DETECTION, ESTIMATION AND FORECAST FOR NONLINEAR SYSTEMS

By

Asim Zaheer Ud Din

Reg. No. NUST201590326PSMME2715F

Thesis Supervisor

Dr Yasar Ayaz

Department of Robotics and Intelligent Machine Engineering

School of Mechanical and Manufacturing Engineering

National University of Sciences &Technology

Islamabad, Pakistan

2022

DETECTION, ESTIMATION AND FORECAST FOR NONLINEAR SYSTEMS

By

Asim Zaheer Ud Din

Reg. No. NUST201590326PSMME2715F

A thesis submitted to the National University of Sciences & Technology
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Robotics and Intelligent Machine Engineering

Thesis Supervisor

Dr Yasar Ayaz

Department of Robotics and Intelligent Machine Engineering

School of Mechanical and Manufacturing Engineering

National University of Sciences &Technology

Islamabad, Pakistan

2022

THESIS ACCEPTANCE CERTIFICATE

It is certified that final copy of PhD thesis written by Asim Zaheer Ud Din (Registration No. NUST201590326PSMME2715F), of Department of Robotics and Intelligent Machine Engineering (SMME) has been vetted by undersigned, found complete in all respects as per NUST Statutes / Regulations / PhD Policy, is free of plagiarism, errors, and mistakes and is accepted as partial fulfillment for award of PhD degree. It is further certified that necessary amendments as pointed out by GEC members and foreign / local evaluators of the scholar have also been incorporated in the said thesis.

Signature: _____

Name of Supervisor: Dr. Yasar Ayaz

Date: _____

Signature (HOD): _____

Date: _____

Signature (Dean/Principal): _____

Date:_____

AnnexL
FormPhD-7
DOCTORAL PROGRAMME OF STUDY
(Must be type written)

National University of Sciences & Technology
REPORT OF DOCTORAL THESIS DEFENCE

Name: **Asim Zaheer Ud Din** NUST Regn No: **NUST201590326PSMME2715F**

School/College/Centre: **School of Mechanical & Manufacturing Engineering**

Title: **Detection, estimation, and forecast for nonlinear systems**

DOCTORAL DEFENCE COMMITTEE

Doctoral Defence held on **11.11.2021**

		QUALIFIED	NOT QUALIFIED	SIGNATURE
GEC Member-1:	**Dr Muhammad Jawad Khan**	✓		
GEC Member-2:	**Dr Hasan Sajid**	✓		
GEC Member (External):	**Dr Noman Naseer**	✓		
Supervisor:	**Dr Yasar Ayaz**	✓		
Co Supervisor (if appointed):	**NA**			
External Evaluator-1: (Local Expert)	**Dr Faraz Akram**	✓		
External Evaluator-2: (Local Expert)	**Dr Adnan Daud Khan**	✓		
External Evaluator-3: (Foreign Expert)	**Dr Hendrik Santosa**	✓		
External Evaluator-4: (Foreign Expert)	**Dr Ahsan Saadat**	✓		

FINAL RESULT OF THE DOCTORAL DEFENCE
(Appropriate box to be signed by HOD)

[✓] PASS [ ] FAIL

The student **Asim Zaheer Ud Din** Regn No **NUST201590326PSMME2715F** is / is NOT accepted for Doctor of Philosophy Degree.

Dated: 23-11-2021

Dr Javaid Iqbal
Principal & Dean
School of Mechanical & Manufacturing Engineering
(SMME) NUST, Islamabad

Dean/Commandant/Principal/DG

Distribution:
01 x original copy each for PGP Dte, Exam Branch Main Office NUST and Student's dossier at the School/College/Centre.
01 x photocopy each for HoD, Supervisor, Co-Supervisor (if appointed), sponsoring agency (if any) and 05 copies for insertion in Dissertation.
Note:* Decision of External Evaluators (Foreign Experts) will be sought through video conference, if possible, on the same date and their decision will be intimated (on paper) to HQ NUST at a later date.

CERTIFICATE OF APPROVAL

This is to certify that the research work presented in this thesis titled 'Detection, estimation and forecast for nonlinear systems' was conducted by Mr. Asim Zaheer ud Din under the supervision of Dr. Yasar Ayaz. No part of this thesis has been submitted anywhere else for any degree. This thesis is submitted to the School of Mechanical and Manufacturing Engineering in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the field of Robotics and Intelligent Machine Engineering.

Student: Asim Zaheer Ud Din Signature: _____

i. External examiner 1: Dr. Faraz Akram, Assistant Professor, Faculty of Engineering and Applied Sciences, Riphah University, Islamabad, Pakistan Signature: _____

ii. External examiner 2: Dr. Adnan Daud, Associate Professor, Faculty of Renewable Energy Engineering, UET Peshawar, Pakistan Signature: _____

iii. GEC 1: Dr. Muhammad Jawad Khan, Assistant Professor, SMME, NUST Pakistan Signature: _____

iv. GEC 2: Dr. Hasan Sajid, Assistant Professor, SMME, NUST Pakistan Signature: _____

Supervisor: Dr Yasar Ayaz Signature: _____

Principal: Dr Javaid Iqbal Signature: _____

Author's Declaration

I, Asim Zaheer Ud Din here by state that my PhD thesis titled "Dectection, estimation and forecast for nonlinear systems" is my own work. The work has not been submitted previously for any degree from National University of Sciences and Technology or anywhere else. The material used from other sources has been properly acknowledged / referred.

Student: Asim Zaheer Ud Din

Date:_____

Plagiarism Undertaking

I solemnly declare that research work presented in the thesis titled “Detection, estimation and forecast for nonlinear systems” is solely my research work with no significant contribution from any other person. Small contribution / help wherever taken has been duly acknowledged and that complete thesis has been written by me.

I understand the zero tolerance policy of the HEC and National University of Science & Technology towards plagiarism. Therefore, I as an author of the above titled thesis declare that no portion of my thesis has been plagiarized and any material used as reference is properly referred / cited.

I undertake that if I am found guilty of any formal plagiarism in the above titled thesis even after award of PhD degree, the university reserves the right to withdraw / revoke my PhD degree, and that HEC and the university has the right to publish my name on the HEC/university website on which names of students are placed who submitted plagiarized thesis.

Student Name & Signature:

Asim Zaheer Ud Din

Signature: _____

. . . correct intent is great strength

ACKNOWLEDGEMENT

It is the guidance and support from my elders, teachers, family and friends that made this work achievable.

Asim Zaheer Ud Din
December 1st, 2021
mr.a.z@ieee.org

Preface

This thesis proposes and implements solutions in context to smart infrastructure system for smart grid. The thesis is organized as:

Chapter 1 discusses benefits and technologies of smart grid. Also in this chapter, the smart grid is presented as a combination of three subsystems.

Chapter 2 proposes a telemetry solution based on radio frequency (RF) and general packet radio service (GPRS) technologies to monitor total power consumption and power flow in electricity distribution company network. Chapter also presents deisgn of a over-load detector based on Generalized Likelihood Ratio Test for output feeder of grid station. Results of implementation of telemetry system and over-load detector in form of reduction in load-shedding for electricity distribution company of Pakistan are also included.

In Chapter 3 multivariate multistep short-term power forecasting framework for electricity distribution company has been presented. Discussion on how 24 hours ahead power forecasting is helpful in achieving economic dispatch from power generation facility is made. Impact assessment on application of proposed forecasting framework for electricity distribution company of Pakistan is also included.

Chapter 4 demonstrates bivariate 24 hours ahead forecasting of power imported and power exported by electricity distribution company using Long-short-term-memory network (LSTM). Forecasting performance of LSTM network for changes in input features, network settings, and training algorithms is also explored.

Chapter 5 introduces estimation and output feedback control scheme using higher order Sliding Mode observer (SMO) for minimum phase nonlinear systems. This presented robust feedback linearization scheme is demonstrated in computer simulations for magnetic Levitation (Maglev) and DC motor systems.

In Chapter 6 Cubature Kalman filter (CKF) estimator for output feedback control is introduced. Tracking performance of estimator based output feedback control scheme is compared with tracking performance of state feedback control. Tracking performances of Cubature Kalman filter, Unscented Kalman filter, Extended Kalman filter, State-Space

Recursive Least Squares and SMO based output feedback control schemes in presence of external disturbance and parameters perturbation is also compared.

In Chapter 7 a neuro-estimator based output feedback control scheme is discussed. An Emulation Design based discrete robust feedback linearization controller is achieved on employing neural network-aided dual Unscented Kalman filter estimation algorithm. Simulation results demonstrate robustness of presented scheme.

Chapter 8 discusses output feedback control scheme for minimum phase nonlinear systems with unknown plant (system) parameters. Dual Unscented Kalman filter estimation and Emulation Design based discrete feedback linearization controller are employed in this sampled-data control configuration.

Chapter 9 concludes this thesis, and suggests few recommendations for future work.

ABSTRACT

DETECTION, ESTIMATION AND FORECAST FOR NONLINEAR SYSTEMS

By

ASIM ZAHEER UD DIN

This thesis presents and implements detection, estimation, and forecast algorithms in context of smart infrastructure system for smart grid. A novel application of radio frequency wireless mesh network and general packet radio service technologies in a telemetry solution has been proposed. The telemetry solution measures power flow in the energy network of an electricity distribution company. The solution utilizes some selected circuits of grid stations, and calculates total power consumed, total power imported and total power exported by the distribution company. The selection of circuits for sensors installation is the key for reducing solution cost as compare to the case when sensors are installed on all the power output points. The framework involves installation of specially developed energy sensors (smart energy meters) and data concentrator units at the selected grid stations. The approach has been tested on two electricity distribution companies of Pakistan: Islamabad Electric Supply Company and Peshawar Electric Supply Company. Also in this work, result of over-load detection based on generalized likelihood ratio test for an industrial feeder of Islamabad Electric Supply Company is included. Detection probability of 0.96 with a false alarm probability of 0.04 has been achieved for 30 minutes data interval.

Further, 4 years power data obtained from above mentioned system is utilized in a multivariate multi-step-ahead short-term forecasting formulation. The formulation operates on multiple inputs from multiple variables, and provides multi-step-ahead forecasts by generating multiple outputs for multiple variables. The presented framework is effective for large forecasting horizons since it forecasts for temporally dependent sub-

intervals called runs from large horizon. Thus the framework forecasts are less biased and suffer low variance, as compared with direct method and iterated method estimators respectively. Feedward neural network and Long-short-term-memory network models have been evaluated in presented framework. The proposed framework has demonstrated forecasts of power import and power export with a horizon value of 48 for Peshawar Electric Supply Company (PESCO), Pakistan. The averaged mean absolute percentage error of two forecasted time series is 12.76 %, whereas, 24 hours ahead power consumption of PESCO total consumers has been forecasted with mean absolute percentage error of 8.6%. Furthermore, exploiting 24 hours ahead power consumption forecasts has resulted in better power dispatch for PESCO grid stations by reducing mean absolute error by 11.52 times between PESCO power allocated and PESCO power consumed.

Next the thesis presents an Euler approximate discrete-time Sliding Mode observer (SMO) which simultaneously estimates states and combined effect of unmodeled system dynamics and disturbances. Emulation Design procedure is employed in designing of discrete feedback linearization controller. Computer simulations demonstrate performance of presented output feedback scheme for tracking applications of magnetic levitation and DC motor systems. Results illustrate that reducing sampling period more adversely affects Euler approximate discrete observer performance for faster changing system dynamics than for slower changing dynamics. The proposed scheme also exhibits good performance in presence of disturbances and parameters perturbation.

Furthermore, it is demonstrated via simulations that robust tracking control is achived on using estimator (e.g Kalman filter, SMO, SSRLS filter) in sampled-data output feedback configuration, as compared to performing tracking using sampled-data state feedback scheme. Simulation results show that SMO based output feedback tracking is most robust, followed by CKF and EKF based output feedback scheme. UKF based output feedback scheme is robust against external disturbance; but for case of system parameter perturbation, UKF tracking error takes longer time to converge. State-Space Recursive Least Squares (SSRLS) based scheme behaves poorly in presence of external disturbance. This is because SSRLS estimation is based on constant velocity model and not on actual nonlinear system model.

Finally, output feedback control scheme for case of unknown system parameters has been presented. The scheme employs dual UKF estimation algorithm and Emulation Design based discrete feedback linearization controller. Implementation results exhibit that presented output feedback control scheme demonstrates better tracking performance and parameter estimation error when parameter estimate is initialized with a value (in dual estimation algorithm) which is closer to actual system parameter value.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ALGORITHMS

LIST OF PUBLICATIONS

- **Publication 1:** Asim Zaheer Ud Din, Yasar Ayaz, Mian Ilyas Ahmad, Momena Hasan, Salman Masaud and Naveed Muhammad, "Cost-Effective Telemetry for Energy Network of an Electricity Distribution Company: Part I", Turkish Journal of Electrical Engineering & Computer Sciences, vol. 27, no. 2, pp. 889–902, 2019. **(Presented in Chapter 2)**
- **Publication 2:** Asim Zaheer Ud Din and Yasar Ayaz, "Multivariate Multistep Short-Term Power Forecasting for Electricity Distribution Company", Journal of King Saud University – Computer and Information Sciences (in review), 2021. **(Presented in Chapter 3)**
- **Publication 3:** Asim Zaheer Ud Din, Yasar Ayaz, Momena Hasan, Jawad Khan, and Muhammad Salman, "Bivariate short-term electric power forecasting using LSTM network", 3rd International Conference on Robotics and Automation in Industry, 2019, Rawalpindi, Pakistan. **(Presented in Chapter 4)**
- **Publication 4:** Asim Zaheer Ud Din, Yasar Ayaz, Muhammad Salman, Momena Hasan and Niaz Mehdi, "Sampled-data robust feedback linearization using higher order sliding mode observer", 15th International Bhurban Conference on Applied Sciences & Technology (IBCAST), 2018, Islamabad, Pakistan. **(Presented in Chapter 5)**
- **Publication 5:** Asim Zaheer, Yasar Ayaz, Momena Hasan and Muhammad Salman, "Sampled-data Robust Feedback Linearization Using Estimator", IEEE 14th International Workshop on Advanced Motion Control, 2016, Auckland, New Zealand. **(Presented in Chapter 6)**
- **Publication 6:** Asim Zaheer Ud Din and Yasar Ayaz, "Sampled-data Output Feedback Control Using Dual Estimation", International Journal of Industrial and Systems Engineering (in review), 2021. **(Presented in Chapter 7 & Chapter 8)**

CHAPTER 1

THESIS INTRODUCTION - SMART GRID

Electricity system that features all or any of the following four operations: electricity generation, electricity control, electricity distribution, and electricity transmission, is called a grid. On April 24th, 2007 at International Data Corporation, Energy Conference in Chicago, Andres E. Carvallo brought forward the term 'Smart Grid', as a blend of energy, hardware, software, and communications. He defined smart grid as a combination of an electric grid with hardware, software, and communication networks to monitor, control and manage the generation, distribution, consumption and storage of energy.

Alternatively stated, 'Smart Grid' may be considered as a modernization effort of twentieth century power network *"so it monitors, protects, and automatically optimizes the operation of its interconnected elements — from the central and distributed generator through the high-voltage network and distribution system, to industrial users and building automation systems, to energy storage installations and to end-use consumers and their thermostats, electric vehicles, appliances, and other household devices"* [1].

In a smart grid, an automated, broadly distributed power network is achieved via two-way flow of information and electricity. There is a real-time energy supply-demand balance using the benefits of distributed communication and computing.

Summarizing; in addition to the enhanced customer services and grid operations, the smart grid also offers environmental benefits. And all this realized by integration of information technology and communication in all phases of network, starting from point of power generation to the point of power consumption.

1.1 Benefits of smart grid

Several definitions of smart grid have been proposed by various government organizations and authors. Initially smart grid aim was to improve energy efficiency, power demand management, and construction of a reliable grid protection capable of self healing.

However with time, growing requirements made, governments, electricity industries, and research organizations to revise initially defined scope of smart grid.

The requirements and expected benefits of smart grid according to a report from National Institute of Standards and Technology [2] are listed below:

1. Efficiency and capacity enhancement of existing power networks
2. Improvement in power quality and reliability
3. Economic dispatch
4. Automated operation and maintenance
5. Better grid security
6. Reduction in inefficient power generation during peak hours using oil
7. Inclusion of distributed power system (sources)
8. Inclusion of renewable sources
9. Self healing and predictive maintenance
10. Improvement in resilience against power interruptions
11. Provision of new energy storage mechanism and facilitation for charging of electric vehicles
12. Reduction in greenhouse gas emissions using new power sources and electric vehicles
13. Enhanced options for consumer
14. New enhanced markets, services, and products

1.2 Literature review & motivation - Technologies in smart grid

To comprehend above mentioned features of smart grid paradigm, and to standardize smart grid work being carried out for different parts of electrical power system, National Institute of Standards and Technology also proposed a reference conceptual model [2]. The smart grid has been divided into seven domains namely, customers, markets, service providers, operations, electricity generation, electricity distribution, and electricity transmission in this conceptual model. To materialize a smart grid, technologies have to be developed and applied in these seven domains.

Brief reviews on technologies for smart grid have been presented in [3 – 5]. In [6, 7] current standards for smart grid have been reviewed; and recommendations have been

proposed for future smart grid standards. Jiang in [8] has presented industry point of view on smart power distribution system; and has also identified future technologies for distribution network of smart grid.

An optimum and complete smart grid is a vision. It may be considered as a designed integration of services, functions, sub-systems, and complementatary systems, under an intelligent autonomous monitoring, and control. Recognizing the broad scope of smart grid, researchers and industrialists have presented different opinions based on their priorities and plans. In [9] authors have presented smart grid system, as a system of systems, and have divided it in three majors areas (systems) on technical grounds, namely: smart infrastructure system, smart management system, and smart protection system. These three systems are briefly discussed below. Detailed references to thesis work are presented in later chapters of this thesis.

1.2.1 Smart infrastructure system

The smart infrastructure system encompasses information, energy, and communication frameworks of smart grid. For both the information and energy, the flow of data is bidirectional. i.e. not only the information data and energy flow forward from source of generation (network) to end-user; but also flow backward from end-user to the network. Furthermore, different systems, devices, and applications operating within smart grid regime are interconnected using communication framework.

1.2.2 Smart management system

In smart grid the objectives of advanced management, and control functionalities are achieved via smart management system. Provision for distributed and centralized control is made available. Objectives of advanced management may include, supply-demand balance, utility optimization, energy efficiency, and emissions reduction.

1.2.3 Smart protection system

The smart protection system includes all software and hardware of smart grid that offers energy network predictive maintenance, fault detection, and self-healing features. The smart protection framework incorporates network reliability analysis; and also addresses security and privacy concerns.

1.3 Thesis contribution on smart infrastructure system

The work presented in this thesis contributes to smart infrastructure sub-system of smart grid. As discussed above the smart infrastructure sytem includes all the hardware and software working in their particular frameworks, relating to information, energy, and communication. The working frameworks and advanced features of intelligent infrastructure system in smart grid proposed and implemented in this thesis are briefly discussed below.

1.3.1 Advanced energy metering, monitoring, and management

This work proposes and implements a telemetry solution for electricity distribution company (EDC) operating in smart grid regime. The telemetry solution architecture utilizes both RF and GPRS technologies. It utilizes sensors installed on power import circuits and power export circuits of an EDC energy network grid stations, to calculate the real-time total power consumed by EDC region. The result is a reduced number of required sensors (and also solution cost); as compared to the case, when sensors are installed on every output feeder of EDC network. The framework comprises installation of specially designed energy sensor (smart energy meter, class 0.5) and data concentrator unit (DCU) at the selected grid stations, for measurement of energy data that includes active energy, reactive energy, active power, apparent power, current, voltage and power factor. A developed computer server application is also installed in power control center of distribution company. The developed solution provides both, telemetry of distribution network, and real-time total power consumption value of complete region serviced by EDC.

Also in this work, to achieve advanced load (energy) management features, a over-load detection algorithm based on generalized likelihood ratio test for an industrial feeder of a grid station has been proposed and implemented. Over-load detector results, and impact of thesis work on areas of advanced energy metering, monitoring, and management for Islamabad Electric Supply Company of Pakistan are also included.

The above stated work is **presented in Chapter 2** of this thesis.

1.3.2 Advanced technologies in communication

The work presented in this thesis proposes and implements a novel application of RF wireless mesh network and GPRS technologies in a telemetry solution to measure

power flow in the energy network of an electricity distribution company in smart grid. This hybrid combination of RF and GPRS technologies have been used in remote energy metering of industrial, commercial, and residential consumers throughout the world. The proposed and implemented telemetry application for EDC energy network has not been considered before to the best of our knowledge.

The telemetry solution utilizes energy data from some selected circuits of grid stations, and calculates total power consumed, total power imported and total power exported by the distribution company as also discussed above. Within a grid station the measured data reaches DCU using 433 MHz wireless mesh network, and is transmitted to a remote power control center using GPRS. Energy data from different grid stations across the energy network of EDC is collected at power control center, and utilized in real-time calculations of total power consumed, total power import and total power export. The approach has been tested on two electricity distribution companies of Pakistan: Islamabad Electric Supply Company and Peshawar Electric Supply Company.

This solution is also **dicussed in Chapter 2** of thesis.

1.3.3 Advanced electricity utilization

Optimal operations and utilization of electricity generation facilities is one of the major objectives in operations of smart grid. Optimal operation of electricity generation facility requires that power generated balances consumers demand in real-time. This demand-supply balance is also referred to as economic dispatch. To assist economic dispatch an effective approach [10] is to model consumer demand, and then employ this model to forecast consumer demand on short-term, medium-term, and long-term bases. These forecasted power demand values are helpful in optimal scheduling of power plants. The work presented in this thesis discusses demand model learning for EDC consumer; and demonstrates 24 hours ahead short-term power forecasting results using learned model. Impact of short-term power forecasting for Peshawar Electric Supply Company of Pakistan is also discussed quantitatively.

The work on advanced electricity utilization is **presented in Chapter 3 and Chapter 4** of this thesis.

1.3.4 Advanced electricity generation

In smart grid and conventional electrical grid, power stations, electrical substations, transmission lines, and distribution lines are interconnected to form a network for delivering electricity from power producers to power consumers. Power stations or generating stations are facilities generating three-phase electricity. The electricity generation at power plant is the process of generating electric power from primary energy sources and motiv power. Most often in power plant this electricity is generated by electromechanical generators that are driven by water turbines, steam turbines, wind turbines, internal combustion engines, etc.

Electromechanical generator and motor are machines used most regularly in electrical power generation industry to convert mechanical energy into electrical energy, and vice versa. These electromagnetic system consist of rotor and stator. Usually bearings are mounted on each end of rotor to support rotor alignment, and for attaching load to rotor shaft. In context of smart grid, one effort towards an advanced and efficient electricity generation can be to use active magnetic bearings instead of conventional bearings. Magnetic bearings use magnetic levitation to support load, and are able to hold moving parts without physical contact. Also they provide high relative motion with very low friction, no mechanical wear, and no lubrication. This thesis presents state and parameter estimation algorithms, and output feedback control configurations for magnetic levitation system. The idea is to further employ these configurations in design and development of active magnetic bearings.

Sampled-data output feedback control (OFC) of magnetic levitation system for four cases along with following techniques are presented:

a. Case-I: Unknown plant model
    i. OFC based on sliding mode observer (information about input function required only) **– discussed in Chapter 5**
    ii. OFC based on state space recursive least squares filter (using constant velocity model) **– discussed in Chapter 6**

b. Case-II: Known plant model
    i. OFC based on Kalman (EKF, UKF, CKF) estimators **– discussed in Chapter 6**
c. Case-III: Plant model mismatch
    i. OFC based on NN-aided UKF **– discussed in Chapter 7**
d. Case-IV: Unknown plant parameters
    i. OFC based on dual UKF **– discussed in Chapter 8**

Figure 1-1 demonstrates different reading outlines for thesis chapters.

1.4 Conclusion

A brief overview on benefits and technologies related to smart grid has been presented. Furthermore, smart grid system is discussed as an integration of three sub-systems, smart infrastructure system, smart management system, and smart protection system.

This chapter has set direction for upcoming thesis discussion. Later chapters investigate detection, estimation, and forecast in context of nonlinear systems and processes of smart grid.

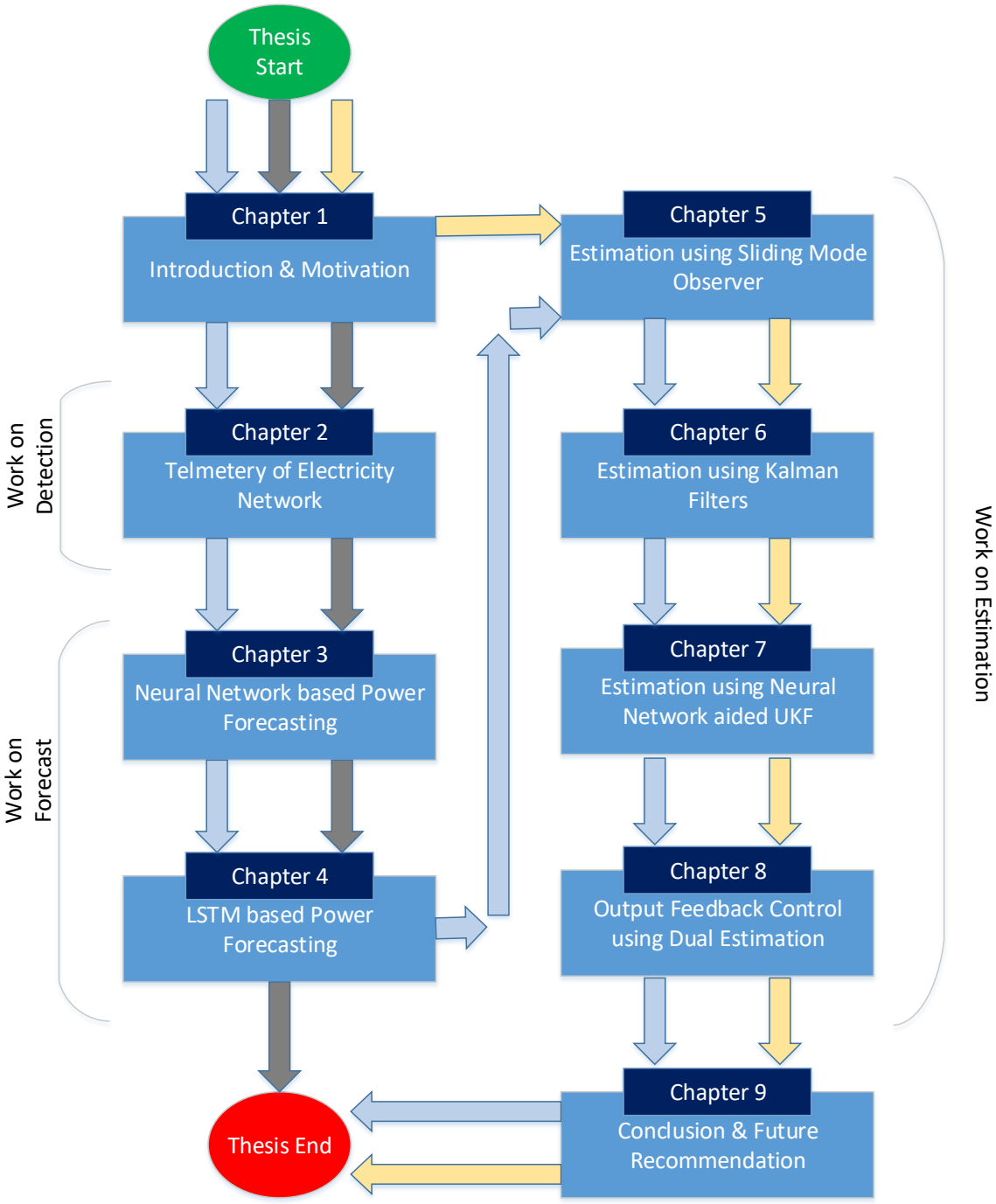

Figure 1-1. Different reading outlines for thesis chapters.

CHAPTER 2

TELEMETRY OF ELECTRICITY DISTRIBUTION COMPANY ENERGY NETWORK

We present a novel application of radio frequency wireless mesh network and general packet radio service technologies in a telemetry solution to measure power flow in the energy network of an electricity distribution company. The telemetry solution utilizes some selected circuits of grid stations, and calculates total power consumed, total power imported and total power exported by the distribution company. The selection of circuits for sensors installation is the key for reducing solution cost as compare to the case when sensors are installed on all the power output points. The framework involves installation of specially developed energy sensors (smart energy meters) and data concentrator units at the selected grid stations, for measurement of energy data that includes active energy, reactive energy, active power, apparent power, current, voltage and power factor. The measured data reach data concentrator unit using 433 MHz wireless mesh network, and is transmitted to a remote power control center using general packet radio service. Energy data from different grid stations across the energy network is collected at power control center, and utilized in calculation of total power consumed, total power import and total power export. The approach has been tested on two electricity distribution companies of Pakistan: Islamabad Electric Supply Company and Peshawar Electric Supply Company. Also in this work, result of over-load detection based on generalized likelihood ratio test for an industrial feeder of Islamabad Electric Supply Company is included. Detection probability of 0.96 with a false alarm probability of 0.04 has been achieved for 30 minutes data interval.

2.1 Introduction

Most of utilities (electricity distribution companies) in developing countries are neither able to remotely measure power usage of individual consumer, nor able to calculate real time total power consumption of complete region serviced by that utility [11]. Chun-Hao Lo in [11] has presented a detailed survey on progressive smart grid system. The

author reports that smart energy meters have multiple built-in interfaces, supporting various wireless and wired communication protocols. These smart meters, together with RF and power line communication technologies, may constitute a mesh communication network, and get utilized in smart grid solutions. The author also reports that WMN is considered as one of the foreseeable approaches to support smart grid applications; and can also manage other non-smart grid uses at the same time. Further the author has included result of a survey; that telecommunication technologies are quite mature and readily deployed these days. Therefore, Global System for Mobile communications (GSM) and GPRS have become technology candidates for supporting wireless communication in grid stations. They can be used for purposes of remote monitoring and control of substations and distributed energy sources. Also from work of [12-14], it can be concluded that present electric power system can be modernized into next generation power system (smart grid), by incremental inclusion of technology and intelligence.

In smart grid stations, real-time system monitoring and load control can also be achieved using any WMN, like ZigBee. These WMNs are also being recognized as a promising technology for present grid stations. In [15], ZigBee deployment guidelines, under interference of wireless local area network have been presented. [16] presents an experimental study on communication channel of wireless sensor networks in environments of industrial power control room and 500 kV substation.

In [17], 2.4 GHz ISM band parameters (RMS delay, coherence bandwidth and electromagnetic interferences) under different conditions of line-of-sight and polarizations are investigated for a 400 KV substation. In [18], a real-time monitoring system is presented, that monitors electrical quantities of nationwide electricity transmission network. The system was developed through National Power Quality Project of Turkey. Detailed study on wireless communication system for grid station is also presented in [19-21]. Whereas, work on modeling of customer load consumption patterns in power distribution system is discussed in [22-24].

In this chapter we present a telemetry solution for electricity distribution company. The telemetry solution architecture utilizes both RF and GPRS technologies. This hybrid combination of RF and GPRS technologies have been used in remote energy metering of

industrial, commercial, and residential consumers throughout the world. The proposed telemetry application for EDC energy network has not been considered before to the best of our knowledge. It utilizes sensors installed on power import circuits and power export circuits of an EDC energy network, to calculate the total power consumed by EDC region. The result is a reduced number of required sensors (and also solution cost); as compared to the case, when sensors are installed on every output feeder of EDC network. The solution comprises specially designed energy sensor (smart energy meter) and DCU. A developed computer server application is also installed in power control center of distribution company. The developed solution provides both, telemetry of distribution network, and real-time total power consumption value of complete region serviced by EDC. Due to brevity, designs of energy sensor, and DCU are not included in this chapter. This chapter only discusses the communication architecture. The proposed solution has been deployed for two electricity distribution companies of Pakistan under USAID Energy Policy Program. These projects were part of an effort to modernize the present power system of Pakistan.

The contributions of the chapter include a) the implementation of a real time system for an EDC that measures and monitors power import, power export and power consumed, b) the utilization of both RF and GPRS technologies in the proposed framework, and c) the identification of efficient sensor installation points resulting in lower system cost. With these contributions, logged power consumption profiles are available for analysis that helps in controlling, forecasting and planning the consumption patterns, which results in buying cheaper power from generation company.

2.2 Background

Many of the developing countries are unable to completely meet their growing energy (power) needs. In these countries, there may be a shortage of power generation due to different reasons. In such situations (countries), there is government or privately administered office, which regulates and monitors total power (megawatt) generated on national level. And distributes this generated power on allocated quota to country's electricity distribution companies.

As power generation is limited, and electricity distribution companies are provided allocated quota of this power; often a circumstance comes up, when electricity distribution companies face power shortfall. Power shortfall situation arises when power demand from EDC's end consumer exceeds EDC allocated quota, or the quota allocated (distributed) to EDC is already inadequate to fulfill its end consumer actual power need (demand).

To manage EDCs power shortfalls, the power monitoring office issues load-shedding schedules on daily or weekly basis, in which, each EDC disconnects its end consumer power (electricity) supply for duration of load-shedding. This electricity shutdown (disconnection) helps EDC to remain within received allocated power quota.

An accurate load-shedding schedule may be formulated using EDC allocated power quota and real-time total power consumption values. This schedule can help EDC to remain compliant with allocated power quota; and also not to shutdown consumer electricity for unrequired (extra) time durations.

For the case, when no automatic, real-time mechanism is in-place in EDC; that could measure and monitor actual, real-time EDC total power consumption (i.e. power consumed by EDC end consumer). EDC employs a manual mechanism, in which energy readings are read from electromechanical and digital energy meters (installed in different grid stations of EDC distribution network) by human meter readers. Human operators from different grid stations locations, via telephone calls convey these energy readings to power control center (PCC). In PCC, all the energy readings are manually summed up to calculate EDC total power consumption. For example, this manual mechanism of EDC total power consumption measurement and calculation, in some EDCs of Pakistan involve receiving around 40 telephone calls at PCC from different grid stations. Also in some cases, a total time period of 30-45 minutes is required to manually read all energy meters from different grid stations, and then to calculate one value of EDC total power (MW) consumption in PCC.

EDC power consumption value worked out by this manual mechanism do not represents a true real-time value, and also at times is inaccurate, due to human involvement at different stages of this power measurement-calculation mechanism.

Eventually, as EDCs manually-calculated power consumption values are often inaccurate and delayed, the power monitoring office ends up preparing load-shedding schedules, which either disconnects end consumer power for extra (unjustified) time durations, or make EDCs to go beyond their allocated power quotas.

The above discussed mechanism for EDC power (MW) consumption measurement and calculation; and load-shedding schedules formulation, can be improved, by developing a solution, that remotely measures power flow in EDC energy (distribution) network. As the developed solution is to be implemented in developing countries, so it should be cost-reasonable and swift in installation. With the primary objective to obtain real-time EDC total power consumption in PCC.

This work presents a telemetry solution for electricity distribution company energy network. Chapter discusses implementation and solution architecture (model) of wireless telemetry solution. The solution operates in real-time and accurately measures, both power flow in EDC distribution network and EDC total power (MW) consumption.
Details of developed solution are presented in subsequent sections of this chapter.

2.3 Proposed Solution

Before presenting architecture and functionality, solution constituent modules for EDC energy network are described:

2.3.1 Import circuit

A circuit within a grid station of EDC distribution network; from (using) which electrical power is imported (taken in) by EDC from power generation facility, or from another electricity distribution company. An EDC distribution network may have more than one import circuit, and also multiple import circuits may exist in one grid station.

An import circuit that delivers power to EDC end consumer through an output feeder is categorized as import circuit of type I. Whereas, an import circuit that steps down voltage, and distributes power to another grid station of the EDC is termed as import circuit of type II.

For example, in Pakistan, power transformer of required specification is installed in an import circuit, that steps down voltage either from 220 kV to 132 kV, or 132 kV to 66 kV, or 66 kV to 33 kV.

2.3.2 Output feeder

A circuit with power transformer, that steps down voltage and delivers power to EDC end consumer. For example, in Pakistan, power is delivered to EDC end consumer through output feeder at 11KV.

2.3.3 Export circuit

A circuit within an EDC distribution network grid station, that steps down and delivers power to grid station of another electricity distribution company. An EDC distribution network may have more than one export circuit, and also multiple export circuits may exist in one grid station. Also, both import circuit and export circuit may be present in a single grid station.

2.3.4 Common distribution point

An EDC grid station is termed as Common Distribution Point (CDP), if it distributes power to another two, or more than two grid stations. The other grid stations may be both / all of same EDC, or both / all of another EDC, or one / some of same EDC and one / some of another EDC. i.e. CDP may consist all import circuits-type II, or all export circuits, or a combination of both import circuit-type II and export circuit.

2.3.5 Power flow in an electricity distribution network

Figure 2-1 demonstrates general flow of electric power from power generation facility (PGF) to end consumer through EDCs energy (distribution) networks. Figure 2-1 shows grid stations of two electricity distribution companies, namely, electricity distribution company 'A' (EDC'A'), and electricity distribution company 'B' (EDC'B'). Grid station of EDC'A' is represented by G1; whereas, grid station of EDC'B' is represented by G2. Common Distribution Points (CDPs), power import, and power export circuits (branches) in distribution network of EDC'A' are also highlighted in Figure 2-1.

2.4 Solution Architecture and Functional Description

In this section, overall architecture and functional description of developed telemetry solution are presented. As discussed earlier, primary objective is to remotely measure from PCC, real-time EDC total power (MW) consumption (i.e. real-time total power consumed by EDC end consumers). Whereas, secondary requirement from

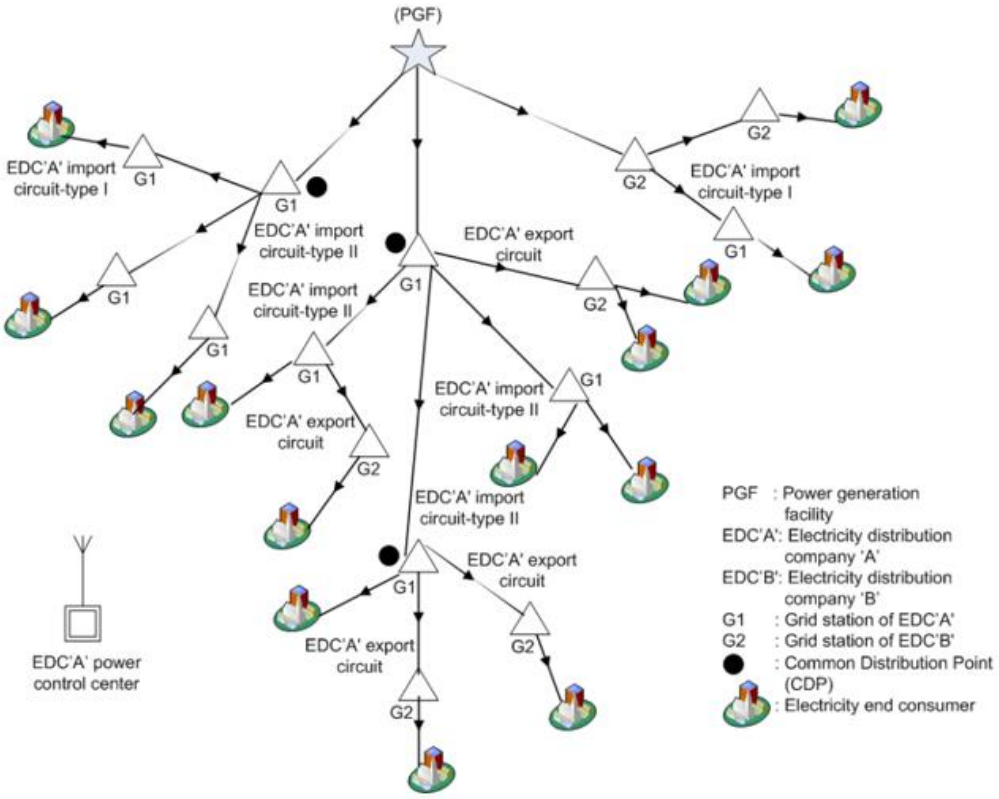

Figure 2-1. A general flow of electric power from power generation facility to end consumer of EDC 'A' and EDC 'B'.

telemetry solution, is to provide in PCC, energy parameters of power flowing in different circuits of EDC distribution network.

In the proposed developed solution, energy sensor (smart energy meter) is installed on every import circuit and export circuit of EDC distribution network. This results into a lesser number of circuits, and reduced solution cost; as compared to the case if every output feeder circuit of EDC was selected for sensor installation. Figure 2-2 demonstrates circuits selected for an EDC. Whereas, Figure 2-3 shows energy sensors installed on circuits within one grid station. The energy sensor has been designed, developed, and qualified against standards, i) WAPDA specification for Solid State TOD Energy Meters DDS-50:2007, ii) PEPCO specification Advance Metering Infrastructure (AMI) DDS-98:2011, iii) NTDC

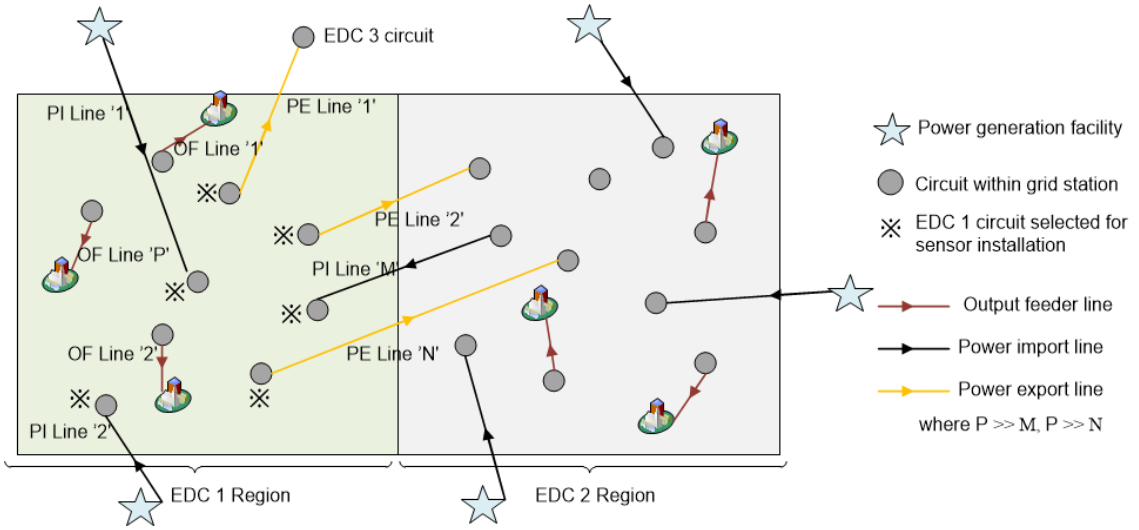

Figure 2-2. Illustration of selected circuits for DISCO '1'.

specification Meter Data Collection Server to be used in AMI DDS-110:2012. Energy parameters (active energy, reactive energy, active power, apparent power, current, voltage, power factor) of EDC import / export circuits, measured by each sensor, are wirelessly transmitted to EDC power control center via DCU. The DCU is also installed in grid station.The solution utilizes a hybrid combination of RF 433MHz WMN and GPRS technologies.

In PCC, real-time total EDC active power (MW) consumption is calculated by obtaining difference between EDC real-time values of total active power (MW) import and total active power (MW) export. i.e. let:

Active power (MW) imported by EDC via 1st import circuit at time instant $t = P_{i,1|t}$

Active power (MW) imported by EDC via nth import circuit at time instant $t = P_{i,n|t}$

Active power (MW) exported by EDC via 1st export circuit at time instant $t = P_{e,1|t}$

Active power (MW) exported by EDC via mth export circuit at time instant $t = P_{e,m|t}$

Real-time power imports ($P_{i,1|t}, P_{i,1|t}, \cdots, P_{i,n|t}$) and exports ($P_{e,1|t}, P_{e,1|t}, \cdots, P_{e,m|t}$) values are remotely obtained at EDC power control center, where real-time total power import and total power export are calculated as:

$$\text{Total power import} = \sum_{a=1}^{n} P_{i,a|t} \; ; \; \text{Total power export} = \sum_{b=1}^{m} P_{e,b|t}$$

Figure 2-3. Energy sensors installed on circuits in one grid station.

Whereas, real-time total active power (MW) consumption by EDC is calculated using real-time values of: total power (MW) consumption $=\sum_{a=1}^{n} P_{i,a|t} - \sum_{b=1}^{m} P_{e,b|t}$

Parameter *t*, i.e. time interval after which energy parameters (active energy, reactive energy, active power, apparent power, current, voltage, power factor) from each sensor are remotely retrieved, is programmable, and can have any value greater than 10 seconds.

All import and export circuits in EDC distribution network are identified with help of related EDC office.For telemetry of EDC distribution network, energy sensors are installed on circuits (of importance), which are also selected with helpof EDC office. These circuits, may include circuits, which deliver power to industries, or deliver power to populated areas of high power consumption. Energy measurements from these sensors also reach EDC power control center (via DCU), and are utilized in power monitoring of distribution network.

2.4.1 Architecture description

Figure 2-4 presents architecture of developed telemetry solution. A grid station may have any number of import circuits, export circuits and other circuits of interest. Within a grid station, the installed sensors and DCU form a wireless mesh network. Each DCU in a

grid station acquires energy data from sensors (installed on circuits) using RF 433 MHz, and uploads it on server in EDC power control center via GPRS provided by cellular company. Energy data from all sensors is acquired by DCU, and after packet formation, is uploaded on PCC server. The DCU continues with this dedicated task of data acquisition, packet formation, and uploading on PCC server.

As import / export circuits and other circuits (of interest) are present in more than one grid station, and also at different locations in EDC distribution network. Therefore, a DCU is installed in each grid station, from where circuit energy data is intended to be remotely read from EDC power control center.

Figure 2-4 demonstrates telemetry of 'N' grid stations of an EDC distribution network. A grid station may have 'N' number of circuits (sensors). Each grid station DCU acts as a gateway for sensor energy data, and uploads sensors data on PCC server.

In PCC, this sensor data is monitored and logged in files with date-time stamp. Also, this data is used to calculate real-time EDC total power (MW) consumption, as discussed earlier. As shown in Figure 2-4, monitoring of energy data is also possible from remote multiple client (node) positions. These multiple clients access PCC server data via internet.

Figure 2-5 demonstrates installation of energy sensors and DCU within an EDC grid station. Sensor is installed in grid station circuit to output of current transformer (CT) and potential transformer (PT). Whereas, DCU operates on single phase AC voltage of 100-250 V. The two types of import circuit and export circuit are also highlighted in Figure 2-5.

2.5 Detector Implementation

Each sensor (energy meter) also provides over-load protection at installed circuit. The circuit may be of any type; import, export, or a branch of EDC energy network. As the energy sensor measures power (kW) consumed through each circuit, a detector is implemented in energy meter software; which generates an alarm signal when power consumed reaches a pre-defined threshold value. The alarm signal triggers a specific circuit which disconnects all, or particular output load of that circuit, depending on EDC requirements. EDC concerned authorities set pre-defined threshold value for each energy sensor detector.

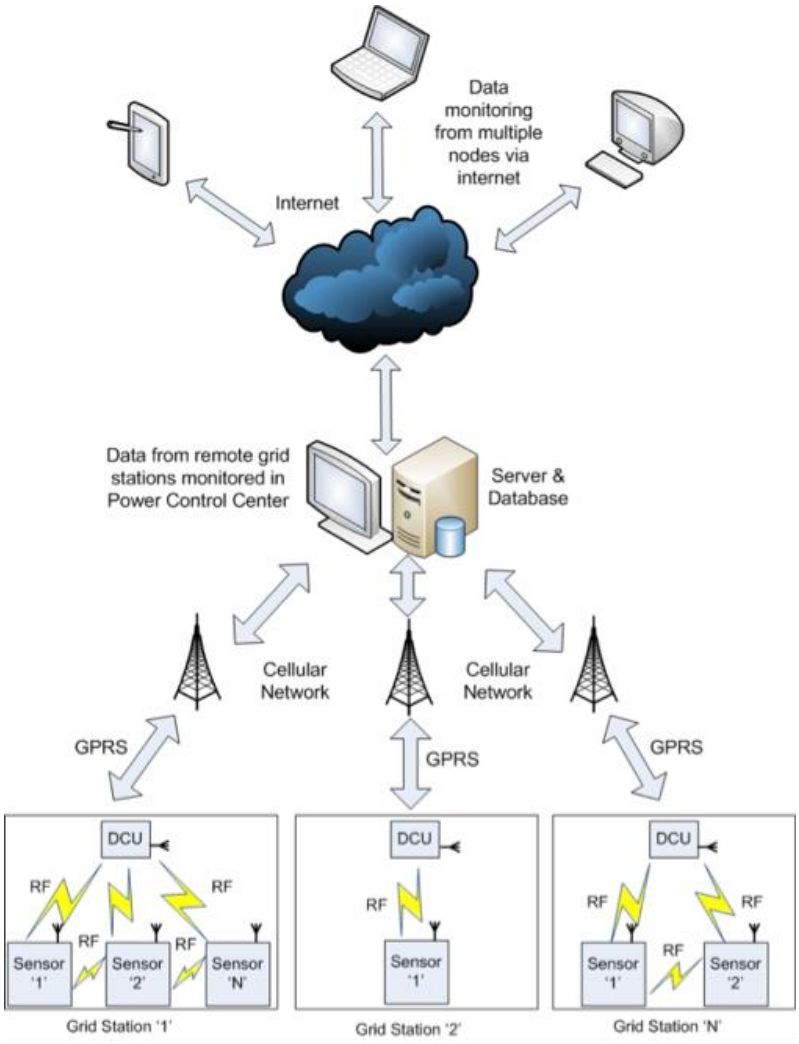

Figure 2-4. Architecture of developed telemetry solution.

2.5.1 Load consumption model for detector implementation

The measured power consumed is modeled to have an average or mean value A_N, for an interval N, with load variations (random values over a mean) modeled as additive zero mean white Gaussian noise with variance σ^2, represented by $\mathcal{N}(0, \sigma^2)$. The notation $\mathcal{N}(\mu, \sigma^2)$ denotes a Gaussian probability density function with mean μ and variance σ^2. Load consumption model utilized by sensor detector: $x[n] = A_N + w[n]$,

where, $x[n]$ represents load consumption values at time step n; $n = 0,1,2, \ldots N-1$, N is maximum number of samples. The circuit power readings are being measured at 2 $msec$

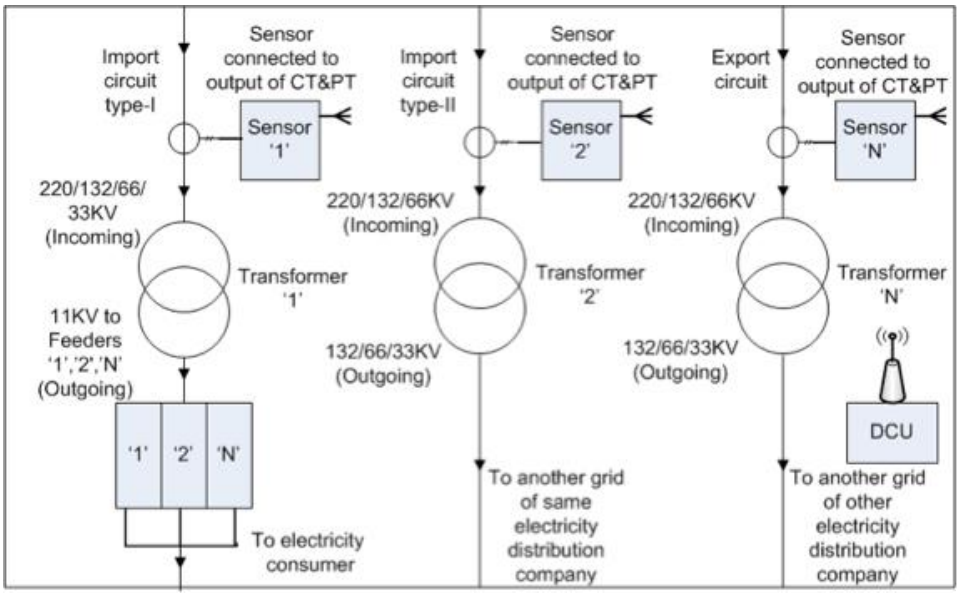

Figure 2-5. Energy sensors installed on circuits within 220/132/66/33 kV grid station.

sample time by each energy sensor; for detector implementation power readings in sensor memory are updated after every 2 seconds.

2.5.2 Problem formulation

The signal (over-load) detection problem is a binary hypothesis test, i.e. a single hypothesis must be chosen between two competing hypotheses H_0 (null hypothesis) and H_1(alternative hypothesis). The goal of detector is to decide either H_0 or H_1 based on the observed (measured) set of data $\{x[0], x[1], \dots, x[N-1]\}$. This is a mapping from each possible data set value into a decision. Following detector hypotheses are defined [25, 26]:

$$H_0: x[n] = A_N + w[n] < \gamma \; ; \; H_1: x[n] = A_{th} + w[n] > \gamma \qquad (2\text{-}5.1)$$

where, $n = 0,1,2, \dots N-1$. A_N and A_{th} are load mean value for hypothesis H_0 and H_1 respectively. γ is threshold. $A_{th}$is set by EDC on basis of circuit load consumption history. The situation of deciding H_1 when H_0 is true, is termed as false alarm. $p(H_1; H_0)$ is referred to as the probability of false alarm, and is denoted by P_{FA}. Over-load detection is defined as the case when measured load consumption value reaches the predefined value of A_{th}, i.e. deciding H_1 when H_1 is true in actual. $p(H_1; H_1)$ is referred to as the probability of detection, and is denoted by P_D. The design goal for required detector is to maximize $P_D = p(H_1; H_1)$ subject to constraint $P_{FA} = p(H_1; H_0) = \alpha$.

2.5.3 Neyman-Pearson theorm

The Neyman-Pearson (NP) approach [27, 28] has been used, which gives a test statistic, that is a function of measured load consumption $x[n]$; and a threshold value. The detector decides H_1, and a signal is generated (which is used to disconnect output load) when test statistic value exceeds the threshold value. NP approach to signal detection (hypothesis testing) states: To maximize P_D for a given $P_{FA} = \alpha$ decide H_1 if:

$$L(x) = \frac{p(x; H_1)}{p(x; H_0)} > \gamma \qquad (2\text{–}5.2)$$

where the threshold γ is found for a given value of P_{FA} from: $P_{FA} = \int_{\{x:L(x)>\gamma\}} p(x; H_0)\, dx = \alpha$ (right tail probability), and complementary cumulative distribution function [25] $Q(x)$, given by: $\frac{1}{\sqrt{2\pi}x} e^{-\frac{1}{2}x^2}$.

The function $L(x)$ is termed likelihood ratio since it indicates for each value of x the likelihood of H_1versus the likelihood of H_0. $p(x; H_0)$ and $p(x; H_1)$ are the probability density functions of $x[n]$ under hypothesis H_0 and H_1 respectively.

As in Equation (2-5.1) A_N and random variation in load (variance) σ_0^2 are unknown for hypothesis H_0, and change with time. For hypothesis H_1, A_{th} is set by EDC, and therefore known; whereas, random variation in load (variance) σ_1^2 is unknown. So, detector is to be designed when signal (load mean) is unknown, but deterministic for hypothesis H_0; and known for hypothesis H_1. Noise (load variations) is Gaussian for both hypotheses, with unknown variances.

The problem thus becomes to decide between H_0 and H_1 when the PDFs depend on a set of unknown parameters. These parameters are not the same under each hypothesis. Under H_0 the vector θ_0 represents the unknown parameters; whereas, θ_1 represents the unknown parameters under hypothesis H_1. θ_0 and θ_1 are defined as $\theta_0 = \{\sigma_0^2,\ A_N\}$, $\theta_1 = \sigma_1^2$.

The PDFs of H_0 and H_1 are represented by $p(x; \theta_0, H_0)$ and $p(x; \theta_1, H_1)$ respectively. The approach now is to estimate the unknown parameters for use in a likelihood ratio test.

2.5.4 Generalized Likelihood Ratio Test (GLRT)

The GLRT [29, 30] replaces the unknown parameters by their maximum likelihood estimates (MLEs). From Equation (2-5.2), the GLRT decides H_1 if:

$$L_G(x) = \frac{p(x; \hat{\theta}_1, H_1)}{p(x; \hat{\theta}_0, H_0)} > \gamma \qquad (2\text{–}5.3)$$

where $\hat{\theta}_1$ is the MLE of θ_1 assuming H_1 is true ($maximizes\ p(x; \theta_1, H_1)$), and $\hat{\theta}_0$ is the MLE of θ_0 assuming H_0 is true ($maximizes\ p(x; \theta_0, H_0)$). The GLRT approach has been used, as it provides information about the unknown parameters. So in this work MLEs are determined before determining $L_G(x)$.

2.5.5 Detector design

The probability density function of $x[n]$ under hypothesis H_0 and H_1 with unknown parameters have distributions, $\mathcal{N}(\hat{A}_N, \hat{\sigma}_0^2)$ and $\mathcal{N}(A_{th}, \hat{\sigma}_1^2)$ respectively, and are found as:

$$p(x; \hat{\theta}_0, H_0) = \frac{1}{(2\pi\hat{\sigma}_0^2)^{N/2}} \exp\left[-\frac{1}{2\hat{\sigma}_0^2}\sum_{n=0}^{N-1}\left(x[n] - \hat{A}_N\right)^2\right]$$

$$p(x; \hat{\theta}_1, H_1) = \frac{1}{(2\pi\hat{\sigma}_1^2)^{N/2}} \exp\left[-\frac{1}{2\hat{\sigma}_1^2}\sum_{n=0}^{N-1}(x[n] - A_{th})^2\right]$$

Where Gaussian probability density function for a scalar random variable x is defined as; $p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2\sigma^2}(x-\mu)^2\right]$, with μ and σ^2 as mean and variance. $\hat{A}_N$ estimated value of mean A_N is calculated as: $\hat{A}_N = \frac{1}{N}\sum_{n=0}^{N-1} x[n]$

The load consumption $x[n]$ variances ($\sigma_0^2,\ \sigma_1^2$) under hypothesis H_0 and H_1 are estimated as:

$$\hat{\sigma}_0^2 = \frac{1}{N}\sum_{n=0}^{N-1}\left(x[n] - \hat{A}_N\right)^2$$

$$\hat{\sigma}_1^2 = \frac{1}{N}\sum_{n=0}^{N-1}(x[n] - A_{th})^2$$

The PDF under each hypothesis have been presented above, with the difference in means causing the PDF under $H_1$to be shifted right (as $A_{th} > A_N$). Evaluating Equation (2-5.3) using expressions for $p(x; \hat{\theta}_0, H_0)$ and $p(x; \hat{\theta}_1, H_1)$:

$$\frac{\frac{1}{\left(2\pi\hat{\sigma}_1^2\right)^{N/2}}\exp\left[-\frac{1}{2\hat{\sigma}_1^2}\sum_{n=0}^{N-1}(x[n]-A_{th})^2\right]}{\frac{1}{\left(2\pi\hat{\sigma}_0^2\right)^{N/2}}\exp\left[-\frac{1}{2\hat{\sigma}_0^2}\sum_{n=0}^{N-1}\left(x[n]-\hat{A}_N\right)^2\right]} > \gamma$$

Simplifying:

$$\left(\frac{\hat{\sigma}_0^2}{\hat{\sigma}_1^2}\right)^{N/2} > \gamma \qquad (2\text{–}5.4)$$

2.6 Telemetry of IESCO and PESCO Energy Networks

The deployment of proposed telemetry solution completed in September, 2015 and October, 2016 for IESCO and PESCO respectively. To date the developed solution is effectively providing both IESCO PCC and PESCO PCC, telemetry of energy network and real-time total power (MW) consumption of EDC consumers.

This section also provides a comparison between the implemented telemetry solution and power monitoring system implemented in Turkey [18].

2.6.1 Telemetry of IESCO energy network vs Turkey power monitoring system

The electricity distribution network of IESCO delivers power to over 2.4 million consumers (www.iesco.com.pk). 52 import circuits, 11 export circuits, and 20 other circuits (of importance) in 49 grid stations were identified to be used by the proposed solution to calculate the total power consumed by IESCO region consumers. For these 83 circuits in 49 grid stations of IESCO distribution network; 83 energy sensors and 49 DCUs are installed.

Whereas IESCO region electricity consumers draw power through 929 11 KV output feeders; and if sensors were to be installed on output feeders like the Turkey power monitoring system [18], then a total of 929 sensors would have been required. This would have increased the solution cost. However, the implemented solution of this work is more cost effective, fast deployable, scalable, and more suitable for developing countries.

Values of energy parameters, power import, power export and power consumed are updated after every 15 seconds for telemetry solution deployed in IESCO. Figure 2-6 demonstrates detector performance of test statistics Equation (2-5.4) via receiver operating characteristics (ROC) curves for an output feeder of IESCO. Using Equation (2-5.4) and hypotheses definition Equation (2-5.1) relationship equation of P_D dependent on P_{FA}, test

statistics and threshold was obtained. In Figure 2-6 each point on the curve corresponds to a value of (P_{FA}, P_D) for a given threshold γ. As γ increases P_{FA} decreases and so does P_D, and vice versa. Result shows that detection probabilities of 0.96, 0.91 and 0.85 are obtained with a constraint of 0.04 false alarm probability at measured data (power) intervals of 30 minutes, 20 minutes and 10 minute respectively.

Figure 2-7 briefly presents impact of proposed telemetry and detector work on IESCO scheduled and unscheduled power (load) shedding. Approximated averaged number of hours per day load-shedding for months, January to December are shown. Power shutdown time due to faults and maintenance of IESCO distribution network has been excluded.

2.6.2 Telemetry of PESCO energy network vs Turkey power monitoring system

The electricity distribution network of PESCO delivers power to over 2.6 million consumers (www.pesco.gov.pk). 82 import circuits, 30 export circuits, and 42 other circuits (of importance) in 68 grid stations were identified to be used by the proposed solution to calculate the total power consumed by PESCO region consumers. For these 154 circuits in 68 grid stations of PESCO distribution network, 154 energy sensors and 68 DCUs are installed.

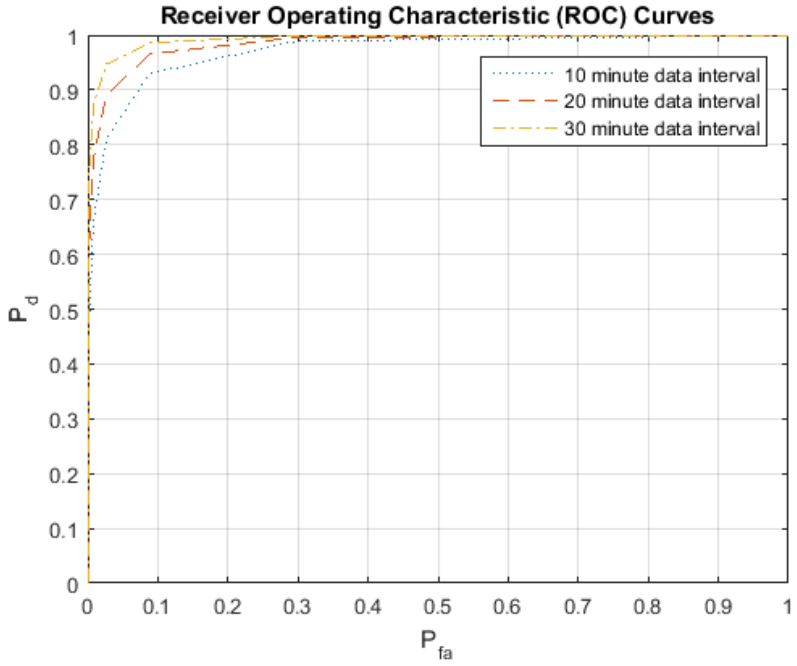

Figure 2-6. Detector receiver operating characteristics.

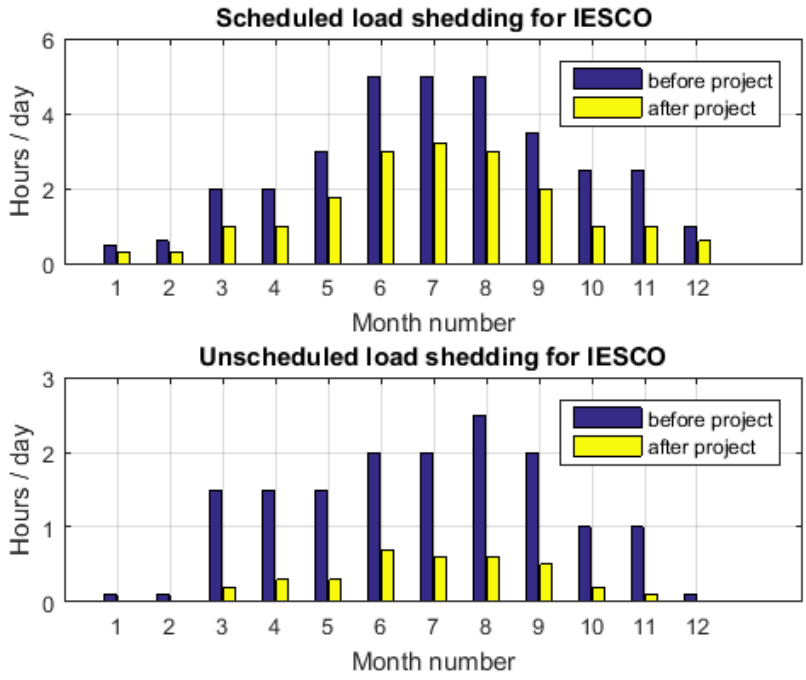

Figure 2-7. Impact of proposed solution on load-shedding schedules of IESCO.

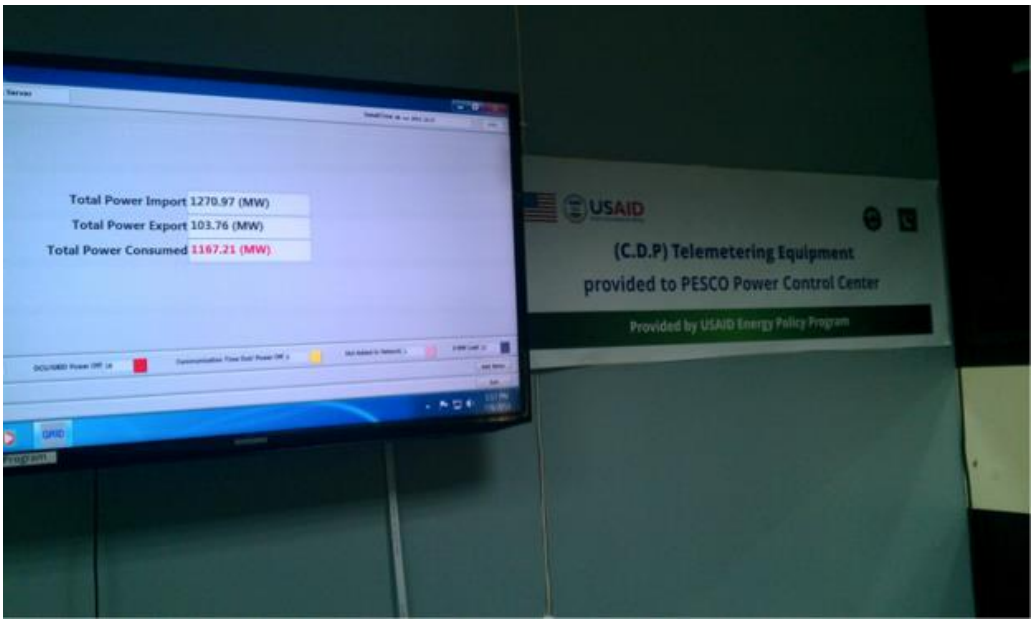

Figure 2-8. Photograph of PESCO Power Control Center server LCD screen.

Whereas PESCO region electricity consumers draw power through 969 11 KV output feeders; and if sensors were to be installed on output feeders, like the Turkey power monitoring system [18], then a total of 969 sensors would have been required. This would have increased the solution cost. Figure 2-8 shows picture of PESCO PCC server LCD screen. Real-time values of total power import, power export and power consumed by PESCO are displayed. These values are updated after every 15 seconds.

2.7 Conclusion

This work has discussed installation of energy sensors on import and export circuits, as well as DCU installation in grid stations to remotely monitor real-time total power (MW) consumed by the region serviced by an electricity distribution company. Within a grid station, the installed energy sensors and DCU form an RF 433 MHz wireless mesh network. The proposed solution via DCU offers GPRS based telemetry of EDC energy circuits. The deployed telemetry solutions for IESCO and PESCO have helped in circuits (power transformers) load balancing and quick identification of malfunctioning grid station equipment.

Next chapter will discuss impact of telemetry solution and load detection on power distribution system of Pakistan. Also, application of time series forecasting techniques for multi-variable multistep power demand forecasting will be presented.

Supporting publication: Asim Zaheer Ud Din, Yasar Ayaz, Mian Ilyas Ahmad, Momena Hasan, Salman Masaud and Naveed Muhammad, “Cost-Effective Telemetry for Energy Network of an Electricity Distribution Company: Part I”, Turkish Journal of Electrical Engineering & Computer Sciences, vol. 27, no. 2, pp. 889–902, 2019.

CHAPTER 3

MULTIVARIATE MULTISTEP SHORT-TERM POWER FORECASTING FOR ELECTRICITY DISTRIBUTION COMPANY

In this chapter, a multivariate multi-step-ahead short-term forecasting formulation has been presented. The formulation operates on multiple inputs from multiple variables, and provides multi-step-ahead forecasts by generating multiple outputs for multiple variables. The presented framework is effective for large forecasting horizons since it forecasts for temporally dependent sub-intervals called runs from large horizon. Thus the framework forecasts are less biased and suffer low variance, as compared with direct method and iterated method estimators respectively. The proposed framework has demonstrated forecasts of power import and power export with a horizon value of 48 for PESCO, Pakistan. The averaged mean absolute percentage error of two forecasted time series is 12.76 %, whereas, 24 hours ahead power consumption of PESCO total consumers has been forecasted with mean absolute percentage error of 8.6 %. Furthermore, exploiting 24 hours ahead power consumption forecasts has resulted in better power dispatch for PESCO grid stations by reducing mean absolute error by 11.52 times between PESCO power allocated and PESCO power consumed.

3.1 Introduction

Efficient storage of electricity on large scale is not always feasible. It is therefore vital that power generated at a given instance matches power requirement from consumer. Accurate forecasting of power demand is an important tool that can assist power generation company in load planning and precise load dispatch for power distribution companies [31].

The electric power forecasting is classified into four categories depending on duration of the forecasting horizon. These are very short-term power forecasting, short-term power forecasting, medium-term power forecasting, and long-term power forecasting [32]. In short-term forecasting, power is forecasted from several hours to several weeks in future using time series data of observed power [32]. These short-term power forecasts are essential for economic dispatch. In economic dispatch optimal output for electricity

generation facilities is determined fulfilling consumer power demand, subject to operational and transmission constraints, and at the lowest possible cost [33].

3.1.1 Background and motivation

Before 2014 power dispatch control in Pakistan was executed using undeveloped methods utilizing dated or unreliable data. Also, dispatch controller in National Power Control Center (NPCC) had no or little knowledge about real time operating status at power generation and power distribution sides. The dispatch operations relied mainly on controller experience and on rough estimates of available power capacity and demand at any given time instance. At times, load assessments were done on manually recorded historical data. This caused conservative power allocation to electricity distribution companies. Information about operations were verbally communicated which delayed reaction times to generation changes and dynamic load. Simple voltage frequency meter was installed in power control center, which provided indication on instantaneous power supply-demand imbalance to dispatch controller. These practices led to controller reactive actions rather than predictive decisions. As a result of this inefficient power dispatch, electricity distribution companies had to resort to unplanned or planned power outages for its consumers [34].

To resolve these issues a couple of prominent projects were initiated in 2014 under USAID (United States Agency for International Development) Power Distribution Program (PDP) and Energy Policy Program (EPP). The project under PDP involved installation of automatic meter reading devices at grid substations of 10 electricity distribution companies of Pakistan. The purpose was to monitor in near real time, power and status of 11 kV circuit breakers. In the other project funded under EPP, a telemetry solution [35] was deployed on common distribution points of electricity distribution companies, as also discussed in previous chapter. The solution provided power import and power export by distribution company. The impact of deployed solution in reducing incidents of planned and unplanned load-shedding has been reported in [35]. This work is in continuation to exploit archival data from solution [35], and move forward towards achieving a tool, that can be employed for control and coordination of power from generation facility to all the electricity distribution companies of country.

In this chapter we have proposed a multivariate multistep forecasting framework. The framework, using artificial neural network based hypothesis has been tested on power data from Peshawar Electric Supply Company, Pakistan, to produce 24 hours ahead short-term power forecasts. The framework not only offer provision to input multiple independent time series, but can also forecast multiple variables with long forecasting horizon. The proposed scheme has been evaluated for forecasting of two independent variables having a forecasting horizon of 48, while employing five independent variables (time series) for input.

3.1.2 Related work

A few relevant work relating to power forecast for nationwide and utility company consumers, recently completed in China, France, and Turkey have been reported in [36–40]. [36] presents a novel forecasting model based on an improved particle swarm optimization-extreme learning machine. The model performance comparison with back propagation neural network for China utility company is also included. In [37] long-term local power forecast has been performed using multi-level data model simulator. The forecasting method is to be further utilized for ten projects in France.

Multivariate multistep time series forecasting frameworks have recently received considerable attention due to its broad applications in power, transportation, finance, environment and etc. Authors in [41] have proposed a multivariate very short-term forecasting framework based on empirical dynamic modeling. [42, 43] report on multivariate time series forecasting employing deep learning and LSTM networks. [44] have investigated use of autoencoder (artificial neural network) based multi-to-multi mapping network for wind power forecasting for a time horizon of 24–72 hours. In [45] a hybrid forecasting method combining seasonal autoregressive integrated moving average model and neural network for multistep ahead forecasting has been demonstrated.

Artificial neural networks (ANNs) have become popular short-term power forecasting method in past decade. ANNs are data driven and can approximate any nonlinear forecasting function by learning patterns from recorded observations [46]. A few relevant references to ANN applications to short-term forecasting are such as, radial basis

function [47, 48], generalized regression neural network [49], nonlinear autoregressive neural network with exogenous input [50], multilayer perceptron [51], etc.

In most of reported recent work on ANN based short-term power forecasting, values are forecasted at one hour resolution, and for a time horizon of several hours to 24 hours [52]. Also the forecasting models are designed for multiple inputs, but they provide single output for single variable at one forecasting instance, like in [53, 54]. Further investigation on enhancing forecasting performance of model includes study on features engineering. Short-term power demand has been shown to be significantly dependent upon weather parameters like dry bulb temperature and dew point temperature as also discussed in [55, 56]. Furthermore, power requirements are driven by human activities which are normally cyclic. Therefore, providing time information to power forecasting model reduces forecasting error as also highlighted in [57].

3.1.3 Chapter contribution

The contributions of this chapter are discussed below.

1. The work in this chapter extends on univariate formulation presented in [58] to propose a multivariate multistep ahead forecasting framework. The framework utilizes one forecasting model that processes multiple inputs, and achieves multistep ahead forecast by generating multiple outputs in single forecasting instance. For case of large forecasting horizon, proposed scheme forecasted values are less biased, as compared with direct estimator method, which utilizes many forecasting models [59]. Also in comparison with iterated method for large forecasting horizon case, forecasted values in proposed scheme suffer low variance [59].
2. Most of reported work on short-term power forecasting, as also highlighted in previous section, has demonstrated forecasting performance at a resolution of 60 minutes. Whereas in our work, we have trained forecasting model on data logged at 30 minutes resolution. Training of forecasting model on high frequency and noisy data is more challenging as compared to low frequency, less noisy data [46]. Also in this work, multiple variables forecasts are generated at 30 minutes resolution for forecasting horizon.
3. Furthermore this work contributes by employing heat index as an external feature for time series forecasting to model power consumption dependence on weather. Recent

reported work in literature has mostly modeled weather dependence using dry bulb temperature and dew point temperature. Our investigation reveals that single weather parameter heat index is adequate for modeling weather influence on consumer power consumption requirement.

4. Another salient feature of proposed framework is its ability to generate multivariate forecasts for large forecasting horizons. In comparison with seasonal, stationary, and linear time series, short-term power consumption time series are non-seasonal, non-stationary, and non-linear [32]. Developing an accurate model for large forecasting horizons for such a time series requires effort. In our work, this has been achieved by dividing the forecasting horizon into sub-horizons called runs, and using a 30 minutes resolution data for training instead of 60 minutes resolution. Using a higher frequency sampled data with prior known sub-interval (run) aids in identifying periodicity in power consumption time series.

5. The overall impact of presented work for PESCO has been in form of improvement of power dispatch from power generation facility to PESCO grid stations. Before exploiting forecasted power values, the mean absolute error between PESCO power allocated and power consumed was 95.57 MW. After considering 24 hours ahead forecasts during power dispatch planning, this error has reduced by 11.52 times to an error of 8.29 MW.

3.2 Problem Formulation

Consider an univariate electric power time series S_1 with l observations $\{s_{1,t-1}, s_{1,t-2}, \dots, s_{1,t-l}\}$. To forecast single step power value in future for time series S_1, the Takens theorem [60] implies that between state of system and finite window of the time series there exists a one to one differential mapping function f, such that $R^l \rightarrow R$. This univariate multi-input single-output forecast for time series S_1 with dimension l is formulated as:

$$s_{1,t+1} = f(s_{1,t-d}, s_{1,t-d-1}, \dots, s_{1,t-d-l+1}) \quad (3\text{-}2.1)$$

Here d is lag time. Formulation (3-2.1) assumes that function f can accurately model the time series S_1. However, in most cases complete information about model is unknown, so Equation (3-2.1) is extended to formulation of stochastic time series [61]:

$$s_{1,t+1} = f(s_{1,t-d}, s_{1,t-d-1}, \dots, s_{1,t-d-l+1}) + w(t) \quad (3\text{-}2.2)$$

$w(t)$ is zero mean noise term which represents unknown information about the accurate model of function f. To forecast power values of H steps in future for time series S_1; the model Equation (3-2.2) generates one-step prediction, and predicts iteratively H times with predicted output fed as input for the next prediction till data points $\{s_{1,t+1}, s_{1,t+2}, \dots, s_{1,t+H}\}$ are obtained. Here H is forecasting horizon with value, $H > 1$.

The model Equation (3-2.2) preserves conditional dependencies between forecasted values of correlated noisy terms in time series S_1. However this recursive technique is sensitive to accumulation and propagation of prediction errors, resulting in degradation of forecasting accuracy for large value of H [59]. In [58] a multi-input multi-output formulation for univariate multistep forecasting is proposed to avoid this prediction error accumulation:

$$O_1 = F(S_1) + W \tag{3-2.3}$$

Here $F: R^l \rightarrow R^H$, uses recorded l observations $S_1 = \{s_{1,t-1}, s_{1,t-2}, \dots, s_{1,t-l}\}$ and H future data points $O_1 = \{s_{1,t+1}, s_{1,t+2}, \dots, s_{1,t+H}\}$ are forecasted. W is noise vector whose covariance is not essentially diagonal.

3.2.1 System and dataset description

This work is an extension to our work [35] also presented in previous chapter. In [35] we presented implementation of a telemetry solution for energy network of an electricity distribution company. The implemented solution wirelessly transmitted energy data (active energy, reactive energy, active power, apparent power, current, voltage, and power factor) measured by multiple energy sensors (smart energy meter) to remote PCC using hybrid combination of RF and GPRS technologies. We also showed that total power consumption of complete region serviced by an electricity distribution company can be measured at low cost (using reduced number of sensors), if sensors are installed on power import points (circuits) and power export points (circuits) rather than on all output feeders of EDC network. The proposed solution has been working in IESCO and PESCO for last around four years. The project was part of an effort to upgrade power system infrastructure of electricity distribution companies in Pakistan.

The developed solution provides provision to monitor energy data of all power import and power export points in EDC network at an interval of 10 seconds from PCC.

Also this data is logged with date and time stamp at an interval of 30 minutes in data base for analysis and planning for power consumption profiles.

In PCC real-time total active power consumption of complete region serviced by EDC is calculated using real-time values of total active power imported and total active power exported:

$$\text{Total power (MW) consumption} = \sum_{a=1}^{p} P_{i,a|t} - \sum_{b=1}^{q} P_{e,b|t} \qquad (3\text{–}2.4)$$

Where $\sum_{a=1}^{p} P_{i,a|t}$ and $\sum_{b=1}^{q} P_{e,b|t}$ are sum of power imported and power exported by EDC via p import points and q export points at time instance $t$respectively.

This chapter describes work done for PESCO. PESCO is located in Peshawar, Pakistan, and provides service of power distribution to over 2.6 million consumers of approximately 12000 square kilometers area. The data from telemetry system installed in PESCO includes parameters active energy, reactive energy, active power, apparent power, currents of three phases, voltages of three phases, and power factor of p import points and q export points. For PESCO $p = 82$ and $q = 30$, and these 112 power circuits are present at 68 different grid stations locations in PESCO region [35].These logged parameters are available for duration June 1st, 2014 to August 31st, 2018 at a resolution of 30 minutes, and therefore comprises 74544 observations. The 74544 values of both total power import P_I and total power export $P_E$for PESCO are calculated by summing active power measured from 82 import circuits and 30 export circuits respectively. The two time series are represented as:

$$Total\ power\ import\ (MW), P_I = \{p_{I,1}, p_{I,2}, \dots, p_{I,74544}\}$$

$$Total\ power\ export\ (MW), P_E = \{p_{E,1}, p_{E,2}, \dots, p_{E,74544}\}$$

The total power consumption $P_C$of PESCO is then calculated using Equation (3-2.4) to obtain time series:

$$Total\ power\ consumption, P_C = \{p_{C,1}, p_{C,2}, \dots, p_{C,74544}\}$$

3.2.2 Problem statement

The task is to forecast 24 hours ahead values of power imported $\hat{P}_I$ and power exported $\hat{P}_E$ by electricity distribution company. This objective can be achieved by

formulating a multivariate multistep forecast problem, and then generating multiple outputs for each variable to be forecasted. The number of outputs for each variable is equal to forecasting horizon H. i.e. to learn a mapping function f from n time series, with each time series having l observations; and then forecast m time series each with H outputs. Where m may not be equal to n. From Equation (3-2.3) following is formulated:

$$\mathrm{O} = \mathbf{F}(\mathrm{S}) + \mathrm{W} \tag{3-2.5}$$

Here W is noise covariance matrix, **F** is mapping hypothesis such that: $R^{(n\times l)\times 1} \rightarrow R^{(m\times H)\times 1}$; time series vectors matrix, S and output vectors matrix,O are as:

$$\mathrm{S} = [S_1 \quad S_2 \quad \cdots \quad S_n]^T ; \quad O = [O_1 \quad O_2 \quad \cdots \quad O_m]^T$$

Where, $S_1 = \{s_{1,t-d}, s_{1,t-d-1}, \dots, s_{1,t-d-l+1}\}$, $S_n = \{s_{n,t-d}, s_{n,t-d-1}, \dots, s_{n,t-d-l+1}\}$, $O_1 = \{o_{1,t+1}, o_{1,t+2}, \dots, o_{1,t+H}\}$,$O_m = \{o_{m,t+1}, o_{m,t+2}, \dots, o_{m,t+H}\}$.

The mapping hypothesis **F** in multivariate forecasting formulation Equation (3-2.5) utilizes only same time series data as input, which are to be forecasted. However in many cases, variables to be forecasted are also influenced by exogenous variables (features) obtained from other sources [62], and that are not extracted from input time series (which is to be forecasted). To further generalize, and also to improve forecasting performance of model Equation (3-2.5), matrix U is introduced to obtain:

$$\mathrm{O} = \mathbf{F}(\mathrm{U}) + \mathrm{W} \tag{3-2.6}$$

Where,

$$\mathrm{U} = [U_1 \quad U_2 \quad \cdots \quad U_{n+j}]^T = [S_1 \quad \cdots \quad S_n \quad A_1 \quad \cdots \quad A_j]^T$$
$$A_1 = \{a_{1,t-d}, a_{1,t-d-1}, \dots, a_{1,t-d-l+1}\},$$
$$A_j = \{a_{j,t-d}, a_{j,t-d-1}, \dots, a_{j,t-d-l+1}\}$$

Here $A_1, A_2, \dots, A_j$ represents j different exogenous features which influence the m forecasted time series $O_1, O_2, \dots, O_m$. To remain consistent with number of elements of each vector in matrix S, each external feature is required to have l observations in matrix U.

3.3 Problem Solution

This work extends univariate multiple output forecasting formulation Equation (3-2.3) [58] to multivariate multiple output model Equation (3-2.6) that also includes external features for its input. However for cases when the forecasting horizon H is large, order of

matrix O becomes large due to increase in dimensionality of forecasted multivariate time series. This makes multivariate estimation susceptible to large variance [44]. To countermeasure such effect different strategies including combination of models (predictors) have been reported in literature [45].

We propose a forecasting framework for multivariate multiple output problem that utilizes one model to forecast even for the cases of large forecasting horizon. The framework operates by dividing large forecasting horizon into Rnumber of sub-intervals of forecasting horizon. These sub-intervals are called *runs*. The total number of samples (forecasts) in R runs equal forecasted samples in horizon H. In this scheme, the R runs are conditionally independent, and the mapping function Equation (3-2.6) maps R different sets of input to forecast samples of R runs.

This scheme exploits the fact that for cases where time series data comprises variety of temporal patterns, conditional dependence among group of neighboring observations within each pattern is high. Whereas the conditional dependence between observations of two patterns is minimum, and decreases further as number of observation samples between two patterns increases. The number of runs R in a forecasting horizon is selected by user having a priori knowledge of system generating the time series data. i.e. the user divides the horizon H into R runs of neighboring forecast samples which form a pattern (interval) of interest to user. In the proposed scheme, it is assumed that each output variable to be forecasted has same value of H and R.

In following sections an insight into working, advantage, and application of proposed multivariate multiple output forecast framework is presented.

3.3.1 Probabilistic graphical model representation

The problem of forecasting future H values for time series $O_1, O_2, \dots, O_m$ is a task to estimate the distribution of m, H dimensional random vectors$O_1, O_2, \dots, O_m$conditional on values of U. i.e. the stochastic dependence (3-2.2) between forecasted value $o_{i,t}$, (where $i = 1, \dots, m,$) and observed input data $U_1, U_2, \dots, U_{n+j}$ generates existence of $p(O|U)$, a multivariate conditional probability with $O \in R^{(m\times H)\times 1}$ and $U \in R^{((n+j)\times l)\times 1}$. In this work, the proposed multistep forecasting model has implicitly implemented an estimator for this multivariate conditional distribution.

The distribution $p(O|U)$ becomes highly complex for cases of large values of $n, m,$ and H. To reduce complexity, and gain further benefits over the iterated and direct methods for multistep prediction; the proposed framework divides forecasting horizon H into R runs. For an insight into multivariate multiple output conditional distribution, and to highlight implemented framework benefits; Figure 3-1 demonstrates probabilistic graphical model [63] representation of distribution modeled by proposed approach for r^{th} run with $d = 0,\ m = 2,\ r^{th}\ run\ elements = 3$.

The graphical model in Figure 3-1 is equivalent to conditional independence assumption given in Equation (3-3.1). In machine learning this assumption is utilized to simplify multivariate classification problems by Naïve Bayes classifier.

$$p_i(O|U) = \prod_{a=1}^{m} p(O_m|U) \qquad (3-3.1)$$

where,

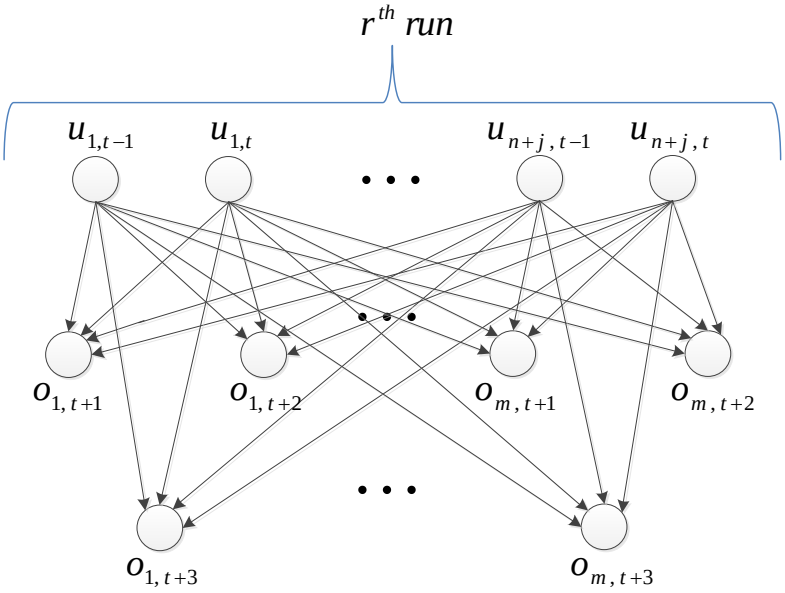

Figure 3-1. Probabilistic graphical modeling representation of r^{th} run conditional distribution $p_i(O|U)$ with $d = 0, m = 2, r^{th}$ run elements $= 3$ for multi-input multivariate multiple output prediction problem.

$$p(O_a|U) = \prod_{b=1}^{n+j} p(O_a|U_b)$$

$$= \left(\prod_{b=1}^{n} p(O_a|S_b)\right)\left(\prod_{c=1}^{j} p(O_a|A_c)\right)$$

The multivariate multiple output model Equation (3-2.6) returns multivariate estimation of joint distribution $p(O|U)$ Equation (3-3.1) for each of the runs. The idea of dividing forecasting horizon into runs, accounts for the temporal dependencies and patterns existing between components of each of the output $O_1, O_2, \dots, O_m$. This consideration results in achieving a low biased estimate as compared to the direct estimator scheme. Although the iterated method (which employs one-step ahead prediction criteria,) returns non biased estimator of the conditional distribution. But its drawback is high variance due to propagation of prediction error [64].

3.3.2 Feature selection and representation

In this section validation of proposed multivariate multiple output forecast framework is presented for short-term forecast of power imported $\hat{P}_I$ and power exported $\hat{P}_E$ by electricity distribution company. The forecasted power import and power export values are then used to calculate short-term forecast for total power consumed $\hat{P}_C$ Equation (3-2.4) by complete region serviced by electricity distribution company.

To perform short-term forecast of 24 hours ahead at a resolution of 30 minutes (for $\hat{P}_I$ and $\hat{P}_E$), for system and dataset described in Section 3.2 using formulation Equation (3-2.6) requires forecasting horizon, $H = 48$. This value of H makes forecasted values consistent with time interval of logged observations of power data for the duration June 1st, 2014 to August 31st, 2018. For system description in Section 3.2 values of $m = 2$ and $n = 2$ gives output vectors that are to be forecasted:

$$O = [O_1 \quad O_2]^T = [\hat{P}_I \quad \hat{P}_E]^T \tag{3-3.2}$$

With, $\hat{P}_I = \{p_{I,t+1}, p_{I,t+2}, \dots, p_{I,t+48}\}$, $\hat{P}_E = \{p_{E,t+1}, p_{E,t+2}, \dots, p_{E,t+48}\}$.

Further, taking into account the a priori knowledge of PESCO officials about power consumption patterns of PESCO region; the horizon $H = 48$ is divided into 12 runs, $R =$

12, each run of 2 hours (4 forecast elements) for the proposed forecasting framework of this work. Values of $d = 49, l = 2$ are selected for the input time series $\{S_1, S_2\} = \{P_I, P_E\}$ of given system. This work only addresses problem ofdetermining mapping function **F** for multivariate multipleoutput framework, and does not include value selection process for variables d and l. Good references on lag time d and order selection l are given in [51, 65]. The values $d = 48,\ l = 2$ correspond to 24 hours ago observations of power imported and power exported by PESCO region demonstrate forecasting performance acceptable to PESCO officials.

Forecasting model Equation (3-2.6) also offers provision to take into account exogenous features obtained from other sources. Our study showed that load demand of a region is affected by weather condition (heat index), and also by time and calendar variables (time-of-day and day-of-week).

Most of the work on short-term load forecasting in literature, reports use of weather parameters like dry bulb temperature, and dew point temperature in determining

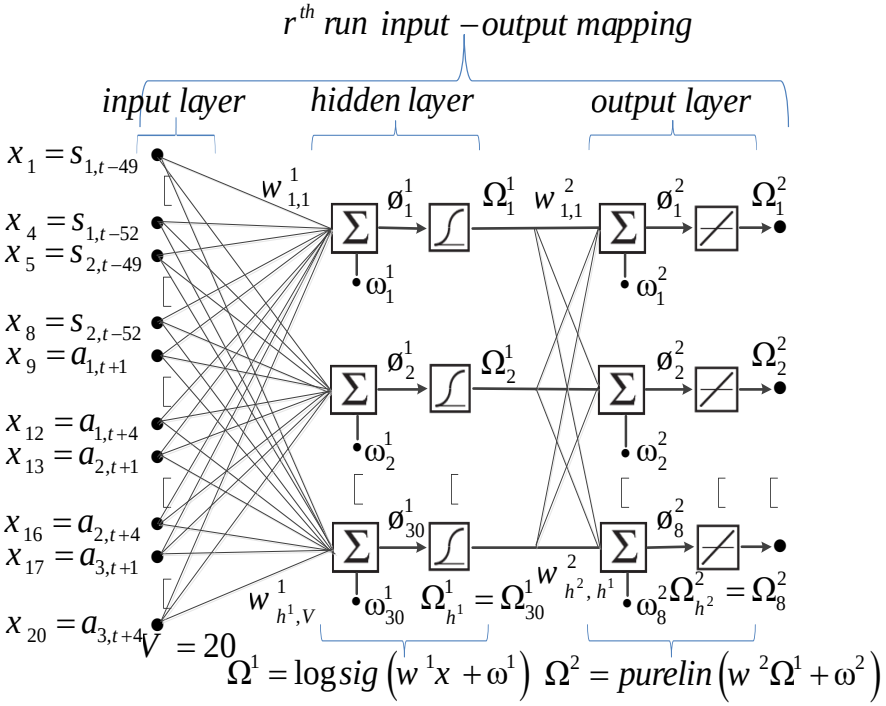

Figure 3-2. Two layers neural network implementation for multi-input multivariate multiple outputs mapping function for r^{th} run with $d = 0, m = 2, i^{th}$ run elements $= 3$.

the load demand of region for a particular time. However, we have contributed by utilizing weather parameter heat index (apparent temperature), as an input for modeling weather influence on power import and power export by PESCO region. To measure heat index we installed 100 smart sensors all across PESCO region. We then averaged the 100 heat index values from different locations to obtain a single heat index value HI, representing PESCO region. We recorded 74544 observations for HI at a resolution of 30 minutes for duration June 1st, 2014 to August 31st, 2018.

$$Heat\ index\ (℃), HI = \{HI_1, HI_2, \dots, HI_{74544}\}$$

The two other external features, time-of-day indicator, TOD and day-of-week indicator, DOW are important in determining power imported and power exported owing to routine lifestyle of human beings [57]. Therefore, at a resolution of 30 minutes, we have for TOD and DOW features:

$$Time - of - day, TOD = \{TOD_1, TOD_2, \dots, TOD_{74544}\}$$

$$Day - of - week, DOW = \{DOW_1, DOW_2, \dots, DOW_{74544}\}$$

Considering above discussed features, input vectors matrix U, with $n = 2$ and $j = 3$, for i^{th} run of $r = 12$ becomes:

$$\mathrm{U} = [U_1 \quad U_2 \quad U_3 \quad U_4 \quad U_5]^T$$

Where, $U_1 = \{s_{1,t-49}, s_{1,t-50}, s_{1,t-51}, s_{1,t-52}\}$,$U_2 = \{s_{2,t-49}, s_{2,t-50}, s_{2,t-51}, s_{2,t-52}\}$,

$U_3 = \{a_{1,t+1}, a_{1,t+2}, a_{1,t+3}, a_{1,t+4}\}$,$U_4 = \{a_{2,t+1}, a_{2,t+2}, a_{2,t+3}, a_{2,t+4}\}$,

$U_5 = \{a_{3,t+1}, a_{3,t+2}, a_{3,t+3}, a_{3,t+4}\}$

Table 1 summarizes features utilized in the proposed multivariate multiple output forecasting framework that has been implemented to generate 24 hours ahead forecast for PESCO power import and power export.

3.3.3 Forecasting model and learning algorithm

This work presents and implements ANN for modeling of hypothesis function **F** (3-2.6), which represents multivariate multiple outputs forecasting formulation. We implemented a 20 inputs ANN, $V = 20$, with 2 layers, $D = 2$, shown in Figure 3-2. The 20 inputs $\boldsymbol{x} = \{x_1, x_2, \dots, x_{20}\}^T$ for ANN were selected as there are 5 input vectors $\{U_1, U_2, U_3, U_4, U_5\}$, and each vector with 4 elements, for the forecasting problem at hand.

TABLE 1. FEATURES SUMMARY FOR MULTIVARIATE FORECASTING FRAMEWORK

Symbol	Description	Value
S_1	Power import, P_I values from past 2 hours duration (2200 – 2400) at 30 minutes interval	Measured using sensor
S_2	Power export, P_E values from past 2 hours duration (2200 – 2400) at 30 minutes interval	Measured using sensor
A_1	Heat index, HI values of 2 hours duration at 30 minutes interval, for which power forecast is required	Obtained from weather forecast
A_2	Time-of-day, TOD indicator of 2 hours duration at 30 minutes interval, for which power forecast is required	Encoded: 0000 hours $\rightarrow$ 1, 0030 hours $\rightarrow$ 1.5, 0100 hours $\rightarrow$ 2, …, 2330 hours $\rightarrow$ 24.5
A_3	Day-of-week, DOW indicator of 2 hours duration at 30 minutes interval, for which power forecast is required	Encoded: Monday $\rightarrow$ 1, Tuesday $\rightarrow$ 2, Wednesday $\rightarrow$ 3, Thursday $\rightarrow$ 4, Friday $\rightarrow$ 5, Saturday $\rightarrow$ 6,Sunday $\rightarrow$ 7

ANN first layer (hidden layer), has 30 neurons, $h^1 = 30$, with log-sigmoid transfer function for each hidden layer neuron, given by:

$$neuron\ output = \frac{1}{1 + e^{-(neuroninput)}}$$

The output of first layer is given by $\Omega^1 = \{\Omega_1^1, \Omega_2^1, \dots, \Omega_{30}^1\}^T$. The design variables, number of hidden layer neurons and log-sigmoid function were selected via experiment and literature review [46–48], [51].

ANN second layer has 8 neurons, $h^2 = 8$, with linear transfer function for each second layer neuron. The output of second layer is given by $\Omega^2 = \{\Omega_1^2, \Omega_2^2, \dots, \Omega_8^2\}^T$. $(\Omega_1^2, \Omega_2^2, \Omega_3^2, \Omega_4^2)$ are 4 forecast elements for power import $\hat{P}_I$, whereas $(\Omega_5^2, \Omega_6^2, \Omega_7^2, \Omega_8^2)$ are 4 forecast elements for power export $\hat{P}_E$. In Figure 3-2, w^1 and w^2 are weight matrices for hidden layer and output layer respectively. Also, ω^1 and ω^2 are neuron bias vectors for hidden layer and output layer respectively. The mapping hypothesis **F** in (3-2.6) for r^{th} run can be written from Figure 3-2 as:

$$\mathbf{F} = \Omega^{\mathbf{2}} = \text{purelin}(\text{w}^2\text{logsig}(\text{w}^1\boldsymbol{x} + \omega^1) + \omega^2) \qquad \text{(3-3.3)}$$

For the unknown parameters in Equation (3-3.3), parameter vector is written as following, which is to be learned:

$$\theta = [\theta_1\, \theta_2 \dots \theta_b]^T$$

$$\theta = \left[w_{1,1}^1 w_{1,2}^1 \dots w_{h^1,V}^1 \omega_1^1 \dots \omega_{h^1}^1 w_{1,1}^2 \dots \omega_{h^D}^D\right]^T \qquad \text{(3-3.4)}$$

here, $b = h^1(V+1) + h^2(h^1+1) + \cdots + h^D(h^{D-1}+1)$, $h^1 = 30$, $h^2 = 8$, $V = 20$ and $D = 2$.

In this work, the unknown parameters of mapping hypothesis Equation (3-3.3) have been determined by considering this a training problem, and applying Levenberg-Marquardt algorithm to the multilayer neural network shown in Figure 3-2. The performance index for this multilayer network training is sum of squared errors over the Gtargets in the training set $\{(\boldsymbol{x}_g, \boldsymbol{y}_g); g = 1,2, \dots, G\}$:

$$E(\theta) = \sum_{g=1}^{G} (\boldsymbol{y}_g - a_g^2)^T (\boldsymbol{y}_g - a_g^2)$$

$$= \sum_{g=1}^{G} (\boldsymbol{e}_g)^T (\boldsymbol{e}_g) = \sum_{g=1}^{G} \sum_{i=1}^{h^D} (e_{i,g})^2 = \sum_{i=1}^{B} (v_b)^2 \qquad \text{(3-3.5)}$$

Here, $e_{i,g}$is the i^{th} element of the error for g^{th}input-target pair, and $B = G \times h^D$. From Equation (3-3.5) error vector is:

$$\boldsymbol{v} = [v_1 v_2 \ \ldots v_B]^T = \left[e_{1,1}\, e_{2,1} \ldots\ e_{h^D,1} e_{1,2} \ \ \ldots\ \ e_{h^D,G}\right]^T$$

Thus the parameter update equation as per Levenberg-Marquardt algorithm for mapping hypothesis (3-3.3) is [66]:

$$\theta_{k+1} = \theta_k - [J^T(\theta_k)J(\theta_k) + \mu_k I]^{-1} J^T(\theta_k)\boldsymbol{v}(\theta_k)$$
$$\Delta\theta_k = -[J^T(\theta_k)J(\theta_k) + \mu_k I]^{-1} J^T(\theta_k)\boldsymbol{v}(\theta_k) \qquad (3\text{-}3.6)$$

I is identity matrix. The useful feature of Levenberg-Marquardt algorithm is, as μ_k is decreased to zero the algorithm becomes Gauss-Newton; whereas, the algorithm approaches steepest descent algorithm with small learning rate as μ_k is increased. $J(\theta_k)$ is Jacobian matrix given by:

$$J(\theta) = \begin{bmatrix} \frac{\partial e_{1,1}}{\partial w^1_{1,1}} & \frac{\partial e_{1,1}}{\partial w^1_{1,2}} & \cdots & \frac{\partial e_{1,1}}{\partial w^1_{h^1,V}} & \frac{\partial e_{1,1}}{\partial \omega^1_1} & \cdots & \frac{\partial e_{1,1}}{\partial \omega^1_{h^1}} & \frac{\partial e_{1,1}}{\partial w^2_{1,1}} & \cdots & \frac{\partial e_{1,1}}{\partial \omega^D_{h^D}} \\ \frac{\partial e_{2,1}}{\partial w^1_{1,1}} & \frac{\partial e_{2,1}}{\partial w^1_{1,2}} & \cdots & \frac{\partial e_{2,1}}{\partial w^1_{h^1,V}} & \frac{\partial e_{2,1}}{\partial \omega^1_1} & \cdots & \frac{\partial e_{2,1}}{\partial \omega^1_{h^1}} & \frac{\partial e_{2,1}}{\partial w^2_{1,1}} & \cdots & \frac{\partial e_{2,1}}{\partial \omega^D_{h^D}} \\ \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ \frac{\partial e_{h^D,1}}{\partial w^1_{1,1}} & \frac{\partial e_{h^D,1}}{\partial w^1_{1,2}} & \cdots & \frac{\partial e_{h^D,1}}{\partial w^1_{h^1,V}} & \frac{\partial e_{h^D,1}}{\partial \omega^1_1} & \cdots & \frac{\partial e_{h^D,1}}{\partial \omega^1_{h^1}} & \frac{\partial e_{h^D,1}}{\partial w^2_{1,1}} & \cdots & \frac{\partial e_{h^D,1}}{\partial \omega^D_{h^D}} \\ \frac{\partial e_{1,2}}{\partial w^1_{1,1}} & \frac{\partial e_{1,2}}{\partial w^1_{1,2}} & \cdots & \frac{\partial e_{1,2}}{\partial w^1_{h^1,V}} & \frac{\partial e_{1,2}}{\partial \omega^1_1} & \cdots & \frac{\partial e_{1,2}}{\partial \omega^1_{h^1}} & \frac{\partial e_{1,2}}{\partial w^2_{1,1}} & \cdots & \frac{\partial e_{1,2}}{\partial \omega^D_{h^D}} \\ \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ \frac{\partial e_{h^D,G}}{\partial w^1_{1,1}} & \frac{\partial e_{h^D,G}}{\partial w^1_{1,2}} & \cdots & \frac{\partial e_{h^D,G}}{\partial w^1_{h^1,V}} & \frac{\partial e_{h^D,G}}{\partial \omega^1_1} & \cdots & \frac{\partial e_{h^D,G}}{\partial \omega^1_{h^1}} & \frac{\partial e_{h^D,G}}{\partial w^2_{1,1}} & \cdots & \frac{\partial e_{h^D,G}}{\partial \omega^D_{h^D}} \end{bmatrix} \qquad (3\text{–}3.7)$$

Here $V = 20, D = 2, h^1 = 30$ and $h^2 = 8$. G is number of targets in training set. The terms of Jacobian matrix have been computed as [66]: $[J]_{c,L} = \tilde{s}^d_{i,c} \times \Omega^{d-1}_{J,g}$ or if $\theta_L$is a bias, then:$[J]_{c,L} = \tilde{s}^d_{i,c}$. Where $c = (g-1)h^D + k$ and $\tilde{s}^d_{i,c}$ is Marquardt sensitivity, given by:

$$\tilde{s}^d_{i,c} \equiv \frac{\partial v_c}{\partial \emptyset^d_{i,g}} = \frac{\partial e_{k,g}}{\partial \emptyset^d_{i,g}}$$

$\emptyset^d_{i,g}$ is i^{th} element of net input at layer d on g^{th} input-target pair.

3.3.4 Training and forecasting framework

The dataset obtained from telemetry system installed in PESCO is for the duration June 1st, 2014 to August 31st, 2018at a resolution of 30 minutes, and comprises 74544 observations. The data for interval June 1st, 2014 to May 22nd, 2017 is categorized as training set, whereas the data for interval May23rd, 2017 to August 31st, 2018 is test set

used for evaluating 24 hours ahead forecast generated by proposed framework. The outputs (observations) in training set are used to learn unknown parameters of forecasting model.

This section presents the training of unknown parameter vector Equation (3-3.4) of mapping hypothesis Equation (3-3.3) for multiple input, multivariate, multiple output forecasting formulation Equation (3-2.6). Before generating forecast, training of hypothesis Equation (3-3.3) is performed by LM algorithm using features identified in Section 3.3 from data for duration June 1st, 2014 to May 22nd, 2017. This training dataset represented by $\{(\boldsymbol{x}_g, \boldsymbol{y}_g); g = 1,2, \dots, G\}$ comprises of 52181 training instances, $G = 52181$.

Figure 3-3 briefly highlights steps involved in LM algorithm. The algorithm provides a balance between guaranteed convergence of steepest descent and speed of Gauss-Newton's method. Implementation of LM backpropagation algorithm is summarized as:

a. The training set $\{\boldsymbol{x}_g; g = 1,2, \dots, G\}$ is presented to network Equation (3-3.3). Each training vector has 20 elements $\{x_1, x_2, \dots, x_{20}\}^T$. For each vector of training set the network output and error $\boldsymbol{e}_g = \boldsymbol{y}_g - a_g^2$is computed. Then sum of squared errors $E(\theta)$for complete training set is computed using Equation (3-3.5).
b. The Jacobian matrix $J(\theta)$ is computed using Equation (3-3.7).
c. $\Delta\theta_k$ is computed using Eqaution (3-3.6) and the sum of squared errors Eqaution (3-3.5) is recomputed using $\theta_k + \Delta\theta_k$.
d. If the new error is less than error computed in step (a), compute updated parameters $\theta_{k+1} = \theta_k + \Delta\theta_k$, divide μ by ζ and return to step (a). If the new error is greater than error computed in step (a), then multiply μ by ξ and return to step (c).
e. The algorithm is continued till any of condition mentioned in Figure 3-3 is fulfilled.

On training completing, the hypothesis Equation (3-3.3) is used to forecast 24 hours ahead values for two variable, power import $\hat{P}_I$ and power export $\hat{P}_E$ for PESCO. The mapping hypothesis generates 8 outputs for every test vector; 4 forecasts each for both variables $\hat{P}_I$ and $\hat{P}_E$in all of the 12 runs. The hypothesis is tested with test set $\{(\tilde{\boldsymbol{x}}_{\tilde{g}}, \tilde{\boldsymbol{y}}_{\tilde{g}}); \tilde{g} = 1,2, \dots, \tilde{G}\}$ for $\tilde{G} = 22363$ test vectors. Each test vector has 20 inputs

$\{\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_{20}\}^T$. The test set is prepared using features identified in Section 3.3 from data for duration May 23rd, 2017 to August 31st, 2018. $\tilde{\boldsymbol{y}}_{\tilde{g}}$ is target output for $\tilde{g}^{th}$ test vector (instance) given by:

$$\tilde{\boldsymbol{y}}_{\tilde{g}} = [P_{I,\tilde{g}} \quad P_{E,\tilde{g}}]^T$$

Where $P_{I,\tilde{g}}$ and $P_{E,\tilde{g}}$ are power imported and power exported by PESCO at $\tilde{g}^{th}$ time instance respectively. Henceforth, $\tilde{\boldsymbol{y}}$ represents actual values (power) for 24 hours duration (one day) for test set data May 23rd, 2017 to August 31st, 2018 given by:

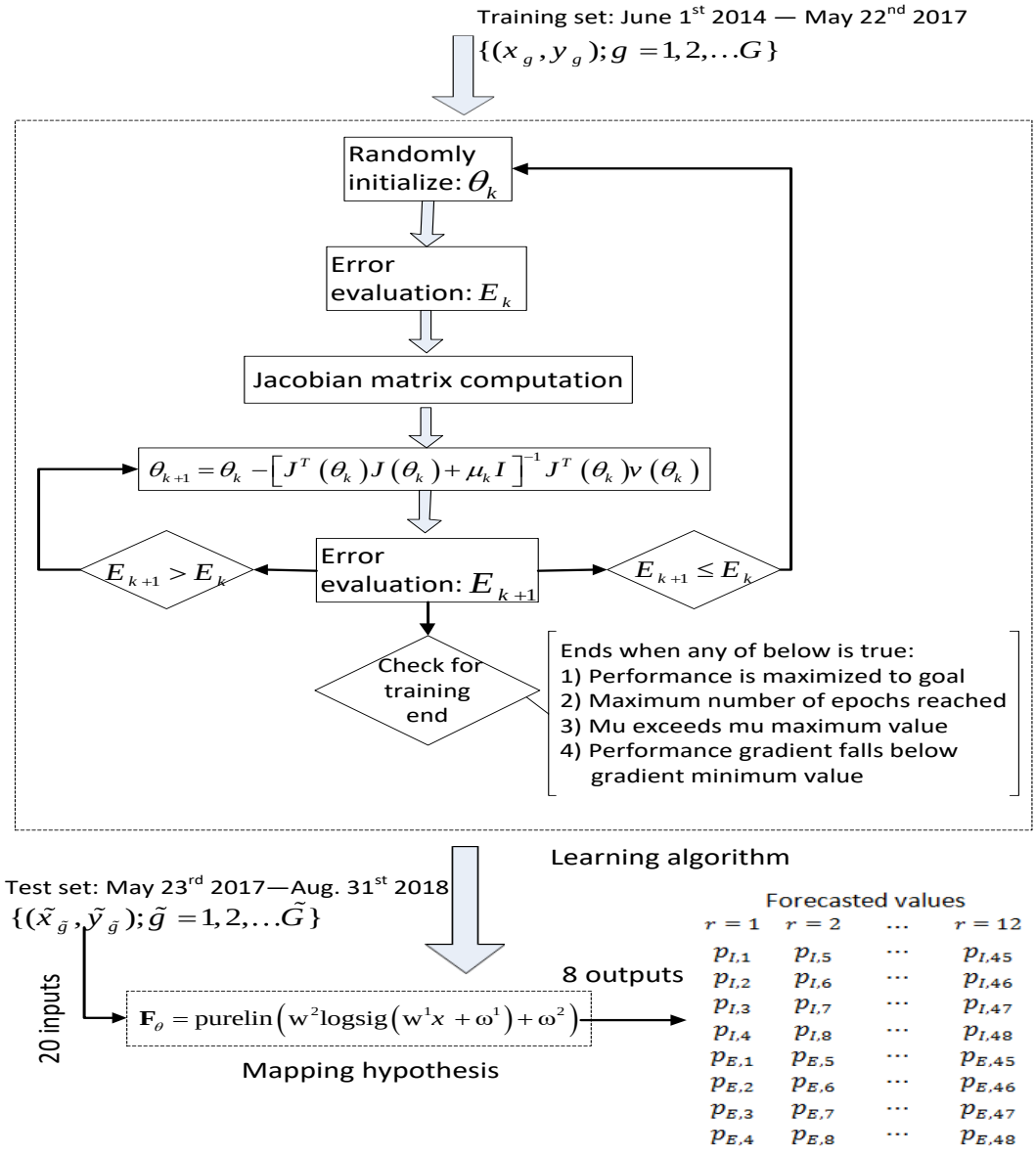

Figure 3-3. Brief overview of learning and forecasting framework.

$$\tilde{\boldsymbol{y}} = [P_I \quad P_E]^{\boldsymbol{T}}$$

With, $P_I = \{p_{I,t+1}, p_{I,t+2}, \dots, p_{I,t+48}\}$, $P_E = \{p_{E,t+1}, p_{E,t+2}, \dots, p_{E,t+48}\}$ are actual power imported and power exported respectively for 24 hours at 30 minutes interval.

The hypothesis Equation (3-3.3) with learned parameter θ Equation (3-3.4) is utilized in formulation Equation (3-2.6) to generate 24 hours ahead forecast $\mathrm{O} = [\hat{P}_I \quad \hat{P}_E]^T$ Equation (3-3.2) for 466 days of test set data May 23rd, 2017 to August 31st, 2018. The forecasted output O is compared against actual (target) $\tilde{\boldsymbol{y}}$ power values for these 466 days of test set data, to establish performance of proposed multivariate multiple output forecasting framework.

The 24 hours ahead forecast for $\hat{P}_I$ and $\hat{P}_E$ has been achieved by dividing the forecasting horizon, $H = 48$ into 12 runs, $R = 12$, as shown in Figure 3-3. In every run, 2 hours ahead forecast for both variables $\hat{P}_I$ and $\hat{P}_E$ is generated. The selection of 4 samples (2 hours) forecast for both variables $\hat{P}_I$ and $\hat{P}_E$ in a single run has been decided on basis of prior knowledge about PESCO region power consumption pattern.

3.3.5 Performance metric

To evaluate forecasting accuracy, the output $\mathrm{O} = [\hat{P}_I \quad \hat{P}_E]^T$ of hypothesis Equation (3-3.3) is compared against output $\tilde{\boldsymbol{y}} = [P_I \quad P_E]^{\boldsymbol{T}}$ of test set. As discussed earlier the test set comprises of power imported and power exported by PESCO for 466 days, $D = 466$. In literature various performance metrics (measures) have been proposed [67]. This work has utilized two performance measures namely, mean absolute error (MAE) and mean absolute percentage error (MAPE). Their definitions and salient features follow:

3.3.5.1 Mean absolute error

The performance metric $(MAE)_I$ evaluates power import forecast against actual (target) power imported, and is defined as:

$$(MAE)_I = \frac{1}{D}\sum_{c=1}^{D}(\varepsilon_I)_c$$

Where $D = 466$, and ε_I given by:

$$\varepsilon_I = \frac{1}{48}\sum_{c=1}^{48}\left|p_{I,c} - \hat{p}_{I,c}\right|$$

Similarly $(MAE)_E$ is calculated to evaluate power export forecast against actual (target) power exported using $D = 466$, and ε_E:

$$\varepsilon_E = \frac{1}{48}\sum_{c=1}^{48}\left|p_{E,c} - \hat{p}_{E,c}\right|$$

The performance metric MAE measures average absolute deviation of forecast from output, and shows magnitude of cumulative forecasting error. Neither direction of error is known, nor large forecast errors are penalized by MAE. MAE depends on data transformation and scale of measurement. Smaller the value of MAE, better is the forecast performance.

3.3.5.2 Mean absolute percentage error

The performance metric $(MAPE)_I$ also evaluates power import forecast against actual (target) power imported, and is given by:

$$(MAPE)_I = \frac{1}{D}\sum_{c=1}^{D}(\epsilon_I)_c$$

Where $D = 466$, and ϵ_I given by:

$$\epsilon_I = \frac{1}{48}\sum_{c=1}^{48}\left|\frac{p_{I,c} - \hat{p}_{I,c}}{p_{I,c}}\right| \times 100$$

Similarly $(MAPE)_E$ is calculated to evaluate power export forecast against actual (target) power exported using $D = 466$, and ϵ_E:

$$\epsilon_E = \frac{1}{48}\sum_{c=1}^{48}\left|\frac{p_{E,c} - \hat{p}_{E,c}}{p_{E,c}}\right| \times 100$$

The performance metric MAPE represents percentage of average absolute error occurred. Like MAE, in MAPE neither direction of error is known, nor large forecast errors are penalized. MAPE is affected by data transformation but is independent of scale of measurement.

3.4 Results

This section presents results of proposed multivariate multistep forecasting framework (shown in Figure 3-3) for data set which has been obtained from telemetry system [35] installed in Peshawar region of Pakistan under USAID EPP for Pakistan. The framework forecasts 24 hours ahead values of power imported and power exported by region, serviced by Peshawar electricity distribution company. These forecasted powers are then used to calculate 24 hours ahead power consumption of PESCO region using Eqaution (3-2.4). The benefits of using power import and power export to forecast and calculate power consumption is discussed in our work [35].

Figure 3-4 qualitatively demonstrates the heat index, power import, and power export training data at 30 minutes resolution, which has been used to train hypothesis Equation (3-3.3) using Levenberg-Marquardt algorithm Equation (3-3.6). The Levenberg-Marquardt algorithm parameters $\mu = 0.001$ (for $k = 0$), $\zeta = 10$, $\xi = 10$ are set for learning unknown variables Equation (3-3.4) of hypothesis. The values for other two external features, day-of-week and time-of-day were encoded as shown in Table 1 at 30 minutes resolution for training data duration June 1st, 2014 – May 22nd, 2017. From Figure 3-4, it is noticeable that there are more high frequency components in power export training data as compared to power import training data. These high frequency components have adversely affected forecasting performance for power export in this work. For a multivariate forecasting problem taking inputs with different band of frequency components, such a behavior in forecasting performance has been reported [68].

In Figure 3-5 qualitative curves for actual power import P_I and forecasted power import $\hat{P}_I$ are plotted at a resolution of 30 minutes for test interval May 23rd, 2017 – August 31st, 2018. On the horizontal axis the months are categorized in quarters as: Q1 is (January – March), Q2 is (April – June), Q3 is (July – September), and Q4 is (October – December). The proposed forecasting framework with ANN based hypothesis has been validated for test dataset, and generated 466 forecast of 24 hours ahead values for power import. The performance measures MAE and MAPE for forecast of power import have values of 98.72 (MW) and 8.9 % respectively. In Figure 3-5, residual plot, $Error = P_I - \hat{P}_I$ has also been plotted at 30 minutes resolution for test dataset duration.

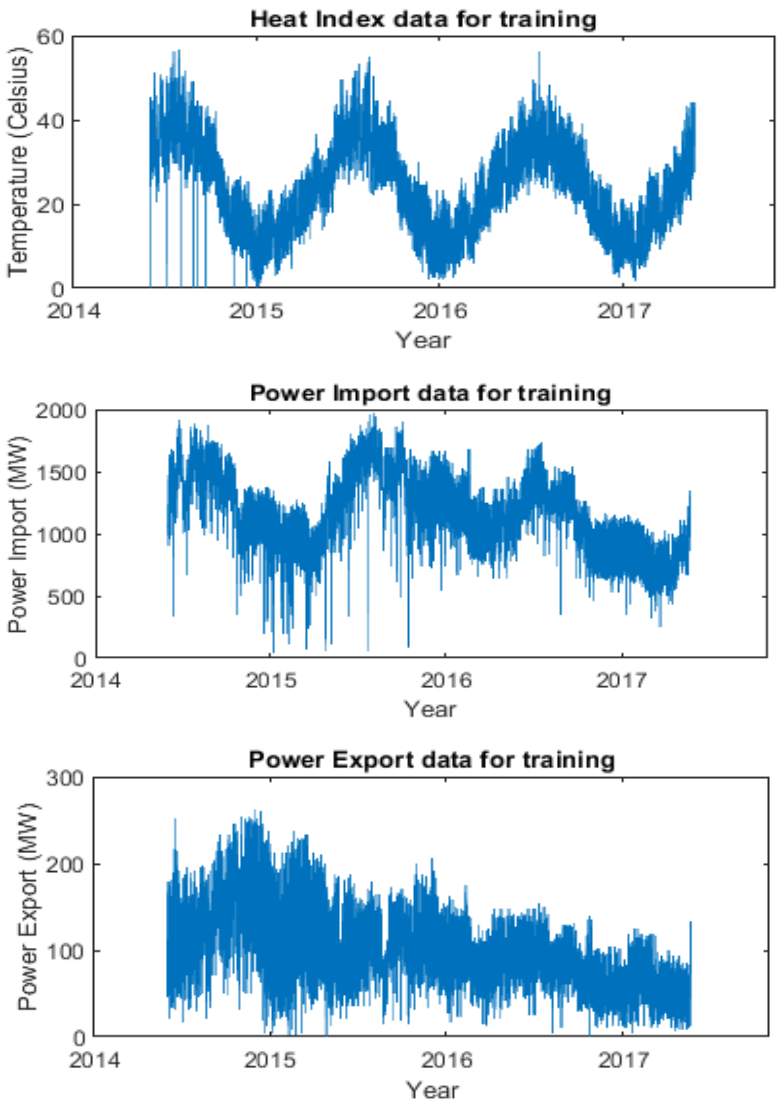

Figure 3-4. Qualitative representation of heat index, power import, and power export data used for learning forecasting hypothesis.

Figure 3-6 represents qualitative curves for actual power export P_E and forecasted power export $\hat{P}_E$ plots at a resolution of 30 minutes for test interval May 23rd, 2017 – August 31st, 2018. The implemented framework has been validated for test dataset, and generated 466 forecast of 24 hours ahead values for power export at 30 minutes resolution. The performance measures are $MAE = 12.56$ (MW) and $MAPE = 16.63$ % for forecast of power export. The values are also tabulated in TABLE 2. As indicated earlier the values of MAE and MAPE for power export forecast are higher as compared

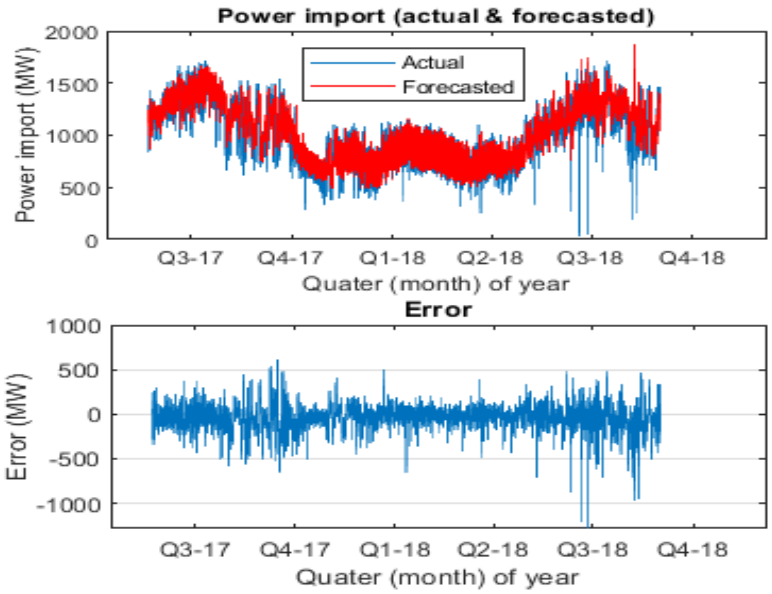

Figure 3-5. Qualitative representation of actual and forecasted power import for test dataset.

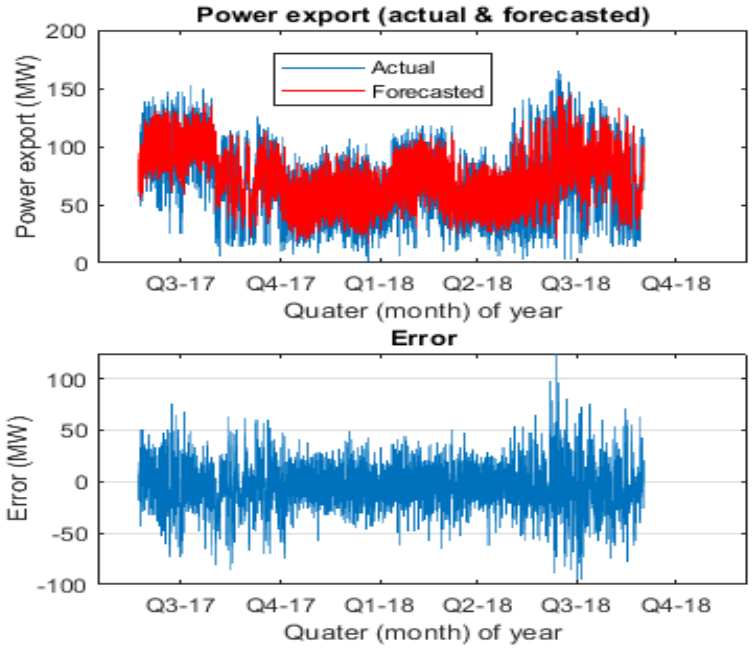

Figure 3-6. Qualitative representation of actual and forecasted power export for test dataset.

TABLE 2. RESULTS SUMMARY FOR MULTIVARIATE FORECASTING FRAMEWORK

Variable	Mean Absolute Error	Mean Absolute Percentage Error
Power Consumed	91.75 (MW)	8.6 %
Power Import	98.72 (MW)	8.9 %
Power Export	12.56 (MW)	16.63 %

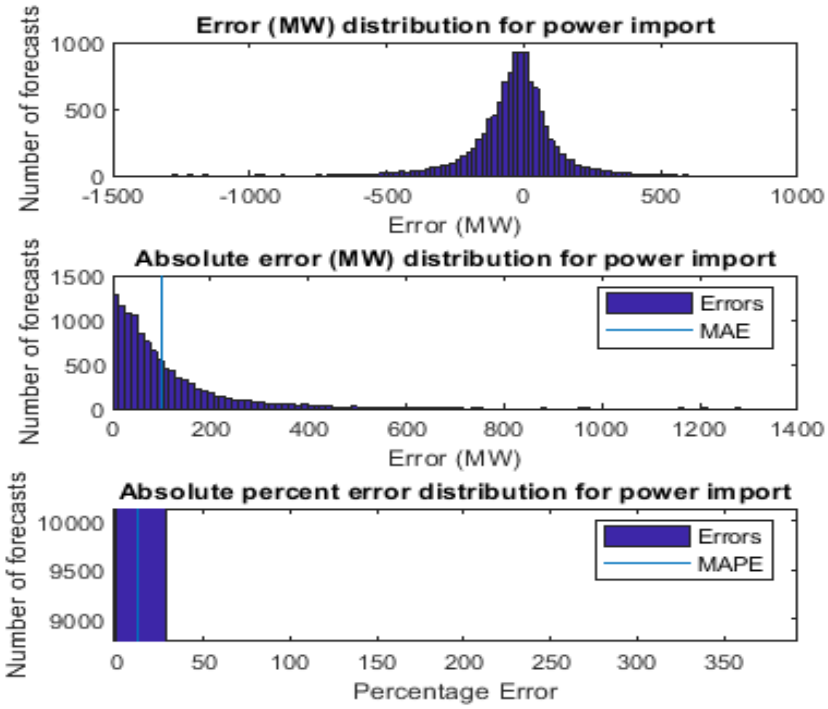

Figure 3-7. Error distribution of power import forecasts for test dataset.

with MAE and MAPE of power import forecast. In Figure 3-6, residual plot, $Error = P_E - \hat{P}_E$ has also been plotted at 30 minutes resolution for test dataset duration.

For further qualitative analysis of forecasting performance, Figure 3-7 and Figure 3-8 present error distribution (via histogram plots) for power import and power export forecasts respectively. Absolute error distribution and absolute percent error distribution plots for power import and power export forecasts are also shown.

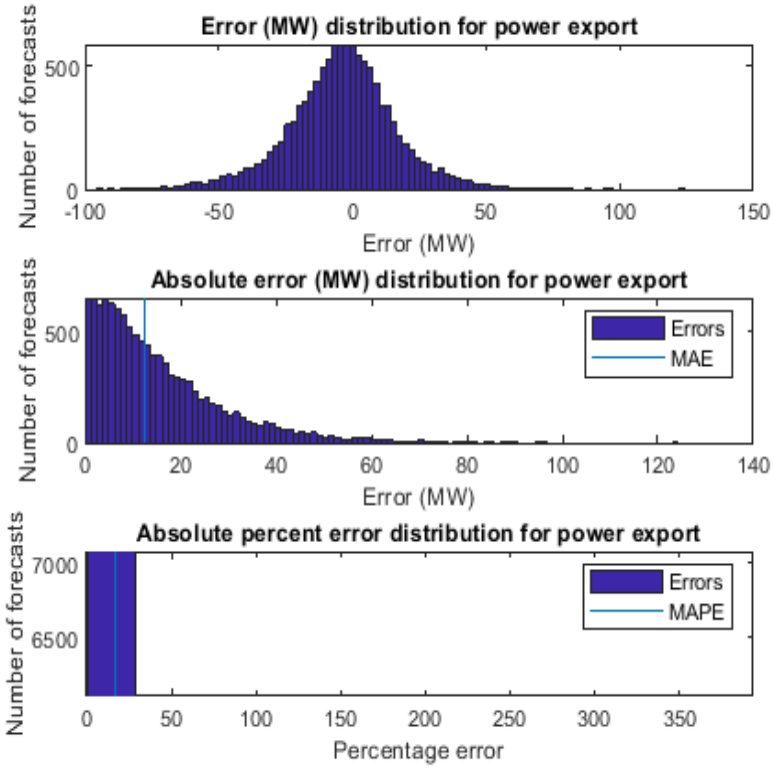

Figure 3-8. Error distribution of power export forecasts for test dataset.

Additional investigation into forecasting performance is presented by randomly selecting three weeks from the test dataset duration May 23rd, 2017 – August 31st, 2018. For each of the three weeks actual powers (P_I, P_E) and forecasted powers $(\hat{P}_I, \hat{P}_E)$ are plotted along their residual plots $(P_I - \hat{P}_I)$ and $(P_E - \hat{P}_E)$. MAPE for 24 hours ahead forecast for 7 days of week is also included. Figure 3-9 randomly shows actual and forecasted plots along their error for following three weeks of test dataset duration: 31/10/17 – 11/6/17, 1/2/18 – 1/8/18, and 4/24/18 – 4/30/18. MAPE values of forecasted power import for these three weeks are 16.21%, 5.77% and 6.81% respectively. Whereas three weeks MAPE values of 32.20%, 23.44% and 19.97% are obtained for forecasted power export.

Table 2 exhibits one year average MAPE values of 8.9% and 16.63% for forecasted power import and forecasted power export for test duration respectively. The MAPE of forecasted power import is less compared to MAPE of forecasted power export.

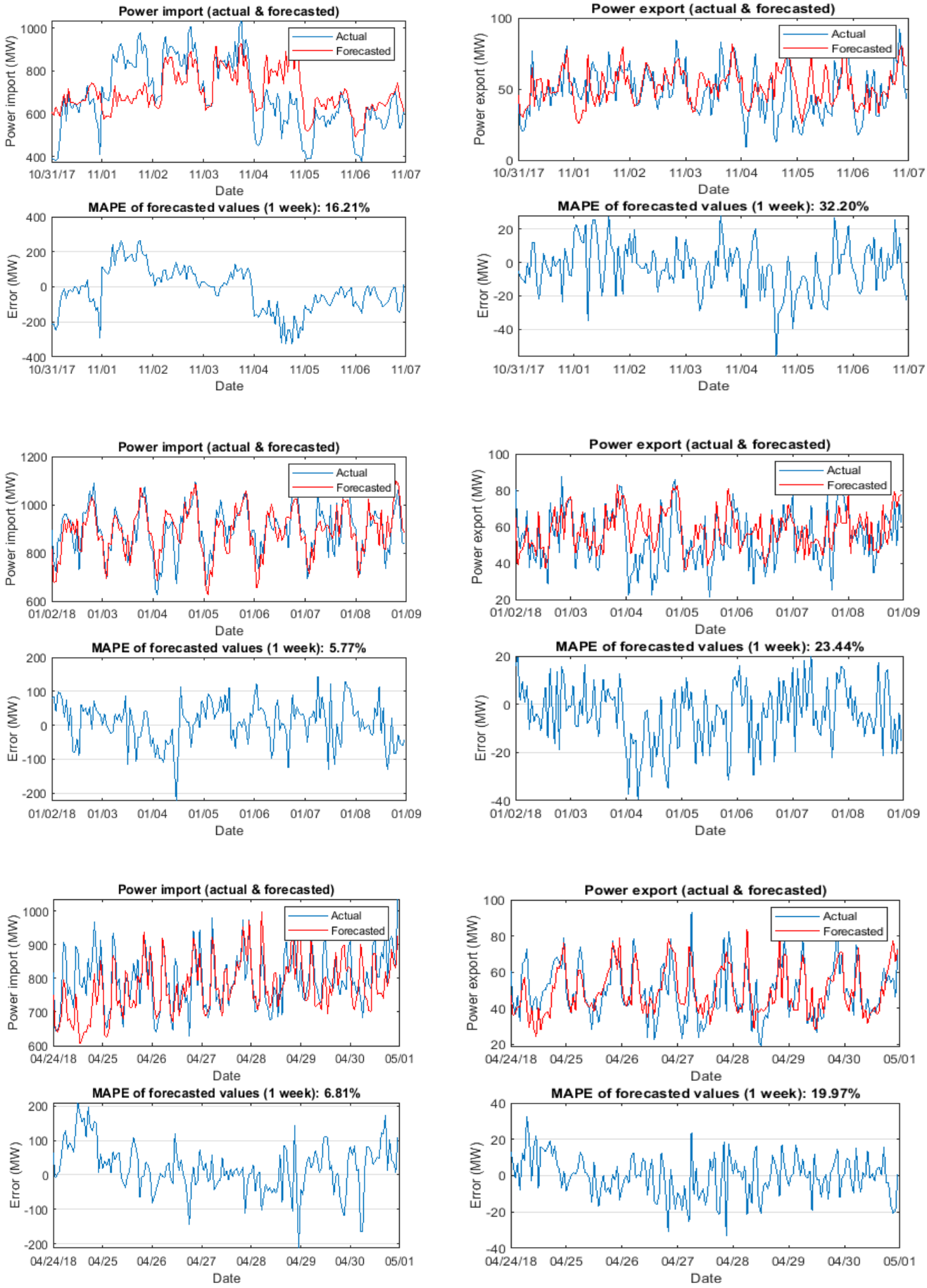

Figure 3-9. Qualitative representation of actual & forecasted power import and actual & forecasted power export for three randomly selected weeks from test dataset.

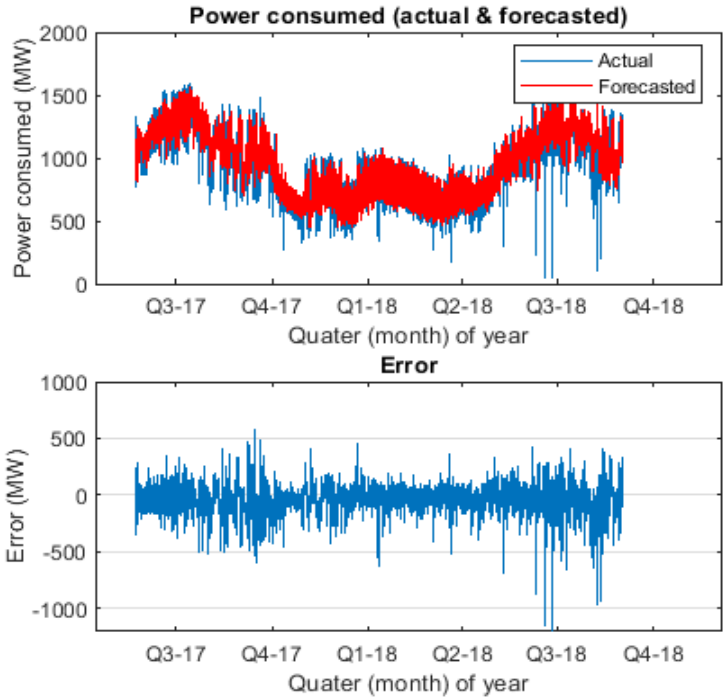

Figure 3-10. Qualitative representation of actual and forecasted power consumption for test dataset.

This result is also reflected in Figure 3-9, where MAPE values of forecasted power import are less than forecasted power export for randomly plotted weeks. This superior forecasting performance for power import is due to better training of ANN forecasting model for power import as compared to power export. This occurred because training data available for power import shows more variety and variance as compared to power export data. This is qualitatively evident from Figure 3-4. The training instances for power export show more repetition, and hence forecasting model not trained enough to forecast for new instances of test duration. Furthermore, the variation in weekly MAPE values of forcasted power import and power export, are due to variation in periodicity and similarity of power values among days of a week. For a week that have days when actual power import or actual power export shows more periodicity, the MAPE values of forcast are less for that week. This exhibited from plots of Figure 3-9.

In Section 3.2 a brief description of system [35] was given. The system provides solution to compute real-time power consumption of a region by employing, real-time

power imported by region EDC, and real-time power exported by region EDC. As discussed in introduction section, 24 hours prior information about power requirement of a region is helpful for processes of power generation, planning, and distribution. Figure 3-10 demonstrates both actual and forecasted values of total power consumption by PESCO region for test duration May 23rd, 2017 – August 31st, 2018, at 30 minutes resolution. Actual power consumption, P_C has been computed using values of (P_I, P_E) for test duration in (3-2.4). Whereas forecasted power consumption, $\hat{P}_C$ has been computed using values of $(\hat{P}_I, \hat{P}_E)$ for test duration in (3-2.4). Figure 3-10 also exhibits residual $(P_C - \hat{P}_C)$ at a resolution of 30 minutes. The performance metric for test dataset is $MAE = 91.75$ (MW) and $MAPE = 8.6\ \%$ for 24 hours ahead forecast of total power consumption by PESCO region. Results are also summarized in Table 2.

3.5 Power Forecasts Impact Assessment

The developed telemetry solution [35] for EDC (also discussed in previous chapter) was designed to provide real-time monitoring of important power circuits within PESCO energy network; and also monitoring of interface (import & export) circuits between EDCs energy networks and national power transmission grid, operated byNational Transmission Dispatch Company (NTDC) Pakistan. [35] has discussed the effectiveness of deployed solution in terms of reduction in incidents of unplanned and planned load-shedding. This work is in continuation to our previous work [35]. This section briefly assesses the impact of utilizing logged power data from PESCO telemetry system, for NTDC generation and PESCO distribution system.

Information sources for power forecast impact assessment include NPCC and EPP offices. These sources provided data on national power grid, e.g. load-shedding schedules, power generation, dispatch, and distribution. Whereas, half hourly logged power import, power export, and power consumption values of complete region serviced by PESCO, was obtained from PESCO power control center.

The impact of short-term (24 hours ahead) power forecasting on basis of logged data from PESCO telemetry system has shown both indirect and direct benefits. These benefits (impacts) have been evaluated at EDC and NPCC levels; and quantified wherever possible, given data and confidentiality constraints.

Logged power data in PCC of PESCO, and demand supply data for PESCO provided by NTDC for duration January 1st, 2018 – February28th, 2019 was analyzed to assess the direct impacts. The telemetry solution was primarily designed to monitor total power consumption of EDC region. However, additional features of tracking, analyzing, and forecasting power consumption values, have provided NTDC, and PESCO the ability to better match generated power dispatch with actual power consumption of PESCO. Resulting in more precise load dispatch.

Furthermore, power forecasting has considerably assisted in reducing requirement for holding back large percentage of generated power. These high operating reserves were previously maintained to respond to unforeseen sudden increase in power demand at short notices. The operator in PCC can now more correctly set spinning margins (reserves), as compared to past cases of making conservative estimates. Therefore, a noticeable impact of increase in power availability for dispatch has been achieved via reduction of operating reserves requirement.

In Figure 3-11, power difference between power allocated and power consumed by PESCO for duration September 1st, 2018 – November30th, 2018 is plotted at 30 minutes resolution. Here an adjustment factor previously used by NPCC has been added to PESCO power allocation. Quantifying power difference curve of these three months gave $MAE = 95.57$ (MW). Also in Figure 3-11, power difference between power allocated and power consumed with forecasting aid has been plotted for duration December 1st, 2018 – February28th, 2019. This time an adjustment factor derived from 24 hours ahead forecast was added to PESCO power allocation. Again quantifying power difference curve for these three months gave $MAE = 8.29$ (MW) [49]. Thus utilizing 24 hours ahead power consumption forecasted values, generated a more precise load dispatch for PESCO electricity consumers, and reduced the MAE by 11.52 times.

For the indirect benefits, deployment of telemetry solution along with power forecasting provision has already started to exhibit its economic impacts on industrial, commercial, and residential sectors of PESCO [50]. The improvement in PESCO's power supply situation is also creating a positive social impact on its population. The reduction

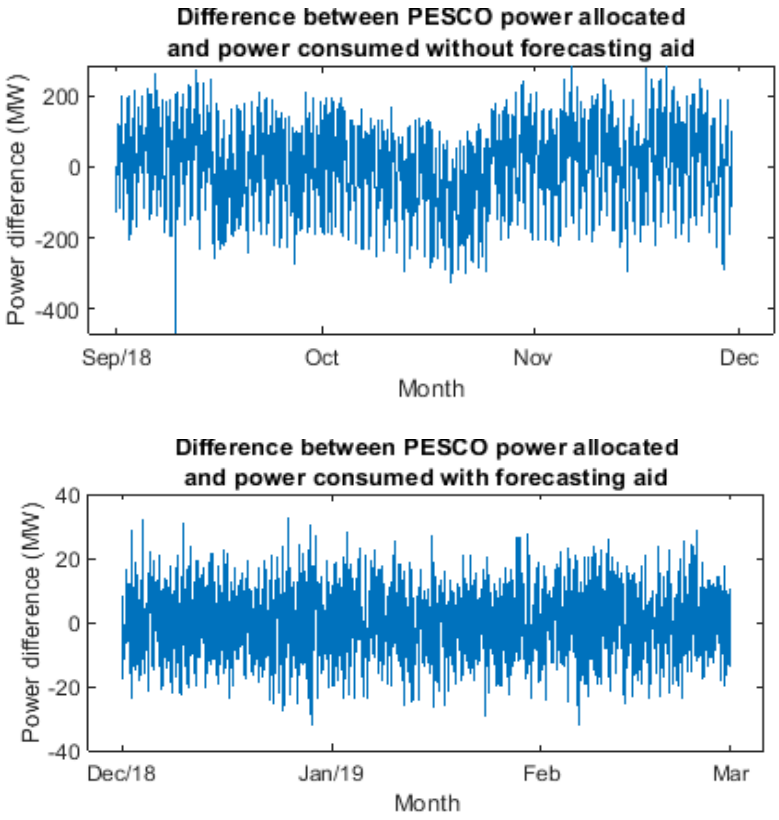

Figure 3-11. Power forecast impact assessment in terms of power dispatch to PESCO electricity consumer.

in unplanned load-shedding, and provision of more power supplies is directly helping elevate Human Development Index values [71].

The deployed telemetry system is assisting PESCO officials in making preemptive actions and maintenances, while monitoring system imbalances. The system is also providing true archival data. The data can be used by electricity distribution company to understand their consumer power demand profiles. And then make accurate power obligation from generation facility accounting seasonal and daily demand variations.

3.6 Conclusion

Back in year 2010-2011, and its subsequent years, Pakistan faced a persistent power crisis, which was a function of faulty transmission system, insufficient power generation, and increasing demand. There was no system intact to get real time grid performance data. Inefficient power dispatch and suboptimal utilization of generated power capacity were

common practices. This resulted into nationwide electricity shortages. The work presented in this chapter is part of an effort to upgrade Pakistan power system.

In next chapter, short-term forecasts for power import and power export by EDC using LSTM network in multivariate, multistep forecasting framework is discussed.

Supporting publication: Asim Zaheer Ud Din and Yasar Ayaz, "Multivariate Multistep Short-Term Power Forecasting for Electricity Distribution Company", Journal of King Saud University – Computer and Information Sciences (in review), 2021.

CHAPTER 4

BIVARIATE SHORT-TERM ELECTRIC POWER FORECASTING USING LSTM NETWORK

In this chapter, we have utilized LSTM network to generate short-term 24 hours in advance forecast for two (bivariate) independent time series. The work presents LSTM forecasting performance for three different weight optimizing algorithms, namely, Adaptive moment estimation, Root mean square propagation, and Stochastic gradient descent with momentum. Also, investigation into forecasting performance on changes in LSTM network and training options has been made. Furthermore, effects of different input features on LSTM short-term forecasts are demonstrated. The presented work has been employed for PESCO 4 years electric power data, recorded at 30 minutes resolution. From all the forecasting test cases of import power and export power for PESCO; the lowest values obtained are MAPE = 9.47 % and MAPE = 12.37 % for import power and export power respectively.

4.1 Introduction

Short-term power forecasting is key research area for assisting in economic electricity dispatch of power systems. Major benefits of economic dispatch include optimal utilization of available power generation facilities, reduced electricity cost, and minimum or no requirement for power outages (load-shedding) [72].

4.1.1 Background and motivation

The work presented in this chapter is in continuation to our earlier work [35, 73], discussed in previous chapters. [35] discusses development of a wireless telemetry solution deployed in Islamabad and Peshawar Electric Supply Companies of Pakistan. Whereas [73] demonstrates 24 hours in advance forecast of total power consumption by Peshawar electricity consumer, based on data from system [35]. This work is an extension to prevous chapter and [73]. All the three work are part of an effort to propose enhanced solutions for advance infrastructure system in smart grid regime; and also to bring betterment in power generation and distribution mechanism of Pakistan.

As also briefed in previous chapters Pakistan power system before 2014 was mostly manually driven with no real-time status information about power generated and power consumed. This led to sub-optimal utility of power generation facilities and high frequency scheduled and unscheduled load-shedding. The system [35] not only has been effectively functioning for last around 5 years, but also providing important data for consumer electricity profiling.

In this chapter we have demonstrated bivariate 24 hours ahead short-term forecasting using LSTM network. The forecasting performance for two variables, import power and export power of PESCO is investigated using three training algorithms, Adaptive moment estimation, Stochastic gradient descent with momentum, and Root mean square propagation. Effect on performance due to changes in LSTM network and training options are also reported. Furthermore, LSTM network forecasting for different sets of input features is also discussed.

4.1.2 Related work

LSTM network [74] a type of recurrent neural network has been reported for short-term electric power forecasting applications, such as [75, 76]. LSTM networks have shown better learning ability for temporal sequence as compared to deep neural networks [77]. A variety of optimization algorithms exist which are used for training of LSTM networks. [78] reports Adaptive moment estimation (Adam) optimized LSTM network for electricity price forecasting. The work demonstrates better performance than stochastic gradient based optimization on dataset from New South Wales, Australia. In [79] authors have proposed an adaptive rate learning algorithm, and a variance reduction technique to speed up convergence of stochastic gradient descent (SGD) algorithms for deep neural networks. Author in [80] proposed Root mean square propagation (RMSProp) algorithm for neural network optimization for cases of big dataset. The algorithm works by dividing dataset into mini-batches, and uses recent magnitude of gradient running average for dividing current gradient.

Most of the reported work on 24 hours ahead electric power forecasting have performed single variable prediction and that too at a resolution of 60 minutes, like in [57].

Furthermore, reported literature also highlights enhancement in short-term power forecasting performance on utilizing weather and time information [57, 62].

4.1.3 Chapter contribution

The contributions of this chapter are briefly highlighted as:

1. This chapter has performed short-term electric power forecasting for two independent variables, whereas, most of reported work has presented single variable power forecasting. We have demonstrated 24 hours in advance forecasts for total import power and total export power of PESCO using LSTM. Contrary to short-term electric power forecasts at 60 minutes resolution; we trained our LSTM network with 30 minutes resolution dataset, and also generated bivariate forecasts at 30 minutes interval.
2. Furthermore in this work a comparison of LSTM forecast performance is also included. The comparison is built for different cases. The cases are prepared on basis:
    i. Changes in LSTM network.
    ii. Changes in training options.
    iii. Change in training algorithm (Adam, SGD, RMSProp).
    iv. Different sets of feature from electric power time series, weather information, and time indication.
3. To achieve 24 hours ahead forecast with horizon $H = 48$, two forecasting methods have beenemployed. First method uses recursive strategy, while second uses direct strategy.
4. Another contribution of 24 hours forecasts of import power and export power is assistance to PESCO in its planning for i) consumer power requirement, and ii) power export to another distribution company.

4.2 Problem Formulation

This section mathematically formulates problem discussed in previous section. Consider *n* time series $\{I_1, I_2, \dots, I_n\}$ each with *l* recorded observations. Whereas two time series represented by $O_1$and O_2 are of interest and to be computed. Each of the time series $O_1$and O_2 is required to be known for H observations in future. Where H is called

forecasting horizon. The output time series $O_1$and O_2 are obtained by deriving a transformation function f that utilizes time series $\{I_1,\ I_2, \dots, I_n\}$ as input:

$$\{O_1,\ O_2\} = f\{I_1,\ I_2, \dots, I_n\} \quad (4\text{-}2.1)$$

such that f is mapping function: $R^{(n\times l)} \rightarrow R^{(2\times H)}$. And: $I_1 = \{i_{1,t-1}, i_{1,t-2}, \dots, i_{1,t-l}\}$, $I_2 = \{i_{2,t-1}, i_{2,t-2}, \dots, i_{2,t-l}\}$, $I_n = \{i_{n,t-1}, i_{n,t-2}, \dots, i_{n,t-l}\}$,

$O_1 = \{o_{1,t+1}, o_{1,t+2}, \dots, o_{1,t+H}\}$, $O_2 = \{o_{2,t+1}, o_{2,t+2}, \dots, o_{2,t+H}\}$.

4.2.1 System and dataset description

The work presented in this chapter is based on our telemetry system [35] which has been deployed in Pakistan two electricity distribution companies namely, IESCO and PESCO. Details about system can be read from [35] and Chapter 2. In addition to other features, the system provides provision for monitoring and logging of following important parameters of import and export circuits of EDC energy network.

i. Active energy (kWh)
ii. Reactive energy (kVArh)
iii. Active power (kW)
iv. Apparent power (kVA)
v. Current (A)
vi. Voltage (V)
vii. Power factor

These parameters are logged at a resolution of 30 minutes. One of our system working in PESCO has recorded above data (parameters) for duration June 1st, 2014 to August 31st, 2018 with 74544 observations. Let P_I and P_E represent total import power and total export power of PESCO for these 74544 observations:

$$Total\ import\ power\ (MW), P_I = \{p_{I,1}, p_{I,2}, \dots, p_{I,74544}\}$$

$$Total\ export\ power\ (MW), P_E = \{p_{E,1}, p_{E,2}, \dots, p_{E,74544}\}$$

Calculation of total import power P_I and total export power P_E is discussed in our work [35]. Other then recorded P_I and P_E time series, our another work [73] provides heat index of PESCO region for duration June 1st, 2014 to August 31st, 2018 at 30 minutes resolution. The time series is represented as:

$$Heat\ index\ (℃), \qquad HI = \{HI_1, HI_2, \dots, HI_{74544}\}$$

4.2.2 Problem statement

The objective of current work is to forecast 24 hours in advance two independent variables $\hat{P}_I$ (import power) and $\hat{P}_E$ (export power) for PESCO at a resolution of 30 minutes. This has been achieved using formulation (4-2.1), and utilizing three time series P_I, P_E and HI described in previous section. From (4-2.1):

$$\{\hat{P}_I,\ \hat{P}_E\} = f\{I_1,\ I_2, I_3, \dots, I_n\} \tag{4-2.2}$$

Where $(I_1,\ I_2, I_3, ...)$ represents features derived from time series (P_I, P_E, HI), and I_n in Equation (4-2.2) represents exogenous features (variables). The transformation function f performance is dependent on these features for generating forecast.

4.3 Problem Solution

This work proposes to determine a single model for 24 hours ahead forecast as compared to direct method [59]. In direct method multiple models are derived to generate forecast for large forecasting horizon *H*. This work utilizes one model to forecast two independent time series $\hat{P}_I$ and $\hat{P}_E$ for $H = 48$ using different sets of input. The input sets and forecasting methods are elaborated later in this section. The unknown mathematical model f in Equation (4-2.2) has been determined from training data comprising both the inputs and the desired output via supervised learning technique.

In following sections features selection (input sets), forecasting model and methods, learning algorithms, and performance metric for forecasting evaluation are briefly presented.

4.3.1 Feature selection and representation

This section discusses features which have been used in current bivariate forecasting work. Related work on power forecasting has reported use of additional data features derived from observed time series [75]. Also use of time information, e.g. hour of day information, and day of week information have shown profound influence on

TABLE 3. FEATURES SUMMARY FOR LSTM FORECASTING MODEL

Symbol	Description	Value
I_1	One sample (30 minutes) ago power import value	Obtained from recorded observations of power import, P_I time series
I_2	One sample (30 minutes) ago power export value	Obtained from recorded observations of power export, P_E time series
I_3	48 samples (24 hours) ago power import value	Obtained from recorded observations of power import, P_I time series
I_4	48 samples (24 hours) ago power export value	Obtained from recorded observations of power export, P_E time series
I_5	336 samples (168 hours) ago power import value	Obtained from recorded observations of power import, P_I time series
I_6	336 samples (168 hours) ago power export value	Obtained from recorded observations of power export, P_E time series
I_7	Heat index value of time instance for which forecast is required	Obtained via weather forecast
I_8	Time of day information of time instance for which forecast is required	Encoding: 0000 hours $\rightarrow$ 1, 0030 hours $\rightarrow$ 1.5, 0100 hours $\rightarrow$ 2, …, 2330 hours $\rightarrow$ 24.5
I_9	Day of week information of time instance for which forecast is required	Encoding: Monday $\rightarrow$ 1, Tuesday $\rightarrow$ 2, Wednesday $\rightarrow$ 3, Thursday $\rightarrow$ 4, Friday $\rightarrow$ 5, Saturday $\rightarrow$ 6, Sunday $\rightarrow$ 7

power forecasts [62]. This is because human (consumer) power consumption activities are normally periodic.

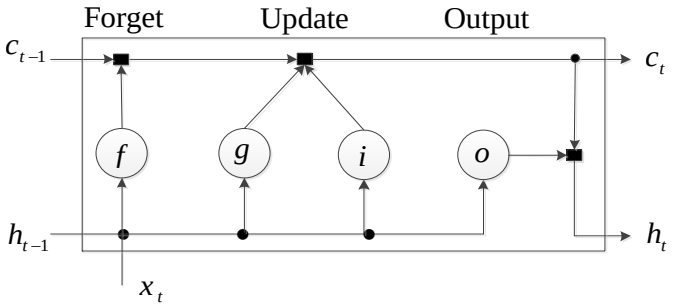

Figure 4-1. Flow of data in a LSTM cell.

Reported work on short-term forecasting has also considered dependence of power forecast on weather parameters [57]. Our work has considered heat index formodeling forecast dependence on weather. Whereas most of reported work has used dew point temperature, and dry bulb temperature for modeling forecast dependence on weather. Table 3 highlights features that have been employed in our work to generate 24 hours in advance power import and power export values for PESCO. Result section demonstrates forecasting performance (comparison) for different sets of input (features) described in Table 3.

4.3.2 Forecasting model and learning algorithm

Our work has employed LSTM network to model unknown transformation function f in Equation (4-2.2). LSTM network has been shown to learn long-term dependencies between samples of a time series [74]. Figure 4-1 highlights data flow in a LSTM cell at time instance t. In Figure 4-1, h_t and c_t are output state and cell state respectively for input sample x_t at time instance t. Whereas h_{t-1} and c_{t-1} are output state and cell state respectively for one previous time instance. In LSTM f, g, i, o represents LSTM cell components forget gate, layer input, input gate, and output gate respectively, with following equations at time instance t:

$$
\begin{aligned}
f_t &= \sigma\left(W_f x_t + R_f h_{t-1} + b_f\right) \\
g_t &= tanh\left(W_g x_t + R_g h_{t-1} + b_g\right) \\
i_t &= \sigma(W_i x_t + R_i h_{t-1} + b_i) \\
o_t &= \sigma(W_o x_t + R_o h_{t-1} + b_o)
\end{aligned}
\tag{4-3.1}
$$

Where, $\sigma(x) = (1 + e^{-x})^{-1}$.

Modeling of function f by LSTM network requires weights computation of LSTM network. The weight matrices of single LSTM cell from (4-3.1) can be written as:

$$W = \begin{bmatrix} W_i \\ W_f \\ W_g \\ W_o \end{bmatrix}, R = \begin{bmatrix} R_i \\ R_f \\ R_g \\ R_o \end{bmatrix}, b = \begin{bmatrix} b_i \\ b_f \\ b_g \\ b_o \end{bmatrix}$$

The weights of complete LSTM network have been determined via supervised learning technique using features discussed in Table 3. Effects of changing LSTM network architecture on 24 hours ahead forecast are demonstrated in results section of this chapter.

Further, to determine weights of LSTM network via supervised learning technique, following three learning algorithms have been used:

4.3.2.1 Stochastic gradient descent with momentum (SGDM)

The SGDM algorithm adds a momentum term to weight update equation of stochastic gradient descent, which helps in reducing oscillations along path of steepest descent towards optimum [79].

4.3.2.2 Root mean square propagation (RMSProp)

RMSProp algorithm uses different learning rate for different parameters which are automatically adapted to reduce loss function [80].

4.3.2.3 Adaptive moment estimation (Adam)

Adam algorithm adds a momentum term to parameter update equation of RMSProp [78].

The performance of LSTM network for 24 hours in advance, short-term forecasting with weights learning using above three different learning algorithms is highlighted in results section.

4.3.3 Training and forecasting methods

Section 4-2 discussed dataset obtained from telemetry system installed in PESCO comprising 74544 observations for duration June 1st, 2014 to August 31st, 2018. The dataset also contained 74544 observations of total power import P_I and total power export P_E by PESCO. This work utilizes 70% of 74544 observations as training set, i.e. data for duration 01.06.2014 to 22.05.2017 is used to learn weights of LSTM network. Whereas, data for

duration 23.05.2017 to 31.08.2018 is test dataset employed to evaluate forecasting performance of LSTM model for f in Equation (4-2.2).

The unknown weights of LSTM model for f in Equation (4-2.2) have been learnt via three learning algorithms SGDM, RMSProp, and Adam while using training dataset. The training data is given by $\{(\boldsymbol{x}_p, \boldsymbol{y}_p); p = 1,2, \dots, P\}$ with $P = 52181$ training instances. And test data given by $\{(\tilde{\boldsymbol{x}}_q, \tilde{\boldsymbol{y}}_q); q = 1,2, \dots, Q\}$ with $Q = 22363$ test instances. $(\boldsymbol{x}_p, \boldsymbol{y}_p)$ represents p^{th} instance input and desired output of training data. $(\boldsymbol{x}_q, \boldsymbol{y}_q)$ represents q^{th} instance test input and test output of test data.In result section effects of changing training options for learning algorithms for test data are discussed.

Once training is complete, the trained LSTM network generates 24 hours in advance forecasts of import power $\hat{P}_I$ and export power $\hat{P}_E$ for PESCO. The LSTM network 24 hours forecasts are evaluated for test cases of total 466 days of test data duration 23.05.2017 to 31.08.2018.

Where for 24 hours duration,

$$\begin{aligned} P_I &= \{p_{I,t+1}, p_{I,t+2}, \dots, p_{I,t+48}\} \\ P_E &= \{p_{E,t+1}, p_{E,t+2}, \dots, p_{E,t+48}\} \\ \hat{P}_I &= \{\hat{p}_{I,t+1}, \hat{p}_{I,t+2}, \dots, \hat{p}_{I,t+48}\} \\ \hat{P}_E &= \{\hat{p}_{E,t+1}, \hat{p}_{E,t+2}, \dots, \hat{p}_{E,t+48}\} \end{aligned} \quad (4\text{-}3.2)$$

After model training two different methods have been used to generate 24 hours in advance forecast of $\hat{P}_I$ and $\hat{P}_E$ Equation (4-3.2). For 24 hours ahead forecast at a resolution of 30 minutes requires forecasting horizon, $H = 48$ Equation (4-3.2).

4.3.3.1 First method

In first method with data features of 24 hours ago from P_I and P_E time series, forecast is performed as, from Equation (4-2.2):

for $i = 1$ to 48

$$(\hat{p}_{I,t+i}, \hat{p}_{E,t+i}) = f\{I_3, I_4 \dots\}$$

end for

Also in first method with data features of 168 hours ago from P_I and P_E time series, forecast is performed as, from Equation (4-2.2):

for $i = 1$ to 48

$\left(\hat{p}_{I,t+i}, \hat{p}_{E,t+i}\right) = f\{I_5, I_6 \dots\}$

end for

4.3.3.2 Second method

In second method with data features of 30 minutes ago from P_I and P_E time series, forecast is performed as, from Equation (4-2.2):

for $i = 1$ to 48

if $i = 1$ then

$\left(\hat{p}_{I,t+i}, \hat{p}_{E,t+i}\right) = f\{I_1, I_2 \dots\}$

else

$\left(\hat{p}_{I,t+i}, \hat{p}_{E,t+i}\right) = f\left\{\hat{p}_{I,t+(i-1)}, \hat{p}_{E,t+(i-1)} \dots\right\}$

end

end for

In both forecasting method exogenous features are also used along data features from recorded time series. In next section results are demonstrated for both forecasting methods with different sets of feature as described in Table 3.

4.3.4 Performance metric

Short-term bivariate forecasting performance of LSTM network for 466 days, $D = 466$, of test dataset with different sets of feature, different training options, and different learning algorithms have been evaluated using mean absolute error (MAE) and mean absolute percentage error (MAPE).

The MAE and MAPE for forecasted import power are given as:

$$(MAE)_I = \frac{1}{D}\sum_{c=1}^{D}\left(\frac{1}{48}\sum_{c=1}^{48}\left|p_{I,c} - \hat{p}_{I,c}\right|\right)_c$$

$$(MAPE)_I = \frac{1}{D}\sum_{c=1}^{D}\left(\frac{1}{48}\sum_{c=1}^{48}\left|\frac{p_{I,c} - \hat{p}_{I,c}}{p_{I,c}}\right| \times 100\right)_c$$

Similarly$(MAE)_E$ and $(MAPE)_E$ are calculated for forecasted export power.

4.4 Results

In this section quantitative and qualitative results of 24 hours in advance bivariate forecast of PESCO import power and export power for 466 days is represented. Table 4 presents quantitative results for 14 test cases with different settings of training algorithms, LSTM network, training options, and input features set. In all test cases of

TABLE 4. RESULTS SUMMARY FOR LSTM FORECASTING MODEL

Serial #	LSTM network	Training Options	Training Algorithm	Features Set	Forecasting Error (Import power)	Forecasting Error (Export power)
1	Sequence input layer with 2 inputs–LSTM layer with 200 units–Completely joined layer with 2 outputs–Regression layer	Training iteration = 130 Initial learn rate = 0.005 Method for rate learning = piecewise Period for reducing learn rate = 80 Factor for reducing learn rate = 0.2 Shuffle =every iteration Gradient threshold =1	Adam	(I_3, I_4)	MAE =88.01 (MW) MAPE =9.69 %	MAE =11.20 (MW) MAPE =13.70 %
2	Sequence input layer with 2 inputs–LSTM layer with 200 units– Completely joined layer with 2 outputs–Regression layer	Training iteration = 130 Initial learn rate = 0.005 Method for rate learning = piecewise Period for reducing learn rate = 80 Factor for reducing learn rate = 0.2 Shuffle =every iteration Gradient threshold =1	SGDM	(I_3, I_4)	MAE =92.48 (MW) MAPE =10.19 %	MAE =11.80 (MW) MAPE =14.29 %
3	Sequence input layer with 2 inputs–LSTM layer with 200 units– Completely joined layer with 2 outputs–Regression layer	Training iteration = 130 Initial learn rate = 0.005 Method for rate learning = piecewise Period for reducing learn rate = 80 Factor for reducing learn rate = 0.2 Shuffle =every iteration Gradient threshold =1	RMSProp	(I_3, I_4)	MAE =87.99 (MW) MAPE =9.80 %	MAE =11.39 (MW) MAPE =14.40 %
4	Sequence input layer with 2 inputs–LSTM layer with 200 units– Completely joined layer with 2 outputs–Regression layer	Training iteration = 130 Initial learn rate = 0.005 Method for rate learning = piecewise Period for reducing learn rate = 80	Adam	(I_5, I_6)	MAE =113.04 (MW) MAPE =12.00 %	MAE =13.67 (MW) MAPE =19.33 %

		Factor for reducing learn rate = 0.2 Shuffle =every iteration Gradient threshold =1				
5	Sequence input layer with 2 inputs–LSTM layer with 200 units– Completely joined layer with 2 outputs–Regression layer	Training iteration = 130 Initial learn rate = 0.005 Method for rate learning = piecewise Period for reducing learn rate = 80 Factor for reducing learn rate = 0.2 Shuffle =never Gradient threshold =1	Adam	(I_3, I_4)	MAE =86.55 (MW) MAPE =9.47 %	MAE =11.08 (MW) MAPE =12.37 %
6	Sequence input layer with 2 inputs–LSTM layer with 300 units– Completely joined layer with 2 outputs–Regression layer	Training iteration = 130 Initial learn rate = 0.005 Method for rate learning = piecewise Period for reducing learn rate = 80 Factor for reducing learn rate = 0.2 Shuffle =never Gradient threshold =1	Adam	(I_3, I_4)	MAE =89.28 (MW) MAPE =10.00 %	MAE =12.22 (MW) MAPE =18.29 %
7	Sequence input layer with 2 inputs–LSTM layer with 200 units– Completely joined layer with 2 outputs–Regression layer	Training iteration = 130 Initial learn rate = 0.005 Method for rate learning = piecewise Period for reducing learn rate = 80 Factor for reducing learn rate = 0.2 Shuffle =never Gradient threshold =1	Adam	(I_1, I_2)	MAE =122.82 (MW) MAPE =13.17 % Forecasting Method = 2nd	MAE =13.96 (MW) MAPE =17.33 % Forecasting Method = 2nd
8	Sequence input layer with 3 inputs–LSTM layer with 200 units– Completely joined layer with 2 outputs–Regression layer	Training iteration = 130 Initial learn rate = 0.005 Method for rate learning = piecewise Period for reducing learn rate = 80 Factor for reducing learn rate = 0.2 Shuffle =never Gradient threshold =1	Adam	(I_3, I_4, I_7)	MAE =89.89 (MW) MAPE =10.05 %	MAE =11.36 (MW) MAPE =14.79 %

Serial #	LSTM network	Training Options	Training Algorithm	Features Set	Forecasting Error (Import power)	Forecasting Error (Export power)
9	Sequence input layer with 3 inputs–LSTM layer with 300 units– Completely joined layer with 2 outputs–Regression layer	Training iteration = 130 Initial learn rate = 0.005 Method for rate learning = piecewise Period for reducing learn rate = 80 Factor for reducing learn rate = 0.2 Shuffle =never Gradient threshold =1	Adam	(I_3, I_4, I_7)	MAE =91.26 (MW) MAPE =10.13 %	MAE =11.61 (MW) MAPE =15.42 %
10	Sequence input layer with 3 inputs–LSTM 1st layer with 200 units– LSTM 2nd layer with 100 units– Completely joined layer with 2 outputs–Regression layer	Training iteration = 130 Initial learn rate = 0.005 Method for rate learning = piecewise Period for reducing learn rate = 80 Factor for reducing learn rate = 0.2 Shuffle =never Gradient threshold =1	Adam	(I_3, I_4, I_7)	MAE =96.91 (MW) MAPE =11.14 %	MAE =11.54 (MW) MAPE =14.72 %
11	Sequence input layer with 4 inputs–LSTM layer with 200 units– Completely joined layer with 2 outputs–Regression layer	Training iteration = 130 Initial learn rate = 0.005 Method for rate learning = piecewise Period for reducing learn rate = 80 Factor for reducing learn rate = 0.2 Shuffle =never Gradient threshold =1	Adam	(I_3, I_4, I_8, I_9)	MAE =87.03 (MW) MAPE =9.66 %	MAE =11.02 (MW) MAPE =13.65 %
12	Sequence input layer with 5 inputs–LSTM layer with 200 units– Completely joined layer with 2 outputs–Regression layer	Training iteration = 130 Initial learn rate = 0.005 Method for rate learning = piecewise Period for reducing learn rate = 80 Factor for reducing learn rate = 0.2 Shuffle =never Gradient threshold =1	Adam	$(I_3, I_4, I_7, I_8, I_9)$	MAE =91.19 (MW) MAPE =10.12 %	MAE =11.33 (MW) MAPE =14.13 %
13	Sequence input layer with 5 inputs–LSTM layer with 400 units– Completely joined layer with 2 outputs–Regression layer	Training iteration = 130 Initial learn rate = 0.005 Method for rate learning = piecewise Period for reducing learn rate = 80 Factor for reducing learn rate = 0.2	Adam	$(I_3, I_4, I_7, I_8, I_9)$	MAE =95.53 (MW) MAPE =10.90 %	MAE =11.95 (MW) MAPE =15.50 %

		Shuffle =never Gradient threshold =1				
14	Sequence input layer with 2 inputs–LSTM layer with 200 units– Completely joined layer with 200 neurons–Dropout layer with 0.5 factor–Fully connected layer with 2 outputs–Regression layer	Training iteration = 130 Initial learn rate = 0.005 Method for rate learning = piecewise Period for reducing learn rate = 80 Factor for reducing learn rate = 0.2 Shuffle =never Gradient threshold =1	Adam	(I_3, I_4)	MAE =86.41 (MW) MAPE =9.52 %	MAE =11.18 (MW) MAPE =13.30 %

Table 4 first forecasting method is used except for test case at serial number 7. From Table 4 it is evident that best forecasting performance (MAPE = 9.47 % and MAPE = 12.37 % for import power and export power respectively) is achieved for settings shown at serial number 5. Figure 4-2 to Figure 4-4 presents qualitative results of LSTM forecasting performance for 7 test cases. Actual and forecasted import power and export power curves for test duration (466 days) are plotted. Error curve between the forecasted and actual values is also plotted. On x-axis 'Q' represents a single quarter of year with 3 months duration. Although visually the forecasted power import and forecasted power export value curves appear to be similar in Figure 4-2 to Figure 4-4, but the quantitative performance measure for forecasting, i.e. MAPE is different for all the plotted 7 test cases. From Table 4, the MAPE values of forecasted power import plots for Figure 4-2 to Figure 4-4 test case number 1, 2, 3, 4, 5, 7, 13 are 9.69%, 10.19%, 9.80%, 12.00%, 9.47%, 13.17%, 10.90% respectively. Similarly the MAPE values of forecasted power export for 7 test cases of Figure 4-2 to Figure 4-4 can be seen from Table 4.

4.5 Conclusion

The work presented in this chapter and previous chapters is an effort towards developing a detailed framework for forecasting short-term and long-term power requirements of electricity distribution companies of Pakistan. With the installation of systems (as discussed in earlier section) real-time monitoring and logging of power data for distribution companies is already being carried out. The data from these systems is used to learn model for power forecasting. The forecasted power values are useful for both

National Transmission & Dispatch Company, and for the electricity distribution companies of Pakistan.

From next chapter onwards, state and parameter estimation algorithms application in output feedback control of magnetic levitation and motor systems are discussed.

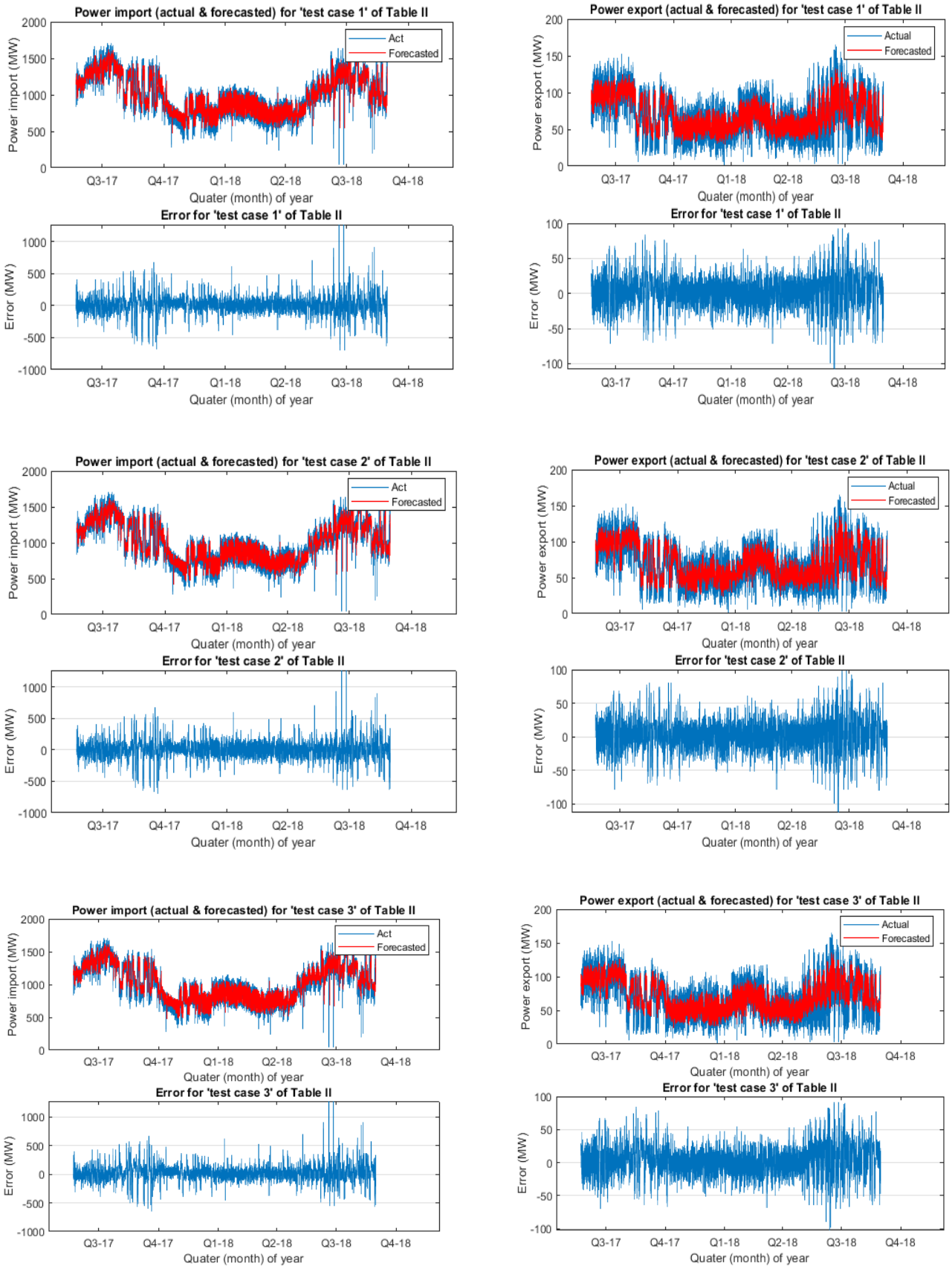

Figure 4-2. Graphical representations of forecasted & actual import power, and forecasted & actual export power for three test cases.

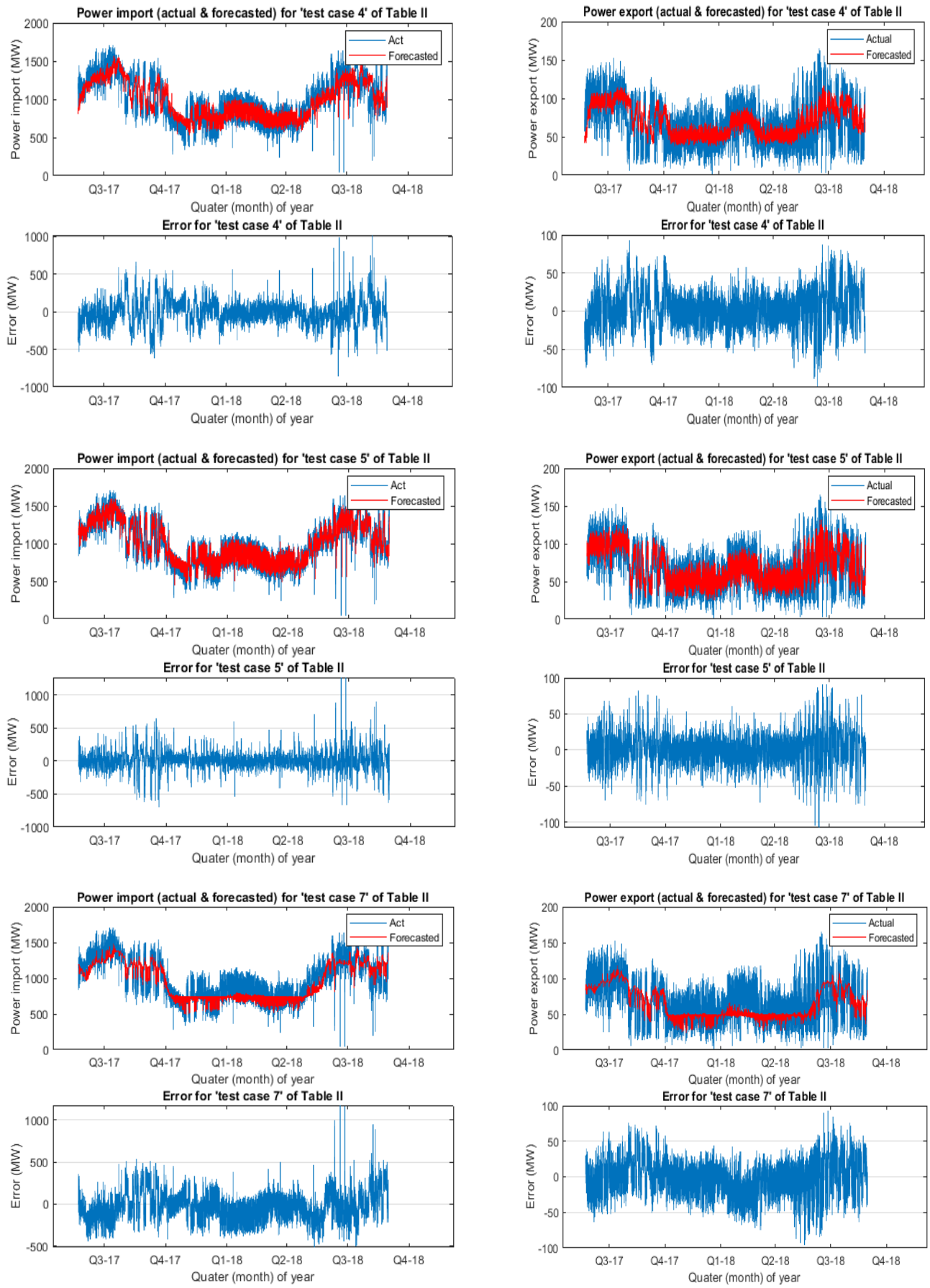

Figure 4-3. Graphical representations of forecasted & actual import power, and forecasted & actual export power for three test cases.

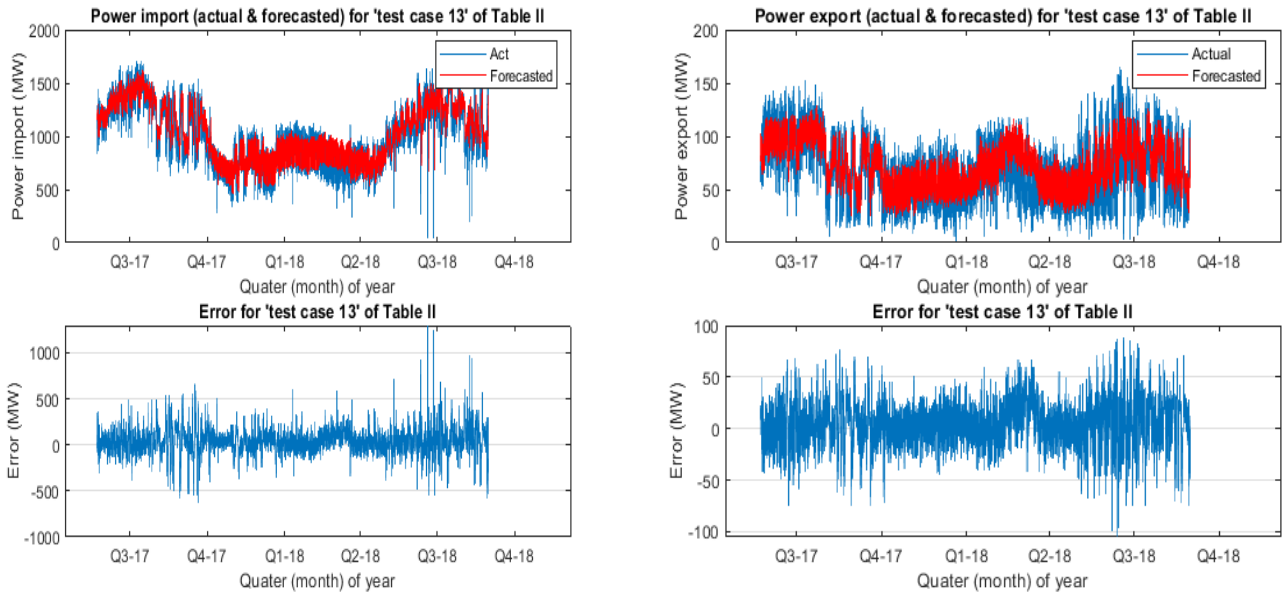

Figure 4-4. Graphical representations of forecasted & actual import power, and forecasted & actual export power for one test case.

Supporting publication: Asim Zaheer Ud Din, Yasar Ayaz, Momena Hasan, Jawad Khan, and Muhammad Salman, "Bivariate short-term electric power forecasting using LSTM network", 3rd International Conference on Robotics and Automation in Industry, 2019, Rawalpindi, Pakistan.

CHAPTER 5
ESTIMATION USING HIGHER ORDER SLIDING MODE OBSERVER

The previous Chapter 2 to Chapter 4 have presented implementation results of proposed detection and forecast algorithms for nonlinear processes in context to smart grid applications. Hereforth, Chapter 5 to Chapter 8, results of estimation algorithms utilized in proposed sampled-data output feedback controller for nonlinear magnetic levitation system in context to smart grid are presented.

As also discussed in Chapter 1, magnetic levitation system are used in active magnetic bearings of electromechanical generators for advanced and efficient electricity generation in smart grid solution. These magnetic bearings use magnetic levitation to support load, and are able to hold moving parts without physical contact. Also they provide high relative motion with very low friction, no mechanical wear, and no lubrication. Chapter 5 to Chapter 8 presents design and simulated results for estimation algorithms, and output feedback control configurations for magnetic levitation system. The idea is to further implement these configurations for development of active magnetic bearings in later stages of efficient electricity generation project.

The problem statement for work discussed in Chapter 5 to Chapter 8 is summarized as:

- To design and simulate sampled-data output feedback control (OFC) for nonlinear magnetic levitation system for following cases:
    i. Case-I: Unknown plant model – discussed in Chapter 5
    ii. Case-II: Known plant model – discussed in Chapter 6
    iii. Case-III: Model mismatch – discussed in Chapter 7
    iv. Case-IV: Unknown plant parameters – discussed in Chapter 8
- To compare state estimation performance of Kalman, SSRLS & sliding mode state estimators.

In this chapter an estimation based output feedback control configuration is presented to achieve robust feedback linearization for nonlinear minimum phase systems.

An Euler approximate discrete-time observer based on Sliding Mode observer is presented. The observer while operating in discrete-time domain simultaneously estimates states and drift term using system output. The estimated drift term is used in feedback loop to compensate unmodeled system dynamics and disturbances. Emulation Design procedure is employed in designing of discrete controller. The performance of presented sampled-data robust feedback linearization scheme is demonstrated for tracking applications of magnetic levitation and DC motor systems in computer simulations. Magnetic bearings using magnetic leviatation, and motor systems are important components for electricity generation industry. Simulation results illustrate that reducing sampling period more adversely affects Euler approximate discrete observer performance for faster changing system dynamics than for slower changing dynamics. The proposed scheme also exhibits good performance in presence of disturbances and parameters perturbation.

5.1 Introduction

Often in control applications it is required to achieve specified closed-loop performance and robustness in presence of uncertainties, using only available system output. One approach is to reconstruct state vector on basis of nominal system model using observer; and then using this estimated state information in a robust or adaptive control scheme to achieve required performance.

Another approach is to estimate drift term which constitutes combined effects of unknown parameters, model uncertainties and disturbances. The idea of estimating drift term and then canceling it via feedback control law has also been discussed in works of [81, 82]. In [82] the authors have considered relative degree one systems, whereas in [83] a high gain observer has been employed to estimate states and uncertainties in the system. In this work sampled-data robust feedback linearization scheme has been proposed based on the robust feedback linearization work presented in [84]. Robust feedback linearization has also been demonstrated in the works of [83, 85]. In [83] an extended high gain observer is employed, whereas higher order SMO is used in [85]. In both works state and drift term estimation is based on nominal plant model. In the work of [84], the authors have employed a higher order SMO to estimate states and drift term for relative degree n minimum phase systems. Euler approximate discrete model of SMO [84] along with discrete controller

(based on Emulation Design) has been used in this work. Detailed work on Euler approximate discrete model is present in [86, 87], whereas Emulation Design procedure for discrete controller designing is discussed in [88].

Sampled-data control configuration has been used in this work as most plants and processes in engineering practice are continuous-time. Controlling a continuous-time plant / process using a computer or a digital controller, forms a sampled-data control system. Therefore, sampled-data control system also known as computer-controlled system involves both continuous-time and discrete-time signals in its operation.

Furthermore, continuous-time nonlinear plant or process controlled by either a nonlinear or linear digital controller is categorized as a nonlinear sampled-data control configuration (system).A nonlinear sampled-data control configuration based on [84] is discussed later in this chapter.

Generally three main techniques are used to design controller for a sampled-data system [89]. First technique is known as Emulation Design. In this technique, the controller is designed in continuous-time domain. Continuous-time controller is then discretized, and digitally implemented. Second technique is called Direct Discrete-time Design. In this technique, a discrete-time controller is designed in discrete-time domain using discrete-time model of plant or process [88]. Inter-sample behavior is not considered during design in this technique. Third technique is Sampled-data Design, which considers inter-sample behavior during design process.

5.2 Problem Formulation

Consider a single-input-single-output nonlinear system in *Generalized Controller Canonical Form* (GCCF) [90]:

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= x_3 \\ &\vdots \\ \dot{x}_r &= f(x,\xi) + g(x)u \\ \dot{\xi} &= \varphi(x,\xi) \\ y &= x_1 \end{aligned} \tag{5-2.1}$$

with states *n* and *relative degree*ρ, such that $\rho \leq n$. $x \in \mathcal{R}^{\rho}$ is observable state sub-vector, $\xi \in \mathcal{R}^{n-\rho}$ is unobservable state sub-vector and $u \in \mathcal{R}$ is control input. $f(\cdot)$ and $g(\cdot)$ are smooth vector fields.

***Assumption*:** System Equation (5-2.1) is *minimum phase system*, as *zero dynamics* $\dot{\xi} = \varphi(0, \xi)$ of system is input-to-state stable [91]. System Equation (5-2.1) can be written in the form:

$$\begin{aligned} \dot{x} &= Ax + B[f(x,\xi) + g(x)u] \\ \dot{\xi} &= \varphi(x,\xi) \\ y &= Cx \end{aligned} \tag{5-2.2}$$

where

$$A = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ 0 & 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, B = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}, C = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}^{T}$$

Let $\hat{x}$ is estimate of state vector *x*, and $g(\hat{x})$ is input gain on basis of $\hat{x}$. System Equation (5-2.2) can be written as:

$$\begin{aligned} \dot{x} &= Ax + B[f(x,\xi,\Delta_u) + g(\hat{x})u] \\ \dot{\xi} &= \varphi(x,\xi) \\ y &= Cx \end{aligned} \tag{5-2.3}$$

where $\Delta_u = g(x) - g(\hat{x})$ is uncertainty of input channel. If in Equation (5-2.3), $f(\cdot)$ is exactly known and $g(\hat{x}) \neq 0$, then feedback linearization control for Equation (5-2.3) is [90]:

$$u = \frac{1}{g(\hat{x})}(-f(x,\xi,\Delta_u) - Kx) \tag{5-2.4}$$

K is gain matrix, which can be designed using any modern or classical technique e.g. LQR, pole placement such that $A - BK$ is Hurwitz and closed-loop system meets required performance objectives. Using control law Equation (5-2.4) in system Equation (5-2.3), yields linear closed-loop dynamics:

$$\dot{x} = (A - BK)x \tag{5-2.5}$$

However in most practical cases function $\boldsymbol{f}(\cdot)$ is not exactly known and only output of system is available. Thus the control law Equation (5-2.4) is not realizable, and control law:

$$\boldsymbol{u} = \frac{\mathbf{1}}{\boldsymbol{g}(\hat{\boldsymbol{x}})}\left(-\hat{\boldsymbol{f}} - \boldsymbol{K}\hat{\boldsymbol{x}}\right) \tag{5-2.6}$$

is used. In Equation (5-2..6), if drift term $\hat{\boldsymbol{f}}$ and states $\hat{\boldsymbol{x}}$ are correctly estimated such that $\hat{\boldsymbol{f}} \to \boldsymbol{f}(\cdot)$ and $\hat{\boldsymbol{x}} \to \boldsymbol{x}$ in finite time, the desired performance of Equation (5-2.5), and effect of control law Equation (5-2.4) can be achieved.

5.3 Robust Observer & Controller

In this section a robust observer [84] is discussed, which estimates the drift term $\hat{f}$ and states $\hat{x}$. Methodology is similar in both [84] and [83], except a High Gain Observer is proposed in [83]. With $\hat{f} \equiv \hat{f}_1$, the observer structure [84] follows:

$$\begin{gathered}
\dot{\hat{x}}_1 = -\lambda_{n+1}L^{1/(n+1)}|\hat{x}_1 - x_1|^{n/(n+1)}sign(\hat{x}_1 - x_1) + \hat{x}_2 \\
\dot{\hat{x}}_2 = -\lambda_n L^{1/n}|\hat{x}_2 - \hat{x}_1|^{(n-1)/n}sign(\hat{x}_2 - \hat{x}_1) + \hat{x}_3 \\
\vdots \\
\dot{\hat{x}}_r = \hat{f}_1 + g(\hat{\boldsymbol{x}})u \\
\hat{f}_1 = -\lambda_{n-\rho+2}L^{1/(n-\rho+2)}\left|\hat{x}_\rho - \hat{x}_{\rho-1}\right|^{(n-\rho+1)/(n-\rho+2)} \\
sign\left(\hat{x}_\rho - \hat{x}_{\rho-1}\right) + \hat{x}_{\rho+1} \\
\vdots \\
\dot{\hat{x}}_n = \hat{f}_{n-\rho+1} \\
\hat{f}_{n-\rho+1} = -\lambda_2 L^{1/2}\left|\hat{x}_n - \hat{f}_{n-\rho}\right|^{1/2}sign\left(\hat{x}_n - \hat{f}_{n-\rho}\right) + \hat{x}_{n+1} \\
\dot{\hat{x}}_{n+1} = -\lambda_1 L sign\left(\hat{x}_{n+1} - \hat{f}_{n-\rho+1}\right)
\end{gathered} \tag{5-3.1}$$

$\boldsymbol{\lambda_i}$ can be chosen as suggested in [92], and $\boldsymbol{L} > \mathbf{0}$ is *Lipschitz constant*.

5.3.1 Euler approximate discrete-time robust observer

In this section Euler approximate discrete-time model of robust observer Equation (5-3.1) is presented. Euler approximation is used, as it is the simplest approximation that preserves the structure of continuous-time robust observer Equation (5-3.1) [86]. Discrete-time observer model follows:

$$\hat{x}_{1,k} = \hat{x}_{1,k-1} + T\{-\lambda_{n+1}L^{1/(n+1)}\left|\hat{x}_{1,k-1} - x_{1,k-1}\right|^{n/(n+1)}$$
$$sign\left(\hat{x}_{1,k-1} - x_{1,k-1}\right]) + \hat{x}_{2,k-1}\}$$
$$\hat{x}_{2,k} = \hat{x}_{2,k-1} + T\{-\lambda_{n}L^{1/n}\left|\hat{x}_{2,k-1} - \hat{x}_{1,k-1}\right|^{(n-1)/n}$$
$$sign\left(\hat{x}_{2,k-1} - \hat{x}_{1,k-1}\right) + \hat{x}_{3,k-1}\}$$
$$\vdots$$
$$\hat{x}_{\rho,k} = \hat{x}_{\rho,k-1} + T\{\hat{f}_{1,k-1} + g(\hat{\boldsymbol{x}}_{k-1})u\}$$
$$\hat{f}_{1,k-1} = -\lambda_{n-\rho+2}L^{1/(n-\rho+2)}\left|\hat{x}_{\rho,k-1} - \hat{x}_{\rho-1,k-1}\right|^{(n-\rho+1)/(n-\rho+2)}$$
$$sign\left(\hat{x}_{\rho,k-1} - \hat{x}_{\rho-1,k-1}\right) + \hat{x}_{\rho+1,k-1}$$
$$\vdots$$
$$\hat{x}_{n,k} = \hat{x}_{n,k-1} + T\{\hat{f}_{n-\rho+1,k-1}\}$$
$$\hat{f}_{n-\rho+1,k-1} = -\lambda_{2}L^{1/2}\left|\hat{x}_{n,k-1} - \hat{f}_{n-\rho,k-1}\right|^{1/2} sign\left(\hat{x}_{n,k-1} - \hat{f}_{n-\rho,k-1}\right)$$
$$+\hat{x}_{n+1,k-1}$$
$$\hat{x}_{n+1,k} = \hat{x}_{n+1,k-1} + T\{-\lambda_{1}Lsign\left(\hat{x}_{n+1,k-1} - \hat{f}_{n-\rho+1,k-1}\right)\} \qquad (5\text{-}3.2)$$

$\boldsymbol{T}$ is sampling period. Value of $\boldsymbol{T}$ is close to zero. The discrete-time observer Equation (5-3.2) is used in configuration as shown in Figure 5-1 to achieve sampled-data robust feedback linearization for minimum phase systems. Plant operates in continuous-time, where $\boldsymbol{w}(\boldsymbol{t})$ represents process noise. Measurement noise $\boldsymbol{v}(\boldsymbol{t})$ is added to output $\boldsymbol{y}(\boldsymbol{t})$ of plant. Continuous-time plant output $\boldsymbol{y}(\boldsymbol{t})$ is discretized (Sampling) using zero-order hold approach [87]. The observer Equation (5-3.2) uses measured state (output $\boldsymbol{y}[\boldsymbol{k}]$) and input $\boldsymbol{u}[\boldsymbol{k}]$ to estimate the drift term $\hat{\boldsymbol{f}}$ and state $\hat{\boldsymbol{x}}$ in Equation (5-2.6). Measured state (output $\boldsymbol{y}[\boldsymbol{k}]$) and estimated state vector $\hat{\boldsymbol{x}}$ are used to generate $\boldsymbol{u}[\boldsymbol{k}]$. Controller gain matrix $\boldsymbol{K}$ is designed using any modern or classical technique e.g. LQR, pole placement such that $\boldsymbol{A} - \boldsymbol{BK}$ is Hurwitz. Control input $\boldsymbol{u}[\boldsymbol{t}]$ is obtained using zero-order hold approach (Hold) [87], and is applied to the plant.

5.3.2 Discrete-time robust controller

In this section discrete-time robust feedback linearization controller based on Emulation Design procedure is presented. From Equation (5-2.6), discrete robust controller follows:

$$u[k] = \frac{1}{g(\hat{x}[k])}\left(-\hat{f}[k] - K\hat{x}[k]\right) \tag{5-3.3}$$

The controller Equation (5-3.3) along with observer Equation (5-3.2) is used in configuration shown in Figure 5-1 to achieve robust feedback linearization for sampled-data nonlinear minimum phase system.

5.4 Computer Simulations: magnetic levitation system

This section presents computer simulation results for the proposed Euler approximate discrete-time robust observer Equation (5-3.2) and controller Equation (5-3.3) for magnetic levitation system. Observer Equation (5-3.2) estimates the drift term and states for given magnetic levitation system, and feedback linearization control law Equation (5-3.3) achieves tracking of reference signal. The reference signal is a bidirectional chirp, and levitated magnet disc of Maglev system tracks the reference chirp signal.

5.4.1 Maglev model

Magnetic levitation system with following model equation [93] and parameters (listed in Table 5) have been used in this work:

$$m\ddot{y}_1 + c_1\dot{y}_1 = F_{u11} - mg \tag{5-4.1}$$

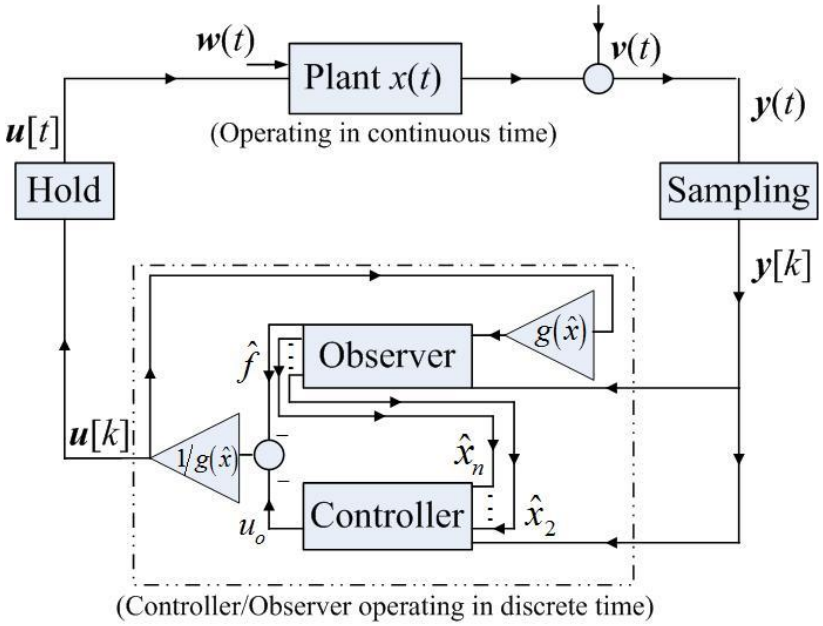

Figure 5-1.Sampled-data robust feedback linearization

where $\boldsymbol{y_1}$ is position of levitated magnet disc, and $\boldsymbol{F_{u11}}$ is repulsive magnetic force [93] acting on disc due to current coil:

$$\boldsymbol{F_{u11}} = \frac{\boldsymbol{u}}{\boldsymbol{1,65(y_1 + 6.2)^4}}$$

u is input voltage applied to coil. State space representation of Equation (5-4.1) is:

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

$$\dot{x}_2 = \frac{u}{0.198(x_1 + 6.2)^4} - 9.8 - 9.7275x_2$$

$$\boldsymbol{y = x_1} \tag{5-4.2}$$

$\boldsymbol{x_1}$ and $\boldsymbol{x_2}$ (position and velocity) are two states of Maglev system ($\boldsymbol{n = 2}$). $\boldsymbol{y = x_1}$ is the output of system. For given output and Maglev system model Equation (5-4.2), *relative degree* $\boldsymbol{\rho = 2}$. As $\boldsymbol{n = \rho}$, no unobservable state exists for Equation (5-4.2), and hence it has no zero dynamics and is *minimum phasesystem* by default [91]. The system Equation (5-4.2) is transformed into state space error model for tracking control. With

$$e_1 = x_1 - r$$

$$\boldsymbol{e_2 = x_2 - \dot{r}} \text{ (as } \boldsymbol{\dot{x}_1 = x_2}) \tag{5-4.3}$$

error model obtained is:

$$\dot{e}_1 = e_2$$

$$\dot{e}_2 = \frac{u}{0.198(e_1 + r + 6.2)^4} - 9.8 - 9.7275(e_2 + \dot{r}) - \ddot{r}$$

where r is the reference (chirp) signal, $f(e)$ and input gain $g(e)$ is:

$$f(e) = -9.8 - 9.7275(e_2 + \dot{r}) - \ddot{r} \tag{5-4.4}$$

$$g(e_1) = \frac{1}{0.198(e_1 + r + 6.2)^4}$$

TABLE 5. MAGNETIC LEVITATION SYSTEM PARAMETERS

Name	*Symbol*	*Value / Unit*
Mass of magnet	m	0.12 kg
Acceleration due to gravity	g	9.8 ms^{-2}
Coefficient of viscosity	c_1	1.1673

5.4.2 Observer, controller and plant initialization

The discrete-time robust observer Eqaution (5-3.2) uses discrete input $u[k]$ and $e_1[k]$ ($e_1[k] = y[k] - r[k]$) in the configuration shown in Figure 5-1, and estimates drift term $\hat{f}$ and $\hat{e}_2$ such that $\hat{f} \to f(\cdot)$ in Equation (5-4.4), and $\hat{x}_2 \to x_2$ using Equation (5-4.3) in finite time.

From Equation (5-3.2), Euler approximate discrete-time robust observer for given error model of Maglev system is:

$$\hat{e}_{1,k} = \hat{e}_{1,k-1} + T\{-\lambda_3 L^{1/3}\left|\hat{e}_{1,k-1} - e_{1,k-1}\right|^{2/3} sign\left(\hat{e}_{1,k-1} - e_{1,k-1}\right]) + \hat{e}_{2,k-1}\}$$

$$\hat{e}_{2,k} = \hat{e}_{2,k-1} + T\{\hat{f}_{1,k-1} + g\left(e_{1,k-1}\right)u\}$$

$$\hat{f}_{1,k-1} = -\lambda_2 L^{1/2}\left|\hat{e}_{2,k-1} - \hat{e}_{1,k-1}\right|^{1/2} sign\left(\hat{e}_{2,k-1} - \hat{e}_{1,k-1}\right) + \hat{e}_{3,k-1}$$

$$\hat{e}_{3,k} = \hat{e}_{3,k-1} + T\{-\lambda_1 L sign\left(\hat{e}_{3,k-1} - \hat{f}_{1,k-1}\right)\} \quad (5\text{-}4.5)$$

Discrete robust feedback linearization control law implemented for Maglev system based on Emulation Design procedure:

$$u[k] = \frac{1}{g(e_1[k])}\left(-\hat{f}[k] - k_1 e_1[k] - k_2 \hat{e}_2[k]\right)$$

where $\hat{f}[k]$ and $\hat{e}_2[k]$ are estimated by observer and

$$g(e_1[k]) = \frac{1}{0.198(e_1[k] + r[k] + 6.2)^4}$$

For Euler approximate discrete-time robust observer (5-4.5) $\lambda_1 = 10$, $\lambda_2 = 3200$ and $\lambda_3 = 3200$ (designed as suggested in [94]). *Lipschitz constant* $L = 1$ and

$$\hat{e}_{1,k=0} = 0,\ \hat{e}_{2,k=0} = 0,\ \hat{e}_{3,k=0} = 0 \quad (5\text{-}4.6)$$

The performance of discrete robust observer Equation (5-4.5) for sampling period $T = 0.001sec.$ and $T = 0.01sec.$ is studied in computer simulations. Controller gains $k_1 = 1.72$ and $k_2 = 0.065$ are designed using pole placement method such that $A - BK$ is Hurwitz. Matrices A and B are obtained by linearization of Equation (5-4.2) at 0.05 (bias value of reference chirp signal).

Maglev system in Equation (5-4.2) is initially at (position) $x_1 = 0.03$ *meter* and (velocity) $x_2 = 0$. This initial condition of system is different from observer initialization Equation (5-4.6). Output $y = x_1$, i.e. position of levitated magnet disc tracks reference

bidirectional chirp signal of 0.03 *meter* amplitude, 0.05 bias and frequency 37.68 *rad/sec*. Additive white Gaussian noise of variance 10^{-7} and 10^{-5} is added as process noise and measurement noise respectively.

5.4.3 Results discussion

Reference trajectory and position of Maglev levitated magnet disc, both are plotted in Figure 5-2 for a simulation duration of 80 seconds. The tracking simulation is performed with a sampling period $T = 0.001sec.$ for discrete observer and controller.

The levitated magnet disc considerably tracks bidirectional chirp trajectory of 0.03*meter* amplitude, 0.05 bias and 37.68 *rad/sec* frequency. Initial tracking error (value of 10^{-1}), which is due to different initial condition of observer Equation (5-4.5) and Maglev Equation (5-4.2), (and also because of different bias value for reference trajectory,) settles to a constant reduced value as can be seen in Figure 5-2 (simulation duration $72-80$ seconds).

Tracking error plot is shown in Figure 5-3. Error plot is obtained by difference of reference chirp trajectory and position of levitated magnet disc. Tracking error remains within range of 10^{-2} for sampling period $T = 0.001sec.$ as shown in Figure 5-3.

Figure 5-4 demonstrates effect of increasing sampling period T on performance of discrete observer Equation (5-4.5). In general, Euler approximate discrete-time model is a good approximation of exact model typically only for smaller values of T [89]. The tracking error increases and reaches value of 10^{-1} when sampling period is increased from $T = 0.001sec.$ to $T = 0.01sec.$ Also, as output / system dynamics of Maglev system are changing at a higher rate (tracking chirp trajectory of 37.68 *rad/sec* frequency) as compared with DC motor system (discussed in next section), increasing sampling period adversely affects performance of discrete observer Equation (5-4.5).

5.5 Computer Simulations: DC Motor System

As a second example to demonstrate the performance of observer Eqaution (5-3.2) and control law Equation (5-3.3), a DC motor system is considered. The motor shaft speed tracks a sinusoidal reference shaft speed signal (trajectory).

5.5.1 DC motor model

DC motor with following model equation and parameters (listed in Table 6) have been currently used. A nonlinear DC motor model given in appendix A has also been used in another work.

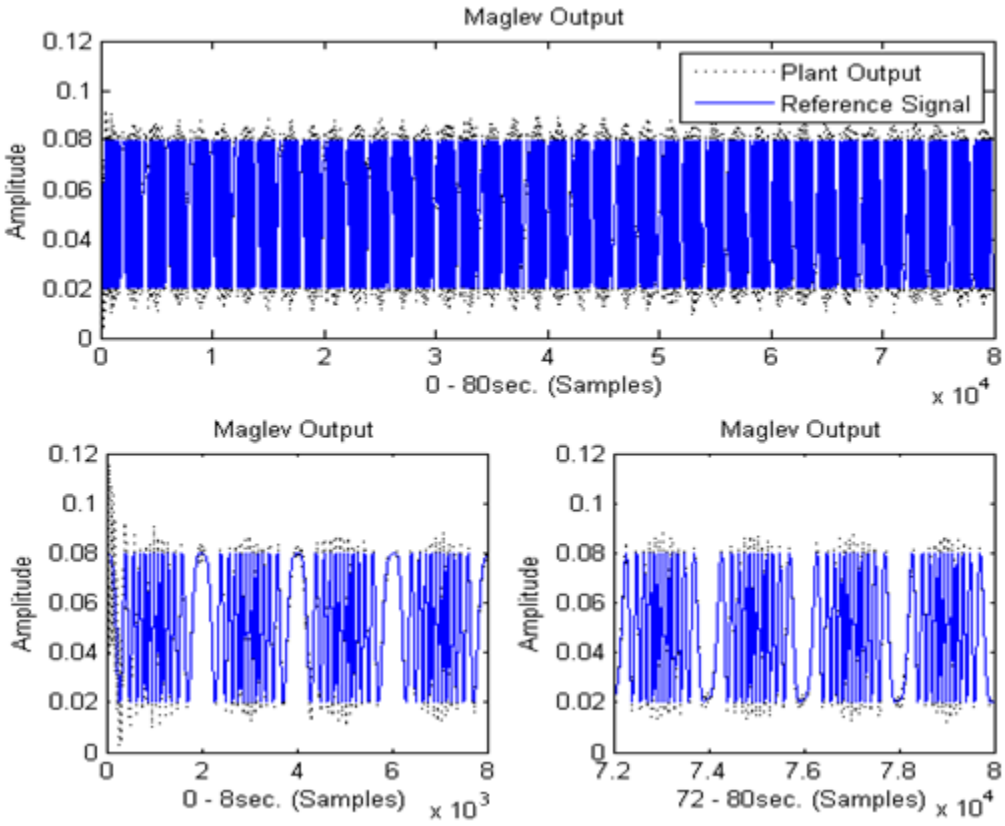

Figure 5-2. Levitated magnet disc of Maglev tracking a reference bidirectional chirp (37.68 *rad*/sec) trajectory. Sampling period T = 0.001sec.

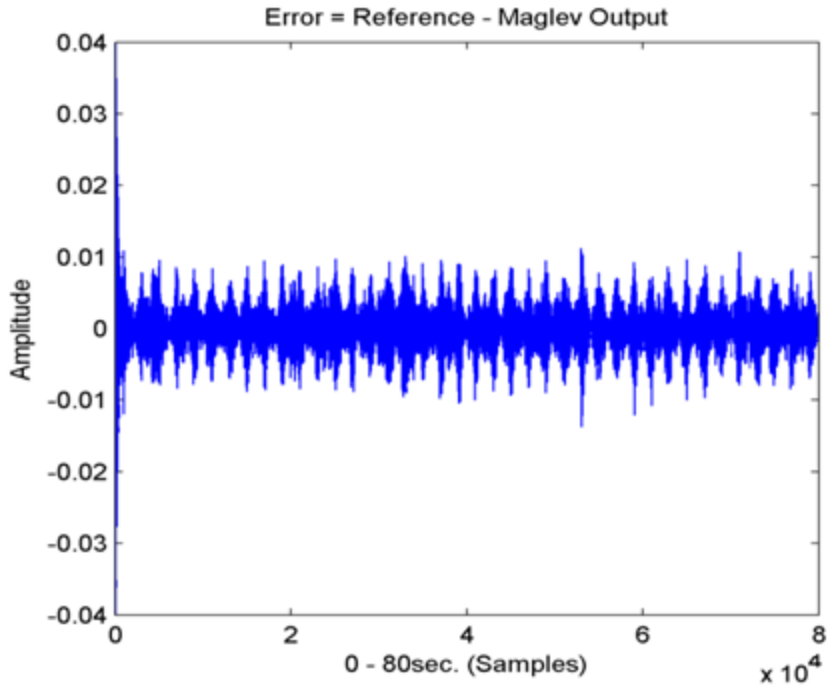

Figure 5-3. Maglev tracking error plot for reference bidirectional chirp (37.68 *rad*/sec) trajectory. Sampling period T = 0.001sec.

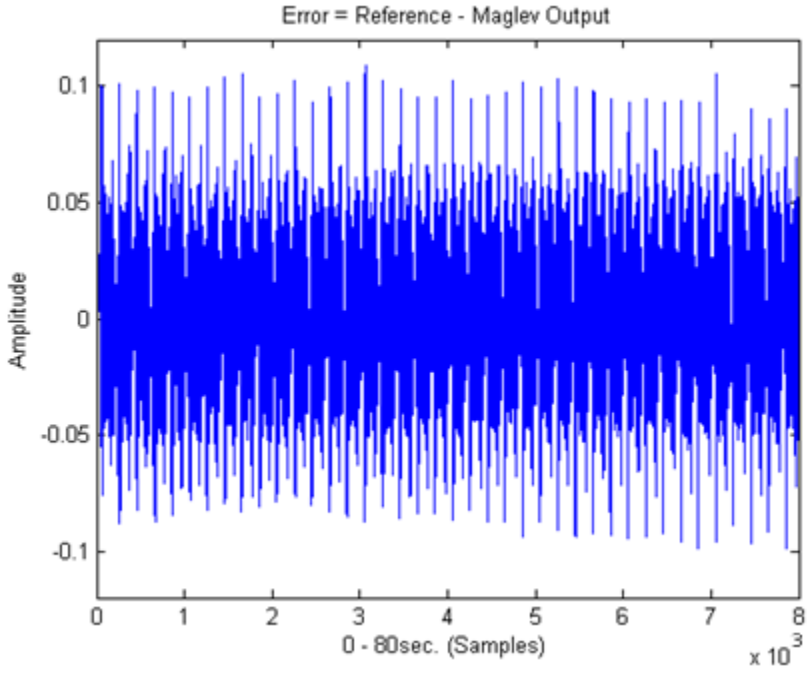

Figure 5-4. Maglev tracking error plot for reference bidirectional chirp (37.68 *rad*/sec) trajectory. Sampling period T = 0.01sec.

$$L_o \frac{di}{dt} = u - Ri - K_e\omega$$

$$J \frac{d\omega}{dt} = K_t i - \tau_l \quad (5\text{-}5.1)$$

where u is input voltage, ω and i are shaft speed and armature current respectively. $\tau_l = B\omega$ is motor load torque. State space representation of Equation (5-5.1) is:

$$\dot{x}_1 = 8x_2 - 10x_1$$

$$\dot{x}_2 = 66.67u - 33.33x_2 - 0.067x_1$$

$$y = x_1 \quad (5\text{-}5.2)$$

x_1 and x_2 (shaft speed and armature current) are two states of DC motor ($n = 2$). $y = x_1$ is the output of system. For given output and DC motor model Equation (5-5.2), *relative degree*$\rho = 2$.As $n = \rho$, no unobservable state exists for Equation (5-5.2), and hence it has no zero dynamics and is *minimum phasesystem* by default [91].

The system Equation (5-5.2) is transformed into GCCF error model to achieve tracking using control law Equation (5-3.3). With

$$e_1 = r - x_1$$

$$e_2 = \dot{r} - 8x_2 + 10x_1 \ \ (\text{as } \dot{x}_1 = 8x_2 - 10x_1) \quad (5\text{-}5.3)$$

GCCF error model obtained is:

$$\dot{e}_1 = e_2$$

$$\dot{e}_2 = -333.836e_1 - 43.33e_2 + 333.836r + 43.33\dot{r} + \ddot{r} - 533.36u$$

where r is the sinusoidal reference shaft speed, $f(e)$ and input gain b_o is:

TABLE 6. DC MOTOR PARAMETERS

Name	*Symbol*	*Value / Unit*
Inertia of motor rotor and load	J	0.001 kgm^2
Armature resistance	R	0.5 Ω
Armature inductance	L_o	15 mH
Back-emf constant	K_e	0.001 $Vrad^{-1}$
Torque constant	K_t	0.008 NmA^{-1}
Coefficient of viscous friction	B	0.01 $Nmsrad^{-1}$

$$f(e) = -333.836e_1 - 43.33e_2 + 333.836r + 43.33\dot{r} + \ddot{r} \quad (5\text{-}5.4)$$

$$b_o = -533.36$$

5.5.2 Observer, controller and plant initialization

The discrete-time robust observer Equation (5-3.2) uses discrete input $u[k]$ and $e_1[k]$ ($e_1[k] = r[k] - y[k]$) in the configuration shown in Figure 5-1, and estimates drift term $\hat{f}$ and $\hat{e}_2$ such that $\hat{f} \rightarrow f(\cdot)$ in (5-5.4), and $\hat{x}_2 \rightarrow x_2$ using Equation (5-5.3) in finite time.

From Equation (5-3.2), Euler approximate discrete-time robust observer for given error model of DC motor:

$$\hat{e}_{1,k} = \hat{e}_{1,k-1} + T\{-\lambda_3 L^{1/3}\left|\hat{e}_{1,k-1} - e_{1,k-1}\right|^{2/3} sign\left(\hat{e}_{1,k-1} - e_{1,k-1}\right) + \hat{e}_{2,k-1}\}$$

$$\hat{e}_{2,k} = \hat{e}_{2,k-1} + T\{\hat{f}_{1,k-1} + b_o u\}$$

$$\hat{f}_{1,k-1} = -\lambda_2 L^{1/2}\left|\hat{e}_{2,k-1} - \hat{e}_{1,k-1}\right|^{1/2} sign\left(\hat{e}_{2,k-1} - \hat{e}_{1,k-1}\right) + \hat{e}_{3,k-1}$$

$$\hat{e}_{3,k} = \hat{e}_{3,k-1} + T\{-\lambda_1 L sign\left(\hat{e}_{3,k-1} - \hat{f}_{1,k-1}\right)\} \quad (5\text{-}5.5)$$

Discrete robust feedback linearization control law implemented for DC motor based on Emulation design procedure:

$$u[k] = \frac{1}{b_o}\left(-\hat{f}[k] - k_1 e_1[k] - k_2 \hat{e}_2[k]\right)$$

where $\hat{f}[k]$ and $\hat{e}_2[k]$ are estimated by observer and

$$b_o = -533.36$$

For Euler approximate discrete-time robust observer Equation (5-5.5) $\lambda_1 = 1$, $\lambda_2 = 9$ and $\lambda_3 = 25$ (designed as suggested in [94]). These observer parameters are tuned at smaller values as compared with parameters ($\lambda_1 = 10,\ \lambda_2 = 3200,\ \lambda_3 = 3200$) of observer Equation (5-4.5), because for the case of Maglev system, the discrete-time observer estimates for a faster changing output / system dynamics (chirp trajectory of 37.68 *rad/sec*frequency). Whereas for DC motor, discrete-time observer estimates and shaft speed (output) tracks reference speed signal of comparatively lower frequency (1.57 *rad/sec.)*. Also for Equation (5-5.5) *Lipschitz constant*$L = 1$ and

$$\hat{e}_{1,k=0} = 0,\ \hat{e}_{2,k=0} = 0,\ \hat{e}_{3,k=0} = 0 \quad (5\text{-}5.6)$$

The performance of discrete robust observer Equation (5-5.5) for sampling period $T = 0.0001sec.$, $T = 0.001sec.$ and $T = 0.01sec.$ is investigated in computer simulations. Controller gains $k_1 = 9$ and $k_2 = 6$ are designed using pole placement (two poles are placed at $-3, -3$) method such that $A - BK$ is Hurwitz. Matrices A and B are obtained by linearization of Equation (5-5.2) at 0 (bias value of reference shaft speed).

DC motor in Equation (5-5.2) is initially at (shaft speed) $x_1 = 0$ and (armature current) $x_2 = 0$. Output $y = x_1$, i.e. motor shaft speed tracks sinusoidal reference shaft speed of 0.02 amplitude, 0 bias and frequency 1.57 *rad/sec*. Additive white Gaussian noise of variance 10^{-7} and $10^{-5}$is added as process noise and measurement noise respectively.

5.5.3 Results discussion

Reference signal and DC motor shaft speed, both are plotted in Figure 5-5 for asimulation duration of 80 seconds. Tracking simulation is performed with a sampling period $T = 0.001sec.$ for discrete controller and observer Equation (5-5.5). Motor shaft speed tracks reference speed signal (sine wave) of 0.02 amplitude, 0 bias and 1.57 *rad/sec* frequency. Shaft speed tracks reference speed signal from the beginning (simulation duration $0-8$ seconds) to the end (simulation duration $72-80$ seconds) of simulation as can be seen in Figure 5-5.

Tracking error plot is shown in Figure 5-6. Error plot is obtained by difference of reference speed signal and motor shaft speed. Tracking error remains within range of 10^{-3} for sampling period $T = 0.001sec.$ as shown in Figure 5-6.

Euler approximate discrete observer Equation (5-5.5) for DC motor, when tuned demonstrates lower tracking error as compared to tuned discrete observer Equation (5-4.5) for Maglev system at same sampling period $T = 0.001sec.$ This is because, observer Equation (5-5.5) estimates system dynamics which change at a lower rate (reference signal frequency 1.57 *rad/sec*) as compared with observer Equation (5-4.5), which estimates faster changing (reference signal frequency 37.68 *rad/sec*) output / system dynamics.

Figure 5-7 demonstrates effect of increasing sampling period T on performance of discrete observer Equation (5-5.5). The tracking error only slightly increases from 2×10^{-3} to 4×10^{-3} when sampling period is increased from $T = 0.001sec.$ to $T = 0.01sec.$ As the motor shaft speed is changing at a lower rate, increasing sampling period

slightly affects the performance of observer Equation (5-5.5). On further increasing sampling period of observer

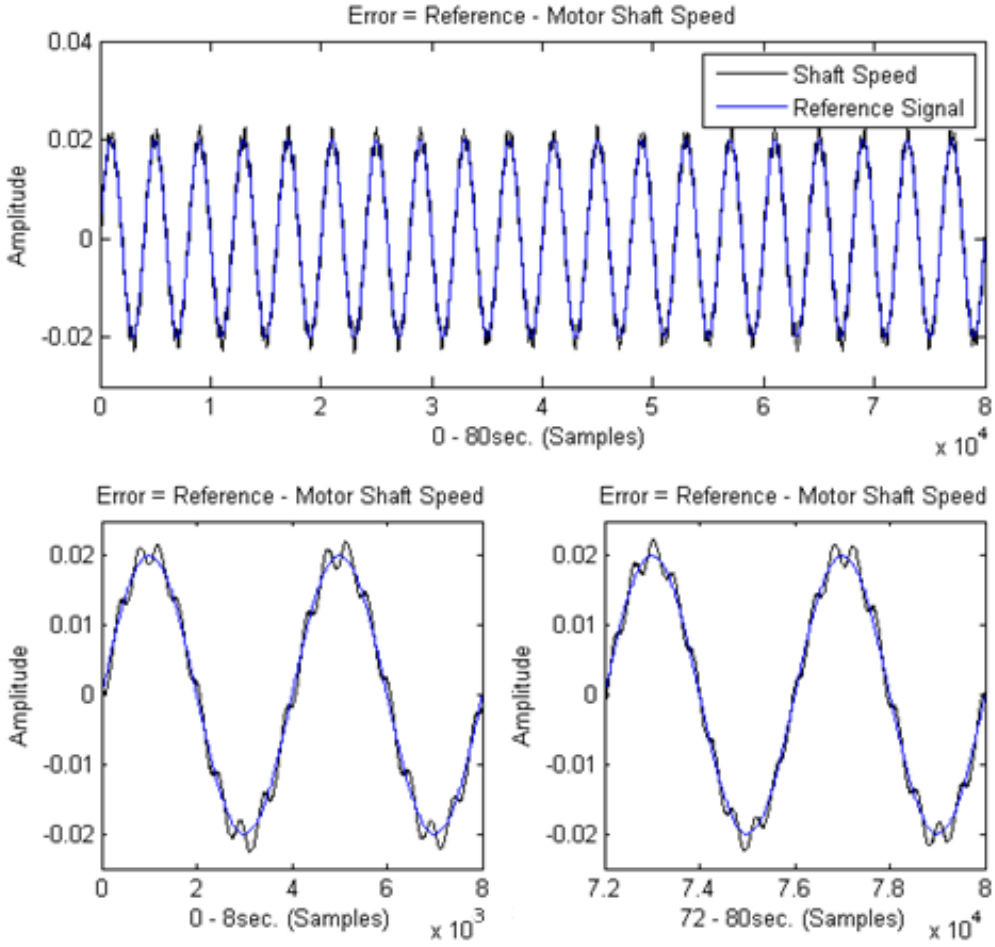

Figure 5-5. DC motor shaft speed tracking a reference sine (1.57 *rad*/sec) signal. Sampling period *T* = 0.001sec.

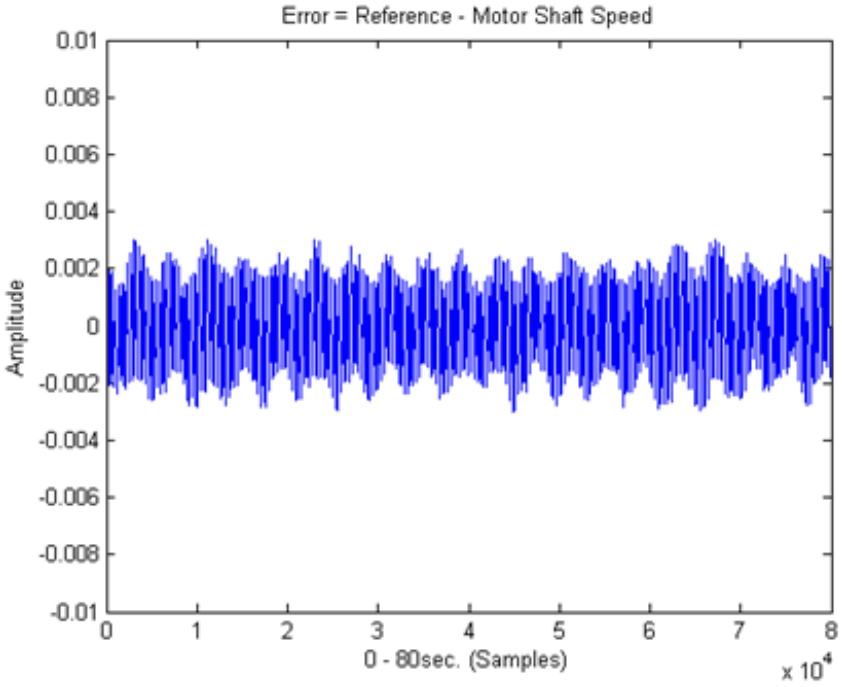

Figure 5-6. Shaft speed tracking error plot for reference sine (1.57 *rad*/sec) signal. Sampling period T = 0.001sec.

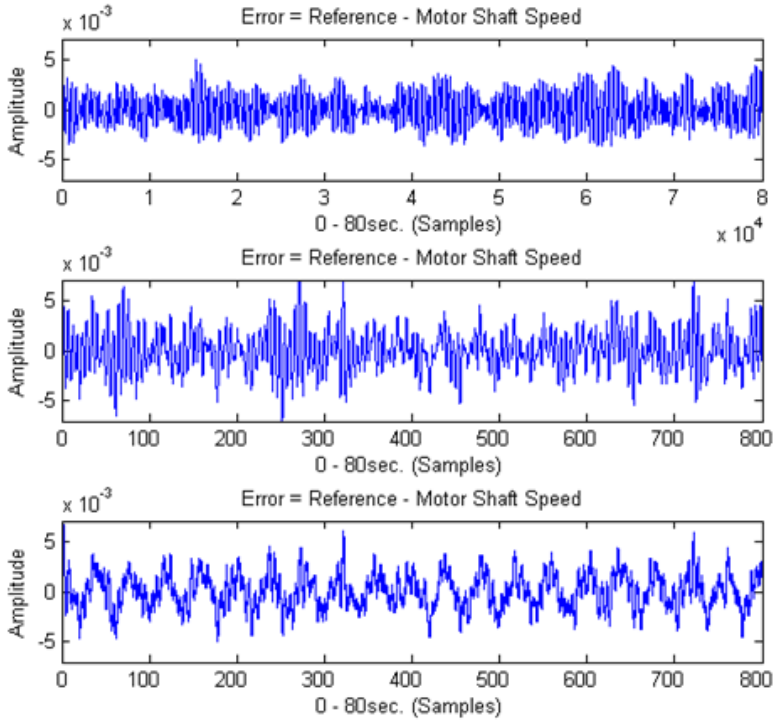

Figure 5-7. Shaft speed tracking error plot for reference sine signal with i) T = 0.01sec. ii) T = 0.1sec. iii) T = 0.1sec. and retuned $\lambda_1, \lambda_2, \lambda_3$.

Equation (5-5.5) and controller to$T = 0.1sec.$, the tracking error crosses value of 5×10^{-3}. This tracking error is reduced when observer Equation (5-5.5) parameters ($\lambda_1 = 1$, $\lambda_2 = 45$, $\lambda_3 = 90$) are retuned, and simulation is performed again at sampling period $T = 0.1sec.$ (as can be seen in Figure 7-7).

Simulation results suggests that increasing sampling period more adversely affects tracking results of Maglev system than that of DC motor. This is because of faster changing output / system dynamics of Maglev system as compared with DC motor.

5.6 Tracking in Presence of Disturbance & Parameters Perturbation

In this section tracking performance is demonstrated for robust feedback linearization controller Equation (5-3.3) and Euler approximate discrete robust observer Equation (5-3.2) in presence of disturbance and parameter perturbation for Maglev system Equation (5-4.1).

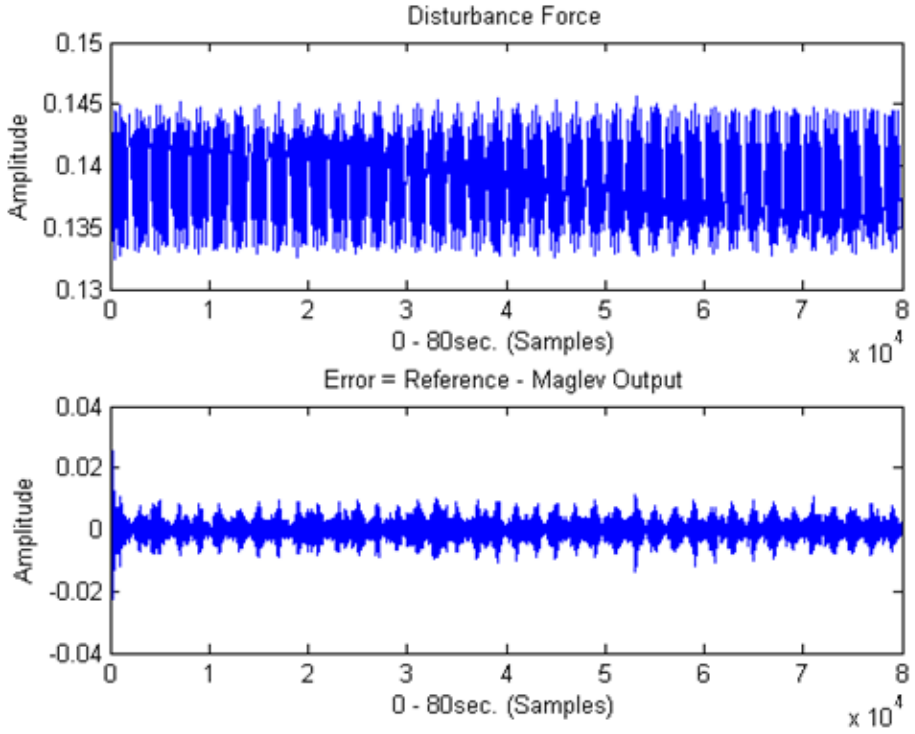

Figure 5-8. i) Disturbance force acting on lower magnet disc. ii) Maglev tracking error plot for reference bidirectional chirp (37.68 *rad*/sec) trajectory in presence of disturbance. Sampling period T = 0.001sec.

5.6.1 Tracking in presence of disturbance

Maglev system dynamics as modeled in [93] is given by Equation (5-4.1). Where y_1 is position and m is mass of levitated magnet disc. $m\ddot{y}_1$ is force (in direction of $\ddot{y}_1$) acting on disc due to acceleration $\ddot{y}_1$, mg is weight (force) acting downwards, and force due to friction is given by $c_1\dot{y}_1$.

The levitated magnet disc moves (levitates) on a vertical glass guide rod. A repulsive magnetic force F_{u11} acts on levitated disc by energizing (lower) drive coil. Height of levitated disc increases on increasing drive coil current and vice versa.

A disturbance (repulsive) force acts on the levitated magnet disc (lower) if a second (upper) magnet disc is stacked on top of lower disc. Both discs move (levitate) on the vertical glass guide rod. Upper magnet disc is stacked on top of lower such that both magnet discs have similar pole close to one another. Upper disc levitates and moves under influence of magnetic force from upper (second) drive coil.

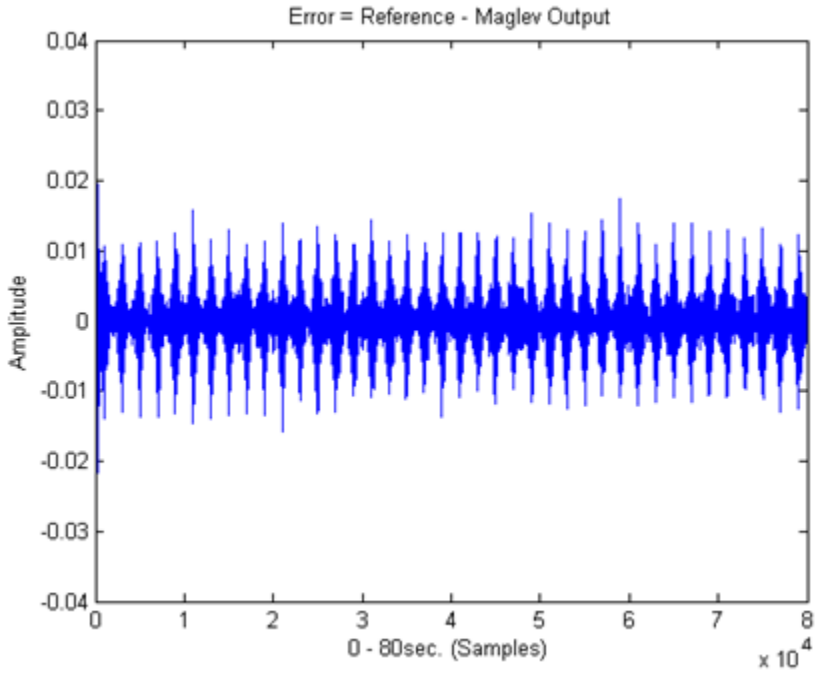

Figure 5-9. Maglev tracking error plot for reference bidirectional chirp (37.68 *rad*/sec) trajectory in face of parameters perturbation. Sampling period T = 0.001sec.

The disturbance (repulsive) force acting on lower disc due to upper disc levitation (movement) is given by [93]:

$$F_{m12} = \frac{2.69}{(0.12 + y_2 - y_1 + 4.2)^4} \tag{5-6.1}$$

which depends on positions ($y_1$and y_2) of lower and upper levitated magnet discs.

Furthermore in this work, performance of robust feedback linearization controller and observer Equation (5-4.5) is demonstrated for tracking (sampling period $T = 0.001sec.$) a chirp trajectory (0.03*meter* amplitude, 0.05 bias and 37.68 *rad/sec*frequency) by lower (first) magnet disc in presence of disturbance force Equation (5-6.1). Upper (second) magnet disc is made to track a sine trajectory (0.01*meter* amplitude and 18.84 *rad/sec*frequency) using second drive coil; and upper disc position y_2 (while tracking sine trajectory) determines F_{m12} in Equation (5-6.1).

Disturbance force acting on lower levitating magnet disc is shown in Figure 5-8, which has approximate mean and frequency of 0.14 *N* and 37.68 *rad*/sec respectively. Tracking error of lower levitating magnet disc in presence of this disturbance force is also shown in Figure 5-8. It can be seen that tracking error value in both Figure 5-8 and Figure 5-3 is considerably same, which demonstrates robust behavior of observer Equation (5-4.5) and controller Eqaution (5-3.3) in presence of disturbance force.

5.6.2 Tracking in face of parameters perturbation

Equation (5-4.1) models Maglev system dynamics, where mass of levitated magnet disc is $m = 0.12kg$ and coefficient of viscosity $c_1 = 1.1673$ (from Table 5). Figure 5-9 shows tracking error for chirp trajectory (37.68 *rad*/sec) at sampling period $T = 0.001sec$. for $+50\%$ mass perturbation and $+50\%$ perturbation in value of coefficient of viscosity. Tracking error in both Figure 5-9 and Figure 5-3 is considerably same which demonstrates performance of presented sampled-data robust feedback linearization scheme.

5.7 Conclusion

In this chapter a robust feedback linearization scheme for nonlinear minimum phase systems has been presented. Approximate discrete model (based on Euler approximation) of higher order SMO has been successfully used in simulation (tracking applications) for

magnetic levitation and DC motor systems. Simulation results show that performance of approximate discrete observer depends on sampling period and system dynamics.

Next chapter presents CKF estimator based output feedback control scheme. Comparison of tracking performances exhibited by both sampled-data state feedback control, and sampled-data output feedback control in presence of external disturbance and parameters perturbation is also presented. Also in next chapter, tracking performance of SMO is compared with tracking performances of Kalman and SSRLS estimators based output feedback control schemes.

Supporting publication: Asim Zaheer Ud Din, Yasar Ayaz, Muhammad Salman, Momena Hasan and Niaz Mehdi, "Sampled-data robust feedback linearization using higher order sliding mode observer", 15th International Bhurban Conference on Applied Sciences & Technology (IBCAST), 2018, Islamabad, Pakistan.

CHAPTER 9
CONCLUSION

- Back in year 2010-2011, and its subsequent years, Pakistan faced a persistent power crisis, which was a function of faulty transmission system, insufficient power generation, and increasing demand. There was no system intact to get real time grid performance data. Inefficient power dispatch and suboptimal utilization of generated power capacity were common practices. This resulted into nationwide electricity shortages. The current work funded under USAID Energy Policy program to upgrade Pakistan power system has been successfully completed. It has been accepted by power sector stakeholders as an important tool for grid monitoring, and electricity supply to consumer. The work has demonstrated its effectiveness in load-shedding reduction [35], and in optimal utilization of power generation capacity for PESCO as discussed in thesis.
- The deployed telemetry solutions for IESCO and PESCO have helped in circuits (power transformers) load balancing and quick identification of malfunctioning grid station equipment.
- The work presented in this thesis is an effort towards developing a detailed framework for forecasting short-term and long-term power requirements of electricity distribution companies of Pakistan. With the installation of system real-time monitoring and logging of power data for distribution companies is already being carried out. The data from these systems can be used to learn model for power forecasting. The forecasted power values are useful for both National Transmission & Dispatch Company, and for the electricity distribution companies of Pakistan.
- This project is a forward step towards achieving a planned transition for formation of a autonomous central power purchasing agency and an open power trading market in Pakistan.
- Furthermore, the thesis has presented estimation based output feedback control schemes for nonlinear minimum phase systems, e.g magnetic levitation system, and DC motor system. The purpose of estimation work done in thesis is to implement

these state estimation and control schemes for active magnetic bearing and motor systems in electricity generation facilities of smart grid regime. Output feedback tracking schemes using following different estimators have been presented:

1. Kalman estimators (Extended Kalman filter, Unscented Kalman filter and Cubature Kalman filter)
2. State-Space Recursive Least Squares filter
3. Sliding Mode observer
4. Neural network-aided Unscented Kalman filter (neuro-estimator)
5. Dual Unscented Kalman filter

- The presented controllers are Emulation Design based discrete feedback linearization controller, and discrete-controller obtained by eigenvalues placement.
- It has been demonstrated via computer simulations that output feedback control based on estimator is more robust to external disturbances and parameters perturbation as compared to state feedback control.
- Output feedback tracking scheme for case of unknown plant model of nonlinear system have been presented using State-Space Recursive Least Squares filter.
- Two robust feedback linearization schemes based on estimation have been presented. First scheme employs Euler approximate discrete-time Sliding Mode observer, whereas second scheme is based on neural network-aided Unscented Kalman filter. Simulation results demonstrate that feedback linearization based on Sliding Mode observer is more robust than feedback linearization scheme using neural network-aided Unscented Kalman filter.
- Dual Unscented Kalman filter based output feedback control scheme has been presented for nonlinear systems with unknown system parameters.

6.1 Recommendations for future work

- Development of anomaly detection algorithm for electricity distribution company energy network.
- Development of multivariate medium-term and long-term power forecasting frameworks.

- Implementation of state and parameter estimation algorithms, and output feedback controllers for active magnetic bearings in generator and motor of electricity generation facility.
- Implementation of estimation based control schemes presented in Chapter 7 may be performed on magnetic levitation lab test bench.
- In Chapter 6, Chapter 7 and Chapter 8 control schemes employ Kalman estimation and Emulation Design based discrete feedback linearization controller. Performances of control schemes based on Kalman estimation and other controller designing techniques may be explored.
- Stability and convergence of Euler approximate discrete-time Sliding Mode observer discussed in Chapter 5, may be studied and proofed.
- Other approximation techniques to obtain discrete-time observer from [89] may be studied.
- Stability and convergence of NN-UKF estimator discussed in Chapter 7, may be studied. NN-UKF estimator employing Joint UKF technique for robust output feedback configuration may be explored. Also performance of NN-UKF estimator with control techniques other than feedback linearization be studied; and dynamic neural network may also be employed in NN-UKF implementation.
- In Chapter 8, output feedback control employing dual UKF estimation is presented. Joint UKF estimation for state and parameter estimation in output feedback control be studied, and compared with dual UKF estimation. Also performance of dual UKF estimation with control techniques other than feedback linearization be studied; and dynamic neural network may also be employed in dual UKF state-parameter estimation implementation.
- Control schemes discussed in Chapter 5, Chapter 6, Chapter 7, and Chapter 8 are based on Emulation Design technique for discrete controller designing. These control schemes may be designed using Direct Discrete-time Design and Sampled-data Design techniques for discrete controller designing.

APPENDIX A

NONLINEAR DC MOTOR MODEL

Dynamics of DC motor with a rigid arm attached are given by [84]:

$$J\ddot{\theta} = k_c u - F$$

where, θ represents shaft position, J is total moment of inertia of both rotor and arm, k_c is motor input constant, and u is applied terminal voltage. F is unknown frictional torque, and can be represented by dynamical Lugre model [116]:

$$F = z + \sigma_1(\varepsilon_0 \dot{z}) + \sigma_2 \dot{\theta}$$

$$\varepsilon_0 \dot{z} = \dot{\theta} - \frac{|\dot{\theta}|}{s(\dot{\theta})} z$$

where, $\boldsymbol{\varepsilon_0}$ is reciprocal of average stiffness of bristles. $\boldsymbol{\sigma_1}$ is damping coefficient of bristles, whereas, $\boldsymbol{\sigma_2}$ is viscous friction. Stribeck curve, $\boldsymbol{s(\dot{\theta})}$, is given as:

$$s(\dot{\theta}) = \begin{cases} F_{c^+} + (F_{s^+} - F_{c^+}) e^{-\left(\dot{\theta}/v_s\right)^2} & \dot{\theta} > 0 \\ F_{c^-} + (F_{s^-} - F_{c^-}) e^{-\left(\dot{\theta}/v_s\right)^2} & \dot{\theta} < 0 \\ \left(s(\theta_+) + s(\theta_-)\right)/2 & \dot{\theta} = 0 \end{cases}$$

$\boldsymbol{F_{s^\pm}}$ and $\boldsymbol{F_{c^\pm}}$ are static and coulomb frictions respectively. $\boldsymbol{v_s}$ is Stribeck velocity of motor. DC motor parameters along with nominal values are listed in following table.

TABLE. DC MOTOR PARAMETERS

Symbol	**Value / Unit**
J	0.095 kgm^2
k_c	2.5
σ_1	1.5*Ns/m*
σ_2	0.004*Ns/m*
ε_0	0.01
F_{c^+}	0.023 k_c *N*
F_{c^-}	0.021 k_c *N*
F_{s^+}	0.058 k_c *N*
F_{s^-}	0.052 k_c *N*
v_s	0.01 *m/s*

APPENDIX B

QUANTITATIVE PERFORMANCE COMPARISON OF ESTIMATION BASED OUTPUT FEEDBACK CONTROL SCHEMES

Quantitative performance analysis of earlier discussed output feedback control schemes is obtained by implementing tracking control for Maglev system. A step reference of zero bias, and 10 cm (0.1m) amplitude is applied (for Maglev lower levitating disc), and performance of each scheme (estimator + feedback linearization controller) is observed. Response analysis parameters (% overshoot, settling time, mean squared error) for Maglev output (position of lower levitating disc) to step reference are obtained from pictorial view. In simulation, initial condition (values) for both plant states and estimator are assumed zero. Measurement noise ± 0.75 cm is simulated as additive white Gaussian noise (AWGN). % overshoot, settling time and mean squared error for previously discussed schemes is tabulated in Table 1. Whereas, % overshoot, settling time and mean squared error in presence of disturbance is tabulated in Table 2. Disturbance force is shown in Figure 2. In Table 3, % overshoot, settling time and mean squared error are tabulated for output feedback control schemes using UKF and neural network-aided UKF estimators in presence of disturbance force shown in Figure 3.

TABLE 1

Feedback linearization controller (FLC)	Incorporates noise information	% Over shoot	Settling time (ms). 3% criteria	Mean squared error (cm)
FLC + State feedback	NA	2.3†	~ 0†	0.161†
FLC + Sliding mode observer	No	1.8**	~ 0**	0.231**
FLC + SSRLS estimator	No	2.4	2.2	0.390
FLC + EKF estimator	Yes	1.8	~ 0	0.162
FLC + UKF estimator	Yes	3	2.5	0.479
FLC + CKF estimator	Yes	1.8	~ 0	0.169
FLC + (State-parameter[ɣ] estimation)	Yes	4.6 ($m_{\circ}$=0.09) 6.8 ($m_{\circ}$=0.05) 9.1 ($m_{\circ}$=0.01)	5 8.3 11	0.676 0.921 1.258

† similar simulation conditions, except, measurement noise ± 0.1 cm simulated as additive white Gaussian noise (AWGN).

** for step reference. See Figure 1.

[ɣ] actual mass of Maglev disc 0.12kg, $m_{\circ}$= parameter (mass) initial estimate.

TABLE 2

Feedback linearization controller (FLC)	Incorporates noise information	% Over shoot	Settling time (ms). 3% criteria	Mean squared error (cm)
FLC + State feedback	NA	-70†	> 20†	1.50019
FLC + Sliding mode observer	No	1.8**	~ 0**	0.232**
FLC + SSRLS estimator	No	-80	> 20	1.82
FLC + UKF estimator	Yes	5.5	8.75	0.769

† similar simulation conditions, except, measurement noise ± 0.1 cm simulated as additive white Gaussian noise (AWGN).

** for step reference. See Figure 1.

TABLE 3

Feedback linearization controller (FLC)	Incorporates noise information	% Over shoot	Settling time (ms). 3% criteria	Mean squared error (cm)
FLC + UKF estimator	Yes	58	> 20	3.18
FLC + NN-aided UKF	Yes	6.2	9.5	0.98

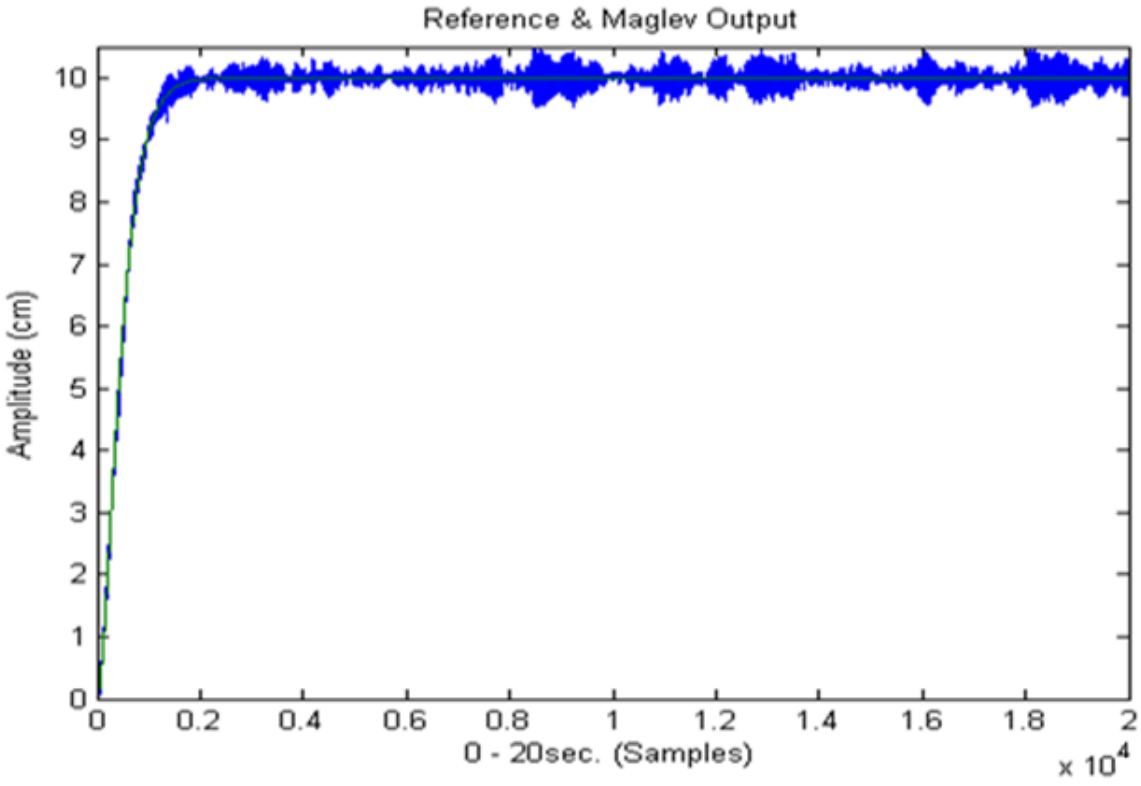

Figure 1. Step reference & Maglev output for sliding mode observer

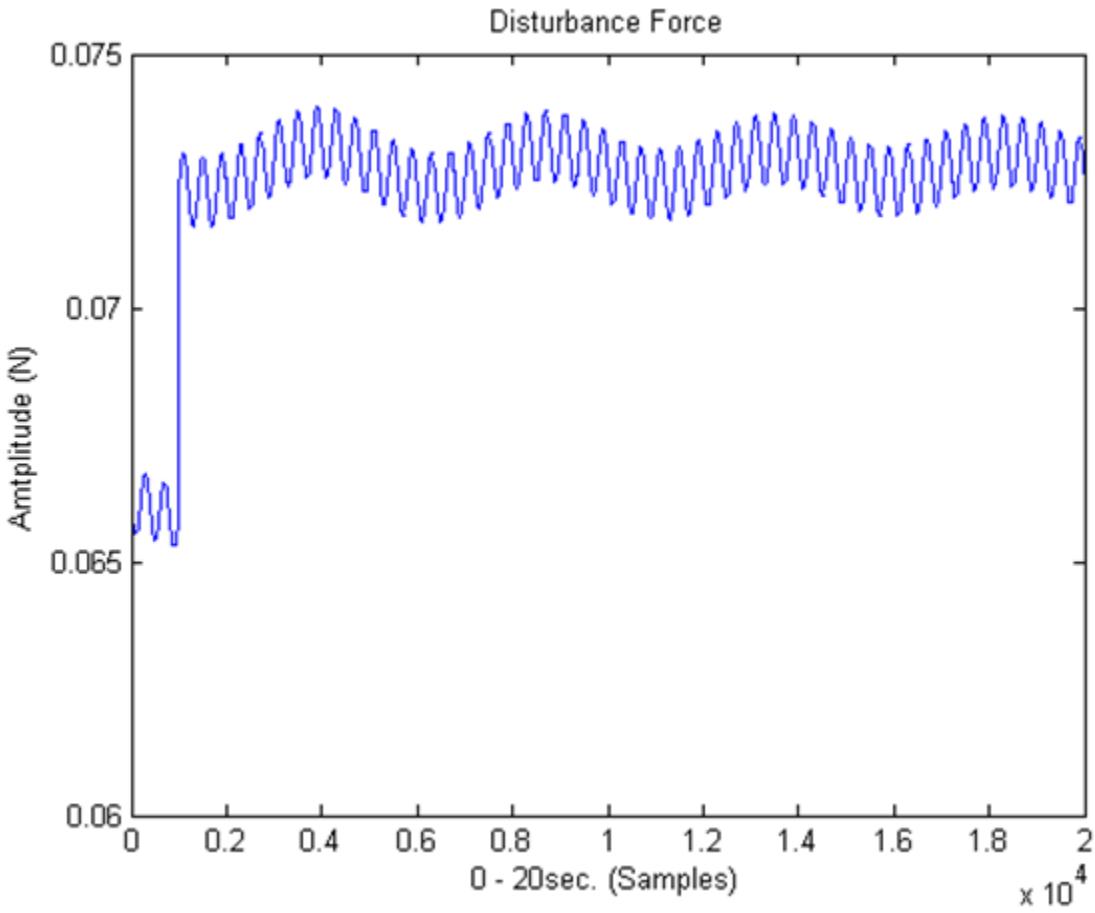

Figure 2. Disturbance force for Table 2 simulation results

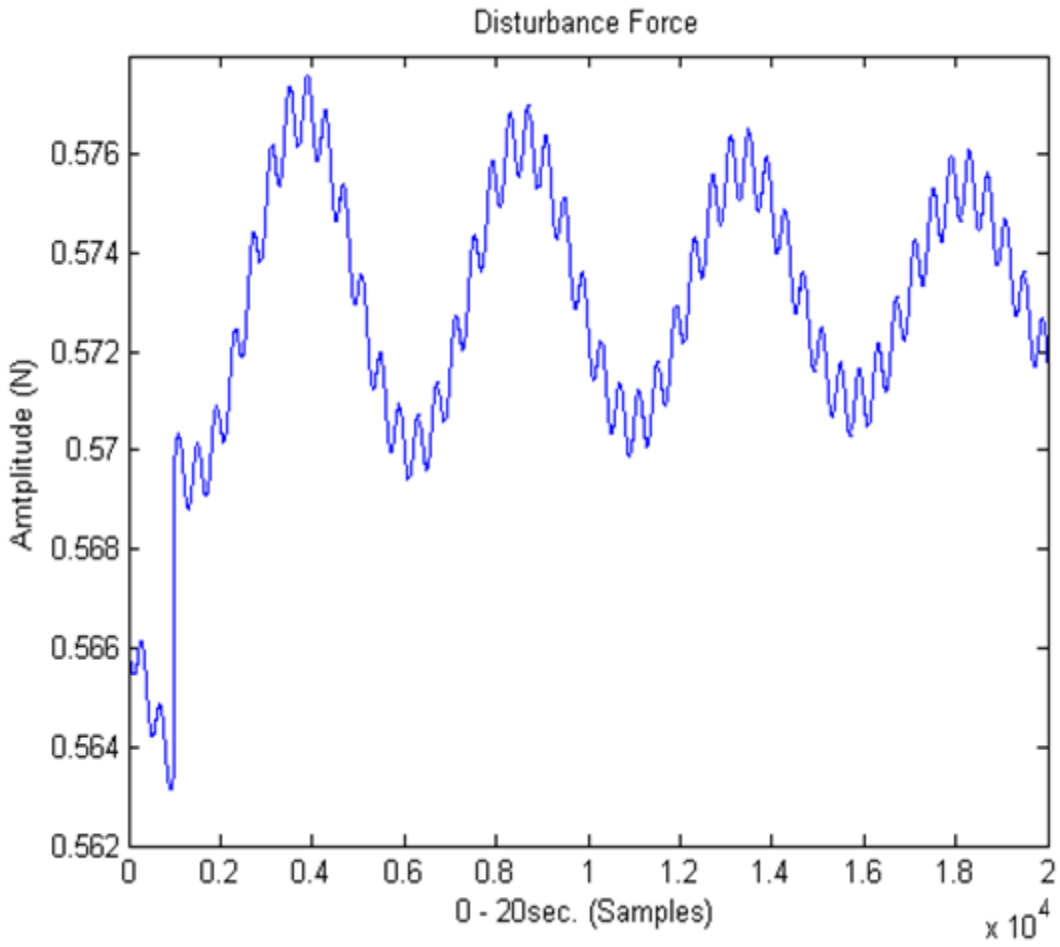

Figure 3. Disturbance force for Table 3 simulation results

REFERENCES

[1] U.S. Department of Energy, "Smart Grid", Avalilabe: https://www.smartgrid.gov/the_smart_grid/index.html. Accessed Jan, 2018.

[2] National Institute of Standards and Technology, "NIST framework and roadmap for smart grid interoperability standards, release 1.0", Available: http://www.nist.gov/publicaffairs/releases/upload/smartgridinteroperabilityfinal.pdf. Accessed Jan, 2018.

[3] G. Dileep, "A survey on smart grid technologies and applications," *Renewable Energy*, vol. 146, pp. 2589-2625, 2019.

[4] R. Kappagantu and S. A. Daniel, "Challenges and issues of smart grid implementation: A case of Indian scenario," *Journal of Electrical Systems and Information Technology*, vol. 5, no. 3, pp. 453-467, 2018.

[5] M. H. F. Wen, K. Leung, V. O. K. Li, X. He, "A survey on smart grid communication system," *APSIPA Transactions on Signal and Information Processing*, vol. 4, 2015.

[6] S. N. Nazmus, A. Khandakar, A. G. Mark and M. Datta, "A survey of smart grid architectures, applications, benefits and standardization," *Journal of Network and Computer Applications*, vol. 76, pp. 23-36, 2016.

[7] M. Kim, "Toward Smart Microgrid with Renewable Energy: An Overview of Network Design, Security, and Standards," *International Conference on Computational Science and its Applications*, vol. 7971, pp. 142-156, 2017.

[8] Y. Jiang, C. Liu and Y. Xu, "Smart Distribution Systems," *Energies*, vol. 9, no. 297, 2016.

[9] D. M. Souran, H. H. Safa, B. G. Moghadam, M. Ghasempour and P. T. Heravi, "Smart grid technology in power systems," *Proceedings of the 6th International Workshop Soft Computing Applications*, vol. 357, pp. 1367−1381, 2017.

[10] T. Broeer, "Analysis of smart grid and demand response technologies for renewable energy integration: operational and environmental challanges," Ph.D. Dissertation, Department of Mechanical Engineering, University of Victoria, 2015.

[11] C. Lo and N. Ansari, "The progressive smart grid system from both power and communications aspects," *IEEE Communications Surveys & Tutorials*, vol. 14, pp. 799-821, 2018.

[12] H. Miao, G. Chen and Z. Dong, "Enhanced evolutionary heuristic approaches for remote metering smart grid networks," *IET Networks*, vol. 5, pp. 153-161, 2016.

[13] K. Mets, J. Ojea and C. Develder,"Combining power and communication network simulation for cost-effective smart grid analysis, "*IEEE Communications Surveys & Tutorials*," vol. 16, pp. 1771-1796, 2017.

[14] X. Luan, Z. Zheng and T.Wang,"Hybrid cooperation for machine-to-machine data collection in hierarchical smart building networks," *IET Communication*, vol. 9, pp. 421-428, 2018.

[15] F. M. Sallabi, A. M. Gaouda, A. H. EI-Hag and M. M. A.Salama, "Evaluation of ZigBee Wireless Sensor Networks under high power disturbances," *IEEE Transactions on Power Delivery*, vol. 29, pp. 13-20, 2017.

[16] H. Gharavi and B. Hu,"Wireless infrastructure M2M network for distributed power grid monitoring," *IEEE Network*, vol. 31, pp. 122-128, 2017.

[17] R. M. Sandoval, A. Garcia-Sanchez, J. Molina-Garcia-Pardo, F. Garcia-Sanchez, J. Garcia-Haro, "Radio-channel characterization of smart grid substations in the 2.4-GHz ISM band," *IEEE Transactions on Wireless Communications*, vol. 16, pp. 1294-1307, 2017.

[18] T. Demirci, A. Kalaycioglu, D. Kucuk, O. Salor, M. Guderi, S. Pakhuylu, T. Atalik, T. Inan, I. Cadirci, Y. Akkaya, et al, "Nationwide real-time monitoring system for electrical quantities and power quality of the electricity transmission system," *IET Generation, Transmission & Distribution*, vol. 5, pp. 540-550, 2019.

[19] P. Parikh, T. Sidhu and A. Shami, "A comprehensive investigation of wireless LAN for IEC 61850-based smart distribution substation applications," IEEE Transactions on Industrial Informatics, vol. 9, pp. 1466-1476, 2017.

[20] W. N. Gallo, T. Heimfarth, D. D. Ferreira and T. M. Mendes, "Real-time detection system of electrical disturbances for remote communication stations and smart grid," *IEEE Latin America Transactions*, vol. 15, pp. 1894-1900, 2017.

[21] M. Yigit, V. C. Gungor and P. Boluk, "Performance analysis of Hamming code for WSN-based smart grid applications," *Turkish Journal of Electrical Engineering & Computer Sciences*, vol. 26, pp. 125-137, 2018.

[22] K. Samarakoon, J. Ekanayake and N.Jenkins, "Reporting available demand response," *IEEE Transactions on Smart Grid*, vol. 4, pp. 1842-1851, 2018.

[23] I. Kocaarslan, M. T. Akcay, S. E. Ulusoy, E. Bal and H.Tiryaki, "Creation of a dynamic model of the electrification and traction power system of a 25 kV AC feed railway line together with analysis of different operation scenarios using MATLAB/Simulink," *Turkish Journal of Electrical Engineering & Computer Sciences*, vol. 25, pp. 4254-4267, 2017.

[24] L. Chuan and A.Ukil, "Modeling and validation of electrical load profiling in residential buildings in Singapore," *IEEE Transactions on Power Systems*, vol. 30, pp. 2800-2809, 2018.

[25] *Fundamental of Statistical Signal Processing*, 2nd ed., Prentice Hall, New Jersey, USA, 1998.

[26] *Advances in Modern Blind Signal Separation Algorithms: Theory and Applications*, Morgan & Claypool, Williston, USA, 2010.

[27] C. Scott and R.Nowak, "A Neyman-Pearson approach to statistical learning," *IEEE Transactions on Information Theory*, vol. 51, pp. 3806-3819, 2005.

[28] S. Bayram and S. Gezici,"On the restricted Neyman-Pearson approach for composite hypothesis-testing in presence of prior distribution uncertainty," *IEEE Transactions on Signal Processing*, vol. 59, pp. 5056-5065, 2019.

[29] P. Iervolino and R.Guida, "A novel ship detector based on the generalized-likelihood ratio test for SAR imagery," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 10, pp. 3616-3630, 2017.

[30] A. Moschitta, P.Carbone and C. Muscas,"Generalized Likelihood Ratio test for voltage dip detection," *IEEE Transactions on Instrumentation and Measurement*, vol. 60, pp. 1644-1653, 2017.

[31] Y. Zhang, R. Yang, K. Zhang, H. Jiang and J. J. Zhang, "Consumption behavior analytics-aided energy forecasting and dispatch," *IEEE Intelligent Systems*, vol. 32, no. 4, pp. 59–63, 2017.

[32] Y. Goude, R. Nedellec and N. Kong, "Local short and middle term electricity load forecasting with semi-parametric additive models," *IEEE Transactions on Smart Grid*, vol. 5, no. 1, pp. 440–446, 2018.

[33] J. Lin. (2017, September 22). *Electricity Markets: Theories and Applications*. (1st ed.) [Online]. Available: https://onlinelibrary.wiley.com/doi/book/10.1002/9781119179382.

[34] M. Shahbaz. (2018, Oct.). *Measuring Economic Cost of Electricity Shortage: Current Challenges and Future Prospects in Pakistan*. [Online]. Available: http://www.mpra.ub.uni-muenchen.de/67164/.

[35] A. Z. Uddin, Y. Ayaz, M. I. Ahmad, M. Hasan, S. Masaud and N. Muhammad, "Cost-effective telemetry for energy network of an electricity distribution company: part I," *Turkish Journal of Electrical Engineering and Computer Sciences*, vol. 27, no. 2, pp. 889–902, 2019.

[36] Y. He, F. Guang and R. Chen, "Prediction of electricity demand of China based on the analysis of decoupling and driving force," *IET Generation, Transmission & Distribution*, vol. 12, no. 13, pp. 3375–3382, July 2018.

[37] N. Kong, M. Bocquel, T. Barbier, R. Girard, E. Magliaro, G. Kariniotakis, G. Pelton and P. Cauchois, "Long-term forecast of local electrical demand and evaluation of future impacts on the electricity distribution network," *CIRED – Open Access Proceedings Journal*, vol. 2017, no. 1, pp. 2401–2405, Oct 2017.

[38] E. Terciyanli, T. Demirci, D. Kucuk, M. Sarac and I. Cadirci, "Enhanced nationwide wind-electric power monitoring and forecast system," *IEE Transactions on Industrial Informatics*, vol. 10, no. 2, pp. 1171–1184, May 2017.

[39] J. Kim, D. Kim, W. Yoo, J. Lee and Y. Kim, "Daily prediction of solar power generation based on weather forecast information in Korea," *IET Renewable Power Generation*, vol. 11, no. 10, pp. 1268–1273, August 2017.

[40] J. Esteves, R. Pastor, N. P. Silva, R. Pestana and Z. Chen, "Forecasting photovoltaics/concentrated photovoltaics at national level – Portugal experience," *IET Renewable Power Generation*, vol. 13, no. 5, pp. 703–709, April 2019.

[41] J. Ma, M. Yang, X. Han and Z. Li, "Ultra-short-term wind generation forecast based on multivariate empirical dynamic modeling," *IEEE Transaction on Industry Applications*, vol. 54, no. 2, pp. 1029–1038, March 2018.

[42] J. Toubeau, J. Bottieau, F. Vallee and Z. D. Greve, "Deep learning-based multivariate probabilistic forecasting for short-term scheduling in power markets," *IEEE Transactions on Power Systems*, vol. 34, no. 2, pp. 1203–1215, March 2019.

[43] F. Liu, M. Cai, L. Wang and Y. Lu, "An Ensemble model based on adaptive noise reducer and over-fitting prevention LSTM for multivariate time series forecasting," *IEEE Access*, vol. 7, pp. 26102–26115, Feb 2019.

[44] J. Yan, H. Zhang, Y. Liu, S. Han L. Li and Z. Lu, "Forecasting the high penetration of wind power on multiple scales using multi-to-multi mapping," *IEEE Transactions on Power Systems*, vol. 33, no. 3, pp. 3276–3284, May 2018.

[45] D. B. Alencar, C. M. Affonso, R. C. L. Oliveira and J. C. R. Filho, "Hybrid approach combining SARIMA and neural networks for multi-step ahead wind speed forecasting in Brazil," *IEEE Access*, vol. 6, pp. 55986–55994, Oct 2018.

[46] A. Baliyan, K. Gaurav and S. K. Mishra, "A Review of Short Term Load Forecasting using Artificial Neural Network Models," *Procedia Computer Science*, vol. 48, pp. 121–125, 2018.

[47] C. Cecati, J. Kolbusz, P. Rozycki, P. Siano and B. M. Wilamowski, "A novel RBF training algorithm for short-term electric load forecasting and comparative studies," *IEEE Transactions on Industrial Electronics*, vol. 62, no. 10, pp. 6519–6529, October 2017.

[48] B. Zhao, Y. Liang, X. Gao and X. Liu, "Short-Term Load Forecasting Based on RBF Neural Network," *Journal of Physics Conference Series*, vol. 1069, no. 1, 2018.

[49] R. Hu, S. Wen, Z. Zeng and T. Huang, "A short-term power load forecasting model based on the generalized regression neural network with decreasing step fruit fly optimization algorithm," *Neurocomputing*, vol. 221, pp. 24–31, 2017.

[50] L. Hanshen, Z. Yuan, H. Jinglu and L. Zhe, "A localized NARX Neural Network model for Short-term load forecasting based upon Self-Organizing Mapping," *IEEE 3rd International Future Energy Electronics Conference and ECCE Asia*, 2017.

[51] J. Moon, Y. Kim, M. Son, and E. Hwang, "Hybrid Short-Term Load Forecasting Scheme Using Random Forest and Multilayer Perceptron," *Energies*, vol. 11, no. 3283, 2018.

[52] K. B. Sahayand and M. M. Tripathi, “Day ahead hourly load forecast of PJM electricity market and iso New England market by using artificial neural network,” *Proc. ISGT*, 2018, pp.1–5.

[53] H. Butwiengpun, S. Tunyasrirut, S. Wangnippantoand and W. Permpoonsinsup, “Hourly load forecasting of power system over Northeastern Thailand using artificial neural network,” in *Proc. IEECON*, Pattaya, 2017, pp.1–5.

[54] A. Yadav, A. Sahay, M. R. Yadav, S. Bhandari, A. Yadav and K. B. Sahay, “One hour ahead short-term electricity price forecasting using ANN algorithms,” in *Proc. ICUE*, Phuket, 2018, pp.1–5.

[55] H. A. Shaikh, M. A. Rahman and A. Zubair, “Electric load forecasting with hourly precision using long short-term memory networks,” in *Proc. ECCE*, Cox’s Bazar, 2019, pp.1–5.

[56] S. Singh, S. Hussain and M. A. Bazaz, “Short term load forecasting using artificial neural network,” in *Proc. Fourth International Conference on Image Information Processing*, Shimla, India, 2017.

[57] A. R. Kham, S. Razzaq, T. Alquthami, M. R. Moghal, A. Amin and A. Mahmood, “Day ahead load forecasting for IESCO using artificial neural network and bagged regression tree,” in *Proc. ICPESG*, Mirpur, 2018, pp.1–5.

[58] S. B. Taieb, A. Sorjamma and G. Bontempi, “Multiple-output modeling for multi-step-ahead time series forecasting,” *Neurocomputing*, vol. 73, no. 10, pp. 1950–1957, Jun. 2014.

[59] S. B. Taieb, G. Bontempi, A. F. Atiya and A. Sorjamaa, “A review and comparison of strategies for multi-step ahead time series forecasting based on the nn5 forecasting competition,” *Expert Systems with Applications*, vol. 39, no. 8, pp. 7067–7083, 2015.

[60] N. H. Packard, J. P. Crutchfeld, J. D. Farmer and R. S. Shaw, “Geometry from a time series,” *Physical Review Letters*, vol. 45, no. 9, pp. 712–716, 1980.

[61] J. Zheng, C. Xu, Z. Zhang, and X. Li, “Electric load forecasting in smart grids using long-short-term-memory based recurrent neural network,” in *Proc.51st Annual CISS*, Baltimore, MD, 2017, pp.1–6.

[62] J. Fiot and F. Dinuzzo, “Electricity demand forecasting by multi-task learning,” *IEEE Transactions on Smart Grid*, vol. 9, no. 2, pp. 544–551, 2018.

[63] J. Zheng, C. Xu, Z. Zhang and X. Li, “Tutorial and Survey on Probabilistic Graphical Model and Variational Inference in Deep Reinforcement Learning,” in *Proc. IEEE Symposium Series on Computational Intelligence*, Xiamen, China, 2019.

[64] T. Mcelroy, “When are direct multi-step and iterative forecasts identical?,” *Journal of Forecasting*, vol. 34, pp. 315–336, 2018.

[65] A. Sorjamaa, J. Hao, N. Reyhani, Y. Ji and A. Lendasse, "Short-Term and long-term forecasting for the 3D point position changing by using artificial neural networks," *International Journal of Geo-information*, vol. 7, no. 86, 2018.

[66] M. T. Hagan (2017, September 14). Neural Network Design. (2nded.) [Online]. Available: https://hagan.okstate.edu/nnd.html.

[67] R. Adhikari (2018, February 26). An Introductory Study on Time Series Modeling and Forecasting. (1sted.) [Online]. Available: https://arxiv.org/abs/1302.6613.

[68] S. F. Crone and N. Kourentzes, "Forecasting high—frequency time series with neural networks – an analysis of modeling challenges from increasing data frequency," in *Proc. DMIN*, 2018, pp.1–7.

[69] Spreadsheets containing demand, generation, scheduled and unscheduled load shedding statistics for September 2018, October 2018, November 2018, December 2018, January 2019, & February 2019. NPCC (electronic communication), 2019

[70] WAPDA, PESCO's Performance Statistics, 2016, 2017 and 2018, Statistics Directorate WAPDA, Government of Pakistan, Lahore, 2016-18

[71] A. D. Pasternak, "Global Energy Futures and Human Development: A Framework for Analysis," Lawrence Livermore National Laboratory, University of California, US Department of Energy, Washington, D.C., October 2000.

[72] P. T. Mana and J. Hansen, "Distributed economic dispatch using framework for network co-simulation," in *Proc. 2018 IEEE Power & Energy Society General Meeting*, Portland, OR, USA, 2018, pp.1–5.

[73] A. Z. Uddin, Y. Ayaz, H. Sajid and M. Jawad, "Multivariate multistep short-term-power-forecasting for electricity distribution company," to be published.

[74] J. Gonzalez and W. Yu, "Non-linear system modeling using LSTM neural networks," *International Federation of Automatic Control*, vol. 51, no. 13, pp. 485-489, 2018.

[75] N. Kim, M. Kim and J. K. Choi, "LSTM based short-term electricity consumption forecast with daily load profile sequences," in *Proc. 2018 IEEE 7th Global Conference on Consumer Electronics*, Nara, Japan, 2018, pp.1–5.

[76] H. A. Shaikh, M. A. Rahman and A. Zubair, "Electric load forecasting with hourly precision using long short-term memory networks," in *Proc. 2019 International Conference on Electrical, Computer and Communication Engineering*, Cox's Bazar, Bangladesh, 2019, pp.1–4.

[77] Y. Liu, Y. Wang, X. Yang and L. Zhang, "Short-term travel time prediction by deep learning: A comparison of different LSTM-DNN models," in *Proc. 2017 IEEE 20th International Conference on Intelligent Transportation Systems*, Yokohama, Japan, 2017.

[78] Z. Chang, Y. Zhang and W. Chen, "Effective Adam-optimized LSTM neural network for electricity price forecasting," in *Proc. 2018 IEEE 9th International Conference on Software Engineering and Service Science*, Beijing, China, 2018.

[79] C. Gulcehre, J. Sotelo, M. Moczulski and Y. Bengio, "A robust adaptive stochastic gradient method for deep learning," in *Proc. 2017 International Joint Conference on Neural networks*, Anchorage, AK, USA, 2017.

[80] G. Hinton, "Neural Networks for Machine Learning," Available: https://www.cs.toronto.edu/~tijmen/csc321/slides/lecture_slides_lec6.pdf, Accessed Mar 2017.

[81] C. Liu, G. Liu and J. Fang, "Feedback linearization and extended state observer-based control for Rotor-AMBs system with mismatched uncertainties", *IEEE Transactions on Industrial Electronics*, vol. 64, no. 2, pp. 1313–1322, 2017.

[82] S. Mahjoub, F. Mnif and N. Derbel, "Second-order sliding mode approaches for the control of a class of underactuated systems", *International Journal of Automation and Computing*, vol. 12, no. 2, pp. 134–141, 2018.

[83] H. K. Khalil, "Extended High-Gain Observers as Disturbance Estimators", *Journal of Control, Measurement, and System Integration*, vol. 10, no. 3, pp. 125–134, 2017.

[84] S. Iqbal, C. Edwards and A. I. Bhatti, "Robust feedback *linearization* for minimum phase systems", in *Proceedings of the 18th International Federation of Automatic Control World Congress*, vol. 18, no. 1, pp. 11151–11156, 2014.

[85] X. Lu, Y. Xie and L. Chen, "Feedback linearization and sliding mode control for VIENNA rectifier based on differential geometry theory", *Mathematical Problems in Engineering*, 2015.

[86] F. M. Malik, M. B. Malik and K. Munawar, "Discretization error bounds for sampled-data control of nonlinear systems", *Arabian Journal for Science and Engineering*, vol. 38, pp. 3429–3435, 2014.

[87] *Discrete-time control system design with applications*, Springer-Verlag, New York, 2014.

[88] H. Beikzadeh, "Multirate nonlinear sampled-data systems: analysis and design," Ph.D. Dissertation, Department of Electrical and Computer Engineering, University of Alberta, 2018.

[89] H. Lee and Y. Fujisaki, "Output-feedback control for sampled-data systems with variable sampling rate", *International Journal of Control*, vol. 91, no. 4, pp. 897–906, 2017.

[90] *Analysis and Design of Nonlinear Control Systems: In honor of Alberto Isidori*, Springer-Verlag Berlin and Heidelberg GmbH & Co., 2010.

[91] *Nonlinear Systems*, 3rd Edition, Prentice Hall, 2002.

[92] C. Edwards and Y. B. Shtessel, “Adaptive continuous higher order sliding mode control”, *Automatica*, vol. 65, , pp. 183–190, 2016.

[93] *Manual for magnetic levitation system*, Educational Control Products, 2009.

[94] S. Mobayen and F. Tchier, “Robust global second-order sliding mode control with adaptive parameter-tuning law for perturbed dynamical systems”, *Transactions of the Institute of Measurement and Control*, vol. 40, no. 6, 2017.

[95] I. Arasaratnam and S. Haykin, “Cubature Kalman Filters”, *IEEE Transactions on Automatic Control*, vol. 54, no. 6, pp. 1254−1269, 2018.

[96] B. Zang, J. Han, H. Yun and X. Chen, “Nonlinear aeroelastic system identification based on neural network”, *Applied Sciences*, vol. 8, no. 10, 2018.

[97] M. A. Arain, H. V. H. Ayala and M. A. Ansari, “Nonlinear system identification using neural network”, in *International Conference on Emerging Trends and Applications in Information Communication Technologies*, vol. 281, pp. 122–131, 2016.

[98] K. Gaurav, A. K. Sahoo and S. K. Mishra, “Nonlinear system identification using functional link multilayer perceptron neural networks”, *International Journal of Applied Engineering Research*, vol. 10, no. 44, pp. 31542–31546, 2015.

[99] Y. Huang, Y. Zhang, B. Xu, Z. Wu and J. A. Chambers, “A new adaptive Extended Kalman filter for cooperative localization”, *IEEE Transactions on Aerospace and Electronic Systems*, vol. 54, no. 1, 2018.

[100] J. Ni, C. Wang, X. Fan and S. X. Yang, “A bioinspired neural model based Extended Kalman Filter for robot SLAM”, *Mathematical Problems in Engineering*, vol. 2014, 2016.

[101] L. Yao, T. Li, Y, Li, W. Long and J. Yi, “An improved feed-forward neural network based on UKF and strong tracking filtering to establish energy consumption model for aluminum electrolysis process”, *Neural Computing and Applications*, vol. 31, no. 8, pp. 4271−4285, 2018.

[102] D. Sierociuk, M. Macias, W. Malesza and G. Sarwas, “Dual estimation of fractional variable order based on the Unscented Fractional order Kalman filter for direct and networked measurements”, *Circuits, Systems, and Signal Processing*, vol. 35, pp. 2055−2082, 2016.

[103] S. R. Jondhale and R. S. Deshpande, “Kalman filtering framework-based real time target tracking in wireless sensor networks using generalized regression neural networks”, *IEEE Sensors Journal*, vol. 19, no. 1, pp. 224−233, 2019.

[104] S. Baldi, S. Yuan, P. Endel and O. Holub, “Dual estimation: constructing building energy models from data sampled at low rate”, *Applied Energy*, vol. 169, pp. 81−92, 2016.

[105] H. Guo, Z. Wang, Y. Li, D. Wang and G. Wang, "State of charge and parameters estimation for Lithium-ion battery using dual adaptive Unscented Kalman Filter", in *29th Chinese Control and Decision Conference*, 2017.

[106] M. Abdollahpouri, M. Haring, T. A. Johansen, et al, "Nonlinear State and Parameter Estimation using Discrete-Time Double Kalman Filter", in *International Federation of Automatic Control*, vol. 50, no. 1, pp. 11632-11638, 2017.

[107] S. Aslan, A. T. Cemgil, M. S. Aslan, B. U. Toreyin and A. Akin, "Joint parameter and state estimation of the hemodynamic model by iterative Extended Kalman smoother", *Biomedical Signal Processing and Control*, vol. 24, pp. 47−62, 2016.

[108] C. Lin, X. Gong, R. Xiong, and X. Cheng, "A novel H-infinity and EKF joint estimation method for determining the center of gravity position of electric vehicles", *Applied Energy*, vol. 194, pp. 609−616, 2017.

[109] S. Damodharao and T. S. L. V. Ayyarao, "Joint state and parameter estimation by Extended Kalman Filter technique", *International Journal of Engineering Research and Development*, vol. 11, pp. 42−52, 2018.

[110] I. Klein, I. Rusnak and Y. Bar-Shalom, "Joint Kalman filter for formation moving with wiener process acceleration", in *2014 IEEE 28th Convention of Electrical & Electronics Engineers*, Israel, 2016.

[111] H. Khodadadi and H. Jazayeri-Rad, "Applying a dual Extended Kalman Filter for the nonlinear state and parameter estimations of a continuous stirred tank reactor", *Computers and Chemical Engineering*, vol. 35, no. 11, pp. 2426−2436, 2014.

[112] Z. Xian, J. Lian, M. Shan, L. Zhang, X. He and X. Hu, "A square root Unscented Kalman filter for multiple view geometry based stereo cameras / inertial navigation", *International Journal of Advanced Robotic Systems*, vol. 13, no. 5, 2016.

[113] D. Lee, G. Vukovich, and R. Lee, "Robust adaptive Unscented Kalman Filter for spacecraft attitude estimation using quaternion measurements", *Journal of Aerospace Engineering*, vol. 30, no. 4, 2017.

[114] A. Yu, Y. Liu, J. Zhu and Z. Dong, "An improved dual Unscented Kalman Filter for state and parameter estimation", *Asian Journal of Control*, vol. 18, no. 4, pp. 1427−1440, 2016.

[115] Z. Zhou. L. Yang, and Y. Li, "An adaptive dual Kalman filtering algorithm for Locata/GPS/INS integrated navigation", in *2013 Proceedings of China Satellite Navigation Conference*, vol. 245, pp. 527−541, 2016.

[116] S. Goyat and M. Tushir, "Analysis DC motor nonlinear behavior using modeling and simulation", *International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering*, vol. 5, no. 1, pp. 263−268, 2016.