MACHINE LEARNING-BASED LOAD AND PRICE FORECASTING FOR ENERGY MANAGEMENT SYSTEMS

SUPERIOR UNIVERSITY

PH.D. ELECTRICAL ENGINEERING

By:

ADNAN YOUSAF

ROLL NO: PHEE-S15-001

SESSION: 2015-18

DEPARTMENT OF ELECTRICAL ENGINEERING

SUPERIOR UNIVERSITY

Lahore, Pakistan

MACHINE LEARNING-BASED LOAD AND PRICE FORECASTING FOR ENERGY MANAGEMENT SYSTEMS

SUPERIOR UNIVERSITY

A thesis submitted in partial fulfillment for the

degree of Doctor of Philosophy in

Electrical Engineering

by

ADNAN YOUSAF

ROLL NO: **PHEE-S15-001**

SESSION**: 2015-18**

SUPERVISOR:

DR. MUSTAFA SHAKIR

DEPARTMENT OF ELECTRICAL ENGINEERING

SUPERIOR UNIVERSITY

Lahore, Pakistan

Adnan Yousaf

Dated:

DEDICATION

To our last prophet Muhammad (Peace be upon him), my great hero and messenger of Allah Almighty. One of the most influential people that humanity has ever witnessed.

I dedicate the thesis to all my family members and my spouse.

RESEARCH COMPLETION CERTIFICATE

It is certified that the research work contained in this dissertation titled **"Machine Learning-based Load and Price Forecasting for Energy Management Systems"** has been investigated and carried out by **Mr. Adnan Yousaf** Roll No. PHEE-S15-001 under my supervision. Therefore, the undersigned hereby certify that they have read and recommended the dissertation entitled for the degree of **Philosophy in Electrical Engineering**.

EXAMINATION JURY/BOARD APPROVAL:

Sr. No	Name	Role	Signature
1	Dr. Arfan Jaffar	Chair	
2	Dr. Mustafa Shakir	Thesis Supervisor	
3	Dr. Intesar Ahmed	External Examiner-1	
4	Dr. Muhammad Kamran Khan	External Examiner-2	
5	Dr. Saif Ur Rehman	Internal Examiner-1	

AUTHOR'S DECLARATION

I, **Adnan Yousaf** Roll No. **PHEE-S15-001** student of Department of Electrical Engineering, Superior University, Lahore in the subject of **Ph.D. Electrical Engineering** Session 2015-18, hereby declare that the matter printed in the dissertation as **"Machine Learning-based Load and Price Forecasting for Energy Management Systems"** is my research work. The text and results mentioned in this dissertation have not been printed, published, or submitted in any form at any national or international organization. I hereby certify that this research does not involve any plagiarized material or results that another person has published. My colleagues and friends assisted me while carrying out this research; I identified their contribution as well. If it contains any plagiarized information, I bear full responsibility.

Signature:

Date: ___________________

____
Adnan Yousaf

PLAGIARISM UNDERTAKING

I solemnly state that the research work illustrated in this dissertation titled **"Machine Learning-based Load and Price Forecasting for Energy Management Systems"** is my work without any remarkable contribution of any other person. Small contribution taken from any colleague has been duly acknowledged and stated clearly.

I also recognize the zero-tolerance policy of the HEC and Superior University, Lahore, towards plagiarism. Therefore, as an author of the above-titled dissertation, I declare that this research does not involve any plagiarized material or results that another person has published, and any material used as a reference is correctly cited in the bibliography section.

I bear complete responsibility that if I am found guilty of any formal plagiarism in the above-titled dissertation even after the award of my degree, the University reserves the right to revoke my degree at any stage. Furthermore, HEC and the University have the right to declare blacklisted at any forum.

Adnan Yousaf
Dated:

LIST OF PUBLICATIONS

It is certified that following publication(s) have been made of the research work that has been carried out for this thesis:

Journal Articles:

1. **Adnan Yousaf,** Rao Muhammad Asif, Mustafa Shakir, Ateeq Ur Rehman, and Mohmmed S. Adrees **"An Improved Residential Electricity LF Using a Machine-Learning-Based Feature Selection Approach and a Proposed Integration Strategy"**, ***Sustainability, 2021***, Vol. 13, Issue 11, 6199. DOI: 10.3390/su.13116199 **(IF: 2.576) (Category: W).**
2. **Adnan Yousaf,** Mustafa Shakir, et al. **"An Improved Residential EPF Using a Machine-Learning-Based Feature Selection Approach and a Proposed Integration Strategy"**, Submitted in "***Sustainability.*** **(IF: 2.576) (Category: W).**

Conference Proceedings:

1. **A. Yousaf**, B. A. Khan, U. Bashir, and F. Ahmad, **"Overview of Implementing Microgrid, Its Policies, Incentives and Challenges in Pakistan,"** 2019 6th International Conference on Electrical and Electronics Engineering (ICEEE), 2019, pp. 6-11, DOI: 10.1109/ICEEE2019.2019.00009.
2. M. Ali, **A. Yousaf,** and F. Usman, **"Designing and simulation of load control & monitoring system through DSM technique,"** 2017 8th International Renewable Energy Congress (IREC), 2017, pp. 1-4, DOI: 10.1109/IREC.2017.7926042.

Adnan Yousaf

Dated:

ACKNOWLEDGEMENTS

In the name of Allah Almighty, the Most Beneficent, the Most Merciful. First and foremost, I would like to thank Almighty Allah for making all this possible for me and bestowing me with uncountable blessings. I express my sincere gratitude to Prof. Chaudry **Dr. Abdur Rehman** (Chairman Superior Group), **Dr. Sumaira Rehman** (Rector Superior University), **Pirzada Sami Ullah Sabri**, and **Capt. Sajjad Mehmood** for their invaluable leadership and inspiration throughout my academic tenure at Superior University.

I am extremely grateful and would like to express my sincere gratitude to **Dr. Mustafa Shakir** for his continuous encouragement and motivation towards pursuing a Doctor of Philosophy Degree in Electrical Engineering. It was enjoyable and a great experience to work, learn from someone very keen, dedicated, and have excellent command over the subject. Whenever I needed his guidance regarding any problem with the research, his office doors were always open for me. He was eager to address and resolve those queries on a priority basis. Moreover, his teaching methodology is a source of great inspiration for me.

I want to acknowledge all faculty and staff members of Superior University, Lahore, to provide a supportive and working environment.

In the end, I am very thankful to my parents, who have always inspired me with their inspiration and praise. They also provided me a carefree environment to help me maintain focus on my study with complete concentration. My father is the most enthusiastic about supporting all the decisions which I made. May Allah Almighty give them good health and keep them safe. Finally, to every person who directly or indirectly contributed to this work, their kindness means a lot to me. Thank you all very much.

Adnan Yousaf

Dated:

ABSTRACT

An intelligent load forecasting (LF) model is proposed for residential loads using a novel Machine Learning (ML)-based approach, achieved by assembling an integration strategy model with the Mean Absolute Percentage Error (MAPE) optimizer. In this proposed method, the time-series-based auto-regression schemes were carried out to collect historical data and set the objective functions of the proposed model. An algorithm with seven different auto-regression models was also developed and validated through a feed-forward adaptive-network-based fuzzy inference system (ANFIS) model based on the ML approach.

Moreover, a binary genetic algorithm (BGA) was deployed for the best feature selection, and the best fitness score obtained with Principal Component Analysis (PCA). A unique decision integration strategy is presented that led to a remarkably improved transformation in reducing MAPE. The model then tested by using a one-year Pakistan Residential Electricity Consumption (PRECON) dataset, and the attained results verify the validity of the proposed model with promising values in MAPE of 1.70%, 1.77%, 1.80%, and 1.67% for summer, fall, winter, and spring seasons, respectively. The overall improvement percentage is 17%, which represents a substantial increase for small-scale decentralized generation units.

In the second part, a novel and improved technique is introduced to forecast electricity prices. The data of various power producers, Capacity Purchase Price (CPP), Power Purchase Price (PPP), Tariff rates, and load demand from National Electric Power Regulatory Authority (NEPRA) is considered for MAPE reduction in PF. Eight-time series and auto-regression algorithms are developed for data fetching and setting the objective function. The feed-forward ANFIS based on the ML approach and space vector regression (SVR) is introduced to forecast price forecasting (PF) by taking input from time series and auto-regression algorithms. Best feature selection is made by adopting the BGA-PCA approach, which reduces the repeated, irrelevant, and unnecessary data that ultimately minimizes the complexity and computational time of the model. The proposed integration strategy computes the MAPE according to the above-mentioned steps, which exhibited significant improvement of the system.

Finally, the third part presents demand-side management (DSM) based on the Firefly algorithm (FA) has been presented. The peak load exerts extra stress on the grid; therefore, the FA has been established to forecast load to minimize peak load demand. By implementing FA, the cost of electricity has been reduced 21%, 19%, and 20% for building 1, 2, and 3, respectively. It utilizes the different types of the forecasted load for three residential buildings to verify the system's capability under different load scenarios. Similarly, it has also reduced the cost of electricity by implementing the data set of PF very efficiently.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATION

AI	Artificial Intelligence
AMI	Advanced Metering Infrastructure
ACO	Ant Colony Optimization
ANN	Artificial Neural Network
ANFIS	Adaptive- Network based Fuzzy Inference System
ATPS	Actual Time Pricing Scheme
BGA	Binary Genetic Algorithm
CPP	Capacity Purchase Price
PPP	Power Purchase Price
CTS	Consumption Time Scheme
DR	Demand Response
DSM	Demand Side Management
EHO	Elephant Herding Algorithm
EMC	Energy Management Controllers
EPF	Electricity Price Forecasting
EPS	European Power System
EU-NITE	European Network on Intelligent Technologies
FA	Firefly Algorithm
FS	Feature Selection
GA	Genetic Algorithm
GRU	Gated Recurrent Units
GRNN	Generalized Regression Neural Network
HPCCs	High Power Consumption Class

HEMC	Home Energy Management Controller
HVAC	Heating, Ventilation and Air-conditioning
ISO	Independent System Operator
IoT	Internet of Things
KF	Kalman Filter
LF	Load Forecasting
LED	Light Emitting Diode
LSTM	Long Short-Term Memory
LPCCs	Less Power Consumption Class
MAPE	Mean Absolute Percentage Error
ML	Machine Learning
MFFNN	Multi-Layer Feed Forward Neural Network
MLNN	Multi-Layer Neural Network
MLP	Multi-Layer Perceptron
NEPRA	National Electric Power Regulatory Authority
NEM	National Electricity Market
NFN	Neural Fuzzy Network
PAR	Peak to Average Ratio
PCA	Principal Component Analysis
PF	Price Forecasting
PRECON	Pakistan Residential Electricity Consumption Dataset
PSO	Particle Swarm Optimization
PMS	Price Management System
PV	Photo Voltaic

RBF-NN	Radial Basic Function-Neural Network
RCGA	Real-Coded Genetic Algorithm
RLNG	Re-Gasified Liquefied Natural Gas
RNN	Recurrent Neural Network
RT	Real Time
STLF	Short Time Load Forecasting
SVR	Space Auto Regression
SVM	Support Vector Machine
SOFM	Self-Organizing Feature Mapping
ToU	Time of Use
UPS	Uninterruptable Power Supply
EP	Energy Pricing
RES	Renewable Energy Source
FL	Fuzzy Logic
FNN	Fuzzy Neural Network
HSA	Harmony Search Algorithm

LIST OF SYMBOLS

Symbols	Meaning/Description
A	Set of appliances to be shifted
$\boldsymbol{a_i}$&$\boldsymbol{c_i}$	Set of premise parameters
$(\boldsymbol{ast})$ & $(\boldsymbol{\beta end})$	Start and end times of each appliance set by users
$\boldsymbol{D}$	Actual data
Ds	Discomfort level
$\boldsymbol{E}$	Overall improved percentage
$\boldsymbol{E_1}$	With proposed integration, improved percentage
$\boldsymbol{F_{t-1}}$	Load forecasting demand
$\boldsymbol{fit_{fun}}$	Feature selection function
$\boldsymbol{f}$	Forecast weights
$\boldsymbol{k_n, l_n, \text{and} m_n}$	Parametric set
$\boldsymbol{M_1}$	Average MAPE without a proposed integration strategy
$\boldsymbol{M_2}$	Average MAPE with only the proposed integrated approach without feature selection
$\boldsymbol{M_3}$	MAPE of 10 houses computed by our proposed model
m	Number of slots per hour
L	Time duration of each appliance
P	Power ratings of each appliance
$\boldsymbol{p_i, q_i, r_i}$	Parameters of $\overline{w}_n$
$\boldsymbol{P_{t-1}}$	Previous period load demand

$\mathbf{P_t}$	The regression forecast
$\boldsymbol{P_{y-i}}, \boldsymbol{P_t'}, \boldsymbol{P_t''}$ **and** $\boldsymbol{P_{t,SVR}}$	ANFIS inputs to nodes
TS	Switching on of appliances
T	Time slots in 24 hours
T1	Peak hours
T2	Off-peak hours
$\boldsymbol{\mu A_i}$	A membership function
$\boldsymbol{\bar{w}_n}$	The output of stage 3 of the ANFIS model
$\boldsymbol{w_s}$	Weeks per season
$\boldsymbol{y(t)}$	Auto regression objective function
$\boldsymbol{y'(t)}\&\boldsymbol{y''(t)}$	Level-11 auto regression objective function
$\boldsymbol{\gamma_k Y_{k,t}}$	Extra auto regression term for a year of 4 season

CHAPTER 1

INTRODUCTION

The growing global need for electricity has given rise to an intelligent solution termed a Smart Grid. A Smart Grid system is an intelligent system that can sense and control loads to avoid power outages and are flexible enough to incorporate other energy resources and loads. End users can alter their energy consumption according to their preference of time, price, and significance in various aspects. This approach also enables the customers to minimize their bills, and the grid is capable of rapid outage detection. The Smart Grid covers multiple functionalities, including secure two-way communication of information between the electric suppliers and users; provides measurements via Smart meters [1], [2]. Controllable power generation groups include biomass and hydropower, while wind and solar are categorized in the variable and intermittent power generation groups. Load Forecasting (LF) is especially required when both groups of energy resources are integrated into the utility grid. In the conventional power management system, only the supply side is managed, while it is required to conduct demand-side management via load LF and demand response (DR) management. Smart appliances reduce the residential peaks by alerting consumers to shift the load to low peak periods or by shedding the load automatically [3]. Currently, LF and the implementation of DR are used to improve efficiency, ensure system reliability, and achieve the balance of load and supply. This approach is much better than conventional schemes as it offers improved response time, reduced cost, and fewer environmental hazards and runs smoothly if integrated with distributed systems [4]. It also increases energy efficiency and helps National Electricity Markets (NEMs) grow by incorporating control and communication systems to maintain reliable power systems. However, the physical implementation of the smart grid is facing both technical and financial challenges.

The electrical grid is an enormous, interconnected network that delivers electricity to customers. In recent years the electricity delivery system has gone under many technological transformations, and it couldn't keep its pace with such innovations. The sudden blackout confirms traditional grid inefficiency. Therefore, the future of the traditional grid is the vision of Smart Grid. The idea of SG is widely considered as an up-gradation to electricity generation, transmission, and distribution and is shown in Figure 1.1.

Smart Grid is defined as an electrical system that uses ICTs to collect bidirectional information, i.e., from the generation and Load side, to stabilize the system, enhance efficiency, reliability, and economies. The critical components of the Smart Grid are renewable energy systems (RESs) and Advanced metering infrastructure (AMI). Besides, with the penetration of RES into a microgrid, a small decentralized system is considered a viable trend nowadays. The microgrid operates in two modes, i.e., grid-connected and stand-alone mode. The Smart Grid enables the microgrid to be a prosumer (producer and consumer)[5].

The microgrid is generally divided into two groups, i.e., the integration and its operating mechanisms. The integrated study deals with its implementation as testbeds with the electrical system technology, while the operation mechanism study emphasizes stabilizing the supply and demand. For this, DR in Smart Grid applications is considered [6].

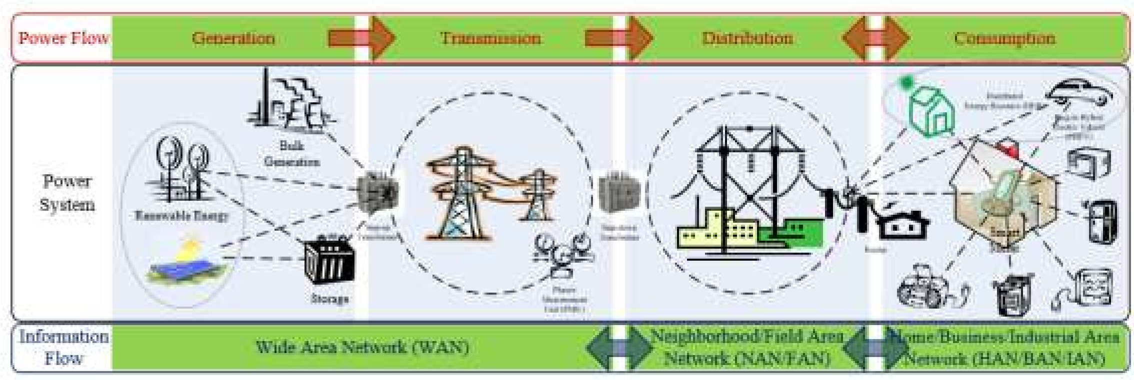

Figure 1.1 *The idea of Smart Grid.*

The provision of continuous supply to meet the growing demand is the main task. Therefore, Smart grid technologies are developed to evolve the traditional grid and mitigate these challenges. PF method plays a significant role in today's updated electricity market as well as Smart Grid operation. The electricity price forecasting (EPF) can be classified into short-term, medium-term, and long-term forecasting. This method of forecasting facilitates every individual generator to control the optimal bidding strategies.

Moreover, a long-time decision of joint agreement and capital investments in advanced generation units depends on the Price forecasting (PF) method. These features make the PF process a difficult task for researchers. Energy Pricing (EP) is used as essential information by electricity providers and end-users. The Smart Grid is now considered a platform that provides an opportunity to end-users and electricity providers to regulate their biding strategies w.r.t demand-side management (DSM) models. Also, the operational stability of the system and capital investments can be enhanced by increasing the rapid response of electricity providers [7]. The PF method offers the robust, stable, and routine operation of power markets.

The electric power system is more than 134 years old. It has functioned with the monopoly as if a single company owns generation transmission and distribution sectors. These are being restructured to operate under a competitive environment. Day by day, the size and complexity of the system have increased as an alternating current transmission system planned at 1100/1200kV generating unit sizes have gone up to 800/1000MW. Numerous schemes have been proposed today to improve energy consumption because the increment in energy consumption increases power demand. Due to this, the load on the existing electric power grid is constantly growing [8].

The smart grid is the future power grid that can increase security, reliability, and control to improve existing electric grid generations, transmissions, and distribution systems. Achieving this requires the LF and PF to play a vital role in the current evolution of the smart grid, and even we cannot do better planning without considering them. The utilities understand the ability of the demand side actions in decreasing the peak electricity load. Peak load reduction is the hour of the need by the electricity sector by using any of the strategies [9]. Conventionally, when demand-side issues are considered, the focus was on more significant industrial users, but they abide by demand-side strategies due to stringent norms. The computation of load and PF is not an easy task due to integrating different power generation units and loads, as illustrated in Figure 1.1.

Furthermore, an increased number of electric appliances led to shortages when used simultaneously. Even the residential and commercial users need to think about their energy usage pattern for a well-functioning electricity market. Moreover, the residential and commercial load demand variations are not taken for observations due to measuring devices. With the advent of new technologies in Smart metering and infrastructure development, supply and demand matching is possible concerning the pricing and different control strategies. Research and practical testing need to be encouraged with both the load and PF. LF will become an essential part of the power system soon. In addition to the smart grid, which faces the technical and financial challenges described above, LF also needs to ensure accurate predictions with minimal Mean Absolute Percentage Error (MAPE). There are two types of concerns in the existing LF model, and the model we propose is better than previous models.

1.1 Background

The current energy market depends on renewable energy resources because these are playing a leading role in fulfilling the energy demand. In rising energy demand, an economical solution is needed to address persistent issues [10]. An exhaustive work concludes that renewable energy generation resources have an extensive impact on energy unit prices. The generation side and demand side companies are in contact to develop a steadiness between supply and demand. An optimized home energy management system proposed in [11] computes the unit price of electricity by utilizing a RES and battery storage system. Energy unit price is determined by correspondence from generation groups with the requirement from load allocation bodies. The subsequent energy demand is calculated from load allocation bodies and fed to the supply side to avoid inconvenience such as electricity shutdown.

On the other hand, the dependability and protection of the electricity system are managed by a regulatory authority to harmonize with local distribution companies. This regulatory authority also determines the energy unit price [12]. The cost-effective solution in any domain is a considerable factor. Recently, energy unit PF has been presented, and diverse performances have also been suggested. As a precipitate, research work is also significant. Numerous areas, including signal processing, statistical modeling, Machine Learning (ML), big-data analysis, and Artificial Intelligence (AI), are anticipated tools [13]. The overview of the latest trends in power systems is presented in Figure 1.2.

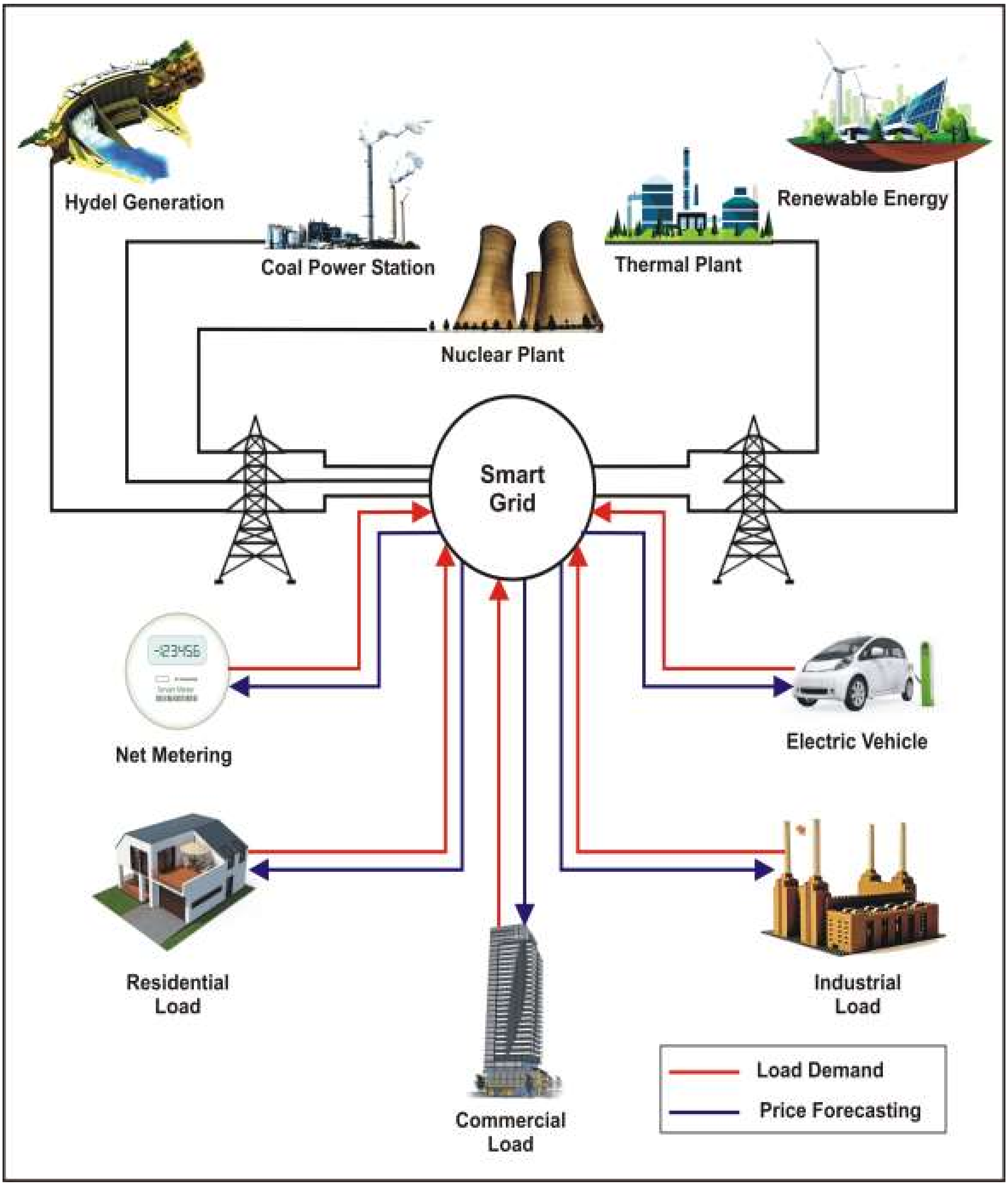

Figure 1.2 *Overview of latest trends in power system.*

Several methods have been proposed to forecast load consumption and per-unit prices, such as engineering and statistical methods. AI applications permeate our lives and become a dominant force in both fundamental approaches and technological advancement by using in large-scale ML, deep learning, reinforcement learning, robotics, computer vision, natural language processing, collaborative systems, crowdsourcing, and human computation, algorithmic game theory, computational social choice, internet of things

(IoT) and neuromorphic computing [14]. Many researchers are working on the AI applications such as Artificial Neural Network (ANN), Support Vector Machine (SVM), evolutionary programming, expert system, fuzzy logic (FL), and LF-related problems. In this regard, many hybrid AI models have been developed to monitor, schedule, and forecast electric load, including neural networks, fuzzy expert systems, fuzzy neural networks (FNN), neural expert systems, neural-GA, and fuzzy expert [15], [16], [17], [18], [19], [20], [21]. On the other hand, the primary attention has been made to ANN undoubtedly as the LF application of ANN was published in the 1980s, and eventually, the applications are growing steadily [22]. AI is a part of many techniques through ANN and SVM [23], [24]. The other two methods, including engineering and statistical approaches, are also applied and used, although they have deficiencies such as complexity, accuracy, and non-flexibility [25]. ANNs have also gained popularity for their simplicity and heftiness. Mainly, it focuses on the back propagation method while it is a slow learning process, and there is no exact rule for over and under fitting [26]. FL is applied to collect historical load data updating the current load by necessary compensation. Later self-organizing feature mapping (SOFM) is used to load profiling. The memory of ANN forecasts variables such as holidays, day type, and weather [27].

Various steps and efforts are proposed in the latest research to meet the large gap between demands and supply. Moreover, electricity cannot be stored in large amounts. The construction of newer power plants is not an ultimate solution as it may lead to carbon emissions and the depletion of resources [28]. Electric utilities have realized that consumer demands cannot be met satisfactorily by adding new generating capacity alone [29]. The concept of DSM is managing energy consumption on the consumer side for energy management and DR, as presented in Figure 1.3. DSM and DR are the terms used for energy management. These deal with actions that influence consumers' energy usage patterns in times of peak load [30]. It also includes the decision-making, implementation, and follow of the activities of utilities to motivate the consumers to modify their level and pattern of electricity consumption according to retail price [31]. Hence, the present focus of researchers is on the domain of energy management.

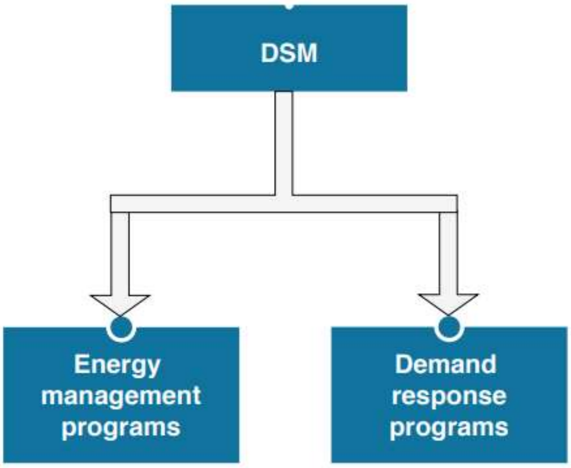

Figure 1.3 *Research strength DSM classification.*

1 MW of power consumption can be considered approximately by 1.4MW-1.6MW of power generation. DSM implementation costs only about 25-30% of the cost of increasing the cost of generation. Stakeholders of DSM are society, utilities, and consumers, which are highly benefitted by DSM activities [12]. Reduction of capital investment, reduced use of fuels in electric vehicles, deferred capacity expansion, better utilization of generating plants, low cost of service, improvement in operating efficiency and flexibility, increased efficiency in the use of existing assets are some of the benefits of the utility side [32]. Conservation of natural resources, reduced environmental degradation, increased economic viability of industry and households, and maximized consumer welfare benefit society.

1.2 Objectives

This thesis has the following four objectives:

1. The first objective is to improve Residential Electricity LF using a Machine-Learning-Based Feature Selection Approach and a Proposed Integration Strategy.
2. This research aims to improve residential EPF using an ML-based feature selection approach and a proposed integration strategy.
3. The third objective is to optimize DSM using FA, where LF and PF data are considered, where per unit cost and load demand of different buildings have to be examined.

1.3 Thesis Contributions

The contributions are well accomplished and summarized as:

In the first part of the thesis, data of 10 houses from the Pakistan Residential Electricity Consumption Dataset (PRECON) has considered where the LF has optimized by reducing the MAPE. Time series and auto-regression algorithms have been developed to fetch the data set and set the objective function, while; proposed equations set the base for further validation verified by improved results. Similarly, a feed-forward Adaptive Network-based Fuzzy Inference System (ANFIS) based on the ML approach is proposed, which is better than the previous one. The binary genetic algorithm-principal component analysis (BGA-PCA) approach for attaining the best feature selection is evaluated for further improvement. MAPE calculations for the integration strategy are formulated for individual building and obtained remarkable improvement. Finally, our results validate the proposed model by providing 17% overall system improvement compared to past work. Moreover, the most optimized value of MAPE is obtained for the four seasons.

Similarly, in the second part, the data of various power producers, CPPs, PPPs, Tariff rates, and load demand from NEPRA is considered for MAPE reduction in PF. Eight-time series and auto-regression algorithms are developed for data fetching and setting the objective function. The feed-forward ANFIS is based on the ML approach to forecasting the PF by taking the inputs from time series and auto-regression algorithms. Best feature selection is made by adopting the BGA-PCA approach. It reduces the repeated, irrelevant and unnecessary data and ultimately minimizes the model's complexity and computational time. The proposed integration strategy computes the MAPE according to the steps mentioned above and obtains significant improvements.

In the third part, demand-side management based on the FA has been presented. The peak load exerts extra stress on the grid; therefore, to minimize peak load, the FA has been used to reschedule load timing so that the Peak to Average Ratio (PAR) is minimized. The proposed algorithm has chosen the load of three residential buildings forecasted in Chapter 3, containing different load types. One of the buildings was given an industrial/commercial load to verify the system's ability to handle different conditions. As a result, the proposed system has efficiently reduced the PAR by LF, and the cost of electricity is reduced by PF (Chapter 4).

1.4 Used Environment

This complete model is simulated in a MATLAB simulation environment on a research workstation with Intel Core i5-4300CU, 8GB RAM, 64-bit operating system, and 2.50 GHz Processor. The proposed model of 10 houses with a one-year hourly database is simulated in 50 min. A general overview of LF, PF, and power management systems and researchers' latest work is briefly introduced in this chapter. Overall thesis contribution and highlights of each chapter are also discussed.

1.5 Structure of Dissertation

In Chapter 2, a survey of related work is carried out by reviewing the past work detail. The key feature of the previous work is summarized in tabular form. The latest trends and different methods used for the load and PF are reviewed critically for better understandings. The power management system techniques are also critically analyzed, and crucial factors are elaborated as well. In the end, the latest articles related to LF, PF, and PMS are summarized.

Chapter 3 of this thesis is based on the framework for LF's time series and auto-regression algorithms. In the second part of the chapter, the model is further validated by the ANFIS technique based on ML. The third part is a novel feature selection model based on BGA-PCA algorithms. The fourth part validates the model by the proposed integration strategy. The last two sections are the model validation, results in discussion, and conclusion, respectively.

In Chapter 4, the same method as Chapter 3 for fetching the data set of PF by adding space vector regression is utilized. Similarly, this model is also further validated by ANFIS approach in the next part. The feature selection is also utilized using BGA-PCA algorithms, same as in chapter 3. The proposed integration strategy obtains the final MAPE. In the end, the validation of the model results in discussion, and the conclusion is briefly explained.

In Chapter 5, the power management system is modeled using forecasted data of load and price. A day ahead load shifting technique is formulated mathematically for optimization of the power system. The evolutionary FA is implemented to solve the DSM algorithm. In the next part, three buildings have been selected for the PMS having different types of load. Forecasted values of load and price are used as input parameters for the proposed system. In the last part, the simulation results and conclusion are discussed.

In Chapter 6, the conclusion of the discussion is drawn from this research, and extended future work is proposed.

CHAPTER 2
LITERATURE REVIEW

2.1 Introduction

The introduction of a Smart Grid requires a more detailed load modeling in which the behavior of individual appliances should also be considered [33], [34], [35]. Load modeling and power scheduling for DR has become more technical art than first science principles because of the vast complexity and variety of features that define final-use electrical system components [36]. The services are planned to endure peak demand times with demand growth forecasts from community distribution lines, via the semi-voltage range feeders and channels, to the large electrical transmission network at the different levels of the supply system [13], [37]. Simplifying observations and deterministic forecasts based on historical calculation instead of high precision simulations, regularly tuned to specific system circumstances, represent the expectations about demand variability and association with system voltage and frequency shifts. For obtaining better results under performance constraints, profit margins are then imposed on such designs. Usually, the load is handled similar to climate, where historical evidence is evaluated to associate leading factors with actual behavior, such as the moment of the day or time of year, weather, etc. [34], [38], [39], [40]. The essence of preparing load models needs to shift when the load starts to engage in system activities using knowledge sharing with process and business providers. Distribution circuit designers ought to determine with load involvement what end-use services would respond to cut peak flows so that capacity limits are not broken. Current load models need to represent specific devices and electrical systems turning on and off to adapt appropriately to voltage and frequency requirements. Once those signals are withdrawn, they need to show the expected reaction to signals that support reduction and describe how the load reacts. Loads needed to drop out for a duration of time, for instance, will produce a recovery in energy demand if they are not adequately handled, just like production ramps [41], [42], [43]. This condition will minimize the load behavior individuality or unpredictability and add a cyclic component to the load curve that can take a bit of time to go down. To achieve a clearer view of the overall dynamics of an area, simulations that represent particular devices are required because they need to be optimized by using an optimal load schedule algorithm for actual data. The modeling issue is compounded when attempting to anticipate improvements in end-use schemes. In addition to automation for residential and industrial buildings, improvements will continue to be implemented with modern loads, like large, flat-screen displays, hybrid cars, and LED lighting that saves energy efficiently [19]. Moreover, hybrid energy production and energy systems, e.g., batteries and thermal energy, would become widely available. The rate of absorption of these modifications and national and regional exposure to the product offering for such modifications greatly complicates the issue of the delivery plan [44]. Decision-makers ought to be part of the energy decision phase more than ever before. Policymakers and administrators need to realize the economic and effective implications of these actions directly and quickly. Transmission stage design is based on the transmission feeders' simpler, consolidated load models [45]. Generally, it is also presumed lumped inputs for devices susceptible to amplitude and current and a smoothing factor induced by the range of equipment going up and down. A distribution network designer has been a core component of the framework of energy regulation. Still, now the consequences of these new features must be explicitly articulated to policymakers and authorities [46]. The versatility in modeling multiple load amplification strategies would be an essential feature of load modeling. Retail arms of owned services, municipal utilities, or cooperatives are the large proportion of modern demand aggregators. Any major commercial clients, however, are acquiring their control from other aggregators. The aggregated load will be distributed across a wide region in this situation, likely through multiple delivery networks and operational agencies of the power system [47], [48]. Two characteristics of the consolidated load are relevant from a scheduling point of view: the volume and its position [49]. Although aggregators would guarantee production for their consumers in marketplaces, they may need to work with electricity distribution providers due to their organizational limitations. It is important to coordinate consolidated, final-use involvement to apply to multiple markets and support transmission networks. A crucial aspect for the efficient deployment of the network is system LF [50]. Network controllers using ML approaches can estimate the energy requirements with great precision with decades of work expertise and assisted by advanced load prediction systems [51]. Analysts forecast, log, and adjust hourly needs and peaks using measures such as temperature, climate, day of the week, and vacations [52], [53], [54], [55]. Methods for calculating the spectrum of load demand for the analysis period of concern have also been developed by timely information. When load appears something of a participant in system processes, these projections remain essential to operators. However, current approaches still need to be reconfigured as huge load chunks can dynamically adapt to operational conditions. As distribution and transmission network administrators need to place the supply infrastructure to handle alternative energy sources and a flexible load curve efficiently, the number of unknown parameters continues to increase [56]. Wisely incorporated into system processes, load intervention becomes a new weapon for system designers to reach healthy margins while maintaining productivity powers and reducing environmental effects.

2.2 Related Works

The forecasting medium supports establishing system models, behavior, and plan by reviewing current and past market trends. Load and PF are the main factors of this medium, demonstrating generation capabilities, resource handling, capital cost, profit analysis, and other system plans [57]. In these viable power energy souks, load estimation and price prognoses are significant factors for prime maneuver development. Diverse methodologies exist for price and load estimation, but without feature selection, skill methodologies are inadequate. The mentioned feature selection skill deals with the modeling of intermingling structures and nonlinearities of forecast progressions. By considering technicalities and generation site plans, the electricity price estimation is a more valuable chore. The energy rate is estimated with an optimistic approach based on market-side demand and predictive forecasting. Renewable energy backs the economical energy unit price. Detailed scrutiny of market data in an arithmetical way is a fundamental requirement in this approach [58].

The current energy market is much dependent on renewable energy resources. PV and wind are widely used and playing a leading role to fulfill the energy demand. In rising energy demand, an economical solution is acceptable. An exhaustive work concludes that renewable energy generation resources have an extensive impact on energy unit prices. The generation side and demand side companies are in contact to develop a steadiness between supply and demand. Energy unit price is determined by correspondence from generation groups with the requirement from load allocation bodies. The next day energy demand is calculated from load allocation bodies and fed to the supply side to avoid inconvenience like shutdown. On the other side, for a deregulated energy marketplace, dependability and protection of the electricity system are managed by a regulatory authority established to harmonize with local broadcast companies. This regulatory authority also determines the energy unit price [59].

The cost-effective solution in any domain is a considerable factor. Recently, energy unit PF fascinated too much worth, and diverse performances suggested. Therefore, numerous areas, including signal processing, statistical modeling, ML, big-data analysis, and AI, are anticipated tools [60].

The consumer side is classified into two classes: less power consumption class (LPCCs) and high power consumption class (HPCCs). These consumers consume electricity according to their class. The heavy electric operated appliances required more power like Microwave Oven, Iron, etc. The demand side peak load foundation is built for heavy electric appliances operate from HPCCs. Beforehand, to calculate the price, altered schemes, including Consumption Time Scheme (CTS) and actual-time pricing scheme (ATPS), were proposed. In CTS, the calculation of price concerns consumption time. The unit price is maximum for peak energy consumption time, the unit price is minimum for off-peak energy consumption time, and the moderate price for mid energy consumption time. The calculated energy price reasonable for all residential, commercial, and industrial users as well. Considering the fact, HPCCs consume more power in peak time hours and responsible for peak time load curve formation. By laws, all maximum price units must be charged from HPCCs, but the current billing system is not working according to this approach . LPCCs claim this irregularity in the billing system. The unfair price calculating scheme CTS in the billing system is notified in this work. Illegally this billing system exerts a financial burden for LPCCs. To avoid this irregularity, an early load measuring-based mechanism was designed to calculate next-day load side demand. From this mechanism, the LF data sheet supports the calculation of the actual user load consumption during peak and off-peak hours for LPCCs and HPCCs. As a result, we will calculate the accurate billing cost for both users according to the consumption [61].

2.2.1 Load Forecasting

The introduction of a smart grid requires a more detailed load modeling in which the behavior of individual appliances should also be considered [62]. Household electricity consumption in the US was analyzed, and it is found that air-conditioning, water, and space heating comprise 66% of energy consumption [63], [64]. Residential modeling of load profiles can be top-down, bottom-up, or hybrid [65]. The top-down model considers the residential load a large energy pool in which individual household consumption is not considered. In a bottom-up approach, data from each appliance is considered while the hybrid model combines both the features of top-down and bottom-up [66]. It is investigated in [67] that the top-down model uses past measures and doesn't take future changes in load; therefore, it is suitable for finding the supply side requirements. Load modeling is done by taking aggregated data from the meters of residential, commercial, and industrial users [68], [69], [70]. Response of the customers and grid reliability are the critical factors in achieving advantages of a smart grid, which results in spurring the effective decision by the supply providers and end-users. Users can effectively utilize and save electricity with the use of demand-side management [41], [71], [72], [73].

A hybrid model can distribute the electrical load into two parts, where the first is a scaled curve load having five ANNs and the second is a day maxima and minima having both the FL and ANNs for LF. The benefit of the proposed hybrid structure was to use both FL and ANNs for uncertainty handling [74], [75]. In [76], a hybrid approach is proposed for daily LF using FNN and an improved GA. The GA has been used to compute the optimal policy of fuzzy rules, and FL has been used to deal with variable linguistic information in LF. This approach has reduced the everyday problems of convergence in maxima and minima methods to initial values. An expert system known as LoFy was developed for short time load forecasting (STLF), and it has three models, including daily, weekly, and special days for forecasting.

Moreover, it was concluded that no technique could be evaluated for all types of days [77]. Another AI technique, multilayer perceptron (MLP), has also been implemented for STLF. It includes a data mining methodology to construct rules for STLF. The optimal tree regression is also helpful in defining a relation between input and output variables. It is also helpful for the classification of input data into clusters for pre-filtering. Hence MLP is easy to handle because of the similarity of classified input data [78], [79]. ANN models can also be merged with time-series models for better policy development of hourly loads for all types of days. Two parts are included in this approach. One technique is based on correlation techniques for the selection of input variables and training sets. Moreover, the second technique was used for weekdays and one for the weekend and obtained an improved MAPE [80].

SVMs are the most recent regression and classification problem analysis [81], [82], [83], [84]. In [35], the Statistical Learning theory approach has been used to solve the same problem. It is different from neural networks in their nonlinear mapping of data and uses simple linear functions to develop linear decision boundaries in the new space. Therefore, due to its advancements, SVMs are also implemented for short-term scheduling [85]. Its performance, when; compared with the autoregressive methods, indicates that SVM has better forecasting. It gives the result by considering past and current data points while ignoring other influential elements. The daily load demand of the month can also be computed using SVM. This idea was also encouraged by the European Network on Intelligent Technologies (EU-NITE) network [86].

A Multi-Layer Neural Network (MLNN) model is presented in [87] for LF while the computational time is very high. A model of Long Short-Term Memory (LSTM) and RNN in [88] evaluated the accuracy rate by labeling the model's name of Gated Recurrent Units (GRU). The load is forecasted in [89] by a hybrid technique of three different schemes as Cuckoo Search, Singular Spectrum Analysis, and SVM to increase the accuracy in forecasting. Despite the hybrid model, it cannot also reduce the computational time as well. The models mentioned earlier forecast the load in an optimized manner but originated a problem of reducing the computational time, data storage setup, and irrelevant data. In this regard, feature selection and feature extraction techniques have been presented in recent years. The authors in [90] computed the accuracy by data preprocessing through feature selection and feature extraction while the authors removed the irrelevant data to predict accurate forecasting.

GA is a commonly used algorithm for feature selection as a hybrid model of Ant Colony Optimization (ACO), and GA for LF is presented in [91], [92]. ANFIS is used for the best prediction. In [93], a combination of GA and support vector regression (SVR) is proposed for the LF of hotels while GA is deployed for feature selection and SVR to optimize LF. Similarly, some other GA-based schemes [94] for LF of industries and health care centers are also carried out where the feature selection approach shows accuracy in the prediction of LF. The trend of integration strategy becomes more valuable due to its high accuracy rate of LF. The authors proposed an integrated strategy to forecast the power consumption and the model hybridized with GA-ANFIS for LF. Similarly, the integration methods proposed for the combined decision of different schemes together and finalized the final decision [95]. The highlights of the related work are summarized in Table 2.1.

Table 2.1 A comparison of related work of LF.

Work	Key Contribution	AI Approach	FS	Limitations
[27]	LF	Fuzzy logic and ANNs	NO	They have deficiencies, including complexity and non-flexibility
[28]	Peak LF	Adaptive backpropagation learning-based ANN	NO	The slow learning process and no exact rule for over and underfitting
[30]	LF	Fuzzy logic and ANNs	NO	Mainly it focuses on the load curve, not on MAPE calculation
[31]	GA based daily LF	NFN	NO	Only deals with variable linguistic information in LF
[32]	STLF one for weekdays and weekend	ANNs	NO	No technique can be evaluated for all types of days
[33]	Time-series models for policy development of hourly loads	ANN	NO	Very complex as no technique was used for eliminating irrelevant data
[35]	Regression and classification problem analysis for LF	SVM	NO	Only computed STLF
[37]	Daily LF of the month	EU-NITE network	NO	Only the daily load demand of the month is computed
[87]	LF of Wind Power	MLNN	NO	Computational time is very high.
[88]	Price Forecast by GRU	RNN	NO	Not reduced the irrelevant data.
[89]	STLF	SVM	NO	Computational time is very high
[90]	Load and Price Forecast	NO	Yes	Low computational time but fewer inputs are taken into account.
[91]	Hourly LF	ANFIS	Yes	Comparatively high MAPE
[93]	Hotel room LF	SVM	Yes	The model is less complex, but fewer inputs are considered.
[94]	Industrial LF	Radius Basis Function Neural Network	Yes	Not work well for large datasets
[95]	Survey on an integrated strategy to LF	ANFIS	Yes	Only suitable of the given scenario
Proposed Work	Residential Electricity LF with a novel Proposed Integration Strategy	ANFIS	Novel BGA-PCA	A new variety of data input and complex auto-regression methods are not analyzed.

2.2.2 Price Forecasting

Many researchers forecasted the price of electricity using ANN strategy in traditional ways. In this regard, a multi-layer feed-forward Neural Network (MFFNN), Neuro-Fuzzy, Fuzzy Neural Network, and Adaptive Wavelet Neural Network approaches based on the different analysis used to extract selection features of price and load signals in each hour [39], [57], [61], [96], [97], [98], [99], [100], [101], [102]. The electricity price varies due to dynamic variations in load demand throughout the day; therefore, this strategy minimizes the MAPE in load and price on hourly, daily, and weekly intervals [103]. With the dynamic variations in load containing high-frequency features, an error is considerably increased; therefore, this research recommends Wavelet Transform to resolve this issue.

Similarly, an order book-based forecasting by implementing ML-based ANN is introduced in [58]. It employed the Austrian electricity spot market prices using bidding methodology and order book data. For optimization, it selects appropriate, limited features by computing the actual and expected demand of the power. The fundamental principle is to conduct an auction for electricity daily by presenting the supply and demand curve in a specific period. Another approach Hybrid Wavelet ANFIS Approach used in [104], formulates the relation between the purchase and demand of electric power and then similarly describes the relation between supply and sold electricity. The forecasted price of electricity almost tallies with the generation cost of various power plants. Although it extracts significant results and better performance than linear regression methods due to limited feature selection, the proposed algorithm does not contain any statistical data, optimized model, and demand curves.

In [90], a methodology based on feature selection with mathematical and statistical modeling is proposed, utilizing the minimum block set from the input and applying a computational filtering strategy to formulate optimized PF. A Real-Coded Genetic Algorithm (RCGA) is implemented to adjust appropriate filter settings for achieving the required forecasted price of electricity. The inclusion of theoretical statistics of interaction gain in the proposed methodology increases the significant efficiency of the system to process PF of electrical power. It continuously visualizes the relevancy and redundancy of the selected feature, but it has a limited environment for feature selection to minimize the validation error occurred by wrappers.

Various parameters and features should be taken into account while optimizing the load and PF of the electrical power system. For this purpose, an approach based on ANN utilizes input parameters for tuning the load and price of the electricity [105], [106] adequately. While predicting these values, the algorithm processes tuned parameters to make decisions by collecting required information. The proposed methodology collects a substantial amount of information to adjust appropriate features using Recurrent Neural Network (RNN). After selecting the suitable parameter forecasting engine processes, it is to declare the estimated data set in terms of load and price. An algorithm based on fuzzy set theory is implemented on a distribution network system to optimize the cost of uncertain load growth [105]. Although, it presented a better response with increased accuracy in load and PF of electricity. However, it contains nominal error rate deficiency concerning other systems.

The research illustrated in [106] describes a trained RNN algorithm based on Kalman Filter (KF) to forecast the load and price of electricity using one-step and n^{th} step prediction. The proposed study practiced informative data of the European Power System (EPS) to validate the execution of the scheme. Using the time series-based multilayer ANN-KF model increases forecasting speed and tracking behavior, but the implementation of KF in the ANN algorithm relatively rigid than traditional strategies.

Binary Genetic Algorithm VS Kalman Filter Algorithm:

Binary genetic algorithm is cast-off to choose significant forecasters that knowingly impact the load pattern between a numeral of input variable quantity, but it will not update. On the other side in the Kalman filter algorithm comprises of two stages: forecast and update. In all the previous work these terms are also known as “Propagation” and Correction”. The Kalman filter algorithm is summarized as follows:

Prediction:

The Predicted state estimation is given as

$$\check{x}^{-} = F * \check{x}^{+}(k-1) + B * u(k-1)$$

The Predicted error covariance can by found by

$$P(K)^{-} = F * \check{P}^{+}(k-1) * F^{T} + Q$$

Update:

For the Measurement residual we will use this equation

$$\widetilde{\boldsymbol{Y}(\boldsymbol{k})} = \boldsymbol{Z}(\boldsymbol{K}) - \boldsymbol{H}\breve{x}^{-}(k)$$

The Gain of the Kalman algorithm is given as

$$\boldsymbol{K}(\boldsymbol{k}) = P(K)^{-}\, H^{T}(R + HP(K)^{-}\, H^{T})^{-1}$$

The state estimate can be updated by this

$$\breve{x}^{+} = \breve{x}^{-} + K(k)\widetilde{\boldsymbol{Y}}$$

The error covariance after the update become this

$$P(K)^{+} \; = (1 - K(k)H)P(K)^{-}$$

The estimation of a variable is denoted by the hat operator (^). The superscript (-) will show the Prior prediction and the (+) show the Updated version. The term P will use to encrypts the error covariance.

While in the updating stage, Y(k) is coming first. Innovation is the measurement residual because it is the difference between the predicted value and the actual value. The correction can be found with the help of the product of Gain and the residual. After that the updated error covariance can be found. The predicted covariance is greater than the updated error covariance.

Integration of wind power in traditional energy sources reduces the electricity generation cost and takes part in electricity consumers' well-being. A real-time PF bivariate distribution scheme is implemented on the wind power system to minimize forecasting errors [107]. While the correlation factor remains negative, the suggested price of this distribution is less than the marginal distribution; however, it provides more efficient estimation under the conditional probability of forecasting fault slips.

In previous researches, the AI technique is implemented for the long-term load and PF of electricity conventionally. In [108], [109], [110], [111], [112], [113], a short-term load and PF using NN fitting tool to compute hourly and daily basis data of weather temperature and electricity load as input features. Furthermore, generalized-RNN indicates temperature statistics and price signals as input parameters. Both techniques estimate accurately prescribed variables, but the integration of Radial Basic Function-Neural Network (RBF-NN) and GRNN have mutual effects that declare some computational errors. A comparative analysis of different AI approaches compiled in Table 2.2, including their limitations.

Table 2.2 *A comparison of related work of PF.*

Work	Key Contribution	AI Approach	FS	Limitations
[57]	Price and LF	MFFNN	No	The dynamic variations in load contain high-frequency features which considerably increases error
[58]	PF	ML based ANN	No	Due to limited feature selections it does not contains any statistical data, optimized model, and demand curves.
[90]	Load and Price Forecast	RCGA	Yes	Partial environment for feature selection to minimize the validation error occurred by wrappers
[114]	Price and LF	ANN-RNN	Yes	Nominal error rate deficiency occurs in other systems
[106]	PF	RNN-KF	No	The implementation of KF in the ANN algorithm relatively rigid than traditional strategies
[107]	PF	Bivariate Distribution Scheme	No	At negative correlation factor, the forecasted price of this distribution is less than the marginal distribution
[108], [109]	Price and LF	ANN-RNN	No	The integration of RBFNN and GRNN have mutual effects which declare some sort of computational errors
Proposed Work	PF with a novel Proposed Modified Strategy	ANFIS-SVR	Novel BGA-PCA	A new variety of data input and complex auto-regression methods are not analyzed.

2.2.3 Power Management System

Various techniques presented in [11], [115], [116], [117] presented a lot of techniques regarding DSM used in Smart Grid and also studied the features of techniques, its benefits, and architecture. The use of many appliances in homes is increasing; therefore, residential users have a critical role in improving appliance scheduling strategies. Consumers need flexibility and awareness in these strategies. Improving the demand for electrical energy for appliances in homes is a challenge that faces both consumers and utilities, especially when the consumption of electric power is high during peak loads. The purpose of this thesis is to provide an overview of optimization techniques for residential scheduling devices. There are two traditional techniques: metaheuristic algorithm and heuristic - algorithms reviewed in this study.

In [13] proposes a user-aware DR approach to manage residential loads. User comfort and savings are considered mainly, and user comfort is modeled as a weight factor before comfort over savings. The game is based on a modified regret matching procedure which has the advantage of centralized and decentralized schemes. The work can be extended with multiple energy resources in the future.

Two optimization techniques, such as the FA and harmony search algorithm, are presented in this paper. These heuristic techniques are applied to different types of appliances based on their energy consumption. Single and multiple home appliances are considered in this research. ToU is used as a pricing signal in the proposed technique to calculate electrical cost [118].

This paper examines the performance of EMC using met–heuristic algorithms. The scheme used to calculate the electrical cost in this proposed system is CPP. Different appliances based on power consumption are categorized into three groups. This article minimizes the cost of electrical energy and peak load shift from peak time to off-peak time. The simulation result shows that the average ratio of total electricity cost to the peak has decreased to a certain level. In addition, the harmony search algorithm (HSA) performs better than FA in terms of PAR and electricity costs [119].

This article focuses on maximizing the use of renewable energy resources. Three different issues, namely without the battery, with the battery, and the cost of its life, are considered and analyzed. Different heuristic techniques, including FA and particle swarm optimization (PSO), are used to establish optimal processing of wind, battery, hydro based micro grid [120].

Farahani [121] suggested algorithms of various types to improve classic FA performance. The FA in first-class used a learning automaton to absorb and adapt random parameters. The second class was to balance the exploration and exploitation features of the proposed meta-heuristic over time, for which the GA was hybridized with the FA. The FA in the third class used a random walk based on Gauss's division to move fireflies to the search area. The results obtain from experiments show that the algorithms proposed in this research were highly applicable with the PSO algorithm and classical FA.

In the previous research, the FA was used to resolve cost-effective sending and has managed to find a more accurate solution than another meta-heuristic algorithm. A swarm-based FA is a metaheuristic with a higher conversion rate and shorter implementation duration than other metaheuristic techniques while solving economic load delivery. Given that the FA lacks local optimism. Many researchers have suggested improvements and improve the FA performance to achieve faster and more efficient global solutions. In this article, three of these recent additions were adopted to address the economic remittance of sic generating units [122].

In [123], a new algorithm introduces this study to solve and reduce the electricity cost problem using two metaheuristics techniques FA and Elephant herding optimization (EHO). A scheduling process is used as a HEMC to maintain the balance between the demand and load side. This study aims to determine the low cost by considering the maximum factors of user comfort and the average-to-peak ratio. The scheduling process and performance are evaluated by the comparative analysis of these two meta-heuristic techniques.

A comprehensive overview of this living and developed the discipline of swarm intelligence to show that the FA can be applied to virtually every problem that arises. On the other hand, it encourages new researchers and algorithm developers to use this simple yet highly efficient algorithm to solve problems. The primary purpose of this research is the successful implementation of FA in different areas and thus the widening scope of its potential users [124]. Table 2.3 presents a comparative analysis of different methodologies related to DSM.

Table 2.3 *Comparative Analysis of related work to DSM.*

Work	Key Contribution	Limitations
[7]	The techniques of scheduling residential appliances are presented in this study. This study is classified into two techniques such as meta-heuristic techniques and heuristic techniques.	With the widespread use of home appliances, residential consumers need to improve the strategies for scheduling appliances.
[11]	This paper presents DSM architecture models and algorithms based on customers behavior and smart appliances integration	Implementations need to be done
[13]	Proposes autonomous Game theory DR systems to minimize the cost of power generation.	Game theory can be an alternate to DSM
[121]	This study aims to consider the factor of maximizing user comfort, peak-to-average ratio (PAR), and the low cost by considering firefly optimization (FF).	Both algorithms performed a comparative analysis of the separately implemented version to evaluate the process of scheduling devices and improve their performance.
[122]	A swarm FA implemented except meta-heuristic algorithm for a cost-effective solution	The algorithm lacks local optimism
[123]	Two meta-heuristics techniques FA and EHO is utilized for the scheduling process in a HEMC to maintain the balance between the demand and load side	Due to the integration of two meta-heuristics techniques, the processing rate is comparatively reasonable, but the scheduling rate is low.
[124]	A comprehensive overview of the swarm intelligence of FA is performed to express performance capabilities to encourage new researchers and developers to implement it in future works.	The implementation of FA is simple and highly efficient in solving problems.
[125]	The scheme used to calculate the electrical cost in this proposed system is CPP using a Meta-heuristic algorithm	HSA perform better than FA in terms of PAR and electricity costs
[126]	Applied on different types of appliances in single or multiple homes based on their energy consumption using FA and harmony search algorithm	ToU is used as a pricing signal in the proposed technique for the calculation of the electrical cost

[127]	The proposed strategy reduces peak load and the cost by compromising the comfort of the household consumer. Optimization is performed using a multi-objective GA	Decreases the peak load and also reduces the utility bill
[128]	This article focuses on maximizing the use of renewable energy resources using different heuristic techniques	Provide consistency in daily profit and low runs.
[129]	A concessional and maximum energy consumption schedule is focused on using the GA to reduce and production costs and improve the efficiency factor at the utility and consumer level.	It is used over the single objective, which benefits consumers with intelligent load scheduling
[130]	The author compares the performance of the HEMC, and this controller is designed for energy consumption scheduled based on the heuristic algorithm.	The heuristic technique, such as the GA-based energy management controller, performs more efficiently than other heuristic techniques.

2.3 Summary

In this chapter, various research is reviewed critically to analyze the computational abilities of different AIs regarding load and PF. In this regard, many ANN methodologies are studied, such as GA, NFN, NFN for a short time, long time, and hourly basis LF. Similarly, many kinds of research are assessed regarding the PF using ANN strategy with MFFNN, Neuro-Fuzzy, FNN, and Adaptive Wavelet Neural Network approach based on the different analysis used to extract selection features price and load signals in each hour. Similarly, the power management system is modeled using forecasted data of load and price by researchers. A long-term, very long-term, short-term, very short-term load shifting techniques are also the part of the literature.

CHAPTER 3
AN IMPROVED RESIDENTIAL ELECTRICITY MACHINE LEARNING-BASED LOAD FORECASTING

3.1 Introduction

The first concern of existing models is that the feature selection approach is not employed for analysis, as shown in Table 1. Furthermore, the mentioned methods have the problems of high computational time, complex data, repeated data, and issues with extracting relevant data from a massive quantity of data. In our proposed model (BGA-PCA), historical data of four seasons based on different inputs for 10 houses were exploited, whereas the duplicated, irrelevant, and unimportant data were removed without affecting the information. This contribution reduces the computational time and makes it less complex than the other existing models that do not employ the feature selection approach.

The second concern is that the different models existing in the literature have drawbacks under diverse conditions, as demonstrated in Table 1. Our proposed integration strategy aims to improve the success rate by using a combined strategy of time series auto-regression, the ANFIS model, and the BGA-PCA model in a single collaborative method. In our model, the final decision is not made autonomously considering all algorithms; only the best model is considered, based on the historical data for MAPE calculation of each season.

We investigated that some researchers have worked on a GA-based feature selection approach to reduce complexity and computational time. To the best of our knowledge, the proposed hybrid model of BGA-PCA for feature selection purposes has never been deployed yet in electricity LF. Further improvement in MAPE calculation is made by integration strategy where the optimistic integration model takes maximum inputs compared with other models that have limitations (demonstrated in Table 1) of fewer inputs. Its a novel attempt to integrate a time series auto-regression algorithm, ANFIS based algorithm, and BGA-PCA algorithms for optimal MAPE calculation in LF.

This research work presents a novel approach for residential LF using an ML-based feature selection approach and proposed integration strategy. Most of the articles in the literature have improved the accuracy rate of MAPE using AI approaches, including ANN, SVM, ANFIS, evolutionary programming, expert system, fuzzy logic, and LF-related problems, etc. These approaches have the limitation of high computational time and complexity as it includes a high amount of data. So, reducing irrelevant and repeated data without affecting the essential information is a vital step in LF. On the other hand, the research work on feature selection-based GA algorithms has become the latest trend in LF. In this regard, we proposed a hybrid model of BGA-PCA for feature selection purposes, and it has never been deployed in electricity LF. The other novelty of our research work is an integration of a time series auto-regression algorithm, ANFIS based algorithm, and BGA-PCA algorithms for optimal MAPE calculation in LF. This model not only reduces the unnecessary inputs through ML-based feature selection (BGA-PCA) and also improved a remarkable transformation in reducing MAPE.

3.2 Proposed Framework

This section is based on the proposed LF techniques and framework such as:

3.2.1 Time Series and Auto-Regression Method

The time series forecasting models have a standard approach of collecting historical data for predicting load, and nine different regression algorithms are taken in [131]. In our model, we have computed seven estimation models and three auto-regression models with one exponential smoothing. The auto-regression models are further estimated at level-II with exponential smoothing. The algorithm with symbolic representation is elaborated in Table 3.1, and parameters that have been taken into account are defined in Table 3.2.

Table 3.1 *Symbolic representation.*

Symbols	Description
$y(t)$	Auto-regression objective function
$y'(t)\&y''(t)$	Level-II Auto-regression objective function
P_{t-1}	Previous period load demand
F_{t-1}	LF demand
$\gamma_k Y_{k,t}$	Extra auto-regression term for a year of 4 season
fit_{fun}	Feature selection function
$a_i\&c_i$	Set of premise parameters

***Table 3.2** Time series and auto-regression algorithm.*

Models	Formulations	Description of Variable
Level I Auto-regression Model (AR1)	$y_t = \beta_0 + \beta_1 y_{t-1} + \beta_2 y_{t-2} + \beta_3 y_{t-3} + u_t$	Variation in Weekly load demand: Special holidays or weekend $y(t-1)$, peak hour $y(t-2)$, off-peak hour $y(t-3)$
Level I Auto-regression Model (AR2)	$y_t = \beta_0 + \beta_1 y_{t-1} + \beta_2 y_{t-2} + \beta_3 y_{t-3} + \sum_{k=1}^{4} \alpha_k S_{k,t} + u_t$	AR2 is modeled by adding season auto-regression term $S_{k,t}$where k = 1,…,4. Pakistan has four seasons such as summer, winter, fall, and spring.
Level I Auto-regression Model (AR3)	$y_t = \beta_0 + \beta_1 y_{t-1} + \beta_2 y_{t-2} + \beta_3 y_{t-3} + \gamma_k Y_{k,t} + \sum_{k=1}^{4} \alpha_k \bar{S}_{k,t} + u_t$	AR3 model has an extra auto-regression term for a year $\gamma_k Y_{k,t}$ and for the seasons of the year $\bar{S}_{k,t}$ where k = 1,…,4.
Exponential Smoothing Method	$y_t = \alpha P_{t-1} + (1-\alpha) F_{t-1} + u_t$	This model is used for the purpose where there is no seasonality in the load. Where α = smooth constant, P_{t-1}= previous period load demand, and F_{t-1}= previous period LF demand
Level II Auto-regression Model (AR1)	$y'_t = \alpha P_{t-1} + (1-\alpha) F_{t-1} + u_t$ $y''_t = y_t = \beta_0 + \beta_1 y_{t-1} + \beta_2 y_{t-2} + u_t$ $y_t = y'_t - y''_t + u_t$	In level-II, the model is applied on the data of the same series in AR1 but the demand values and residential values from customer demand modeled in the Exponential Smoothing Method are added.
Level II Auto-regression Model (AR2)	$y'_t = \alpha P_{t-1} + (1-\alpha) F_{t-1} + u_t$ $y''_t = \beta_0 + \beta_1 y_{t-1} + \beta_2 y_{t-2} + \beta_3 y_{t-3} + \sum_{k=1}^{4} \alpha_k S_{k,t} + u_t$ $y_t = y'_t - y''_t + u_t$	In level-II, the model is applied on the data of the same series in AR2 but the demand values and residential values from customer demand modeled in the Exponential Smoothing Method are added.
Level II Auto-regression Model (AR3)	$y'_t = \alpha P_{t-1} + (1-\alpha) F_{t-1} + u_t$ $y''_t = \beta_0 + \beta_1 y_{t-1} + \beta_2 y_{t-2} + \beta_3 y_{t-3} + \gamma_k Y_{k,t} + \sum_{k=1}^{4} \alpha_k \bar{S}_{k,t} + u_t$ $y_t = y'_t - y''_t + u_t$	In level-II, the model is applied on the data of the same series in AR3 but the demand values and residential values from customer demand modeled in the Exponential Smoothing Method are added.

3.2.2 ML Approach of Proposed Feed-forward ANFIS

The model is further validated by the ML approach for implementing ANN, while it is a computer model influenced by the brain and nervous system in animals or humans. Each network is a "neuron" integration that can generate values from sources that feed data across to the network. It provides a systematic analysis of neural networks. The majority of the research articles indicated that ANNs could be divided into two categories. 30 US power companies used a software-based ML platform in 1998 [132]. The radial base feature network, self-organizing maps, grouping, and recursive cognitive network are a few of the neural forecasting models used by LF. The ANFIS model is presented in [133] to forecast the potential load, and we have modified it according to our requirement as five stages computed in this section. An algorithm uses an unmonitored training data principle to construct a load-temperature relationship to forecast 24 hours of load [134]. This method uses real-time data as error-correcting function input, and simulation findings reveal that acceptable percent mean absolute error percentage is better than the conventional method. In this regard, the proposed system validates the Feed-forward ANN of one-year LF data. The proposed Feed-forward ANFIS model consists of five stages where Random Forest training schemes are taken into account, and the training process is given as follows:

Stage 1: This layer is known as the fuzzy fiction layer. At this layer, the received signal of any node will transmit to the other layer. Each node in this layer produces a membership function of the linguistic variable. The outputs are given in the equations where the linear data containing weekly, seasonally, and per year demand is illustrated by considering Equation 3.1. The non-linear data include ambient temperature, dew pony, solar irradiation, peak loads time, off-peak loads time, energy price, sunshine duration, fog duration in winter, and robust weather conditions. Table 3.2 defines it with the exponential method and level-2 AR methods considered in Equations 3.2 and 3.3.

$$O_i^1 = \mu A_i\left(Y_{y-i}\right) \text{ for I = 1,2,3} \tag{3.1}$$

$$O_2^1 = \mu A_2(Y't) \tag{3.2}$$

$$O_3^1 = \mu A_3(Y''t) \tag{3.3}$$

where: $Y(t-1)$, y_t', and Y_t'' are the inputs to the nodes as shown in Figure 3.1, μA_i is a member function.

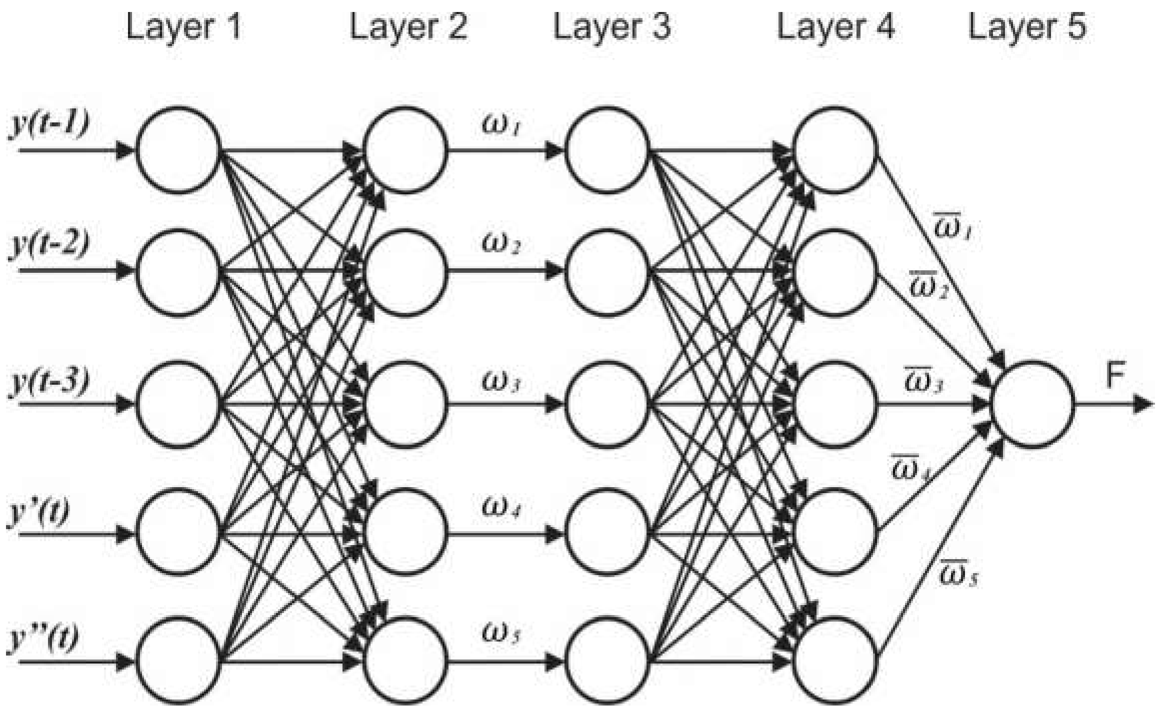

***Figure 3. 1** Flowchart for illustration of training and forecasting processes of the proposed feedback ANFIS algorithm.*

A is a linguistic variable that is linked with the node function. The μA in the model is chosen based on Equation 3.4

$$\mu A_i(x) = exp\left[-\frac{Y_t - c_i}{2a_i}\right]^2 \qquad (3.4)$$

where a_i and c_i are the set of premise parameters.

Stage 2: This layer 2 is the rule layer. It is used as a simplification of the predicter training scheme and minimizes the burden. The output of each node demonstrates the firing strength of each rule obtained by the membership functions given in Equation 3.5.

$$w_n = \mu A_i(Y_{t-i}) \times \mu A_2(Y't) \times \mu A_3(Y''t) \text{ for } \quad n = 1,2,3,4,5 \qquad (3.5)$$

Stage 3: Each node is a fixed node and is named N. This layer is known as the normalization layer because it calculates a ratio of firing strength of the rule relative and the sum of firing strengths of previous rules. The output is given by Equation 3.6.

$$\overline{w}_n = \frac{w}{w_1 + w_2 + w_3 + w_4 + w_5} \text{ n } = 1,2,3,4,5 \qquad (3.6)$$

Stage 4: The nodes at this stage depend on adaptive nodes. The output for each rule is calculated by taking the product of normalized firing strength from the previous phase and the first-order Sugeno Model. This layer is known as the defuzzification layer.

$$\theta_n^4 = \bar{w}_n f_n = \bar{w}_n (p_n + q_n + r_n) \tag{3.7}$$

where $\bar{w}_n$is the output of stage 3, and $p_i, q_i, and\ r_i$ are the parametric set.

Stage 5: The total output for the proposed model is calculated at this layer. The calculation is done based on the output values of each rule. That's why this layer is called the sum layer. There is a single node here that computes the overall output. The overall output is computed in a single node as illustrated in Figure 1 and given in Equation 3.8.

$$\theta_n^5 = \sum_n \bar{w}_n f_n = \frac{\sum_n \bar{w}_n f_n}{\sum_n \bar{w}_n} \tag{3.8}$$

Using the ANFIS model, the final output for given premise parameters can be represented as a linear combination of consequent parameters.

$$\begin{gathered} f = \bar{w}_1(p_1 + q_1 + r_1) + \bar{w}_2 (p_2 + q_2 + r_3) + \bar{w}_3(p_3 + q_3 + r_3) + \\ \bar{w}_4(p_4 + q_4 + r_4) + \bar{w}_5(p_5 + q_5 + r_5) \end{gathered} \tag{3.9}$$

3.3 BGA-PCA for Feature Selection

Feature selection is used in this research for eliminating the extraneous and redundant data that can improve accuracy and reduce computation time as applied to ML in [135]. In this regard, the BGA is based on the evolutionary perception of genetics and natural selection as it is considered a global solution for optimization problems. In this proposed model, the feature selection is based on BGA and the blend of PCA, as shown by the flowchart in Figure 2. The best feature selection is evaluated in five key steps. The steps are:

3.3.1 Initialization of Data Inputs and Model Paraments

The model starts with objective functions as BGA works on binary data set of special values for variables (chromosomes) as a GA starts with these variables to indicate the optimal variables for the problem [136]. The model paraments of this section are "1" and "0," where "1" means that the feature selected for the fitness evaluation and "0" is for the feature that is going to be not selected. After defining the objective functions and model parameters, the next step is creating the initial population.

3.3.2 Initial Population

As this feature selection is a GA-based solution, there is a need for two matrices for creating the initial population that is "k" matrix and "l" matrix where k defines the number of chromosomes and l indicates the length. Quantity and length of chromosomes are computed by population size and the number of genes, respectively. The suggestion for every population is given [137].

3.3.3 Determining Fitness from PCA

Although the time series and auto-regression method is used in this research for assessing the deployed fitness function, the objective functions are linked with parameters in Table 3.3 to evaluate the means square error of the auto-regression method. After MSE evaluation and for each subset hare is a need for data reduction. All forecasting work is consisting a massive amount of data and facing difficulties of a lot of computational work. PCA has been used to overcome the computational complexity by reducing the number of features and processing time for intrusion detection [138]. The auto-regression methods are a function of $y(t)$ and two matrices k and l are proposed then the PCA the linear combination for required value forms as:

$$PC_1 = \emptyset_1 y_1 + \emptyset_2 y_2 + \emptyset_3 y_3 + \ldots + \emptyset_k y_l \tag{3.10}$$

And the significant sample variance is reduced by:

$$\sum_{j=1}^{l} \emptyset_{j1}^2 = 1 \tag{3.11}$$

It indicates and optimizes the problem that is needed as a maximize function formed as:

$$\frac{1}{n}\sum_{j=1}^{k}\left(\sum_{j=1}^{l} \emptyset_k y_l\right)^2 \tag{3.12}$$

The data is further minimized by:

$$\frac{1}{n}\sum_{j=1}^{n} \emptyset_{j1}^{2} = 0 \tag{3.13}$$

On the other hand, the main objective of BGA is a minimization of MSE as a loss function for the ML approach that is formulated by:

$$fit_{fun} = \frac{i}{n}\sum_{i=1}^{n}(Ti - yt)^2 \tag{3.14}$$

where 'T' is a vector function of load demand and n is a training sample or assumptions for forecasting. As earlier mentioned, the term 1 defines the selected values while; 0 defines the non-selected values for the assessment of chromosomes. Each irritation of the BGA-PCA feature selection method decreases the MSE and finds the best objective function value.

3.3.4 Selection or Reproduction

The parameters considered for BGA are given in Table 3.3. The chromosome length is 48, and the number of iterations set to be 50. The new population is further; produced by crossover and mutation sections. The crossover and mutation functions use two chromosomes as parent-1 and parent-2 in this model, while the crossover function is the child selection.

Table 3.3 *The parameters considered for BGA.*

Model Parameters	**Considered Values**
Population Size	48
Selection probability	1
Selection Mechanism	Tournament selection
Crossover probability	0.95
Mutation Probability	0.20
Maximum iteration	50
Stopping criteria	1.When reached max no. of iteration 2.If fitness values are not better than the previous

In the parents' chromosomes, the child must be better than the parent. The XOR function is applied for the binary form where the crossover probability is 0.95. Here are two options, one is yes for the violation of any constraints, and the second is no for evaluating the new population. This crossover function can operate single time or multiple times base on the maximum length of chromosomes. In the mutation process, each bit is checked individually for searching for the best possible solution where its operating function single/multiple times is the same as the crossover function. The mutation probability is 0.20 as it is a genetic disorder of chromosomes, and it creates a random number which uniformly distributed with a maximum length of chromosome size. The proposed BGA-PCA model is chosen by setting the initial population, various probability values for mutation and crossover, maximum iterations, and the stopping criteria. It helps in the selection of a new population ultimately for the best feature subset.

3.3.5 Convergence Condition and Final Feature Subset

The new population is further tested in one step if it succeeds in finding the best chromosome within the iteration limit, decode it, and consider the best feature subset. On the other hand, the new population might not be better than the previous then it is suggested that keep the initial value as new. The convergence limits are following:

- When reached max no. of iteration
- If fitness values are not better than the previous

After finding the best chromosome, encode the best fitness score, and finally, choose the best feature subset. This flowchart of this complete process is shown in Figure 3.2 and is ready for further validation.

3.4 Proposed Integration Strategy

LF approach is based on time series and auto-regression model, ML method, and proposed feature selection scheme. In this section, the forecasting is further optimized by reducing MAPE. In this regard, this section integrates all approaches, where data fetching and modeling of time series and auto-regression model are considered. One-year data having four seasons with the transformation of per week data, ML, and feature selection are not considered as a final decision because the integration strategy looks toward the best algorithm of the related week of each season. Moreover, the MAPE calculations are modified as computed for the integration strategy in [131] for individual buildings as fellows.

$$MAPE_i = \frac{1}{W_s}\sum_{t=1}^{W_s}\frac{\left|fit_{fun}(t) - D(t)\right|}{D(t)} \tag{3.15}$$

where $it_{fun}(t)$ is a function taking from feature selection concerning time t, D is actual data, and w_s shows weeks per season. The average MAPE is then further calculated as:

$$MAPE_{avg} = \sum_{i=1}^{W_s} C_i * MAPE_i \tag{3.16}$$

Table 3.4 *The sequence of simulation.*

Proposed Algorithm	
Step 1:	According to Table 3, Adjust the simulation parameters.
Step 2:	Compute y(t) and $y'(t)$ and $y''(t)$ for different objective functions.
Step 3:	Generate random estimation for feedback ANFIS model consists of five stages.
Step 4:	Evaluate feature selection with BGA-PCA
Step 5:	Compute $fit_{fun} = \frac{i}{n}\sum_{i=1}^{n}(Ti - yt)^2$ for the different objective function of $y(t)$.
Step 6:	Compute $MAPE_i$ for 10 houses
Step 7:	Compute $MAPE_{avg}$
Step 8	A plot of Figures for MAPE of four seasons

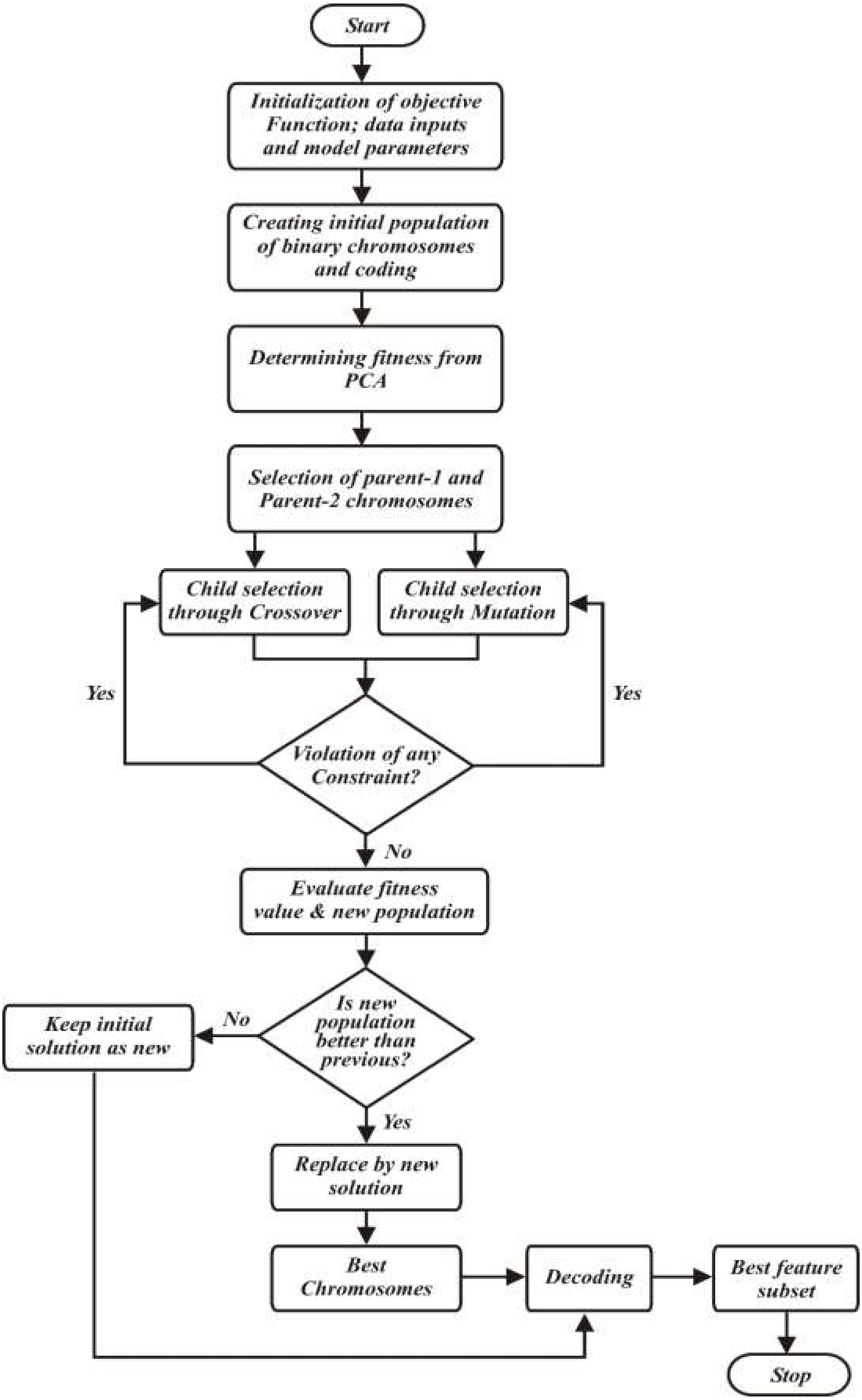

Figure 3. 2 *Flowchart of the BGA-PCA-based feature selection scheme.*

The PRECON data is modeled by the proposed time series and auto-regression system, where it has seven different estimation models. These consist of three AR models, one exponential smoothing, and three AR models with exponential smoothing. The algorithm is elaborated in Table 3 and validated with a feed-forward ANN of one-year LF data. The proposed feedback ANFIS model consists of five stages and uses real-time data as error-correcting function input. The simulation findings reveal that an acceptable mean absolute error percentage is better than the conventional method. The next step for the model evaluation starts from the feature selection, and a novel BGA-PCA model is presented in the thesis for the best fitness score. Finally, the best feature subset is chosen as indicated in Equation 3.14. Moreover, the seasonal and average MAPE calculation for the integration strategy is formulated for an individual building in Equations 3.15 and 3.16, respectively. This complete sequence of simulations is shown in Table 3.4.

3.5 Model Evaluation

In this thesis, the time series and auto-regression models are considered for ML-based feature selection named BGA-PCA. It is further validated using the integration strategy technique. For this purpose, 10 residential houses are taken into consideration where building area in sqr. Ft, numbers of floors, building year, the ceiling height in feet, total numbers of rooms, bedrooms, living rooms, drawing rooms, kitchen, connection type, numbers of people, numbers of adults (14 to 60), number of children (0–13), permanent residents, temporary residents, numbers of ACs, numbers of refrigerators, numbers of washing machines, numbers of electronic devices, numbers of fans, numbers of water dispensers, numbers of water pumps, numbers of electric heaters, numbers of irons, numbers of lighting devices, and numbers of UPS are the parts of data set.

As earlier discussed, the data of 10 houses from PRECON are considered for the LF. The proposed work and Random Forest training scheme are shown in Figure 3.3 systematized as follows:

Time series and auto-regression algorithms are developed to fetch the data set and set the objective function while; the symbolic representation of proposed equations is stated in Table 3.1 set the base for further validation as improved results validate it. Table 3.2, the data set fetching with seven AR methods are computed by different linear and non-linear data.

The preprocessing of the AR method is illustrated in Figure 3.3, where the 3 types of data are fed to feedback ANFIS. It contains linear data of Level-1 AR models, non-linear data of smooth exponential method, and Level-II AR models as computed in Table 3.3.

The time series AR data has been considered the inputs of proposed Feed-forward ANFIS as formulated in Equations 3.1-3.3. The overall output of Feed-forward ANFIS is computed in Equations 3.8 and 3.9. $\overline{w}_1$, $\overline{w}_2$, $\overline{w}_3$, $\overline{w}_4$, $\overline{w}_5$ are the five best weights obtained from the ANFIS model; For example, weights are $\overline{w}_1 = 50\%$, $\overline{w}_2 = 40\%$, $\overline{w}_3 = 30\%$, $\overline{w}_4 = 20\%$, $\overline{w}_5 = 10\%$ respectively. Then the new weights according to Equation 3.6 could be: $\overline{w}_1 = \frac{50}{150} = 33.3\%$, $\overline{w}_2 = \frac{40}{150} = 26.6\%$, $\overline{w}_3 = \frac{30}{150}$ 20%, $\overline{w}_4 = \frac{20}{150} = 13.3\%$, $\overline{w}_5 = \frac{10}{150} = 6.6\%$. The forecasted weight according to Equations 3.8 and 3.9, where f_n is the forecasting parameter of each weight according to week, weekend, month, season, and year as parameters defined in Table 3.3. Assume that f_n are 30, 25, 20, 15, 10 than the forecast weights are computed as: $f = 0.33 \times 30 + 0.26 \times 25 + 0.2 \times 20 + 0.13 \times 15 + 0.06 \times 10 \approx 23$.

BGA-PCA for the proposed feature selection used in this work eliminates the extraneous and redundant data that can improve accuracy and reduce computation time. The main objective of BGA is to minimize the loss function as computed in Equation 3.14. The extra parameters, repeated terms, and irrelevant data have been removed by the BGA-PCA algorithms, as shown in the flowchart of the algorithm in Figure 3.2 and simulation parameters in Table 3.3. It also shows that the child's chromosomes are better than the parent's. In the next step, the XOR function is applied where the crossover probability is 0.95. Whereas there are two possibilities, the first one is yes if the violation of any constraints is found, and the second one is no when the new population is evaluated. This crossover function can operate single time or multiple times based on the maximum length of chromosomes. In the mutation process, each bit is checked individually for searching for the best possible solution where is operating function single/multiple times is the same as the crossover function. The mutation probability is 0.20 as it is a genetic disorder of chromosomes, and it creates a random number that is distributed uniformly with a maximum length of chromosome size. The proposed BGA-PCA model is chosen by setting the initial population, various probability values for mutation and crossover, maximum iterations, and the stopping criteria that ultimately help select a new population for the best feature subset.

MAPE calculation for the integration strategy is formulated for an individual building, and average MAPE calculation is done in Equations 3.15 and 3.16, respectively. As per MAPE calculation in Table 3.5, the average MAPE of House.1 for the summer season is 1.72 as calculated from Equation 3.16. The weekly coefficient from different algorithms C_i are 30%, 25%, 20%, 15%, 10% and $MAPE_i$ of house.1 according to different algorithms are 1.70, 1,77,1.68,1.78, 1.69 as per Equation 3.15. The integration strategy computed the minimum MAPE as per Equation 3.16 is $MAPE\ of\ house.1 = 0.30 \times 1.70 + 0.25 \times 1.77 + 0.20 \times 1.68 + 0.15 \times 1.78 + 0.10 \times 1.69 = 1.72$. Similarly, the MAPE of each house for four seasons is calculated, and the average is computed as given in Table 3.6.

The MAPE improvements are discussed in Equations 3.17–3.19, and the calculated results are demonstrated in Table 3.6 and Figures 3.4–3.7 validate the proposed model. The MAPE has improved the system by 0.89 for the summer seasons of house 1, as its value was 2.97 with MAPE and 2.08 without MAPE, taking the MAPE value equal to 1.72. Then E_1 according to Equation (17) is $\frac{2.97-2.08}{2.97}$= 29.99% and after four-season the house1 MAPE average becomes 39%, shown in Table 3.5. The further improvement is calculated in Equation 3.18 as $\frac{2.08-1.72}{2.08}$ = 17%, where after four seasons, the house1 MAPE average becomes 23% shown in Table 3.6. The overall MAPE improvement of house1 is calculated using Equation 3.19 as $39\% - 23\% = 16\%$, and the average of 10 houses is 17%.

The highest load is in the summer season, in which July and August were at their peak, and the minimum load is in the winter season, in which January and February were on their minimum value. Besides, the feature selection approach has the following features: day hours, day of the weeks, weekend, holidays, weeks in a month, month of a season, seasons in a year, ambient temperature, dew pony, solar irradiation, peak loads time, off-peak loads time, energy price, sunshine duration, fog duration in winter, and robust weather conditions. One year of data from July 2018 to June 2019 is taken into account. After deploying the complete model modeled in Section 3 and the complete sequence of simulation illustrated in Table 3.5, the MAPE for the LF improvement, MAPE of each house is calculated where the average MAPE without proposed integration strategy is (M_1). The average MAPE with only the proposed integrated approach without feature selection is (M_2). This improvement percentage is calculated by:

$$E_1 = \frac{(M_1 - M_2)}{M_2} \tag{3.17}$$

The average MAPE of 10 houses computed by our proposed model is (M_3). Then the improvement is formulated by:

$$E_2 = \frac{(M_2 - M_3)}{M_3} \tag{3.18}$$

The overall improvement Percentage can be calculated by:

$$E = E_1 - E_2 \tag{3.19}$$

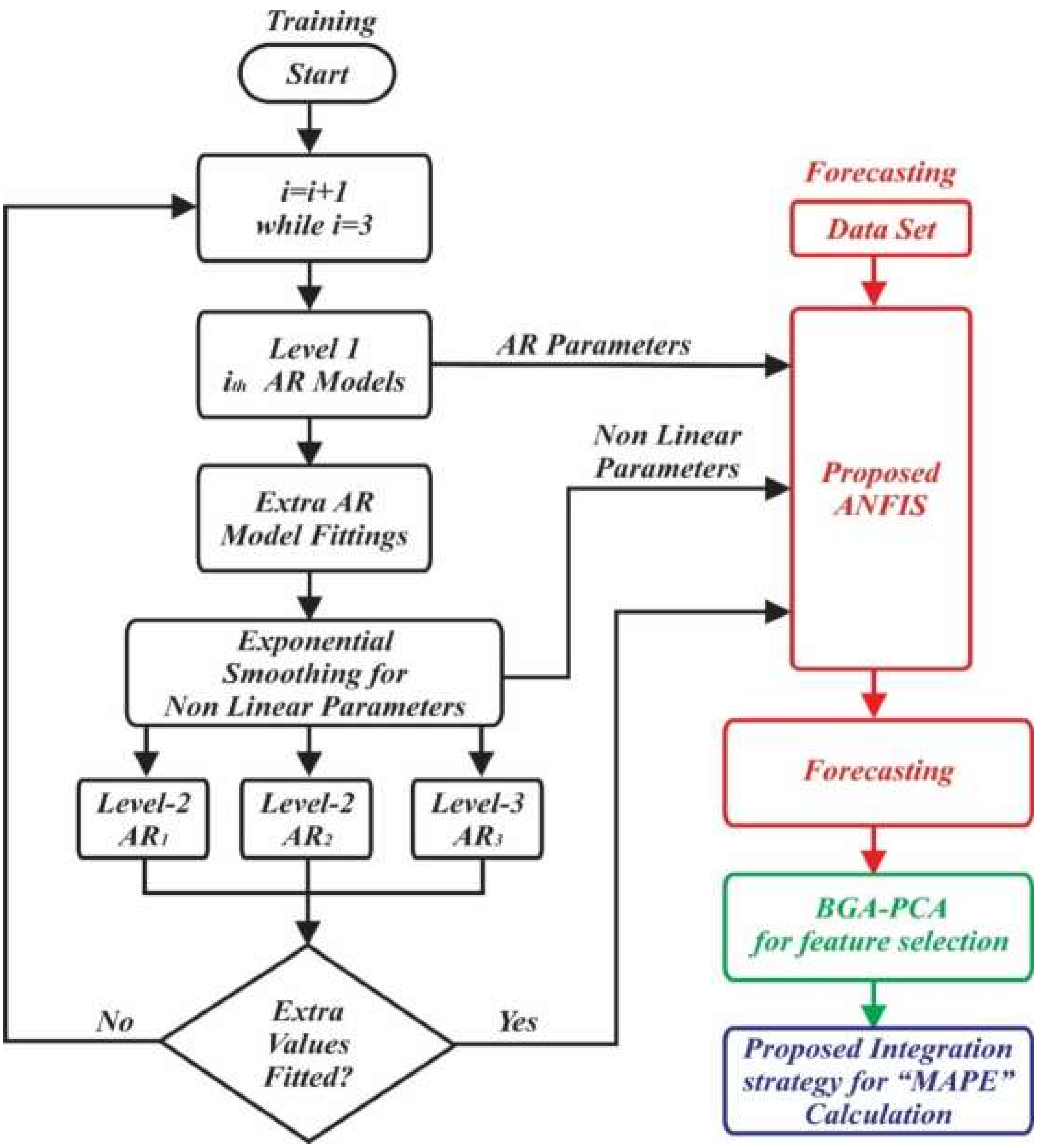

Figure 3.3 *Flowchart of preprocessing for time series auto-aggregation method with the proposed framework*

Table 3.5 *LF improvement per building.*

Customer Type.	Without Proposed Integration M_1 MAPE Percentage				With Proposed Integration M_2 MAPE Percentage				With Proposed Integration & BGA-PCA feature selection M_3 MAPE Percentage				With Proposed Integration Improvement Percentage	With Proposed Integration & ML Improvement Percentage	Over All Improvement Percentage
	Summer MAPE %	Fall MAPE %	Winter MAPE %	Spring MAPE %	Summer MAPE %	Fall MAPE %	Winter MAPE %	Spring MAPE %	Summer MAPE %	Fall MAPE %	Winter MAPE %	Spring MAPE %	$E_1 = \frac{(M_1 - M_2)}{M_2}$, %	$E_2 = \frac{(M_2 - M_3)}{M_3}$, %	$E = E_1 - E_2$, %
H 1	2.97	3.01	3.07	2.83	2.08	2.19	2.23	2.06	1.72	1.79	1.76	1.71	39%	23%	16%
H 2	2.88	3.02	3.06	2.67	2.10	2.19	2.20	2.02	1.69	1.67	1.84	1.67	37%	24%	13%
H 3	2.87	2.98	3.08	2.84	2.16	2.17	2.29	2.02	1.82	1.75	1.81	1.71	36%	22%	14%
H 4	2.84	2.94	3.09	2.89	2.10	2.19	2.18	2.08	1.73	1.80	1.84	1.69	38%	21%	16%
H 5	2.91	2.96	3.09	3.01	2.20	2.14	2.16	2.07	1.72	1.72	1.77	1.75	40%	23%	17%
H 6	2.81	2.94	2.99	2.98	2.19	2.11	2.16	2.02	1.81	1.78	1.81	1.64	38%	20%	18%
H 7	2.88	3.08	2.98	2.94	2.06	2.13	2.17	2.06	1.66	1.77	1.82	1.60	41%	23%	18%
H 8	2.9	3.04	2.97	2.85	2.01	2.12	2.15	2.06	1.67	1.85	1.76	1.65	41%	20%	21%
H 9	2.94	3.06	3.06	2.84	2.10	2.16	2.11	2.07	1.60	1.76	1.79	1.69	41%	23%	18%
H 10	2.95	2.99	3.08	2.83	2.04	2.14	2.30	2.04	1.60	1.83	1.82	1.61	39%	24%	15%
Avg	2.90	3.00	3.05	2.87	*2.10*	*2.15*	*2.20*	*2.05*	***1.70***	***1.77***	***1.80***	***1.67***	39%	22%	**17%**

This complete model is simulated in a MATLAB simulation environment on a research workstation with Intel Core i5-4300CU, 8GB RAM, 64-bit operating system, and 2.50 GHz Processor. The proposed model of 10 houses with a one-year hourly database is simulated in 50 min.

3.6 Results Discussion

LF improvement for MAPE of each house is calculated in Table 3.5. The results of M_1 without proposed integration strategy is 2.90, 3.00, 3.05, and 2.87 percentage for Summer, Fall, Winter, and Spring, respectively. The results are improved with only the proposed integrated approach without feature selection: 2.10, 2.15, 2.20, and 2.05 percentage for Summer, Fall, Winter, and Spring, respectively. The results are more improved with the proposed model illustrated in Table 3.6, and the average MAPE of 10 houses M_3 is calculated with the help of proposed equations and sequence simulation.

The proposed model attained 1.70, 1.77, 1.80, and 1.67 percentages for Summer, Fall, Winter, and Spring, respectively. The overall improvement based on 10 buildings is 17% computed in Table 3.6.

Randomly 8 weeks of each season are selected to illustrate the forecasting results, where Sunday is considered a weekend and the seasons are: summer (July–August 2018), Fall (October–November 2018), Winter (January–February 2019), and Spring (March–April 2019). The actual and forecasted results of electricity demand with minimum MAPE are depicted in Figure 4–7 for the Summer, Fall, Winter, and Spring seasons, respectively. Besides the overall improvement with our proposed model is 17% computed in Table 3.6, the proposed model has some better LF results than LF models given in [138], [139], [140].

Although the results of our research work are promising and desirable, some limitations of this study are that a new variety of data input like the probability of load shedding, faults occurring, and power failure of the region can also be considered. In addition, seven regression methods are adopted in time series and auto-regression algorithms, but this model can consider more AR methods for accurate data preprocessing. Despite these trivial limitations, our trial with a large dataset is speed-up and beneficial for the researchers as it obtained a MAPE of less than 2%.

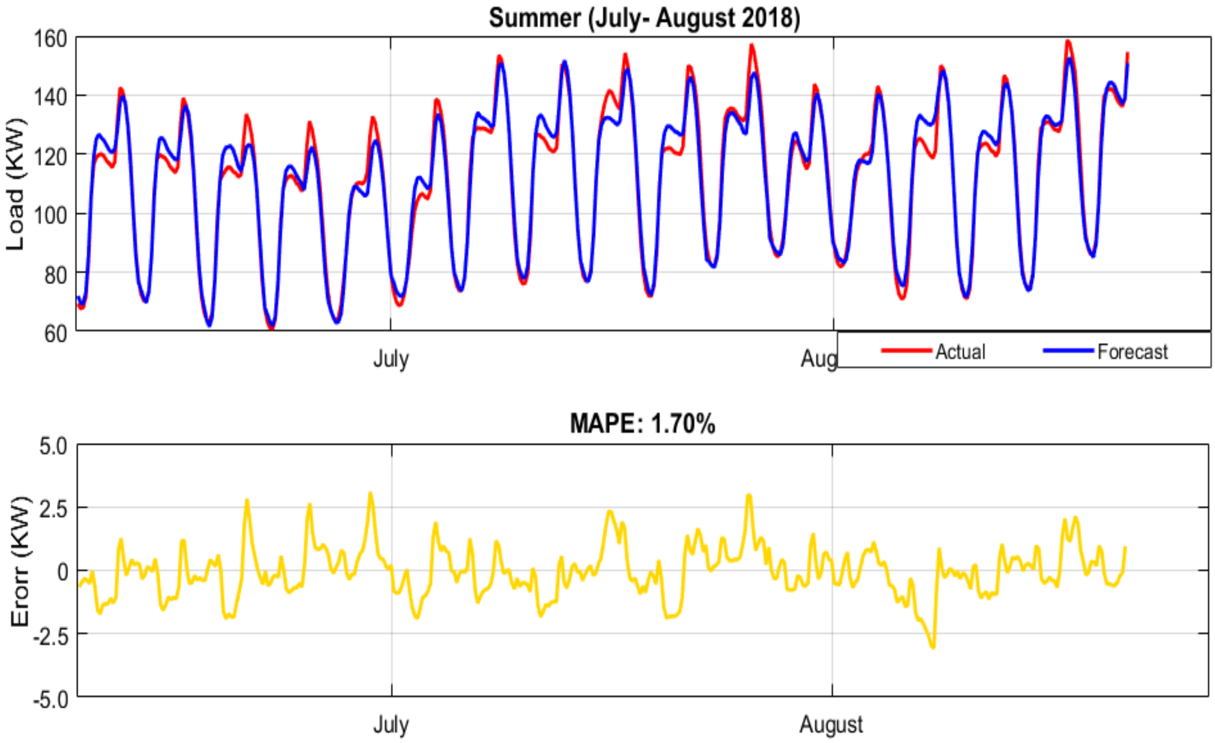

Figure 3.4 *Actual vs. forecasted residential consumption demand of Summer season and average MAPE of a season.*

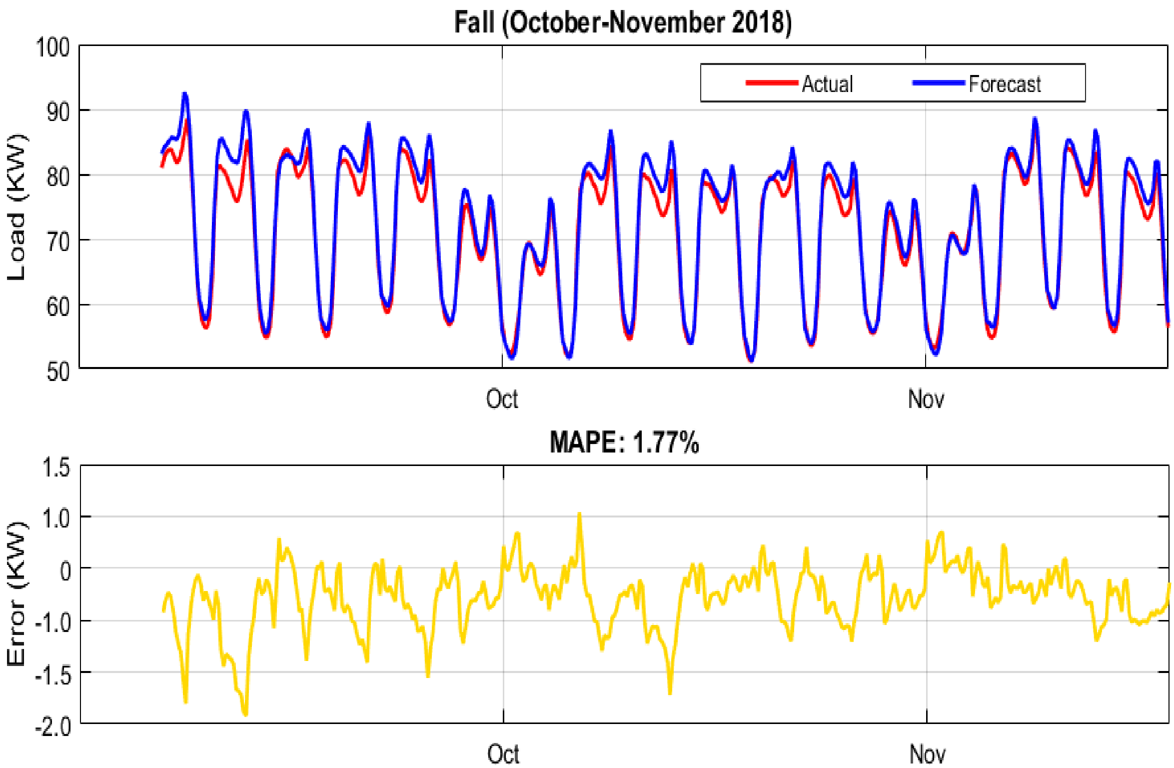

Figure 3.5 *Actual vs. forecasted residential consumption demand of Fall season and average MAPE of a season.*

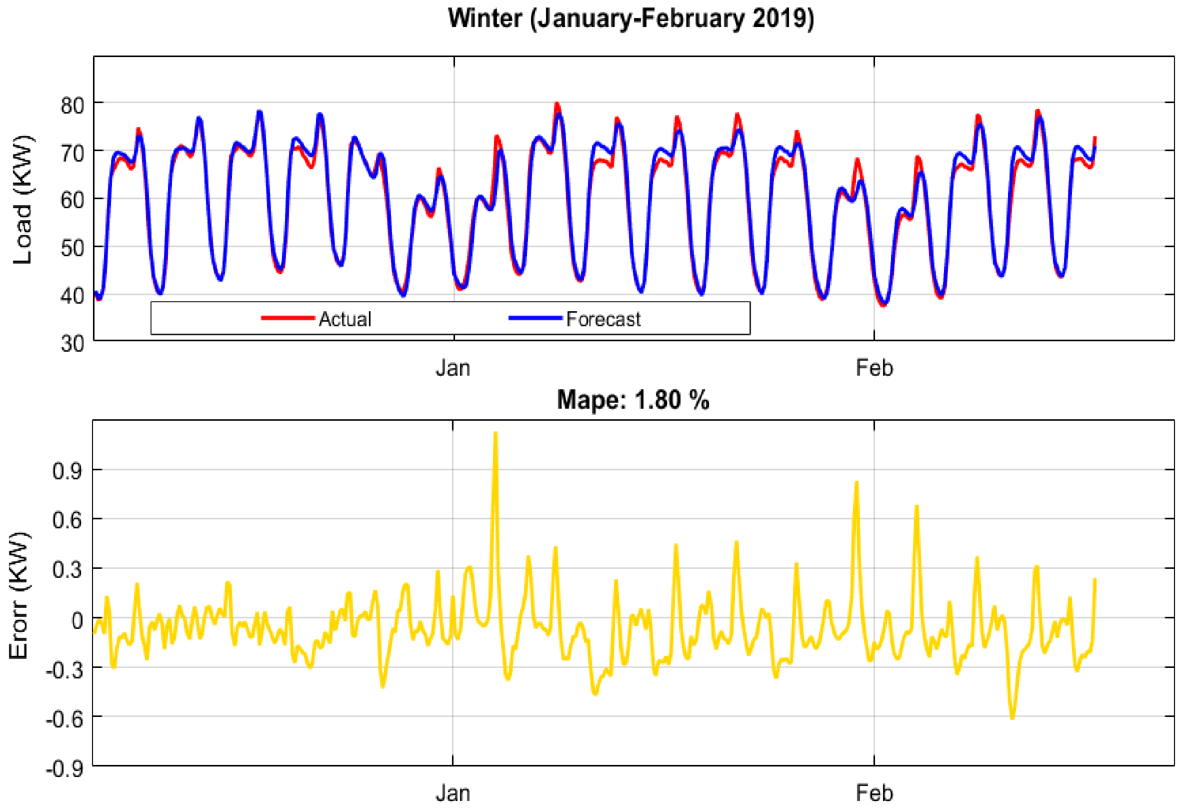

Figure 3.6 *Actual vs. forecasted residential consumption demand of winter season and average MAPE of a season.*

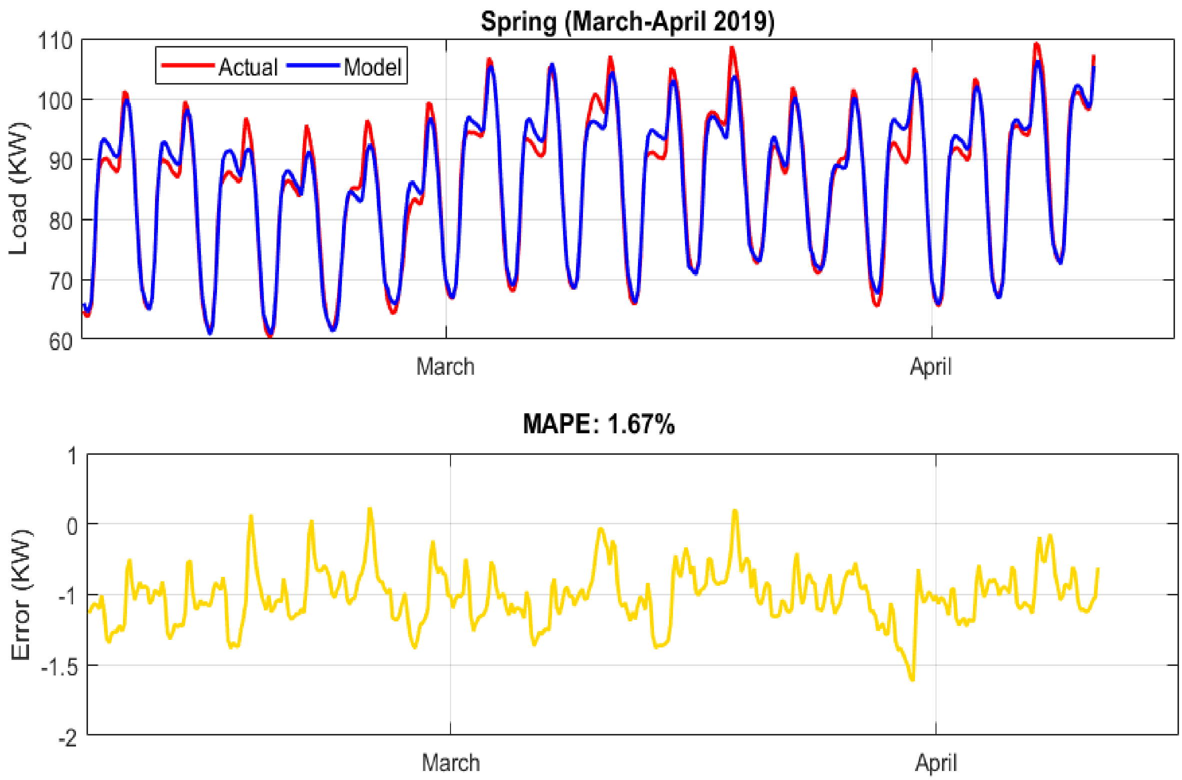

Figure 3.7 *Actual vs. forecasted residential consumption demand of Spring season and average MAPE of a season.*

Table 3.6 *Comparison Results in Figures 3.5-3.8 with past works.*

Results[139] **Novel Hybrid STLF Model Five Regions in Australia**		**Results**[140] **Second Learning of Error Trend Model on Different Regions**		**Results (Proposed Model) 10 Residential Building**	
Season	Min MAPE %	Season	Min MAPE %	Season	Min MAPE %
Summer (July–Sep)	1.68	Summer	1.84	Summer (July–August)	1.70
Fall (Oct–Dec)	2.71	Fall	1.87	Fall (Oct–Nov)	1.77
Winter (Jan–Mar)	2.28	Winter	1.66	Winter (Jan–Feb)	1.80
Spring (April–June)	2.29	Spring	1.80	Spring	1.67

3.7 Summary

Load Forecasting is one of the Smart Grid applications which supports power generation system for its optimal operation. Therefore, in this chapter, a novel combinational forecasting model is proposed. A new time series and auto-regression algorithm for analyzing the more complex and massive historical data are presented. First, the validation of the time series AR model is proved using the ML-based feedback ANFIS model. After that, for best feature selection BGA-PCA approach is used. A proposed integration strategy showed an annual improvement in MAPE value and an average MAPE of 17%. Hence, our proposed model can optimally enhance the efficiency of the power generation system by predicting optimized LF.

CHAPTER 4

AN IMPROVED RESIDENTIAL MACHINE LEARNING-BASED ELECTRICITY PRICE FORECASTING

4.1 Introduction

It is necessary to forecast electricity prices for the Independent System Operator (ISO) and the end-users and marketers. Generally, the profit bidders in the potential electricity market need the future electricity prices to earn a good profit; however, the existing electricity market has made PF more complex as they are extremely deregulated and non-linear. Due to the system's non-linear and unstable behavior, an accurate prediction of prices has become more complex. It also affects the bidding policies in the electricity market. The PF techniques are categorized into 3 classes according to their specific forecasting model as Statistical Model, time series model, and AI-based models [141]. The AI-based forecasting approach has acquired a lot of attention in recent years as it assures a convinced level of accuracy of price estimation compared to unstable variations of dependent or independent variables in the statistical model.

This thesis collects data from the NEPRA dataset, as illustrated in Table 4.1 where PF is proposed for 1 year. Power producers that are generating power locally here in Pakistan, CPPs, PPPs, and Tariff rates are considered data input. Finally, the MAPE of PF is optimized for the Lahore electric supply company (LESCO). The prices of each power producer CPPs and PPPs show variation during the different months and according to load demand. The proposed system is based on twelve months as events while PF is optimized and tries to reduce the MAPE. The eight models of time-series and auto-regression approach are adopted for fetching the data to optimize PF, and for feature selection, the BGA-PCA technique is also proposed.

4.2 Proposed Strategy

This section describes the complete framework of Proposed PF.

4.2.1 Time Series and Auto-Regression Method

The time series forecasting model is proposed in [131], the nine regression algorithms are introduced for the LF. In this model, time series and auto-regression methods are taken into consideration for PF. Seven models are computed where three AR level-1, one exponential smoothing, and three AR level-II.

Level I Auto-regression Model (AR1)

Variation in Price demand: Peak hour $\mathrm{P(t-1)}$, off-peak hour $\mathrm{P(t-2)}$, Special holidays or weekend $\mathrm{P(t-3)}$:

$$\mathrm{P_t} = \beta_0 + \beta_1 P_{t-1} + \beta_2 \mathrm{P_{t-2}} + \beta_3 \mathrm{P_{t-3}} + \mathrm{u_t} \tag{4.1}$$

Level I Auto-regression Model (AR2)

AR2 is modeled by adding month auto-regression term $\mathrm{M_{k,t}}$where k=1,…,12. As 12 months in a year.

$$\mathrm{P_t} = \beta_0 + \beta_1 P_{t-1} + \beta_2 \mathrm{P_{t-2}} + \beta_3 \mathrm{P_{t-3}} + \sum_{\mathrm{k=1}}^{12} \alpha_\mathrm{k} \mathrm{M_{k,t}} + \mathrm{u_t} \tag{4.2}$$

Level I Auto-regression Model (AR3)

AR3 model has an extra auto-regression term in annual,$\gamma_k A_{k,\mathrm{t}}$ and for the seasons of the year $M_{\mathrm{k,t}}$where k=1,…,12.

$$\mathrm{P_t} = \beta_0 + \beta_1 P_{t-1} + \beta_2 \mathrm{P_{t-2}} + \beta_3 \mathrm{P_{t-3}} + \gamma_k A_{k,\mathrm{t}} + \sum_{\mathrm{k=1}}^{12} \alpha_\mathrm{k} \overline{\mathrm{M}}_{\mathrm{k,t}} + u_t \tag{4.3}$$

Support Vector Regression (SVR)

This regression forests $\mathrm{P_t}$by considering the inputs $\mathrm{P_{t-1}}, \mathrm{P_{t-2}},$ and $\mathrm{P_{t-3}}$ for each month $\sum_{\mathrm{k=1}}^{12} \alpha_\mathrm{k} \mathrm{M_{k,t}}$ of the annual forecasting.

$$P_{t,SVR} = SVR\ (\mathrm{P_{t-1}}, \mathrm{P_{t-2}}, \mathrm{P_{t-3}}) \tag{4.4}$$

Exponential Smoothing Method

This model is used for the purpose where there is no monthly pricing. Where α= smooth constant, P_{t-1}= previous pricing, and F_{t-1}= PF demand

$$P = \alpha E_{t-1} + (1 - \alpha)F_{t-1} + u_t \tag{4.5}$$

Level II Auto-regression Model (AR1)

In level-II, the model is applied on the data of the same series in AR1; however, the price values and residential values from customer demand modeled in the Exponential Smoothing Method are added.

$$P_t' = \alpha E_{t-1} + (1 - \alpha)F_{t-1} + u_t \tag{4.6}$$

$$P_t'' = y_t = \beta_0 + \beta_1 P_{t-1} + \beta_2 P_{t-2} + \beta_3 P_{t-3} + u_t \tag{4.7}$$

$$P_t = y_t' - y_t'' + u_t \tag{4.8}$$

Level II Auto-regression Model (AR2)

In level-II, the model is applied on the data of the same series in AR2; however, price values and residential values from customer demand modeled in the Exponential Smoothing Method are added.

$$P_t' = \alpha P_{t-1} + (1 - \alpha)F_{t-1} + u_t \tag{4.9}$$

$$P_t'' = \beta_0 + \beta_1 P_{t-1} + \beta_2 P_{t-2} + \beta_3 P_{t-3} + \sum_{k=1}^{4} \alpha_k M_{k,t} + u_t \tag{4.10}$$

$$P_t = P_t' - P_t'' + u_t \tag{4.11}$$

Level II Auto-regression Model (AR3)

In level II, the model is applied to the data of the same series in AR3; however, the demand values and residential values from customer demand modeled in the Exponential Smoothing Method are added.

$$P_t' = \alpha E_{t-1} + (1 - \alpha)F_{t-1} + u_t \tag{4.12}$$

$$P_t'' = \beta_0 + \beta_1 P_{t-1} + \beta_2 P_{t-2} + \beta_3 P_{t-3} + \gamma_k Y_{k,t} + \sum_{k=1}^{4} \alpha_k \overline{M}_{k,t} + u_t \tag{4.13}$$

$$P_t = P_t' - P_t'' + u_t \tag{4.14}$$

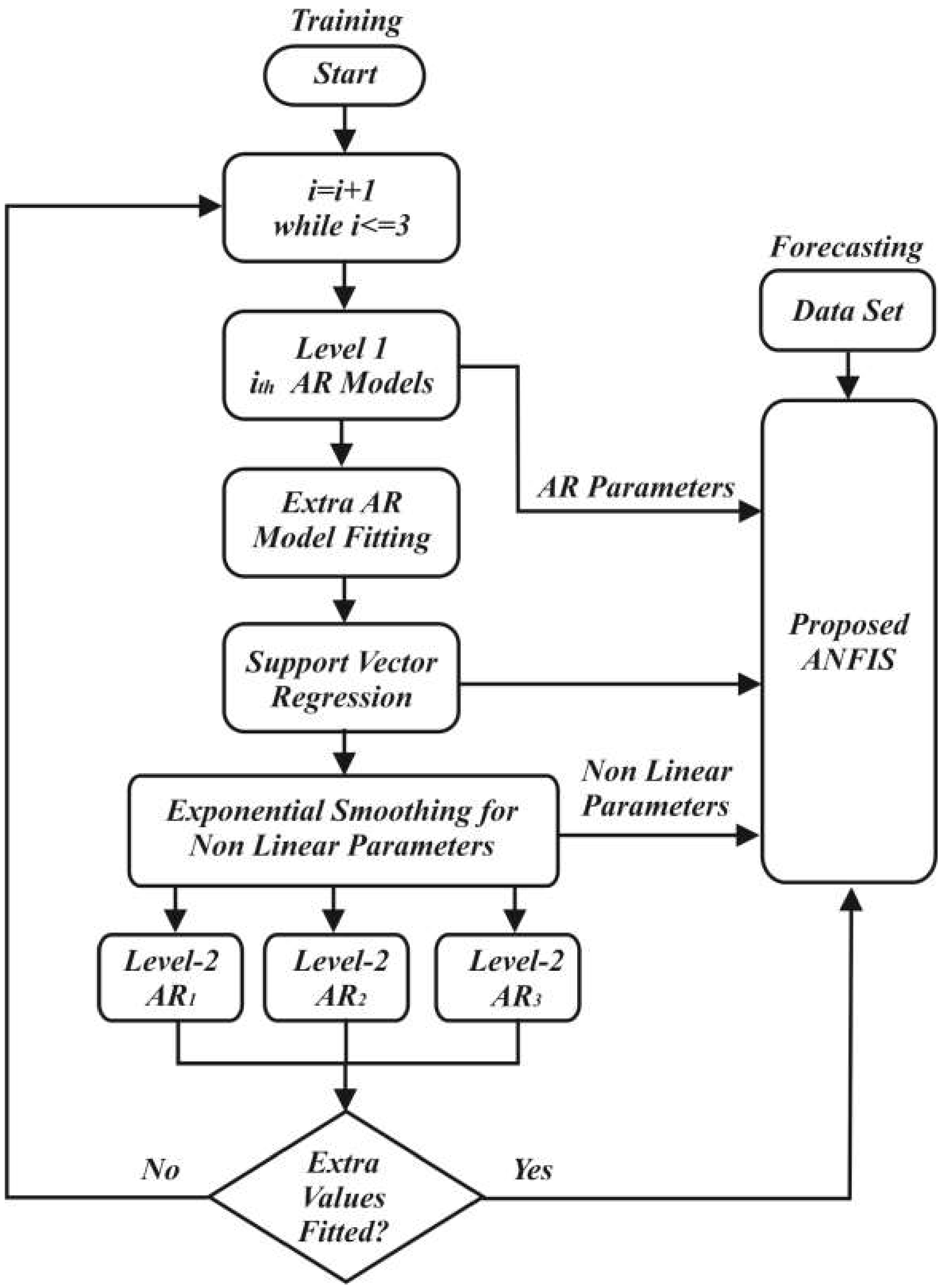

Figure 4.1 *Illustration of input parameters of ANFIS.*

4.2.1 Machine Learning Approach of Proposed Feed-forward ANFIS

The multi-layer feed-forward Neural Network (MLFF-NN) approach based on Fourier analysis was used to extract selection features of price and load signals in each hour [142]. The electricity price varies due to dynamic variations in load demand throughout the day; therefore, this strategy minimizes the MAPE in load and price on hourly, daily, and weekly intervals. However, with the dynamic variations in load containing high-frequency features, an error is considerably increasing; therefore, this research recommends Wavelet Transform to resolve this issue. In [133], the ANFIS model for LF forecast the electrical load demand and modified it as per PF. This method utilizes the real-time data of NEPRA extracted from time series and the auto-regression approach. The simulation results express realistic and acceptable outcomes with minimum MAPE. Therefore, the proposed methodology is implemented in 5 layers: fuzzy fiction, fuzzy rules, normalization, defuzzification, and sum. The training procedure modeling is given as illustrated in Figure 4.2.

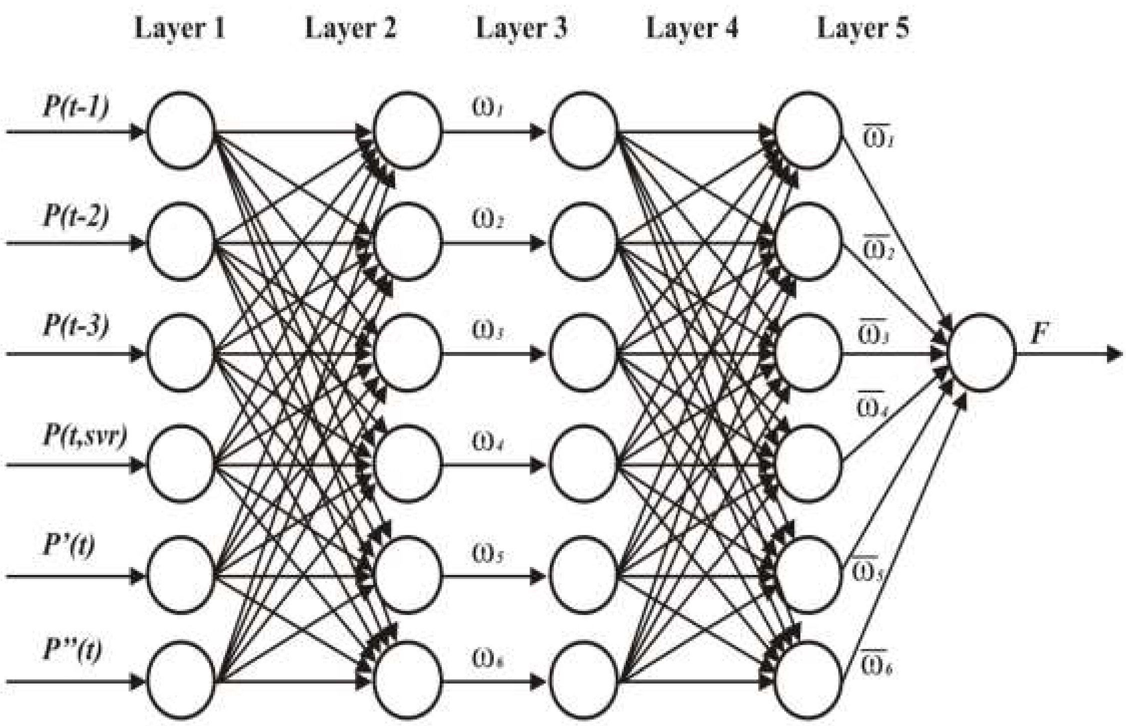

Figure 4.2 *Flowchart for illustration of PF processes of the proposed feedback ANFIS algorithm.*

Stage 1: In this stage, the fuzzy fiction layer transmits the received signal of any node simply to another layer. Each node for the linguistic parameter declares a member function. The output estimation of electricity price on an hourly, weekly, and monthly basis. The price estimation mainly includes power consumption statistics during the peak load time, off-peak load time, and dynamic weather conditions. The exponential methodology and level-2 AR methods are considered in Equations (4.15)-(4.18).

$$J_i^1 = \mu A_i\left(P_{y-i}\right) \text{ for i=1,2,3} \tag{4.15}$$

$$J_2^1 = \mu A_2(P't) \tag{4.16}$$

$$J_3^1 = \mu A_3(P''t) \tag{4.17}$$

$$J_4^1 = \mu A_4(P_{t,SVR}) \tag{4.18}$$

Where: $P_{y-i}, P_t', P_t''$and$P_{t,SVR}$ are the inputs to the nodes as shown in Figure 1 $\mu A_i$is a member function.

A is a linguistic parameter that is linked with the node function.

The µA in the model is chosen based on Equation (4.19)

$$\mu A_i(x) = exp\left[-\frac{P_t - c_i}{2a_i}\right]^2 \tag{4.19}$$

where a_i and c_i are the set of premise parameters.

Stage 2: At this stage, the rule layer is employed, which simplifies the member function of PF and minimizes the burden on nodes by processing the training data set. The node's output expresses the firing strength of each rule and is obtained by the membership functions illustrated by Equation 4.20.

$$w_n = \mu A_i(P_{t-i}) \times \mu A_2(P't) \times \mu A_3(P''t) \times \mu A_4\left(P_{t,SVR}\right) \text{ for n} = 1,2,3,4,5,6 \tag{4.20}$$

Stage 3: Here the normalization layer computes the ratio of firing strength of the relative rule to the sum of firing strengths of preceding rules. Each node at this stage is a fixed node and is named N. The output is given by Equation 4.21.

$$\bar{w}_n = \frac{w}{w_1 + w_2 + w_3 + w_4 + w_5 + w_6} \text{ for } \quad n = 1,2,3,4,5,6 \tag{4.21}$$

Stage 4: At this stage, the defuzzification layer evaluates the output for each rule by taking the product of normalized firing strength from the previous stage and first-order Sugeno Model. The nodes at this stage are based on adaptive nodes.

$$\theta_n^4 = \bar{w}_n f_n = \bar{w}_n(k_n + l_n + m_n) \tag{4.22}$$

where $\bar{w}_n$ is the output of stage 3, and k_n, l_n, and m_n are the parametric set.

Stage 5: Finally, the sum layer computes the total output for the proposed model. The calculation is done based on the output values of each rule. There is a single node here that computes the overall output. The overall output is computed in a single node shape, as illustrated in Figure. 4.1 and given in Equation 4.23.

$$\theta_n^5 = \sum_n \bar{w}_n f_n = \frac{\sum_n \bar{w}_n f_n}{\sum_n \bar{w}_n} \tag{4.23}$$

It has been observed that the ANFIS model declares the final output for a given premise by representing a linear combination of consequent parameters.

$$f = \bar{w}_1(k_1 + l_1 + m_1) + \bar{w}_2(k_2 + l_2 + m_2) + \bar{w}_3(k_3 + l_3 + m_3) + \bar{w}_4(k_4 + l_4 + m_4) + \bar{w}_5(k_5 + l_5 + m_5) + \bar{w}_6(k_6 + l_6 + m_6) \tag{4.24}$$

4.2.2 BGA-PCA for Feature Selection

In section 3.2 of the previous chapter, feature selection eliminates the extraneous and redundant data that can improve accuracy and reduce computation time for PF. In this proposed model, the feature selection is based on BGA by the blend of PCA; the flow chart is presented in Figure. 3.2. Five critical steps in section 3.2 evaluate the best feature selection while equations 3.10-3.13 are used. On the other hand, the main objective of BGA is a minimization of MSE as a loss function for the ML approach that is formulated Equation 4.25 as:

$$fit_{fun} = \frac{i}{n}\sum_{i=1}^{n}(Ti - Pt)^2 \qquad 4.25$$

where T is a vector function of price estimation and n is a training sample or assumptions for forecasting. The term 1 defines the values that have to be chosen, and 0 defines the values that have not to be chosen for the assessment of chromosomes. Thus, each irritation of the BGA-PCA feature selection method decreases the MSE and finds the best objective function value.

The parameters considered for BGA are presented in Table 4.1.

Table 4.1 *The parameters considered for BGA-PCA.*

Sr. No	Model Parameters	Considered values
1.	Population Size	72
2.	Selection probability	1
3.	Selection Mechanism	Tournament selection
4.	Crossover probability	0.90
5.	Mutation Probability	0.20
6.	Maximum iteration	48
7.	Stopping criteria	1. When reached max no. of iteration 2. If fitness values are not better than the previous

4.2.3 Proposed Integration Strategy

PF approach is based on time series and auto-regression model, ML method, and proposed feature selection scheme. In this section, the forecasting is further optimized by reducing MAPE. Therefore, all approaches are integrated into this part where data fetching and modeling of time series and auto-regression model are considered. One-year data having twelve months with the transformation of per week data, ML, and feature selection are not considered as a final decision because the integration strategy looks toward the best algorithm of the related week of each month. Moreover, the MAPE calculations are modified as computed for the integration strategy in Chapter 3. We computed the MAPE of PF for each month as fellows.

$$MAPE_i = \frac{1}{W_m}\sum_{t=1}^{W_m}\frac{\left|fit_{fun}(t) - P_a(t)\right|}{P_a(t)} \qquad 4.26$$

wherefit_{fun} (t)is a function taking from feature selection concerning time t, P_a is the actual price, and w_m is weeks per month. The average MAPE is then further calculated as:

$$MAPE_{avg} = \sum_{i=1}^{W_m} C_i * MAPE_i \qquad 4.27$$

Table 4.2 *Sequence of simulation.*

Proposed Algorithm	
Step 1:	According to Equations 4.1-4.13, adjust the Simulation parameters.
Step 2:	Compute P(t), $P'(t), P''(t)$, and $P(t, svr)$ for different objective functions.
Step 3:	Generate random estimation for feedback ANFIS model consists of five stages.
Step 4:	Evaluate feature selection with BGA-PCA as per Equation 4.11
Step 5:	Compute $fit_{fun} = \frac{i}{n}\sum_{i=1}^{n}(Ti - Pt)^2$ for the different objective function of $P(t)$.
Step 6:	Compute $MAPE_i$ for 12 months
Step 8:	Compute $MAPE_{avg}$
Step 9:	Plot the Figures for MAPE of 12 months

4.3 Model Evaluation and Result Discussion

In this section, the time series and auto-regression models are considered for ML-based feature selection, BGA-PCA. Furthermore, the integration strategy technique further validates it. For this purpose, 12 power producers such as RES, Hydel Power, Gas Power, RNLG, imported coal, local coal, residual fuel oil, bagasse, uranium, RNLG new, imported power, CPP per kWh of each month, PPP per kWh of each month and load demand of each month are taken into consideration as input data set.

The data from table 4.1 is taken for estimation of PF for twelve months. The proposed work and Random Forest training scheme are shown in Figure 4.3 as systematized as follows:

Time series and auto-regression algorithms are developed to fetch the data set and set the objective function while proposed equations 4.1-4.13. According to these equations, eight AR methods fetch data set by different linear and non-linear parameters.

The preprocessing of the AR method where 4 types of data are fed to feedback ANFIS contains linear data of Level-1 AR models as shown in Figure 4.1.

The time series AR data has been considered as an input to the proposed Feed-forward ANFIS, and the overall output of Feed-forward ANFIS is computed in Equations 4.9-4.10. $\overline{w}_1$, $\overline{w}_2$, $\overline{w}_3$, $\overline{w}_4$, $\overline{w}_5$ and$\overline{w}_6$are the six best weights obtained from the ANFIS Model. For example, weights are $\overline{w}_1 = 60\%$, $\overline{w}_2 = 50\%$, $\overline{w}_3 = 40\%$, $\overline{w}_4 = 30\%$, $\overline{w}_5 = 20\%$and $\overline{w}_6 = 10\%$respectively. Then the new weights according to Equation 4.7 could be: $\overline{w}_1 = \frac{60}{210} = 28.5\ \%$, $\overline{w}_2 = \frac{50}{210} = 23.8\ \%$, $\overline{w}_3 = \frac{40}{210} = 19\ \%$, $\overline{w}_4 = \frac{30}{210} = 14.2\%$, $\overline{w}_5 = \frac{20}{210} = 9.5\%$ and $\overline{w}_6 = \frac{10}{210} = 4.7\%$ The forecasted weight according to equations 4.9 and 4.10, where f_n is the forecasting parameter of each weight according to week, weekend, month, and year as parameters computed in Equations 4.1-4.13. Assume that f_n are 30, 25, 20, 15, 10 and 5 then the forecast weights are computed as: $f =$ 0.28*30+ 0.23*25+ 0.19*20 + 0.14*15+ 0.09*10+0.04*5 ≈ 21.15

BGA-PCA for feature selection is used to eliminate the extraneous and redundant data that can improve accuracy and reduce computation time. The main objective of BGA is minimization as the loss function is computed in Equation 4.13. The BGA-PCA algorithms have removed the extra parameters, repeated terms, and irrelevant data with given simulation parameters in Table 4.1.

MAPE calculation for the integration strategy is computed for each month, and the average MAPE calculation is done in Equations 4.12 and 4.13, respectively. As per MAPE calculation in Table. 4.2, the MAPE of July 2019 is 5.56 as calculated from Equation 4.13. The weekly coefficient from different algorithms C_i are 30%, 24%, 19%, 15%, 8% and 4%. The integration strategy computed the minimum MAPE as per Equation (16) is $MAPE\ of\ July\ 2019\ = 0.3 * 5.56 + 0.24 * 5.52 + 0.19 * 5.6 + 0.15 * 5.54 + 0.08 * 5.59 + 0.04 * 5.57 = 5.55$. Similarly, the MAPE of each month is calculated as given in Table.4.3. The flow chart of the proposed framework PF is presented in Figure 4.3.

This complete model is simulated in a MATLAB simulation environment on a research workstation with Intel Core i5-4300CU, 8GB RAM, 64-bit operating system, and 2.50 GHz Processor. The database is simulated in 40 minutes on the workstation mentioned above.

PF improvement for MAPE of each month is calculated in Table 4.3. The proposed model attained MAPE of 5.56,6.26,4.61,4.87,4.94,4.92,4.74,6.14,6.27,4.47,6.32 and 6.41 percentage for July 2019 to June 2020, respectively. Thus, the overall significant improvement based on 12 months is less than 7%, compared to Table 4.3.

Table 4.3 *Comparison: Results in Figures 4.4 - 4.7 with past works.*

Results [57] **Multi-layer feed-forward NN (MFFNN) in England**		**Results** [143] **Hybrid ARMA-FLNN Model in USA**		**Results (Proposed Model) NEPRA-LESCO in Pakistan**	
Year	Avg. MAPE % For 1 year	Year	Avg. MAPE % for 1 year	Year	Avg. MAPE % for 1 year
2012	7.0231	Jan 14 - Dec 14	5.46	Jul 19 – June 20	**5.45**

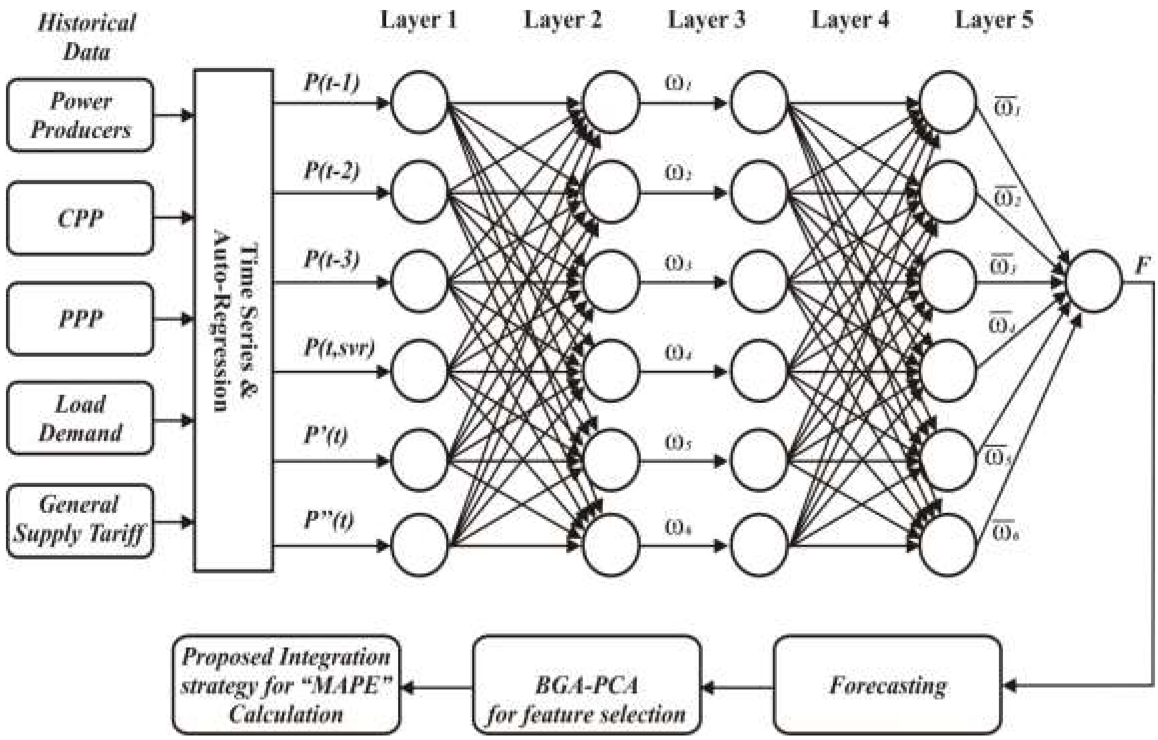

Figure 4.3 *Flowchart of the proposed framework.*

Table 4.4 *MAPE calculation for PF improvement per month.*

Customer Type	**Without Proposed Integration M_1 MAPE Percentage**	**With Proposed Integration M_2 MAPE Percentage**	**With Proposed Integration & BGA-PCA feature selection M_3 MAPE Percentage**	**Percentage Overall Improvement**
	Each Month MAPE %	**Each Month MAPE %**	**Each Month MAPE %**	**Each Month Improvement %**
July 2019	6.61	6.12	5.56	9.15%
August 2019	7.42	6.87	6.26	8.88%
Septembe r2019	5.51	5.10	4.61	9.61%
October 2019	5.81	5.38	4.87	9.48%
Novembe r 2019	5.89	5.46	4.94	9.52%
December 2019	5.87	5.44	4.92	9.56%
January 2020	5.66	5.24	4.74	9.54%
February 2020	7.28	6.74	6.14	8.90%
March 2020	7.43	6.88	6.27	8.87%
April 2020	5.35	4.95	4.47	9.70%
May 2020	7.49	6.93	6.32	8.80%
June 2020	7.59	7.03	6.41	8.82%
Average			**5.45**	**9.24%**

4.4 Results Discussion

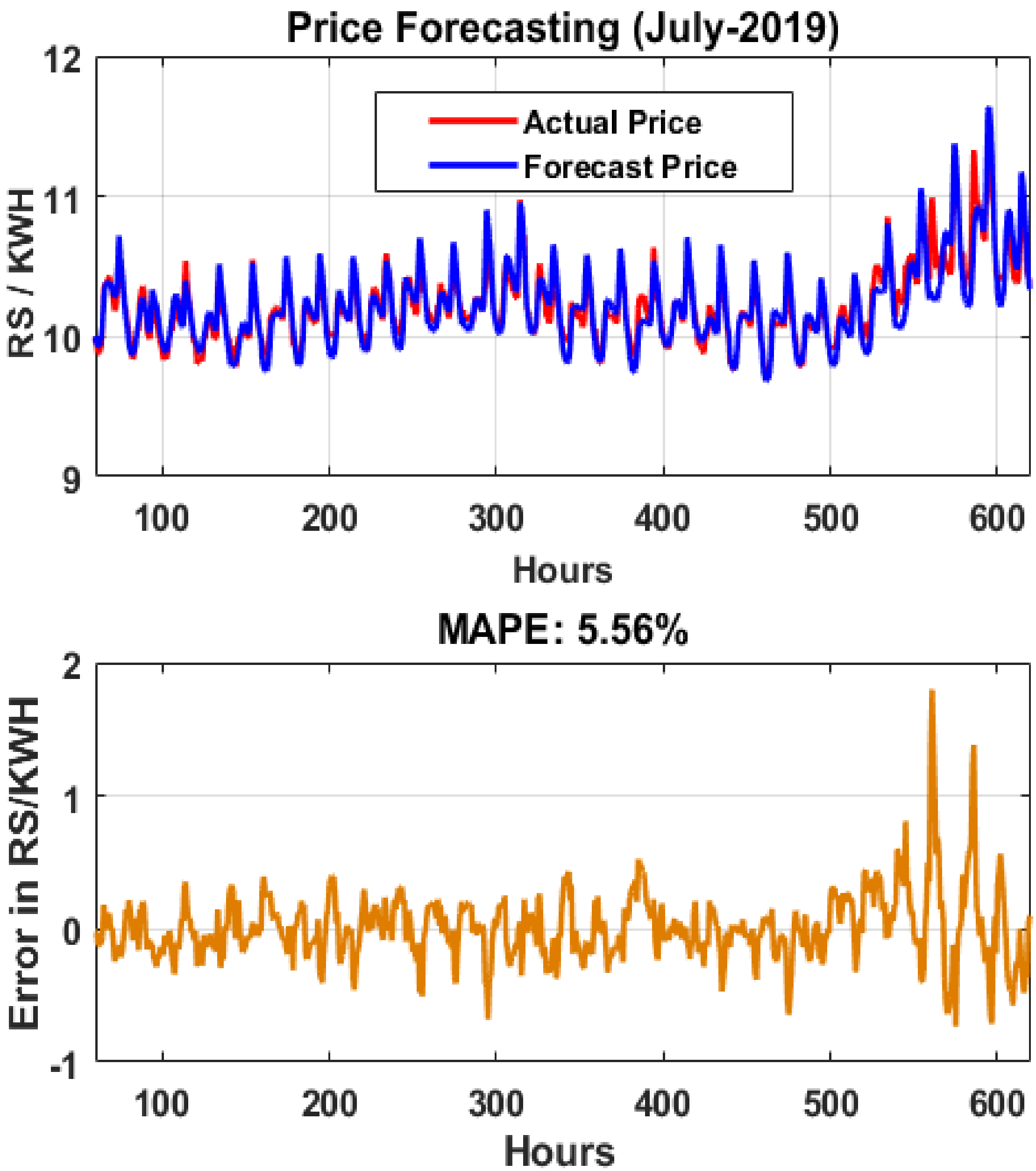

Figure 4.4 *Actual vs. forecasted price of electricity for July 2019 and average MAPE of the month.*

The MAPE of July-2019 without proposed integration is 6.12% after feature selection and the proposed integration model MAPE is optimized to 5.56%. Thus, the overall improvement of the proposed model is 9.15%.

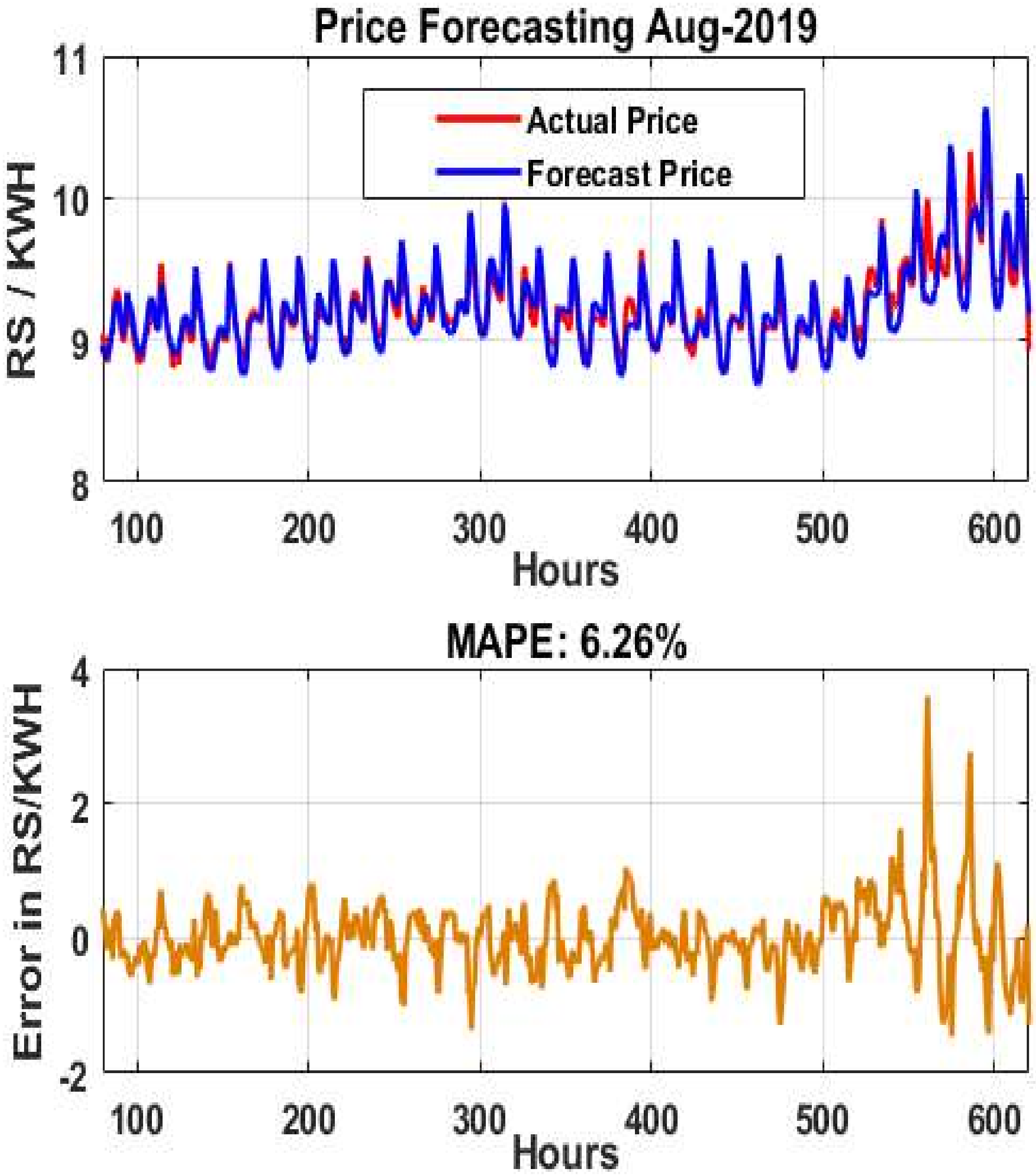

Figure 4.5 *Actual vs. forecasted price of electricity for August 2019 and average MAPE of the month.*

The MAPE of August-2019 without proposed integration is 6.87% after feature selection, and the proposed integration model, the MAPE, is optimized to 6.26%. Thus, the overall improvement of the proposed model is 8.88%.

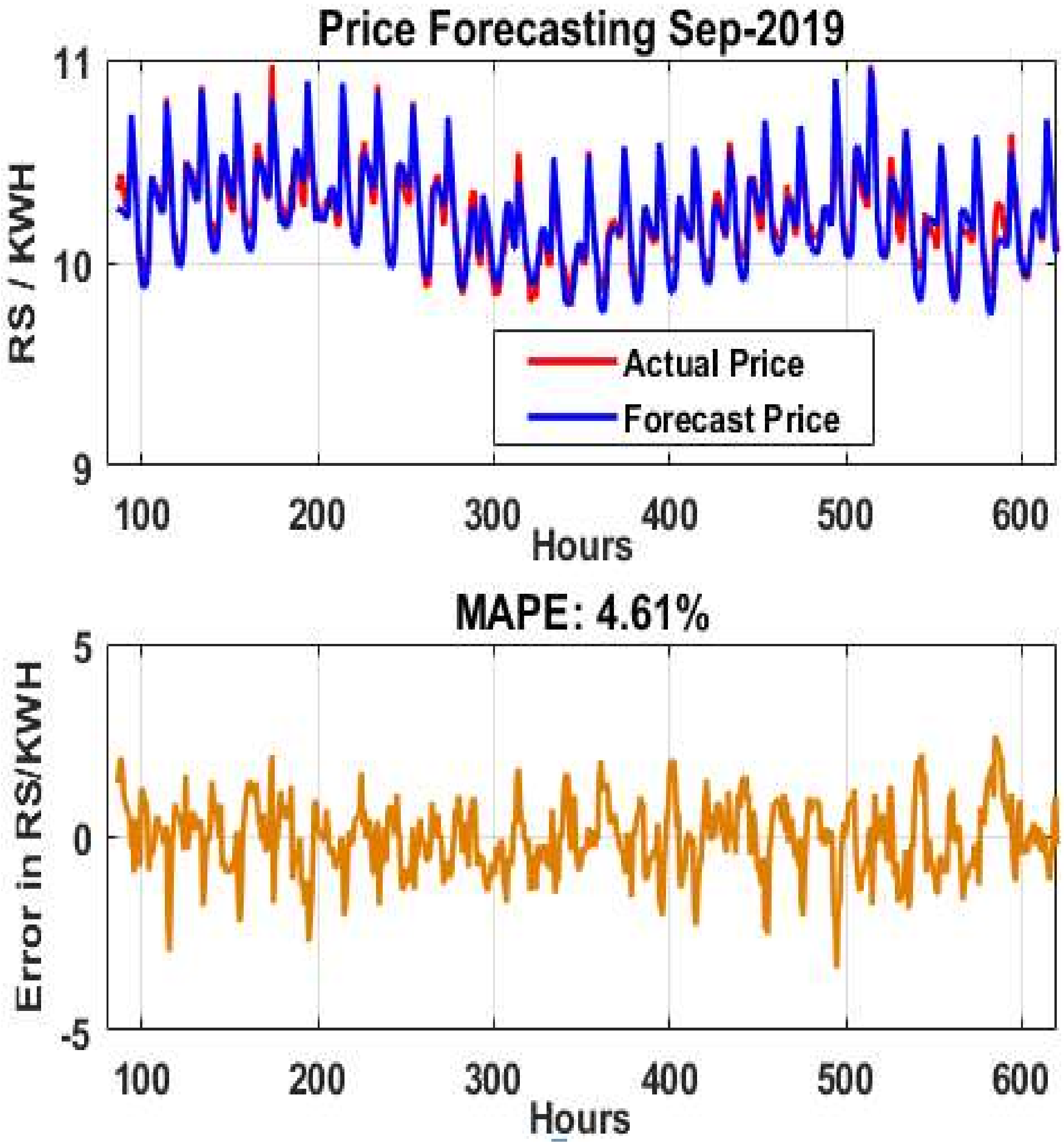

Figure 4.6 *Actual vs. forecasted price of electricity for September 2019 and average MAPE of the month.*

The MAPE of September-2019 without proposed integration is 5.10% after feature selection, and the proposed integration model, the MAPE, is optimized to 4.61%. Thus, the overall improvement of the proposed model is 9.61%.

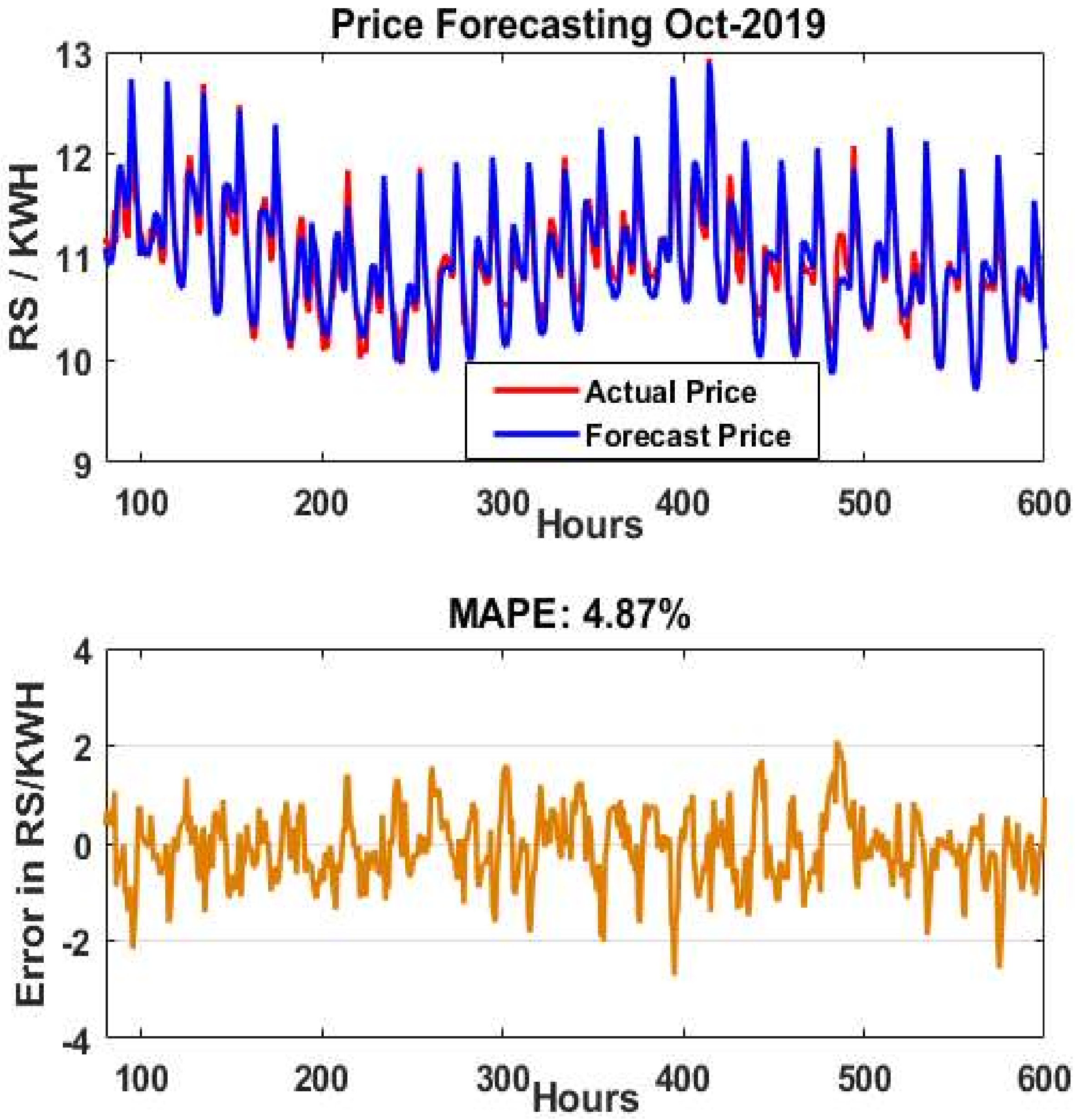

Figure 4.7 *Actual vs. forecasted price of electricity for October 2019 and average MAPE of the month.*

The MAPE of October-2019 without proposed integration is 5.38% after feature selection and the proposed integration model MAPE is optimized to 4.87%. Thus, the overall improvement of the proposed model is 9.48%.

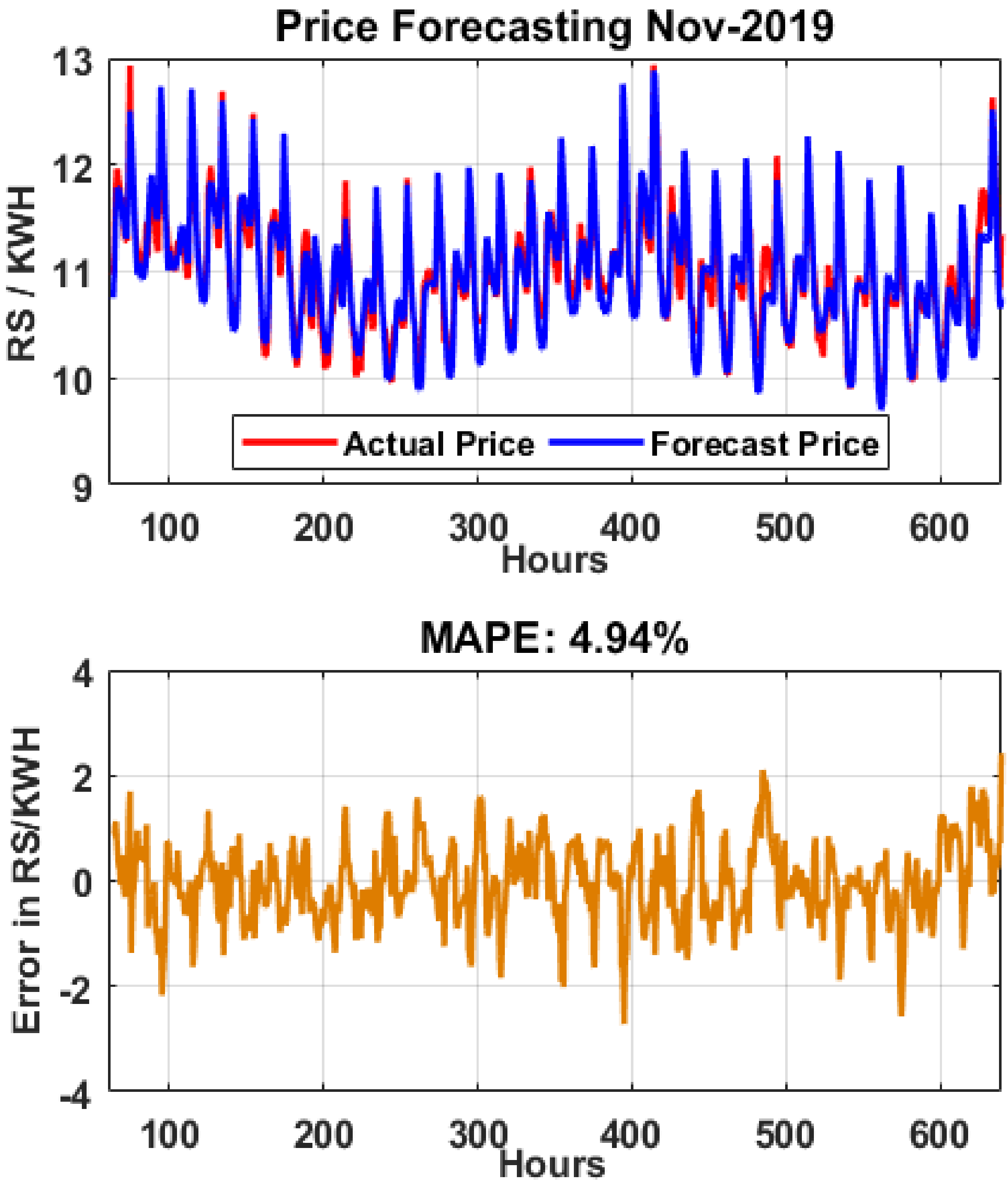

Figure 4.8 *Actual vs. forecasted price of electricity for November 2019 and average MAPE of the month.*

The MAPE of November-2019 without proposed integration is 5.46%; after feature selection and proposed integration model, the MAPE is optimized to 4.94%. Thus, the overall improvement of the proposed model is 9.52%.

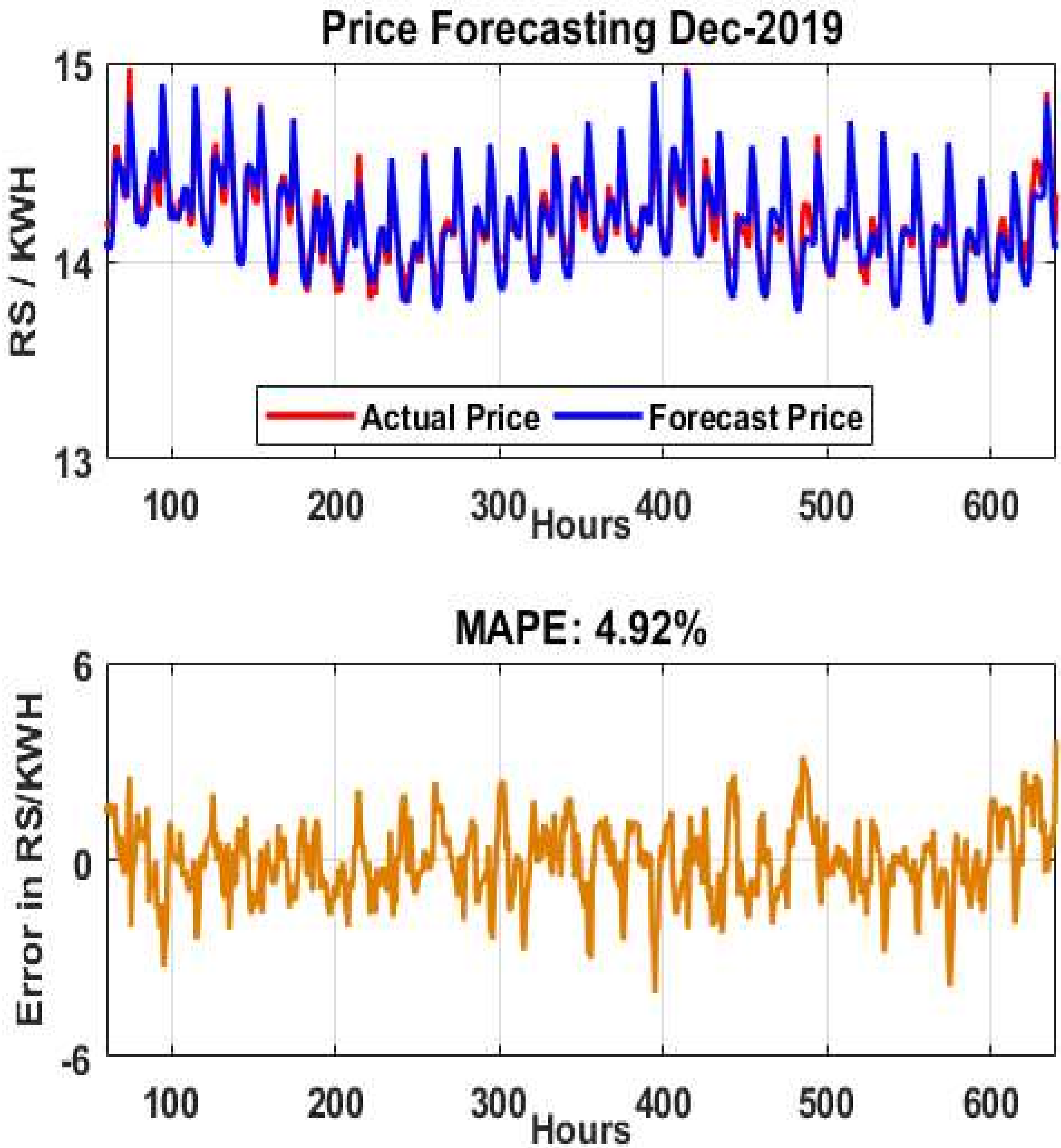

Figure 4.9 *Actual vs. forecasted price of electricity for December 2019 and average MAPE of the month.*

The MAPE of December-2019 without proposed integration is 5.44% after feature selection and the proposed integration model MAPE is optimized to 4.92%. Thus, the overall improvement of the proposed model is 9.56%.

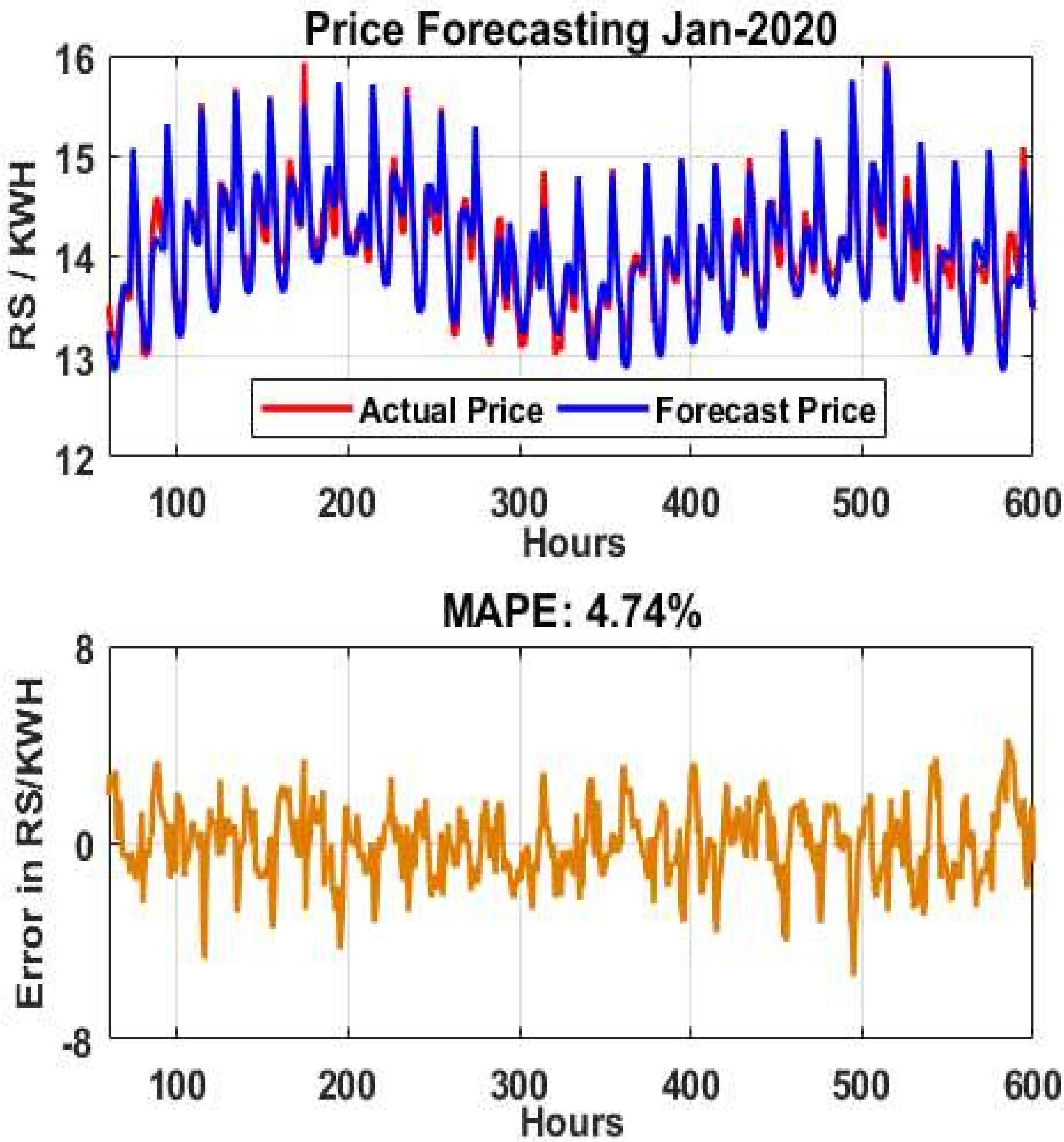

Figure 4.10 *Actual vs. forecasted price of electricity for January 2020 and average MAPE of the month.*

The MAPE of January-2020 without proposed integration is 5.24%; after feature selection and the proposed integration model, the MAPE is optimized to 4.74%. Thus, the overall improvement of the proposed model is 9.54%.

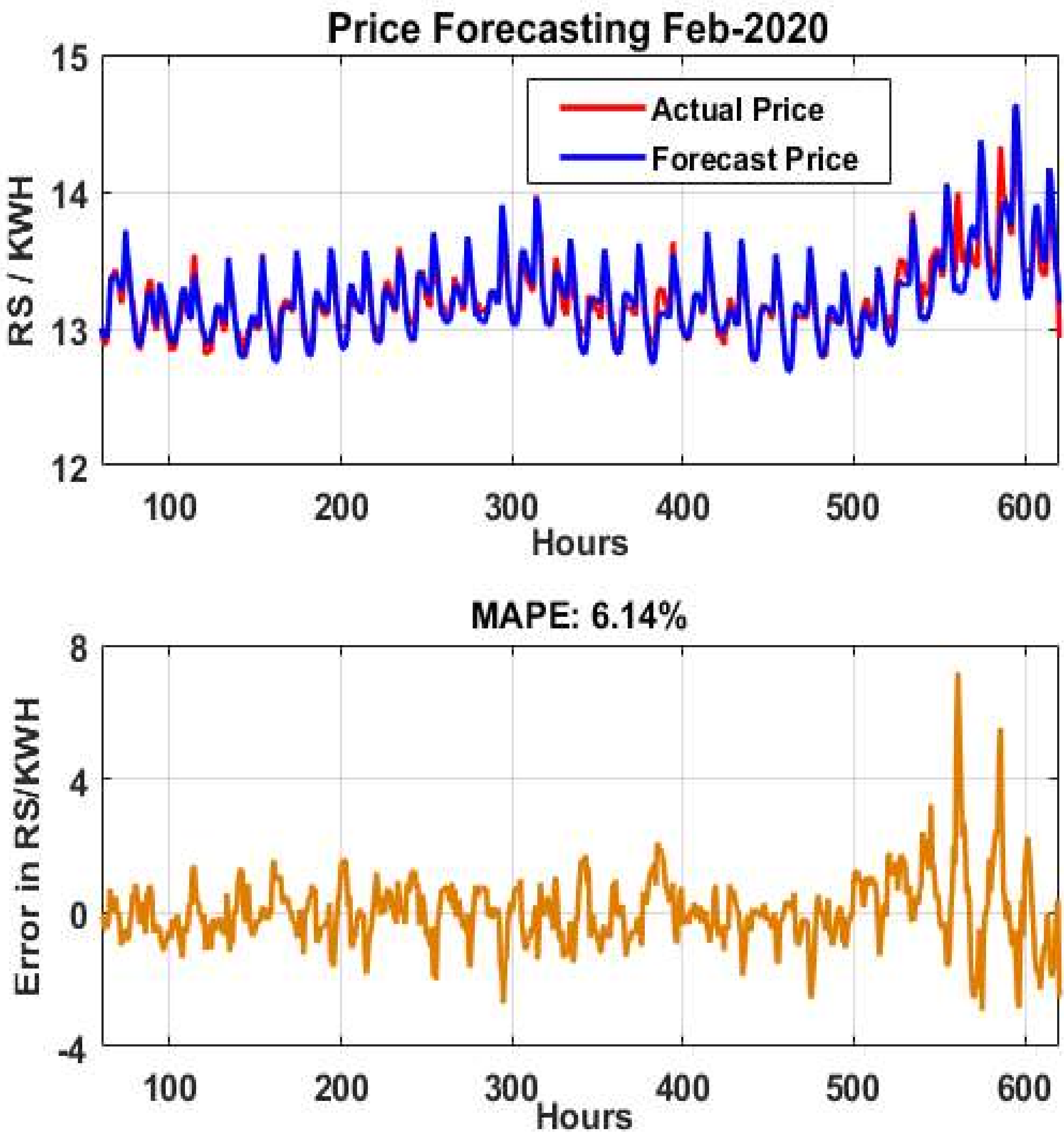

Figure 4.11 *Actual vs. forecasted price of electricity for February 2020 and average MAPE of the month.*

The MAPE of February-2020 without proposed integration is 6.74% after feature selection, and the proposed integration model, the MAPE, is optimized to 6.14%. Thus, the overall improvement of the proposed model is 8.90%.

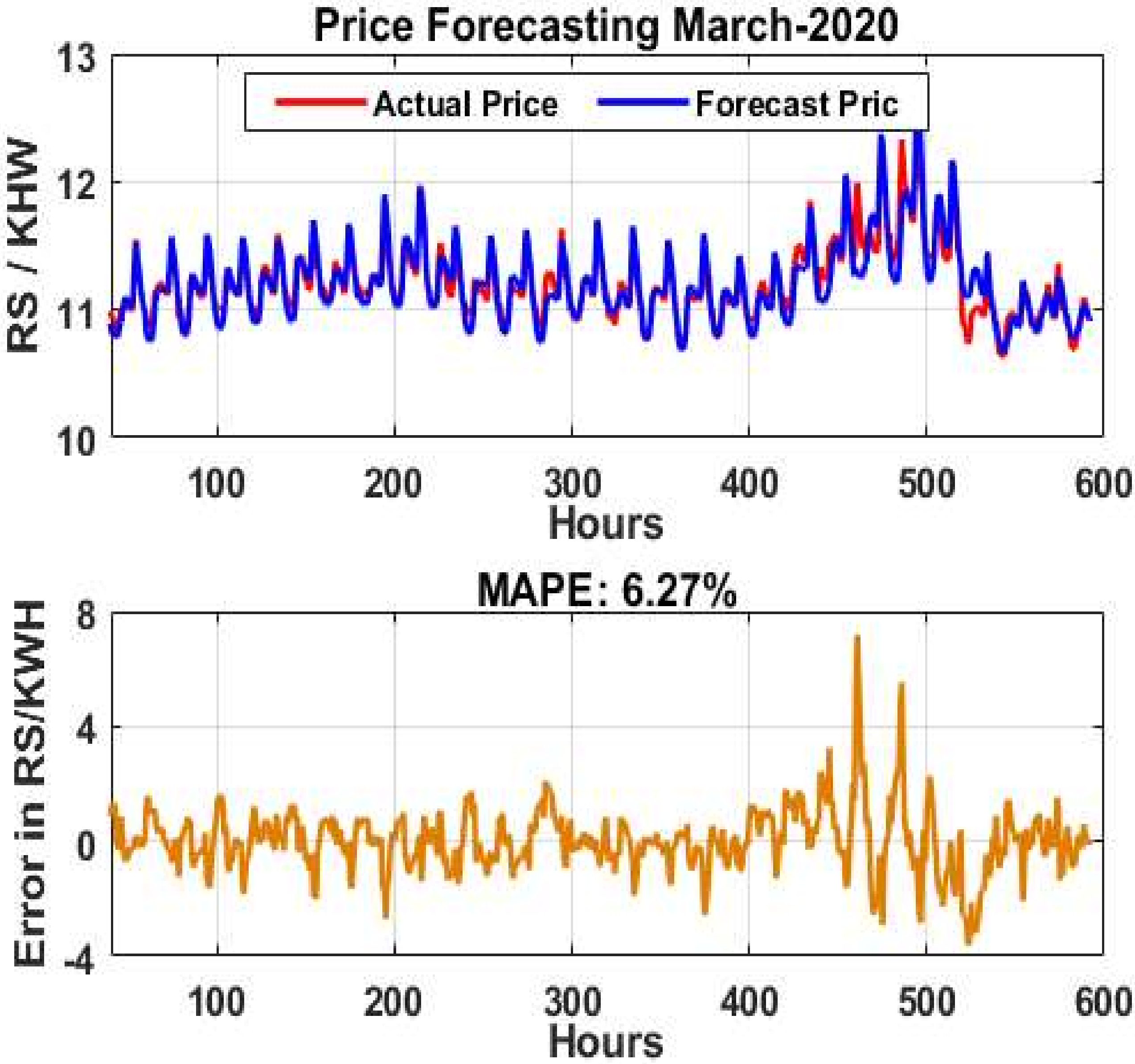

Figure 4.12 *Actual vs. forecasted price of electricity for March 2020 and average MAPE of the month*

The MAPE of March-2020 without proposed integration is 6.88% after feature selection, and the proposed integration model, the MAPE, is optimized to 6.27%. Thus, the overall improvement of the proposed model is 8.87%.

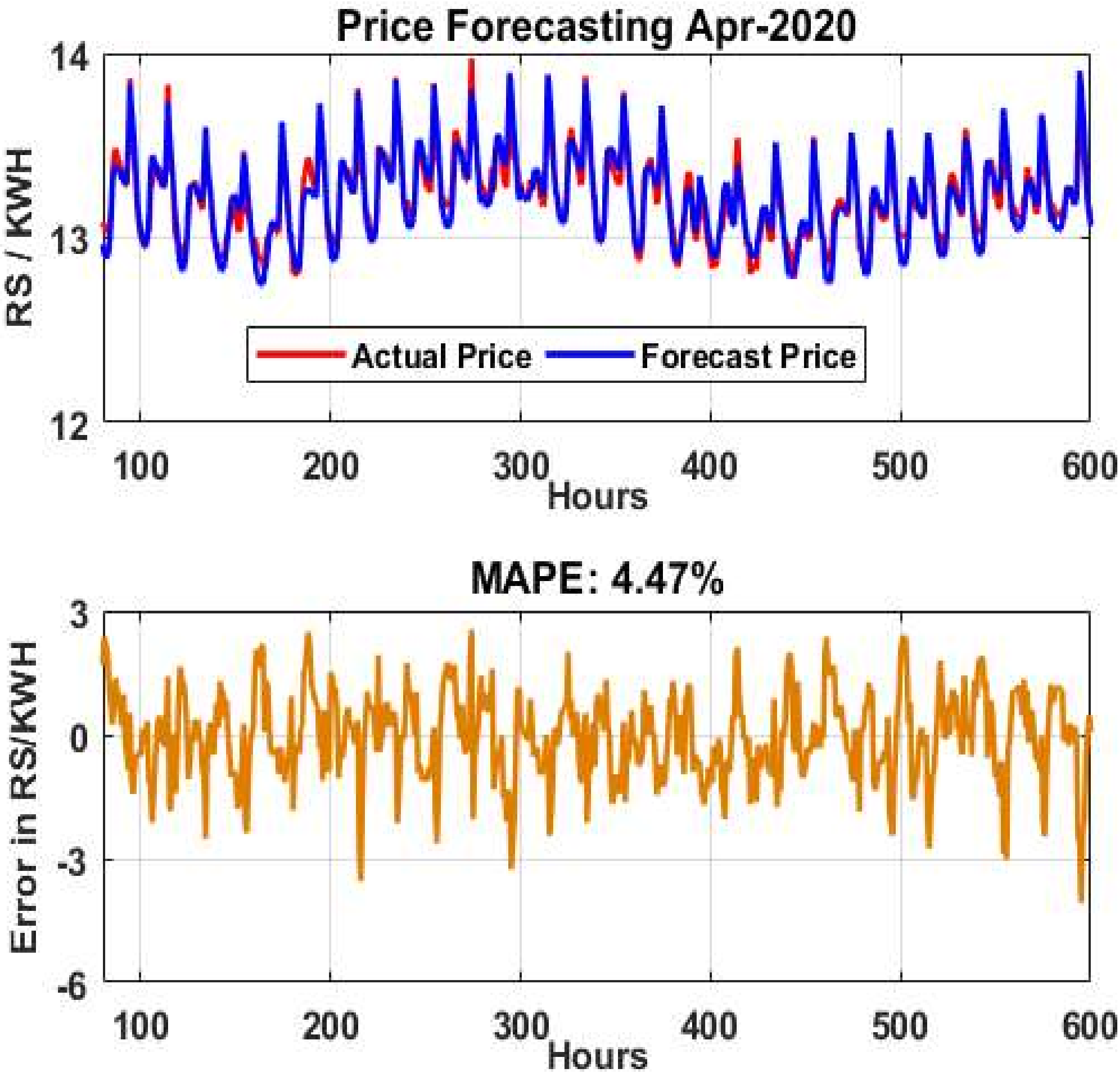

Figure 4.13 *Actual vs. forecasted price of electricity for April 2020 and average MAPE of the month*

The MAPE of April-2020 without proposed integration is 4.95% after feature selection and the proposed integration model, the MAPE is optimized to 4.47%. Thus, the overall improvement of the proposed model is 9.70%.

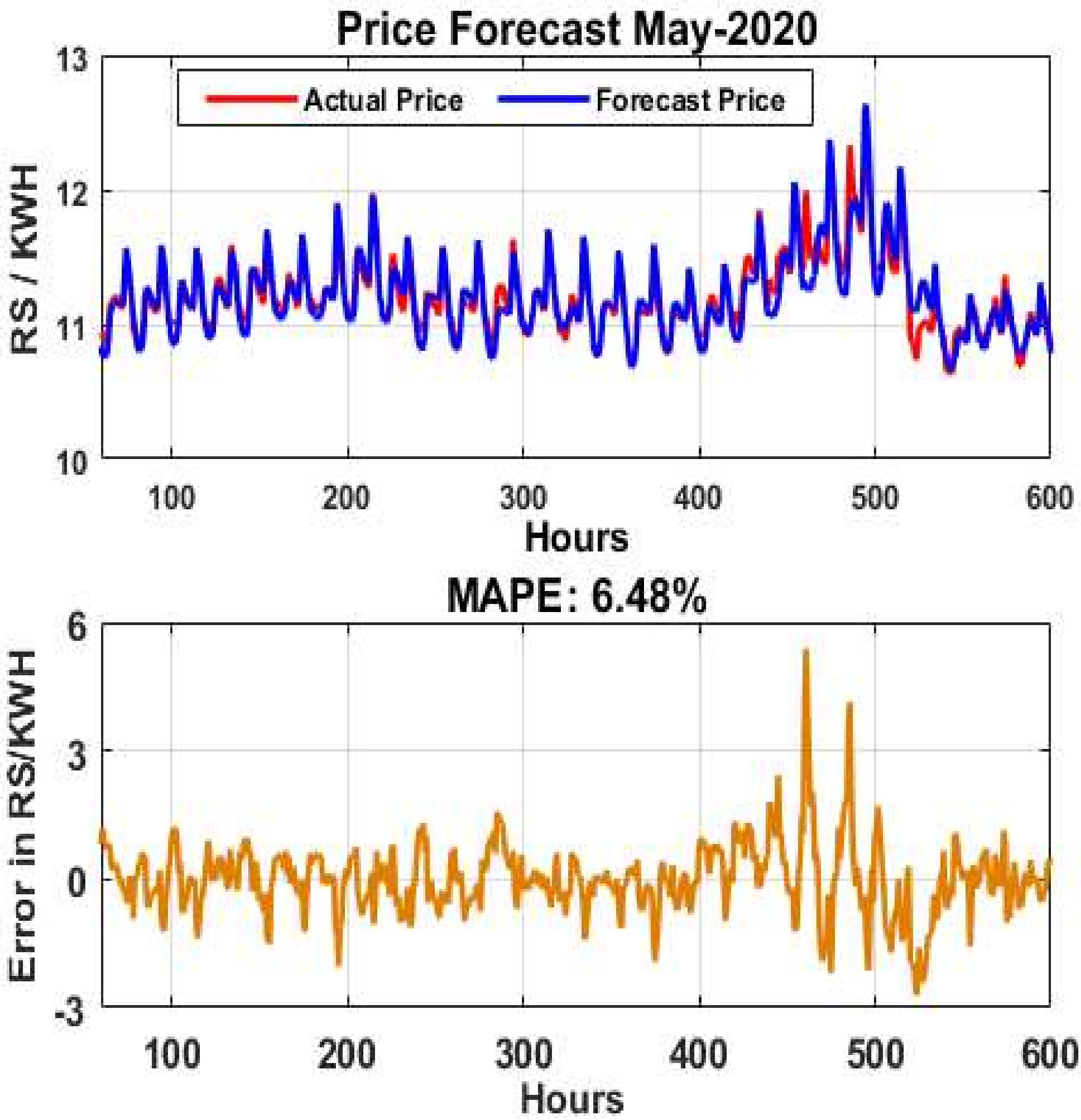

Figure 4.14 *Actual vs. forecasted price of electricity for May 2020 and average MAPE of the month.*

The MAPE of May-2021 without proposed integration is 6.93% after feature selection, and the proposed integration model, the MAPE, is optimized to 6.32%. Thus, the overall improvement of the proposed model is 8.80%.

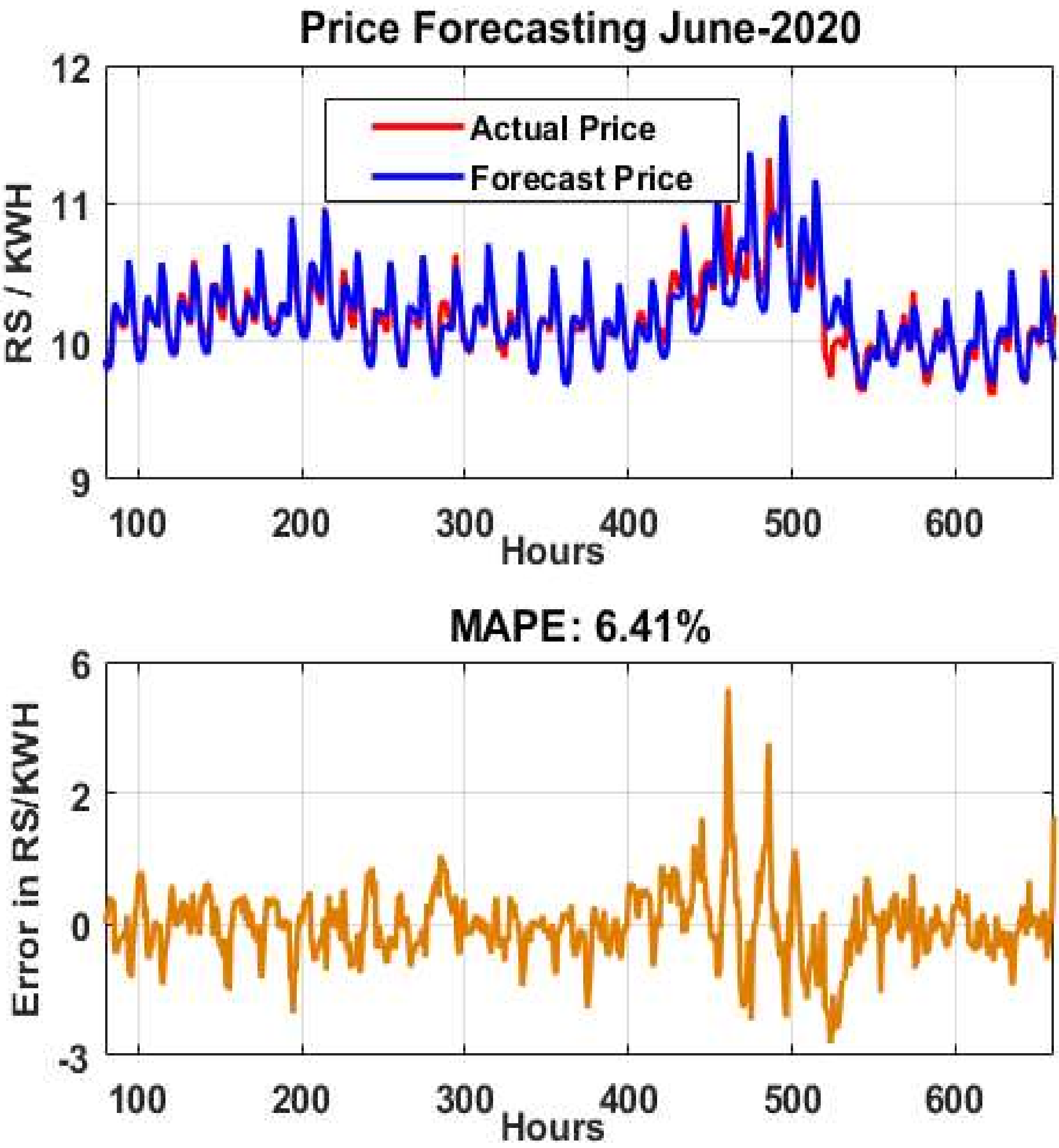

Figure 4.15 *Actual vs. forecasted price of electricity for June 2020 and average MAPE of the month*

The MAPE of June-2020 without proposed integration is 7.03% after feature selection, and the proposed integration model the MAPE is optimized to 6.41%. Thus, the overall improvement of the proposed model is 8.82%.

4.5 Summary

PF is considered an essential application of the Smart Grid System. Therefore, to forecast the price of electricity, a novel combinational Forecasting Method is presented. The proposed method is based on Time Series and AR algorithm, ML-based feedback ANFIS, BGA-PCA algorithm, and an integration strategy. In our research, the complex and big historical data are analyzed using a new time-series and AR algorithm and validated by the ML-based feedback ANFIS model, in which 6 stages are computed. Additionally, with the help of the BGA-PCA approach, the best feature selection is attained. As a result, an improvement in MAPE is recorded as less than 7%, and an average MAPE is recorded as 5.45, which is achieved by the proposed Integration strategy. The results validate the performance of our model, and hence we can conclude that the optimal PF can overcome the problems related to the planning and operating of the smart grid.

CHAPTER 5

LOAD FORECASTING AND PRICE FORECASTING BASED POWER MANAGEMENT SYSTEM

5.1 Introduction

In this chapter, FA implementation in MATLAB for the DSM will be presented. Proposed demand-side management has been designed based on the mathematical model presented in Chapter 3. All the algorithms and their coding revolve around two main objectives: cost reduction and power management. In the proposed technique, three buildings have been selected for the DSM.

These buildings have been given name as

1. Building 1 with 2500 devices used by consumers.
2. Building 2 with 109 devices used by consumers.
3. Building 3 with 808 devices used by consumers.

For the implementation of DSM, different kinds of loads have been selected for these buildings. These loads are commonly used appliances in houses, industrial machines, or commercial instruments. Then the time of operation is selected by turning devices on and off in each interval.

5.1.1 Building 1

This building contains mainly household appliances. Here the First column contains the name of the load. The sec contains watt rating, the 3rd contains the probability of on-time of load, the 4th column contains off time of load, 5th column contains the number of times devices turn on in each house. For example, a washing machine is mainly used once a day in the morning for washing clothes. The same case is with the dryer. But the oven may be used many times.

The percentage of generation by different power producers with CPPs, PPPs, and load demand is presented in Table 5.1.

Table 5. 1 *Power producers with CPPs, PPPs, and load demand.*

	Categories		Generation (%)
Power Producers	Renewable Energy Resources		2.79%
	Hydro Power Generation		26.86%
	Gas Production		16.41%
	RLNG (RNLG)		4.46%
	Imported Coal		17.73%
	Local Coal		0.00%
	Residual fuel oil		3.09%
	Bagasse		0.84%
	Uranium		5.98%
	RNLG New		21.18%
	Imported Power		0.34%
	Mix (Captive)		0.33%
Months	**Capacity Purchase Price (CPP) Per kWh**	**Power Purchase Price (PPP) Per kWh**	**Load Demand GWh**
July 2019	4.58	9.44	14275
August 2019	4.87	9.05	14450
September 2019	5.93	9.74	13230
October 2019	7.43	11.77	10107
November 2019	8.96	11.91	8306
December 2019	8.60	12.54	8493
January 2020	8.24	13.86	8753
February 2020	9.45	13.47	7686
March 2020	7.85	12.95	9556
April 2020	6.84	12.94	10974
May 2020	5.81	11.93	12901
June 2020	5.61	11.42	13525

The details of loads used by consumers in building 1 are presented in Table 5.2.

Table 5. 2 *Load details of building 1.*

Device Type	**Hourly Consumption of Device (kW)**				**Number of Devices**
	watts	**Start Time**	**Off time**	**Count of usage in a day**	
Dryer	500	8	10	1	190
Dish Washer	1000	18	20	4	290
Washing Machine	1000	9	11	1	265
Oven	4000	7	22	8	280
Iron	1200	6	8	2	341
Vacuum Cleaner	1000	1	4	1	158
Fan	500	1	24	Continuous	288
Kettle	1000	5	22	4	406
Toaster	1000	7	9	2	48
Rice-Cooker	2000	10	20	2	59
Hair Dryer	1500	6	23	2	58
Blender	400	8	15	3	66
Frying Pan	1000	12	22	4	101
Coffee Maker	1500	1	24	10	56
	Total				2500

5.1.2 Building 2

Building 2 has been selected as the commercial industry. This is because it contains all high-power loads. The details of loads used by consumers in building 2 are expressed in Table 5.3.

Table 5.3 *Load of details of building 2.*

Device Type	Hourly Consumption of Device (kW)				**Number** Devices
	Watts	Start Time	Off Time	Count of usage in a day	
Water Heater	2000	1	24	Continuous	39
Welding Machine	4000	7	22	20	35
Fan/AC	1000	1	24	Continuous	16
Arc Furnace	10000	8	20	30	8
Induction Motor	2000	1	24	20	5
DC Motor	1000	1	24	20	6
		Total			109

5.1.3 Building 3

Building 3 has been selected as the residential Building. It contains all midrange loads. Following are the details of the loads.

Table 5.4 *Load details of building 3.*

Device Type	Hourly Consumption of Device (kW)				Number
	Watts	**Start**	**Off Time**	**Count of usage**	**Devices**
Water Dispenser	500	1	24	10	156
Dryer	1000	8	9	2	117
Kettle	2000	7	10	5	123
Oven	4000	2	23	10	77
Coffee Maker	2000	6	14	10	99
Fan/AC	1500	1	24	Continuous	93
Air Conditioner	2000	1	24	Continuous	56
Lights	300	18	7	Continuous	87
		Total			808

This building contains 808 devices. Therefore, to create the composite system, a large number of devices have been used. Since a large building may contain these devices, this table has been used as input data for the proposed FA. Following is a flowchart of the algorithm.

Here a total of 12 blocks are shown. Each block has been implemented in coding. Most of the input data is in the form of vectors. Following is details of the flowchart

1) First of all, selection of appliances is done
2) Then it captures the On-time of each appliance after initialization.
3) The duration of the running of every appliance is load duration.
4) Load power rating is also used as input in vector form
5) Cost of electricity so that per month bill can be calculated. This cost varies from country to country.
6) Peak hours detail is also necessary. This is because peak hour cost affects 3 phase connection customers, not single-phase connection customers in Pakistan. But still, it is required because most of the users are 3 phase connection customers and load on the national power grid is high at peak hours.

7) All the above parameters are used as input to the FA. These variables are modeled based on model equations presented in Chapter 3.

8) Now the output of the firefly block is generated. It is useless until it is proved fruitful. So to create a closed-loop system, unscheduled load parameters are generated.

9) Unscheduled load, cost, and peak values are calculated. These values are now compared with the firefly block if these values have two properties.

   a) Firefly Load Peak < Unscheduled Load Peak

   b) Firefly Load Bill < Unscheduled Load Bill

   Then proposed scheduling DSM is correct. If not, then Firefly bock is again called back to recalculate parameters and reschedule load.

10) If (9) is satisfied, then output is generated. This output contains DSM data which can be used to reduce electricity bills and also power peaks in demand.

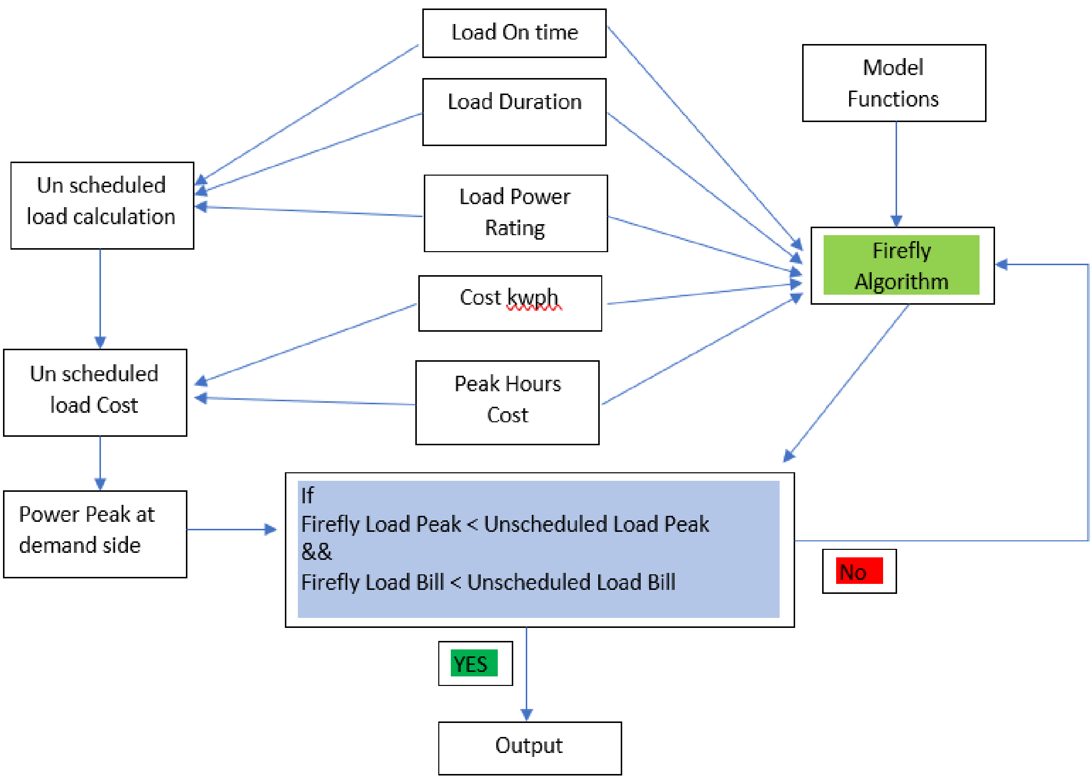

Figure 5.1 *The flowchart of the proposed firefly algorithm.*

5.2 Methodology

The proposed system is based on demand-side management using firefly. The demand-side management model presented in this thesis is deterministic. The basic parameters used in this thesis are fixed, such as power ratings of every piece of equipment used in the houses, peak hours in which the price of every power drawn is higher, and the price of electricity. Following is modeling done in this thesis.

All the appliances that cannot be controlled, for example, lights, etc., cannot be scheduled because if the time of higher electricity cost is ongoing, users cannot turn off all their lights for the fans. The only freedom, in this case, is there turn off the lights and fans in those areas where no one is sitting, and maximum conservation policies are adopted.

The second category of the load is air conditioning appliances in which our power management and proposed techniques can be applied. These appliances can also not be shifted temporarily. Usually, there are two types of air conditioning systems. One is HVAC, and the other one is local air conditioning units installed on each house. Usually, HVAC control is difficult compared to individual units because it draws a large current and will cool an entire building even if someone does not need it. But on the other hand, the tiny air conditioning units can be e used to conserve the Power by shifting their cooling temperature levels, and in this way, they will trip and conserve the power.

Another category of the appliances can be shifted at any other time in which peak load is reduced or peak hours are not applicable. These loads are dishwasher dryers, water pumps, etc. These kinds of loads have a fixed price cost ratio as compared to their electricity consumption. Therefore, a calculation is required that assesses the operational time of the water pump filling the water tank every time because the water tank capacity is fixed. For example, if a motor fills a water tank in 30 minutes regarding the time, it will always take 30 minutes.

This thesis considers the following vectors to design the model equation for the Firefly-based power management system proposed in this thesis.

Set of appliances which can be shifted have been represented as

$$A = [A_1 + A_2 + \cdots \dots \dots \dots \dots \dots \dots \dots \dots \dots + A_k] \tag{5.1}$$

Power ratings of each appliance are expressed as

$$P = [P_1 + P_2 + \cdots \dots \dots \dots \dots \dots \dots \dots \dots \dots + P_k] \tag{5.2}$$

X = P/m is the power rating of each device & “m” is the number of slots per hour.

The time duration of each appliance is given as

$$L = [L_1 + L_2 + \cdots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots + L_k] \tag{5.3}$$

When each appliance is turned on is expressed as

$$T_s = [t_1 + t_{s2} + \cdots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots + t_{sk}] \tag{5.4}$$

For the representation of the time slots in 24 hours, the following vector has been used

$$T = [1 + 2 + 3 + \cdots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots + n] \tag{5.5}$$

The above Vectors represent the model presented in this thesis for demand-side management using the FA.

$$\text{Power Requirement} = \sum \mathrm{n} \notin T \sum a \notin A\, Xn, a \tag{5.6}$$

It's a straightforward model that can be used for power optimization. For managing the peak hour issues in the proposed system, a separate vector has been created in which 2 elements exist as this vector is

$$\mathrm{T} = [\mathrm{T1}, \mathrm{T2}] \tag{5.7}$$

Here

T1 = Peak Hours

T2 = Off-peak hours

The concept of peak hours has been introduced in the proposed system to encourage people to save electricity. During peak hours, all residential users utilize electricity, and therefore, the electrical stress on the power grid increases in this time. The higher cost is applied to all the power consumed in peak hours to discourage users from turning on all the high-power devices at this time.

For the electricity cost reduction for the consumers, the following model equation will be used in the proposed thesis.

$$\text{Total Cost reduction in Bill} = \sum_{n=1}^{n=N} Power\ Requirement - n.\, Cn \tag{5.8}$$

So if we minimize the above equation, then the total cost will be reduced. Here Cn is cost vectors which consist of bills of individual appliances. Now when we schedule loads, then the customer must not be dissatisfied with arrangements. So there must be a satisfaction model. So if we minimize $\sum a \notin A\ Ds$ (Ds = Discomfort level) then the customer will bear arrangements. So

$$Ds = \frac{(Ts - ast)}{(\beta end - \ L\text{-}\ ast)Va} \in A \tag{5.9}$$

Now here, two parameters can be operated by the customer. One is the start time, and the other is off time. Here (ast) and (βend) are the starting and ending times of each appliance set by users to finish the task. Here Ts is the time of start physically, and L is the time interval for each appliance to complete the task.

Now the main objective is to minimize peak load. For this purpose, the following model can be used.

$$\sum_{n=1}^{n=N} Power\ Requirement(Scheduled) <= \sum_{n=1}^{n=N} Actual\ Power\ Requirement \tag{5.10}$$

It means that after scheduling, the load must be less or equal to the actual load. But the following condition must be validated for the objective to complete

$$Max(T * Pscheduled) < Max\ (T * Pactual) \tag{5.11}$$

The above condition validates that if schedule power drawn is less than actual load in peak hours, it reduces overall cost and system capacity to handle the maximum load.

5.2.1 Firefly Algorithm

Xin-she-young develop FA in [144], [145]. FA is a specific rules algorithm that effectively seeks the finest solution in local research. There are two basic variables of FA as attractiveness and emission of light. There are three basic steps are followed in FA:

1. At the same time, the firefly is attracted to the most attractive FA.
2. In the second step, the attraction is proportional to the flashing light.
3. In the last step, coefficient value is used to control the intensity of the light

The firefly is attracted to another firefly with a brighter flashlight than its own, as presented in Figure 5.2. Explain the Pseudocode of the FA in [146] and specify the light intensity value, initialization of light intensity, light absorption stability, and distance between two fireflies. The flowchart of FA is shown in the following figure.

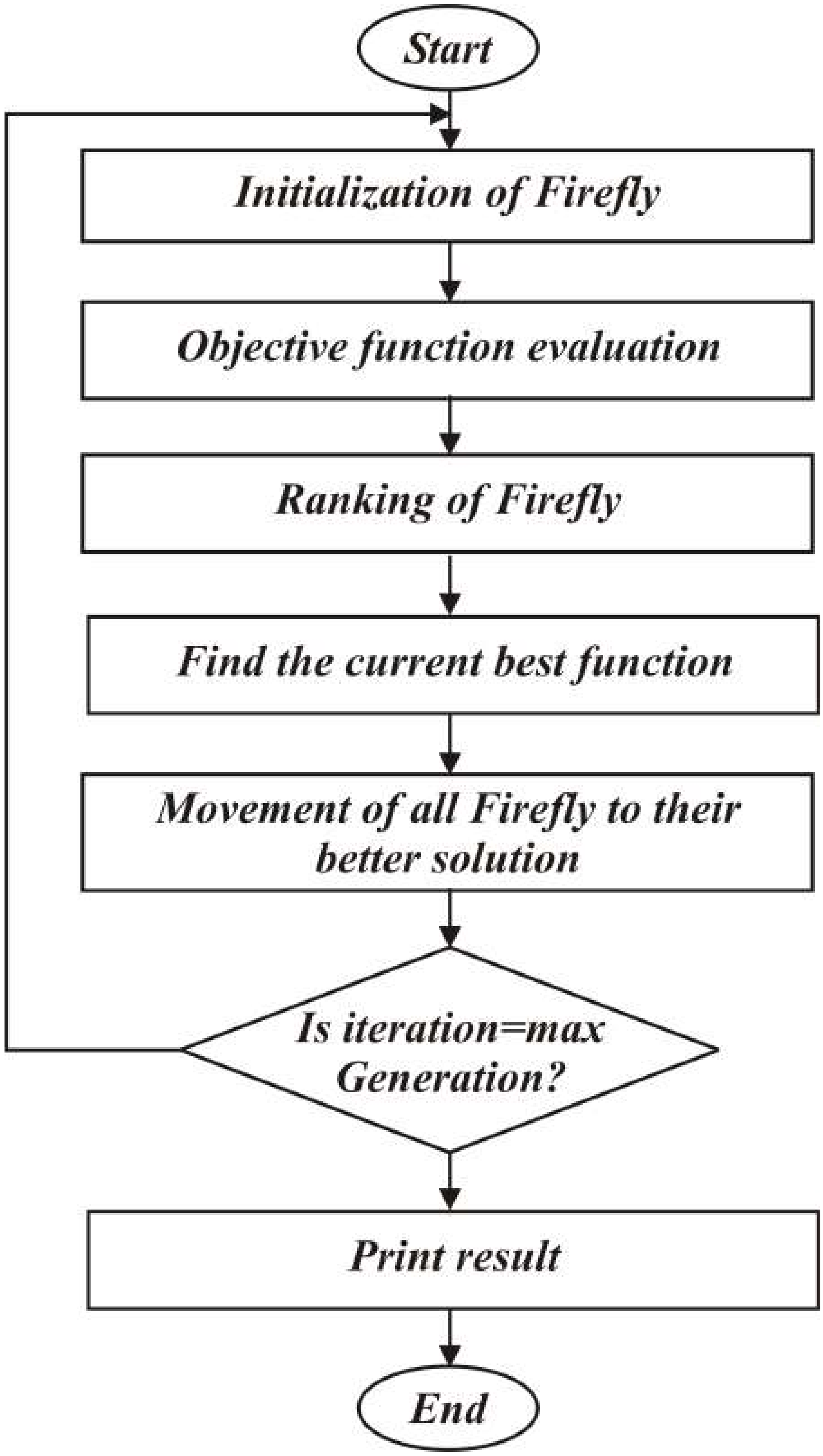

Figure 5.2 *Flow chart of firefly algorithm.*

5.3 Result Discussion

In this section result, a discussion of the proposed system has shown. This thesis has presented the FA-based demand-side management technique and has approved all the objectives presented in Chapter 1.

5.3.1 Tariff Plan

Figure 5.2 presents the overview of the FA regarding the cost of electricity related to its time slot. Here two-stage tariff plan has been used.

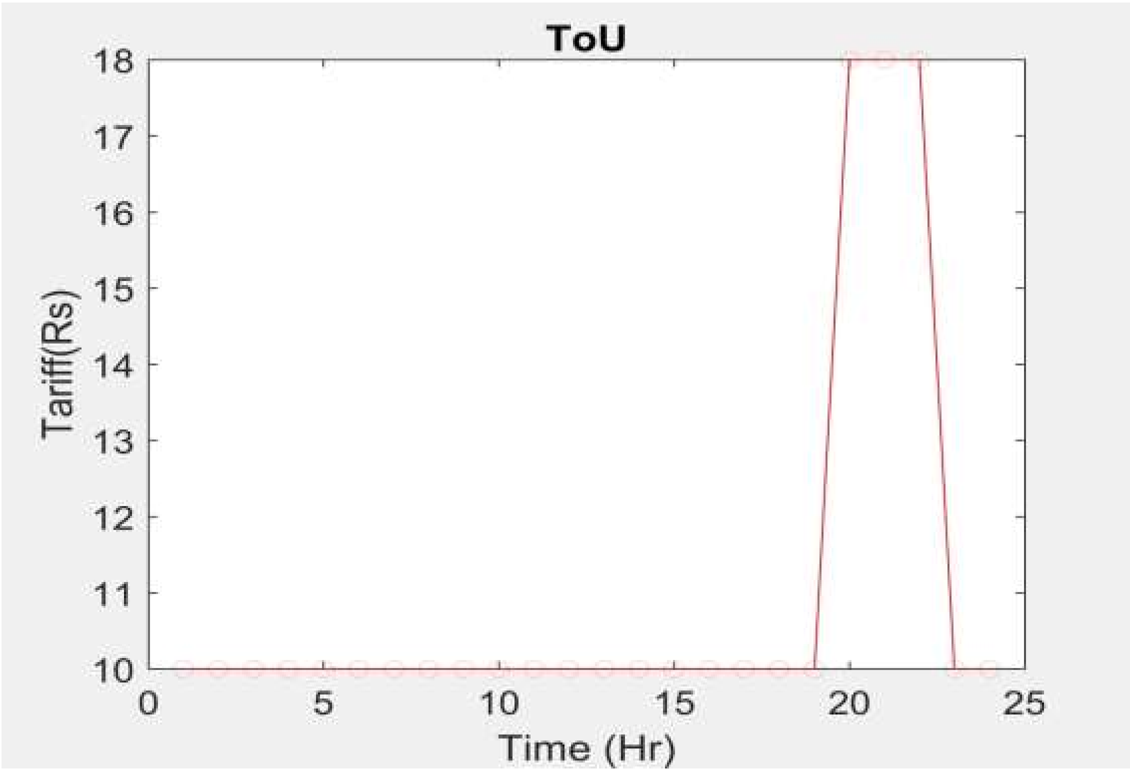

Figure 5.3 *Tariffs Plan.*

Figure 5.3 shows the tariff plan implemented in the proposed system. The y-axis shows the tariff rate in Pkr, and the x-axis shows the time slots. It is a standard 2 Level tariff plan. In 24 hours, approximately 3 to 4 hours are peak hours, from 7 pm to 11:00 p.m. In these peak hours, the price is approximately double. Due to this, all the consumers are utilizing maximum appliances at this stage. Therefore, electricity demand is also higher at this time. For encouraging power conservation, these tariff plans have been introduced so that in peak hours' customers must conserve electricity so that cost of electricity and electrical stress on the national grid can be reduced.

5.3.2 Load Management

Load Management Building 1

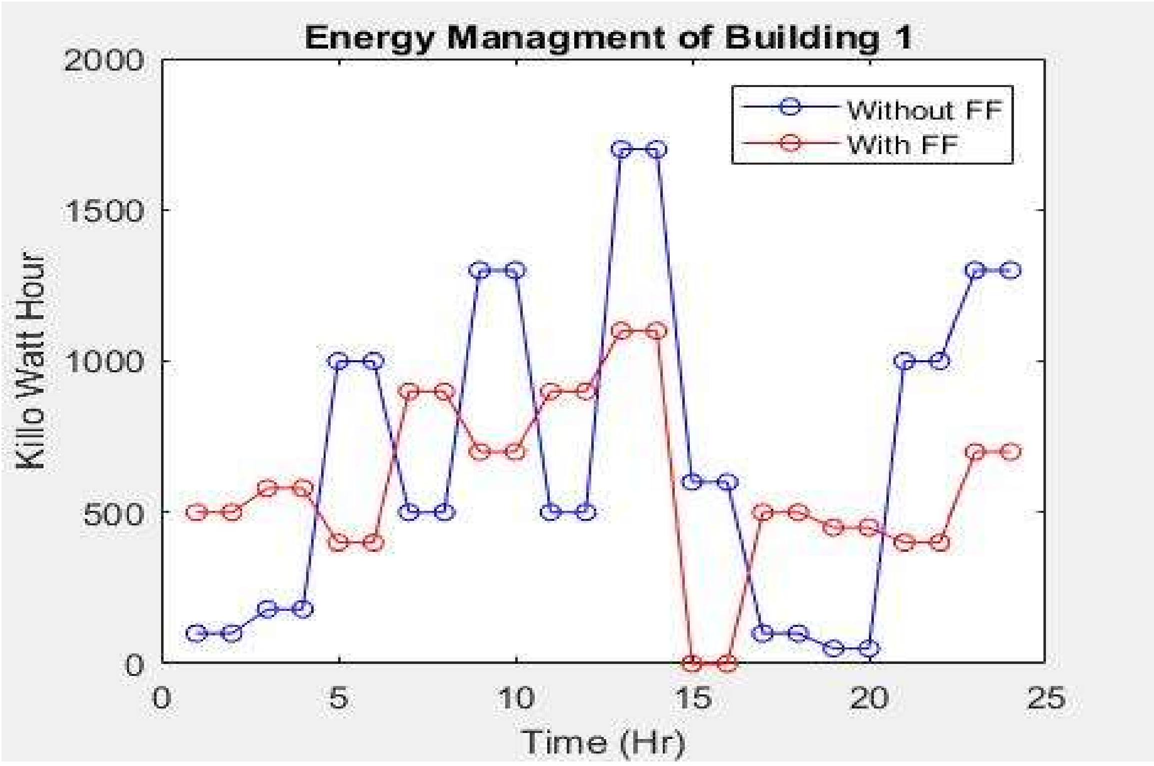

Figure 5.4 *Results of building 1 load.*

Figure 5.4 presents the building 1 load management with and without FF. The y-axis shows the power consumption in kilowatt-hour, and the x-axis shows the time slots. Here blue graph is about unscheduled load, while the red one is scheduled to load with firefly. In this graph, peak load has been shifted from 1700kw to 1100kw. It reduces peak load by 35%. It will also reduce the demand for power source capacity.

Load Management Building 2

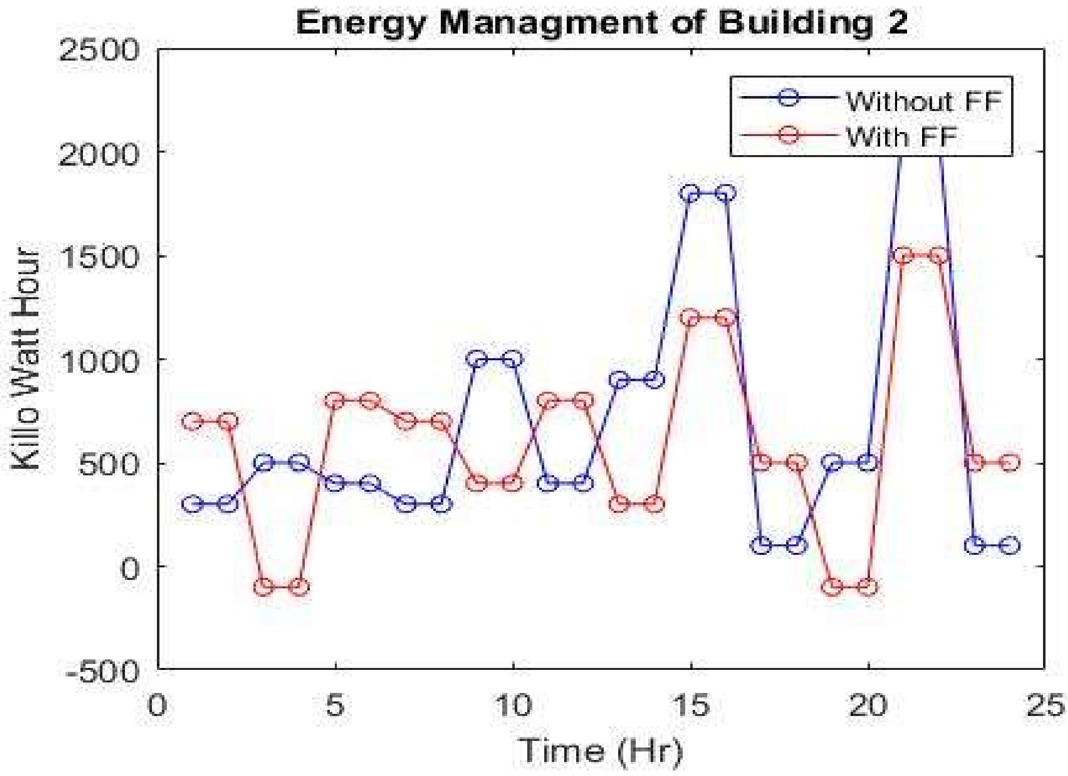

Figure 5.5 *Results of building 2 load.*

Figure 5.5 illustrates the building 2 load management with and without FF. The y-axis shows the power consumption in kilowatt-hour, and the x-axis shows the time slots. Here it can be seen that the peak of the load was at 2100kw, but the proposed system shifted it to 1600kw. So, in building 2, an overall 24% reduction in peak load has been observed.

Load Management Building 3

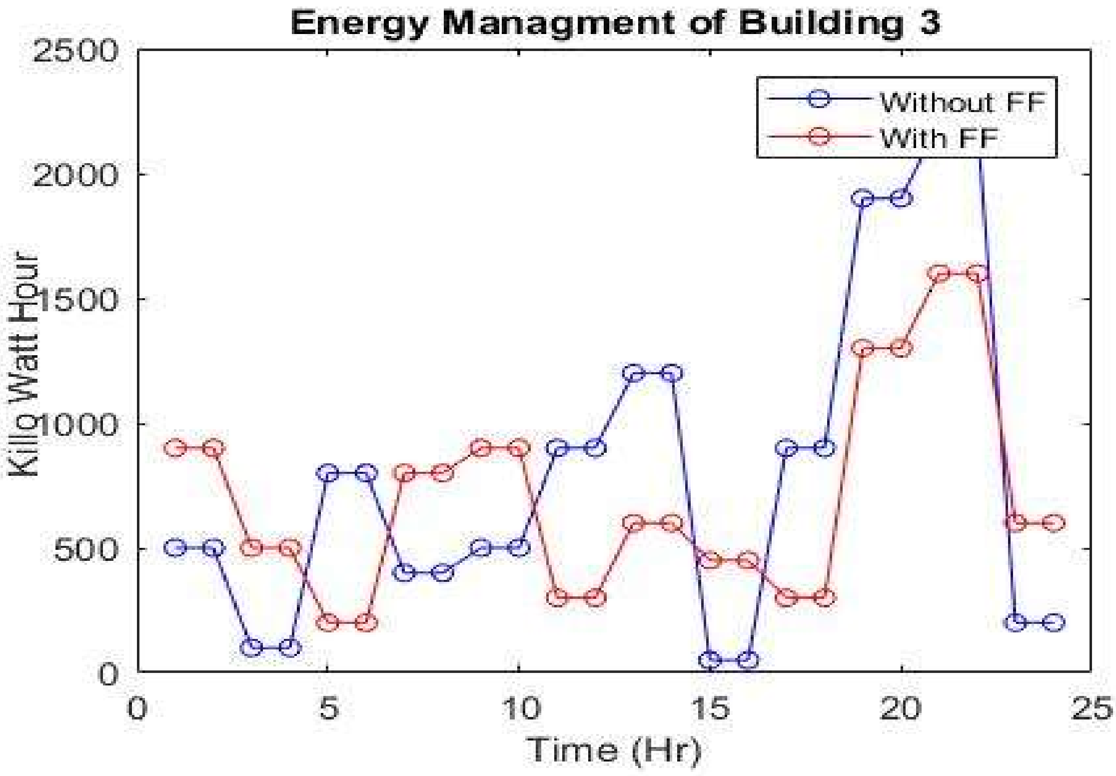

Figure 5.6 *Results of building 3 load.*

Figure 5.6 shows the building 3 load management with and without FF. The y-axis shows the power consumption in kilowatt-hour, and the x-axis shows the time slots. Here it can be seen that the maximum peak in building 3 is occurring at 2300kw. Therefore, the proposed system has rescheduled load and shifts it to 1600kw. It is a 30% reduction in peak demand.

Total Load Management

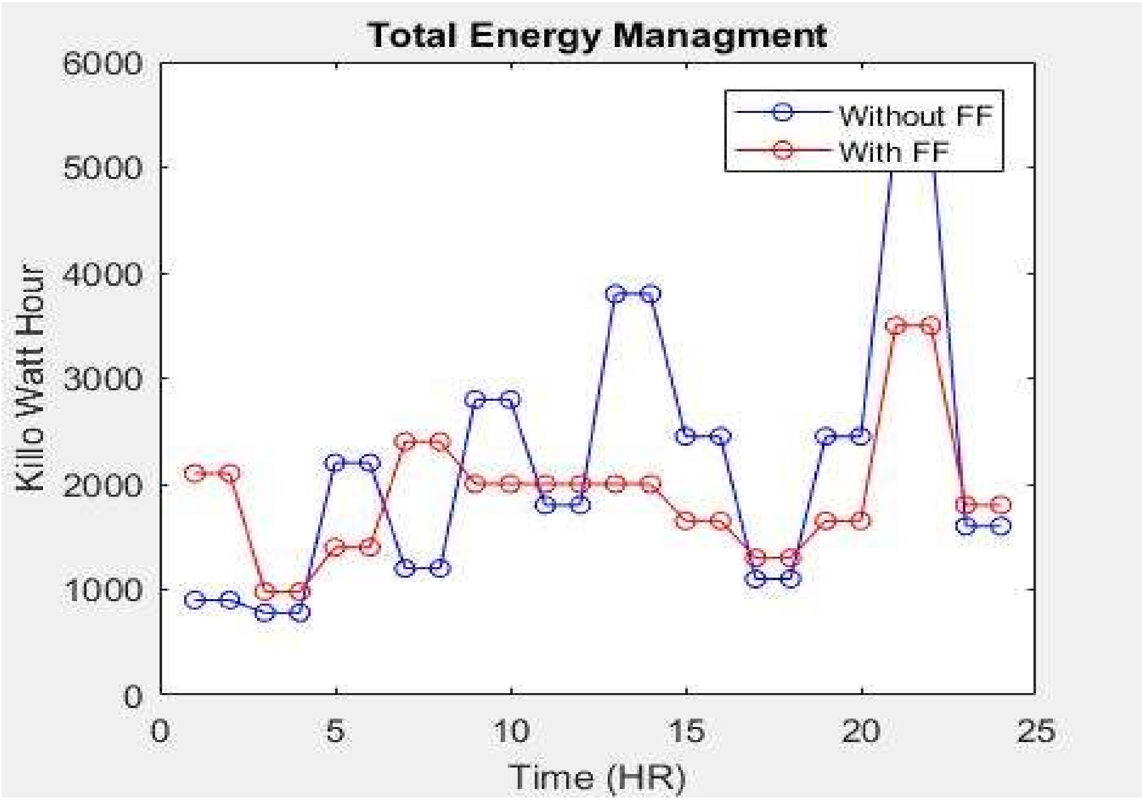

Figure 5.7 *Buildings total DSM.*

Figure 5.7 exhibits the main output and objective of the proposed system. The y-axis shows the power consumption of all buildings in kilowatt-hour, and the x-axis shows the time slots. Here, the unscheduled load is blue colored, while the second graph is an FA-based managed graph. As discussed earlier, the main objectives of the proposed system are to reduce electricity cost and peak reduction. The figure clearly explains that unscheduled load has various peaks. So, it has a higher cost per kilowatt-hour. However, the FA has managed this unscheduled load in such a way that its peaks are reduced.

Although some of the load peaks are greater than the unscheduled load, the overall effect of the proposed system has reduced the electricity cost. An unscheduled load requires a 5400 KW power source, while DSM-based load requires only a 3500 KW source. It means a total 35% peak reduction has been done. Also, cost reduction & electrical stress on the national grid have been reduced.

5.3.3 Per day Electricity Bill

Per day Electricity Bill Building 1

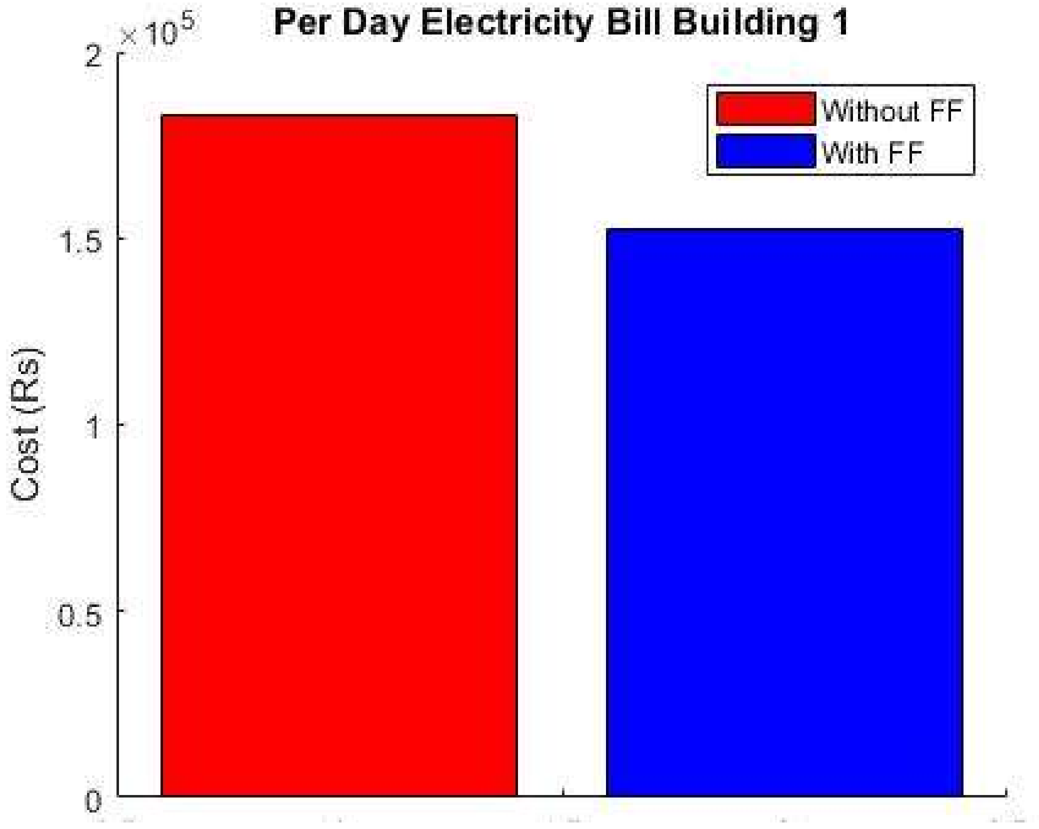

Figure 5.8 *Building 1 per day bill.*

Figure 5.8 presents the power cost comparison of building 1. The blue color represents a firefly-based scheduled load in this graph, while the red-colored bar represents an unscheduled load. It is visible that the FA has reduced the cost of electricity per day from 190000Rs to 150000Rs. It is a 21% reduction in the bill for building 1.

Per day Electricity Bill Building 2

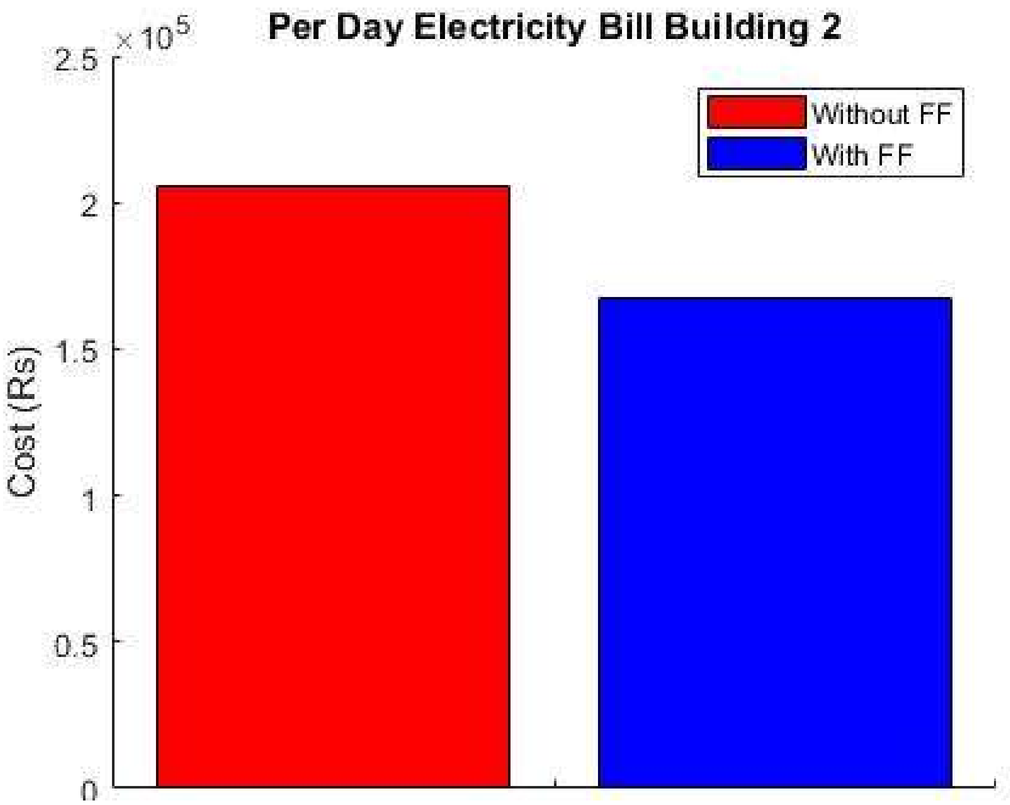

Figure 5.9 *Building 2 per day bill.*

Figure 5.9 shows the power cost comparison of building 2. The y-axis shows the Cost, and the x-axis shows the time slots. It is visible that the FA has reduced the cost of electricity per day from 210000Rs to 170000Rs. It is a 19% reduction in the per-day bill.

Per day Electricity Bill Building 2

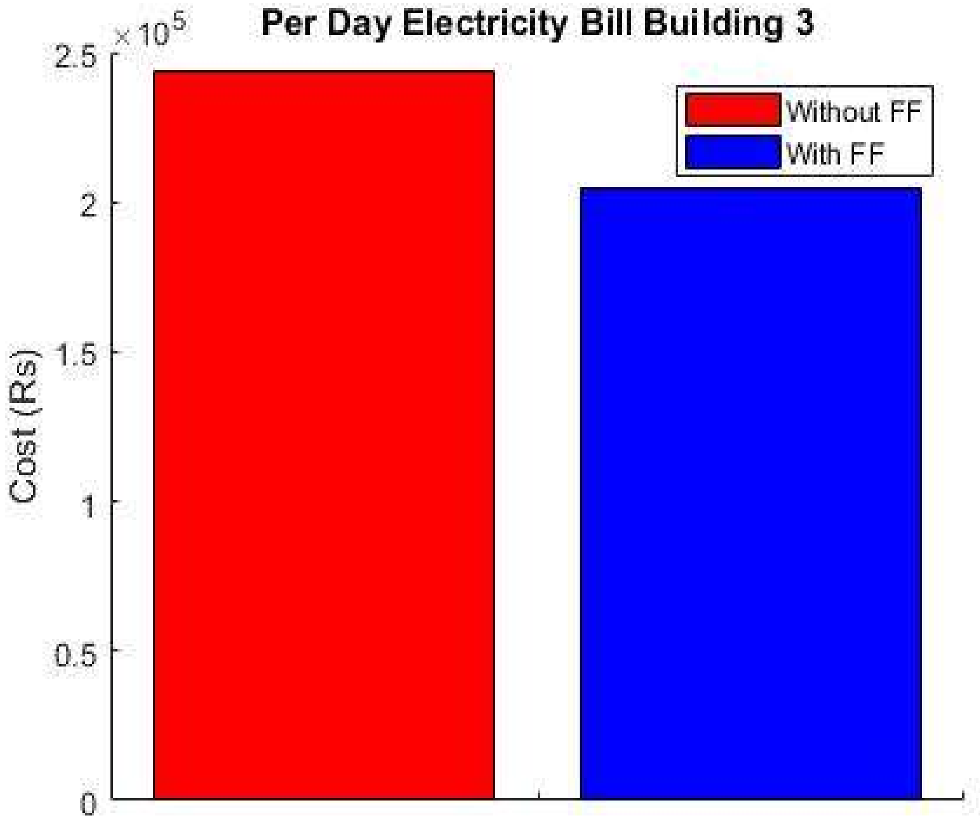

Figure 5.10 *Building 3 per day bill.*

Figure 5.10 illustrates the power cost comparison of building 3. The y-axis shows the Cost, and the x-axis shows the time slots. It is visible that the FA has reduced the cost of electricity per day from 250000Rs to 200000Rs. It is a 20% reduction in the bill for building 3.

Per day Electricity Total Bill

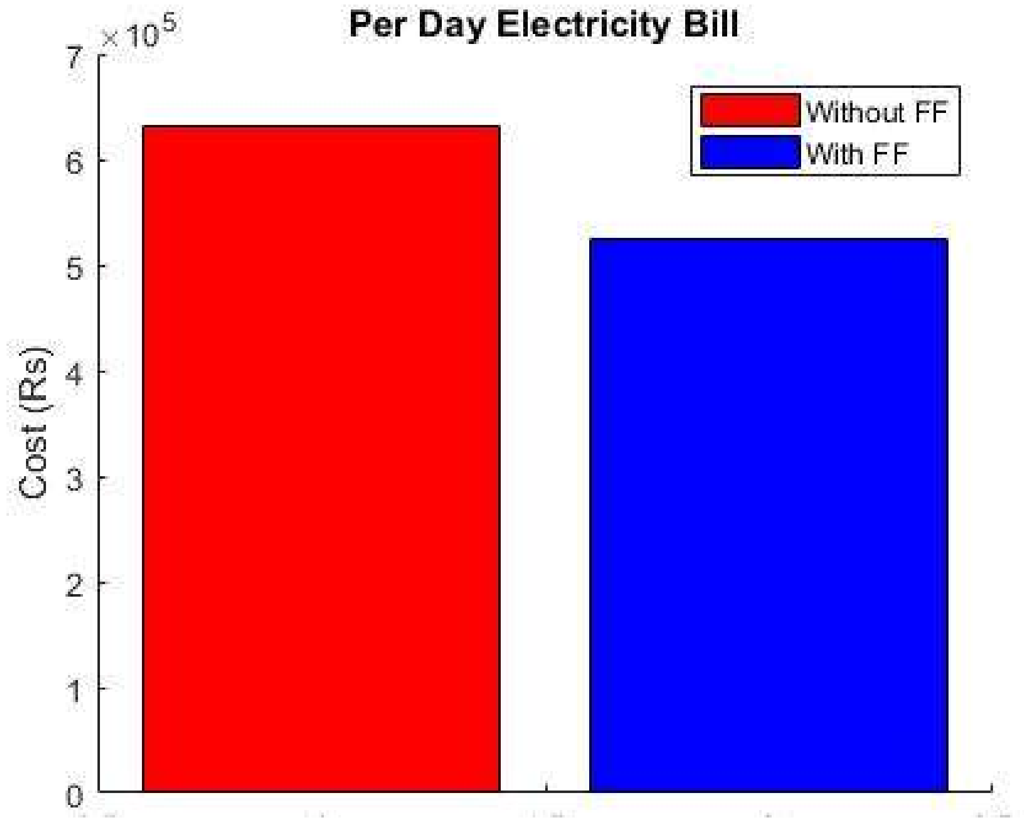

Figure 5.11 *Total building bill per day.*

Figure 5.11 presents the cost of electricity off three buildings in a single day. The y-axis shows the Cost, and the x-axis shows the time slots. The application of the proposed system for the DSM has reduced the electricity cost from 650000Rs to 520000Rs approximately. This is approximately 20% cost reduction and power saving.

5.4 Summary

The Peaks of power affect the distribution side and the generation side, which usually occurs during the summertime because of high electricity demands. For reducing the excessive load on generating plant, load shedding occurs. If load shedding is not occurred, the generation system can collapse, which happens when the demand gets higher than the supply. Therefore, to tackle this issue, the proposed model has been constructed based on the FA. This algorithm helps in minimizing peak load and costs.

CHAPTER 6

CONCLUSIONS AND FUTURE RECOMMENDATIONS

6.1 Conclusions and Recommendations

In the first part of the thesis, LF is vital for better planning and operating power generation systems in Smart Grid applications. This thesis proposes a new combinational forecasting model based on the following: time series and auto-regression algorithm for data fetching, ML-based feedback ANFIS, BGA-PCA algorithm for best feature selection, and proposed integration strategy for MAPE calculation. We presented a new time series and auto-regression algorithm for analyzing the more complex and massive historical data. The time series AR model is further validated using the ML-based feedback ANFIS model, where five stages are computed. BGA-PCA approach for attaining the best feature selection is evaluated and attained more optimized results. A remarkable improvement in MAPE and average MAPE have been achieved by a proposed integration strategy which gives a 17% annual improvement better than previous ones. Mainly, the attained MAPE: 1.70, 1.77, 1.80, and 1.67 percentage for Summer, Fall, Winter, and Spring, respectively. This model can optimize the performance of power grids by predicting the optimized LF and can overcome the problems related to the planning and operation of the smart grid.

PF is a crucial aspect for better planning and operating power generation systems as a Smart Grid application in the second part of the thesis. A new combinational forecasting model based on time series and auto-regression algorithm, ML-based feedback ANFIS, BGA-PCA algorithm is proposed for best feature selection. In this thesis, a novel time-series and auto-regression algorithm for analyzing the more complex and massive historical data is presented. The time series AR model is further validated by the ML-based feedback ANFIS model, where six stages are computed. BGA-PCA approach for attaining the best feature selection is evaluated and attained more optimized results. A remarkable improvement in MAPE of less than 7% in overall performance and an average MAPE of 5.45 has been achieved by a proposed integration that is significantly improved compared to previous ones. This model can optimize the performance of power grids by predicting the optimized PF and can overcome the problems related to the planning and operating of the smart grid.

In the third part, Peaks of power badly affect the power distribution network and the generation side. Usually, load shedding is done to reduce the peak loads in the summer. However, if load shedding does not occur, the system can collapse because power generation is not infinite. For example, if the entire power requirement of a house is 5 kilowatts but at 8:00 pm, the entire family watches television and air conditioning system are running at maximum power & due to this factor the Power requirement increase is to 6-kilo watts. Now even this increase in power lasts only for 5 minutes or 10 minutes, but WAPDA has to specify 6 kilowatts for the house or else the system will collapse. But if we reschedule the loads so that peak load does not occur and is distributed among average, then generation capacity can also be minimized, and the cost is reduced.

The case mentioned above is the main summary of the proposed system. The proposed system has been constructed on the FA to minimize peak load and costs. The chapter has presented objective functions and equations to design a base for the FA.

In the future, this proposed model can be implemented for LF with various time horizon locations to validate its effectiveness.

- By taking a new variety of data input for economic studies, the model can be implemented for PF in financial applications.
- This model can be checked by our model through various data sizes of statistically significant data.
- Price and Load forecasting are considered the main feature of the revolutionary smart grid system, which could help to enhance electricity generation in a decentralized grid network.
- To provide reliable and uninterruptable electricity supply during peak hours, it could facilitate the generation system by providing essential information on load and price forecasting.

- LF and PF based EMS can be minimized peak load demand by load and price forecasting analysis through Machine Learning approach. However, the current study can be extended further into more smart grid applications.

REFERENCES

[1] C. W. Gellings, "Then and now: The perspective of the man who coined the term 'DSM,'" *Energy Policy*, vol. 24, no. 4, pp. 285–288, Apr. 1996, doi: 10.1016/0301-4215(95)00134-4.

[2] F. H. Malik and M. Lehtonen, "A review: Agents in smart grids," *Electr. Power Syst. Res.*, vol. 131, pp. 71–79, Feb. 2016, doi: 10.1016/J.EPSR.2015.10.004.

[3] M. A. B. Siddique, M. A. Khan, A. Asad, A. U. Rehman, R. M. Asif, and S. U. Rehman, "Maximum power point tracking with modified incremental conductance technique in grid-connected PV array," Nov. 2020, doi: 10.1109/CITISIA50690.2020.9371803.

[4] S. N. Fallah, M. Ganjkhani, S. Shamshirband, and K. wing Chau, "Computational intelligence on short-term load forecasting: A methodological overview," *Energies*, vol. 12, no. 3. MDPI AG, p. 393, Jan. 27, 2019, doi: 10.3390/en12030393.

[5] R. M. Asif *et al.*, "Design and analysis of robust fuzzy logic maximum power point tracking based isolated photovoltaic energy system," 2020, doi: 10.1002/eng2.12234.

[6] R. Bayindir, I. Colak, G. Fulli, and K. Demirtas, "Smart grid technologies and applications," *Renewable and Sustainable Energy Reviews*, vol. 66. Elsevier Ltd, pp. 499–516, Dec. 01, 2016, doi: 10.1016/j.rser.2016.08.002.

[7] A. Shewale, A. Mokhade, N. Funde, and N. Bokde, "An overview of demand response in smart grid and optimization techniques for efficient residential appliance scheduling problem," *Energies*, vol. 13, p. 4266, 2020, doi: 10.3390/en13164266.

[8] Z. Bojkovic and B. Bakmaz, "Smart grid communications architecture: a survey and challenges," 2012, pp. 83–89.

[9] R. Dharani, M. Balasubramonian, T. S. Babu, and B. Nastasi, "Load shifting and peak clipping for reducing energy consumption in an indian university campus," *Energies*, vol. 14, no. 3, p. 558, Feb. 2021, doi: 10.3390/en14030558.

[10] M. N. Q. Macedo, J. J. M. Galo, L. A. L. de Almeida, and A. C. de C. Lima,

"Demand side management using artificial neural networks in a smart grid environment," *Renew. Sustain. Energy Rev.*, vol. 41, pp. 128–133, 2015, doi: https://doi.org/10.1016/j.rser.2014.08.035.

[11] A. Ahmad *et al.*, "An optimized home energy management system with integrated renewable energy and storage resources," *Energies*, vol. 10, no. 4, 2017, doi: 10.3390/en10040549.

[12] T. Logenthiran, D. Srinivasan, and T. Z. Shun, "Demand side management in smart grid using heuristic optimization," *IEEE Trans. Smart Grid*, vol. 3, no. 3, pp. 1244–1252, 2012, doi: 10.1109/TSG.2012.2195686.

[13] M. A. Khan, N. Javaid, A. Mahmood, Z. A. Khan, and N. Alrajeh, "A generic demand-side management model for smart grid," *Int. J. Energy Res.*, vol. 39, no. 7, pp. 954–964, 2015, doi: https://doi.org/10.1002/er.3304.

[14] F. Khan, M. A. B. Siddiqui, A. U. Rehman, J. Khan, M. T. S. A. Asad, and A. Asad, "IoT based power monitoring system for smart grid applications," Feb. 2020, doi: 10.1109/ICEET48479.2020.9048229.

[15] H. B. Jond, R. Benveniste, and V. V. Nabiyev, "Peak load appliance scheduling for demand-side management in the future smart grids," Dec. 2015, doi: 10.1109/SGCF.2015.7354925.

[16] S. Moon and J. W. Lee, "Multi-residential demand response scheduling with multi-class appliances in smart grid," *IEEE Trans. Smart Grid*, vol. 9, no. 4, pp. 2518–2528, Jul. 2018, doi: 10.1109/TSG.2016.2614546.

[17] A. Safdarian, M. Fotuhi-Firuzabad, and M. Lehtonen, "Optimal residential load management in smart grids: a decentralized framework," *IEEE Trans. Smart Grid*, vol. 7, no. 4, pp. 1836–1845, Jul. 2016, doi: 10.1109/TSG.2015.2459753.

[18] S. Salinas, M. Li, and P. Li, "Multi-objective optimal energy consumption scheduling in smart grids," vol. 4, no. 1, pp. 341–348, 2013.

[19] and T. P. I. A. Remani T , E. A. Jasmin, "Residential load scheduling with renewable generation in the smart grid: A reinforcement learning approach," *IEEE Journals Mag.*, vol. 13, no. 3, pp. 3283–3294, 2018, Accessed: Jun. 06, 2021. [Online]. Available: https://ieeexplore.ieee.org/document/8426048.

[20] Y. H. Lin and M. S. Tsai, “An advanced home energy management system facilitated by nonintrusive load monitoring with automated multiobjective power scheduling,” *IEEE Trans. Smart Grid*, vol. 6, no. 4, pp. 1839–1851, Jul. 2015, doi: 10.1109/TSG.2015.2388492.

[21] D. Liu, J. Xiao, J. Liu, X. Yuan, and S. Zhang, “Dynamic energy trading and load scheduling algorithm for the end-user in smart grid,” *IEEE Access*, vol. 8, pp. 189632–189645, 2020, doi: 10.1109/ACCESS.2020.3031325.

[22] R. C. Bansal and J. C. Pandey, “Load forecasting using artificial intelligence techniques: A literature survey,” *International Journal of Computer Applications in Technology*, vol. 22, no. 2–3. Inderscience Publishers, pp. 109–119, 2005, doi: 10.1504/IJCAT.2005.006942.

[23] A. J. Del Real, F. Dorado, and J. Durán, “Energy demand forecasting using deep learning: Applications for the French grid,” *Energies*, vol. 13, no. 9, p. 2242, May 2020, doi: 10.3390/en13092242.

[24] B. Masood *et al.*, “Investigation of deterministic, statistical and parametric NB-PLC channel modeling techniques for advanced metering infrastructure,” *Energies*, vol. 13, no. 12, Jun. 2020, doi: 10.3390/en13123098.

[25] D. Niu, Y. Wang, and D. D. Wu, “Power load forecasting using support vector machine and ant colony optimization,” *Expert Syst. Appl.*, vol. 37, no. 3, pp. 2531–2539, Mar. 2010, doi: 10.1016/j.eswa.2009.08.019.

[26] L. M. Saini, “Peak load forecasting using bayesian regularization, resilient and adaptive backpropagation learning based artificial neural networks,” *Electr. Power Syst. Res.*, vol. 78, no. 7, pp. 1302–1310, Jul. 2008, doi: 10.1016/j.epsr.2007.11.003.

[27] B. Yoon, C. Yoon, and Y. Park, “On the development and application of a self–organizing feature map–based patent map,” *R&D Manag.*, vol. 32, no. 4, pp. 291–300, Sep. 2002, doi: 10.1111/1467-9310.00261.

[28] G. T. Costanzo, G. Zhu, M. F. Anjos, and G. Savard, “A system architecture for autonomous demand side load management in smart buildings,” *IEEE Trans. Smart Grid*, vol. 3, no. 4, pp. 2157–2165, 2012, doi: 10.1109/TSG.2012.2217358.

[29] K. Long and Z. Yang, “Model predictive control for household energy management

based on individual habit," in *2013 25th Chinese Control and Decision Conference (CCDC)*, 2013, pp. 3676–3681, doi: 10.1109/CCDC.2013.6561587.

[30] V. Stavrakas and A. Flamos, "A modular high-resolution demand-side management model to quantify benefits of demand-flexibility in the residential sector," *Energy Convers. Manag.*, vol. 205, p. 112339, Feb. 2020, doi: 10.1016/j.enconman.2019.112339.

[31] J. H. Yoon, R. Baldick, and A. Novoselac, "Dynamic demand response controller based on real-time retail price for residential buildings," *IEEE Trans. Smart Grid*, vol. 5, no. 1, pp. 121–129, 2014, doi: 10.1109/TSG.2013.2264970.

[32] B. Gao, W. Zhang, Y. Tang, M. Hu, M. Zhu, and H. Zhan, "Game-theoretic energy management for residential users with dischargeable plug-in electric vehicles," *Energies*, vol. 7, no. 11, pp. 7499–7518, 2014, doi: 10.3390/en7117499.

[33] G. Yan, T. Han, W. Zhang, and S. Zhao, "Short-term load forecasting of smart grid based on load spatial-temporal distribution," in *2019 IEEE Innovative Smart Grid Technologies - Asia (ISGT Asia)*, 2019, pp. 781–785, doi: 10.1109/ISGT-Asia.2019.8881259.

[34] A. Cataliotti, V. Cosentino, D. Di Cara, P. Russotto, E. Telaretti, and G. Tinè, "An innovative measurement approach for load flow analysis in mv smart grids," *IEEE Trans. Smart Grid*, vol. 7, no. 2, pp. 889–896, 2016, doi: 10.1109/TSG.2015.2430891.

[35] H. Yang, J. Zhang, J. Qiu, S. Zhang, M. Lai, and Z. Y. Dong, "A practical pricing approach to smart grid demand response based on load classification," *IEEE Trans. Smart Grid*, vol. 9, no. 1, pp. 179–190, 2018, doi: 10.1109/TSG.2016.2547883.

[36] K. Ma, T. Yao, J. Yang, and X. Guan, "Residential power scheduling for demand response in smart grid," *Int. J. Electr. Power Energy Syst.*, vol. 78, pp. 320–325, 2016, doi: https://doi.org/10.1016/j.ijepes.2015.11.099.

[37] Q. Xu, Y. Ding, Q. Yan, A. Zheng, and P. Du, "Day-ahead load peak shedding/shifting scheme based on potential load values utilization: Theory and Practice of Policy-Driven Demand Response in China," *IEEE Access*, vol. 5, pp. 22892–22901, 2017, doi: 10.1109/ACCESS.2017.2763678.

[38] Y. K. Bichpuriya, S. A. Soman, and A. Subramanyam, "Robust probability density forecasts of yearly peak load using non-parametric model," in *2016 IEEE Power and Energy Society General Meeting (PESGM)*, 2016, pp. 1–5, doi: 10.1109/PESGM.2016.7741892.

[39] H. Jahangir *et al.*, "A novel electricity price forecasting approach based on dimension reduction strategy and rough artificial neural networks," *IEEE Trans. Ind. Informatics*, vol. 16, no. 4, pp. 2369–2381, 2020, doi: 10.1109/TII.2019.2933009.

[40] M. Zawad Ali, N. Sakib Khan Shabbir, M. Sifatul Alam Chowdhury, A. Ghosh, and X. Liang, "Regression models of critical parameters affecting peak load demand forecasting," in *2018 IEEE Canadian Conference on Electrical Computer Engineering (CCECE)*, 2018, pp. 1–4, doi: 10.1109/CCECE.2018.8447786.

[41] A. Mahmood, F. Baig, N. Alrajeh, U. Qasim, Z. A. Khan, and N. Javaid, "An enhanced system architecture for optimized demand side management in smart grid," *Appl. Sci.*, vol. 6, no. 5, 2016, doi: 10.3390/app6050122.

[42] A. Garulli, S. Paoletti, and A. Vicino, "Models and techniques for electric load forecasting in the presence of demand response," *IEEE Trans. Control Syst. Technol.*, vol. 23, no. 3, pp. 1087–1097, 2015, doi: 10.1109/TCST.2014.2361807.

[43] J. Xie, T. Hong, and J. Stroud, "Long-term retail energy forecasting with consideration of residential customer attrition," *IEEE Trans. Smart Grid*, vol. 6, no. 5, pp. 2245–2252, 2015, doi: 10.1109/TSG.2014.2388078.

[44] B. Zhao *et al.*, "Energy management of multiple microgrids based on a system of systems architecture," *IEEE Trans. Power Syst.*, vol. 33, no. 6, pp. 6410–6421, 2018, doi: 10.1109/TPWRS.2018.2840055.

[45] H. Jain, B. Bhatti, T. Wu, B. Mather, and R. Broadwater, "Integrated transmission-and-distribution system modeling of power systems: state-of-the-art and future research diretions," *Energies*, vol. 14, p. 12, 2020, doi: 10.3390/en14010012.

[46] D. Zhang, J. Li, and D. Hui, "Coordinated control for voltage regulation of distribution network voltage regulation by distributed energy storage systems," *Prot. Control Mod. Power Syst.*, vol. 3, no. 1, pp. 1–8, Dec. 2018, doi:

10.1186/s41601-018-0077-1.

[47] E. O. Kontis, T. A. Papadopoulos, A. I. Chrysochos, and G. K. Papagiannis, "Measurement-based dynamic load modeling using the vector fitting technique," *IEEE Trans. Power Syst.*, vol. 33, no. 1, pp. 338–351, 2018, doi: 10.1109/TPWRS.2017.2697004.

[48] N. Huang, W. Wang, S. Wang, J. Wang, G. Cai, and L. Zhang, "Incorporating load fluctuation in feature importance profile clustering for day-ahead aggregated residential load forecasting," *IEEE Access*, vol. 8, pp. 25198–25209, 2020, doi: 10.1109/ACCESS.2020.2971033.

[49] W. Gang, W. Ke, L. Jun, J. Wenlu, and Z. Hong, "The comprehensive optimal scheduling model of power supply and demand considering network constraints," in *2016 IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC)*, 2016, pp. 2050–2054, doi: 10.1109/APPEEC.2016.7779846.

[50] G. Zhang and J. Guo, "A novel method for hourly electricity demand forecasting," *IEEE Trans. Power Syst.*, vol. 35, no. 2, pp. 1351–1363, 2020, doi: 10.1109/TPWRS.2019.2941277.

[51] A. Filali, Z. Mlika, S. Cherkaoui, and A. Kobbane, "Preemptive SDN load balancing with machine learning for delay sensitive applications," *IEEE Trans. Veh. Technol.*, vol. 69, no. 12, pp. 15947–15963, 2020, doi: 10.1109/TVT.2020.3038918.

[52] B. Goehry, Y. Goude, P. Massart, and J.-M. Poggi, "Aggregation of multi-scale experts for bottom-up load forecasting," *IEEE Trans. Smart Grid*, vol. 11, no. 3, pp. 1895–1904, 2020, doi: 10.1109/TSG.2019.2945088.

[53] H. Dong, Y. Gao, X. Meng, and Y. Fang, "A multifactorial short-term load forecasting model combined with periodic and non-periodic features - a case study of Qingdao, China," *IEEE Access*, vol. 8, pp. 67416–67425, 2020, doi: 10.1109/ACCESS.2020.2986031.

[54] B. A. Høverstad, A. Tidemann, H. Langseth, and P. Öztürk, "Short-term load forecasting with seasonal decomposition using evolution for parameter tuning," *IEEE Trans. Smart Grid*, vol. 6, no. 4, pp. 1904–1913, 2015, doi:

10.1109/TSG.2015.2395822.

[55] E. Gonzalez-Romera, M. A. Jaramillo-Moran, and D. Carmona-Fernandez, "Monthly electric energy demand forecasting based on trend extraction," *IEEE Trans. Power Syst.*, vol. 21, no. 4, pp. 1946–1953, 2006, doi: 10.1109/TPWRS.2006.883666.

[56] D. Lu, Y. Liu, Q. Liao, B. Wang, W. Huang, and X. Xi, "Time-domain transmission line fault location method with full consideration of distributed parameters and line asymmetry," *IEEE Trans. Power Deliv.*, vol. 35, no. 6, pp. 2651–2662, 2020, doi: 10.1109/TPWRD.2020.2974294.

[57] A. Agarwal, A. Ojha, S. C. Tewari, and M. M. Tripathi, "Hourly load and price forecasting using ANN and fourier analysis," in *2014 6th IEEE Power India International Conference (PIICON)*, 2014, pp. 1–6, doi: 10.1109/POWERI.2014.7117736.

[58] S. Schnürch and A. Wagner, "Electricity price forecasting with neural networks on EPEX order books," *Appl. Math. Financ.*, vol. 27, no. 3, pp. 189–206, May 2020, doi: 10.1080/1350486X.2020.1805337.

[59] M. Alam, "Day-ahead electricity price forecasting and scheduling of energy storage in LMP market," *IEEE Access*, vol. 7, pp. 165627–165634, 2019, doi: 10.1109/ACCESS.2019.2952451.

[60] M. Alamaniotis, D. Bargiotas, N. G. Bourbakis, and L. H. Tsoukalas, "Genetic optimal regression of relevance vector machines for electricity pricing signal forecasting in smart grids," *IEEE Trans. Smart Grid*, vol. 6, no. 6, pp. 2997–3005, 2015, doi: 10.1109/TSG.2015.2421900.

[61] K. Aurangzeb, S. Aslam, M. S. Muhammad, and M. Alhussein, "A fair pricing mechanism in smart grids for low energy consumption users," *IEEE Access*, vol. PP, p. 1, 2021, doi: 10.1109/ACCESS.2021.3056035.

[62] J. Widén, M. Lundh, I. Vassileva, E. Dahlquist, K. Ellegård, and E. Wäckelgård, "Constructing load profiles for household electricity and hot water from time-use data—Modelling approach and validation," *Energy Build.*, vol. 41, no. 7, pp. 753–768, 2009, doi: https://doi.org/10.1016/j.enbuild.2009.02.013.

[63] R. Bartels and D. G. Fiebig, “Metering and modelling residential enduse electricity load curves,” *J. Forecast.*, vol. 15, no. 6, pp. 415–426, 1996, doi: 10.1002/(SICI)1099-131X(199611)15:6<415::AID-FOR636<3.0.CO;2-J.

[64] A. U. Rehman, R. A. Naqvi, A. Rehman, A. Paul, M. T. Sadiq, and D. Hussain, “A trustworthy siot aware mechanism as an enabler for citizen services in smart cities,” *Electronics*, vol. 9, no. 6, p. 918, Jun. 2020, doi: 10.3390/electronics9060918.

[65] A. S. Alahmed and M. M. Almuhaini, “Hybrid top-down and bottom-up approach for investigating residential load compositions and load percentages,” *arXiv*, 2020.

[66] N. Narayan, Z. Qin, J. Popovic-Gerber, J. C. Diehl, P. Bauer, and M. Zeman, “Stochastic load profile construction for the multi-tier framework for household electricity access using off-grid DC appliances,” *Energy Effic.*, vol. 13, no. 2, pp. 197–215, 2020, doi: 10.1007/s12053-018-9725-6.

[67] E. (2021). Proedrou, “A comprehensive review of residential electricity load profile models. I,” *EEE Access.*, vol. 9, no. 1, pp. 12114–12133, 2021, doi: 10.1109/ACCESS.2021.3050074.

[68] B. Masood, S. Guobing, R. A. Naqvi, M. B. Rasheed, J. Hou, and A. U. Rehman, “Measurements and channel modeling of low and medium voltage NB-PLC networks for smart metering,” *IET Gener. Transm. Distrib.*, vol. 15, no. 2, pp. 321–338, 2021, doi: https://doi.org/10.1049/gtd2.12023.

[69] R.-F. Yuan, Q. Ai, and X. He, “Research on dynamic load modelling based on power quality monitoring system,” *IET Gener. Transm. \& Distrib.*, vol. 7, no. 1, pp. 46–51, 2013.

[70] Z. A. Khan, D. Jayaweera, and M. S. Alvarez-Alvarado, “A novel approach for load profiling in smart power grids using smart meter data,” *Electr. Power Syst. Res.*, vol. 165, pp. 191–198, 2018.

[71] M. S. Hossain, N. A. Madlool, N. A. Rahim, J. Selvaraj, A. K. Pandey, and A. F. Khan, “Role of smart grid in renewable energy: An overview,” *Renew. Sustain. Energy Rev.*, vol. 60, pp. 1168–1184, 2016, doi: https://doi.org/10.1016/j.rser.2015.09.098.

[72] Z. Cao, J. Lin, C. Wan, Y. Song, Y. Zhang, and X. Wang, “Optimal cloud

computing resource allocation for demand side management in smart grid," *IEEE Trans. Smart Grid*, vol. 8, no. 4, pp. 1943–1955, 2016.

[73] M. A. López, S. De La Torre, S. Mart\'\in, and J. A. Aguado, "Demand-side management in smart grid operation considering electric vehicles load shifting and vehicle-to-grid support," *Int. J. Electr. Power \& Energy Syst.*, vol. 64, pp. 689–698, 2015.

[74] P. Brandstetter, P. Chlebis, and P. Simonik, "Active power filter with soft switching," *Int. Rev. Electr. Eng.*, vol. 5, no. 6, pp. 2516–2526, 2010.

[75] N. Ammar, M. Sulaiman, and A. F. M. Nor, "Analysis load forecasting of power system using of fuzzy logic and artificial neural network," *J. Telecommun. Electron. Comput. Eng.*, vol. 9, no. 3, pp. 181–192, 2017.

[76] K. F. Leung, F. H. F. Leung, H. K. Lam, and S. H. Ling, "Application of a modified neural fuzzy network and an improved genetic algorithm to speech recognition," *NEURAL Comput. Appl.*, vol. 16, no. 4–5, pp. 419–431, May 2007, doi: 10.1007/s00521-006-0068-4.

[77] A. Moradzadeh, S. Zakeri, M. Shoaran, and B. Mohammadi-ivatloo, "Short-term load forecasting of microgrid via hybrid support vector regression and long short-term memory algorithms," 2020.

[78] G. Dudek, "Multilayer perceptron for short-term load forecasting: from global to local approach," *Neural Comput. Appl.*, vol. 32, no. 8, pp. 3695–3707, 2020, doi: 10.1007/s00521-019-04130-y.

[79] N. Shabbir, R. Ahmadiahangar, L. Kütt, and A. Rosin, "Comparison of machine learning based methods for residential load forecasting," in *2019 Electric Power Quality and Supply Reliability Conference (PQ) \& 2019 Symposium on Electrical Engineering and Mechatronics (SEEM)*, 2019, pp. 1–4.

[80] I. Khandelwal, R. Adhikari, and G. Verma, "Time series forecasting using hybrid arima and ann models based on DWT Decomposition," *Procedia Comput. Sci.*, vol. 48, no. C, pp. 173–179, 2015, doi: 10.1016/j.procs.2015.04.167.

[81] J. Zhang, Y.-M. Wei, D. Li, Z. Tan, and J. Zhou, "Short term electricity load forecasting using a hybrid model," *Energy*, vol. 158, pp. 774–781, 2018.

[82] L. C. P. Velasco, D. L. L. Polestico, G. P. O. Macasieb, M. B. V Reyes, and F. B. Vasquez Jr, "Load forecasting using autoregressive integrated moving average and artificial neural network," *Int. J. Adv. Comput. Sci. Appl*, vol. 9, pp. 23–29, 2018.

[83] K. Jaskie, D. Smith, and A. Spanias, "Deep learning networks for vectorized energy load forecasting," in *2020 11th International Conference on Information, Intelligence, Systems and Applications (IISA*, 2020, pp. 1–7.

[84] T. Liu, Y. Jin, and Y. Gao, "A new hybrid approach for short-term electric load forecasting applying support vector machine with ensemble empirical mode decomposition and whale optimization," *Energies*, vol. 12, no. 8, p. 1520, 2019.

[85] and I. U. Muhammad Abu Bakar Siddique, Adeel Asad, Rao M. Asif, Ateeq Ur Rehman, Muhammad Tariq Sadiq, "Implementation of Incremental Conductance MPPT Algorithm with Integral Regulator by using Boost Converter in Grid Connected PV Array," *IETE J. Res.*, vol. 67, no. 1, 2021, doi: 10.1080/03772063.2021/1920481.

[86] M.-W. Chang and B. Chen, "EUNITE Network Competition: Electricity Load Forecasting," 2001.

[87] H. Wang, G. Li, G. Wang, J. Peng, H. Jiang, and Y. Liu, "Deep learning based ensemble approach for probabilistic wind power forecasting," *Appl. Energy*, vol. 188, pp. 56–70, 2017, doi: https://doi.org/10.1016/j.apenergy.2016.11.111.

[88] U. Ugurlu, I. Oksuz, and O. Tas, "Electricity Price Forecasting Using Recurrent Neural Networks," *Energies*, vol. 11, no. 5, 2018, doi: 10.3390/en11051255.

[89] X. Zhang, J. Wang, and K. Zhang, "Short-term electric load forecasting based on singular spectrum analysis and support vector machine optimized by Cuckoo search algorithm," *Electr. Power Syst. Res.*, vol. 146, pp. 270–285, 2017, doi: https://doi.org/10.1016/j.epsr.2017.01.035.

[90] O. Abedinia, N. Amjady, and H. Zareipour, "A new feature selection technique for load and price forecast of electrical power systems," *IEEE Trans. Power Syst.*, vol. 32, no. 1, pp. 62–74, 2017, doi: 10.1109/TPWRS.2016.2556620.

[91] M. Sheikhan and N. Mohammadi, "Neural-based electricity load forecasting using hybrid of GA and ACO for feature selection," *Neural Comput. Appl.*, vol. 21, no.

8, pp. 1961–1970, 2012, doi: 10.1007/s00521-011-0599-1.

[92] M. D. Olagoke, A. A. Ayeni, and M. A. Hambali, "Short term electric load forecasting using neural network and genetic algorithm," *Int. J. Appl. Inf. Syst.*, vol. 10, no. 4, pp. 22–28, 2016.

[93] R. Urraca, A. Sanz-Garcia, J. Fernandez-Ceniceros, A. Pernia-Espinoza, and F. J. Martinez-De-Pison, "Improving hotel room demand forecasting with a hybrid GA-SVR methodology based on skewed data transformation, feature selection and parsimony tuning," *Log. J. IGPL*, vol. 25, no. 6, pp. 877–889, 2017, doi: 10.1093/jigpal/jzx029.

[94] Y. Liu, Y. Yin, J. Gao, and C. Tan, "Wrapper feature selection optimized svm model for demand forecasting," in *2008 The 9th International Conference for Young Computer Scientists*, 2008, pp. 953–958, doi: 10.1109/ICYCS.2008.151.

[95] U. G. Mangai, S. Samanta, S. Das, and P. R. Chowdhury, "A survey of decision fusion and feature fusion strategies for pattern classification," *IETE Tech. Rev.*, vol. 27, no. 4, pp. 293–307, 2010, doi: 10.4103/0256-4602.64604.

[96] D. Singhal and K. S. Swarup, "Electricity price forecasting using artificial neural networks," *Int. J. Electr. Power Energy Syst.*, vol. 33, no. 3, pp. 550–555, 2011, doi: https://doi.org/10.1016/j.ijepes.2010.12.009.

[97] M. Ranjbar, S. Soleymani, N. Sadati, and A. M. Ranjbar, "Electricity price forecasting using artificial neural network," in *2006 International Conference on Power Electronic, Drives and Energy Systems*, 2006, pp. 1–5, doi: 10.1109/PEDES.2006.344294.

[98] A. Khotanzad, E. Zhou, and H. Elragal, "A neuro-fuzzy approach to short-term load forecasting in a price-sensitive environment," *IEEE Trans. Power Syst.*, vol. 17, no. 4, pp. 1273–1282, 2002, doi: 10.1109/TPWRS.2002.804999.

[99] N. M. Pindoriya, S. N. Singh, and S. K. Singh, "An adaptive wavelet neural network-based energy price forecasting in electricity markets," *IEEE Trans. Power Syst.*, vol. 23, no. 3, pp. 1423–1432, 2008, doi: 10.1109/TPWRS.2008.922251.

[100] A. Pourdaryaei, H. Mokhlis, H. A. Illias, S. H. R. A. Kaboli, S. Ahmad, and S. P. Ang, "Hybrid ANN and artificial cooperative search algorithm to forecast short-

term electricity price in de-regulated electricity market," *IEEE Access*, vol. 7, pp. 125369–125386, 2019, doi: 10.1109/ACCESS.2019.2938842.

[101] C. Wan, Z. Xu, Y. Wang, Z. Y. Dong, and K. P. Wong, "A Hybrid approach for probabilistic forecasting of electricity price," *IEEE Trans. Smart Grid*, vol. 5, no. 1, pp. 463–470, 2014, doi: 10.1109/TSG.2013.2274465.

[102] D. Bissing, M. T. Klein, R. A. Chinnathambi, D. F. Selvaraj, and P. Ranganathan, "A hybrid regression model for day-ahead energy price forecasting," *IEEE Access*, vol. 7, pp. 36833–36842, 2019, doi: 10.1109/ACCESS.2019.2904432.

[103] T. Kristiansen, "Forecasting Nord Pool day-ahead prices with an autoregressive model," *Energy Policy*, vol. 49, pp. 328–332, 2012.

[104] J. P. S. Catalão, H. M. I. Pousinho, and V. M. F. Mendes, "Hybrid wavelet-pso-anfis approach for short-term electricity prices forecasting," *IEEE Trans. Power Syst.*, vol. 26, no. 1, pp. 137–144, 2011, doi: 10.1109/TPWRS.2010.2049385.

[105] C. Gu, W. Yang, Y. Song, and F. Li, "Distribution network pricing for uncertain load growth using fuzzy set theory," *IEEE Trans. Smart Grid*, vol. 7, no. 4, pp. 1932–1940, 2016, doi: 10.1109/TSG.2016.2518175.

[106] A. Y. Alanis, "Electricity prices forecasting using artificial neural networks," *IEEE Lat. Am. Trans.*, vol. 16, no. 1, pp. 105–111, 2018, doi: 10.1109/TLA.2018.8291461.

[107] D. Lee, H. Shin, and R. Baldick, "Bivariate probabilistic wind power and real-time price forecasting and their applications to wind power bidding strategy development," *IEEE Trans. Power Syst.*, vol. 33, no. 6, pp. 6087–6097, 2018, doi: 10.1109/TPWRS.2018.2830785.

[108] K. B. Sahay and M. M. Tripathi, "Day ahead hourly load forecast of PJM electricity market and ISO new england market by using artificial neural network," in *2013 IEEE Innovative Smart Grid Technologies-Asia (ISGT Asia)*, 2013, pp. 1–5, doi: 10.1109/ISGT-Asia.2013.6698744.

[109] K. Upadhyay, M. Tripathi, and S. Singh, "An approach to short term load forecasting using market price signal," *19th Int. Conf. Electr. Distrib.*, 2007.

[110] T. Abreu, C. Minussi, M. Lopes, U. Alves, and A. Lotufo, "Electrical Customer

Profile Using Fuzzy Logic Theory," *IEEE Lat. Am. Trans.*, vol. 18, no. 08, pp. 1353–1361, 2020, doi: 10.1109/TLA.2020.9111670.

[111] A. T. Al-Awami, N. A. Amleh, and A. M. Muqbel, "Optimal demand response bidding and pricing mechanism with fuzzy optimization: Application for a virtual power plant," *IEEE Trans. Ind. Appl.*, vol. 53, no. 5, pp. 5051–5061, 2017, doi: 10.1109/TIA.2017.2723338.

[112] E. E. Elattar, S. K. Elsayed, and T. A. Farrag, "Hybrid local general regression neural network and harmony search algorithm for electricity price forecasting," *IEEE Access*, vol. 9, pp. 2044–2054, 2021, doi: 10.1109/ACCESS.2020.3048519.

[113] H. Mosbah and M. El-hawary, "Hourly electricity price forecasting for the next month using multilayer neural network," *Can. J. Electr. Comput. Eng.*, vol. 39, no. 4, pp. 283–291, 2016, doi: 10.1109/CJECE.2016.2586939.

[114] M. Usman, Z. Ali Khan, I. U. Khan, S. Javaid, and N. Javaid, "Data analytics for short term price and load forecasting in smart grids using enhanced recurrent neural network," in *2019 Sixth HCT Information Technology Trends (ITT)*, 2019, pp. 84–88, doi: 10.1109/ITT48889.2019.9075065.

[115] M. Alamaniotis and L. H. Tsoukalas, "Implementing smart energy systems: Integrating load and price forecasting for single parameter based demand response," in *2016 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe)*, 2016, pp. 1–6, doi: 10.1109/ISGTEurope.2016.7856299.

[116] L. Wang, Y. Wang, J. Chen, and B. Wen, "Power purchase model based on interruptible load management and balancing market forecast," in *2017 4th International Conference on Information Science and Control Engineering (ICISCE)*, 2017, pp. 1596–1600, doi: 10.1109/ICISCE.2017.333.

[117] F. Luo, X. Yang, W. Wei, H. Lu, T. Zhang, and J. Shao, "Data driven load forecasting method considering demand response," in *2020 IEEE Power Energy Society General Meeting (PESGM)*, 2020, pp. 1–5, doi: 10.1109/PESGM41954.2020.9281779.

[118] G. Cau, D. Cocco, M. Petrollese, S. Knudsen Kær, and C. Milan, "Energy management strategy based on short-term generation scheduling for a renewable

microgrid using a hydrogen storage system," *Energy Convers. Manag.*, vol. 87, pp. 820–831, 2014, doi: https://doi.org/10.1016/j.enconman.2014.07.078.

[119] M. Marzband, M. Ghadimi, A. Sumper, and J. L. Domínguez-García, "Experimental validation of a real-time energy management system using multi-period gravitational search algorithm for microgrids in islanded mode," *Appl. Energy*, vol. 128, pp. 164–174, 2014, doi: https://doi.org/10.1016/j.apenergy.2014.04.056.

[120] J. Pascual, J. Barricarte, P. Sanchis, and L. Marroyo, "Energy management strategy for a renewable-based residential microgrid with generation and demand forecasting," *Appl. Energy*, vol. 158, pp. 12–25, 2015, doi: https://doi.org/10.1016/j.apenergy.2015.08.040.

[121] B. Zhao, M. Xue, X. Zhang, C. Wang, and J. Zhao, "An MAS based energy management system for a stand-alone microgrid at high altitude," *Appl. Energy*, vol. 143, pp. 251–261, 2015, doi: https://doi.org/10.1016/j.apenergy.2015.01.016.

[122] J. Xiao, P. Wang, L. Setyawan, and Q. Xu, "Multi-Level Energy Management System for Real-Time Scheduling of DC Microgrids With Multiple Slack Terminals," *IEEE Trans. Energy Convers.*, vol. 31, no. 1, pp. 392–400, 2016, doi: 10.1109/TEC.2015.2488639.

[123] T. Lv and Q. Ai, "Interactive energy management of networked microgrids-based active distribution system considering large-scale integration of renewable energy resources," *Appl. Energy*, vol. 163, pp. 408–422, 2016, doi: https://doi.org/10.1016/j.apenergy.2015.10.179.

[124] F. Farzan, M. A. Jafari, R. Masiello, and Y. Lu, "Toward optimal day-ahead scheduling and operation control of microgrids under uncertainty," *IEEE Trans. Smart Grid*, vol. 6, no. 2, pp. 499–507, 2015, doi: 10.1109/TSG.2014.2368077.

[125] A. U. Rehman *et al.*, "Efficient energy management system using firefly and harmony search algorithm," 2017.

[126] A. Saba *et al.*, "Home energy management using firefly and harmony search algorithm," 2017.

[127] O. Oladeji and O. Olakanmi, "A genetic algorithm approach to energy consumption

scheduling under demand response," *IEEE Int. Conf. Adapt. Sci. Technol. ICAST*, vol. 2015, 2015, doi: 10.1109/ICASTECH.2014.7068096.

[128] U. Asgher *et al.*, "Smart energy optimization using heuristic algorithm in smart grid with integration of solar energy sources," *Energies*, vol. 11, p. 3494, 2018, doi: 10.3390/en11123494.

[129] M. Latifi, A. Khalili, A. Rastegarnia, S. Zandi, and W. M. Bazzi, "A distributed algorithm for demand-side management: Selling back to the grid," *Heliyon*, vol. 3, no. 11, p. 457, Nov. 2017, doi: 10.1016/j.heliyon.2017.e00457.

[130] A. Ahmed *et al.*, "Performance Measurement of Energy Management Controller Using Heuristic Techniques," 2017, pp. 181–188, doi: 10.1007/978-3-319-61566-0_17.

[131] Z. H. Kilimci *et al.*, "An improved demand forecasting model using deep learning approach and proposed decision integration strategy for supply chain," *Complexity*, vol. 2019, p. 9067367, 2019, doi: 10.1155/2019/9067367.

[132] G. Nguyen *et al.*, "Machine learning and deep learning frameworks and libraries for large-scale data mining: a survey," *Artif. Intell. Rev.*, vol. 52, no. 1, pp. 77–124, Jun. 2019, doi: 10.1007/s10462-018-09679-z.

[133] A. Jadidi, R. Menezes, N. De Souza, and A. C. De Castro Lima, "Short-term electric power demand forecasting using NSGA II-ANFIS Model," *Energies*, vol. 12, no. 10, p. 1891, May 2019, doi: 10.3390/en12101891.

[134] J. Cai, J. Luo, S. Wang, and S. Yang, "Feature selection in machine learning: A new perspective," *Neurocomputing*, vol. 300, pp. 70–79, Jul. 2018, doi: 10.1016/j.neucom.2017.11.077.

[135] J. Cai, J. Luo, S. Wang, and S. Yang, "Feature selection in machine learning: A new perspective," *Neurocomputing*, vol. 300, pp. 70–79, 2018, doi: https://doi.org/10.1016/j.neucom.2017.11.077.

[136] A. T. Eseye and M. Lehtonen, "Short-term forecasting of heat demand of buildings for efficient and optimal energy management based on integrated machine learning models," *IEEE Trans. Ind. Informatics*, vol. 16, no. 12, pp. 7743–7755, 2020, doi: 10.1109/TII.2020.2970165.

[137] A. Hassanat, K. Almohammadi, E. Alkafaween, and E. Abunawas, “Choosing mutation and crossover ratios for genetic algorithms — a review with a new dynamic approach,” 2019, doi: 10.3390/info10120390.

[138] K. Keerthi Vasan and B. Surendiran, “Dimensionality reduction using principal component analysis for network intrusion detection,” *Perspect. Sci.*, vol. 8, pp. 510–512, 2016, doi: https://doi.org/10.1016/j.pisc.2016.05.010.

[139] M. R. Haq and Z. Ni, “A new hybrid model for short-term electricity load forecasting,” *IEEE Access*, vol. 7, pp. 125413–125423, 2019, doi: 10.1109/ACCESS.2019.2937222.

[140] M. F. Elahe, M. Jin, and P. Zeng, “An adaptive and parallel forecasting strategy for short-term power load based on second learning of error trend,” *IEEE Access*, vol. 8, pp. 201889–201899, 2020, doi: 10.1109/ACCESS.2020.3036096.

[141] Peter Stone, Rodney Brooks, Erik Brynjolfsson, Ryan Calo, Oren Etzioni, and Greg Hager, “Artificial Intelligence and Life in 2030." one hundred year study on artificial intelligence,” Stanford, CA, Sep. 2016. Accessed: Jun. 05, 2021. [Online]. Available: https://ai100.stanford.edu/2016-report.

[142] M. Erol-Kantarci and H. T. Mouftah, “Using wireless sensor networks for energy-aware homes in smart grids,” in *The IEEE symposium on Computers and Communications*, 2010, pp. 456–458, doi: 10.1109/ISCC.2010.5546765.

[143] S. K. Dash and P. K. Dash, “Short-term mixed electricity demand and price forecasting using adaptive autoregressive moving average and functional link neural network,” vol. 7, pp. 1241–1255, 2019.

[144] X.-S. Yang and X. He, “Firefly algorithm: Recent advances and applications,” *Int. J. Swarm Intell.*, vol. 1, no. 1, pp. 36–50, 2013, doi: 10.1504/IJSI.2013.055801.

[145] X. S. Yang, “Firefly algorithms for multimodal optimization,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, Mar. 2009, vol. 5792 LNCS, pp. 169–178.

[146] L. Zhang, L. Liu, X. S. Yang, and Y. Dai, “A novel hybrid firefly algorithm for global optimization,” *PLoS One*, vol. 11, no. 9, p. e0163230, Sep. 2016, doi: 10.1371/journal.pone.0163230.