

South Dakota State University

Open PRAIRIE: Open Public Research Access Institutional Repository and Information Exchange

Electronic Theses and Dissertations

2020

Distributed Machine Learning Approach to Fast Frequency Response-based Inertia Estimation in Low Inertia Grids

Abodh Poudyal

South Dakota State University

Follow this and additional works at: <https://openprairie.sdstate.edu/etd>

 Part of the Power and Energy Commons

Recommended Citation

Poudyal, Abodh, "Distributed Machine Learning Approach to Fast Frequency Response-based Inertia Estimation in Low Inertia Grids" (2020). *Electronic Theses and Dissertations*. 4088.
<https://openprairie.sdstate.edu/etd/4088>

This Thesis - Open Access is brought to you for free and open access by Open PRAIRIE: Open Public Research Access Institutional Repository and Information Exchange. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of Open PRAIRIE: Open Public Research Access Institutional Repository and Information Exchange. For more information, please contact michael.biondo@sdstate.edu.

DISTRIBUTED MACHINE LEARNING APPROACH TO FAST FREQUENCY
RESPONSE-BASED INERTIA ESTIMATION IN LOW INERTIA GRIDS

BY

ABODH POUDYAL

A thesis submitted in partial fulfilment of the requirements for the

Master of Science

Major in Electrical Engineering

South Dakota State University

2020

THESIS ACCEPTANCE PAGE

Abodh Poudyal

This thesis is approved as a creditable and independent investigation by a candidate for the master's degree and is acceptable for meeting the thesis requirements for this degree. Acceptance of this does not imply that the conclusions reached by the candidate are necessarily the conclusions of the major department.

Timothy M. Hansen

Advisor

Date

Siddharth Suryanarayanan

Department Head

Date

Dean, Graduate School

Date

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude towards my thesis advisors Dr. Timothy M. Hansen, Dr. Reinaldo Tonkoski, and Dr. Robert Fourney for their continuous guidance, support, and encouragement. This work would not have been possible without their valuable insights, motivation, and enthusiasm.

Besides my advisors, I would like to thank the graduate faculty representative Dr. Jonathan Jenks for his participation and interest in my work.

I am very thankful to Dr. Ujjwol Tamrakar for his continuous guidance and help, that pushed me towards the end goal of this project. Additionally, I would like to thank Dr. Rodrigo Daniel Trevizan for guiding me in different parts of this work. Special thanks to Niranjan and Fernando for sharing their work and helping me in completing my research. I am grateful to my friends Deepak, Kapil, Dr. Venkat, Samitnjaya, Bikram, Romancha, Ibin, Prateek, Rupak, Dr. Shiva, Sunil, Nischal, Manisha, and Avijit for bolstering my efforts and research goals.

Last but not the least, I would like to express my gratitude to my mother, Aruna Upadhyay, my brother Dr. Prasanta Poudyal, my sister Prasamsa Poudyal, and my brother-in-law Dr. Nimish Kumar Acharya for providing me with continuous encouragement and support throughout my study and research. I would also like to convey my deepest love towards my nephews Nihit and Parth.

CONTENTS

ABBREVIATIONS	ix
LIST OF FIGURES	xiv
ABSTRACT	xv
CHAPTER 1 INTRODUCTION	1
1.1 Background	1
1.2 Objective	6
1.3 Contribution	6
1.4 Thesis Outline	6
CHAPTER 2 CHALLENGES OF COMBINED VOLTAGE AND FREQUENCY CONTROL IN LOW VOLTAGE GRIDS USING POWER ELEC- TRONIC CONVERTERS	8
2.1 Background	8
2.2 Related Works	9
2.3 Line Parameter Impact on Grid Operation	10
2.3.1 High Voltage Transmission Grid	10
2.3.2 Low Voltage Distribution Grid	12
2.4 Grid Codes for Low Voltage Grids	13
2.5 Fast Frequency Response using Distributed Energy Resources	14
2.6 Related Works on Recent IEEE Grid Standards	15
2.7 Combined Voltage and Frequency Control	18
2.7.1 Simulation Setup	18
2.7.2 Results and Analysis	19
2.7.3 Market for Combined Voltage and Frequency Control	21

2.8 Conclusions	23
CHAPTER 3 SINGLE AREA INERTIA ESTIMATION USING NEURAL NET- WORKS	24
3.1 Introduction	24
3.2 Related Works	24
3.3 Proposed Work	26
3.4 Traditional Power System Frequency Dynamics	26
3.4.1 Characterizing a Frequency Event	26
3.4.2 Swing Equation and Inertia Constant	27
3.4.3 Power System Frequency Control Loops	28
3.5 Inertia Estimation using Neural Networks	29
3.5.1 System Perturbation using Excitation Signals	29
3.5.2 Estimating Inertia Constant using Multilayer Perceptron	30
3.5.3 Estimating Inertia Constant using CNN	32
3.6 Simulation Setup	35
3.6.1 Overview	35
3.6.2 Simulation Benchmark	35
3.6.3 Neural Network Architecture and Hyperparameters Selection	36
3.7 Results and Analysis	37
3.7.1 Estimating Damping Constant using Neural Network	37
3.7.2 Inertia Estimation using Neural Network	40
3.8 Conclusion	43
CHAPTER 4 MULTI-AREA INERTIA ESTIMATION USING DISTRIBUTED MACHINE LEARNING	47
4.1 Introduction	47
4.2 Related Works	48

4.3	Proposed Work	49
4.4	Two-Area Power System	50
4.5	Multi-Area Inertia Estimation using Federated Learning	52
4.5.1	System Perturbation using Excitation Signal	52
4.5.2	Independently and Identically Distributed Data	53
4.5.3	Federated Learning-based Inertia Estimation	55
4.6	Simulation Setup	59
4.6.1	Overview	59
4.6.2	Simulation Benchmark	60
4.6.3	Data Distribution and Hyperparameters Selection for FL	61
4.7	Results and Analysis	62
4.7.1	Performance metrics for IID data	62
4.7.2	Performance metrics for non-IID data	63
4.7.3	Communication Cost	65
4.8	Conclusion	67

CHAPTER 5	INERTIA ESTIMATION OF A SYSTEM WITH NON-SYNCHRONOUS UNIT	69
5.1	Introduction	69
5.2	Proposed Work	70
5.3	Inertia Estimation of a System with MPC Unit	71
5.3.1	MPC-based Single Area System	71
5.3.2	Inertia Estimation Using Pre-Trained CNN	73
5.4	Simulation Setup	74
5.4.1	Case 1: Simulation on a Single Frequency Snapshot	75
5.4.2	Case 2: Simulation on Multiple Frequency Snapshots	75
5.5	Results and Analysis	77
5.5.1	Case 1: Simulation on a Single Frequency Snapshot	77

5.5.2 Case 2: Simulation on Multiple Frequency Snapshots	79
5.6 Conclusion	81
CHAPTER 6 CONCLUSIONS	82
REFERENCES	84

ABBREVIATIONS

ACE	area control error
AGC	automatic generation control
ANN	artificial neural network
AWGN	add white Gaussian noise
CNN	convolutional neural network
DER	distributed energy resource
ERCOT	Electricity of Reliability Council of Texas
ESS	energy storage system
FERC	Federal Energy Regulatory Commission
FFR	fast frequency response
FL	federated learning
HV	high voltage
IEEE	Institute of Electrical and Electronics Engineers
IID	independently and identically distributed
ISO	independent system operator
LV	low voltage
MLP	multilayer perceptron
MPC	model predictive controller
MSE	mean squared error
MV	medium voltage
NERC	North American Electric Reliability Corporation
PCC	point of common coupling
PHEV	plug-in hybrid electric vehicles
PLL	phase-locked-loop
PMU	phasor measurement unit

PV	photovoltaic
ReLU	rectified linear unit
RES	renewable energy source
ROCOF	rate-of-change-of-frequency
SNR	signal to noise ratio
UFLS	under frequency load shedding
VI	virtual inertia

LIST OF FIGURES

Figure 1.1.	Distributed energy resources.	1
Figure 1.2.	Installed capacity and forecast of distributed solar PV (in MW_{dc}) per year in the US [4].	2
Figure 1.3.	Transition of power grid from conventional synchronous generator-dominated system to inverter-dominated system [6].	3
Figure 1.4.	Change in frequency versus generation loss for different period of time at ERCOT [8].	4
Figure 1.5.	Trigerring of UFLS at the National Grid of the Great Britain due to high penetration of RES [9]. The upper figure shows the frequency response of the grid following an event whereas the lower figure represents the generation mix of different resources by year.	5
Figure 2.1.	General transmission model of a grid [33].	11
Figure 2.2.	Different stages of frequency control [38].	15
Figure 2.3.	Power flow in a grid by aggregating DERs [41]	16
Figure 2.4.	Frequency-watt droop control in LV grids [38]	17
Figure 2.5.	Simulation benchmark for combined voltage and frequency control.	19
Figure 2.6.	Effect of change in active power on voltage with varying R/X ratios.	21
Figure 2.7.	Interaction between voltage and frequency dynamics when inertia constant is 2 s.	22
Figure 3.1.	System frequency response following an event.	27
Figure 3.2.	Isolated power system equivalent generator transfer function model.	29
Figure 3.3.	A sample of excitation signal, frequency, and ROCOF measurements for $M = 2s$ and $\Delta P_L = 10^{-3}$ p.u.	31
Figure 3.4.	General architecture of a multilayer perceptron network.	32
Figure 3.5.	General architecture of a 1-D convolutional neural network.	34
Figure 3.6.	Schematic of inertia estimation using CNN.	34

Figure 3.7.	Change in frequency for different values of damping constant [71].	38
Figure 3.8.	Frequency sensitivity with respect to damping constant [71].	38
Figure 3.9.	Change in frequency for changing values of M and D with a constant ΔP_L . Here, $M = 1\text{s}, 5\text{s}, \text{ and } 9\text{s}$ whereas $D = [0.5, 1.0, 1.5, 2.0, 2.5]$. The sampling time for the snapshot is taken as 0.02 s.	39
Figure 3.10.	ROCOF for changing values of M and D with a constant ΔP_L . Here, $M = 1\text{s}, 5\text{s}, \text{ and } 9\text{s}$ whereas $D = [0.5, 1.0, 1.5, 2.0, 2.5]$. The sampling time for the snapshot is taken as 0.02 s.	40
Figure 3.11.	MSE loss (upper) and natural logarithm of MSE loss (lower) of inertia estimates with the number of epochs for MLP. Here, natural logarithm of MSE loss will amplify the values of MSE that are in close proximity of 0 for better visibility and analysis.	43
Figure 3.12.	MSE loss (upper) and natural logarithm of MSE loss (lower) of inertia estimates with the number of epochs for CNN.	44
Figure 3.13.	Evolution of training weights from the hidden layer to output layer of CNN.	45
Figure 3.14.	Model accuracy on the inertia estimates from CNN. The estimates within 10% of the actual values are considered to be accurate.	45
Figure 3.15.	Inertia estimates (in seconds) on the validation set after completing the training process.	46
Figure 4.1.	Basic electrical structure of bulk power system.	47
Figure 4.2.	Two-area power system. The generator at each of the areas represent an equivalent generator of multiple generators serving an equivalent load at that bus.	50
Figure 4.3.	Transfer function-based model of a two-area system	52

Figure 4.4. A sample of excitation signal, frequency, and ROCOF measurements at area 1 when $M_1 = 9s$, $M_2 = 5s$, $\Delta P_{L1} = 2 \times 10^{-3}$ p.u., and $\Delta P_{L2} = 0$ p.u. Only the area in which the frequency snapshots are to be collected is perturbed via excitation signal. In this case, area 1 is perturbed via $\Delta P_{L1} = 2 \times 10^{-3}$ p.u. whereas $\Delta P_{L2} = 0$ p.u.	53
Figure 4.5. General example of MNIST IID (left) and non-IID (right) dataset. IID data are identically distributed i.e. the probability distribution of all of the random samples (digits) are equal. On the other hand, for non-IID case the probability distribution of few samples dominate the others. . .	54
Figure 4.6. FL-based framework for inertia estimation in multi-area power system. The cloud storage is a non-volatile storage device that stores the aggregated CNN model and information about selected clients.	57
Figure 4.7. Overall schematic of FL-based inertia estimation in multi-area power system. It is to be noted that only one of the areas is perturbed at a time via ΔP_L to collect the snapshots.	60
Figure 4.8. Evolution of aggregated weight, w_{t+1} , using federated averaging for each communication round i between the server and ESS clients for MLP-based training(left) and CNN-based training(right). The plots represent the IID case when $\mathbb{E} = 1$	63
Figure 4.9. Aggregated model accuracy on validation set for each communication round i between the server and ESS clients for MLP-based training(left) and CNN-based training(right). The plots represent the IID case when $\mathbb{E} = 1$	64
Figure 4.10. Evolution of aggregated weight, w_{t+1} , using federated averaging for each communication round i between the server and ESS clients for MLP-based training(left) and CNN-based training(right). The plots represent the non-IID case when $\mathbb{E} = 1$	65

Figure 4.11. Aggregated model accuracy on validation set for each communication round i between the server and ESS clients for MLP-based training(left) and CNN-based training(right). The plots represent the non-IID case when $\mathbb{E} = 1$	65
Figure 4.12. Comparison of accuracy for IID and non-IID data when $i = 40$ and $\mathbb{E} = 5$	66
Figure 4.13. Validation accuracy of a CNN-based model on IID data with respect to i for different values of \mathbb{E} . On varying the values of \mathbb{E} , the number of communication rounds can be controlled in FL.	67
Figure 5.1. Frequency response of a power system with VI support from additional non-synchronous unit.[6].	69
Figure 5.2. Transfer function-based single area power with additional MPC-based VI unit	71
Figure 5.3. ΔP_m (upper) and $\Delta P_m + \Delta p_{vi}$ (lower) observed as a response to an excitation signal, ΔP_L , in the form of a pulse train with amplitude of 0.05 p.u. and time period of 5 s fed to a system with $M = 5$ s.	72
Figure 5.4. Schematic of inertia estimation in an MPC-based power system using pre-trained CNN. The CNN has been pre-trained on a power system with 1700 frequency snapshots corresponding to different values of M and ΔP_L	73
Figure 5.5. $\Delta\omega$ and $\Delta\dot{\omega}$ response for system with and without MPC-based VI unit. For the system with VI unit, $M = 5$ s whereas for the system without VI unit $M = M_x = 6.3708$ s. ΔP_L is used as a pulse train signal of amplitude 0.05 p.u.	77

Figure 5.6. $\Delta\omega$ and $\Delta\dot{\omega}$ response for system with and without MPC-based VI unit. For the system with VI unit, $M = 5$ s whereas for the system without VI unit $M = M_x = 6.3708$ s. ΔP_L is used as a step change signal of amplitude 0.05 p.u. introduced at 5 s.	78
Figure 5.7. Frequency responses for different values of M on a single area syn- chronous unit-based system compared with a similar system with an additional non-synchronous unit. Here, $M_{vi} = M_x = 6.3708$ s repre- sents the equivalent inertia constant of the system with additional VI unit estimated by a pre-trained CNN. ΔP_L is used as a step change signal of amplitude 0.05 p.u. introduced at 5 s.	78
Figure 5.8. Evolution of weights (left) and accuracy (right) of a CNN model trained for 200 epochs on frequency snapshots of single area system with sam- pling rate of 50 Hz.	80
Figure 5.9. Distribution of equivalent inertia estimates (M_x) estimated by a pre- trained CNN model. The estimated values are for 1700 snapshots trained on an equivalent single area system with MPC-based VI unit. .	80

ABSTRACT

DISTRIBUTED MACHINE LEARNING APPROACH TO FAST FREQUENCY RESPONSE-BASED INERTIA ESTIMATION IN LOW INERTIA GRIDS

ABODH POUDYAL

2020

Recent updates to the IEEE 1547-2018 standard allow active participation of distributed energy resources (DERs) in power grid services with the goal of increased grid reliability and resiliency. With the rapid growth of DERs towards a low inertia converter-dominated grid, the DERs can provide fast frequency response (FFR) services that can quickly counteract the change in system frequency through inertial support. However, in low voltage grids, frequency and voltage face dynamics coupling due to a high resistance to reactance ratio and cannot be controlled separately as in the bulk electric grid. Due to the coupling effect, the control of one parameter also affects the dynamics of the other parameter. A part of this work highlights the role of DERs to provide grid ancillary services underscoring the challenges of combined voltage and frequency control in low voltage grids.

Increasing penetration of renewable energy sources (RES) also decreases the power system inertia, there by affecting the stability of bulk grid. The stochastic nature of RES makes the power system inertia a time-varying quantity. Furthermore, converter-dominated grids have different dynamics compared to conventional grids and therefore estimates of the inertia constant using existing dynamic power system models are unsuitable. This work proposes a novel inertia estimation technique based on convolutional neural networks (CNN) that use local frequency measurements. The model uses a non-intrusive excitation signal to perturb the system and measure frequency using a phase-locked loop. The estimated inertia constants, have significant accuracy for the training, validation, and testing sets. Additionally, the proposed approach can be applied over traditional inertia estimation methods that do not incorporate the dynamic impact of

renewable energy sources.

The frequency response of power systems changes drastically when multi-area power systems with interconnected tie-lines are considered. Furthermore, higher penetration of RES increases the stochasticity in interconnected power systems. Hence, it is important to estimate the multi-area parameters ensuring communication and coordination between each of the areas. A robust and secure client-server-based distributed machine learning framework is used to estimate power system inertia in a two-area system. The proposed approach can be efficiently optimized to increase the training performance.

It is important to analyze the performance of a trained machine learning model in a real-world scenario with unknown dynamics. A pre-trained CNN is tested on a system with model predictive controller (MPC)-based virtual inertia (VI) unit. Results show that the frequency and inertial response of conventional synchronous generators-based system differs drastically as compared to the system with non-synchronous generator-based VI support.

CHAPTER 1 INTRODUCTION

1.1 Background

Considering the environmental impacts and unsustainability of non-renewable electric energy resources, the world is transitioning towards a renewable-dominated electric power system. Many countries are setting renewable energy source (RES) goals, some of which are closing in on 100% — Norway (97.9%), Brazil (82.5%), Colombia (75.7%), and Canada (65.9%) [1]. However, with the increase in inverter-based resources, such as solar photovoltaics (PV), wind, and energy storage systems (ESS), the equivalent inertia of the power system decreases [2], [3].

RESs have also been introduced in the low voltage (LV) side with the concept of DERs. Some of the common DERs are shown in Fig. 1.1.

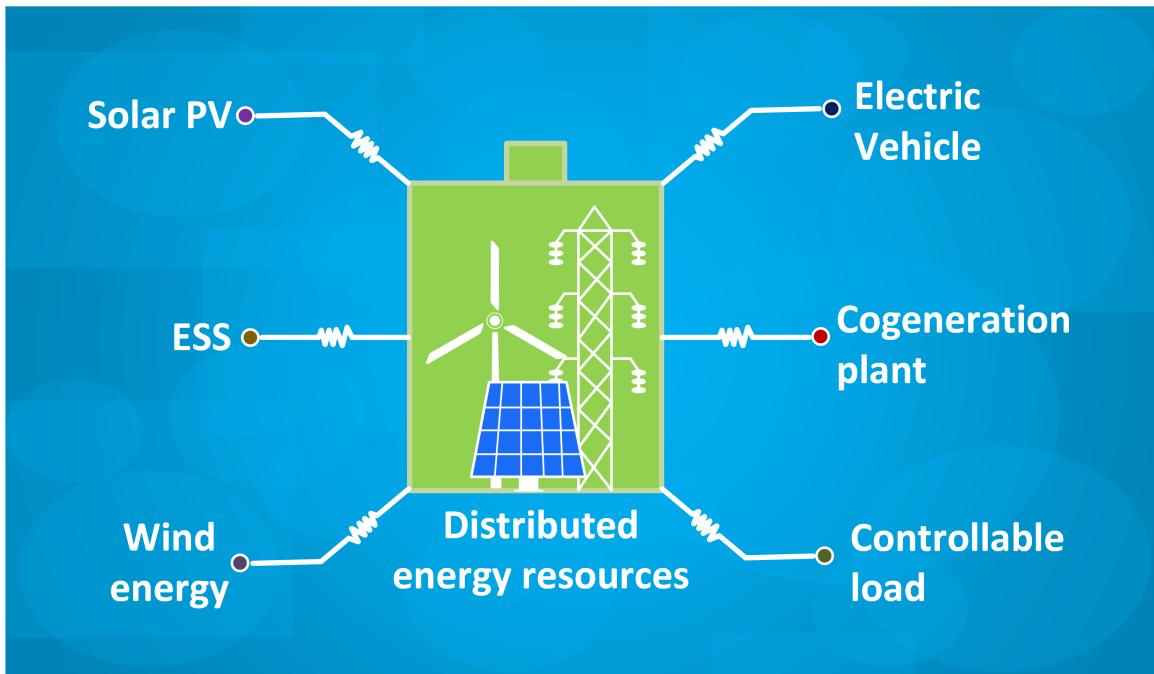


Figure 1.1. Distributed energy resources.

Among all of the DERs, distributed solar PV has shown a rapid year-by-year installation growth. Fig. 1.2 illustrates the current and future capacity of distributed solar PV for residential, non-residential, and utility sector of the US. The residential PV has increased

by 8% since the second quarter of 2018 and is expected to rise upto 19% by 2021 [4].

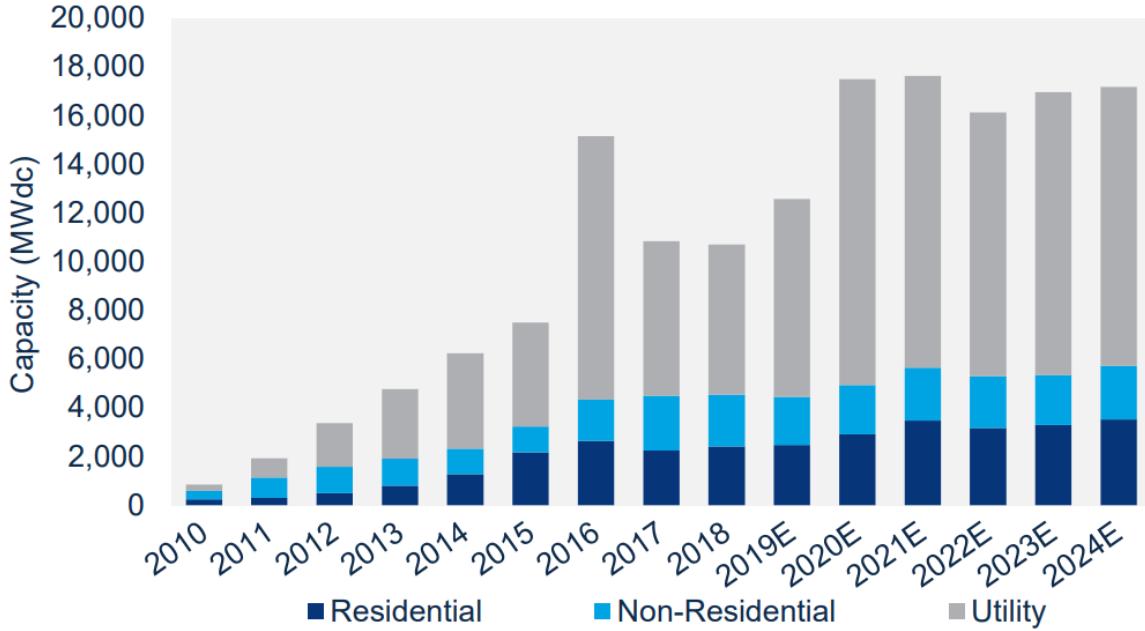


Figure 1.2. Installed capacity and forecast of distributed solar PV (in MW_{dc}) per year in the US [4].

RESs are connected to the grid via power electronic converters. Therefore, with the integration of large percentages of RES, the grid is becoming converter-dominated (i.e., instantaneous penetration of RES in the grid is more than 50% [5]). Fig. 1.3 shows the transition of conventional synchronous generators-based grid to a non-synchronous inverter-based grid.

The inertial response (absorption or release of instantaneous power opposing the power imbalance) of a system with a high number of synchronous units is a function of the equivalent inertia constant of the connected units at different parts of the system. This makes the inertia constant dependent on the number and size of actively connected synchronous units at any given time. Additionally, RES and other inverter-based resources are considered passive in terms of inertial response as they do not possess any rotary components to provide the necessary response. With the increasing penetration of RES, the number of conventional synchronous generators in the grid reduces. As an effect of instantaneous mismatch in generation and load, and without adequate inertial support, the

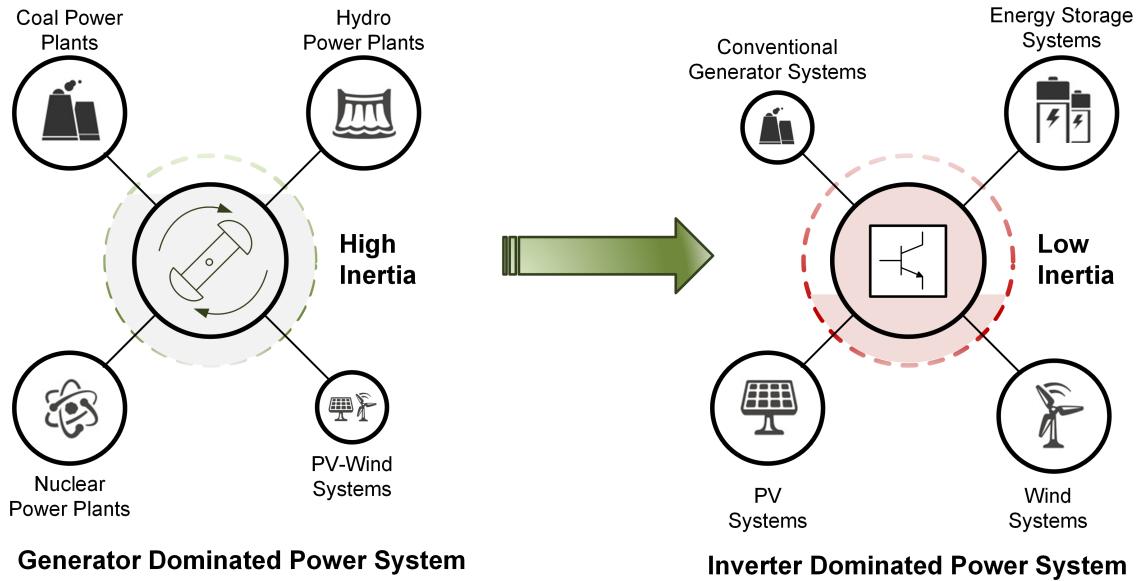


Figure 1.3. Transition of power grid from conventional synchronous generator-dominated system to inverter-dominated system [6].

power system can experience large frequency deviations and rate-of-change-of-frequency (ROCOF) that can lead to under-frequency load shedding (UFLS) and cascading failures [7]. In the past, power system stability issues were solved using a combination of the physical properties and control mechanisms of the synchronous generators. Unlike the conventional synchronous generator-dominated system, converter-dominated systems are unable to provide inertial support to the grid that is part of the system's FFR .

Research has shown that the reduction of system inertia raises a serious concern on the frequency stability of a power grid. In a technical report published by Electricity of Reliability Council of Texas (ERCOT) , one of the independent system operators (ISOs) of the United States, a same outage in two different time periods induced a change in frequency of varying nature in the system [8]. Figure 1.4 shows the change in frequency for a similar generation loss at two different time period at ERCOT. In the technical report, ERCOT has also stated that such change in the response for a similar kind of outage is due to decreasing inertial response which is directly related to the increasing penetration of RES in the grid.

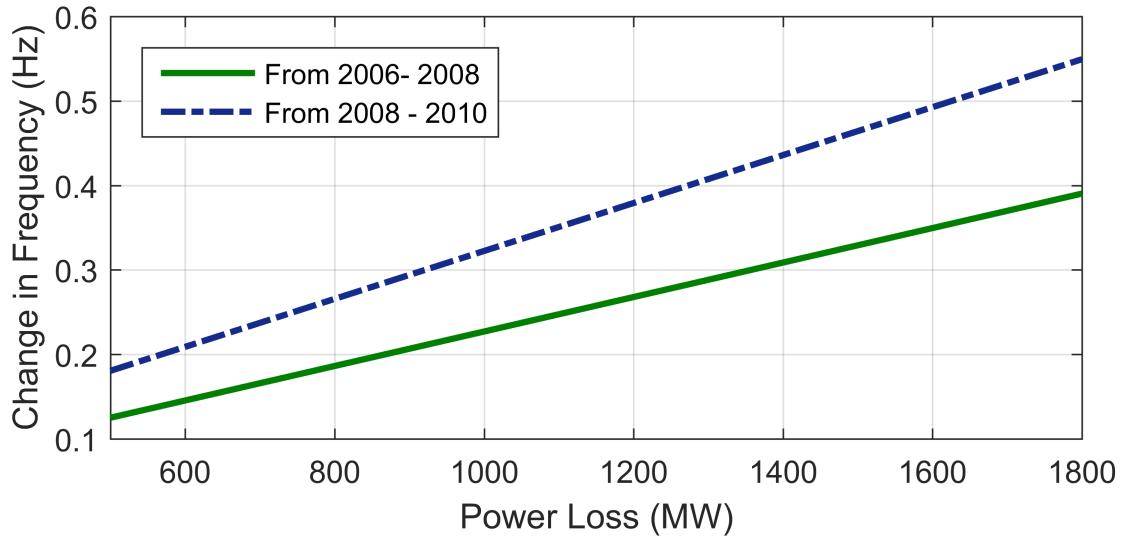


Figure 1.4. Change in frequency versus generation loss for different period of time at ERCOT [8].

On August 9, 2019 Great Britain faced a blackout following a 5% loss in generation [9]. The major reason for the blackout is considered to be low system inertia due to high penetration of RES. The blackout affected around 1 million people for approximately an hour. Fig. 1.5 (upper) shows the frequency deviation due to loss of generation in the National Grid of the Great Britain. It can be seen that, when the frequency started to recover, following an outage of a combined cycle gas turbine of capacity 680 MW, an additional 750 MW of wind farm tripped causing UFLS due to the tripping of frequency relays. The cause of low inertia for this outage is further supplemented by Fig. 1.5 (lower) which shows the yearly record of percentage of total generation. It can be seen that at the time of the outage, the total percentage of generation from synchronous unit decreased below 50% whereas the penetration of wind energy, RES, increased over 25%. Hence, the penetration of RES is highly responsible for decreasing system inertia, and thereby increasing the risk of possible high impact outages.

Recent advancements in control techniques have shown that RES can also provide inertial response to the system; to replace the lost system inertia, FFR through converter-based VI was designed to reduce the ROCOF of the power system [6],

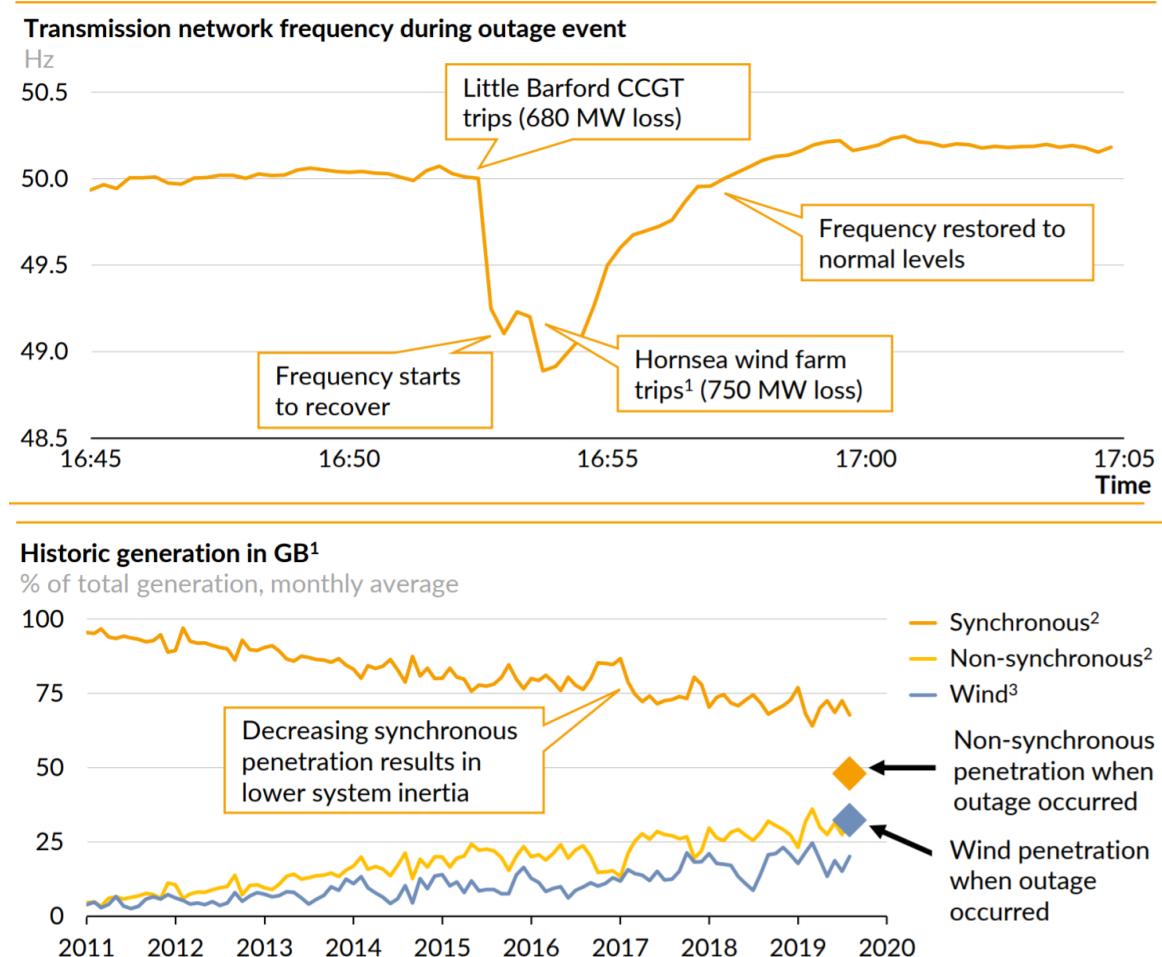


Figure 1.5. Triggering of UFLS at the National Grid of the Great Britain due to high penetration of RES [9]. The upper figure shows the frequency response of the grid following an event whereas the lower figure represents the generation mix of different resources by year.

[10]–[15]. However, the stochastic nature of RES creates unknowns in the time-varying system inertia constant [8]. Although RES can provide the inertial response in the form of VI, the time varying nature of inertia constant will make it difficult to identify the current status of inertia in the system. For instance, the system operators will not be able to dispatch or have proper plan to dispatch the required amount of FFR support in case of an outage or possible outage. Therefore, the time-varying inertia constant must be estimated for grid stability analysis. Furthermore, inertia estimation can have the following benefits:

1. robust grid frequency control

2. controlled penetration of renewable energy
3. improved power system reliability through fast frequency response analysis and procurement
4. better market design for grid ancillary services

1.2 Objective

The primary objective of this thesis was to propose a model-free data-driven distributed estimation technique to identify the power system inertia in order to facilitate the fast-frequency control of low inertia power systems.

1.3 Contribution

The major contributions of this thesis are:

- a A review on the challenges of combined voltage and frequency control in LV grids using power electronic converters.
- b A CNN-based method to estimate power system inertia in a single area system using local frequency measurements from a phase-locked-loop (PLL) of an ESS.
- c A federated learning approach to estimate the power system inertia in multi-area system with independently and identically distributed (IID) and non-IID data.
- d An approach towards power system inertia estimation for a system with additional non-synchronous unit for FFR support.

1.4 Thesis Outline

This thesis is organized as follows. Chapter 2 describes the challenges of combined voltage and frequency control in a low voltage grid with power electronic converters. The problems due to dynamics cross-coupling effects are discussed, and opportunities for new market design are provided. Chapter 3 presents a CNN-based inertia estimation technique

using local frequency measurements. Details on the power system frequency dynamics, CNN formulation, and system perturbation using excitation signals are presented in this Chapter. Chapter 4 describes a client-server-based FL framework for multi-area inertia estimation. Similarly, the effect of varying communication rounds and local training epochs is also discussed. Inertia estimation on a system with additional non-synchronous unit is discussed in Chapter 5. Chapter 6 presents the concluding remarks of the thesis.

CHAPTER 2 CHALLENGES OF COMBINED VOLTAGE AND FREQUENCY CONTROL IN LOW VOLTAGE GRIDS USING POWER ELECTRONIC CONVERTERS

2.1 Background

Power system stability issues are not usually considered at the low voltage (LV) level, but the proliferation of DERs raises a question on converter interactions and grid stability. Conventionally, DERs were treated as negative loads and were not allowed to participate in market and grid operations [16]. When encountering abnormal grid conditions, DERs were mandated to disconnect from the grid. As more DERs are installed, such disconnections led to more adverse grid conditions [17]. These conditions led to the update of the IEEE 1547-2018 standard, which allows DERs to participate in grid ancillary services during normal and transient conditions [18].

Additionally, it is necessary to incentivize DERs for providing grid ancillary services. Following the standards of Federal Energy Regulatory Commission (FERC) Orders No. 755, 784, 841, and 890, PJM, MISO, and CAISO became the first U.S. ISOs to introduce ESS to regulate system frequency [19], [20]. With several other DERs being connected to the LV grid, there are opportunities to transform existing and design new U.S. electricity markets. Therefore in LV grids it is necessary to simultaneously control the voltage and the frequency for grid ancillary services.

In LV grids, active and reactive power injection/absorption are not completely decoupled [21], [22]. The case is significantly different from the conventional power systems, where the voltage is controlled through the excitation system of the generator and the speed governor is responsible for frequency control [23]. Hence, the control is decoupled where the voltage is controlled using reactive power and frequency is controlled using active power. However, such decoupled control is not possible in an inverter-based system, particularly in LV grids. The increased installation of DERs in LV

grids can increase feeder voltage during high injection and low load times [24], [25]. The problem is further exacerbated when active power is modulated to control frequency which can severely impact the grid voltage in LV grids. It is also known multi-timescale control is desired for voltage and frequency, as they have multi-timescale dynamics [26]. Simultaneous control can affect the dynamics of one parameter while controlling the other. Therefore, it is necessary to develop a combined voltage and frequency control strategy in the LV system considering their multi-timescale dynamics to maintain overall system stability and power quality.

2.2 Related Works

Existing literature reviews on voltage and frequency control are more focused on microgrids rather than large-scale integration of DERs in the distribution system. With the recent amendment to the IEEE 1547 standard, the direction of research on combined voltage and frequency control in LV grids with the problem of coupling effects has changed drastically. Voltage and frequency control of a microgrid using plug-in hybrid electric vehicles (PHEV) as the energy medium was presented in [27]. The authors aggregated PHEV, independent of network parameters, and controlled the voltage and frequency at the point of common coupling (PCC) of a microgrid. The authors in [28] proposed an optimal frequency response algorithm using distributed demand response. The authors adhered to the constraint of frequency control through a limited number of DERs as imposed by the Belgian federation of electrical and gas network operators. In [29], a decentralized control framework was presented that manages voltage issues due to the regulation of frequency through active power injection. An active power sharing module tracks the active power change through coordination among the DERs, and a voltage regulation module regulates the bus voltage within specified limits through reactive power control based on each DER. The authors in [12] presented frequency control in an isolated microgrid using a voltage regulation approach. Such control mechanisms only require a local signal that avoids communication overhead and can

facilitate fast response.

Furthermore, from the impedance ratio point of view, several other control strategies were proposed to alleviate the problem of coupling in LV grids. Virtual impedance was introduced in [30], [31] using a feed-forward controller to emulate a resistive or inductive output impedance in LV grids. The addition of virtual impedance in the control loop changed the output impedance of the converter to meet the desired resistance to reactance ratio and the validity of conventional droop control strategies was justified. A control scheme to decouple power in LV grids was proposed in [32]. Although LV grids behave differently than HV grids, restricting the impedance range of the converter can decouple active and reactive power, however a change in impedance might result in further instability in the converter's voltage. Hence, with the increasing penetration of RES, the control of voltage and frequency is no longer decoupled. Hence, it required to build a control strategy that incorporates the coupling effect as well as the multi-time scale dynamics of voltage and frequency.

2.3 Line Parameter Impact on Grid Operation

2.3.1 High Voltage Transmission Grid

Fig. 2.1 shows a general transmission model of a grid. Here, P_s and P_r are the sending and receiving end active power, Q_s and Q_r are the sending and receiving end reactive power, S_s and S_r are the sending and receiving end apparent power, V_s and V_r are the sending and receiving end voltage, respectively. Further, δ is the power angle, ϕ is the phase angle, I_s is the sending end current, and R and X are the resistance and reactance of the transmission line, respectively. Assuming the power flow from the sending end to the receiving end, considering V_r as the reference voltage, the total apparent power injected from the sending end is given by:

$$S_s = P_s + jQ_s = V_s I_s^* \quad (2.1)$$

where I_s^* is the complex conjugate of the sending end voltage. The impedance of the generator is neglected here for the sake of simplicity. After some calculation, the

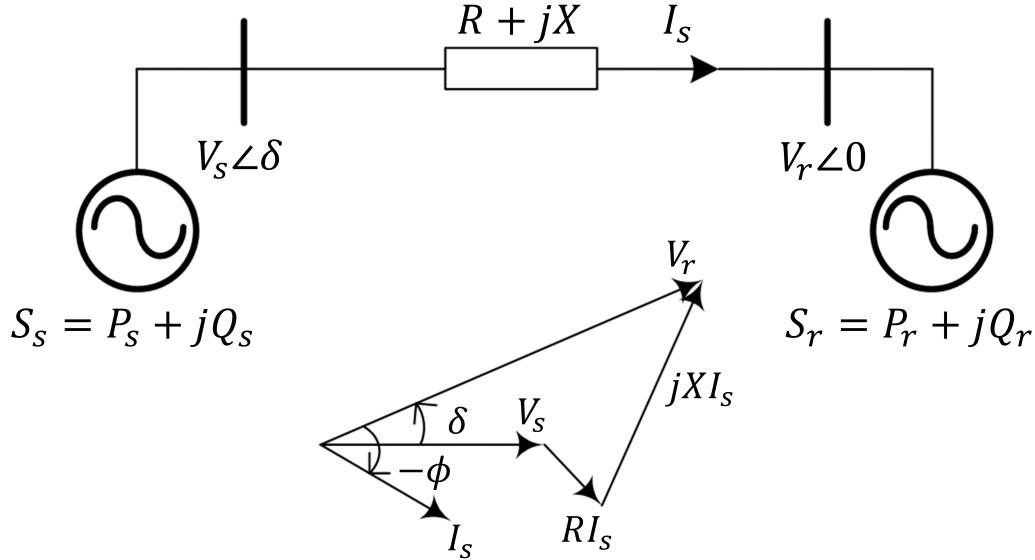


Figure 2.1. General transmission model of a grid [33].

equations for P_s and Q_s are given by:

$$P_s = \frac{V_s[R(V_s - V_r \cos\delta) + X V_r \sin\delta]}{R^2 + X^2} \quad (2.2)$$

$$Q_s = \frac{V_s[-R(V_r \sin\delta) + X(V_s - V_r \cos\delta)]}{R^2 + X^2}. \quad (2.3)$$

General transmission line parameters for high voltage (HV), medium voltage (MV), and LV grids are shown in Table 2.1. In HV grids, the transmission line reactance is significantly larger than the resistance (i.e., low R/X ratio), so the effect of resistance can be excluded to approximate P_s and Q_s . Considering a small power angle ($\sin\delta \approx \delta$ and $\cos\delta \approx 1$), P_s and Q_s can be reduced to the following equations:

$$P_s \approx \frac{V_s V_r}{X} \delta \quad (2.4)$$

$$Q_s \approx \frac{V_s(V_s - V_r)}{X}. \quad (2.5)$$

For HV grids with small δ and low R/X ratio, P_s controls δ , which is directly related to the system's frequency, and Q_s controls the generator terminal voltage. Based on these relations, P - f and Q - V droop control techniques were proposed for HV transmission networks [21].

Table 2.1. General transmission line parameters [21]

Type of network	$R(\Omega/km)$	$X(\Omega/km)$	R/X
HV	0.060	0.191	0.31
MV	0.161	0.190	0.85
LV	0.642	0.083	7.7

2.3.2 Low Voltage Distribution Grid

LV grids have a significantly larger R/X ratio than HV grids due to the resistive nature of the network. The diameter of the conductors for LV distribution are smaller than those used in HV transmission; X increases with geometric mean radius of the conductor and R increases with the decreasing size of the conductor. Neglecting X and considering a small δ for LV grids, (2.2) and (2.3) can be approximated as:

$$P_s \approx \frac{V_s(V_s - V_r)}{R} \quad (2.6)$$

$$Q_s \approx -\frac{V_s V_r}{R} \delta. \quad (2.7)$$

For LV grids with small δ and large R/X ratio, P_s controls the terminal voltage of the grid and Q_s controls the frequency. In the LV grids, decoupling P and Q is challenging and in some cases conventional P - f and Q - V droop control strategies are infeasible. Although partially non-linear P - V and Q - f droop control strategies were proposed in [34], the coupling can affect the transient performance of the grid and can result in inaccurate

power sharing due to incongruous line impedances [35].

2.4 Grid Codes for Low Voltage Grids

IEEE 1547 is the standard used for the interconnection of DERs to the U.S. power grid. Recent changes in the IEEE standard mandated a number of ancillary services from grid-connected inverters. The first iteration of the standard for the interconnection of DERs (IEEE 1547-2003) had no provisions for grid support functions — in the case of a frequency/voltage event the DER units had to mandatorily disconnect. As the standard evolved to IEEE 1547-2018, due the growing trend in converter-based renewable generation, DER installations are required to ride-through faults in the system. In addition, dynamic voltage and frequency response are required from DER inverters to maintain stability.

Based on the new standard, DER units should be capable of providing grid support functions, such as constant reactive power mode and voltage-reactive power (volt/VAr) mode, where the output reactive power is regulated based on voltage of the PCC. Apart from these basic functions, when the DER penetration is significant other advanced functions such as voltage-active power and active-reactive power mode are also desired. Dynamic frequency support functions are mandated from the DER units. A droop-based frequency-Watt ($f\text{-}W$) frequency control, where the active power of the DER unit is controlled based on the operating frequency, is required from the inverters to reduce frequency variations due to the stochastic nature of RES. The standard also allows inertial response from the inverters, where the active power is regulated based on the ROCOF.

IEEE 1547-2018 specifies the performance requirements from DERs, but does not provide guidelines for the design and operation of these units equipped with advanced grid-support functions. Considering most DER units are deployed in LV grids, the voltage-frequency coupling will impact how different grid functions behave or interact. The controllers designed to deploy these grid-support functions must consider the fact that control of one variable impacts the other (e.g., providing inertial response will also impact

the PCC voltage).

2.5 Fast Frequency Response using Distributed Energy Resources

FFR is the self-deployed and controlled contribution of electrical power from a unit to quickly counteract the initial frequency change in the system as a result of reduced inertial response [36]. The FFR from DERs can be in the form of VI that limits the ROCOF and frequency nadir, thereby providing the necessary frequency response to facilitate grid stability. In case of HV grids, the North American Electric Reliability Corporation (NERC) set standard BAL-003-1 that limits a frequency deadband of ± 36 mHz [37] during the steady state operation. However, during transients it would be difficult to maintain frequency within the given band especially in systems with high DER penetration. Therefore, with low-inertia systems, it is essential to install FFR units to impede the ROCOF to maintain frequency within limits. The integrated DER units must provide FFR within a fraction of seconds after a frequency event has occurred, illustrated in Fig. 2.2.

Deloading of distributed solar PV, vehicle-to-grid (V2G) operation of PHEV, ESS, and demand response can also be viable options for FFR [39]. With the burgeoning number of installed DERs, their FFR capabilities should be properly utilized. U.S. ISOs have sparse visibility and control over the distribution system, including DERs. The concept of Distribution System Operators (DSOs) have been implemented in Ireland, and are being discussed in NYISO and CAISO [16]. It will be necessary to have a coordinated control mechanism on DERs through DSOs or other distributed control approaches, highlighting the opportunity for new robust market design options for DER ancillary services.

Frequency response is provided by changing active power injection. The coupling nature of LV grids, however, also changes the terminal voltage due to the change in active power. A small mismatch in voltage can combine with the large impedances provided by the distribution line resulting in large circulating reactive current in the inverter [40].

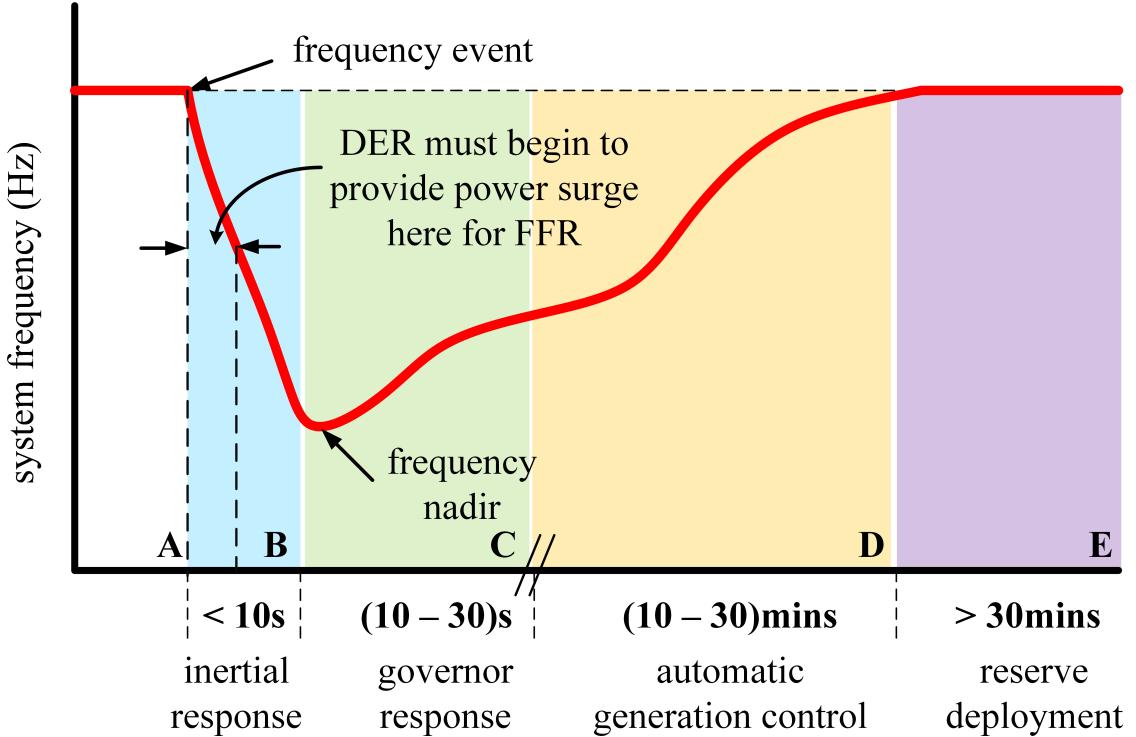


Figure 2.2. Different stages of frequency control [38].

Furthermore, uncoordinated frequency response from the DERs cause congestion and voltage issues in the distribution grid. Considering this situation, a new regulation in Belgium allows only ten connection points within a radius of 100 m to provide frequency control at any time [28]. This highlights the need for a coordination mechanism for voltage and frequency control of LV grids using DERs.

2.6 Related Works on Recent IEEE Grid Standards

NERC defines DER aggregation (DERA) as a virtual resource formed by aggregating multiple DERs at the different locations of a distribution system [41]. Because this is an aggregated model, the resistance of the source impedance is neglected and the reactance is set to a large value. The voltage is controlled using a volt/VAr mode, and the frequency is controlled using f -W droop, both with an asymmetric deadband [42]. Spatial variation of DERs and coordinating different bus voltage levels have not been considered in the DERA models. In a system with a large number of DERs, it is

unsuitable to observe system stability through the aggregation of DERs as the aggregated model does not represent the actual scenario of DER interconnection. In a system with large number of distributed connections, a decentralized strategy to simultaneously control the PCC voltage and frequency is desired. Fig. 2.3 illustrates the power flow representation including DERA from the bulk grid point of view.

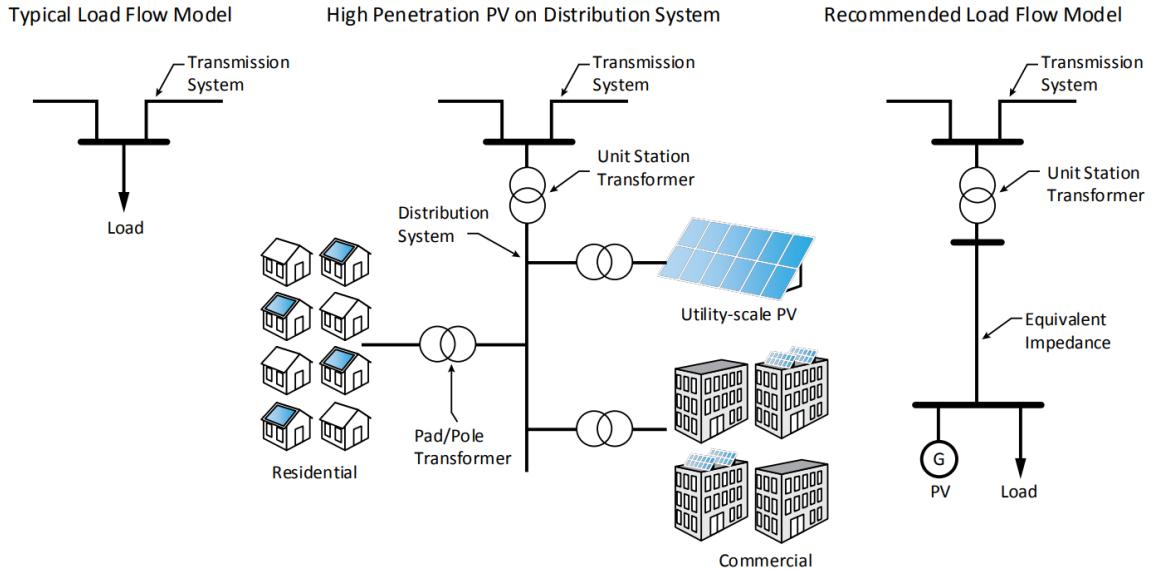


Figure 2.3. Power flow in a grid by aggregating DERs [41]

A team including the National Renewable Energy Lab and the Sandia National Laboratories developed a real-time hybrid power hardware-in-the-loop inverter model that captures both the distribution and bulk grid dynamics [38]. The inverter was designed for Hawaiian grid that contains a relatively high number of distributed solar PVs. The designed PV inverter can address dynamic and steady-state voltage and frequency issues. The voltage is controlled using both the volt/VAr and the volt/Watt methods. The volt/Watt control is used when the voltage crosses the ANSI C84.1 voltage limit of 1.05 pu, and volt/VAr control mode is used when the voltage significantly deviates from its nominal value [43]. Overfrequency events are controlled using a f -W droop as illustrated in Fig. 2.4 with a deadband frequency of ± 36 mHz (db_{of}) and droop slope of 0.04 pu/Hz (k_{of}) as defined in NERC BAL-003-1.

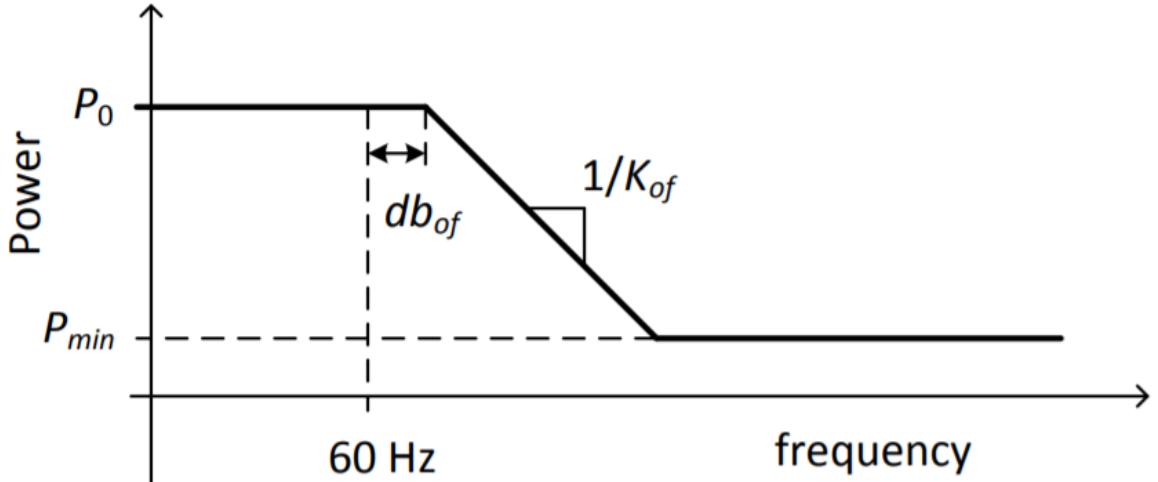


Figure 2.4. Frequency-watt droop control in LV grids [38]

Due to the lack of communication signals and finite response time between spatially distributed inverters, it is challenging to quantify the exact response time. Adding a communication system allows control of the f - W droop remotely [38]. In that study, only downward response (overfrequency event) is considered as most of the PVs operate at their maximum power point. To respond to both overfrequency and underfrequency events, it may be desired to deload the solar PV or in coordination with an ESS. The work considered the separate effect of voltage and stability control on PV energy production, however the effect of changing one parameter due to the control on the other in LV grids has not been discussed.

Similarly, a team at Argonne National Lab developed a dynamic solar PV model for distribution systems in the phasor domain [44] that captures the dynamics of distributed PV. The work adheres to the IEEE 1547-2018 standard and implements volt-VAR, volt-Watt, and voltage/frequency ride through. However, the authors do not discuss frequency control mechanisms in addition to voltage control issues, nor the coupling issues that may arise in LV grids.

Although volt/Watt and volt/VAr controls are not mandatory in IEEE 1547-2018, they will have a significant presence in future DER integration [11], [22] and need to

consider P and Q coupling in LV grids, including limitations on the hosting capacity of the distribution grid. Moreover, aggregated DER models were considered that are unsuitable for large networks where it is essential to involve each of the DERs in the control loop. Some of the literature above consider the coupling effect in LV grids, but fail to address the recent changes in the grid standards. Therefore, it is necessary to establish a combined voltage and frequency control strategy to integrate spatially distributed DERs in LV grids to provide grid ancillary services.

2.7 Combined Voltage and Frequency Control

This section presents an analysis on the challenge of combined voltage and frequency control in LV grids. Due to the coupling effect in LV grids, the voltage and the frequency controller design is completely different from that of a HV grid. The design is even more involved while controlling both voltage and frequency, as most of the control is achieved solely through active power modulation. Furthermore, the impact of coupling effect is large in low inertia grids, where the frequency deviation is large. Therefore, to analyze the effect of frequency dynamics on the grid voltage, it is assumed that the proposed benchmark demonstrates the characteristics of a LV low inertia grid.

2.7.1 Simulation Setup

Fig. 2.5 shows the benchmark used for simulation of combined voltage and frequency control in LV low inertia grids. The inverter is modeled as average current controlled voltage source while grid is modeled as thevenin equivalent voltage source of voltage v_g , resistance R and inductance L . Capacitor of capacitance C , which is a part of inverter filter is connected at PCC. Let ω be the system frequency, i_g be the current flowing to the grid, and v_c be the voltage at PCC. Let v_{cd} and v_{cq} represent direct axis (d) and quadrature axis (q) component of v_c respectively. Let i_{inv} be the inverter current and i_L be the load current. The modeling and simulation of the benchmark was conducted in

MATLAB/Simulink 2018b¹

The simulation parameters for the benchmark in Fig. 2.5 are given in Table 2.2. Here, $X = 2\pi fL$ is the inductive reactance of the transmission line in Ω . A new variable, K , is introduced as the R/X ratio. To perform multiple simulation studies, five different values of K are selected as [0.25, 0.5, 1.0, 2.0, 5.0], where the lowest value of 0.25 represents HV grid and the highest value of 5.0 represents a LV grid [25]. For a given value of K and X , R is then calculated as $R = XK$ for each simulation.

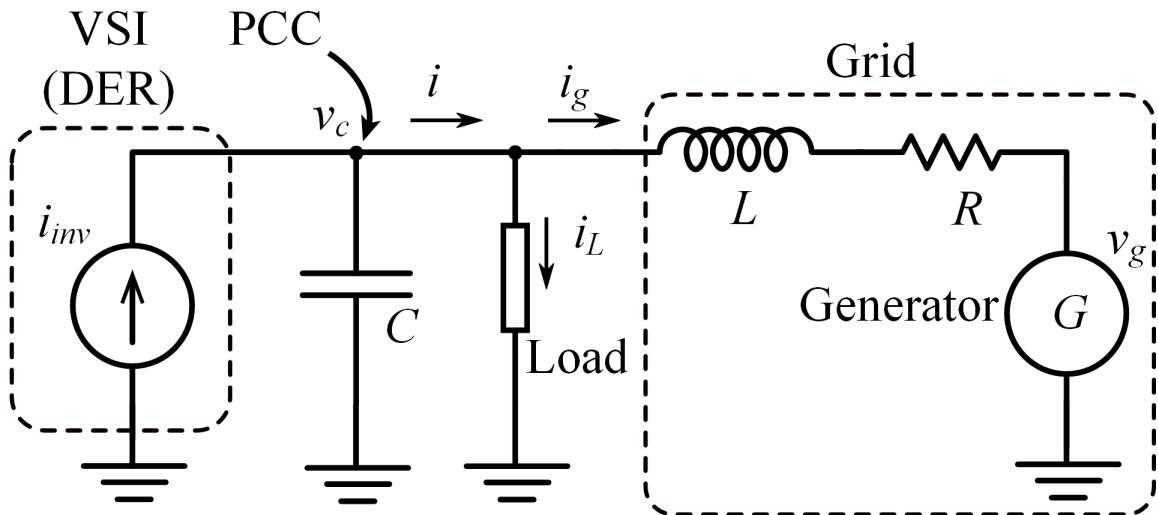


Figure 2.5. Simulation benchmark for combined voltage and frequency control.

Here, the active power of the inverter (ΔP_{inv}) is chosen arbitrarily to see the open-loop response of the system. At 0.5 s, ΔP_{inv} of magnitude 0.5 p.u. is manually injected to the system and the dynamics of the direct component of v_c , v_{cd} , for the corresponding value of K is observed. Only v_{cd} needs to be observed since we are taking v_c as reference and are neglecting the dynamics of PLL, which makes v_{cq} zero.

2.7.2 Results and Analysis

1. Case 1: Effect of active power on voltage dynamics:

It is desired to observe the relation of v_{cd} and change in active power, as in the case

¹The implementation and analysis is a joint work with Niranjan Bhujel, who is a graduate student at South Dakota State University as of July 2020.

Table 2.2. Parameters for Combined Voltage and Frequency Control

Parameters	Values
grid voltage (v_g)	208 V
grid frequency (ω)	60 Hz
reactance (X)	0.0433 Ω
capacitance (C)	220 μF
base power	100 kVA
base voltage (dq-frame)	170 V

of a HV grid v_{cd} is unaffected by the active power injected in the grid. Fig. 2.6 shows the voltage dynamics with different values of K . From the plot, we can see that after injecting ΔP_{inv} at 0.5 s, the change in voltage, Δv_{cd} , is insignificant for $K = 0.25$ which shows the case of a HV grid where active power has minimal effect on the voltage. On the other hand, when $K = 5.0$, Δv_{cd} reaches a value of 0.2 p.u. which means that when $\Delta P_{inv} = 0.5$ p.u. is injected to the system, the voltage changes by 0.2 p.u. for $K = 5.0$. This shows that v_{cd} can be modulated via ΔP_{inv} when K is high.

2. Case 2: Interaction of voltage and frequency dynamics:

Fig. 2.7 shows the frequency and voltage dynamics of the above system when an active load, ΔP_L , changes by 0.5 p.u. Here, the inverter control is not incorporated in the system and the change in active power is only due to the change in ΔP_L . To emulate low inertia grid, the inertia constant of G is set as 2 s. From Fig. 2.7(a) we can see that when ΔP_L changes by 0.5 p.u. at 0.5 s, frequency goes back to nominal value very slowly, which is governed by its primary and secondary controllers.

From Fig. 2.7(b), it can be seen that voltage initially has much faster dynamics, however the dynamics of voltage follows that of the frequency after approaching a certain value. Thus, the change in frequency also affects voltage dynamics. It is also interesting to see that the voltage shows both fast and slow dynamics. Therefore, in a low inertia system, frequency deviation is very large that can also significantly

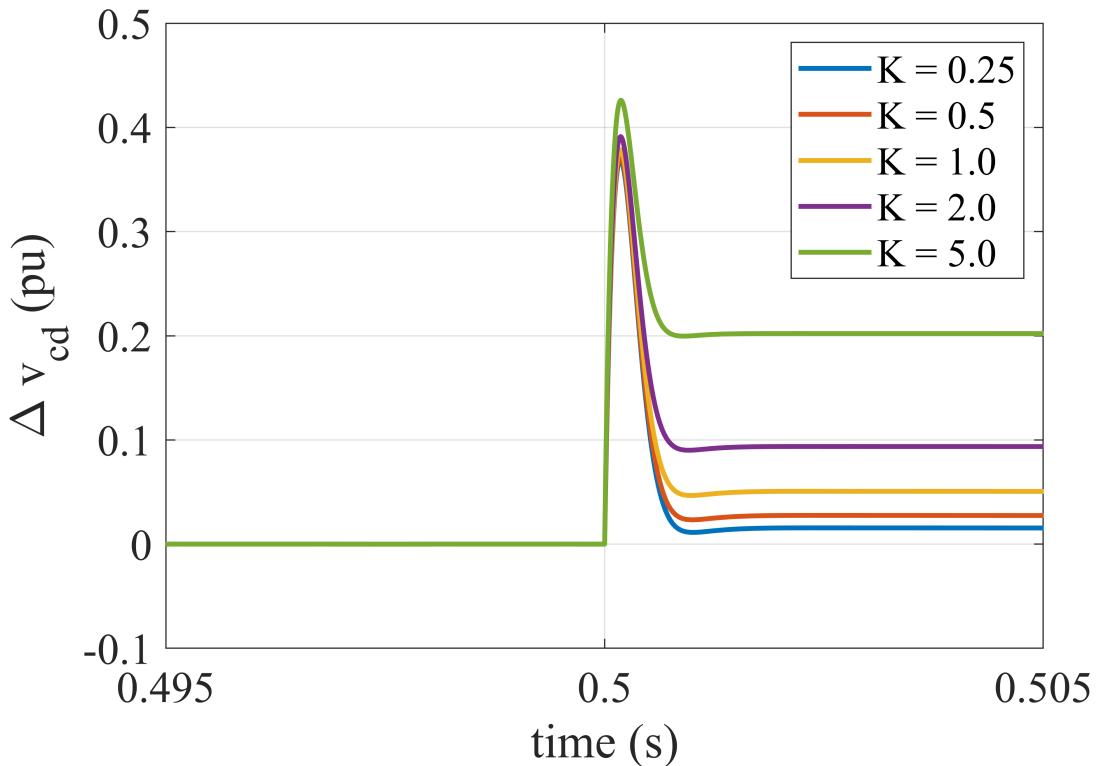


Figure 2.6. Effect of change in active power on voltage with varying R/X ratios.

impact voltage dynamics. Also, we can see that value of K has little impact on frequency dynamics but the effect is large on voltage dynamics. When K is large as in the case of LV grids, voltage deviates by large value.

From the above analysis, we can posit that in a LV low inertia power system, voltage is sensitive to active power change and the impact of frequency deviation is also seen in voltage dynamics. Thus when the objective is to control combined voltage and frequency, it is necessary to consider their coupling along with their multi-timescale dynamics unlike in a HV power system.

2.7.3 Market for Combined Voltage and Frequency Control

Envisioning the future voltage and frequency regulation ancillary market for LV brings many questions, such as: (i) how are participants going to be compensated?; (ii) are resources exclusive for voltage or frequency regulation?; and (iii) should the regulation be

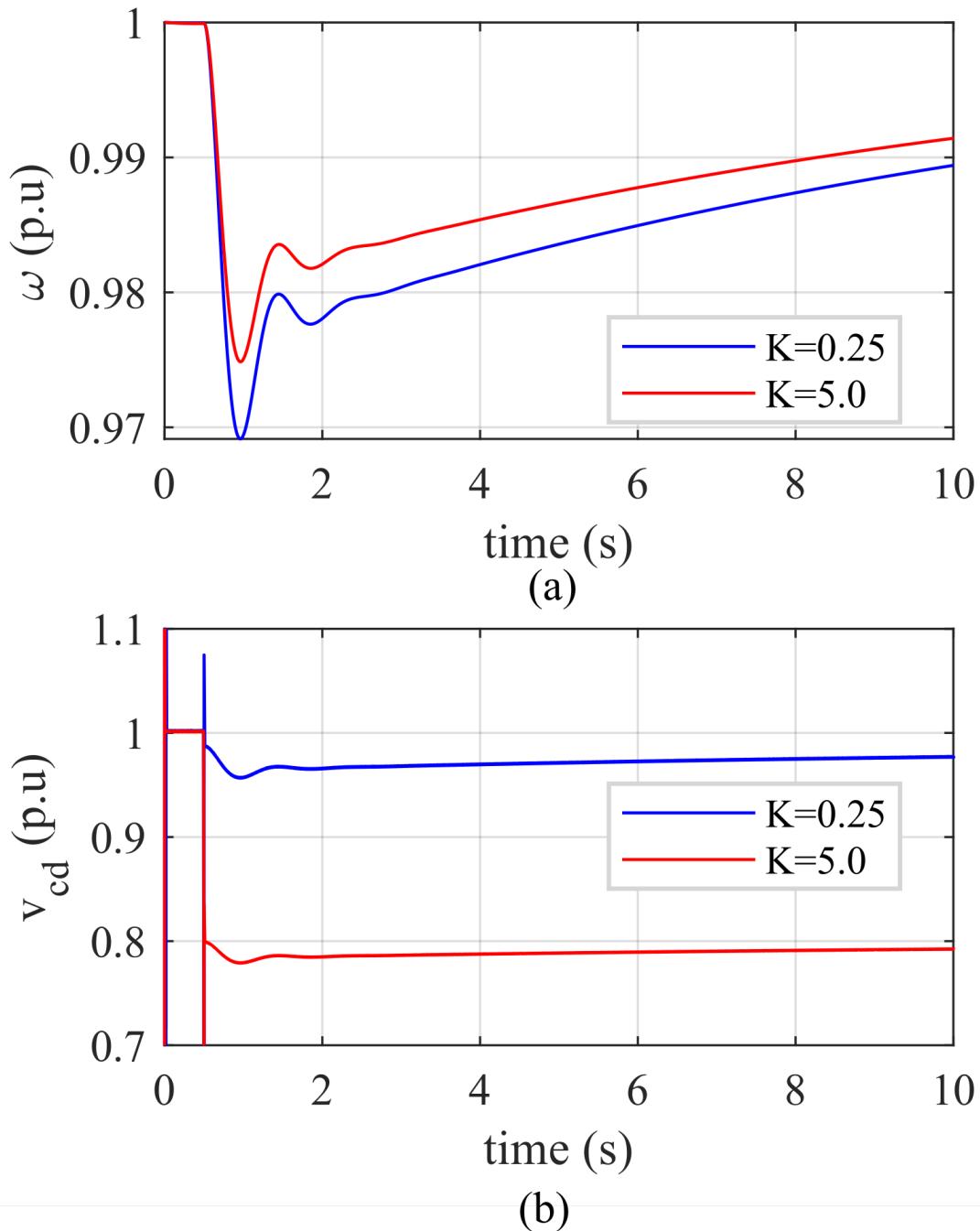


Figure 2.7. Interaction between voltage and frequency dynamics when inertia constant is 2 s.

centralized or distributed? A detailed discussion presented in [45] compared centralized versus distributed voltage and frequency regulation in LV grids. The distributed approach

can provide the same control with greater scalability and reliability, but frequency regulation requires a centralized market. The FFR in LV networks can be aggregated in clustered voltage regions (with voltage regulation considerations) to a centralized market. Furthermore, a hybrid approach would be capable of interacting with the existing markets while reducing the communication requirements and improving stability.

2.8 Conclusions

The evolution of power electronic converters and rapid growth of DER integration in the grid resulted in the amendment of IEEE 1547-2003 to IEEE 1547-2018 standard. The proliferation of DERs can increase grid vulnerability towards voltage and frequency fluctuations. In LV grids, it is necessary to have combined voltage and frequency control, as changing one parameter can create instability in the other. The future goal of modeling a LV grid converter should include a combined voltage and frequency control strategy addressing the IEEE 1547 standard and the coupling effect of voltage and frequency dynamics in LV grids. Additionally, DERs are anticipated to have a major role in grid resiliency based on their distributed nature, flexibility in converter control software, and proliferation. Necessary market designs considering the control topology should be proposed to incentivize DERs participating in grid ancillary services.

CHAPTER 3 SINGLE AREA INERTIA ESTIMATION USING NEURAL NETWORKS

3.1 Introduction

Chapter 2 focused on the challenges of combined voltage and frequency control in LV grid. With the advancement in inverter-based technologies the control of the two parameters are coupled. Furthermore, power electronic devices decreases the overall inertia of the power system. Identification of system inertia is of great importance for FFR support during frequency events. Frequency related issues and system inertia estimation will be discussed more in the remaining chapters.

Non-synchronous RES installations are accelerating over most other traditional synchronous generator-based resources due to environmental benefits, decreasing costs, and other technological advancements [46]. Recent advancements in control techniques have shown that RES can provide VI to the system. Hence, as mentioned in Chapter 1, the stochastic nature of RES creates unknowns in the time-varying system inertia constant. Therefore, the inertia constant is no longer known and must be estimated for grid stability analysis. This chapter focuses on the system inertia estimation for single area system using CNN.

3.2 Related Works

Inertia estimation using polynomial approximation of frequency transients was proposed in [47]. A polynomial approximation technique was applied to the frequency transient waveforms to mitigate oscillatory components on ROCOF, but the proposed approach only considers synchronous generation and the selection of polynomial order was specific to the dataset used. An improved polynomial approximation technique using phasor measurement units (PMUs) was proposed in [48], and other PMU-based inertia estimation techniques were presented in [49]–[53]. However, in case of microgrids, such PMU measurements may not be readily available which makes the above techniques

unsuitable for a microgrid environment. Furthermore, PMU measurements are only available at the PCC, making the above techniques suitable only for limited scenarios. The ERCOT uses a real-time inertia monitoring tool based on the unit commitment plans submitted by the interconnected generators. However, the tool does not consider non-synchronous generating units. In the modern grid with interconnected RES with VI capabilities, the inertia constant estimation based purely on synchronous generators is inaccurate [8].

Most methods described above are based on a mathematical model of the power system. However, it is difficult to represent the entire power system is extremely complex to be approximated via mathematical modeling. As RES installations increase, the linearized reduced models may no longer represent the system dynamics [54]. Swing equation-based models work well for synchronous generator-based power systems, but it may be preferable to develop a model-free technique to better model the uncertain dynamic behavior of converter-interfaced RES.

A model-free Markov Gaussian approach is used to dynamically estimate system inertia in [55], but this approach required a large amount of historical data (i.e., two years) that may be unavailable or cumbersome for training considering a moderate sampling time. A neural network-based inertia estimation technique is proposed in [56]. The proposed method uses inter-area modal information as neural network inputs and estimates the inertia constant as an output of the network. The modal information typically represents to the frequency of oscillation and damping in a power system. Such modes of oscillation occur due to the slow mechanical oscillations of number of synchronous units, that are responsible for the system inertia, located at different areas. However, the proposed approach only estimates the inertia constant for large systems with limited accuracy, and only works with traditional synchronous generation.

3.3 Proposed Work

In this work, a model-free data-driven approach is presented to estimate the system inertia constant. A CNN was chosen to identify spatial features (i.e., time-varying) of the input data. The proposed method can estimate the system inertia with significant accuracy even with noisy frequency measurements obtained from the PLL of an ESS that is used to perturb the system. Hence, the major contribution here is the design of a model-free inertia constant estimation technique using CNN based only on local frequency measurements from a PLL.²

3.4 Traditional Power System Frequency Dynamics

The equivalent generator model provides a basis for deriving power system frequency dynamics via the swing equation [57]. The equivalent generator concept holds true when the synchronous units are taken as the primary generation source supplemented by RES as a secondary source. This section describes the relevant dynamic equations used to model system inertial response. Although a model-free approach is proposed in this work, it is important to analyze the theory behind frequency and inertial response in a power system. Such a single equivalent generator model is an approximated model of frequency dynamics that involves suitable parameter approximations. This work uses a similar equivalent generator model for simulation and analysis.

3.4.1 Characterizing a Frequency Event

Fig. 3.1 gives a general overview of the frequency response of a power system during a frequency event (e.g., increase in load or loss in generation). When electrical generation is less than load consumption, the instantaneous frequency of the rotor is lower than the nominal frequency until the primary controllers act in the system. At the point of a frequency event, the instantaneous ROCOF — determined by the inertial response of the system opposing the change in frequency — is at its maximum value. For low inertia

²Source code available at https://github.com/abodh/Neural_Network_Inertia_Estimation

systems, the instantaneous ROCOF may be large enough to trigger UFLS relays and other protection schemes. The system inertia acts approximately between time frame A and B in Fig. 3.1. The inertial response occurs in a short period of time, after which the synchronous generator governors act in time frame B to C. The system operator will then re-dispatch using automatic generation control (AGC) in time frame C to D, and if required procure reserves to maintain frequency balance.

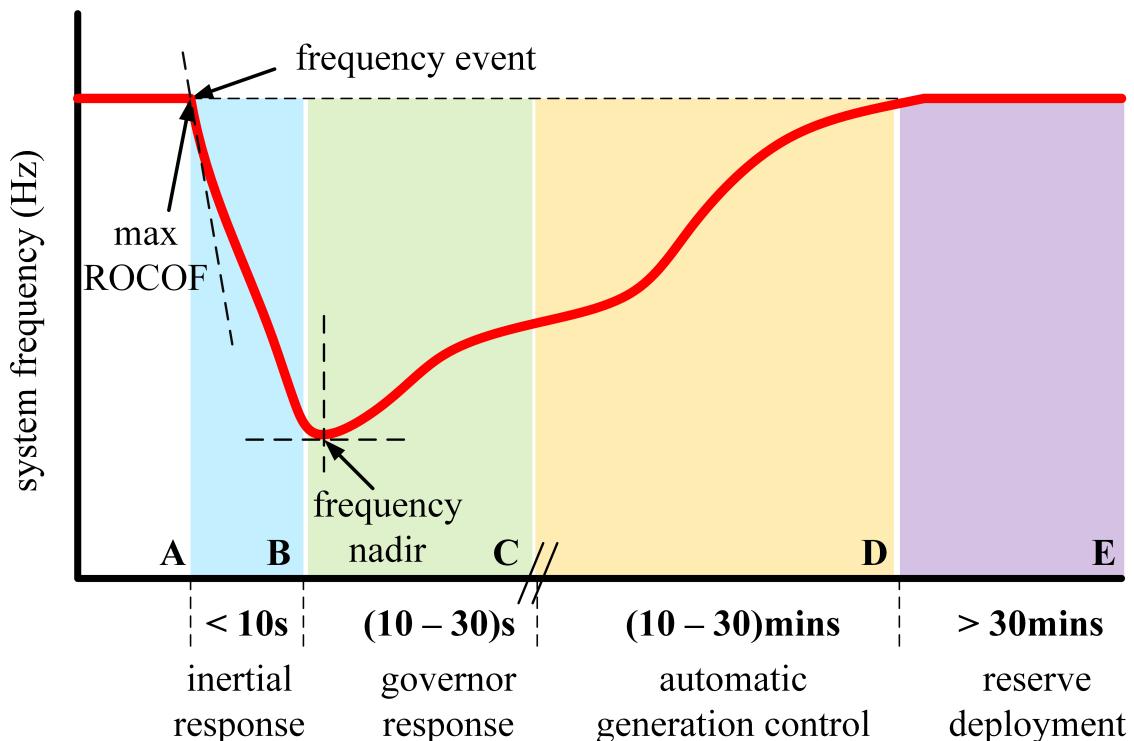


Figure 3.1. System frequency response following an event.

3.4.2 Swing Equation and Inertia Constant

When multiple generators connected in a bus are serving the load, such as in a microgrid, the system frequency, ω , is evaluated using the center of inertia frequency, that is given by the following standard equation:

$$\omega = \frac{\sum_{i=1}^n H_i \omega_i}{\sum_{i=1}^n H_i} \quad (3.1)$$

where H_i and ω_i are the inertia constant and angular frequency of the i^{th} generator.

Similarly, the inertia constant of the system, known as the center of inertia, is given by the following standard equation:

$$H = \frac{\sum_{i=1}^n H_i S_i}{\sum_{i=1}^n S_i} \quad (3.2)$$

where S_i is the apparent power rating of the i^{th} generator. More details on these parameters can be found in [23], [58], [59].

If we consider frequency-dependent loads with damping constant D and change in system frequency (in p.u.) $\Delta\omega$, a linearized form of the swing equation for the dynamic response of synchronous generation in the power system (in p.u.) is given as [15], [23]:

$$M\Delta\dot{\omega} + D\Delta\omega = \Delta P_m - \Delta P_e \quad (3.3)$$

where $\Delta\dot{\omega}$ is the ROCOF, ΔP_m is the total change in mechanical power, ΔP_e is the total change in electrical power, and $M = 2H$ is the inertia constant. This notation of inertia constant is commonly used in literature and the rest of the work follows this convention. Furthermore, to model any change in load, ΔP_L , ΔP_e is replaced by ΔP_L in (3.3).

3.4.3 Power System Frequency Control Loops

An isolated power system with frequency control loops is shown in Fig. 3.2. There are two different frequency control loops — primary control loop and secondary control loop. The frequency droop governors of the system are responsible for the primary control (i.e., time frame B to C in Fig. 3.1 as described above). The secondary control loop represents the AGC from the previous section and is responsible for removing the steady-state error. The control loops are also associated with the dynamics of the turbine-governor, that is well-explained by the following equations [60]:

$$T_g \Delta \dot{P}_m + \Delta P_m = -R^{-1} \Delta \omega + \Delta P_s \quad (3.4)$$

$$\Delta P_s = -K\Delta\delta \quad (3.5)$$

where T_g is the turbine-governor time constant of the equivalent generator, $\Delta\delta$ is the change in phase angle of the equivalent generator, R is the equivalent droop constant, K is the integral control gain, and ΔP_s is the additional power from the secondary controller. In this work, we are assuming that the AGC only comprises of integral control action and the effect of tie-line is neglected. Furthermore, the turbine-governor dynamics used in this work is a simplified single order block. However, in the real-world scenario, the dynamics of the turbine and the governor are different and we need to define separate block to show their dynamics. This simplified model could represent the turbine-governor dynamics of low order system such as steam turbines. However, further modification is required to incorporate the dynamics of higher order system such as hydro generation unit.

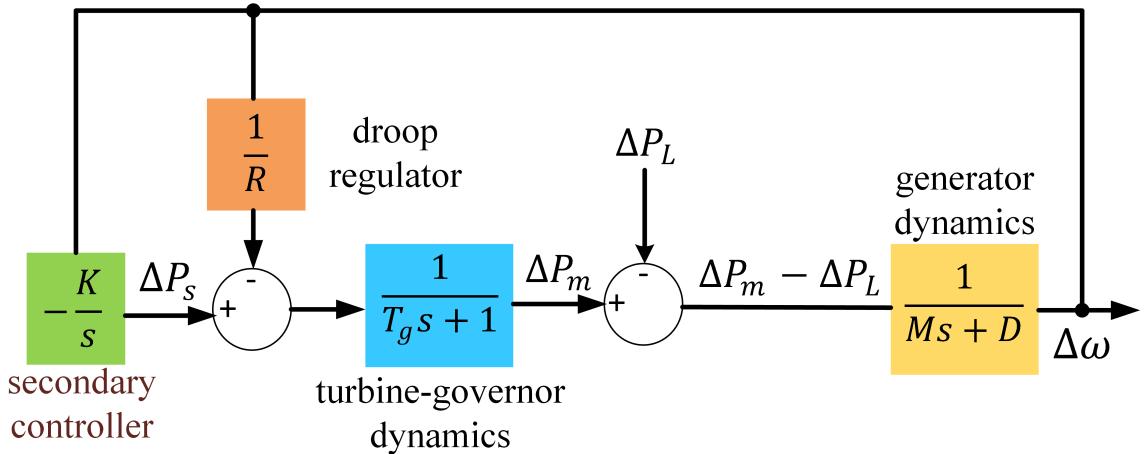


Figure 3.2. Isolated power system equivalent generator transfer function model.

3.5 Inertia Estimation using Neural Networks

3.5.1 System Perturbation using Excitation Signals

Excitation signals are a method to conduct power system dynamic studies in situations involving system perturbation without affecting system stability [61]. Power electronic devices embedded in the ESS are flexible enough to generate such excitation

signals. A sample excitation signal, fed to the system in Fig. 3.2 with an amplitude of ΔP_L , and corresponding measurements of $\Delta\omega$ and $\Delta\dot{\omega}$ is shown in Fig. 3.3. A number of snapshots of $\Delta\omega$ and $\Delta\dot{\omega}$ can be collected by varying M and ΔP_L , as shown in Fig. 3.3. The sampling time and sampling frame of the measurements can be defined after collecting the snapshots from the local ESS PLL. The noise in $\Delta\omega$ and $\Delta\dot{\omega}$ represents the measurement noise at PLL, which is discussed in the sections below. Because we are interested in estimating the inertia constant, we only consider the sampling frame in which the inertial response is prominent, similar to time frame A to B in Fig. 3.1.

3.5.2 Estimating Inertia Constant using Multilayer Perceptron

Multilayer perceptron (MLP) is a class of artificial neural network (ANN), or feedforward neural network, with multiple layers. It contains an input layer, one or more hidden layers, and an output layer with varying number of neurons. More details on MLP and its architecture can be found in [62].

An MLP can be used to estimate the power system inertia constant using the frequency measurements, $\Delta\omega$ and $\Delta\dot{\omega}$. Fig. 3.4 shows a general architecture of a MLP. For this experiment, both $\Delta\omega$ and $\Delta\dot{\omega}$ were stacked vertically to form a column vector of size z . The network contains two hidden layers with h_1 and h_2 hidden units respectively. Since we are estimating the value of inertia constant, the output layer has a single neuron with the estimated value. Before running any simulation, the entire dataset is splitted into training and testing set where the optimization on the loss function occurs during the training process. The training can also be performed in batches. If trained in batches of size b , an additional dimension is added to the input and output layer with the estimates for corresponding input in the batch.

The MLP is trained to minimize the mean squared error (MSE) between actual and predicted outputs via backpropagation [63]. MSE, parameterized by the weight of the network (w), is then given by:

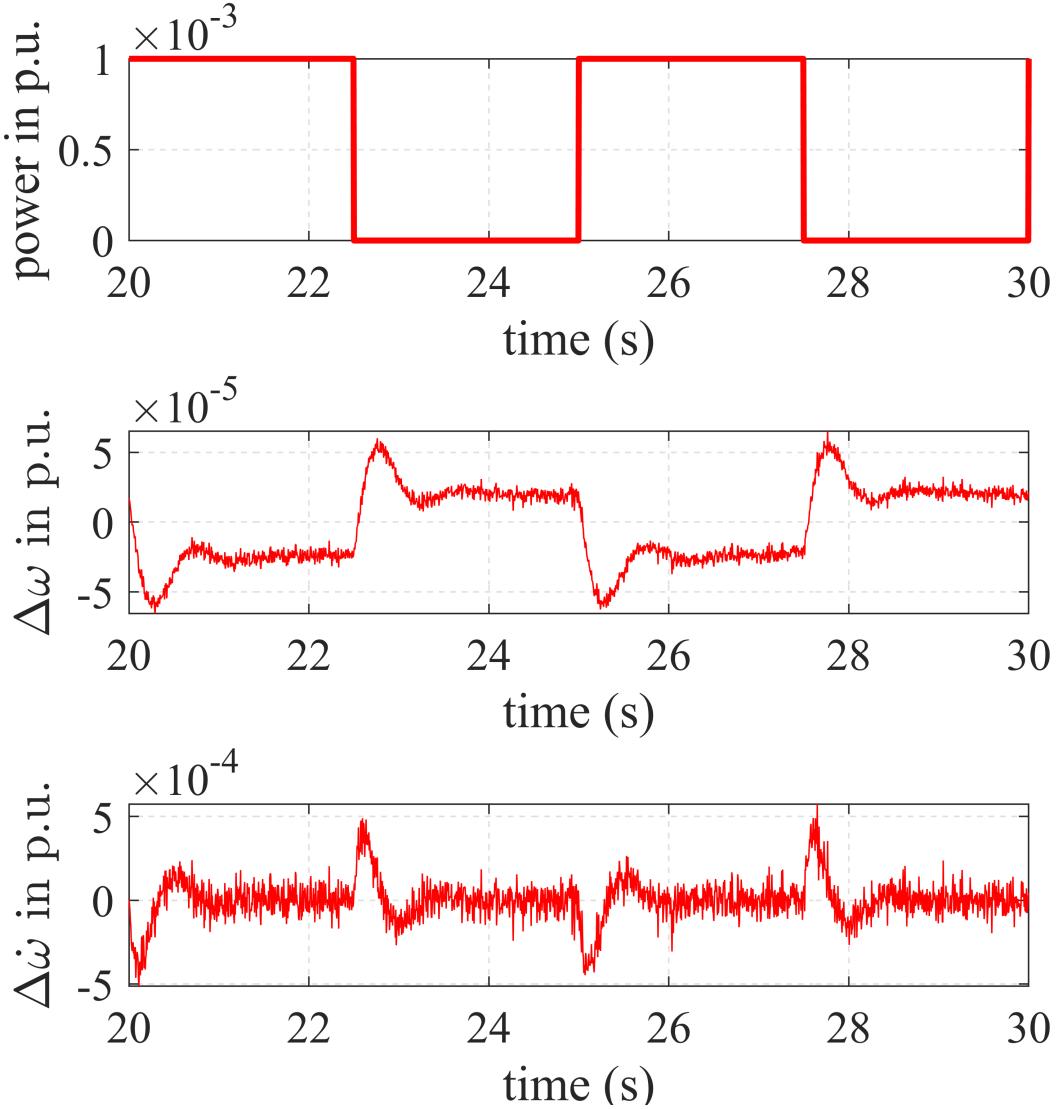


Figure 3.3. A sample of excitation signal, frequency, and ROCOF measurements for $M = 2s$ and $\Delta P_L = 10^{-3}$ p.u.

$$MSE(w) = \frac{1}{n} \sum_{i=1}^n (y_i - \tilde{y}_i) \quad (3.6)$$

where n is the total number of output samples, y_i is the 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 backpropagation can be written as:

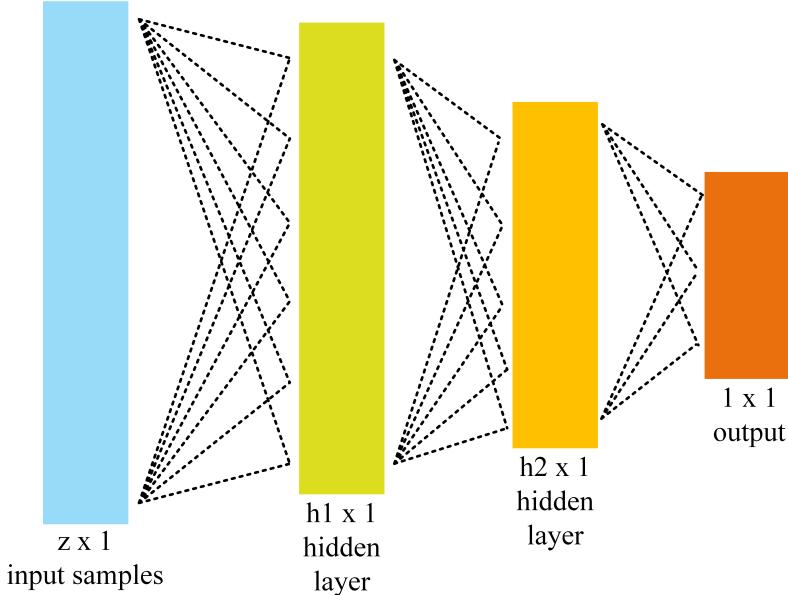


Figure 3.4. General architecture of a multilayer perceptron network.

$$w_{t+1} = w_t - \alpha \frac{\partial E_{MSE}}{\partial w_t} \quad (3.7)$$

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 (3.6). There are several other variants of (3.7) that include additional terms such as momentum and weight decay. The analysis with additional parameters can be found in [64]. Both MSE and root MSE (RMSE) can be used to assess the performance of a neural network-based estimator. MSE computes the error on the square of the target value, whereas RMSE computes the error on a scale similar to the target value. This work shows both metrics while analyzing the results.

3.5.3 Estimating Inertia Constant using CNN

CNNs have been successfully used in computer vision, image processing, and other fields in signals and time-series analysis [63]. CNN is a variant of the feedforward neural network, but with additional *convolution layers* that model the spatial input features (i.e., time-series) that make it ideal for inertia estimation.

A 1-dimensional (1-D) CNN, Fig. 3.5, can be trained to estimate system inertia from frequency measurements, $\Delta\omega$ and $\Delta\dot{\omega}$, obtained from the ESS PLL. Unlike the case of MLP, both $\Delta\omega$ and $\Delta\dot{\omega}$ were stacked horizontally to create an input in the form of a row vector of size c so that the CNN kernels, also known as convolution filters, can effectively extract the spatial features from the input. CNNs are expected to perform well when trained in batches [65]. However, the concept of batch validation or batch testing will not affect the performance of the model significantly as the network parameters are not updated while validating or testing the model. During each training iteration, a batch of size b is selected in random from the entire dataset without replacement. Before selecting a batch, the entire dataset has to be split in two sets — training and testing. The process of batch training repeats until all of the batches (i.e., the entire training dataset) have been trained. One epoch of training is completed when the entire data samples (i.e., snapshots) have each been trained. Because a supervised learning approach is taken, the snapshots should be trained multiple times. At each training iteration, the CNN input size is $b \times c$, and the output will be a column vector of size b with inertia estimates for corresponding input snapshots in the batch.

The kernels of each convolution layer is a vector, represented by R and S in Fig. 3.5, that slides through the respective input samples to get the activation for the respective convolution layer. The size of the kernel is a hyperparameter (parameter whose value is set before the training process begins), and can be tuned to get the best CNN performance. The convolution layers also contain a third dimension known as channels. The number of channels for each convolution layer, represented by p and q in Fig. 3.5, is also a network hyperparameter. Because we have a single dimensional input, the number of input channels will be 1. At the end of the convolution layers, all of the channels are flattened to a single-channel column vector, Y . This flattened layer is now the input to a feedforward neural network that is trained to minimize the MSE, using (3.6), between actual and predicted outputs via backpropagation using (3.7). The fully connected feedforward

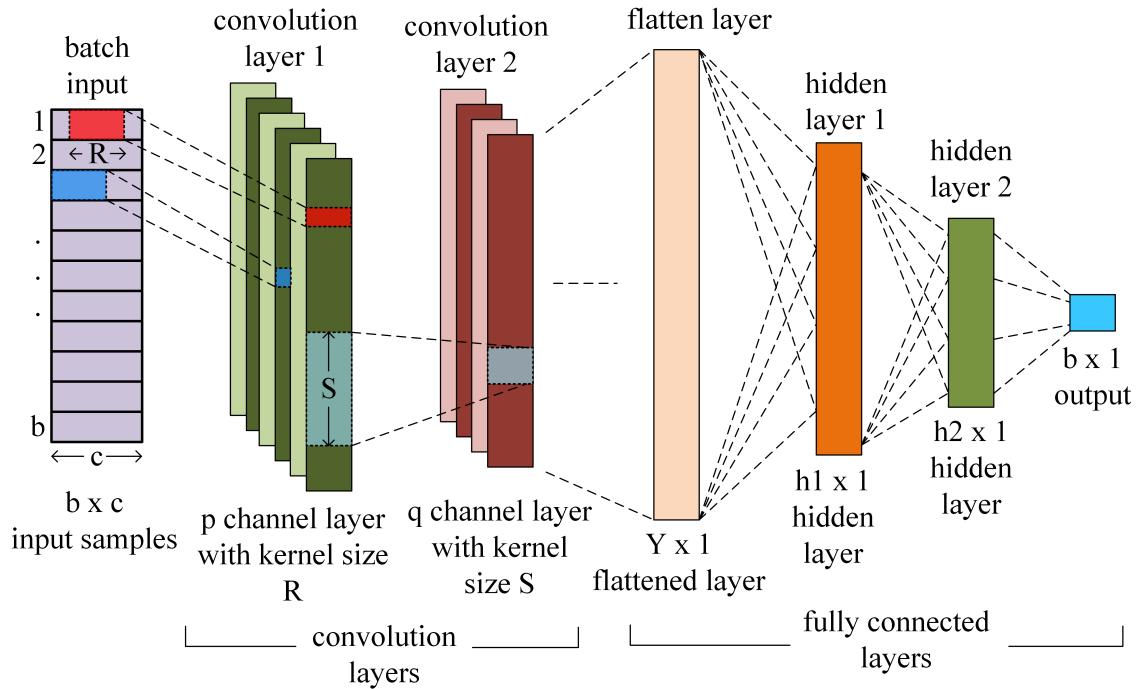


Figure 3.5. General architecture of a 1-D convolutional neural network.

network used in this CNN has two hidden layers with h_1 and h_2 hidden units, respectively.

A schematic of the overall framework of inertia estimation using CNN is shown in Fig. 3.6. Here, the additional Gaussian noise signal is added in the measurement to mimic noisy PLL measurements. CNN improves the estimation by minimizing MSE between actual value (M) and estimated value (\tilde{M}) using (3.6), and updates the model parameters using (3.7).

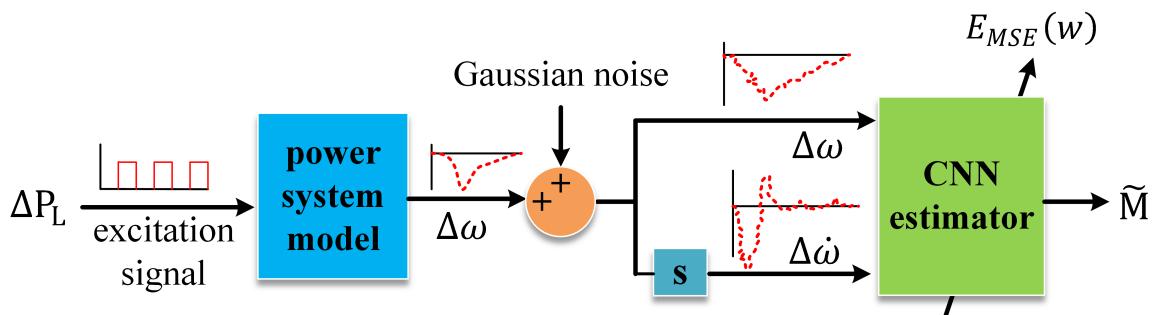


Figure 3.6. Schematic of inertia estimation using CNN.

3.6 Simulation Setup

3.6.1 Overview

The modeling and simulation of the power system, along with data collection and pre-processing, was conducted in MATLAB/Simulink 2018b. The MLP and CNN models were developed in Python using PyTorch, an open-source library for deep learning studies [66]. To leverage the fast computing abilities of PyTorch, the machine learning models were trained on South Dakota State University's Roaring Thunder cluster on NVIDIA Tesla P100/V100 GPUs. Although GPUs were used to train this model for speed of analysis, modern microcontrollers with ARM cortex cores have been successful in training deep CNN and MLP architectures and can be used in real-world implementations [67].

3.6.2 Simulation Benchmark

The equivalent generator model given in Fig. 3.2 was used for the experiment to collect the frequency snapshots, and the respective simulation parameters are given in Table 3.1. Values of M typically lie from 2 s to 9 s [68]. Hence, the frequency snapshots were collected for 17 different values of M from 2 s to 10 s with an increment of 0.5 s. Similarly, excitation signals with 100 different values of ΔP_L from 10^{-3} p.u. to 0.1 p.u. with an increment 10^{-3} p.u. were used.

Table 3.1. Simulation Parameters for Isolated Power System

Parameters	Values
inertia constant (M)	2 s – 10 s with a step of 0.5 s
change in load (ΔP_L)	10^{-3} p.u. – 0.1 p.u. with a step of 10^{-3} p.u.
damping coefficient (D)	1.5
speed regulation droop(R)	5%
turbine-governor time constant(T_g)	0.2 s
secondary controller gain (K)	2

To collect realistic data samples, white Gaussian noise was introduced in the signal

using add white Gaussian noise (AWGN) block in MATLAB/Simulink. The signal-to-noise ratio (SNR) of 60 dB with a co-variance of 1e-6 was found to be appropriate for our setup as described in [69]. A total of 1,700 snapshots were collected for training. A sampling frequency of 200 Hz was selected as mentioned in the IEEE standard for frequency measurements [70]. Similarly, a sampling frame (inertial response time frame) of 1 s was used — from 31 s to 32 s — as the system took some time to show the steady-steady response towards the excitation signal.

3.6.3 Neural Network Architecture and Hyperparameters Selection

A sampling time of 200 Hz gives 200 data points ($c = z = 400$) for each of the snapshots of $\Delta\omega$ and $\Delta\dot{\omega}$, extracted at a sampling frame of 1 s. Before using the snapshots as neural network inputs, data normalization techniques should be used as machine learning models might not converge when the magnitudes of input samples are significantly different. Min-max scalar technique was used that bounded the input values within [0,1]. For MLP, a brute-force search algorithm was applied to find the best hyperparameters and network architecture . It was determined that the network performs best when $h_1 = h_2 = 25$. The snapshots were trained in batches with $b = 30$. Additionally, to add non-linearity in the neural network mapping, the output of neural networks in each of the layers should be passed through an activation function [64]. In this work, tangent hyperbolic (tanh) activation function has been used in each of the layers of MLP.

For CNN, we experimentally determined that two convolution layers with $p = 10$ and $q = 20$, and kernels with sizes $R = S = 3$ gave the best results. Similarly, for the feed forward layers, the size of the flattened layer depends on the architecture of the convolution layers [63]. For this specific architecture, with $c = 400$ and $q = 20$, the value of Y was calculated to be 7960, whereas selecting $h_1 = 800$ and $h_2 = 50$ improved the performance of the network. The training was performed in batches with $b = 30$. In this work, rectified linear unit (ReLU) activation was used for convolution layers and tanh for feed forward network. Because we are estimating continuous values, the activation

function for the final output layer is omitted.

For both of the networks, a modified version of weight update algorithm was used with $\alpha = 1e-3$, momentum 0.5, and weight decay of 5e-4. The simulation was carried out for 1000 epochs in case of MLP and 200 epochs for CNN by dividing the entire dataset into two parts — 1360 snapshots ($\sim 80\%$) for training and 340 snapshots ($\sim 20\%$) for validation.

3.7 Results and Analysis

3.7.1 Estimating Damping Constant using Neural Network

As an additional experiment, the damping constant from (3.3) was estimated using the network shown in Fig. 3.4. The backpropagation was performed on the actual value of D using (3.7). The simulation was conducted on the system shown in Fig. 3.2. For this experiment, the simulation parameters were selected from Table 2.2 and the value of D was varied as [0.5, 1.0, 1.5, 2.0, 2.5]. However, the estimated values of damping constant, (\tilde{D}) , deviated from D by a significant margin. After a few epochs, the MLP estimated the average value of D over the entire snapshots.

Fig. 3.7 shows the analysis of change in frequency with respect to time for different values of D [71]³. It can be seen that the effect of changing D is only prominent in the region of frequency nadir. This is further supplemented through the sensitivity analysis shown in Fig. 3.8. The frequency sensitivity is the highest or most prominent at the time of frequency nadir. At all other time period the frequency sensitivity with respect to D is insignificant.

To analyze the combined effect of change in M and change in D , the simulation was carried for three different values of $M = 1s$, $5s$, and $9s$, a constant value of $\Delta P_L = 0.05p.u.$, and five different values of D as mentioned above. Fig. 3.9 and Fig. 3.10 shows the effect of change in M and D on $\Delta\omega$ and $\Delta\dot{\omega}$ respectively. It can be seen that the

³Fig. 3.7 and Fig. 3.8 are the results taken from [71]. The results are referenced in this work to better describe the problem of damping constant estimation using the proposed approach.

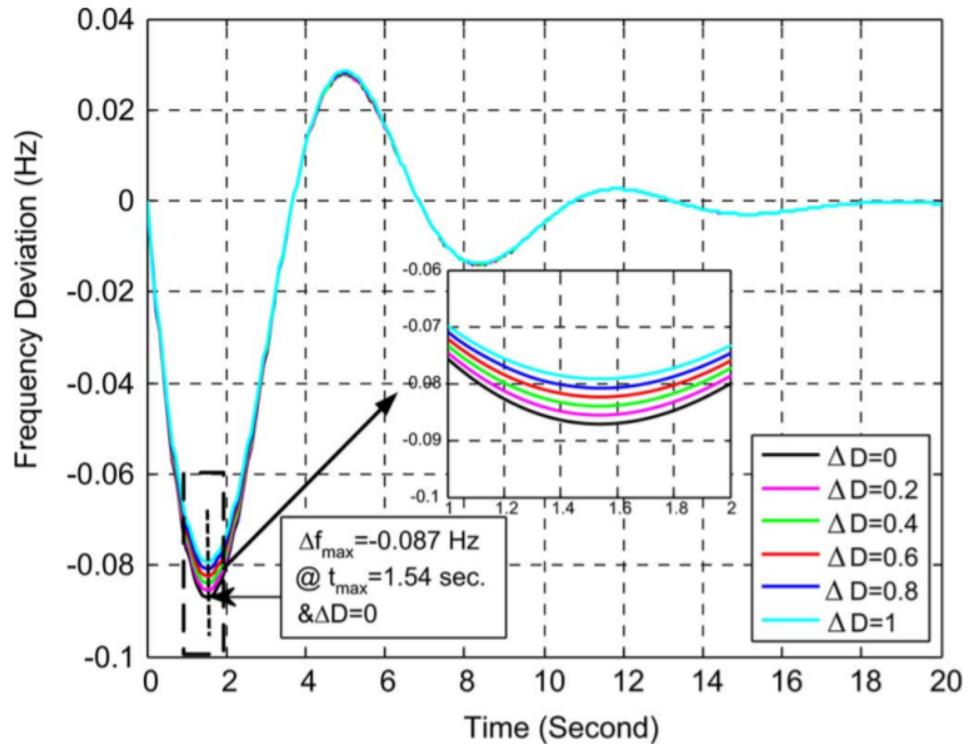


Figure 3.7. Change in frequency for different values of damping constant [71].

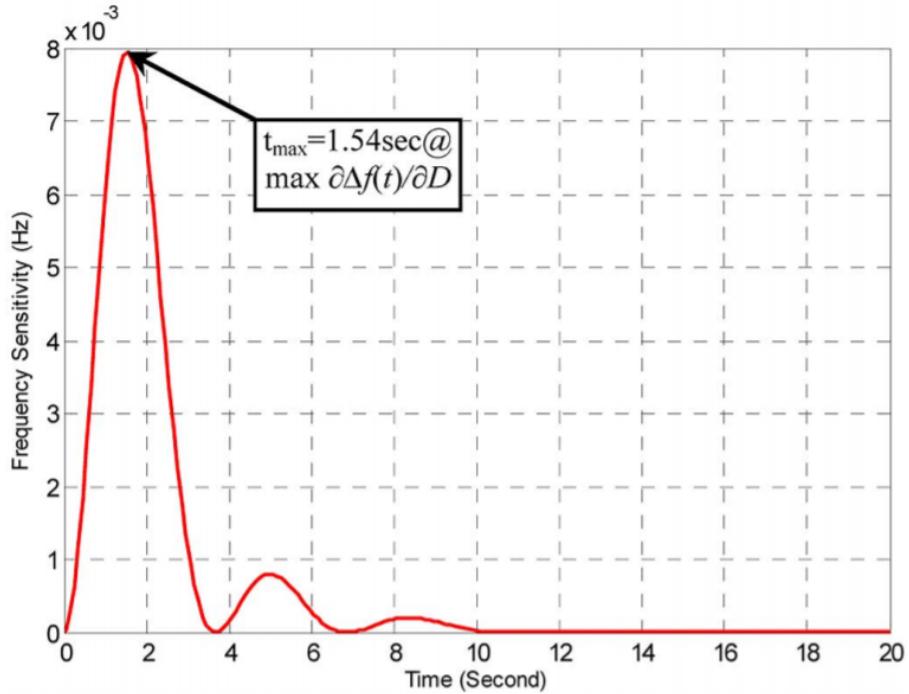


Figure 3.8. Frequency sensitivity with respect to damping constant [71].

effect of M is highly prominent on $\Delta\omega$ and $\Delta\dot{\omega}$, whereas the effect of D is only prominent on the region of frequency nadirs or general point of inflection.

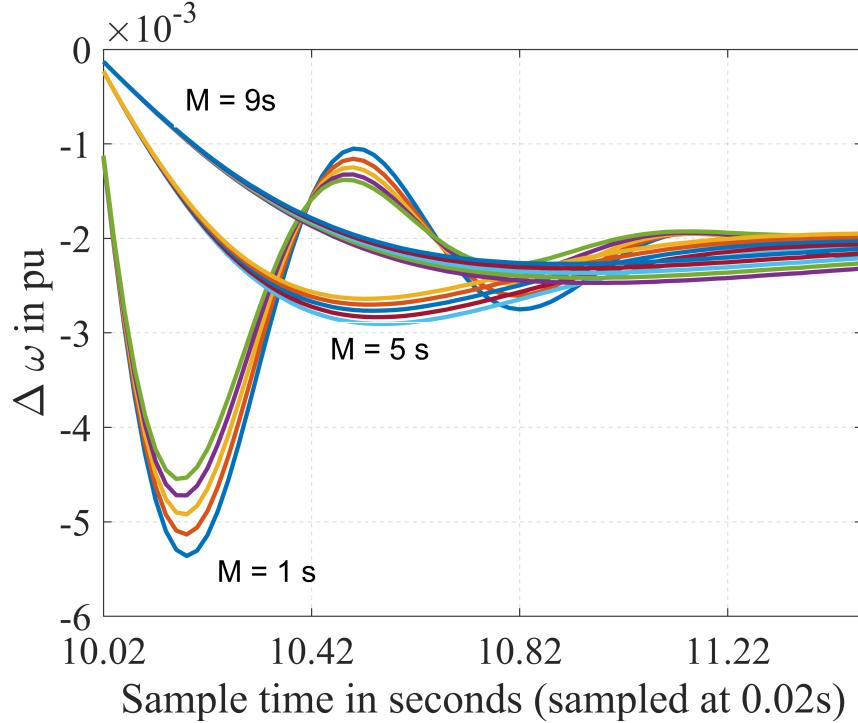


Figure 3.9. Change in frequency for changing values of M and D with a constant ΔP_L . Here, $M = 1\text{ s}, 5\text{ s}$, and 9 s whereas $D = [0.5, 1.0, 1.5, 2.0, 2.5]$. The sampling time for the snapshot is taken as 0.02 s .

Table 3.2 shows five randomly selected values of \tilde{D} estimated by the neural network. The estimator depicts the case of an average estimator model. This may be due to the fact that D only affects the region of frequency nadir and the sampling frame taken for the estimation is of the region in which the effect of M is prominent, which might not incorporate the measurements of frequency nadir.

Table 3.2. Actual and Estimated values of Damping Constant

Actual values (D)	Estimated Values(\tilde{D})
0.5	1.1528
1.0	1.2078
1.5	1.533
2.0	1.6199
2.5	1.3697

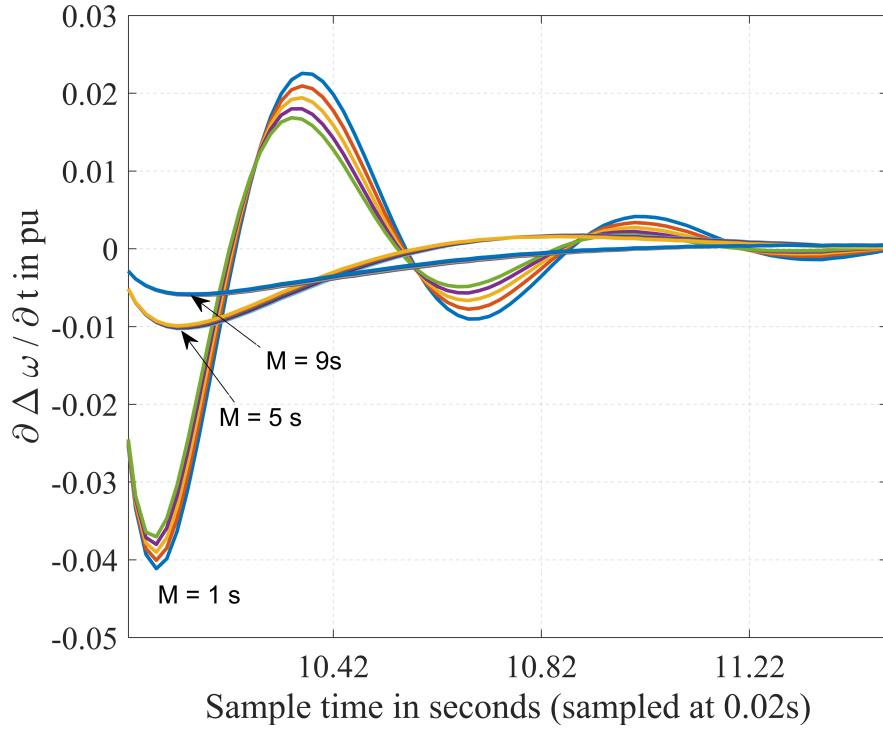


Figure 3.10. ROCOF for changing values of M and D with a constant ΔP_L . Here, $M = 1\text{ s}$, 5 s , and 9 s whereas $D = [0.5, 1.0, 1.5, 2.0, 2.5]$. The sampling time for the snapshot is taken as 0.02 s .

Furthermore, even after selecting the sampling frame that corresponds to the region of frequency nadir or point of inflection, the proposed approach might not be viable approach to estimate the damping constant. From Fig. 3.10 and Fig. 3.9 we can see that the point of inflection occurs at different point of time for $\Delta\omega$ and $\Delta\dot{\omega}$. Hence, selecting a specific sampling frame to train the neural network might not be an appropriate method to estimate the damping constant.

3.7.2 Inertia Estimation using Neural Network

3.7.2.1 Effect of multicollinearity due to ΔP_L

Initially ΔP_L was considered as one of the inputs to the neural network. However, after further investigation, it was found that ΔP_L and $\Delta\omega$ are strongly correlated, with coefficients between ΔP_L , $\Delta\omega$, and $\Delta\dot{\omega}$ shown in Table 3.3. Because inertia estimation is a regression problem, multicollinearity — collinearity between the inputs — can severely

impact the estimated values. If two of the inputs are highly correlated, then the effect of one input is redundant in the model. After removing ΔP_L as an input to CNN, the overall RMSE was reduced from 0.419 to 0.2309. Thus, the multicollinearity due to the inclusion of ΔP_L as an input affected the overall performance of CNN. Additionally, in real-world scenarios, it may be infeasible to exactly know the value of ΔP_L and the fact that it can be excluded is an important result for inertia estimation.

Table 3.3. Correlation coefficients between $\Delta\omega$, $\Delta\dot{\omega}$, and ΔP_L

	$\Delta\omega$	$\Delta\dot{\omega}$	ΔP_L
$\Delta\omega$	1	-0.069	-0.85
$\Delta\dot{\omega}$	-0.069	1	-0.127
ΔP_L	-0.85	-0.127	1

3.7.2.2 Analyzing the Performance Metrics

As the number of training epochs increase, MSE decreases. After a certain number of epochs, the MSE between consecutive epochs seem unchanged as the value of MSE will be extremely small and hence it is not obvious to see if the model is improving. The natural logarithm function gives a high negative value when the input lies in close proximity of 0. Hence, observing the natural logarithm of MSE shows if the model is improving further. Fig. 3.11 and Fig. 3.12 shows the MSE and natural logarithm of MSE of training and validation sets. The validation loss for MLP is highly fluctuating as the training is done in batches and some of the validation batches show high MSE on testing phase. However, the curve is significantly better for CNN as the process of convolution facilitates the identification of useful features from the input snapshots. Furthermore, to get a similar level of accuracy, MLP was trained for 1000 epochs whereas CNN was trained only for 200 epochs. Hence, CNN performed well in terms of inertia estimation as compared to MLP. The analysis of rest of the performance metrics in this chapter are shown in terms of CNN estimator.

From Fig. 3.12 we can see that the MSE is decreasing as the logarithm of MSE gives high negative values with the increase in number of epochs. It is interesting to notice

that the logarithm of validation MSE deviates at around epoch 190, and hence the training was stopped to obtain the best model (i.e., the early stopping technique in neural network literature). In validation, the errors fluctuate more than the training errors. Because the neural network weights are not updated after computing the loss for validation sets, some of the validation batches may significantly differ from the actual value which creates the observed fluctuations.

Fig. 3.13 represents the evolution of weights from the final hidden layer to the output layer for a random sample from the batch. Because $h_2 = 50$ and we are estimating a single value of inertia, we will observe 50 different weights that saturate near epoch 190. This infers that the weights are no longer updated in the backpropagation due to minimal MSE.

Fig. 3.14 shows the accuracy of the CNN model with a tolerance of 10%. The model has a validation accuracy of 97.35%, with an RMSE of 0.2309, after training for 190 epochs. Fig. 3.15 shows the violin plot of the estimated values of the validation set. The straight line represents a perfect estimation case, whereas the violins represent the distribution of the specific estimates. It can be noticed that the distribution of most of the estimated values are mostly within the tolerance level. Some of the estimated values deviate from the actual value, which are prominent at the tail region of the violin plot, and contribute to the lower accuracy of the estimator network.

Finally, the model was tested on a different set of snapshots that were not used for training or validation. In this phase, CNN is considered to be fully trained and is assessed on an unknown environment. From the previous experiment, the model was saved at epoch 190 and was used on the test set. A different set of 25 snapshots were collected by varying M from 2 s to 10 s with an increment of 2 s and ΔP_L from 0.01 p.u. to 0.1 p.u. with an increment of 0.02 p.u. The trained CNN model estimated the inertia constant on an unknown test set with an RMSE of 0.1763.

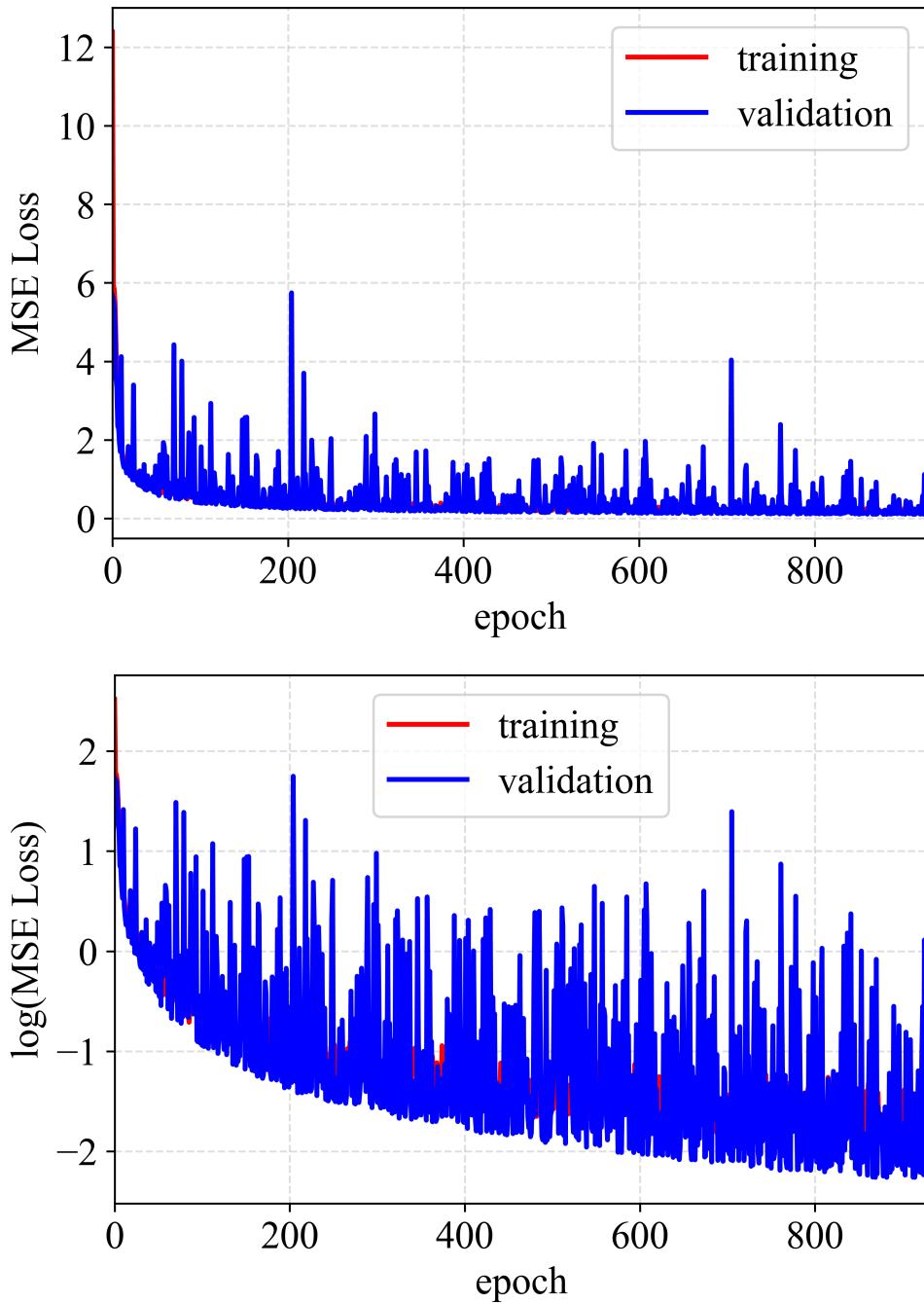


Figure 3.11. MSE loss (upper) and natural logarithm of MSE loss (lower) of inertia estimates with the number of epochs for MLP. Here, natural logarithm of MSE loss will amplify the values of MSE that are in close proximity of 0 for better visibility and analysis.

3.8 Conclusion

In this work, the inertia constant of a power system was estimated using a neural network — MLP and CNN. To avoid system imbalance, a non-intrusive excitation signal

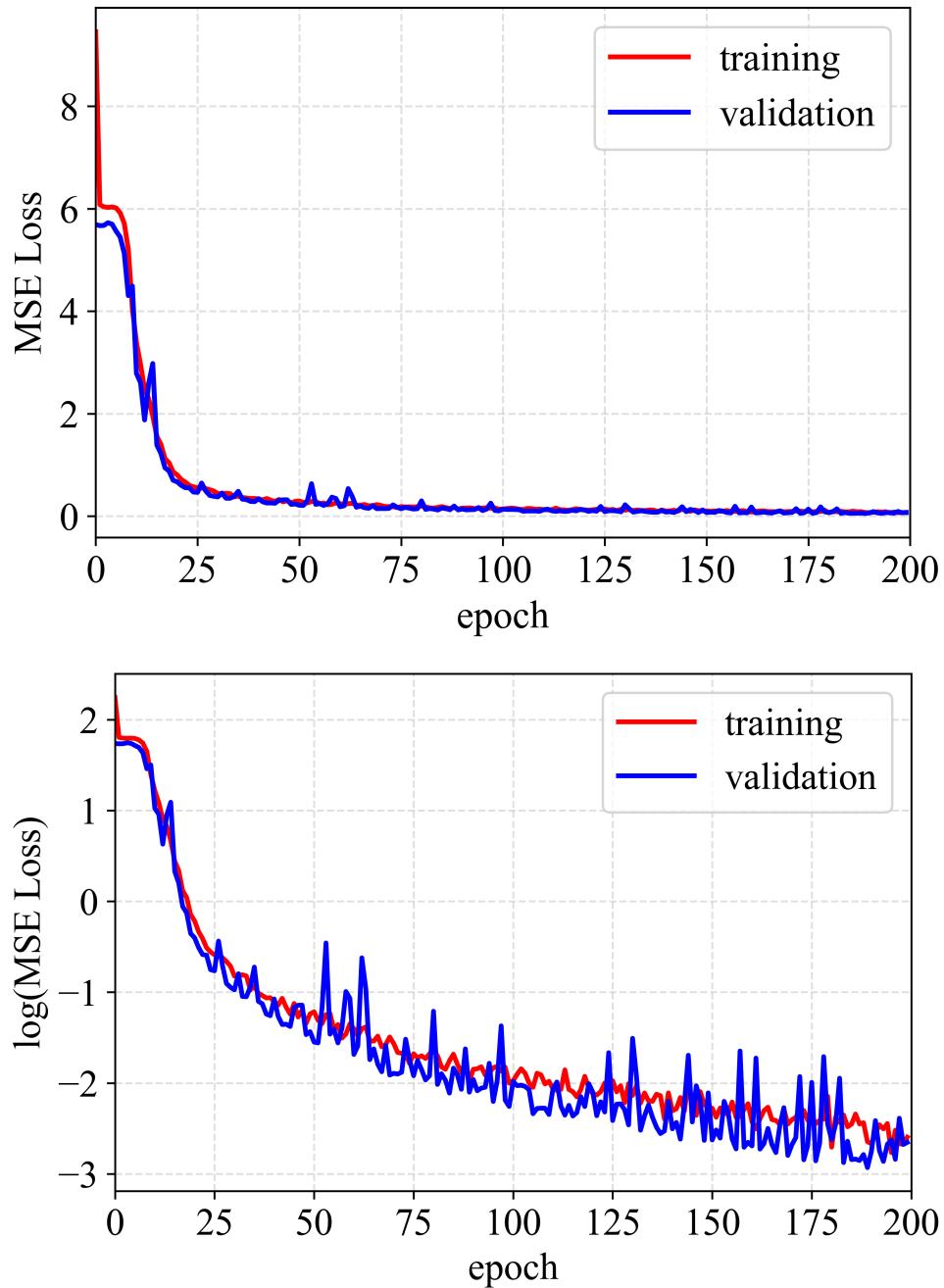


Figure 3.12. MSE loss (upper) and natural logarithm of MSE loss (lower) of inertia estimates with the number of epochs for CNN.

was used to collect frequency snapshots. The model-free approach to estimate the inertia constant depended only on frequency measurements, while not requiring ΔP_L as an input. It was found that MLP takes longer epochs to converge and has higher fluctuations in validation accuracy as compared to CNN. The CNN-based inertia estimator showed good

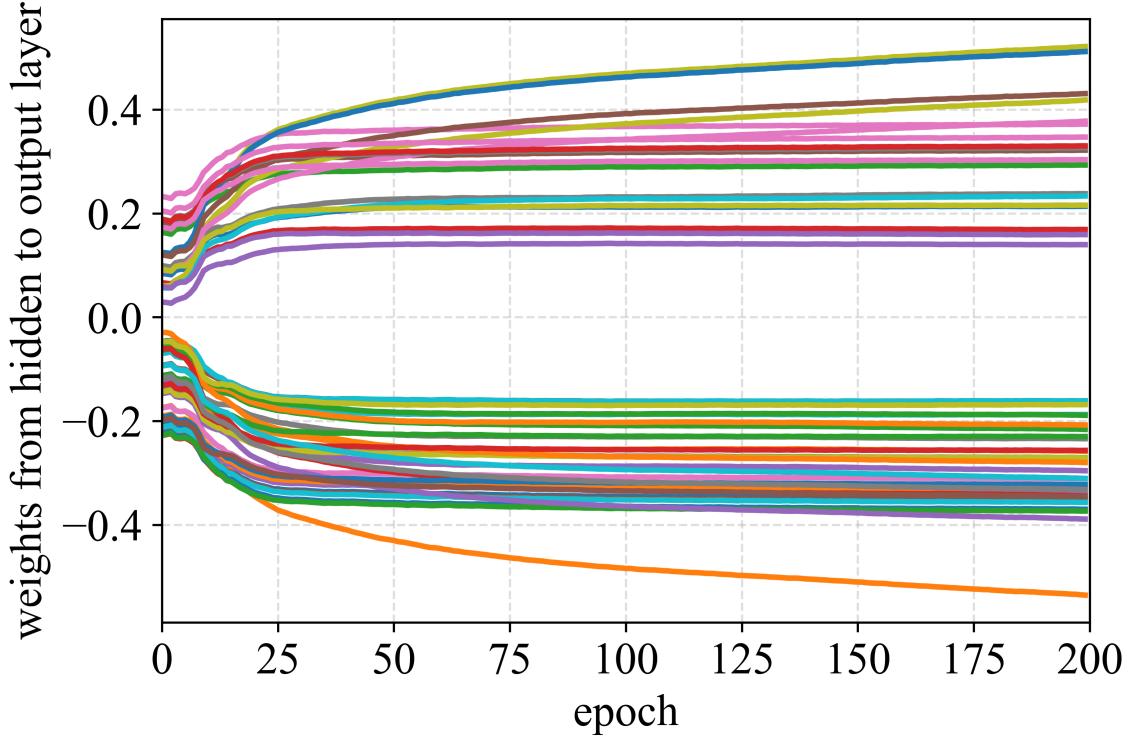


Figure 3.13. Evolution of training weights from the hidden layer to output layer of CNN.

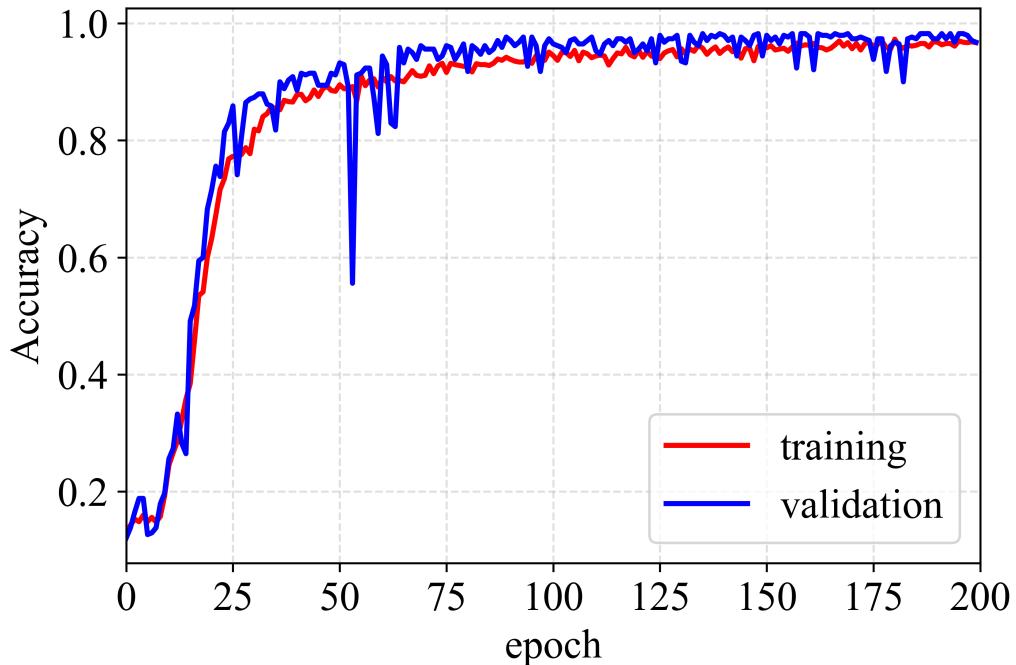


Figure 3.14. Model accuracy on the inertia estimates from CNN. The estimates within 10% of the actual values are considered to be accurate.

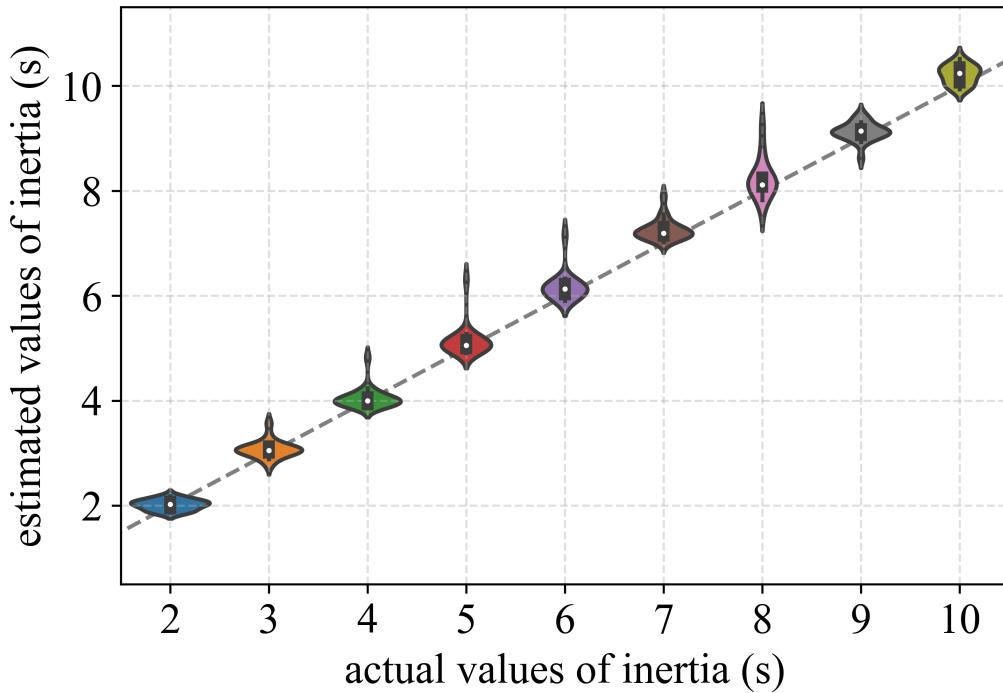


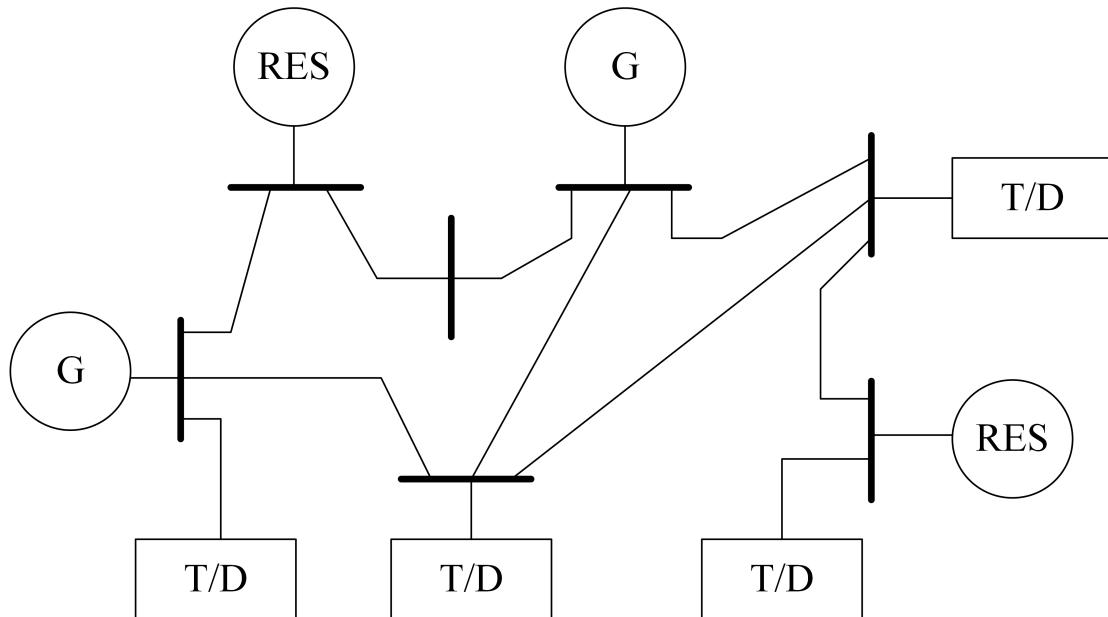
Figure 3.15. Inertia estimates (in seconds) on the validation set after completing the training process.

performance even with noisy input samples. Considering the stochastic nature of renewable energy sources and the ability of power converters to provide VI, such a model-free approach can be applied to estimate unknown power system inertia in current and future converter-dominated grids.

CHAPTER 4 MULTI-AREA INERTIA ESTIMATION USING DISTRIBUTED MACHINE LEARNING

4.1 Introduction

The electric grid in bulk power system follow mesh topology — there exists multiple path from one node to the other [72]. In such architecture, power system contains different sub-systems, also known as areas or regions, interconnected via tie-lines. Typically, a power system is divided into areas to have multiple control zones. Additionally, each of the areas has several interconnected generators. Therefore, in multi-area power system several generators can support the system via tie-lines [23], [59]. Fig. 4.1 shows a general schematic of a bulk power grid. The bulk grid is meshed whereas the grid follows radial architecture while going deep on to the distribution side. The areas or control zones in bulk grid are determined by techniques explained in [23], [73].



G : generator

T/D: transmission/distribution

RES : renewable energy resources

Figure 4.1. Basic electrical structure of bulk power system.

Due to the reasons explained in Chapter 1, power system inertia is a time-varying parameter in the modern grid. The task of inertia estimation becomes even more challenging when considering its meshed architecture as the inertial support can also come from interconnected areas via tie-lines. Hence, the frequency dynamics in multi-area system will be different than that in single area system. A preliminary work on CNN-based inertia estimation for single area system is presented in Chapter 3. This chapter focuses on an extended version of the work presented in Chapter 3, that incorporates the tie-line power flow in multi-area power system while estimating the inertia constant. A distributed machine learning technique — federated learning (FL) — is applied in order to estimate the inertia constant for multiple areas.

4.2 Related Works

There are very few works that consider multi-area power system to estimate the inertia constant. The authors in [52] estimate equivalent inertia of individual areas in a multi-area power system using PMU measurements. The proposed estimation technique requires that the area buses have frequencies close to each other. However, PMU measurements might not be readily available at all areas and such strict conditions restrict the proposed approach to specific cases. An inertia estimation method for a multi-area interconnected electric power system using electromechanical oscillation mode has been discussed in [74]. The relation between inertia and the frequency and damping of a mode from the active tie-line power flow was developed. However, the penetration of RES in any one of the areas can change the dynamics of electromechanical modes drastically. Hence, such model-based methods might not be feasible to estimate inertia with RES based generation. A model-free approach to estimate multi-area inertia constant has been presented in [56]. The authors use inter-area modal information, particularly frequency and damping of oscillation mode, to estimate the inertia constant. However, the training data is centralized and trained in a single neural network. This adds an additional cost to centrally collect the distributed data.

4.3 Proposed Work

With the increasing penetration of RES in the power grid, it is desired to have an inertia estimation method that is independent of the complex power system model. Furthermore, the model-free method should be cost efficient, accurate and computationally effective. In this work, a distributed machine learning approach, FL, is applied to estimate the inertia constant in multi-area power system using local frequency measurements from the PLL of ESS that is used to perturb the system. Similar to the work in Chapter 3, a CNN was chosen to identify spatial features of the input data. However, in this work, a consensus-based *federated averaging algorithm* [75] averages the CNN models of individual power system areas. The ESS units located at different power system areas are designated as clients where the CNN models train the frequency snapshots to estimate the inertia constant locally. On the other hand, a central device, selected arbitrarily, acts as a server for federated averaging. Therefore, the major contribution here is the design of a model-free FL-based inertia estimation approach in multi-area power system using local frequency measurements from a PLL of ESS. The proposed approach is verified for both IID and non-IID data in the test system with significant accuracy.

The proposed work is based on the following assumptions:

1. For the preliminary analysis and simplicity, the proposed approach is tested on a transfer function-based two-area system. The number of areas determines the number of clients for FL. Hence, the proposed approach is implemented on a system with two clients.
2. The dynamics and contribution of multiple generators in an area can be represented by a single equivalent generator.
3. The system has a pre-existing communication protocol between the clients and the server. The design and description of such protocol is not within the scope of this work.

4.4 Two-Area Power System

This section provides a detailed theory behind the dynamics of a two-area power system. Although a model-free approach is proposed in this work, it is important to analyze the theory behind the dynamics of a multi-area power system to gain full understanding of this problem.

A general schematic of a two-area power system is shown in Fig. 4.2. Here, multiple generators serving an equivalent load at each of the areas is represented by an equivalent generator, and is modeled as a voltage source.

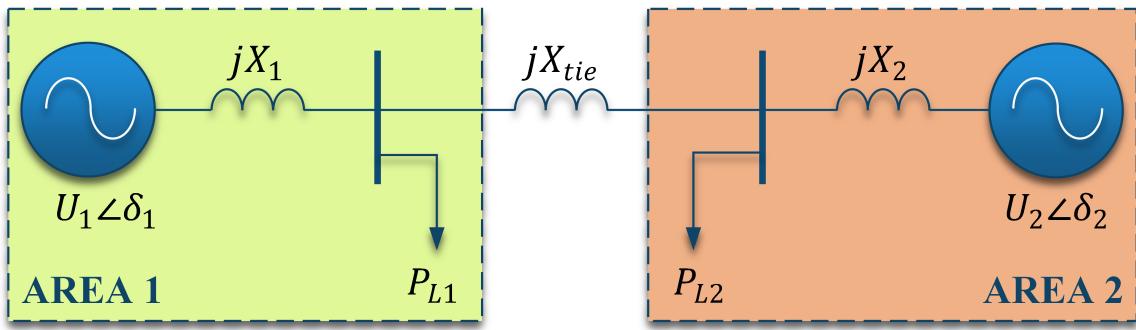


Figure 4.2. Two-area power system. The generator at each of the areas represent an equivalent generator of multiple generators serving an equivalent load at that bus.

Let us consider that the bus at area 1 is at higher potential than the bus at area 2 and all other electromagnetic dynamics are ignored⁴. Hence, the tie-line power flow from the sending end at area 1 to the receiving end at area 2 is given by [23], [59], [76]:

$$P_{tie} = \frac{U_1 U_2}{X} \sin(\delta_1 - \delta_2) \quad (4.1)$$

where P_{tie} is the tie-line power flow from area 1 to area 2, U_1 and U_2 are the emfs of the voltage sources, δ_1 and δ_2 represent the rotor angle of equivalent generator, and X is the equivalent reactance of the system including the line reactance X_{tie} and synchronous reactances X_1 and X_2 .

⁴Subscript 1 denotes area 1 and subscript 2 denotes area 2, unless stated otherwise, throughout this chapter.

For small deviations, U_1 and U_2 are constant and the changes in synchronous reactances, X_1 and X_2 , are also minimal. Hence, (4.1) becomes:

$$\Delta P_{tie} = T_0(\Delta\delta_1 - \Delta\delta_2) \quad (4.2)$$

where $T_0 = \frac{U_1 U_2}{X} \cos(\delta_1^0 - \delta_2^0)$ is the tie-line coefficient. Here, the superscript 0 in δ_i^0 refers to the initial condition.

The power exported by or imported to an area is the summation of the tie-line power flow and the power consumed by an equivalent load at that particular area. Thus, with the help of above equations, the linearized model of electrical power corresponding to respective areas can be formulated as:

$$\begin{aligned} P_{e1} &= T_0(\Delta\delta_1 - \Delta\delta_2) + P_{L1} \\ P_{e2} &= -T_0(\Delta\delta_1 - \Delta\delta_2) + P_{L2} \end{aligned} \quad (4.3)$$

where P_e is the electrical power. Here, negative sign on the second equation denotes that the tie-line power flows from area 1 to area 2.

Using the models above, the two-area power system can be extended to a transfer function-based model as shown in Fig. 4.3.

The dynamics related to individual areas, excluding the dynamics due to tie-line power flow, has been well explained in Chapter 3. The model in Fig. 4.3, however, contains an additional signal termed as area control error (ACE). ACE reflects the power imbalance in each of the areas and is given by:

$$ACE = B\Delta\omega + \Delta P_{tie} \quad (4.4)$$

where B is the bias factor. Typically, the bias factor for any area is governed by the equivalent droop constant and damping constant of that particular area, i.e.

$$B = 1/R + D [59].$$

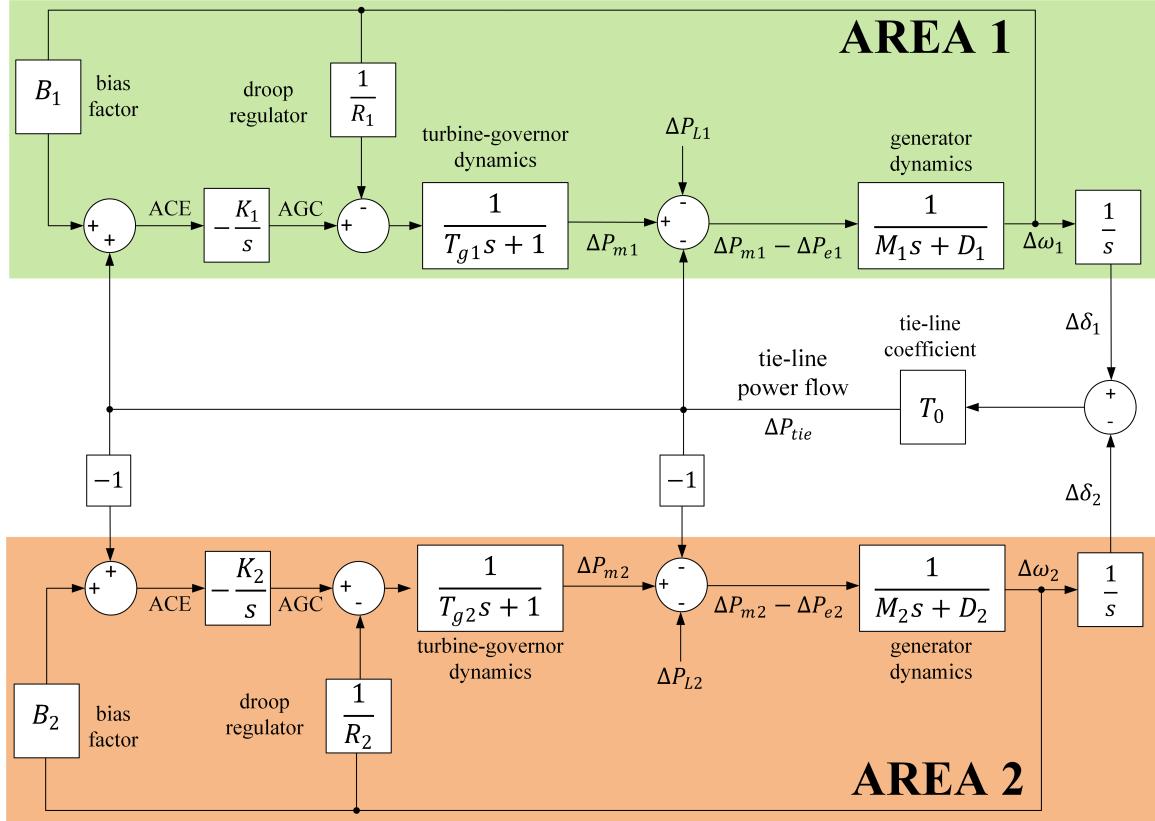


Figure 4.3. Transfer function-based model of a two-area system

4.5 Multi-Area Inertia Estimation using Federated Learning

4.5.1 System Perturbation using Excitation Signal

The technique of system perturbation using excitation signal has been well explained in Chapter 3. In this work, similar perturbation technique has been used. However, it should be noted that in order to collect the snapshots, only the area in which the snapshots are to be collected is perturbed while the change in load for the other area is assumed to be zero. A sample excitation signal, fed to area 1 of the system in Fig. 4.3 with an amplitude of ΔP_{L1} , and corresponding measurements of $\Delta\omega_1$ and $\Delta\dot{\omega}_1$ is shown in Fig. 4.4.

The sampling time and sampling frame of the measurements can be defined after collecting the snapshots from the PLL of local ESS. The noise in $\Delta\omega_1$ and $\Delta\dot{\omega}_1$ represents the Gaussian measurement noises discussed in Chapter 3. Because we are interested in

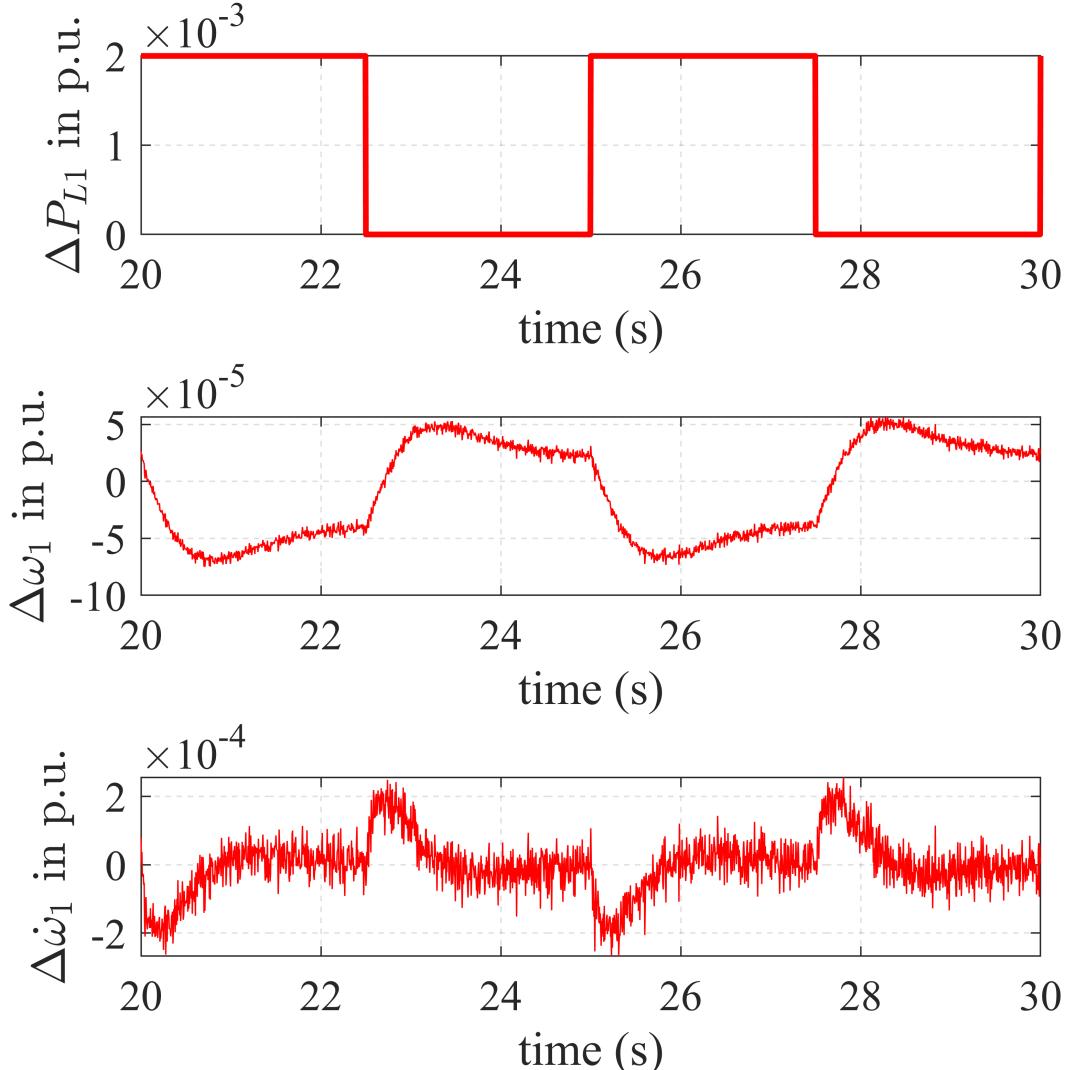


Figure 4.4. A sample of excitation signal, frequency, and ROCOF measurements at area 1 when $M_1 = 9s$, $M_2 = 5s$, $\Delta P_{L1} = 2 \times 10^{-3}$ p.u., and $\Delta P_{L2} = 0$ p.u. Only the area in which the frequency snapshots are to be collected is perturbed via excitation signal. In this case, area 1 is perturbed via $\Delta P_{L1} = 2 \times 10^{-3}$ p.u. whereas $\Delta P_{L2} = 0$ p.u..

estimating the inertia constant, we only consider the sampling frame in which the inertial response is prominent, similar to time frame A to B in Fig. 3.1. By varying M and ΔP_L for individual areas, multiple snapshots can be collected.

4.5.2 Independently and Identically Distributed Data

In the context of probability and statistics, IID data are random variables whose probability distribution are equal and mutually independent; the converse being true for

non-IID data. Fig. 4.5 shows a general example consisting of MNIST dataset [77] that clarifies the difference between IID and non-IID data. For IID case (left), it can be seen that the dataset has almost equal probability distribution of digits from 0 to 9. On the other hand, for non-IID case (right), digits 1 and 7 have high probability distribution.



Figure 4.5. General example of MNIST IID (left) and non-IID (right) dataset. IID data are identically distributed i.e. the probability distribution of all of the random samples (digits) are equal. On the other hand, for non-IID case the probability distribution of few samples dominate the others.

Typically, distributed data is non-IID and hence any particular client's sample distribution will not represent the population distribution. Furthermore, the clients could also have unequal number of data samples depending on the application [75]. In multi-area power system, the inertia constant is non-identically distributed as it could be different for different areas. In this work, for IID-case, the frequency snapshots corresponding to different values of M and ΔP_L are randomly distributed to the client ESS. This will ensure that the probability distribution of data samples are even. To mimic non-IID case, firstly, the overall snapshots are collected and arranged in ascending order of M . Then the entire dataset is divided into equal number of parts so that each of the clients get one part of data. This ensures that each of the clients has non-identically distributed

data. The manual distribution of data is only conducted for experimental purpose. In real-world power system cases, the clients will have its own data (most likely non-IID) and it is not required, and viable, to combine and re-distribute the data as discussed here.

4.5.3 Federated Learning-based Inertia Estimation

Existing centralized machine learning approaches required collected the training data in a centralized location and perform the prediction on a single model based on the aggregated training data [78]. This method is expensive and inefficient from both communication and memory point of view — the central server needs to have enormous storage capacity to accommodate entire training data from the clients and the clients have to communicate time and again to update the training data.

FL is a secure and robust framework that facilitates distributed machine learning. The training is performed remotely on individual clients in a distributed fashion. However, after the training has been completed a central server aggregates the trained weights from each of the clients and then re-distributes the aggregated weights to the clients. FL is highly efficient and robust than conventional machine learning techniques due to following reasons:

1. Only model weights are communicated from the clients to server and vice-versa.

Although the model weights are floating point values, the information can be further segregated using an encoding-decoding technique at the client and the server side. This ensures data privacy (differential privacy [75]) and discourages possible cyber-attacks, which is a serious concern in the field of power systems.

2. The training process at each of the clients is offline whereas the communication between the clients and server occurs only during the weight aggregation and distribution phase, if the clients are available. Communication cost is important in any online optimization task as the bandwidth of communication could be limited.

In FL, the number of local training epochs can be varied to improve the model at

client's level. This can reduce the communication cost drastically.

3. Since the weights of the neural network are just floating point values, the server does not require excess memory to store and aggregate the weights of the clients. This is an important aspect of FL that makes it more robust as compared to other machine learning algorithms.

FL can be applied in a multi-area system to estimate the inertia constant of individual areas. It is to be noted that FL does not estimate the inertia but provides a framework to do so in distributed fashion. In FL the training data — frequency snapshots — reside at the client ESS location making it a distributed machine learning algorithm. Each of the areas will have a shared CNN model to estimate the inertia constant at that particular area. A detailed description on the CNN model is provided in Chapter 3. Fig. 4.6 shows a general framework of FL-based inertia estimation in multi-area power system. It is assumed that there is a central server somewhere within the interconnected system⁵.

Each of the communication rounds between the server and the clients can be represented into three processes — check-in, configuration and training, and weight aggregation. Let N be the total number of ESS clients, each belonging to an area n (for two-area system $n = 1$ and 2). During check-in, only C fraction of N clients are selected by the server. To be selected, the clients must be connected via a communication channel. The red cross denotes the clients that are not selected during the check-in at the particular round. Let Ψ_i be the set of m ESS clients for round i . Here, $m = \max(C \times N, 1)$ is the number selected clients and the expression on the right side ensures that at least one client is selected. During the configuration and training process, the server initializes the weight of the shared CNN model, $w_0 \leftarrow w_t$, and distributes it to the selected clients in Ψ_i .

⁵The server can be selected arbitrarily and any one of the ESS clients, with a small non-volatile storage, can act as a server. As an assumption, both the clients and the server have communication devices and are guided via some communication protocol — for eg. network interface card with internet protocol.

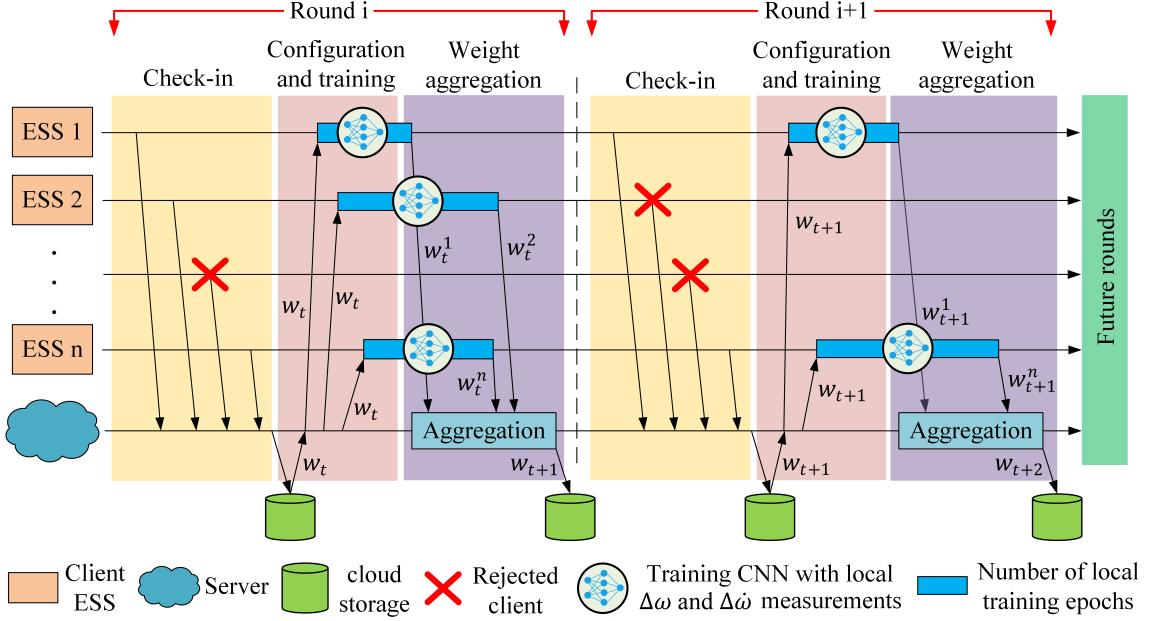


Figure 4.6. FL-based framework for inertia estimation in multi-area power system. The cloud storage is a non-volatile storage device that stores the aggregated CNN model and information about selected clients.

Initializing a common weight in the server is found to be effective than random initialization of weights in each of the clients [75], [79].

The learning rate (α) and local mini-batch size (b) are defined at each of the ESS clients n . The number of local epochs, \mathbb{E} , can be varied to achieve the best performance. Furthermore, \mathbb{E} can be varied for individual clients i.e. CNN at different clients can be trained, on local frequency measurements, for different number of epochs. Since, the server does not keep track of number of training epochs at the CNN of each client n , it is reasonable to posit that in real world scenario \mathbb{E} will be different for different clients. However, for simplicity, in this work \mathbb{E} is consistent for each of the clients. Let \mathbb{P}_n be the set of training snapshots at ESS client n , \mathbb{B} represents the batches of training snapshots where each batch $\beta \in \mathbb{B}$ is of size b . At each ESS clients n , the CNN model then estimates the inertia constant by minimizing MSE via backpropagation, using Eqs. (3.6) and (3.7). Finally, when the clients are online, the trained weights of each of the clients, w_{t+1}^n , are sent back to the server for aggregation. In this work, the aggregation is performed once

the server receives the trained weights of all selected clients.

Weight aggregation is the most important process in a communication round that makes FL different than other machine learning techniques. After collecting the trained weights, w_{t+1}^n , from all clients in Ψ_i , a weighted average method based on the number of data samples on each of the clients is given by:

$$w_{t+1} = \sum_{n=1}^N \frac{\mu_n}{\mu} w_{t+1}^n \quad (4.5)$$

where μ_n is the number of training snapshots at client n , and μ is the total number of training snapshots for N clients.

The server stores the aggregated weight, w_{t+1} , in its persistent storage as a checkpoint and the entire process is repeated for the next round, $i + 1$. The algorithm for FL is known as federated averaging due to its unique weight averaging method shown in (4.5). The pseudocode for federated averaging is given in Algorithm 1.

A schematic of the overall framework of inertia estimation in multi-area system using FL is shown in Fig. 4.7. A two-area power system model connected by a tie-line is used in this work. The power system at each area is perturbed by excitation signal ΔP_L ; non-simultaneous. The perturbation is fed only at one of the areas in which the frequency snapshots are to be observed while keeping the perturbation at the other area to zero i.e. in Fig. 4.7 ΔP_{L1} is an excitation signal with a given amplitude and frequency whereas $\Delta P_{L2} = 0$. The additional Gaussian noise signal is added in the measurement to mimic noisy PLL measurements. CNN located at each of ESS clients improves the estimation by minimizing MSE between actual value (M) and estimated value (\tilde{M}) using (3.6), and updates the model parameters using (3.7). However, the local weight updates corresponding to each of the clients, represented by w_t^1 and w_t^2 , are different. The trained updates are sent to the server for aggregation via a secured communication channel. The server then aggregates the weight and distributes the shared model to each of the ESS clients. This process repeats for several rounds until the global model converges.

Algorithm 1: Federated Averaging for Inertia Estimation

```

1 initialize the number of ESS clients  $N$ , fraction of ESS clients to be selected  $C$ ,
   number of local epochs at each of the ESS clients  $\mathbb{E}$ , local mini-batch size  $b$ , and
   learning rate  $\alpha$ 
2 Server:
3   initialize the weight of CNN  $w_0 \leftarrow w$ 
4   for each round  $i = 1, 2, \dots$  do
5     maximum ESS clients selected ( $m$ )  $\leftarrow \max(C \times N, 1)$ 
6      $\Psi_i \leftarrow$  random set of  $m$  ESS clients for each round
7     for each ESS client  $n \in \Psi_i$  do
8       // call function at line 14
9       gradient ( $w_{t+1}^n$ )  $\leftarrow$  ESS_Client( $n, w$ )
10    end
11    aggregate weight,  $w_{t+1} \leftarrow \sum_{n=1}^N \frac{\mu_n}{\mu} w_{t+1}^n$ 
12    distribute aggregated weight to the clients in  $\Psi_i$ 
13  end
14 Function ESS_Client( $n, w$ ):
15   // this function will run on each of the ESS Clients  $n$ 
16    $\mathbb{B} \leftarrow$  splitting entire training snapshots,  $\mathbb{P}_n$ , into batches of size  $b$ 
17   for  $l$  to  $\mathbb{E}$  do
18     for batch  $\beta \in \mathbb{B}$  do
19       // perform back propagation
20        $w_{t+1}^n \leftarrow w_t^n - \alpha \nabla E_{MSE}$ 
21     end
22   end
23   return  $w_{t+1}^n$ 

```

4.6 Simulation Setup

4.6.1 Overview

The modeling and simulation of the multi-area power system, along with data collection and pre-processing, was conducted in MATLAB/Simulink 2018b. The CNN model and FL framework was developed in Python using PyTorch, an open-source library for deep learning studies [66]. To leverage the fast computing abilities of PyTorch, the machine learning model was trained on South Dakota State University's Roaring Thunder cluster on NVIDIA Tesla P100/V100 GPUs. Although GPUs were used to train this model for speed of analysis, modern microcontrollers with ARM cortex cores have been

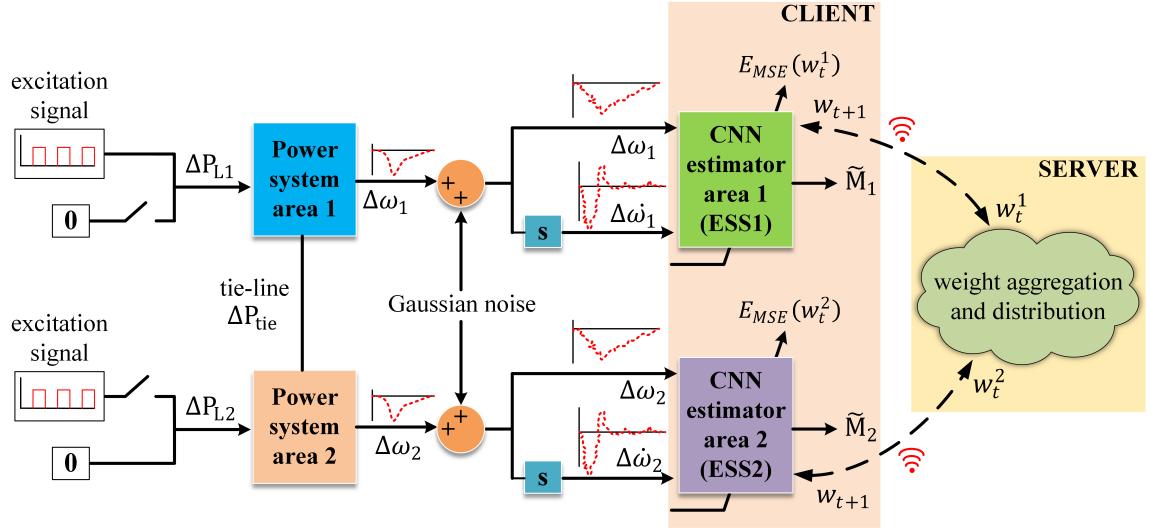


Figure 4.7. Overall schematic of FL-based inertia estimation in multi-area power system. It is to be noted that only one of the areas is perturbed at a time via ΔP_L to collect the snapshots.

successful in training deep CNN architectures and can be used in real-world implementations [67]. Furthermore, the microcontrollers can be extended with a wireless communication module to facilitate client-server-based communication for FL [80].

4.6.2 Simulation Benchmark

The transfer function-based two area system with an equivalent generator model given in Fig. 3.2 was used for the experiment to collect the frequency snapshots, and the respective simulation parameters are given in Table 4.1 [23], [59], [76]. In order to have some variation in the dataset, the snapshots were collected from both the areas with different values of M_1 and M_2 . Similarly, for each of the areas, excitation signals with 100 different values of ΔP_L from 10^{-3} p.u. to 0.1 p.u. with an increment 10^{-3} p.u. were used. As mentioned before the excitation signals were fed only at the area in which the snapshots were collected at that time, as shown in Fig 4.7.

To collect realistic data samples, white Gaussian noise was introduced in the signal using add white Gaussian noise (AWGN) block in MATLAB/Simulink. The signal-to-noise ratio (SNR) of 60 dB with a co-variance of 1e-6 was found to be

Table 4.1. Simulation Parameters for Multi-Area Power System

Parameters	Values
inertia constant for area 1 (M_1)	[2, 4, 6, 8] s
inertia constant for area 1 (M_2)	[3, 5, 7, 9, 10] s
change in load ($\Delta P_{L1}, \Delta P_{L2}$)	10^{-3} p.u. – 0.1 p.u. with a step of 10^{-3} p.u.
damping coefficient (D_1, D_2)	1.5
speed regulation droop for area 1 (R_1)	5%
speed regulation droop for area 2 (R_2)	10%
turbine-governor time constant (T_{g1}, T_{g2})	0.2 s
secondary controller gain (K_1, K_2)	$4.65e - 2$
bias factor for area 1 (B_1)	21.5
bias factor for area 2 (B_2)	11.5
tie-line coefficient (T_0)	1

appropriate for our setup as described in [69]. A total of 900 snapshots were collected from both of the areas. A sampling frequency of 200 Hz was selected as mentioned in the IEEE standard for frequency measurements [70]. Similarly, a sampling frame (inertial response time frame) of 1 s was used — from 31 s to 32 s — as the system took some time to show the steady-steady response towards the excitation signal.

4.6.3 Data Distribution and Hyperparameters Selection for FL

A sampling time of 200 Hz gives 200 data points ($c = 400$) for each of the snapshots of $\Delta\omega$ and $\Delta\dot{\omega}$, extracted at a sampling frame of 1 s. To have a comparison on the basis of communication rounds and architecture, the simulation was carried on both CNN and MLP, for which the architecture and hyperparameters were selected as described in Chapter 3⁶. This section describes a general way of data partition for IID and non-IID case and hyperparameters selection for FL framework.

To test the global model on validation set for each communication rounds, the overall data was splitted into two parts — 720 snapshots ($\sim 80\%$) for training and 180 snapshots ($\sim 20\%$) for validation. The general way of having an IID case would be to

⁶The architecture and hyperparameters of CNN and MLP are similar to the ones used in Chapter 3 and are properly defined in this work in case of any changes.

distribute the snapshots to each of the areas so as to have near equal probability distribution for each value of inertia constant. To achieve this, the training dataset was randomly shuffled and re-distributed so that each ESS clients contains 350 random snapshots. For non-IID case, the 720 training snapshots were arranged in ascending order of M and distributed in equal parts to the ESS clients. This ensure that the two ESS clients have snapshots corresponding to different values of M , for eg. the snapshots corresponding to $M = 2$ s is in area 1 but not in area 2.

The effectiveness of federated averaging algorithm depends on the three hyperparameters — C , \mathbb{E} , and b [75]. Since $N = 2$, we select the value of $C = 1$, which means that all of the clients are selected during any communication round. Furthermore, we experimentally verified that $b = 10$ works well for both CNN and MLP. When some level of accuracy is desired from the global model, the algorithm can be stopped at the particular communication round in which the global accuracy is obtained. However, for inertia estimates, we assume that the estimated value within 10% of the actual value is a correct value. Hence, we pre-defined i before conducting the simulation in this work. The simulation conducted for different values of \mathbb{E} and i for MLP and CNN are presented in Section 4.7. The combinations of i and \mathbb{E} are so chosen to get a similar value of RMSE on validation set.

4.7 Results and Analysis

4.7.1 Performance metrics for IID data

For IID case, the RMSE values on validation set of 180 snapshots for MLP and CNN is shown in Table 4.2. When $\mathbb{E} = 1$, it takes 200 communication rounds between the server and the ESS clients to achieve an RMSE of 0.3652 when trained on CNN. However, when trained on MLP an RMSE of 0.4387 is obtained with 1000 communication rounds. To get the similar level of model performance MLP requires higher number of client-server communication than CNN.

Fig. 4.8 shows the evolution of aggregated weights of the global model, obtained via

Table 4.2. RMSE of estimated values of inertia on validation set. The presented value is the RMSE observed for i^{th} communication round where the neural networks at the ESS clients are trained for \mathbb{E} number of local epochs on IID data.

MLP			CNN		
$i = 1000$	$i = 200$	$i = 100$	$i = 200$	$i = 100$	$i = 40$
$\mathbb{E} = 1$	$\mathbb{E} = 5$	$\mathbb{E} = 10$	$\mathbb{E} = 1$	$\mathbb{E} = 2$	$\mathbb{E} = 5$
0.4387	0.4482	0.4484	0.3652	0.3534	0.3914

federated averaging, for IID case when $\mathbb{E} = 1$. It can be seen that the weights are saturated at $\sim i = 175$ for CNN (approximately 5.7 times less than MLP) whereas some of the weights do not converge for MLP-based model even when $i = 1000$.

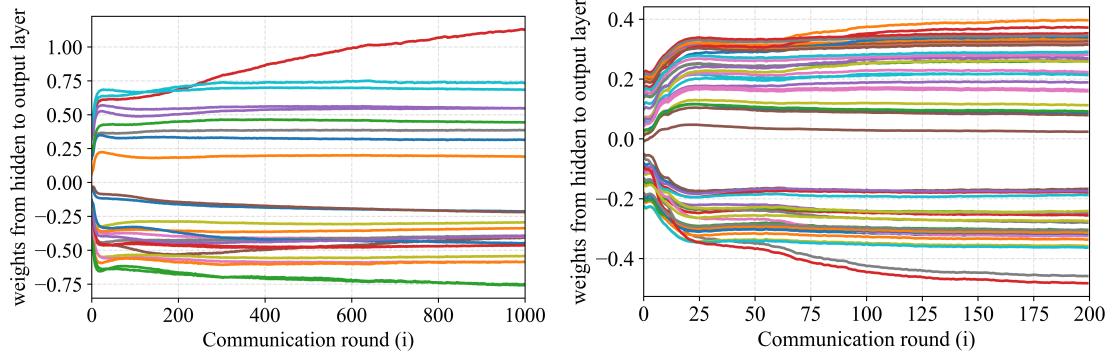


Figure 4.8. Evolution of aggregated weight, w_{t+1} , using federated averaging for each communication round i between the server and ESS clients for MLP-based training(left) and CNN-based training(right). The plots represent the IID case when $\mathbb{E} = 1$.

Similarly, Fig. 4.9 shows the accuracy of the aggregated model on validation set at the end of each communication rounds for IID case when $\mathbb{E} = 1$. The MLP-based model gave a validation accuracy of 95% at $i = 1000$ and CNN gave a validation accuracy of 96.67% at 200 epochs. Hence, based on the results above, CNN based estimator outperforms MLP in terms of communication cost and RMSE for IID data.

4.7.2 Performance metrics for non-IID data

For non-IID case, the RMSE values on validation set of 180 snapshots for MLP and CNN is shown in Table 4.2. When $\mathbb{E} = 1$, it takes 200 communication rounds between the server and the ESS clients to achieve an RMSE of 0.372 when trained on CNN. However,

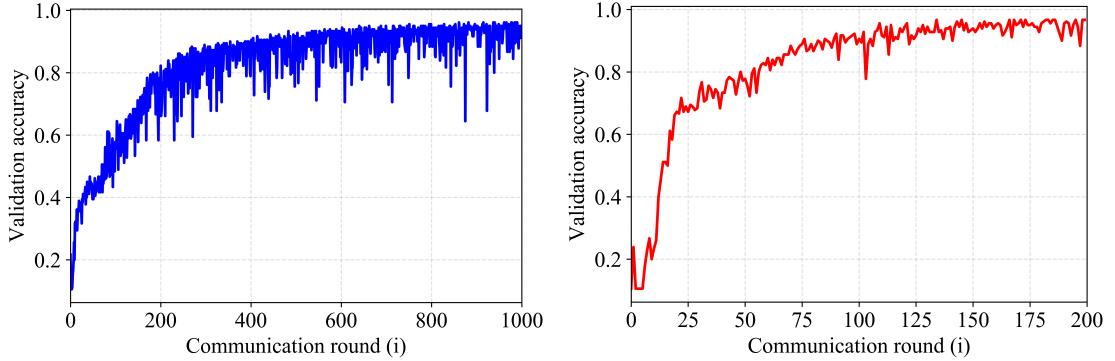


Figure 4.9. Aggregated model accuracy on validation set for each communication round i between the server and ESS clients for MLP-based training(left) and CNN-based training(right). The plots represent the IID case when $\mathbb{E} = 1$.

on MLP, an RMSE of 0.3851 is obtained with 1000 communication rounds.

Table 4.3. RMSE of estimated values of inertia on validation set. The presented value is the RMSE observed for i^{th} communication round where the neural networks at the ESS clients are trained for \mathbb{E} number of local epochs on non-IID data.

MLP			CNN		
$i = 1000$	$i = 200$	$i = 100$	$i = 200$	$i = 100$	$i = 40$
$\mathbb{E} = 1$	$\mathbb{E} = 5$	$\mathbb{E} = 10$	$\mathbb{E} = 1$	$\mathbb{E} = 2$	$\mathbb{E} = 5$
0.3851	0.372	0.3876	0.3982	0.5372	0.401

Fig. 4.10 shows the evolution of aggregated weights of the global model, obtained via federated averaging, for non-IID case when $\mathbb{E} = 1$. Similarly, Fig. 4.11 shows the accuracy of the aggregated model on validation set at the end of each communication rounds. Similar to the IID case, it can be seen that the weights are saturated at $\sim i = 175$ for CNN whereas some of the weights do not converge for MLP-based model even when $i = 1000$.

From the above analysis, it is interesting to observe that for two-area ESS clients, the performance metrics do not have much difference for the IID and the non-IID case. Fig. 4.12 shows the comparison of validation accuracy for IID and non-IID case when $\mathbb{E} = 5$. The convergence for IID case is slightly better than the non-IID case. However, it is also important to notice that in this work we have distributed the data to the ESS clients by manually separating the data to IID and non-IID fashion in a controlled environment.

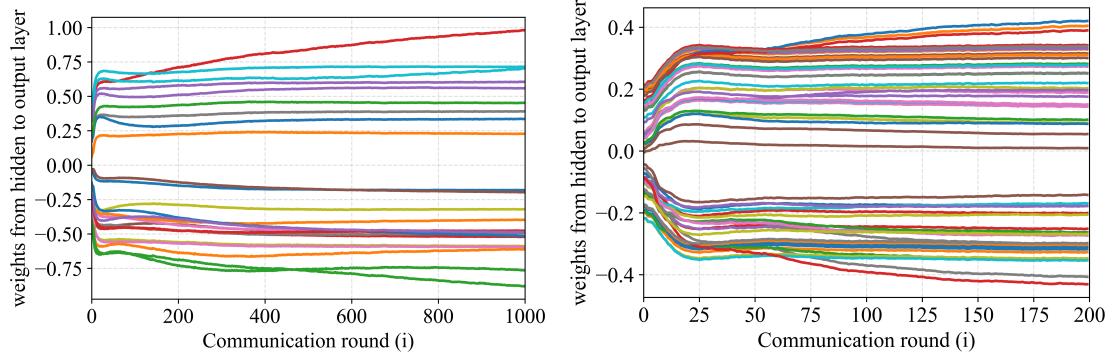


Figure 4.10. Evolution of aggregated weight, w_{t+1} , using federated averaging for each communication round i between the server and ESS clients for MLP-based training(left) and CNN-based training(right). The plots represent the non-IID case when $\mathbb{E} = 1$.

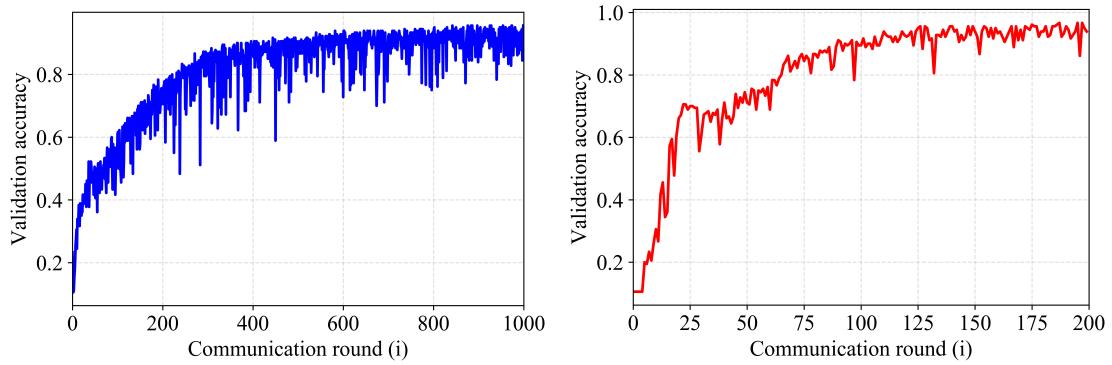


Figure 4.11. Aggregated model accuracy on validation set for each communication round i between the server and ESS clients for MLP-based training(left) and CNN-based training(right). The plots represent the non-IID case when $\mathbb{E} = 1$.

Such scenario is not possible in the real-world scenarios and hence, the results might be significantly different when hundreds of clients are considered with highly non-IID data. However, it is important to analyze that the global model was successful to estimate the inertia constant without being trained on several snapshots.

4.7.3 Communication Cost

Communication overhead incurs the highest optimization cost in FL [75]. Although FL rejects the clients that are unable to provide the update or are offline during particular instant of communication, the cost of communication overhead still overpowers the individual computational cost on the client as well as the cost to add an additional client in

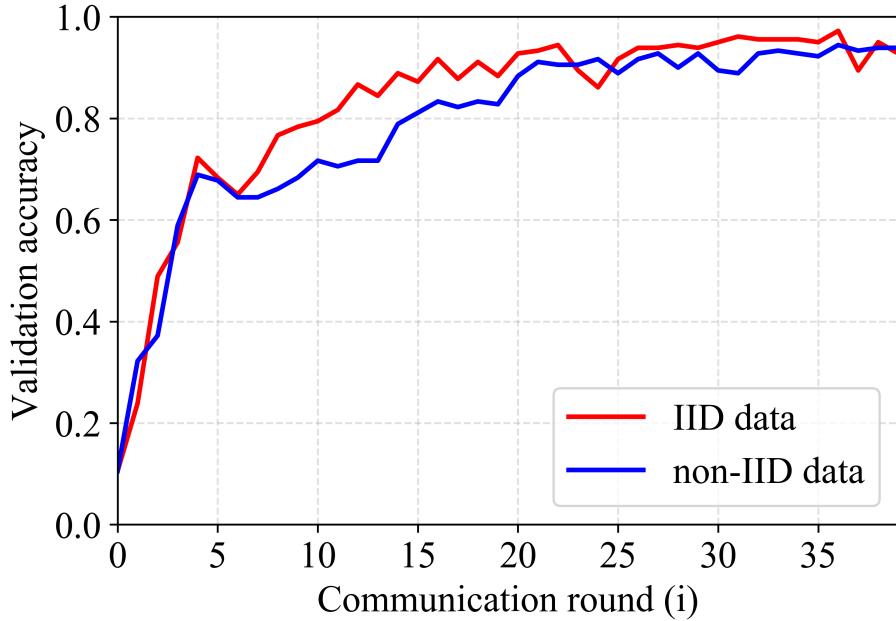


Figure 4.12. Comparison of accuracy for IID and non-IID data when $i = 40$ and $\mathbb{E} = 5$.

the framework [75], [78]. In FL the communication cost can be drastically reduced by increasing the number of \mathbb{E} to certain extent. Fig. 4.13 shows the validation accuracy of a CNN model for IID data with respect to i for different values of \mathbb{E} . It can be seen that when the value of \mathbb{E} is increased from 1 to 5, the number of communication round to achieve desired accuracy reduces drastically. When $\mathbb{E} = 5$ only 40 communication rounds would suffice to achieve an accuracy beyond 95%. This is a decrease in communication round by a multiple of 5 as compared to the case when $\mathbb{E} = 1$. Furthermore, the CNN model with $\mathbb{E} = 5$ was approximately 6 times computationally efficient than MLP counterpart. Therefore, when $E = 5$, CNN-based FL framework is 30 times more efficient than the MLP-based framework. Hence, the FL learning framework can be made more efficient by decreasing the number of communication rounds between the server and the ESS clients and simultaneously increasing the value of \mathbb{E} .

Furthermore, early stopping technique, as discussed in Chapter 3, can be used when the desired accuracy is obtained. In Fig. 4.13, the training at the client ESS and the communication between the client ESS and the server can be stopped when $i = 36$

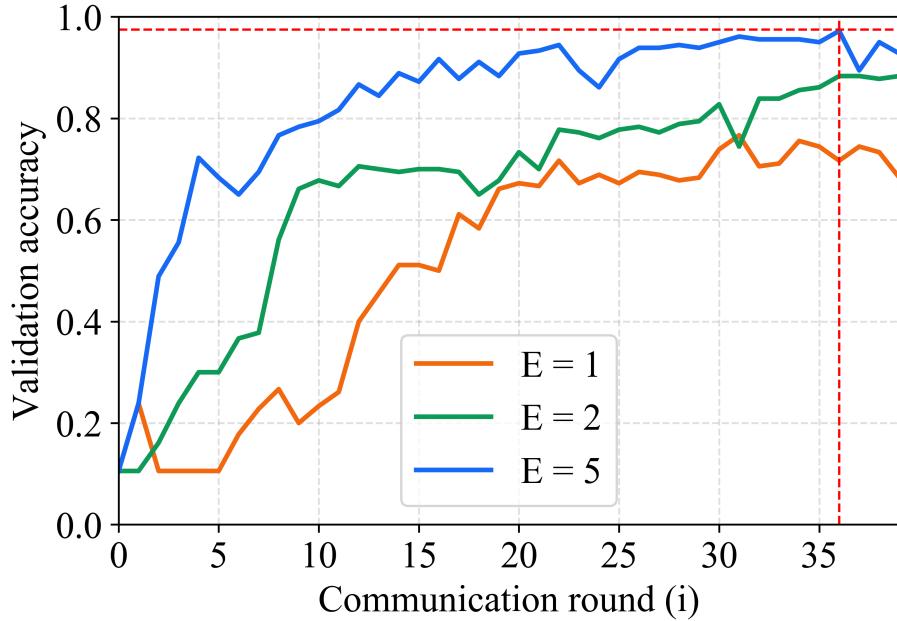


Figure 4.13. Validation accuracy of a CNN-based model on IID data with respect to i for different values of \mathbb{E} . On varying the values of \mathbb{E} , the number of communication rounds can be controlled in FL.

(represented by a dashed vertical red line with an identical intersecting line showing the equivalent accuracy) in order to get the maximum accuracy.

4.8 Conclusion

In this work, the inertia constant was estimated in a multi-area power system using federated averaging algorithm. The simulation was conducted and verified for two neural network architectures — MLP and CNN. The frequency snapshots were collected at PLL of each of the ESS clients using non-intrusive excitation signals. It was found that MLP takes greater number of communication rounds as compared to CNN to get a similar level of accuracy. Furthermore, the framework was verified to perform well for both IID and non-IID data with significant accuracy, which is important in the field of power systems that contains highly non-IID data. Both the MLP and CNN-based inertia estimators showed good performance even with noisy input samples. It was also verified that the number of communication rounds between the ESS clients and the server can be drastically reduced by increasing the number of local epochs, \mathbb{E} at each of the clients. The

communication cost was improved by 30 times while using a CNN model with $\mathbb{E} = 5$ as compared to the MLP model with $\mathbb{E} = 1$.

CHAPTER 5 INERTIA ESTIMATION OF A SYSTEM WITH NON-SYNCHRONOUS UNIT

5.1 Introduction

As discussed in Chapter 1 and Chapter 3, due to the penetration of non-synchronous generations units, such as RES-based generation, the power system inertia has become a time-varying quantity and continuously decreasing. However, recent advancements in power systems control leverage RES to provide VI to the grid [6], [81]–[84]. Furthermore, incentivizing the grid-connected devices for providing the grid ancillary services, like frequency regulation and VI, is also an active area of research [85], [86]. Fig. 5.1 shows a general frequency response of a power grid corresponding to a loss in generation or increase in load. The solid plot in green represents the response with VI and the dashed plot corresponds to no VI case. It can be seen that with VI support, the ROCOF can be controlled thereby reducing the frequency nadir.

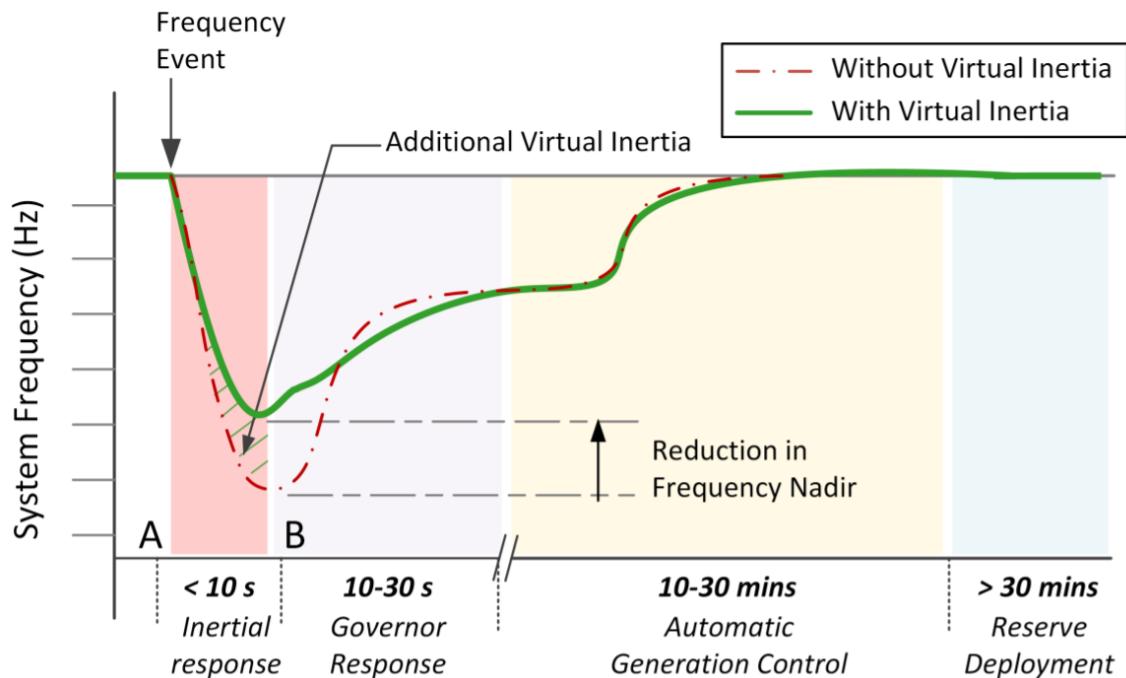


Figure 5.1. Frequency response of a power system with VI support from additional non-synchronous unit.[6].

FFR is a self-deployed and controlled contribution of electrical power, generally rapid increase or decrease in active power, from a unit to quickly counteract the initial frequency change in the system as a result of reduced inertial response [36]. However, until now, the quantification of inertia from FFR units have been an open area of research. Although the FFR units supply necessary power to slow down the initial ROCOF during a frequency event, it is necessary to quantify the support in terms of value for optimized control from system operator's point of view.

Recently, model predictive controller (MPC) has gained popularity over traditional PI-based frequency control devices [15]. MPC-based devices can be termed as FFR units that can quickly supply the inertial response to maintain the frequency stability in a grid. Furthermore, the such devices are non-synchronous in nature and hence traditional swing-equation based power systems model can not capture the effect of inertial response due to non-synchronous units.

5.2 Proposed Work

This work is an extension of the work presented in Chapter 3. In this work, a pre-trained CNN model trained on frequency snapshots of a single area power system is used to estimate the equivalent inertia constant of a system with an additional non-synchronous unit. An MPC-based VI unit developed in [15] is used as a plug-and-play device to represent the non-synchronous unit. The pre-trained CNN model is used to estimate the inertia for different combinations of known inertia constant of synchronous units and excitation signals. Therefore, the major contribution of this work is to estimate the power system inertia incorporating the effect of non-synchronous units in the frequency response. The proposed approach is verified and analysed in a transfer function-based single area power system with additional MPC-based VI unit.

5.3 Inertia Estimation of a System with MPC Unit

5.3.1 MPC-based Single Area System

Fig. 5.2 shows a modified transfer function-based single area power system with a MPC-based VI unit for FFR support. Unlike the system in Fig. 3.2, the system shown in Fig. 5.2 contains an additional MPC unit that takes the frequency measurements as input and provides the necessary VI power, Δp_{vi} , to the grid. The problem is formulated as an optimization problem on some weighted cost function. In this work, the weights of the cost functions are defined to penalize the estimate of measured frequency from the power system model and power output from the MPC unit. Furthermore, the performance of the MPC unit depends on the prediction and control horizon along with sampling time. In this work, such parameters are carefully selected to match the need of the simulation. More details on the design and formulation of MPC, used in this work, can be found in [15].

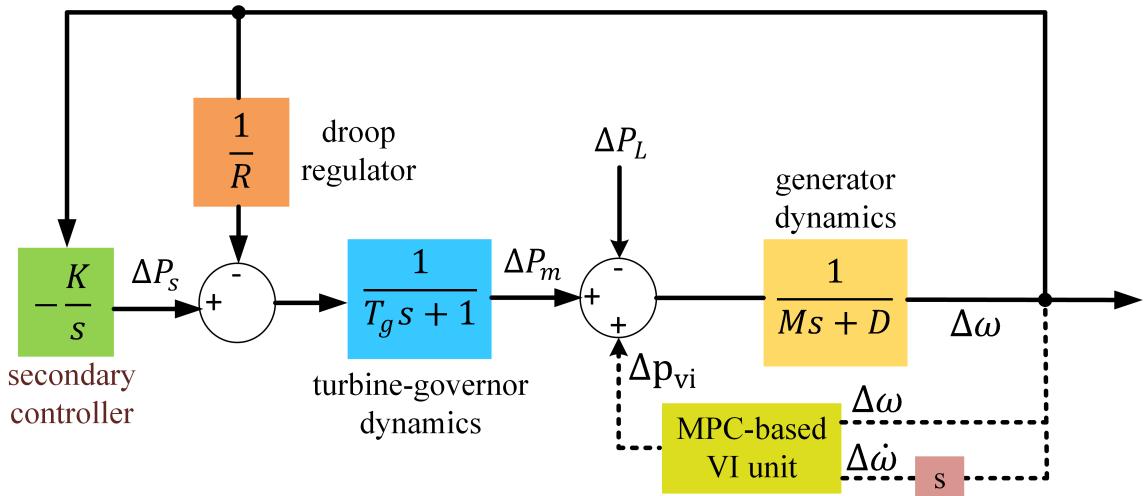


Figure 5.2. Transfer function-based single area power with additional MPC-based VI unit

In this work, the MPC unit has been used as a black-box unit so as to generalize the functioning of the CNN estimator towards any other FFR units. The MPC unit emulates the inertial response during first few seconds of a frequency event. The controller acts as an FFR unit that quickly counteracts the change in frequency. Fig. 5.3 shows a general response of the turbine-governor block and MPC unit of the system towards ΔP_L .

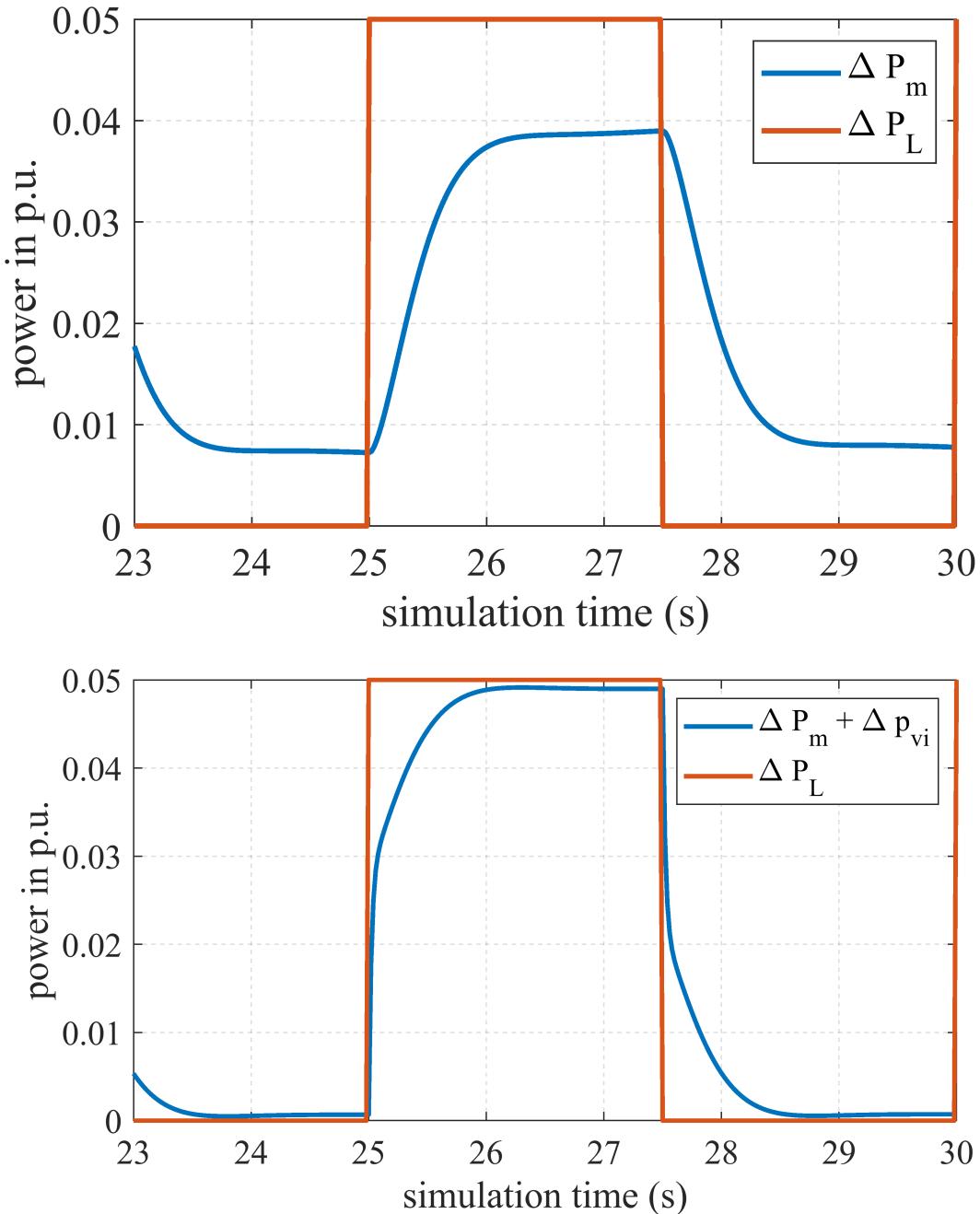


Figure 5.3. ΔP_m (upper) and $\Delta P_m + \Delta p_{vi}$ (lower) observed as a response to an excitation signal, ΔP_L , in the form of a pulse train with amplitude of 0.05 p.u. and time period of 5 s fed to a system with $M = 5\text{s}$.

The responses of different power signals are observed for the system in Fig. 5.2 with $M = 5\text{s}$ and $\Delta P_L = 0.05\text{ p.u.}$ It is clear from former discussions that perturbing the system with ΔP_L triggers a change in frequency which is counteracted by the primary and

secondary frequency control loops of the system. However, Fig. 5.3 shows that the control loops response solely cannot fulfill the overall loss in generation at a cycle of ΔP_L . On the other hand, MPC-based VI unit can provide the FFR support almost instantaneously, which is shown by $\Delta P_m + \Delta p_{vi}$ curve in the figure. Hence, a drastic change in the frequency response of power systems can be expected due to added response from the integrated non-synchronous units.

5.3.2 Inertia Estimation Using Pre-Trained CNN

Fig. 5.4 shows a general schematic of inertia estimation in a system with additional non-synchronous unit — MPC-based VI unit. The CNN estimator is pre-trained on number of snapshots, by varying M and ΔP_L , on a single area system with synchronous units as described in Chapter 3. Here, ΔP_L is a pulse train of varying amplitude. To mimic noisy measurements in the PLL, white Gaussian noise has been added to $\Delta\omega$. However, in the later part of the experiment the simulation has also been conducted with ΔP_L as a step change instead of a pulse train without considering any noisy frequency measurements. The MPC-unit continuously takes $\Delta\omega$ and $\Delta\dot{\omega}$ as inputs and provides the control signal in the form of Δp_{vi} as discussed in the former sections.

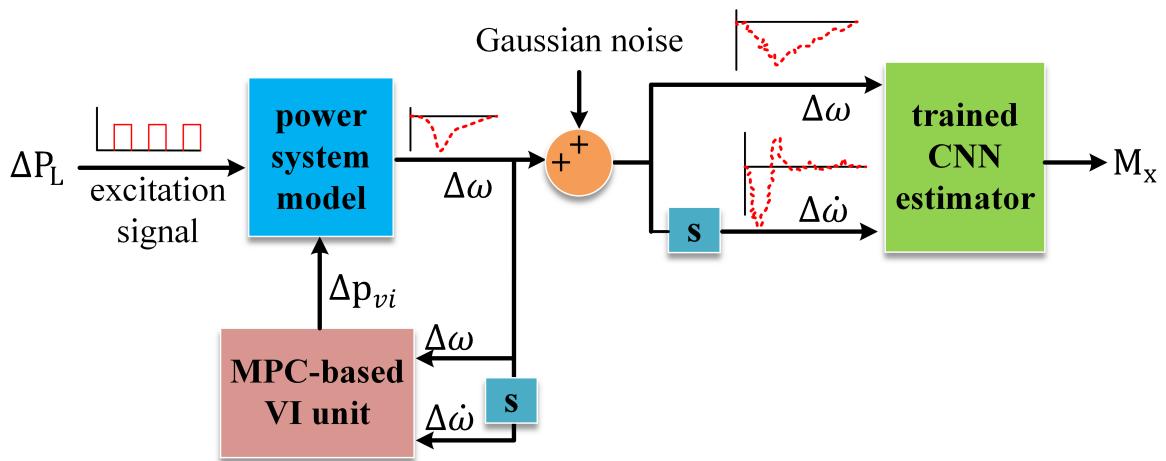


Figure 5.4. Schematic of inertia estimation in an MPC-based power system using pre-trained CNN. The CNN has been pre-trained on a power system with 1700 frequency snapshots corresponding to different values of M and ΔP_L .

Here, $M_x = M + M_{vi}$ is the equivalent inertia of the system with MPC-based VI unit,

where M is the inertia constant of the equivalent synchronous units and M_{vi} is the inertia constant associated with the VI unit. Hence, the pre-trained CNN estimates M_x from which one can calculate the value of M_{vi} .

The analysis is done for two cases on which the equivalent inertia of the system, M_x , is estimated via a pre-trained network. In the first case, the simulation is performed for a specific value of M and ΔP_L in the form of pulse train for which M_x is estimated. Furthermore, the single area system in Fig. 3.2 is re-simulated with $M = M_x$ to compare the frequency response with and without the additional VI unit. In the second case, Fig. 5.4 is simulated with parameter sweep over M and ΔP_L to get a series of M_x from a pre-trained CNN. It was observed that when ΔP_L was used as a pulse train then there is a mismatch in the steady state values of the frequency responses for the system with VI and without VI unit. Hence, to avoid using any offset and to ensure a common trigger point for all of the snapshots, the simulation in the second case was conducted with ΔP_L as step change. Additionally, the simulation was conducted on a high sampling rate and hence the Gaussian measurement noise was removed to avoid major inaccuracies in the estimates due to high sampling rate. Finally, the estimates are analyzed to get a better understanding of the response provided by the MPC-based VI unit for both cases.

5.4 Simulation Setup

The modeling and simulation of the power system, along with data collection and pre-processing, was conducted in MATLAB/Simulink 2018b. The MPC was formulated and implemented in MATLAB/Simulink 2018b using CVXGEN that generates C codes for convex optimization⁷. The CNN model was developed in Python using PyTorch, an open-source library for deep learning studies [66]. To leverage the fast computing abilities of PyTorch, the machine learning model was trained on South Dakota State University's Roaring Thunder cluster on NVIDIA Tesla P100/V100 GPUs. Although GPUs were used

⁷The modeling and formulation of MPC-based VI unit has been done by Dr. Ujjwol Tamrakar, who is a former graduate student at SDSU. In this work, the MPC-unit is used as a plug-and-play device and is taken from [15].

to train this model for speed of analysis, modern microcontrollers with ARM cortex cores have been successful in training deep CNN architectures and can be used in real-world implementations [67].

5.4.1 Case 1: Simulation on a Single Frequency Snapshot

A pre-trained CNN model, trained on 1700 frequency snapshots obtained from a single area system given in Fig. 3.2, is used to estimate M_x based on the model given in Fig. 5.2. A detail description on training the CNN model is given in Chapter 3. The typical parameter values used in the simulation are given in Table 5.1. To collect realistic data samples, white Gaussian noise was introduced in the signal using AWGN block in MATLAB/Simulink. The SNR of 60 dB with a co-variance of 1e-6 was found to be appropriate for our setup as described in [69]. The sampling frequency 200 Hz and a sampling frame (inertial response time frame) of 1 s was used — from 31 s to 32 s — as the system took some time to show the steady-steady response towards the excitation signal. Furthermore, the prediction and control horizon of 1 s performed well for this case. As the MPC unit is used as a plug-and-play device, the weights associated with the cost function are not altered in this work.

Table 5.1. Case 1: Simulation on a Single Snapshot
Summary of Simulation Parameters

Parameters	Values
inertia constant (M)	5 s
change in load (ΔP_L)	0.05 p.u.
damping coefficient (D)	1.5
speed regulation droop(R)	5%
turbine-governor time constant(T_g)	0.2 s
secondary controller gain (K)	2

5.4.2 Case 2: Simulation on Multiple Frequency Snapshots

To estimate the multiple values of M_x , the CNN model, described in Chapter 3, was trained with sampling rate of 50 Hz that gives 50 data points ($c = 100$) for each of the snapshots of $\Delta\omega$ and $\Delta\dot{\omega}$. Unlike previous simulations, a step change of ΔP_L is introduced

at 5 s for each simulation. The frequency snapshots, with sampling frame from 5 s – 7 s, were collected by simulating the model in Fig. 3.2 with typical parameter values given in Table 5.2. The sampling frame of 2 s is used to ensure that the inertial response for all different values of M are captured.

Table 5.2. Case 2: Simulation on Multiple Snapshots
Summary of Simulation Parameters

Parameters	Values
inertia constant (M)	2 s – 10 s with a step of 0.5 s
change in load (ΔP_L)	10^{-3} p.u. – 0.1 p.u. with a step of 10^{-3} p.u.
damping coefficient (D)	1.5
speed regulation droop(R)	5%
turbine-governor time constant(T_g)	0.2 s
secondary controller gain (K)	2

For CNN, we experimentally determined that two convolution layers with $p = 10$ and $q = 20$, and kernels with sizes $R = S = 3$ gave the best results. Similarly, for the feed forward layers, the size of the flattened layer depends on the architecture of the convolution layers [63]. For this specific architecture, with $c = 200$ and $q = 20$, the value of Y was calculated to be 3960, whereas selecting $h_1 = 500$ and $h_2 = 50$ improved the performance of the network. The training was performed in batches with $b = 30$. In this work, rectified linear unit (ReLU) activation was used for convolution layers and tanh for feed forward network. A modified version of weight update algorithm was used with $\alpha = 1e-3$, momentum 0.5, and weight decay of 5e-4. The simulation was carried out for 200 epochs for CNN by dividing the entire dataset into two parts — 1360 snapshots ($\sim 80\%$) for training and 340 snapshots ($\sim 20\%$) for validation.

Simultaneously, the system in Fig. 5.4 was simulated with a similar parameters given in Table 5.2 to collect another set of frequency snapshots. These snapshots, however, include the additional response from the MPC-based VI unit. Finally, the trained CNN model on above parameters estimated M_x for 1700 different snapshots.

5.5 Results and Analysis

5.5.1 Case 1: Simulation on a Single Frequency Snapshot

Based on the simulation parameters for $M = 5$ s and ΔP_L as a pulse train with an amplitude of 0.05 p.u., the pre-trained CNN estimated M_x to be 6.3708 s. In order to validate the inertial response, a single area system was simulated with $M = M_x = 6.3708$ for same ΔP_L . Fig. 5.5 shows the frequency and ROCOF response for the system with and without VI unit. However, due to different steady-state values for the two systems the point of trigger at each cycle of ΔP_L is different. Furthermore, the response seems to match for an initial fraction of seconds which is more prominent for $\Delta\dot{\omega}$ in Fig. 5.5.

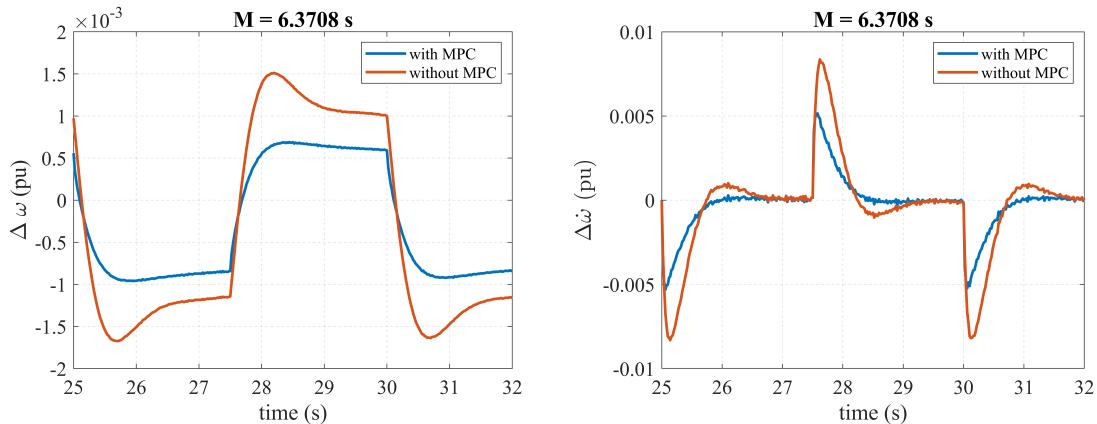


Figure 5.5. $\Delta\omega$ and $\Delta\dot{\omega}$ response for system with and without MPC-based VI unit. For the system with VI unit, $M = 5$ s whereas for the system without VI unit $M = M_x = 6.3708$ s. ΔP_L is used as a pulse train signal of amplitude 0.05 p.u.

To observe the response at same trigger point, the simulation was re-conducted for same value of $M = 6.3708$ but with a step change ΔP_L of amplitude 0.05 p.u. at 5 s. Although the response is more clear for a step change ΔP_L , the conclusion on the mismatch in response for two system remains unchanged. Although a close response is seen for a fraction of seconds, the overall response for a system with and without non-synchronous unit differs by a great extent.

Fig. 5.7 shows the frequency responses for different values of M on a single area synchronous unit-based system compared to a similar system with an MPC unit.

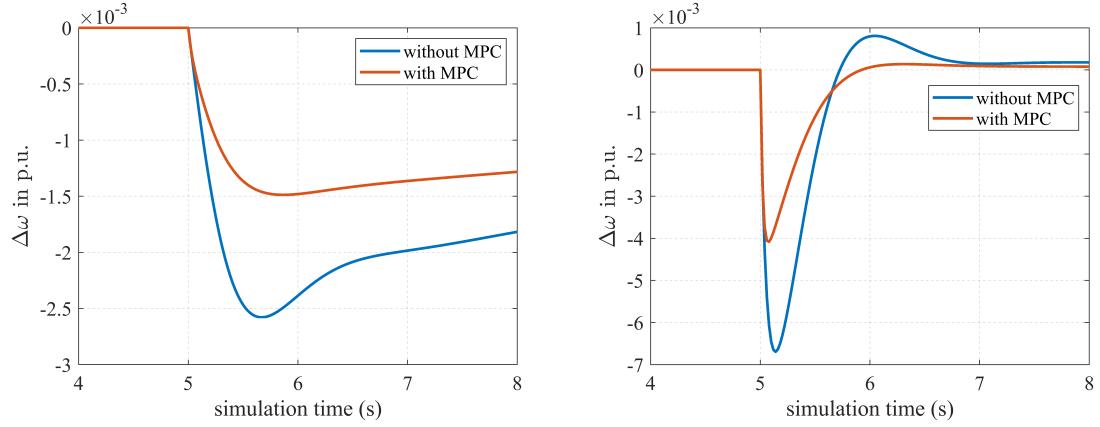


Figure 5.6. $\Delta\omega$ and $\Delta\dot{\omega}$ response for system with and without MPC-based VI unit. For the system with VI unit, $M = 5$ s whereas for the system without VI unit $M = M_x = 6.3708$ s. ΔP_L is used as a step change signal of amplitude 0.05 p.u. introduced at 5 s.

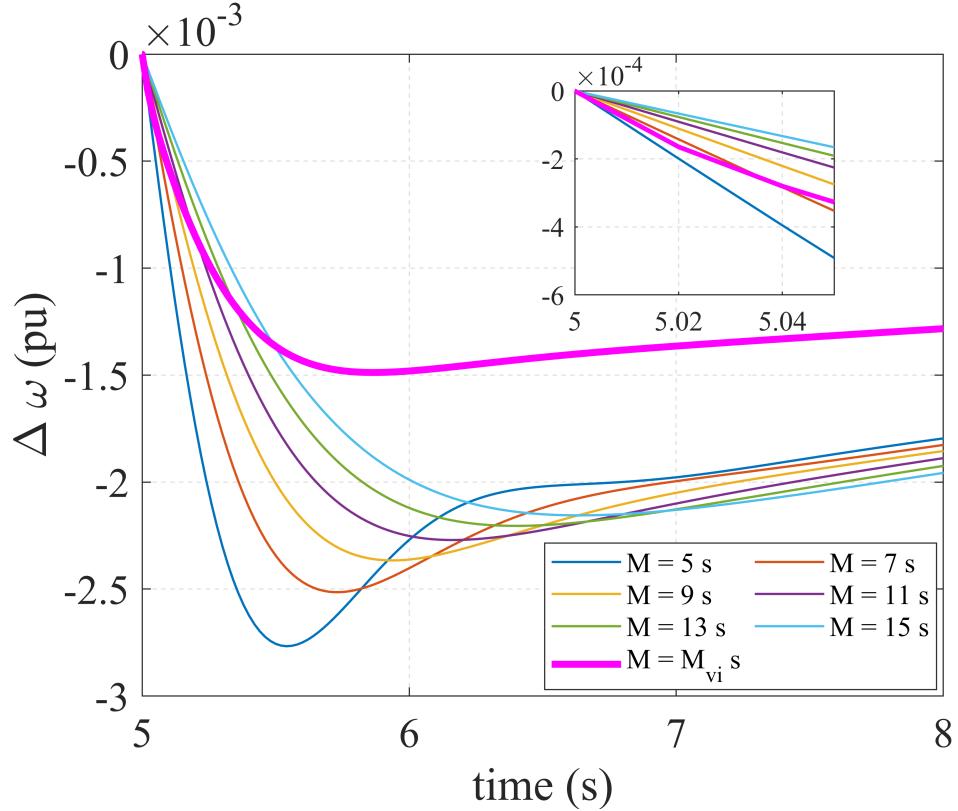


Figure 5.7. Frequency responses for different values of M on a single area synchronous unit-based system compared with a similar system with an additional non-synchronous unit. Here, $M_{vi} = M_x = 6.3708$ s represents the equivalent inertia constant of the system with additional VI unit estimated by a pre-trained CNN. ΔP_L is used as a step change signal of amplitude 0.05 p.u. introduced at 5 s.

It is interesting to see that the response corresponding to $M_{vi} = M_x = 6.3708$ s matches with the frequency response of $M = 7$ s for a fraction of seconds before deviating by a huge margin. However, the response with additional non-synchronous unit does not match with any of the other frequency responses corresponding to $M = 5$ s to $M = 5$ s with an increment of 2 s. This concludes that the additional MPC-based VI unit provides some additional response in addition to VI. Furthermore, the inertial response of conventional synchronous generator-based system differs significantly as compared to the response observed on a system with additional non-synchronous units with FFR support. Hence, it is desired to analyze some other parameters associated with the VI unit in addition FFR support. In Chapter 3 we discussed that the change in generator's damping can affect the frequency nadir. It can be seen in Fig. 5.7 that the response with additional non-synchronous unit also adds the damping to the system. Hence, as an important analysis, the damping from the additional unit, in addition to the inertial response, should also be considered.

5.5.2 Case 2: Simulation on Multiple Frequency Snapshots

Fig. 5.8 shows the evolution of weights and validation accuracy of CNN model trained on frequency snapshots taken at a sampling rate of 50 Hz. The weights are fairly saturated with a reasonable validation accuracy of 94% at the end of 200 epochs. The training accuracy, however, is slightly lower than the former cases. It is to be noted that a sampling frame of 2 s is for different values of M and ΔP_L . A sampling frame of 2 s could capture an entire sample from the point of event to restoration for cases with high M and ΔP_L . Hence, this could saturate the training accuracy to some extent as unwanted input features might mislead the CNN estimator.

Fig. 5.9 shows the distribution of equivalent inertia (M_x) estimated by the pre-trained CNN model, described above. The x-axis represents the inertia constant value of equivalent synchronous units (M) where y-axis represents the equivalent inertia constant of the system with synchronous as well as non-synchronous VI units i.e.,

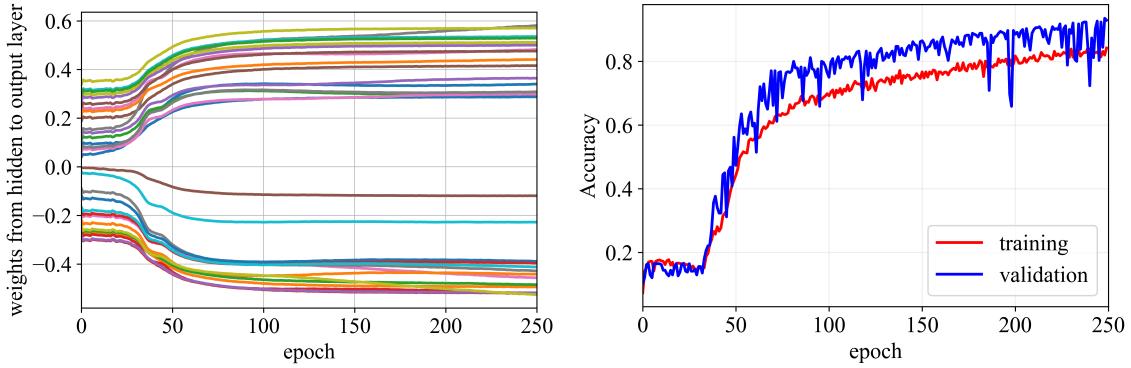


Figure 5.8. Evolution of weights (left) and accuracy (right) of a CNN model trained for 200 epochs on frequency snapshots of single area system with sampling rate of 50 Hz.

$M + M_{vi}$. It is interesting to see that for the value of $M \geq 9.5$ s, the value of $M_x \leq M$, which shows that the MPC-based VI unit provides the additional response to counteract the decreasing inertia. This is even more prominent for lower values of M . However, for $M \leq 9.5$ s, the distribution of M_x is similar to the corresponding value of M . This could be due to the fact that at higher values of M , the frequency control loops in the power system would handle the decrease in inertia with limited VI support.

Additionally, the estimates are mostly clustered with a few outliers. However, the outliers are not much affecting the distribution as the interquartile range of each of the estimates are almost equally distributed and are close to the median value.

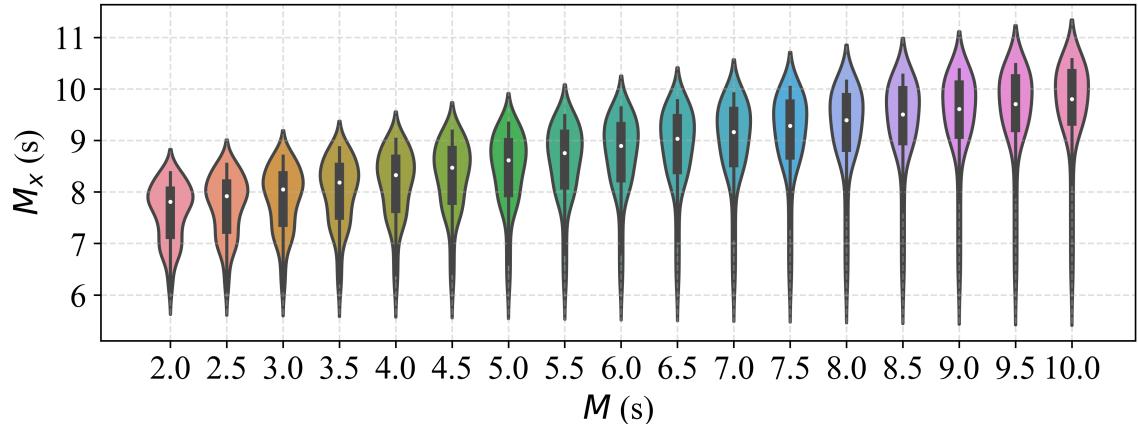


Figure 5.9. Distribution of equivalent inertia estimates (M_x) estimated by a pre-trained CNN model. The estimated values are for 1700 snapshots trained on an equivalent single area system with MPC-based VI unit.

5.6 Conclusion

In this work, a pre-trained CNN was leveraged to estimate the inertia constant for a system with additional non-synchronous VI units. an MPC-based VI unit was used as a plug-and-play unit to observe the frequency response due to non-synchronous unit. The simulation was conducted for two different cases. In the first case, the simulation was conducted for a single snapshot and the frequency response was compared between the system with and without VI unit. It was found that the response changes drastically with an additional VI unit and it is required to analyze some other parameters associated with the VI unit in addition to VI. In the second case, the equivalent inertia constant with additional non-synchronous unit was estimated using a pre-trained CNN for number of snapshots. The snapshot were collected by varying M and Δ_P . It was found that for higher value M in the system, the response from VI units is minimal as compared to the response when M is low.

CHAPTER 6 CONCLUSIONS

The research work proposed a general model-free data-driven approach to estimate power system inertia in a multi-area system incorporating the response due to interconnected non-synchronous units. Considering the complexity of power system models with variable RES, the proposed approach can be beneficial to estimate the inertia constant in multi-area system with interconnected tie-lines. In addition, the challenges of combined voltage and frequency control in LV grids were presented. In LV grids, the modulation of active power changes the grid voltage drastically in addition to frequency. With unprecedented growth in the number of DERs, the problem of coupling effect is even more prominent that can increase the grid vulnerability towards voltage and frequency fluctuations. Hence, the combined voltage and frequency control techniques in LV grids should consider the coupling effect and multi-timescale dynamics between the two parameters. Additionally, grid-connected DERs should be incentivized considering their future roles in grid ancillary services.

Furthermore, with increasing penetration of RES the power system inertia is constantly decreasing. The stochastic RES induces the variability in the inertia constant which can create issues for the system operators to procure fast frequency services on time. Additionally, RES penetration increases the complexity in modeling a power system. Hence, data-driven approaches would be helpful in cases where the complexity in developing a model is extremely high. Unlike the traditional neural network models, CNN-based model can identify the spatial information CNN estimated the inertia constant with just frequency measurements taken from the PLL of ESS, while not requiring ΔP_L as an input. The CNN-based estimator estimated the power system inertia with an accuracy of accuracy of 97.35% and root mean square error of 0.2309.

Apart from the single area inertia estimation, the CNN estimator worked well for two-area system on client-server-based FL framework. The ESS on specific areas were

selected as clients whereas a random server communicated with the clients to update a global aggregated model through federated averaging. It was found that CNN-based client models outperform traditional MLP-based estimator models significantly in terms of communication rounds. Furthermore, a controlled increase in number of local epochs can change the performance of the global model drastically. While increasing the number of local epochs on ESS clients, the global model was improved by 30 times over the conventional MLP-based estimator model. The proposed framework works well for both IID and non-IID data which is an important aspect of FL on power systems.

Finally, testing the pre-trained CNN to estimate the inertia constant on a system with additional non-synchronous VI units provided an important pathway in this research direction. A pre-trained CNN estimated the inertia constant for a system with additional MPI-based VI unit. However, the frequency response of a system with the estimated value of inertia did not match perfectly with any of the frequency responses for a list of other values of inertia constant. This leads to a conclusion that the inertial and frequency response of a system with non-synchronous VI unit is significantly different than that of the one with all synchronous units. Furthermore, it is important to analyse additional parameters from the non-synchronous unit that contribute in counteracting the change in frequency at the grid. It was observed that out of other possible responses, the VI unit also provides some additional damping response to the system and needs further investigation. Lastly, it was observed that synchronous units with high value of inertia constant require minimal VI support from the non-synchronous units.

REFERENCES

- [1] “Global energy statistical yearbook 2019,” Enerdata, Tech. Rep., Jul. 2019.
- [2] F. Milano, F. Dörfler, G. Hug, D. J. Hill, and G. Verbič, “Foundations and challenges of low-inertia systems,” in *2018 Power Systems Computation Conference (PSCC)*, Jun. 2018, pp. 1–25.
- [3] A. Robb, “Grid inertia: Why it matters in a renewable world,” *Renewable Energy World*, Oct. 2019. [Online]. Available: <https://www.renewableenergyworld.com/2019/10/25/grid-inertia-why-it-matters-in-a-renewable-world/#gref> [Accessed: %204-May-2020].
- [4] Wood Mackenzie Power & Renewables, “U.S. Solar Market Insight Executive summary,” Solar Energy Industries Association, Tech. Rep., Sep. 2019.
- [5] B. Kroposki, B. Johnson, Y. Zhang, V. Gevorgian, P. Denholm, B. Hodge, and B. Hannegan, “Achieving a 100% renewable grid: Operating electric power systems with extremely high levels of variable renewable energy,” *IEEE Power and Energy Magazine*, vol. 15, no. 2, pp. 61–73, Mar. 2017.
- [6] U. Tamrakar, D. Shrestha, M. Maharjan, B. Bhattachari, T. Hansen, and R. Tonkoski, “Virtual inertia: Current trends and future directions,” *Applied Sciences*, vol. 7, no. 7, p. 654, Jun. 2017.
- [7] F. Milano, F. Dörfler, G. Hug, D. J. Hill, and G. Verbič, “Foundations and challenges of low-inertia systems (invited paper),” in *2018 Power Systems Computation Conference (PSCC)*, Jun. 2018, pp. 1–25.
- [8] ERCOT, “Inertia: Basic Concepts and Impacts on the ERCOT Grid,” ERCOT, Tech. Rep., Apr. 2018.
- [9] H. Sturgess, W. Mak, and A. Bourn, “Initial analysis and thoughts on the power outages on 9th August,” Aurora Energy Research, Tech. Rep., Aug. 2019.
- [10] A. Luna, U. Tamrakar, T. M. Hansen, and R. Tonkoski, “Frequency response in grids with high penetration of renewable energy sources,” in *2018 North American Power Symposium (NAPS)*, Sep. 2018, pp. 1–5.
- [11] M. Farrokhabadi et al., “Microgrid stability definitions, analysis, and examples,” *IEEE Transactions on Power Systems*, vol. 35, no. 1, pp. 13–29, Jun. 2019.
- [12] M. Farrokhabadi, C. A. Cañizares, and K. Bhattacharya, “Frequency control in isolated/islanded microgrids through voltage regulation,” *IEEE Transactions on Smart Grid*, vol. 8, no. 3, May 2017.
- [13] A. Ingallali, A. Luna, V. Durvasulu, T. M. Hansen, R. Tonkoski, D. A. Copp, and T. A. Nguyen, “Energy storage systems in emerging electricity markets: Frequency regulation and resiliency,” in *2019 IEEE Power Energy Society General Meeting (PESGM)*, Aug. 2019, pp. 1–5.

- [14] Y. Fu, Y. Wang, and X. Zhang, “Integrated wind turbine controller with virtual inertia and primary frequency responses for grid dynamic frequency support,” *IET Renewable Power Generation*, vol. 11, no. 8, pp. 1129–1137, Jul. 2017.
- [15] U. Tamrakar, T. M. Hansen, R. Tonkoski, and D. A. Copp, “Model predictive frequency control of low inertia microgrids,” in *2019 IEEE 28th International Symposium on Industrial Electronics (ISIE)*, Jun. 2019, pp. 2111–2116.
- [16] M. Ahlstrom, E. Ela, J. Riesz, J. O’Sullivan, B. F. Hobbs, M. O’Malley, M. Milligan, P. Sotkiewicz, and J. Caldwell, “The evolution of the market: Designing a market for high levels of variable generation,” *IEEE Power and Energy Magazine*, vol. 13, no. 6, pp. 60–66, Nov. 2015.
- [17] K. Grimes, “California power outages: Extreme weather, extreme energy shortages, or extreme litigation?” *California Globe*, Oct. 2019. [Online]. Available: <https://californiaglobe.com/section-2/california-power-outages-extreme-weather-extreme-energy-shortages-or-extreme-litigation/> [Accessed: %2011-Nov-2019].
- [18] “IEEE Standard for Interconnection and Interoperability of Distributed Energy Resources with Associated Electric Power Systems Interfaces,” IEEE Std 1547-2018 (Revision of IEEE Std 1547-2003), Tech. Rep., Apr. 2018.
- [19] R. Manghani, *The state of u.s. storage market: Opportunities for customers and utilities*. [Online]. Available: https://aeenewengland.org/images/downloads/Past_Meeting_Presentations/ravi_manghani__gtm_research__utility_of_the_future.pdf [Accessed: %204-Nov-2019].
- [20] M. Brown, *Miso moves forward to further integrate energy storage resources*. [Online]. Available: <https://www.misoenergy.org/about/media-center/miso-moves-forward-to-further-integrate-energy-storage-resources/> [Accessed: 5-Nov-2019].
- [21] A. Engler and N. Soultanis, “Droop control in lv-grids,” in *2005 International Conference on Future Power Systems*, Nov. 2005, p. 6.
- [22] H. Sun, Q. Guo, J. Qi, V. Ajjarapu, R. Bravo, J. Chow, Z. Li, R. Moghe, E. Nasr-Azadani, U. Tamrakar, G. N. Taranto, R. Tonkoski, G. Valverde, Q. Wu, and G. Yang, “Review of challenges and research opportunities for voltage control in smart grids,” *IEEE Transactions on Power Systems*, vol. 34, no. 4, pp. 2790–2801, Jul. 2019.
- [23] P. Kundur, N. J. Balu, and M. G. Lauby, *Power system stability and control*. McGraw-hill New York, 1994, vol. 7.
- [24] K. Duwadi, R. Fourney, R. Tonkoski, and T. M. Hansen, “Sustainability metrics for inverter-based voltage regulation methods in pv-rich low voltage grids,” in *2019 IEEE International Conference on Electro Information Technology (EIT)*, May 2019, pp. 441–446.

- [25] R. Tonkoski and L. A. C. Lopes, “Voltage regulation in radial distribution feeders with high penetration of photovoltaic,” in *2008 IEEE Energy 2030 Conference*, Nov. 2008, pp. 1–7.
- [26] X. Zhang, M. Farina, S. Spinelli, and R. Scattolini, “Multi-rate model predictive control algorithm for systems with fast-slow dynamics,” *IET Control Theory & Applications*, vol. 12, no. 18, pp. 2468–2477, Dec. 2018.
- [27] B. Pournazarian, P. Karimyan, G. Gharehpetian, M. Abedi, and E. Pouresmaeil, “Smart participation of PHEVs in controlling voltage and frequency of island microgrids,” *International Journal of Electrical Power & Energy Systems*, vol. 110, pp. 510–522, Sep. 2019.
- [28] J. Engels, B. Claessens, and G. Deconinck, “Grid-constrained distributed optimization for frequency control with low-voltage flexibility,” *IEEE Transactions on Smart Grid*, Jul. 2019[Early Access].
- [29] Z. Tang, T. Liu, C. Zhang, Y. Zheng, and D. J. Hill, “Distributed control of active distribution networks for frequency support,” in *2018 Power Systems Computation Conference (PSCC)*, Jun. 2018, pp. 1–7.
- [30] J. He, Y. W. Li, J. M. Guerrero, F. Blaabjerg, and J. C. Vasquez, “An islanding microgrid power sharing approach using enhanced virtual impedance control scheme,” *IEEE Transactions on Power Electronics*, vol. 28, no. 11, pp. 5272–5282, Nov. 2013.
- [31] J. Matas, M. Castilla, L. G. d. Vicuña, J. Miret, and J. C. Vasquez, “Virtual impedance loop for droop-controlled single-phase parallel inverters using a second-order general-integrator scheme,” *IEEE Transactions on Power Electronics*, vol. 25, no. 12, Dec. 2010.
- [32] P. Zhang, H. Zhao, H. Cai, J. Shi, and X. He, “Power decoupling strategy based on ‘virtual negative resistor’ for inverters in low-voltage microgrids,” *IET Power Electronics*, vol. 9, no. 5, pp. 1037–1044, Apr. 2016.
- [33] H. Laaksonen, P. Saari, and R. Komulainen, “Voltage and frequency control of inverter based weak lv network microgrid,” in *2005 International Conference on Future Power Systems*, Nov. 2005, p. 6.
- [34] K. D. Brabandere, A. Woyte, R. Belmans, and J. Nijs, “Prevention of inverter voltage tripping in high density pv grids,” in *19th Photovoltaic Solar Energy Conference*, Jun. 2004.
- [35] H. Han, Y. Liu, Y. Sun, M. Su, and J. M. Guerrero, “An improved droop control strategy for reactive power sharing in islanded microgrid,” *IEEE Transactions on Power Electronics*, vol. 30, no. 6, pp. 3133–3141, Jun. 2015.

- [36] R. Eriksson, N. Modig, and K. Elkington, “Synthetic inertia versus fast frequency response: A definition,” *IET Renewable Power Generation*, vol. 12, no. 5, pp. 507–514, Mar. 2018.
- [37] NERC, “Reliability guideline: Primary frequency control,” NERC, Tech. Rep., Dec. 2015.
- [38] A. Hoke, *Fast grid frequency support from distributed inverter-based resources*. [Online]. Available: <https://www.nrel.gov/docs/fy18osti/71461.pdf> [Accessed: %204-Nov-2019].
- [39] Z. A. Obaid, L. M. Cipcigan, L. Abraham, and M. T. Muhssin, “Frequency control of future power systems: Reviewing and evaluating challenges and new control methods,” *Journal of Modern Power Systems and Clean Energy*, vol. 7, no. 1, pp. 9–25, Aug. 2018.
- [40] J. W. Simpson-Porco, F. Dörfler, and F. Bullo, “Voltage stabilization in microgrids via quadratic droop control,” *IEEE Transactions on Automatic Control*, vol. 62, no. 3, pp. 1239–1253, Mar. 2017.
- [41] NERC, “Distributed energy resources: Connection modeling and reliability considerations,” NERC, Tech. Rep., Feb. 2017.
- [42] I. Green, *Der_a modeling: Der_a as a part of composite load*. [Online]. Available: <https://www.wecc.org/Administrative/DER%5C%20Models%5C%20CAISO-%5C%20Green.pdf> [Accessed: %204-Nov-2019].
- [43] J. Giraldez, A. Hoke, P. Gotseff, N. Wunder, M. Blonsky, M. Emmanuel, A. Latif, E. Ifuku, M. Asano, and T. Aukai, “Advanced inverter voltage controls: Simulation and field pilot findings,” National renewable Energy Lab, Tech. Rep. NREL/TP-5D00-72298, Oct. 2018, p. 67.
- [44] S. J. Plathottam, S. Abhyankar, and P. Hazra, “Dynamic modeling of solar pv systems for distribution system stability analysis,” in *2019 IEEE Power Energy Society Innovative Smart Grid Technologies Conference (ISGT)*, Feb. 2019, pp. 1–5.
- [45] Z. Cheng, J. Duan, and M. Chow, “To centralize or to distribute: That is the question: A comparison of advanced microgrid management systems,” *IEEE Industrial Electronics Magazine*, vol. 12, no. 1, pp. 6–24, Mar. 2018.
- [46] Z. Ma, A. Pesaran, V. Gevorgian, D. Gwinner, and W. Kramer, “Energy storage, renewable power generation, and the grid: Nrel capabilities help to develop and test energy-storage technologies,” *IEEE Electrification Magazine*, vol. 3, no. 3, pp. 30–40, 2015.
- [47] T. Inoue, H. Taniguchi, Y. Ikeguchi, and K. Yoshida, “Estimation of power system inertia constant and capacity of spinning-reserve support generators using measured frequency transients,” *IEEE Transactions on Power Systems*, vol. 12, no. 1, pp. 136–143, Feb. 1997.

- [48] M. Shamirzaee, H. Ayoubzadeh, D. Farokhzad, F. Aminifar, and H. Haeri, “An improved method for estimation of inertia constant of power system based on polynomial approximation,” in *2014 Smart Grid Conference (SGC)*, Dec. 2014, pp. 1–7.
- [49] P. Wall, F. Gonzalez-Longatt, and V. Terzija, “Estimation of generator inertia available during a disturbance,” in *2012 IEEE Power and Energy Society General Meeting*, Jul. 2012, pp. 1–8.
- [50] M. A. M. Ariff, B. C. Pal, and A. K. Singh, “Estimating dynamic model parameters for adaptive protection and control in power system,” *IEEE Transactions on Power Systems*, vol. 30, no. 2, pp. 829–839, Jul. 2015.
- [51] P. M. Ashton, C. S. Saunders, G. A. Taylor, A. M. Carter, and M. E. Bradley, “Inertia estimation of the GB power system using synchrophasor measurements,” *IEEE Transactions on Power Systems*, vol. 30, no. 2, pp. 701–709, Jul. 2015.
- [52] 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, no. 6, Jun. 2018.
- [53] L. Lugnani, D. Dotta, C. Lackner, and J. Chow, “Armax-based method for inertial constant estimation of generation units using synchrophasors,” *Electric Power Systems Research*, vol. 180, p. 106 097, Mar. 2020.
- [54] F. Allella, E. Chiodo, G. M. Giannuzzi, D. Lauria, and F. Mottola, “On-line estimation assessment of power systems inertia with high penetration of renewable generation,” *IEEE Access*, vol. 8, pp. 62 689–62 697, Mar. 2020.
- [55] X. Cao, B. Stephen, I. F. Abdulhadi, C. D. Booth, and G. M. Burt, “Switching Markov Gaussian models for dynamic power system inertia estimation,” *IEEE Transactions on Power Systems*, vol. 31, no. 5, pp. 3394–3403, Dec. 2016.
- [56] 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.
- [57] J. Schiffer, P. Aristidou, and R. Ortega, “Online estimation of power system inertia using dynamic regressor extension and mixing,” *IEEE Transactions on Power Systems*, vol. 34, no. 6, May 2019.
- [58] M. Farrokhabadi, C. A. Cañizares, J. W. Simpson-Porco, E. Nasr, L. Fan, P. A. Mendoza-Araya, R. Tonkoski, U. Tamrakar, N. Hatzigyriou, D. Lagos, R. W. Wies, M. Paolone, M. Liserre, L. Meegahapola, M. Kabalan, A. H. Hajimiragha, D. Peralta, M. A. Elizondo, K. P. Schneider, F. K. Tuffner, and J. Reilly, “Microgrid stability definitions, analysis, and examples,” *IEEE Transactions on Power Systems*, vol. 35, no. 1, pp. 13–29, Jun. 2020.

- [59] L. Fan, *Control and dynamics in power systems and microgrids*. CRC Press, 2017, vol. 1.
- [60] U. Markovic *et al.*, “LQR-based adaptive virtual synchronous machine for power systems with high inverter penetration,” *IEEE Transactions on Sustainable Energy*, vol. 10, no. 3, pp. 1501–1512, Dec. 2019.
- [61] 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.
- [62] T. Mitchell, *Machine Learning: A Guide to Current Research*. McGraw-hill New York, 1997.
- [63] M. Z. Alom *et al.*, “A state-of-the-art survey on deep learning theory and architectures,” *Electronics*, vol. 8, no. 3, Mar. 2019.
- [64] O. I. Abiodun, A. Jantan, A. E. Omolara, K. V. Dada, N. A. Mohamed, and H. Arshad, “State-of-the-art in artificial neural network applications: A survey,” *Heliyon*, vol. 4, no. 11, e00938, Nov. 2018.
- [65] E. Hoffer, I. Hubara, and D. Soudry, “Train longer, generalize better: Closing the generalization gap in large batch training of neural networks,” in *Advances in Neural Information Processing Systems*, May 2017, pp. 1731–1741.
- [66] A. Paszke, S. Gross, S. Chintala, G. Chanan, E. Yang, Z. DeVito, Z. Lin, A. Desmaison, L. Antiga, and A. Lerer, “Automatic differentiation in PyTorch,” in *Proceedings of Neural Information Processing Systems*, Dec. 2017.
- [67] N. Suda and D. Loh, “Machine learning on Arm Cortex–M microcontrollers,” ARM, Tech. Rep., 2019.
- [68] A. Ulbig, T. S. Borsche, and G. Andersson, “Impact of low rotational inertia on power system stability and operation,” *IFAC Proceedings Volumes*, vol. 47, no. 3, pp. 7290–7297, Aug. 2014.
- [69] 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.
- [70] “IEEE standard for synchrophasor measurements for power systems,” *IEEE Std C37.118.1-2011 (Revision of IEEE Std C37.118-2005)*, pp. 1–61, Dec. 2011.
- [71] H. Huang and F. Li, “Sensitivity analysis of load-damping characteristic in power system frequency regulation,” *IEEE Transactions on Power Systems*, vol. 28, no. 2, pp. 1324–1335, May 2013.
- [72] J. Taft and A. Becker-Dippmann, “Grid architecture,” Pacific Northwest National Laboratory, Tech. Rep. PNNL-24044, Jan. 2015, p. 115.

- [73] H. Mehrjerdi, S. Lefebvre, M. Saad, and D. Asber, “A decentralized control of partitioned power networks for voltage regulation and prevention against disturbance propagation,” *IEEE Transactions on Power Systems*, vol. 28, no. 2, pp. 1461–1469, May 2013.
- [74] G. Cai, B. Wang, D. Yang, Z. Sun, and L. Wang, “Inertia estimation based on observed electromechanical oscillation response for power systems,” *IEEE Transactions on Power Systems*, vol. 34, no. 6, pp. 4291–4299, Nov. 2019.
- [75] H. B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. Arcas, “Communication-efficient learning of deep networks from decentralized data,” in *20th International Conference on Artificial Intelligence and Statistics (AISTATS)*, Mar. 2017.
- [76] G. Andersson, *Dynamics and Control of Electric Power Systems*. ETH Zurich, Feb. 2012.
- [77] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, “Gradient-based learning applied to document recognition,” *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, Nov. 1998.
- [78] B. McMahan, D. Ramage, and et al., *Federated learning: Collaborative machine learning without centralized training data*, Apr. 2017. [Online]. Available: <https://ai.googleblog.com/2017/04/federated-learning-collaborative.html>.
- [79] I. J. Goodfellow, O. Vinyals, and A. M. Saxe, “Qualitatively characterizing neural network optimization problems,” in *2015 International Conference on Learning Representations (ICLR)*, Jan. 2015.
- [80] Li Li, Xiaoguang Hu, and Weicun Zhang, “Design of an arm-based power meter having wifi wireless communication module,” in *2009 4th IEEE Conference on Industrial Electronics and Applications*, Jun. 2009, pp. 403–407.
- [81] B. K. Poolla, S. Bolognani, and F. Dörfler, “Optimal placement of virtual inertia in power grids,” *IEEE Transactions on Automatic Control*, vol. 62, no. 12, pp. 6209–6220, Dec. 2017.
- [82] J. Fang, H. Li, Y. Tang, and F. Blaabjerg, “Distributed power system virtual inertia implemented by grid-connected power converters,” *IEEE Transactions on Power Electronics*, vol. 33, no. 10, pp. 8488–8499, Oct. 2018.
- [83] W. Wu, Y. Chen, A. Luo, L. Zhou, X. Zhou, L. Yang, Y. Dong, and J. M. Guerrero, “A virtual inertia control strategy for dc microgrids analogized with virtual synchronous machines,” *IEEE Transactions on Industrial Electronics*, vol. 64, no. 7, pp. 6005–6016, Jul. 2017.
- [84] T. Kerdphol, F. S. Rahman, Y. Mitani, M. Watanabe, and S. K. Küfeoğlu, “Robust virtual inertia control of an islanded microgrid considering high penetration of renewable energy,” *IEEE Access*, vol. 6, pp. 625–636, Nov. 2018.

- [85] B. K. Poolla, S. Bolognani, N. Li, and F. Dörfler, “A market mechanism for virtual inertia,” *IEEE Transactions on Smart Grid*, vol. 11, no. 4, pp. 3570–3579, Jul. 2020.
- [86] T. A. Nguyen, R. H. Byrne, R. J. Concepcion, and I. Gyuk, “Maximizing revenue from electrical energy storage in miso energy frequency regulation markets,” in *2017 IEEE Power Energy Society General Meeting*, Jul. 2017, pp. 1–5.