarios that belong to different classes (coherent scenarios) but have similar feature values are identified. Next, the degree of dissimilarity between such scenarios is determined by the number of buses that do not match between equivalent coherent areas. Scenarios whose degree of dissimilarity falls below a predetermined threshold are "forced" to belong to the same class (same bus voltage coherency); b) in stage two, scenarios whose frequency of class is equal to one are recognized. Then, such "individuals" are matched in class to groups of scenarios as long as the distance between the individual and the group under analysis falls below a predefined threshold or if one of them is enclosed by the other. Here, the Euclidean distance is used for this purpose and; c) the last stage consists of a filter that removes from the dataset those scenarios whose frequency of class falls below a threshold, which is defined by the kind and size of the dataset.

The last step involves the definition of the PMU topology. The number of units and their location in the grid are determined by means of an iterative exploring process in which the number of PMUs increases from a pre-defined minimum to a maximum that is equal to the largest number of coherent areas in the dataset. At each iteration and for a given number of units, candidate topologies are found through a combinatorial process so that a certain pre-defined portion of the grid is observed. Here, buses that are medoids a number of times over the total amount of simulated scenarios will be the candidate buses under consideration. Each candidate topology is evaluated with respect to voltage coherency over the dataset in order to obtain the

number of simulated scenarios whose voltage coherent areas (in magnitude and angle) include at least one PMU bus, like in the example shown in Figure 3.4. The developed methodology takes into account the fact that for a given loading condition and kind of contingency coherency in voltage magnitude and angle may be different. The use of data-mining tools to find the PMU location under such condition represents a contribution of this thesis.

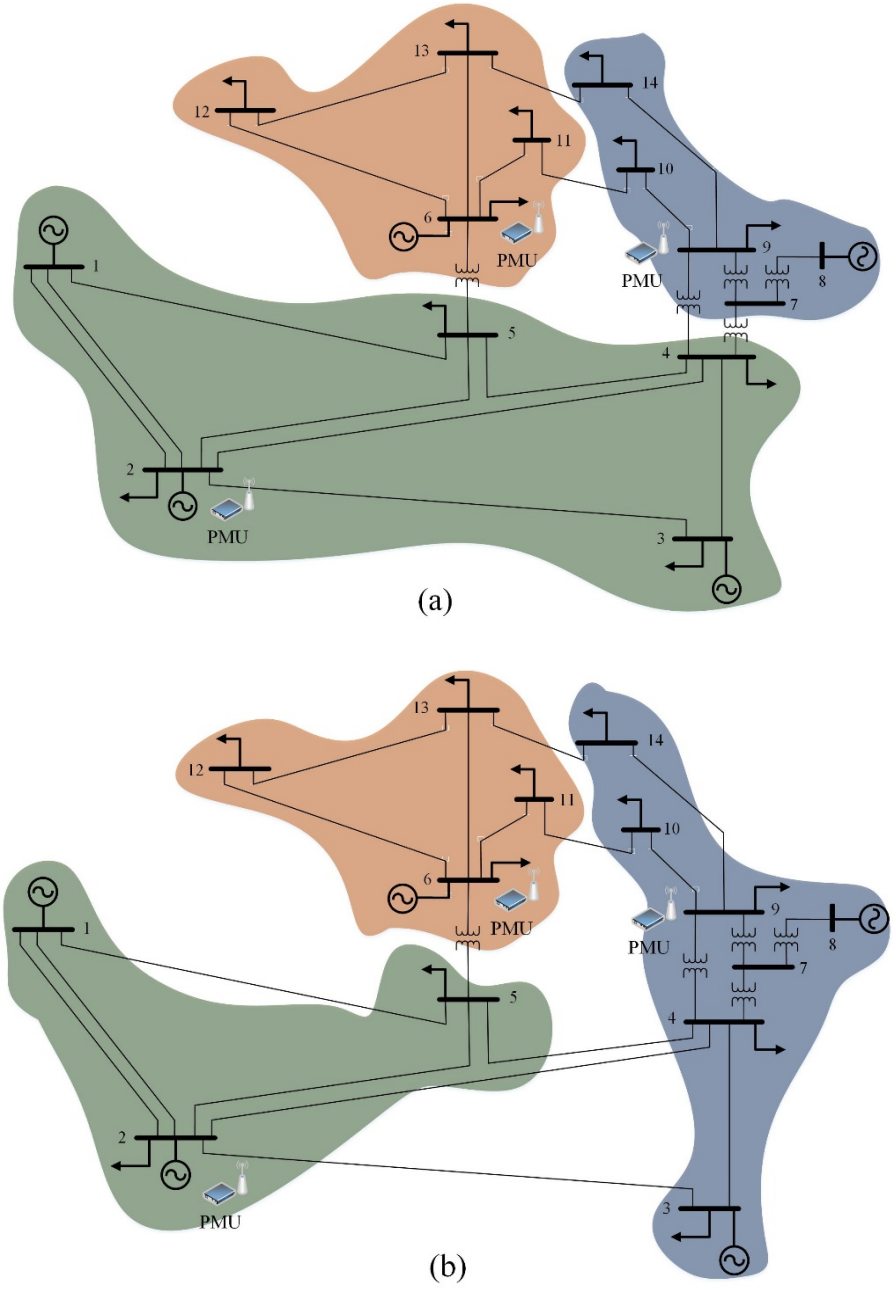

Fig. 3.4. Example of post-contingency bus voltage coherent areas. PMU location
(a) Voltage magnitude, (b) Voltage angle

The candidate who observes all coherent areas (magnitude and angle simultaneously) in most scenarios is stored. The process is repeated iteratively while increasing the amount of PMUs until the maximum is reached. The final solution will be the best between all the stored topologies considering the relationship between amount of units and number of observed scenarios.

This final step is represented with the flowchart in Figure 3.5.

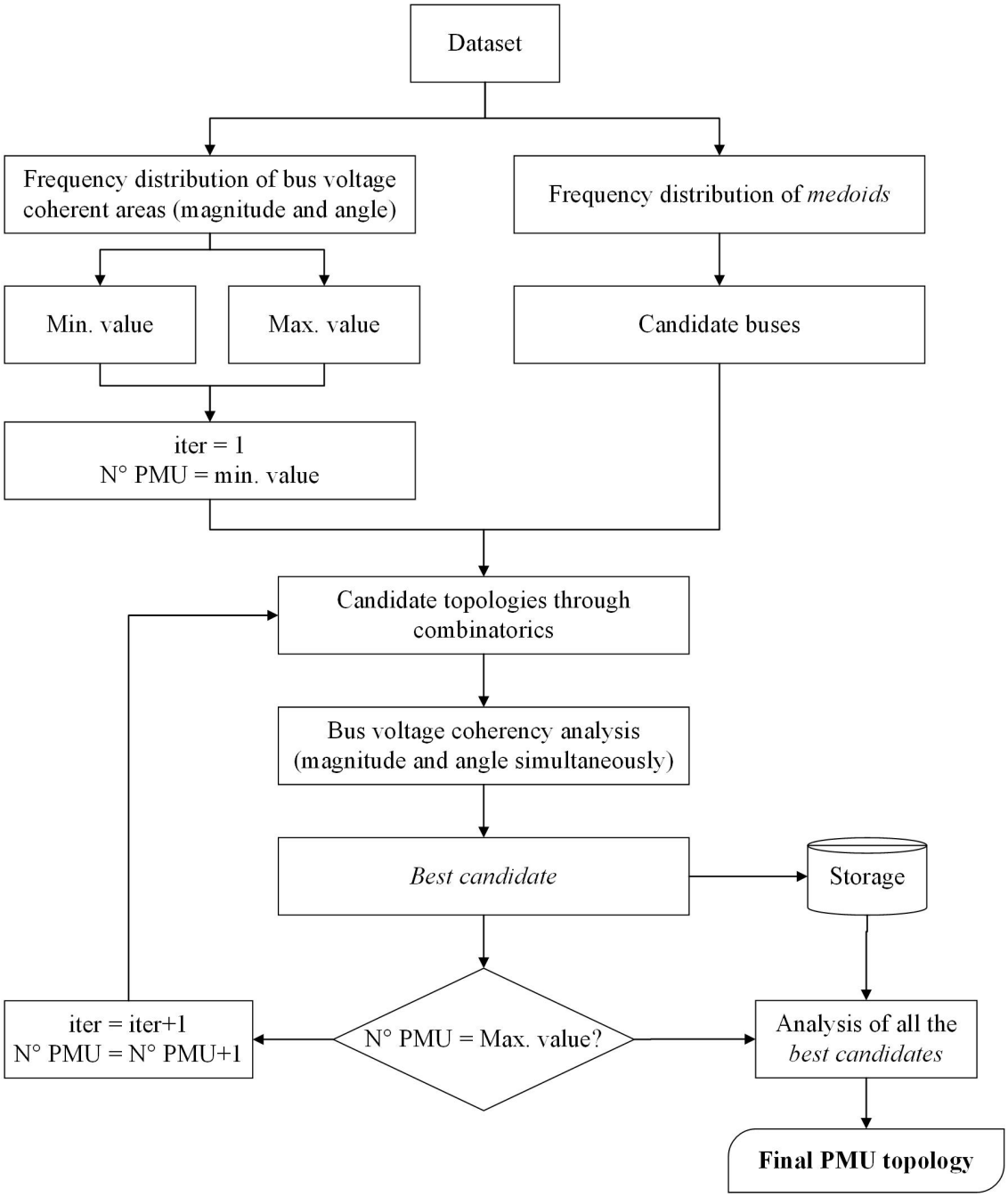

Fig. 3.5. PMU topology: iterative exploring process.

3.3 Classifier and generation of bus voltage pseudo-measurements

The classifier based on the k-nearest neighbors algorithm presented in this work is capable of operating at high speed and founds its performance on a classification matrix that is generated offline from the preprocessed dataset in the PMU topology approach.

The construction of the classification matrix involves two steps: first, scenarios are selected so that the sample includes all the classes. Second, for each scenario an assignment vector that shows the connection between observed and non-observed buses is defined. Here, class and PMU location are involved. In cases where there is the possibility of assigning more than one observed bus, the final solution will be the one with the minimum electrical distance. Finally, the classification matrix is built from the selected scenarios and the assignment vectors.

In the presence of a new and unknown scenario, the classifier is able to predict the class and the assignment vector at high speed. To achieve this, it uses the pre-contingency system loading condition provided by the estimator and the type of contingency (inputs). This information is pre-processed following the same rules as in the second step of the PMU topology approach, i.e. variables are normalized, dimension is reduced and the scenario is classified according to the pre-contingency system loading condition with a previously trained Random Forest classifier. Then, once the information that belongs to the new scenario has been adjusted, it is compared with the information inside the classification matrix to identify the element with highest similarity by means of the Euclidean distance between inputs. The chosen element defines the class and the assignment vector for the new scenario, see Figure 3.6.

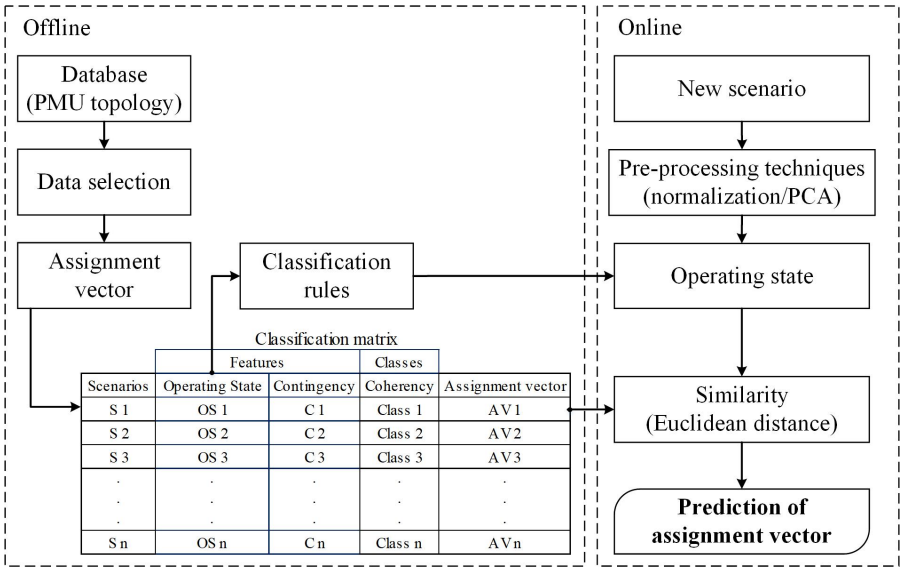

Fig. 3.6. Classifier.

Finally, the bus voltage pseudo-measurements at PMU non-observed buses are obtained from results delivered by the classifier, the information provided by PMUs and the estimated states at previous time using the following expression:

$$\widetilde{vps}_i^U = \hat{x}_{k-1}^U + \widetilde{\Delta V}_{PMU}^{O \in U} \tag{3.1}$$

where $\widetilde{vps}_i^U$ is the calculated bus voltage pseudo-measurement at the non-observed bus U at time i, $\hat{x}_{k-1}^U$ is the estimated static state at non-observed bus U at time k-1 and $\widetilde{\Delta V}_{PMU}^{O \in U}$ is the voltage deviation registered from the PMU measurements at the observed bus O that is assigned to the non-observed bus U according to the assignment vector.

Since bus voltage pseudo-measurements are calculated from PMU measurements at coherent buses and the estimated state at previous time, they are considered less accurate than real measurements. Therefore, a standard deviation equal to ten times the standard deviation of phasor measurements is considered.

3.4 Summary

This chapter presents the methodology for generating bus voltage pseudo-measurements that allows the hybrid state estimator to monitor the power system behavior in transient regime. This approach is based on the concept of bus voltage coherency. Because coherency changes all the time depending on the kind of contingency and the system loading condition, the first part of this chapter focuses on defining the number and location of PMUs so that all coherent areas in voltage magnitude and angle can be observed regardless the kind of contingency and the system loading condition. This is achieved through an offline data-mining based approach and a large set of scenarios that is obtained from numerical simulation by means of the Monte Carlo technique. Then, a classifier based on the k-nearest neighbors algorithm will predict bus voltage coherency at high speed and with that an assignment vector that dictates the relation between observed and non-observed buses. The classifier relies on a classification matrix that is created offline from the data used to define the PMU topology. Finally, bus voltage pseudo-measurements are generated using bus voltage measurements (magnitude and angle) from PMUs, the system loading condition before the contingency that is obtained from the previous estimated static state and the assignment vector predicted by the classifier.

4 Proposed hybrid state estimator

This chapter presents the hybrid state estimator developed in this thesis. First, an overview of the estimation procedure is given. Next, its operation is explained in detail. Algorithms employed and mathematical models used by the estimator to represent the static and dynamic behavior of the power system are described. The approach used for detecting and identifying large errors is presented.

4.1 Overview

The hybrid state estimation procedure described below consists of two phases depending on the system operating conditions as shown in the flowchart of the Figure 4.1. When the power system is in stationary regime only phase one takes place and the static state together with all the variables that come from it such as line currents, power flows and power injections, are estimated at SCADA speed by a WLS based static state estimator. Conventional as well as PMU measurements (phasor of current and voltage) are used together in one single estimation run. Once the estimation is completed, a post-estimation method is carried out in order to detect and identify bad measurements. Weights corresponding to bad measurements are decreased in order to minimize their effect on the result and the estimation is rerun. Input in phase one includes the grid topology, parameters of lines and transformers, mathematical models that represent the network behavior in stationary regime and both kind of measurements.

Phase two is activated when a physical disturbance happens and the system is in transient regime. In this case, two estimators work in sequence at PMU speed making use of voltage and current measurements (magnitude and angle) provided by such units. This allows estimating the power system dynamics associated with any kind of transient event. The kind of transient event will define how long stage two will be active. First, a static state estimator is used to estimate bus voltages as soon as the PMU measurement set arrives. Here, a major drawback is the lack of available PMUs that makes full grid observability unfeasible. This issue is solved through the novel data-mining based methodology for restoring full PMU observability explained in chapter 3. This approach uses as input the estimated static state at the previous time, phasors of bus voltage provided by PMUs and information regarding the contingency such as failed element and kind of fault and generates voltage pseudo-measurements. Since a post-estimation bad data processing technique is unfeasible at fast scan rates such as those of PMU

measurements, the static estimator in phase two uses the information provided by the bad data processor from phase one to adjust the measurement weights if needed. Additionally, in order to make phase two robust against PMU's gross errors, a WLAV based linear estimator is used. Second, dynamic variables such as rotor angle and speed of generators and speed of motors are estimated by means of an EKF based dynamic state estimator. The estimator uses as input the estimated static state, mathematical models that represent the dynamic behavior of the systems under consideration and their parameters. Phase two will run until the power system reaches the stationary regime again.

The hybrid state estimation approach proposed in this work was thought and designed under the following assumptions:

- Measurement and process errors follow a normal distribution, with standard deviation σ and mean $\mu = 0$.
- Uncertainty of networks parameters is considered and modelled by using a uniform distribution function with specified lower and upper bounds.
- In stationary regime (phase one in the estimation procedure) conventional measurements as well as phasor synchronized measurements are used together in the same estimation run. The time-skew between measurements is usually tolerated due to the slowly varying operating conditions of power systems under normal operating conditions [25].
- The contingency together with its time of occurrence is known. Despite the fact this assumption represents an appreciable simplification of the problem, advances in computer science have lately led to several promising works [84–86] in the field.
- The power system operates under balanced condition and the load composition at each bus (static–dynamic load ratio) is known.

The sequence of operation of the proposed estimation methodology is shown in Figure 4.2.

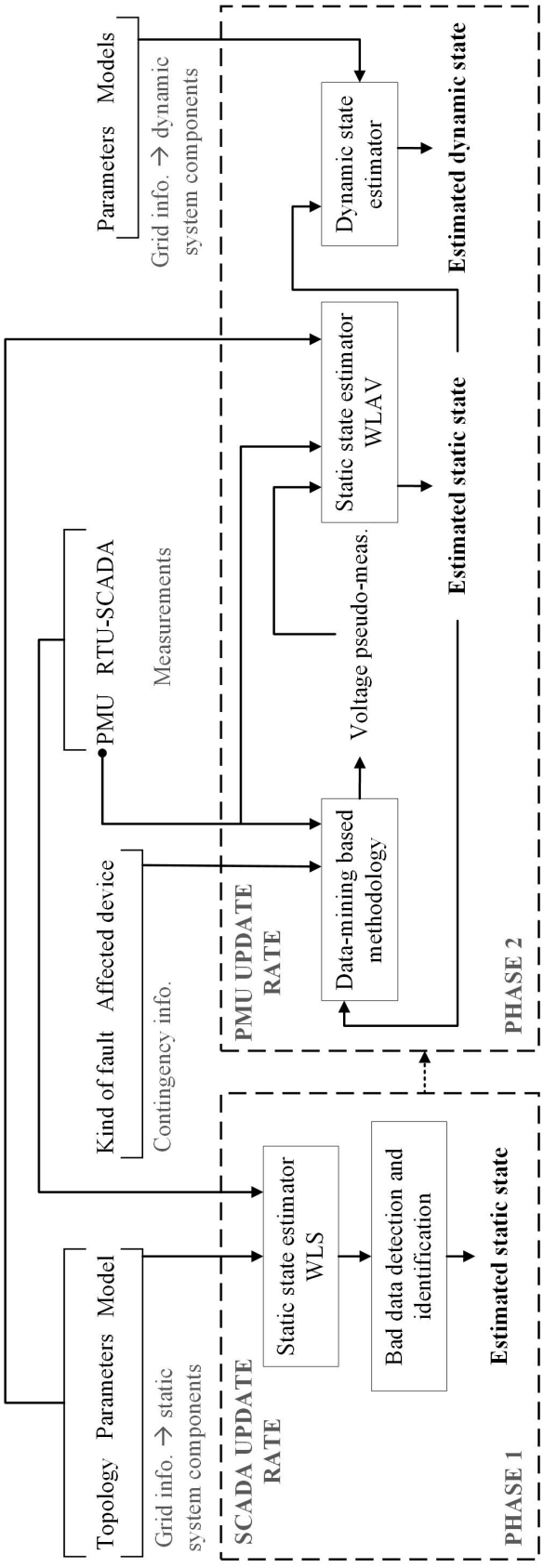

Fig. 4.1. Estimation procedure followed by the proposed hybrid state estimator.

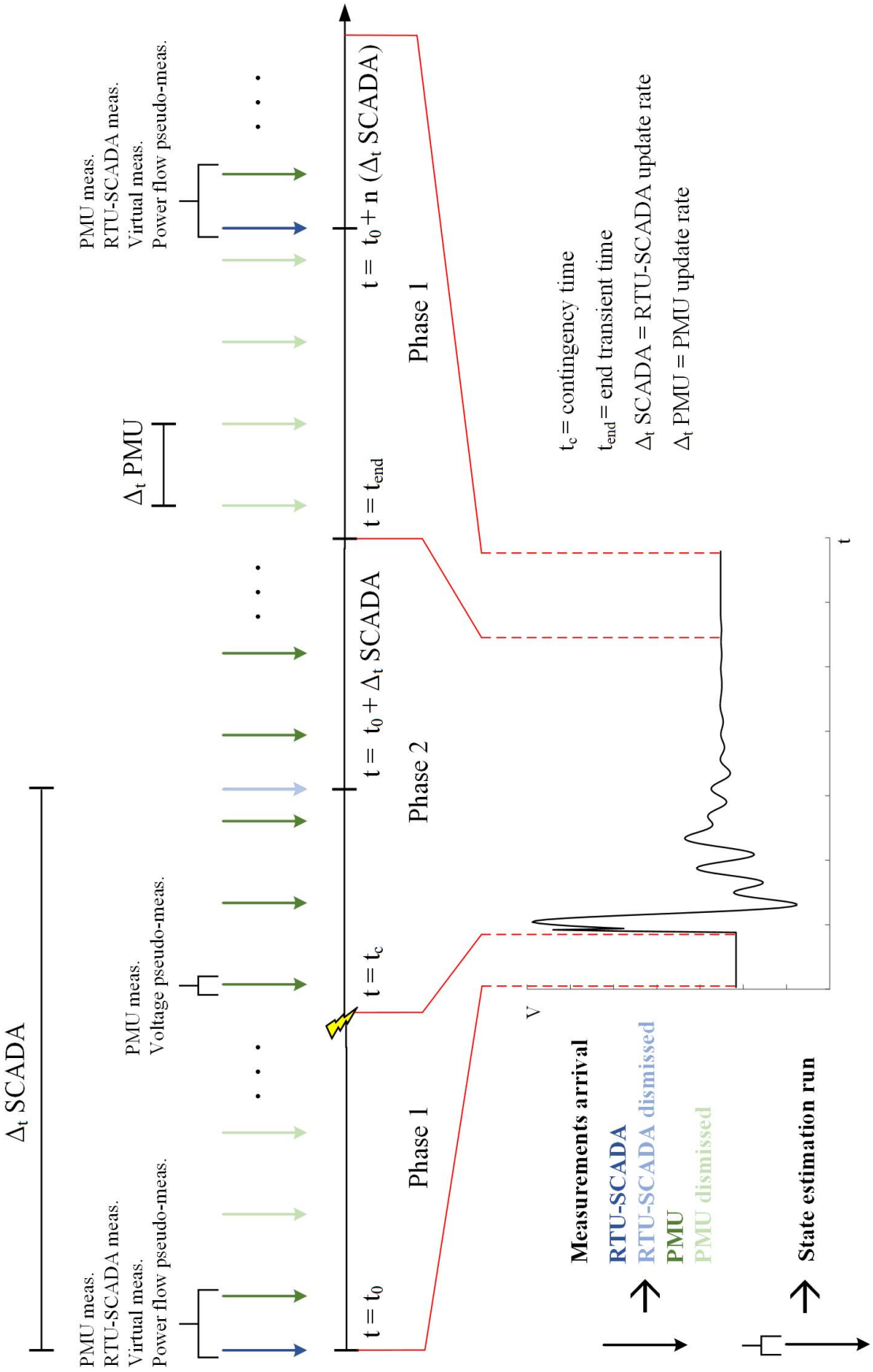

Fig. 4.2. Sequence of operation of the proposed hybrid state estimator.

4.2 Static state estimation

This subsection describes the static state estimation approaches used by the hybrid estimator in both operating phases.

Firstly, mathematical models used by the estimation algorithms to represent the network behavior are presented and described. Then, static estimators employed in each phase are presented and explained in detail.

4.2.1 Component modeling

It is considered that the power system operates under balanced conditions. Loads and branch power flows are three phase and balanced, all transmission lines are fully transposed, and compensation devices are symmetrical in the three phases. Such assumptions allow the use of single phase positive sequence equivalent circuit for modeling purpose [25].

Transmission lines

Transmission lines are represented by a two-port pi-model whose parameters correspond to the positive sequence equivalent circuit of transmission lines. A transmission line with a positive sequence series admittance of $\tilde{y}_{ij} = g_{ij}+jb_{ij}$ and a shunt admittance between bus i and the ground of $\tilde{y}_{sij} = g_{sij}+jb_{sij}$, will be modelled by the following equivalent circuit [25]:

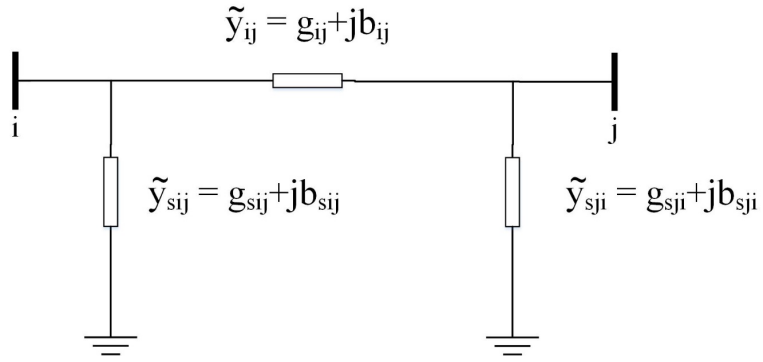

Fig. 4.3: Equivalent circuit for a transmission line.

Shunt capacitors or reactors

Shunt capacitors or reactors which may be used for voltage and/or reactive power control, are represented by their per phase susceptance at the corresponding bus.

Tap changing and phase shifting transformers

Transformers with off-nominal but in-phase taps, can be modeled as series impedances in series with ideal transformers [25]:

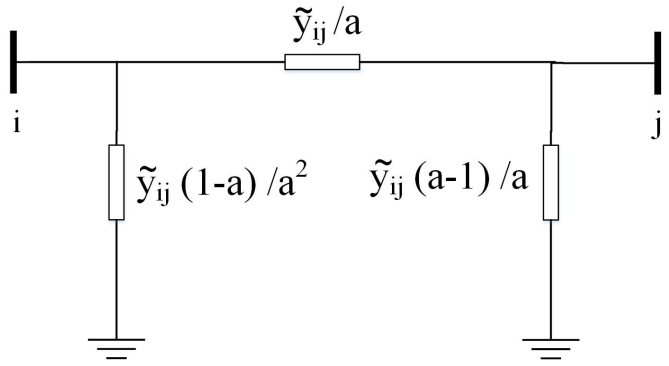

Fig. 4.4: Equivalent circuit of an in-phase tap changer transformer.

where $\tilde{y}_{ij}$ is the positive series admittance of the branch i-j and a is the in-phase tap ratio. For a phase shifting transformer where the off-nominal tap value is complex, the admittance matrix will not be longer symmetrical. Therefore, a passive equivalent circuit can no longer be realized for the phase shifting transformer [25]. The following set of nodal equations is used to model the transformer behavior in such case [25]:

$$\begin{bmatrix} \tilde{\mathbf{i}}_{ij} \\ \tilde{\mathbf{i}}_{ji} \end{bmatrix} = \begin{bmatrix} \frac{\tilde{y}}{|a|^2} & \frac{-\tilde{y}}{a^*} \\ \frac{-\tilde{y}}{a} & \tilde{y} \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{v}}_i \\ \tilde{\mathbf{v}}_j \end{bmatrix} \tag{4.1}$$

Loads

Loads models are traditionally classified into two categories: static models and dynamic models. In this thesis a composite load model which includes both kind of models is taken into consideration for load modeling.

In phase one (stationary regime) loads and generators are modeled as constant power injections. When the system is in phase two (transient regime) generators are dynamically modelled and loads are represented using a composite load model which involves a static and a dynamic component (motor). Since dynamic models are employed by the dynamic estimator they will be introduced and explained later in detail.

The static component of the composite model represents the dependency at a given time of the active and reactive powers on the connection bus voltage and the system frequency in quasi-

static state. This state is characterized by slow modifications of the operating conditions, and the transition from one state to another can be considered a series of steady states. The modeling of these characteristics involves only algebraic equations.

The polynomial or ZIP model is the one used in phase two to represent the voltage and frequency dependency of loads. It is composed of constant power ($\bar{V}^o$), constant current ($\bar{V}^1$) and constant impedance ($\bar{V}^2$) components [24]:

$$\begin{aligned} p_{st} &= p_{st0}[p_1\bar{V}^2+p_2\bar{V}+p_3](1+K_{pf}\Delta f) \\ q_{st} &= q_{st0}[q_1\bar{V}^2+q_2\bar{V}+q_3](1+K_{qf}\Delta f) \\ &\text{where } \bar{V}=\frac{v}{v_0},\ \Delta f=f-f_0 \end{aligned} \tag{4.2}$$

where p_{st0}, q_{st0}, v_0 and f_0 represent power consumption of the static load (active and reactive power respectively), bus voltage magnitude and frequency at the initial operating condition. Coefficients p_1 to p_3 and q_1 to q_3 define the proportion of each component in the model. Typically, K_{pf} ranges from 0 to 0.3, and K_{qf} ranges from -0.2 to 0 [24]. The bus frequency $\hat{f}$ is obtained either directly from PMUs when the bus is measured by such devices or computed as follows:

$$\hat{f}=\frac{\Delta\theta}{2\pi\,\Delta t_{PMU}}+f_{PMU\_next} \tag{4.3}$$

where $\Delta\theta=\theta-\theta_0$ is the bus voltage angle deviation and Δt_{PMU} is the PMU refresh rate. f_{PMU_next} is the frequency reported by the closest PMU to the bus under analysis.

Network model

The described component models can be used to make the static network model for the entire system. This is done by writing a set of nodal equations which are obtained using Kirchhoff's current law at each bus [25]:

$$\mathbf{I}=\begin{bmatrix}\tilde{i}_1\\ \tilde{i}_2\\ \vdots\\ \tilde{i}_N\end{bmatrix}=\begin{bmatrix}\tilde{y}_{11} & \cdots & \tilde{y}_{1N}\\ \tilde{y}_{21} & \cdots & \tilde{y}_{2N}\\ \vdots & \vdots & \vdots\\ \tilde{y}_{N1} & \cdots & \tilde{y}_{NN}\end{bmatrix}\begin{bmatrix}\tilde{v}_1\\ \tilde{v}_2\\ \vdots\\ \tilde{v}_N\end{bmatrix}=\mathbf{YV} \tag{4.4}$$

where $\tilde{i}_k$ is the current injection at bus k, $\tilde{v}_k$ is the voltage at bus k and $\tilde{y}_{km}$ is the (k,m)-th element of the bus admittance matrix **Y**. Admittance matrix is in general complex and symmetric. The use of nodal equations facilitates the modification of equations in case of topology changes [25].

4.2.2 Phase one

In phase one the static state, i.e. bus voltage magnitudes and angles, and all the algebraic variables that come from it such as power flows, power injections and current magnitudes and angles by branches, will be estimated at SCADA speed using a WLS based static state estimator (see subsection 2.2.2, Weighted least squares). Despite the fact the non-linear WLS method is iterative, it performs quite fast thanks to its efficient implementation [28, 68].

The estimator employed in this case has been modified in order to include PMU measurements together with conventional measurements. This allows, first, improving the accuracy of the estimated state [14] and, second, detecting and identifying bad PMU data so that this information can be used later in phase two. Time-skew errors will not be an issue here since the system is assumed in stationary regime.

The measurement vector consists of conventional measurements of power injection $\tilde{\mathbf{s}}_{\mathbf{inj}}$, power flow $\tilde{\mathbf{s}}_{\mathbf{fl}}$ and voltage magnitude **v**, PMU voltage measurements $\tilde{\mathbf{v}}_{\mathbf{PMU(M<\theta)}}$ in polar coordinates, virtual measurements $\tilde{\mathbf{s}}_{\mathbf{virt}}$ (zero injection), and power flow pseudo-measurements $\tilde{\mathbf{s}}_{\mathbf{ps}}$:

$$\mathbf{z_1}=\left[\mathbf{v}\ \tilde{\mathbf{v}}_{\mathbf{PMU(M<\theta)}}\ \tilde{\mathbf{s}}_{\mathbf{inj}}\ \tilde{\mathbf{s}}_{\mathbf{fl}}\ \tilde{\mathbf{s}}_{\mathbf{virt}}\ \tilde{\mathbf{s}}_{\mathbf{ps}}\right]^T \tag{4.5}$$

The incorporation of current measurements in traditional estimators introduces ill-conditioning problems. That is why, PMU current measurements are replaced by power flow pseudo-measurements, which are obtained as follows:

$$\tilde{s}_{ps.\ i\text{-}j} = \tilde{v}_{PMU(M<\theta)\ i} \times \tilde{i}^{*}_{PMU(M<\alpha)\ i\text{-}j} \tag{4.6}$$

where $\tilde{s}_{ps.\ i\text{-}j}$ is the power flow pseudo-measurement by branch i-j, $\tilde{v}_{PMU(M<\theta)\ i}$ is the PMU voltage measurement at bus i and $\tilde{i}^{*}_{PMU(M<\alpha)\ i\text{-}j}$ is the conjugate of the PMU current measurement by branch i-j.

Standard deviations of power flow pseudo-measurements are calculated using the following expression:

$$u_{P\_Q_{ps.\ i\text{-}j}} = \sqrt{\sum_{k=1}^{4} \left(\frac{\partial P\_Q_{ps.\ i\text{-}j}}{\partial p_k} \right)^2 \times u_{p_k}^2} \qquad (4.7)$$

$$\mathbf{p} = (\tilde{v}_{PMU(M<\theta)\ i}, \tilde{i}^{*}_{PMU(M<\alpha)\ i\text{-}j})$$

in (4.7) $u_{P\_Q_{ps.\ i\text{-}j}}$ is the standard deviation of power flow pseudo-measurement (active or reactive power) by branch i-j, **p** is a vector that contains the PMU measurements of voltage and current involved in the calculation of power flow pseudo-measurements, p_k represents the k-th element of such vector, and u_{pk} is the standard deviation of the element p_k.

Considering the measurement model in (2.1) and the network model given above (subsection 4.2.1 Component modeling), power injection and power flow measurements can be expressed in terms of state variables as follows:

$$p_i = v_i^2 \sum_{j\in\vartheta(i)} (g_{ij}+g_{sij}) - v_i \sum_{j\in\vartheta(i)} v_j \alpha_{ij}$$

$$q_i = -v_i^2 \sum_{j\in\vartheta(i)} (b_{ij}+b_{sij}) - v_i \sum_{j\in\vartheta(i)} v_j \beta_{ij} \qquad (4.8)$$

$$p_{ij} = v_i^2 (g_{ij}+g_{sij}) - v_i v_j \alpha_{ij}$$

$$q_{ij} = -v_i^2 (b_{ij}+b_{sij}) - v_i v_j \beta_{ij} \qquad (4.9)$$

where $\vartheta(i)$ is the set of buses connected to bus i. Elements α_{ij} and β_{ij} are defined as follows:

$$\alpha_{ij} = g_{ij} \cos(\theta_i-\theta_j) + b_{ij} \sin(\theta_i-\theta_j)$$

$$\beta_{ij} = g_{ij} \sin(\theta_i-\theta_j) - b_{ij} \cos(\theta_i-\theta_j) \qquad (4.10)$$

The static state is estimated iteratively using the Normal equations in (2.15) (subsection 2.2.2, Weighted least squares).

Despite its efficiency, the WLS approach is non-robust and will be biased even in the presence of a single bad measurement. That is why, the estimation process is followed by a bad data processor method. Error detection is performed by assessing the target function $J(\hat{\boldsymbol{x}})$ by means of the Chi-square test. Then, the measurement with the largest normalized residual is identified

and its weight is modified in order to reduce its effect on the result, after which the estimator reruns (see subsection 2.2.4, Bad data processing).

The static estimator in phase one is depicted by the following flowchart:

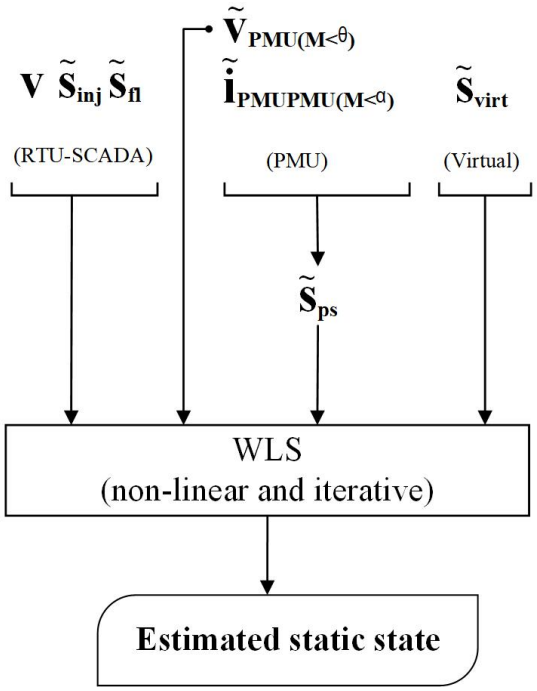

Fig. 4.5. Static state estimator (phase one)

4.2.3 Phase two

During stage one static states are estimated by a WLS based estimator at SCADA speed whenever conventional measurements arrive. However, unlike phase one, in phase two only PMU measurements together with bus voltage pseudo-measurements are processed since the system is no longer in stationary regime and therefore the use of measurements non-synchronized and at different reporting rate could constitute a source of error. Pseudo-measurements are obtained by means of the data-mining based approach presented and explained in the previous chapter. The problem is transformed from polar to rectangular coordinates in order to get a constant and numerical Jacobian (**H**) of the functions **h**. The measurement vector can be expressed as follows:

$$\mathbf{z_2} = \left[\tilde{\mathbf{v}}_{\mathbf{PMU(x,y)}}\ \tilde{\mathbf{i}}_{\mathbf{PMU(x,y)}}\ \widetilde{\mathbf{vps}}_{\mathbf{(x,y)}}\right]^{T} \tag{4.11}$$

Current by branch i-j can be written in rectangular coordinates as a function of state variables as follows:

$$\begin{aligned} i_{i\text{-}j,r} &= v_i \left[(g_{ij}+g_{sij})\cos\theta_i-(b_{ij}+b_{sij})\sin\theta_i\right] - v_j \left[g_{ij}\cos\theta_j-b_{ij}\sin\theta_j\right] \\ i_{i\text{-}j,i} &= v_i \left[(b_{ij}+b_{sij})\cos\theta_i-(g_{ij}+g_{sij})\sin\theta_i\right] - v_j \left[b_{ij}\cos\theta_j-g_{ij}\sin\theta_j\right] \end{aligned} \tag{4.12}$$

where subscripts “r” and “i” represent the real and imaginary parts of the current phasor. Standard deviations of the measurements which initially are expressed in polar coordinates must be also adapted to rectangular coordinates. This is accomplished using the classical theory of propagation of uncertainty [56]:

$$\begin{aligned} u_{V_{i,r}} &= \sqrt{(\cos\theta_i)^2 u_{V_i}^2+(v_i\sin\theta_i)^2 u_{\delta_i}^2} \\ u_{V_{i,i}} &= \sqrt{(\sin\theta_i)^2 u_{V_i}^2+(v_i\cos\theta_i)^2 u_{\delta_i}^2} \end{aligned} \tag{4.13}$$

where $u_{V_{i,r}}$ and $u_{V_{i,i}}$ are standard deviations of voltage measurement at bus i real and imaginary part and, $u_{V_i}^2$ and $u_{\delta_i}^2$ are standard deviations of voltage measurement at bus i magnitude and angle respectively.

As mentioned before, despite the WLS technique is computationally efficient, it does not provide any robustness against gross errors. As a result, a post-estimation bad data processing method for detecting, identifying and eliminating bad data when using WLS approaches becomes mandatory. The main drawback of such a technique when applied into stage two is its high computational cost, which makes it inefficient and unfeasible at fast scan rates. That is why, aiming at improving the algorithm robustness in transient regime when PMUs are affected by large errors, a robust weighted least absolute value (WLAV) based linear estimator is used instead of the conventional WLS to estimate grid bus voltages since it can automatically reject bad measurements during the estimation process [29, 87] without any additional post-estimation processing.

This method has not been implemented in the past mainly for two reasons: first, when using RTU-SCADA measurements the problem becomes nonlinear and computationally slow and second, the existence of “leverage points” causes the algorithm fail to automatically eliminate large errors. However, it has been recently shown that these disadvantages could be removed if the network were fully observed only by PMUs. Under such condition, the problem becomes linear eliminating the need for repeated solutions and the vulnerability of leverage points can

be fixed by simple use of scaling [29]. Scaling can be carried out by dividing each column of the Jacobian matrix, **H**, by the largest entry of that column, then the same procedure is applied to the rows [88].

When only phasor measurements are available the measurement equation takes the following form [28]:

$$\mathbf{z}=\mathbf{H}\hat{\mathbf{x}}+\mathbf{r} \tag{4.14}$$

where **r** is the measurement residual and the Jacobian **H** is a constant matrix. The state estimated $\hat{\mathbf{x}}$ will be given by the solution of the following optimization problem [28]:

$$\begin{aligned} &\text{Min. } \mathbf{m}^{T}\,|\mathbf{r}| \\ &\text{s.t. } \mathbf{z}-\mathbf{H}\hat{\mathbf{x}}=\mathbf{r} \end{aligned} \tag{4.15}$$

where **m** is a weight vector. The optimization problem shown above is expressed as an equivalent linear programming (LP) problem [28] and solved using the interior point solver [89].

4.3 Dynamic state estimation

So far the proposed hybrid estimator is able to provide information about power system static state in stationary as well as in transient regime at high speed. However, nothing has been said in relation to dynamic variables that play a major role in power system stability and dynamic security assessment.

This subsection focus on the proposed dynamic estimation approach in charge of estimating dynamic variables. In first place, the models used to represent the behavior of dynamic grid components such as generators and motors are presented and described. Next, the dynamic estimator used in the second phase of the proposed hybrid estimator is explained in detail.

4.3.1 Component modeling

The definition of power system stability and its classification presented above (see subsection 2.1.2, Power system stability) allows identifying those components whose operating conditions affect power system stability at most. Since power systems rely mostly on synchronous machines for generation of electrical power, a necessary condition for satisfactory system operation is that all synchronous machines remain in synchronism. This aspect is influenced by the dynamics of generator rotor angles and power-angle relationships.

Nevertheless, instability may also occur without loss of synchronism. A system based on synchronous machine supplying an induction motor through transmission line can become unstable

due to the collapse of load voltage. Keeping machines in synchronism is not an issue in such a case; instead the concern is stability and control of voltage [24]. Even though voltage instability is a local phenomenon, its consequences have a great effect on the system operation, leading sometimes to a voltage collapse. Voltage collapse is usually associated to the reactive power of load not being met because of limitations on the provision and transmission of reactive power. Voltage stability is "strongly influenced by the static characteristics and the dynamic response of the load" [23]. Furthermore, when motors are subjected to low voltage, they tend to stall and consume excessive reactive power from the network. As a result, the equilibrium between supply and demand cannot be kept, causing voltage stability problems. Thus, induction motors are usually "the driven force for voltage instability". Their rotation speed can be used as indicator of short-term voltage stability. Usually, conventional under voltage load shedding schemes used to mitigate short-term voltage stability problems do not take into account the dynamics of motors. The triggering signal for load shedding is the voltage magnitude alone, which is not always a reliable indicator [90].

That is why, in this thesis generators and motors are dynamically modelled with the aim of estimating dynamic variables that are of interest in power system stability assessment such as rotor angle and speed of generators and speed of induction motors.

Synchronous generator

For stability analysis of large systems the stator transients of generators are usually neglected. By doing this, the stator quantities contain only fundamental frequency components and the stator voltage equations appear as algebraic equations. This allows the use of steady-state relationships for representing the interconnecting transmission network. The transients associated with the network decay very rapidly and there is little justification for modeling their effects in stability studies. "The network transients cannot be neglected unless machine stator transients are also neglected; otherwise equations representing various elements of the power system will be inconsistent" [24].

In addition, it is assumed that speed changes are small and do not have significant effect on voltage. This assumption "counterbalances the effect of neglecting the stator transients so far as the low-frequency rotor oscillations are concerned" [24].

Based on the assumptions made before, the fourth-order model in d-q reference frame is used by the dynamic state estimator to represent the synchronous machine behavior which can be expressed by means of the following differential-algebraic system of equations in per unit [23]:

$$
\begin{aligned}
t'_{d0}\frac{de'_q}{dt} &= -e'_q-(x_d-x'_d)i_d+e_{fd}\\
t'_{q0}\frac{de'_d}{dt} &= -e'_d-(x_q-x'_q)i_q\\
\frac{d\delta}{dt} &= \omega_0\omega\\
2H\frac{d\omega}{dt} &= p_m-p_e-D\omega\\
v_d=v\sin(\delta-\theta),\ v_q &= v\cos(\delta-\theta)\\
i_d=\frac{e'_q-v_q}{x'_d},\ i_q &= \frac{v_d-e'_d}{x'_q}\\
p_e=v_di_d+v_qi_q,\ q_e &= -v_di_q+v_qi_d
\end{aligned}
\tag{4.16}
$$

where t'_{d0} and t'_{q0} are the open-circuit transient time constants. e'_d, e'_q and e_{fd} are the transient voltages along the d and q axes and the field voltage. x_d, x_q, x'_d, and x'_q are synchronous and transient reactance. i_d and i_q are the d and q axis current. δ is the angle giving the position of rotor, ω_0 is the nominal synchronous speed and ω is the speed deviation. H is the inertia constant and D is the damping coefficient. p_m, p_e and q_e are the mechanical and electrical power (active and reactive) of the generator. v and θ are the terminal bus voltage magnitude and phase angle and v_d and v_q are the terminal bus voltage in the d and q axes.

Dynamic load model (induction motor)

Typically, motor consumption represents 60 to 70 % of the total energy supplied by a power system. Therefore, the dynamics attributable to motors are usually the most significant aspects of dynamic characteristics of system loads.

The acceleration of the induction motor can be represented in per unit by [90]:

$$
2H\frac{d\omega}{dt}=t_e-t_m \tag{4.17}
$$

where:

$$t_m = t_{m0}\left(\frac{\omega}{\omega_0}\right)^2$$
$$p_{mot} = \frac{t_e\left(r_s + \frac{r_r}{s}\right)}{\left(\frac{r_r}{s}\right)},\ s = 1-\omega \tag{4.18}$$

in equations (4.17) and (4.18) ω is the rotation speed; H is the inertia constant; t_e and t_m are the electromagnetic and mechanical torque; p_{mot} is the power consumption of the motor; r_s and r_r are the stator and rotor resistance; s is the slip; p_{mot0}, t_{m0}, and ω_0 are active power, mechanical torque and rotation speed under nominal condition. t_e is assumed to be equal to t_m in steady state.

4.3.2 EKF based dynamic state estimator in phase two

In the second phase of the hybrid estimator an EKF based dynamic state estimator is in charge of estimating dynamic variables of generators and motors. The algorithm is chosen from different alternatives for nonlinear systems (such as UKF, PF and ensemble KF) because its high computational efficiency makes it suitable for use in real-time monitoring and control functions of transient phenomena, since it is required that the estimation of current states must be finished before the next set of measurement arrives. In other words, the state estimation algorithm must be able to keep up with the measurement data flow.

On the contrary, when a system is highly nonlinear, EKF tends to have poor estimation accuracy and even diverge. This is because first order Taylor approximation used in this algorithm introduces too much error [9]. However, it has been shown that increasing the measurement sampling rate through interpolation methods can improve the estimation accuracy but at the expense of increasing the computation time [9, 91].

The dynamic estimator is implemented in the hybrid scheme using a model decoupling technique called “event play back”, which has the potential to decouple the EKF problem to better use of the measurement data at terminal buses. Measurement data, which in the proposed hybrid scheme are the estimated static state (WLAV estimator in phase two), are separated in two groups. One group is treated as input signals to the dynamic model and the other group is treated as measurements or output in the estimation problem. Using such a technique, there is no longer the need to deal with the dynamic model of an entire power system [45]. Rather, small-scale

dynamic models of generators and motors will be used in the EKF in parallel. Estimators and measurements involved in phase two are depicted in the flowchart in Figure 4.6.

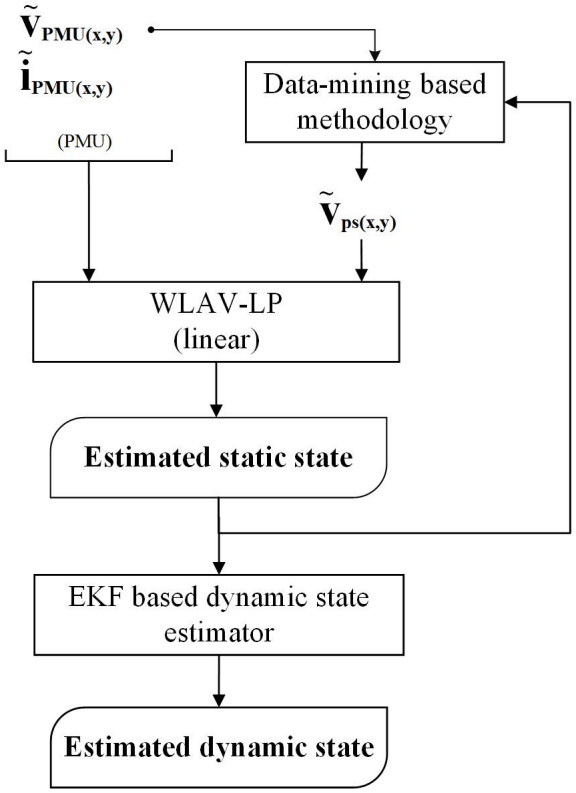

Fig. 4.6. Static and dynamic state estimators (phase 2).

The dynamic model of generator in (4.16) can be represented on state-space model following the structure in (2.4) and (2.5) as follows:

$$\mathbf{x}=[x_1\ x_2\ x_3\ x_4]^t=[\delta\ \omega\ e'_d\ e'_q]^t$$

$$\mathbf{u}=[u_1\ u_2\ u_3\ u_4]^t=[e_{fd}\ v\ \theta\ p_m]^t$$

$$\dot{x}_1=\omega_0 x_2$$

$$\dot{x}_2=\frac{p_m-v\sin(\delta-\theta)\frac{x_4-v\cos(\delta-\theta)}{x'_d}-v\cos(\delta-\theta)\frac{v\sin(\delta-\theta)-x_3}{x'_q}-Dx_2}{2H} \qquad (4.19)$$

$$\dot{x}_3=\frac{-x_3-(x_q-x'_q)\frac{v\sin(\delta-\theta)-x_3}{x'_q}}{t'_{q0}}$$

$$\dot{x}_4=\frac{-x_4-(x_d-x_d')\dfrac{x_4-v\cos(\delta-\theta)}{x_d'}}{t_{d0}'}$$

$$\mathbf{z}=[z_1\ z_2\ ]^t=[p_e\ \ q_e]^t$$

$$z_1=v\sin(\delta-\theta)\frac{x_4-v\cos(\delta-\theta)}{x_d'}+v\cos(\delta-\theta)\frac{v\sin(\delta-\theta)-x_3}{x_q'}$$

$$z_2=-v\sin(\delta-\theta)\frac{v\sin(\delta-\theta)-x_3}{x_q'}+v\cos(\delta-\theta)\frac{x_4-v\cos(\delta-\theta)}{x_d'}$$

In equation (4.19) the bus voltage at terminal bus (magnitude v and angle θ) together with measurements of electrical power (active p_e and reactive q_e) are obtained from the linear static estimator. The field voltage (e_{fd}) and the mechanical power (p_m) are assumed measurable and therefore known. Nevertheless, in practice these quantities may not be accessible. In such case alternative approaches could be used so that the solution can be reached even if the inputs are not given. Among possible alternatives, the following can be mentioned:

a) The inclusion of the governor and the automatic voltage regulator (AVR) models in the machine formulation. Under such condition, the unknown inputs become internal variables and the reference values for voltage and power, which are known, will be the new inputs
b) The use of alternative estimation algorithms that consider unknown inputs, such the one proposed in [46]

To derive the discrete-time EKF algorithm, the system of equations presented above must be discretized and rewritten properly as follows:

$$\begin{aligned}\mathbf{x}_k&=\mathbf{f}_{k-1}(\mathbf{x}_{k-1},\mathbf{u}_{k-1},\mathbf{w}_{k-1})\\ \mathbf{z}_k&=\mathbf{h}_k(\mathbf{x}_k,\mathbf{u}_k,\mathbf{e}_k)\end{aligned}\qquad(4.20)$$

In equation (4.20) k and k-1 indicate the instant of time t=kΔt and t=(k-1)Δt respectively; $\mathbf{x}_k$ is the state vector; $\mathbf{u}_k$ is the input vector; $\mathbf{f}_k$ is the system function; $\mathbf{z}_k$ the measurable output; $\mathbf{h}_k$ the output function; $\mathbf{w}_k$ represents either the process noise or inaccuracies in the system model and $\mathbf{e}_k$ is the measurement noise. Both noises as mentioned before are supposed white, Gaussian with zero mean and independent of each other. The dynamic state in generators is estimated following the two-step prediction-correction process given in detail in equations (2.26)-(2.29)

(subsection 2.2.1, Discrete-time extended Kalman Filter). The overall structure of the estimation process in generators is illustrated in Figure 4.7. Variables with a hat operator represents estimated quantities.

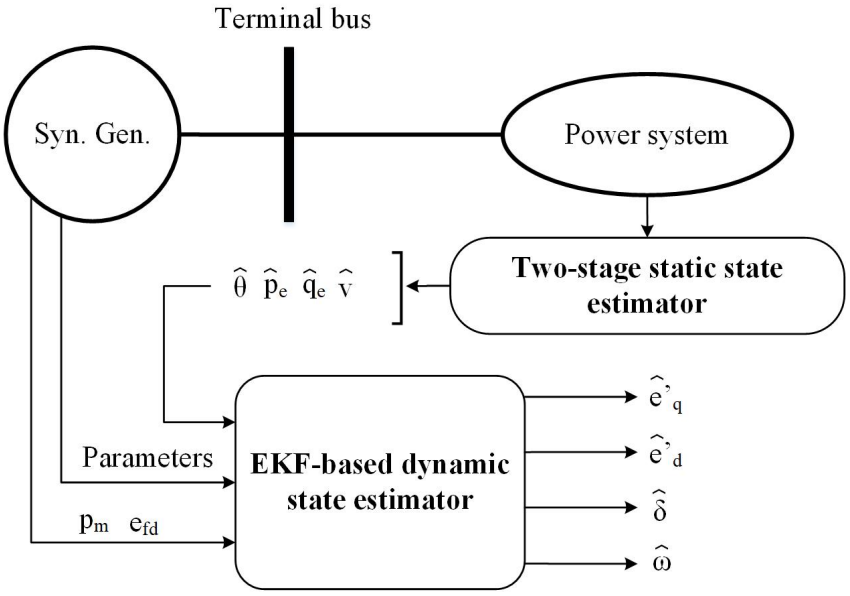

Fig. 4.7. Dynamic state estimation process for generators (phase two).

The dynamic model of the motor in equations (4.17) and (4.18) can be represented on state-space model as follows:

$$\mathbf{x}=x_1=\omega$$

$$\mathbf{u}=u_1=t_e$$

$$\dot{x}_1=\frac{t_e - t_{m0}\left(\frac{x_1}{\omega_0}\right)^2}{2H} \tag{4.21}$$

$$z_1=p_{mot}$$

$$z_1=t_e\,\frac{r_s+\frac{r_r}{1-x_1}}{\frac{r_r}{1-x_1}}$$

As mentioned before, in this work a composite load model is considered. It is made up of two parts: an induction motor and a static load. The power contribution of each one in stationary state regime is assumed known. During a transient event, the power consumption of the static load (active p_{st} and reactive q_{st}) is approximated from the estimated static state and the PMU

measurement of frequency using the expressions (4.2) and (4.3). The power injection at terminal bus (active p_e and reactive q_e) is estimated by the static estimator. Then, outputs, i.e. the power demanded by motors during a transient event, can be determined as the difference between power injection and power consumption of the static load as follows:

$$p_{mot}=p_e-p_{st}$$
$$q_{mot}=q_e-q_{st} \tag{4.22}$$

As with generators the system of equations must be discretized in order to derive the discrete-time EKF algorithm. After that, the dynamic variable is estimated following the two-step prediction-correction process given in detail in equations (2.26)-(2.29) (subsection 2.2.1, Discrete-time extended Kalman Filter).

The estimation process of motors is depicted by the flowchart in Figure 4.8. Variables with a hat operator represents estimated quantities.

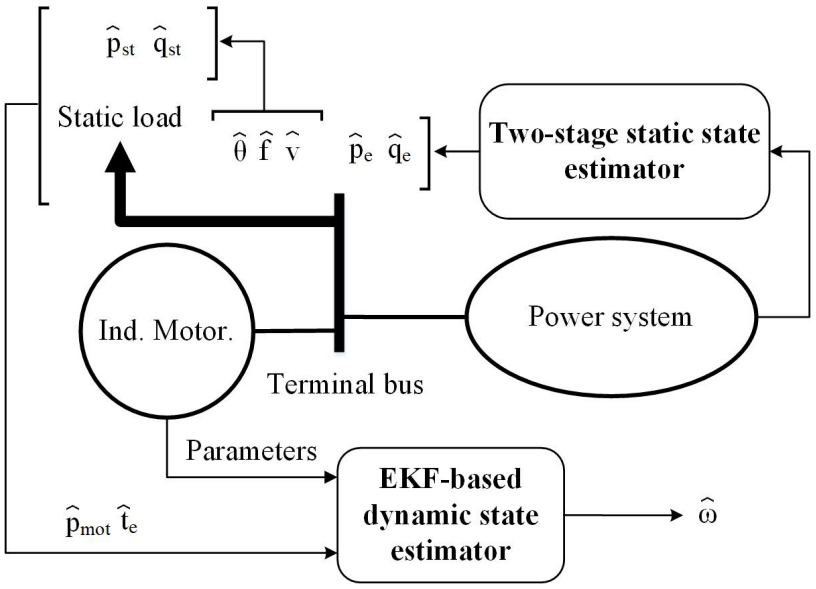

Fig. 4.8. Dynamic state estimation process for motors (phase two).

4.4 Summary

This chapter presents the hybrid state estimator developed in this thesis for monitoring the power system dynamics associated to slow and fast transient events. First a short description about the estimation procedure is given. Then, the features of the estimator are presented and described according to the kind of variable to be estimated, either static or dynamic. Models used by the estimator to represent the behavior of the power system on static and dynamic condition as well as the estimation algorithms employed in each phase are explained in detail.

The hybrid estimator basically consists of two phases. In stationary regime phase one takes place and the static state is estimated at SCADA using a WLS approach. Conventional as well as PMU measurements are used together in one single estimation run. Once the estimation is completed, a post-estimation method based on residuals is carried out with the aim of filtering gross measurement errors. Phase two occurs when the system operates in transient regime. This time, two estimators work in sequence at PMU speed. First, a WLAV estimator is used to estimate bus voltages as soon as the PMU measurement set arrives. Since a post-estimation bad data processing technique is unfeasible at fast scan rates, the static estimator in phase two uses the information provided by the bad data processor from phase one to adjust the measurement weights. Finally, dynamic variables such as rotor angle and speed of generators and speed of motors are estimated by means of an EKF based dynamic state estimator.

Unlike the PMU-only state estimators proposed in the literature, the hybrid approach developed in this thesis does not require a PMU installed at each generator or motor terminal bus to estimate the dynamic state in such buses. The WLAV based linear estimator in phase two is in charge of delivering the necessary information to the EKF so that it can carry out its work regardless of local PMU availability.

5 Test results and analysis

In this chapter the performance of the proposed hybrid state estimator together with the methodology for generating bus voltage pseudo-measurements are tested and assessed. The New England benchmark system is used to this end. Performance parameters such as accuracy, computing time, convergence and robustness are evaluated. In addition, it is analyzed how the redundancy of pseudo-measurements and the use of different estimation algorithms in phase two affects the estimated result. The analysis of results highlights the ability of the proposed estimator to accurately monitor power system dynamics associated to slow and fast transient events.

5.1 Description of the test system

The hybrid state estimator together with the methodology for generating bus voltage pseudo-measurements are tested on the New England benchmark system. It consists of 39 buses and 10 generators. The nominal frequency is 60 Hz and the main voltage level is 345 kV. Some changes were introduced into the grid with the aim of satisfying the N-1 security criterion [69]. In addition, motors have been included together with static loads [69]. The test system single-line diagram is depicted in Figure 5.1 and its data is presented in detail in Appendix A.

The software “PowerFactory” is used to generate simulation data that imitate responses of a real system. Dynamic simulations are run using an integration step size of 10 ms. An update rate of 2.4 s for conventional measurements and 40 ms for synchronized measurements is considered.

In order to obtain more realistic scenarios and to test the estimator under more real conditions, model inadequacy and process noises are considered. This is accomplished by using more complex models to generate the simulation data in relation to the models used by the hybrid state estimator to estimate the states. Generators are represented in detail using a full sixth-order electromechanical model including transient and sub-transient dynamics in addition to governor and excitation systems. Motors are modelled with a seventh-order double cage rotor model. Static loads are represented using a polynomial model where voltage and frequency dependencies are considered (details can be found in Appendix A).

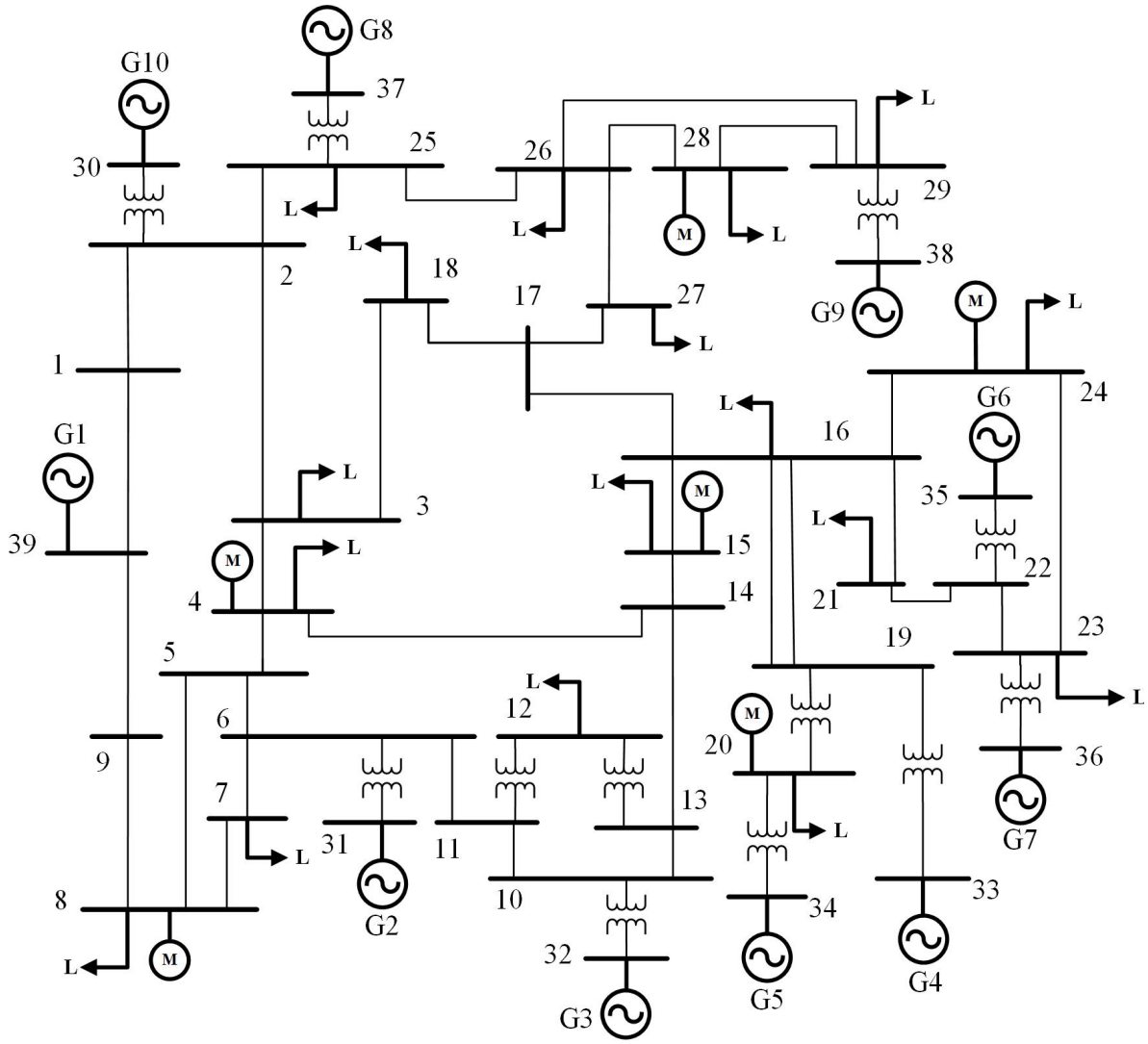

Fig. 5.1. New England benchmark system.

Measurements are obtained by adding a random error with normal distribution and zero mean to results from simulation. Standard deviations are computed from the maximum uncertainty for each kind of measurement (see Table 5.1).

Transmission line parameters are usually assumed to be exact and known. However, in practice this assumption do not hold true. In this work, the uncertainty of line parameters is considered and modelled using a uniform distribution function with typical upper and lower bounds of ± 2%.

Table 5.1 Maximum measurement uncertainties [13]

Conventional measurements			PMU measurements		
p_i	q_i	p_{fl} q_{fl} v	v i	θ α	f
2 %			0.5 %	0.02°	0.02 %

5.2 Measurement topology

In this work it is assumed that the conventional measurement system ensures the estimation problem solution in phase one, i.e. the power grid is observable. Location and amount of conventional measurements are defined in order to get a maximum spanning tree (topological approach). A measurement redundancy of 1.2 is achieved.

In order to determine the PMU topology, a database made up of 10,800 operating scenarios is generated from numerical simulations using the Monte Carlo method. As mentioned before, three kinds of N-1 contingencies are simulated: three-phase fault, load curtailment and generator outage. The database comprises stable and unstable scenarios. A fault clearing time of 80 ms and a simulation time of 10 s are considered.

Then, coherency is evaluated in each scenario using the k-medoids based double-loop routine explained above (subsection 3.2, Probabilistic approach for defining the PMU topology), which is applied to time series of voltage magnitude and angle. Similarity is calculated through the Euclidean distance between Wavelet coefficients. The choice of the clustering method and the distance metric used in this work is based on an appropriate robustness-processing time ratio. As an example, Figures 5.2 and 5.3 depict the clusters of voltage magnitude and angle defined by the proposed approach when a three-phase fault on line 23-24 occurs. Bus voltage magnitudes and angles are grouped together in 5 and 4 clusters respectively.

Each coherent scenario for bus voltage magnitude and angle is identified through the use of integers. The clustering routine determines 4,419 different coherent scenarios for voltage angle and 6,019 different coherent scenarios for voltage magnitude over the total 10,800 simulated scenarios. It is found that the great majority of classes or coherent scenarios (71.7% and 77.2% for voltage angle and magnitude respectively) have just one example, namely the rate of occurrence is one, which clearly conducts to an imbalanced multi-class dataset. Besides, most scenarios are grouped in 4 and 5 clusters. The class distribution as a function of the number of coherent areas before adapting the dataset is depicted in Figure 5.4.

Pre-processing techniques are applied on the dataset with the aim of making them appropriate for analysis and mining:

1) First, the data dimension is reduced by means of PCA considering the pre-contingency system loading condition. The number of principal components is chosen so that to achieve a cumulative proportion of variance explained of at least 90%. This means that the new set of

variables will represent the variability of the original dataset in the above mentioned proportion. This is achieved by using the first four principal components (see Figure 5.5);

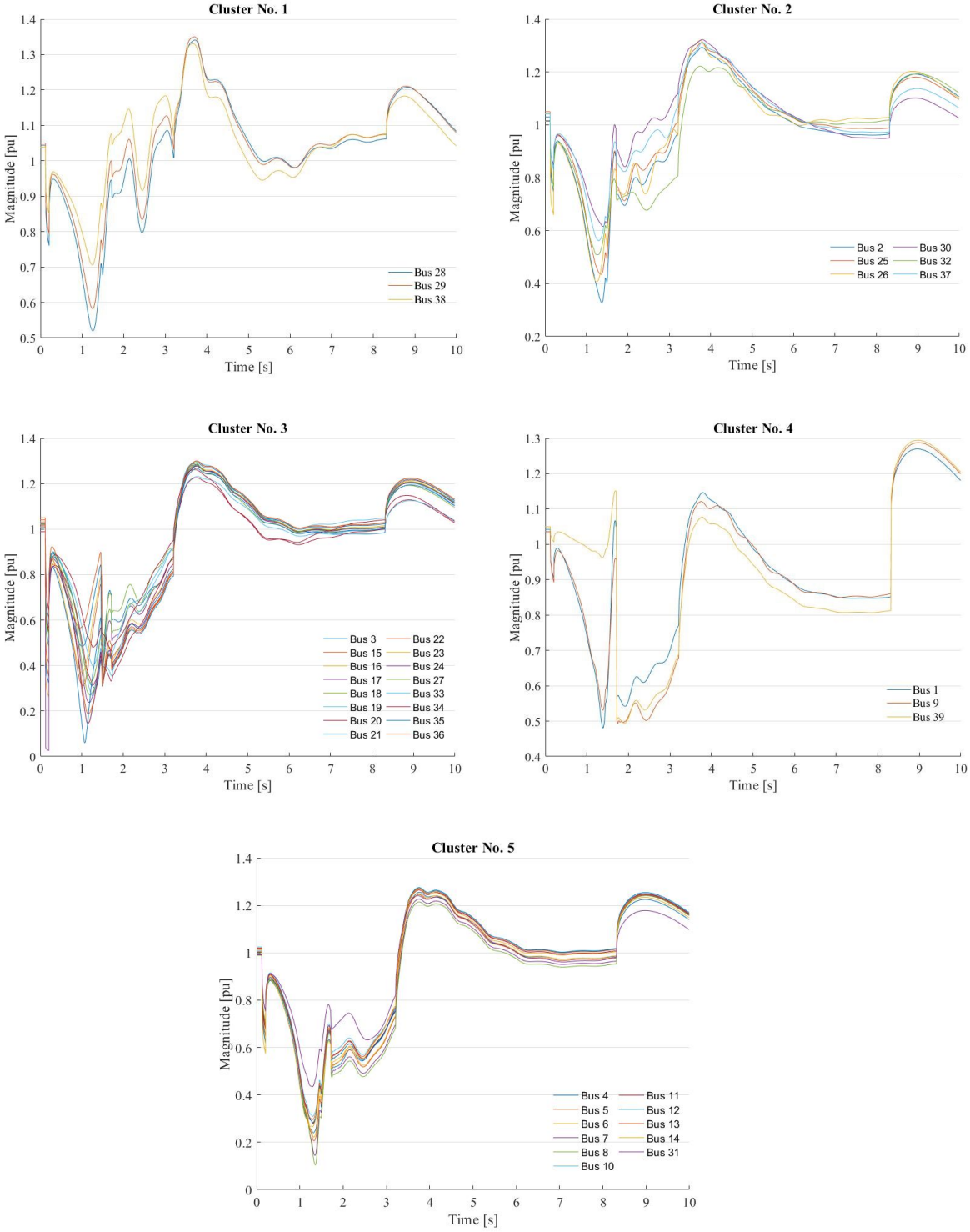

Fig. 5.2. Coherent areas under a three-phase fault on line 23-24.
Voltage magnitude

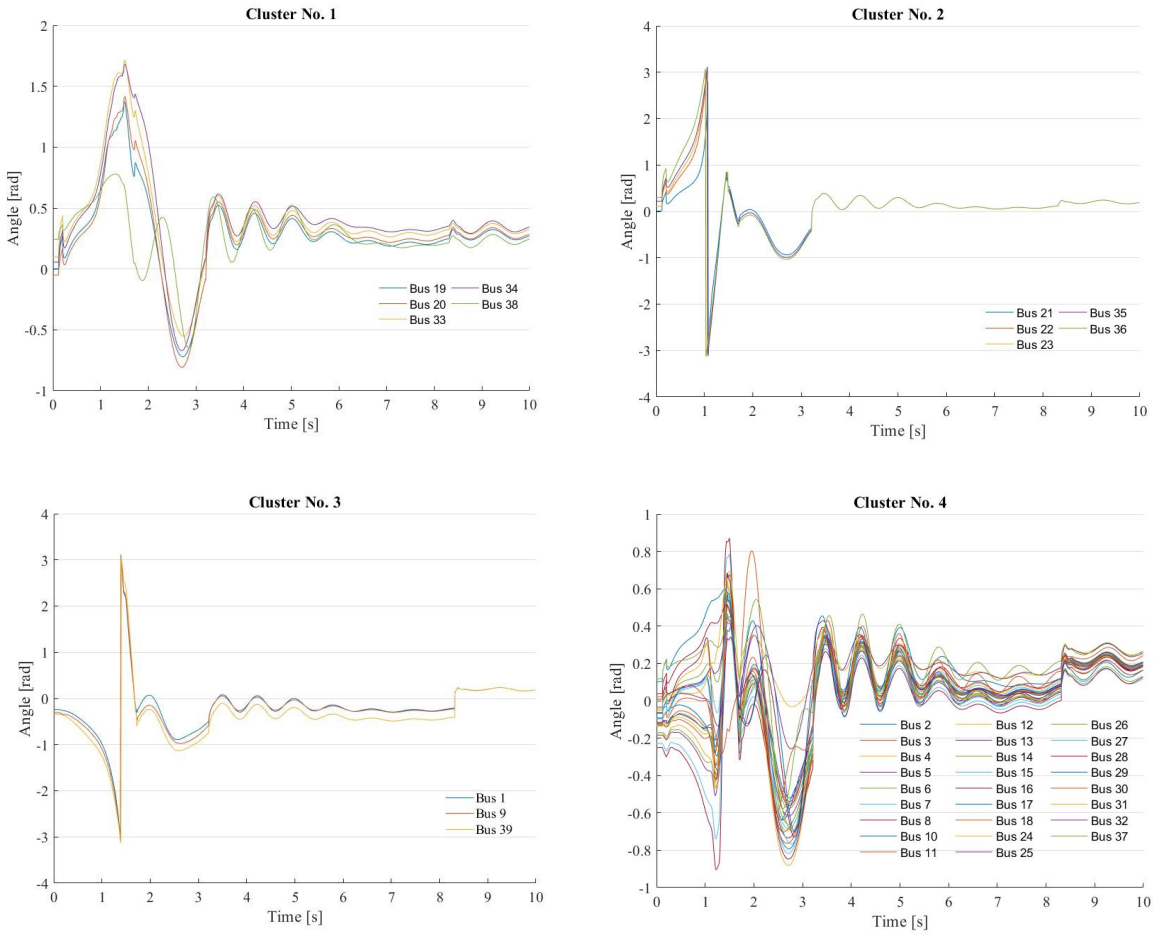

Fig. 5.3. Coherent areas under a three-phase fault on line 23-24.
Voltage angle

2) Next, variables are normalized with the aim of adjusting values measured on different scales to a common scale and scenarios are grouped together according to the system loading condition. The k-means with Euclidean distance is used to this end. The number of clusters is defined by plotting the change of the intra-cluster variance as a function of the number of clusters and then choosing that point from where there is no significant change in variance. Scenarios are grouped into eight clusters. Figure 5.6 shows the clusters obtained as a function of the three first components;
3) An analysis on the reduced and normalized dataset reveals scenarios with similar feature values but belonging to different classes or coherent scenarios, which constitutes a clear symptom of overlapping classes. Further investigation into classes allows detecting small differences between coherent areas in each of them. Here, the three-step filtering process explained above is implemented. First, “boundary” scenarios with similar feature values but different classes are identified. Then, the degree of dissimilarity between pairs of boundaries scenarios is calculated by measuring the number of buses that contribute to having different

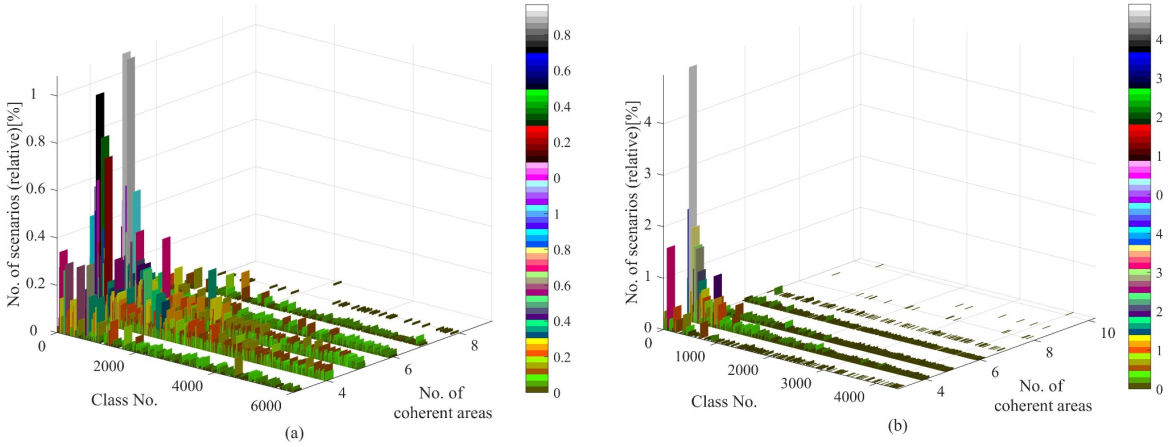

Fig. 5.4. Class distribution before filtering process.
(a) Voltage magnitude, (b) Voltage angle

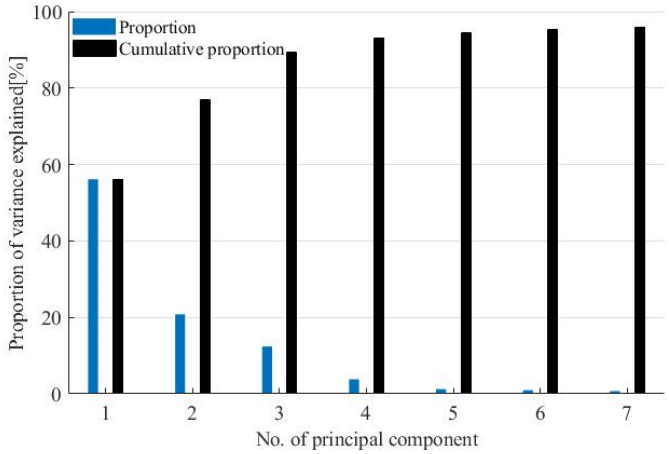

Fig. 5.5. Variance explained (PCA)

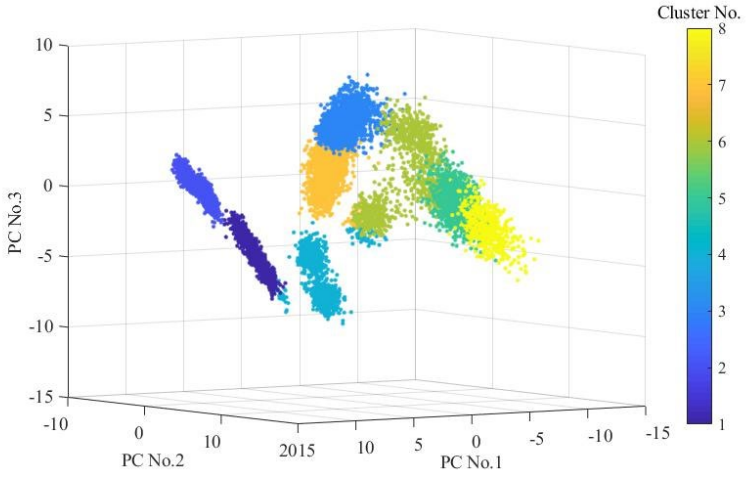

Fig. 5.6. Clustering according to pre-contingency system loading condition

coherent areas in both scenarios. When the number of buses falls below 10% of the total amount of buses in the grid under analysis both scenarios adopt the same class. This process

is carried out until all the scenarios that meet this condition are “absorbed” by their corresponding pairs. The direction of absorption is defined so that the larger clusters absorb scenarios from the smaller ones. As a result of this first filtering stage the number of classes (coherent scenarios) is reduced by 71.8% and 70.8% for voltage angle and magnitude respectively. The second stage operates only on coherent scenarios with a rate of occurrence of one. The main idea here is to approximate these scenarios to clusters with similar feature values found in the immediate vicinity of the scenario under analysis. This time, the number of classes is further reduced by 5.5% and 5.2% for voltage angle and magnitude respectively. Finally, the last stage dismisses coherent scenarios whose rate of occurrence is equal to one, two and/or three. As a result, 1,804 simulated scenarios (16.7%) are dismissed. The last filter reduces the number of classes by another 11.7% and 14.5% for voltage angle and magnitude respectively. Figure 5.7 depicts using a bar plot the evolution of the number of classes according to each of the filtering steps mentioned above. In addition, Figure 5.8 shows the new distribution of classes. As can be seen, the filtering process leads to a more uniform distribution of class avoiding the well-known class imbalance problem. Figure 5.9 depicts the dataset distribution as a function of the pre-contingency system loading condition and the kind of contingency before and after the last step of the filtering process. From the histograms it can be noticed that the removal of scenarios performed in the last step of the filtering process does not significantly affect the distribution of the original data. Moreover, the removed scenarios are less likely to occur.

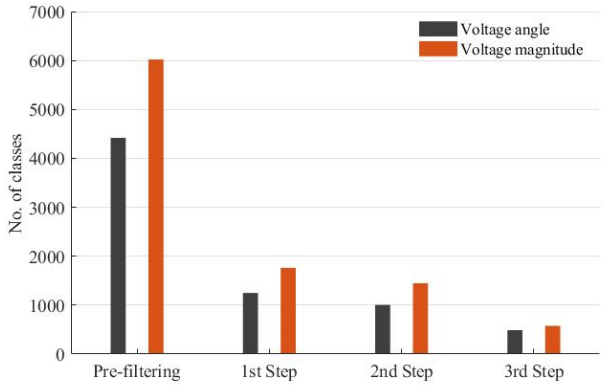

Fig. 5.7. Effect of the filtering process on the number of classes (coherent scenarios)

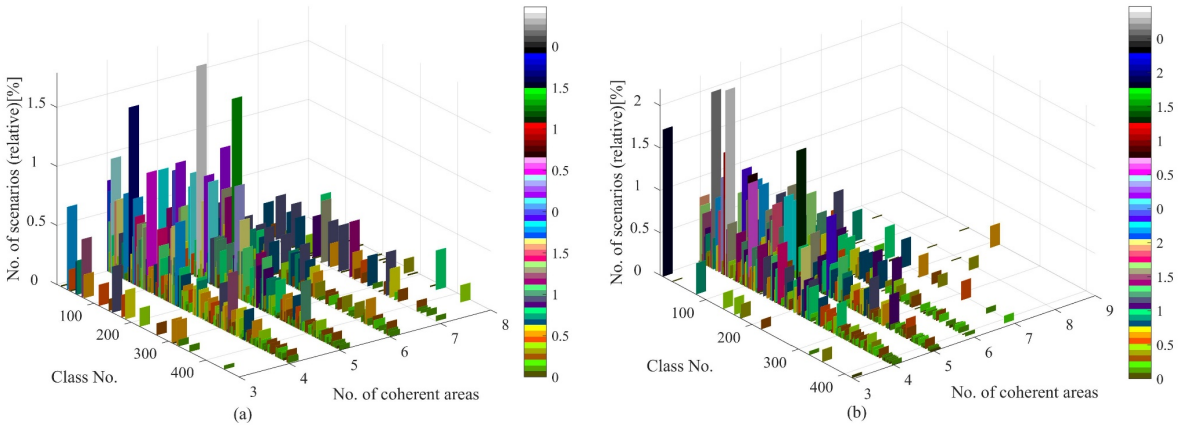

Fig. 5.8. Class distribution after filtering process.
(a) Voltage magnitude, (b) Voltage angle

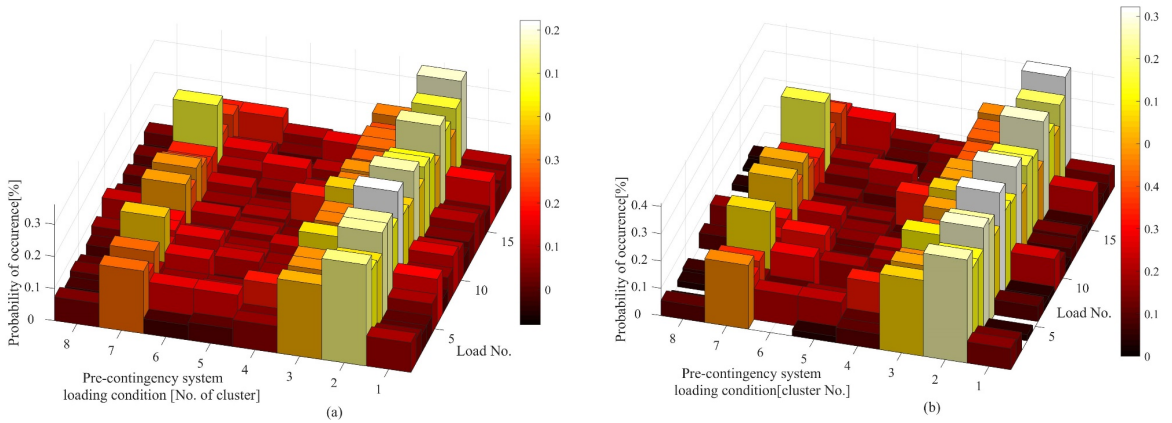

Fig. 5.9. Dataset distribution according to the kind of contingency.
Load curtailment. (a) before and (b) after filtering process

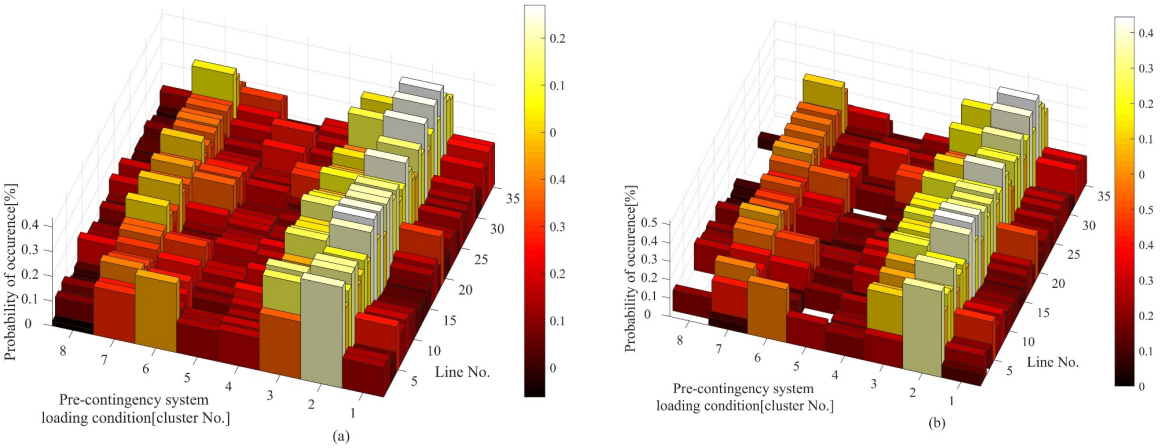

Fig. 5.10. Dataset distribution according to the kind of contingency.
Three-phase fault (a) before and (b) after filtering process

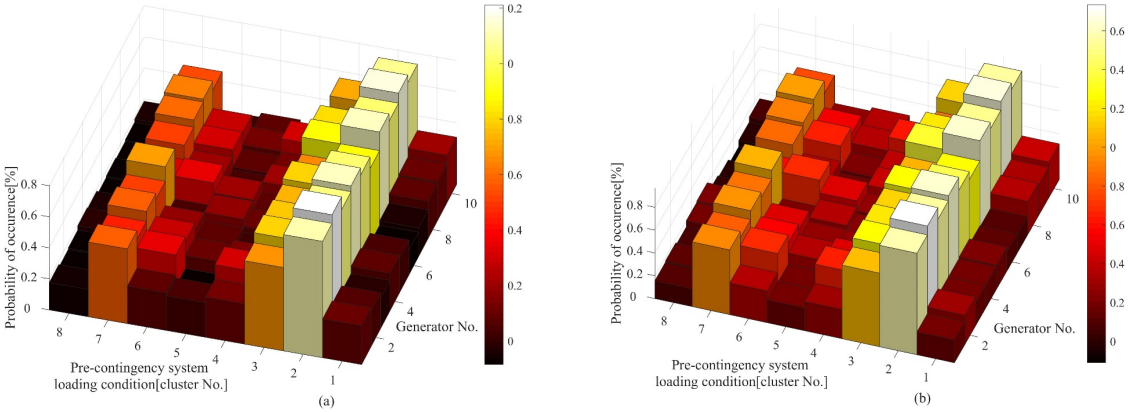

Fig. 5.11. Dataset distribution according to the kind of contingency. Generator outage (a) before and (b) after filtering process

Finally, the number of PMUs and their location is found through the iterative exploring process explained in subsection 3.2. Buses that are medoids simultaneously in voltage magnitude and angle in at least 10% of the total simulated scenarios are taken into consideration when generating the PMU candidate topologies by means of combinatorics. According to the class distribution shown in Figure 5.8 topologies ranging from a minimum of 3 to a maximum of 9 PMUs are analyzed. In order to speed up the process, only topologies that offer a higher degree of network observability are considered. Configurations for which the degree of observability falls by 10% or more of its maximum value are discarded from the analysis. Table 5.2 shows a summary of the results obtained. It contains the number of candidate topologies according to the amount of installed PMUs and the best solution for each case defined as the one with the least number of non-observed scenarios over the total amount of simulated scenarios. The solution chosen in this work for the New England system involves 7 PMUs at buses 2, 6, 9, 14, 20, 23 and 29. This decision is based on the fact that from 7 units an increase in the number of PMUs does not lead to an appreciable improvement in the number of observed scenarios. The solution results in 18% of measured buses, 72% of observed buses and 28% of non-observed buses. The PMU topology adopted allows observing all the coherent areas in voltage magnitude and angle simultaneously in 92.8% of all simulated scenarios. A key aspect to be emphasized from the obtained results and the selected PMU topology is that the number of units involved in it represents a reduction of 40% in the minimum amount of units required to observe the whole system if only PMU measurements were considered. This is not a minor issue when trying to find a way of dealing with overload on communication and data processing systems in large-scale applications due to the incorporation of PMUs. Figure 5.12 depicts the measurement topology (PMU and SCADA measurements) adopted in the New England system.

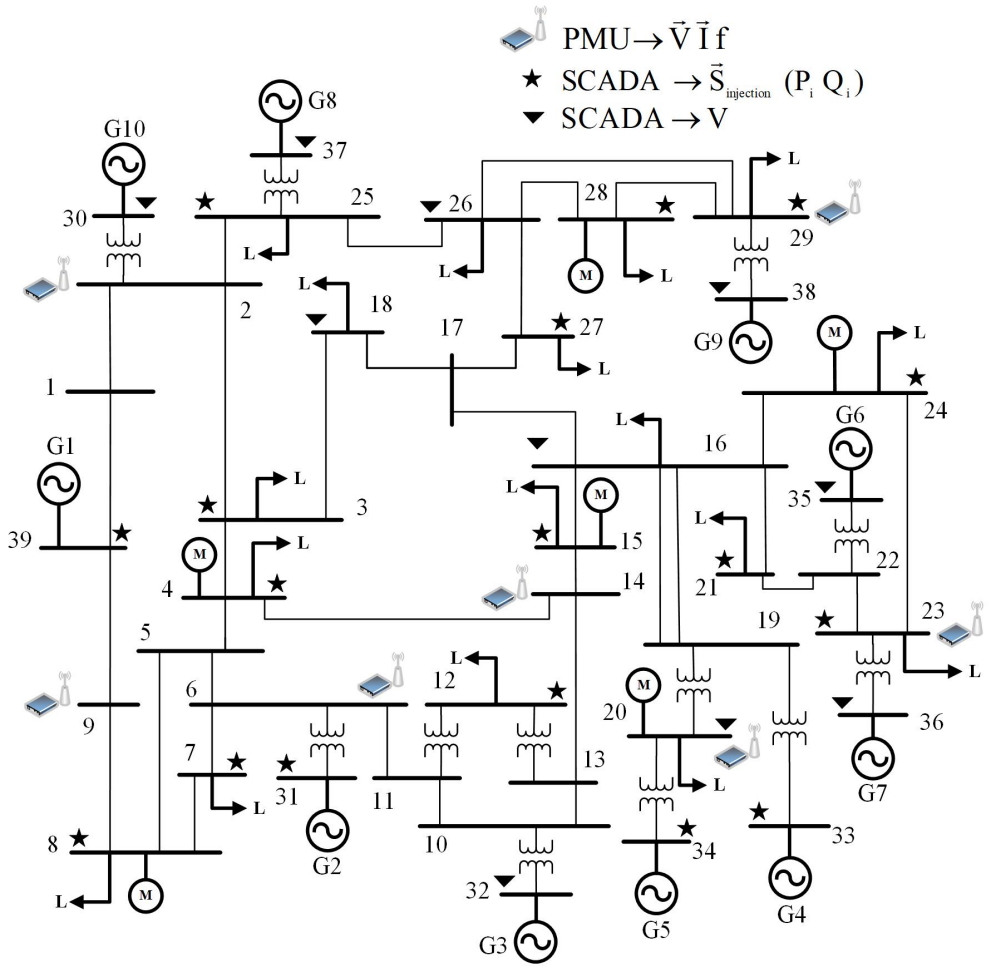

Fig. 5.12. Measurement topology adopted in the New England system

Table 5.2 Summary of results (iterative exploring process)

No. of PMUs	3	4	5	6	7	8	9
No. of candidate topologies	13	131	137	325	376	1100	1698
Best candidate [Buses]	2, 16, 29	2, 6, 19, 29	2, 6, 19, 22, 29	2, 6, 14, 20, 23, 29	2, 6, 9, 14, 20, 23, 29	2, 6, 9, 13, 14, 20, 23, 29	2, 5, 6, 9, 13, 16, 20, 23, 29
Success rate [%]	9.5	66.9	78.7	86.7	92.8	93.7	94.6

5.3 Coherency forecasting and accuracy analysis of bus voltage pseudo-measurements

First, a classification matrix is built from a sample that includes approximately 60% of simulated scenarios, i.e. 5,188 out of 8,996 scenarios. Scenarios are selected so that the sample includes all the classes or coherent scenarios (420 classes for voltage angle and 489 classes for voltage magnitude). For each scenario in the matrix, an assignment vector is determined. After that, a Random Forest classifier is trained using data from the classification matrix with the aim of predicting the cluster associated to the pre-contingency system loading condition of a new and unknown scenario. Test scenarios used to assess the classifier performance are obtained from data that have not been used before in the classification matrix. Almost 90% of test data are correctly predicted considering voltage magnitude and angle simultaneously. The average computing time for each scenario is 12 ms. As can be seen in Figure 5.13, misclassified scenarios are usually scenarios with a low frequency of class in the training sample. The classifier fails in these cases because the class imbalance, which, despite having been partially solved by the three-step filtering process explained above, is still present.

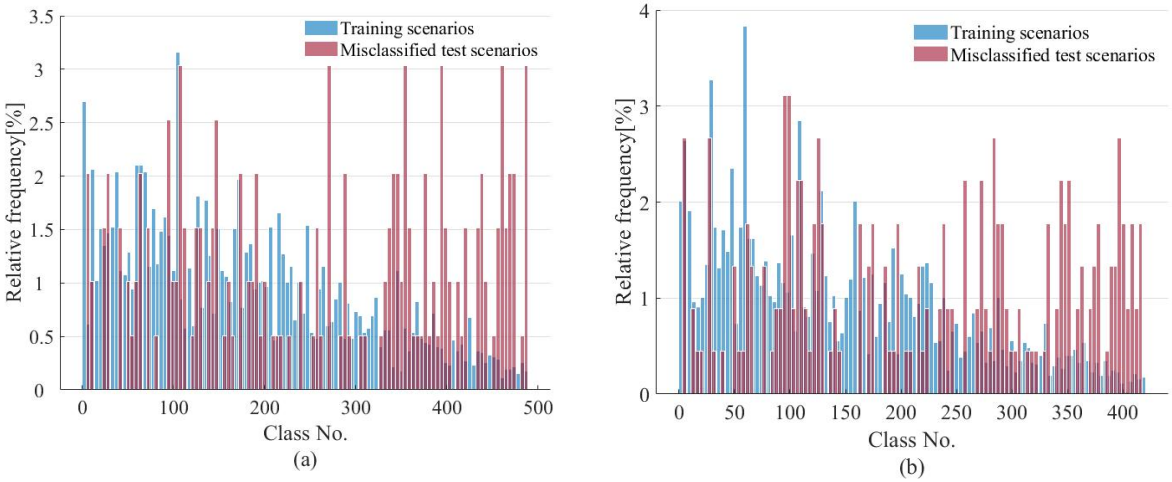

Fig. 5.13 Data distribution. Training scenarios vs test scenarios incorrectly predicted (a) voltage magnitude, (b) voltage angle

The mean absolute deviation is used as a measure to evaluate the accuracy of voltage pseudo-measurements:

$$Err_{Vps}=\frac{\sum_{i=1}^{i=N}\left|vps_i-v_i\right|}{N} \tag{5.1}$$

where Err_{Vps} is the voltage pseudo-measurement error (magnitude or angle), N is the number of time periods under analysis, vps_i and v_i are the voltage pseudo-measurement and the real voltage (magnitude or angle) at time i.

Figures 5.14, 5.15 and 5.16 depict with the help of box-plots the distribution of the voltage pseudo-measurement error at each non-observed bus and group the results together according to whether the coherent scenario is correctly or incorrectly predicted. From the point of view of the kind of contingency, pseudo-measurements are less accurate in those events that result in a system response with a higher degree of nonlinearity, such three-phase faults or generator outages. In this case the error remains in most scenarios below 1.5 % and 1º for voltage magnitude and angle respectively. When analyzing scenarios according to the classifier performance, in most cases pseudo-measurement errors from misclassified scenarios show a higher degree of dispersion reaching higher maximum values, which indicates they are less accurate. Nevertheless, the difference of accuracy between both kind of scenarios is minor.

As an example, Figure 5.17 shows the bus voltage pseudo-measurement at the non-observed bus 35 considering the three kind of contingencies under analysis, i.e. three-phase fault, generator outage and load curtailment. Bus voltage pseudo-measurements approach the real value.

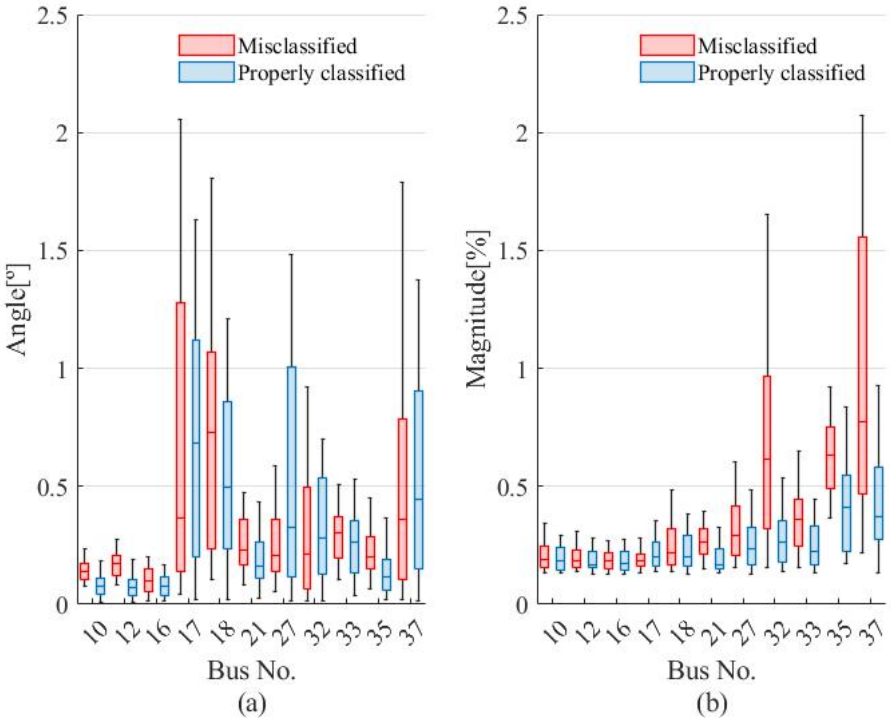

Fig. 5.14 Distribution of the voltage pseudo-measurement error at non-observed buses when a generator outage takes place (a) voltage angle, (b) voltage magnitude

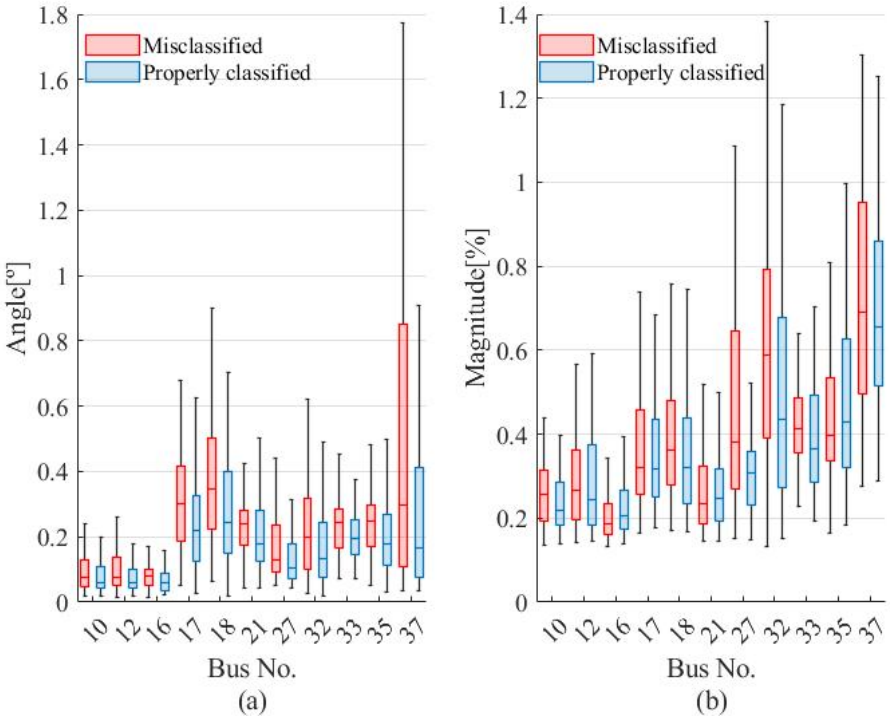

Fig. 5.15 Distribution of the voltage pseudo-measurement error at non-observed buses when a three-phase fault takes place (a) voltage angle, (b) voltage magnitude

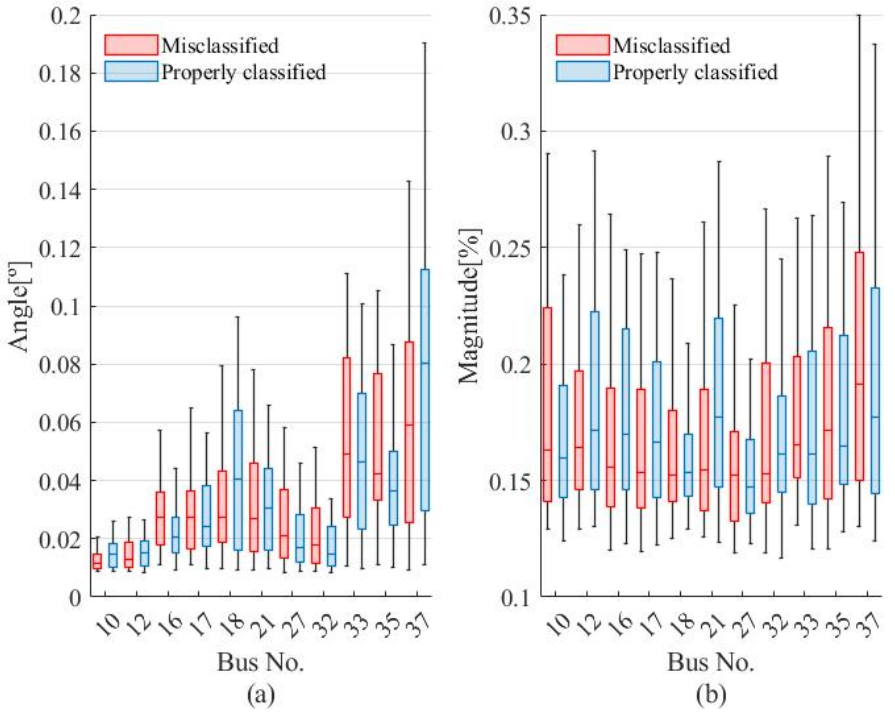

Fig. 5.16 Distribution of the voltage pseudo-measurement error at non-observed buses when a load curtailment takes place (a) voltage angle, (b) voltage magnitude

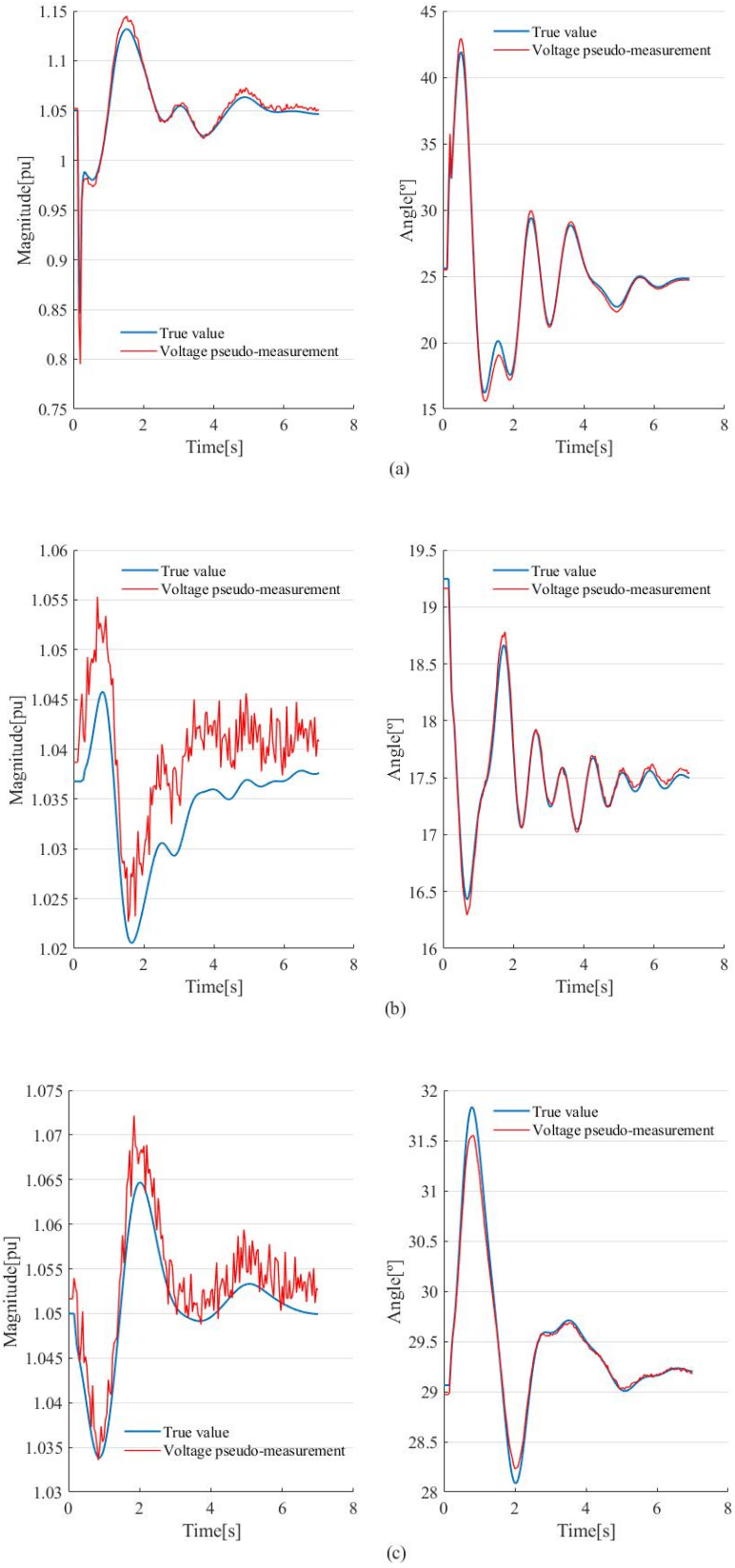

Fig. 5.17 Pseudo-measurement vs true voltage at non-observed bus 35 (a) three-phase fault on line 25-26, (b) generator outage at bus 38 (G9), (c) load curtailment at bus 24 (load step of -37.3%)

5.4 Analysis of performance of the proposed hybrid state estimator

In this section the performance of the hybrid state estimator is investigated by means of four case studies. The main goal is to characterize the algorithm capacity with respect to its accuracy, processing time and robustness. The estimator is tested under different operating conditions regarding the classifier response, the redundancy of bus voltage pseudo-measurements, the algorithm employed in phase two for estimating the static state and the grid parameters and PMU measurement errors.

Static estimator in phase one initializes assuming flat start condition. The initial dynamic state vector is set from the estimated static state in phase one. In addition, speed of machines will likely be close to its nominal value when the power system is in stationary regime and under normal operating conditions. The initial covariance matrix for generators and motors is set to be $\mathbf{P}_0$=diag([5 5 5 5]) and P_0=5 respectively.

As mentioned before, in this work measurement as well as structure (line parameters) uncertainties are taken into consideration. Furthermore, process noises from model inadequacy and linearization are also included. The process and measurement covariance matrices for generators are defined as $\boldsymbol{Q}=0.06^2\times(\boldsymbol{I}_{4x4})$ and $\boldsymbol{R}_{obs}=0.06^2\times(\boldsymbol{I}_{4x4})$ or $\boldsymbol{R}_{unobs}=0.6^2\times(\boldsymbol{I}_{4x4})$ respectively depending on whether the bus associated with the machine is observable by a PMU or not. Process and measurement variances for motors are set to be $w_k=0.9^2$ and $v_k=0.06^2$ respectively. Uncertainties of measurable inputs in generators (field voltage and mechanical power) are modelled in the same way as conventional measurements.

The mean absolute deviation is employed as a measure of error to evaluate the estimation accuracy, see equation (5.3.1).

Because in phase two estimators operate in sequence, computing times introduce a delay of the order of a few tens or milliseconds. In order to take this aspect into consideration, the dynamic estimator in phase two uses as input the estimated static state at the previous PMU time step, i.e. the dynamic estimator uses estimated data from 40 ms ago.

Case I (base case)

The first case is also defined as base case because the algorithm operates following the assumptions made above as regards noises, algorithms involved in the estimation process and number of bus voltage pseudo-measurements.

The estimator accuracy is statically characterized and classified according to whether the coherency of the simulated scenario has been properly or misclassified. Figures 5.18 to 5.23 show the distribution of the estimated static state error according to the kind of contingency and the kind of bus from a PMU observability point of view. In addition, Figures 5.24 to 5.27 depict the error distribution of the estimated dynamic state for generators at buses 33, 34 and 37 and motors at buses 4 and 20. Because bus 34 is observed by PMUs the degree of dispersion of the estimation error at this bus is remarkably smaller. As expected, misclassified scenarios tend to present a higher degree of dispersion of the estimation error reaching higher values. Nevertheless, the difference between both cases is rather small and what is more, in some cases the error distribution for misclassified scenarios is slightly less dispersed, as it is the case with scenarios that entail three-phase fault at buses 33, 34 and 37 (see Figures 5.22 and 5.25). This is due, first, to the fact that the algorithm for voltage coherency definition is prone to error when dealing with scenarios whose coherent areas exhibit almost the same bus voltage behavior but still remain as individual areas, and second, to an uneven number of scenarios with strong nonlinearities in each of the test samples under analysis, which contributes to higher estimation errors and eventually to more dispersed error distributions.

Considering only properly classified scenarios and according to the kind of contingency and the kind of bus, as expected, the estimator is less accurate in scenarios that entail stronger nonlinearities and at non-observed buses. Considering only PMU-non observed buses, the median of the estimated static state error is in the range of 0.01 to 0.7º and 0.2 to 0.7% for voltage magnitude and angle respectively. In relation to the dynamic state, the median of the error goes from 0.5 to 3.8º for rotor angle of generators and from 0.006 to 0.07% for rotation speed of generators. Regarding motors, the median of the error is about 0.4% for rotation speed.

Convergence is analyzed by measuring the time evolution of the estimation error. The dynamic estimator based on the EKF fails to find a solution in about 11% of the test scenarios. This amount is reduced to 4% if the time delay between static and dynamic estimators in phase two is removed. In addition, it was found that most cases with convergence problems imply unstable scenarios.

The average computing time is determined in order to check whether the proposal can keep up with PMUs' update rate of 25 samples per second. Tests are performed in a computer with an Intel Core i7 2.3 GHz processor and 16 GB RAM. Results in table 5.3 prove that the execution time in phase two is lower than the PMU sampling period, enabling the proposed scheme to track fast dynamics. Although a centralized approach like the one proposed in this thesis might

not work in large scale power grids due to the higher computing time involved, it can be extended to a decentralized scheme typically employed for monitoring and controlling of large scale systems, where the system is divided into smaller areas and local estimators run in each of them. In addition, the problem can be adapted to be solved by means of parallel computing methods. Successful applications in this field can be found in [29, 53].

Figures 5.28 and 5.29 depict the time series of the estimated static and dynamic state at buses 37 and 24 respectively. The first scenario involves a three-phase fault on line 3-18. The error of the estimated static state in such case is close to the median (0.5% and 0.3º for voltage magnitude and angle respectively). The second scenario is a generator outage at bus 39. As in the case before, the error of the estimated static state is close to the median (0.17% and 0.04º for voltage magnitude and angle respectively). In both scenarios the bus voltage coherency and the assignment vector that dictates the relation between PMU observed and non-observed buses have been correctly predicted.

Table 5.3 Average computational time (ms)

WLS	WLAV	EKF
330	15	0.2

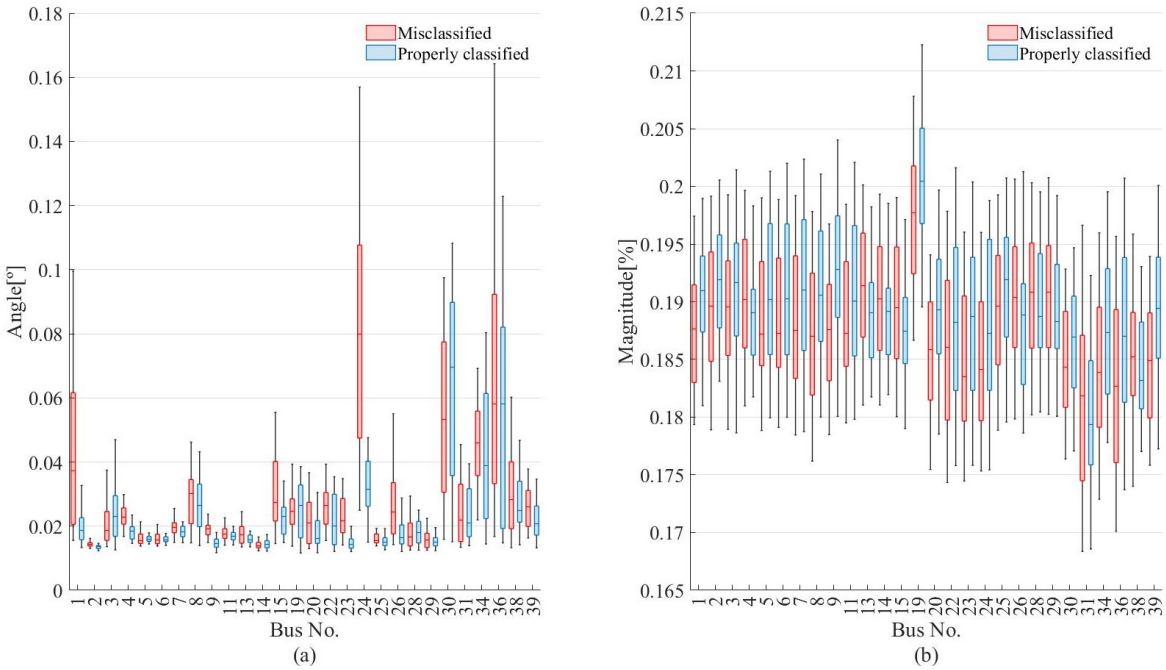

Fig. 5.18 Distribution of the estimated static state error at observed buses when a generator outage takes place (a) voltage angle, (b) voltage magnitude

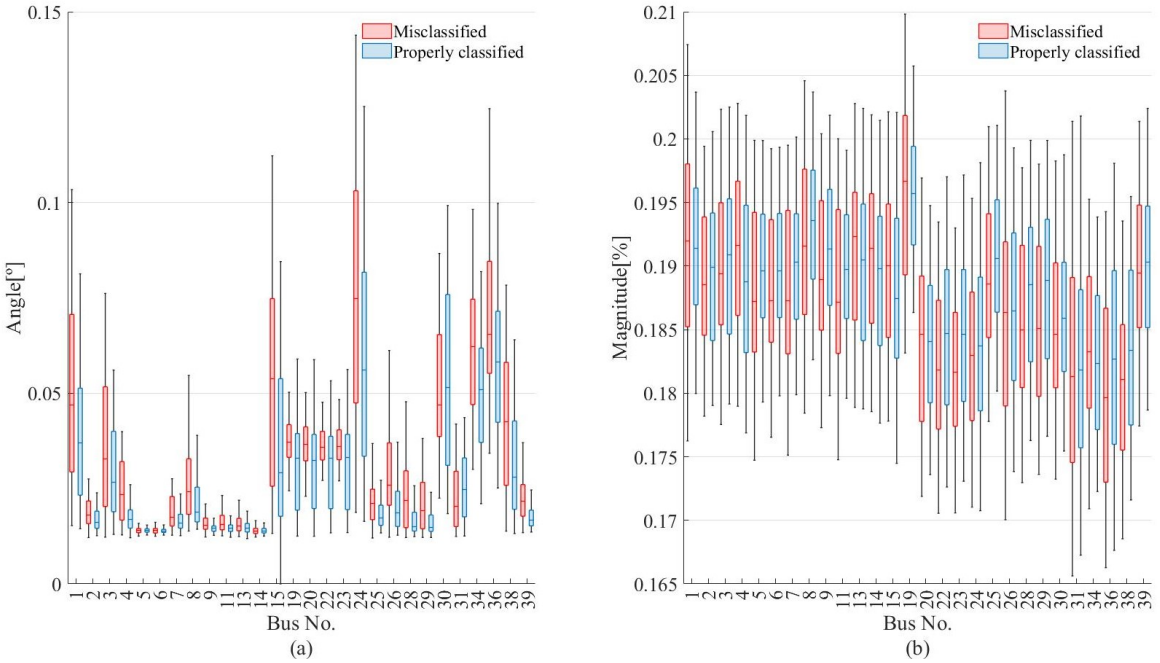

Fig. 5.19 Distribution of the estimated static state error at observed buses when a three-phase fault takes place (a) voltage angle, (b) voltage magnitude

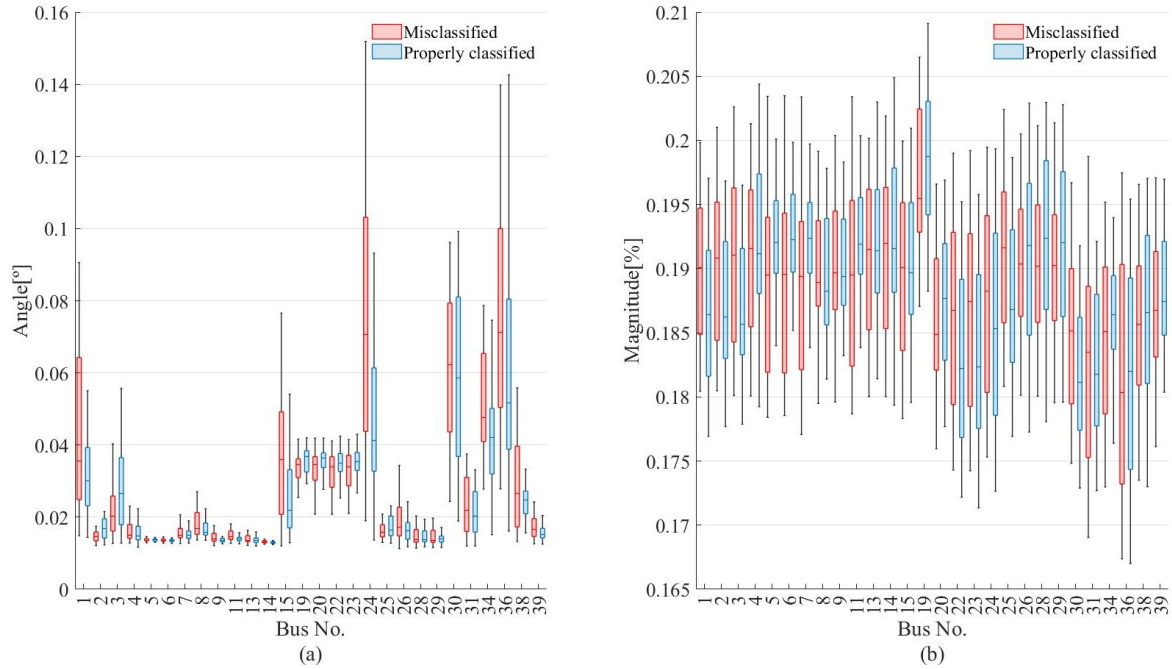

Fig. 5.20 Distribution of the estimated static state error at observed buses when a load curtailment takes place (a) voltage angle, (b) voltage magnitude

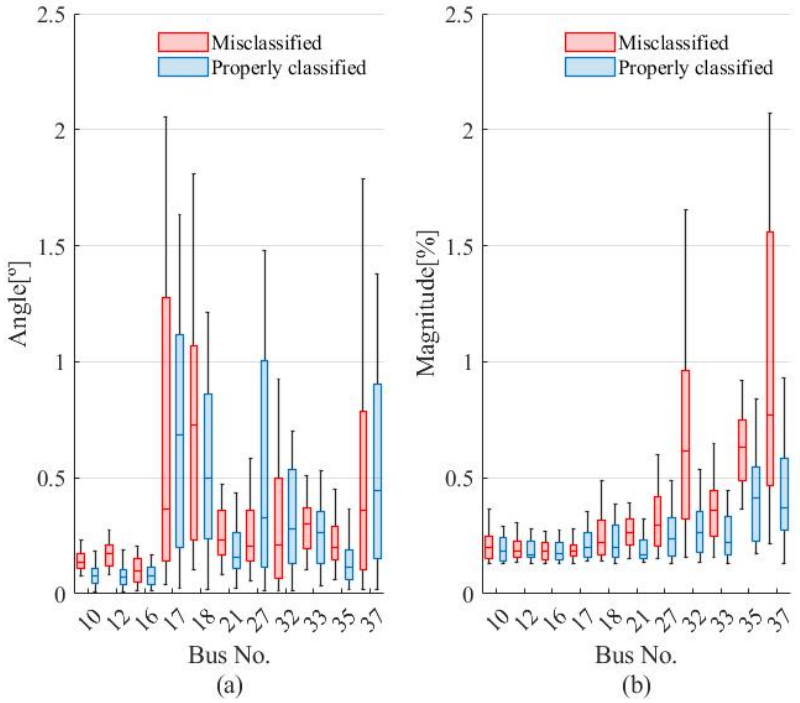

Fig. 5.21 Distribution of the estimated static state error at non-observed buses when a generator outage takes place (a) voltage angle, (b) voltage magnitude

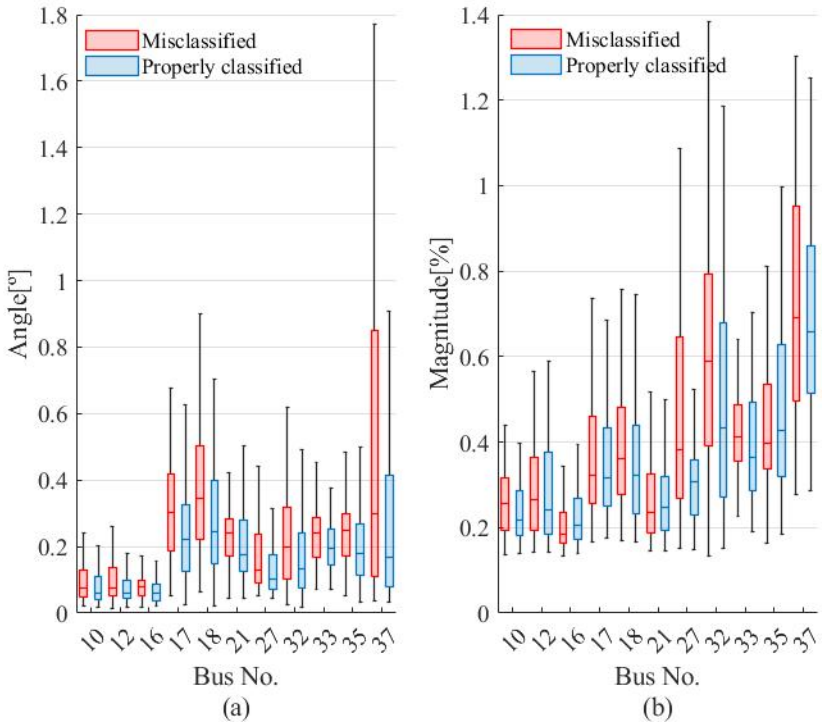

Fig. 5.22 Distribution of the estimated static state error at non-observed buses when a three-phase fault takes place (a) voltage angle, (b) voltage magnitude

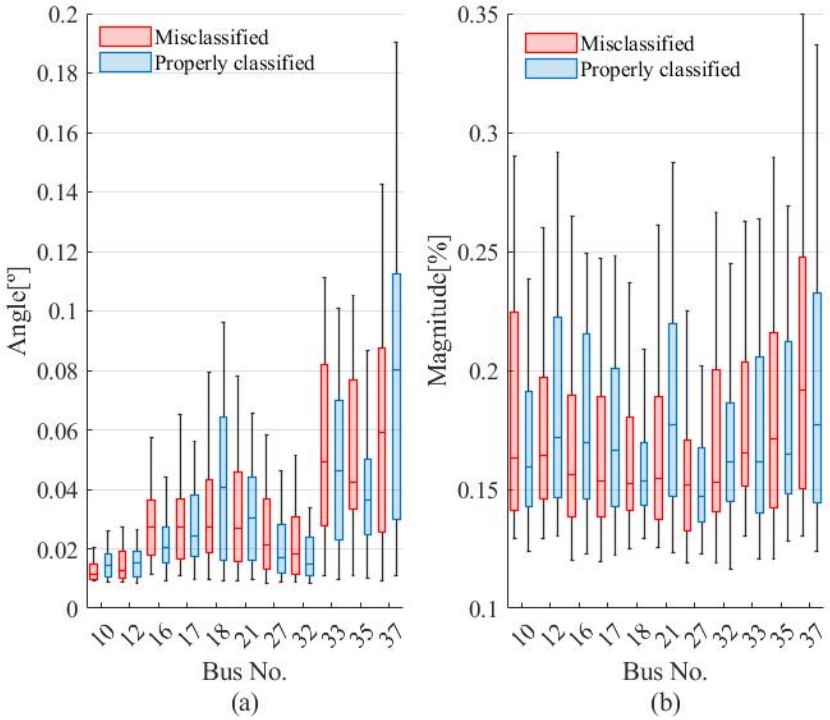

Fig. 5.23 Distribution of the estimated static state error at non-observed buses when a load curtailment takes place (a) voltage angle, (b) voltage magnitude

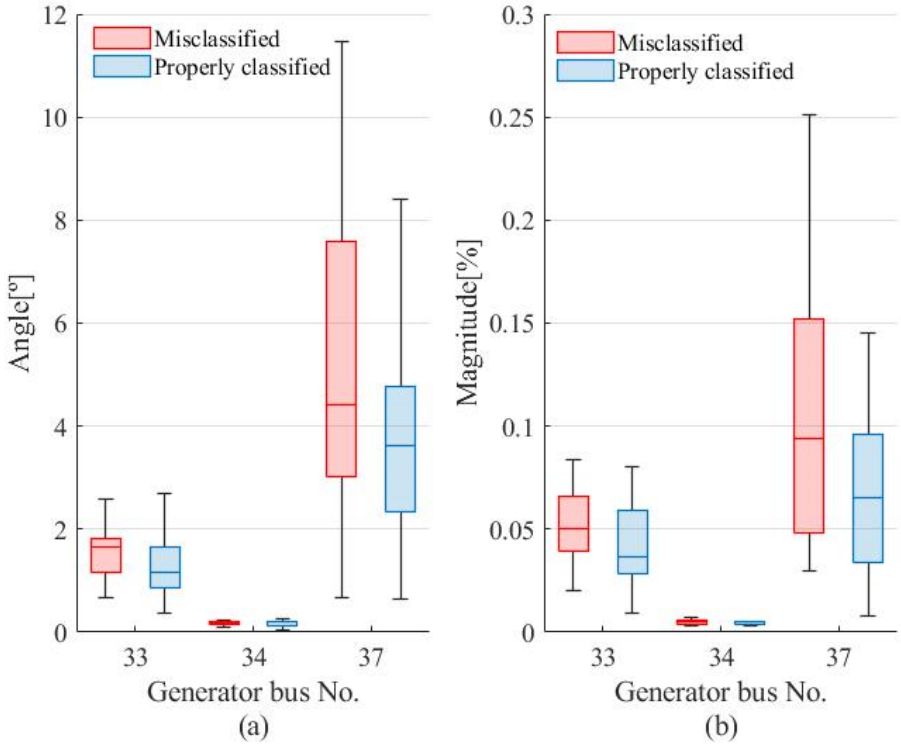

Fig. 5.24 Distribution of the estimated dynamic state (generators) error when a generator outage takes place (a) rotor angle of generator, (b) rotation speed of generator

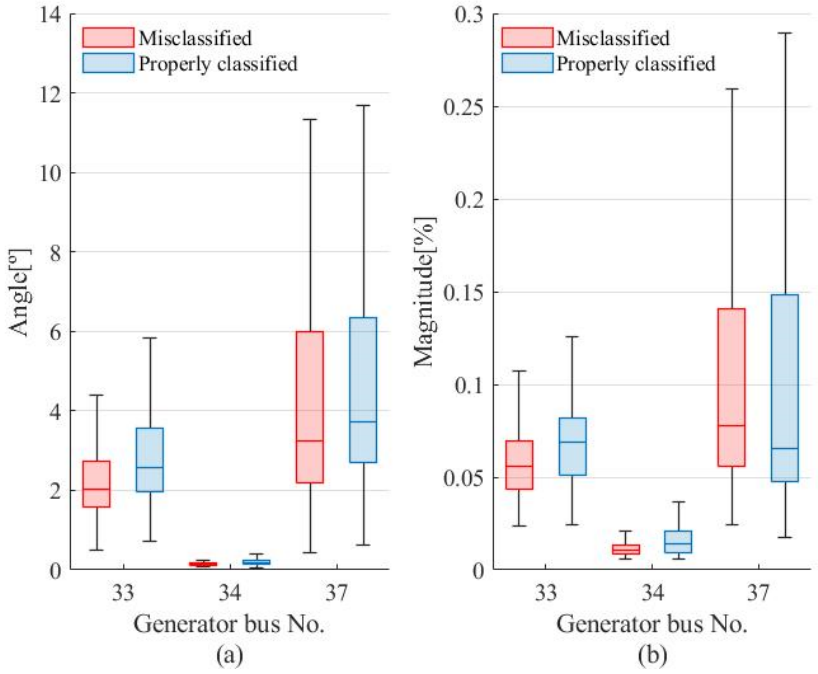

Fig. 5.25 Distribution of the estimated dynamic state (generators) error when a three-phase fault takes place (a) rotor angle of generator, (b) rotation speed of generator

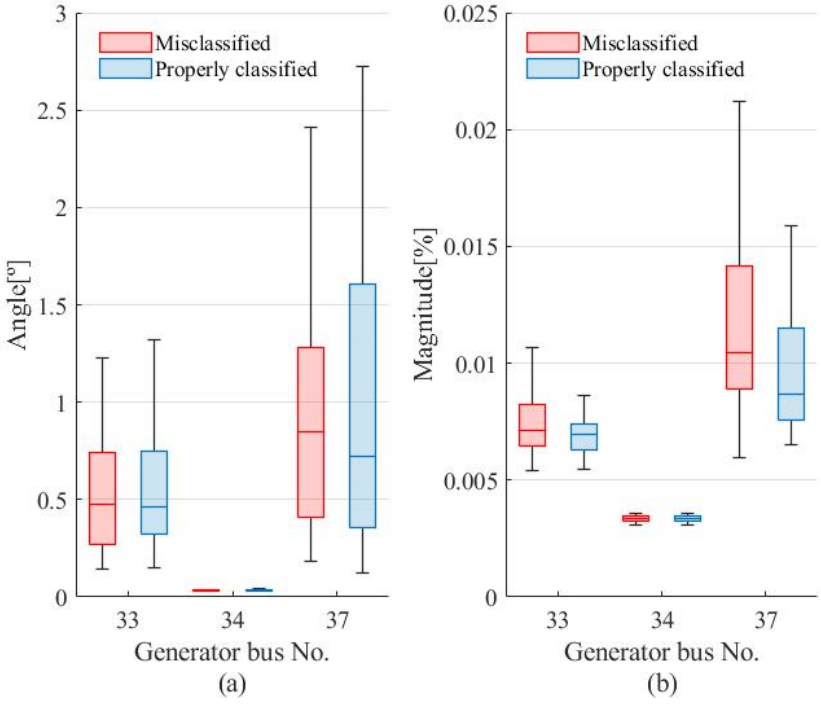

Fig. 5.26 Distribution of the estimated dynamic state (generators) error when a load curtailment takes place (a) rotor angle of generator, (b) rotation speed of generator

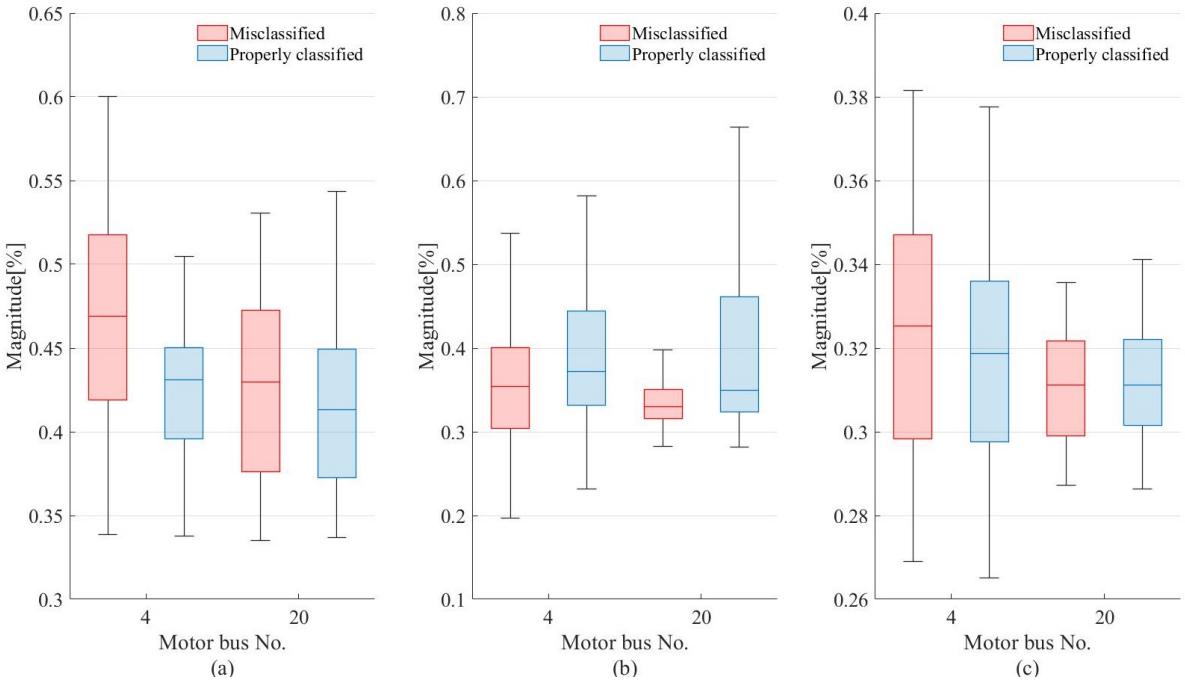

Fig. 5.27 Distribution of the estimated dynamic state error (rotation speed of motor) (a) generator outage, (b) three-phase fault, (c) load curtailment

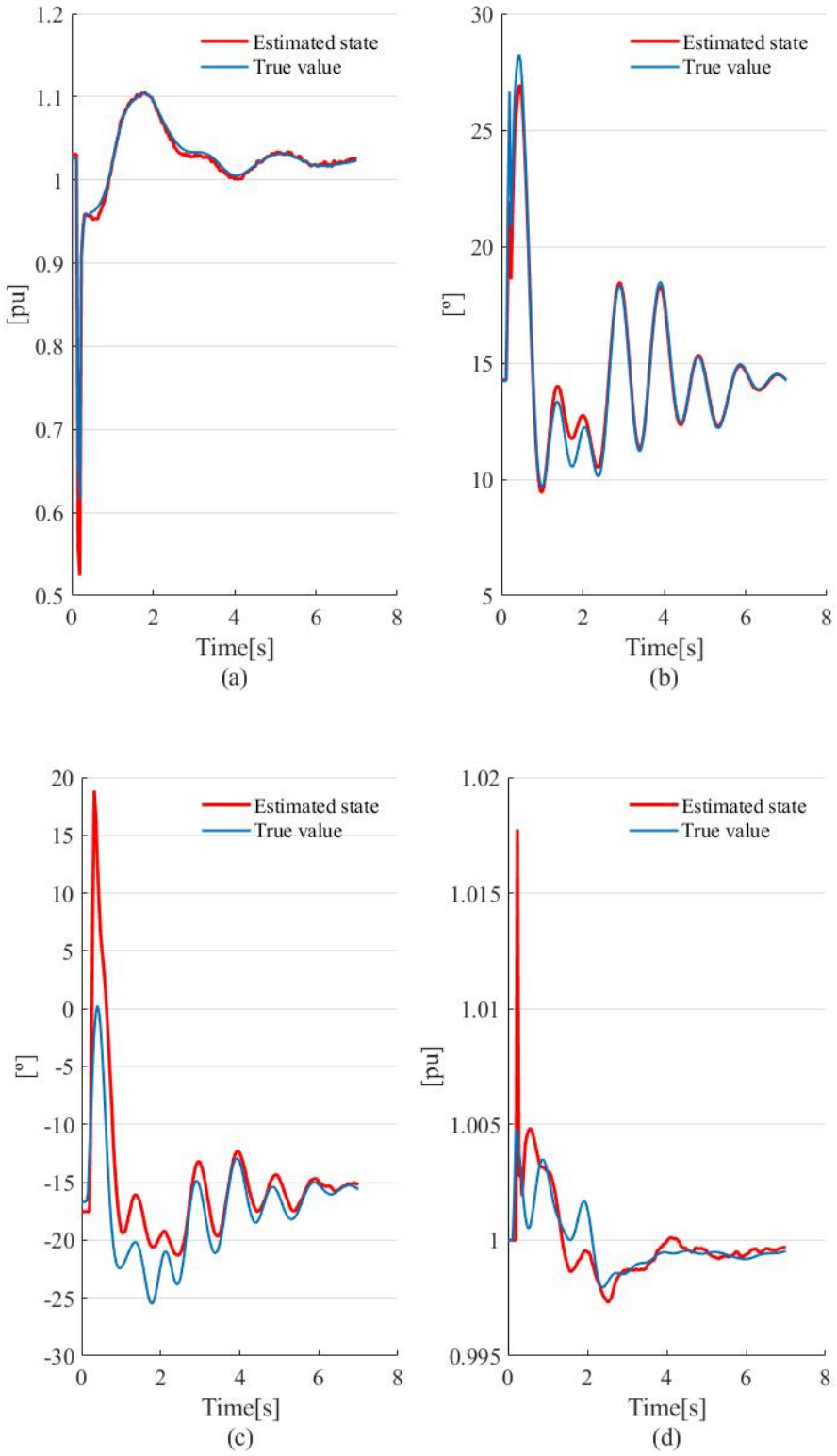

Fig. 5.28 Estimated static and dynamic state at bus 37 under a three-phase fault on line 3-18 (a) voltage magnitude, (b) voltage angle, (c) rotor angle of generator, (d) rotation speed of generator

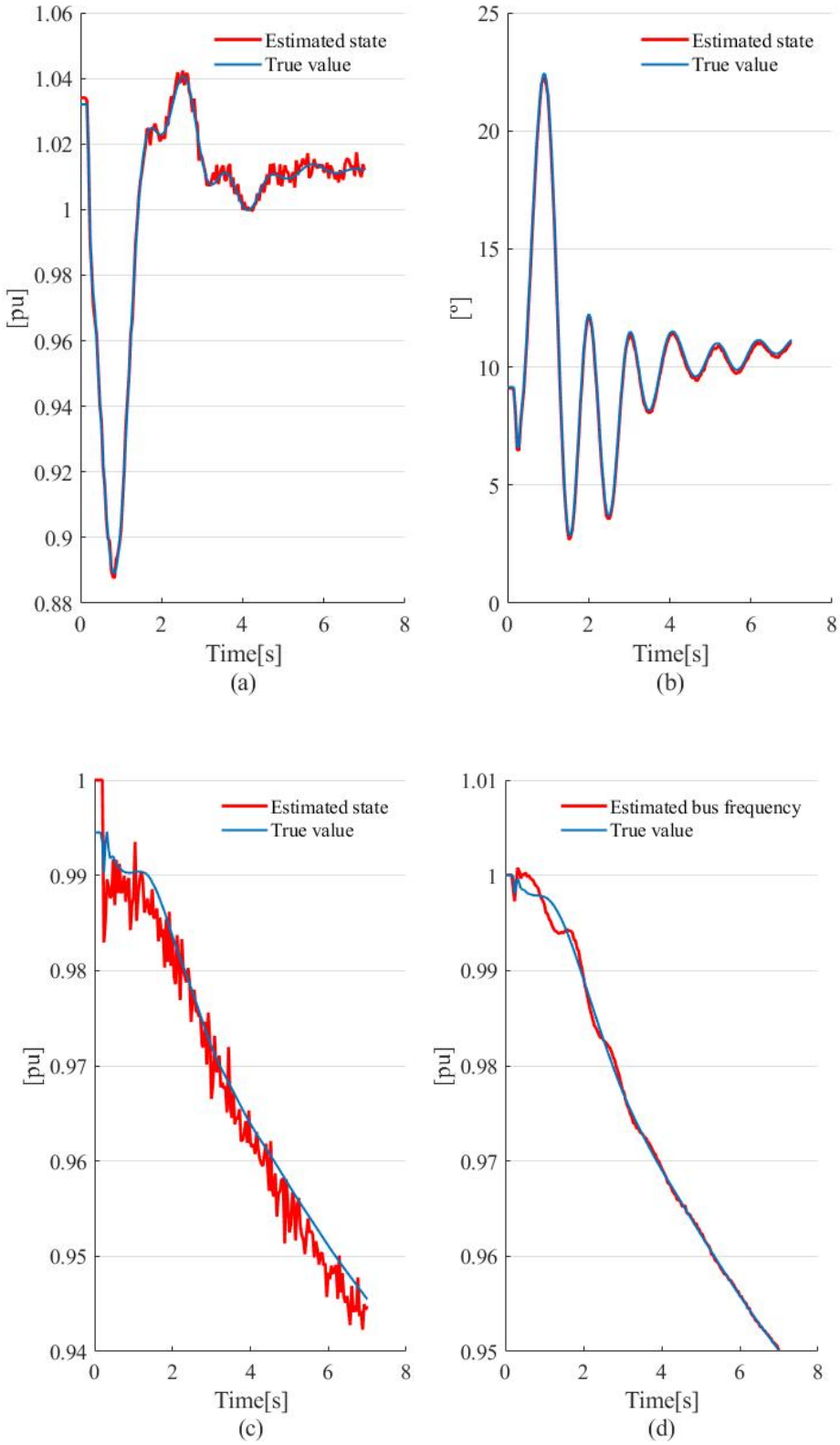

Fig. 5.29 Estimated static and dynamic state at bus 24 under a generator outage at bus 39 (a) voltage magnitude, (b) voltage angle, (c) rotation speed of motor, (d) frequency of bus voltage

Results presented above prove that the proposed scheme is able to inform the system operating conditions with reasonable accuracy. Nevertheless, the performance of the estimator is severely degraded in the presence of gross PMU errors. The main reason behind this issue lies in the lack of a post-estimation bad data processor. Such a scheme becomes unfeasible at high scan rates. For this reason, estimators in phase two use the information provided by the bad data processing step in phase one to adapt the measurement weights in phase two. Nevertheless, and as it will be shown below, this does not solve large error issues when there is no redundancy among measurements. The developed methodology for generating bus voltage pseudo-measurements assigns only one voltage pseudo-measurement (magnitude and angle) to each non-observed bus.

With the aim of enhancing the measurement redundancy in phase two and thus the algorithm filtering capacity and its robustness against large errors, a modified version of the already proposed estimator is proposed. The alternative employs an ancillary set of bus voltage pseudo-measurements, which is obtained using the same coherency concept explained before, but this time without considering the criterion of minimum electrical distance employed when assigning observed to non-observed buses. Because of that, the additional set of pseudo-measurements is considered even more inaccurate with a standard deviation 10% greater than its counterpart used in the initial approach.

Case II

The second case attempts to prove the advantages that an additional set of bus voltage pseudo-measurements at the observed buses gives to the performance of the proposed hybrid state estimation approach in the presence of large PMU errors. A large error is simulated by adding a constant signal to the voltage magnitude of the PMU measurement at bus 20 on t = 0 s with a value equal to 30 % of the bus nominal value. Both alternatives are analyzed and compared with each other regarding their accuracy.

Figure 5.30 to 5.36 shows the estimated static and dynamic state at buses 33, 34 and 20 when the generator at bus 37 goes out. Here can be noticed that, although the inclusion of additional pseudo-measurements in some cases may result in a slight loss of performance, as is the case with the voltage angle at bus 20 (see Figure 5.35b), the estimator gains the ability to filter out the large error. This is visible in the estimated voltage magnitudes at buses 34 and 20 and power injections at buses 33 and 34 (see Figures 5.33-a, 5.35-a, 5.31-a-b and 5.33-c-d). In the presence of a PMU failure, the filtering capacity results in a remarkable improvement of the algorithm

accuracy not only when estimating the static but also the dynamic state (see Figures 5.32-a-b, 5.34-a-b and 5.36).

Figures 5.37 to 5.42 depict distribution of the estimated static state error for observed and non-observed buses. Results are grouped together according to the kind of contingency. As expected, the large error affects to a larger extent buses that are near the failed PMU. This can be easily observed in the distribution of the estimated static state error at buses 19, 20, 33 and 34. In addition, it can be seen that the ancillary set of pseudo-measurements results in a loss of performance that becomes more evident in voltage angles, but it gives the algorithm the ability to filter the large error. This leads eventually to better results as regards the dynamic state, as is the case with generators at buses 33 and 34 and the motor at bus 20 (as can be seen in boxplots depicted in Figures 5.43 to 5.45).

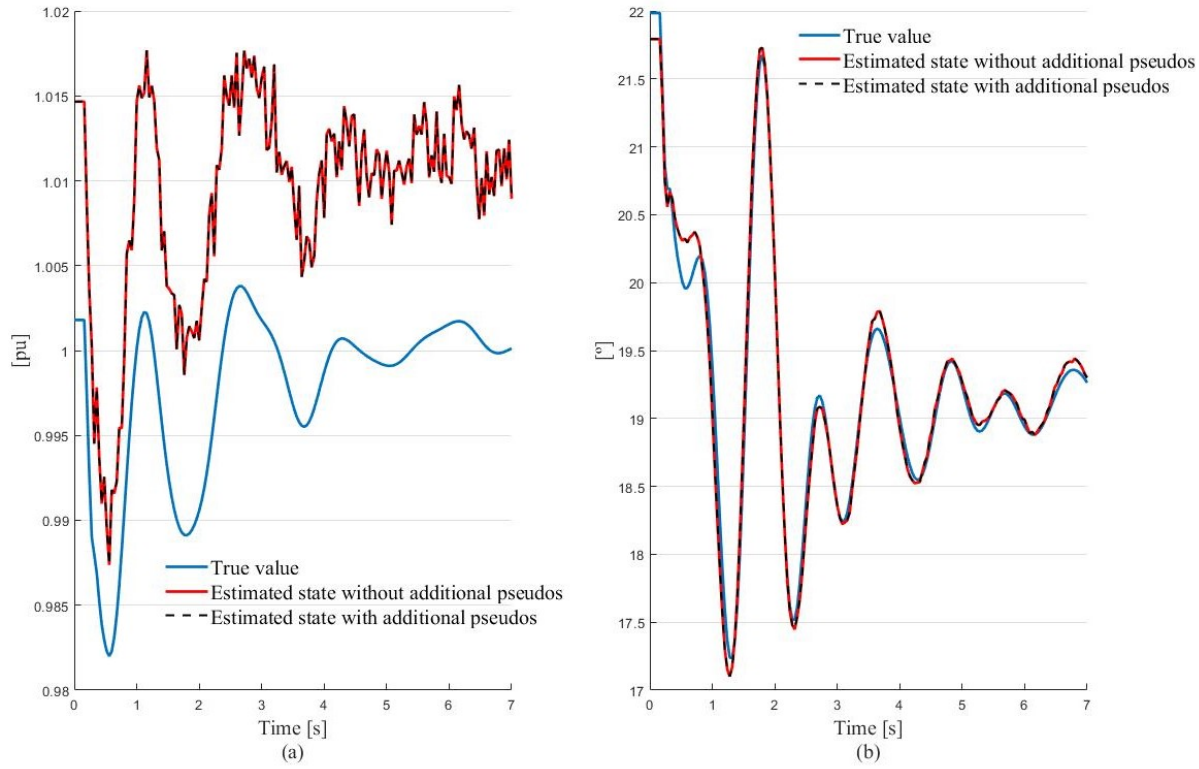

Fig. 5.30 Estimated static state at bus 33 under a generator outage at bus 37 (a) voltage magnitude, (b) voltage angle

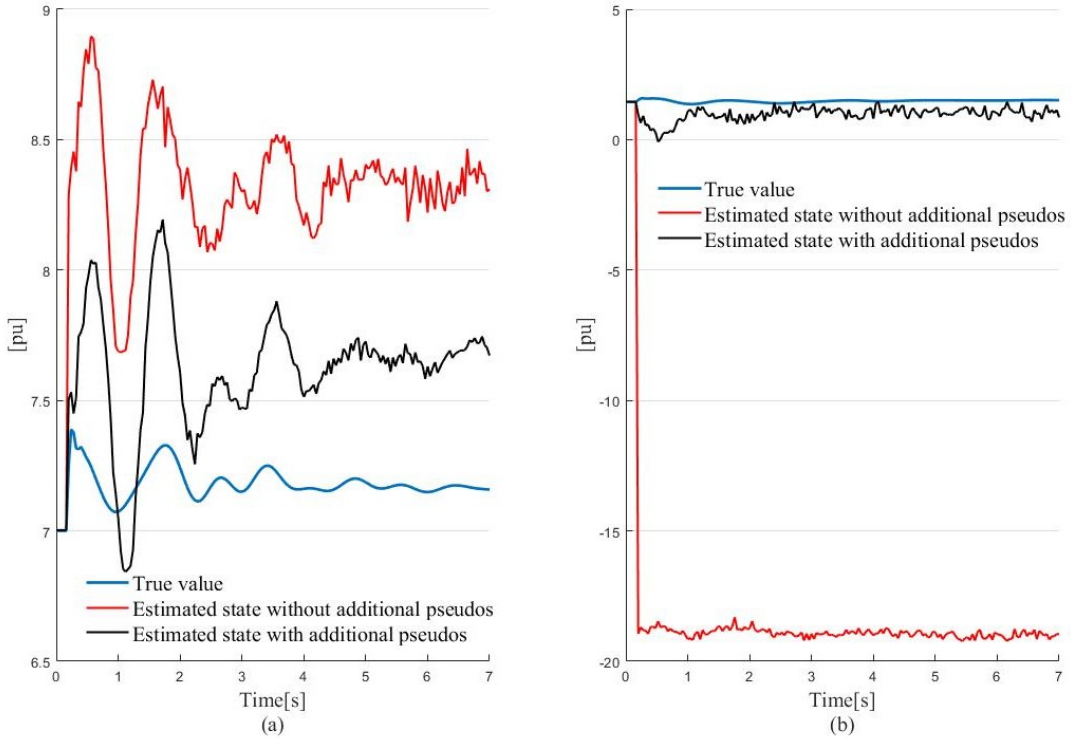

Fig. 5.31 Estimated static state at bus 33 under a generator outage at bus 37 (a) active power injection, (b) reactive power injection

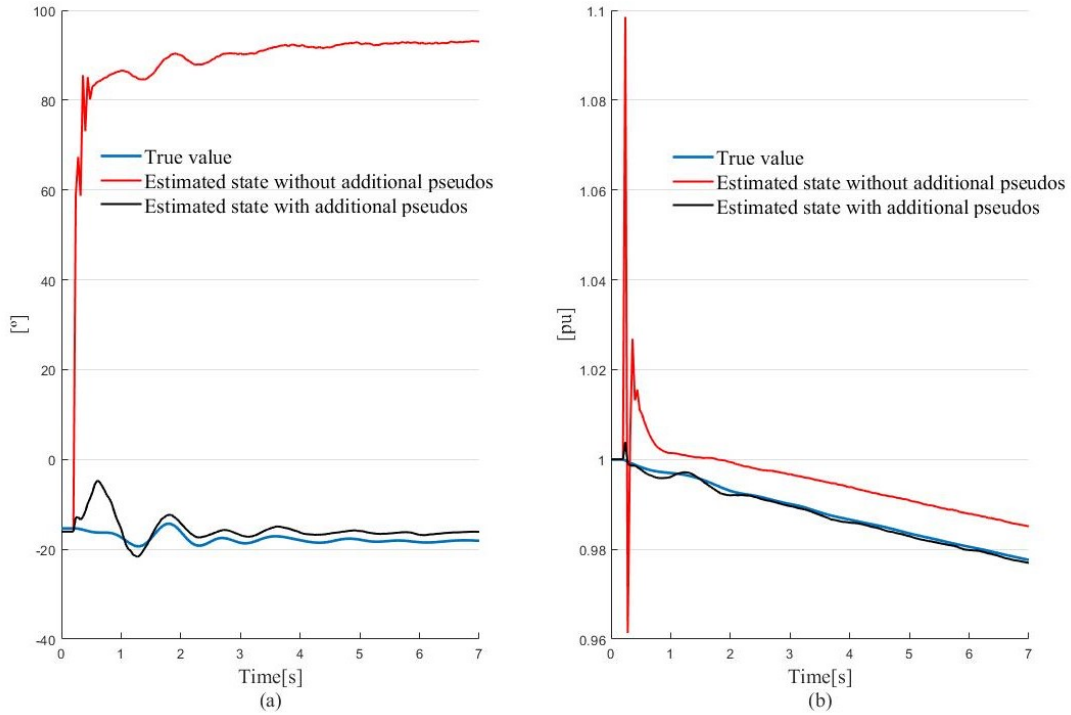

Fig. 5.32 Estimated dynamic state at bus 33 under a generator outage at bus 37 (a) rotor angle of generator, (b) rotation speed of generator

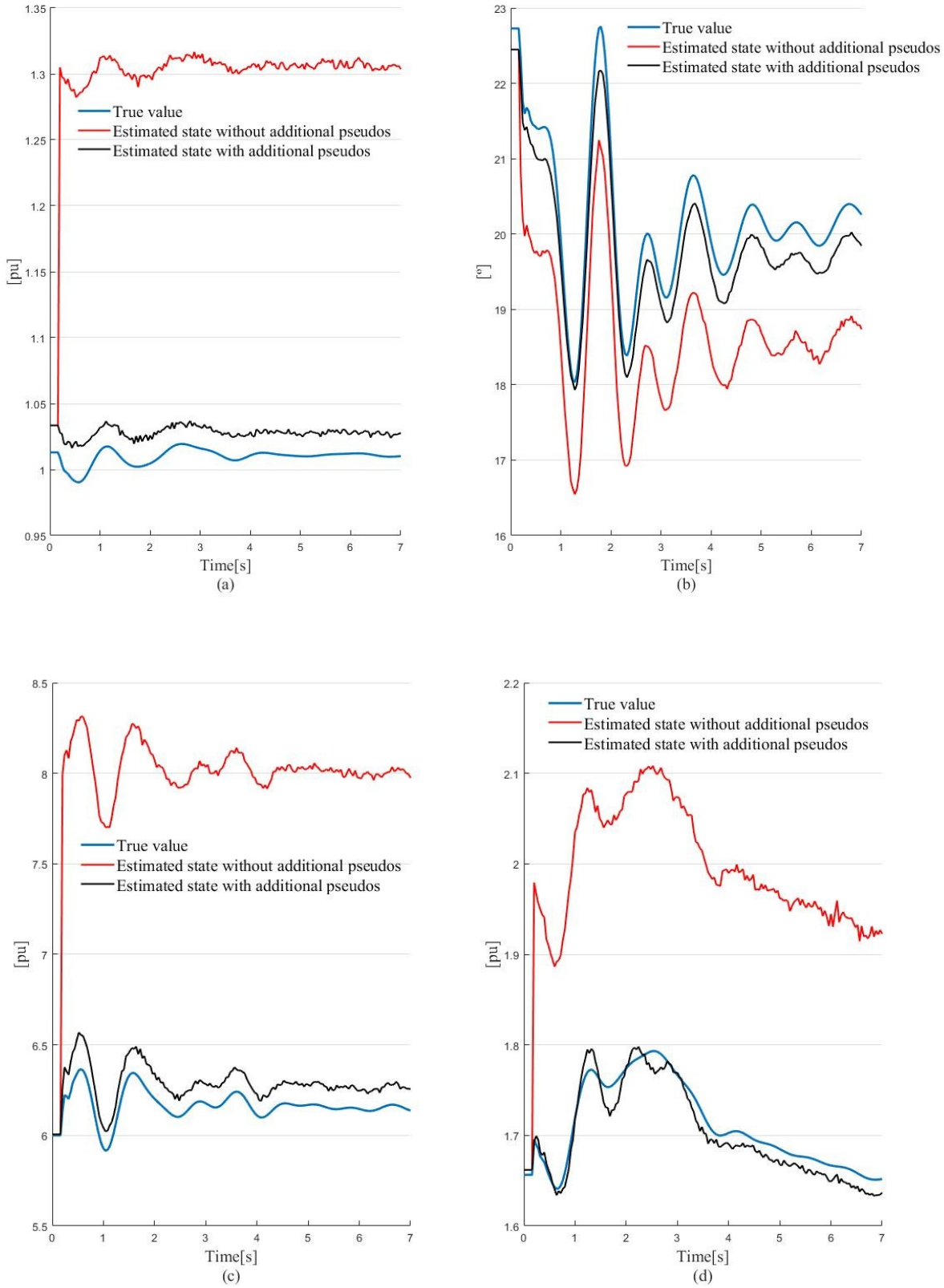

Fig. 5.33 Estimated static state at bus 34 under a generator outage at bus 37 (a) voltage magnitude, (b) voltage angle, (c) active power injection, (d) reactive power injection

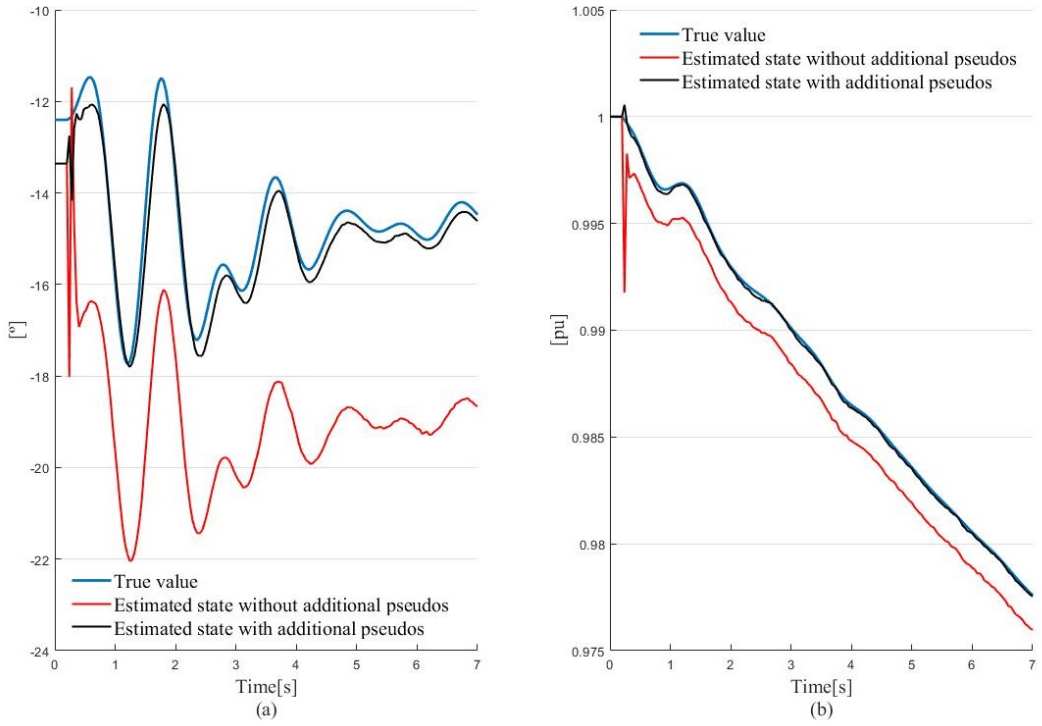

Fig. 5.34 Estimated dynamic state at bus 34 under a generator outage at bus 37 (a) rotor angle of generator, (b) rotation speed of generator

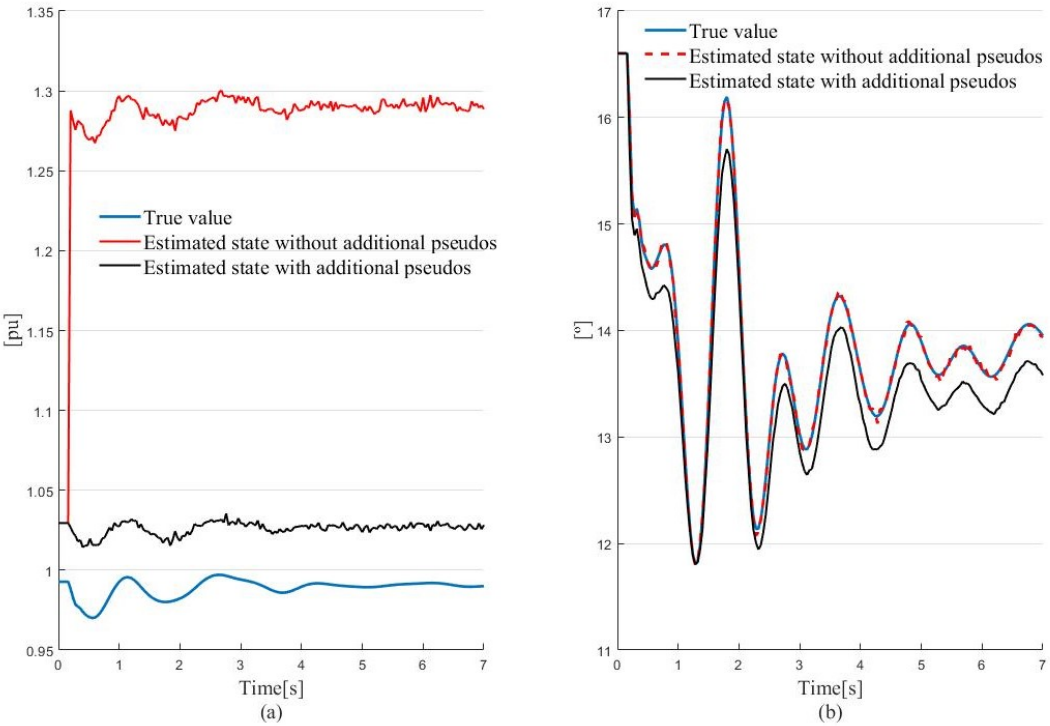

Fig. 5.35 Estimated static state at bus 20 under a generator outage at bus 37 (a) voltage magnitude, (b) voltage angle

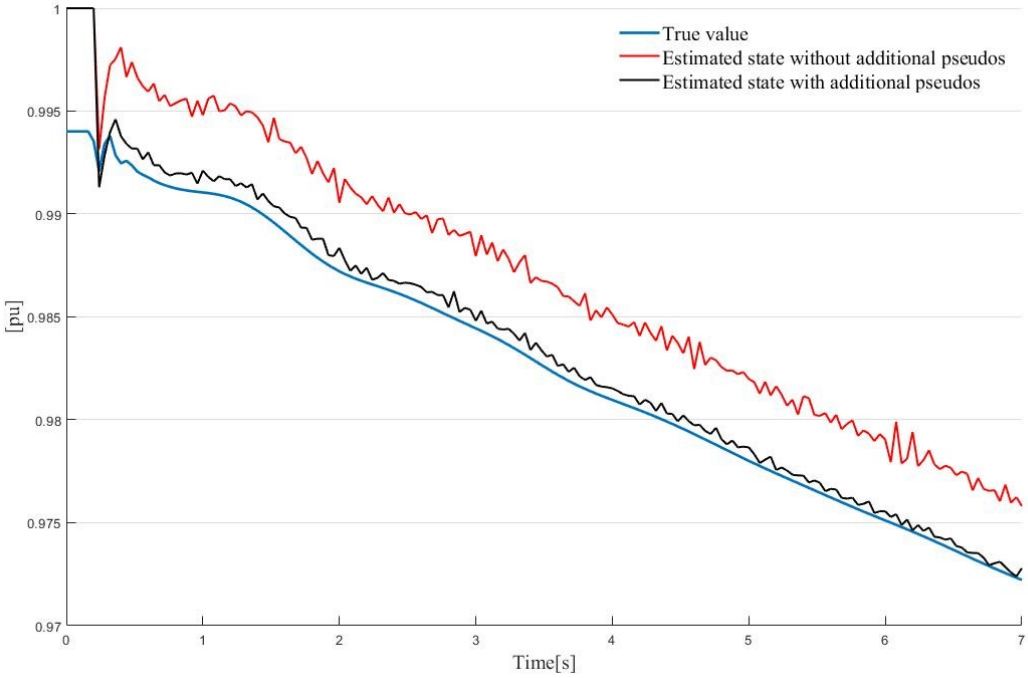

Fig. 5.36 Estimated dynamic state (rotation speed of motor) at bus 20 under a generator outage at bus 37

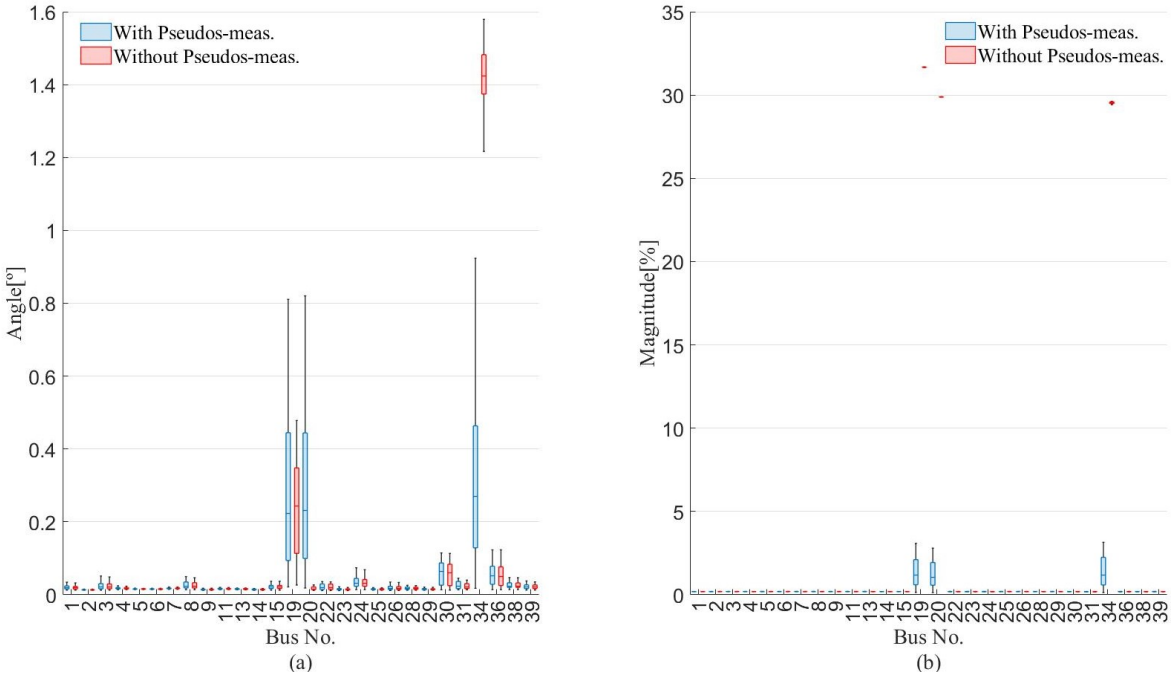

Fig. 5.37 Distribution of the estimated static state error at observed buses when a generator outage takes place (a) voltage angle, (b) voltage magnitude

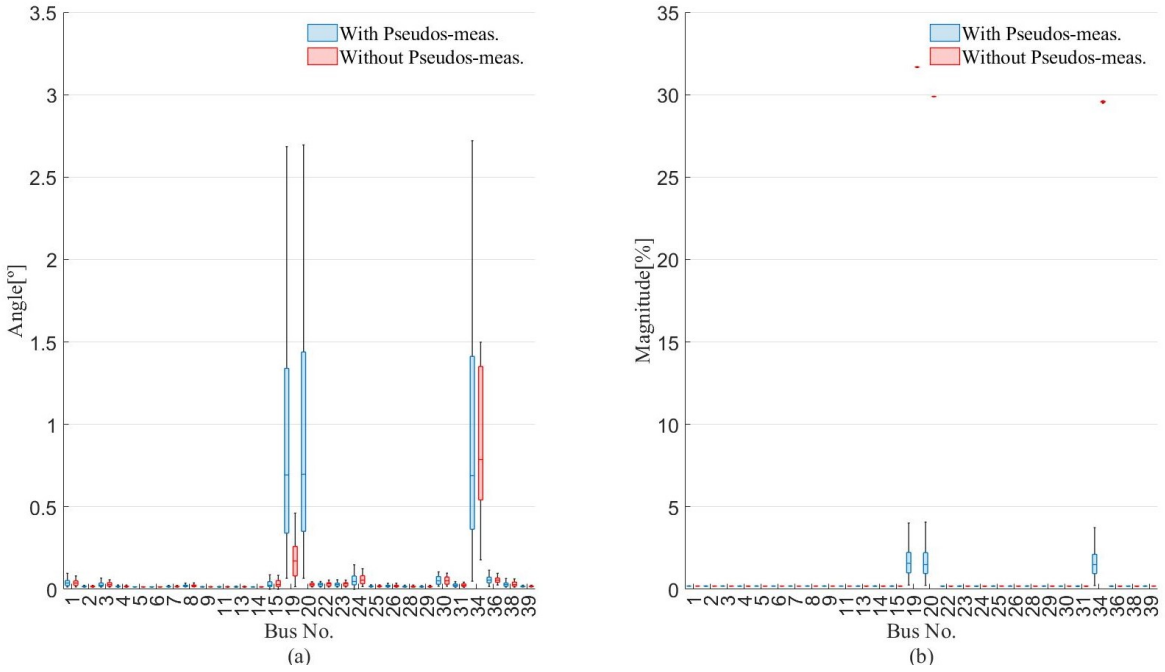

Fig. 5.38 Distribution of the estimated static state error at observed buses when a three-phase fault takes place (a) voltage angle, (b) voltage magnitude

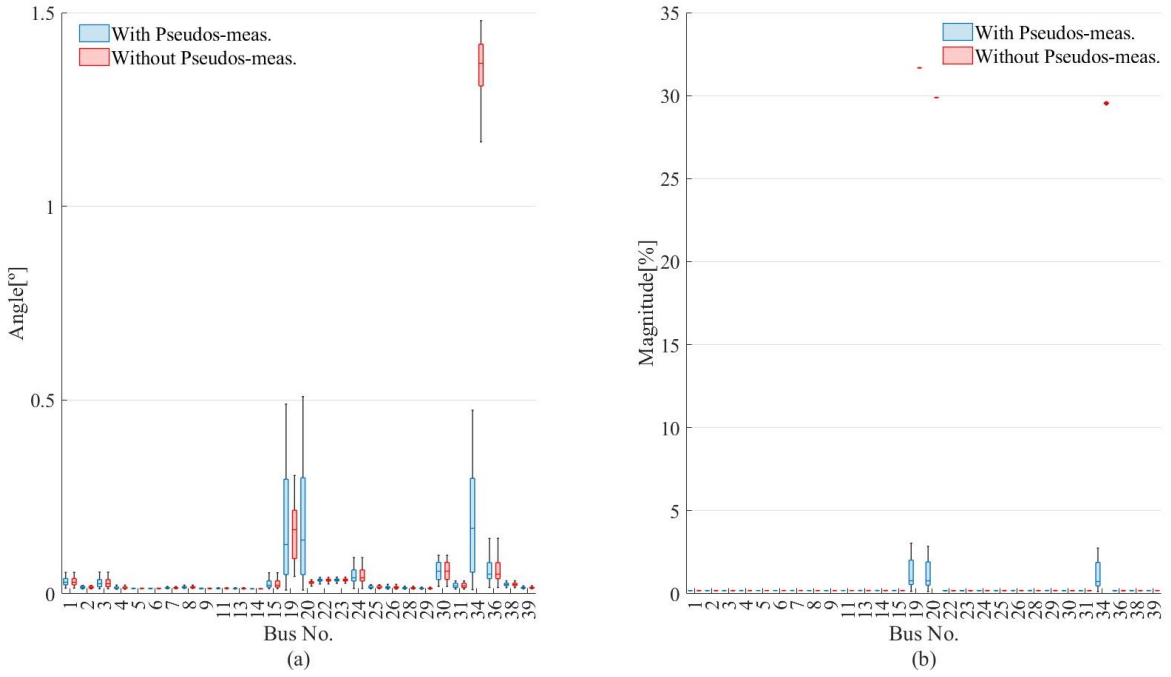

Fig. 5.39 Distribution of the estimated static state error at observed buses when a load curtailment takes place (a) voltage angle, (b) voltage magnitude

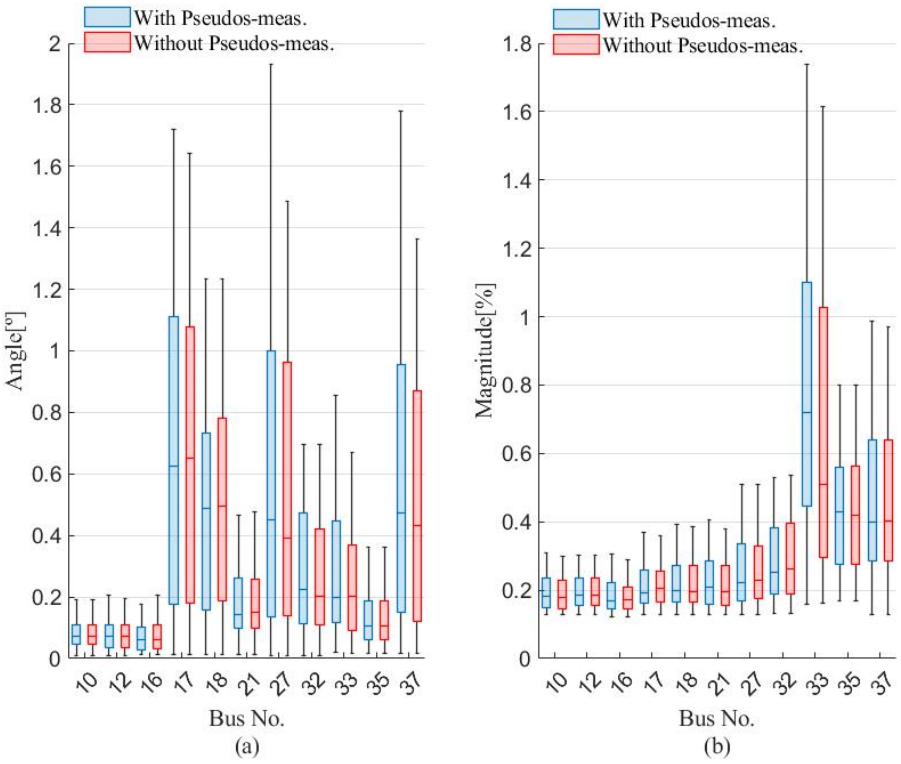

Fig. 5.40 Distribution of the estimated static state error at non-observed buses when a generator outage takes place (a) voltage angle, (b) voltage magnitude

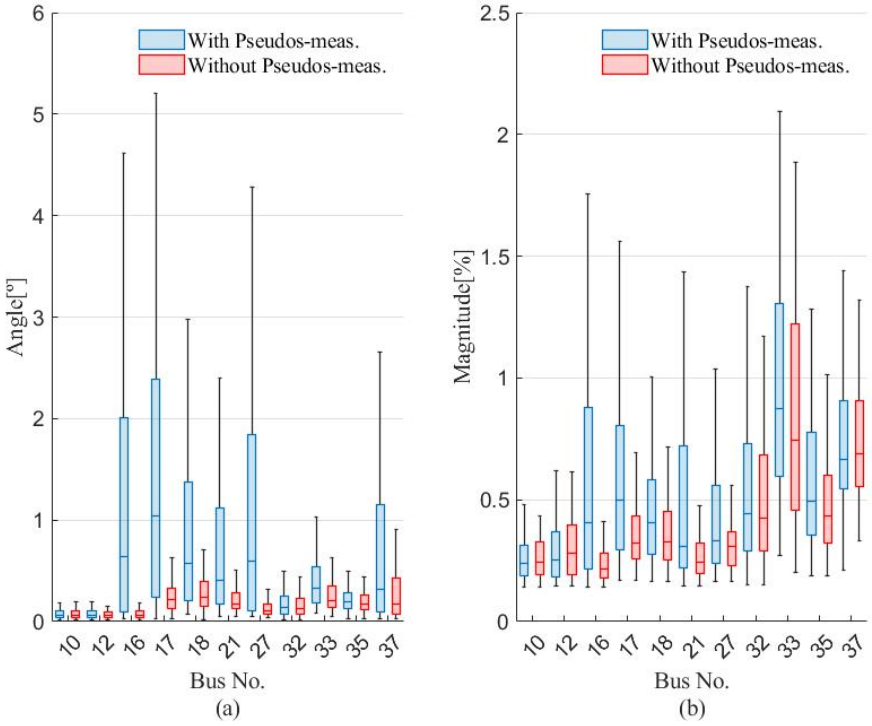

Fig. 5.41 Distribution of the estimated static state error at non-observed buses when a three-phase fault takes place (a) voltage angle, (b) voltage magnitude

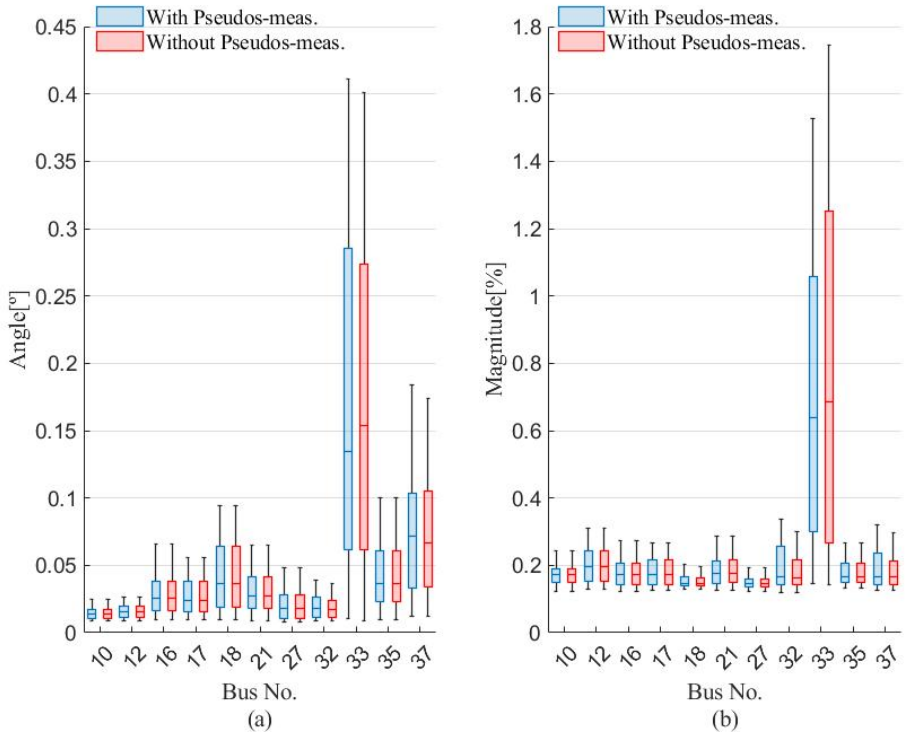

Fig. 5.42 Distribution of the estimated static state error at non-observed buses when a load curtailment takes place (a) voltage angle, (b) voltage magnitude

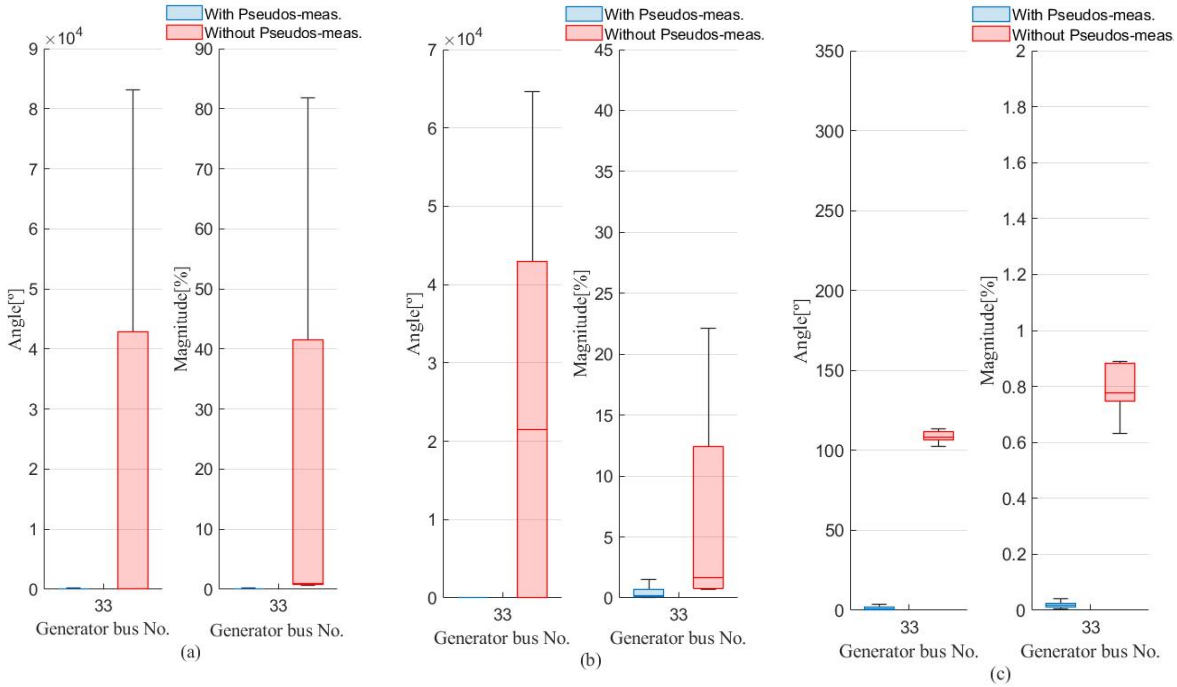

Fig. 5.43 Distribution of the estimated rotor angle and rotation speed error (generator at bus 33) according to the kind of contingency (a) generator outage, (b) three-phase fault, (c) load curtailment

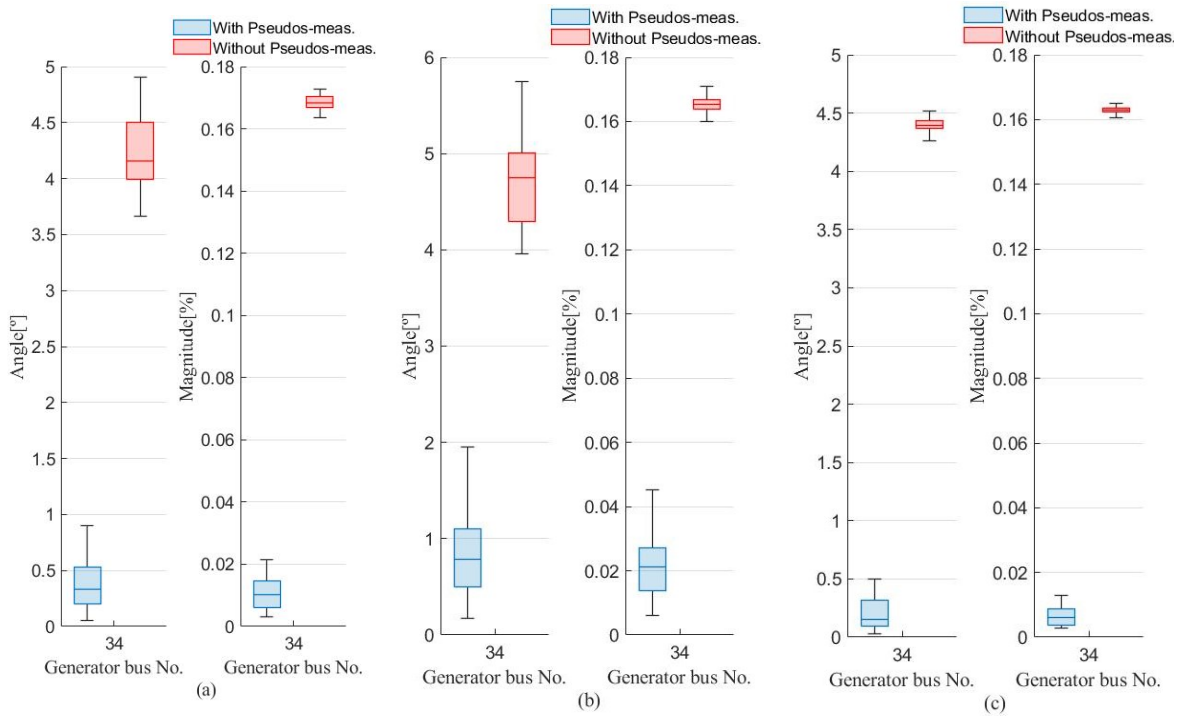

Fig. 5.44 Distribution of the estimated rotor angle and rotation speed error (generator at bus 34) according to the kind of contingency (a) generator outage, (b) three-phase fault, (c) load curtailment

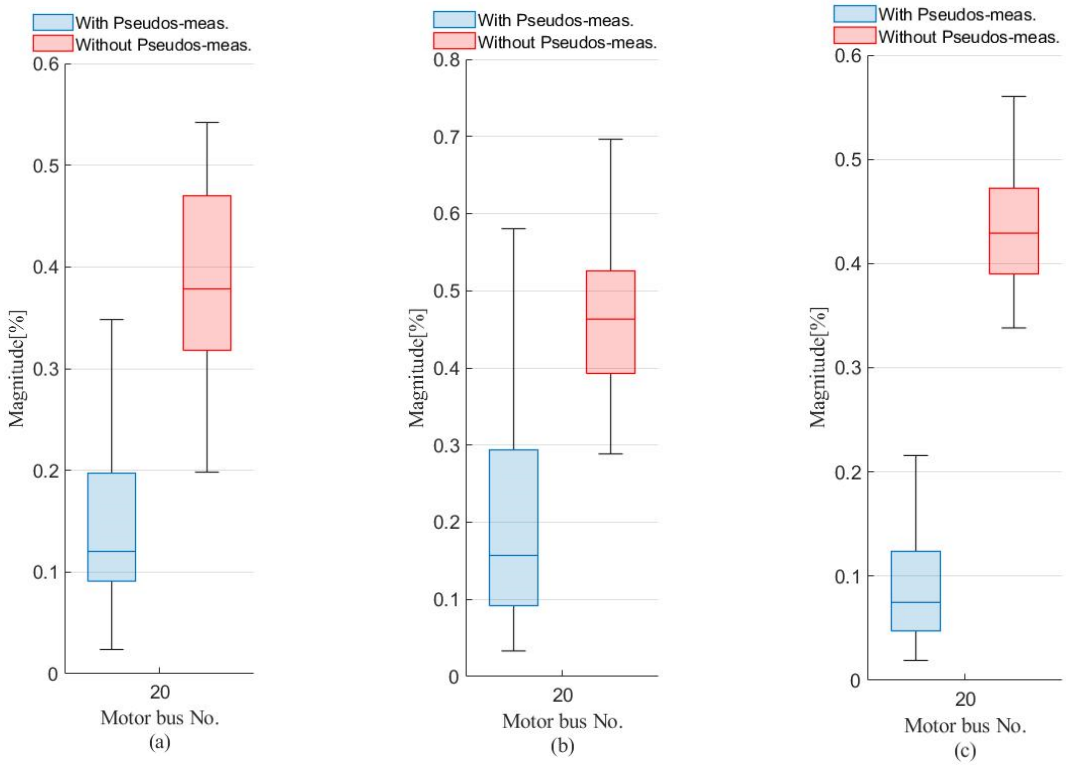

Fig. 5.45 Distribution of the estimated rotation speed error (motor at bus 20) according to the kind of contingency (a) generator outage, (b) three-phase fault, (c) load curtailment

Case III

As mentioned before, despite the WLS method is computationally efficient, it does not provide any robustness against large errors. A post-estimation bad data processing step becomes mandatory for such a technique. The main drawback of this method when applied in phase two is its high computational cost that makes it unfeasible at high scan rates. That is why, a robust WLAV based linear estimator is proposed instead the conventional WLS to estimate grid bus voltages in phase two.

This case aims to demonstrate the improvement in the algorithm performance that comes from such alternative. Both options, i.e. the WLS based scheme and its more robust alternative based on the WLAV, are evaluated under different operating scenarios. In the first place, no PMU gross error is considered. Then, as in the previous case, a malfunction of the PMU at bus 20 is assumed and simulated by adding a constant signal equal to 30% of the reported value on t = 0 s.

Figures 5.46 to 5.51 depict the distribution of the estimated bus voltage error at observed and non-observed buses and according to the kind of contingency. As can be noticed, at some buses the estimation error is slightly more dispersed and reaches higher values in the WLAV variant than in its counterpart based on the WLS technique. This difference is more pronounced in voltage magnitudes. Thus, the WLS based option is a little more accurate when PMUs do their job without any inconvenience. Figures 5.52 to 5.55 show the estimated dynamic states at buses 4, 32, 33, 34 and 20. There is no appreciable difference in performance between both options when dealing with dynamic states. This situation is reversed when having a PMU malfunction. Under such a circumstance, the alternative based on the WLAV is more effective than its counterpart to filter out the large error. This can be observed in the distribution of the estimated static state error at buses that are located in the immediate vicinity of the bus with the failed PMU, as happens with buses 19, 20, 33 and 34 (see boxplot in Figures 5.56 to 5.61). As expected, this difference in performance between both options is also reflected in the estimated dynamic states (see Figures 5.62 and 5.63).

As an example, Figures 5.64 to 5.70 depict the estimated static and dynamic state at buses 33, 34 and 20 when almost 60% of the load at bus 24 is suddenly removed from the grid. Results from both options are plotted together for comparative purposes. As can be seen, the WLAV based variant outperforms the WLS based option when assuming a large error in the reported PMU measurement at bus 20.

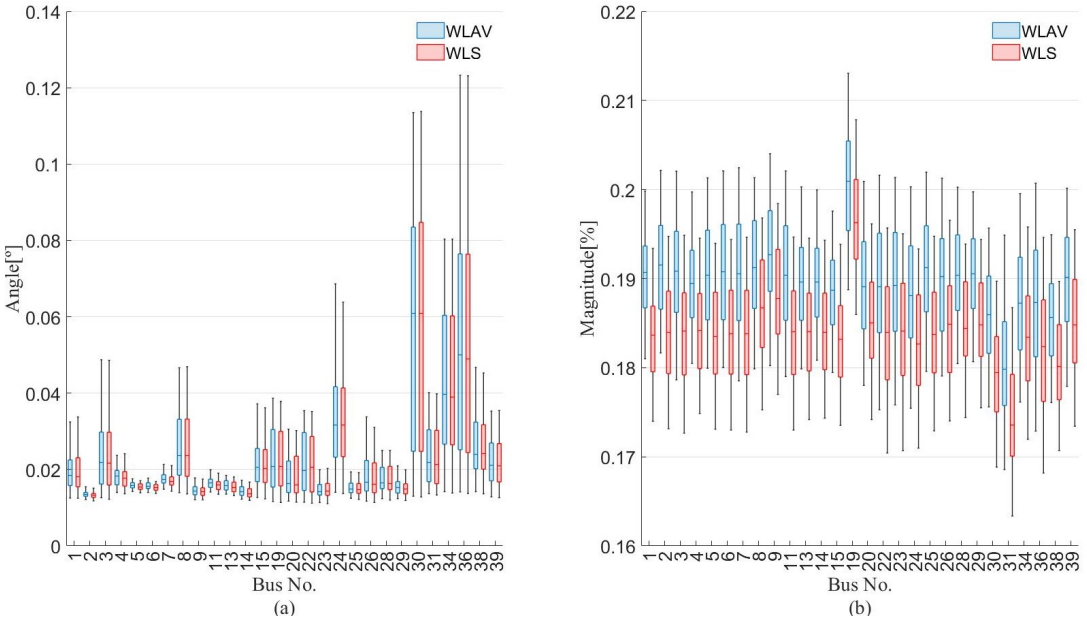

Fig. 5.46 Distribution of the estimated static state error at observed buses when a generator outage takes place (a) voltage angle, (b) voltage magnitude

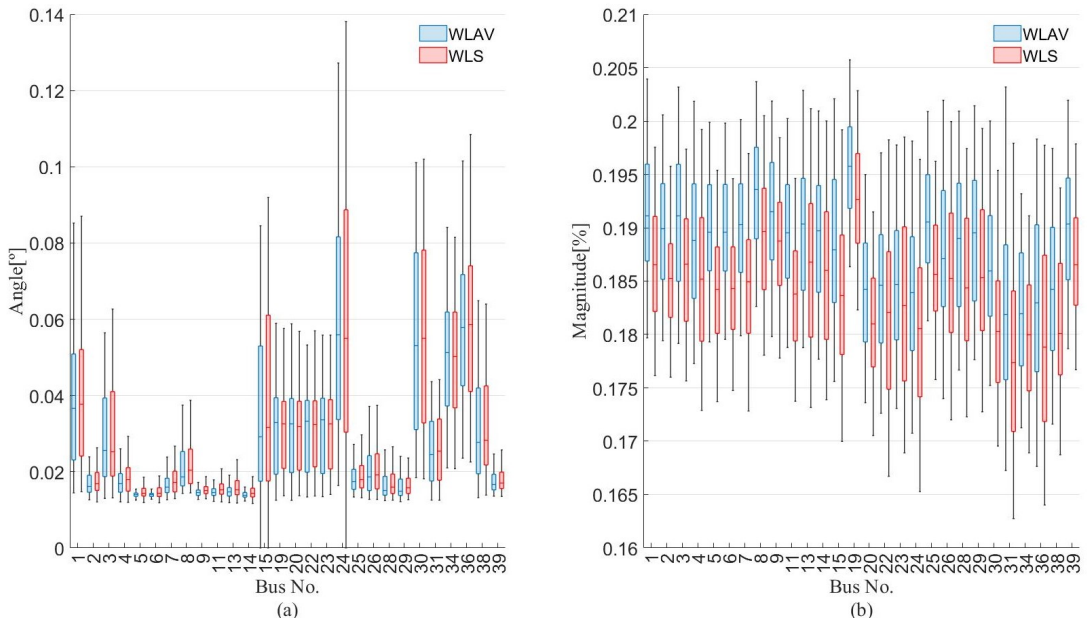

Fig. 5.47 Distribution of the estimated static state error at observed buses when a three-phase fault takes place (a) voltage angle, (b) voltage magnitude

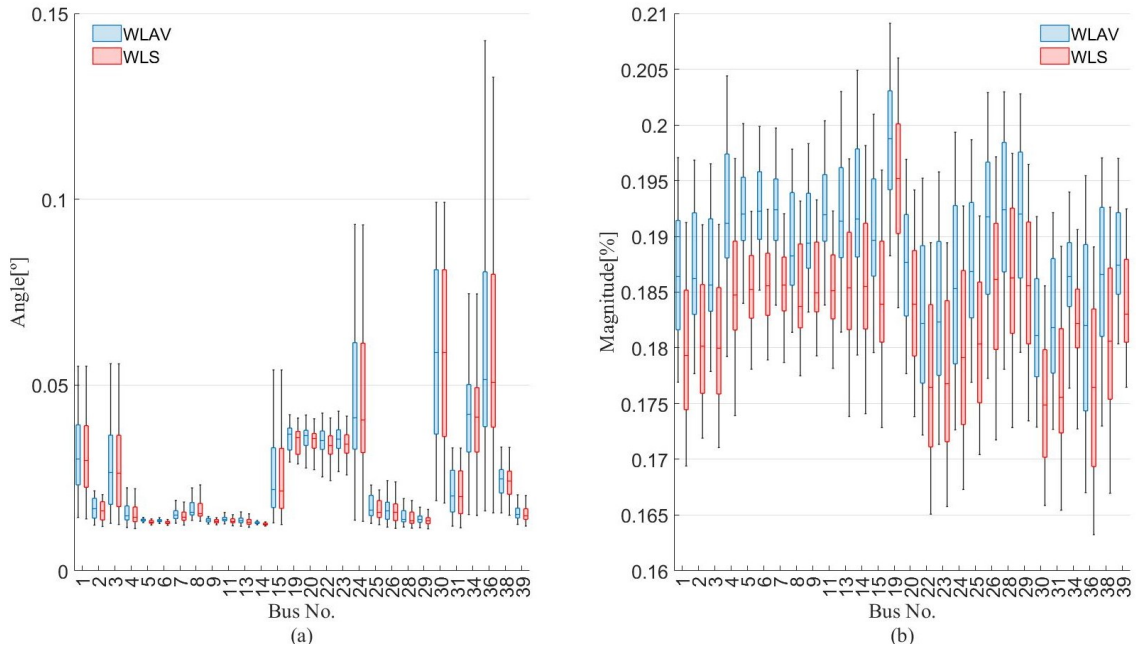

Fig. 5.48 Distribution of the estimated static state error at observed buses when a load curtailment takes place (a) voltage angle, (b) voltage magnitude

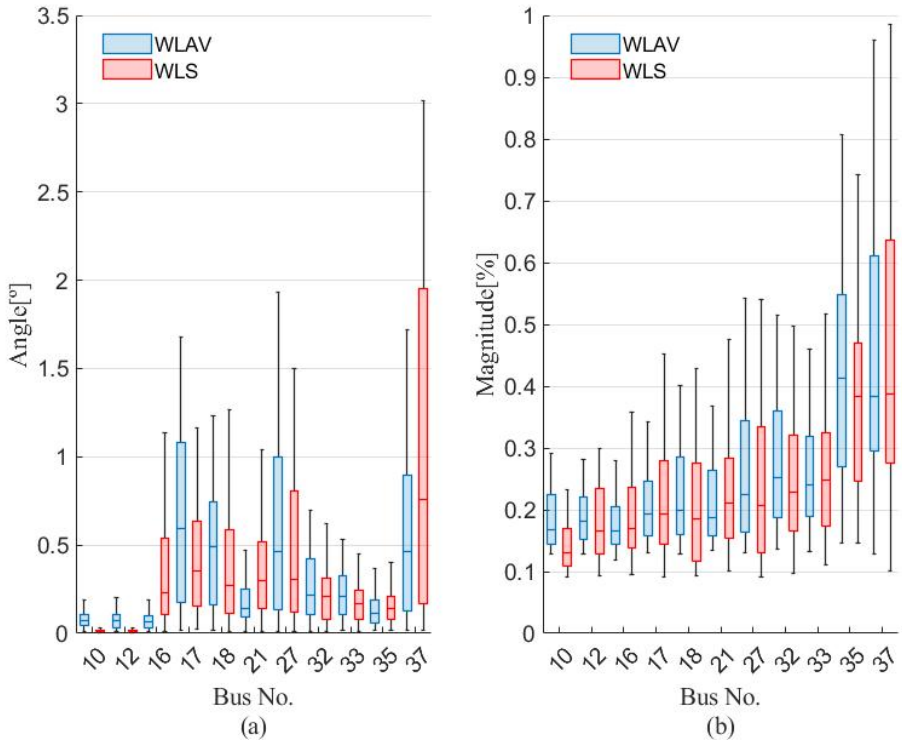

Fig. 5.49 Distribution of the estimated static state error at non-observed buses when a generator outage takes place (a) voltage angle, (b) voltage magnitude

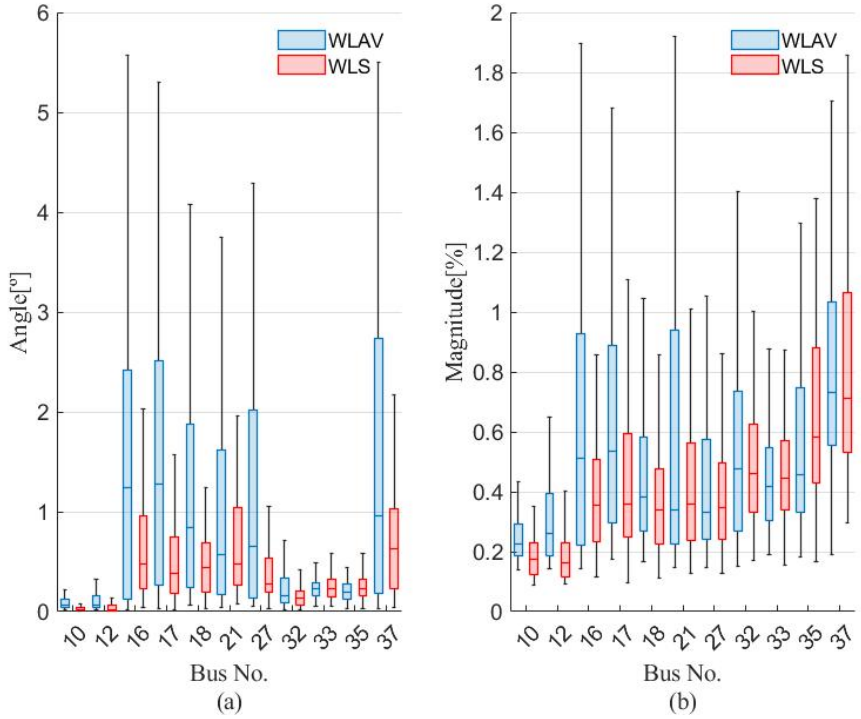

Fig. 5.50 Distribution of the estimated static state error at non-observed buses when a three-phase fault takes place (a) voltage angle, (b) voltage magnitude

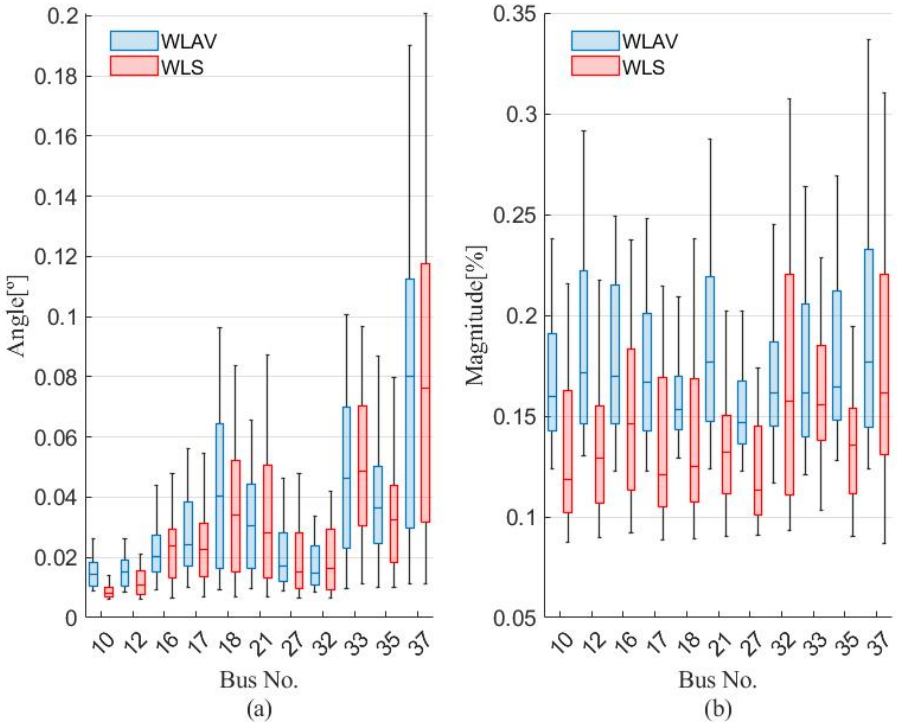

Fig. 5.51 Distribution of the estimated static state error at non-observed buses when a load curtailment takes place (a) voltage angle, (b) voltage magnitude

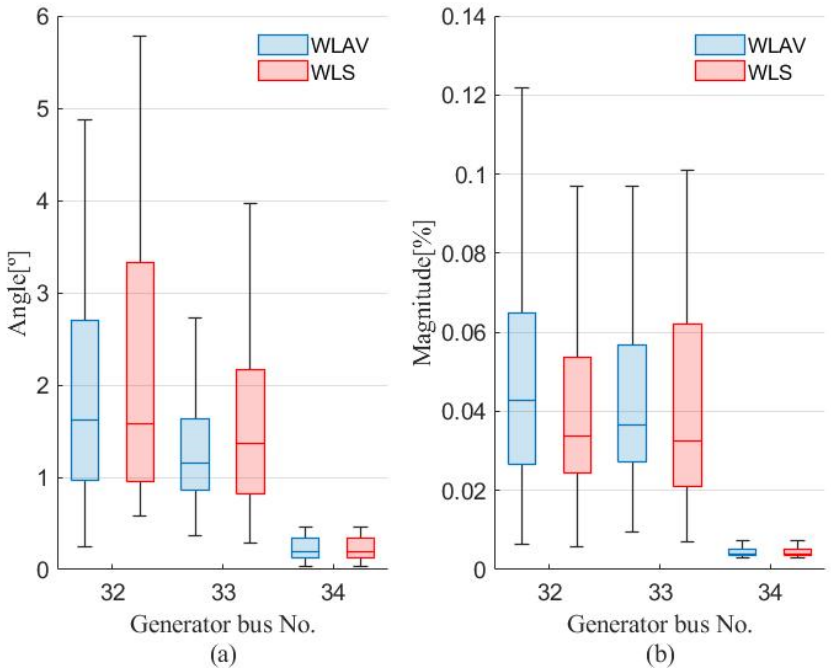

Fig. 5.52 Distribution of the dynamic state (generators at buses 32, 33 and 34) error when a generator outage takes place (a) rotor angle of generator, (b) rotation speed of generator

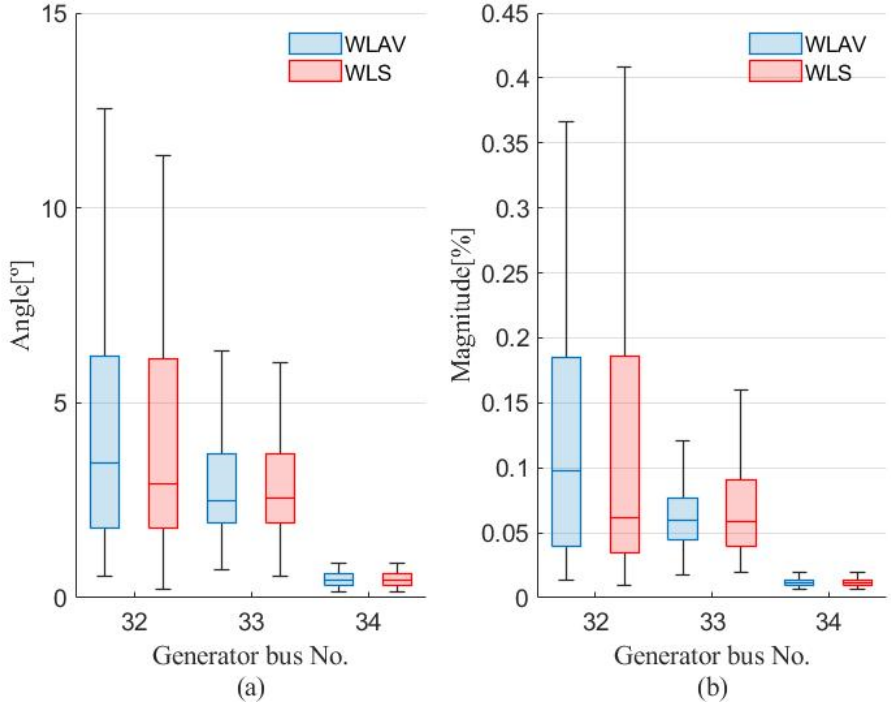

Fig. 5.53 Distribution of the dynamic state (generators at buses 32, 33 and 34) error when a three-phase fault takes place (a) rotor angle of generator, (b) rotation speed of generator

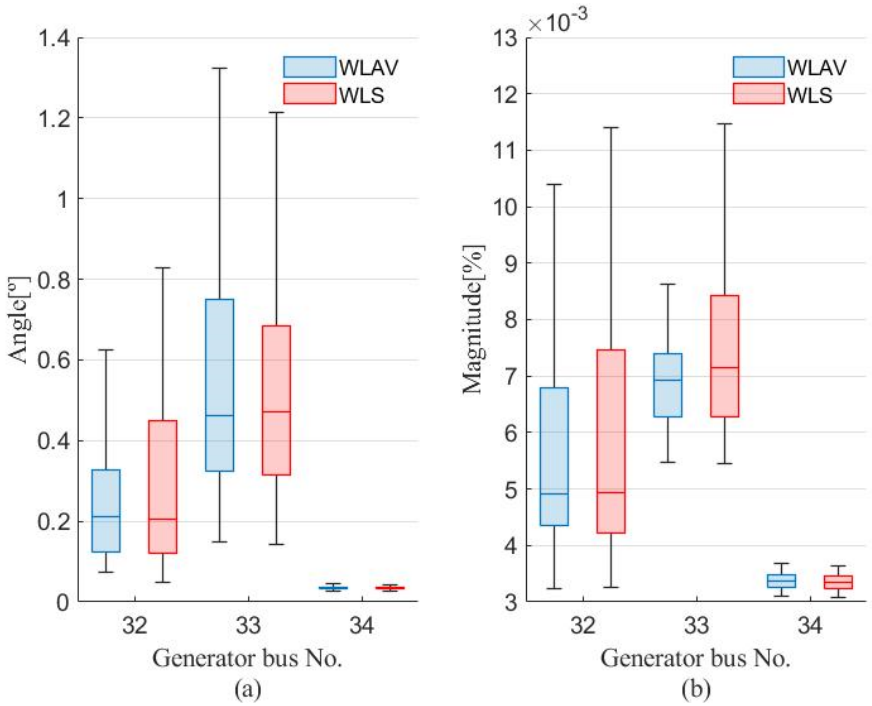

Fig. 5.54 Distribution of the dynamic state (generators at buses 32, 33 and 34) error when a load curtailment takes place (a) rotor angle of generator, (b) rotation speed of generator

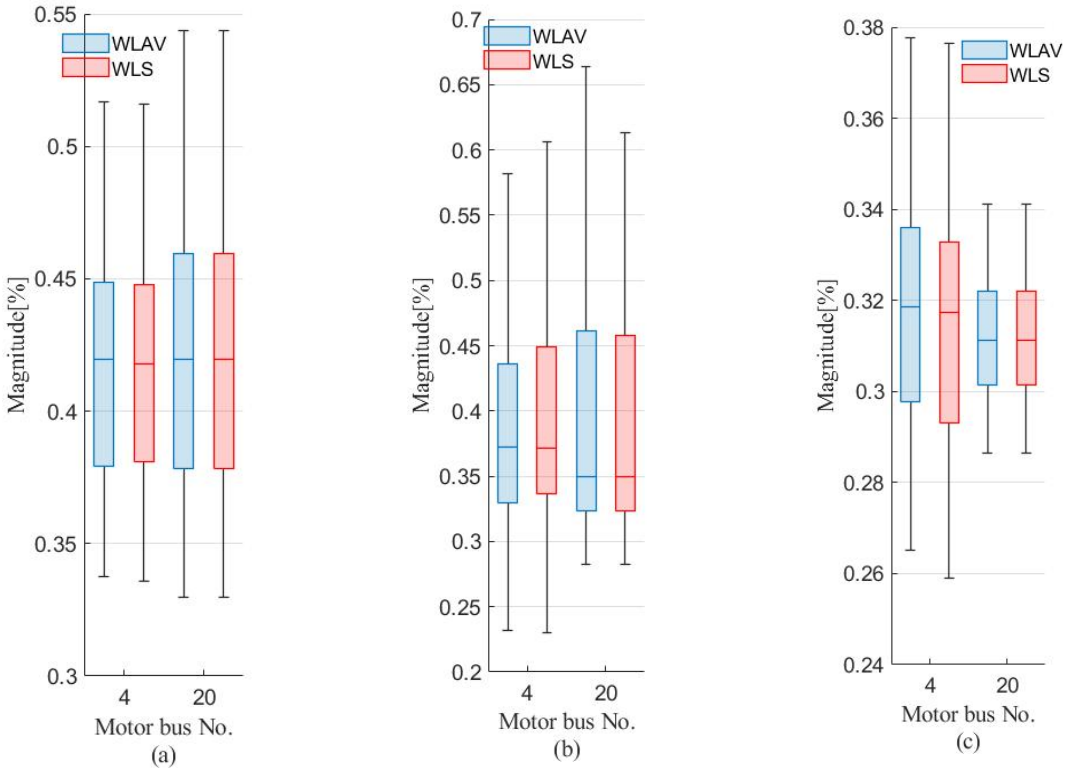

Fig. 5.55 Distribution of the estimated rotation speed error (motors at buses 4 and 20) according to the kind of contingency (a) generator outage, (b) three-phase fault, (c) load curtailment

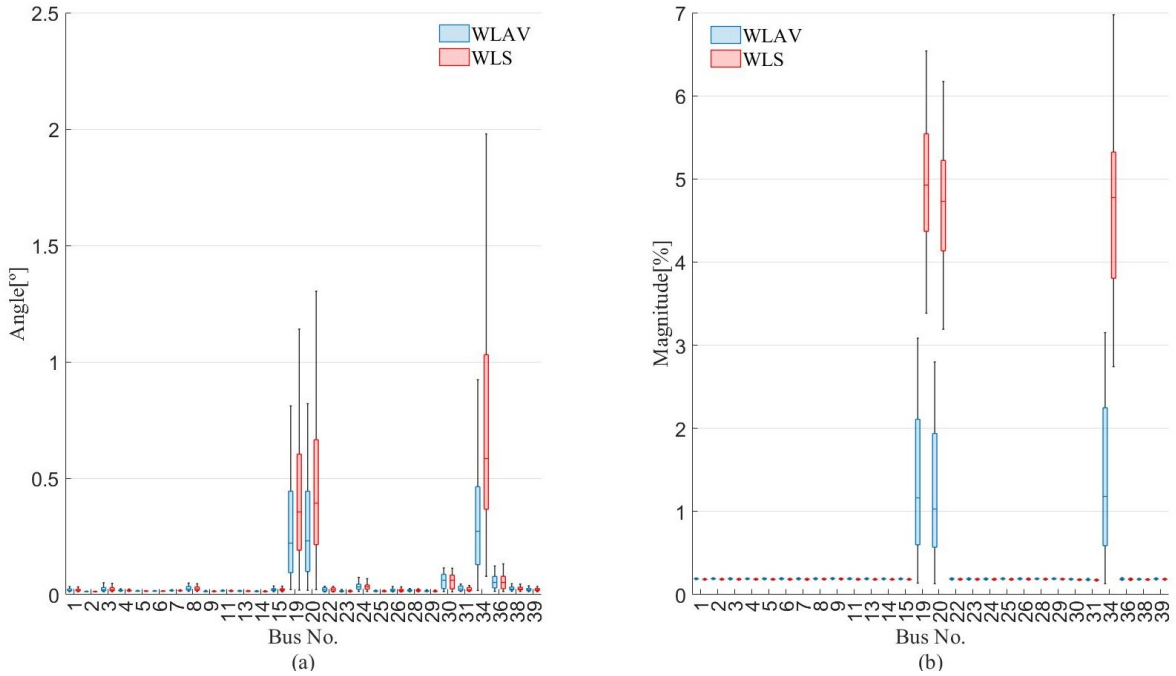

Fig. 5.56 Distribution of the estimated static state error at observed buses when a generator outage takes place (a) voltage angle, (b) voltage magnitude

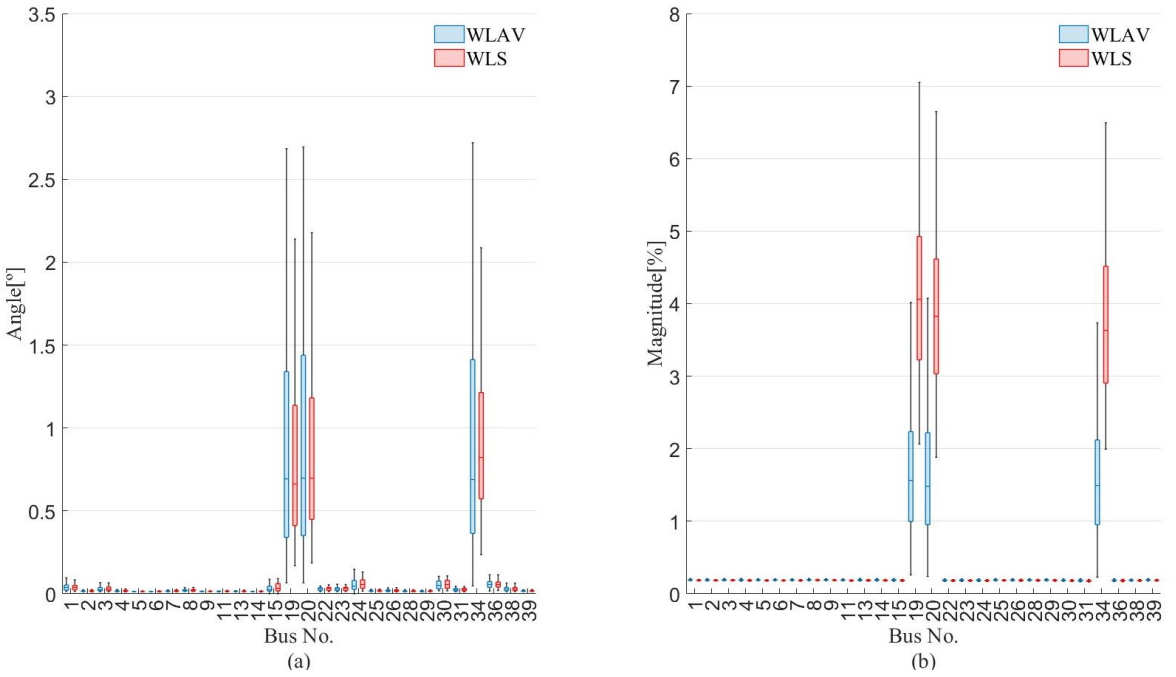

Fig. 5.57 Distribution of the estimated static state error at observed buses when a three-phase fault takes place (a) voltage angle, (b) voltage magnitude

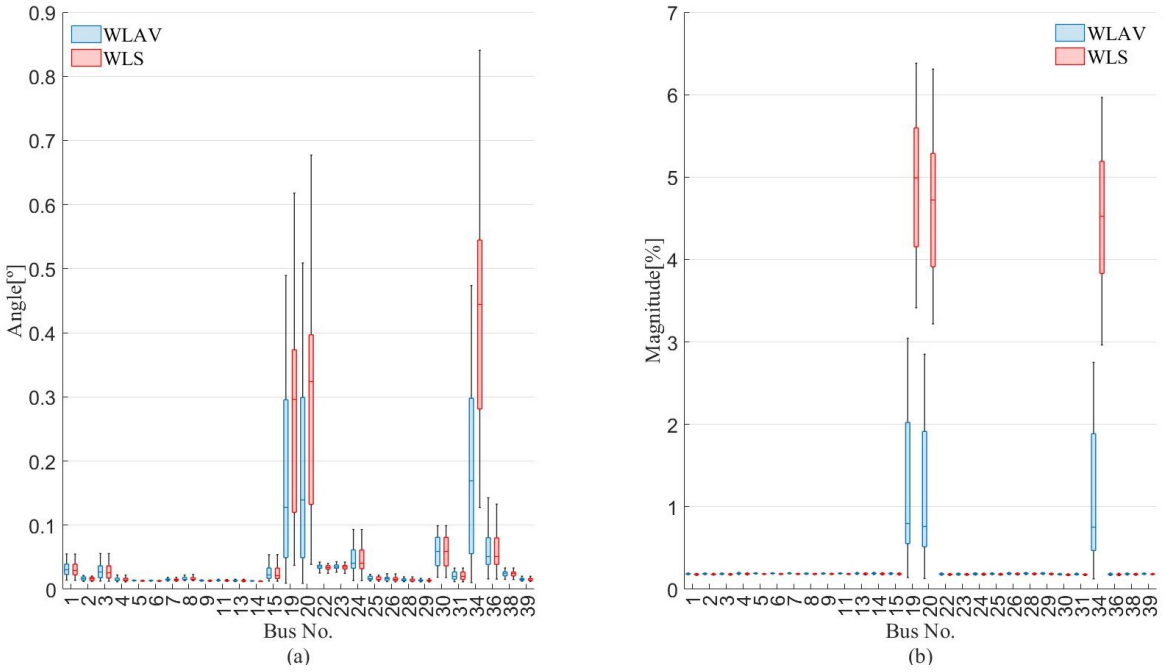

Fig. 5.58 Distribution of the estimated static state error at observed buses when a load curtailment takes place (a) voltage angle, (b) voltage magnitude

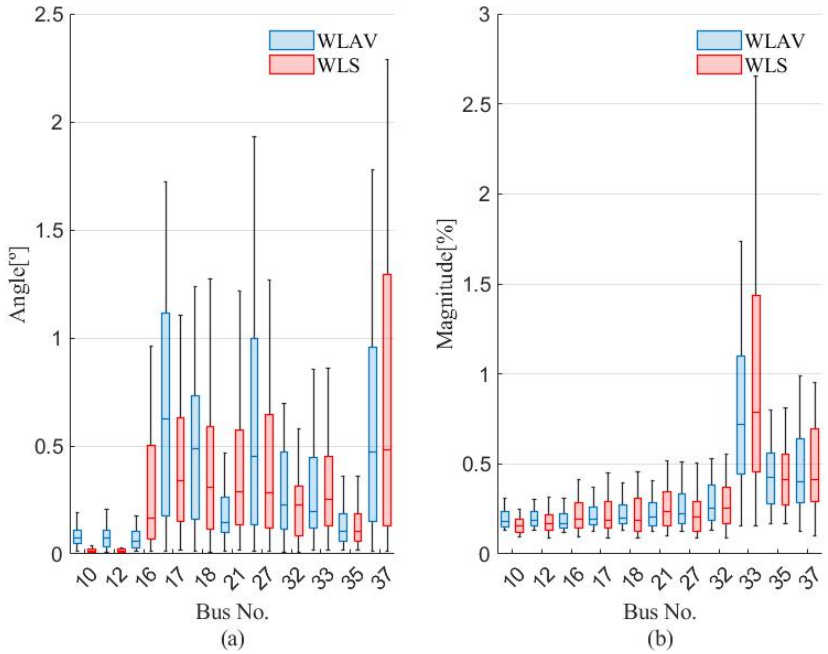

Fig. 5.59 Distribution of the estimated static state error at non-observed buses when a generator outage takes place (a) voltage angle, (b) voltage magnitude

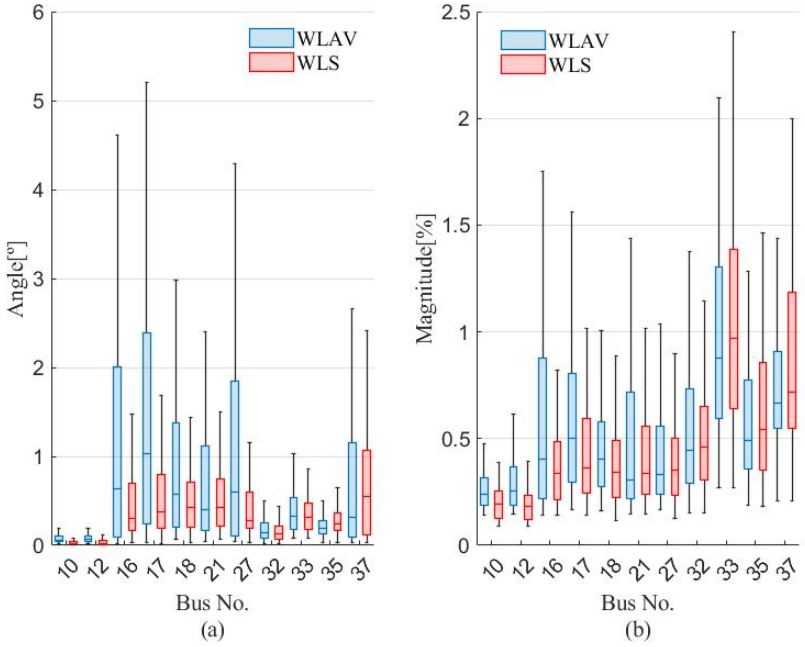

Fig. 5.60 Distribution of the estimated static state error at non-observed buses when a three-phase fault takes place (a) voltage angle, (b) voltage magnitude

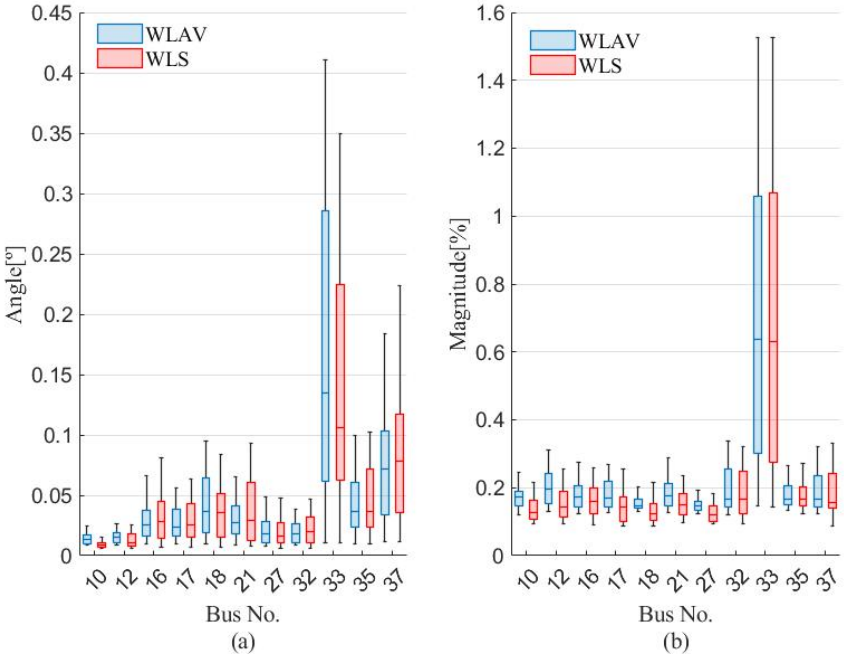

Fig. 5.61 Distribution of the estimated static state error at non-observed buses when a load curtailment takes place (a) voltage angle, (b) voltage magnitude

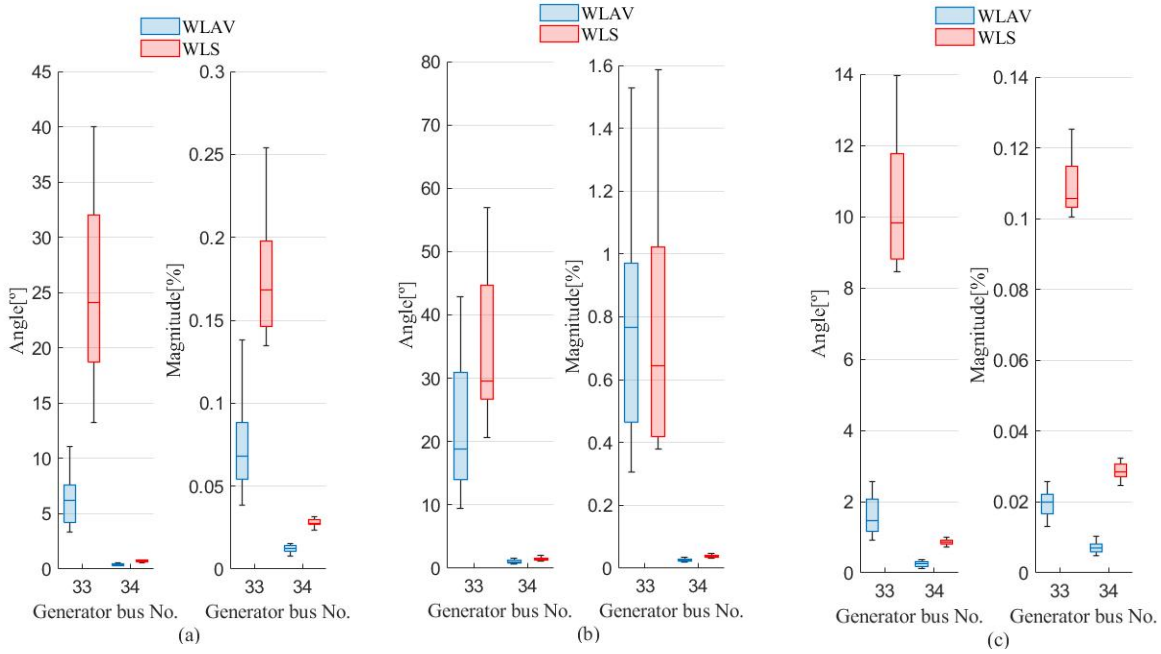

Fig. 5.62 Error distribution of the estimated rotor angle and rotation speed (generators at buses 33 and 34) according to the kind of contingency (a) generator outage, (b) three-phase fault, (c) load curtailment

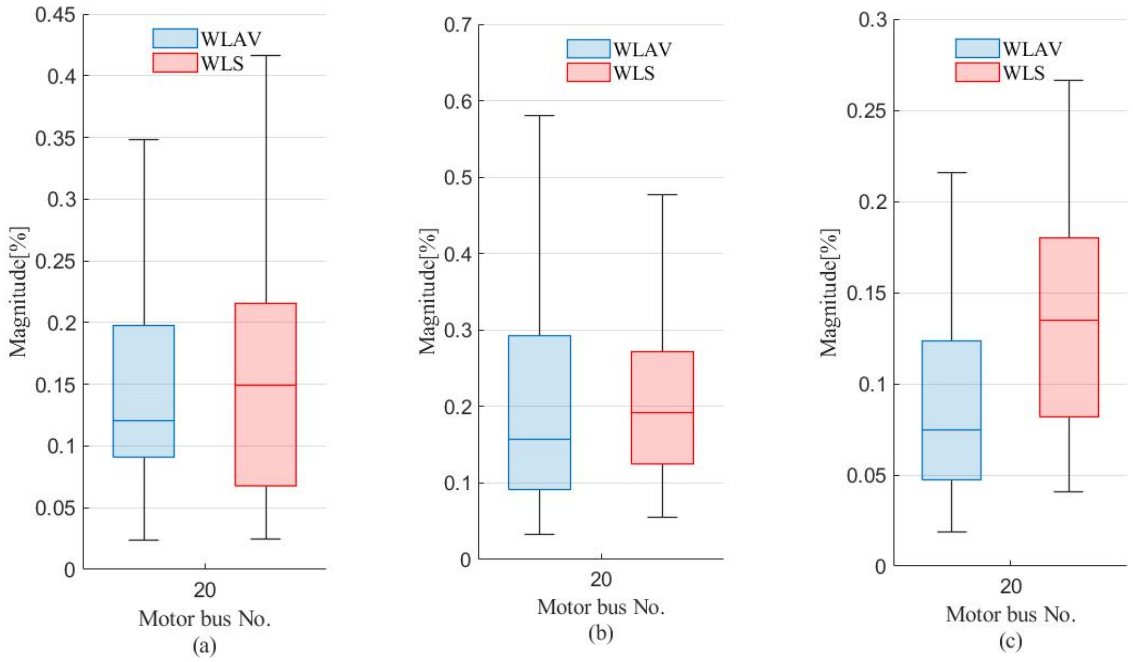

Fig. 5.63 Error distribution of the estimated rotation speed (motors at bus 20) according to the kind of contingency (a) generator outage, (b) three-phase fault, (c) load curtailment

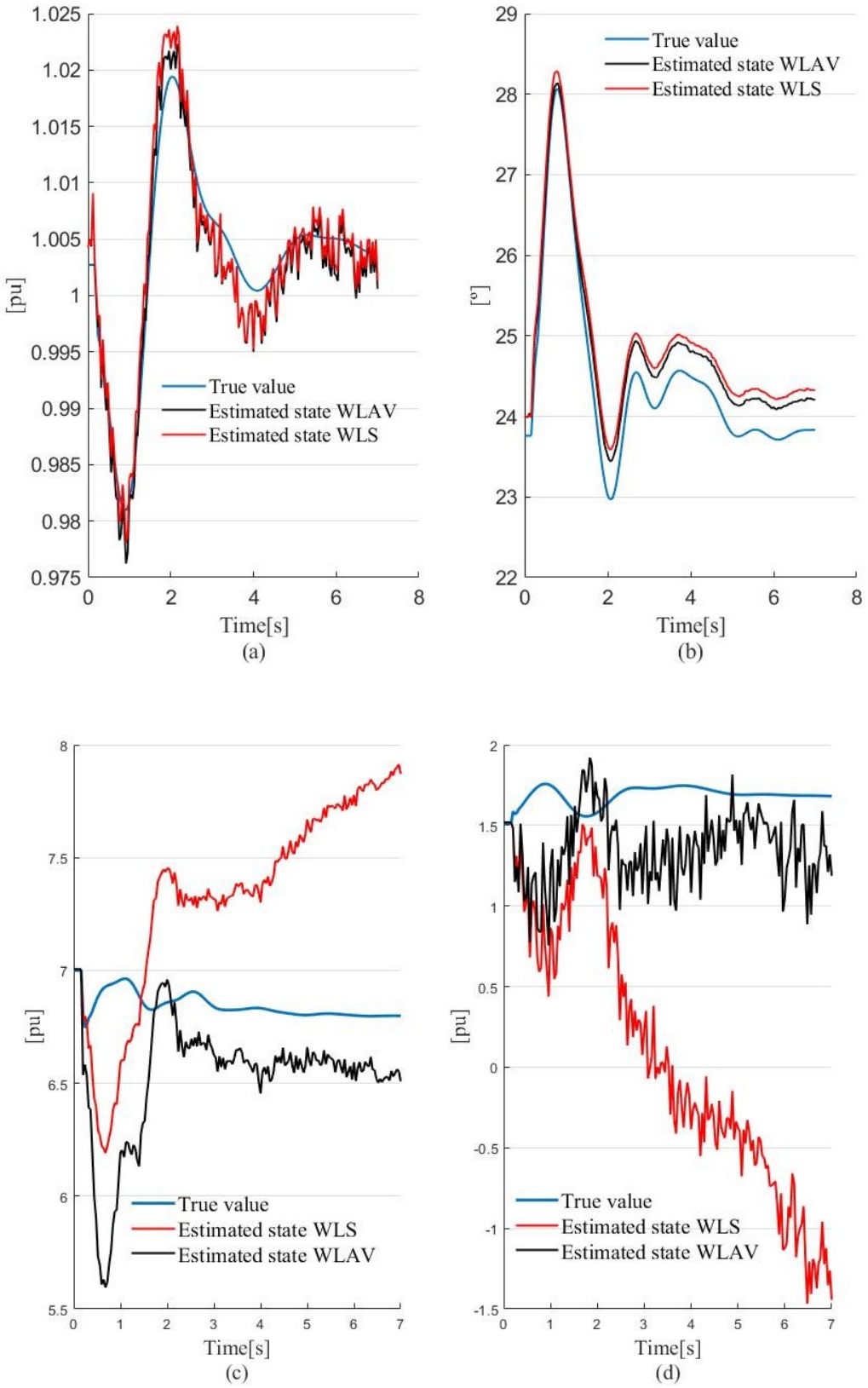

Fig. 5.64 Estimated static state at bus 33 under a load curtailment at bus 24 (a) voltage magnitude, (b) voltage angle, (c) active power injection, (d) reactive power injection

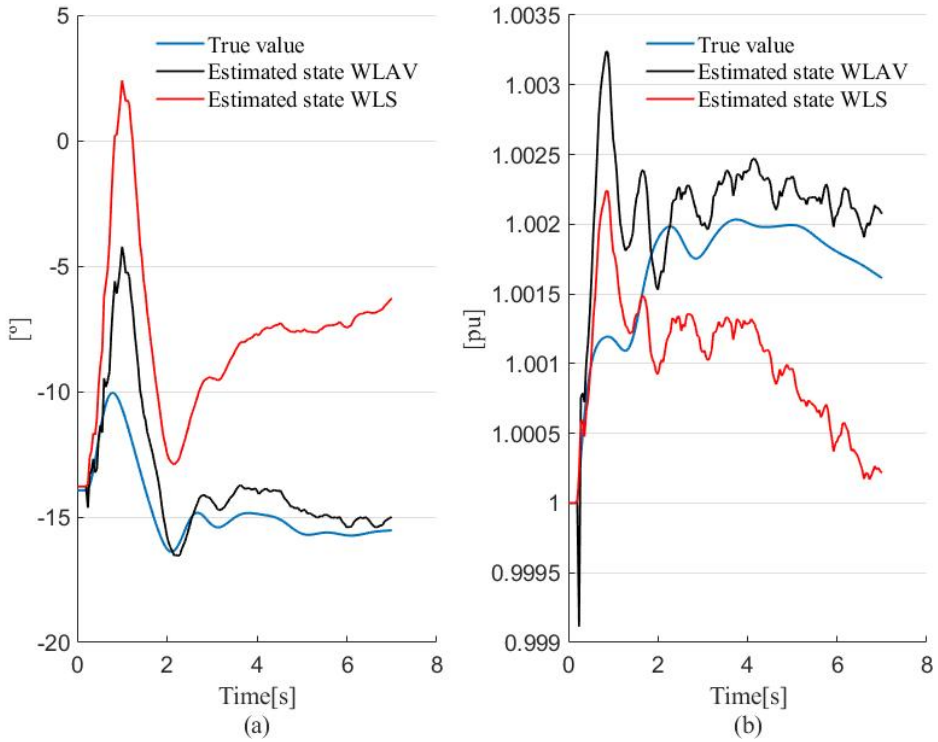

Fig. 5.65 Estimated dynamic state (generator at bus 33) under a load curtailment at bus 24 (a) rotor angle of generator, (b) rotation speed of generator

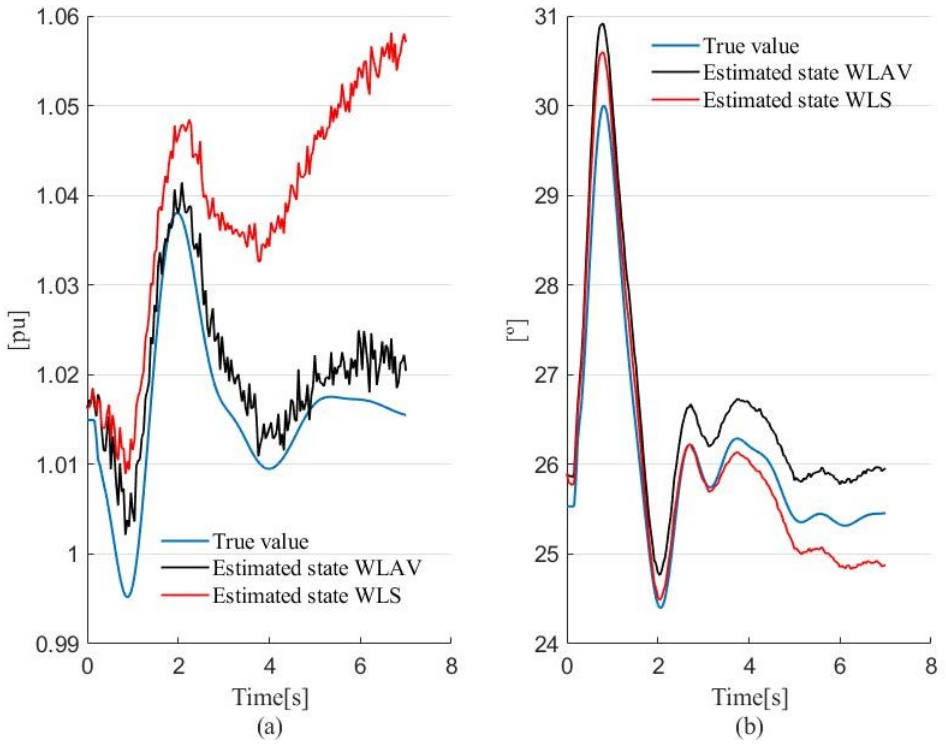

Fig. 5.66 Estimated static state at bus 34 under a load curtailment at bus 24 (a) voltage magnitude, (b) voltage angle

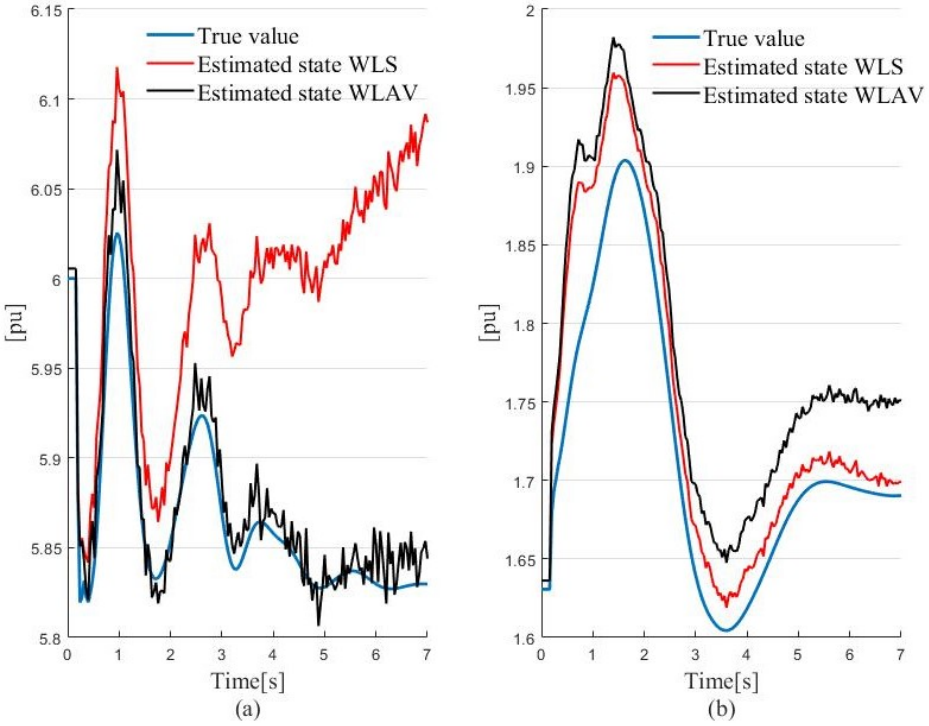

Fig. 5.67 Estimated static state at bus 34 under a load curtailment at bus 24 (a) active power injection, (b) reactive power injection

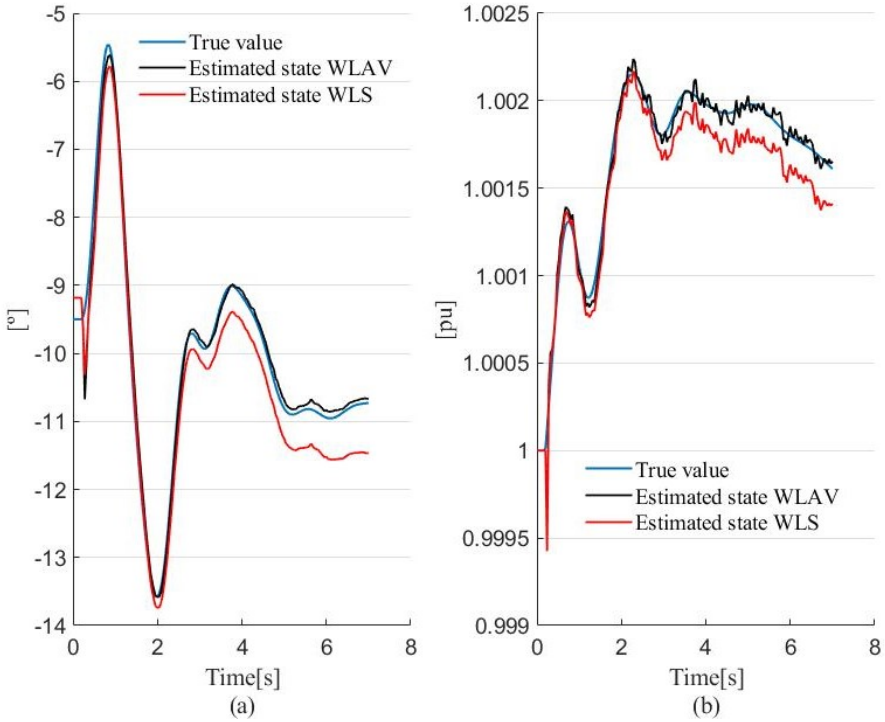

Fig. 5.68 Estimated dynamic state (generator at bus 34) under a load curtailment at bus 24 (a) rotor angle of generator, (b) rotation speed of generator

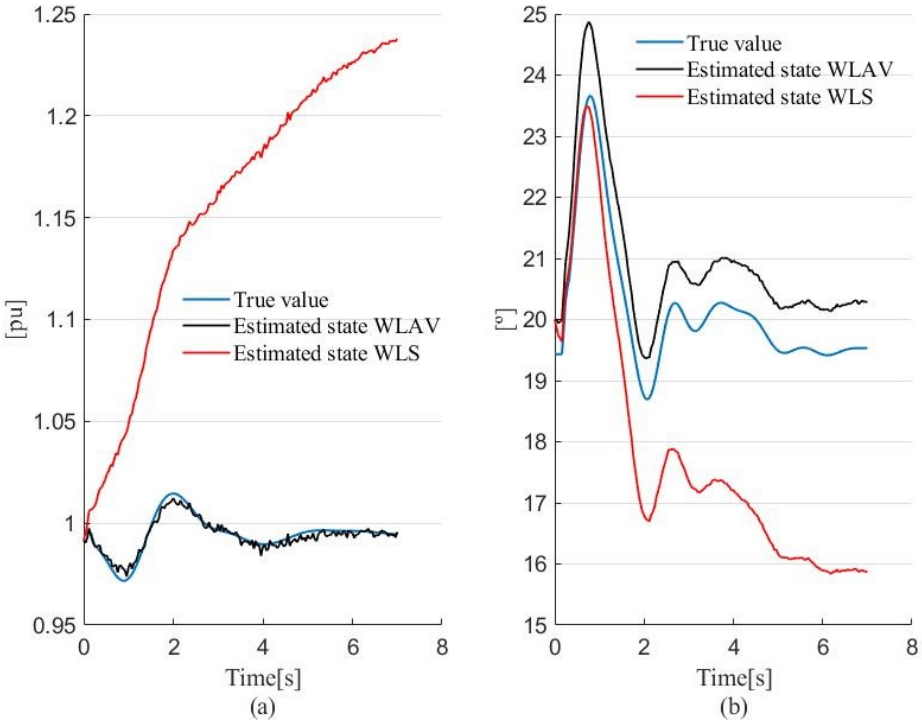

Fig. 5.69 Estimated static state at bus 20 under a load curtailment at bus 24 (a) voltage magnitude, (b) voltage angle

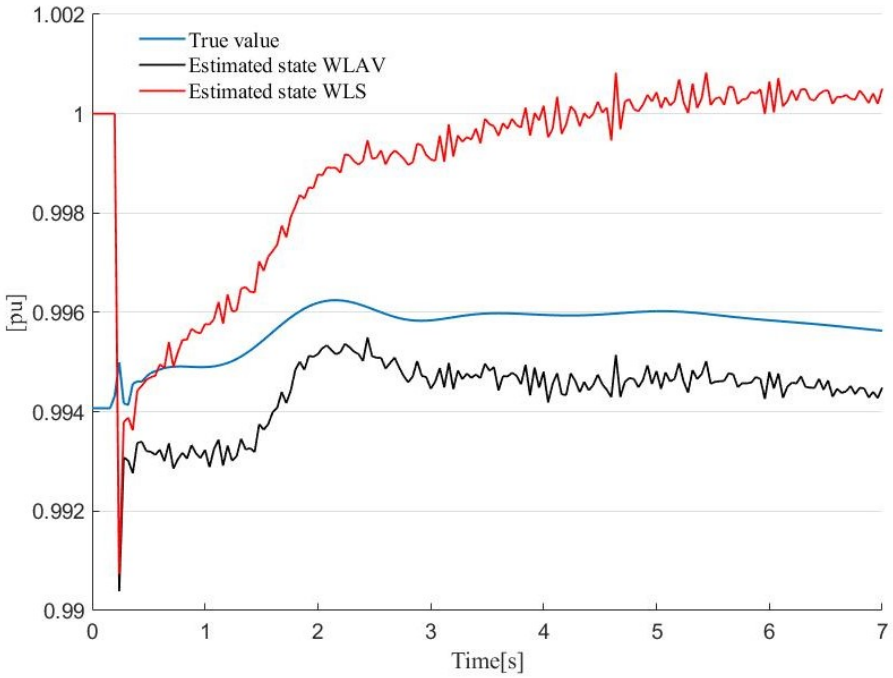

Fig. 5.70 Estimated dynamic state (rotation speed of motor at bus 20) under a load curtailment at bus 24

Case IV

This case aims to test the estimator performance against higher noise levels. For this purpose measurement variances and noise boundaries in grid parameters are increased one at a time by a factor of five with respect to their initial values. Convergence and estimation accuracy are evaluated under the new operating conditions.

In relation to convergence, higher noise levels lead to an increase in the number of cases where the algorithm does not approach a finite value. It has been found that a noise level growth in the measurement variance by a factor of 5 increases by about 8% the number of scenarios where the EKF fails to converge. This amount falls to 2% if the additional noise affects only the grid parameters.

Figures 5.71 to 5.76 show the distribution of the estimated static state error at observed and non-observed buses. Results are grouped together according to the kind of contingency and kind of noise under consideration. It can be observed from the error distribution depicted by the boxplots that measurement noises have a major impact on the estimated voltage magnitude. On the other hand, parameter noises have a greater influence on the estimated voltage angle.

Figures 5.77 to 5.80 depict the distribution of the dynamic state error at buses 4, 20, 33, 34 and 35. As before, results are grouped together according to the kind of contingency and the kind of noise under analysis. The distribution of the estimation error becomes usually a little more dispersed and reaches higher values when adding additional noise in measurements. This is consistent with the increases in the number of scenarios for both cases, i.e. 8% when having higher noise levels in measurements and 2% for higher noise levels in grid parameters, where the EKF fails to converge.

As expected, higher noise levels leads to greater estimation errors, either in static or dynamic variables. Nevertheless, the increase in the median of the estimated static and dynamic error is minor. Then, it can be affirmed that the estimator remains robust under such new operating conditions.

In addition, it can be observed that there is a clear relationship between the degree of dispersion of the estimation error and it highest value at a certain bus and the bus location relative to a PMU bus. The greater the distance between a certain bus and a PMU bus, the greater the degree of dispersion of the estimation error and its maximum value.

As an example, Figures 5.81 to 5.104 show the estimated static and dynamic states at buses 4, 20, 33, 34 and 35 when a generator outage at bus 39 occurs and considering higher noise levels

in measurements and grid parameters as well. As can be observed, under any of the assumed conditions the algorithm is capable of estimating the static and dynamic state with an adequate accuracy.

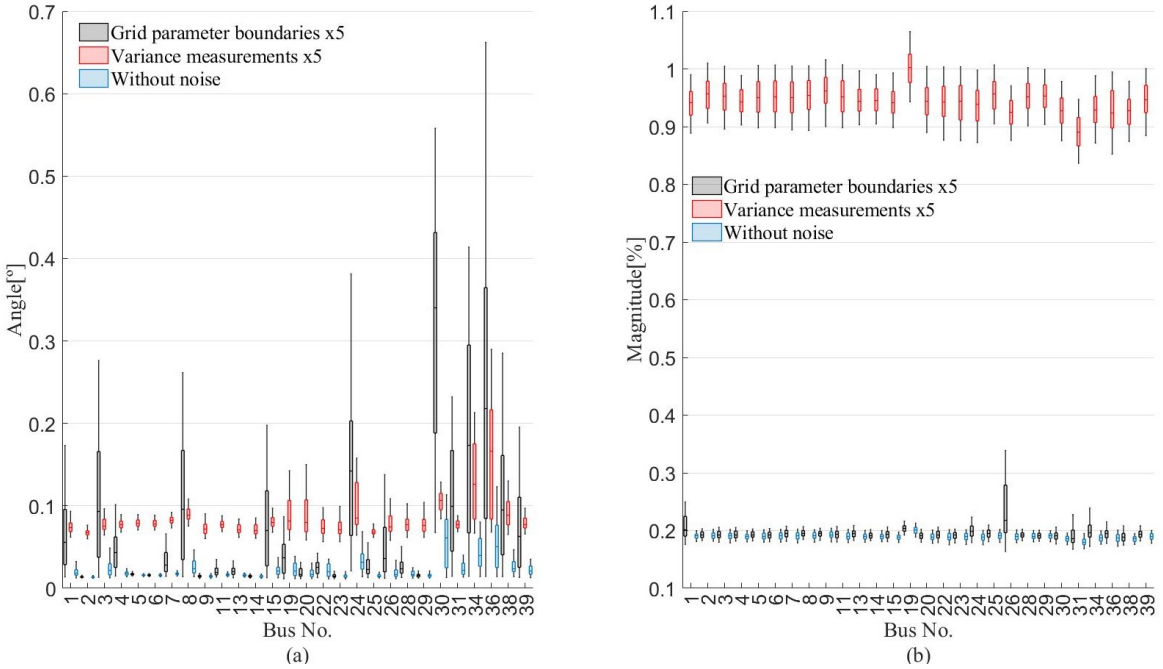

Fig. 5.71 Distribution of the estimated static state error at observed buses when a generator outage takes place (a) voltage angle, (b) voltage magnitude

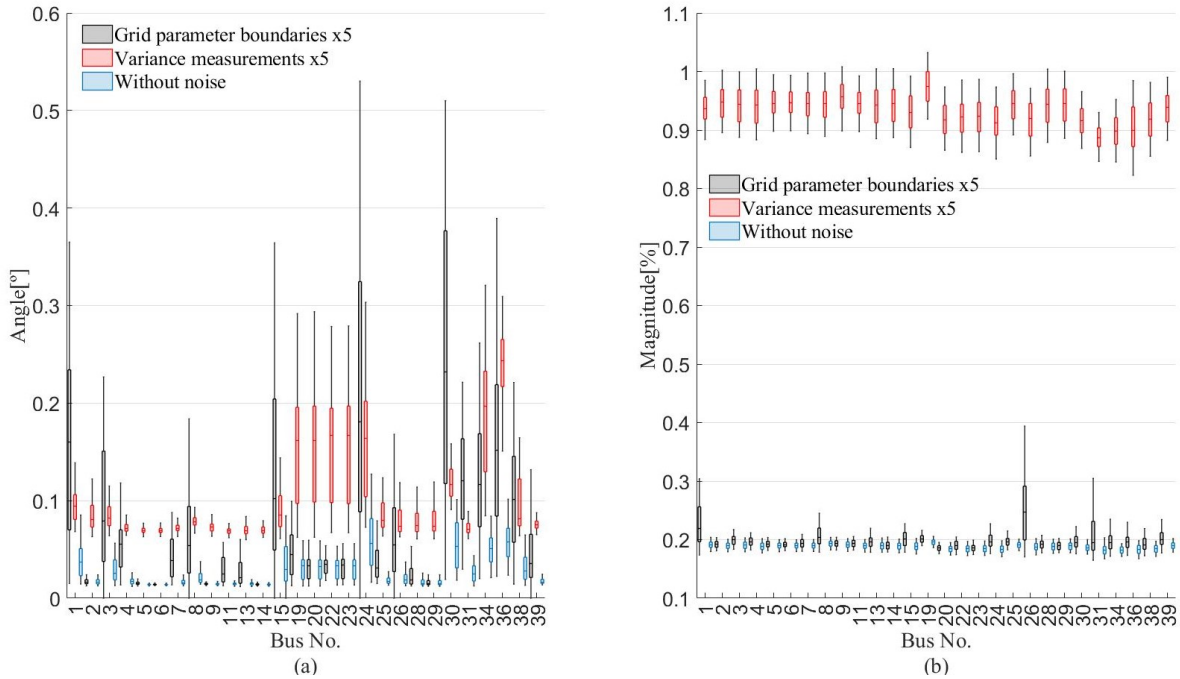

Fig. 5.72 Distribution of the estimated static state error at observed buses when a three-phase fault takes place (a) voltage angle, (b) voltage magnitude

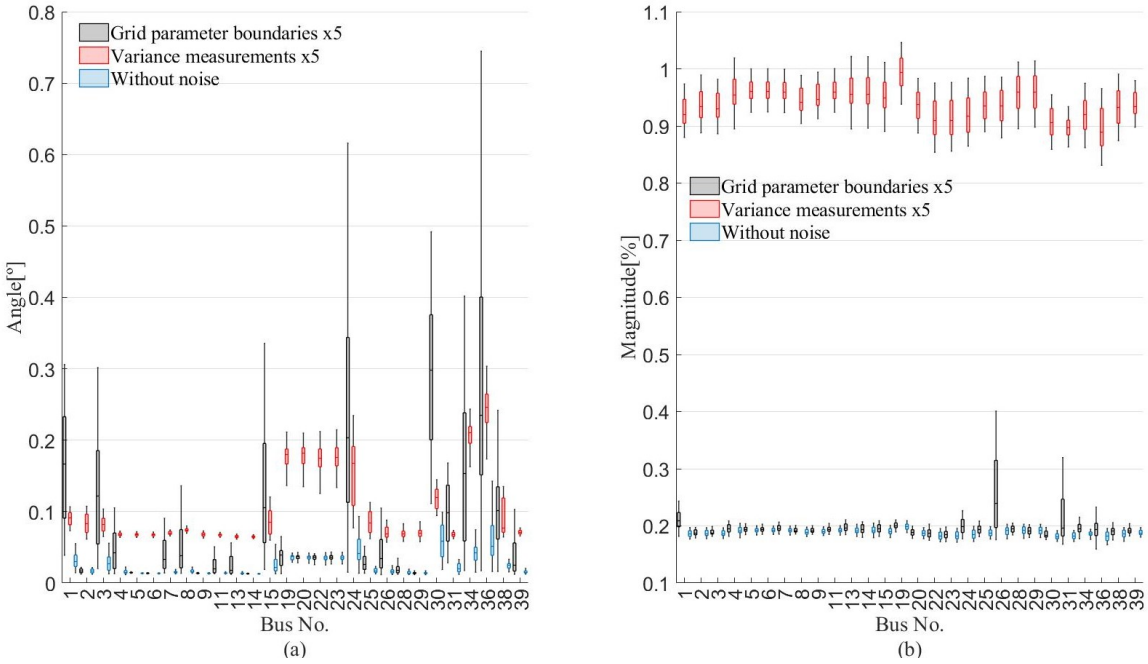

Fig. 5.73 Distribution of the estimated static state error at observed buses when a load curtailment takes place (a) voltage angle, (b) voltage magnitude

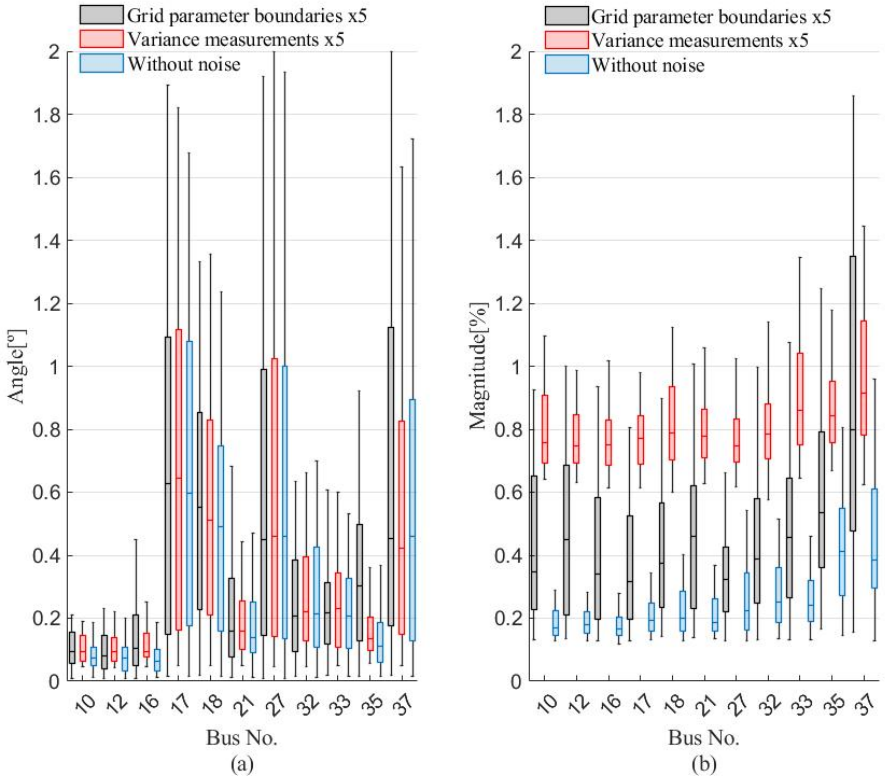

Fig. 5.74 Distribution of the estimated static state error at non-observed buses when a generator outage takes place (a) voltage angle, (b) voltage magnitude

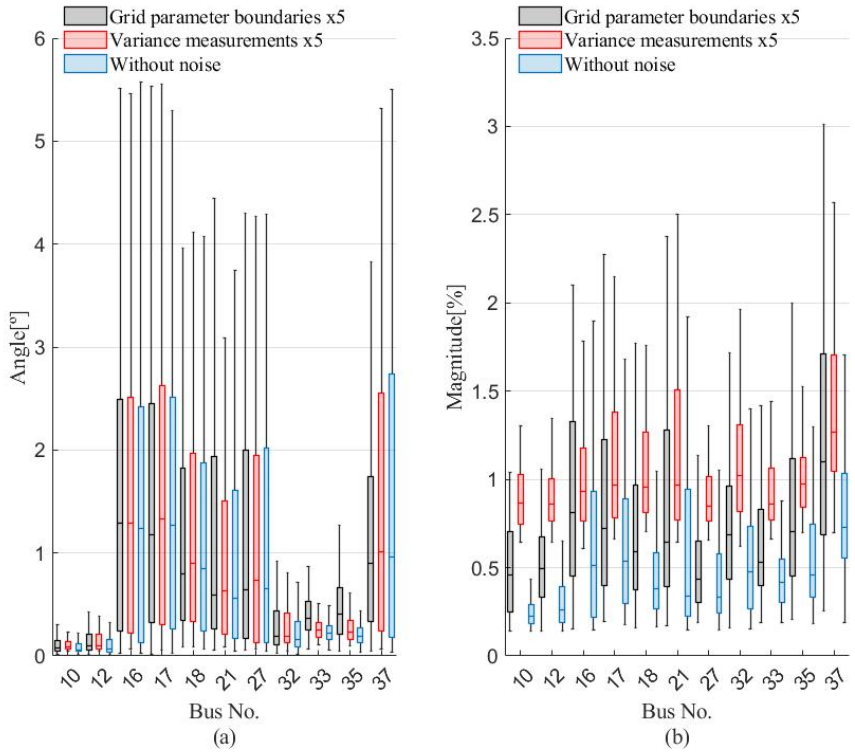

Fig. 5.75 Distribution of the estimated static state error at non-observed buses when a three-phase fault takes place (a) voltage angle, (b) voltage magnitude

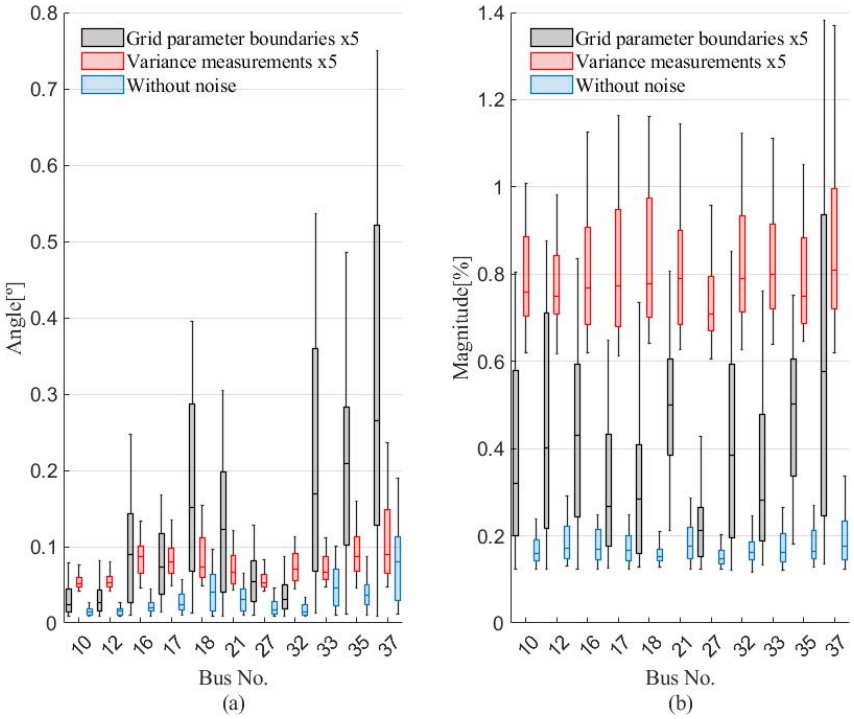

Fig. 5.76 Distribution of the estimated static state error at non-observed buses when a load curtailment takes place (a) voltage angle, (b) voltage magnitude

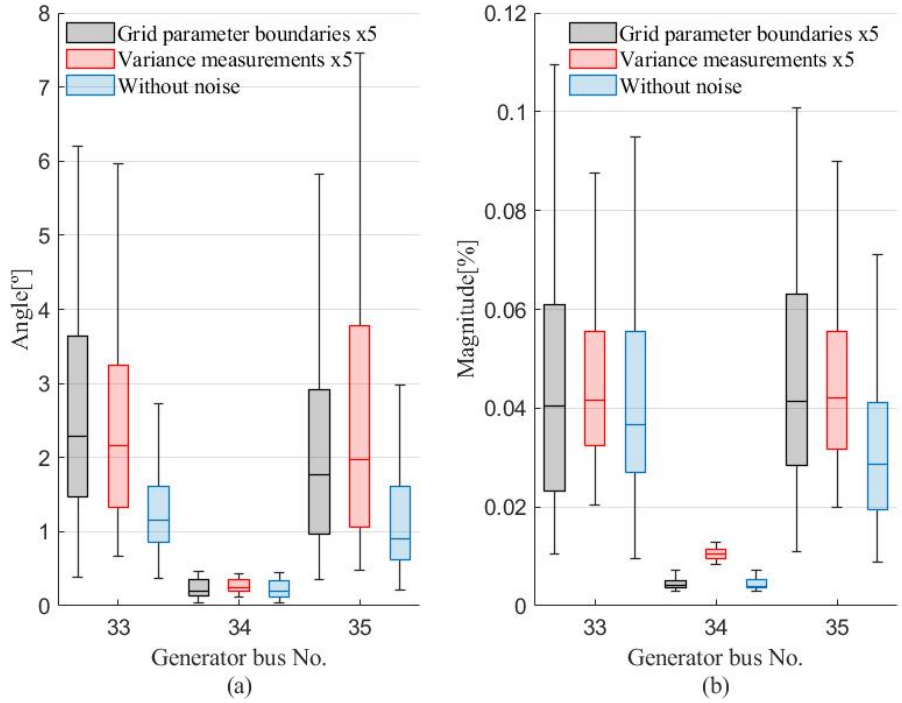

Fig. 5.77 Distribution of the estimated dynamic state error (generator at buses 33, 34, and 35) when a generator outage takes place (a) rotor angle of generator, (b) rotation speed of generator

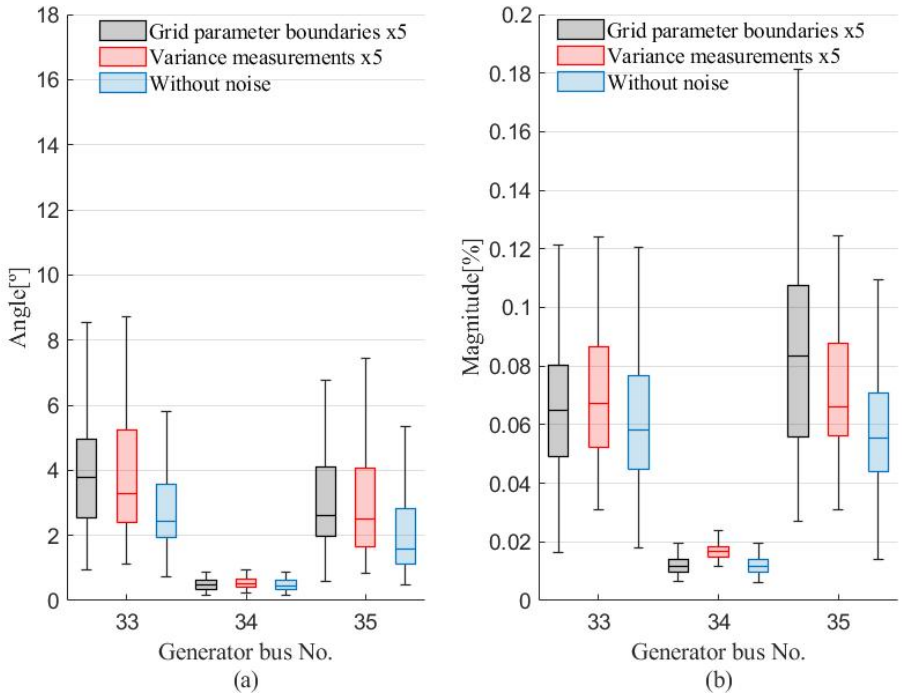

Fig. 5.78 Distribution of the estimated dynamic state error (generator at buses 33, 34, and 35) when a three-phase fault takes place (a) rotor angle of generator, (b) rotation speed of generator

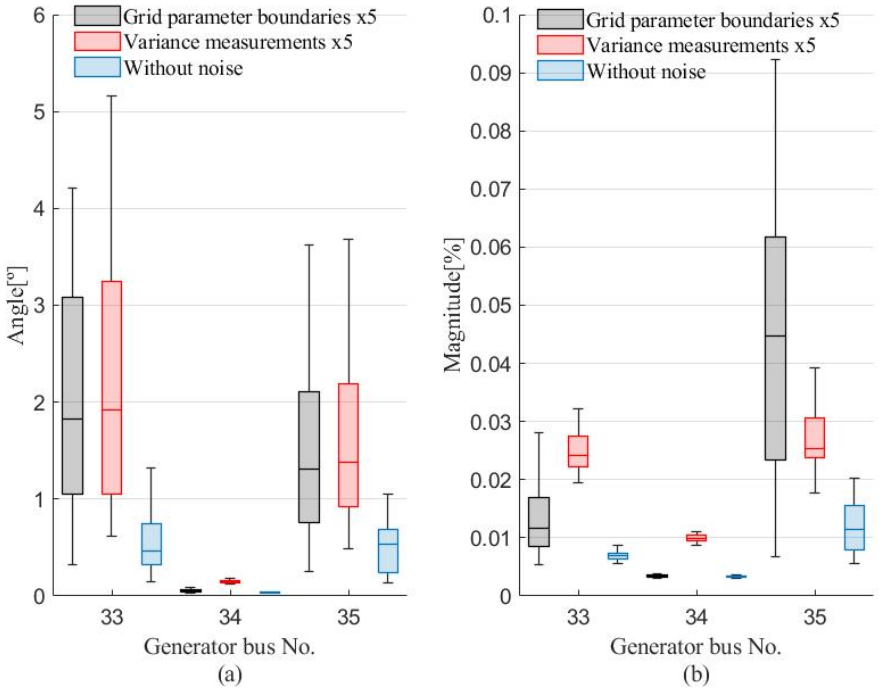

Fig. 5.79 Distribution of the estimated dynamic state error (generator at buses 33, 34, and 35) when a load curtailment takes place (a) rotor angle of generator, (b) rotation speed of generator

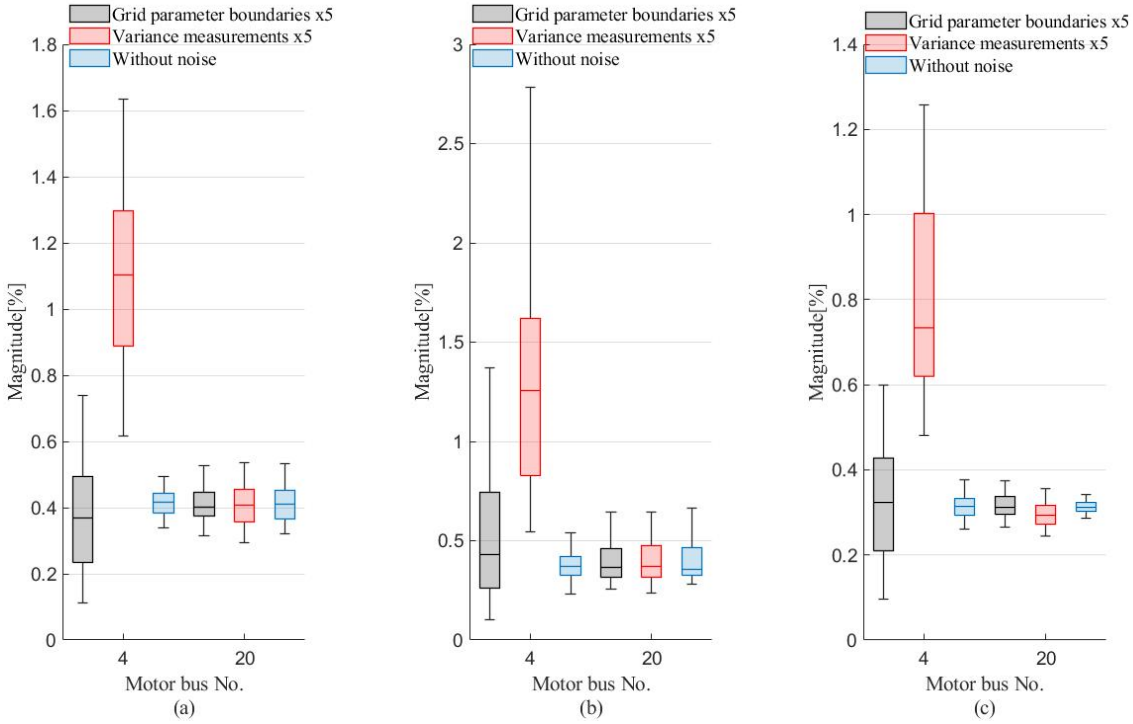

Fig. 5.80 Distribution of the estimated rotation speed error (motors at buses 4 and 20) according to the kind of contingency (a) generator outage, (b) three-phase fault, (c) load curtailment

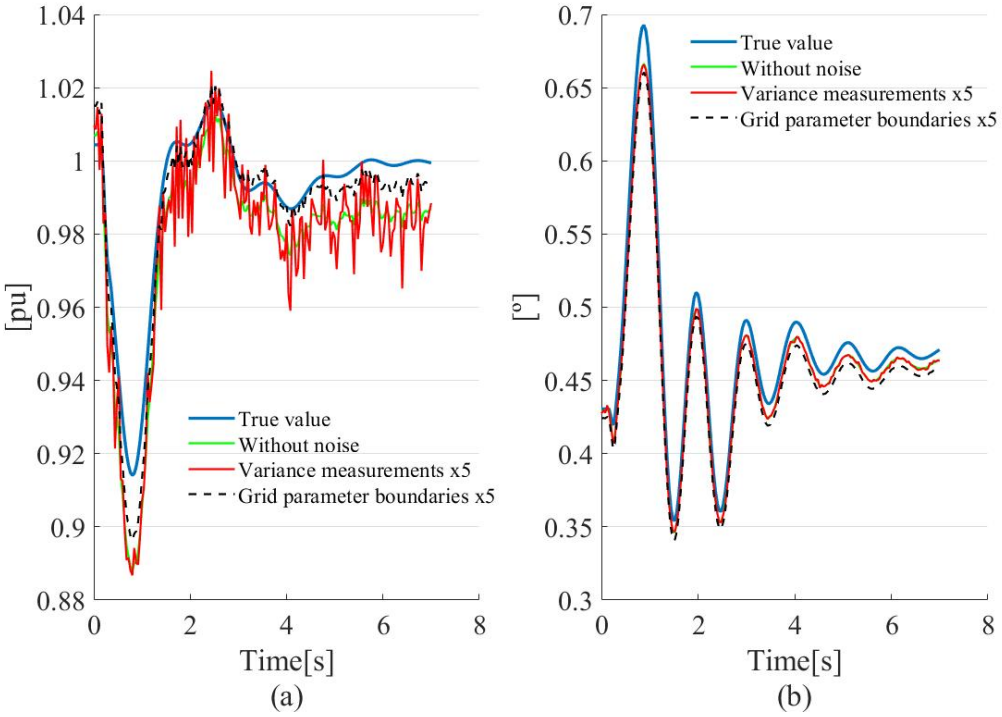

Fig. 5.81 Estimated static state at bus 33 under a generator outage at bus 39 (a) voltage magnitude, (b) voltage angle

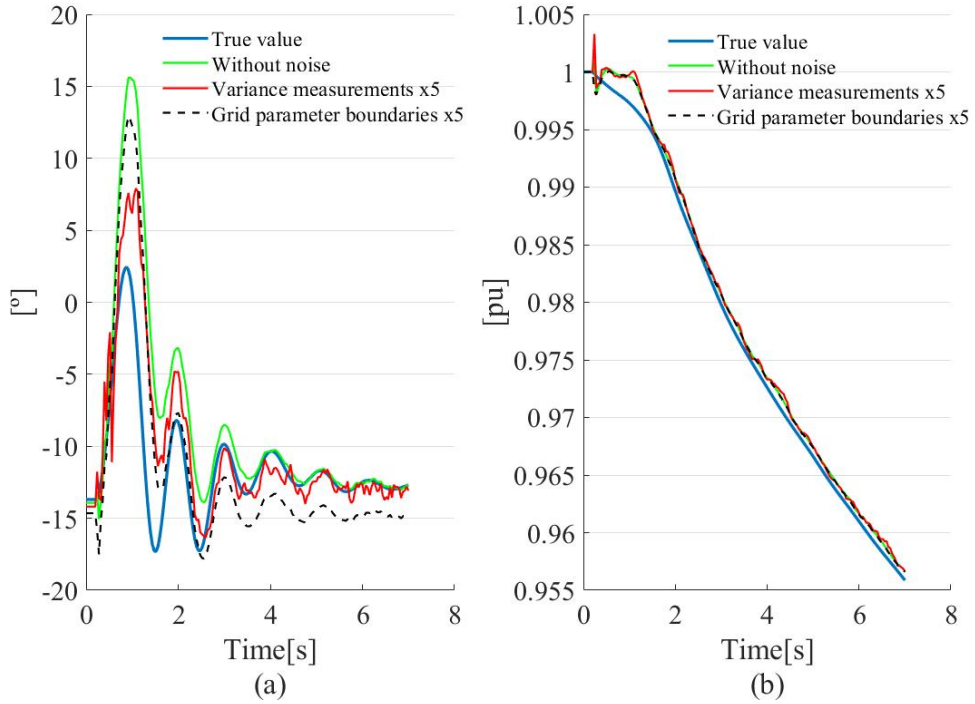

Fig. 5.82 Estimated dynamic state (generator at bus 33) under a generator outage at bus 39 (a) rotor angle of generator, (b) rotation speed of generator

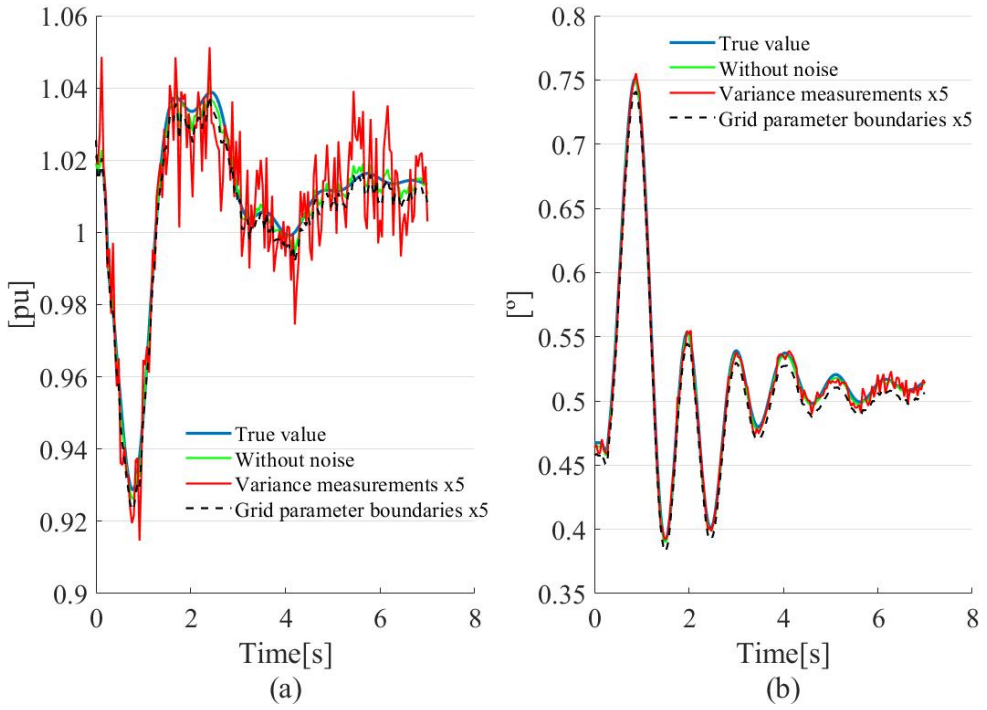

Fig. 5.83 Estimated static state at bus 34 under a generator outage at bus 39 (a) voltage magnitude, (b) voltage angle

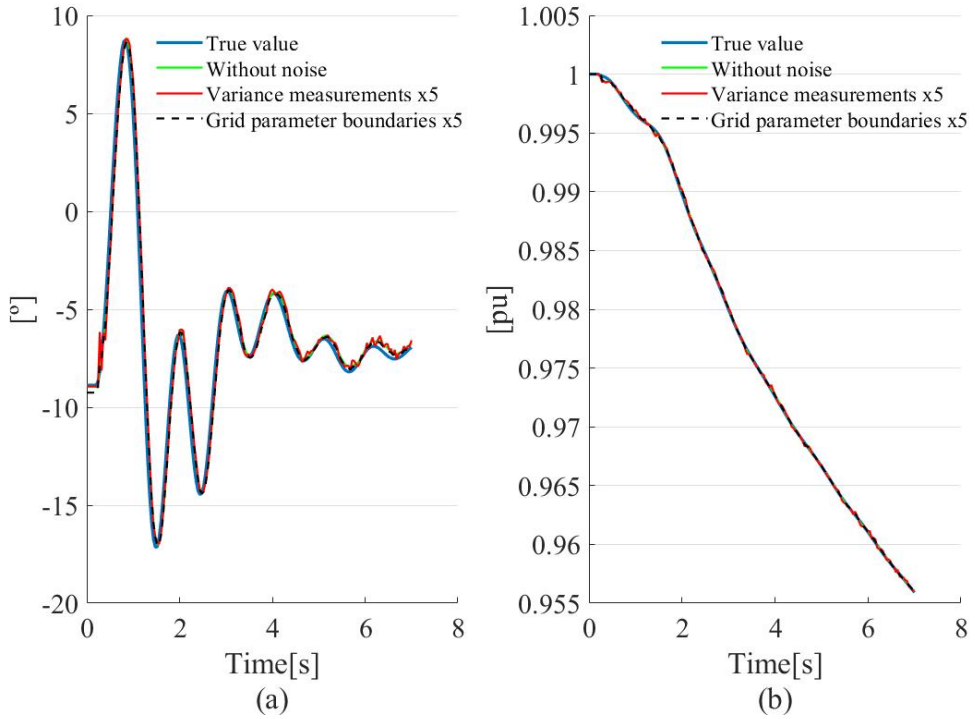

Fig. 5.84 Estimated dynamic state (generator at bus 34) under a generator outage at bus 39 (a) rotor angle of generator, (b) rotation speed of generator

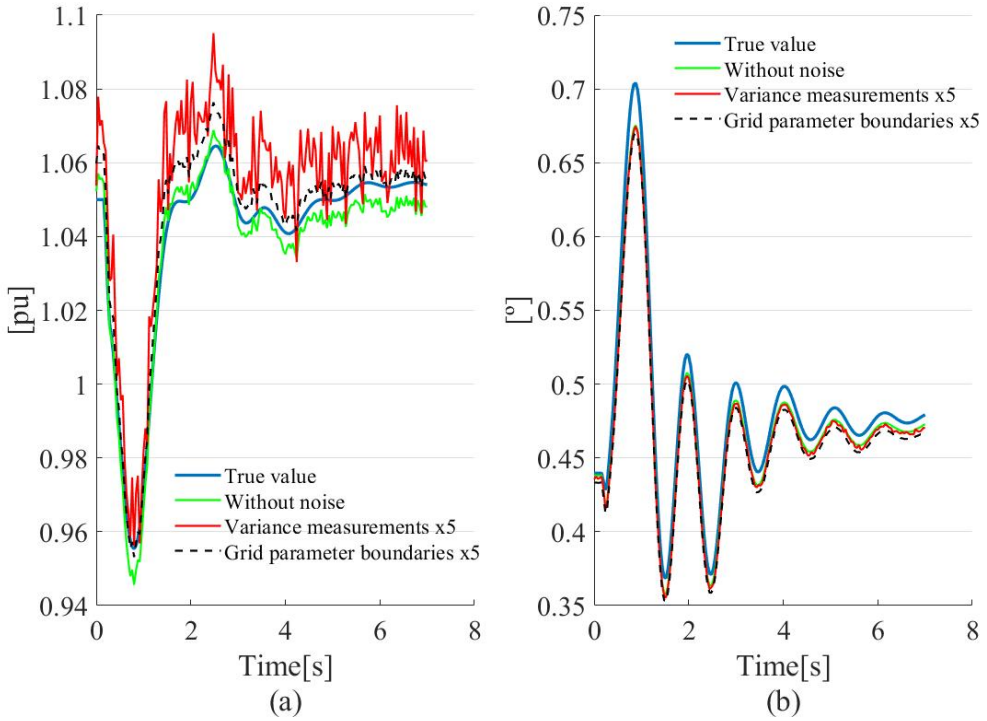

Fig. 5.85 Estimated static state at bus 35 under a generator outage at bus 39 (a) voltage magnitude, (b) voltage angle

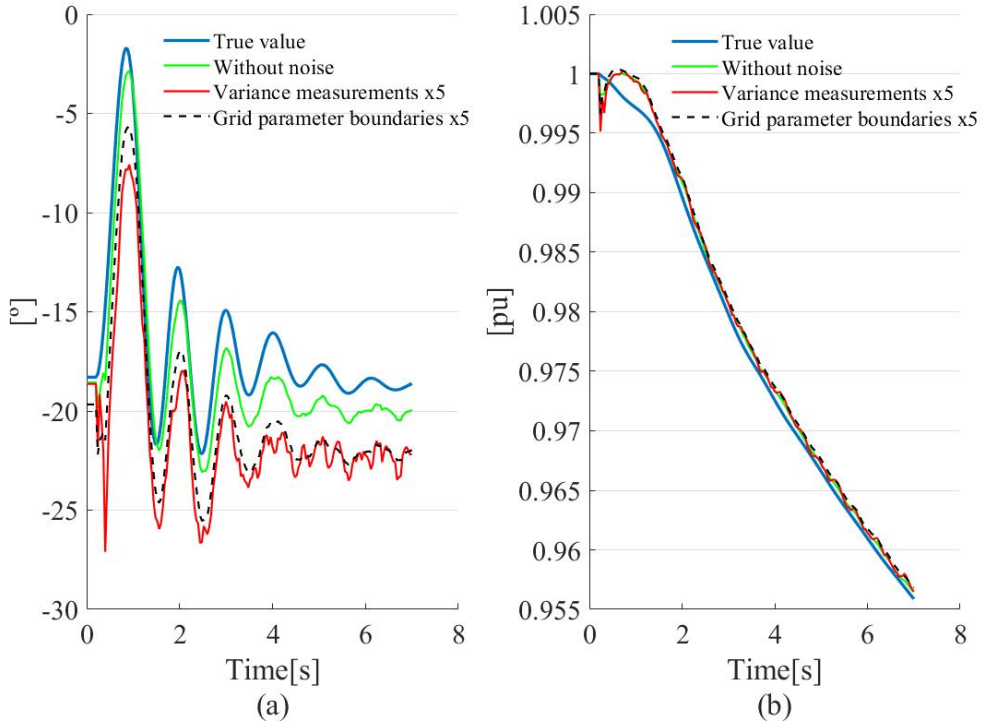

Fig. 5.86 Estimated dynamic state (generator at bus 35) under a generator outage at bus 39 (a) rotor angle of generator, (b) rotation speed of generator

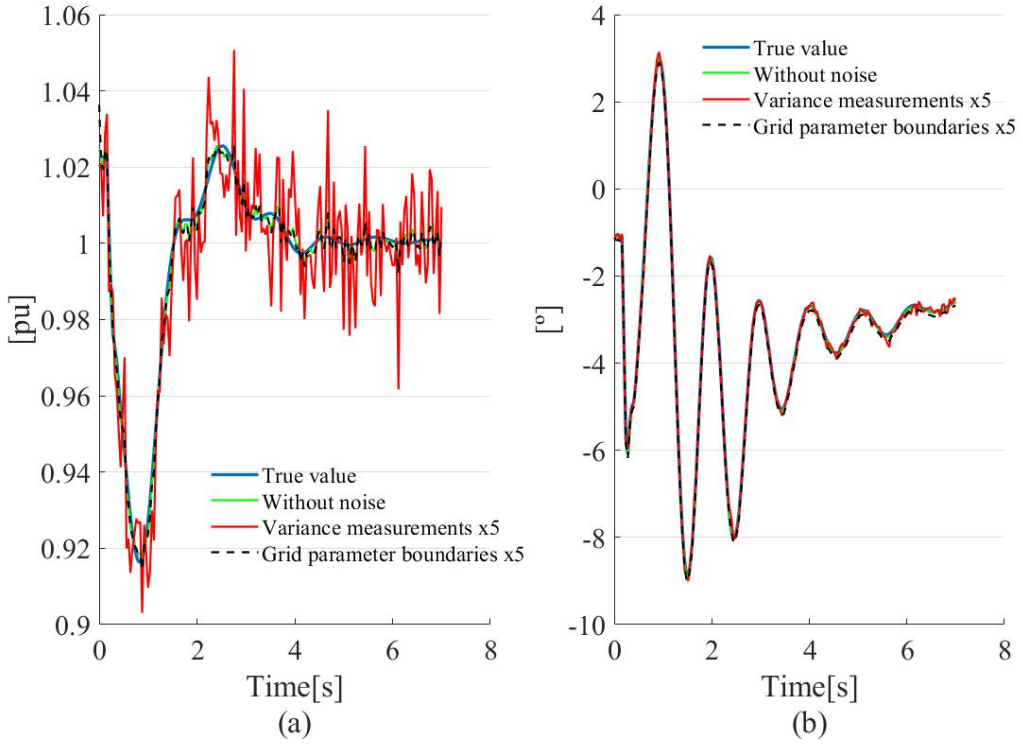

Fig. 5.87 Estimated static state at bus 4 under a generator outage at bus 39 (a) voltage magnitude, (b) voltage angle

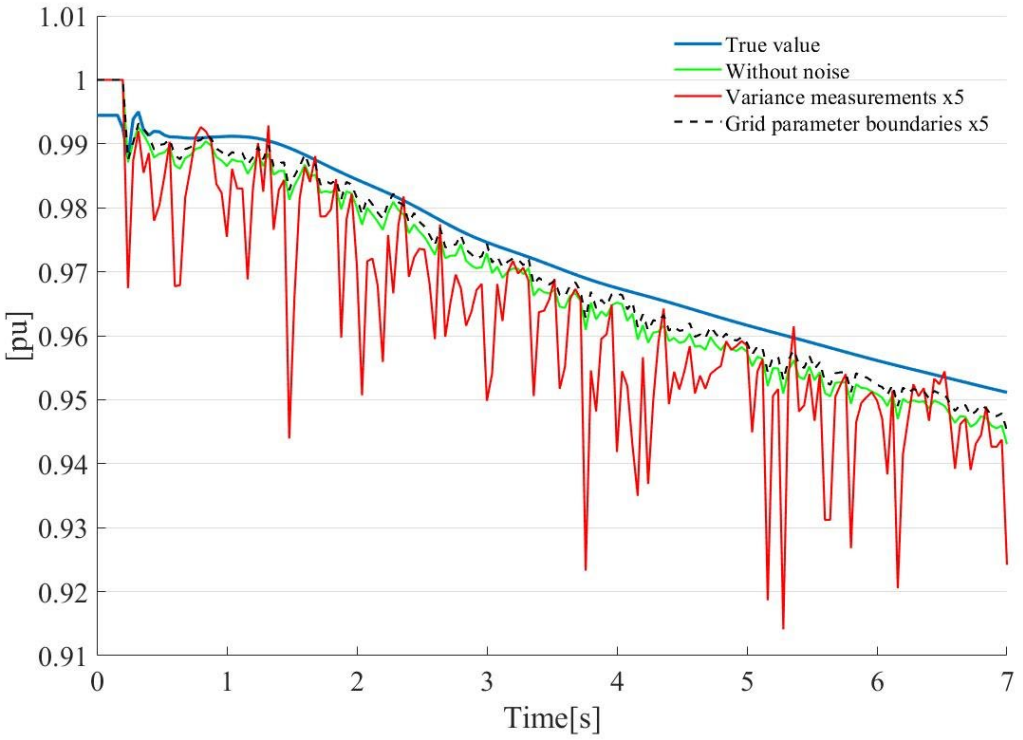

Fig. 5.88 Estimated speed of rotation (motor at bus 4) under a generator outage at bus 39

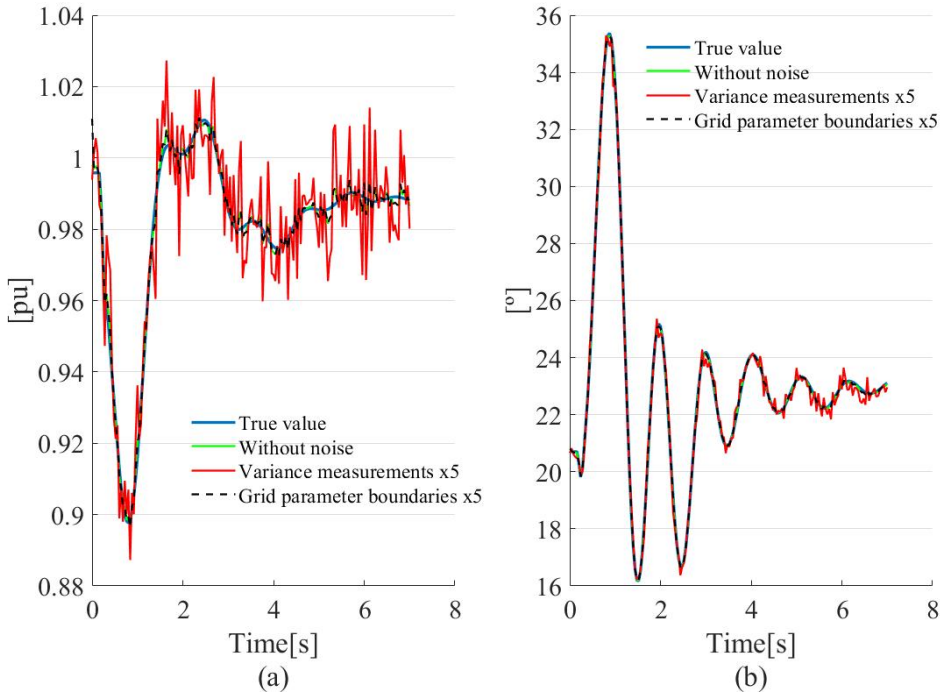

Fig. 5.89 Estimated static state at bus 20 under a generator outage at bus 39 (a) voltage magnitude, (b) voltage angle

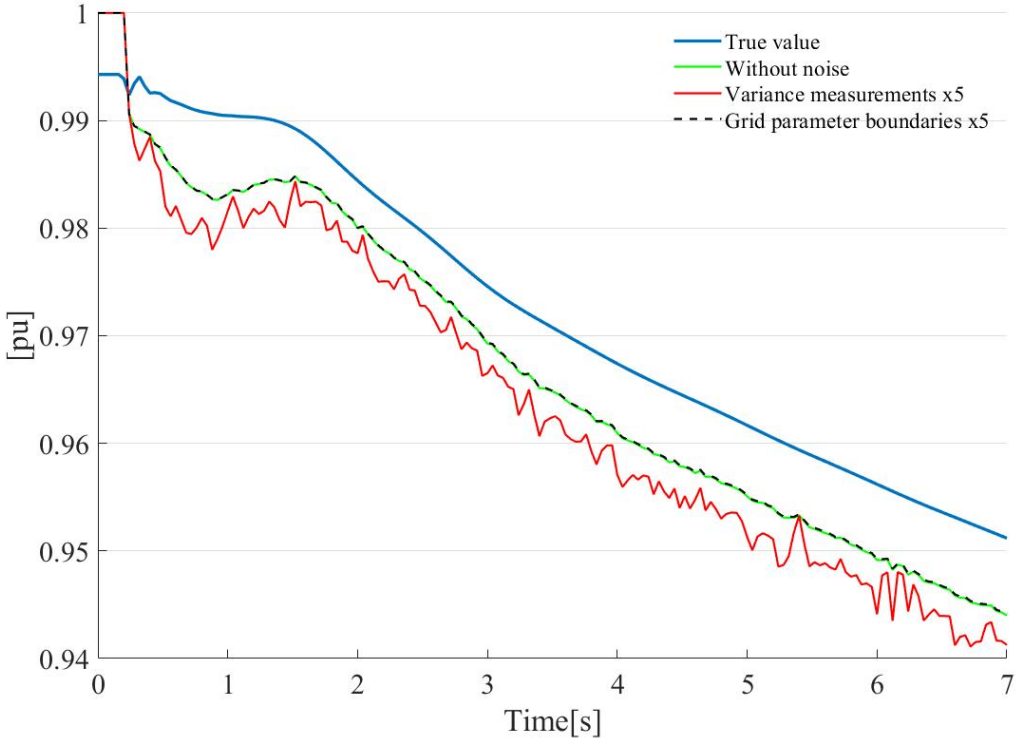

Fig. 5.90 Estimated speed of rotation (motor at bus 20) under a generator outage at bus 39

This page intentionally left blank

6 Conclusion and outlook

Nowadays, power systems are operating close to their capacity limits due to the increasing electricity demand and deregulated electricity market. This situation becomes even worse due to the large-scale integration of distributed energy resources and new technologies in the demand. Such a change results in more uncertainties over the system dynamics leading to new stochastic operating behaviors and dynamics. If the system is subjected to a disturbance under such stressed conditions, instability is more likely to occur.

In order to increase the system stability and security a new monitoring tool capable of informing the system dynamics and enhancing the power system situational awareness is needed. An evolution in this sense will allow carrying out an online dynamic security assessment clearing the way for new and more effective local and global control schemes aimed at improving system stability and security. Nevertheless, the conventional monitoring system of the Energy Management System, i.e. the state estimator, is unable to provide such information, mainly because it depends on data from Remote Terminal Units whose sampling rate is too low to reveal electromechanical dynamic responses. In addition, this data is not well synchronized, thus making global monitoring and controlling tasks unfeasible.

The arrival of PMUs has opened up the possibility for a new dynamic wide-area monitoring function. Their high accuracy, reporting speed and synchronization capacity make them suitable for global monitoring of power system dynamics. In addition, if enough PMUs were deployed, the measured voltage and current phasors could make the entire network observable, thus making state estimation feasible. Issues like low data scan and time skew could be eliminated. Such PMU-only state estimator is linear and can run at PMU speed since no iterations are involved. A truly dynamic state would be available under such conditions. However, the deployment of such a scheme is expensive and far from becoming real in large scale. The major limiting factor for the wide-scale deployment of PMUs is the cost and availability of the associated communication and data storage infrastructure. A more realistic and feasible deployment of PMUs in state estimation is using a hybrid approach that combines synchronized phasor measurements with conventional measurements. That is why, this thesis focuses on the development of a hy-

brid state estimator capable of monitoring power system dynamics related to slow and fast phenomena but without leaving aside the current limitation as regards the reduced amount of PMUs that can be installed in power grids.

6.1 Summary of results

The proposed hybrid state estimator together with the methodology for restoring full PMU observability were tested in the New England benchmark system under several operating conditions. Results are summarized as follows:

- The proposed approach for defining the PMU topology, i.e. the number of units and their location, found for the New England benchmark system that a topology involving only 7 PMUs at buses 2, 6, 9, 14, 20, 23 and 29 allows observing all the coherent areas in voltage magnitude and angle simultaneously with 92.8 % of probability. This amount of units represents a reduction of 40% in the minimum amount of units required to observe the whole system if only PMU measurements were considered. Such a reduction in the amount of necessary equipment will undoubtedly contribute to mitigate the effects of large amounts of additional data coming from PMUs on communication, data storage and processing systems.
- The classifier developed in this thesis for predicting bus voltage coherency is capable of working at high speed with an average computing time of 12 ms per scenario. Additionally, 90% of simulated test scenarios are well predicted considering voltage angle and magnitude simultaneously. It was found that misclassified scenarios are usually scenarios with a low frequency of class in the training sample. This is a clear symptom of class imbalance.
- The accuracy of bus voltage-pseudo measurements is evaluated by using as indicator the mean absolute deviation. Considering the kind of contingency, pseudo-measurements are less accurate in events that result in a system response with a higher degree of nonlinearity, such three-phase faults or generator outages. In such cases, the error remains in most scenarios below 1.5 % and 1º for voltage magnitude and angle respectively. When analyzing scenarios according to the classifier performance, in most cases pseudo-measurement errors from misclassified scenarios show a higher degree of dispersion reaching higher maximum values, which indicates they are less accurate. However, the difference of accuracy between both kind of scenarios is minor.
- The performance of the hybrid estimator has been investigated by means of four case studies. Measurements as well as structure uncertainties have been taken into consideration. What is more, process noises from model inadequacy and linearization are also included.

Computing times in the second phase introduce a delay of order of a few milliseconds. This is because in phase two static and dynamic estimators operate in sequence. This issue is taken into consideration. Next, the results from the four case studies are briefly presented:

i. In the first case the estimator operates following base guidelines as regards initialization, number of bus voltage pseudo-measurements and algorithms involved in the estimation process. The estimator is capable of estimating with accuracy the static and dynamic state, even in those buses where there is no PMU.

   In addition, it was found that the EKF in phase two fails to find a solution in about 11% of the test scenarios. This amount is reduced to 4% if the time delay between static and dynamic estimators in phase two is removed. Most cases with convergence problems imply unstable scenarios.

   The average computing time of 15 ms in phase two makes the proposed scheme capable of keeping up with the PMU update rate and tracking fast dynamics.

   Although a centralized approach like the one proposed in this thesis might not work in large scale power grids due to the higher computing times involved, it can be extended to a decentralized scheme typically employed for monitoring and controlling of large scale systems.

   The performance of the estimator is severely degraded in the presence of gross PMU errors. The main reason behind this issue lies in the lack of a post-estimation bad data processor in phase two.

ii. The second case study introduces a modified version of the hybrid estimator presented in the first case study. This time, an ancillary set of bus voltage pseudo-measurements is included in the formulation. The main goal behind this change aims at improving the estimator performance under the presence of large PMU errors. Results show that, although the inclusion of additional pseudo-measurements in some cases may result in a slight loss of performance, the estimator gains the ability to filter out large errors.

iii. The third case study investigates the effect on the estimator performance of using in phase two a robust WLAV estimator instead the conventional WLS. Both options are evaluated under several operating conditions. First, no gross PMU error is considered. Then, as in the second case study, a malfunction of a PMU is assumed and simulated. The WLS based option is a little more accurate when PMUs work well. This situation is reversed when having a PMU malfunction. Under such a circumstance, the alternative based on the WLAV is more effective than its counterpart to filter out the large error.

iv. In the last case study the robustness of the hybrid state estimator is assessed. Measurement variances and noise boundaries in grid parameters are increased one at a time by a factor of five with respect to their initial values. It was found that higher noise levels lead to an increase in the number of cases where the algorithm does not approach a finite value. A noise level growth in the measurement variance by a factor of 5 increases by about 8% the number of scenarios where the EKF fails to converge. This amount falls to 2% if the additional noise affects only the grid parameters.
As expected, higher noise levels leads to greater estimation errors, either in static or dynamic variables. Nevertheless, the increase in the median of the estimated static and dynamic error is minor. Then, it can be affirmed that the estimator remains robust under such new operating conditions.

6.2 Contributions

The main contributions of this thesis are:

- A data-mining based methodology for defining a PMU topology, i.e. number of units and their location in the grid, so that all the coherent areas are observed with high probability. This is accomplished by means of an offline data-mining based knowledge discovery process which consists of three steps. First, a large set of scenarios that are representative of normal, during fault and post-fault behavior in a broad range of conditions is obtained from numerical simulations. Then, data pre-processing techniques are carried out. The main goal here is to make data analysis and comprehension easier and to obtain data that is suitable for creating good predictive models. Finally, the number of PMUs and their location is determined through an iterative exploring process where candidates topologies are evaluated with respect to bus voltage coherency. Candidates who observe all coherent areas in most scenarios are stored. The final solution will be the best one between all the stored topologies considering the relationship between amount of units and number of observed scenarios.
- A classifier based on the k-nearest neighbors algorithm that is designed offline and forecast coherency at high speed whenever a contingency happens and with that an assignment vector that shows the connection between observed and non-observed buses. The classifier is supported by a classification matrix that is generated offline from the pre-processed data during the definition of the PMU topology.
- A novel methodology for generating bus voltage pseudo-measurements that is based on the concept of bus voltage coherency and allows restoring the PMU observability. Buses that are coherent with each other show similar voltage behavior during a transient event and,

therefore, the behavior in one of these buses represents approximately the behavior of the remaining buses. The proposed approach relies on this idea to define bus voltage pseudo-measurements by means of voltage magnitudes and angles reported by PMUs, the pre-contingency system loading condition and the assignment vector predicted by the classifier proposed in this thesis.

- A novel hybrid state estimator that is capable of accurately monitoring the power system dynamics associated to fast and slow phenomena using a reduced number of PMUs. The proposed can estimate with accuracy even the dynamic states of units which are not equipped with a PMU at their terminal. To circumvent the problem of inadequate PMU observability, it uses the information provided by the methodology for generating bus voltage pseudo-measurements.

  It consists of two phases: when the power system is in stationary regime phase one takes place and the static state is estimated at SCADA speed by a WLS based estimator. Conventional as well as PMU measurements are used together in one single estimation run. Once the estimation is completed, a post-estimation method based on residuals is carried out in order to detect and identify bad measurements. Phase two is activated when a physical disturbance happens and the system is in a transient regime. In this case, two estimators work in sequence making use of all the information provided by the PMUs. First, a static state estimator is used to estimate bus voltages as soon as the PMU measurement set arrives. Since a post-estimation bad data processing technique is unfeasible at fast scan rates, the static estimator in phase two uses the information provided by the bad data processor from phase one to adjust the measurement weights. Additionally, in order to make phase two robust against PMU gross errors, a WLAV based linear estimator is used. Finally, dynamic variables such as rotor angle and speed of generators and speed of motors are estimated by means of an EKF based dynamic state estimator.

6.3 Future work

For follow-up research, the following extensions and directions of additional research have been found to be promising:

- The proposed approach has been evaluated in a small benchmark system. Its implementation in large scale interconnected power grids represents a challenge due to the increased size and complexity of the system model and measurement volume. Nevertheless it can be adapted to work in a decentralized scheme. In addition, high performance computing and parallel computing techniques could be employed to make possible its application in large

systems. In this regard, delays associated to communication systems, which have not been considered in this thesis, should also be taken into account.

- The performance of the EKF is degraded when the system behavior becomes unstable. The strongly nonlinear behavior of the system could be the reason behind the convergence issues. However, a detailed analysis should be perform on this point with the aim of discarding other causes. More robust options like the Unscented Kalman Filter (UKF) and the Particle Filter (PF) to cite a few should be tested and evaluated. Data driven methods represent another direction for further research in this sense.
- In the proposed approach, measurement and process errors were assumed to have zero mean with known covariance matrices and to follow Gaussian distribution. Algorithms used by the hybrid estimator to estimate the state perform well under such conditions. However, in the reality, for most practical power systems, these assumptions do not hold true. Even though additional source of errors, like model inadequacy and noises in grid parameters, have been taken into account in this thesis, the Kalman filter performance could be greatly degraded in the presence of non-Gaussian system uncertainties. Therefore, future research about enhancing the robustness of the dynamic estimator in the presence of non-Gaussian noise is required.
- In the proposed hybrid state estimator mechanical power and field voltage of the synchronous machine are assumed to be measurable when estimating the dynamic state. Nevertheless, in reality these signals may not be accessible. Alternative approaches, like the Extended Kalman Filter with unknown inputs (EKF-UI) developed in the field of civil engineering for earthquake damage estimation studies must be considered and analyzed. Another option with potential in this regard entails the inclusion of the governor and the AVR models in the generator formulation. By doing this the unknown inputs become internal variables and reference quantities will be the new inputs.
- In this work, dynamic state estimation in motors demands knowing in detail the load composition at each bus, i.e. the static-dynamic load ratio. In this sense, load composition needs to be reliably estimated. A proposal with potential in this regard may involve the use of artificial intelligence together with dynamic post-contingency data from the hybrid state estimator in order to approximate the bus load composition. Such an approach would be of great interest for load modelling and stability analysis purposes.

References

[1] J. -P. Heckel and C. Becker.: Investigation of the Voltage Stability in the Integrated Energy System of Northern Germany. In NEIS 2020; Conference on Sustainable Energy Supply and Energy Storage Systems, 2020, pp. 1–6.

[2] S. Birk, C. Brosig, E. Waffenschmidt, and T. Schneiders.: Influence of Sector Coupling in Future Inner City Low Voltage Grids. In 2018 7th International Energy and Sustainability Conference (IESC), 2018, pp. 1–8.

[3] C. Liu et al.: A systematic approach for dynamic security assessment and the corresponding preventive control scheme based on decision trees. IEEE Trans Power Syst, vol. 29, no. 2, pp. 717–730, 2014.

[4] Y. Zhang, Y. Xu, Z. Y. Dong, Z. Xu, and K. P. Wong.: Intelligent Early Warning of Power System Dynamic Insecurity Risk: Toward Optimal Accuracy-Earliness Tradeoff. IEEE Transactions on Industrial Informatics, vol. 13, no. 5, pp. 2544–2554, 2017.

[5] G. Andersson et al.: Causes of the 2003 major grid blackouts in North America Europe, and recommended means to improve system dynamic performance," IEEE Trans Power Syst, vol. 20, no. 4, pp. 1922–1928, 2005.

[6] E. Ghahremani and I. Kamwa.: Online state estimation of a synchronous generator using unscented Kalman filter from phasor measurements units. IEEE Trans Energy Convers, vol. 26, no. 4, pp. 1099–1108, 2011.

[7] A. K. Singh and B. C. Pal.: Decentralized dynamic state estimation in power systems using unscented transformation. IEEE Trans Power Syst, vol. 29, no. 2, pp. 794–804, 2014.

[8] J. Zhao et al.: Roles of Dynamic State Estimation in Power System Modeling, Monitoring and Operation. IEEE Trans Power Syst, p. 1, 2020.

[9] 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.

[10] N. Zhou, D. Meng, and S. Lu.: Estimation of the dynamic states of synchronous machines using an extended particle filter. IEEE Trans Power Syst, vol. 28, no. 4, pp. 4152–4161, 2013.

[11] A. G. Phadke, J. S. Thorp, R. F. Nuqui, and M. Zhou, Eds.: Recent developments in state estimation with phasor measurements. In: 2009 IEEE/PES Power Systems Conference and Exposition, 2009.

[12] Z. Jin, P. Wall, Y. Chen, J. Yu, S. Chakrabarti, and V. Terzija.: Analysis of Hybrid State Estimators: Accuracy and Convergence of Estimator Formulations. IEEE Trans Power Syst, vol. 34, no. 4, pp. 2565–2576, 2019.

[13] G. N. Korres and N. M. Manousakis.: State estimation and bad data processing for systems including PMU and SCADA measurements. Electr Power Syst Res, vol. 81, no. 7, pp. 1514–1524, 2011.

[14] M. Zhou, V. A. Centeno, J. S. Thorp, and A. G. Phadke.: An alternative for including phasor measurements in state estimators. IEEE Trans Power Syst, vol. 21, no. 4, pp. 1930–1937, 2006.

[15] E. Farantatos, G. K. Stefopoulos, G. J. Cokkinides, and A. P. Meliopoulos.: PMU-based dynamic state estimation for electric power systems. In: 2009 IEEE Power and Energy Society General Meeting. Calgary, AB, Canada, July 2009.

[16] DOE, Factors Affecting PMU Installation Costs. [Online]. Available: https://www.energy.gov/sites/prod/files/2014/11/f19/SG-PMU-cost-study-Oct2014.pdf

[17] Y. F. Huang, S. Werner, J. Huang, N. Kashyap, and V. Gupta.: State estimation in electric power grids: Meeting new challenges presented by the requirements of the future grid. IEEE Signal Process Mag, vol. 29, no. 5, pp. 33–43, 2012.

[18] M. Asprou, S. Chakrabarti, and E. Kyriakides.: A two-stage state estimator for dynamic monitoring of power systems. IEEE Syst. J., vol. 11, no. 3, pp. 1767–1776, 2017.

[19] A. Monticelli, State Estimation in Electric Power Systems - A Generalized Approach. Springer, Berlin (2012).

[20] P. Kundur et al.: Definition and classification of power system stability. IEEE Trans Power Syst, vol. 19, no. 3, pp. 1387–1401, 2004.

[21] K. Morison, L. Wang, and P. Kundur.: Power system security assessment. IEEE Power and Energy Magazine, vol. 2, no. 5, pp. 30–39, 2004.

[22] M. Shahidehpour, W. F. Tinney, and Y. Fu.: Impact of security on power systems operation. Proc. IEEE, vol. 93, no. 11, pp. 2013–2025, 2005.

[23] Mohammad Shahidehpour, Handbook of electric power system dynamics. IEEE Press, (2013).

[24] P. Kundur, Power System Stability and Control. McGraw-Hill, Inc., 1994.

[25] A. Abur and A. Gomez-Exposito, Power System State Estimation: Theory and Implementation. Marcel Dekker, Inc., New York (2004).

[26] Do Coutto Filho, M. B. and J. C. Stacchini de Souza.: Forecasting-aided state estimation - Part I: Panorama. IEEE Trans Power Syst, vol. 24, no. 4, pp. 1667–1677, 2009.

[27] J. Zhao et al.: Power System Dynamic State Estimation: Motivations, Definitions, Methodologies, and Future Work. IEEE Trans Power Syst, vol. 34, no. 4, pp. 3188–3198, 2019.

[28] M. Göl and A. Abur.: LAV based robust state estimation for systems measured by PMUs. IEEE Trans. Smart Grid, vol. 5, no. 4, pp. 1808–1814, 2014.

[29] C. Xu and A. Abur.: A fast and robust linear state estimator for very large scale interconnected power grids. IEEE Trans. Smart Grid, vol. 9, no. 5, pp. 4975–4982, 2018.

[30] H. Ma, L. Yan, Y. Xia, and M. Fu, Kalman Filtering and Information Fusion. Science Press Beijing, 2020.

[31] D. Simon, Optimal state estimation: Kalman, H, and nonlinear approaches. Wiley-Interscience, 2006.

[32] A. Monticelli.: Electric power system state estimation. Proc. IEEE, vol. 88, no. 2, pp. 262–282, 2000.

[33] M. Panteli and D. S. Kirschen.: Situation awareness in power systems: Theory, challenges and applications. Electr Power Syst Res, vol. 122, pp. 140–151, 2015.

[34] X. Qing, H. R. Karimi, Y. Niu, and X. Wang.: Decentralized unscented Kalman filter based on a consensus algorithm for multi-area dynamic state estimation in power systems. Int J Electr Power Energy Syst, vol. 65, pp. 26–33, 2015.

[35] G. N. Korres, A. Tzavellas, and E. Galinas.: A distributed implementation of multi-area power system state estimation on a cluster of computers. Electr Power Syst Res, vol. 102, pp. 20–32, 2013.

[36] A. Gómez-Expósito, A. Abur, P. Rousseaux, A. de la Villa Jaen, C. Gómez-Quiles.: On the use of PMUs in power system state estimation. In: 17th Power Systems Computation Conference. Stockholm, Sweden (2011).

[37] A. G. Phadke and T. Bi.: Phasor measurement units, WAMS, and their applications in protection and control of power systems. Journal of Modern Power Systems and Clean Energy, vol. 6, no. 4, pp. 619–629, 2018.

[38] V. Kirincic, S. Skok, and V. Terzija.: A two-step hybrid power system state estimator. Int. Trans. Elecr. Energy Sys., vol. 25, no. 7, pp. 1158–1172, 2015.

[39] G. Valverde, S. Chakrabarti, E. Kyriakides, and V. Terzija.: A constrained formulation for hybrid state estimation. IEEE Trans Power Syst, vol. 26, no. 3, pp. 1102–1109, 2011.

[40] L. Zhang, A. Bose, A. Jampala, V. Madani, and J. Giri.: Design, Testing, and Implementation of a Linear State Estimator in a Real Power System. IEEE Trans. Smart Grid, 2016.

[41] C. Rakpenthai, S. Uatrongjit, and S. Premrudeeprechacharn.: State estimation of power system considering network parameter uncertainty based on parametric interval linear systems. IEEE Trans Power Syst, vol. 27, no. 1, pp. 305–313, 2012.

[42] G. N. Korres and N. M. Manousakis.: State estimation and observability analysis for phasor measurement unit measured systems. IET Gener. Transm. Distrib., vol. 6, no. 9, pp. 902–913, 2012.

[43] F. Aminifar, M. Shahidehpour, M. Fotuhi-Firuzabad, and S. Kamalinia.: Power system dynamic state estimation with synchronized phasor measurements. IEEE Trans. Instrum. Meas., vol. 63, no. 2, pp. 352–363, 2014.

[44] A. Del Angel, P. Geurts, D. Ernst, M. Glavic, and L. Wehenkel.: Estimation of rotor angles of synchronous machines using artificial neural networks and local PMU-based quantities. Neurocomputing, vol. 70, 16-18, pp. 2668–2678, 2007.

[45] L. Fan and Y. Wehbe.: Extended Kalman filtering based real-time dynamic state and parameter estimation using PMU data. Electr Power Syst Res, vol. 103, pp. 168–177, 2013.

[46] E. Ghahremani and I. Kamwa.: Dynamic state estimation in power system by applying the extended kalman filter with unknown inputs to phasor measurements. IEEE Trans Power Syst, vol. 26, no. 4, pp. 2556–2566, 2011.

[47] Z. Huang, K. Schneider, and J. Nieplocha.: Feasibility studies of applying Kalman Filter techniques to power system dynamic state estimation. In: International Power Engineering Conference (IPEC 2007). Singapore (2007).

[48] H. Tebianian and B. Jeyasurya.: Dynamic state estimation in power systems: Modeling, and challenges. Electr Power Syst Res, vol. 121, pp. 109–114, 2015.

[49] A. Vahidnia, G. Ledwich, E. Palmer, and A. Ghosh.: Dynamic equivalent state estimation for multi-area power systems with synchronized phasor measurement units. Electr Power Syst Res, vol. 96, pp. 170–176, 2013.

[50] J. Zhang, G. Welch, G. Bishop, and Z. Huang.: A two-stage kalman filter approach for robust and Real-Time power system State Estimation. IEEE Trans. Sustainable Energy, vol. 5, no. 2, pp. 629–636, 2014.

[51] E. Ghahremani and I. Kamwa.: Local and Wide-Area PMU-Based Decentralized Dynamic State Estimation in Multi-Machine Power Systems. IEEE Trans Power Syst, 2016.

[52] M. Rostami and S. Lotfifard.: Distributed dynamic state estimation of power systems. IEEE Transactions on Industrial Informatics, vol. 14, no. 8, pp. 3395–3404, 2018.

[53] H. Karimipour and V. Dinavahi.: Parallel relaxation-based joint dynamic state estimation of large-scale power systems. IET Gener. Transm. Distrib., vol. 10, no. 2, pp. 452–459, 2016.

[54] I. Kamwa, B. Baraboi, and R. Wamkeue.: Sensorless ANN-based speed estimation of synchronous generators: Improved performance through physically motivated pre-filters. In: The 2006 IEEE International Joint Conference on Neural Network Proceedings. Vancouver, BC, Canada (2006).

[55] J. Zhao and L. Mili.: Power System Robust Decentralized Dynamic State Estimation Based on Multiple Hypothesis Testing. IEEE Trans Power Syst, vol. 33, no. 4, pp. 4553–4562, 2018.

[56] S. Chakrabarti, E. Kyriakides, G. Ledwich, and A. Ghosh.: Inclusion of PMU current phasor measurements in a power system state estimator. IET Gener. Transm. Distrib., vol. 4, no. 10, pp. 1104–1115, 2010.

[57] T. S. Bi, X. H. Qin, and Q. X. Yang.: A novel hybrid state estimator for including synchronized phasor measurements. Electr Power Syst Res, vol. 78, no. 8, pp. 1343–1352, 2008.

[58] R. F. Nuqui, A. G. Phadke.: Hybrid linear state estimation utilizing synchronized phasor measurements. In: 2007 IEEE Lausanne POWERTECH. Lausanne, Switzerland, July 2007.

[59] Y. Guo, W. Wu, B. Zhang, and H. Sun.: A distributed state estimation method for power systems incorporating linear and nonlinear models. Int J Electr Power Energy Syst, vol. 64, pp. 608–616, 2015.

[60] H. Karimipour and V. Dinavahi.: Parallel domain decomposition based distributed state estimation for large-scale power systems. IEEE Trans Ind Appl, vol. 2015, 2015.

[61] X. Yang, X. P. Zhang, and S. Zhou.: Coordinated algorithms for distributed state estimation with synchronized phasor measurements. Appl. Energy, vol. 96, pp. 253–260, 2012.

[62] L. Zhao and A. Abur.: Multiarea state estimation using synchronized phasor measurements. IEEE Trans Power Syst, vol. 20, no. 2, pp. 611–617, 2005.

[63] W. Jiang, V. Vittal, and G. T. Heydt.: A distributed state estimator utilizing synchronized phasor measurements. IEEE Trans Power Syst, vol. 22, no. 2, pp. 563–571, 2007.

[64] W. Jiang, V. Vittal, and G. T. Heydt.: Diakoptic state estimation using phasor measurement units. IEEE Trans Power Syst, vol. 23, no. 4, pp. 1580–1589, 2008.

[65] J. Zhao et al..: Power system real-time monitoring by using PMU-based robust state estimation method. IEEE Trans. Smart Grid, vol. 7, no. 1, pp. 300–309, 2016.

[66] H. Karimipour and V. Dinavahi.: Extended Kalman Filter-Based Parallel Dynamic State Estimation. IEEE Trans. Smart Grid, vol. 6, no. 3, pp. 1539–1549, 2015.

[67] A. Sharma, S. C. Srivastava, and S. Chakrabarti.: Multi-agent-based dynamic state estimator for multi-area power system. IET Gener. Transm. Distrib., vol. 10, no. 1, pp. 131–141, 2016.

[68] M. Gol and A. Abur.: A hybrid state estimator for systems with limited number of PMUs. IEEE Trans Power Syst, vol. 30, no. 3, pp. 1511–1517, 2015.

[69] J. Cepeda.: Evaluación de la vulnerabilidad del sistema electrico de potencia en tiempo real usando tecnología de medición sincrofasorial. Revista Técnica Energía, 2014.

[70] D. T. Larose and C. D. Larose, Discovering Knowledge in Data: An Introduction to Data Mining: Second Edition, 2014.

[71] J. Han, M. Kamber, J. Pei, Data Mining Concepts and Techniques. The Morgan Kaufmann Series in Data Management Systems, 2012.

[72] X. Wu et al..: Top 10 algorithms in data mining. Knowledge and Information Systems, vol. 14, no. 1, pp. 1–37, 2008.

[73] J. C. Cepeda, J. L. Rueda, I. Erlich, and D. G. Colome.: Probabilistic approach-based PMU placement for real-time power system vulnerability assessment. IEEE PES Innovative Smart Grid Technologies Conference Europe. Berlin, Germany, October 2012.

[74] B. K. Saha Roy, A. K. Sinha, and A. K. Pradhan.: An optimal PMU placement technique for power system observability. Int J Electr Power Energy Syst, vol. 42, no. 1, pp. 71–77, 2012.

[75] A. Ahmadi, Y. Alinejad-Beromi, and M. Moradi.: Optimal PMU placement for power system observability using binary particle swarm optimization and considering measurement redundancy. Expert Systems with Applications, vol. 38, no. 6, pp. 7263–7269, 2011.

[76] R. Dubey.: A novel technique for optimal PMU placement with effects on dynamic state estimation. In: *2016 2nd International Conference on Control, Instrumentation, Energy & Communication (CIEC)*. Kolkata, India (2016).

[77] N. H. Abbasy and H. M. Ismail.: A unified approach for the optimal PMU location for power system state estimation. IEEE Trans Power Syst, vol. 24, no. 2, pp. 806–813, 2009.

[78] A. Gómez-Expósito, A. J. Conejo, and C. Cañizares, Electric energy systems: Analysis and operation, 2016.

[79] Alexis Sarda Espinosa.: Comparing Time-Series Clustering Algorithms in R Using the dtwclust Package. http://cran.r-project.org/web/packages/dtwclust/vignettes/dtwclust.pdf, 2017.

[80] D. J. Ketchen Jr. and C. L. Shook.: The application of cluster analysis in strategic management research: An analysis and critique. Strategic Management Journal, vol. 17, no. 6, pp. 441–458, 1996.

[81] S. Wang and X. Yao.: Multiclass imbalance problems: Analysis and potential solutions. IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, vol. 42, no. 4, pp. 1119–1130, 2012.

[82] E. Fernandes and A. de Carvalho.: Evolutionary inversion of class distribution in overlapping areas for multi-class imbalanced learning. Information Sciences, vol. 494, pp. 141–154, 2019.

[83] J. A. Saez, M. Galar, and B. Krawczyk.: Addressing the Overlapping Data Problem in Classification Using the One-vs-One Decomposition Strategy. IEEE Access, vol. 7, pp. 83396–83411, 2019.

[84] K. Chen, C. Huang, and J. He.: Fault detection, classification and location for transmission lines and distribution systems: A review on the methods. High Voltage, vol. 1, no. 1, pp. 25–33, 2016.

[85] Z. He, L. Fu, S. Lin, and Z. Bo.: Fault detection and classification in EHV transmission line based on wavelet singular entropy. IEEE Trans Power Delivery, vol. 25, no. 4, pp. 2156–2163, 2010.

[86] K. M. Silva, B. A. Souza, and N. S. D. Brito.: Fault detection and classification in transmission lines based on wavelet transform and ANN. IEEE Trans Power Delivery, vol. 21, no. 4, pp. 2058–2063, 2006.

[87] A. Rouhani and A. Abur.: Linear phasor estimator assisted dynamic state estimation. IEEE Trans. Smart Grid, vol. 9, no. 1, pp. 211–219, 2018.

[88] M. K. Çelik and A. Abur.: Use of scaling in wlav estimation of power system states. IEEE Trans Power Syst, vol. 7, no. 2, pp. 684–692, 1992.

[89] F. S. Hillier and G. J. Lieberman, Introduction to Operations Research: McGraw-Hill Higher Education, 2010.

[90] Y. Dong, X. Xie, K. Wang, B. Zhou, and Q. Jiang.: An Emergency-Demand-Response Based under Speed Load Shedding Scheme to Improve Short-Term Voltage Stability. IEEE Trans Power Syst, vol. 32, no. 5, pp. 3726–3735, 2017.

[91] Y. Huang, G. V. Reklaitis, and V. Venkatasubramanian.: A heuristic extended Kalman filter based estimator for fault identification in a fluid catalytic cracking unit. Industrial and Engineering Chemistry Research, vol. 42, no. 14, pp. 3361–3371, 2003.

[92] PowerFactory.: Technical reference: General load. Heinrich-Hertz-Straße 9. Gomaringen, Germany, 2020.

[93] PowerFactory.: Technical reference: Asynchronous Machine. Heinrich-Hertz-Straße 9. Gomaringen, Germany, 2020.

[94] PowerFactory 2020.: Technical reference: Synchronous machine. Heinrich-Hertz-Straße 9. Gomaringen, Germany, 2020.

Appendix A. Test system data

This appendix contains details about the IEEE New England benchmark system used in this thesis for developing and testing the proposed hybrid state estimator and the methodology for generating voltage pseudo-measurements. The parameters were extracted from [69] and are used to perform simulations in stationary and transient regime.

Table A.1. Buses

Bus	Type	Voltage	LOAD				GEN
		(kV)	MW	MVAr	Type	Motor	MW
01	PQ	345	-	-	-	-	-
02	PQ	345	-	-	-	-	-
03	PQ	345	322	2.4	C	-	-
04	PQ	345	500.0	184.0	I	40%	-
05	PQ	345	-	-	-	-	-
06	PQ	345	-	-	-	-	-
07	PQ	345	233.8	84.0	C	-	-
08	PQ	345	522.0	176.0	R	25%	-
09	PQ	345	-	-	-	-	-
10	PQ	345	-	-	-	-	-
11	PQ	345	-	-	-	-	-
12	PQ	69	7.5	88.0	I	-	-
13	PQ	345	-	-	-	-	-
14	PQ	345	-	-	-	-	-
15	PQ	345	320.0	153.0	C	30%	-
16	PQ	345	329.0	32.3	R	-	-
17	PQ	345	-	-	-	-	-
18	PQ	345	158.0	30.0	R	-	-
19	PQ	345	-	-	-	-	-
20	PQ	69	628.0	103.0	R	25%	-
21	PQ	345	274.0	115.0	C	-	-
22	PQ	345	-	-	-	-	-
23	PQ	345	247.5	84.6	I	-	-
24	PQ	345	308.6	92.2	I	40%	-
25	PQ	345	224.0	47.2	R	-	-
26	PQ	345	139.0	17.0	C	-	-
27	PQ	345	281.0	75.5	R	-	-
28	PQ	345	206.0	27.6	R	25%	-
29	PQ	345	283.5	26.9	I	-	-
30	PV	13.8	-	-	-	-	250
31	SL	13.8	9.2	4.6	I	-	-
32	PV	13.8	-	-	-	-	650

Table A.1. Continued*

Bus	Type	Voltage (kV)	LOAD MW	LOAD MVAr	LOAD Type	LOAD Motor	GEN MW
33	PV	13.8	-	-	-	-	632
34	PV	13.8	-	-	-	-	508
35	PV	13.8	-	-	-	-	650
36	PV	13.8	-	-	-	-	560
37	PV	13.8	-	-	-	-	540
38	PV	13.8	-	-	-	-	830
39	PV	345	1104.0	250.0		-	1000

Table A.2. Lines*

Bus i	Bus j	Resistance	Reactance	Susceptance	Rated Power (MVA)
01	02	0.0035	0.0411	0.6987	1,000
01	39	0.0010	0.0250	0.7500	1,000
02	03	0.0013	0.0151	0.2572	1,200
02	25	0.0070	0.0086	0.1460	1,000
03	04	0.0013	0.0213	0.2214	1,000
03	18	0.0011	0.0133	0.2138	1,000
04	05	0.0008	0.0128	0.1342	1,000
04	14	0.0008	0.0129	0.1382	1,000
05	06	0.0002	0.0026	0.0434	1,000
05	08	0.0008	0.0112	0.1476	1,000
06	07	0.0006	0.0092	0.1130	1,000
06	11	0.0007	0.0082	0.1389	1,000
07	08	0.0004	0.0046	0.0780	1,000
08	09	0.0023	0.0363	0.3804	1,000
09	39	0.0010	0.0250	1.2000	1,000
10	11	0.0004	0.0043	0.0729	1,000
10	13	0.0004	0.0043	0.0729	1,000
13	14	0.0009	0.0101	0.1723	1,000
14	15	0.0018	0.0217	0.3660	1,000
15	16	0.0009	0.0094	0.1710	1,200
16	17	0.0007	0.0089	0.1342	1,200
16	19**	0.0032	0.0390	0.1520	1,200
16	21	0.0008	0.0135	0.2548	1,000
16	24	0.0003	0.0059	0.0680	1,000
17	18	0.0007	0.0082	0.1319	1,000
17	27	0.0013	0.0173	0.3216	1,000
21	22	0.0008	0.0140	0.2565	1,200
22	23	0.0006	0.0096	0.1846	1,000
23	24	0.0022	0.0350	0.3610	1,000
25	26	0.0032	0.0323	0.5130	1,000

Table A.2. Continued*

Bus i	Bus j	Resistance	Reactance	Susceptance	Rated Power (MVA)
26	27	0.0014	0.0147	0.2396	1,000
26	28	0.0043	0.0474	0.7802	1,000
26	29	0.0057	0.0625	1.0290	1,000
28	29	0.0014	0.0151	0.2490	1,000

*Parameters in pu (100 MVA base)
**Double circuit

Table A.3. Transformers*

Bus i	Bus j	Resistance	Reactance	Transformer Ratio	Rated Power (MVA)
02	30	0.0000	0.0181	1.0250	1,000
06	31	0.0000	0.0250	1.0700	1,200
10	32	0.0000	0.0200	1.0700	1,200
12	11	0.0016	0.0435	1.0060	1,000
12	13	0.0016	0.0435	1.0060	1,000
19	20	0.0007	0.0138	1.0060	1,000
19	33	0.0007	0.0142	1.0700	1,000
20	34	0.0009	0.0180	1.0090	1,000
22	35	0.0000	0.0143	1.0250	1,200
23	36	0.0005	0.0272	1.0000	1,000
25	37	0.0006	0.0232	1.0250	1,000
29	38	0.0008	0.0156	1.0250	1,200

*Parameters in pu (100 MVA base)

Load models

Three kind of consumers were considered by load modeling: Residential, Industrial and Commercial. Their behavior is represented using a combination of static load and dynamic loads at 50% each as shown in [92]. The dynamic load is represented using the nonlinear voltage and linear frequency dependence model. Table A.4 shows the parameters used in such a model [69].

Table A.4. Parameters (dynamic load)

Kind of consumer	Frequency dependence		Voltage dependence			
	kpf	kqf	aP	e_aP	aQ	e_aQ
Commercial	1.3	-1.9	1.0	0.7	1.0	2.5
Industrial	2.6	1.6	1.0	0.1	1.0	0.6
Residential	0.7	-2.3	1.0	1.2	1.0	2.8

In addition, it is considered that a percentage of the load is composed by motors. The standard seventh-order model presented in [93] is used. The motor parameters are given in Table A.5 [69].

Table A.5. Motors

Parameter	Value
Rated power (kVA)	400
No. of pole pairs	4
Rotor	Double cage
Stator resistance	0.0358222
Stator reactance	0.01
Magnitude reactance	1.554626
Operating rotor resistance	0.01080755
Operating rotor reactance	0.4210244
Starting rotor resistance	0.06276684
Starting rotor reactance	0.2291248
Inertia ($kg.m^2$)	21.62247

*Parameters without unit are in pu (the base is the unit nominal power and voltage)

Generator models

The dynamic behavior of generators is represented with the standard sixth-order model of the synchronous machine as shown in [94]. In addition, classic models of AVR and GOV have been considered. The data associated with the parameters is given below [69].

Table A.6. Generators

Parameter \ Unit	01	02	03	04	05	06	07	08	09	10
No. of units	10	20	20	20	20	20	20	20	20	20
Rated power (MVA)	114.1	49.25	40.75	38.3	32.7	43.1	40.05	37.7	51.75	26.1
Rated voltage (kV)	345	13.8	13.8	13.8	13.8	13.8	13.8	13.8	13.8	13.8
Power factor	0.876	0.964	0.920	0.914	0.917	0.928	0.936	0.928	0.918	0.958
Pmax (MW)	100	47.5	37.5	35.0	30.0	40.0	37.5	35.0	47.5	25.0
Pmin (MW)	30.0	7.5	2.5	6.0	7.5	8.0	3.5	5.0	9.0	4.0
Qmax (MVAr)	55.0	13.0	16.0	15.5	13.0	16.0	14.	14.0	20.5	7.5
Qmin (MVAr)	-40.0	-10.0	-12.5	-12.5	-10.0	-12.5	-10.0	-10.0	-15.0	-5.0
H (s)	43.82	3.08	4.39	3.73	3.98	4.04	3.30	3.22	3.33	8.05
R_a	0	0	0	0	0	0	0	0	0	0
X_d	0.228	2.906	2.033	2.007	2.747	2.189	2.363	2.187	2.180	0.522
X_q	0.217	2.778	1.932	1.976	2.616	2.077	2.339	2.111	2.122	0.360
X'_d	0.068	0.687	0.433	0.334	0.523	0.431	0.392	0.430	0.590	0.162
X'_q	0.091	1.675	0.714	1.272	1.086	0.702	1.490	0.687	0.608	0.042
X''_d	0.055	0.512	0.310	0.245	0.262	0.302	0.304	0.324	0.445	0.115
X''_q	0.055	0.512	0.310	0.245	0.275	0.302	0.304	0.324	0.445	0.151
X_l	0.034	0.345	0.248	0.226	0.229	0.193	0.258	0.211	0.308	0.065
T'_{d0} (s)	7.00	0.656	5.70	5.69	5.40	7.30	5.66	6.70	4.79	10.20
T'_{q0} (s)	0.70	1.50	1.50	1.5	0.44	0.40	1.50	0.41	1.96	0.00
T''_{d0} (s)	0.035	0.033	0.029	0.028	0.027	0.037	0.028	0.034	0.024	0.050
T''_{q0} (s)	0.035	0.033	0..029	0.028	0.027	0.037	0.028	0.034	0.024	0.090
SG10	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
SG12	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

*Parameters without unit are in pu (the base is the unit nominal power and voltage)

Table A.7. AVR ESDC1A

Par. \ Unit	01	02	03	04	05	06	07	08	09	10
Tr	0.02	0	0	0	0	0	0	0	0	0
Ka	200.0	6.2	5.0	5.0	40.0	5.0	40.0	5.0	40.0	5.0
Ta	0.005	0.05	0.06	0.06	0.02	0.02	0.02	0.02	0.02	0.06
Tc	3.1	0	0	0	0	0	0	0	0	0
Tb	40.0	0	0	0	0	0	0	0	0	0
Te	0.05	0.405	0.5	0.5	0.785	0.471	0.73	0.528	1.4	0.25
Kf	0.01	0.057	0.08	0.08	0.03	0.075	0.03	0.085	0.03	0.04
Tfl	0.3	0.5	1.0	1.0	1.0	1.246	1.0	1.26	1.0	1.0
Ke	0.5	-0.633	-0.02	-0.053	1.0	-0.042	1.0	-0.047	1.0	-0.049
E1	3.9	3.036	2.342	2.868	3.927	3.587	2.802	3.191	4.257	3.546
Se1	0.0001	0.66	0.13	0.08	0.07	0.064	0.53	0.072	0.62	0.08
E2	5.2	4.049	3.123	3.824	5.236	4.782	3.736	4.255	5.676	4.728
Se2	0.001	0.88	0.34	0.314	0.91	0.251	0.74	0.282	0.85	0.26
Vrmin	-5.0	-5.0	-10.0	-1.0	-10.0	-1.0	-6.5	-1.0	-10.5	-1.0
Vrmax	9.0	5.0	10.0	1.0	10.0	1.0	6.5	1.0	10.5	1.0

Table A.8. HYGOV

Parameter \ Unit	10
r	0.1
Tr	10.0
Tf	0.05
Tg	0.3
Tw	0.51
At	1.15
cosn	0.9579
Dturb	0
qnl	0.08
R	0.04
Gmax	0.1533
Velm	0.29
Gmin	0.9579

Table A.9. TGOV1

Par. \ Unit	02	03	04	05	06	07	08	09
T3	5.0	5.0	5.0	5.0	5.0	5.0	5.0	5.0
T2	1.25	1.25	1.25	1.25	1.25	1.25	1.25	1.25
At	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Dt	0	0	0	0	0	0	0	0
R	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
T1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
cosn	0.964	0.920	0.914	0.917	0.928	0.936	0.928	0.867
Vmin	0.152	0.061	0.157	0.229	0.186	0.087	0.133	0.174
Vmax*	0.159	0.068	0.914	0.917	0.928	0.936	0.928	0.867

Table A.10. Generation cost

Unit	c_2	c_1	c_0
01	0.0042	32.950	4,403.431
02	0.0109	89.702	10,948.005
03	0.0109	89.702	10,948.005
04	0.0109	16.298	1,948.189
05	0.0109	16.298	1,948.189
06	0.0109	16.298	1,948.189
07	0.0109	16.298	1,948.189
08	0.0044	32.956	4,181.673
09	0.0044	32.956	4,181.673
10	0	0	0

*Cost function: $f(p)=c_2p^2+c_1p+c_0$, $f\left(\frac{\$}{h}\right)$, p(MW)

This page intentionally left blank

List of symbols

General notation rules as follows.

- Scalar functions, variables and parameters expressed in p.u. are in lower case Latin fonts.
- Function and variable vectors are in lower case, bold fonts.
- Matrices are in upper case, bold fonts.
- The subscript 0 indicates initial value.
- A hat operator indicates an estimated quantity.
- A tilde operator indicates a phasor.
- A superscript asterisk indicates the conjugate of complex quantities.
- A superscript T indicates transpose.
- A superscript + denotes a posteriori.
- A superscript O denotes PMU observed bus.
- A superscript U denotes PMU non-observed bus.
- A subscript - denotes a priori.
- A subscript k indicates time instant.
- Time derivatives appear with dot or ∂.
- Subscripts (M<θ) and (M<α) denotes polar coordinates.
- A subscript (x,y) denotes rectangular coordinates.

z	Measurement
x	System state (also y)
e	Measurement error
h	Measurement function (matrix is **D**)
r	Measurement residual
σ^2	Variance
μ	Mean
p	Model parameter
u	System input
f, g	Nonlinear functions (system model)
A	Transition matrix
ζ	Trend vector

w	Model error
J	Objective function for the weighted least square problem
R	Covariance of measurement error
G	Gain Matrix
H	Jacobian of measurement function
C	Unit vector
F	Transition matrix (Kalman filter)
Q	Covariance of model error (Kalman filter)
P	Covariance of the estimated state
K	Kalman filter gain
L, M	Partial derivatives of nonlinear functions f and h (Kalman filter)
I	Identity matrix
χ^2	Chi-square distribution
S	Sensitivity matrix
t	Tuple vector (classification)
B, at	Attribute matrix and element (classification)
cl_j	Cluster j
$\mathbf{ce}_j$	Centroid of cluster j
$\mathbf{o}_i$	Object i (clustering)
WSS	Within sum of squares
v	Bus voltage magnitude
θ	Bus voltage angle
f	Bus voltage frequency
I	Current magnitude
α	Current angle
$\widetilde{vps}$	Calculated bus voltage pseudo-measurement
ΔV_{PMU}	Bus voltage deviation registered from PMU measurements
Y	Admittance matrix
Δt_{PMU}	PMU refresh rate
p_{st}, q_{st}	Active and reactive power consumption (static load)
f_{PMU_next}	Frequency reported by the closes PMU to the bus under analysis
p_i, q_i	Active and reactive power injection at bus i
$\tilde{s}_{inj}$	Conventional measurement of power injection

$\tilde{s}_{fl}$	Conventional measurement of power flow
$\tilde{v}_{PMU}$	Voltage PMU measurement
$\tilde{i}_{PMU}$	Current PMU measurement
$\tilde{s}_{virt}$	Virtual measurement (zero injection)
$\tilde{s}_{ps}$	Power flow pseudo-measurement
t'_{d0}, t'_{q0}	Open-circuit transient constants
e'_d, e'_q, e_{fd}	Transient voltages along the d and q axes and the field voltage
x_d, x_q, x'_d, x'_q	Synchronous and transient reactance
δ, ω	Rotor angle and speed deviation
H, D	Inertia constant and damping factor
p_m, p_e, q_e	Mechanical and electrical power (active and reactive) of the generator
v_d, v_q	Terminal bus voltage in the d and q axes
t_e, t_m	Electromagnetic and mechanical torque
p_{mot}	Power consumption of motor
r_s, r_r	Stator and rotor resistance
s	slip

This page intentionally left blank

List of acronyms

AI	Artificial Intelligence
ANN	Artificial Neural Network
CIGRE	International Council on Large Electric Systems
CKF	Cubature Kalman Filter
CVI	Cluster Validity Index
DAE	Differential Algebraic Equations
DFT	Discrete Fourier Transform
DSA	Dynamic Security Assessment
EKF	Extended Kalman Filter
EMS	Energy Management System
GPS	Global Positioning System
HVDC	High Voltage Direct Current
IED	Intelligent Electronic Device
IEEE	Institute of Electrical and Electronics Engineers
IES	Integrated Energy Systems
LAV	Least Absolute Value
MASE	Multi-area State Estimation
MIP	Mixed-Integer Programming
PCA	Principal Component Analysis
PDC	Phasor Data Concentrator
PDF	Probability Distribution Function
PF	Particle Filter
PMU	Phasor Measurement Unit
RTU	Remote Terminal Unit
SCADA	Supervisory Control Data Acquisition System
UKF	Unscented Kalman Filter
WLAV	Weighted Least Absolute Value
WLS	Weighted Least Square
WSS	Within Sum of Squares

This page intentionally left blank

Evidence of scientific activity

During my activities as a research associate, first, in the Institute of Electrical Energy (U.N.S.J. – C.O.N.I.C.E.T.) in San Juan, Argentina, and then, in the Institute of Energy Systems, Energy Efficiency and Energy Economics (TU Dortmund), in Dortmund, Germany, I gained experience in the field of monitoring systems based on PMU technology. The following are publications on my research work.

Publications 2020

[E7] Ortiz, G.A., Rehtanz, C., Colome, D.G.: Monitoring of power system dynamics under incomplete PMU observability condition. IET Gener. Transm. Distr. 2020;1-16. https://doi.org/10.1049/gtd2.12111

Publications 2019

[E5] Ortiz, G.A., Rehtanz, C., Colome, D.G.: Real-time State Estimation in a System Partially Observed by PMUs: A Coherency Data-mining Based Approach. In: *2019 9th ICPES*. Perth, Australia (2019).

[E6] Ortiz, G.A., Rehtanz, C., Colome, D.G.: Determinación de Pseudomediciones Dinámicas de Tensión en un Sistema Eléctrico de Potencia Parcialmente Observado por Equipos PMU. In: *Décimo Oitavo Encontro Regional Ibero-Americano do CIGRE (XVIII)*. Foz do Iguaçu, Brasil (2019).

Publications 2017

[E3] Ortiz, G.A., Colome, D.G.: Tendencias en Estimadores de Estado con Uso de Mediciones Sincrofasoriales. In: *Decimoséptimo Encuentro Regional Ibero-Americano del CIGRÉ (XVII ERIAC)*. Ciudad del Este, Paraguay (2017).

[E4] Ortiz, G.A., Colome, D.G.: Dynamic State Estimation of an Electric Power System. In: *2017 IEEE PES Innovative Smart Grid Technologies Conference - Latin America (ISGT Latin America)*. Quito, Ecuador (2017).

Publications in 2016

[E2] Ortiz, G.A., Colome, D.G., Quispe Puma, J.J.: State estimation of power system based on SCADA and PMU measurements. In: *2016 IEEE ANDESCON*. Arequipa, Perú (2016).

Publications in 2015

[E1] Ortiz, G.A., Colome, D.G.: Tendencias en el Diseño e Implementación de Plataformas de Prueba de Dispositivos PMU. In: *Décimo Sexto Encuentro Regional Ibero-Americano del CIGRÉ (XVI ERIAC)*. Puerto de Iguazú, Argentina (2015).