Enhancing Frequency Stability of Low-Inertia Grids with Novel Security Constrained Unit Commitment Approaches

by
Mingjian Tuo

A report submitted to the Department of Electrical and Computer Engineering,

Cullen College of Engineering

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in Electrical and Computer Engineering

Chair of Committee: Xingpeng Li

Committee Member: Kaushik Rajashekara

Committee Member: Zhu Han

Committee Member: Lei Fan

Committee Member: David R. Jackson

University of Houston

August 2023

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to Dr. Xingpeng Li for providing me the opportunity to conduct my Ph.D. studies at the Renewable Power Grid lab at the University of Houston. His mentorship, advice, and encouragement throughout the past four years have been invaluable for the development of this research. I would like to thank Prof. Kaushik Rajasekhara for the continuous support and encouragement through my program.

I would also like to thank Dr. Zhu Han, together with Dr. Lei Fan, and Dr. David R. Jackson, for being part of my dissertation committee and providing valuable insights on how to improve this dissertation. I would also like to thank already graduated students Dr. Arun Venkatesh Ramesh, Dr. Amin Sadat, Dr. Yu Yao, and Dr. Shilei Jiao together with current students Cunzhi Zhao, Jin Lu, Thuan Pham, Jesus Silva, Qiushi Wang, for the supportive discussions and camaraderie while pursuing this program.

I would like to thank my parents, Jun Tuo and Fuying Liu, for all their love, support, and patience in the last four years of my academic life. They always encouraged me to pursue academic excellence.

ABSTRACT

Conventional synchronous generators are gradually being replaced by low-inertia inverter-based resources. Such transition introduces more complicated operation conditions. Insufficient system inertia would lead to dramatical change in rate of change of frequency (RoCoF) and further results in under frequency load shedding as well as tripping of generator protection devices.

Inertia estimation can ensure the accountability and reliability of inertia response through implementation of frequency control ancillary services. Existing power system inertia estimation approaches do not take such heterogeneity into consideration, and the risk of estimation error introduced by uniform model-based approach is underestimated. Hence, strategies such as dynamic inertia estimation method and machine learning-assisted inertia estimation approaches are introduced.

Frequency related constraints have been imposed in the conventional security-constrained unit commitment (SCUC) model to keep the minimum amount of synchronous inertia online and secure system frequency stability. However, in a large system inertia distribution may vary significantly, areas of low inertia are more susceptible to frequency deviation following G-1 contingency, posing risks of load shedding and generation trip. To address these issues in frequency constrained SCUC, ensuring locational frequency stability, two novel strategies are proposed in this thesis:

(1) A novel location based RoCoF constrained SCUC (LRC-SCUC) model that can capture the locational frequency response characteristics and counteract the impact of system oscillation, guaranteeing the RoCoF security following a G-1

disturbance. A multiple-measurement-window method is introduced in this work to constrain highest value during oscillation. Simulation results demonstrate the effectiveness of proposed LRC-SCUC model. The results also show that deploying virtual inertia techniques not only reduces the total cost, but also improves the system market efficiency.

(2) A generic data-driven frequency and RoCoF predictor is first trained to predict maximal frequency deviation and the highest locational RoCoF simultaneously based on a high-fidelity simulation dataset. And the derived frequency related constraints are then incorporated into machine learning assisted SCUC to guarantee frequency stability following the worst generator outage case while ensuring operational efficiency. In addition, sparse computation and an active rectified linear unit (ReLU) linearization method are implemented to further improve the algorithm efficiency while retaining solution quality.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

NOMENCLATURE

Sets and indices:

G	Set of generators.
G_n	Set of generators connected to bus n.
K	Set of lines.
$K^+(n)$	Set of lines with bus n as receiving bus.
$K^-(n)$	Set of lines with bus n as sending bus.
T	Set of time periods.
N	Set of buses.
N_G	Set of generator buses.
$N_{loc,g}$	Set of local generator buses for generator g.
$N_{n-loc,g}$	Set of non-local generator buses for generator g.
g	Generator g.
k	Line k.
t	Time t.
n	Bus n.
NS	Set of samples.
NL	Set of neural network layers.
$\mathcal{H}$	Set of active selected neurons.
$\bar{\mathcal{H}}$	Set of not selected neurons.
q	Neural network layer q.

l	l-th neuron of a neural network layer.
s	Sample s.

Parameters:

c_g	Linear operation cost for generator g.
P_g^{min}	Minimum output limit of generator g.
P_g^{max}	Maximum output limit of generator g.
P_k^{max}	Long-term thermal line limit for line k.
b_k	Susceptance of line k.
$D_{n,t}$	Forecasted demand at bus n in period t.
$E_{n,t}$	Forecasted renewable generation at bus n in period t.
R_g^{hr}	Ramping limit of generator g.
R_g^{re}	Reserve ramping limit of generator g.
H_g	Inertia constant of conventional generator g.
c_g^{NL}	No load cost for generator g.
c_g^{SU}	Startup cost of generator g.
c_g^{RE}	Reserve cost of generator g.
c_n	Nodal cost for additional inertia at bus n.
Δt	Frequency monitoring window.
nT	Number of time periods.

M_t^{Total} Virtual inertia budget.

Ω A big real number.

T_1 First frequency measuring point.

T_2 Second frequency measuring point.

$RoCoF_{lim}$ Pre-specified RoCoF threshold.

DT_g Minimum down time of generator g.

UT_g Minimum on time of generator g.

nT Number of time periods.

A A big real number.

f_{nom} System frequency nominal value.

f_{lim} Pre-specified minimal frequency threshold.

W_q Weights matrix of layer q.

b_q Bias matrix of layer q.

$W_{N_L+1}^{dev}$ Weights matrix of last layer for maximal frequency deviation prediction.

$W_{N_L+1}^{rcf}$ Weights matrix of last layer for maximal RoCoF prediction.

$b_{N_L+1}^{dev}$ Bias matrix of last layer for maximal frequency deviation prediction.

$b_{N_L+1}^{rcf}$ Bias matrix of last layer for maximal RoCoF prediction.

$UB_{q(l)}$ Upper bound of preactivated value of l-th neuron of layer q.

$LB_{q(l)}$ Lower bound of preactivated value of l-th neuron of layer q.

Variables:

$P_{g,t}$ Output of generator g in period t.

$m_{g,t}$ Average nodal inertia in period t after loss of generator g.

$\Delta m_{g,t}$ Change in average nodal inertia in period t after loss of generation.

m_t Average nodal inertia in period t.

$r_{g,t}$ Reserve from generator g in period t.

$u_{g,t}$ Commitment status of generator g in period t.

$v_{g,t}$ Start-up variable of generator g in period t.

M_t System aggregated inertia in period t.

M_t^{VI} Virtual inertia in period t.

$P_{k,t}$ Flow online k in period t.

$\theta_{n,t}$ Phase angle of bus n in period t.

$\theta_{m,t}$ Phase angle of bus m in period t.

$\mu_{g,t}$ Maximal output indicator of generator g in period t.

$\rho_{g,t}$ Maximal output value of generator g in period t.

$z_{q(l),t}$ Preactivated value of l-th neuron of layer q in period t.

$\hat{z}_{q(l),t}$ Activated value of l-th neuron of layer q in period t.

$a_{q(l),t}$ Activation status l-th neuron's ReLU of layer q in period t.

$z_{i,j,\epsilon,t}^{q}$ Preactivated value of input patch centered at location (i,j) of kernel ϵ of layer q in period t.

$\hat{z}_{i,j,\epsilon,t}^{q}$ Activated value of input patch centered at location (i, j) of layer q in period t.

$a_{i,j,\epsilon,t}^{q}$ Activation status of neuron's ReLU corresponding to input patch centered at location (i, j) of kernel ϵ of layer q in period t.

1. INTRODUCTION

1.1 Background

Power system is an electrical network which consists of generation, distribution and transmission system. Main components within a power system include, but not limited to, generators, loads, transmission lines, transformers, phase shifters, circuit breakers, and shunts. As energy cannot be economically stored on a large scale. Electricity must be generated, transported, and consumed simultaneously within a power system. Therefore, it is very challenging to maintain reliable real-time operations of modern power systems.

Power system operations falls into several time horizons, including day-ahead scheduling and real-time system control. The energy production is a multi-period problem, typically spanning 24 hours, to meet the demand at each hourly interval in advance. The formulations of the model are convex and linear expressions to ensure that a global optimal solution can be obtained. Given the load forecast and renewable generation prediction, the production of energy can be obtained by solving an economical operational problem subject to multiple physical restrictions such as Kirchhoff's laws, and network limits. Generators have constraints related to ramping rates, minimum up-down times, and reserve margins. This process would provide information on the status and dispatch points of generators within each time interval, which is also known as unit-commitment (UC).

Contingency may result in system violations, islanding, irreversible damage to electrical equipment, and in the worst-case a blackout. Therefore, it’s important to

ensure system stability under the condition of potential contingency. The guidelines of NERC (North American Reliability Council) standard 51 describes that the solution of the UC requires an N-1 reliability criterion [1]. To ensure uninterrupted power supply in the event of generator or line failures, it is necessary to plan and operate the power system in a manner that can accommodate contingencies. To meet this requirement, system operators typically employ the security-constrained unit commitment (SCUC) approach for generator dispatch and commitment. However, the inclusion of security constraints in SCUC leads to higher operational costs compared to the standard unit commitment (UC) approach.

1.2 Motivations

Governments around the world are adopting clean energy goals to limit carbon dioxide emissions and combat climate change. Due to the characteristics of environmental benefits and growing cost-competitiveness, the penetration level of renewable energy resources (RES) has significantly increased in the last decades. The total global installed capacity has increased by a factor of about 6 for wind power and a factor of 40 for solar power in the past decade [2]. Such transition has also accelerated the retirement of traditional coal-based plants and changed the investments in energy portfolios.

Synchronous generators used to play an important role in system operation, a number of mandatory standards for power system reliability have been developed based on the characteristics of synchronous generator dominant system. Complying with these standards secure the system stability. However, with the influx of inverter-

based RES, the dynamics of traditional system have been totally changed. With the uncertainty nature of RES such as intermittency in wind and solar generation, modern power systems are more complex and dynamical. Traditionally, synchronous generators could provide inertia to the system through stored kinetic energy in rotating mass which can counteract frequency excursion during disturbances and thus enhance the frequency stability. However, RESs are interfaced to the grid through converters which electrically decouples the rotor's inertia from the whole system [3]. Such inverter-based resources contribute little synchronous inertia to the whole system, which is even true for wind power plants taking advantage of kinetic energy stored in wind turbines. With more generation coming from converter-based resources, insufficient inertia would be a main challenge for power systems stability.

Frequency and related metrics have been used to estimate the stability of the system, while closed-form stability conditions for ancillary services procuring could deduced from the swing equation [4]. The system equivalent model for frequency analysis is a simplification of the actual frequency dynamics in a power grid, where the one-machine swing equation is extended to the whole system by using the system average inertia constant H (Center of Inertia concept), and load damping is modelled as a single constant D. The parameters in the model are typically represented by an aggregated of all components within the system, not including any effect of network topology, generator location, the impedance of tie-lines and RES generator and control characteristics. However, only considering the system equivalent model would underestimate the actual need for frequency ancillary services, leading to unexpected

contingency. Different from focusing on the collective performance of the power system, the RoCoF experienced by each bus is distinct. Relative location of measurement point to disturbance is a pertinent factor in system inertial response and the RoCoF is usually higher for location where networks are weakly interconnected [5]. The impact of disturbance propagation on post-contingency frequency dynamics and locational frequency security against N-1 contingency can be derived and incorporated into SCUC formulations.

Energy storage technologies, such as largescale batteries system, have been considered to provide ancillary services, such as virtual inertia, to address the aforementioned issues in the power system. By emulating the characteristics of traditional mechanical inertia using power electronic devices and control strategies, power electronic devices, such as grid-connected inverters can mimic the inertial response of synchronous generators. These control strategies involve monitoring the system frequency and responding to frequency deviations by dynamically adjusting the power output. By implementing virtual inertia, power systems can improve their ability to respond to disturbances and maintain stable operation. The synthetic inertia provided by renewable energy sources and other grid-connected devices helps to dampen frequency fluctuations and enhance system resilience. It helps maintain grid stability and support the reliable operation of the overall power system while reducing the preventive operational cost. In addition, the demand side also experiences significant changes over the past years with increased demand patterns, demand response strategies have made this transfer bi-directional and could potentially

improve system security. Therefore, more advanced operational strategies and system analysis approaches are required to handle such transition in the system and ensure the system stability.

The development of PMU based wide area measurements systems (WAMS) have enabled data-driven techniques in power system analysis [6]. Through leveraging collected data, deep learning algorithms have been widely used on works such as decision-support, prediction and automation to help operators improve system performance. Applying machine learning (ML) to solve these challenging situations is critical for development of clean and green energy of the future and it could help achieve a reasonable statistical evaluation of the risk. Applications of ML toward renewable energy have been widely researched and studied in recent years.

Compared to traditional computational approaches, machine learning algorithms have an intrinsic generalization capability with greater computational efficiency and scalability. Since machine learning algorithms have the ability to learn complex nonlinear input-output relationships, reflexively adapting themselves to the data, they have been used to predict variables within the power system. There are two broad types of problems namely, regression and classification. The regression models estimate continuous values whereas classification models approximate a mapping function from input variables to identify classifications and categories. Modern power systems are connected to different devices which provide frequency regulation service, the estimation of frequency related metrics based on low-order model approximation

may be inaccurate. Regression models have been used to capture the entire characteristics and track frequency related metrics for system state estimation.

1.3 Summary of contents

The rest of this dissertation is structured as follows. A thorough literature review is presented in Chapter 2. In this chapter, the specifics of system identification and extraction of inertia values is described. Then, the method of dynamic inertia distribution estimation and total system inertia calculation are detailed. At the end of this chapter, the simulation results on the test system and the concluding remarks are presented.

In Chapter 3, the frequency dynamics of power systems are first described. The proposed inertia estimation algorithms using long-term recurrent convolutional network (LRCN) and graph convolutional neural networks (GCN) techniques are compared. Simulation setup is detailed. Case studies show that that the proposed models have better performances than the benchmark deep neural models, and the proposed models also show high robustness under conditions with higher noises. The proposed optimal phasor measurement units (PMUs) placement method has been proved to be capable of improving the performance of all implemented models.

In Chapter 4, the fundamentals of the post-contingency frequency dynamics of reduced model are first derived, and the corresponding analytic expressions are then incorporated into the SCUC model. $G-1$ contingency at different locations is investigated to implement the non-linear location based RoCoF constraints, and the PWL method is utilized to linearize those non-linear location based RoCoF

constraints. Case studies show that imposing location based RoCoF constraints in the SCUC model can ensure the locational frequency security during worst-case contingency event.

Chapter 5 first discusses the power system mathematical based model. The methodology of model-based data generation is detailed. Training dataset is generated from models over various scenarios, and deep neural network based RoCoF predictors are trained based on the high-fidelity simulation dataset. The locational RoCoF-limiting constraints are then derived based on the RoCoF predictor. And a set of mixed-integer linear constraints are used to reformulate the nonlinear constraints. Time domain simulation results on PSS/E demonstrate the effectiveness of the proposed algorithms.

Chapter 6 studies both the frequency deviation stability and RoCoF security. DNN-based frequency metrics predictor is used to track the locational RoCoF and maximal frequency deviation simultaneously. The corresponding frequency constraints are incorporated into the traditional SCUC model. The efficiency of the machine learning assisted FCUC is discussed. The concept of sharing parameters and sparse neural network are introduced to improve the efficiency of proposed model. Besides, an active rectified linear unit (ReLU) linearization method is implemented to further improve the algorithm efficiency while retaining solution quality.

Finally, Chapter 7 concludes this dissertation and presents potential future work.

2. PHYSICS-BASED INERTIA ESTIMATION

Inertia plays a critical role in the stability and operation of power systems, as it provides a buffer against disturbances and helps maintain the frequency within acceptable limits. Modern systems are required to accommodate an increasing volume of RES, these changes have led to a decrease in the overall system inertia, which is primarily attributed to the reduced rotational inertia provided by traditional synchronous generators. In addition, due to the physical characteristics of these resources, their interaction with the grid is substantially different from the interaction of the traditional synchronous generators which have steam and hydro turbines. While the rotating parts of the synchronous generators inherently contribute inertia to the system, converter-based resources are decoupled from the system via power electronics.

In order to maintain operational security of the power system, its frequency must be within a predefined range, and the deviation should be not far from the nominal frequency. In the case of large production units tripping and supply/demand imbalances, the inertia that originally slows down the frequency deviations is significantly reduced, resulting in much higher rate of change of frequency (RoCoF) in systems. The reduction in system inertia also reduces the time for primary control to take actions. Frequency would then drop to an unacceptable level, further causing in the disconnection of production units and loads, producing a cascade effect, which may lead to widespread power outages.

Since the frequency ancillary services are becoming more complicated and less predictable. It is therefore necessary to gain an improved understanding of both the inertial frequency response of the power system and the security of the system in near to real-time. Inertia estimation could ensure the impact of incidents to specific areas of the network is understood, facilitating more economically efficient operation of the power system. Dynamic analysis is used to study the response of the power system to disturbances and analyzing the frequency deviations. By observing the frequency response, system operators and analysts can estimate the effective inertia in the system. Reference [7] proposed an inertia estimation approach which divides the system into multiple subareas and estimates inertia of each subarea separately, but the approximation made in mathematical model introduces additional errors. The Electric Reliability Council of Texas (ERCOT) uses a real-time sufficiency monitoring tool to monitor inertia based on the operating plans submitted by the generation resources [8].

2.1 Power System Inertia

Power system inertia estimation is the process of quantifying the total inertia present in a power system. Inertia is a measure of the rotating mass and energy stored in synchronous generators, and it plays a crucial role in the stability and dynamic behavior of the system. Traditionally, the inertia of synchronous generators in a power system can be estimated based on their physical characteristics, such as their moment of inertia and rated power. These values are typically provided by manufacturers or obtained through measurements and can be used to calculate the total inertia of the synchronous generators connected to the system.

2.1.1 Inertia of a single machine

Conventional synchronous generators play a significant role in system frequency regulation; the inertia provided by synchronous generators help the power system to resist changes in system frequency. For a single rotating machine, the nominal inertia of it is equal to its kinetic energy in megawatt seconds (MWs) at rated speed, which is determined by the moment of inertia and rotational speed,

$$E_i = \frac{1}{2} J_i {\omega_i}^2 \tag{2.1}$$

where J_i is the moment of inertia of a generator and turbine [$\mathrm{kg \cdot m^2}$], ω_n is the rate mechanical angular velocity of the rotor [$\mathrm{rad/s}$],

The inertia of a single rotating shaft is commonly measured by its inertia constant, which is the per-unit value of inertia depending on the base value of the rated apparent power. Which can be express as,

$$H_i = \frac{J_i {\omega_i}^2}{2S_{B_i}}, \tag{2.2}$$

where S_{B_i} is the rate apparent power if the generator [VA],

Sometimes moment of inertia of a turbine-generator is given in gravimetric units GD^2. Gravimetric units can be converted to joules using,

$$J_i = \frac{GD^2}{4}, \tag{2.3}$$

where GD^2 is the moment of inertia [$\mathrm{kg \cdot m^2}$].

2.1.2 Inertia of the system

The frequency response of the power system is determined by the total stored kinetic energy within the system. All synchronously connected generators contribute to this resistance with rotational energy stored in rotating mass. System inertia is then defined as the summation of kinetic energy stored in connected synchronous units, which is defined as follows,

$$E_{sys} = \sum_{i=1}^{N} \frac{1}{2} J_i {\omega_i}^2. \tag{2.4}$$

Synchronous condensers (SC) and induction motor provide a viable alternative to the inertia provided by synchronous generation [2]. Because synchronous condenser is a rotational machine, the kinetic energy stored in its rotating mass is available as an active power source [9]. SC can be used to inject or absorb the rotational energy to enhance the system frequency stability for the system frequency regulation in the event of contingency, the inertia of the SC can be calculated in a similar way as generators. Induction motor loads also plays an important role in modern power system. They connected synchronously to power system also contribute to system inertia and can be taken into account.

2.2 System Frequency Stability

As a fundamental health measure of power systems, the nominal grid frequency of North American is 60 Hz. In a power grid, hundreds or thousands of generators are synchronized, and grid frequency is determined by how fast these conventional generators spin, which reflects the ability of a grid to balance generation

and demand. The ability of a system to maintain within a narrow range under normal conditions and to recover from disturbance is often referred to as frequency stability. Ever-changing load and generation profiles will cause power imbalance and frequency deviations. When system frequency drops to a low level, it may impose vibratory stresses on the turbine blades and reduce the outputs of boiler feed pumps or fans [10]. To protect generators from low-frequency operational conditions, under-frequency protective relays are involved so that generators will be tripped off when the frequency falls below a predefined level (59.5 Hz in most of United States and 59.3 Hz in Texas). Thus, the operation of frequency related control is of great importance and has been identified as a high priority area by many power system operators [11]-[14].

The inertial frequency response is influenced by various factors, including the total system inertia, the rate of change of power injections, and the system's control and protection mechanisms. Analyzing and understanding the inertial frequency response is crucial for power system operators and engineers to ensure stable and reliable system operation. It helps in designing appropriate control strategies, determining system requirements, and assessing the impact of changes in the generation mix on frequency stability.

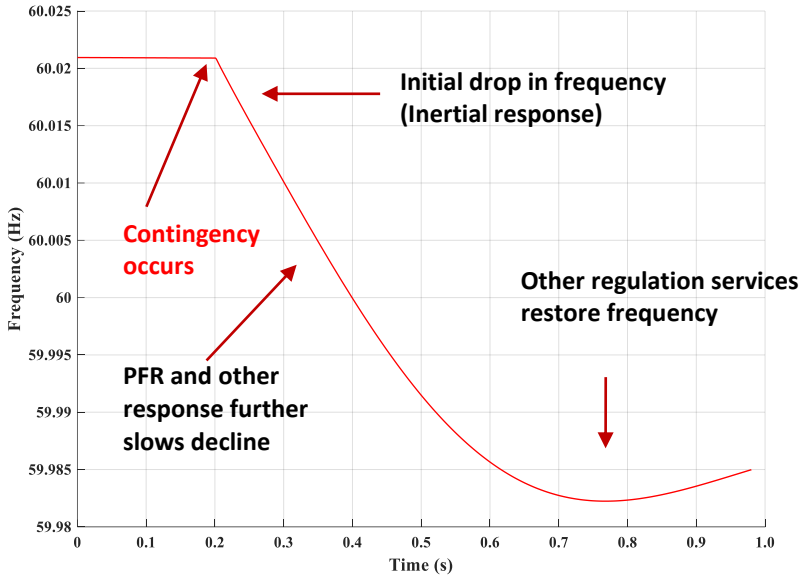


Fig. 2.1 Frequency evolution during an event.

Traditionally, synchronous machines are used to generate active power, regulate the frequency, and provide inertial response. During an event, the inertia within rotating mass of a synchronous machine would response instantaneously to compensate for fluctuations and disturbances in the short term shown in Figure 2.1. After that, the primary frequency response (PFR) would get involved. PFR detects changes in frequency and automatically adjusts operations of synchronized generators to arrest the frequency decline before reaching the frequency nadir. This usually takes place on time scales of tens of seconds.

After that secondary frequency regulations take over by varying the active power generated by the machines. This takes place on time scales of minutes. The tertiary control, when implemented, and generator rescheduling are slower and take place on time scales of the order of tens of minutes and hours, respectively. The summary of frequency regulation schemes is listed in Table 2.1.

Table 2.1 Frequency Regulation Schemes

No.	Control Name	Responding Time	Objectives
1	Inertial Response	< 1 seconds	Power balancing, transient frequency. (automatic)
2	Primary Control, governor	3 seconds-10 minutes	Power balancing, transient frequency. (automatic)
3	Secondary Control (AGC)	30 seconds-20 minutes	Power balancing, steady-state frequency. (automatic)
4	Tertiary Control (SCED)	5 minutes to 1 hour	Power balancing via economic dispatch.

2.3 Inertia Estimation using Measured Frequency Disturbance

2.3.1 Frequency Dynamics

Power variations in consumption and production affect the frequency continuously. Small power disturbance is only visible as noise in the frequency due to the mass of the synchronous generators in the power system. For large power imbalances, the system frequency excursion from the nominal value would become larger. Following a sudden power deficit, for example load step increase or a trip of a generator, the kinetic energy stored in synchronized rotors would be discharged to compensate the power mismatch, the rotating speed of the generator would be decreased and thereby the frequency. The dynamic behavior a single machine can be described using the motion equation of a rotating mass as,

$$\Delta P_m - \Delta P_e = M_i \frac{d\Delta\omega_i}{dt} + D_i \Delta\omega_i, \tag{2.5}$$

where ΔP_m is the mechanical power and ΔP_e is the electrical output power, with M_i and $D_i$denoting the normalized inertia and damping constants.

How fast the frequency changes depend on the change in active power and the system inertia. In systems with high inertia, the frequency response is slower, meaning that the frequency changes gradually over time. This slower response allows for more time to implement corrective actions and restore the balance between generation and load. On the other hand, in systems with low inertia, such as those with a high penetration of renewable energy sources or reduced rotating masses, the frequency response is faster, resulting in more rapid frequency deviations. System equivalent frequency response model has been widely used for system operations. The characteristics of power systems are very complex due to the existence of multiple electromechanical oscillation modes, system control noise, and variant distribution of inertia throughout the grid. The principal frequency dynamics can be described by the evolution of the center of inertia (COI) speed [15]-[16], which is defined as,

$$\omega_{COI} = \frac{\sum_{i=1}^{N} H_i \omega_i}{\sum_{i=1}^{N} H_i}, \tag{2.6}$$

where N is the total number of synchronous generators, H_i is the i^{th} unit's inertia constant, and ω_i is the angular frequency of the rotor of the i^{th} generator. In this way, a system could be considered as a single equivalent center of inertia. And the simplified swing equation of the whole grid is considered as the extension of one machine model [17],

$$\Delta P_{m,sys} - \Delta P_{e,sys} = M \frac{d\Delta\omega}{dt} + D\Delta\omega, \tag{2.7}$$

where $\Delta P_{m,sys}$ is the system mechanical power change and $\Delta P_{e,sys}$ is the system electrical output power change, with M and D denoting the system inertia and damping constants.

2.3.2 Inertial Frequency Response Estimation

Inertial frequency response refers to the behavior of the power system's frequency following a disturbance or sudden change in power generation or load. It is directly related to the presence of inertia in the system. When a disturbance occurs, such as a sudden loss of generation or a large increase in load, the system experiences an imbalance between power supply and demand. The period immediately following an in-feed loss, up to one second, is considered as the inertial response of the system, before the primary frequency response reserves are activated [18]. During this time, the system frequency response is unregulated, and its behavior is dictated primarily by the inertia present in the power system. Then the equation relating the maximum post-contingency RoCoF to the total system inertia can be expressed as,

$$\frac{df}{dt} = \frac{-\Delta P}{2H_{sys}S_{Base}} f_0, \tag{2.8}$$

where H_{sys} is the system equivalent inertia constant, S_{Base} is the system rate apparent power. f_0 is the system frequency when the disturbance occurs. The relationship of the synchronous speed and the system frequency can be calculated as,

$$\omega_0 = 2\pi f_0. \tag{2.9}$$

Most inertia estimation approaches rely on event transient measurement of collective system model following recorded disturbances. From the swing equation,

given the certain level of disturbance, the system inertia can be estimated based on the swing equation (2.8).

2.3.3 Impact of RES on Frequency Response

Renewable energy resources have been recognized as the most promising low carbon generations. In recent years, wind and photovoltaic (PV) power plants have witnessed a significant growth. Giga Watts (GW) wind and PV generation have been installed in many countries. For some countries in Europe, the wind or PV generation may even be able to meet most of electricity demand [19]. However, the increasing integration of renewable energy sources (RES) has a significant impact on the inertial response of power systems. Traditional power systems rely on the inertia provided by synchronous generators, which helps stabilize the system's frequency during disturbances. However, RES, such as solar and wind power, do not inherently possess the same level of inertia as synchronous generators.

Renewable energy sources typically exhibit minimal or significantly lower rotational inertia compared to synchronous generators. This reduction in system inertia can lead to more rapid frequency deviations after disturbances, as there is a decreased amount of rotational energy available to mitigate frequency fluctuations. Abrupt changes in RES generation or sudden loss of RES output can result in more substantial frequency excursions. The lower system inertia associated with RES presents challenges in maintaining stable frequency. The limited inertial response from RES can lead to frequency deviations that exceed acceptable limits, potentially

compromising the reliable operation of the power system. To address this, additional measures are often required to ensure frequency control and stability.

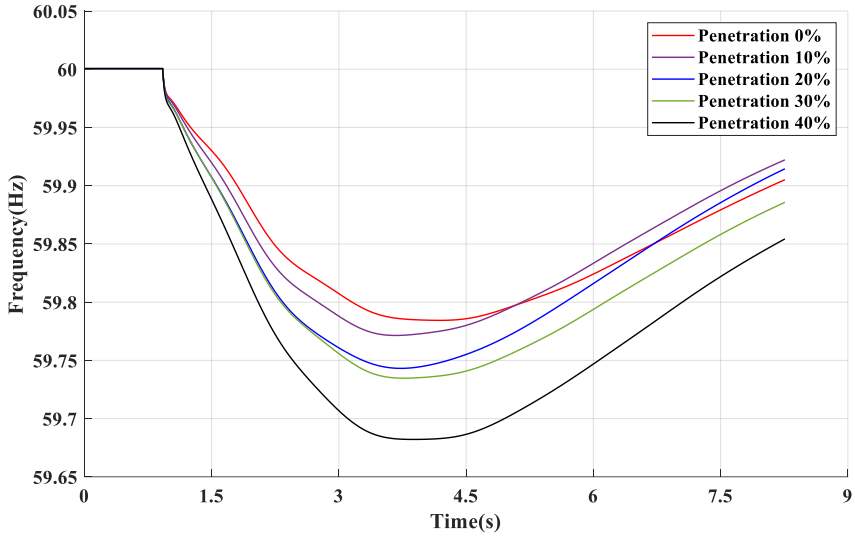


Fig. 2.2 Frequency response under different penetration levels of renewable generation on the IEEE 24-bus system.

Fig. 2.2 shows the simulation results of IEEE 24-bus system with different penetration levels of renewable generation in Transient Security Analysis Tools (TSAT). System frequency responses are compared regarding to the penetration rates at 0%, 10%, 20%, 30% and 40%, respectively. Results indicate that for systems with higher RES penetration level, the degradation of system frequency response leads to lower frequency nadir and higher initial RoCoF values.

Compensating for the reduced inertial response from RES necessitates the integration of supplementary ancillary services, such as frequency regulation and fast-acting reserves, within power systems. These services play a crucial role in managing frequency deviations by swiftly adjusting the power output from controllable resources or utilizing energy storage systems. As the penetration of RES increases, the

provision of these ancillary services becomes increasingly vital to maintain system stability and ensure reliable power system operation.

2.4 Dynamic Inertia Estimation

Dynamic inertia estimation refers to the process of continuously and adaptively estimating the effective inertia of a power system in real-time or near real-time. Unlike traditional static inertia estimation methods that rely on fixed values or assumptions, dynamic inertia estimation takes into account the dynamic behavior and changing conditions of the system. By leveraging measurements and data from the power system, such as frequency deviations, and applying mathematical models or algorithms, the effective inertia of the system can be estimated dynamically. Dynamic inertia estimation is particularly important in modern power systems that incorporate a significant amount of renewable energy sources and inverter-based resources. These resources have limited or negligible inherent inertia, and their dynamic response to system disturbances can significantly impact the overall system inertia.

Accurate and real-time estimation of dynamic inertia helps system operators and planners in assessing and managing the stability and reliability of the power system. It enables them to make informed decisions regarding control actions, grid operation, and resource dispatch to maintain system stability and mitigate the potential impacts of changing generation profiles.

2.4.1 Frequency Measurement

Power system frequency measurement involves monitoring and quantifying the frequency of the alternating current (AC) in a power system. Frequency is an essential parameter in power systems as it reflects the balance between power supply and demand. The nominal frequency in most power systems is 50 Hz or 60 Hz, depending on the region. For north America the nominal frequency is 60 Hz.

Phasor Measurement Unit (PMU) is specialized device that being used in the power industry to measure and monitor electrical quantities such as voltage, current, and frequency. PMUs provide highly accurate time-synchronized data on the state of an electrical power system. Most PMUs can calculate up to 30 to 60 samples per cycle with the GPS time stamp provided by hardware that has an accuracy of millisecond or higher [20], which can be used to improve the reliability, stability, and efficiency of the system.

The ability to estimate the inertia of the system using measured frequency through PMU during disturbances are affected by multiple factors including: precise data on the size of loss; identification of event start time; accuracy of frequency measurement; and location of measurement point relative to in-feed loss. Fig. 2.3 shows the distinct frequency response on different buses within IEEE-24 bus system. It also has been proved [21] that a method of curve fitting is required to mitigate the impact of measured transients in frequency following a loss, otherwise the calculated RoCoF may be significantly larger than the true value.

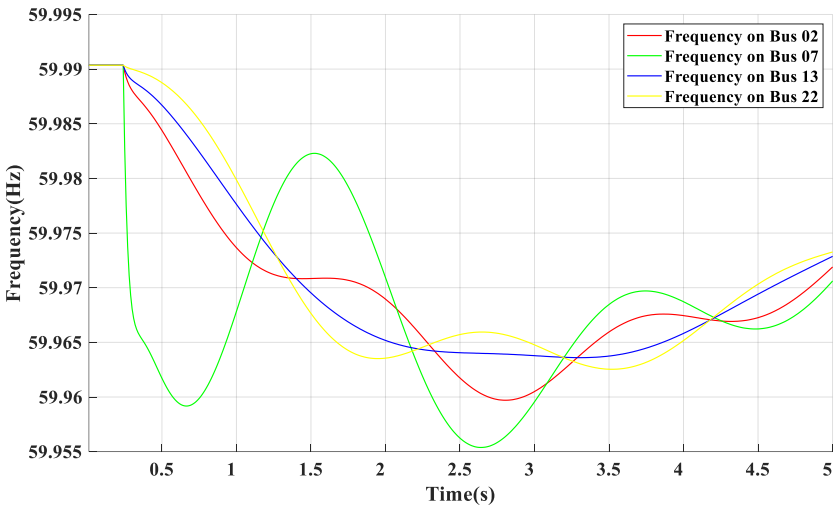


Fig. 2.3 Frequency measurements on different buses.

The cost of PMUs depends on the count of measurement channels available. Due to the high cost of having a PMU at each node, it is important to select the limited measurement taken from a location in the network which is suitable for the estimate of RoCoF.

2.4.2 Inertia Distribution Estimation

Power system inertia distribution refers to the presence of diverse elements and inertia related characteristics within a power system. It implies that the power system is composed of different types of generation sources, transmission lines, distribution networks, and loads that exhibit various responses in their inertia related behaviors. Power system inertia distribution estimation poses challenges in terms of system planning, operation, control, and stability. It involves coordination and integration of various generation sources and the implementation of different control schemes

Thus, for an equivalent COI model of a large system, the neglection of power swing and oscillation dynamics would introduce high approximation errors into the dynamic analysis. The research presented in [22] demonstrated that a power system can be considered as multiple centers of inertia, coupled through the network. To locate the equivalent center of inertia and estimate the system inertia accurately, an inertia distribution index (IDI) is introduced, which has been proved to be highly linear correlation with system transfer function residue. Availability of measurements from PMU makes it possible to evaluate the deviation of bus frequency from COI frequency in real time. The electrical distance from the monitored bus to COI location can be defined as follows,

$$dist(f_k, f_{COI}) = \int_{T_0+t_d}^{T+T_0+t_d} \left(f_k(\tau) - f_{COI}(\tau)\right)^2 d\tau, \tag{2.10}$$

where T_0 is the time when a disturbance is detected, t_d is the dead time considering the dead band of frequency, T is the time length of the integration period to be determined, and n is the total number of buses. Normalized inertia distribution index following a disturbance can be calculated as,

$$IDI_k = \frac{dist(f_k, f_{COI})}{\max\limits_{k\in\{1,..,n\}} dist(f_k, f_{COI})}. \tag{2.11}$$

2.4.3 COI based Inertia Estimation

The value of IDI_k reflects the electrical distance from bus k to the COI location, the closest bus to the COI is determined as,

$$k_{COI} = \arg \min_{k\in\{1,..,n\}} IDI_k. \tag{2.12}$$

However, the COI location may not always be located at a particular bus; and IDI_k may also vary under different disturbance events. Thus, to accurately estimate the system inertia, a clustering approach is proposed in this paper to determine the multi-bus COI area. Following a disturbance event, the bus k_{COI} is selected as the initial mean of points in the COI area cluster, which represents that this bus is the most stable bus under the specific event. The electrical distance from an estimated bus to bus k_{COI} can then be calculated below,

$$dist\left(f_k, f_{k_{COI}}\right) = \int_{T_{0+t_d}}^{T+T_0+t_d} \left(f_k(\tau) - f_{k_{COI}}(\tau)\right)^2 d\tau, \tag{2.13}$$

where $f_{k_{COI}}$ is the measured frequency of the bus nearest to COI location.

The COI area S_{COI} consists of buses that have electrical distances less than the pre-determined threshold value δ,

$$S_{COI} = \left\{k: dist\left(f_k, f_{k_{COI}}\right) \leq \delta\right\}. \tag{2.14}$$

It is known that the location of disturbance is a key factor in the inertia estimation. During a normal operation period, disturbance on different buses may cause distortion in bus frequency. To mitigate the impact of disturbance location on system inertia estimation, a dynamic COI area estimation method is proposed. During a specific time period, the system inertia is assumed to be stable under normal operation. We set a system observation window, within which events are detected while the system remains stable. The set of buses, S_{COI}^T, identified within the COI area over a period T_{win} is defined. The COI bus over period T_{win}, k_{COI}^T,, is defined in (2.15),

$$S_{COI}^{T_{win}} = \left\{k^t : dist\left(f^t{}_k, f^t{}_{k_{COI}}\right) \leq \delta, t \in T_{win}\right\}, \tag{2.15}$$

$$k_{COI}^{T_{win}} = \arg \max_{k \in \{1,..,n\}} C_k, \tag{2.16}$$

where C_k is the count of bus k identified as a bus of the COI area over a period of T_{win}, t indicates the event time within the observation window. Highest C_k means bus k is closest to the COI location and its changes in angle and frequency are minimal over period T_{win}. The impact of bus location on system inertia can be ranked by sorting C_k in descending order. Here, C_p is defined as the second highest index, which indicates that to some extent bus p may represent the dynamics of system. When C_p/C_k is larger than a threshold, which is set as 0.6, it means the contribution from bus p cannot be neglected; if $dist(f_k, f_{k_{COI}})$ also satisfies the criteria, the new estimated RoCoF $\frac{\mathrm{d}f_{est}}{\mathrm{d}t}$ then can be re-estimated as follows,

$$\frac{\mathrm{d}f_{est}}{\mathrm{d}t} = \frac{C_k}{C_p + C_k} \cdot \frac{\mathrm{d}f_k}{\mathrm{d}t} + \frac{C_p}{C_p + C_k} \cdot \frac{\mathrm{d}f_p}{\mathrm{d}t}, \tag{2.17}$$

where ($\frac{\mathrm{d}f_k}{\mathrm{d}t}$ is the measured RoCoF on bus k, and $\frac{\mathrm{d}f_p}{\mathrm{d}t}$ is the measured RoCoF on bus p. If there is no feasible bus p, C_p is set to 0.

The system inertia can be then estimated following the procedures shown in Fig. 2.4. The proposed dynamic inertia estimation method can effectively detect events and estimate the inertia accurately using the data extracted from WAMS system. If the size of loss is accurately known, then the total system inertia E_{est} can be estimated,

$$E_{est} = \frac{f_0 \Delta P}{2 \frac{\mathrm{d}f_{est}}{\mathrm{d}t}}. \tag{2.18}$$

In a large system, inertia of a regional area can be estimated following a disturbance where the loss occurs outside the area [23], then "Δ" P can be extended to cover the total power crossing the area boundary,

$$\Delta P = \sum_{i \in B} \Delta P_i, \tag{2.19}$$

where ΔP_i is the change in boundary exchange power in MW, B is the set of boundary transmission lines.

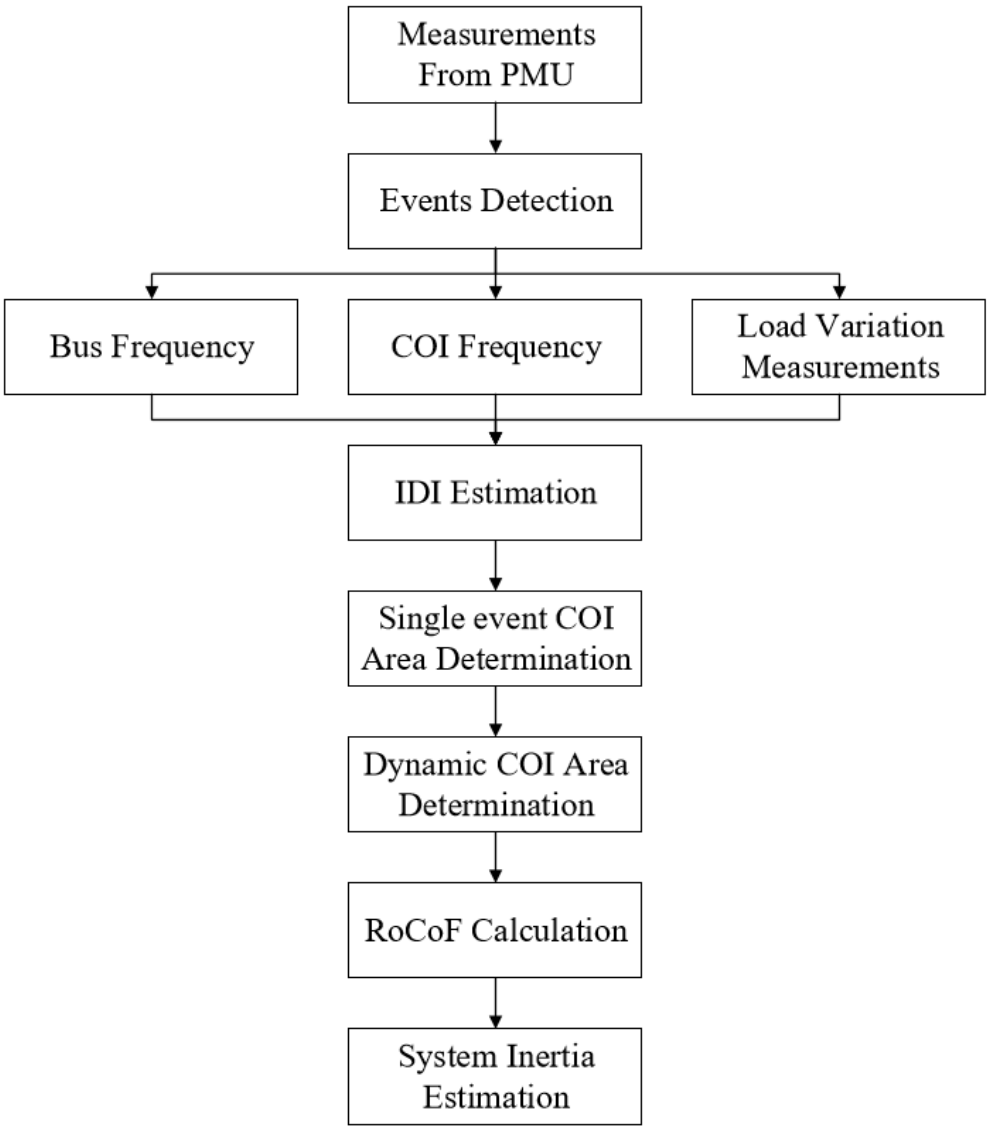


Fig. 2.4 System inertia estimation process on events.

2.4.4 Results Analysis

The dynamic system inertia estimation approach is evaluated on the IEEE 24-bus test system. The system has 24 buses (17 buses with loads), 38 branches, 33 generators. The total system load is 1,684 MW. The simulation model was implemented using TSAT, which is a core module of DSATools [24].

For an event where the disturbance appears as two distinct loss events, a non-monotonic frequency deviation may occur leading to erroneous IDI values. Integration period less than 0.5s can avoid the frequency distortion and make sure event detector captures more events. To determine the optimal integration period and ensure the IDI of the bus closest to COI location reaches the lowest value, sensitivity of integration period T is tested on the base model. Fig. 2.5 shows the results of the sensitivity test of integration period. It can be observed that the IDI on bus 18, 21, 22 and 23 reach the lowest value at the integration period of 0.2 s. Thus, the integration period is set to 0.2 s.

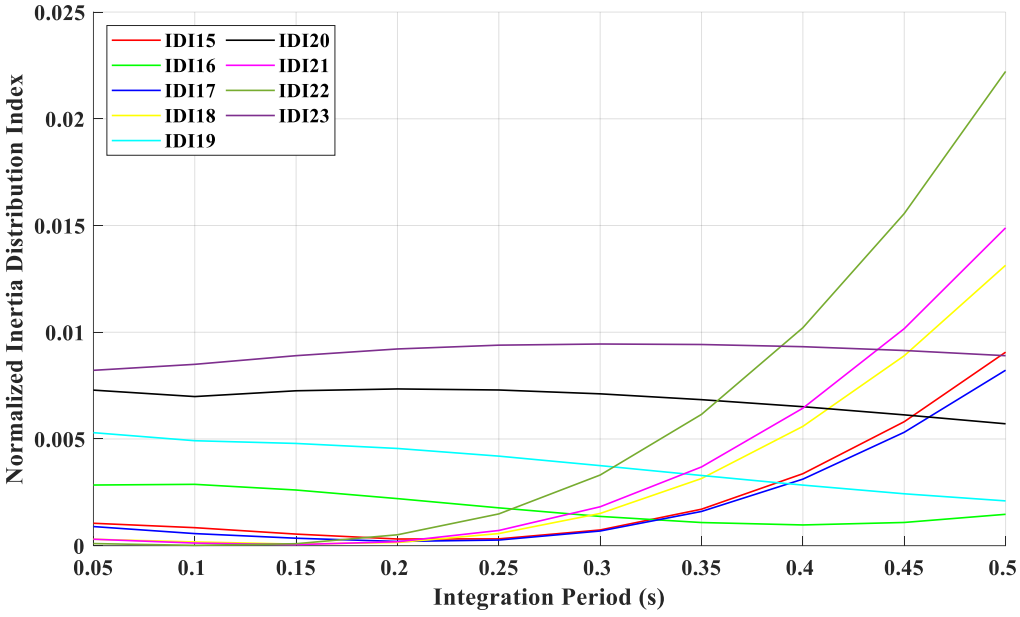


Fig. 2.5 Sensitivity of integration period.

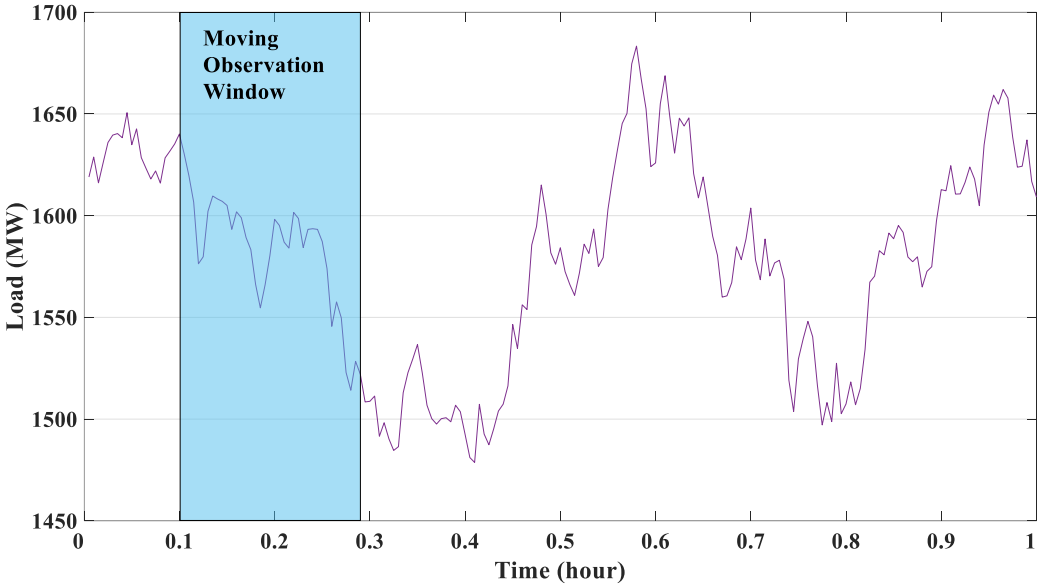


Fig. 2.6 Load variation profile.

The system total load profile is simulated for 60 minutes, as shown in Fig. 2.6 100 perturbations occur evenly on each bus with same probability, and a moving 10-minute observation window is applied. In the first 10-minute window, 16 events are detected. Fig. 2.7 shows the result of the identified COI area buses in the 24-bus system using the proposed method, larger yellow circle means higher C_k value of bus k. It is observed from Fig. 2.8 that bus 13 is identified as the COI bus of period T_{win}. While bus 23 also shows its close electrical distance to the COI location based on the proposed method. The results also show that frequency on bus 2, bus 7 and bus 22 contains harmonic waves. Inertia distribution indexes on these buses are estimated between 0.9 and 1, which indicates that these buses are relatively unstable and RoCoF measurements on these buses could suffer high bias.

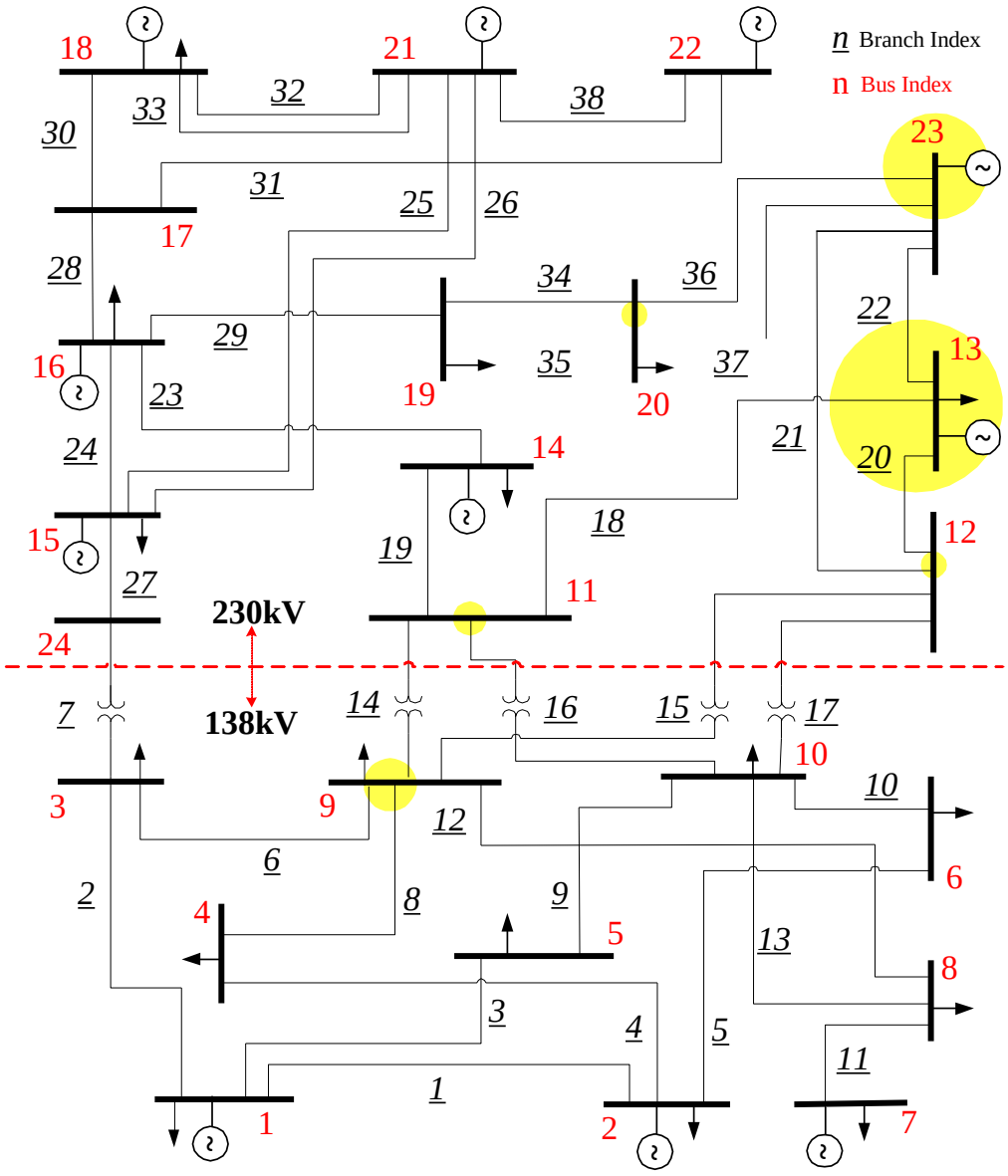


Fig. 2.7 Center of inertia area estimation.

Method of curve fitting is used to mitigate the impacts of measured transients in frequency following a loss which leads to significant large RoCoF value. Fig. 2.8 shows the measured RoCoF on the determined COI bus, the measured RoCoF is corrected from -0.074 Hz/s to -0.046 Hz/s. The results of system inertia estimation, under a selected event, obtained with the proposed method are displayed in Table 2.2.

For a single detected event, the system inertia H_{COI} is estimated as 30,044.6 MWs^2 using the traditional method, while the real system inertia is 31,525 MWs^2; The corresponding percentage estimation error $\%H_{dif}^{COI}$ is -4.70%. The system inertia estimated with the proposed dynamic inertia estimation method, H_{prop}, is 30,600.9 MWs^2 that corresponds to an estimation error $\%H_{dif}^{prop}$ of -2.93%, which shows substantial improvement over the results for traditional single event estimation method: the inertia estimation error dropped by 37.6% from 1,480.4 MWs^2 to 924.1 MWs^2.This corresponds to an overall estimation improvement of 1.77%.

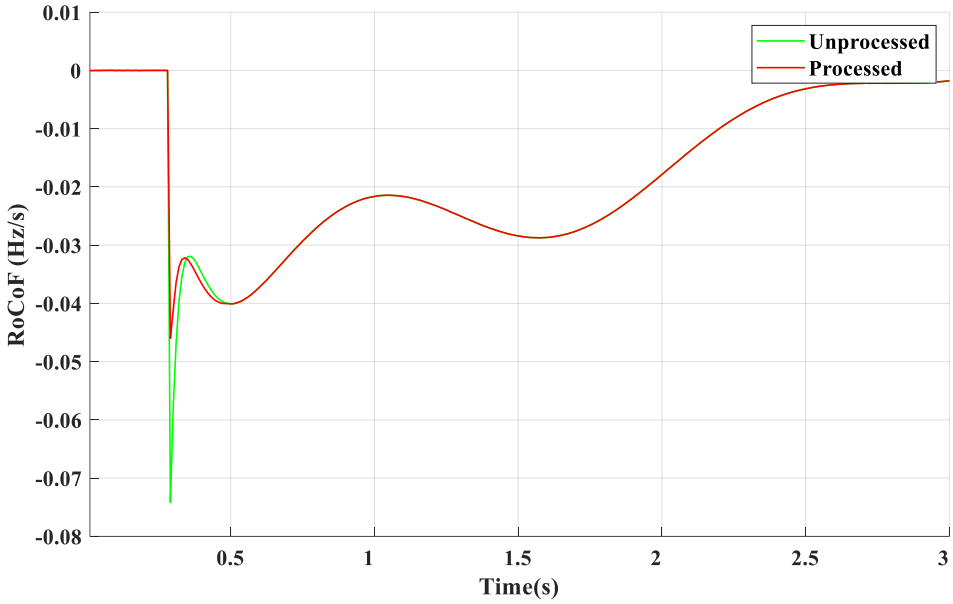


Fig. 2.8 RoCoF measurement in Center of inertia area.

Table 2.2 Results of Inertia Estimation with Various Methods

ΔP (MW)	H_{real} (MWs)	H_{COI} (MWs)	$\%H_{dif}^{COI}$	H_{prop} (MWs)	$\%H_{dif}^{prop}$
52.56	31,525.0	30,044.6	-4.70%	30,600.9	-2.93%

To evaluate the impact of renewable penetration on inertia distribution, another emulation was conducted under scenario of 20% wind penetration level: generators on bus 2, bus 7 and bus 13 are replaced with wind generators. Fig. 2.9 shows the results under scenario of 20% wind penetration level. A significant excursion of COI location can be observed due to installation of wind plants. It shows that the COI location shifts towards the area where many synchronous generators are located and synchronized online.

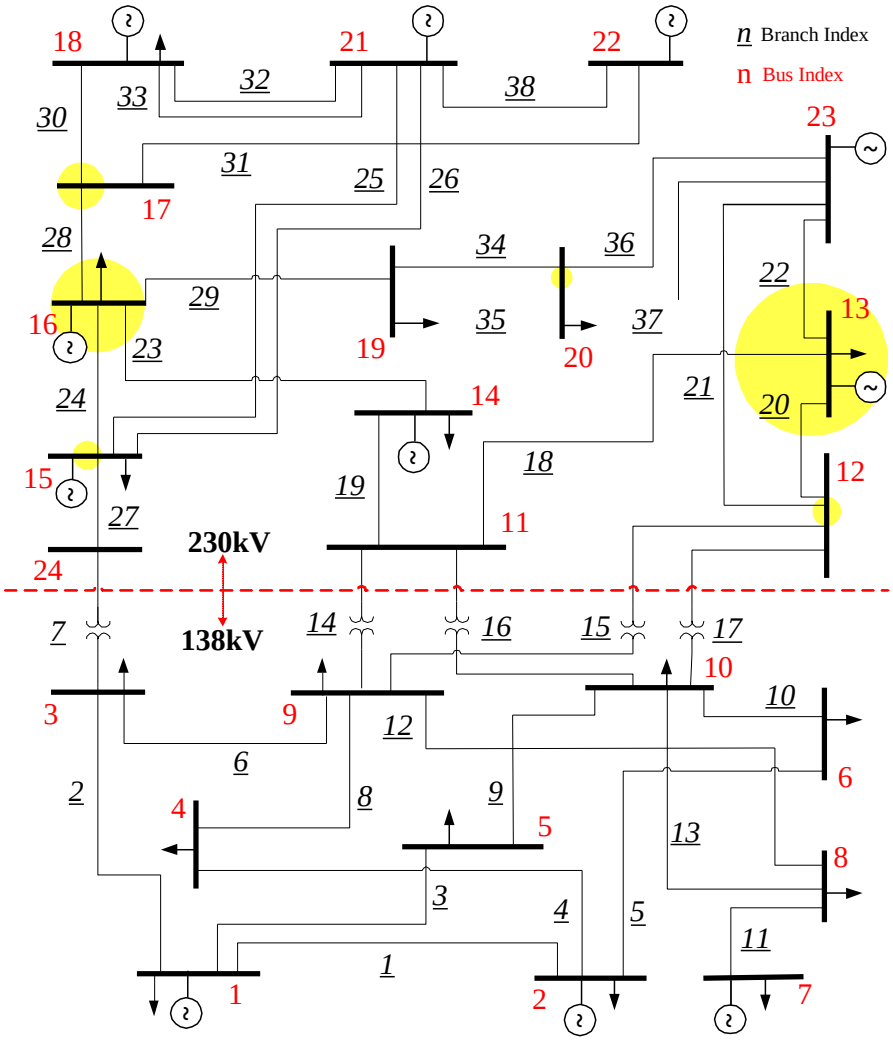


Fig. 2.9 Center of inertia area estimation with 20% wind generation penetration.

2.5 Summary

As the integration of variable renewable generation increases, the reduction in the system inertia poses a serious challenge for frequency regulation. Evaluation of the system inertia distribution traditionally based on a single disturbance event may be susceptible to power swings and oscillation between machines, which could deteriorate the accuracy of measurements and lead to high biased estimation.

This chapter proposes a proposed a PMU measurements-driven method which estimates the dynamic system inertia distribution and determines the center of inertia (COI) area. The frequency response under different renewable generation penetration levels is first tested. Then, an index based on electrical distance is used to estimate the inertia distribution over the entire grid. Butterworth filter is introduced to mitigate the impact of noise-induced measurement errors. To reduce the bias from location of measurements relative to the location of in-feed loss, disturbances on different buses over an observation window are combined; then a clustering algorithm based on electrical distance is utilized to accurately estimate the location of COI suitable for measurements. Areas with different inertia distribution levels are proposed to provide useful information to generation dispatch and frequency control.

The simulation results on the IEEE 24-bus system indicate that the power system with lower RES penetration shows a better frequency response, where the nadir is relatively higher and the RoCoF is less steep. The sensitivity test is then conducted to determine the optimal time length of integration period. The results also show that the proposed dynamic inertia estimation method utilizing the proposed

clustering algorithm has a better performance on system inertia estimation by incorporating the impact of perturbation location and oscillation between machines. Buses within COI area show relative stability comparing to the neighbor areas, measurements on these buses are relative robust and authentic. Unstable buses, which suffer harmonic waves, are also determined during the estimation process. Finally, the impact of geographic location of RES on COI area is examined. Overall, the proposed method is more robust and accurate for estimating system inertia distribution. Potential applications using the concept of inertia distribution estimation would be explored in the future.

3. MACHINE LEARNING ASSISTED INERTIA ESTIMATION

Inertia plays a crucial role in maintaining system stability. It provides a buffer against sudden changes in power supply or demand. Inertia estimation can ensure the accountability and reliability of inertia response through implementation of frequency control ancillary services [25]-[26]. Accurate estimation of power system inertia also helps system operators and planners assess the system's stability margins and ensure that sufficient inertia is available to maintain stable frequency and voltage levels. Power system inertia estimation is crucial for grid planning and operation. It provides insights into the system's dynamic behavior and helps in identifying areas where additional inertia resources may be required. This information aids in determining optimal generator placement, scheduling of generation resources, and the design of frequency control strategies to ensure reliable and efficient operation of the power system. Accurate estimation of system inertia helps in understanding the impact of RES on the overall system dynamics and enables the development of strategies to maintain grid stability and reliable operation. Protective relays and devices rely on accurate knowledge of the system's inertia to trigger appropriate actions during faults or abnormal conditions. Proper coordination of protective devices based on accurate inertia estimation helps in minimizing the impact of faults and ensures the safety and reliability of the power system.

In large-scale deregulated interconnection power systems, inertia information is only available within operators' own territories. Thus, system-wide inertia estimation is important for operators to provide frequency regulation services.

Traditionally, system frequency response is analyzed by looking at the collective performance of all generators using a system equivalent model. Based on event measurements and mathematical model, the system inertia could be estimated by the number and size of actively connected synchronous units.

Inertia estimation method based on mathematical model is highly dependent on accuracy of measurements from phasor measurement units (PMUs) or equivalent devices. However, modern power systems are connected to different devices which provide frequency regulation service, and inertia constant estimation purely based on synchronous generators is inaccurate [27]. Moreover, RES and other inverter-based sources are traditionally considered passive in terms of inertial response. The variability nature of RES also imports uncertainties into the system inertial response as well as system inertia constant [28]. Recent study in [29] shows that control schemes emulating synchronous machine response can be used to contribute system inertia. Therefore, the swing equation-based models may not be able to capture the entire characteristics. In addition, nonlinearities in system frequency response such as deadbands and saturations cannot be taken into considerations either. Thus, the estimated value based on mathematical model may suffer inaccuracy in various conditions.

The invention and development of PMU based wide area measurements systems (WAMS) enable the application of data-driven techniques in power system analysis [30]. The advantage of data-driven inertia estimation is its ability to capture the dynamic and time-varying nature of the power system. It can account for changes

in generation mix, renewable energy integration, and load variations, which have a direct impact on the effective inertia. A neural network-based inertia estimation technique is proposed in [31], which utilizes inter-area model information as neural network inputs and estimates the inertia constant as an output of the network. However, this approach only estimates the inertia constant for large systems with only traditional synchronous generation. A convolutional neural network (CNN) based model is proposed in [32], which estimates the system inertia through frequency response and RoCoF data. Graphs are a kind of data structure which models a set of objects (nodes) and their relationships (edges) [33]. Recent advances in deep neural network (DNN) offer an opportunity to integrate graph topology into a neural network, creating a graph neural network (GNN) model [34]. Power system can be represented as a graph with high dimensional features and interdependency among buses. This perspective may offer a better state of the art machine learning for power systems analysis.

3.1 System Perturbation using Probing Signal

The power system with frequency control loops is shown in Fig. 3.1. For primary frequency control, once the power mismatch event has occurred and frequency has started to drop from nominal value, the deviation is fed into closed control droop where turbine-governor counteracts the power mismatch [35]. Focusing on the short period following the disturbance, the model is simplified further when analyzing frequency dynamics before any secondary control gets involved.

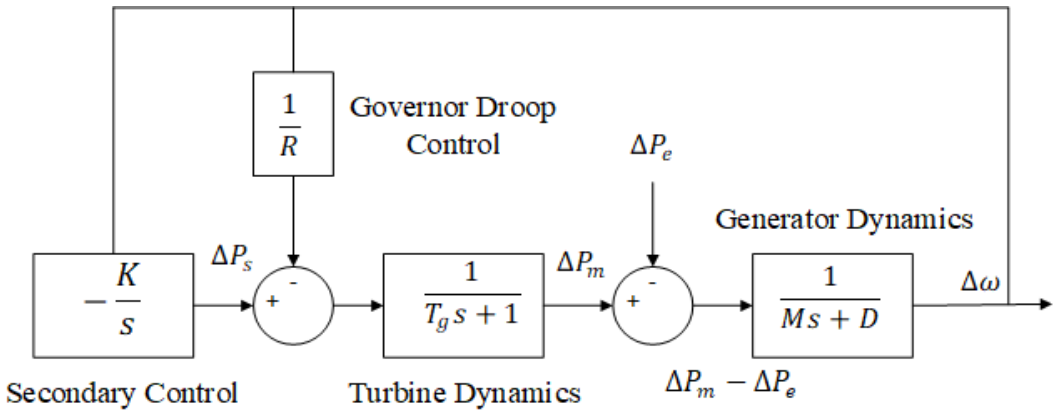


Fig. 3.1 Generator transfer function model.

Synchronized measurement technology makes it possible to sample analogue voltage and current wave data in synchronism with a global positioning system (GPS) clock, it can also record the corresponding frequency related data from widely distributed locations.

With PMUs widely used in modern power systems. Fig. 3.2 shows the topology of wide area monitoring system (WAMS). Measurements from PMUs are obtained from widely distributed locations, and synchronized with respect to a GPS clock [36]. Synchrophasor technologies allow direct measurement of frequency and bus voltages. With the development of PMU based WAMS, the accuracy of measurements improves significantly.

Most PMUs can calculate up to 30 to 60 samples per cycle with the GPS time stamp provided by hardware that has an accuracy of millisecond or higher [37], reporting rates of 10 - 240 samples per second are allowed. In this paper, the sampling rate of PMU is set to 200 per second.

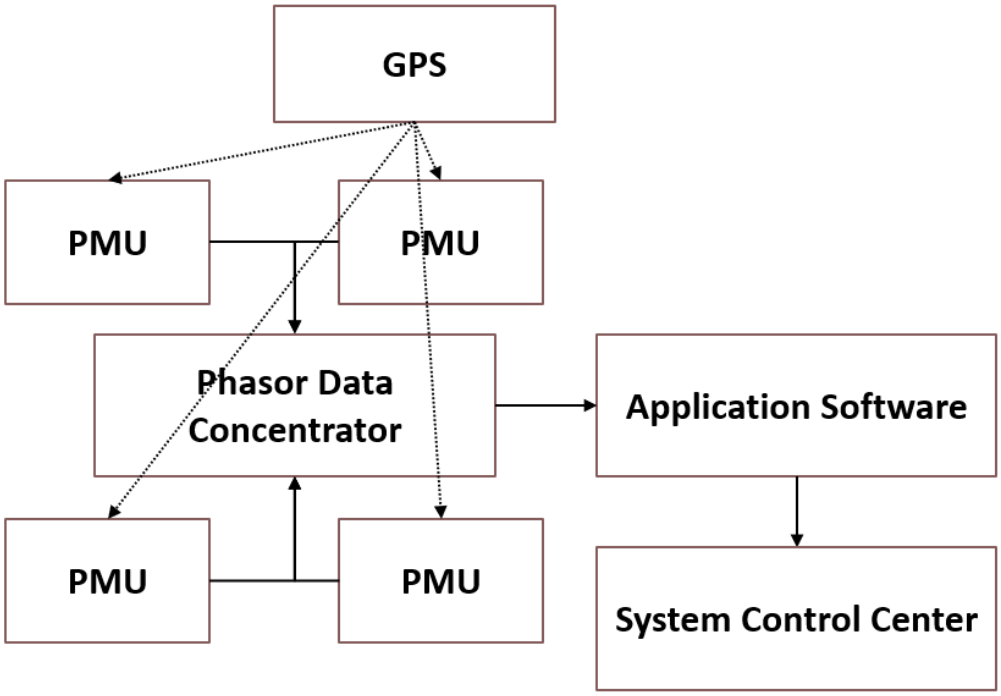


Fig. 3.2 Wide area measurement system.

Low level probing signal method has been conventionally used for generator dynamic studies. A modified form of detrended fluctuation analysis has been introduced to determine the event suitability for probing signal method [32]. With PMUs installed throughout the system providing highly accurate measurements and test improvement such as microperturbation method (MPM), the probing signal method can provide effective approach for system inertia identification without affecting system stability [38].

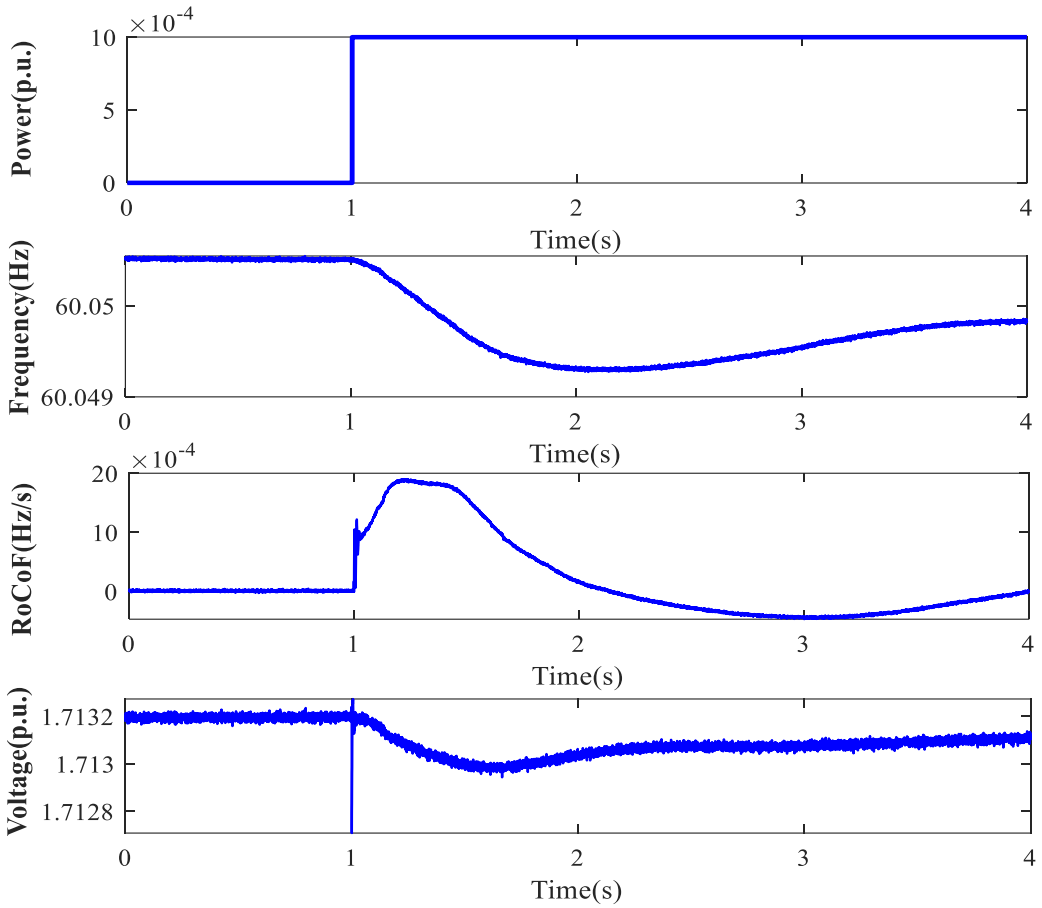


Fig. 3.3 A sample of probing signal, ambient measurements for P_E=0.001 p.u.

A sample probing signal, fed to the system with an amplitude of P_E, and corresponding PMU measurements are shown in Fig. 3.3. With varying system inertia and probing signal amplitude, a number of ambient measurements of $\Delta\omega$, $\Delta\dot{\omega}$ and ν can then be collected.

3.2 Neural Network based Inertia Estimation

3.2.1 Inertia Estimation using LRCN

Motivation of CNNs roots in the history of neural networks for graph data processing, recurrent neural networks (RNN) are utilized on graphs and cycles. Study

in [39] has shown that CNNs have the ability to extract spatial features and compose them to construct expressive representations. An example of a convolutional neural network is shown in Fig. 3.4 [40].

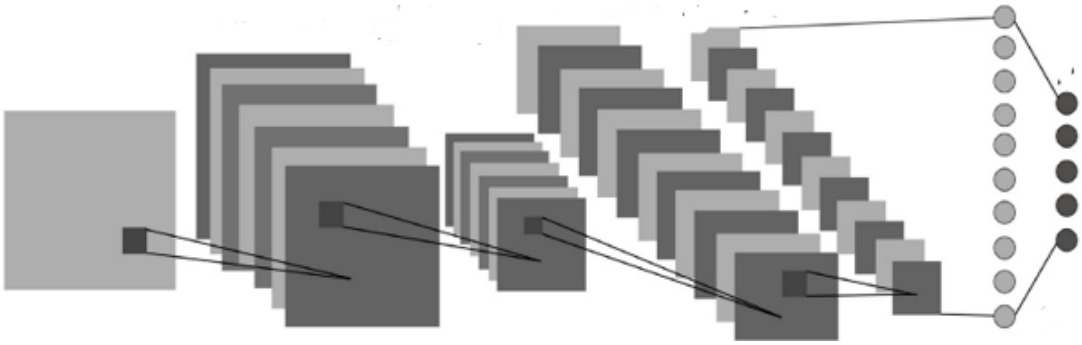


Fig. 3.4 Illustration of convolutional neural network architecture [40].

Long short-term memory (LSTM) is an extended frame of RNN which can exhibit temporal behavior of time-series input data. An LSTM cell typically compromises three gates: input, forget and output gates [41].

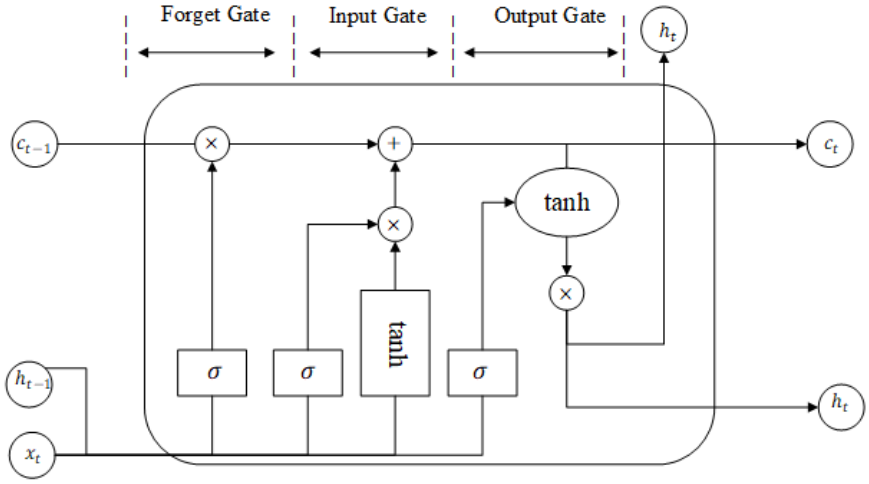


Fig. 3.5 Illustration of an LSTM cell.

The fundamental equations of LSTM network can be rep-resented as follows,

$$h_t = (1 - z_t) * h_{t-1} + z_t * h_t, \tag{3.1}$$

$$z_t = \sigma\left(W_f[h_{t-1}, x_t] + b_f\right), \tag{3.2}$$

where, x_t is the network input; h_t is the output state of the neuron from LSTM network; h_{t-1} is the previous state of the neuron; z_t computes the necessary information and removes the irrelevant data; σ is the sigmoid function; W_f is the weight and b_f is the bias.

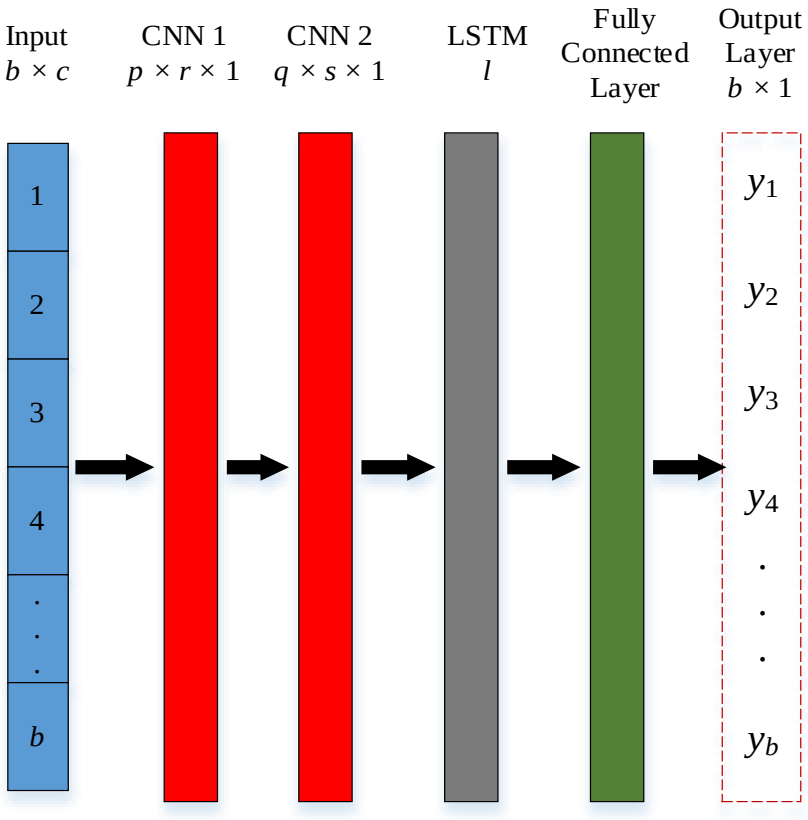


Fig. 3.6 General architecture of proposed LRCN model.

Since LRCN leverages the strength of rapid progress in CNN and has the ability to capture the dependencies in a sequence, it has been successfully used in computer vision, image processing, and other fields in signals and time-series analysis

[42]. The architecture of the proposed LRCN model is illustrated in Fig. 3.6 and can be trained to estimate system inertia from ambient measurements obtained from the PMU.

The proposed LRCN model first processes the measurements input with 1-D-CNN layers, whose outputs are then fed into the LSTM recurrent sequence model; and the fully connected layer finally produces the estimated inertia constant. The samples in training set are defined in batches which will be propagated through the networks. One epoch of training is completed when all the training samples have been passed forward and backward once. The number of iterations is defined as the number of passes, and each pass uses the same batch size that is the number of samples. At each training iteration, the LRCN model input size is $b \times c$, and the output will be a column vector of size b with inertia estimates for corresponding input in the time period. While the dimension of c is determined by the set of features and feature sampling rate.

The mean squared error (MSE) measures the average squared difference between actual and predicted outputs. The goal of training is to minimize MSE via back propagation which will provides best estimator [43]. The fully connected network used in this model includes one flatten layer and two hidden layers. MSE is defined as,

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \tilde{y}_i)^2, \tag{3.3}$$

where n is the total number of training samples, y_i is the actual value of i^{th} output, and $\tilde{y}_i$ is the estimated value corresponding to the i^{th} output. Similarly, the weight update equation via back propagation is expressed as,

$$w_{t+1} = w_t - \alpha \frac{\partial E_{MSE}}{\partial w_t}, \tag{3.4}$$

where w_t is the weight for current iteration, w_{t+1} is the updated weight for next iteration, α is the learning rate, and E_{MSE} is the MSE obtained from expression (3.4).

3.2.2 GCN based Inertia Estimation

Geometric deep learning is a recent emerging field. Traditional CNNs have limitations in processing graphical data which have explicit topological graph correlation embedded [44]. Recent advancement of CNN results in the rediscovery of GNNs. GCN has been developed by extending the convolution operation onto graphs and in general onto non-Euclidean spaces. Previous studies in [45]-[46] have proved that GCN provides state-of-arts performance in graph analysis tasks.

Power system is an interconnected network of generators and loads. The graph structure of the power system consists of nodes (buses) and edges (branches). The branches in the power system are undirected, such graphs provide information on buses and their connections. The convolution operator in propagation module is used to aggregate information from neighbours. Considering $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ as an undirected graph representing a power system, where $\mathcal{V} \in \mathbb{R}^N$ denotes its nodes and $\mathcal{E} \in \mathbb{R}^K$ denotes its edges. Let $A \in \mathbb{R}^{N \times N}$ be the adjacency matrix of $\mathcal{G}$, we can define a renormalization trick as,

$$V = \widetilde{D}^{-\frac{1}{2}}\tilde{A}\widetilde{D}^{-\frac{1}{2}}, \tag{3.5}$$

where $\tilde{A} = A + I_N$ represents an adjacency matrix with self-connections. Typically, the element at (i, j) of the adjacency matric A is defined as follows,

$$A_{ij} = \begin{cases} 1; \text{ if } \mathcal{V}_i, \mathcal{V}_j \in \mathcal{V}, (\mathcal{V}_i, \mathcal{V}_j) \in \mathcal{E} \\ 0; \text{ if } \mathcal{V}_i, \mathcal{V}_j \in \mathcal{V}, (\mathcal{V}_i, \mathcal{V}_j) \notin \mathcal{E}' \end{cases} \tag{3.6}$$

where $(\mathcal{V}_i, \mathcal{V}_j)$ denotes the branches from i to j. The diagonal degree matrix $\widetilde{D}$ for $\mathcal{G}$ is defined as $\widetilde{D}_{ii} = \sum_j \tilde{A}_{ij}$.

The graph convolutional layer is defined as follows,

$$F^l(X, A) = \sigma\left(VF^{(l-1)}(X, A)W_k^l + b^l\right), \tag{3.7}$$

where F^l is the convolutional activations and b^l is the bias matrix at the l-th layer; $F^0 = X$ is the input matrix. Fig. 3.7 demonstrates the message passing mechanism in forward propagation, a target node (bus 8) receiving information from its neighboring nodes.

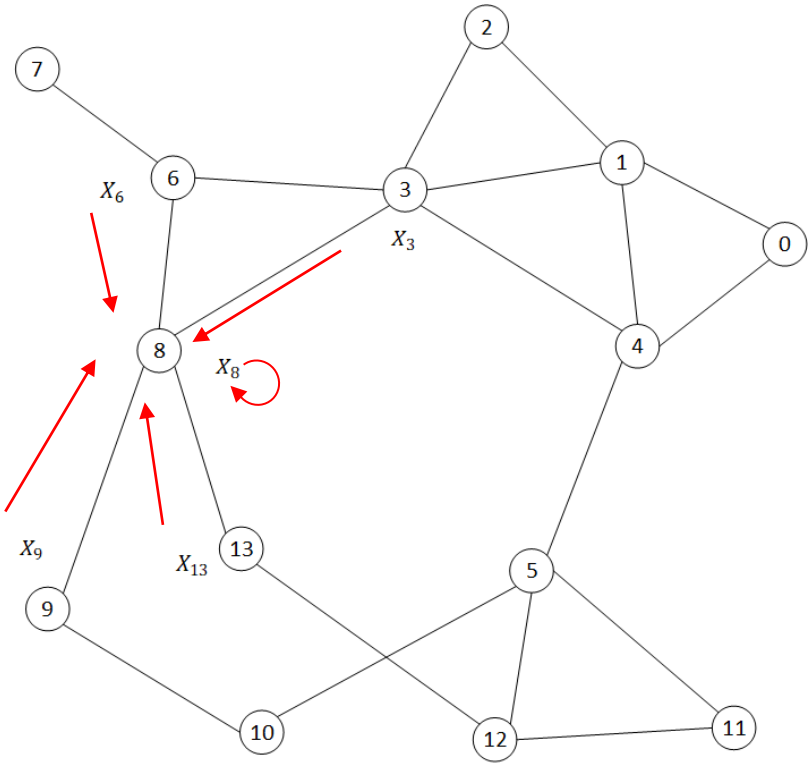


Fig. 3.7 Example of message passing mechanism.

3.2.3 Optimal PMU Allocation

Due to the high cost of having a PMU at each node, various techniques are used to solve the optimal PMU placement (OPP) problem for observability of the system for static and dynamic state estimators [47]-[48]. Traditionally OPP is a binary optimization problem, the objective functions considered in OPP problems in previous studies are mainly for minimization of the number of PMUs (3.8), maximization of observability (3.9), or both as a multi-objective function,

$$Min\ F_1 = \sum_{i=1}^{N} x_i, \tag{3.8}$$

$$Max\ F_2 = \sum_{i=1}^{N} o_i, \tag{3.9}$$

where $x_i, i = 1,2, \dots, N$ is elements of the PMU installation indication row vector X. For the base case, o_i denotes the observability of each bus. The complete topological observability can be expressed as follows,

The complete topological observability can be expressed as follows,

$$O = A \cdot X, \tag{3.10}$$

$$O \geq u, \tag{3.11}$$

where u is a row vector $N \times 1$ consisting of binary variables, representing that the monitored bus is observed by PMU.

3.3 Simulation Setup

The IEEE 24-bus system [49] was used for the experiment to collect the training data. The system has 24 buses (17 buses with loads), 38 branches, and 38 generators. The system inertia *M* typically ranges from 3s to 8s. Hence, to ensure the practicality of the proposed model, the measurements snapshots were collected for 11 different values of *M* from 3s to 8s with an increment of 0.5s following daily dispatches considering the RES penetration. Similarly, probing signals with 100 different values of P_E from 0.001 p.u. to 0.01 p.u. with an increment 0.0001 p.u. were used.

The modeling and simulation of the power system, along with data collection, were conducted in MATLAB/Simulink 2019b. The data pre-processing was conducted in both MATLAB and Python. The proposed LRCN and GCN based models were developed in Python using Keras and PYG. The initial data analyzed in this study

were acquired from PMUs with a sampling rate larger than 200 Hz; nodal measurements of $\Delta\omega$ and $\Delta\dot{\omega}$ are obtained. By using only one second sampling frame for normalization, the real-time applicability of this method is maintained. Similarly, following the same pattern we obtained nodal voltage measurement v. Since the training data come from the ambient measurements of all PMUs may suffer different sampling rate which is larger than 200 Hz, without dimension reduction process the original training data would increase the complexity of the model and may also lead to overfitting. In addition, there are bad data in the raw measurements which would introduce high error in analysis results. Therefore, the bad data points are first discarded from raw measurements, and then we downsample the measurements to 200 Hz for next step. Additionally, Gaussian noise signal is added to the constituent tonic to mimic the noisy measurements. Different signal-to-noise ratios (SNR) are investigated in this paper. The data are collected between multiple sessions, all the measurements are normalized by employing min/max normalization between [0, 1].

To find the best time frame of data extraction, different time windows of the ambient measurements are determined: (1) the time frame is first chosen from 0s to 1s following the perturbation, where initial RoCoF is included; (2) the second time frame is from 0.5s to 1.5s after the signal infeed. With $\Delta\omega$ and $\Delta\dot{\omega}$ as basic features combination, the coefficient of determination and validation accuracy are used as evaluating metrics.

Wrapper feature selection method is utilized to determine the optimal combination of features for inertia estimation model training: (i) the proposed LRCN

model is used as the inertia estimator, (ii) accuracy score is used as the evaluation metric, (iii) greedy forward selection as the subset selection policy. The specific metrics for feature evaluations is expressed as follows,

$$ACC = \frac{|\{y \in TD|\ |y - \tilde{y}| \leq \mu\}|}{|\ TD|}, \tag{3.12}$$

where ACC denotes the proportion of the correctly predicted values with μ tolerance. TD denotes the validation dataset, and μ is the predetermined threshold.

Constraints of PMU locations are not considered in the base case. With a resampling rate of 200Hz, measurements on each PMU node gives 200 data points at a sampling frame of 1s. In this paper, ambient measurements of $\Delta\omega$, $\Delta\dot{\omega}$ and v on generator buses are assumed as available measurements.

Availability of PMU data are affected by several realistic factors, which have been studied in [18]. Given limited resources, the objective function (16) of OPP is applied in this work to maximize the system observations. Zero injection bus effect is also added. Traditionally, zero injection bus (ZIB) means no load or generator is connected to it. Since we are more interested in the statement of generators, the impact of zero generation injection bus (ZGIB) is considered in this work where generator is the only factor. The topological observability constraint of each of the bus connected to a ZGIB is updated by introducing virtual connections to every other non-ZGIB bus connected to that same ZGIB,

$$\tilde{o}_i = \sum_{j}^{N} A_{ij}x_j + \sum_{j}^{N} w_{ij}x_j \ \forall i = 1,2,\dots,N, \tag{3.13}$$

where w_{ij} is an auxiliary binary variable: 1, if buses i and j are both connected to the same ZGIB; otherwise set zero. The GCN model consists of one GNN layer and two hidden fully connected layer. The dimension of input features is defined based on the number of available PMUs which gives $n \times d$, where d is the nodal features of GCN layer. The fully connected layers are set as 64 and 128 respectively.

3.4 Results Analysis

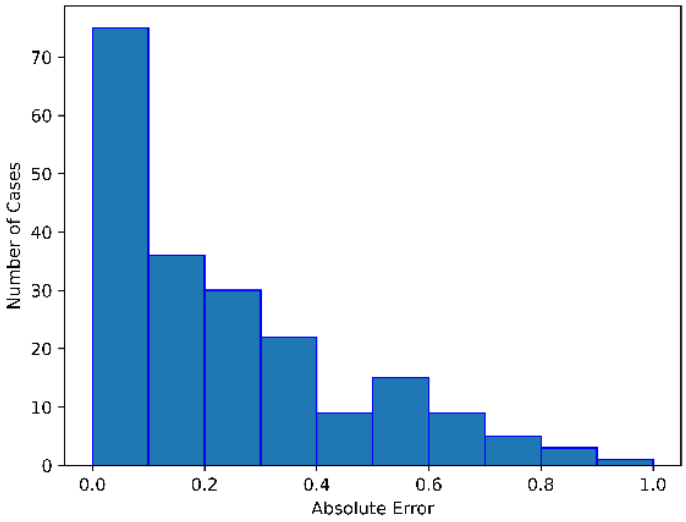
To investigate the best time period for feature extraction, all PMUs are considered available in base case. Measurement $\Delta\dot{\omega}$ is selected as training feature combination; training data extracted from two periods are then fed into the aforementioned models, including the proposed LRCN and GCN models.

Performances of different inertia estimation models are summarized in Table 3.1. As it can be observed, the DNN model has the lowest validation accuracy for both scenarios. CNN based model has a relatively higher validation accuracy, showing advantage in processing spatial data. The proposed LRCN model has a validation accuracy of 96.89% with a tolerance of 0.5s for features extracted from period 0.0 - 1.0s period, while it is only 82.57% for the use of features extracted from period 0.5 - 1.5s. GCN based model has the highest validation accuracy in both scenarios at 98.15% and 83.19% respectively. On the whole, the features extracted from the time frame following the disturbance contain prominent inertial response information, and accordingly have a positive impact on the overall performance of inertia constant estimation model.

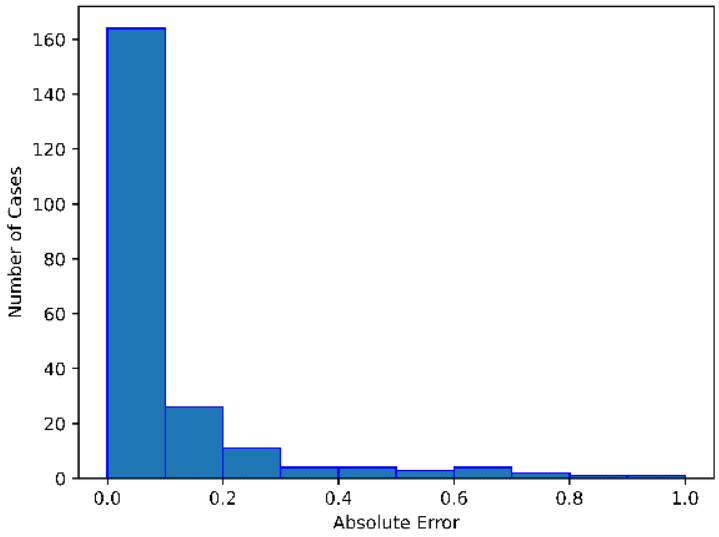
Table 3.1 Performance Comparison for Different Models

Models	Period (0.0-1.0s)		Period (0.5-1.5s)	
	Validation Accuracy	MSE	Validation Accuracy	MSE
DNN	93.27%	0.065	78.18 %	0.314
CNN	94.46%	0.052	80.36%	0.236
LRCN	96.89%	0.028	82.57%	0.215
GCN	97.48%	0.022	83.19%	0.203

Fig. 3.8 compares the distribution of absolute prediction error for the LRCN model with features extracted from 0.5s - 1.5s and 0.0s - 1.0s respectively. Using features extracted from period 0.0 - 1.0s, the coefficient of determination of the proposed LRCN model is 0.9625, which is higher the use of features extracted from period 0.5 - 1.5s at 0.7619.



(a) Features extracted from 0.5s - 1.5s.



(b) Features extracted from 0.0s - 1.0s.

Fig. 3.8 Absolute error of prediction with LRCN model using features extracted from different time periods.

Table 3.2 Comparison of Different Features Sets for LRCN Model

Features Set	$\Delta\omega$	$\Delta\dot{\omega}$	$\Delta\omega + \Delta\dot{\omega}$	$\Delta\omega + \Delta\dot{\omega} + v$
Validation Accuracy	80.30%	96.89%	97.34%	95.76%
MSE	0.296	0.032	0.025	0.030
Coefficient of Determination	0.8945	0.9585	0.9725	0.9564

Optimal combination of features extracted from 0.5s - 1.5s period is first selected through greedy forward selection. Table 3.2 compares the results with LRCN model as performance estimator. It can be observed that only considering $\Delta\omega$ measurement as input feature provides a validation accuracy of 80.30%. Combination of $\Delta\omega$ and $\Delta\dot{\omega}$ has a highest validation accuracy at 97.34% with 0.5s tolerance, which

outperforms other feature combinations. Thus, $\Delta\omega + \Delta\dot{\omega}$ is selected as the optimal combination for model training.

Fig. 3.9 and Fig. 3.10 depict the evolution of MSE losses on the training and validation datasets over the training process of the proposed models. As it can be seen, for both LRCN and GCN cases, MSE decreases as the number of epochs increases. In terms of MSE, the validation loss of LRCN model drops faster and reaches minimal value at 0.025. It should be noted that a sudden drop is observed in GCN model training, and the GCN based model has a lower validation loss at 0.020.

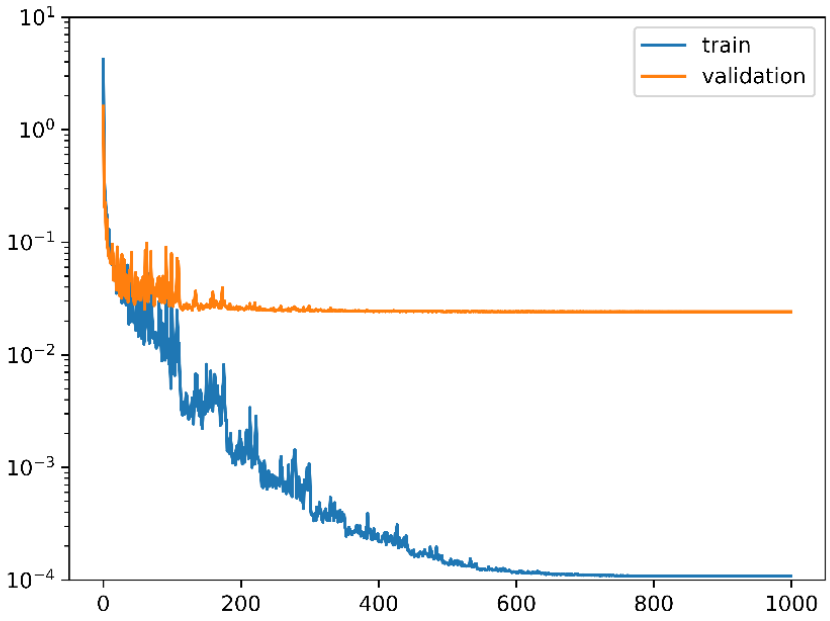


Fig. 3.9 The learning curve of the proposed LRCN model: MSE losses versus the number of epochs.

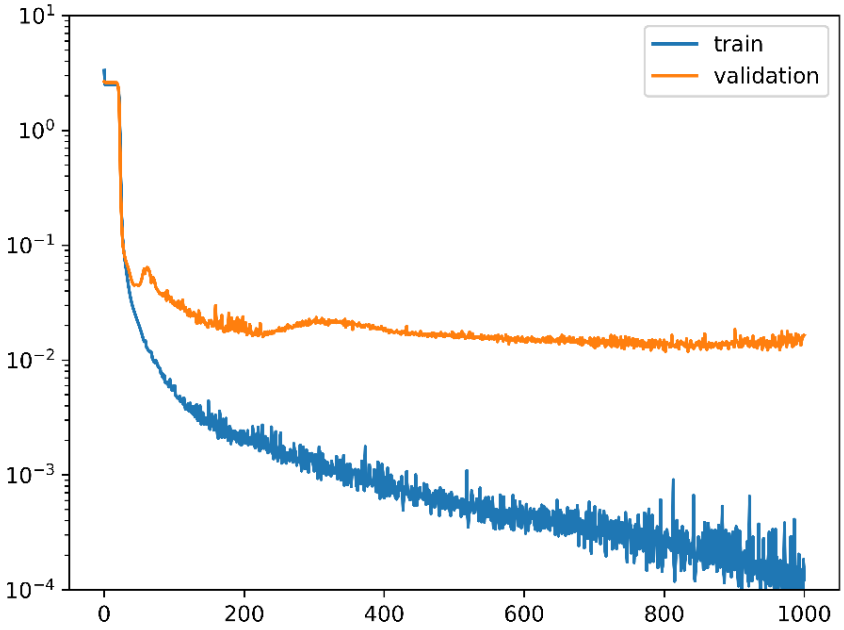


Fig. 3.10 The learning curve of the proposed GCN model: MSE losses versus the number of epochs.

The proposed LRCN and GCN based approaches are then compared with benchmark algorithm [31] in Table 3.3, $\Delta\omega$ and $\Delta\dot{\omega}$ are used as primary features to train the model. Both algorithms are employed to train the inertia constant estimation model. The results show that the CNN model has a validation accuracy of 95.18% with 0.5s tolerance, which is higher than the validation accuracy of DNN model at 93.45%. These results reflect that CNN based model has the better capability to process spatial data comparing to traditional DNN model. Nevertheless, the proposed LRCN model has a validation accuracy of 97.34% with 0.5s tolerance, and GCN model has a validation accuracy at 98.15%. Additionally, the coefficient of determinations of LRCN model and GCN model are 0.9725 and 0.9826 respectively, which are higher than the benchmark CNN and DNN models. An explanation could be that the proposed LRCN and GCN models are more efficient algorithms to identify

critical temporal information and graphical information embedded in the collected power system data. Combining Table III and Figs. 9-10, we can observe that GCN has the highest validation accuracy and the lower MSE, indicating that GCN has a better performance in processing graphical data.

In addition to the ideal condition, the proposed models are compared with the benchmark models under high noise conditions. Study in [50] has shown that a SNR of 45dB is considered as a good approximation of noise power under realistic conditions.

Table 3.3 Comparison of Models with Optimal Feature Combination

Model	Validation Accuracy	Coefficient of Determination	MSE
DNN	93.45%	0.9224	0.058
CNN	95.18%	0.9369	0.045
LRCN	97.34%	0.9725	0.025
GCN	98.15%	0.9826	0.020

Table 3.4 shows the inertia estimation accuracy of all models with combination of $\Delta\omega$ and $\Delta\dot{\omega}$ as training features. After adding additional Gaussian noise signal with a SNR of 45dB to the ambient measurements, the overall MSE increases for both models, while the validation accuracy reduces accordingly. Understandably, a significant reduction in validation accuracy can be observed in both cases.

Table 3.4 Comparison of Different models with SNR at 45dB

Model	w/o SNR	w/ SNR at 45dB
DNN	93.45%	90.84%
CNN	95.18%	92.13%
LRCN	97.34%	93.25%
GCN	98.15%	93.87%

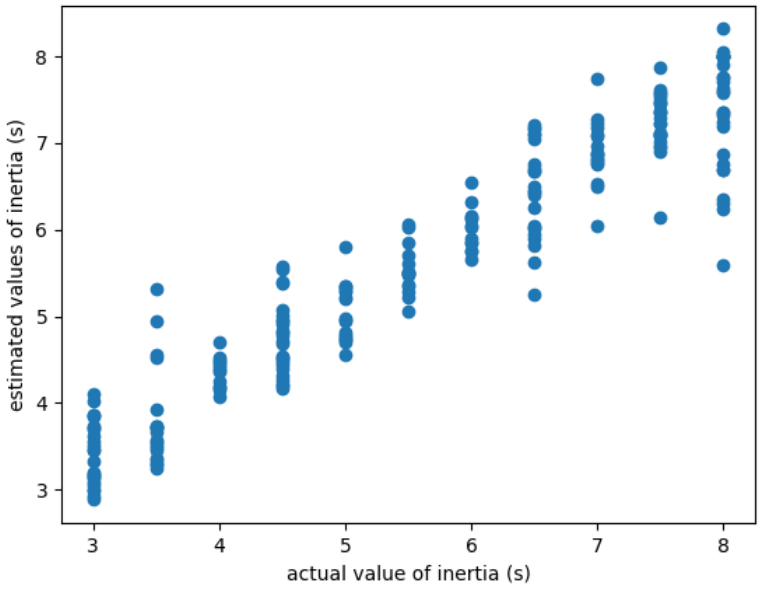


Fig. 3.11 Prediction results of the benchmark CNN model with SNR at 45dB.

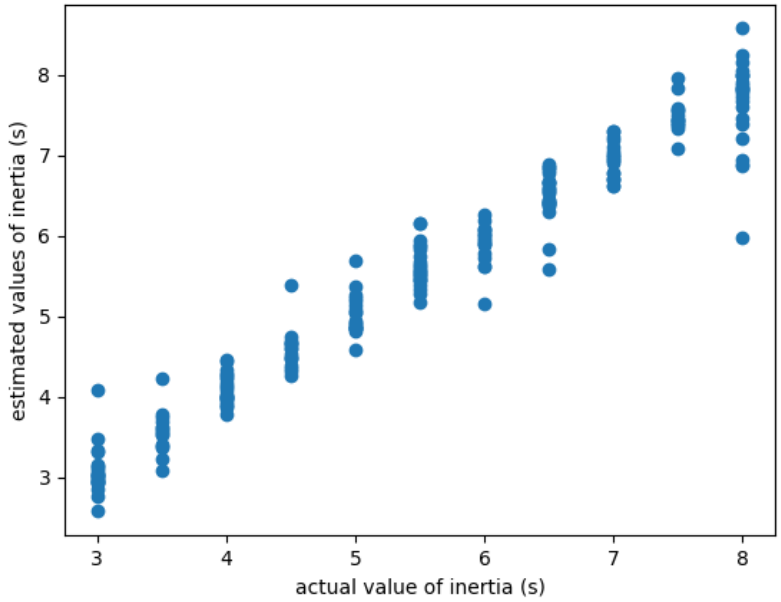


Fig. 3.12 Prediction results of the proposed LRCN model with SNR at 45dB.

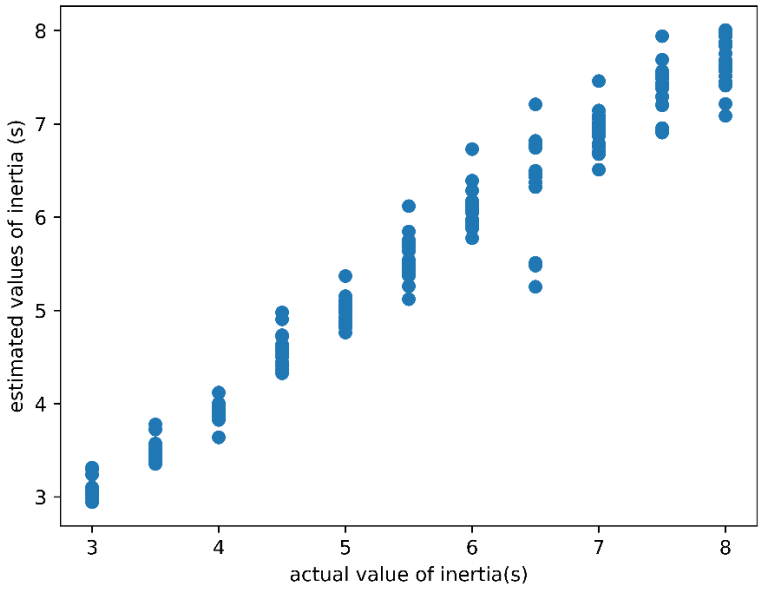


Fig. 3.13 Prediction results of the proposed GCN model with SNR at 45dB.

The scatter points of CNN model, the proposed LRCN and GCN models when SNR of 45dB is applied are depicted in Figs. 3.11-3.13. Results show that only

considering measurements of frequency related data may suffer high error when noise is applied.

The method described in this research uses a wrapper feature selection process and then. Results of different feature combinations are listed in Tables 3.5 and 3.6. It can be observed that the proposed LRCN model has a validation accuracy of 93.25% with 0.5s tolerance under the condition of SNR at 45dB. With voltage measurements added, the validation accuracy is improved to 93.87% with 0.5s tolerance.

Table 3.1 Comparison of LRCN Model (SNR 45 dB)

Features Set	$\Delta\omega + \Delta\dot{\omega}$	$\Delta\omega + \Delta\dot{\omega} + v$
Validation Accuracy	93.25%	93.76%
MSE	0.119	0.098
Coefficient of Determination	0.9032	0.9156

Table 3.2 Comparison of GCN Model (SNR at 45 dB)

Features Set	$\Delta\omega + \Delta\dot{\omega}$	$\Delta\omega + \Delta\dot{\omega} + v$
Validation Accuracy	93.87%	94.56%
MSE	0.088	0.071
Coefficient of Determination	0.9227	0.9449

In summary, (i) under a low noise condition, measurements of $\Delta\omega$ and $\Delta\dot{\omega}$ are the optimal set of features suitable for inertia estimation; (ii) with SNR of 45dB added, the performance of benchmark model decreases significantly, while the proposed models based on optimal features combination of $\Delta\omega$, $\Delta\dot{\omega}$ and v shows higher robustness and better performance.

The application of proposed ZGIB based OPP approach to IEEE 24-bus system is carried, and the corresponding impact on inertia estimation models are investigated in this section.

Table 3.7 listed the results of PMU location considering the proposed ZGIB-OPP method. Given the total number of PMUs limited by two, the proposed OPP method considering ZGIB suggests locating PMUs at buses 2 and 16 for the best inertia estimation performance. When the total number of PMUs increases, the proposed ZGIB OPP adds more buses into the optimal set, indicating the consistency in optimal PMUs locations.

Table 3.3 Optimal Locations of PMUs given Limited Resources

No. of PMUs	**2**	**3**	**4**	**5**
Bus #	2, 16	2, 16, 21	2, 16, 21, 23	2, 13, 16, 21, 23

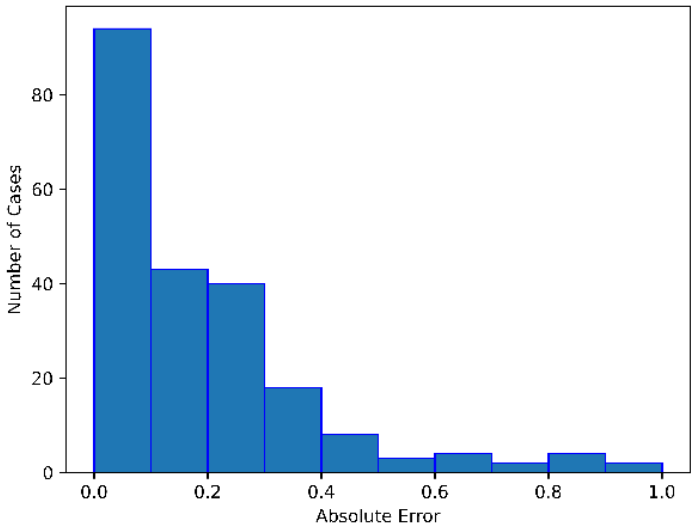


Fig. 3.14 Distribution of absolute values for random PMU locations.

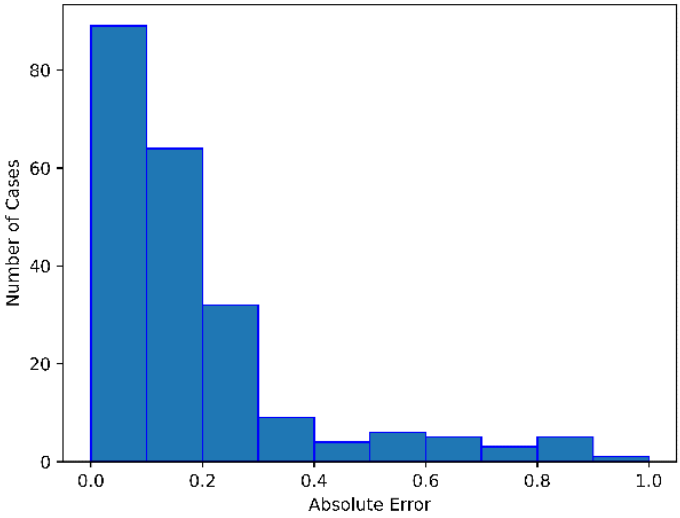


Fig. 3.15 Distribution of absolute values for ZGIB-OPP.

We apply the proposed ZGIB-OPP to IEEE 24-bus system, and the system wide measurements of limited channels are then obtained for model training. Fig. 3.14

and Fig. 3.15 compare the distribution of prediction absolute error for different PMUs settings using LRCN model. Comparing to the random PMU location case, the proposed ZGIB-OPP improves the performance of LRCN model, resulting in more samples with absolute error landed within 0.1-0.2s.

Table 3.4 Comparison of Different Models with Five PMUs Limit

PMUs Locations	ZGIB-OPP	Random PMUs Location	Full PMUs
DNN	91.36%	90.00%	93.45%
CNN	92.73%	90.45%	95.18%
LRCN	95.26	92.36%	97.34%
GCN	95.89%	94.65%	98.15%

Results listed in Table 3.8 also show that with limited PMU channels, as well as combination of $\Delta\omega$ and $\Delta\dot{\omega}$ as input feature, the validation accuracy of all models drops accordingly. For DNN and CNN model, the validation accuracy drops to 90.00% and 90.45% respectively. Validation accuracy of LRCN and GCN model drops slightly when there are limited channels or under random PMU locations. With ZGIB-OPP applied, GCN based model has the highest accuracy of 95.89%. It should be noted that applying ZGIB-OPP improves the performances of all models, indicating that the overall observability of the WAMS is improved through the proposed ZGIB-OPP algorithm.

3.5 Summary

Data-driven inertia estimation refers to the use of historical power system data to estimate the effective inertia of a power system. Instead of relying on static values or assumptions, data-driven methods leverage the availability of system measurements. Machine learning techniques, such as neural networks, support vector machines, or regression models, are commonly applied to develop data driven models that can capture the relationship between the available data and the inertia of the system.

Deep neural networks have been applied for inertia estimation as extensive amounts of data can be obtained through power system digital equipment and advanced measuring infrastructures such as PMUs. In this chapter, LRCN and GCN based learning algorithms are proposed to estimate system inertia constant. System wide ambient measurements based on WAMS are used as candidate features for model training, and a wrapper feature selection is also used to optimize the feature combination. Considering the limitation on PMU settings, an ZGIB-OPP method is proposed to maximize the observability of the WAMS given limited PMU resources. Results indicate that the proposed LRCN and GCN models have better performances than the benchmark DNN and CNN models. The proposed LRCN model and GCN model also show high robustness under conditions with higher noises. The proposed ZGIB-OPP method has been proved to be capable of improving the performance of all implemented models. Considering that the IEEE 24-bus system model used in this research has a mix generation of both synchronous generators and inverter-based

resources, the proposed approach can also be applied to estimate inertia constant in realistic conditions.

4. PHISICS-BASED LOCATIONAL ROCOF-CONSTRAINED UNIT COMMITMENT

4.1 Literature Review

Power system frequency stability refers to the ability of the system to maintain its frequency within an acceptable range under normal operating conditions. The nominal frequency, such as 50 Hz or 60 Hz, is maintained by balancing the power generation and consumption in the grid. Frequency stability is crucial because deviations from the nominal frequency can affect the performance of electrical equipment and disrupt the synchronization of interconnected systems. RoCoF is a measure of the rate at which the frequency is changing over time. It indicates the speed at which the system frequency is deviating from its nominal value. RoCoF is important because it provides information about the dynamic behavior of the power system and helps in assessing the system's response to disturbances. When there is a sudden change in power generation or demand, the RoCoF may deviate from its normal value. A stable RoCoF ensures that the frequency deviation does not escalate rapidly, and that the system can recover to its steady-state condition without instability or severe voltage fluctuations.

When RoCoF or frequency deviation exceeds certain thresholds, protection devices would disconnect generators from the grid [51][52]. In fact, RoCoF related protection was found to be one of the main factors that limit the shift towards a 100% converter-based power system in Ireland [53]. If the frequency drops rapidly due to insufficient inertia, conventional generators may not be fast enough to respond to the deviation resulting in lower nadir and activation of load shedding. On August 9, 2019,

over a million customers were affected by a major power disruption (mainly came from offshore wind farms) that occurred across England and Wales and some parts of Scotland, and the frequency of the system hit 48.8 Hz [54]. The impact of RESs integration on power systems has been studied to address frequency stability challenge. In [55], the authors explored the massive deployment of grid-forming converters and its effects on frequency stability, results show that the system stability is lost when converter-based resources penetration reaches 80%. The Electric Reliability Council of Texas (ERCOT) studied the effect of low inertia on grid security and reliability [56]-[57]. To ensure frequency stability, the enhanced frequency response has been introduced in Great Britain recently which includes technologies like battery storage, interconnectors, and demand response [58]. Frequency control ancillary services have been implemented in the Australian National Electricity Market to maintain the system stability [59].

4.1.1 Operation in Low Inertia Power System

The economic objective function in SCUC takes into account the cost of generation and other factors. Frequency-related constraints can also be included in the objective function to reflect the importance of maintaining frequency stability. This ensures that the optimization process considers the impact of different generation and dispatch decisions on the system frequency.

Several transmission system operators impose extra inertia constraints in the conventional unit commitment model to keep the minimum amount of synchronous inertia online [60]. Federal Energy Regulatory Commission has suggested that

frequency control capabilities will need to be imposed into traditional unit commitment [61]. EirGrid has also introduced a synchronous inertial response (SIR) constraint to ensure that the available inertia does not fall below a static limit of 23,000 MWs in Ireland [62].

The optimal power flow with primary frequency response related constraint is investigated in [63]. Refs. [64] and [65] derived the analytical expression of the system frequency response model, frequency-related constraints are incorporated into traditional unit commitment: enforcing limitations on RoCoF, frequency nadir and steady-state error that are derived from a uniform frequency response model. Ref. [66] studies the system equivalent model based RoCoF constrained SCUC (ERC-SCUC). A novel mixed-integer linear programming (MILP) unit commitment formulation was proposed in [67], which simultaneously optimizes energy production and the allocation of inertia. Despite these great efforts of modeling the classic system frequency response, previous studies rely on a simplification of the actual frequency dynamics and collective performance of system that cannot be able to capture the entire system characteristics.

The distinct frequency responses experienced by different bus have been observed in recent publications. System equivalent model-based operations may fail to supply sufficient ancillary services against contingency. Ref. [68] has shown that generators on the buses adjacent to the event may suffer higher RoCoF comparing to distant buses. Similar conclusion has been made in [18] that the relative location of measurement point to disturbance is a pertinent factor in system inertial response, the

RoCoF is usually higher for location where networks are weakly interconnected. To guarantee frequency stability accounting for spatial variation, a mixed analytical numerical approach based on multi-regions has been studied in [4], which investigated the model combining evolution of the center of inertia and certain inter-area oscillations, stability conditions are proposed to co-optimize the existing ancillary services. Regarding oscillations mitigations, network coherency was considered as an alternative performance metric in [69], and a system dynamic model was formulated to determine the optimal allocation of additional inertia based on the disturbance location against high nodal frequency excursions.

4.1.2 Additional Inertia

Synchronous condenser (SC) traditionally provides reactive power compensation and keeping voltage stability in power systems [70]. Ref. [71] investigates the approach of using SC for decreasing under frequency load shedding. Interaction between active power and reactive power channels of synchronous condensers to improve primary frequency control is studied in [72]. SC is an alternator that can provide inherently inertial response due to the electromechanical coupling with the grid [73]. The kinetic energy stored in its rotating mass can naturally counteract the frequency change during disturbances.

Besides SC and induction motors, wind turbine active control and virtual inertia (VI) wind power has been considered as an important RES in modern power system. Active control of wind turbines enables the synthetic inertia response of the wind power plants [74]. A wind turbine may reduce its rotational speed in order to

release the reserved power generation capacity if it is working in the over-speed zone instead of maximal power points tracking (MPPT). Such inertia control enables wind power plants to increase their power outputs to counteract the power imbalance in the event by discharge the kinetic energy stored in wind turbines.

Recently, a plethora of strategies has been proposed to emulate synchronous machine models of various degrees of fidelity under names such as synthetic inertia or virtual synchronous machine [75]-[77]. These studies have demonstrated the efficacy of virtual inertia (VI) method which imitates the kinetic inertia of synchronous generator to improve the system dynamic behavior [78][102]. There are different implementations for synchronous machine response emulation with varying fidelity. Virtual inertia techniques for solar PV generation have been investigated by [79]. It is noted that the virtual inertia requires fast responsive energy buffer; the kinetic energy in a wind turbine and the energy in a battery are limited energy resources for virtual inertia responses.

4.2 System Dynamic Modeling

The synchronous generator provides inertia to the power system through strongly coupled mechanical dynamics and electrical dynamics. Following a sudden change in load or a generation contingency, the dynamic behavior of the system frequency can be described using the swing equation of system equivalent single-machine representation,

$$P_m - P_e = M_{sys}\frac{\partial \Delta\omega}{\partial t} + D_{sys}\Delta\omega, \tag{4.1}$$

where M_{sys} and D_{sys} are the aggregated system inertia constant and damping coefficient corresponding to the committed synchronous generators respectively. P_m is the mechanical input power change. P_e is the electrical output power change.

The transmission network can be modeled as a graph consisting of nodes (buses) and edges (branches). Using the topological information and the system parameters, the swing equation can then be extended and reformulated to describe the oscillatory behavior of each individual bus as follows [4],

$$m_i\ddot{\theta}_i + d_i\dot{\theta}_i = P_{in,i}^{(0)} - P_{e,i}, \qquad i \in \{1, \dots, n\}, \tag{4.2}$$

where m_i and d_i denote the inertia coefficients and damping ratio for node i respectively, while $P_{in,i}^{(0)}$ and $P_{e,i}$ refer to the power input and power output respectively. The electrical output power $P_{e,i}$ at bus i is only related to the voltage phase angles as illustrated by (4.3),

$$P_{e,i} = \sum_{j=1}^{n} b_{ij} \sin(\theta_i - \theta_j), \ i \in \{1, \dots, n\}. \tag{4.3}$$

By substituting (4.3) into (4.2), the network-coupled dynamical systems defined by sets of differential equations of the form,

$$m_i\ddot{\theta}_i + d_i\dot{\theta}_i = P_{in,i}^{(0)} - \sum_{j=1}^{n} b_{ij} \sin(\theta_i - \theta_j), \quad i \in \{1, \dots, n\}. \tag{4.4}$$

With inertia on certain nodes $m_i > 0$, it is an approximation model for the swing dynamics of high-voltage transmission network within a few seconds after the

event [80]. In this transient time interval, the network is justified as power system where the ratio factor of branch reactance to its resistance is high. Reactive power and voltage magnitude are not of concern, voltage amplitudes of the system are considered constant. The linear approximation $\sin(\theta_i - \theta_j) \approx \theta_i - \theta_j$ can be justified since the angle difference across the branch are small. Then eliminating passive load buses via Kron reduction [81], we can obtain a network-reduced model with N generator buses. The phase angle θ of generator buses can be expressed by the following dynamic equation,

$$M\ddot{\theta} + D\dot{\theta} = P - L\theta, \tag{4.5}$$

where M = diag($\{m_i\}$), D = diag($\{d_i\}$). Thus, for the Laplacian matrix L of the grid, its off-diagonal elements are $l_{ij} = -b_{ij}V_i^{(0)}V_j^{(0)}$ and diagonals are $l_{ij} = \sum_{j=1,j\neq i}^{n} b_{ij}V_i^{(0)}V_j^{(0)}$. In this paper, the Laplacian matrix of the network-reduced model is real and symmetric [82], which has a complete orthogonal set of eigenvectors $\{\beta_\alpha\}$ with eigenvalue $\{\lambda_\alpha\}$. Due to the construction of the Laplacian matrix L we have zero row and column sums, which implies that there is $\lambda_1 = 0$, corresponding to an eigenvector with constant elements, $(\beta_1)^T = \{\frac{1,\dots,1}{\sqrt{N_g}}\}$. Higher eigenvalues λ_α of L are strictly positive for $\alpha = \{2, \dots, N_g\}$, where N is the number of generator buses, and the second-smallest eigenvalue is the algebraic connectivity [81]. Approach proposed in [83] have shown high accuracy and robustness in identifying the critical nodes accounting for spatial inertia distribution. Under the assumption of homogeneous inertia, the frequency deviations at bus i can then be derived,

$$\delta\dot{\theta}_\iota(t) = \frac{\Delta P e^{-\frac{\gamma t}{2}}}{m}\sum_{\alpha=1}^{N_g}\left(\beta_{\alpha i}\beta_{\alpha b}\frac{sin\left(\sqrt{\frac{\lambda_\alpha}{m}-\frac{\gamma^2}{4}}\,t\right)}{\sqrt{\frac{\lambda_\alpha}{m}-\frac{\gamma^2}{4}}}\right), \tag{4.6}$$

where m denotes average inertia distribution on generator buses; and bus b is the location where disturbance occurs. With both damping and inertia being bounded at each bus, the ratio of damping coefficient to inertia coefficient is strictly prescribed within narrow ranges, $\gamma = d_i/m_i$, which is assumed as a constant [84]. As the frequency is monitored at discrete time interval, we consider a measurement window length of Δt, then RoCoF on bus i, $R_i(t)$, can be calculated as,

$$R_i(t) = -\frac{\delta\dot{\theta}_\iota(t+\Delta t) - \delta\dot{\theta}_\iota(t)}{2\pi\Delta t}. \tag{4.7}$$

After substituting (4.6) into (4.7), we can obtain,

$$R_i(t) = \frac{\Delta P e^{-\frac{\gamma t}{2}}}{2\pi m}\sum_{\alpha=1}^{N_g}\frac{\beta_{\alpha i}\beta_{\alpha b}}{\sqrt{\frac{\lambda_\alpha}{m}-\frac{\gamma^2}{4}}\Delta t}\left[\begin{matrix} e^{-\frac{\gamma\Delta t}{2}}sin\left(\sqrt{\frac{\lambda_\alpha}{m}-\frac{\gamma^2}{4}}(t+\Delta t)\right) \\ -sin\left(\sqrt{\frac{\lambda_\alpha}{m}-\frac{\gamma^2}{4}}\,t\right)\end{matrix}\right]. \tag{4.8}$$

Expression (4.9) is the position-independent contribution ${R_i}^{(1)}$, which is the term corresponding to $\alpha = 1$ in (8). It can be observed that term $\alpha = 1$ describes the dynamics of system equivalent model,

$${R_i}^{(1)}(t) = \frac{\Delta P e^{-\gamma t}(1-e^{-\gamma\Delta t})}{2N_g\pi m\gamma\Delta t}. \tag{4.9}$$

The initial inertial response can then be calculated as,

$$\lim_{\Delta t\to 0,\ t\to 0}{R_i}^{(1)}(t) = \frac{\Delta P}{2\pi N_g m}, \tag{4.10}$$

which corresponds to the inertia response of system equivalent model. In (4.8)-(4.10), the inertia coefficient m plays a vital role in the first few seconds which is inversely proportional to the RoCoF; higher inertia coefficient results in mitigated RoCoF value and oscillation amplitude, which allows longer responding time for the primary control to act.

Further study also suggests that amplitude and period of higher oscillations corresponding to terms α>1 all depend on $\sqrt{\lambda_\alpha/m-\gamma^2/4}$. For IEEE 24-bus system, $\sqrt{\lambda_\alpha/m-\gamma^2/4} \in [0.88,\ 14.18]$, $\sqrt{\lambda_{\alpha\geq 3}/m-\gamma^2/4}$ is almost twice large than the second one. High-lying terms with large eigenvalues λ_α contribute much less than the second slowest eigenmode, i.e. the Fielder mode, of the system Laplacian matrix L [83]. Meanwhile higher-lying modes have short-period contributions, therefore only the first few eigenmodes of the network Laplacian, corresponding to its lower eigenvalues, mainly impact the value. In this study, we neglect them in the qualitative discussions and consider the RoCoF contributions from terms $\alpha \leq 2$. The simplified but effective function to calculate nodal RoCoF is given as follows,

$$R_i^{(1,2)}(t) = \frac{\Delta P e^{-\gamma t}(1-e^{-\gamma \Delta t})}{2N_g \pi m \gamma \Delta t} + \frac{\Delta P e^{-\gamma \frac{t}{2}}}{2\pi m} \frac{\beta_{2i}\beta_{2b}}{\sqrt{\frac{\lambda_2}{m}-\frac{\gamma^2}{4}}\Delta t} \left[e^{-\gamma \frac{\Delta t}{2}} sin\left(\sqrt{\frac{\lambda_2}{m}-\frac{\gamma^2}{4}}(t+\Delta t)\right) - sin\left(\sqrt{\frac{\lambda_2}{m}-\frac{\gamma^2}{4}}t\right)\right]. \quad (4.11)$$

By monitoring the average frequency change, the RoCoF relays can make more secure decisions during contingency. Practically, the time interval or measuring

window for calculating RoCoF ranges from 5 cycles to 10 cycles [85]. In this paper, the average frequency change over a period of 100 ms (6 cycles) is defined as the RoCoF value.

4.3 Derivation of Locational RoCoF Constraints

We first focus on the impact of oscillations, RoCoF experienced by different generator buses can be different depending on their inertia coefficient and electrical distance from the disturbance, and the concept of local RoCoF is proposed in [86], generators in some areas would then suffer much higher RoCoF and have a higher chance to get tripped [87].

Fig. 4.1 shows the distribution of Fiedler mode of the Laplacian matrix corresponding to the reduced network of the IEEE 24-bus system via Kron reduction. The original model has nodes where inertia $m_i = 0$ which gives the celebrated Kuramoto model on a network where angle differences become very large, which is not suitable for high-voltage electric power system analysis in this work. Fiedler mode of the reduced model manifests the frequency dynamics of individual generator bus regarding system connectivity and disturbance propagation [88].

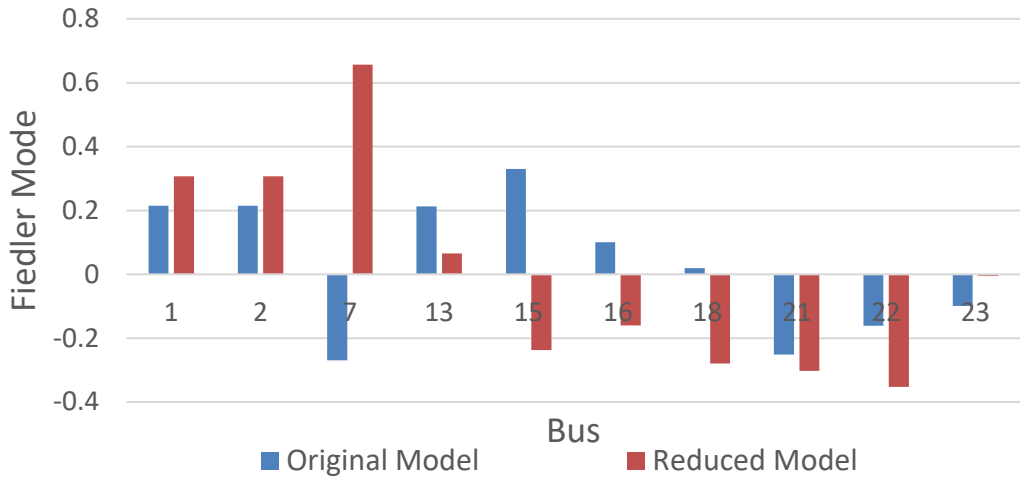
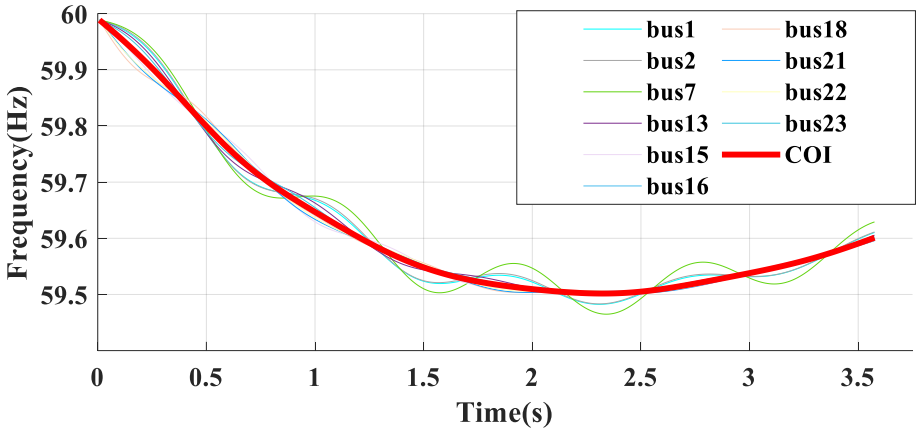


Fig. 4.1 Fiedler mode distribution.

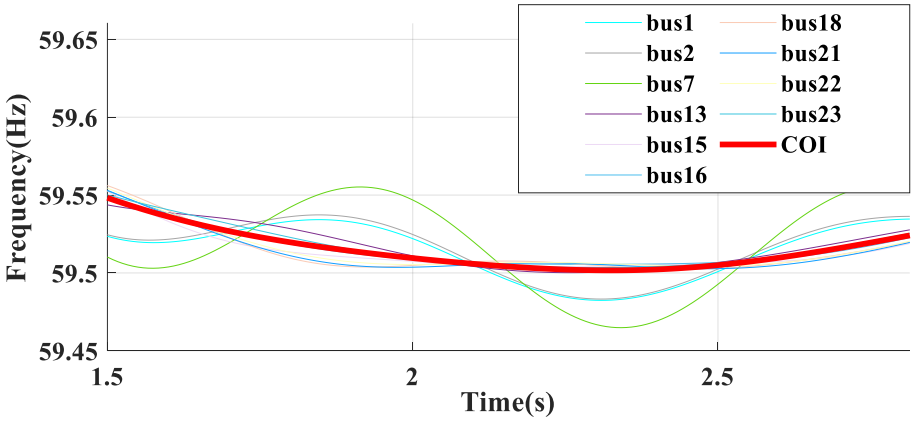
To investigate how Fiedler mode affects the locational inertial response and the operation of relating ancillary services, simulations are conducted on the IEEE 24-bus system with a step increase of 180 MW load on bus 18. Dynamic simulation results are depicted in Fig. 2(a), as we can see that the frequency deviations at bus 13 and bus 23 are close to the deviation at center of inertia (COI) within the whole monitoring course. The largest deviation is observed at bus 7 corresponding to the highest Fiedler mode absolute value.

These deviations in frequency can influence the locational frequency at different buses depending on the size and configuration of the system. Traditionally, generators adjacent to the event is assumed perceive highest RoCoF comparing to the generators on distant buses [55]. While results of [4] and [83] have emphasized the impact of oscillation and disturbance propagation, buses in distance may experience highest RoCoF after a time delay instead of at the initial time point. From the expression (10), it can be inferred that local buses sharing the same positive or

negative Fiedler mode with the event bus would perceive larger deviation before COI frequency, while the non-local generator buses with opposite positive and negative Fiedler mode may experience highest RoCoF later than the COI frequency. Simulation results shown in Fig. 2(b) prove our inference.



(a) Frequency for period between t=0 and t=3.5s.



(b) Frequency for the period between t=1.5s and t=2.5s.

Fig. 4.2 Frequency response following a disturbance on bus 18.

Following a contingency, there are larger frequency excursions and higher RoCoF on non-local generator buses to the disturbance comparing to COI frequency. For these non-local generator buses of opposite positive and negative Fiedler mode

value with disturbance bus, highest RoCoF occurs with a time period delay due to the disturbance propagation.

To guarantee the locational frequency security for all buses and mitigate the highest RoCoF during the oscillation, a multiple-measurement-window method is introduced to capture the highest RoCoF during the oscillation. Compared to the case of a sudden load increase, the loss of generation not only causes mismatch in system power balance, but also the degradation of the system synchronous inertia, resulting in even higher frequency deviation and larger initial RoCoF. Thus, the *G*-1 contingency of largest generation is considered as the worst contingency in this paper. And the aforementioned RoCoF expression in (11) are then incorporated as constraints in SCUC, the resulting locational RoCoF constraint (12) that respects the prescribed threshold $\mathrm{RoCoF}_{\mathrm{lim}}$ is then reformulated as follows with t being set as T_1 andT_2. These two typical setting points are same for all generator buses in this case, which is mainly based on oscillation period of second lowest term $\alpha = 2$, considering trade-off between computational efficiency and measurement accuracy. Note that the responding time of primary control is considered starting from 3 seconds; even with fast frequency response, the general responding time is considered from 0.5s to 1.5s, which is still later than T_2. Hence, primary control is not considered when calculating nodal RoCoF. The derived constraint is then defined as,

$$\frac{\Delta \mathrm{P} e^{-\gamma \frac{t}{2}}(1-e^{-\gamma \Delta t})}{2\mathrm{N}_g \pi m_\Delta \gamma \Delta t} + \frac{\Delta \mathrm{P} e^{-\gamma \frac{t}{2}}}{2\pi m_\Delta} \frac{\beta_{2i}\beta_{2b}}{\sqrt{\frac{\lambda_2}{m_\Delta} - \frac{\gamma^2}{4}}\Delta t} \tag{4.12}$$

$$\begin{bmatrix} e^{-\gamma\frac{\Delta t}{2}}\sin\left(\sqrt{\frac{\lambda_2}{m_\Delta}-\frac{\gamma^2}{4}}(t+\Delta t)\right) \\ -\sin\left(\sqrt{\frac{\lambda_2}{m_\Delta}-\frac{\gamma^2}{4}}t\right) \end{bmatrix} \geq \mathrm{RoCoF}_{\mathrm{lim}},$$

where $m_\Delta = m - \Delta m$, Δm denotes an average loss of inertia coefficient distributed on each bus, N_g is the number of generator buses. Within a plausible range of variables $(\Delta P, m, \Delta m, \gamma)$, the RoCoF on all generator buses following a $\mathsf{G}-1$ contingency can be calculated with the left-hand side terms of (4.12).

4.4 Formulation of LRC-SCUC

The objective of SCUC is to minimize operational cost of generators subject to various constraints. Traditional SCUC (T-SCUC) model does not consider frequency constraints, which may lead to severe system instability issue for lower-inertia power grid with substantial amount of renewable generation. In this section, we first introduce the T-SCUC and then describe how to incorporate RoCoF related constraints into the ERC-SCUC and the proposed LRC-SCUC. Subsequently, we will describe the proposed VI-LRC-SCUC model that utilize virtual inertia.

Objective function (4.13) is shared by three SCUC models: T-SCUC, ERC-SCUC and LRC-SCUC. It is to minimize the total system cost consisting of variable fuel costs, no-load costs, start-up costs, and reserve costs.

The T-SCUC model includes various constraints (4.14)-(4.27). Equation (4.14) enforces the nodal power balance. Network power flows are calculated in (4.15) and are restricted by the transmission capacity as shown in (4.16). The scheduled

energy production and generation reserves are bounded by unit generation capacity and ramping rate (4.17)-(4.22). As defined in (4.20), the reserve requirements ensure the reserve is sufficient to cover any loss of a single generator. The start-up status and on/off status of conventional units are defined as binary variables (4.23)-(4.27),

$$\min_{\Phi} \sum_{g \in G} \sum_{t \in T} (c_g P_{g,t} + c_g^{NL} u_{g,t} + c_g^{SU} v_{g,t} + c_g^{RE} r_{g,t}), \tag{4.13}$$

$$\sum_{g \in G} P_{g,t} + \sum_{k \in K(n-)} P_{g,t} - \sum_{k \in K(n+)} P_{g,t} - D_{n,t} + E_{n,t} = 0, \quad \forall n, t, \tag{4.14}$$

$$P_{k,t} - b_k(\theta_{n,t} - \theta_{m,t}) = 0, \ \forall k, t, \tag{4.15}$$

$$-P_k^{max} \leq P_{k,t} \leq P_k^{max}, \ \forall k, t, \tag{4.16}$$

$$P_g^{min} u_{g,t} \leq P_{g,t}, \quad \forall g, t, \tag{4.17}$$

$$P_{g,t} + r_{g,t} \leq u_{g,t} P_g^{max}, \quad \forall g, t, \tag{4.18}$$

$$0 \leq r_{g,t} \leq R_g^{re} u_{g,t}, \quad \forall g, t, \tag{4.19}$$

$$\sum_{j \in G} r_{j,t} \geq P_{g,t} + r_{g,t}, \quad \forall g, t, \tag{4.20}$$

$$P_{g,t} - P_{g,t-1} \leq R_g^{hr}, \quad \forall g, t, \tag{4.21}$$

$$P_{g,t-1} - P_{g,t} \leq R_g^{hr}, \quad \forall g, t, \tag{4.22}$$

$$v_{g,t} \geq u_{g,t} - u_{g,t-1}, \quad \forall g, t, \tag{4.23}$$

$$v_{g,t+1} \leq 1 - u_{g,t} \quad \forall g, t \leq nT - 1, \tag{4.24}$$

$$v_{g,t} \leq u_{g,t} \quad \forall g, t, \tag{4.25}$$

$$v_{g,t} \in \{0,1\}, \quad \forall g, t, \tag{4.26}$$

$$u_{g,t} \in \{0,1\}, \qquad \forall g,t. \tag{4.27}$$

As mentioned before, the ERC-SCUC model is constrained to guarantee generator frequency stability considering the relative location to the potential $G-1$ contingency. Loss of synchronous generation would lead to a reduction in total system inertia which would cause a higher RoCoF comparing to the event with no synchronous inertia loss. The system RoCoF limit following a $G-1$ contingency is guaranteed by applying the following set of frequency related constraints (4.28)-(4.30). (4.28) defines the rated power of dispatched generators, while (4.30) calculates the system synchronous inertia. (4.30) is derived from the swing equation of system equivalent model, the constraint ensures the system frequency security against loss of generation,

$$k_{g,t} = P_g^{max} u_{g,t}, \qquad \forall\, g,t, \tag{4.28}$$

$$M_t = \frac{\sum_{g \in G} 2H_g k_{g,t}}{\omega_0}, \qquad \forall t, \tag{4.29}$$

$$P_{g,t}/(M_t\omega_0 - 2H_g k_{g,t}) \leq -RoCoF_{lim} \quad \forall g,t. \tag{4.30}$$

If the generation scheduling is not appropriately constrained to reflect the $G-1$ contingency and distinct locational frequency, unexpected tripping of RoCoF relays and cascading contingency may take place. To address this issue, squared Fiedler mode amplitude β_{2b}^2 is introduced in this paper to assess the impact of disturbance, a power loss at bus b, on frequencies in the whole grid. Considering the inter oscillation between area which may cause unexpected tripping of RoCoF relays, we create two locational RoCoF constraints in this paper based on the definition of local buses and

non-local buses. Constraint (4.31)-(4.33) defines the average nodal inertia change over generator buses due to loss of a generator. Constraints (4.34) and (4.35) ensure system stability by imposing limit on locational highest captured RoCoF of buses under possible worst $G-1$ contingency,

$$\Delta m_{g,t} = \frac{2H_g k_{g,t}}{N_g \omega_0}, \qquad \forall g,t, \tag{4.31}$$

$$m_t = \frac{M_t}{N_g}, \qquad \forall t, \tag{4.32}$$

$$m_{g,t} = m_t - \Delta m_{g,t}, \qquad \forall g,t, \tag{4.33}$$

$$\begin{bmatrix} \frac{P_{g,t}e^{-\gamma T_1}(1-e^{-\gamma\Delta t})}{2N_g\pi m_{g,t}\gamma\Delta t} + \frac{P_{g,t}e^{-\gamma\frac{T_1}{2}}}{2\pi m_{g,t}} \frac{\beta_{2n}\beta_{2b}}{\sqrt{\frac{\lambda_2}{m_{g,t}} - \frac{\gamma^2}{4}\Delta t}} \\ e^{-\gamma\frac{\Delta t}{2}} sin\left(\sqrt{\frac{\lambda_2}{m_{g,t}} - \frac{\gamma^2}{4}}(T_1+\Delta t)\right) \\ - sin\left(\sqrt{\frac{\lambda_2}{m_{g,t}} - \frac{\gamma^2}{4}}T_1\right) \end{bmatrix} \leq -\text{RoCoF}_{lim}, \tag{4.34}$$

$\forall n \in \text{N}_{\text{loc,g}}$, g,t,

$$\begin{bmatrix} \frac{P_{g,t}e^{-\gamma T_2}(1-e^{-\gamma\Delta t})}{2N_g\pi m_{g,t}\gamma\Delta t} + \frac{P_{g,t}e^{-\gamma\frac{T_2}{2}}}{2\pi m_{g,t}} \frac{\beta_{2n}\beta_{2b}}{\sqrt{\frac{\lambda_2}{m_{g,t}} - \frac{\gamma^2}{4}\Delta t}} \\ e^{-\gamma\frac{\Delta t}{2}} sin\left(\sqrt{\frac{\lambda_2}{m_{g,t}} - \frac{\gamma^2}{4}}(T_2+\Delta t)\right) \\ - sin\left(\sqrt{\frac{\lambda_2}{m_{g,t}} - \frac{\gamma^2}{4}}T_2\right) \end{bmatrix} \leq -\text{RoCoF}_{lim}, \tag{4.35}$$

$\forall n \in N_{n-loc,g}, g, t.$

There are different implementations for synchronous machine response emulation. The study in [89] shows that control schemes can be utilized to provide equivalent inertia through non-synchronous devices. Such concept introduces techniques like virtual inertia emulation to mimic the behavior of synchronous machines, displacing the inertia provided by synchronous generators with cheap inertia ancillary service provided by various resources such as virtual synchronous machine. To study the effect of virtual inertia on the power grid, an aggregated virtual inertia involved location based RoCoF constrained SCUC model or a VI-LRC-SCUC model is proposed and examined in this paper. Compared to (4.13), this model also considers the cost for virtual inertia provision; the updated objective function for the proposed VI-LRC-SCUC model is shown in (4.35). Moreover, when virtual inertia is considered, constraints (4.29)-(4.30) should be replaced by (4.37)-(4.38). Expression (4.37) describes the system inertial response respect to aggregate inertia contributions from condensers and inverter-based resources. (4.38) defines the change in average nodal inertia while virtual inertia is applied,

$$\min_{\Phi} \sum_{g \in G} \sum_{t \in T} (c_g P_{g,t} + c_g^{NL} u_{g,t} + c_g^{SU} v_{g,t} + c_g^{RE} r_{g,t}) + \sum_{t \in T} c^{VI} M_t^{VI}, \tag{4.36}$$

$$-P_{g,t}/(M_t \omega_0 + M_t^{VI} \omega_0 - 2H_g k_{g,t}) \geq RoCoF_{lim}, \quad \forall g, t, \tag{4.37}$$

$$m_t = \frac{M_t}{N_g} + \frac{M_t^{VI}}{N_g}, \quad \forall t, \tag{4.38}$$

The virtual inertia requires fast responsive energy buffer; the kinetic energy in a wind turbine and the energy in a battery are limited energy resources for virtual inertia responses. Hence, it is practical to set a limit on the total virtual inertia due to budget limit and resource limit. This limit should also be considered as follows,

$$0 \leq M_t^{VI} \leq M_t^{Total}, \quad \forall t, \tag{4.39}$$

We examine five different SCUC models that are summarized in Table 4.1. In this table, the objective functions and constraints enforced are listed for each SCUC model.

Table 4.1 SCUC Formulation of Different Models

Model	Objective Function	Shared Constraints	Unique Constraints
T-SCUC	(4.13)	(4.14)-(4.27)	None
ERC-SCUC			(4.28)-(4.30)
LRC-SCUC			(4.28)-(4.29), (4.31)-(4.35)
VI-ERC-SCUC	(4.36)		(4.28)-(4.29), (4.37), (4.38)
VI-LRC-SCUC			(4.28)-(4.29), (4.31), (4.33)-(4.35), (4.37)-(4.39)

Note that non-linear constraints (4.34) and (3.35) are linearized with the PWL method described in Section 4.5 before solving the associated SCUC models.

4.5 Piecewise Linearization

The constraints on RoCoF for locational frequency dynamics are nonlinear. In order to incorporate these frequency-related constraints into the proposed LRC-SCUC model, a linear approximation method is introduced. [90] proposes a piece-wise linearization technique for obtaining a linearized expression. Since respective damping and droop gains are usually strictly prescribed proportional to the synchronous inertia, the RoCoF expression becomes a function of three variables $R_i^{(1,2)}(\Delta P, m, \Delta m)$. The expression has local convexity within the interval of interest, that can be expressed as $a_v^i \Delta P^\eta + b_v^i m^\eta + c_v^i \Delta m^\eta + d_v^i$. In order to determine these four coefficients, A heuristic least-squares method is proposed in [91] to solve this problem.

The least squares based PWL method aims to minimize the following objective function,

$$\min_{\Psi} \sum_{\eta} \Big(\max_{1 \le v \le \bar{v}} \{ a_v^i \Delta P^\eta + b_v^i m^\eta + c_v^i \Delta m^\eta + d_v^i \} - R_i^{(1,2)}(\Delta P^\eta, m^\eta, \Delta m^\eta) \Big)^2 , \tag{4.40}$$

where $\Psi = \{a_v, b_v, c_v, d_v, \forall v\}$ is the set of parameters to be determined; i denotes the measurement bus; and η denotes the evaluation point; v refers to the index of PWL segments and $\bar{v}$ denotes the number of PWL segments. The problem of fitting $\max_{1 \le v \le \bar{v}} \{ a_v^i \Delta P^\eta + b_v^i m^\eta + c_v^i \Delta m^\eta + d_v^i \}$ to $-R_i^{(1,2)}(\Delta P^\eta, m^\eta, \Delta m^\eta)$ over the plausible range can be considered as minimizing difference between the appropriate PWL segment and the RoCoF function. To solve this min-max problem and eliminate

the inner max operator from the objective function, new variables t_v are introduced and defined as follows,

$$t_1^i = max\{a_1^i \Delta P^\eta + b_1^i m^\eta + c_1^i \Delta m^\eta + d_1^i, \\ a_2^i \Delta P^\eta + b_2^i m^\eta + c_2^i \Delta m^\eta + d_2^i \}, \tag{4.41}$$

$$t_{v-1}^i = max\{t_{v-2}^i, a_v^i \Delta P^\eta + b_v^i m^\eta + c_v^i \Delta m^\eta + d_v^i\}, 3 \leq v \leq \bar{v}. \tag{4.42}$$

Adding new linear inequalities would relieve the objective function from the "max" operator [90]. We basically introduce $\bar{v} - 1$ new binary α_v as well as $\bar{v} - 1$ continuous variables t_v^i, the unconstrained min-max optimization problem (4.40) can be reformulated as a constrained optimization problem (4.43)-(4.47), which is the RoCoF linearization problem,

$$\min_{\Psi} \sum_{\eta} (t_{\bar{v}-1}^i (\Delta P^\eta, m^\eta, \Delta m^\eta) - R_i^{(1,2)} (\Delta P^\eta, m^\eta, \Delta m^\eta))^2, \tag{4.43}$$

subject to the following constraints,

$$a_1^i \Delta P^\eta + b_1^i m^\eta + c_1^i \Delta m^\eta + d_1^i \leq t_1^i \leq a_1^i \Delta P^\eta + b_1^i m^\eta + c_1^i \Delta m^\eta + d_1^i + \alpha_1^i \Omega, \ \forall \eta, \tag{4.44}$$

$$a_2^i \Delta P^\eta + b_2^i m^\eta + c_2^i \Delta m^\eta + d_2^i \leq t_1^i \leq a_2^i \Delta P^\eta + b_2^i m^\eta + c_2^i \Delta m^\eta + d_2^i + (1 - \alpha_1^i)\Omega, \quad \forall \eta, \tag{4.45}$$

$$t_{v-2}^i \leq t_{v-1}^i \leq t_{v-2}^i + \alpha_{v-1}^i \Omega, \quad \forall \eta, 3 \leq v \leq \bar{v}, \tag{4.46}$$

$$a_v^i \Delta P^\eta + b_v^i m^\eta + c_v^i \Delta m^\eta + d_v^i \leq t_{v-1}^i \leq a_v^i \Delta P^\eta + b_v^i m^\eta + c_v^i \Delta m^\eta + d_v^i + \left(1 - \alpha_{v-1}^i\right)\Omega, \quad \forall \eta, 3 \leq \mathrm{v} \leq \bar{\mathrm{v}}, \tag{4.47}$$

where Ω is a sufficiently large positive number.

Upon obtaining the optimal solution $(a_{\bar{v}}^*, b_{\bar{v}}^*, c_{\bar{v}}^*, d_{\bar{v}}^*)$ of the nodal RoCoF linearization model, the nonlinear location based RoCoF constraints (4.34) and (4.35) can be converted into linear constraints in SCUC: $4 * (\bar{v} - 1)$ inequalities that ensures $t_{\bar{v}-1}^i$ is maximum, along with the RoCoF threshold constraints $t_{\bar{v}-1}^i \leq -RoCoF_{lim}$.

A new metric, RoCoF violation gap, is proposed to quantify the performance of different SCUC models in terms of frequency response particularly the inertial response. The RoCoF violation gap is defined as follows,

$$RoCoF_{\ violation\ gap} = RoCoF / RoCoF_{lim} - 1\ ,\ \forall t. \tag{4.48}$$

4.6 Results Analysis

4.6.1 Test System Description

A case study on IEEE 24-bus system is provided to demonstrate the effectiveness of the proposed methods [92]. This test system contains 24 buses, 38 generators and 38 lines, which also considers decarbonized generation characterized by wind power and solar power on generator buses. Fig. 3 shows the renewable generation and load profile for a system scenario with 60% of maximum renewable energy penetration level during peak hour. Electricity demand ranges from 1,432 MW to a peak of 3,222 MW. The mathematical models are implemented in Python using Pyomo [93]-[94] and solved with the Gurobi solver [95], and the optimality gap is set to 0.1%. The computer with Intel® Xeon(R) W-2195 CPU @ 2.30GHz and 128 GB of RAM was utilized to conduct the numerical simulations.

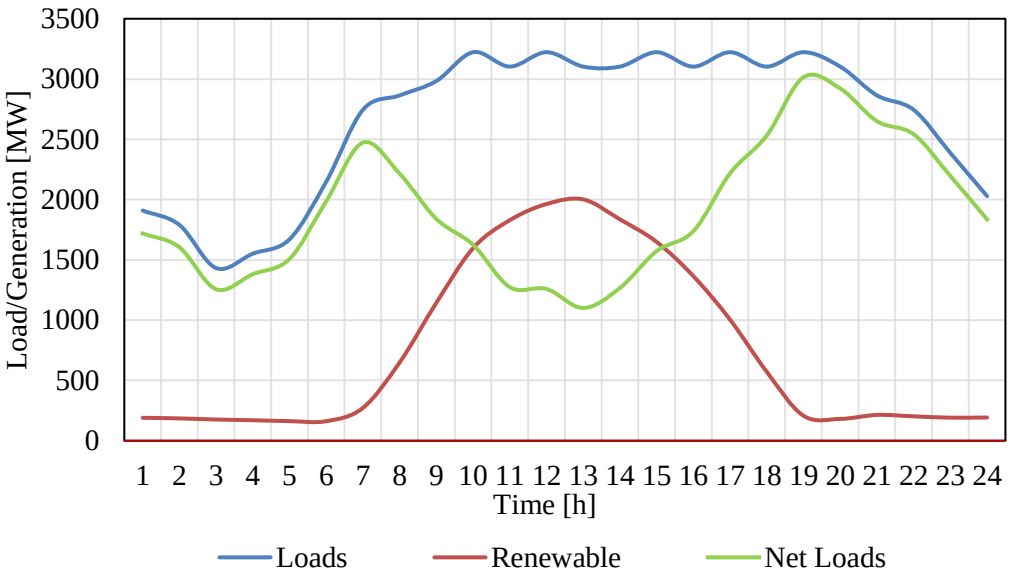


Fig. 4.3 Renewable generation and load profile of the IEEE 24-bus system.

4.6.2 Investigation of Frequency Propagation

In this subsection, the effect of Fiedler mode on disturbance propagation and locational inertial response is simulated and studied. We first investigate the implications of a generation loss in scenarios with different system inertia. The locational RoCoF is numerically calculated based on expression (4.12), RoCoF at t = 0s and Δt = 0.1s on local bus 21 adjacent to event bus 18 that is illustrated in Fig. 4.4 where the loss of largest unit is considered. It can be inferred that the initial frequency oscillations are affected by the location of disturbance.

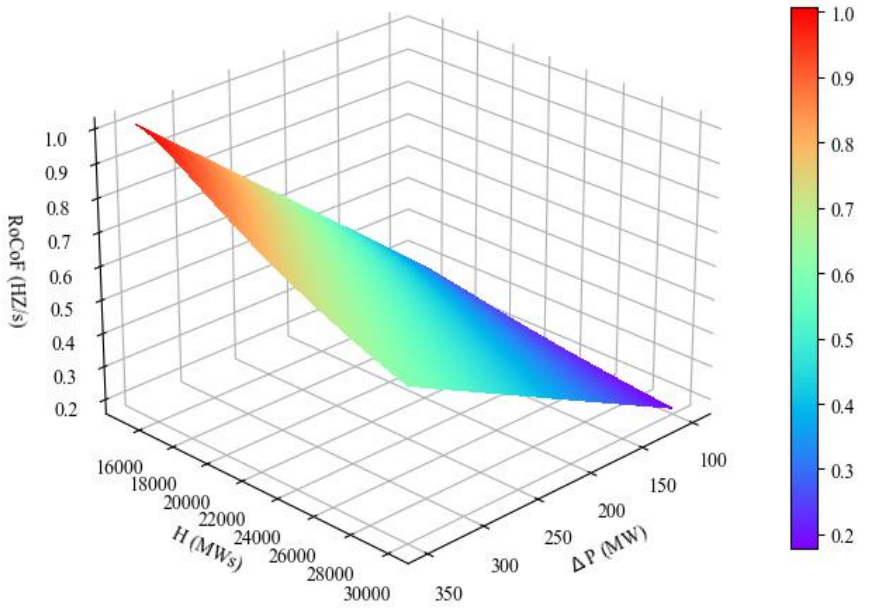


Fig. 4.4 RoCoF of bus 21 at t = 0s and Δt = 0.1s following a *G*-1 contingency.

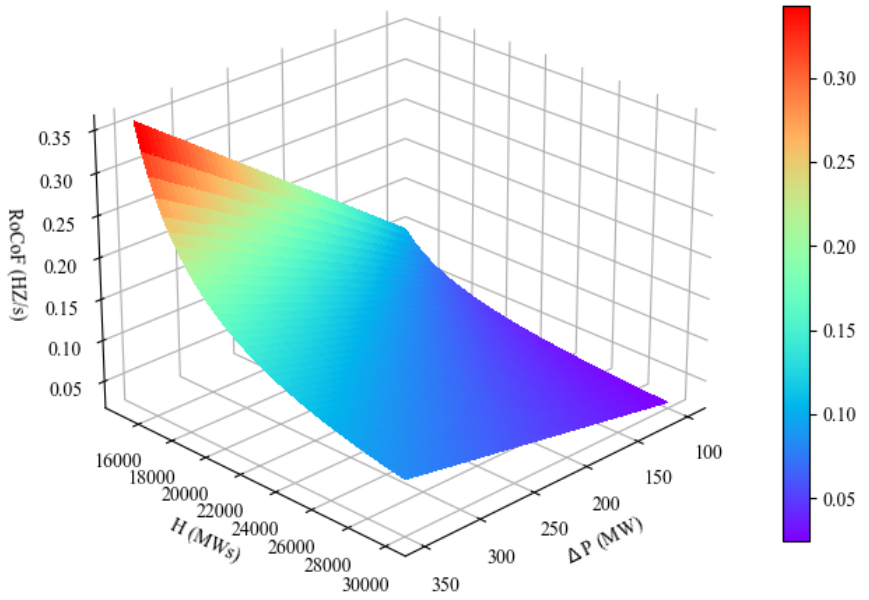


Fig. 4.5 RoCoF of bus 1 at t = 0s and Δt = 0.1s following *G*-1 contingency.

The inertial response on non-local bus 1 is depicted in Fig. 4.5, generator buses with opposite positive and negative Fiedler mode experience a much smaller initial RoCoF at $t = 0s$. Combining results Fig. 4.5 and Fig. 4.6, it can be observed

that local buses adjacent to disturbance substantially experience much higher initial RoCoF than other buses under at $t = 0s$. Hence, the generators buses adjacent to disturbance with large Fiedler mode are more likely to violate the RoCoF limits at the initial time point following a $G - 1$ contingency.

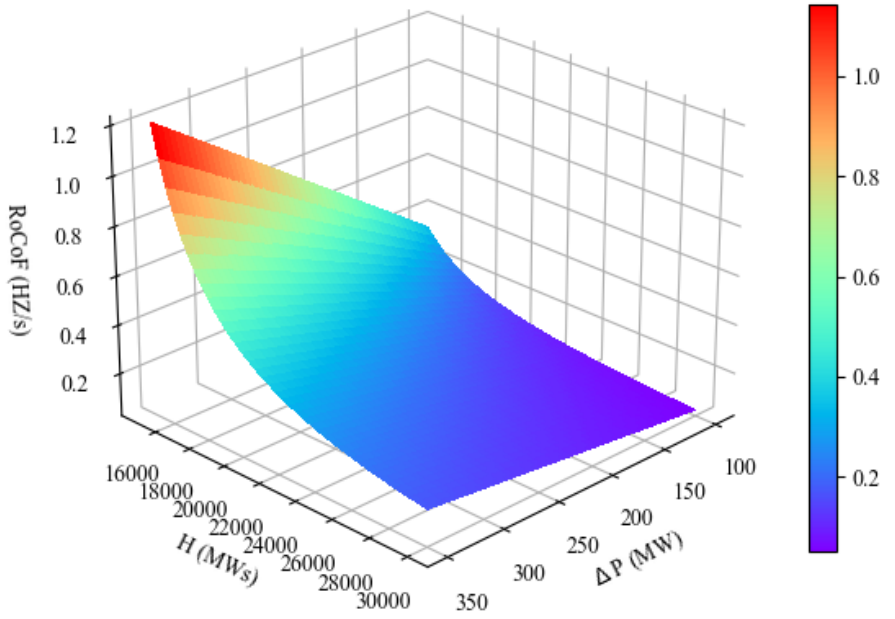


Fig. 4.6 RoCoF of bus 1 at t = 0.4s and Δt = 0.1s following $G - 1$ contingency.

The numerically calculated RoCoF for bus 1 at t = 0.4s and Δt = 0.1s under the same scenario is plotted in Fig. 4.6. It is interesting to find that local bus 1 experiences much large RoCoF at t = 0.4s than same RoCoF at t = 0.0s. The results show that only considering the initial RoCoF may fail to capture highest locational RoCoF value during the oscillation, simulation results indicate the effectiveness of the proposed multiple-measurement-window method which allows us to incorporate the captured the highest RoCoF into LRC-SCUC model and subsequently secure locational frequency stability.

It is also important to note that when system inertia drops below a certain threshold the RoCoF increases drastically especially for the buses have relatively large Fiedler mode absolute value. This subsection demonstrates the effect of inertia and Fiedler mode value on locational RoCoF: with higher system inertia, the amplitude of the oscillation decreases especially for buses with large Fiedler mode absolute value.

4.6.3 LRC-SCUC Simulation

We set the system nominal frequency to 60 Hz. Regarding post-contingency frequency limits, RoCoF must be higher than -0.5Hz/s to avoid the tripping of RoCoF-sensitive protection relays. In [96], generator settings have a typical damping to inertia ratio ranges from [0.086, 0.133]. Similar dynamic parameters are utilized in this work, the ratio is subsequently considered constant as 0.1 in this work. We first conduct the simulation of T-SCUC model for 24-hour period, which serves as a benchmark to show the impact of RoCoF constraints. The evaluation point and the PWL segments are set with $\eta = 3$ and $\bar{v} = 4$ for accuracy and computational efficiency [96]. The computational time for the T-SCUC model is 95.61s while it decreases to 20.91s for ERC-SCUC model. A possible reason is that frequency related constraints reduce the size of feasible solution set, infeasible solutions are discarded earlier because of violating the RoCoF constraints. Meanwhile, most RoCoF related constraints in LRC-SCUC model are not binding constraints. With PWL formulation solved offline, the proposed LRC-SCUC model has a computational time of 75.45s, indicating that the proposed LRC-SCUC model can be solved efficiently.

Fig. 4.7 presents the evolutions of aggregated system inertia over the scheduling horizon. When net load increases during the period of hours 5-8 and hours 16-20, the total system aggregated inertia increases as well for all three cases. Compared to T-SCUC model, imposing RoCoF constraints leads to more synchronous generators scheduled online to ensure minimum synchronous inertia online. The inertia of committed synchronous generators for T-SCUC model is much lower than ERC-SCUC and LRC-SCUC models where RoCoF related constraints are implemented. Note that the total system synchronous inertia of the proposed LRC-SCUC model is the highest among all three models, which reflects the impact of imposing locational RoCoF constraints.

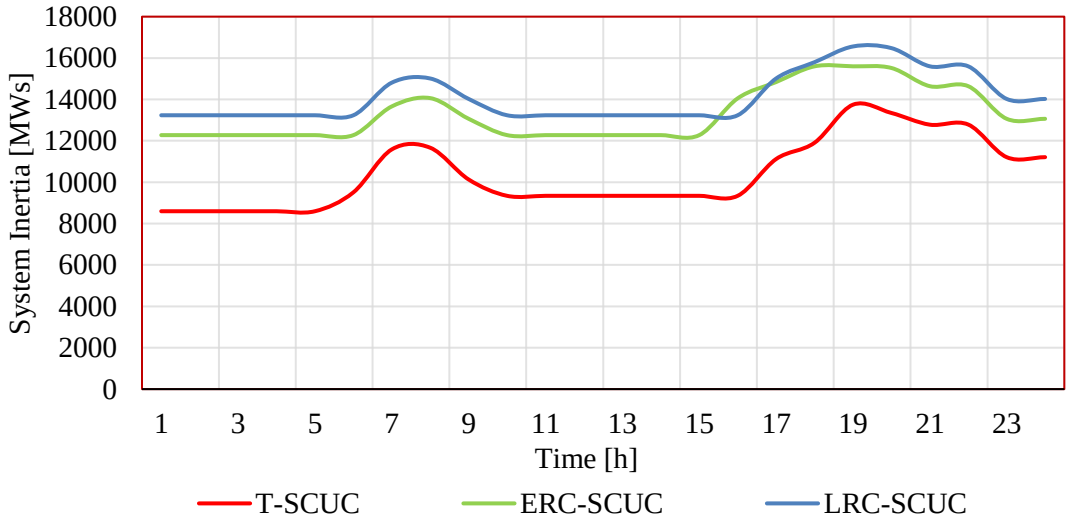


Fig. 4.7 Impact of RoCoF constraints on the total system inertia.

Some other insightful conclusions can also be drawn from Fig. 4.7. When net load decreases dramatically between hours 7-13 due to the increase of RES generation, the system aggregate inertia in T-SCUC case drops significantly subsequently, contingency happens in this time interval may lead to larger frequency deviation and

higher RoCoF. While system inertia maintains relatively high level in two other cases indicating mitigated RoCoF and frequency deviations.

Furthermore, we investigate how commitment scheduling impacts the system frequency dynamics following $G-1$ contingency [97]. The worst-case contingency is assumed to take place in peak hour 12, time-domain simulations are then conducted for all three models on Transient Security Analysis Tools (TSAT) following the loss of largest generator. The detailed model is utilized in this work, and dynamic values are selected within appropriate ranges.

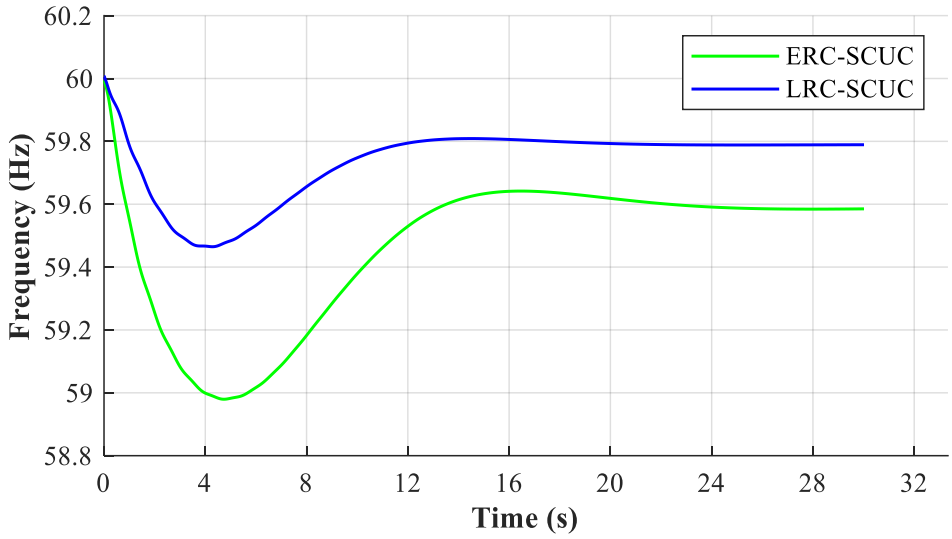
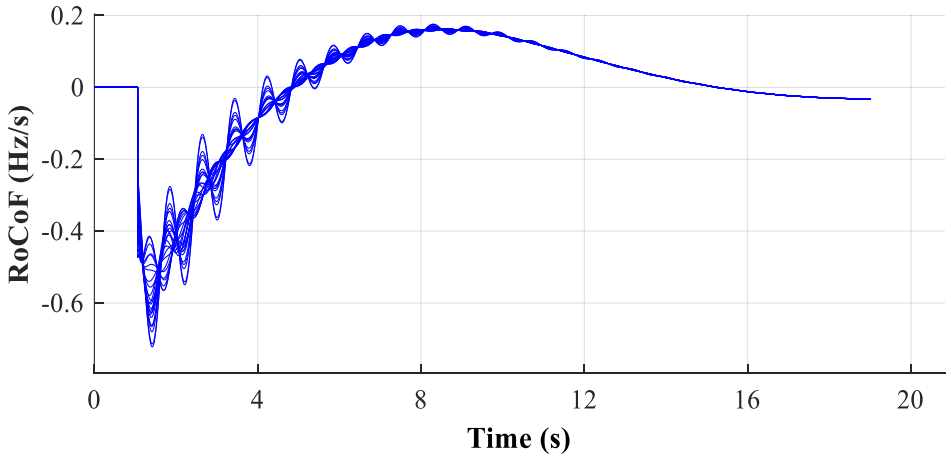


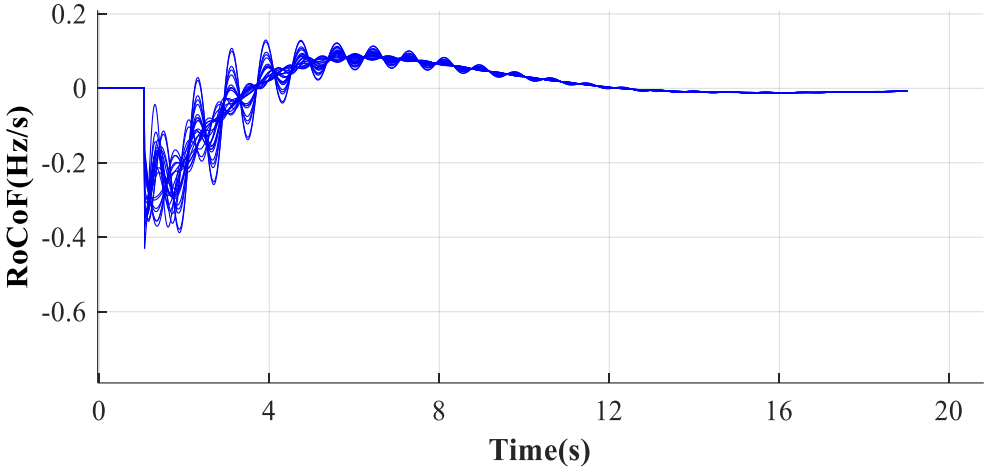
Fig. 4.8 System frequency response after loss of the generator with the largest generation at hour 12.

Results are plotted in Fig. 4.8 comparing the frequency dynamics for ERC-SCUCU and LRC-SCUC. In T-SCUC model without any RoCoF related constraints, loss of the largest committed generation of 400 MW on bus 22 leads to an extremely large frequency excursion at 54.1 Hz. For ERC-SCUC, loss of largest generation at

182.9 MW on bus 22 result in a frequency nadir at 58.9 Hz which is below the threshold of 59.3 Hz. The potential largest $G-1$ contingency level for LRC-SCUC is reduced to 155 MW, which is about 5% of largest demand. The system frequency is maintained above the threshold. Results indicate that the proposed LRC-SCUC model would help following primary frequency response, thus subsequently protect against the worst-case contingency risk.



(a) ERC-SCUC model



(b) LRC-SCUC model.

Fig. 4.9 RoCoF of all buses following the loss of largest generation in different cases.

Fig. 4.9 compares the average RoCoF over a period of 100 ms on generator buses for ERC-SCUC and LRC-SCUC models following the worst-case contingency. For the T-SCUC results, the RoCoF on all buses violate the prescribed limit easily by a large margin especially in time interval from 9h to 16h. For ERC-SCUC model, the highest RoCoF derived from COI frequency can be numerically calculated as -0.50 Hz/s which is above the required value. However, the highest locational RoCoF still violates the prescribed RoCoF limit which reflects that impact of oscillations cannot be handled well by ERC-SCUC model. For the generator commitment and dispatch solution obtained from the proposed LRC-SCUC model, the lowest locational RoCoF is maintained above -0.5 Hz/s, meeting the RoCoF security requirement. It also can be observed that initial RoCoF may not be the lowest RoCoF within the oscillation period, imposing multiple measurement windows can help us capture highest RoCoF. The results indicate the effectiveness and necessity to incorporate locational RoCoF constraints into the scheduling model.

The results of three models are summarized in Table 4.2. Imposing system equivalent model based RoCoF constraints leads to a 10.90% increase in total system cost, the total generation cost increases from $891,391 to $988,524. In the presence of location based RoCoF constraints, an increase of 17.90% in total operational cost is observed. On the other hand, the reserve cost decreases significantly as the worst-case contingency level mitigated. Meanwhile, additional synchronous machines are committed to cover the shortage of inertia to limit locational RoCoF following the loss of largest generation, which accordingly increases the operation cost as well as the

start-up cost.

Table 4.2 SCUC Costs [$] under Different Models

Model	Total	Start-up	Operation	Reserves
T-SCUC	891,391	56,704	744,973	89,714
ERC-SCUC	988,524	59,462	881,633	47,429
LRC-SCUC	1,050,989	66,810	955,983	28,196

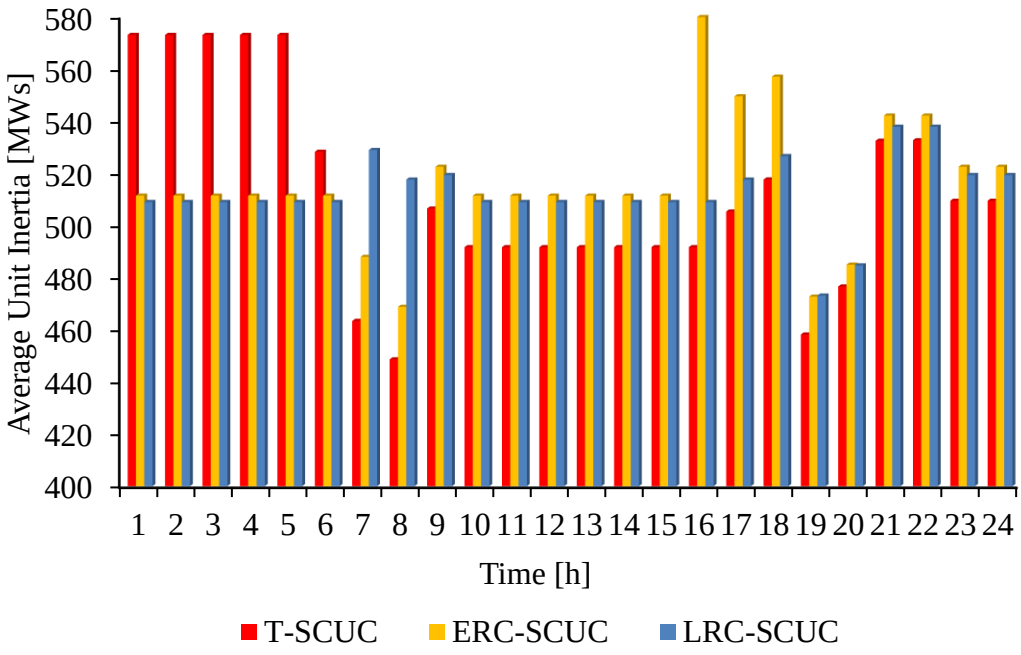


Fig. 4.10 Impact of RoCoF constraints on the average unit inertia contribution.

Furthermore, average unit inertia (AUI) of committed generators is compared in Fig. 4.10 over the whole dispatching horizon. A noticeable large value of AUI between hours 1-5 and hour 16-18 are monitored in T-SCUC and ERC-SCUC cases. Larger AUI indicates that generators of larger rated power and synchronous inertia are scheduled in the system operation, implying potential larger worst-case contingency

level. As shown, the AUI of proposed LRC-SCUC case is relatively stable during the whole horizon. To achieve the energy system with more RES penetration level, re-dispatching existing generators provides options that can help mitigate the frequency deviation and RoCoF violations.

4.6.4 Impact of Synchronous Generator Inflexibility

Conventional generator may take hours to start up again after it shuts down and shut down after it is turned on, which is known as generator minimum ON/OFF time limit constraint; such synchronous generator inflexibility may not be neglected. Thus, this section investigates how such inflexibility affects the performance of proposed LRC-SCUC model.

Combining the results shown in Fig. 4.11 and Fig. 4.12, we can conclude that the proposed LRC-SCUC model outperforms ERC-SCUC and T-SCUC. For ERC-SCUC model, the highest locational RoCoF violates the prescribed RoCoF limit. It could be inferred that initial RoCoF should not be assumed as highest value, and only imposing system equivalent RoCoF constraint cannot secure locational RoCoF stability. The proposed LRC-SCUC model maintains the lowest locational RoCoF above -0.5 Hz/s, and achieves less conservative results when generator min ON/OFF time is considered. It indicates the impact of generators inflexibility would not degrade the conclusion of this paper. Since we focused on the RoCoF stability of system with high RES penetration level, the generator min ON/OFF time limit is not enforced in other parts of this paper.

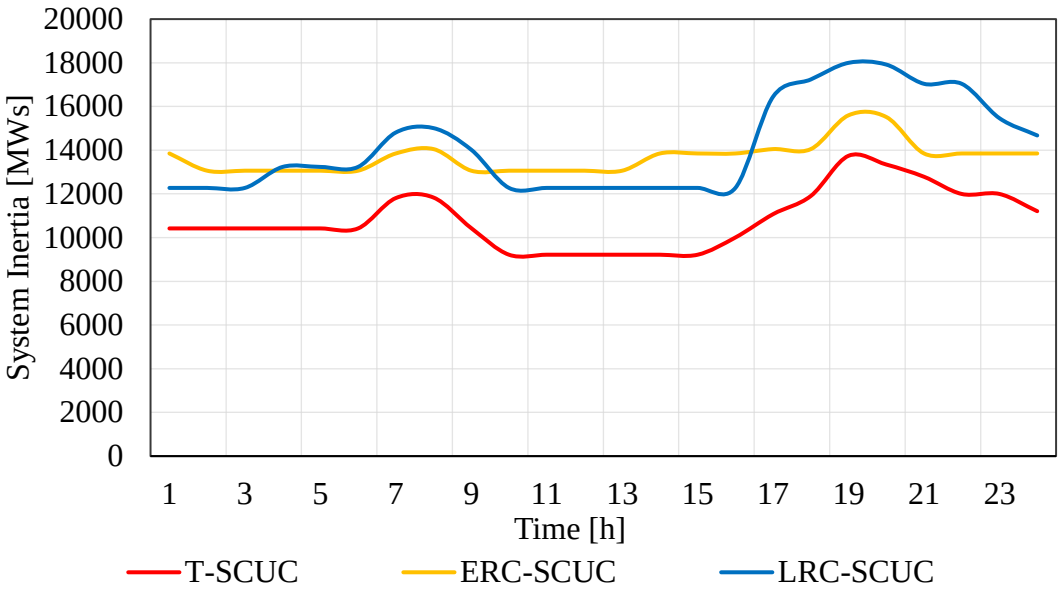
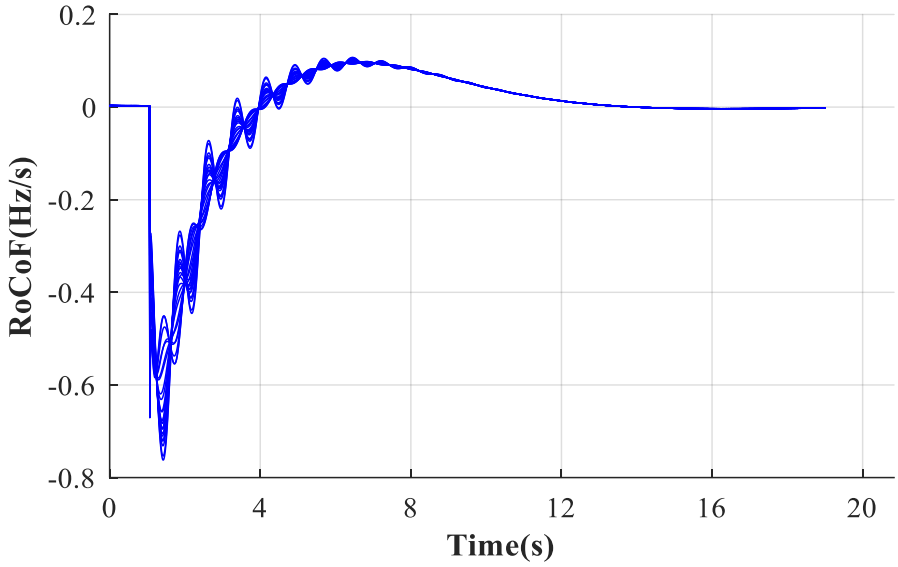
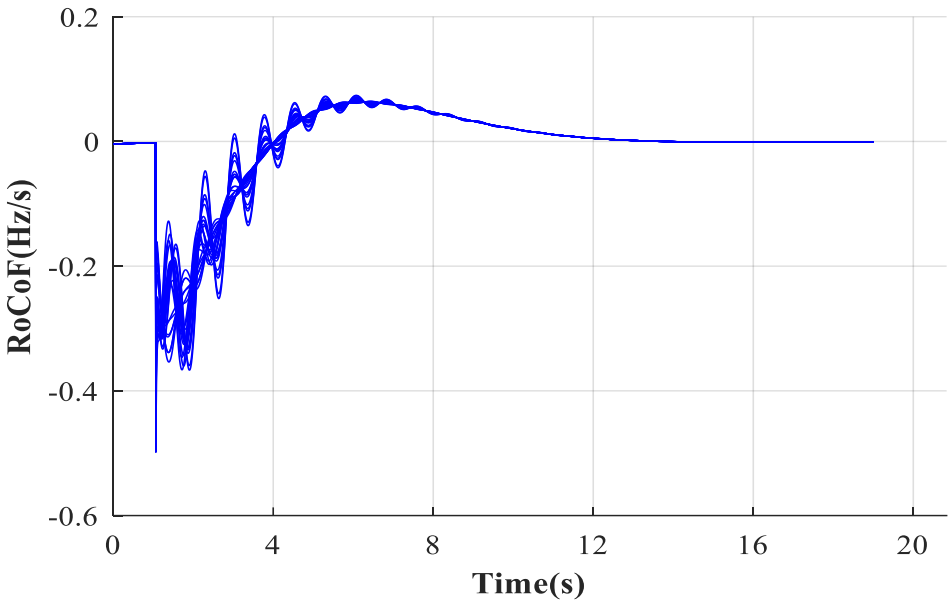


Fig 4.11 Impact of RoCoF constraints on the total system inertia when generator minimum ON/OFF time limits are enforced in SCUC.



(a) ERC-SCUC model.



(b) LRC-SCUC model.

Fig 4.12 RoCoF of all buses following the loss of largest generation in different cases with generator minimum ON/OFF time constraints.

4.6.5 Sensitivity Analysis on RES Penetration Levels

The sensitivity analysis with different renewable penetration levels is conducted in this subsection. Following the daily load profile shown in Fig. 4.3, four scenarios are considered for different RES penetration levels during peak hour from 20% to 80% with an increment of 20%. Given the system peak load is 3,222 MW at hour 12, the corresponding RES generations during peak hour are 644.4 MW, 1,288.8 MW, 1,933.2 MW and 2,577.6 MW respectively.

Table 4.3 presents the aggregated system inertia value at peak hour 12 in different scenarios. As RES penetration level increases from 20% to 80%, the

synchronous inertia of committed synchronous generators based on T-SCUC drops significantly comparing to the LRC-SCUC model.

Table 4.3 System Inertia [MWs] under Different Models

RES Penetration Level	20%	40%	60%	80%
T-SCUC	11,706	10,434	9,342	7,792
ERC-SCUC	14,639	13,063	12,275	10,880
LRC-SCUC	15,599	14,503	13,235	13,095

Table 4.4 Highest ROCOF [Hz/s] under Different Scenarios

RES Penetration Level	20%	40%	60%	80%
T-SCUC	-1.26	-1.32	-1.45	-1.65
ERC-SCUC	-0.75	-0.66	-0.63	-0.62
LRC-SCUC	-0.34	-0.38	-0.42	-0.46

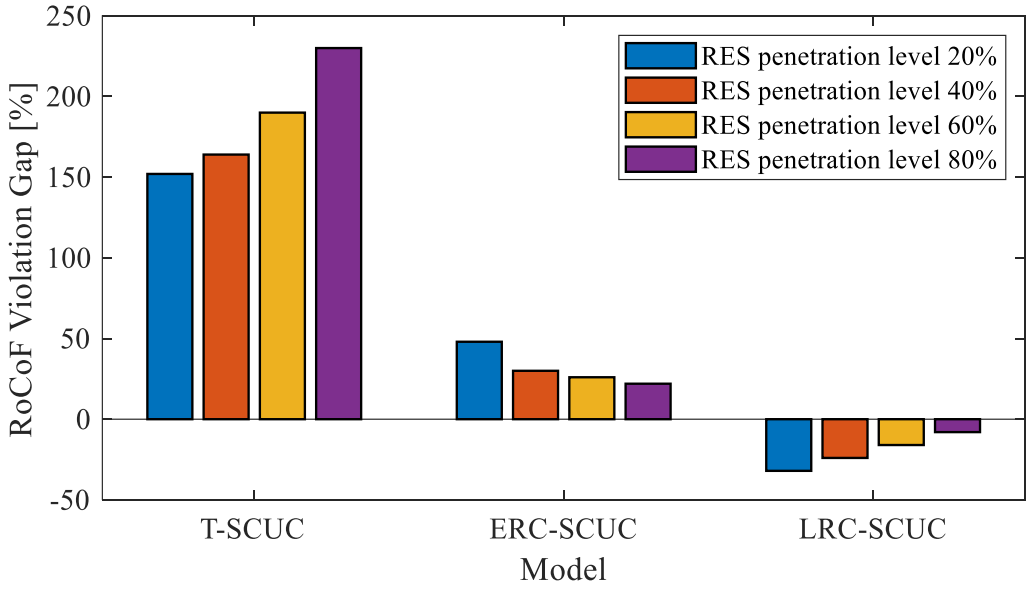


Fig. 4.13 RoCoF violation gaps for different scenarios.

One may be interested to compare the nodal frequency performance of all three SCUC models. Simulations have been conducted in TSAT considering worst-case contingency happens in peak hour 12. Table 4.4 shows the highest RoCoF monitored on generator buses following the $G-1$ contingency under different scenarios. It can be observed that highest RoCoF in both ERC-SCUC and T-SCUC cases violates the limit which may subsequently trip frequency-sensitive protection relays, while the RoCoF is secured above predetermined threshold in all cases with the proposed LRC-SCUC model.

The RoCoF violation gap which indicates difference between the actual values and the thresholds are depicted in Fig. 4.13 [37]. The violation gaps are all positive in T-SCUC and REC-SCUC indicating threshold violations. As RES penetration level increases, less conservative solutions are observed for LRC-SCUC model in this sensitivity test. Understandably, non-uniform distribution caused by generator aggregation is relieved because of less generator aggregation.

Table 4.5 SCUC Costs [$] under Different Scenarios

RES Penetration Level	20%	40%	60%	80%
T-SCUC	1,357,203	1,077,439	891,391	750,199
ERC-SCUC	1,459,983	1,181,493	988,524	845,606
LRC-SCUC	1,546,833	1,253,221	1,050,989	908,353

The costs with those three SCUC models are listed in Table 4.5. Obviously, incorporating frequency related RoCoF constraints in SCUC leads to an increase of system total generation cost in all scenarios. The change in the cost is dependent on the change in penetration level of renewable resources.

4.6.6 Sensitivity Analysis with PWL Evaluation Points

In this section, we compare the proposed method for different evaluation points in terms of computational time and frequency dynamics. Results in Table 4.6 show that more evaluation points decrease the operational cost. The fitting time for 3 evaluation points case is 78.25 s, while it is 165.16 s for 4 evaluation points case. Since fitting process is run off-line, the fitting process wouldn't affect the efficiency the proposed model. It can be observed that computational time for solving proposed LRC-SCUC problem is close for both cases.

Table 4.6 SCUC Results with Different PWL Evaluation Points

PWL	Operational Cost [$]	Computational Time [s]	Fitting Time [s]
$\eta = 3, v = 4$	1,050,989	75.45	78.25
$\eta = 4, v = 4$	1,042,346	77.95	165.16

Dynamic simulation of 4 evaluation points case is conducted on TSAT. From Fig. 4.14 we can observe that the highest RoCoF for 4 evaluation points case is -0.49 Hz/s, implying that solution of the model considering 4 evaluation points is less conservative comparing to the case with 3 evaluation points (-0.42 Hz/s). The

proposed LRC-SCUC model with 4 evaluation points takes slightly more time at least for this test case, providing a less conservative solution.

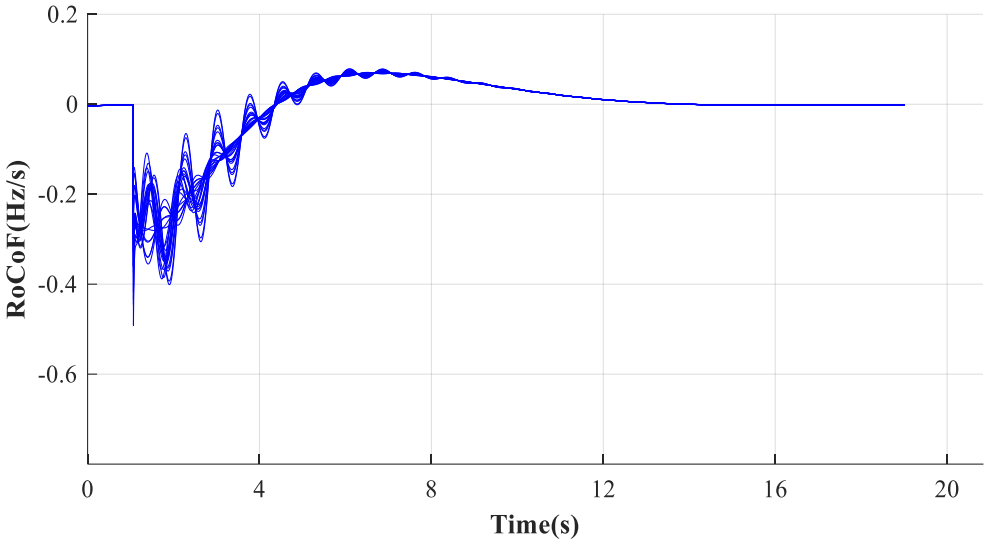


Fig. 4.14 RoCoF evolution of 4 evaluation points case.

4.6.7 Influence of Virtual Inertia on Market Results

In this subsection, the proposed VI-LRC-SCUC model that considers virtual inertia M_t^{VI} is implemented and the effects of virtual inertia on day-ahead scheduling and system stability are examined. We first consider $M_t^{Total} = \infty$ with no upper limit, the cost curves of VI-ERC-SCUC and VI-LRC-SCUC models are depicted in Fig. 4.15. Results show that introducing virtual inertia would potentially reduce the total operational cost of VI-LRC-SCUC model, while cheaper inertia price leads to higher reduction value. With inertia price at 0$/MWs, the total operational cost of VI-LRC-SCUC is equal to the cost of T-SCUC, while it is close to the cost of LRC-SCUC when inertia price is over 0.75$/MWs.

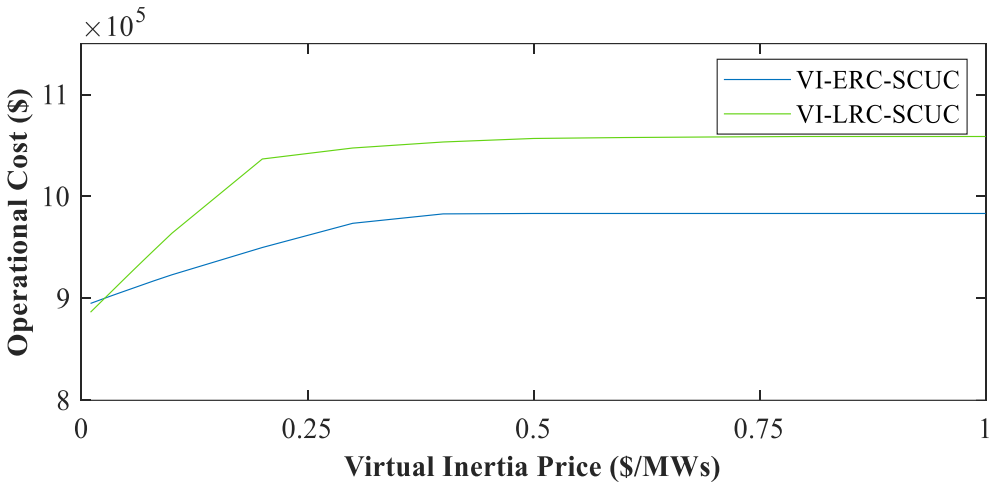


Fig. 4.15 Operational cost curve.

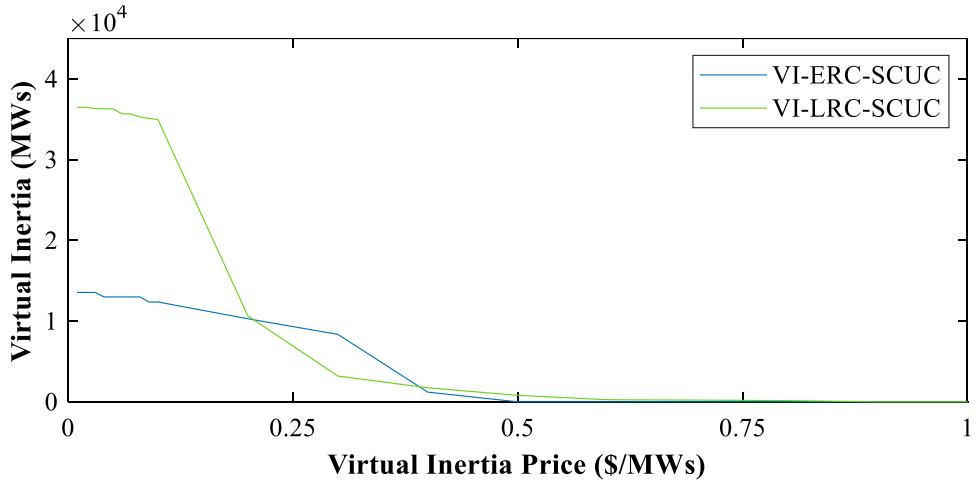


Fig. 4.16 Cost curve of virtual inertia provision.

The cost curve of virtual inertia provision has been depicted in Fig. 4.16. As can be seen, VI-LRC-SCUC is more sensitive to price change of virtual inertia than VI-LRC-SCUC model. The results also indicate that cheaper inertia price doesn't increase the value of needed virtual inertia in both cases. When virtual inertia price increase above 0.5$/MWs, the virtual inertia purchased by the system would decrease

to zero, that means system would rather turn on conventional synchronous generators than purchasing additional virtual inertia.

Accounting for the available capacity and the total cost of the inertia-emulating devices, there is a total budget constraint for additional virtual inertia [31], the total available virtual inertia $\mathsf{M}_t^{\mathsf{Total}}$ is considered as 2,000 MWs. To evaluate the influence of including virtual inertia based RoCoF constraints in SCUC models, we analyze the average market results of five models over 24 hours.

Table 4.7 Average LMP with Different SCUC Models

Model	Average LMP [$/MWh]	Average congestion LMP [$/MWh]	Energy LMP [$/MWh]
T-SCUC	37.56	23.31	14.25
ERC-SCUC	41.44	24.20	17.24
VI-ERC-SCUC	40.41	23.17	17.24
LRC-SCUC	45.89	29.09	16.80
VI-LRC-SCUC	37.88	21.08	16.80

As defined in (4.48), system average locational market price (LMP) over the entire scheduling horizon is proposed to investigate the impact of introducing additional inertia in SCUC on system LMP. Similarly, system average congestion LMP is defined in (4.49),

$$AvgLMP = \frac{1}{\|T\| \cdot \|N\|} \sum_{t \epsilon T} \sum_{n \epsilon N} LMP_{n,t}\,, \qquad \forall n,\, t, \tag{4.49}$$

$$AvgLMP_{cg} = \frac{1}{\|T\| \cdot \|N\|} \sum_{t \epsilon T} \sum_{n \epsilon N} LMP_{cg,n,t}\,, \qquad \forall n,\, t. \tag{4.50}$$

The nodal LMP energy component and congestion component results are presented in Table 4.7. With virtual inertia implemented, the system average LMP and congestion LMP over the scheduling horizon are close for both ERC-SCUC and VI-ERC-SCUC cases. In proposed LRC-SCUC cases, system average congestion LMP decreases from 29.09 $/MWh to 21.08 $/MWh when virtual inertia is included, and the average LMP value decreases accordingly from 45.89 $/MWh to 37.88 $/MWh. This result reflects that imposing virtual inertia can significantly reduce the system average LMP and could be an option that guarantees RoCoF security while avoiding high congestion.

Table 4.8 Average Market Results with Different SCUC Models

Model	Load payment [$/h]	Generator revenue [$/h]	Generator cost [$/h]	Generator rent [$/h]	Congestion revenue [$/h]
T-SCUC	80,476	72,589	37,141	35,448	7,887
ERC-SCUC	82,975	59,651	41,189	18,462	23,324
VI-ERC-SCUC	83,202	62,005	40,475	21,530	21,197
LRC-SCUC	127,770	52,882	43,791	9,091	74,888
VI-LRC-SCUC	81,112	61,827	43,315	18,512	19,285

Table 4.8 details the market results of all cases based on the nodal LMP over the scheduling horizon, and the definitions of the indexes in this table are adapted from [98]. Results show that the congestion revenue is relatively low for both T-SCUC and ERC-SCUC comparing to the proposed LRC-SCUC model. Significant

reduction can be observed in both average congestion and load payment when additional inertia is introduced. Moreover, it is interesting to observe although VI-ERC-SCUC achieves a lower cost than ERC-SCUC, its load payment is even higher. In general, simulation results reflect that introducing virtual inertia could be a cost-effective approach. Such action enables generation flexibility of generators which might be strictly constrained by frequency related constraints, subsequently lowering the burden of transmission lines and reducing the network congestion.

Note that the metrics averaged over the entire scheduling horizon presented in Tables VII&VIII reflect a positive impact of virtual inertia on power system economics. Inertia service pricing mechanism for pricing and payment of inertia for all participating units requires further deep investigation

4.7 Summary

In this chapter, the concept of locational frequency security is introduced. We first investigate the impact of Fiedler mode on locational frequency dynamics, and then the expression of locational frequency dynamics is defined accounting for *G*-1 contingency in multi-machine systems. To capture the highest locational RoCoF during the oscillation, a multiple-measurement-window method is introduced. Furthermore, a piecewise linearization based method is then proposed to convert the non-linear frequency constraints into linear frequency constraints in the proposed LRC-SCUC model, which allows us to optimally schedule the synchronous inertia as well as inertial services provided by non-synchronous resources to meet the minimum system inertia requirement for power systems with higher RES integration.

Simulation results show that imposing location based RoCoF constraints in the SCUC model can ensure the locational frequency security during worst-case contingency event. Such RoCoF-related constraints also significantly affect the scheduling of synchronous generators and consequently the expected system cost. The effect of virtual inertia on inertia pricing and market efficient is examined, results imply that introducing virtual inertia can reduce the total cost by avoiding starting up unnecessary commitment of extra expensive synchronous generators. Compared to VI-ERC-SCUC model, VI-LRC-SCUCU is more sensitive to inertia price.

5. DEEP LEARNING BASED RCUC

5.1 Literature Review

The operation of the power system in the short term is a multifaceted process that commences with day-ahead markets, where the schedules of generators are determined to minimize operational costs. In this stage, unit commitment is employed as an optimization problem to meet the anticipated supply and demand for the following day. The day-ahead market plays a crucial role in scheduling and committing a significant portion of the demand, making it an essential step in power system operations. Given that the optimization problem includes variables and constraints related to the on/off status of generators, it becomes a mixed-integer linear program (MILP).

With the increased penetration of renewable energy sources (RES), maintaining power system frequency stability has become a great challenge for reliable system operations [99]. Traditionally, synchronous generators play an important role in regulating frequency excursion and rate of change of frequency (RoCoF) after a disturbance as it ensures slower frequency dynamics. Due to the retirement and replacement of conventional generation, more generation is coming from converter-based resources such as wind and solar power. Consequently, the system kinetic energy decreases significantly, leaving the system more vulnerable to large variations in load or generation [100]. When RoCoF violates the necessary industrial control and operation standards, protection devices would disconnect generators from the grid. Insufficient inertia exacerbates the need for fast frequency

response services to secure frequency stability [101].

Requirements for adequate frequency control of the electric power system were suggested by Federal Energy Regulatory Commission [102]. The Swedish TSO ordered the Oskarshamn Kraftgrupp to reduce its nuclear power output by 100 MW to mitigate the risk of loss of the power plant in Sweden [103]. Transmission system operators (TSOs) have also suggested to impose extra RoCoF related constraints in the conventional unit commitment (UC) model to keep the minimum amount of synchronous inertia online.

Model-based approaches utilize the swing equation with a piecewise linear mechanical power approximation to analyze the system dynamics, and different versions of RoCoF security conditions can be derived and embedded into the optimization formulation. Several papers have included frequency related constraints into traditional security-constrained unit commitment (SCUC) formulations. In [104], uniform frequency response model was extended by including converter-based control, and constraints on RoCoF are then derived and incorporated into SCUC formulations. However, these approaches oversimplify the problem as they neglect nodal frequency dynamics, and the actual need for frequency ancillary services would be underestimated. Ref. [105] considers the geographical discrepancies and connectivity impacts on nodal frequency dynamics. Results show that model-based approaches may fail to handle higher order characteristics and nonlinearities in system frequency response, approximation in model may also introduce extra errors into the derived constraints, resulting in conservative solutions.

Neural network-based techniques have demonstrated the ability to capture and track system frequency dynamics and associated conditions, thereby improving the performance of conventional SCUC problems. The availability of reliable data and appropriate models for training is a crucial aspect for deep learning (DL) methods to yield high-quality results. Given that the SCUC is performed on a daily basis, historical data can be effectively utilized to learn non-linear correlations between input variables and corresponding outputs. DL has been successfully utilized in the prediction or decision support in complex problems in various power system fields [106]–[109]. A Classification decision tree for frequency security was designed in [110] for traditional unit commitment formulation. A pioneering data-driven approach has been proposed in [111], which incorporates neural network-based constraints against the worst-case contingency into formulations. The overview of the data driven RCUC is described in the following sections using the working pipeline shown in Fig. 5.1.

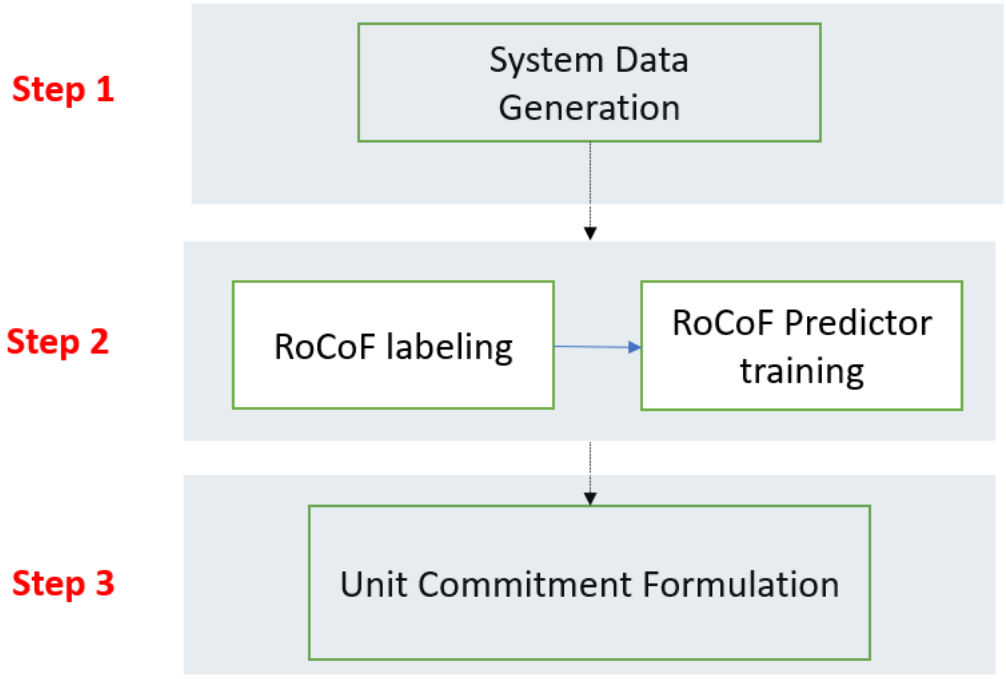


Fig. 5.1 Overview of the data driven RCUC framework.

5.2 System Uniform and Dynamic Models

The frequency of the power system is one of the most important metrics that indicate the system stability. Uniform frequency models are often used in power system studies and analyses, especially for steady-state and simplified calculations. Traditionally, the frequency is treated as unique of the whole system. In these models, 0.the frequency is typically assumed to be equal at all buses or nodes in the power system. The rotating inertia of a synchronous generator is equal to the stored energy E_i in the rotors of the machine at nominal speed. The rotational inertia of a single shaft is commonly defined using its inertia constant and the rated apparent power [112]. For a single machine, the inertia constant is expressed as follows,

$$H_i = \frac{J_i {\omega_i}^2}{2S_{B_i}}, \tag{5.1}$$

where H_i is the inertia constant of the generator in seconds; J_i is the moment of inertia of the shaft in kg·m^{2}s; S_{B_i} is the base power in MVA; and ω_i is the nominal rotational speed instead of the actual speed of the machine.

Power system inertia is defined as the total amount of rotational energy stored in all rotating synchronous units; dynamics of these generators' rotors are directly coupled with the grid electrical dynamics. It can be expressed as follows. For a single generator i, the swing equation is expressed as,

$$\frac{d\omega_i}{dt} = \frac{P_m - P_{load}}{2H_i S_{B_i}} \omega_n, \tag{5.2}$$

where P_m is the mechanical power and P_{load} is the load from the power system, while ω_n is the rated steady state frequency of the system. dω_i/dt is more commonly known as RoCoF. The swing equation of the system equivalent model can be then applied to the whole grid [113]. After a disturbance of power mismatch occurrence, the system RoCoF related to the total system inertia can be defined as,

$$RE^{t_0} = \frac{-\Delta P}{2H_{sys} S_B} \omega_n, \tag{5.3}$$

where ΔP is the sudden change in active power in MW at t=t_0.

However, it is important to note that uniform frequency models have limitations and may not accurately capture the actual frequency response of a power system during transient or dynamic events. They do not consider the effects of system inertia, generator dynamics, and frequency control mechanisms that impact frequency

deviations and response. For more accurate and detailed analysis, especially in scenarios involving significant renewable energy integration or system disturbances, more sophisticated and non-uniform frequency models that account for dynamic effects and control actions are required. Following a sudden change in load or a generation contingency, the dynamic model (5.4) is used to describe the evolution of system level frequency deviation following the maximum infeed generation loss $\Delta\bar{P}$,

$$\Delta\bar{P} = M\frac{\partial\Delta\omega}{\partial t} + D\Delta\omega, \tag{5.4}$$

where M_{sys} and D_{sys} are the aggregated system inertia constant and damping coefficient corresponding to the committed synchronous generators respectively. However, only considering the dynamics of the equivalent model in systems would underestimate the actual need for frequency ancillary services, leading to higher locational RoCoF and larger regional frequency deviation than expected.

Using the topological information and the system parameters, the transmission network can be modeled as a graph consisting of nodes (buses) and edges (branches). The oscillatory behavior of each individual bus can be expressed as follows,

$$m_i\ddot{\theta}_i + d_i\dot{\theta}_i = P_{in,i} - P_{e,i}, \qquad i \in N_G. \tag{5.5}$$

Note that the dynamic model describes the frequency evolution on generator buses, we assumed instantaneous and reliable measurements of frequency and related response are triggered about 100 ms after the fault and they have minor effect on the initial RoCoF values.

5.3 DNN based RCUC Model

5.3.1 DNN based RoCoF Prediction

Prediction of the highest nodal RoCoF is quite challenging due to the non-linear nature of the power system. Although efforts like [96] improve the frequency dynamics model by including the state-of-the-art converter control schemes, model-based approaches are unable to capture the entire characteristics and incorporate high-order models. DNN has shown the ability to amend the limitations of model-based approaches [111].

Given the load forecast d and RES forecast r, following a disturbance ϖ_τ at period τ, the highest locational RoCoF value of the system R_h is a function with respect to the disturbance level, disturbance location, system states, dispatch command, load condition and RES profile,

$$R_h = h^r(s_\tau, u_\tau, d_\tau, r_\tau, \varpi_\tau), \tag{5.6}$$

where s_τ denotes the system states, and u_τ is the generation dispatch at period τ. Compared to the case of a sudden load increase, the loss of generation not only causes mismatch in system power balance but also degrades the system synchronous inertia, resulting in even higher frequency deviation and larger initial RoCoF. Based on our discussion regarding the impact of Fiedler mode on inertial response, disturbance ϖ_τ is defined to be the *G*-1 contingency, the magnitude of event is determined by the generator output power. The DNN based RoCoF predictor for h^r is then expressed as,

$$\hat{R}_h = \hat{h}^r(x, W, b), \tag{5.7}$$

where x is the feature vector, and W and b denote the parameters of a well-trained DNN. Within the transient time period following a disturbance, the characteristics of all generators are assumed unchanged, thus, it is sufficient to select the status of each SG as input features. Both the magnitude and location of the disturbance will affect the inertial response. The generator status feature vector for sample s is defined as follows,

$$u_s = \left[u_{1,s}, u_{2,s}, \cdots, u_{N_G,s}\right]. \tag{5.8}$$

The disturbance feature vector is defined against the loss of largest generation, the magnitude is expressed as,

$$P_s^{\varpi} = \max_{g \in G}\left(P_{1,s}, \cdots, P_{2,s}, \cdots, P_{N_G,s}\right). \tag{5.9}$$

The location of the disturbance is then represented by the index of the generator,

$$g_s^{\varpi} = arg\max_{g \in G}\left(P_{1,s}, \cdots, P_{2,s}, \cdots, P_{N_G,s}\right). \tag{5.10}$$

We encode the information of magnitude and location into the disturbance feature vector as,

$$\varpi_s^G = \left[0, \cdots, 0, \underbrace{P_s^{\varpi}}_{g_s^{\varpi} th\ element}, 0, \cdots, 0\right]. \tag{5.11}$$

Laplacian matrix L of the grid and Fiedler mode value depend on the power-angle characteristics, which are determined by the active power injection. Thus, the active power injection of all SGs will be encoded into the feature vector,

$$P_s = \left[P_{1,s}, \cdots, P_{2,s}, \cdots, P_{N_G,s}\right]. \tag{5.12}$$

The overall feature vector of a sample s can be then defined as follows,

$$x_s = [u_s, \varpi_s^G, P_s]. \tag{5.13}$$

Now consider a fully connected neural network with N_L hidden layer. Each layer uses a ReLU activation function as $\sigma(\cdot) = max\ (\cdot, 0)$ and the output layer is a linear activation function. The predicted RoCoF can be expressed as follows,

$$z_1 = x_s W_1 + b_1, \tag{5.14}$$

$$\hat{z}_q = z_{q-1} W_q + b_q, \tag{5.15}$$

$$z_q = \max(\hat{z}_q, 0), \tag{5.16}$$

$$R_{h,s} = z_{N_L} W_{N_L+1} + b_{N_L+1}, \tag{5.17}$$

where W_q and b_q represent the weight and bias for the m-th hidden layer, and W_{N_L+1} and b_{N_L+1} represent the set of weight and bias of the output layer.

5.3.2 Model based Data Generation

Training samples are generated from models over various load and RES scenarios, T-SCUC models and frequency constrained SCUC models are implemented in this process. Given the load forecast and RES forecast, the T-SCUC is the base model generating dispatching samples. Objective function is to minimize the total system cost consisting of variable fuel costs, no-load costs, start-up costs, and reserve costs.

As mentioned before, two RoCoF constrained SCUC models are utilized. It is worth pointing out that the primary response is neglected in the formulation, without affecting the inertial response and the fundamental findings of this work. For system equivalent model based ERC-SCUC model, constraint (5.18) is introduced to

guarantee generator frequency stability considering the relative location to the potential $\mathsf{G}-1$ contingency,

$$RE^{t_0}\left(\Delta P, H_{sys}, S_B\right) \leq -RoCoF_{lim}, \quad \forall g, t. \tag{5.18}$$

LRC-SCUC introduces locational RoCoF constraints based on the definition of local buses N_{loc} and non-local buses N_{n-loc}. Constraints (5.19) and (5.20) ensure system stability by imposing limit on locational RoCoF over all buses under all $G-1$ contingency,

$$R_n{}^{T_1}\left(\Delta P, H_{sys}, S_B\right) \leq -RoCoF_{lim}, \forall n \in N_{loc}, g, t, \tag{5.19}$$

$$R_n{}^{T_2}(\Delta P, H_{sys}, S_B) \leq -RoCoF_{lim}, \forall n \in N_{n-loc}, g, t. \tag{5.20}$$

The constraints on RoCoF for locational frequency dynamics are nonlinear. In order to incorporate these frequency-related constraints into the proposed LRC-SCUC model, a linear approximation method is introduced. The detail of all models is presented in [105].

5.3.3 DNN Linearization

DNN linearization refers to the process of approximating a deep neural network model with a linear model. It involves simplifying the non-linear behavior of the DNN model. To encode the neural networks into the MILP SCUC problem, decision variables are introduced to build the disturbance feature vector. binary variable $\lambda_{g,t}^{G}$ is used to indicate the status of largest output power of generator g in scheduling period t, a big-M method is introduced to express the disturbance vector,

$$P_{\rho,t} - P_{g,t} \leq M\left(1-\lambda_{g,t}^{G}\right), \forall \rho, g, t, \tag{5.21}$$

$$\sum_{g \in G} \lambda_{g,t}^{G} = 1 \; \forall t \tag{5.22}$$

where M is a big positive number. Equation (5.21) enforces $\lambda_{g,t}^{G}$ to be zero is the dispatched output power of other generators ρ is larger than generator g at period t, while (5.22) limit the number of potential largest generator to be one at one period. When generator g has the largest output power, (5.23) and (5.24) would set the status of largest output power $\lambda_{g,t}^{G}$ to be 1. To express the magnitude of disturbance, variable $\varepsilon_{g,t}$ is defined as the indicators of disturbance,

$$\boldsymbol{\varepsilon}_{g,t} - \mathrm{P}_{g,t} \geq -M\left(1 - \lambda_{g,t}^{G}\right), \forall g,t, \tag{5.23}$$

$$\boldsymbol{\varepsilon}_{g,t} - \mathrm{P}_{g,t} \leq M\left(1 - \lambda_{g,t}^{G}\right), \forall g,t, \tag{5.24}$$

$$0 \leq \boldsymbol{\varepsilon}_{g,t} \leq M\lambda_{g,t}^{G}, \forall g,t. \tag{5.25}$$

Thus, the input feature vector can be expressed as follows,

$$x_t = \left[u_{1,t}, \cdots, u_{N_G,t}, \; \varepsilon_{g,t}, \cdots, \; \varepsilon_{N_G,t}, P_{1,s}, \cdots, P_{N_G,s}\right]. \tag{5.26}$$

Since ReLU activation functions are nonlinear, to include the DNN into the MILP, binary variables $a_{q[l]}$ are introduced which represent the activation status of the ReLU unit at lth neuron of qth layer. For a given sample s , consider A is a big number that is larger than the absolute value of all $\hat{z}_{q[l],s}$. When preactivated value $\hat{z}_{q[l],s}$ is larger than zero, constraints (5.27) and (5.28) will force binary variable $a_{q[l],s}$ to one, and the activated value will be equal to $\hat{z}_{q[l],s}$. When $\hat{z}_{q[l],s}$ is less than or equal to zero, constraints (5.29) and (5.31) will force binary variable $a_{q[l],s}$ to zero. Subsequently, the activated value $z_{q[l],s}$ will be set zero,

$$z_{q[l],s} \leq \hat{z}_{q[l],s} + A(1 - a_{q[l],s}), \forall q, \forall l, \forall s, \tag{5.27}$$

$$z_{q[l],s} \geq \hat{z}_{q[l],s}, \forall q, \forall l, \forall s, \tag{5.28}$$

$$z_{q[l],s} \leq A a_{q[l],s}, \forall q, \forall l, \forall s, \tag{5.29}$$

$$z_{q[l],s} \geq 0, \forall q, \forall l, \forall s, \tag{5.30}$$

$$a_{q[l],s} \in \{0, 1\}, \forall q, \forall l, \forall s. \tag{5.31}$$

5.3.4 Results and Analysis of DNN-RCUC

A case study on IEEE 24-bus system is provided to demonstrate the effectiveness of the proposed DNN-RCUC methods. This test system contains 24 buses, 33 generators and 38 lines, which also considers decarbonized generation characterized by wind power. The base case has a total demand from 1,195 MW to a peak of 2,116 MW. To ensure the practicality of the dataset and the generality of the trained model, load profile and RES profile are sampled based on Gaussian distribution while the deviation of means value ranges from [-20%, 20%] of the based value. The mathematical model-based data generation is operated in Python using Pyomo. Regarding post-contingency frequency limits, RoCoF must be higher than -0.5Hz/s to avoid the tripping of RoCoF-sensitive protection relays, and the optimality gap is set to 0.1%. The PSS/E software is used for time domain simulation and labeling process, and full-scale models with detailed generator dynamics are implemented for more realistic data.

5.3.4.1 DNN Predictor Training

The base vector has a dimension of 99. For the DNN layers, the number of neurons is set 10 for each layer. ReLU is used as the activation function. The architecture of proposed DNN model is showed in Fig. 5.2.

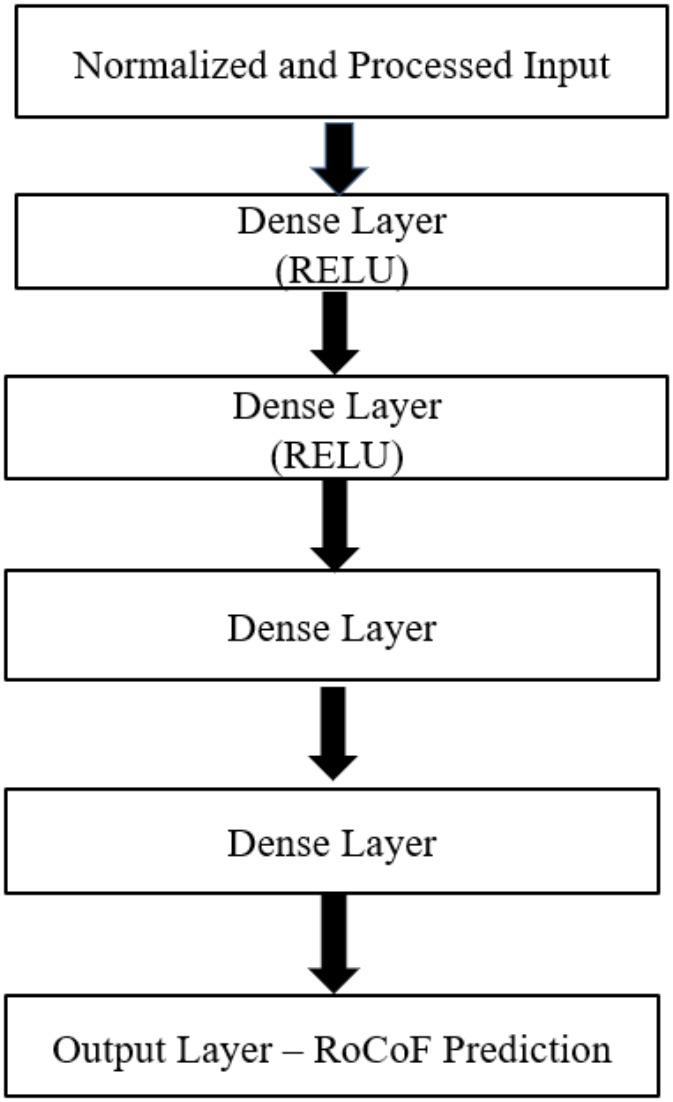


Fig. 5.2 Architecture of proposed DNN model.

The training was operated in batches of 32 data points. An MSE based dynamic learning rate strategy is used for the training. Learning rate schedule is applied in the training process by reducing the learning rate accordingly, the factor by which the learning rate will be reduced is set to 0.5 and the patience value is set to 10 epochs. A total of 9,600 samples were collected based on strategies proposed in previous section. The entire dataset is divided into two subsets: 7,680 samples (80%) for training and 1,920 samples (20%) for validation.

Table 5.3 shows the validation accuracy of the RoCoF predictor under different tolerances. The validation accuracy is 99.27% with 10% tolerance, implying high performance of the trained model. It should be noted that the accuracy is still above 93.55% even with a small tolerance of 5%, indicating the robustness of the trained predictor.

Table 5.1 Validation Accuracy of the proposed DNN-based ROCOF Predictor

Tolerance	10%	9%	8%	7%	6%	5%
Accuracy	99.3%	99.0%	98.5%	96.6%	95.5%	93.6%

5.3.4.2 DNN-RCUC Simulation Results

The forecast load and wind power for test case are plotted in Fig. 5.3 and Fig. 5.4. Due to the computational efficiency, one interval in the simulation process represents 4 hours in periods. The test case has a demand ranging from 1,633 MW to a peak of 1,853 MW. The peak wind generation is 266 MW. All three RoCoF-constrained models are tested on the same test case.

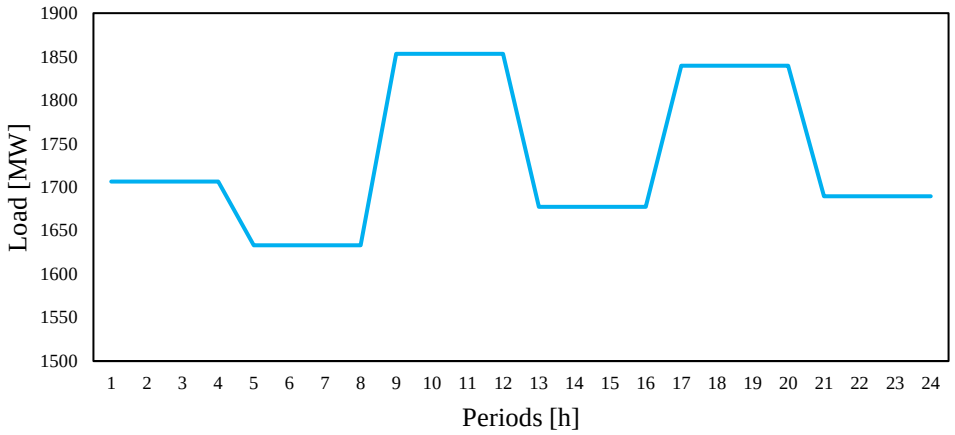


Fig. 5.3 Load profile of the IEEE 24-bus system.

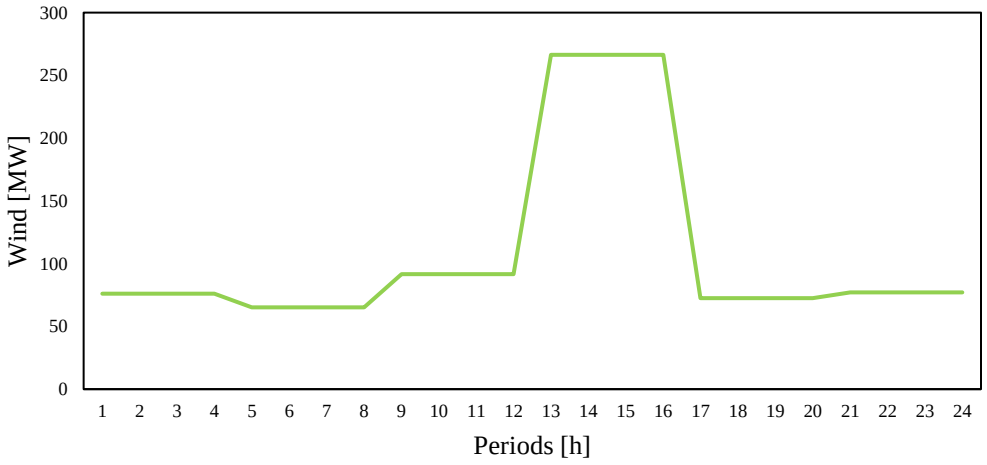


Fig. 5.4 Wind generation of the IEEE 24-bus system.

Table 5.1 SCUC Costs [$] Of Different Models

Model	Total	Start-up	Operation	Reserves
ERC-SCUC	540,128	28,536	443,304	68,288
LRC-SCUC	732,864	40,372	645,884	46,608
DNN-RCUC	1,093,175	48,101	1,033,759	45,260

Table 5.2 compares the unit commitment results of the proposed DNN-based RCUC model and the two mathematical-based SCUC models. It can be observed that the proposed DNN-RCUC has the highest operational cost among all three models; the extra cost results from the efforts in handling the generator aggregation situations, which would be discussed later. On the other hand, the reserve cost is less in LRC-SCUC and DNN-RCUC models. For DNN-RCUC case, the total reserve cost is $46,608, which is slightly lower than the cost of LRC-SCUC model. Additional synchronous machines are committed to cover the loss of largest generation for DNN-RCUC model, which accordingly increases the operation cost as well as the start-up cost.

Additionally, we run the dynamic simulation of G-1 contingency for all three models when the system netload is the lowest. The loss of largest generation at this period is more likely to result in highest RoCoF and largest system deviations due to least synchronous generators online. The highest RoCoF of three cases are listed in TABLE IV. Although with system equivalent model-based RoCoF constraints, ERC-SCUC model still cannot ensure system RoCoF security under such situation. The highest RoCoF of LRC-SCUC model is 0.3920 Hz/s, which gives a relatively high constraints gap at -21.60% below limit. The proposed DNN-RCUC has a highest RoCoF of 0.4952 Hz/s following the loss of largest generation. From TABLE III and Fig. 5.5, it can be concluded that the proposed DNN-RCUC model can secure the system with minimal constraints gap while LRC-SCUC leads to conservative results.

Table 5.2 Highest RoCoF of Different Models

Model	ERC-SCUC	LRC-SCUC	DNN-RCUC
Highest RoCoF [Hz/s]	0.6127	0.3920	0.4952

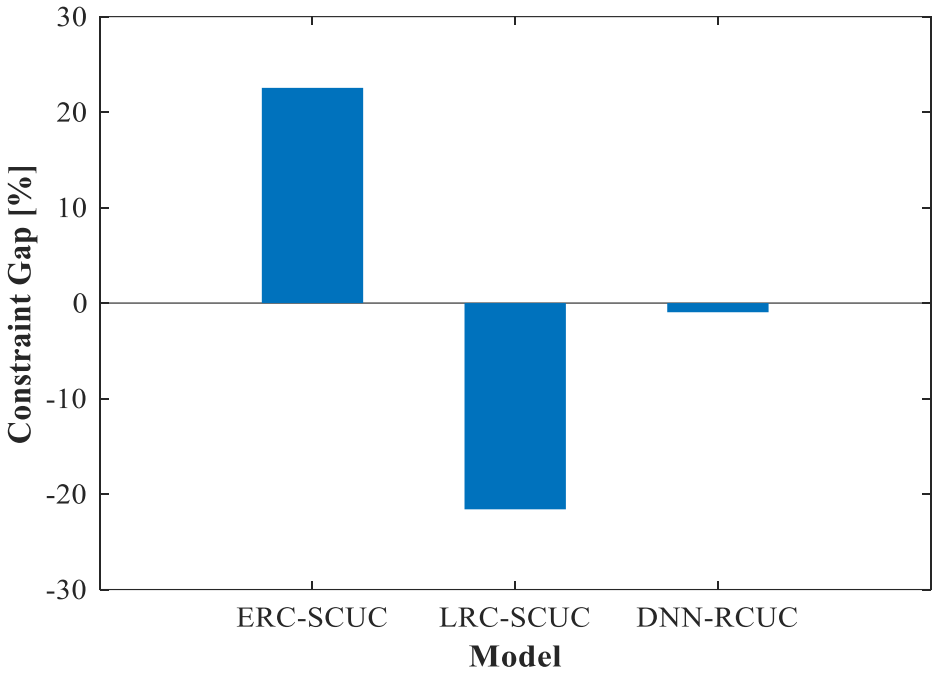


Fig. 5.5 Constraint gaps for different cases.

Table 5.4 Highest RoCoF of Different Models Considering Generator Aggregations

Model	LRC-SCUC	DNN-RCUC
Highest RoCoF [Hz/s]	0.8125	0.4993

Furthermore, the stability issue of generator aggregation is investigated. In this work, we run time domain simulation of two locational RoCoF constrained models in the scenario considering the aggregation of generators on bus 23. Following the loss of the largest generation at hour 11, the highest RoCoF of two models are compared in Table 5.4. It can be observed that in the case of LRC-SCUC model, trip

of the largest generation on bus 23 may cause generator on the same event bus violates the pre-specified RoCoF limit 0.5 Hz/s easily. When the constraints based on DNN are implemented, the system RoCoF can be secured just below the threshold.

5.4 CNN based RCUC

5.4.1 CNN based RoCoF Prediction

CNN is a well-known deep learning architecture inspired by the natural visual perception mechanism of the living creatures [114]. Many methods have been developed to overcome the difficulties encountered in training deep CNNs [115]–[118]. CNNs are widely used in image processing and computer vision tasks, as they are effective in capturing spatial patterns and features. Recent advancement of deep neural network has realized that CNN has the ability to extract multi-scale localized spatial features from the power system and compose them to construct highly expressive representations of the input features x_s, which help improve the performance of machined learning assisted SCUC.

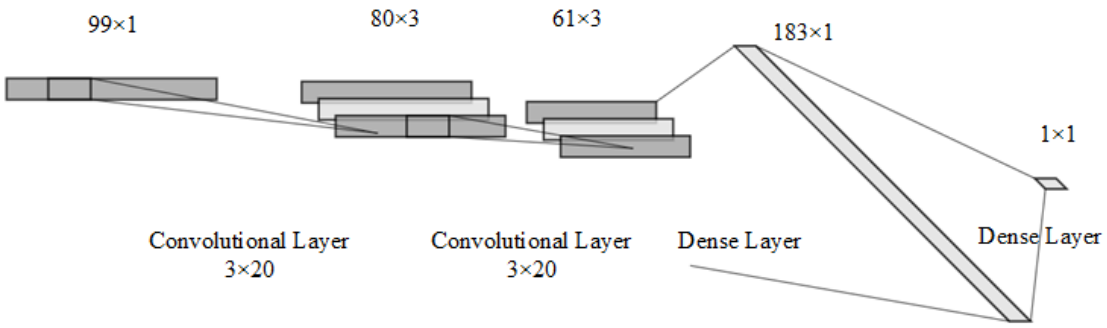


Fig. 5.6 Architecture of proposed CNN model (NN-SVG).

The CNN model used in this study is denoted in Fig. 5.6. The proposed model consists of two types of layers, namely convolutional and fully connected layers. The convolutional layer aim to learn feature representations of the input. The convolutional layer is composed of ϵ convolutional kernels which are used to compute different feature maps. In essence, the neurons within a feature map establish connections with neighboring neurons in the preceding layer, forming a receptive field for each neuron. This receptive field is a region of influence that affects the neuron's input. To generate a new feature map, the input is convolved with a learned kernel and then subjected to an element-wise nonlinear activation function on the convolved outcomes. It's important to note that in the process of generating each feature map, the kernel is utilized across all spatial locations of the input, effectively sharing the same kernel weights throughout the entire input volume. The forward propagation equations are defined as,

$$\hat{z}_{i,j,\epsilon}^{q} = x_{i,j} w_{\epsilon}^{q} + b_{\epsilon}^{q}, \tag{5.32}$$

$$z_{i,j,\epsilon}^{q} = ma\,x\left(\hat{z}_{i,j,\epsilon}^{q}, 0\right), \tag{5.33}$$

where w_{ϵ}^{q} and b_{ϵ}^{q} are the weight vector and bias term of the ϵ-th filter of the q-th layer respectively, and $x_{i,j}$ is the input patch centered at location (i,j) of the q-th layer. It should be noted that the kernel w_{ϵ}^{q} that generates the feature map is shared, such mechanism can reduce the model complexity and improve the efficiency of the model. ReLU is used as the activation function for introducing nonlinearities to CNN.

Before feeding into the fully connected layer, the convolved features should be flattened in advance. Denoting the flatten function as $flatten(\cdot)$, for the generated feature map $z_{i,j,\epsilon}^{q}$ we have,

$$z_{full} = flatten\left(z_{i,j,\epsilon}^{Con}\right), \tag{5.34}$$

$$z_{full+1} = z_{full}W_{full} + b_{full}, \tag{5.35}$$

$$R_{h,s} = z_{N_L}W_{N_L+1} + b_{N_L+1}, \tag{5.36}$$

where Con is the last convolutional layer, W_{full} and b_{full} represent the weight and bias for the fully connected layer. W_{N_L+1} and b_{N_L+1} represent the set of weight and bias of the output layer.

5.4.2 CNN Linearization

Nonlinearity in CNN cannot be directly handled by MLIP solver. CNN linearization is to replace non-linear activation functions, such as ReLU, with linear functions. This effectively removes the non-linearities and further incorporated it into RCUC model. In order to include the CNN into the MILP, binary variables $a_{i,j,\epsilon}^{q}$ are introduced which represent the activation status of the ReLU unit at lth neuron of qth layer. For a given sample s , consider A is a big number that is larger than the absolute value of all $\hat{z}_{i,j,\epsilon}^{q}$ generated by kernel ϵ. When preactivated value $\hat{z}_{i,j,\epsilon}^{q}$ is larger than zero, constraints (5.37) and (5.38) will force binary variable $a_{i,j,\epsilon}^{q}$ to one, and the activated value will be equal to $\hat{z}_{i,j,\epsilon}^{q}$. When $\hat{z}_{i,j,\epsilon}^{q}$ is less than or equal to zero,

constraints (5.39) and (5.41) will force binary variable $a_{i,j,\epsilon}^{q}$ to zero. Subsequently, the activated value $z_{i,j,\epsilon}^{q}$ will be set zero,

$$z_{i,j,\epsilon,s}^{q} \leq \hat{z}_{i,j,\epsilon,s}^{q} + A(1 - a_{i,j,\epsilon,s}^{q}), \forall q, \forall \epsilon, \forall s, \tag{5.37}$$

$$z_{i,j,\epsilon,s}^{q} \geq \hat{z}_{i,j,\epsilon,s}^{q}, \forall q, \forall \epsilon, \forall s, \tag{5.38}$$

$$z_{i,j,\epsilon,s}^{q} \leq A a_{i,j,\epsilon,s}^{q}, \forall q, \forall \epsilon, \forall s, \tag{5.39}$$

$$z_{i,j,\epsilon,s}^{q} \geq 0, \forall q, \forall \epsilon, \forall s, \tag{5.40}$$

$$a_{i,j,\epsilon,s}^{q} \in \{0, 1\}, \forall q, \forall \epsilon, \forall s. \tag{5.41}$$

5.4.3 Results Analysis

The data creation and verification steps are implemented using Pyomo and PSSE. The SCUC model is solved using Gurobi solver. For machine learning step is implemented in Python 3.6. A computer with Intel® Xeon(R) W-2295 CPU @ 3.00GHz, 192 GB of RAM and NVIDIA GeForce RTX 2060, 6GB GPU was utilized.

Table 5.4 Validation Accuracy of the CNN based RoCoF Predictor

Tolerance	5%	10%	15%
DNN	92.78	97.49	98.93
CNN	93.53	97.96	99.17

Several common classification methods are first compared with proposed CNN model on the IEEE 24-Bus system data. The proposed CNN based RoCoF predictor is compared with benchmark DNN in Table 5.5. The results show that with a tolerance of 5%, the proposed CNN model has a validation accuracy of 92.78%. For the benchmark DNN model, the validation accuracy with 5% tolerance is calculated as

93.53% on the same validation dataset, which is relatively lower than the proposed CNN model. Fig. 5.7 presents the evolution of MSE losses on the training and validation sets over the training process of the proposed LRCN model. As can be seen MSE decreases as the number of epochs increases.

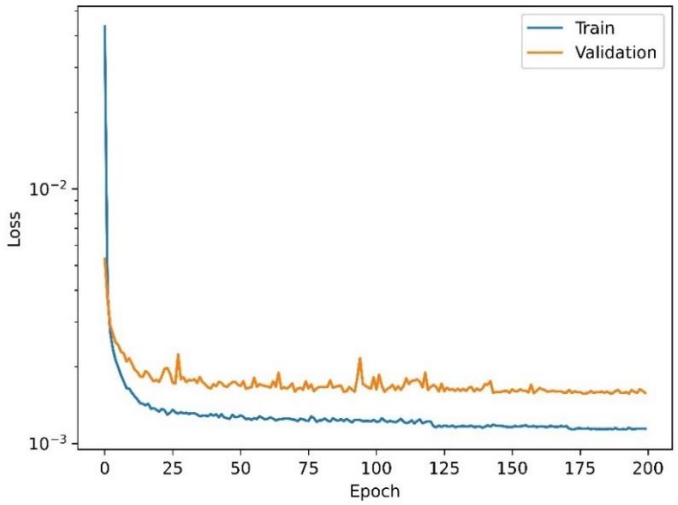


Fig. 5.7 Learning curve of CNN model.

Additionally, we validate performance of CNN predictor using the following metrics to demonstrate the prediction accuracy: (1) median absolute error (MED-E), (2) mean absolute error (MEA-E), and (3) R2 score. From Table 5.7 we can observe that CNN based RoCoF predictor has lower MSE as well as MEA-E, indicating that CNN model has a better performance in processing power system data with graphical information embedded.

Table 5.4 Comparison of Different Models

Model	MED-E	MEA-E	R^2
DNN	0.0072	0.0039	0.9828
CNN	0.0055	0.0021	0.9893

IEEE 24-bus system for 24-hour period is used for UC validation. Electricity demand ranges from 1,300 MW to a peak of 1,786 MW. The mathematical models are implemented in Python using Pyomo and solved with the Gurobi solver, and the optimality gap is set to 0.1%. The results presented in Table 5.7 show that the all RoCoF constrained SCUC models alleviate the reserve cost over T-SCUC. Especially for CNN-RCUC, the reserve cost is reduced from $83,475 to $61,882, giving a reduction of 25.87%. An increase in startup cost can also be observed in Table 5.7 when RoCoF related constraints are applied on period 19, which indicates extra cost is introduced due to improvement of generator flexibility.

Table 5.7 Comparison of Different Models' Costs [$]

Model	Total	Operational	Startup	Reserve
T-SCUC	442,978	336,846	22,657	83,475
ERC-SCUC	456,045	369,820	23,782	62,443
LRC-SCUC	456,122	369,084	23,782	62,256
DNN-RCUC	456,178	370,241	23,782	62,155
CNN-RCUC	456,294	370,630	23,782	61,882

We assume the worst-case contingency takes place in period 19, and the generator outputting the largest power is tripped. The system uniform RoCoF responses of different schedule cases are shown in Fig. 5.8. Fig. 5.8 (a) shows the uniform RoCoF response of ERC-SCUC model. With system equivalent RoCoF constraints incorporated, the highest RoCoF absolute value of ERC-SCUC is strictly

0.5 Hz/s, which satisfies the RoCoF constraint. Fig. 5.8 (b), (c) and (d) show that system uniform RoCoF doesn't violate the threshold in both LRC-SCUC and DNN-RCUC cases. It should be noted that the proposed CNN-RCUC model has the lowest RoCoF value.

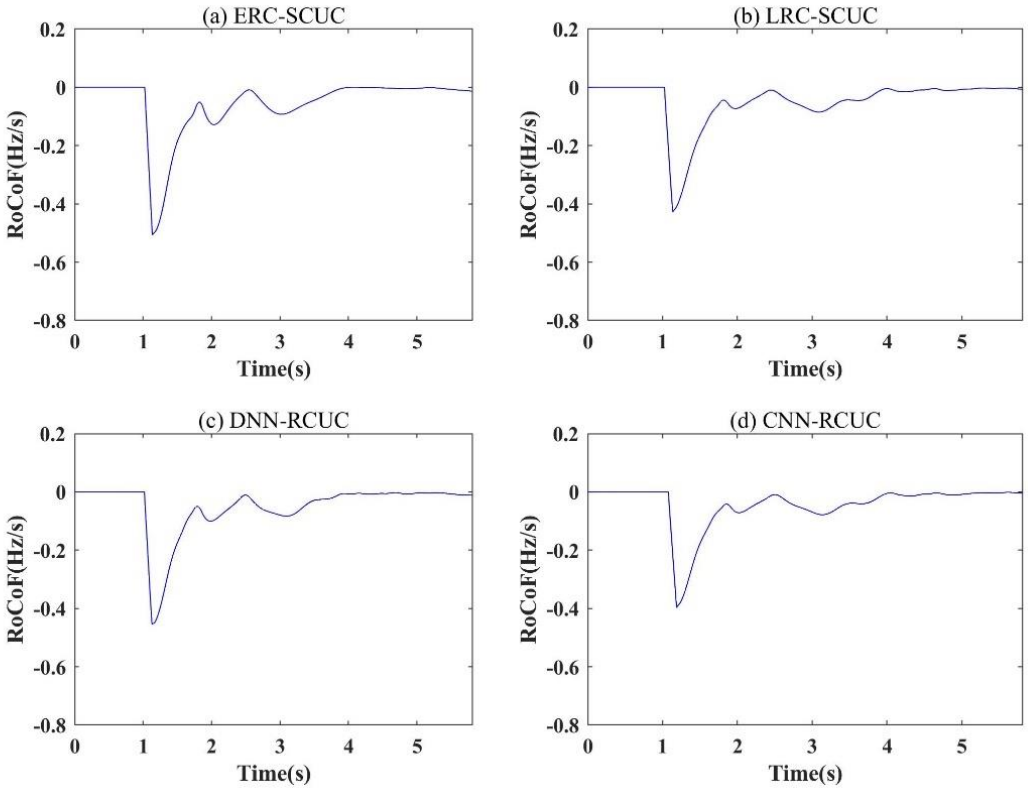


Fig. 5.8 Uniform RoCoF curves of different model following worst contingency case.

The locational RoCoF dynamics of all models are plotted in Fig. 5.9. Combining Fig. 5.8 (a) and Fig. 5.9 (a) we can observe that even though the system uniform RoCoF doesn't violate the threshold when equivalent RoCoF constraint is applied, locational RoCoF on several generator buses violate its threshold due to oscillations. The ERC-SCUC is insecure as dispatched condition cannot withstand the trip of the largest generator, cascaded generator contingency may occur under such

condition. From Fig. 5.8 (b) we can find that the highest RoCoF is mitigated for LRC-SCUC schedule, however it still violates the threshold due to approximation error introduced in the system dynamic model and PWL method. For data driven methods, Fig. 5.9 (b) and (c) show better RoCoF dynamics. As we can observe, DNN-RCUC can mitigate the highest RoCoF to 0.55Hz/s which slightly violate the threshold. The highest RoCoF if the proposed CNN-RCUC method is 0.52 Hz/s which outperforms all other UC schedules.

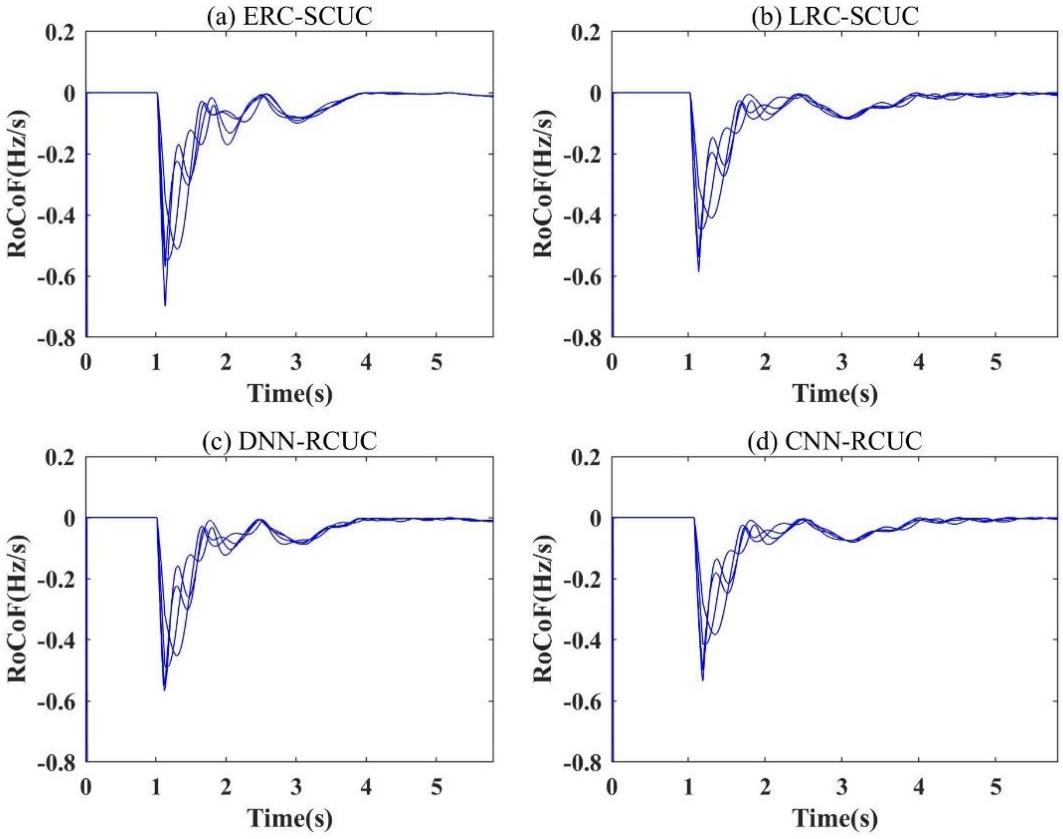


Fig. 5.9 Locational RoCoF curves of different model following worst contingency case.

5.5 Summary

This chapter proposes data driven frame works to secure system RoCoF securities under UC schedules. Initially we use model-based method to generate dataset which covers all conditions. Model-based approaches such as T-SCUC, ERC-SCUC and LRC-SCUC that enforcing system locational frequency security are proposed to efficiently generate realistic data for predictor training. DNN and CNN based RoCoF predictors are then constructed to track the highest RoCoF value of low inertia power system under worst contingency on the generated dataset.

Following this, the trained well RoCoF predictor is reformulated and incorporated into the linear SCUC model. A PWL method is applied to linearize the ReLU functions on each neuron of the RoCoF predictor. Along with this, a DNN-RCUC and CNN-RCUC models are proposed to ensure system post contingency frequency stability. Several SCUC models namely T-SCUC, ERC-SCUC, LRC-SCUC, DNN-RCUC and CNN-RCUC are examined on IEEE 24-bus system. Simulation results on Pyomo show that the proposed DNN-RCUC and CNN-RCUC can reduce the reserve cost while also resulting operational cost increasement.

Verifications on PSS/E show that model-based approaches can handle system uniform frequency stability under worst contingency, however they all fail to secure locational RoCoF stability due to approximation error. And the proposed DNN-RCUC and CNN-RCUC models can significantly mitigate the highest system locational RoCoF value, securing system frequency stability without conservativeness.

6. ACTIVE SPARSE NEURAL NETWORK BASED FCUC

The Security-Constrained Unit Commitment (SCUC) problem is a fundamental optimization task in power system operation. Its objective is to determine the optimal commitment and dispatch of power generating units while satisfying various operational and security constraints. The SCUC problem takes into account factors such as power demand, generator availability, transmission network limitations, reserve requirements, and contingencies. By considering security constraints, SCUC ensures that the power system operates reliably under normal and abnormal conditions, mitigating the risk of blackouts and maintaining system stability. By incorporating DNNs, the SCUC problem can benefit from their ability to capture intricate relationships within input data, which traditional optimization algorithms may struggle to capture.

DNNs can effectively learn complex patterns in power system data, such as historical load profiles, generator characteristics, and frequency response. Ref. [119] proposes a DNN-RCUC which incorporates DNN-based RoCoF predictor into formulations by introducing a set of mixed-integer linear constraints. However, MILP makes the problem harder to solve typically for larger systems. Moreover, there are several security constraints and physical constraints to adhere with to ensure reliable and low-cost solutions. However, DNN can have large group of parameters, leading to large model sizes. The computational efficiency of data-driven based approaches has not been investigated thoroughly yet. Such data-driven approaches may increase computational burden due to dense matrix multiplications [120]. Such data-driven

approaches may increase computational burden due to dense matrix multiplications, this poses challenges in terms of storage, memory, and computational requirements.. On the other hand, reformulation of DNN would also introduce extra group of binary variables and further degrade the computational efficiency. Pruning allows for the removal of unnecessary parameters, resulting in a more compact model that requires fewer computational resources.

6.1 Overview of Solutions

6.1.1 Model based Approaches

The frequency of the power system is one of the most important metrics that indicate the system stability. Traditionally, the frequency of the system is treated as a single-bus representation or center of inertia (COI) representation; the total power system inertia is considered as the summation of the kinetic energy stored in all dispatched generators synchronized with the power system,

$$E_{sys} = \sum_{i=1}^{N} 2H_i S_{B_i}, \tag{6.1}$$

where S_{B_i} is the generator rated power in MVA and H_i denotes the inertia constant of the generator.

Assuming a disturbance in the electrical power, the dynamics between power and frequency can be modeled by the swing equation described in (6.2) with H denoting the normalized inertia constant and D denoting overall damping constant respectively,

$$\Delta P_m - \Delta P_L = 2H\frac{d\Delta\omega}{dt} + D\Delta\omega, \tag{6.2}$$

where ΔP_m is the total change in mechanical power and ΔP_L is the total non-frequency sensitive change, and $D\Delta\omega$ is the frequency-sensitive load change. $d\Delta\omega/dt$ is commonly known as RoCoF. During a short period of time following a disturbance, we can derive the constraints on initial RoCoF used for system uniform model,

$$f_{rcf} = \frac{\Delta P}{2HS_B}\omega_n \leq -RoCoF_{lim}. \tag{6.3}$$

However only considering system uniform metrics neglects the geographical discrepancies in locational frequency dynamics on each bus, which has imposed risks on power system stability. Dynamic model is preferred in modern power system for frequency oscillation analysis. The topological information and system parameters can be embedded into the model by using swing equation on each individual bus to describe the oscillatory behavior within the system. A network-reduced model with N generator buses can be obtained by eliminating passive load buses via Kron reduction. By focusing on the network connectivity's impact on the power system nodal dynamics, the phase angle θ of generator buses can be expressed by the following dynamic equation,

$$M\ddot{\theta} + D\dot{\theta} = P_{in} - L\theta, \tag{6.4}$$

Following a loss of generator on bus b, the steady state evolves under (6.4) can be derived. And the RoCoF value at bus i can then be expressed as,

$$f_{rcf,i}(t_0) = \frac{\Delta P e^{-\frac{\gamma t}{2}}}{2\pi m} \sum_{\alpha=1}^{N_g} \frac{\beta_{\alpha i}\beta_{\alpha b}}{\sqrt{\frac{\lambda_\alpha}{m} - \frac{\gamma^2}{4}}\Delta t} \begin{bmatrix} e^{-\frac{\gamma \Delta t}{2}} \sin\left(\sqrt{\frac{\lambda_\alpha}{m} - \frac{\gamma^2}{4}}(t_0 + \Delta t)\right) \\ -\sin\left(\sqrt{\frac{\lambda_\alpha}{m} - \frac{\gamma^2}{4}} t_0\right) \end{bmatrix}, \quad (6.5)$$

where λ_α is the eigenvalue of matrix L, $\beta_{\alpha i}$ is the eigen vector value; and m denotes average inertia distribution on generator buses; and bus b is where disturbance occurs; Δt is the frequency monitoring window, and t_0 is the measuring time. N_g denotes the set of generator buses in the reduced model. The ratio of damping coefficient to inertia coefficient $\gamma = d_i/m_i$ is assumed as a constant [100]. We can then derive the dynamic RoCoF related constraints for different time period as,

$$f_{rcf,i}(t_1) \leq -\text{RoCoF}_{lim}, \forall\, i \in N_G. \quad (6.6)$$

$$f_{rcf,i}(t_2) \leq -\text{RoCoF}_{lim}, \forall\, i \in N_G. \quad (6.7)$$

6.1.2 Frequency Metrics Constrained Data Driven Approach

Given a system with N generators, the objective goal of the ordinary UC model is to minimize the total operating cost subject to various system operational constraints**,**

$$\begin{aligned} &min. \quad \mathcal{C}(s_t, u_t), \\ &s.t. \quad \mathcal{F}(s_t, u_t, d_t, r_t) = 0, \mathcal{G}(s_t, u_t, d_t, r_t) \leq 0, \forall t, \end{aligned} \quad (6.8)$$

where $\mathcal{F}$ and $\mathcal{G}$ are the equality and inequality constraints respectively; $\boldsymbol{s}_t$ denotes the system states, and $u_t$is the generation dispatch at period t. d_t and r_t denote the load profile and renewable forecast, respectively. Assume a potential disturbance $\boldsymbol{\varpi}_t$ occurs in period t. With system nominal frequency $f_n = 60$ Hz being the base

frequency, the model based RoCoF constraints (6.6) and (6.7) could be derived and then added into the ordinary UC formulation to secure the system frequency stability.

Assumptions made during the system model analysis may introduce approximation error, subsequently lead to unsecure results or conservative results. The idea of data driven approach is to replace model-based constraints by DNN formulations,

$$\hat{h}^f(s_t, u_t, r_t, \varpi_t) \leq \varepsilon, \tag{6.9}$$

where $\hat{h}^f$ is the nonlinear DNN-based frequency metrics predictor, including system wide maximal frequency deviation and maximal locational RoCoF; ε is the vector of predefined threshold.

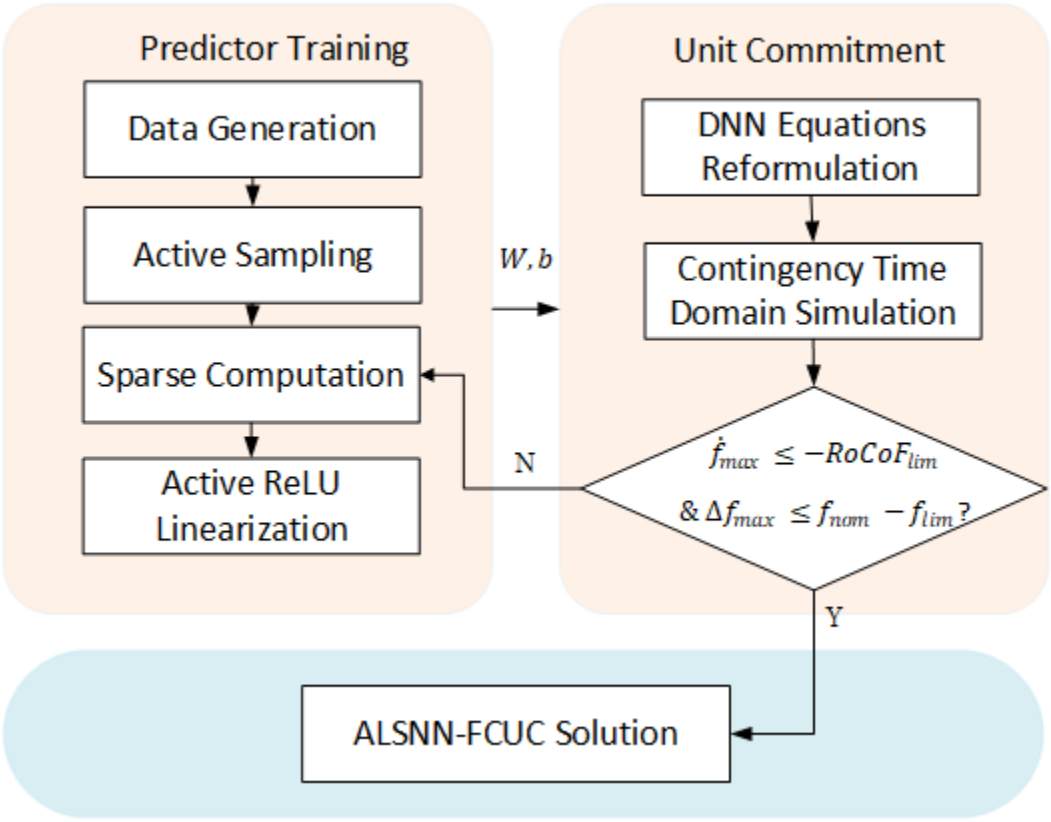


Fig. 6.1 Overview of the proposed approach.

The loss of generation not only results in the largest power outage level, but also leads to the reduction in system inertia, which further leads to the highest RoCoF value and frequency deviation; thus, G-1 event is considered as the worst contingency in this study. It should be noted that the output vector of the DNN-based predictor includes multiple elements which could be generalized for other constraints without including more parameters. The overview of the working pipeline is shown in Fig. 6.1, in which Δf_{max} is the maximal frequency deviation, and $\dot{f}_{max}$ is the highest RoCoF value.

6.2 Model based Dataset Generation

Wide-range space of all power injections is utilized in [111] to ensure reliability under vast ranges of operating conditions for a small group of generators. However, considering a power system consists of a large group of generators, the dimension of x_t would increase accordingly, which would further cause exponentially increase in the dimension of injections state space for practical system condition [121]. Such wide-range injections may lead to divergence during the simulation initialization process and the system would also be subject to post-contingency stability issues.

Unlike randomly data generation considering wide-range space of dispatching, a model-based systematic data generation approach is proposed to generate reasonable and representative data that will be used to train RoCoF Predictors with much less computational burden and compromise of efficiency. Traditional SCUC (T-SCUC) models and RoCoF constrained SCUC models are implemented in

this process to generate training samples over various load and RES scenarios. Objective function of three models is to minimize the total system cost consisting of variable fuel costs, no-load costs, start-up costs, and reserve costs. RoCoF related constraints based on equivalent model and dynamic model are added into the formulation of system equivalent model based RoCoF constrained SCUC (ERC-SCUC) and location based RoCoF constrained SCUC (LRC-SCUC) respectively. The constraints on maximal frequency deviation RoCoF for locational frequency dynamics are nonlinear. In order to incorporate the nonlinear RoCoF-related constraints into the MILP model, a piecewise linear programming method is introduced, and the detail of all models is presented in [105].

6.3 Deep Learning-based Frequency Constraints

6.3.1 Power System Feature Definition

The highest locational frequency deviation and highest system locational RoCoF are considered as functions with respect to the contingency level, contingency location, system states, unit dispatch. Since both the magnitude and location of the contingency will have impact on the locational frequency deviation and locational inertial response, the generator status and dispatching values are encoded in to feature vectors [122]. In order to track the system frequency related metrics at period *t*, here we define the feature vector used for frequency related metrics tracking. The generator status vector is first defined as follow,

$$u_t = \left[u_{1,t}, u_{2,t}, \cdots, u_{NG,t}\right]. \tag{6.10}$$

The disturbance feature vector at period t is introduced for frequency related metrics tracking,

$$P_t^{con} = \max_{g \in G}\left(P_{1,t}, P_{2,t}, \cdots, P_{NG,t}\right). \tag{6.11}$$

The location of the disturbance is represented by the index of the generator producing maximum power,

$$g_t^{con} = arg\max_{g \in G}\left(P_{1,t}, P_{2,t}, \cdots, P_{NG,t}\right). \tag{6.12}$$

The information of magnitude and location is then encoded into the disturbance feature vector as [111],

$$\varpi_t^G = \left[0, \cdots, 0, \underbrace{P_t^{con}}_{g_t^{con} th\ element}, 0, \cdots, 0\right]. \tag{6.13}$$

The active power dispatching of all synchronous generator at period t is then expressed as feature vector,

$$P_t = \left[P_{1,t}, \cdots, P_{2,t}, \cdots, P_{NG,t}\right]. \tag{6.14}$$

By combing the feature vectors mentioned above, the overall feature vector $\boldsymbol{x}_t$ at period t can be then defined as follows,

$$\boldsymbol{x}_t = \left[u_{1,t}, \cdots, u_{NG,t},\ \varepsilon_{1,t}, \cdots,\ \varepsilon_{NG,t}, P_{1,t}, \cdots, P_{NG,t}\right]. \tag{6.15}$$

6.3.2 DNN-based Frequency Metrics Predictor

Regarding frequency related constraints, both model-based approach and data-driven approach share the same variables which are determined by the generator status and disturbance information. Thus, A DNN-based frequency metrics predictor is proposed, which combines the functions of locational frequency deviation and highest

locational RoCoF in frequency-related constraints together into a nonlinear DNN with multiple outputs. The combination of these functions could significantly decrease the number of DNN parameters thus reducing additional variables introduced into the FCUC formulation.

In the aftermath of a large $G-1$ contingency, the frequency metrics predictor for h can be expressed as,

$$\begin{bmatrix}\hat{f}_{dev}\\ \hat{f}_{rcf}\end{bmatrix} = \hat{h}(x_t, W^f, b^f), \tag{6.16}$$

where W^f and b^f denote the well-trained neural network parameters for frequency metrics predictor. Let $\hat{f}_{dev}$ and $\hat{f}_{rcf}$ denote the predicted frequency deviation and highest RoCoF values respectively for case $\boldsymbol{x}_t$. Considering a DNN with N_L hidden layers, where ReLU is used as activation function for each hidden layer and the output layer is linear activation function. The forward propagation of each layer is then expressed as follows,

$$z_1 = x_t W_1 + b_1, \tag{6.17}$$

$$\hat{z}_q = z_{q-1} W_q + b_q, \tag{6.18}$$

$$z_q = \max(\hat{z}_q, 0), \tag{6.19}$$

$$\hat{f}_{dev} = z_{N_L} W_{N_L+1}^{dev} + b_{N_L+1}^{dev}, \tag{6.20}$$

$$\hat{f}_{rcf} = z_{N_L} W_{N_L+1}^{rcf} + b_{N_L+1}^{rcf}, \tag{6.21}$$

where W_q and vector b_q for $q \in N_L$ represent the set of weight and bias across all hidden layers; $W_{N_L}^{dev}$ and $b_{N_L}^{dev}$ represent the set of weight and bias of the output layer

for maximal frequency deviation prediction, and $W_{N_L}^{rcf}$ and $b_{N_L}^{rcf}$ represent the output layer for system highest locational RoCoF prediction. The training process is to minimize the total mean squared error between the predicted output and the labeled outputs of all training samples as follows,

$$\min_{\Phi} \frac{1}{N_S} \sum_{s=1}^{N_S} \left(\Delta f_{max} - \hat{f}_{dev}\right)^2 + \left(\dot{f}_{max} - \hat{f}_{rcf}\right)^2, \tag{6.22}$$

where $\Phi = \left\{W_q, b_q, W_{N_L+1}^{dev}, W_{N_L+1}^{rcf}, b_{N_L+1}^{dev}, b_{N_L+1}^{rcf}\right\}$ represents the set of optimization variables.

Although the aforementioned reformulation of DNN introduces no approximation error, such linearization progress would introduce dense parameter matrix and additional binary variables into MILP model. Thus, the computational burden for the DNN-MILP would increase especially for the condition where multiple scheduling periods are limited by DNN constraints. To handle that problem, sparse computation and active ReLU linearization are proposed to obtain optimal efficiency.

6.3.3 Sparse Computation

Conversion of DNN network into MILP has introduced group parameters. However, not all of these parameters are required to achieve high performance of the predictor. Instead of connecting every pair of neurons in each layer, we prune the connections and redundant parameters from networks to decrease the density of weights matrix without affecting performance of the frequency metrics predictor [123]. The example of training algorithm for sparse computation is shown in Fig. 6.2.

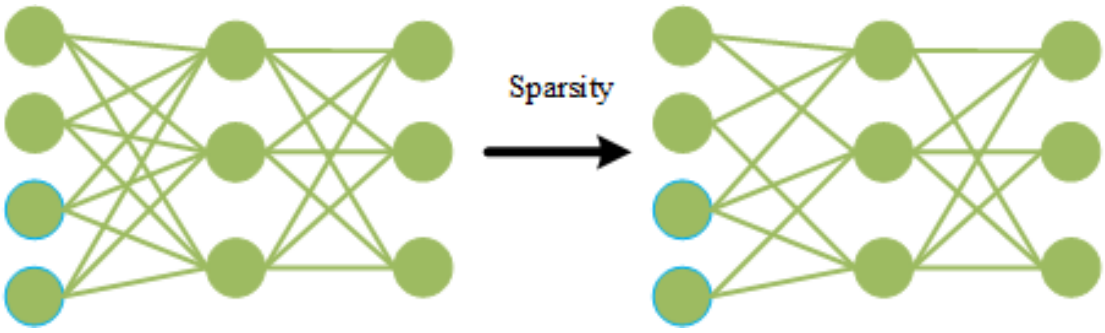


Fig. 6.2 Example of sparse fully connected neural network.

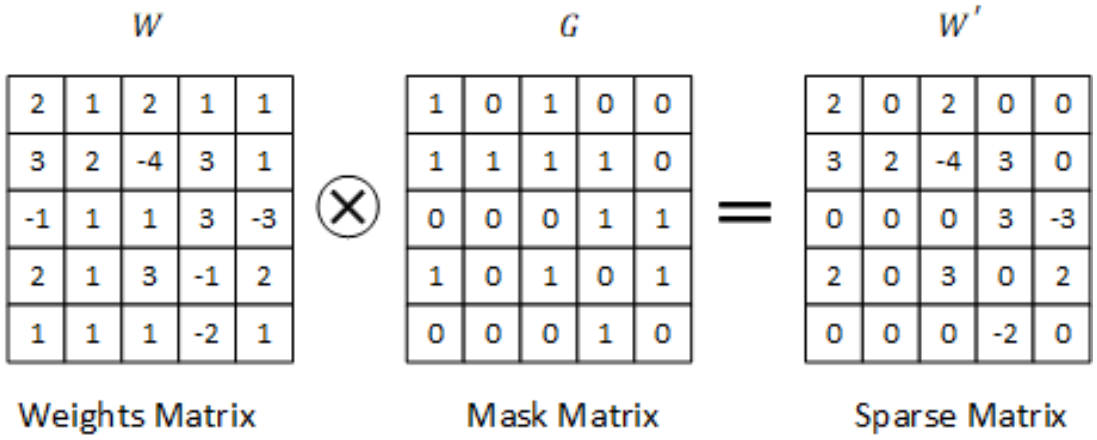


Fig. 6.3 Example of sparse computation process.

In this work, the network's connections are pruned during the training process by applying a mask [124]. First, we initialize the network and pre-trained the parameters of the frequency metrics predictor. For each layer chosen to be pruned, a binary mask variable is added which is of the same size and shape as the layer's weight matrix and determines which of the weights participate in the forward execution of the graph as shown in Fig. 6.3. The weights in each layer are sorted by their absolute values and mask the smallest magnitude weights to zero. The sparsity of the parameter matrix is increased from an initial sparsity value s_0 to a final sparsity value s_{final} over a span of μ pruning steps with pruning frequency Δe,

$$s_e = s_{final} + (s_0 - s_{final})\left(1 - \frac{e - e_0}{\mu \Delta e}\right)^3, \tag{6.28}$$

for $e \in \{e_0, e_0 + \Delta e, \ldots, e_0 + \mu \Delta e\}$.

Gradually increasing the sparsity of the network allows the network training steps to recover from any pruning-induced loss in accuracy. Similar binary masks are applied to the back-propagated gradients, and the weights that were masked in the forward execution are not updated in the back-propagation step, the overall strategy is introduced as follows,

Algorithm 1 Sparse Neural Network Training

Input: Training Dataset $\Theta = \{(x_1, y_1), ..., (x_n, y_n)\}$, $\theta = \{W, b\}$, Mask generator $Sp(\cdot)$, Final Sparsity s_{final}, Initial epoch e_0, Pruning frequency Δe, Pruning times Q, Total Training Epochs E, Batch size B

Output: optimal sparse W, b

$W \leftarrow W_0$ ▷ Initialize W with Pretrained W_0
$b \leftarrow b_0$ ▷ Initialize b with Pretrained b_0
for $e = 1, 2, ..., E$ **do**
  if $e \neq e_0 + \mu \Delta e (\mu \in Q)$ **then**
    $h_1, ..., h_B \leftarrow SNN(\Theta_B, \theta)$
    $\Delta_\theta \leftarrow BP(\Theta_B, h_1, ..., h_B, \theta)$
    $\theta \leftarrow LearningRule(\Delta_\theta, \theta)$
  else
    $W \leftarrow W \odot Sp(s_0, s_{final}, \mu, e)$
    $h_1, ..., h_B \leftarrow SNN(\Theta_B, \theta)$
    $\Delta_\theta \leftarrow BP(\Theta_B, h_1, ..., h_B, \theta)$
    $\theta \leftarrow LearningRule(\Delta_\theta, \theta, Sp(s_0, s_{final}, \mu, e))$
  end if
end for
return θ

6.3.4 Active Sampling

Even though the sparse computation and ReLU linearization have been proved to be able to improve the computational efficiency of the model, there are other important aspects that we should consider. However, with sparse computation introduced, the compressed network of predictor tuned for the data near the critical threshold using region-of-interest active sampling method may fail to approximate the output of the original network. Subsequently, the related frequency constraints may not bind during the unit commitment process. Thus, it is reasonable to consider the tradeoff between the sparsity rate and the prediction error.

Model-based data generation is utilized to generate a group of samples which are labeled with maximal frequency deviation and the highest RoCoF following the worst contingency. Labeling the RoCoF value of adequate samples to achieve desired distributions requires significant computation resources and time, which is known as the labeling bottleneck. In other words, how to sample the cases to be labeled is essential. In order to generate and label desired samples, an active sampling (AS) method is proposed to actively select unlabeled samples.

A frequency security discriminator is first trained based on the small group of samples $\mathcal{F}$ with highest RoCoF and frequency deviation labeled, the average value in 0.1s measuring interval is used. Security label f_{sec} is then applied to the training dataset: if the Δf_{max} and $\dot{f}_{max}$ of a sample are all within the limit, f_{sec} is labeled as 1. Otherwise, if one of the metrics violate the threshold, f_{sec} is labeled as 0. And then we perform the active sampling strategy on unlabeled dataset $\mathcal{U}$ using the discriminator.

The unlabeled samples whose posterior probability provided by the discriminator are then selected as training dataset for RoCoF and frequency deviation predictor. The general sampling strategy uses entropy as measure [125],

$$x_{sec} = \underset{\mathcal{U}}{\operatorname{argmax}} \left(-\sum_{\mathcal{U}}^{1} p(y_i|\mathcal{U}) \log p(y_i|\mathcal{U}) \right), \tag{6.29}$$

$$x_{insec} = \underset{\mathcal{U}}{\operatorname{argmin}} \left(-\sum_{\mathcal{U}}^{1} p(y_i|\mathcal{U}) \log p(y_i|\mathcal{U}) \right), \tag{6.30}$$

where y_i ranges over all possible labelings, and $p(y_i|\mathcal{U})$ is the predicted posterior probability of class membership y_i. This sorting process finds the largest and lowest from the finite set of numerical values. The selected x_{sec} is considered as sample distribute close the threshold which can improve the accuracy of predictor, while x_{insec} denotes samples which help the output of the tuned frequency metrics predictor approximate to the original. In the other word, they help keep the constraints binding after the sparse computation. The selected sample x is then added into Θ and labeled with highest RoCoF and frequency deviation for frequency metrics predictor training.

6.3.5 Active ReLU Linearization

Activation unit has been set on the neurons of hidden layer to realize the nonlinearization of a neural network; activation function, such as ReLU, is applied to the result of linear combination of values from neuron nodes [126]. The incorporation of ReLU function in a neural network into MILP without no approximation would introduce multiple extra binary variables. Subsequently, increasing the computational

burden of the MILP model and leading to poor efficiency. Therefore, an active linearization of ReLU function has been introduced to reduce the DNN size without too much degradation of prediction accuracy [127].

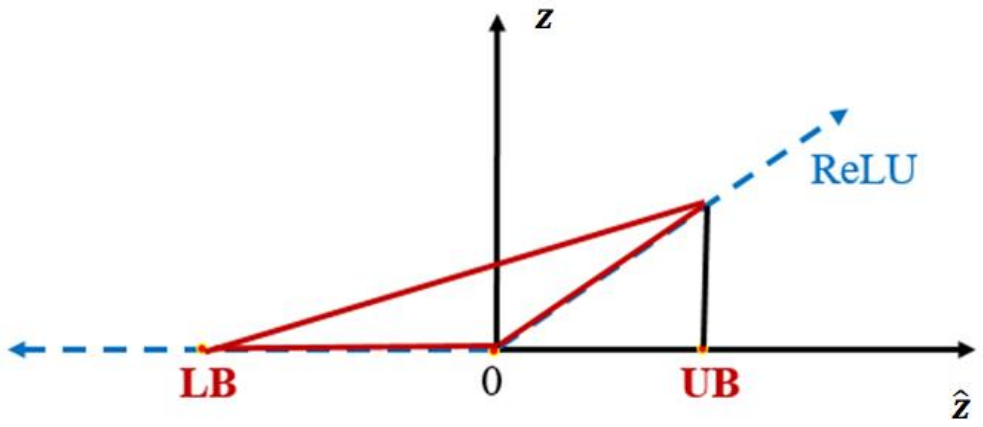


Fig. 6.4 The linear approximation of ReLU activation function.

Approximation of ReLU function is shown in Fig. 6.4. The weighted sum of input signals to the node is denoted as variable $\hat{z}$, and the output of the node is denoted using the variable z. Given the upper and lower bounds $[LB, UB]$ of $\hat{z}$, the relationship of z and $\hat{z}$ can then be approximated by a set of constraints $z \geq 0$, $z \geq \hat{z}$, and $z \leq \frac{UB \cdot (\hat{z} - LB)}{UB - LB}$. These constraints are all linear equations with constant *UB* and *LB*.

Although ReLU linearization can be applied to all the neurons that would convert a DNN-FCUC model into ReLU linearized neural network based FCUC (RLNN-FCUC), this may lead to large approximation error and low prediction accuracy; and the associated frequency constraints may no longer be binding. In this work, we introduce an active selecting method to improve the computational efficiency of the DNN-FCUC model while maintaining the bindingness of derived

constraints. The actively sampled dataset is first fed into the well-trained frequency metrics predictor, a nodal positivity index $\varepsilon_{q[l]}$ is proposed to estimate the percentage of positive pre-activated values of neuron node l in q-th layer,

$$\varepsilon_{q[l]} = \frac{1}{N_S}\left(\sum_{N_S}\hat{z}_{q[l],s} - \sum_{N_S}\left|\hat{z}_{q[l],s} - \frac{1}{N_S}\sum_{N_S}\hat{z}_{q[l],s}\right|\right) \geq \gamma, \tag{6.31}$$

where γ is the threshold set to select nodes suitable for ReLU linearization with less approximation error. For the prediction of a given sample s, equations (6.23)-(6.26) for ReLU function in selected neurons can be replaced by (6.32)-(6.34) as follows,

$$z_{q[l],s} \geq \hat{z}_{q[l],s}, \forall q, \forall l, \forall s, \tag{6.32}$$

$$z_{q[l],s} \leq \frac{UB_{q[l]} \cdot (\hat{z}_{q[l],s} - LB_{q[l]})}{UB_{q[l]} - LB_{q[l]}}, \forall q, \forall l, \forall s, \tag{6.33}$$

$$z_{q[l],s} \geq 0, \forall q, \forall l, \forall s. \tag{6.34}$$

6.4 Mixed-Integer Formulation of ALSNN-FCUC Model

In this section, the proposed ALSNN-FCUC considering frequency related constraints is formulated. The objective of the modified ALSNN-FCUC model is to minimize total operating cost subject to various system operational constraints and guarantee system frequency stability. The formulation is shown below,

$$\min_{\Phi} \sum_{g \in G}\sum_{t \in T}(c_g P_{g,t} + c_g^{NL} u_{g,t} + c_g^{SU} v_{g,t} + c_g^{RE} r_{g,t}), \tag{6.35}$$

$$\sum_{g \in G} P_{g,t} + \sum_{k \in K(n-)} P_{g,t} - \sum_{k \in K(n+)} P_{g,t} - D_{n,t} + E_{n,t} = 0, \quad \forall n, t, \tag{6.36}$$

$$P_{k,t} - b_k(\theta_{n,t} - \theta_{m,t}) = 0, \qquad \forall k,t \tag{6.37}$$

$$-P_k^{max} \le P_{k,t} \le P_k^{max}, \qquad \forall k,t \tag{6.38}$$

$$P_g^{min} u_{g,t} \le P_{g,t}, \qquad \forall g,t, \tag{6.39}$$

$$P_{g,t} + r_{g,t} \le u_{g,t} P_g^{max}, \qquad \forall g,t, \tag{6.40}$$

$$0 \le r_{g,t} \le R_g^{re} u_{g,t}, \qquad \forall g,t, \tag{6.41}$$

$$\sum_{j \in G} r_{j,t} \ge P_{g,t} + r_{g,t}, \qquad \forall g,t, \tag{6.42}$$

$$P_{g,t} - P_{g,t-1} \le R_g^{hr}, \qquad \forall g,t, \tag{6.43}$$

$$P_{g,t-1} - P_{g,t} \le R_g^{hr}, \qquad \forall g,t, \tag{6.44}$$

$$v_{g,t} \ge u_{g,t} - u_{g,t-1}, \quad \forall g,t, \tag{6.45}$$

$$v_{g,t+1} \le 1 - u_{g,t}, \qquad \forall g, t \le nT - 1, \tag{6.46}$$

$$v_{g,t} \le u_{g,t}, \qquad \forall g,t, \tag{6.47}$$

$$\sum_{s=t-UT_g}^{t} v_{g,s} \le u_{g,t}, \forall g, t \ge UT_g, \tag{6.48}$$

$$\sum_{s=t-UT_g}^{t+DT_g} v_{g,s} \le 1 - u_{g,t}, \forall g, t \ge nT - DT_{g,} \tag{6.49}$$

$$u_{g,t}, v_{g,t} \in \{0,1\}, \qquad \forall g,t. \tag{6.50}$$

Equation (6.35) is the objective function, and the basic constraints include (6.36)-(6.50). Equation (6.36) enforces the nodal power balance. Network power flows are calculated in 6.37) and are restricted by the transmission capacity as shown in (6.38). The scheduled energy production and generation reserves are bounded by unit generation capacity and ramping rate (6.39)-(6.44). As defined in (6.42), the reserve requirements ensure the reserve is sufficient to cover any loss of a single generator.

The start-up status and on/off status of conventional units are defined as binary variables (6.45)-(6.47). Minimum-down time before a generator can be started-up and the minimum-up time before a generator can be shutdown are depicted in (6.48) and (6.49), respectively. Indicating variables for generator start-up and commitment status are binary and are represented by (6.50).

Since x_t contains the max operator, and disturbance vector cannot be directly used in the encoding formulation. Supplement variables $\mu_{g,t}$ and $\rho_{g,t}$ are used to indicate the largest generator output for period t. The reformulations are expressed as follows,

$$P_{g,t}^{con} - P_{g,t} \leq A\left(1 - \mu_{g,t}\right), \qquad \forall g, t, \tag{6.51}$$

$$\sum_{g \in G} \mu_{g,t} = 1, \qquad \forall t, \tag{6.52}$$

$$\mu_{g,t} \in \{0,1\}, \qquad \forall g, t, \tag{6.53}$$

$$\rho_{g,t} - P_{g,t} \geq -A\left(1 - \mu_{g,t}\right), \qquad \forall g, t, \tag{6.54}$$

$$\rho_{g,t} - P_{g,t} \leq A\left(1 - \mu_{g,t}\right), \qquad \forall g, t, \tag{6.55}$$

$$0 \leq \rho_{g,t} \leq A\mu_{g,t}, \qquad \forall g, t, \tag{6.56}$$

where A is a big number. The input feature vector is then reformulated as follows,

$$x_t = \left[u_{1,t}, \cdots, u_{N_G,t},\ \rho_{g,t}, \cdots,\ \rho_{N_G,t}, P_{1,s}, \cdots, P_{N_G,s}\right]. \tag{6.57}$$

Then we introduce the formulation of proposed frequency metrics predictor. As discussed before, the neuron of each layer with positivity index $\varepsilon_{q[l]}$ larger than γ is selected out and added into set $\mathcal{H}$. The reformulations of ReLU activation on the selected neurons are expressed as follows,

$$z_1 = \boldsymbol{x}_t W_1 + b_1, \forall t, \tag{6.58}$$

$$z_{q[l],t} \geq \hat{z}_{q[l],t}, \forall q, \forall l \in \mathcal{H}, \forall t, \tag{6.59}$$

$$z_{q[l],t} \leq \frac{UB_{q[l]} \cdot \left(\hat{z}_{q[l],t} - LB_{q[l]}\right)}{UB_{q[l]} - LB_{q[l]}}, \forall q, \forall l \in \mathcal{H}, \forall t. \tag{6.60}$$

The ReLU function on the rest of neuron is reformulated with no approximation as follows. (6.66) and (6.67) constrain the maximal frequency deviation and maximal RoCoF,

$$z_{q[l],t} \leq \hat{z}_{q[l],t} - A(1 - a_{q[l],t}), \forall q, l \in \bar{\mathcal{H}}, t, \tag{6.61}$$

$$z_{q[l],t} \geq \hat{z}_{q[l],t}, \forall q, \forall l \in \bar{\mathcal{H}}, \forall t, \tag{6.62}$$

$$z_{q[l],t} \leq A a_{q[l],t}, \forall q, \forall l \in \bar{\mathcal{H}}, \forall t, \tag{6.63}$$

$$z_{q[l],t} \geq 0, \forall q, \forall l, \forall t, \tag{6.64}$$

$$a_{q[l],t} \in \{0, 1\}, \forall q, \forall l, \forall t, \tag{6.65}$$

$$z_{N_L,t} W_{N_L+1}^{dev} + b_{N_L+1}^{dev} \leq f_{nom} - f_{lim}, \tag{6.66}$$

$$z_{N_L,t} W_{N_L+1}^{rcf} + b_{N_L+1}^{rcf} \leq -RoCoF_{lim}. \tag{6.67}$$

6.5 Results Analysis

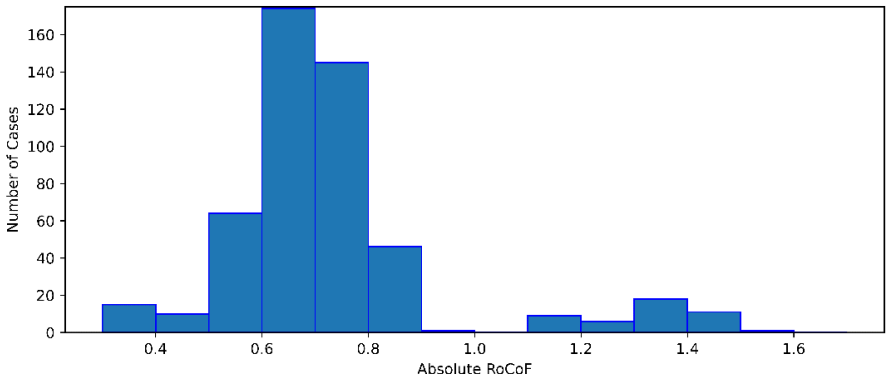
Case study on IEEE 24-bus system [128] is provided to demonstrate the effectiveness of the proposed methods. This test system contains 24 buses, 33 generators and 38 lines, which also has wind power as renewable resources. The mathematical model-based data generation is operated in Python using Pyomo [129]**Error! Reference source not found.**. The PSS/E software is used for time domain simulation and labeling process [130]. We use full-scale models for the

dynamic simulation during the labeling process: GENROU and GENTPJ for the synchronous machine; IEEEX1 for the excitation system; IEESGO for the turbine-governor; PSS2A for the power system stabilizer. Standard WTG and corresponding control modules are employed. The FCUC is performed using Pyomo and Gurobi on a window laptop with Intel(R) Core(TM) i7 2.60GHz CPU and 16 GB RAM.

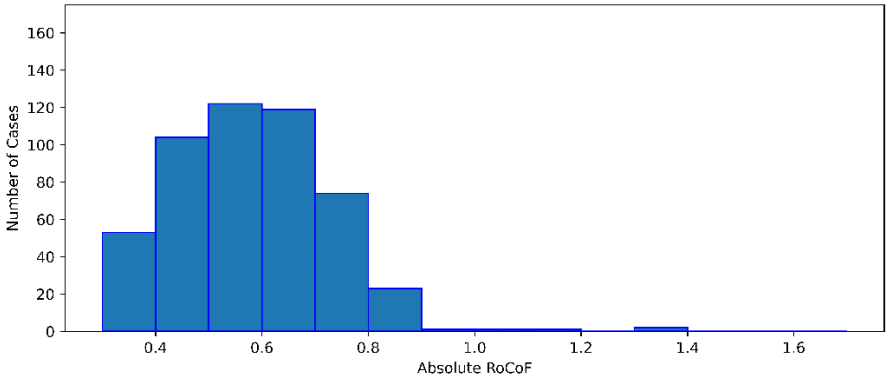
6.5.1 Predictor Training

We first generate 3000 samples as $\mathcal{L}$ for predictor training. Each case is labeled with security status, 0 for insecure and 1 for secure based on post contingency conditions. To ensure the practicality of the dataset and the generality of the trained model, load profile and RES profile are sampled based on Gaussian distribution while the deviation of means value ranges from [-20%, 20%] of the based value. The optimality gap of the solver is set to 0.1%. We assume SGs have adequate reactive power capacity, and WTGs are controlled with a unity power factor.

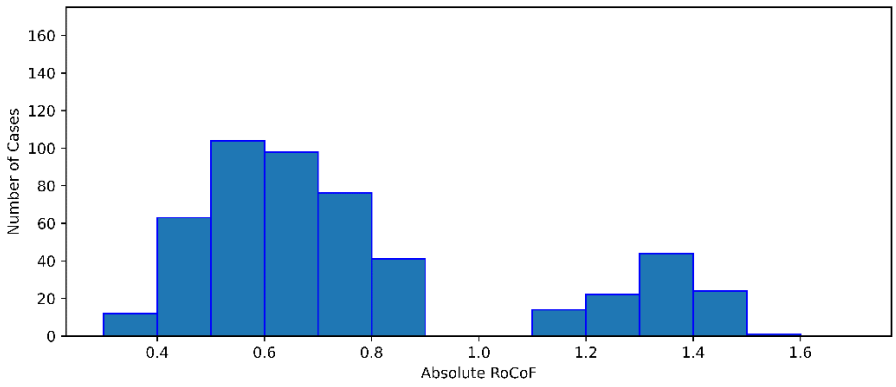
Actively sampling approach is applied to select desired samples from unlabeled samples. Two other datasets are created as benchmark sets: (1) a random dataset is generated using randomly selected (RS) method; (2) a dataset based on uncertainty sampling (US) strategy**Error! Reference source not found.** which queries the instance whose posterior probability of being positive is nearest 0.5.



(a) Randomly selected dataset.



(b) Uncertainty sampling dataset.



(c) Active sampling dataset.

Fig. 6.5 RoCoF value distributions for three different datasets.

For each strategy, we selected 500 samples from the generated training dataset. The distributions of RoCoF values of the sampled cases are depicted in Fig. 6.5. The interval size of the distribution bin is set 0.1, and regions closest to the threshold are [0.4, 0.5] and [0.5, 0.6]. With randomly selected method, most of the cases distribute within [0.6, 0.7] and [0.7, 0.8], while only 15% of the cases are within the range [0.4, 0.6] closest to the threshold. With uncertainty sampling applied, nearly 40% of the cases the selected cases centered around the threshold while cases with RoCoF values larger than 1.0 Hz/s are filtered out. When we use the proposed active learning method, the samples closest to the threshold increase; the number of samples whose RoCoF vales far away from the threshold increases as desired.

Table 6.1 Prediction Accuracy with 1% Tolerance [%]

Metrics	$\dot{f}_{max}$	Δf_{max}
DNN	74.84	60.05
SNN	52.29	59.11

We first compare the prediction accuracy of proposed DNN-based predictor with SNN-based predictor on RS Dataset. The total layer of the DNN-based predictor is four, and the number of neurons is set 10 for each hidden layer. ReLU is used as the activation function. Results in Table 6.1 show that DNN-based predictor has higher prediction accuracy in both RoCoF prediction and frequency deviation.

The proposed frequency metrics predictor with the same DNN structure is then tested on different dataset. As results shown in Table 6.2 and Table 6.3, RoCoF prediction accuracies of all cases are above 93.27% with an error tolerance of 5%,

while the frequency deviation validation accuracies are relatively lower. Predictor trained by uncertainty sampling dataset has the highest prediction accuracy on both RoCoF and frequency deviation values as expected, which gives more accurate frequency related constraints. While the validation accuracy of the proposed active learning dataset is only slightly lower than the uncertainty sampling dataset-based predictor.

Table 6.2 RoCoF Validation Accuracy [%]

Tolerance	10%	9%	8%	7%	6%	5%
RS Dataset	98.39	97.92	97.27	96.05	94.66	93.27
US Dataset	99.26	98.26	97.93	97.23	96.57	96.29
AS Dataset	98.44	98.07	97.60	97.27	96.28	95.11

Table 6.3 Frequency Deviation Validation Accuracy [%]

Tolerance	10%	9%	8%	7%	6%	5%
RS Dataset	97.96	97.44	96.57	93.71	89.50	84.76
US Dataset	99.01	98.73	97.79	96.62	93.36	91.36
AS Dataset	98.88	98.63	97.17	95.18	91.16	88.78

Full-range out-of-sample dataset are used for sparse neural network validation. Sparsity of the neural network is increased from 0% to 90% with interval at 10%. The results of RoCoF prediction accuracy with the tolerance of 5% are depicted in Fig. 6.6. As we can see in both cases, validation accuracy is relatively high at the beginning of the test. When sparsity reaches 60%, the RoCoF prediction accuracy of predictor trained by uncertainty sampling dataset drops significantly, implying low robustness.

For predictor based on randomly selected dataset, the accuracy drops to 80.04% at the sparsity of 50% and gets totally lost at the sparsity of 0.7. Meanwhile, the predictor trained by active sampling dataset significantly outperforms the others at the sparsity of 80%, implying accurate prediction with much less computational burden.

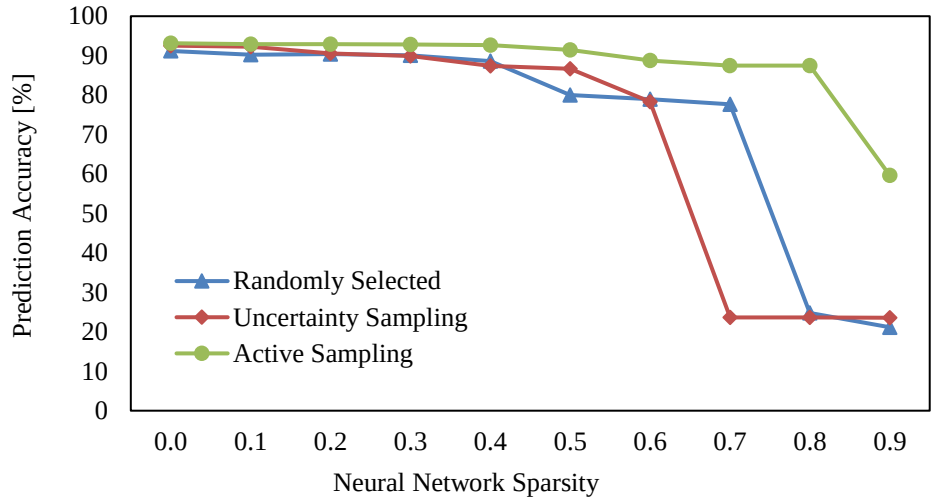


Fig. 6.6 RoCoF prediction accuracy with different NN sparsity.

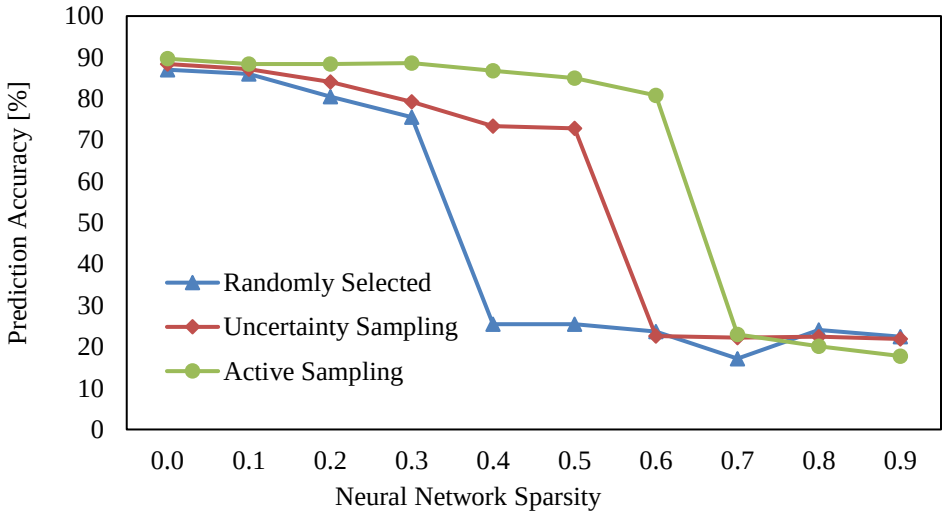


Fig. 6.7 Frequency deviation prediction accuracy with different NN sparsity.

Results in Fig. 6.7 show that the predictor trained by active sampling dataset has higher robustness against sparsity than two other cases. However, comparing to the RoCoF prediction accuracy, frequency deviation prediction accuracy is more sensitive to the change in network sparsity. Although having superiority over two other predictors, the prediction accuracy of predictor trained by active sampling dataset is lost at the sparsity of 0.7.

6.5.2 Simulation Results

The total scheduling horizon is 24 hours, and hours 9-12 are selected to be the time instance where frequency related constraints are applied to secure system stability against generator contingency considering high penetration level of intermittent wind generation and peak hour impact. The test case has a demand ranging from 1,348 MW to a peak of 1,853MW. Regarding post-contingency frequency limits, the maximal frequency deviation should not be larger than 0.5 Hz, and the RoCoF must be higher than -0.5Hz/s to avoid the tripping of RoCoF-sensitive protection relays. The optimality gap is set to 0.1%.

We first investigate the impact of frequency metrics predictor sparsity on efficiency of sparse neural network based FCUC (SNN-FCUC) without ReLU linearization. The heatmap of computational time is shown in Fig. 6.8, and 100% sparsity indicates no frequency related constraints. Computational time less than 5 seconds indicating non-binding of the constraints. As we can see, the computational times of RS and US based predictors drops significantly when sparsity reaches 70%, since frequency related constraints under such sparsity are not binding, implying

frequency stability is not enforced as demonstrated with time-domain simulations. For the case where predictor is trained by AS dataset, the solution is no longer valid in terms of frequency stability when sparsity reaches 90%. It should be noted that although the RS based predictor has an RoCoF prediction accuracy of 77.69% at the sparsity of 70%, frequency requirements are not respected. At the same time, the AS based predictor shows much higher robustness when being incorporated into MILP problem.

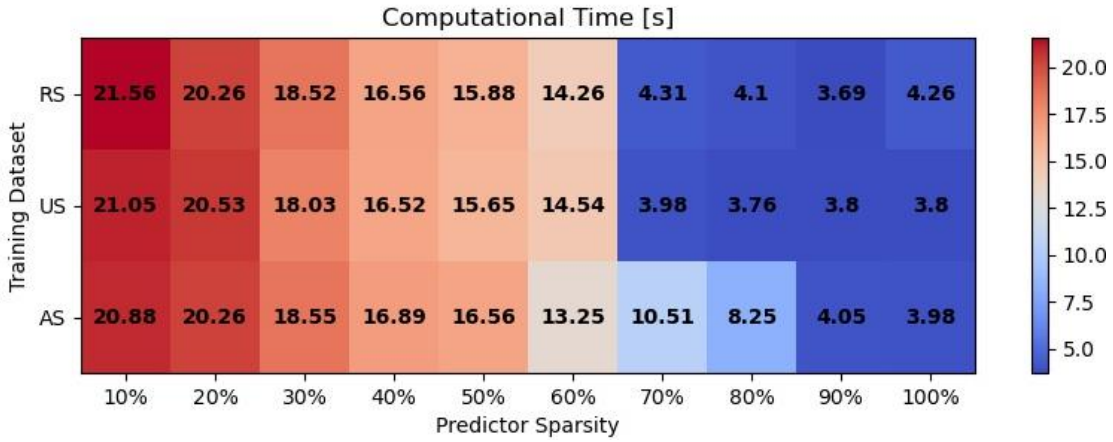


Fig. 6.8. Computational time of SNN-FCUC with different sparsity.

Frequency metrics predictor with sparsity of 0.8 is selected based on the proposed method. The trained well predictor is then incorporated into ALSNN-FCUC models. RoCoF prediction accuracy of the sparse predictor is 87.47%, and the frequency deviation prediction accuracy is 80.80%. γ for active neuron selection is set 0.25. $\dot{f}_{max}$ and Δf_{max} are obtained by conducting time domain simulations under worst contingency on PSS/E at hour 10.

The simulation results of the proposed ALSNN-FCUC model and benchmark models are listed in Table 6.4. Comparing to T-SCUC, all frequency constrained models have relatively higher operational cost, the extra cost comes from the efforts in securing the frequency nadir and RoCoF stability. Solution of LRC-SCUC model is relatively conservative due to approximation error. $\dot{f}_{max}$ of DNN-FCUC model is 0.50 Hz/s, indicating that the solution perfectly satisfies the threshold and there is no conservativeness. It should be also noted that Δf_{max} of models with RoCoF related constraints applied are all within the safe range, implying RoCoF related constraints are more likely to bind comparing to the frequency deviation constraints. And inertia related protections would be a main factor that limits the transition toward RES dominant system.

Table 6.4 Comparison of Different Models

Model	Total Cost [$]	Computational Time [s]	$\dot{f}_{max}$ [Hz/s]	Δf_{max} [Hz]
T-SCUC	419,935	3.89	1.05	0.51
ERC-SCUC	420,171	4.53	0.60	0.29
LRC-SCUC	425,929	6.05	0.44	0.23
DNN-FCUC	422,497	22.56	0.50	0.24
ALSNN-FCUC	421,985	8.56	0.50	0.24

A noticeable increase in computational time can also be observed when there is no approximation in the DNN-FCUC. For the proposed ALSNN-FCUC model, the computational time is reduced by 62% from 22.56 s to 8.56 s when sparse

computation as well as active ReLU linearization is applied. Results show that such algorithm significantly improves the computational efficiency while maintaining the highest post-contingency RoCoF values within an acceptable range.

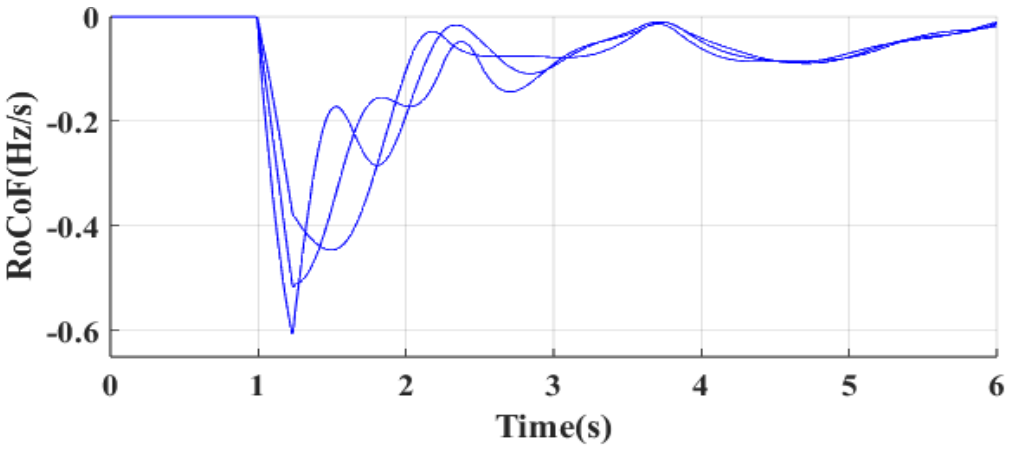
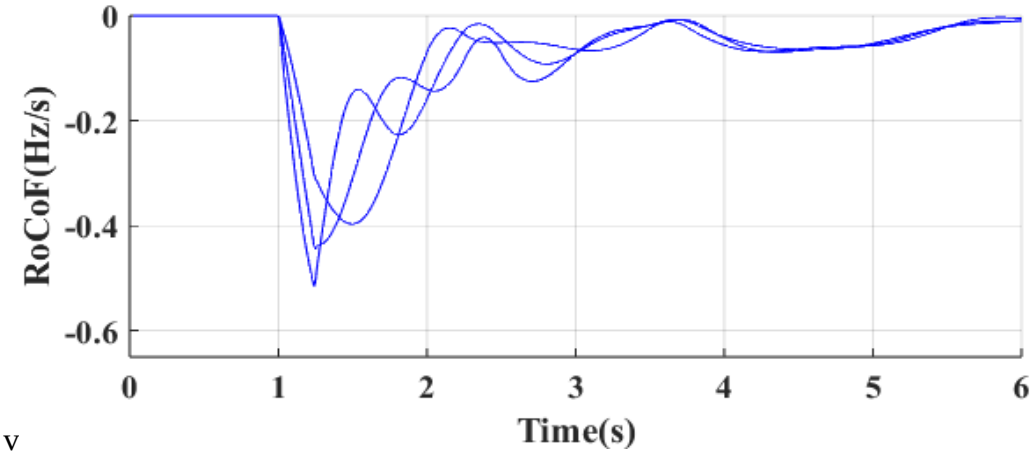


Fig. 6.9 RoCoF evolution of ERC-SCUC model under worst contingency at hour 10.



v

Fig. 6.10 RoCoF evolution of ALSNN-FCUC model under worst contingency at hour 10.

We then compare the time domain simulation results of the proposed ALSNN-FCUC with ERC-SCUC where frequency related constraints are derived from system equivalent model. Worst G-1 contingency is conducted under dispatching

during hour 10. RoCoF evolutions are shown in Fig. 6.9 and Fig. 6.10. As we can see, widely used uniform model-based approach cannot ensure locational RoCoF security due to approximation error of model simplification. With frequency metrics predictor-based constraints added, the proposed ALSNN-FCUC frameworks can secure the system highest locational RoCoF within a safe range.

Table 6.5 Computational Time [s] of Different Constrained Intervals

Total Number of Constrained Hour	4	8	12	16	20	24
DNN-FCUC	23	268	523	NA	NA	NA
ALSNN-FCUC	8	14	50	143	254	1223

Additionally, impact of total number of constrained hours on computational time of FCUC models over 24 hours scheduling horizon is investigated. Results in Table 6.5 show that for DNN-FCUC without sparse computation and active linearization process, computational time for solving FCUC framework increases exponentially as the number of hours that enforce frequency requirements increases. DNN-FCUC reaches the time limit at 3600 seconds when the number of total constrained periods is larger than 16 periods. While the proposed ALSNN-FCUC model has much higher efficiency.

The voltage dynamics of generators are plotted in Fig. 6.11. As we can see, the proposed ALSNN-FCUC dispatch also satisfies the voltage limits for given conditions. It should be noted that voltage and related metrics could be potentially

incorporated into the frequency metrics predictor-based constraints with nearly negligible increase in the computational burden.

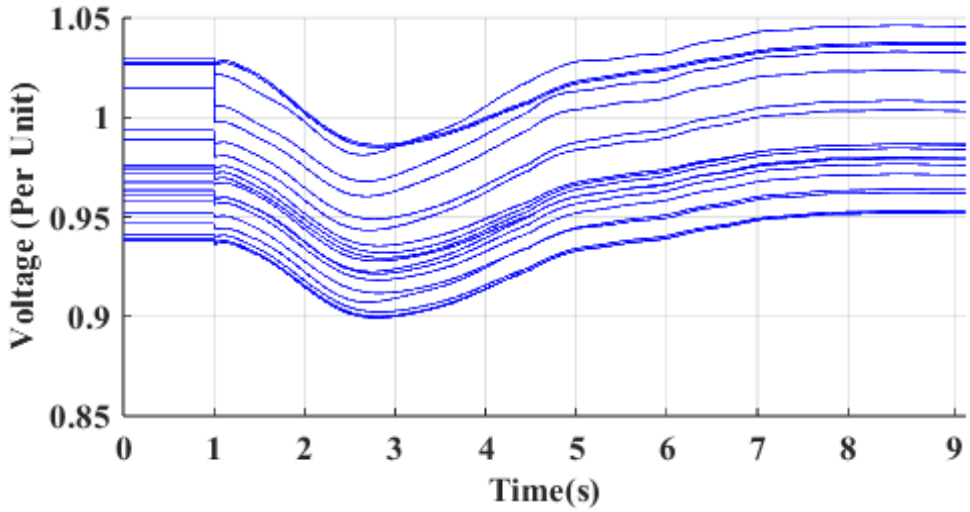


Fig. 6.11 Voltage evolution of ALSNN-FCUC model under worst contingency during period 10.

6.6 Summary

This chapter proposes an ALSNN-FCUC model which incorporates frequency related constraints derived from deep neural networks. A DNN based frequency metrics predictor is constructed to represent maximal frequency and the highest RoCoF value. With concepts of sharing parameters, the proposed approaches are generic and can be implemented to track more potential system metrics without increasing much computational burden.

In addition, the proposed ALSNN-FCUC approach incorporates sparse computations to perform parameter selection and increase neural network sparsity. This proposed sparse DNN subsequently reduces the computational burden of the framework. We propose an active sampling method to improve the robustness of the

trained predictor. This method allows us to increase the sparsity of trained well frequency metric predictor while maintaining the bindingness of these frequency related constraints within FCUC formulations. An active ReLU linearization method is performed over selected neurons to further improve the model efficiency. As compared to traditional DNN based data-driven approach, the proposed ALSNN-FCUC can maintain the system frequency related constraints under worst contingency while reducing the computational time. Verifications on PSS/E show that the proposed ALSNN-FCUC model can efficiently provide high quality solutions which can secure system frequency stability without conservativeness.

7. CONCLUSION AND FUTURE WORK

Power system inertia plays a critical role in maintaining the stability of the power grid. Inertia represents the ability of the power system to respond to sudden changes in the balance between electricity generation and consumption. If the grid experiences a sudden increase or decrease in the amount of electricity being generated, the inertia of the system helps to maintain a stable frequency and prevent power outages. In the recent decades, the increasing integration of variable renewable generation results in the reduction in the system inertia, which poses a serious challenge for frequency regulation. Accurately estimating the inertia of the system is therefore crucial for ensuring that the grid remains stable and reliable under different operating conditions. Additionally, proper estimation of inertia can help grid operators optimize the operation of the power system through frequency related ancillary services, leading to more efficient and cost-effective energy management.

In addition, system inertia is an important parameter that needs to be considered in the formulation of SCUC. This is because system inertia directly affects the response of the power system to disturbances, such as sudden changes in load or generation. Various frequency related constraints are incorporated to ensure that the total system inertia remains above a certain threshold, or to limit the rate of change of frequency (RoCoF) to ensure that the power system remains stable considering the worst contingency. Including inertia related constraints also ensures the power system can respond adequately to disturbances. This is particularly important in power systems with high levels of renewable energy sources, which have lower levels of

inertia compared to traditional fossil-fueled power plants. In such systems, it is necessary to carefully consider the system inertia in order to maintain grid stability.

7.1 Contributions

Evaluation of the system inertia distribution traditionally based on a single disturbance event may be susceptible to power swings and oscillation between machines, which could deteriorate the accuracy of measurements and lead to high biased estimation. Based on the equivalent center of inertia concept, a dynamic system inertia distribution estimation method is proposed in chapter 2. Results also show that power systems with lower renewable energy source (RES) penetration exhibit a better frequency response, with a higher nadir and less steep RoCoF. The conducted sensitivity test determines the optimal integration period for the proposed dynamic inertia estimation method that uses a clustering algorithm to estimate system inertia by accounting for perturbation location and oscillation between machines. The study found that measurements taken from buses within the Center of Inertia (COI) area are relatively stable compared to those from neighboring areas and are therefore more reliable. The estimation process also identified unstable buses experiencing harmonic waves. Additionally, the study examined the impact of the geographic location of RES on the COI area. The proposed method is more robust and accurate in estimating system inertia distribution, and potential applications using the concept of inertia distribution estimation will be explored in the future.

Utilizing neural networks to estimate system inertia constant has become possible due to the vast amounts of data available from power system digital

equipment and advanced measuring infrastructures like PMUs. In chapter 3 we introduce LRCN and GCN based learning algorithms to estimate inertia constant by using system-wide ambient measurements gathered from WAMS as candidate features for model training. A wrapper feature selection is used to optimize the feature combination. Due to the limitations on PMU settings, the authors propose a ZGIB-OPP method to maximize the observability of WAMS with limited PMU resources. Results indicate that the proposed LRCN and GCN models outperform the benchmark DNN and CNN models, even under conditions with high noise levels. The ZGIB-OPP method was also found to improve the performance of all implemented models. This approach is suitable for estimating inertia constant under realistic conditions, considering the IEEE 24-bus system model used in this research features a mix of synchronous generators and inverter-based resources.

In chapter 4, we introduce the concept of locational frequency security and investigate the impact of the Fiedler mode on locational frequency dynamics. The expression of locational frequency dynamics accounting for $G-1$ contingency are derived based on in multi-machine systems reduced model. To capture the highest locational RoCoF during oscillation, a multiple-measurement-window method is proposed tracking the average RoCoF. Additionally, a piecewise linearization-based method is introduced to convert non-linear frequency constraints into linear frequency constraints in the LRC-SCUC model. This enables optimal scheduling of synchronous inertia as well as inertial services provided by non-synchronous resources to meet the minimum system inertia requirement for power systems with higher RES integration.

Simulation results show that incorporating location-based RoCoF constraints in the SCUC model ensures locational frequency security during the worst-case contingency event. These RoCoF-related constraints also significantly affect the scheduling of synchronous generators and the expected system cost. The effect of virtual inertia on inertia pricing and market efficiency is demonstrated, and the results suggest that introducing virtual inertia can reduce the total cost by avoiding starting up unnecessary commitment of extra expensive synchronous generators. Compared to the VI-ERC-SCUC model, the VI-LRC-SCUC model is more sensitive to inertia price.

Chapter 5 proposes two data driven frame works to secure system RoCoF securities under UC schedules. Initially we use model-based method to generate dataset which covers all conditions. Model-based approaches such as T-SCUC, ERC-SCUC and LRC-SCUC that enforcing system locational frequency security are proposed to efficiently generate realistic data for predictor training. DNN and CNN based RoCoF predictors are then constructed to track the highest RoCoF value of low inertia power system under worst contingency on the generated dataset. The well-trained RoCoF predictor is then integrated into the linear SCUC model by applying a PWL method to linearize the ReLU functions of each neuron in the predictor. Furthermore, DNN-RCUC and CNN-RCUC models are proposed to ensure system frequency stability post-contingency.

The effectiveness of several SCUC models, including T-SCUC, ERC-SCUC, LRC-SCUC, DNN-RCUC, and CNN-RCUC, is evaluated on the IEEE 24-bus system

using Pyomo simulations. The results show that the proposed DNN-RCUC and CNN-RCUC models can reduce reserve costs while slightly increasing operational costs.

In addition, the proposed ALSNN-FCUC approach incorporates sparse computations to perform parameter selection and increase neural network sparsity. The proposed sparse DNN subsequently reduces the computational burden of the framework. We propose an active sampling method to improve the robustness of the trained predictor. This method allows us to increase the sparsity of trained well frequency metric predictor while maintaining the bindingness of these frequency related constraints within FCUC formulations. An active ReLU linearization method is performed over selected neurons to further improve the model efficiency. As compared to traditional DNN based data-driven approach, the proposed ALSNN-FCUC can maintain the system frequency related constraints under worst contingency while reducing the computational time. Verifications on PSS/E show that the proposed ALSNN-FCUC model can efficiently provide high quality solutions which can secure system frequency stability without conservativeness.

7.2 Future Work

In chapter 3, LRCN and FCN based inertia estimation algorithms were proposed using system wide measurements from WAMS. In chapter 5 and chapter 6, deep neural networks were leveraged for system frequency metrics tracking, the derived frequency related constraints are incorporated into SCUC thereby securing system frequency stability with smaller RoCoF violation gap comparing to model based SCUC approaches. The linearization of ReLU in neural network has introduced

extra variables, resulting in computational efficiency issues. Along with this preliminary case studies for the proposed NN based RCUC and FCUC models were also presented. The future work can be considered in simultaneously tracking multiple metrics related to system reliability issues such as transient stability, small-signal stability, and voltage stability. Additionally, variable reduction of SCUC could be implemented by predicting generator status and line loading factor using machine learning algorithms, and GNN could be utilized to improve the generator status prediction accuracy by incorporating topography information of the network. The machine learning method can also be applied to power flow optimization problem which, the computational issues could be relieved when the status of generators is not included into the problem. Besides that, the impact of inverter-based sources such as virtual machine and demand side synchronous machines could be modeled into the simulation. Furthermore, data related weather patterns and scenarios can be studied to predict system dynamic rating model and reserve values to handle system with high RES penetration level.

7.3 List of Publications

(1) Mingjian Tuo, Arun Venkatesh Ramesh, Xingpeng Li, "Benefits and Cyber-Vulnerability of Demand Response System in Real-Time Grid Operations", *IEEE Smart Grid Comm*, Nov. 2020, Tempe, AZ, USA.

(2) Mingjian Tuo, "Dynamic Estimation of Power System Inertia Distribution Using Synchrophasor Measurements", *North American Power Symposium (NAPS)*, April. 2021, virtually, Tempe, AZ, USA.

(3) Mingjian Tuo, and Xingpeng Li, “Optimal Allocation of Virtual Inertia Devices for Enhancing Frequency Stability in Low-Inertia Power Systems”, *North American Power Symposium (NAPS)*, Nov. 2021, College Station, TX, USA.

(4) Mingjian Tuo, and Xingpeng Li, “Deep Learning based Security-Constrained Unit Commitment Considering Locational Frequency Stability in Low-Inertia Power Systems”, *North American Power Symposium (NAPS)*, Oct. 2022, Salt Lake City, UT, USA.

(5) Mingjian Tuo, and Xingpeng Li, “Long-term Recurrent Convolutional Networks-based Inertia Estimation using Ambient Measurements”. *IEEE Industry Applications Society Annual Meeting*, Oct. 2022, Detroit, MI, USA.

(6) Vasudharini Sridharan, Mingjian Tuo, Xingpeng Li, " Wholesale Electricity Price Forecasting using Integrated Long-term Recurrent Convolutional Network Model", *Energies*, 15, no. 20: 7606, 2022.

(7) Mingjian Tuo, and Xingpeng Li, "Security-Constrained Unit Commitment Considering Locational Frequency Stability in Low-Inertia Power Grids" *IEEE Transactions on Power Systems*, 2022 (Early Access).

(8) Mingjian Tuo, and Xingpeng Li, " Machine Learning Assisted Inertia Estimation using Ambient Measurements" *IEEE Transactions on Industry Applications*, 2023 (Early Access).

(9) Mingjian Tuo, and Xingpeng Li, " Selectively Linearized Neural Network based RoCoF-Constrained Unit Commitment in Low-Inertia Power Systems " Available: arXiv:2211.08502, Mar. 2023.

REFERENCES

[1] NERC, Transmission System Adequacy and Security, 2005 [Online]. Available: http://www.nerc.com.

[2] F. Milano, F. D¨orfler, G. Hug, D. J. Hill, and G. Verbi˘c, “Foundations and challenges of low-inertia systems,” in *Power Systems Computation Conference (PSCC)*, Jun 2018.

[3] Y. Liu, S. You, and Y. Liu, “Study of wind and PV frequency control in U.S. power grids—EI and TI case studies,” *IEEE Power Energy Technol. Syst.* J., vol. 4, no. 3, pp. 65–73, Sep. 2017.

[4] Luis Badesa, Fei teng, and Goran Strbac, “Conditions for Regional Frequency Stability in Power System Scheduling-Part I: Theory,” *IEEE Trans. Power Syst*., Early Access.

[5] A Adrees, JV Milanović and P Mancarella, “Effect of Inertia Heterogeneity on Frequency Dynamics of Low-inertia Power Systems,” *IET Generc., Transmiss. Distrib.*, vol. 13, no. 14, pp. 2951–2958, Jul. 2019.

[6] R.B. Sharma and G.M. Dhole, “Wide area measurement technology in power systems;” *Procedia Technology*, 25:718–725, 2016.

[7] Y. Cui, S. You, and Y. Liu, "Ambient Synchrophasor Measurement Based System Inertia Estimation," in *2020 IEEE Power Engineering Society General Meeting (PESGM)*, 2020: IEEE, pp. 1-5.

[8] ERCOT, “Inertia: Basic Concepts and Impacts on the ERCOT Grid,” ERCOT, Tech. Rep., Apr. 2018.

[9] N. H. Thi, Y. Guangya, A. H. Nielsen and P. H. Jensen, "Frequency stability improvement of low inertia systems using synchronous condensers," in *2016 IEEE International Conference on Smart Grid Communications*, 2016.

[10] P. Kundur, Power System Stability and Control. New York, NY, USA: McGraw-Hill, 1994.

[11] Electricity Rule Change Proposal, Aust. Energy Market Operator, Melbourne, VIC, Australia, 2017. [Online]. Available: http://www. aemo.com.au.

[12] Fast Frequency Response in the NEM, Aust. Energy Market Operator, Melbourne, VIC, Australia, 2017. [Online]. Available: http://www. aemo.com.au.

[13] The Grid Code, National Grid, London, U.K., 2017. [Online]. Available: http://www.nationalgrid.com.

[14] Enhancement to the Spinning Reserve Requirements for the Singapore Power System, Energy Market Authority, Singapore, 2017. [Online]. Available: http://www.ema.gov.sg.

[15] T. Weckesser and T. Van Cutsem, "Equivalent to represent inertial and primary frequency control effects of an external system," *IET Gener., Transm. & Dis.*, vol. 11, no. 14, pp. 3467–3474, Sep. 2017.

[16] D.P. Chassin, Z. Huang, M.K. Donnelly, C. Hassler, E. Ramirez, C. Ray, "Estimation of WECC system inertia using observed frequency transients," *IEEE Transactions on Power Systems*, vol. 20, pp. 1190-1192, May 2005.

[17] Mingjian Tuo and Xingpeng Li, “Dynamic Estimation of Power System Inertia Distribution Using Synchrophasor Measurements”, in *2020 52nd North American Power Symposium (NAPS)*, Tempe, AZ, USA, Apr. 2021.

[18] Y. Bian, H. Wyman-Pain, F. Li, R. Bhakar, S. Mishra, and N. P. Padhy, “Demand side contributions for system inertia in the GB power system,” *IEEE Transactions on Power Systems*, vol. 33, pp. 3521-3530, Jul. 2011.

[19] Y. Liu, S. You, and Y. Liu, “Study of wind and pv frequency control in u.s. power grids: Ei and ti case studies,” *IEEE Power and Energy Technology Systems Journal*, vol. 4, pp. 65–73, Sep. 2017.

[20] J. Qi, Q. Wu, Y. Zhang, G. Weng, and D. Zhou, “Unified residue method for design of compact wide-area damping controller based on power system stabilizer,’’ *J. Mod. Power Syst. Clean Energy*, vol. 8, no. 2, pp. 367–376, Mar. 2020.

[21] D.P. Chassin, Z. Huang, M.K. Donnelly, C. Hassler, E. Ramirez, C. Ray, “Estimation of WECC system inertia using observed frequency transients,*”* *IEEE Transactions on Power Systems*, vol. 20, pp. 1190-1192, May 2005.

[22] D. Wilson, J. Yu, N. Al-Ashwal, B. Heimission and V. Terzija, “Measuring effective area inertia to determine fast-acting frequency response,” *Electr. Power Energy Syst*., vol. 113, pp. 1-8, Dec. 2019.

[23] D. Wilson, J. Yu, N. Al-Ashwal, B. Heimission and V. Terzija, “Measuring effective area inertia to determine fast-acting frequency response,” *Electr. Power Energy Syst.*, vol. 113, pp. 1-8, Dec 2019.

[24] DSA Tools – Power Labs Inc., British Columbia, Canada, 2011.

[25] S. You, Y. Liu, J. Tan, M. T. Gonzalez, X. Zhang, Y. Zhang and Y. Liu, "Comparative Assessment of Tactics to Improve Primary Frequency Response Without Curtailing Solar Output in High Photovoltaic Interconnection Grids," *IEEE Trans. Sustainable Energy*, vol.10, Issue 2, pp. 718-728, 2019.

[26] Gu, H., Yan, R., & Saha, T. K. "Minimum synchronous inertia requirement of renewable power systems." *IEEE Transactions on Power Systems*. 33(2), 1533–1543, June. 2017.

[27] Federico Milano, and Álvaro Ortega, "A Method for Evaluating Frequency Regulation in an Electrical Grid – Part I: Theory," *IEEE Trans. Power Syst.*, vol. 36, no. 1, pp. 183-193, Jan. 2021

[28] Svenska Kraftn¨at, Statnett, Fingrid and Energinet.dk, "Challenges and Opportunities for the Nordic Power System," *Tech. Rep*., 2016.

[29] Bala Kameshwar Poolla, Saverio Bolognani and Florian Dörfler, "Optimal Placement of Virtual Inertia in Power Grids," *IEEE Transactions on Automatic Control*. vol. 62, no. 12, pp. 6209-6220, Dec. 2017.

[30] A. N. Rao, P. V. Priya, M. Kowsalya, and R. Gnanadass, ''Wide area monitoring for energy system: A review,'' *Int. J. Ambient Energy*, vol. 40, no. 5, pp. 537–553, Jul. 2019.

[31] A. Schmitt and B. Lee, "Steady-state inertia estimation using a neural network approach with modal information," in *2017 IEEE Power Energy Society General Meeting*, Jul. 2017, pp. 1–5.

[32] Abodh Poudyal, Robert Fourney, Reinaldo Tonkoski, Timothy M. Hansen, Ujjwol Tamrakar, and Rodrigo D. Trevizan, “Convolutional Neural Network-based Inertia Estimation using Local Frequency Measurements”, in *Proc. 52nd North Amer. Power Symp.*, April. 2021, virtually, Tempe, AZ, USA.

[33] F. Scarselli, M. Gori, A. C. Tsoi, M. Hagenbuchner and G. Monfardini, "The Graph Neural Network Model," *IEEE Transactions on Neural Networks*, vol. 20, no. 1, pp. 61-80, Jan. 2009.

[34] Thuan Pham and Xingpeng Li, “Reduced Optimal Power Flow Using Graph Neural Network”, in *54th North American Power Symposium*, Salt Lake City, UT, USA, Oct. 2022.

[35] U. Markovic, Z. Chu, P. Aristidou, and G. Hug, “LQR-based adaptive virtual synchronous machine for power systems with high inverter penetration”, *IEEE Trans. Sust. Energy*, 2019, 10, (3), pp. 1501–1512.

[36] K. Tuttelberg, J. Kilter, D. Wilson, and K. Uhlen, “Estimation of power system inertia from ambient wide area measurements,” *IEEE Transactions on Power Systems*, vol. 33, pp. 7249–7257, Nov 2018.

[37] H. Chen, T. Jiang, H. Yuan, H. Jia, L. Bai, and F. Li, ‘‘Wide-area measurement-based voltage stability sensitivity and its application in voltage control,’’ *Int. J. Electr. Power Energy Syst.*, vol. 88, pp. 87–98, Jun. 2017.

[38] J. Zhang and H. Xu, “Online identification of power system equivalent inertia constant,” *IEEE Transactions on Industrial Electronics*, vol. 64, no. 10, pp. 8098–8107, Apr. 2017.

[39] Sridharan V, Tuo M, Li X, “Wholesale electricity price forecasting using integrated long-term recurrent convolutional network model.” *Energies* 15(20):7606.

[40] Saleh Albeiwi and Ausif Mahmood, “A Framework for Designing the Architectures of Deep Convolutional Neural Networks” *Entropy*, 19(6), 242, May 2017.

[41] S. Hochreiter and J. Schmidhuber, ‘‘Long short-term memory,’’ *Neural Comput*., vol. 9, no. 8, pp. 1735–1780, 1997.

[42] Jeff Donahue, Lisa Anne Hendricks, Marcus Rohrbach, Subhashini Venugopalan, Sergio Guadarrama, Kate Saenko, and Trevor Darrell, “Long-term Recurrent Convolutional Networks for Visual Recognition and Description,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 4, pp. 677-691, April 1, 2017.

[43] Akshita Chugh, “MAE, MSE, RMSE, Coefficient of Determination, Adjusted R Squared - Which Metric is Better?”. https://medium.com/analytics-vidhya/mae-mse-rmse-coefficient-of-determination-adjusted-r-squared-which-metric-is-better-cd0326a5697e.

[44] Pope, P.E.; Kolouri, S.; Rostami, M.; Martin, C.E.; Hoffmann, H. Explainability, “Methods for Graph Convolutional Neural Networks,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Long Beach, CA, USA, 16–20 June 2019.

[45] F. Ebrahimzadeh, M. Adeen, and F. Milano, “On the impact of topology on power system transient and frequency stability,” in *Proc. IEEE Int. Conf. Environ. Electr. Eng., IEEE Ind. Commercial Power Syst. Eur. (EEEIC/I&CPS Europe)*, Jun. 2019, pp. 1-5.

[46] T.Weckesser, H. Jóhannsson, M. Glavic, and J. Østergaard, “An improved on-line contingency screening for power system transient stability assessment,” Electr. *Power Compon. Syst*., vol. 45, no. 8, pp. 852-863, May 2017.

[47] Johnson T and Moger T, “A critical review of methods for optimal placement of phasor measurement units,” *Int. Trans. on Elec. Ener. Sys*. 31.

[48] 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, May 2009.

[49] Mingjian Tuo, Arun Venkatesh Ramesh, Xingpeng Li, “Benefits and Cyber-Vulnerability of Demand Response System in Real-Time Grid Operations”, *IEEE Smart Grid Comm*, Nov. 2020, Tempe, AZ, USA.

[50] M. Brown, M. Biswal, S. Brahma, S. J. Ranade, and H. Cao, “Characterizing and quantifying noise in PMU data,” in *2016 IEEE Power and Energy Society General Meeting (PESGM)*, Jul. 2016, pp. 1–5.

[51] L. Ruttledge, N. W. Miller, J. O’Sullivan, and D. Flynn, “Frequency response of power systems with variable speed wind turbines,” *IEEE Transactions on Sustainable Energy*, vol. 3, no. 4, pp. 683–691, 2012.

[52] N. Miller, D. Lew, and R. Piwko, “Technology capabilities for fast frequency response,” *GE Energy Consulting, Tech. Rep*., 2017.

[53] J. O'Sullivan, A. Rogers, D. Flynn, P. Smith, and M. O'Malley,” Studying the maximum instantaneous non-synchronous generation in an Island system—Frequency stability challenges in Ireland,” *IEEE Trans. Power Syst*., vol. 29, no. 6, pp. 2943–2951, Nov. 2014.

[54] Energy Emergencies Executive Committee, “GB power system disruption on 9 August 2019,” 2019. [Online]. Available: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/855767/e3c-gb-power-disruption-9-august-2019-final-report.pdf.

[55] A. Crivellaro, A. Tayyebi, C. Gavriluta, D. Groß, A. Anta, F. Kupzog and F. Dörfler, “Beyond low-inertia systems: Massive integration of grid-forming power converters in transmission grids,” in *2020 IEEE Powe & Energy Society General Meeting (PESGM)*, Montreal, QC, Aug, 2020.

[56] P. Du, R. Baldick, and A. Tuohy, “Integration of Large-Scale Renewable Energy into Bulk Power Systems”. *Springer*, 2017, pp. 1-25.

[57] Jun Cao, Wenjuan Du, Haifeng Wang and Malcolm McCulloch, “Optimal Sizing and Control Strategies for Hybrid Storages System as Limited by Grid Frequency Deviations,” *IEEE Trans. Power Syst.*, vol. 33, no. 5, pp. 5486-5495, Feb. 2018.

[58] Vincenzo Trovato, Agnès Bialecki and Anes Dallagi Demand, “Unit Commitment with Inertia-Dependent and Multispeed Allocation of Frequency Response Services,” *IEEE Trans. Power Syst.*, vol. 34, no. 2, pp. 1537-1548, March. 2019.

[59] J Riesz, J Gilmore, and I MacGill, “Frequency Control Ancillary Service Market Design: Insights from the Australian National Electricity Market,” *Elect. J.*, vol. 28, no. 3, pp. 86–99, 2015.

[60] EirGrid and SONI, “DS3 System Services: Review TSO Recommendations,” EirGrid, Dublin, Ireland, Tech. Rep., May 2013.

[61] Eto, Joseph H., “Use of frequency response metrics to assess the planning and operating requirements for reliable integration of variable renewable generation,” Lawrence Berkeley National Laboratory, Berkeley, CA, USA, Tech. Rep. LBNL-4142E, Jan. 2011.

[62] EirGrid and SONI, “Operational Constraints Update,” *EirGrid, Tech. Rep.*, March 2019.

[63] H. Chavez, R. Baldick, and S. Sharma, “Governor rate-constrained OPF for primary frequency control adequacy,” *IEEE Trans. Power Syst.*, vol. 29, no. 3, pp. 1473–1480, 2014.

[64] Z. Zhang, E. Du, F. Teng, N. Zhang, and C. Kang, ‘‘Modeling frequency dynamics in unit commitment with a high share of renewable energy,’’ IEEE Trans. Power Syst., vol. 35, no. 6, pp. 4383–4395, Nov. 2020.

[65] M. Paturet, Economic valuation and pricing of inertia in inverter-dominated power systems, Master thesis, Swiss Federal Institute of Technology (ETH) Zurich (2019).

[66] L Badesa, F Teng, G Strbac, "Simultaneous scheduling of multiple frequency services in stochastic unit commitment", *IEEE Trans. Power Syst*., vol. 34, no. 5, pp. 3858–3868, Mar. 2019.

[67] T. Borsche and F. Dorfler, "On placement of synthetic inertia with explicit time-domain constraints," May 2017. [Online]. Available: https://arxiv.org/pdf/1705.03244.pdf.

[68] A Adrees, JV Milanović and P Mancarella, "Effect of Inertia Heterogeneity on Frequency Dynamics of Low-inertia Power Systems," *IET Generc., Transmiss. Distrib*., vol. 13, no. 14, pp. 2951–2958, Jul. 2019.

[69] Bala Kameshwar Poolla, Saverio Bolognani and Florian Dörfler, "Optimal Placement of Virtual Inertia in Power Grids," *IEEE Transactions on Automatic Control*. vol. 62, no. 12, pp. 6209-6220, Dec. 2017.

[70] N. Mendis, K. M. Muttaqi, and S. Perera, "Management of battery supercapacitor hybrid energy storage and synchronous condenser for isolated operation of PMSG based variable-speed wind turbine generating systems," *IEEE Trans. Smart Grid*, vol. 5, no. 2, pp. 944–953, Mar. 2014.

[71] N. A. Masood, R. Yan, T. K. Saha, and N. Modi, "Frequency response and its enhancement using synchronous condensers in presence of high wind

penetration," in *Proc. IEEE Power Energy Soc. General Meeting*, Denver, CO, USA, Jul. 26–30, 2015, pp. 1–5.

[72] A. Moeini and I. Kamwa, "Analytical concepts for reactive power base primary frequency control in power systems," *IEEE Trans. Power Syst.*, vol. 31, no. 6, pp. 4217–4230, Nov. 2016.

[73] H. T. Nguyen, G. Yang, A. H. Nielsen, and P. H. Jensen, ''Combination of synchronous condenser and synthetic inertia for frequency stability enhancement in low-inertia systems,'' *IEEE Trans. Sustain. Energy*, vol. 10, no. 3, pp. 997–1005, Jul. 2019.

[74] Y. Liu, S. You, and Y. Liu, "Study of wind and PV frequency control in U.S. power grids—EI and TI case studies," *IEEE Power Energy Technol. Syst. J.*, vol. 4, no. 3, pp. 65–73, Sep. 2017

[75] Q.-C. Zhong and T. Hornik, "Control of Power Inverters in Renewable Energy and Smart Grid Integration". *John Wiley & Sons*, 2012.

[76] H. Bevrani, T. Ise, and Y. Miura, "Virtual synchronous generators: A survey and new perspectives," *International Journal of Electrical Power & Energy Systems*, vol. 54, pp. 244–254, 2014.

[77] S. D'Arco and J. A. Suul, "Virtual synchronous machines - classification of implementations and analysis of equivalence to droop controllers for microgrids," in *IEEE PowerTech Conference*. Grenoble, France, 2013.

[78] Q.-C. Zhong and G. Weiss, “Synchronverters: Inverters that mimic synchronous generators,” *IEEE Trans. Ind. Electron.*, vol. 58, no. 4, pp. 1259– 1267, Apr. 2011.

[79] C. Arghir, T. Jouini, and F. Dörfler, “Grid-forming Control for Power Converters based on Matching of Synchronous Machines”. *Automatica*, vol. 95, pp. 273–282, Sep. 2018.

[80] M. Tyloo, L. Pagnier, and P. Jacquod, “The key player problem in complex oscillator networks and electric power grids: Resistance centralities identify local vulnerabilities,” *Sci. Adv.*, vol. 5, no. 11, 2019, Art. No eaaw8359.

[81] F. D¨orfler and F. Bullo, “Kron reduction of graphs with applications to electrical networks,” *IEEE Trans. Circuits Syst. I: Regular Papers* vol. 60, no. 1, pp. 150–163, Jan. 2013.

[82] Mingjian Tuo, and Xingpeng Li, “Optimal Allocation of Virtual Inertia Devices for Enhancing Frequency Stability in Low-Inertia Power Systems”, in *Proc. 53rd North Amer. Power Symp.*, Nov. 2021, accepted.

[83] L. Pagnier and P. Jacquod, “Inertia location and slow network modes determine disturbance propagation in large-scale power grids,” *PLoS ONE* 14, e0213550 (2019).

[84] Borsche, T.S., Liu, T., Hill, D.J. “Effects of rotational inertia on power system damping and frequency transients,” in *Conf. on Decision and Control*, 2015.

[85] Turaj Amraee, Mohammad Ghaderi Darebaghi, Alireza Soroudi and Andrew Keane, “Probabilistic Under Frequency Load Shedding Considering RoCoF

Relays of Distributed Generators," *IEEE Trans. Power Syst.*, vol. 33, no. 4, pp. 3587-3598, July. 2018.

[86] Bahman Alinezhad Osbouei, Gareth A. Taylor, Olivier Bronckart, Johan Maricq and Martin Bradley, "Impact of Inertia Distribution on Power System Stability and Operation," *2019 IEEE Milan PowerTech*, Milan, Italy, June. 2019.

[87] P. W. Sauer and M. A. Pai, *Power System Dynamics and Stability*. Upper Saddle River, NJ, USA: Prentice-Hall, 1998.

[88] Mingjian Tuo, and Xingpeng Li, "Dynamic Estimation of Power System Inertia Distribution Using Synchrophasor Measurements", in *Proc. 52nd North Amer. Power Symp.*, April. 2021, virtually, Tempe, AZ, USA, pp. 1–6.

[89] Federico Milano, and Álvaro Ortega, "A Method for Evaluating Frequency Regulation in an Electrical Grid – Part I: Theory," *IEEE Trans. Power Syst.*, vol. 36, no. 1, pp. 183-193, Jan. 2021.

[90] H. Ahmadi and H. Ghasemi, "Security-constrained unit commitment with linearized system frequency limit constraints," *IEEE Trans. Power Syst.*, vol. 29, no. 4, pp. 1536-1545, Jul. 2014.

[91] A. Magnani and S. P. Boyd, "Convex piecewise-linear fitting," *Optimize Eng.*, vol. 10, no. 1, pp. 1–17, 2009.

[92] Arun Venkatesh Ramesh, Xingpeng Li, and Kory W Hedman, "An accelerated-decomposition approach for security-constrained unit commitment with corrective network reconfiguration," *IEEE Trans. Power Syst.*, July. 2021, Early Access.

[93] Hart, William E., Jean-Paul Watson, and David L. Woodruff. “Pyomo: modeling and solving mathematical programs in Python.” Mathematical Programming Computation 3, no. 3 (2011): 219-260.

[94] Hart, William E., Carl Laird, Jean-Paul Watson, David L. Woodruff, Gabriel A. Hackebeil, Bethany L. Nicholson, and John D. Siirola. Pyomo – Optimization Modeling in Python. Springer, 2017.

[95] Gurobi Optimization Inc. Gurobi optimizer reference manual. http://www.gurobi.com/, 2014.

[96] Matthieu Paturet, Uros Markovic, Stefanos Delikaraoglou, Evangelos Vrettos, Petros Aristidou and Gabriela Hug, “Stochastic Unit Commiment in Low-Inertia Grids,” *IEEE Trans. Power Syst.*, vol. 35, no. 5, pp. 3448-3458, Sept. 2020.

[97] DSA Tools – Power Labs Inc., British Columbia, Canada, 2011.

[98] Xingpeng Li and Kory W. Hedman, “Enhanced energy management system with corrective transmission switching – Part I: Methodology,” *IEEE Trans. Power Syst.*, vol. 34, no. 6, pp. 4490-4502, Nov. 2019.

[99] P. Du and W. Li, “Frequency response impact of integration of HVDC into a low-inertia AC power grid,” *IEEE Trans. Power Syst*., vol. 36, no. 1, pp. 613-622, Jan. 2021.

[100] A. Sajadi, R. W. Kenyon, and B.-M. Hodge, “Synchronization in electric power networks with inherent heterogeneity up to 100% inverter-based renewable generation,” *Nature Communications*, vol. 13, no. 1, pp. 1–12.

[101] F. Teng, V. Trovato, and G. Strbac, “Stochastic Scheduling With Inertia Dependent Fast Frequency Response Requirements,” *IEEE Transactions on Power Systems*, vol. 31, no. 2, pp. 1557–1566, 2016.

[102] P. Denholm, T. Mai, R. W. Kenyon, B. Kroposki, and M. O’Malley, ‘‘Inertia and the power grid: A guide without the spin,’’ Nat. Renew. Energy Lab. (NREL), Golden, CO, USA, Tech. Rep., 2020.

[103] H. Chavez, R. Baldick, and S. Sharma, “Governor rate-constrained OPF for primary frequency control adequacy,” *IEEE Trans. Power Syst.*, vol. 29, no. 3, pp. 1473–1480, 2014.

[104] Z. Zhang, E. Du, F. Teng, N. Zhang, and C. Kang, ‘‘Modeling frequency dynamics in unit commitment with a high share of renewable energy,’’ IEEE Trans. Power Syst., vol. 35, no. 6, pp. 4383–4395, Nov. 2020.

[105] Mingjian Tuo and Xingpeng Li, “Security-Constrained Unit Commitment Considering Locational Frequency Stability in Low-Inertia Power Grids”, *IEEE Trans. Power Syst.*, Oct. 2022 (Early Access).

[106] Cunzhi Zhao and Xingpeng Li, “Microgrid Day-Ahead Scheduling Considering Neural Network based Battery Degradation Model,” arXiv:2112.08418, Feb. 2022.

[107] Vasudharini Sridharan, Mingjian Tuo and Xingpeng Li, “Wholesale Electricity Price Forecasting using Integrated Long-term Recurrent Convolutional Network Model,” arXiv:2112.13681, Dec. 2021.

[108] Thuan Pham and Xingpeng Li, "Neural Network-based Power Flow Model," IEEE Green Technologies Conference, Houston, TX, USA, Mar. 2022.

[109] Thuan Pham and Xingpeng Li, "Reduced Optimal Power Flow Using Graph Neural Network," arXiv:2206.13591, Jun. 2022.

[110] D. Lagos and N. D. Hatziargyriou, "Data-driven frequency dynamic unit commitment for island systems with high RES penetration," *IEEE Trans. Power Syst.*, vol. 8950, no. c, pp. 1–13, 2021.

[111] Y. Zhang, H. Cui, J. Liu, F. Qiu, T. Hong, R. Yao, and F. Li, "Encoding frequency constraints in preventive unit commitment using deep learning with region-of-interest active sampling," *IEEE Transactions on Power Systems*, vol. 37, no. 3, pp. 1942–1955, 2022.

[112] Y. Bian, H. Wyman-Pain, F. Li, R. Bhakar, S. Mishra, and N. P. Padhy, "Demand side contributions for system inertia in the GB power system," *IEEE Transactions on Power Systems*, vol. 33, pp. 3521-3530, Jul. 2011.

[113] Mingjian Tuo and Xingpeng Li, "Dynamic Estimation of Power System Inertia Distribution Using Synchrophasor Measurements", 2020 52nd North American Power Symposium (NAPS), Apr. 2021.

[114] J. Gu, Z. Wang, J. Kuen, L. Ma, A. Shahroudy, B. Shuai, T. Liu, X. Wang, G. Wang, J. Cai, and T. Chen, "Recent advances in convolutional neural networks," Pattern Recognition, vol. 77, pp. 354–377, 2018.

[115] X.-X. Niu, C. Y. Suen, "A novel hybrid cnn–svm classifier for recognizing handwritten digits", *Pattern Recognition* 45 (4) (2012) 1318–1325.

[116] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, “Imagenet large scale visual recognition challenge”, in *International Journal of Conflict and Violence (IJCV)* 115 (3) (2015) 211–252.

[117] K. Simonyan, A. Zisserman, “Very deep convolutional networks for largescale image recognition”, in *Proceedings of the International Conference on Learning Representations (ICLR)*, 2015.

[118] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich, “Going deeper with convolutions”, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015, pp. 1–9.

[119] Mingjian Tuo and Xingpeng Li, “Deep Learning based Security-Constrained Unit Commitment Considering Locational Frequency Stability in Low-Inertia Power Systems”, in *54th North American Power Symposium*, Salt Lake City, UT, USA, Oct. 2022.

[120] Sheng Liu, Zhihui Zhu, Qing Qu, and Chong You, “Robust training under label noise by over-parameterization,” arXiv preprint arXiv:2202.14026, 2022. 2.

[121] T. T. Nguyen, N. D. Nguyen, and S. Nahavandi, ‘‘Deep reinforcement learning for multiagent systems: A review of challenges, solutions, and applications,’’ *IEEE Trans. Cybern.*, vol. 50, no. 9, pp. 3826–3839, Sep. 2020.

[122] H. Pulgar-Painemal, Y. Wang, and H. Silva-Saravia, “On inertia distribution inter-area oscillations and location of electronically interfaced resources,” *IEEE Transactions on Power Systems*, vol. 33, pp. 995–1003, Jan. 2017.

[123] Torsten Hoefler, Dan Alistarh, Tal Ben-Nun, Nikoli Dryden, and Alexandra Peste. 2021. Sparsity in Deep Learning: Pruning and growth for efficient inference and training in neural networks. J. Mach. Learn. Res. 22 (2021), 241:1–241:124.

[124] M. Zhu and S. Gupta, “To prune, or not to prune: Exploring the efficacy of pruning for model compression,” in *Proc. 6th Int. Conf. Learn. Representations*, 2017.

[125] B. Settles, “Active learning literature survey,” University of Wis-consin Madison Department of Computer Sciences, Tech. Rep., 2009.

[126] W. Xiang, H.-D. Tran, and T. T. Johnson. (Dec. 2017). “Reachable set computation and safety verification for neural networks with ReLU activations.” [Online]. Available: https://arxiv.org/abs/1712.08163.

[127] Liu, B. and Liang, Y. (2021). Optimal function approximation with ReLU neural networks. *Neurocomputing*, 435:216–227.

[128] C. Grigg, P. Wong, P. Albrecht, R. Allan, M. Bhavaraju, R. Billinton, Q. Chen, C. Fong, S. Haddad, S. Kuruganty, W. Li, R. Mukerji, D. Patton, N. Rau, D. Reppen, A. Schneider, M. Shahidehpour, C. Singh, “The IEEE Reliability Test System-1996. A report prepared by the Reliability Test System Task Force of

the Application of Probability Methods Subcommittee," *IEEE Transactions on Power Systems*, vol. 14, no. 3, pp. 1010–1020, Aug. 1999.

[129] Hart, William E., Carl Laird, Jean-Paul Watson, David L. Woodruff, Gabriel A. Hackebeil, Bethany L. Nicholson, and John D. Siirola. Pyomo – Optimization Modeling in Python. Springer, 2017.

[130] Siemens, PSS/E 35.0.00 Model Library, Siemens Industry, Inc., Schenectady, NY, USA, 2019.