SHORT-TERM ELECTRICITY PRICE FORECASTING IN DEREGULATED ELECTRICITY MARKET BASED ON ENHANCED ARTIFICIAL INTELLIGENCE TECHNIQUES

ALIREZA POURDARYAEI

FACULTY OF ENGINEERING
UNIVERSITY OF MALAYA
KUALA LUMPUR

2020

SHORT-TERM ELECTRICITY PRICE FORECASTING IN DEREGULATED ELECTRICITY MARKET BASED ON ENHANCED ARTIFICIAL INTELLIGENCE TECHNIQUES

ALIREZA POURDARYAEI

THESIS SUBMITTED IN FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

FACULTY OF ENGINEERING
UNIVERSITY OF MALAYA
KUALA LUMPUR

2020

UNIVERSITY OF MALAYA

ORIGINAL LITERARY WORK DECLARATION

Name of Candidate: Alireza Pourdaryaei

Registration/Matric No: KHA140078

Name of Degree: Doctor of Philosophy (Ph.D.)

Thesis ("this Work"):

SHORT-TERM ELECTRICITY PRICE FORECASTING IN DEREGULATED ELECTRICITY MARKET BASED ON ENHANCED ARTIFICIAL INTELLIGENCE TECHNIQUES

Field of Study: Power System

I do solemnly and sincerely declare that:

(1) I am the sole author/writer of this Work;
(2) This Work is original;
(3) Any use of any work in which copyright exists was done by way of fair dealing and for permitted purposes and any excerpt or extract from, or reference to or reproduction of any copyright work has been disclosed expressly and sufficiently and the title of the Work and its authorship have been acknowledged in this Work;
(4) I do not have any actual knowledge nor do I ought reasonably to know that the making of this work constitutes an infringement of any copyright work;
(5) I hereby assign all and every rights in the copyright to this Work to the University of Malaya ("UM"), who henceforth shall be owner of the copyright in this Work and that any reproduction or use in any form or by any means whatsoever is prohibited without the written consent of UM having been first had and obtained;
(6) I am fully aware that if in the course of making this Work I have infringed any copyright whether intentionally or otherwise, I may be subject to legal action or any other action as may be determined by UM.

Candidate's Signature: Date:

Subscribed and solemnly declared before,

Witness's Signature Date:

Name:

Designation:

SHORT-TERM ELECTRICITY PRICE FORECASTING IN DEREGULATED ELECTRICITY MARKET BASED ON ENHANCED ARTIFICIAL INTELLIGENCE TECHNIQUES

ABSTRACT

Electricity price forecasting is considered as one of prime factors for operation, planning and scheduling of price-setter market participants. However, possessing time variant, non-linear and non-stationary behaviors make the electricity price a complex signal. The main challenge in this area is providing highly accurate and efficient day-ahead price forecasting. A suitable feature selection technique, which is able to model the interacting features and nonlinearities of the forecast processes, is still required although researches have been performed for day-ahead forecasting. In this research, a hybrid electricity price forecasting methodology is proposed using two-stage feature selection method and optimization using adaptive neuro-fuzzy inference system (ANFIS) technique as a forecasting engine. An important contribution of the proposed method is modeling of interaction in addition to relevancy and redundancy based on information-theoretic criteria for the feature selection. A multi-objective feature technique is developed in this study to extract the most influential subsets of input variables with the maximum relevancy and minimum redundancy. The proposed feature selection technique comprises of Multi-objective Binary-valued Backtracking Search Algorithm (MOBBSA). It is used to search within a number of input variables combinations and to select the feature subsets, which minimizes simultaneously vice-versa the estimation error and the feature numbers. In the developed method of multi-objective feature determination, MOBBSA is used to search within different combinations of input variables and to select the non-dominated feature subsets. ANFIS is applied as an evaluation metric to determine the performance of every feature subset.

The other foremost contribution of the work is proposing a hybrid electricity price forecasting technique to provide more accurate forecasts. This merit is provided by balancing the exploitation of solution structure and exploration of its appropriate weighting factors through the use of Backtracking Search Algorithm (BSA) as an efficient optimization algorithm in learning process of ANFIS approach. Real-world electricity demand and price dataset from Ontario and Australia power markets, which are reported as among the most volatile market worldwide, have been used to validate the performance of the proposed approach. Finally, the obtained results corroborate the premise of the proposed method through the enhanced accuracy compared to the existing artificial intelligence-based models.

Keywords: Adaptive Neuro-Fuzzy Inference System, Backtracking Search Algorithm, Electricity Price Forecasting, Feature Selection

RAMALAN HARGA ELEKTRIK JANGKA PENDEK DALAM PASARAN ELEKTRIK YANG DEREGULASI BERDASARKAN TEKNIK KECERDASAN BUATAN YANG DIPERTINGKATKAN

ABSTRAK

Ramalan beban dan harga perlu dipertimbangkan dalam operasi perancangan yang optimum bagi pasaran elektrik yang berdaya saing. Ramalan harga elektrik merupakan maklumat penting bagi pengurus pasaran elektrik dan pengguna. Walau bagaimanapun, masa varian, tingkahlaku-tidak lelurus dan pegun telah mengakibatkan ramalan harga elektrik menjadi kompleks. Cabaran utama dalam bidang ini adalah menyediakan harga hari akan datang yang tepat dan cekap. Walaupun kajian telah dijalankan bagi ramalan hari akan datang, satu teknik pemilihan ciri-ciri yang cekap dengan keupayaan pemodelan yang tidak-lelurus dan interaksi ciri-ciri proses ramalan masih diperlukan. Dalam kajian ini, harga elektrik gabungan dengan ramalan metodologi adalah dicadangkan. Metodologi yang dicadangkan terdiri daripada dua peringkat ciri-ciri pemilihan kaedah dan dioptimumkan berdasarkan sistem Inferens Neuro-Fuzzy (ANFIS) sebagai enjin ramalan. Sumbangan yang penting bagi kaedah yang dicadangkan adalah pemodelan interaksi perkaitan dan redundansi, berdasarkan kriteria kepada teori maklumat, ciri-ciri pemilihan. Pendekatan pemilihan ciri-ciri pelbagai objektif yang dibangunkan dalam kajian ini untuk pemilihan subsets paling berpengaruh kepada pembolehubah input dengan perkaitan yang maksimum dan minimum redundansi. Teknik pemilihan ciri-ciri yang dicadangkan ini terdiri daripada algoritma multi-objektif nilai binari jejak ke belakang (MOBBSA) sebagai suatu algoritma carian evolusi yang cekap mencari dalam kombinasi yang berlainan untuk pembolehubah input serta untuk memilih ciri-ciri subsets, yang pada masa yang sama mengurangkan bilangan ciri-ciri dan ralat penganggaran. Dalam kaedah pemilihan ciri-ciri pelbagai objektif yang maju, MOBBSA digunakan untuk mencari dalam kombinasi

yang berlainan untuk pembolehubah input dan memilih ciri bebas menguasai subsets, manakala ANFIS akan digunakan sebagai metrik penilaian untuk menentukan prestasi masing-masing mempunyai subset. Satu lagi sumbangan besar adalah teknik ramalan hybrid harga elektrik hibrid yang mampu memberi ramalan yang lebih tepat. Merit ini disediakan dengan menyeimbangkan eksploitasi penyelesaian struktur dan faktor-faktor berwajaran yang sesuai melalui penggunaan algoritma carian jejak ke belakang (BSA) sebagai suatu algoritma pengoptimuman yang cekap dalam proses pendekatan pembelajaran penerokaan bagi ANFIS. Dataset permintaan dan harga elektrik dunia sebenar dari pasaran kuasa Ontario dan Australia yang dilaporkan sebagai antara pasaran tidak menentu yang paling ketara yang di seluruh dunia, telah digunakan untuk menguji prestasi pendekatan yang dicadangkan. Keputusan eksperimen menunjukkan kajian yang dicadangkan telah meningkat ketepatan dalam perbandingan dengan model kepintaran tiruan yang sediada.

Kata kunci: Sistem mudah suai inferens neuro-fuzzy, Algoritma jejak ke belakang, ramalan harga elektrik, Pemilihan ciri-ciri

ACKNOWLEDGEMENTS

All deepest thanks are due to the almighty God, the merciful, and the compassionate for uncountable gifts given to me.

First and foremost, I thank my supervisors, Professor Ir. Dr. Hazlie Bin Mokhlis and Associate Professor Ir. Dr. Hazlee Azil Bin Illias**.** It has been an honor to be one of their PhD students. They have taught me on how good research in power system lab can be done. I appreciate all of their contributions of time, ideas, experience productive and simulating. I thank them for the pleasant and trusting cooperation as well as the possibilities to present the results of my work. They have also been a source of friendships and good advice and collaboration.

Personal thanks to two of my special friends, Dr. Hamid Reza Kaboli and Dr. Shameem Ahmad for supporting my research project and also the members of power systems group (Toufik, Saman, Jabir, Shahrani, Munir, Suzana, Junaid and Abdullah). They have contributed immensely to my personal and professional for their time to discuss with me about different issues related to my work at the University of Malaya. I really enjoyed working with them in my research project.

Lastly, I thank my family for all their love and encouragement, especially to my grandfather, grandmother, aunty, my parents and uncle who have raised me with a love of science and supported me in all of my pursuits.

TABLE OF CONTENTS

Abstract iii

Abstrak v

Acknowledgements vii

Table of Contents viii

List of Figures xii

List of Tables.. xiv

List of Symbols and Abbreviations xvi

CHAPTER 1: INTRODUCTION 1

1.1 Introduction.. 1

1.2 Problem Statement.. 2

1.3 Research Objectives.. 5

1.4 Scope of Study.. 5

1.5 Thesis outline.. 6

CHAPTER 2: LITERATURE REVIEW 9

2.1 Introduction.. 9

2.2 Factors Influencing Price Forecasting 10

2.2.1 Electric Power Demand 10

2.2.2 Weather Conditions 10

2.2.3 Fuel Cost... 10

2.2.4 Available Transmission Capacity.. 10

2.2.5 Generation Reserves 11

2.2.6 Emission Allowances... 11

2.3 Electricity Price Forecasting classification and Methodology 12

2.3.1 Price Forecasting classification 12

2.3.2 Price forecasting Methodology 13

2.4 Electricity Price Forecasting (EPF) approaches 15

2.4.1 Multi-agent models 16

2.4.2 Fundamental models 17

2.4.3 Reduced-form models 17

2.4.4 Statistical models 18

2.4.4.1 Autoregressive Integrated Moving Average 19

2.4.4.2 Generalized autoregressive conditional heteroscedasticity models 21

2.4.5 Computational intelligence models 22

2.4.5.1 Artificial Neural Network 23

2.4.5.2 Support Vector Regression 26

2.4.6 Hybrid models 26

2.5 Fundamental price drivers and input variables 27

2.5.1 Seasonality 27

2.5.2 Variable selection 28

2.6 Progress on electricity price forecasting via artificial intelligence based methods 29

2.7 Summary 37

CHAPTER 3: RESEARCH METHODOLOGY 39

3.1 Introduction 39

3.2 Development of price Forecasting method and enhanced feature selection 39

3.3 Features (price and demand) selection 40

3.3.1 First stage of filtering data using mutual information (MI) 43

3.3.1.1 Mutual information technique 43

3.3.2 Second stage of selected features using (MOBBSA+BSA) hybrid technique..........46

3.3.2.1 Multi-objective backtracking search algorithm 46

3.3.2.1.1 Backtracking Search Algorithm 47

3.3.2.1.2 Multi-objective Backtracking Search Algorithm..........54

3.3.2.2 Adaptive Neuro-Fuzzy Inference System..........59

3.4 The proposed (MI+(MOBBSA+ANFIS)) hybrid feature selection method for short-term electricity price forecasting..........65

3.5 The proposed (ANFIS+BSA) hybrid technique for short-term electricity price forecasting..........68

3.6 Summary..........72

CHAPTER 4: RESULTS AND DISCUSSIONS 73

4.1 Introduction..........73

4.2 Selected Features for Short Term Electricity Price Forecasting..........73

4.3 Simulation Results and Discussions..........82

4.3.1 Validation of the Model Using Machine Learning Methods in Ontario Market..........84

4.3.2 Validation of the Model Using Statistical Methods in Ontario Market..........90

4.4 Australia Day-ahead Electricity Market..........92

4.5 Numerical results..........97

4.5.1 Validation of the Model Using Machine Learning Methods..........97

4.5.2 Validation of the Model Using Statistical Methods..........103

4.6 Comparison of Different Forecasting Methods based on MAPE..........104

4.7 Summary..........106

CHAPTER 5: CONCLUSIONS AND FUTURE WORKS 108

5.1 Conclusions 108

5.2 Future Works 110

REFERENCES 111

LIST OF PUBLICATIONS 121

LIST OF FIGURES

Figure 2.1: Factors Influencing Price Forecasting 11

Figure 2.2: Price Forecasting Time Horizon 13

Figure 2.3: The basic outlook of Flowchart of the EPF 15

Figure 2.4: The general pictorial expression of EPF approach 16

Figure 2.5: Configuration of ANN 23

Figure 2.6: The MLP architecture 25

Figure 3.1: Structure of development of electricity price forecasting 40

Figure 3.2: Graphical representation between mutual information and conditional entropy 45

Figure 3.3: The Pareto optimal set for the two objective functions (A and B are two sample from non-dominated solutions) 54

Figure 3.4: The general structure of ANFIS 61

Figure 3.5: The procedure of feature selection development for short-term electricity price forecasting (EPF) 67

Figure 3.6: The procedure of ANFIS-BSA for short-term electricity price forecasting (EPF) 71

Figure 4.1: The performance of ANFIS-BSA during training of design phase and testing phase and its corresponding error for EPF of Ontario in February 2017 85

Figure 4.2: The performance of ANFIS-BSA during training of design phase and testing phase and its corresponding error for EPF of Ontario in May 2017 87

Figure 4.3: The performance of ANFIS-BSA during training of design phase and testing phase and its corresponding error for EPF of Ontario in August 2017 88

Figure 4.4: The performance of ANFIS-BSA during training of design phase and testing phase and its corresponding error for EPF of Ontario in November 2017 90

Figure 4.5: The performance of ANFIS-BSA during training of design phase and testing phase and its corresponding error for EPF of Queensland in February2018 98

Figure 4.6: The performance of ANFIS-BSA during training of design phase and testing phase and its corresponding error for EPF of Queensland in May2018 100

Figure 4.7: The performance of ANFIS-BSA during training of design phase and testing phase and its corresponding error for EPF of Queensland in August 2018 101

Figure 4.8: The performance of ANFIS-BSA during training of design phase and testing phase and its corresponding error for EPF of Queensland in November 2018............. 102

Figure 4.9: The performance of AI –based method for EPF of Ontario based on MAPE (%) in 2017... 105

Figure 4.10: The performance of AI –based method for EPF of Queensland based on MAPE (%) in 2018 105

LIST OF TABLES

Table 2.1: Review on most recently developed artificial intelligent methods in the literature which have been applied for electricity price forecasting 35

Table 3.1: General structure of BSA 48

Table 3.2: Pseudocode of BSA 53

Table 4.1: The optimal subsets of input variables selected in Ontario by different multi-objective feature selection techniques and their corresponding performances in terms of RMSE value 80

Table 4.2: The optimal subsets of input variables selected in Ontario by different multi-objective feature selection techniques and their corresponding performances in terms of RMSE value 81

Table 4.3: Parameter setting of applied methods 84

Table 4.4: Comparison between forecasting accuracy of studied methods for EPF of Ontario in February 2017 85

Table 4.5: Comparison between forecasting accuracy of studied methods for EPF of Ontario in May 2017 86

Table 4.6: Comparison between forecasting accuracy of studied methods for EPF of Ontario in August 2017 88

Table 4.7: Comparison between forecasting accuracy of studied methods for EPF of Ontario in November 2017.. 89

Table 4.8: Statistical factors of the ANFIS-BSA model for EPF of Ontario in February, May, August and November 2017 92

Table 4.9: The optimal subsets of input variables selected for QLD by different multi-objective feature selection techniques and their corresponding performances in terms of RMSE value 95

Table 4.10: The optimal subsets of input variables selected for QLD by different multi-objective feature selection techniques and their corresponding performances in terms of RMSE value 96

Table 4.11: Comparison between forecasting accuracy of studied methods for EPF of Queensland in February2018 98

Table 4.12: Comparison between forecasting accuracy of studied methods for EPF of Queensland in May2018 99

Table 4.13: Comparison between forecasting accuracy of studied methods for EPF of Queensland in August 2018 101

Table 4.14: Comparison between forecasting accuracy of studied methods for EPF of Queensland in November 2018 102

Table 4.15: Statistical factors of the ANFIS-BSA model for EPF of Queensland in February, May, August and November 2018 104

LIST OF SYMBOLS AND ABBREVIATIONS

k or k'	:	Regression lines
h_i	:	Actual output
(A_i, B_i)	:	Fuzzy sets
(x, y)	:	Inputs to node i
R_o^2	:	The squared correlation coefficient
$R_{o'}^2$	:	The squared correlation coefficient between experimental and predicted values
t_i	:	Predicted output
R_m	:	Confirmed indicator
$\overline{w}$	:	Normalized firing strength of a rule
μ_{Ai}	:	Membership function for Ai fuzzy sets
μ_{Bi}	:	Membership function for Bi fuzzy sets
t	:	Hourly interval
Z	:	Un normalized value
$\bar{Z}$	:	Normalized value
$:=$	:	Update operation
∂ and ε	:	Solutions in Pareto optimal set
$EP\ (t)_{observed}$	:	Observed electricity price at time t
$EP(t)$	:	Predicted electricity price at time t
f	:	Activation function
F	:	Wiener process
$f_i(x, y; p_i, q_i, r_i)$	:	Output of the Sugeno type FIS
f_j^{max}	:	Minimum value of the j^{th} objective function
f_j^{min}	:	Maximum value of the j^{th} objective function

g	:	Transfer function
$g+1$	:	Next generation
g_{best}	:	Overall best value
$e(t)$	:	The whiteness of estimated residuals
h_{0j}	:	Weight assigned to the bias unit of j^{th} neuron in output layer
h_{ij}	:	Connection strength between i^{th} neuron in last hidden layer and j^{th} neuron in output layer
L	:	Hidden layer
low_j	:	Lower search space limits of j^{th} variable
m	:	Number of objectives
map	:	Binary integer-valued matrix
mixrate	:	Control parameter of BSA
Mutant	:	Initial form of trial population
N	:	Standard normal distribution
n_j	:	Neuron in the first hidden layer
n_{Lj}	:	Output signal j^{th} neuron of total (N) neuron in hidden layer
nPop	:	Population size of host nests
nVar	:	Number of respective optimization variable
oldP	:	Historical population
$Ø_p$	:	Parameter of the AR model
p	:	Number of time lags for the autoregressive model
P_{best}	:	Previous best position
permuting	:	Random shuffling function
q	:	Order of the moving average
$Q_{1,i}$	:	Membership grade of a fuzzy set
y_g	:	Global minimum

rand	:	Distributed random numbers
randi	:	Random selection function
T	:	Generated offspring at the end of crossover process
U	:	Uniform distribution function
up_j	:	Upper search space limits of jth variable
P_g	:	Global minimizer
W	:	Weight vector of the regression function
w_{0j}	:	Activation threshold
w_i	:	Initial weight
w_{ij}	:	Connection strength between neurons
w_{max}	:	Maximum boundary of inertia weight
w_{min}	:	Minimum boundary of inertia weight
w_n	:	Inertia weight
w_i	:	Output signal and represent the firing strength of a rule
x_i	:	Input unit
x_n	:	Position of a particel
y_i	:	Productivity of ith individual
a and b in BSA	:	Randomly generated numbers
f_{out}	:	Final output of ANFIS model
Y_{t-p}	:	Time-lagged value
z	:	Output of ANFIS model
e_{t-q}	:	White noise error terms in ARIMA model
U – statistic	:	Thiel's inequality coefficient
σ	:	Width of Guassian MF
c	:	Center of Guassian MF
MI	:	Mutual Information

AI	:	Artificial intelligence
CE	:	Conditional entropy
ANFIS	:	Adaptive neuro-fuzzy inference system
ANN	:	Artificial neural network
AR	:	Auto regressive
ARIMA	:	Auto regressive integrated moving average
BBSA	:	Binary-valued BSA
BP	:	Back propagation
BS	:	Binary search
BSA	:	Backtracking search optimization algorithm
CD	:	Crowding distances
HOED	:	Hourly Ontario electricity demand
DE	:	Differential evolution
DGs	:	Distributed generators
HOEP	:	Hourly Ontario electricity price
EPF	:	Electricity Price Forecasting
ENS	:	Efficient non-dominated sorting
NL_{ED}	:	Number of lag order for electricity demand
NL_{EP}	:	Number of lag order for electricity price
FCM	:	Fuzzy c-means
CNEA	:	Cascaded neuro-evolutionary algorithm
FIS	:	Fuzzy inference system
GA	:	Genetic algorithm
ED(t)	:	Electricity demand at time t
EP(t)	:	Electricity price at time t
GARCH	:	Generalized autoregressive conditional heteroskedasticity

MI	:	Mutual information
P(X)	:	Probability function of X
FSA	:	Fish swarm algorithm
AWNN	:	Adaptive wavelet neural network
OED	:	Orthogonal experimental design
KW	:	Kilowatt
RBFNN	:	radial basis function neural network
VB	:	Virtual Budget
MAPE	:	Mean absolute percentage error
MF	:	Membership function
MLP	:	Multi-layer perceptron
WPT	:	wavelet packet transform
MOBBSA	:	Multi-objective BBSA
MOBSA	:	Multi-objective BSA
MOPSO	:	Multi-objective particle swarm optimization
MW	:	Megawatts
NN	:	Neural network
NSGA	:	Non-dominated sorting genetic algorithm
ISO	:	Independent System Operator
PASA	:	projected assessment of system adequacy
DSM	:	Demand-side management
PrNN	:	Probability neural network
LSSVM	:	Least square support vector machine
TH	:	Threshold
PSO	:	Particle swarm optimization
$H(X\|Y)$	:	Conditional entropy

RACF	:	Residuals autocorrelation function
H(X)	:	Entropy of random variable X
H(X,Y)	:	Joint probability distribution of variable X and Y
RMSE	:	Root mean square error
ACO	:	Ant colony optimization
SC	:	Subtractive clustering
SA	:	Simulated annealing
$/MW.h	:	Dollar per Megawatts hour
DE	:	Differential evolution
GE	:	Genetic evolution
SVM	:	Support vector machine
SVR	:	Support vector regression

CHAPTER 1: INTRODUCTION

1.1 Introduction

As more electricity markets in the world become deregulated, the number of market participants in the electrical power sector is increasing. In this open market, participants need to compete with each other in bidding the best offer of electricity price for trading. In order to offer reasonable bids, forecasting tools are needed to estimate realistic electricity prices for the market. Therefore, accurate electricity price forecasting (EPF) is a key tool for electricity market managers and participants in this deregulated market. Furthermore, with the smart grid technology, forecasting tools become more important since participants able to adjust their bidding strategies with respect to Demand-Side Management (DSM). It is highly crucial for any participants to forecast the volatile wholesale prices with an acceptable level of accuracy to minimize the risk and maximize its profits in day-head trading. For example, improving forecasting accuracy by decreasing 1% of mean absolute percentage error (MAPE) of short-term price will lead to savings of $300,000 per year for a utility with 1 GW peak load (Hong, 2015).

Various methodologies have been introduced for the electricity price forecasting. They vary in the data preprocessing, model selection, calibration and testing phases. More works in this area can be observed due to the electricity market becomes more competitive. Hence, price forecasting is gaining its significance and many experts are beginning to take this factor into consideration. In fact, the energy market is now more open and continuously supporting competition, which makes price forecasting the center of attention of most electricity providers to invest more to introduce efficient and novel approaches. This forecasting helps individual generator to determine the optimal bidding price. Furthermore, decision of joint agreement and investment in a new generation facility in the long run are highly influenced by the price forecasting

(Panapakidis & Dagoumas, 2016). It is imperative to forecast electricity price for the generation companies or Independent System Operator (ISO) as well as different level of customers and investors. Basically, different bidders in competitive electricity market require the future electricity prices to gear up their profit.

Since the recent energy markets are highly becoming deregulated and nonlinear, the price forecasting has become more complex compared to previous days. Due to the nonlinearity and instability of this system, the price forecast has become less accurate. Moreover, it leads to an explosive electricity market by affecting the bidding policies. Due to the uncertainty nature of electricity market price, the supply and demand side managements are experienced with numerous difficulties in day-ahead electricity market (Yang, Ce, & Lian, 2017). The power suppliers may get more privileges in their short-term prediction of their rational offers by knowing the preceding information of electricity market price variations. Moreover, it helps the power suppliers to set up their bidding strategies to enhance their profit in maximum scale. On the other hand, it is very important for the demand side management (DSM) to have the knowledge of market price changes and variations to develop the short-term operational planning. Therefore, in recent years, the researches in electricity market for price forecasting have become more significant.

1.2 Problem Statement

Forecasting of electricity price plays a critical role when it comes to investment decisions and transmission expansion. The forecasting method must be accurate to ensure accurate decision to be made. However, forecasting with high accuracy is relatively a complex task since several exclusive features are needed to be taken into consideration. These are including non-linearity, high volatility, numerous seasonality, mean reversion and price spikes (Yang et al., 2017). In addition, the non-friendly properties of price period line make the forecasting even more difficult (Bask & Widerberg, 2009). Market clearing price (MCP) is categorized by volatility since its hourly determination is kept within a changing and contentious situation (Ziel, Steinert, & Husmann, 2015). A set of various parameters affects MCP evolution. Demand, fuel costing (such as coal and natural gas), hydropower capacity, merit order of engendering manufactories, market participants' tactics and network jam are some of the influential parameters (Singhal & Swarup, 2011). In addition, weather conditions and seasonal variation also affect load profile. On the other hand, electricity prices are determined by a substantial and distinguished set of parameters (Panapakidis & Dagoumas, 2016). Some of these parameters might not be accessible to researchers since they are classified and regulated due to the market competition. Therefore, a number of input combinations should be examined in order to achieve a robust forecast. The inputs, parameters of the model, calibration and the practical implementations should be meticulously selected. This research investigates the day-ahead price forecasting mainly considering typical data such as electricity price and electricity demand. This work proposes a new feature selection technique to extract the features within available data set (typical data), which have the highest effect on the accuracy of the price forecasting for deregulated and competitive market.

Despite the forecasting improvements of metaheuristic methods, some limitations still exist. These include the requirement of predefined knowledge about the structure of the existing relationship between the fitness function variables, which need to be acquired beforehand, excess of control parameters and high sensitivity to initial values of these parameters (Ali Azadeh, Saberi, Ghaderi, Gitiforouz, & Ebrahimipour, 2008). Although the results produced by combinatorial optimization methods are promising, determination of specific point of integration between metaheuristic methods is still a challenge. Furthermore, there is an inescapable increase in effort to achieve proper tuning of control parameters due to the inherent complexity of combinatorial optimization methods. Therefore, different optimization algorithms could be explored in training phase of artificial intelligence method in order to improve the forecasting accuracy.

Remarkable advancements in precise electricity forecasting have been achieved through ANN (Artificial Neural Network), SVR (Support Vector Regression), ANFIS and their hybrid forms. However, a more accurate method to enhance the accuracy of electricity price is still needed. Because AI techniques without combination with metaheuristic techniques do not capable to handle nonlinearity issues related to short term electricity price forecasting as these methods can remove different discriminators in complex environment, and from past experience they can recall, learn and store information which has made them popular in the area of electricity price forecasting. Moreover, all the aforementioned research regarding electricity price forecasting performs well although the forecasting accuracy is still impoverished. In this regard, the application of new forecasting techniques, which can improve the forecasting accuracy of the existing artificial intelligence-based approach, is highly required. Furthermore, the methods are applicable to any deregulated power market to provide realistic forecast prices for a better bidding strategy by generation companies.

The accuracy of the electricity price forecast method needs to be tested to ensure it's able to produce highly accurate results. In order to conduct this test, the most volatile system must be selected and verified through the comprehensive statistical analysis. So far, a comprehensive statistical analysis never been used in electrical price forecasting area. In the past, the work in electrical price forecasting used few indices only namely RMSE, MAE and MAPE. Thus, the accuracy of the developed methods was not fully evaluated.

1.3 Research Objectives

The main aim of this research is to propose electricity price forecasting method based on artificial intelligence (AI) techniques for de-regulated electricity market. The objectives needed to be achieved in this research are:

1. To develop feature selection technique for minimizing number of features and maximizing the forecasting accuracy on electricity price.
2. To improve the forecasting accuracy of AI techniques by optimizing the learning process of artificial neural network (ANN).
3. To propose a hybrid technique for short-term electricity price forecasting based on adaptive neuro-fuzzy inference system (ANFIS) and backtracking search algorithm (BSA).
4. To evaluate the proposed method on competitive electricity markets using comprehensive statistical analysis.

1.4 Scope of Study

The electricity price forecasting is classified as short-term, medium-term and long-term forecasts. However, the pattern of electricity demand depends on seasons. Hence instead of long-term and medium-term price forecasting, short term price forecasting will be more effective in deregulated electricity market for real time decision making. Therefore, this research focuses on short-term price forecasting and the proposed feature selection and forecasting process are developed for short-term electricity price forecasting.

To evaluate the accuracy of electricity price forecasting, competitive electricity markets of Ontario and Queensland electricity markets have been considered as the main case study. In Ontario, Independent Electricity System Operator is bound to the operation of power system, projecting short-term necessity and electricity supply and dealing with the price of real-time spot market electricity for the market of electricity power. Ontario electricity market employs real-time structure, which is also known as single settlement market. Therefore, it leaves a massive challenge in forecasting the electricity price. In Queensland, Australian Energy market Operator (AEMO) delivers a range of planning and forecasting trend to inform decision making including the Electricity Statement of Opportunities (ESOO) and the Integrated System Plan (ISP). The National Electricity Market (NEM) is a single wholesale electricity market for buying and selling electricity between generators and retailers and interconnected regions such as Queensland. Electricity retailers purchase electricity from the wholesale market and consumers in Queensland are able to purchase for the best electricity deal.

In this study, a new feature selection technique is developed to accurately forecast the electricity price. Currently, several techniques have been tried out for electricity prices forecasting. In general, hard and soft computing techniques could be used to forecast electricity prices. To assess the applicability and accuracy of the proposed method, the results are compared with AI-based approaches that have been applied for short-term electricity price forecasting in a successful manner. All the simulations are evaluated in MATLAB environment on a personal computer, with a core-2 quad processor of 2.6 GHz clock speed and 2 GB RAM.

1.5 Thesis outline

This thesis consists of five chapters, where every chapter reports the related topic briefly.

Chapter 1 provides the backgrounds and motivation of the proposed research followed by problem statement. The objectives of the study are presented followed by the scopes of the research. In the end, research methodology and research report outline are given.

Chapter 2 focuses on comprehensive literature review of EPF. The phenomenon of the methodology of the current practice of EPF is discussed with their technical difficulties. At the end of this chapter, recent development and progress on short-term electricity price forecasting for various countries, years, forecasting period and accuracy improvement via various AI based methods are reviewed.

Chapter 3 describes the basic concept of AI-based approaches. Then, most well-known optimization method, which is BSA to improve the accuracy of price forecasting in learning process of AI-based technique, is briefly explained. In addition, the process of multi-objective BSA (MOBBSA) development for selecting the best features is explained after filtering by mutual information (MI) technique in details. The best features are selected through the development of hybrid feature selection technique. In parallel, the new hybrid forecasting method based on the combination of ANFIS-BSA is developed.

Chapter 4 presents the process for developing the multi-objective feature selection to extract the most influential subsets of input variables for short-term EPF. Then, the performance of the optimized ANFIS is compared with other AI-based approaches. Further comparison is provided in this section to specify the most effective structure of input data sets through different feature selection techniques. Moreover, EPF of Ontario mainland is demonstrated hour by hour in preceding year of 2017 by applying different AI forecasting methods and their forecasts are compared with those obtained by proposed approach**.** In this chapter, statistical analysis is provided to show the robustness and applicability of the proposed method to be implemented for future electricity price in competitive electricity market. Also, the validation of the proposed method is presented and discussed for Queensland deregulated market in preceding year of 2018 to show the robustness and effectiveness of developing hybrid technique.

Chapter 5 presents the conclusions and future work that can be conducted for improving the research. A comprehensive list of reference is provided at the end of the thesis.

CHAPTER 2: LITERATURE REVIEW

2.1 Introduction

One of the most important branches of energy forecasting is electricity price forecasting (EPF), which is mainly based on predicting the spot and forward prices in large scale electricity markets. The electricity price forecasts have become an inevitable factor to energy companies over the past 15 years. Since the early 1990s, the electricity market has started to reshape its outdated monopolistic landscape, which is used to be manipulated by the government in the power sectors. This reshaping has been completed with the deregulation and the advent of competitive electricity markets. Thanks to the deregulation of electricity market, the electricity is now traded following with market rules realizing spot and derivative contracts throughout Europe, North America and Australia (Weron, 2014). It should be noted that electricity is a special commodity, which cannot be economically stored and a constant balance between production and user consumption is required for the power system stability. In addition, electricity demands strongly rely on weather conditions (temperature, wind speed, precipitation), daily activities (on-peak vs. off-peak hours, weekdays vs. weekends, holidays) and the intensity of the market. These involving factors make the electricity prices to be dynamic, which is not observed in other markets or commodities. These dynamic prices can be calculated daily, weekly and often annual seasonality and abrupt. It is also under the influence of unpredicted price spikes.

This chapter begins with an overview of EPF raised due to its high penetration. The phenomenon of EPF is discussed with technical difficulties. In the end, this chapter discusses the summary of studies on short-term electricity price forecasting for various countries via AI based methods.

2.2 Factors Influencing Price Forecasting

One of the common behaviors of price is its fluctuations in the deregulated power markets. This fluctuation is mainly due to the different economic and technical factors. In some technical contributions, researchers only considered historical data of prices or both prices and demand in order to project the spot price but they did not take other factors into consideration such as weather, generation reserve, and fuel cost. Figure 2.1 depicts the several factors that impact the spot price.

2.2.1 Electric Power Demand

As an important factor, the system's total demand determines the spot price. According to studies, an increase in system demand imposes a raise in the spot price.

2.2.2 Weather Conditions

Environmental condition also has a direct impact on the electricity demand especially daily temperature. Hence, the fluctuation in weather condition leads to a great impact on the spot price.

2.2.3 Fuel Cost

Another factor that has a great effect on the electricity spot price is the fuel cost, which is the main part of generation cost.

2.2.4 Available Transmission Capacity

Normally, power generators are located relatively far from consumers, which require transmission and can be delivered by transmission network facilities. Some physical constrains in the transmission networks are big hurdles to market participants in order to buy or sell energy. The spot price is also under the influence of these physical constraints.

2.2.5 Generation Reserves

Having the capability of enough energy reserved also affects the electricity spot price. This factor is more significant when the demand for the electricity increases suddenly. With enough reserved energy, the energy generators are able to supply consumers without any difficulty. On the other hand, if there is no adequate reserved electricity energy, consumers will face difficulty in receiving energy and it will lead to an increase on the electricity spot price due to making balance between supply and demand.

2.2.6 Emission allowances:

Emission allowance is another factor that has effect on EPF. Because electricity produced through conventional process causes emission of CO_2, NO_2 and greenhouse gasses. To concern with the environmental protection and to reduce the greenhouse gas, emission allowance is needed to be set. The emission allowance is a type of financial asset that also affects the electricity price.

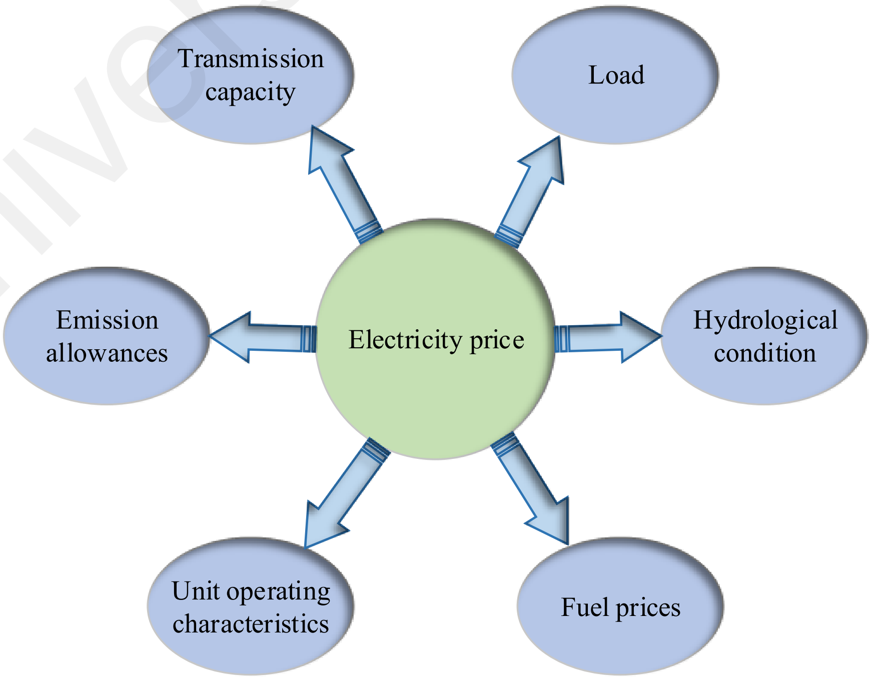

Figure 2.1: Factors Influencing Price Forecasting

From the above discussion it can be summarized that in deregulated electricity markets, electricity price is considered as function of electricity demand. There is high competition for electricity price when electricity demand is so high and the generation is limited. Therefore, with the inherent correlation between electricity price and demand, prediction in smart grid environment such as Ontario electricity and Queensland markets are more complex than the conventional power systems. Therefore, in this work only electricity price and demand in different times are selected for EPF.

2.3 Electricity Price Forecasting classification and Methodology

2.3.1 Price Forecasting classification

Maximizing the profits of numerous market players is the main focus since the introduction of deregulated electricity. When it comes to forecasting, the electricity prices and load are internally correlated and have mutual relationship since they rely on each other and error in one will impose inaccuracy or causing problems to others. Non-storability, seasonal behavior and transportability of the electricity are the main issues, which specifically determine the price. These difficulties bring hindrances and make it complex and do not allow application of forecasting models, which are being widely used in other commodity markets to be adopted for projecting electricity price.

Forecasting the electricity price is commonly classified into three categories as short, medium- and long-term (Weron, 2014), which is shown in Figure 2.2. However, in the literature, there is no any specific border line to differentiate them. In general,

- **Short-term:** It is the most important subcategories for day-to-day market operations and the length of forecasting is from a few minutes up to a few days ahead.

- **Medium-term:** For balancing sheet calculations, risk management and derivatives pricing medium-term forecasting plays a great role. It is from a few days to a few months ahead. Most of the time in electricity price forecasting, the evolution is based on the distribution of prices over certain future time period instead of on the actual point forecasts.
- **Long-term:** This forecasting focuses more on investment profitability analysis and planning and it forecasts for months, quarters, or even years ahead. The information of this forecasting is suitable for determining future sites or fuel sources of power plants.

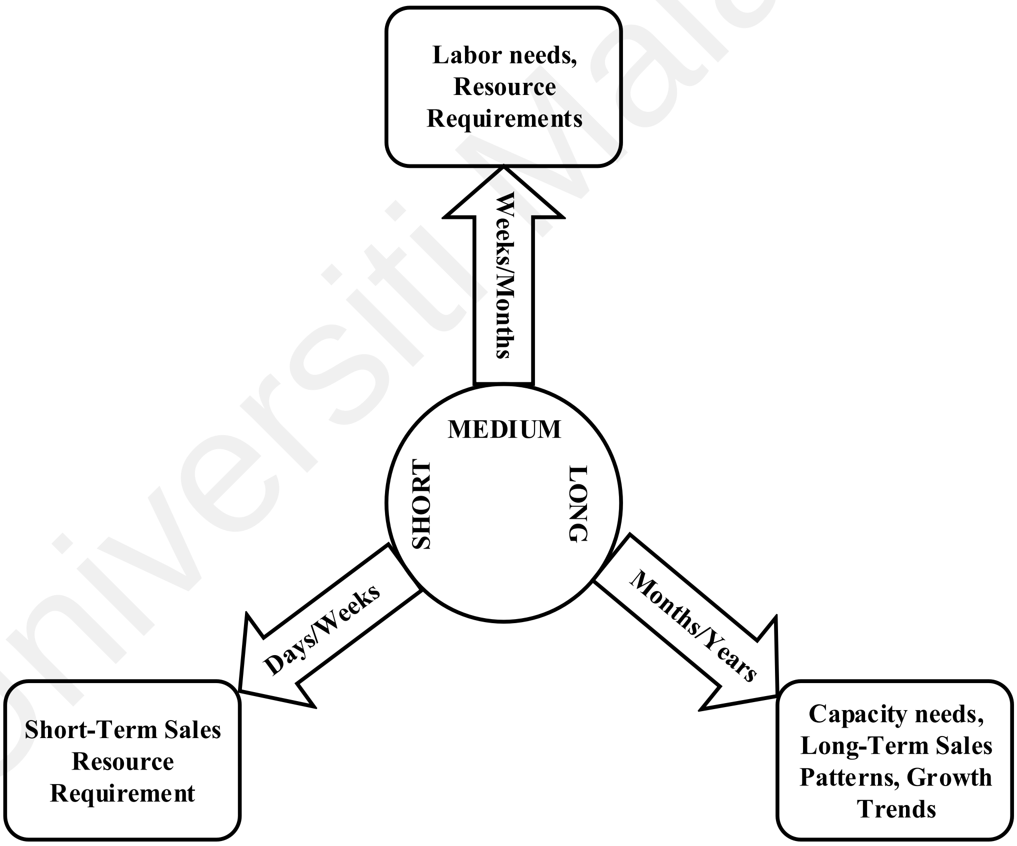

Figure 2.2: Price Forecasting Time Horizon

2.3.2 Price forecasting Methodology

Figure 2.3 illustrates a typical procedure of a price forecasting. The process of time series-based forecasting is also depicted in the flow chart. The process of forecasting is usually initiated by receiving the input data. Past market prices and record of a few weeks to several months are the major information in the input data.

Some other additional inputs are required by other elaborated forecasting models. Such inputs can be the demand factor and/or temperature data. Adopting a modest statistical analysis on the input data set, such as mean and volatility, will aid in model selection and model validation accordingly. In addition, sensitivity analysis of electricity price can also be conducted based on predefined number in threshold relevancy.

One of the most important factors for selecting and design of forecasts models or techniques is the scope of forecasting (e.g. price profile of its volatility). The required precision is also a determinative factor in the selection and design of models or techniques. First, optimization of the models' parameters is usually carried out. The model validation is performed subsequently. If the obtained results are not satisfactory, the process of validation is repeated with different starting parameters. This process will continue until the validation is successful and satisfactory and the model will be adopted to perform the actual forecast.

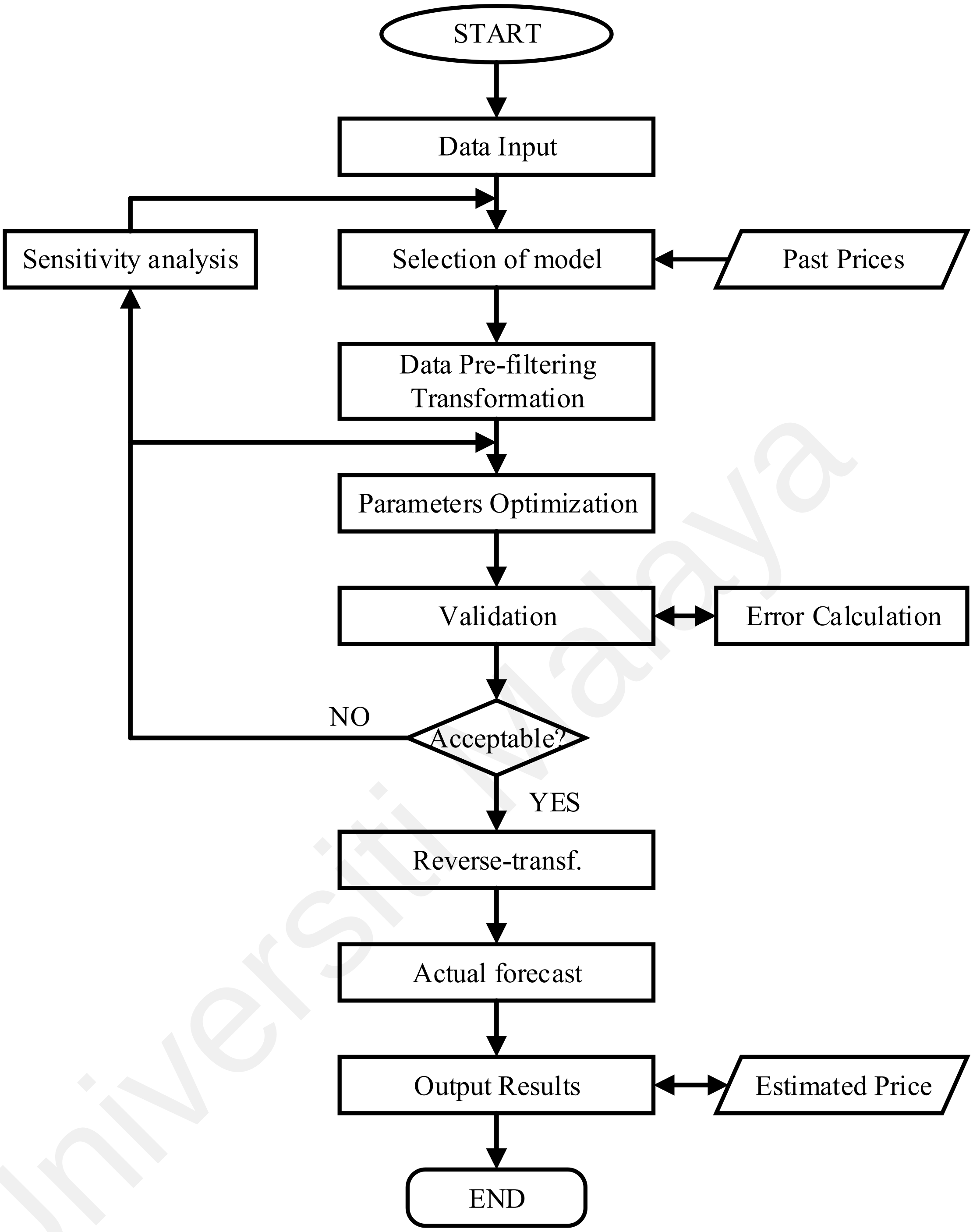

Figure 2.3: The basic outlook of Flowchart of the EPF (Weron, 2014)

2.4 Electricity Price Forecasting (EPF) approaches

Many techniques and methods have been developed or introduced for EPF throughout almost a decade with different rates of success, which can be categorized into six groups (Weron, 2014). Figure 2.4 shows the general EPF methods.

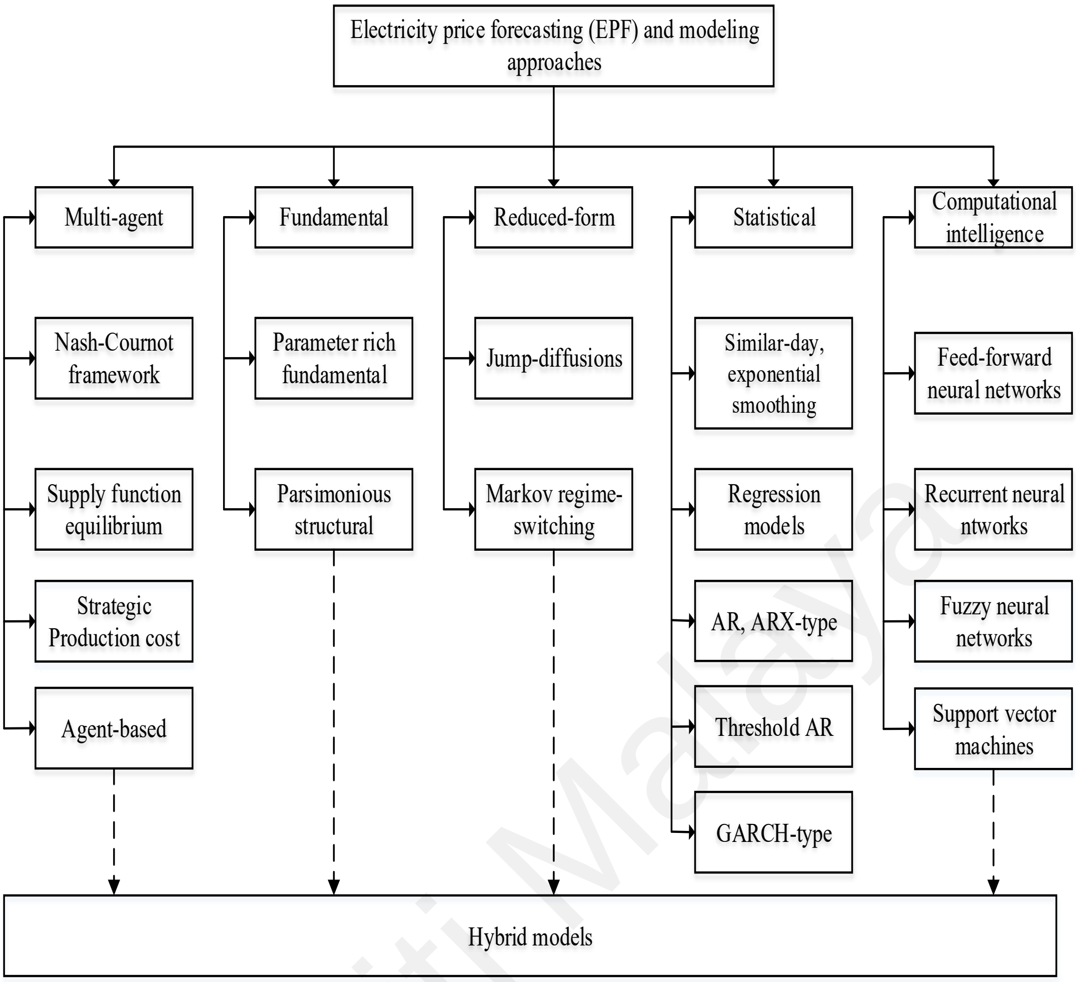

Figure 2.4: The general pictorial expression of EPF approach

2.4.1 Multi-agent models

Multi-agent models such as multi-agent simulation, game theory and equilibrium can simulate the system operation of heterogeneous agents, such as generating units and companies, with ability of interacting with one another and extract the price process by matching the supply and demand in the market (Ventosa, Baıllo, Ramos, & Rivier, 2005). Some of the most common models in this group are cost-based models or production-cost models, PCM, supply function equilibrium (SFE), equilibrium or game theoretic models such as Nash-cournot framework, strategic production-cost models (SPCM) (Ruibal & Mazumdar, 2008) and agent-based models. The main focus of multi-agent models are qualitative problems instead of quantitative results. They are able to give an insights information about whether prices will be above marginal costs or not and how it might affect the players' outcomes. This model also comes with some drawback. When more quantitative results have to be drawn, this model is not a suitable choice especially the prices are expected to be predicted with high level of accuracy.

2.4.2 Fundamental models

Structural or fundamental models capture the basic physical and economic relation, which are accessible in the production and trading of electricity. The fundamental correlations between fundamental drives such as loads, system parameters, weather conditions and others are assumed. In addition, fundamental inputs are normally modeled and predicted separately using statistics, reduced-form or AI. Generally, there are two subclasses of fundamental models, the parameter rich models (Eydeland & Wolyniec, 2003) and parsimonious structure model. Two main challenges in the practical implementation of these models are data availability and incorporation of stochastic changes of the fundamental drivers. Specific assumptions are expected to be made during design of the model, which is about physical and economic associations in the marketplace. As a result, the projected price by this model is highly sensitive to any fluctuations in the assumptions.

2.4.3 Reduced-form models

The statistical properties of electricity prices can be characterized over time by a model named Reduced-form (quantitative, stochastic), which the ultimate objective is derivatives valuation and risk management (Rafal, 2006). The main target of this model is to replicate the main attributes of daily electricity price rather than providing hourly price forecasts with high level of precision. It is relatively comparable to marginal distributions at future time points, price dynamic and associations between commodity prices. If the selected price is not fit for fulfilling the main properties of electricity prices, the extracted results from the model are less likely to be reliable. However, if the model is extremely complicated, it will be prevented by computational burden to be adopted on-line in trading department. Reduced-form models are generally classified into two groups based on the type of market:

- Spot price models provide a frugal representation of the spot prices' dynamics. The major disadvantage of adopting these models is the issue related to pricing derivatives. To be more precise, the identification of the risk premium links spot and forward prices. Jump-diffusion (Benth, Kiesel, & Nazarova, 2012) and Markov regime-switching are the most common subclass of these models.
- Forward price models allow the price of derivatives to be more straightforward but it is limited to those written on the forward price of electricity. These models also come with limitations. They lack the data that can be used for calibration and they are also unable to derive the properties of sport price from forward curve analyses.

2.4.4 Statistical models

A mathematical combination of the preceding prices and/or previous or current value of exogenous factors is used to forecast the current electricity price in statistical methods, which includes econometric and technical analysis. The exogenous factors can be consumption and production figures or weather variable (Weron, 2014).

Additive and multiplicative models are the two popular subclasses of this model. In additive model, the predicted price is the sum of a number of components, while in multiplicative, the anticipated price is the product of a number of factors. The additive model is far more common, but both of them are closely related since a multiplicative model can be transformed into an additive model for log-price. The main reason of popularity of statistical model is that some physical interpretations attached to their components exist, which allow system operators and engineers to comprehend their behavior. However, they also have some limitations in modeling nonlinear behavior of electricity prices and other associated fundamental variable. On the other hand, their performance is not comparatively worse than those of the nonlinear computational intelligence methods. For example, the top winning competitors, among hundreds of participants in the load forecasting track of the Global Energy Forecasting Competition (GEFCom) in 2012, adopted the regression-type models.

Statistical models constitute a very rich class, which include:

- Similar-day and exponential smoothing (Jónsson, Pinson, Nielsen, Madsen, & Nielsen, 2013) methods.
- Regression models (Karakatsani & Bunn, 2008).
- Time series models without (AR, ARMA, ARIMA, Fractional ARIMA - FARIMA, Seasonal ARIMA - SARIMA, Threshold AR-TAR) and models without exogenous variables (ARX, ARMAX, ARIMAX, SARIMAX, TARX).
- Heteroskedastic time series models (GARCH, AR-GARCH) (Koopman, Ooms, & Carnero, 2007)

2.4.4.1 Autoregressive Integrated Moving Average

Among the statistical forecasting techniques, autoregressive integrated moving average (ARIMA) is a high-precision non-structural method for time series forecasting when there is no missing data within the time series. Particularly, in time series analysis, an ARIMA model is considered as a “filter” that tries to separate the time series from the noise and the time series is then extrapolated either to predict future points in the series or to better understand the data. Non-seasonal ARIMA models are generally denoted as ARIMA (p,d,q), where parameters **q**, **d**, and **p** are non-negative integers. The parameter **p** represents the number of time lags for the autoregressive model (i.e. AR (p)), the parameter **d** denotes the number of differences (I) that are needed to make the series stationary and **q** is the order of the moving average (i.e. MA (q)) part (Yuan, Liu, & Fang, 2016).

The autoregressive (AR) part of the ARIMA model with order p is of the form:

$$Y_t = \emptyset_1 Y_{t-1} + \emptyset_2 Y_{t-2} + \cdots + \emptyset_p Y_{t-p} + e_t + c \tag{2.1}$$

where $\emptyset_1, \emptyset_2, \ldots, \emptyset_p$ are the parameters of the model and the independent variables $Y_{t-1}, Y_{t-2} \ldots, Y_{t-p}$ are time-lagged values of the forecast variable. As the forecasts are only dependent on observed values in the previous time periods, this model is named auto regression.

The moving average (MA) part of the ARIMA model consists of the past errors as the explanatory variable. A moving average model with order q is of the form:

$$Y_t = e_t + \theta_1 e_{t-1} + \theta_2 e_{t-2} + \cdots + \theta_q e_{t-q} + c \tag{2.2}$$

where $\theta_1, \theta_2, \ldots, \theta_q$ are the parameters of the model and $e_t, e_{t-1}, e_{t-2}, \ldots, e_{t-q}$ are white noise error terms.

An autoregressive (AR (p)) model is coupled with a moving average (MA (q)) model to form an ARMA (p,q) (autoregressive moving average) model for stationary series. For non-stationary series, differencing is added to ARMA model. Differencing is a method to stabilize the mean of the series, eliminate seasonality and consequently makes the series stationary. The first difference between consecutive observations is calculated according to Eq. (2.3) to mathematically difference the data.

$$\Delta Y_t = Y_t - Y_{t-1} \tag{2.3}$$

Finally, the general form of ARIMA model is formulated in Eq. (2.4), which requires at least $p+d$ pre samples to initialize the time series.

$$\Delta_d Y_t = c + \emptyset_1 \Delta_d Y_{t-1} + \emptyset_2 \Delta_d Y_{t-2} + \cdots + \emptyset_p \Delta_d Y_{t-p} + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q} \quad (2.4)$$

2.4.4.2 Generalized autoregressive conditional heteroscedasticity models

Generalized autoregressive conditional heteroskedasticity (GARCH) is another model and is aimed at modeling the volatility of prices rather than at modeling and forecasting only the changing price in ARIMA model. In this model, the moment of a time series as variant is taken into consideration. In other words, the error term, which is real value minus forecast value, does not have zero mean and constant variance as with an ARIMA process. Then, the error term can be assumed to be related and also be modeled by using an AR process. Hence, a GARCH has the capability of measuring the indirect volatility of a time series because of price spikes. A GARCH (p,q) model can be determined by Assuming a time series with the function x_t and a constant mean offset then:

$$x_t = \mu + \varepsilon_t \quad (2.5)$$

where μ denotes offset and $\varepsilon_t = \sigma_t z_t$

$$\varepsilon_t{}^2 = c + \sum_{i=1}^{q} \varphi_i \varepsilon_{t-i}{}^2 + \sum_{i=1}^{p} \psi_i \varepsilon_{t-i}{}^2 \quad (2.6)$$

where p represents the order of GRACH terms σ^2, and q represents the order of ARCH terms ε^2. It is clear that the Eq. (2.6) with p=0, the GARCH (0,q) turns into an ARCH (q) model. It should be noted that GARCH model is only applicable for stationary time series. Hence, before applying this model, the following equation must be provided for stationary time series:

$$\sum_{i=1}^{q}\varphi_i + \sum_{i=1}^{p}\psi_i < 1 \tag{2.7}$$

2.4.5 Computational intelligence models

Computational intelligence techniques including artificial machine learning, intelligence-based, non-parametric and non-linear statistical combine elements of learning, evolution and fuzziness to design solutions are able to adapt to complicated dynamic systems. They might be considered as "intelligent" in this sense. Some of the subclasses of computational intelligence are artificial neural networks (Keles, Scelle, Paraschiv, & Fichtner, 2016), support vector machines (SVM) (Yan & Chowdhury, 2013), and fuzzy systems (Rodriguez & Anders, 2004), which are widely used in EPF. These models are capable of solving and handling extremely complex and non-linear questions. Generally, the computational intelligence is far better than the statistical techniques at modeling the characteristics of electricity prices. However, this flexibility is not always strength. In fact, sometimes, it turns into a weakness for these models and their adaptability to nonlinear and spiky behaviors will not always end up with better results or probabilistic forecasts. Figure 2.5 demonstrates various types of artificial neural network that can be implemented for different forecasting purposes.

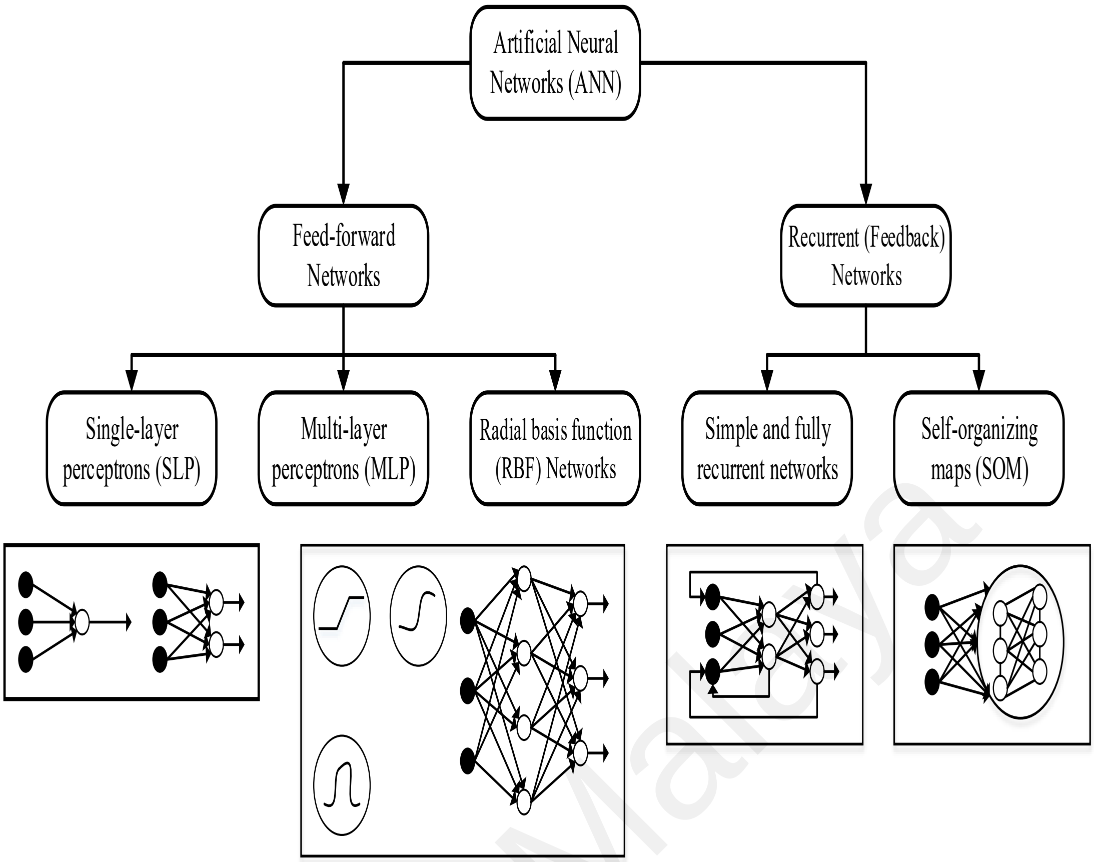

Figure 2.5: Configuration of ANN

2.4.5.1 Artificial Neural Network

Artificial neural networks (ANNs) are artificial networks inspired by nervous system, which are considered as human attempts to understand what goes on in the biological neural networks (in particular the brain). ANNs simulate the learning process provided by nervous system with the hope of capturing the power of biological neural network, collectively performs tasks that even the supercomputers with high-level computational capacity have not been able to process (Landeras, López, Kisi, & Shiri, 2012).

The ANN is analogous to biological neural network, which consists of a highly interconnected network with very simple processors known as neurons. The neurons are linked by weighted connections that communicate to each other by sending the signals from one neuron to other neurons while the strength of weighted connections expresses the importance of each neuron input. Every neuron is associated with a transfer function, which describes how the weighted sum of a neuron input signal is converted to an output signal.

The main characteristic feature of ANNs is that these artificial networks acquire accumulated experience within learning process and respond to new conditions based on the knowledge gained within the learning process. In ANNs, the learning process is provided through repeated adjustments of numerical weights, thus the weighted connections are considered as basic means of long-term memory in these networks.

Various topologies of neural networks (NNs) have evolved based on the training paradigm neuron arrangement and neuron connections. Among the various types of NNs, multilayer perceptron (MLP) and the radial basis function (RBF) network have been the most useful types of NN in different applications. The main differences between these two types of NN reside in the activation functions of the hidden layer. The activation function belongs to the Gaussian family in RBF network whereas, the linear, logistic sigmoid and bipolar sigmoid (hyperbolic tangent) activation functions are used in MLP (Kankal, Akpınar, Kömürcü, & Özşahin, 2011).

Generally, there is a trade-off between higher robustness provides by RBF network and higher accuracy gains by MLP. Due to the non-linear nature of RBF network, it brings much more robustness to adversarial noise. Instead, MLP is an acronym for deep learning in NNs as it has multiple hidden layers to provide higher accuracy.

MLP as a feedforward NN has the generalizing ability to approximate essentially any function with high degree of accuracy, so it is considered as universal approximators.

As shown in Figure 2.6, the MLP architecture is composed of three layers:

i- Input layer, where the data are introduced into the NN (source neurons)

ii- Hidden layer(s) where the data are processed (computational neurons)

iii- Output layer where the results of given inputs are obtained.

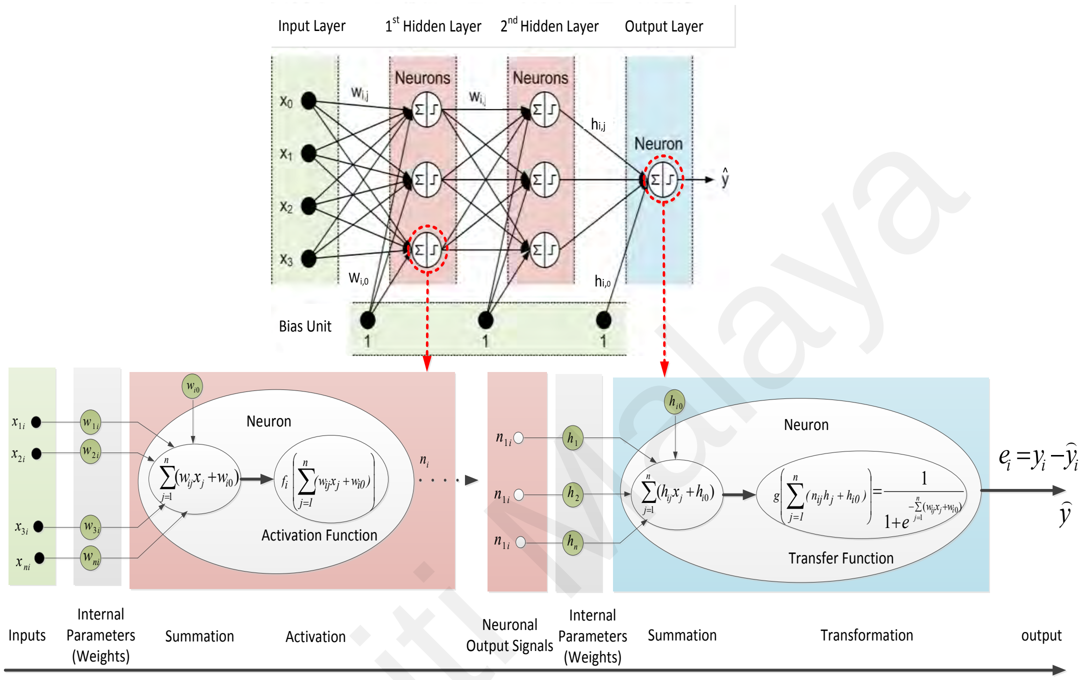

Figure 2.6: The MLP architecture

Each of these layers has several processing units and each unit is fully interconnected with weighted connections to units in the subsequent layer. Each layer contains a number of nodes. Every input is multiplied by each of the nodes using its interconnection weight. The output of each neuron is obtained by passing the sum of the product through an activation function while the bias input (memory) is connected to every neuron to stabilize the origin of activation function for better learning.

MLP networks are usually applied to perform supervised learning tasks, which involve iterative training methods to adjust the connection weights within the network. Generally, several passes are required to attain a desired level of estimation accuracy. Adjustment of the correction weights is carried out using the standard error backpropagation algorithm, which minimizes the total error with the gradient decent method (Raza & Khosravi, 2015).

2.4.5.2 Support Vector Regression

Kernel-based techniques e.g., Kernel principal component analysis (PCA), Kernel Fisher discriminant analysis (KFD), Bayes point machines, Gaussian processes and SVMs (support vector machines) represent a major development in machine learning algorithms. Kernel-based techniques map the data into a higher dimensional feature spaces in the hope that in the higher dimensional spaces the data are linearly separated or have better structure. SVMs as an extension to nonlinear model of the generalized portrait algorithm are the best-known member of Kernel-based techniques, which is able to either classify the input data or capture complex relationships in the input data. SVM that deal with function approximation and forecasting is just termed as SVR and SVM that deal with classification problems is just termed as support vector classification (SVC). With only a few minor modification, SVC is converted to SVR. So, SVM can be promoted to form SVR as a powerful function approximation technique based on statistical learning theory (Bian, Han, Du, Jaubert, & Li, 2016).

2.4.6 Hybrid models

Most of the models and electricity price forecasting techniques in the literature are regarded as hybrid solutions in which two or even more approaches or techniques, in the aforementioned listed previously, are combined. For instance, Neural Networks and Box Jenkins models can be combined and the created model is named AleaModel (AleaSoft).

2.5 Fundamental price drivers and input variables

In this section, two important key parameters should be considered in forecasting process of any competitive market as follows:

2.5.1 Seasonality

A proper treatment of seasonality is one of the most determinative factors in electricity spot price forecasting in the modeling process (Keles et al., 2016). It is also noteworthy to mention that there are three levels of seasonality for the electricity price. These levels are the daily, weekly and to some degree, the annual. The annual or long-term seasonality is commonly negligible in short-term forecasting. The most important factors are the daily and weekly patterns. In this forecasting, for holidays, there is a separate treatment. However, this approach might not be a suitable solution. Disintegrating a series of electricity prices into stochastic and a long-terms seasonal component, modeling them separately and combining the forecasts might result in a high precision gain in comparison to an approach in which a calibration is performed for a particular model to the prices themselves. This approach has recently become well-known as Nowotarski and Weron (Nowotarski & Weron, 2016).

Obviously, the short-term patterns or factors become irrelevant in the mid-term forecasting and most EPF models only consider the average daily prices but the long-term trend-cycle component has an extreme impact on the mid-term forecasting. Any misspecification can be translated into bias, which might make the model draw an inaccurate or even completely wrong estimate in the mean reversion level or the intensity and severity of the price spike. This misspecification will eventually result in an inaccurate risk management strategy. In the long-term forecasting, as the last category, the main concentration is on a range of years, and daily, weekly and even annual seasonality might not be taken into consideration since long-term trends are the predominant factors in this category. Although many technical contributions have been done for the three aforementioned categories, ample treatment, both in-sample and out-of-sample, of the seasonality has not exhaustively been investigated in the literature so far (Weron & Zator, 2015).

2.5.2 Variable selection

Correct and applicable choice of explanatory variables is another vital issue in electricity price forecasting (Ziel et al., 2015). A large set of fundamental drivers determines the current spot price. These drivers include system loads, fuel costs, the reserve margin (i.e., available generation minus/over predicted demand), weather variables and data about ahead-planned maintenance and forced outages. Historical data from the electricity prices also impact the current spot price. Pure price models are often adopted for EPF but most authors, for the day-ahead forecasting, tend to select a combination these fundamental drivers, which are extracted from heuristics and experience of the forecaster (Amjady & Hemmati, 2006). An automated selection or shrinkage procedure has rarely been adopted in EPS, especially for a large set of initial explanatory variables (Uniejewski, Nowotarski, & Weron, 2016). On the other hand, viable tools are available by means of machine learning that in the literature can be classified into two classes (James, Witten, Hastie, & Tibshirani, 2013). The first class is the feature or subset selection, which a subset of predictors believed to be influential, is initially identified and a fitting model on the recued set of variable is determined accordingly. The second class is called shrinkage, which an algorithm is employed to shrink the estimated coefficients around zero, leading to a great reduction in their variance. This class is also known as regularization. Some of the coefficients might be shrunk to zero itself depending of the type of shrinkage. As such, some shrinkage methods such as the lasso-de facto performs variable selection.

It seems that more investigation and studies are required for selecting the most effective input variables. These variables can be the past electricity prices and the past and predicted values of the fundamental drivers. It also plays a great role in those companies at the corporate level when it comes to decision-making mechanisms. In addition, price forecasting is gaining more attention among various market players in the power to adjust their bids in the day-head electricity markets and improve their profits. To be more precise, electricity price is intrinsically unstable but not unpredictable and it is plausible to find the patterns referring back on the historical data and forecast. Therefore, a precise method for price forecasting is a determinative factor for market players since it allows them to maximize their profits by aiding them in selecting correct and wise bidding strategies. Regardless of several proposed models that have been developed over the past decades, they all can generally be categorized into two main types, which are commonly adopted in the electricity price forecasting and they have been named as time series models and simulation based models. Between these two models, time series models are widely adopted for day-ahead forecasting.

In competitive markets, electricity price is a function of electricity demand. However, many factors will effect on short-term electricity price forecasting such as temperature and humidity. These factors will have direct effect on electricity demand. Therefore, with the inherent correlation between electricity price and demand, in this work only electricity price and demand in different times are selected for EPF.

2.6 Progress on electricity price forecasting via artificial intelligence methods

This section provides the recent development on electricity price forecasting based on artificial intelligence methods. Also, summary of the whole methods based on the country, year and accuracy is collected to ensure the coherence and existence of this research from past up to now.

Forecasting techniques can be classified into three categories in accordance to the forecasting framework i.e. statistical models, time series methods and Artificial Intelligence (AI) based approaches. Among different techniques, AI-based electricity price forecasting approaches have gained significant traction in recent years as these approaches offer a remarkable advantage of assuring a certain level of estimation accuracy compared to high fluctuation of independent and dependent variables in the statistical model (Hernandez et al., 2014). Concise reviews of recent methods and techniques on price forecasting based on AI techniques is accessible in (Aggarwal, Saini, & Kumar, 2009; Cerjan, Krželj, Vidak, & Delimar, 2013; Weron, 2014). An extended review including most of the published techniques, such as stochastic models, artificial intelligence (AI) models and the regression models, can be found in (Panapakidis & Dagoumas, 2016).

ARIMA and GARCH are two popular approaches of time series models. They are highly capable and can be a role model for model comparison. In addition, they have a capability to be merged with other models to create a hybrid mode. Applying exhaustive mathematical formulas is the main reason why these approaches have become considerably prevalent. Historical values of the quantity are being used in time series models, which the assumption is that that the quantity progression follows a specific exemplary. (Conejo, Contreras, Espinola, & Plazas, 2005). Moreover, in this model, Pattern's extension is used to pre-define a future time period for as a prediction step. In (Cuaresma, Hlouskova, Kossmeier, & Obersteiner, 2004), AR, ARMA and ARIMA as time series models are compared. Many other sub-models have been introduced (i.e. AR (1) with jumps, AR (1) in logs with jumps, AR (1) with time variant average and rests) and applied in the market of LPX. In (Diongue, Guegan, & Vignal, 2009), a study of comparative simulation has been carried out between a SARIMA-GARCH and a k-factor GIGARCH process. The test study has been carried out on a month EEX market's data and lagged price values were only included models in the study.

Some of the most common ANN-based models in the recent literature are included but not limited to specifically, Multi-Layered Perceptrons (MLPs), Feed Forward Neural Networks (FFNNs), Radial Basis Function Networks (RBFNs), Support Vector Machines (SVMs), Fuzzy Neural Networks (FNNs), Probabilistic Neural Networks (PNN), Recurrent Neural Networks (RNNs) and Self-Organizing Maps (SOMs). MLPs's training speed, simplicity and effectiveness have made them to be the most common network. In (O Abedinia, Amjady, Shafie-Khah, & Catalão, 2015), only MLPs are adopted to forecast while in (Amjady & Daraeepour, 2009b), they are combined with the other models of time series. It is found from the literature that the ANN has been used in some methods to forecast electricity demand and price. Besides, some data mining approach to select particular days for training is combined by the ANN (Amjady & Daraeepour, 2009a).

The most outstanding features of the MLP is that it can be applied in hour-ahead time framework (Amjady & Hemmati, 2009) while most other common studies only refer to day-ahead forecast. The MLP's main function in a hybrid model is to boost the foresight, which is yielded from conventional model of time series (i.e. ARIMA). The ARIMA can be adopted to other variables such as inflow and stored energy in storages, and the entire generation of hydro, system load and price values (Amjady & Keynia, 2008), which the foresights are defined as ANN system input. A combination of two models of time series, ARIMA and GARCH, have been introduced in (Amjady & Keynia, 2009b). A comparison of ANN model with other models such as AR and ARIMA has been studied in (Amjady & Keynia, 2009a). RBFNs models are based on FFNNs models, which has been adopted in (S Anbazhagan & Kumarappan, 2012;

Swaminathan Anbazhagan & Kumarappan, 2013). A hidden output layer is used in an RBF network. In fact, for the activation function of the hidden layer, the RBF is responsible. Such a model is capable to simulate the complex relation holding the data. Rapid adaptability to plausible differences of the relations is one of its features (Lago, De Ridder, & De Schutter, 2018).

A mapping of non-linear original data into high dimensional space is provided by SVMs (S Anbazhagan & Kumarappan, 2014), which a function of linearity is used to define the boundaries of the new space. Unlike MLPs, which is only able to conduct in their objective functions local minima, a solution to a problem can be provided by SVMs. This feature has been adopted in load forecasting area in (Hahn, Meyer-Nieberg, & Pickl, 2009). In (Shiri, Afshar, Rahimi-Kian, & Maham, 2015), in NYISO, between SVM and ANN, a comparison has been studied. In addition, an SVM model is adopted to estimate the prediction intervals in (Shiri et al., 2015), which measures the uncertainty associated to forecasts by approximating the targeted quantities ranges.

The Fuzzy Logic (FL) and ANNs synergetic operation is another model in price forecasting. Here, the literature can be even further categorized. It can be categorized in dual categories. Since, the study that adopts ANN, FL in the system is similar to neuro-fuzzy systems (i.e. ANFIS). Therefore, the studies where ANN together with FN are scattered. Still, they are mingled as a forecaster having two-part (Bigdeli, Afshar, & Amjady, 2009). In the latter approach, a relationship of linguistic description among the input data is used to accomplish the forecasting. The fuzzy logic systems have also been used to forecast electricity price. However, this method of forecasting is based on predefined rules. It is also unable to learn and adapt itself to new conditions. This shortcoming has been moderately addressed in (H. M. I. Pousinho, Mendes, & Catalão, 2012) by combining a fuzzy system and ANN model (Neuro-Fuzzy). The most important feature of a neuro-fuzzy model is its adaptability interface system (ANFIS). This adaptability has made the ANFIS model to be a universal estimator and enabled them to forecast for short, medium and long-term electricity price. In (J. P. d. S. Catalão, Pousinho, & Mendes, 2011), the accuracy of the forecasting and the non-linearity of the forecasted electricity price have been improved by integrating different techniques and forming a hybrid model.

In (Li-xiao, Zheng-fang, Chuan-zhi, Wen-li, & Sheng-shan, 2014), another technique has been presented, which ARIMA procedure and wavelet transform (WT) are combined adopting the power market prediction. To be more precise, in the technique, first historical data is separated by wavelet transform. Then, the ARIMA technique is applied and the inverse wavelet transform is employed respectively to obtain the final prediction outcomes. In another technical contribution (Hamzaçebi, 2008), the price has been forecasted by the ANN technique, which the history and approximated future parameters are used to find the price and quantities. This contribution has also introduced the three-layer back propagation (BP) neural network (NN) adopting demand and fuel cost in the market as entry data.

A combination of the probability neural network (PrNN) and orthogonal experimental design (OED) has been adopted in (Lin, Gow, & Tsai, 2010) to forecast the power price. The PrNN and OED methods have been used for classification and locating the best variable respectively, which eventually increase the forecast precision. The support vector machine (SVM) and the projected assessment of system adequacy (PASA) has been used in (D.-X. Niu, Wang, & Li, 2006) for prediction of price according to the price and load history, and the entry data respectively. The work has also used a regional data of south wales to carry out their evaluations. In (Gong, Che, Wang, & Liang, 2008), the fish swarm algorithm (FSA) has been selected and used as a time series forecast procedure to select the SVM variable. The work used the power price as data entry.

The least square support vector machine (LSSVM) for prediction has been introduced in (Ebrahimian, Barmayoon, Mohammadi, & Ghadimi, 2018) by combining wavelet packet transform (WPT) and feature selection. A probabilistic power price prediction method has been presented in (Wan, Xu, Wang, Dong, & Wong, 2013) and in (Baziar & Kavousi-Fard, 2015), the work presented a combination of SVR and ARIMA method. A combination of three methods as WT, radial basis function neural network (RBFNN) and ARIMA has also been used in (Chang, 2015). In (Amjady & Daraeepour, 2009b), another mixed model has been presented based on interaction of load and price prediction. Another model for price-directed demand response has been introduced in (Alamaniotis, Gatsis, & Tsoukalas, 2018), which Virtual Budget (VB) approach is developed couples' price and load prediction, and lets automated morphing of a consumer's electricity demand.

Another hybrid model has been proposed in (Wu & Shahidehpour, 2014) to predict load and price. The work applied a hybrid time-series and adaptive wavelet neural network (AWNN) to predict the price. In (Gao et al., 2019), different states of multi-block based forecast engine are applied for both price load forecasting purposes. In this work, mechanism of forecasting is consist of multi-block neural network (NN) and optimized by an intelligence algorithm to increase training time and forecasting capabilities. (Oveis Abedinia & Amjady, 2016; O Abedinia et al., 2015) have also proposed other method to model price and demand forecasting.

In (J. P. d. S. Catalão, S. J. P. S. Mariano, V. Mendes, & L. Ferreira, 2007), ANN, ANFIS, and ARIMA are adopted as the forecasters using filter of kalman to calculate the forecast price for a system of Spanish electricity. These three models have also been adopted in (J. Catalão, S. Mariano, V. Mendes, & L. Ferreira, 2007). Here, an algorithm of ordered weighted average, which is modified, is selected for combining the forecast price generated by every model. The modified approach has been examined on quadruple typical weeks in Spain's system. A system that is hybrid is adopted to forecast the prices of electricity in the markets in Spain in (J. Catalão, Pousinho, & Mendes, 2009), which a mutual information method selects the inputs and feeds into the ANFIS. In (J. P. d. S. Catalão et al., 2011), the authors proposed a hybrid approach for electricity price forecasting, which is based on a combination of ANFIS and evolutionary particle swarm optimization method. The forecasting accuracy improvement of ANFIS mode and the ANFIS optimal parameters mode are resolved by the technique, which is evolutionary particle swarm optimization. The recent developed artificial intelligent technique in the literature that have been implemented for forecasting price of electricity are summarized in Table 2.1.

Table 2.1: Review on most recently developed artificial intelligent methods in the literature which have been applied for electricity price forecasting

	Year	Applied for	Characteristics	MAPE *
ARIMA/SVM (L. Zhang, Luh, & Kasiviswanathan, 2003)	2003	New Wales	Application of the discrete wavelet transform Only historical values of price are used 4 test weeks	8.8%
ARIMA (Zareipour, Cañizares, Bhattacharya, & Thomson, 2006)	2006	Ontario	Hourly Ontario electricity price(HOEP) forecasting in Ontario mainland for 6 test week	17.63%
TF (Zareipour et al., 2006)	2006	Ontario	Hourly Ontario electricity price(HOEP) forecasting in Ontario mainland for 6 test week	16.25%
DR (Razak et al., 2016)	2006	Ontario	Hourly Ontario electricity price(HOEP) forecasting in Ontario mainland for 6 test week	16.53%
IESO (Razak et al., 2016)	2006	Ontario	Hourly Ontario electricity price(HOEP) forecasting in Ontario mainland for 6 test week	39.33%
ARIMA/MLP (Amjady, Daraeepour, & Keynia, 2010)	2007	Ontario; New England; Italy	Only historical values of price and load are used 6 test weeks for Ontario; 11 test months for New England; 11 test months for Italy	4.48%
MLP (Aggarwal, Saini, & Kumar, 2008)	2008	Ontario	Application of the discrete wavelet transform 6 test weeks	17.51%
FCM/MLP (Vahidinasab, Jadid, & Kazemi, 2008)	2008	PJM	The clustering method is utilized Only historical values of price and load are used 4 test weeks	4.5%
Heuristic (Aggarwal et al., 2008)	2008	Ontario	Hourly Ontario electricity price(HOEP) forecasting in Ontario mainland for 6 test week	24.89%
NN (Aggarwal et al., 2008)	2008	Ontario	Hourly Ontario electricity price(HOEP) forecasting in Ontario mainland for 6 test week	18.39%

Wavelet+ NN (Aggarwal et al., 2008)	2008	Ontario	Hourly Ontario electricity price(HOEP) forecasting in Ontario mainland for 6 test week	17.51%
MLP(Swaminathan Anbazhagan & Kumarappan, 2013)	2009	New York; Spain	Only historical values of price and load are used 2 test months for NYISO; 4 test weeks for Spain	4.95%
MLP(Amjady & Daraeepour, 2009a)	2009	PJM; Spain; Ontario	Only historical values of price and load are used 4 test weeks for PJM; 4 test weeks for Spain; 3 test months for Ontario	4.55%
Modified Relief + Hybrid NN (HNN) (Amjady et al., 2010)	2010	Ontario	Hourly Ontario electricity price(HOEP) forecasting in Ontario mainland for 6 test week	9.23%
Original Relief + HNN (Amjady et al., 2010)	2010	Ontario	Hourly Ontario electricity price(HOEP) forecasting in Ontario mainland for 6 test week	9.92%
Correlation analysis +HNN (Amjady et al., 2010)	2010	Ontario	Hourly Ontario electricity price(HOEP) forecasting in Ontario mainland for 6 test week	11.16%
Numerical sensitivity analysis +HNN (Amjady et al., 2010)	2010	Ontario	Hourly Ontario electricity price(HOEP) forecasting in Ontario mainland for 6 test week	11.97%
ARIMA/MLP (Amjady & Keynia, 2010)	2010	PJM; Spain	Only historical values of price and load are used 19 test days for PJM; 4 test weeks for Spain	5.14%
MLP (Amjady & Keynia, 2009b)	2010	PJM	17 test days	5.56%
SOM/SVM (Jin, Pok, Paik, & Ryu, 2015)	2010	PJM	The clustering method is utilized for 1 test month	9.96%
ANFIS (H. Pousinho, Mendes, & Catalão, 2010)	2011	Spain	Only historical values of price are used Application of the discrete wavelet transform	6.53%
FL/MLP (Chogumaira & Hiyama, 2011)	2011	Australian New-South Wales	The clustering method is utilized Only historical values of price are used	7.04%
RBF (dos Santos Coelho & Santos, 2011)	2011	Spain	Only historical values of price are used 24 test days	5.5%
ARIMA/RBF (Shafie-Khah, Moghaddam, & Sheikh-El-Eslami, 2011)	2011	Spain	Application of the discrete wavelet transform Only historical values of price are used	4.27%
SVM (H Shayeghi & Ghasemi, 2013)	2013	Iran; Ontario; Spain	Application of the discrete wavelet transform Only historical values of price and load are used 4 test weeks for Iran; 6 test weeks for Ontario; 4 test weeks for Spain	4.41%
LSSVM +CGSA (Shrivastava & Panigrahi, 2014)	2013	Ontario	Hourly Ontario electricity price(HOEP) forecasting in Ontario mainland for 6 test week	15%
Recurrent NN + excitable dynamics (Sharma & Srinivasan, 2013)	2013	Ontario	Hourly Ontario electricity price(HOEP) forecasting in Ontario mainland for 3 test week	12.47%
FTD NN+Wrapper model (Maca & Lhotská, 2013)	2013	Ontario	Hourly Ontario electricity price(HOEP) forecasting in Ontario mainland for 6 test week	12.74%
ANFIS-PSO (H. Pousinho et al., 2010)	2014	PJM; Spain	Application of the discrete wavelet transform Only historical values of price are used 2 test weeks for PJM; 4 test weeks for Spain	3.73%
WELM (Shrivastava & Panigrahi, 2014)	2014	Ontario	Hourly Ontario electricity price(HOEP) forecasting in Ontario mainland for 6 test week	7.88%
WELM with ensembles (Shrivastava & Panigrahi, 2014)	2014	Ontario	Hourly Ontario electricity price(HOEP) forecasting in Ontario mainland for 6 test week	6.66%
MLP (Keynia, 2012)	2015	PJM; Spain	4 test weeks and the last week of all months of 1 year for PJM 4 test weeks for Spain	3.97%
SOM/MLP (D. Niu, Liu, & Wu, 2010)	2015	Spain; New York;	The clustering method is utilized Only historical values of price and load are used	3.22%

		ANEM		
SVM (Hossein Shayeghi, Ghasemi, Moradzadeh, & Nooshyar, 2015)	2015	New York; PJM; New South Wales	Application of the discrete wavelet transform Only historical values of price and load are used 4 test weeks for New York; 1 test week for PJM; 4 test weeks for New South Wales	3.26%
SVM+GA (Razak et al., 2016)	2016	Ontario	Hourly Ontario electricity price (HOEP) forecasting in Ontario mainland for 6 test week (HOEP and Demand are considered as selected features for forecasting analysis.	9.22%
Environmentally adapted generalized neuron (Singh, Mohanty, & Shukla, 2017)	2017	NSW	One week test system of Australian electricity market based on historical price and demand data	2.28%
* Mean absolute percentage errors (MAPEs) are on testing period of each model				

2.7 Summary

An investigation over price forecasting methodologies that have been done in the previous works in the deregulated environment is presented in this chapter. Due to the rapid change in the structure of power markets, forecasting future prices with high level of accuracy is an inevitable routine for market players in order to maximize their profits.

Owing to the dependency of electricity demand and electricity price on various parameters, considering most influential input variables in electricity price is a daunting task. To address the difficulties involved in EPF different techniques have been proposed to achieve a robust model with high accuracy. In this chapter, the statistical methods, artificial intelligence-based approaches and hybrid methods have been employed for short-term electricity price forecasting is explained in details. Although many approaches have been applied to forecast electricity market price in the literature, these approaches have drawbacks. For instance, the major deficiencies of data driven approaches are that there are so many control parameters and they are very sensitive that makes initialization of these parameters value to be very difficult. Although different machine learning methods have been applied for electricity price forecasting, new methodology is still required to provide more accurate electricity price forecast.

Furthermore, in most of the aforementioned literature, long term price forecasting has been considered. However, the pattern of electricity demand depends on the seasons, hence, short term price forecasting will be more effective in deregulated electricity market for real time decision making. Another important factor that influences the efficiency and accuracy of the forecasting is the feature selection. Proper selection of features helps to enhance the efficiency and accuracy of the forecasting. However, an observation has been made that it is very difficult to select the robust feature selection technique for the prediction of electricity price considering non-linearity of the price signal.

CHAPTER 3: RESEARCH METHODOLOGY

3.1 Introduction

This chapter presents the methodology of the forecasting algorithms employed for short-term electricity price forecasting of Ontario and Queensland competitive electricity markets. A significant portion of the chapter is devoted to explaining the conceptual framework of ANFIS, the technique used for improve the accuracy of the electricity price forecasting and the most robust and efficient metaheuristic optimization method, namely backtracking search algorithm (BSA). Then, multi-objective backtracking search algorithm (MOBSA) and its procedure is developed for selecting the best subset of features for short-term electricity price forecasting process after filtering a lot of features by mutual information (MI). The performance of the forecasts is evaluated through a number of measures (indices) that attempt to quantify the applicability of the prediction to its eventual uses and give a full picture of the capabilities of the forecasting algorithm. This chapter also provides a brief summary of alternative and exploratory techniques on feature selection.

3.2 Development of price Forecasting method and enhanced feature selection

The main focus of this research is to improve the forecasting accuracy of the existing artificial intelligence techniques for short-term electricity price forecasting. Therefore, to achieve this goal, the whole methodology is divided into two parts. Firstly, selecting the best features (prices and demands in different time interval) that have the most influential effect on short-term electricity price forecasting through the efficient feature selection method. Secondly, finding robust forecasting method for electricity price in high volatile and competitive markets. This work proposes a short-term day ahead prediction of electricity price forecasting method. At the first step, to reduce redundant features, MI is applied. Then, MOBBSA will search within different combinations of input variables before being evaluated by ANFIS to select the best feature subsets of input variables with maximum relevancy and minimum redundancy. In the second step, the best feature subsets will be fed into the ANFIS again to forecast the price. Here, the backtracking search algorithm (BSA) is utilized to exploit the solution structure and explore its appropriate weighting factors of the learning process of ANFIS in order to increase the forecast accuracy. The forecasting model development is shown in Figure 3.1.

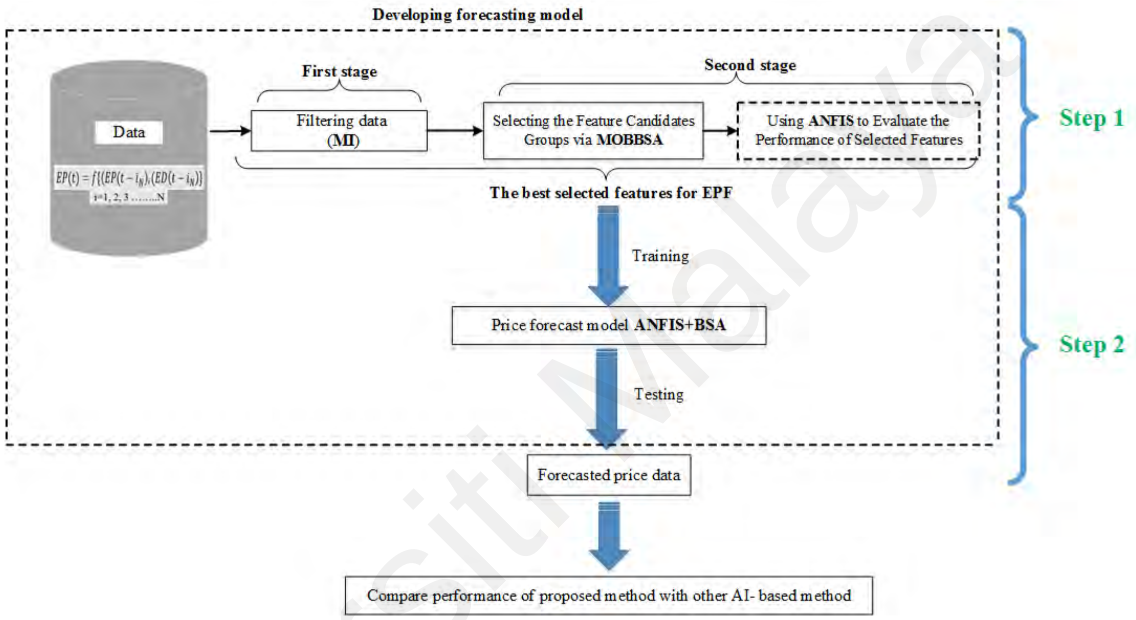

Figure 3.1 Structure of development of electricity price forecasting

3.3 Features (price and demand) selection

In the context of statistics and machine learning, feature (variables, predictors) selection, also known as attribute selection, variable selection or variable subset selection is a method for selecting a subset of relevant features in model construction. In this work, features are prices and demands in different time. The objectives of using feature selection techniques are three-fold:

- Improving the prediction performance of the predictors by reducing overfitting (formally, reduction of variance).

- Providing faster and more cost-effective process to construct the model (facilitate learning process).
- Providing a simplified model that makes it easier to interpret (improving the generalization ability).

The main objective of a feature selection technique is to evaluate the relevancy and redundancy of the input features for selecting the best subset of features representing the most important information of the original feature set. A large number of predictors could result in inferior performance of the extracted models due to the course of dimensionality principle. Removing either redundant or irrelevant features from a data set contains many features without incurring much loss of predictive accuracy is the central premise for using a feature selection technique.

A search strategy along with an evaluation metric is applied in a feature selection algorithm to search candidate subsets of feature and score the performance of these candidates respectively. Testing all possible feature subsets to find one that minimizes the error rate is the simplest feature selection algorithm that provides an exhaustive search of the feature space but it is computationally intractable (Unler & Murat, 2010). Thus, the combination of a search strategy to explore the space of all possible combination of features along with an evaluation measure to assess the quality of features heavily influences on effectiveness of a feature selection algorithm.

According to how a search technique is combined with a learning algorithm (evaluation metric) for construction of the models forms three classes of feature selection methods, namely wrappers, filters, and embedded (Renani, Elias, & Rahim, 2016).

Wrapper (search guided by accuracy) methods use a predictive model to score feature subsets. In wrapper methods, every candidate feature subset is used to train a model that is tested on a holdout set. The score for every candidate subset is provided by evaluating the error rate of the model on testing set. As wrapper methods train a new predictive model for each candidate subsets, they often provide the best performing feature set for that particular type of model at the expense of computationally intensive tasks.

In filter (information gain) methods, instead of the error rate to score a candidate subset of feature, a proxy measure is used. The proxy measures such as the pointwise mutual information, Pearson product-moment correlation coefficient and mutual information are chosen to provide fast computation for capturing the effectiveness of the feature set.

The wrapper methods are higher computationally intensive than filter methods but wrapper methods provide a subset of feature that its performance is evaluated by specific type of learning algorithm. Due to the lack of learning algorithm in filters, a feature set from the filter methods is usually more general and giving lower prediction performance than the set from wrapper methods. The filter methods are widely used to expose the relationships between the variables and provide a feature ranking rather than an explicit best feature subset. Thus, a filter can be used as a preprocessing step for a wrapper to form a hybrid feature selection method. In this hybrid method, filter works as dimensionality reduction method, allowing a wrapper method to be used for appropriate selection of the most relevant features on larger data sets.

Embedded methods learn which subset of features has best contribution to the accuracy of the model while the model is constructed. In embedded methods, the feature selection part and training process cannot be distinguished because selection of the features and model construction procedures are performed simultaneously. Although the embedded methods are less computationally intensive than wrapper methods, the main drawback of these methods is that the selected features are sensitive to the structure of the underlying model. Hence, embedded methods are usually specific to their learning algorithms. Different types of embedded method are classification trees, random forests and regularization approaches. Regularization approaches, also known as penalization approaches, are the most common type of embedded feature selection methods. The penalization approaches add additional constraints into the model construction process, which bias the model toward simplicity by penalizing the model for higher complexity.

3.3.1 First stage of filtering data using mutual information (MI)

Filter type methods are essentially data pre-processing or data filtering methods. Features (prices and demands) are selected based on intrinsic characteristics, which determine their relevance to the target. In filters, the characteristics of the feature selection (EP (t) and ED (t)) are uncorrelated to those of the learning methods; therefore, they have better generalization property. Therefore, this section explains one of the many attribute feature selections, namely mutual information that tries to recover all of the relevant features.

3.3.1.1 Mutual information technique

MI technique has been widely implemented in electricity market price forecasting (Amjady & Hemmati, 2006). However, this technique is facing difficulties due to the lagged values of the candidate inputs comprising of price, load demand and other variables provided by the electricity market. Thus, the individual probability distribution and the joint probability distribution of the candidate inputs are difficult to be obtained (Amjady & Daraeepour, 2009a). Besides, it is noted that the electricity price is a time variant signal. Therefore, long history of the candidate inputs is not relevant to be used as the market conditions are evolving every time. As such, it can mislead or give inaccurate price forecast process due to the lack of information values (Amjady & Daraeepour, 2009a).

The concept of mutual information is to observe the mutual relationship between random features price and demand in different time interval which are denoted by X and Y respectively. It measures the amount of information that is obtained by one variable with respect to the other random variable (Cover & Thomas, 2006). In other words, the mutual information is zero if variable X does not have any information related to variable Y and vice versa. As such, these 2 random variables are independent. High mutual information is obtained if variable X is a deterministic function of variable Y and variable Y as a deterministic function to variable X. It means that all the information conveyed by one variable will be shared by the other variable. Thus, it reduces the uncertainty about one random variable that knows the other random variable. Therefore, the mutual information will be the same as the uncertainty obtained by either one of the variables.

The concept of mutual information is intricately linked by the entropy of the random variable, the mutual information between random variables X and Y; where in this work X represents electricity price and Y represent electricity demand respectively. The mutual information, $MI(X, Y)$ with the joint probability distribution of $P_{XY}(X, Y)$ can be defined as

$$MI(X,Y) = -\sum_{i=1}^{n}\sum_{j=1}^{m} P(X_i, Y_j) log_2(P(X_i, Y_j)) \tag{3.1}$$

Besides the entropy, there is a conditional entropy, which measures the average uncertainty of the first random variable after the second random variable is observed. For example, the conditional entropy CE for random variable X after observing the random variable Y with the conditional probability distribution of $P(X|Y)$ is expressed as:

$$CE(X|Y) = \sum_{j=1}^{m} P(Y_j) \left[- \sum_{i=1}^{n} P(X_i|Y_j) log_2(P(X_i|Y_j)) \right] \tag{3.2}$$

$$CE(X|Y) = - \sum_{i=1}^{n} \sum_{j=1}^{m} P(X_i, Y_j) log_2(P(X_i|Y_j)) \tag{3.3}$$

$$P(X_i|Y_j) = \frac{P(X_i, Y_j)}{P(Y_j)} \tag{3.4}$$

Figure 3.2 shows a graphical representation of the relationship between the MI and CE. If MI is large enough, it indicates that X and Y are closely related and dependent on each other. If X and Y are independent variables, then MI(X, Y) will be equal to zero and thus P(X,Y) = P(X)P(Y).

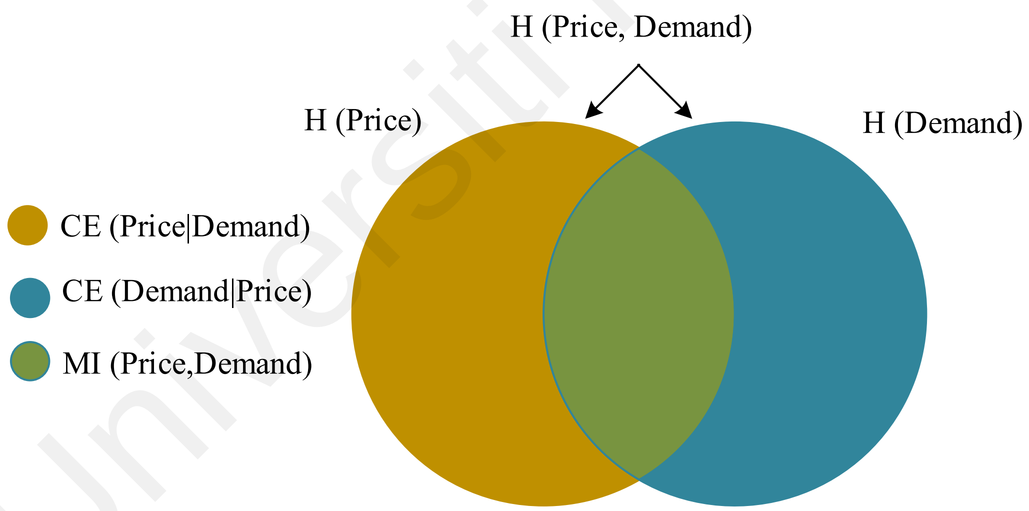

Figure 3.2: Graphical representation between mutual information and conditional entropy

The relationship between MI and CE can be referred as the chain-rule and it can be calculated by

$$MI(X,Y) = H(Y) - CE(Y|X) = H(X) - CE(Y|X) \tag{3.5}$$

$$H(X,Y) = H(X) + CE(Y|X) = H(Y) + CE(X|Y) \tag{3.6}$$

In the first stage of hybrid feature selection, relevancy threshold of TH has been chosen for filtering the redundant features, and after filtering, reminding relevant features are selected for next stage. Input features having lower value of mutual information than the chosen threshold TH presenting less significant influence on the output are eliminated. The variables with mutual information above the given TH are considered as relevant features and form a subset $XS \subset X$which remains for the next stage. Note that, although the threshold TH is provided by a user, a high value of TH results in missing lots of information. Whereas, a low value of TH may include too many features, either relevant or irrelevant leading to a large computational burden. After, computing the mutual information between price and demand based on equation (3.1), the prices and demands are selected to the small group from vast pool for forecasting process.

3.3.2 Second stage of selected features using (MOBBSA+BSA) hybrid technique

In this stage, feature selection is wrapped around a learning method; the usefulness of a feature is directly judged by the estimated accuracy of the learning method. These methods are computationally useful for data with a large number of features, which is the case of electricity price forecast. In the second stage, the prices and demands from filtering are chosen by the proposed hybrid technique or final process. Therefore, this section explains the procedure of development of feature selection in second stage for forecasting price of Ontario and Queensland deregulated markets.

3.3.2.1 Multi-objective backtracking search algorithm

There are two methodologies to solve the EPF problems, which are based on the weighted sum method and non-dominated approach. In this stage, the non-dominated approach is employed for dealing with the EPF problems. The objective function is written as a vector of both objectives and neither is inferior to the other, i.e., minimizing the number of features and maximizing the accuracy simultaneously. This method of optimization is called non-dominated approach (NDA), detailed in this section. The NDA is based on the concept of Pareto front optimal set needs to be explained.

Metaheuristic optimization methods are widely applied in searching for optimal solutions in large-scale problems of engineering and computer science. They work by guiding the searching in a solution space to find the optimal. An evolutionary algorithm (EA) as a subset of the metaheuristic methods emulates the evolutionary process in nature to solve the optimization problems. Such techniques include mutation, crossover, and selection to be applied to each individual of a population. Among these algorithms, BSA is a new evolutionary method developed for forecasting approach to overcome the complexities of highly nonlinear optimization problems. In this research, minimizing the number of features and maximizing the forecasting accuracy are considered as a multi objective.

3.3.2.1.1 Backtracking Search Algorithm (BSA)

BSA is one of the most recently proposed evolutionary algorithms with a simple structure but high effectiveness in solving multimodal functions that enables it to easily adapt to different numerical optimization problems. This merit is provided by balancing exploitation of better results and exploration of the problem's search space through the use of a single control parameter and two advanced crossover and mutation operators. Its strategy contains two advanced crossover and mutation operators for generating a trial population. These operators are unique and quite different from the structures of the crossover and mutation strategies defined in other evolutionary methods (e.g. GA and DE). BSA strategies for generating trial populations and controlling the search-space boundaries and adapting the amplitude of the search-direction matrix provide effective exploration and exploitation capabilities.

BSA has been developed to overcome the drawbacks of metaheuristic methods (e.g. too many control parameters and over sensitivity to initial value of these parameters, premature convergence and time-consuming computation). This is due it has only one control parameter and is not overly sensitive to the initial parameter value. It also possesses a memory that allows it to take advantage of the experience gained from past generations when generating a trial population. In particular, it stores a randomly chosen population from previous generation in its memory for use in generating the search-direction matrix.

Statistical analysis in (Civicioglu, 2013) confirms that BSA is a promising optimization method for solving high multimodal optimization benchmarks over different well-known evolutionary methods. BSA comprises of six stages: initialization, selection-I, mutation, crossover, boundary control and selection-II as presented in Table. 3.1.

Table 3.1: General structure of BSA

1. Initialization

Repeat

    2. Selection I
    Generation of Trial-Population
        3. Mutation
        4. Crossover
    end
    5. Selection II

Until stopping conditions are met;

- **Initialization**

This process of BSA initially scatters the population members in the solution space. The initial value of i th individual in the solution space is defined by using Eq. (3.7).

$$P_{i,j;g=0} \sim U(low_j, up_j) \ \ , y_i = f(P_i) \\ for \\ i = \{1,2,3,\ldots,nPop\}\ ,\ j = \{1,2,3,\ldots,nVar\} \tag{3.7}$$

where, *nPop* is population size, *nVar* is number of respective optimization variable, *U* represents the uniform distribution function. low_j and up_j are lower and upper search space limits of j th variable, y_i is productivity of i th individual and *g* is generation number.

- **Selection-I**

In this stage of BSA algorithm, a historical population (*oldP*), which is utilized to determine the search-direction matrix, is initialized. The historical population is initialized according to Eq. (3.8).

$$oldP_{i,j} \sim U(low_j, up_j) \tag{3.8}$$

Then, the historical population is redefined at every iteration through the 'if-then' decision rule (by comparing two random numbers *a* and *b*) according to Eq. (3.9). Subsequently, in this stage m, the population (*P*) pursues the historical population for a period of time (until it is changed), which provides a memorization process to facilitate the exploration of search space in this algorithm by

$$if\ a < b \Big|_{a,b \sim U(0,1)} then\ oldP := P \tag{3.9}$$

where *P*= is the update operation and *a* and *b* are randomly generated numbers.

Finally, the hierarchical sequencing between the individuals of historical population is permuted through a random shuffling function. The obtained *oldP* is used to determine the search-direction matrix in each generation according to Eq. (3.10),

$$oldP := permuting(oldP) \tag{3.10}$$

where *permuting (oldP)* is a random shuffling function. As a permuting function, it randomly changes the order of the individuals in historical population.

- **Mutation**

In the mutation step, an initial form of trial population (Mutant) is generated through the Eq. (3.11) while the Wiener process (*F*) is applied in mutation process to controls the amplitude of the search-direction matrix (*oldP* - *P*),

$$\begin{aligned} Mutant &= P + F.(oldP - P) \\ F &= 3.rndn \Big|_{rndn \sim N(0,1)} \end{aligned} \tag{3.11}$$

where *N* is the standard normal distribution.

- **Crossover**

A binary integer-valued matrix (*map*) that indicates the active individuals of the trial population is determined for crossover process according to Eq. (3.12). In BSA, active individuals stand for the individuals that participate in a crossover at any time and only they can discover new solutions. The control parameter of BSA determines the probability of corporation between individuals in the crossover process by selecting the active individuals of population. Eq. (3.12) elucidates the unique crossover strategy of BSA:

$$
\begin{aligned}
& map_{i,j} = 1, \\
& if\ a < b \Big|_{a,b \sim U(0,1)}\ then \\
& map_{i,u_{(1:\lceil mixrate.\ rnd\ .\ nVar \rceil\ )}} = 0 \Big|_{rnd \sim U(0,1),\ u = permuting(1,2,3,\ldots,nVar\ )} \\
& else \\
& map_{i,randi(nVar)} = 0 \\
& T := Mutant \\
& if\ \ map_{i,j} = 1\ then\ T_{i,j} := P_{i,j} \\
& for \\
& i = \{1,2,3,\ldots,nPop\}\ ,\ j = \{1,2,3,\ldots,nVar\}
\end{aligned}
\tag{3.12}
$$

where *mixrate* is the control parameter of optimization algorithm. BSA has only one control parameter that controls the number of individuals to be engaged in the crossover process. The control parameter of BSA (*mixrate*) varies from 0% to 100% of population size. The experiments with different values of *mixrate* exposed that BSA algorithm is not too sensitive to initial value of its control parameter. *Permuting(1,2,…,nVar)* is a random permutation that refers to the act of randomly rearranging (reordering) all members. *randi(nVar)* is a random selection function that generates a pseudorandom integer between 1 and *nVar*.

- **Boundary control**

At the end of the crossover process, if an individual in generated offspring (*T*) violates the boundary condition of the optimization problem, the related individual is updated according to the boundary control mechanism developed in this stage.

$$
if\ \ (T_{i,j} < low_j)\ or\ (T_{i,j} > up_j)\ then\ T_{i,j} \sim U(low_j, up_j) \tag{3.13}
$$

- **Selection II**

Finally, the fitness value, which is the difference between the actual and forecasted prices of the best individual, is exported as global minimum and its position is considered as global minimizer according to the Eq. (3.14),

$$\begin{aligned}
&if\ \ f(T_i) < y_i\ then\ ,\ y_i := f(T_i),\ \ P_i := T_i \\
&y_g = min(\,f\,(P_g)) \\
&if\ \ y_g < y_{g-1}\ then\ \ global\,minimum := y_g\ ,\ global\,minimizer := P_g\ , \\
&g = g+1 \\
&for \\
&i = \{1,2,3,\ldots,nPop\}\ ,g = \{1,2,3,\ldots,gMax\}
\end{aligned} \tag{3.14}$$

The pseudocode of BSA algorithm is shown in Table 3.2.

Table 3.2 Pseudocode of BSA

```
Data: nPop, nVar, Max cycle, low, up, f, mixrate
Result: Globalminimizer | globalminimum = f (globalminimizer)
Superorganisms: α, β
  //Initialization
globalminimum g=0 ≈ ∞
for i ← 1 to nPop do
    for j ← 1 to nVar do
        P i,j, oldP i,j ~ U (low j, up j)
    end
    y i = f(P i)
end
for g ← 1 to max cycle do
    // Selection-I
    if rnd < rnd then
        oldP := P
    end
    oldP := permuting (pldP)
    Generation of trial population
        //Mutation
            mutant = P + 3.(normrnd(0,1)). (oldP - P)
        //Crossover
            map 1:nPop, 1:nVar =1
            for i←1 to nPop do
                if rnd < rnd
                    then map i, u (1: [mixrate. rnd. D]) = 0 | u =randperm (nVar)
                    else map i,randi(nVar) = 0
                end
            end
        T := mutant
        for i ← 1 to nPop do
            for j ← 1 to nVar do
                if map i,j = 1
                    then T i,j := P i,j
                end
            end
        end
    end
    // Boundary control
    for i ← 1 to nPop do
        for j ← 1 to nVar do

            if (T i,j < low j) ∨ (T i,j > up j) then
                T i,j := rnd.( up j - low j) + low j
            end
        end
    end
    // Selection-II
    for i ← 1 to nPop do
        if f (T i) < y i then P i := T i , y i; := f(T i) end
    end
    y best = min (f (P)) | best ∈ (1, 2, 3,…, nPop)
    if y best < globalminimum then
        globalminimum := y best
        globalminimizer := P best
    end
end
```

3.3.2.1.2 Multi-objective backtracking search algorithm

In the context of multi-objective optimization, instead of unique solution, there is a Pareto optimal set corresponding to the optimal value of each objective. Considering the two solutions from the Pareto optimal set as denoted by $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_N)$ and $\partial = (\partial_1, \ldots, \partial_N)$ and their corresponding objective functions represented by $f(\varepsilon) = (f_1(\varepsilon), \ldots, f_m(\varepsilon))$ and $f(\partial) = (f_1(\partial), \ldots, f_{\mathrm{m}}(\partial))$. The solution ∂ is dominated by the solution ε, denoted by $f(\varepsilon) \prec f(\partial)$, if and only if the conditions described in Eq. (3.15) are satisfied. Therefore, the solution ε is considered as a non-dominated solution. In this work, the solution ∂ and ε are considered as minimize the number of features and maximize the accuracy between the selected prices and demands in preceding time interval.

$$\begin{aligned} &\forall i \in \{1, \ldots, m\} : \ f_j(\varepsilon) \le f_j(\partial) \\ &\exists i \in \{1, \ldots, m\} : \ f_j(\varepsilon) < f_j(\partial) \end{aligned} \tag{3.15}$$

For two objective functions denoted by f_1 and f_2, the Pareto optimal set is illustrated in Figure 3.3. The dominated solutions are represented by the gray circles and the Pareto optimal set of two objectives (f_1 and f_2) are represented by red circuits.

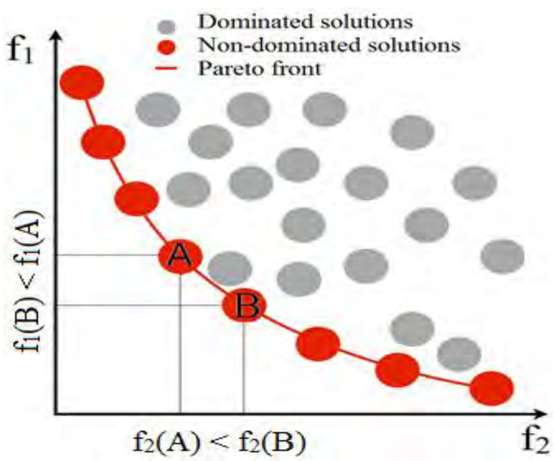

Figure 3.3: The Pareto optimal set for the two objective functions (A and B are two sample from non-dominated solutions)

Multi-objective optimization yields Pareto front as a set of optimal solutions rather than a single optimal. Any solution in the Pareto front is not inferior to another. Improvement to one objective cannot be achieved without sacrificing another. A trade-off between solutions should lead to the best compromise solution. As the objectives of multi-objective optimization are on different scales, a normalization technique is utilized to provide a notionally common scale for all objectives. Then, all the normalized values of each Pareto set are added together and lowest value of Pareto set is selected as the best compromise solution. Therefore, in this section, the features are selected based on minimizing the number of features and maximizing forecasting accuracy simultaneously for final forecasting.

- **External elitist archive**

In the context of evolutionary multi-objective optimization, finding non-dominated solutions within evolutionary process have been retained for different methods. Hence, this sorting method in evolutionary multi-objective optimization algorithms has been enhanced over the last years. Initially, the non-dominated sorting strategy was proposed as an efficient selection strategy in multi-objective optimization (Holland & Goldberg, 1989). Later, Deb et al. (Srinivas & Deb, 1994) applied non-dominated sorting strategy in genetic algorithm and proposed non-dominated sorting genetic algorithm (NSGA) to solve the multi-objective optimization problems. Afterwards, Deb et al. (Deb, Pratap, Agarwal, & Meyarivan, 2002) developed a more efficient non-dominated sorting strategy as an enhanced variant of NSGA known as fast non-dominated sort, which is called NSGA-II.

Necessity to specify the value of sharing parameter (σ_{share}), high computational complexity of non-dominated sorting and lack of elitism are the main criticisms of NSGA approach, which have been solved in NSGA-II. The two novel methods for non-dominated sorting known as deductive sort and climbing sort were presented in (McClymont & Keedwell, 2012). The results demonstrated that deductive sort outperforms the fast non-dominated sort of NSGA-II. Another method for non-dominated sorting named efficient non-dominated sorting (ENS) strategy was proposed in (X. Zhang, Tian, Cheng, & Jin, 2015) that for searching within ENS, two different strategies including binary search (BS) and sequential search (SS) are employed. Although the results show that both ENS-based non-dominated sorting strategies are more efficient than other sorting approaches, the efficiency of the ENS-based sorting strategies decreases as the number of objectives increases (Modiri-Delshad & Rahim, 2016).

Using an elitism strategy to store the non-dominated solutions found within the optimization process is another mechanism proposed in (Zitzler & Thiele, 1999) to update and retain the non-dominated solutions. The main motivation for using elitist reservation mechanism is that a result that is non-dominated to other results in its current generation is not necessarily non-dominated with respect to all other results which optimization algorithm has found so far (Coello, 2006).

An elitist reservation mechanism is adopted in this study as an external elitist archive to update and retain the non-dominated solutions in every generation of BSA. Initially, the external elitist archive is empty; within the optimization progresses, it stores the non-dominated solutions according to the following 'if-then' rules:

i- If the trial pattern (a new generated solution) dominates some of the archived elitist, then all dominated members of the external elitist archive are replaced by non-dominated trial pattern;

ii- If the trial pattern is dominated by at least one member of the external elitist archive, then the trial pattern is disregarded for elitist archive;

iii- If the archived members of the external elitist archive are not dominated by the trial pattern and trial pattern is not dominated by the archived members, then the external elitist archive retains the trial pattern as a new elitist member (non-dominated solution)

As the optimization progresses, the members of external elitist archive increases. Therefore, to prevent overpopulation of the external elitist archive, the crowding distance of all members is measured and the extra members of the archive are removed according to their crowding distance value.

- **Crowding distance**

To keep the external elitist archive to its maximum capacity, the crowding distances (CD) of all solutions in Pareto-front (external elitist archive) are computed and the solution with the lowest CD value is subject to deletion when the archive is overloaded.

The crowding distance is a factor to evaluate the distribution of the solutions in Pareto -front by measuring the density around a solution. The crowding distances computes the distance of two neighbor points around the solution. The crowding distances of i^{th} solution in the Pareto -front is calculated by Eq. (3.16) as

$$CD_i = \sum_{j=1}^{m} \frac{f_j(i+1) - f_j(i-1)}{f_j^{max} - f_j^{min}} \tag{3.16}$$

where CD_i is crowding distance of solution I, f_j is the j^{th} objective function, $f_j^{\,max}$ and $f_j^{\,min}$ are maximum and minimum values of the j^{th} objective function respectively and m is the number of objectives.

For the boundary solutions ($f_j^{\,max}$ and $f_j^{\,min}$), the crowding distance is set to infinite as there is only one neighbor point for those solutions.

- **Procedure of multi-objective BSA (non-dominated approach)**

In the context of multi-objective optimization, instead of unique solution, there is a Pareto optimal set corresponding to the optimal value of each objective. Thus, to extend BSA to a multi-objective optimization approach, the replacement mechanism is adapted according to the concept of Pareto dominance (Modiri-Delshad & Rahim, 2016). Similar to BSA, in multi-objective BSA, mutation and crossover operators are first applied to produce the offspring (T). Then, the comparison in the final step of BSA (export the global minimum) is modified according to the concept of Pareto dominance to replace i^{th} individual of population (P_i) by i^{th} individual in offspring (T_i) if the P_i is dominated by T_i. The consecutive steps in BSA algorithm are followed to form the multi-objective BSA, except the last step (export the global minimum). Instead of exporting a global minimum, a Pareto optimal set is stored.

The following steps represent the sequential procedure of multi-objective BSA with an external elitist archive and crowding distance measure:

Step 1: The initial population (P) equals to the dimension of the optimization variables is randomly generated according to Eq. (3.7).

Step 2: Evaluate the fitness function of each individual of initial population and store the non-dominated solutions among the population into the external elitist archive.

Step 3: The historical population is determined randomly according to Eq. (3.8).

Step 4: The historical population is updated at each iteration through the 'if-then' decision rule according to Eq. (3.9)

Step 5: Apply the mutation operator to generate an initial form of trial population (Mutant) according to Eq. (3.11).

Step 6: Apply the crossover operator over the initial form of trial population obtained in pervious step to generate the final form of offspring (*T*) according to Eq. (3.12).

Step 7: At the end of the crossover process, if an individual in the generated offspring (*T*) violates the boundary condition of the optimization problem, the related individual in offspring is updated according to the boundary control developed by Eq. (3.13).

Step 8: If the i^{th} element of generated offspring (T_i) dominates the i^{th} element of population (P_i), then T_i is replaced by P_i.

Step 9: Update the external elitist archive according to aforementioned 'if-then' rules.

Step 10: If the external elitist archive exceeds its maximum capacity, the crowing distance for all members of elitist archive are computed and the less crowded solutions are removed from the archive one after another.

Step 11: If the stopping criteria are not satisfied, set $g=g+1$ and return to the step 4.

In the final stage of the feature selection technique, MOBBSA will be combined with ANFIS to select the best features for forecasting electricity price. ANFIS is used in the feature selection part to evaluate of selected subset. The next section provides the details of ANFIS.

3.3.2.2 Adaptive Neuro-Fuzzy Inference System

In this stage, the performances of the selected features (prices and demands) are evaluated by ANFIS technique. The fuzzy logic approach is based on the predefined rules (if-then) that lacks the ability to learn and adapt themselves to a new condition. Thus, to overcome this drawback authors in (J. S. R. Jang, 1993) hybridized a fuzzy inference system (FIS) with ANN to form ANFIS. The ANFIS methodology can be considered as an adaptive system in the form similar to ANN, which by training the system the parameters of the fuzzy membership functions (antecedent parameters) and the parameters of the fuzzy system output function (consequent parameters) are adapted. ANFIS possesses the advantage of both FIS and ANN and it has solved the drawbacks of both systems. The complicated procedures of neural networks are bypassed by applying linguistic variables of FIS system and the lack of FIS is solved by applying the neural inference system, which creates the ability to learn and adapt them-self to new condition. Therefore, this approach is capable to simulate complex nonlinear mappings using fuzzy system with ANN learning and it is considered as a universal estimator capable for short, medium and long-term forecasting.

ANFIS was developed as an adaptive system with a set of fuzzy rules (if-then) and tunable membership function (MF) parameters in a training phase. During the training phase of ANFIS, two different parameters should be optimized to provide the learning procedures:

i- Antecedent parameters (the MF parameters)

ii- Consequent parameters (the fuzzy system output function)

As the consequent parameters are linear, to optimize these parameters, the linear least-squares method is applied and to optimize the antecedent parameters similar to neural networks the backpropagation algorithm in conjunction with an optimization method such as gradient descent is applied.

Generally, five different layers construct the ANFIS structure while each layer consists of node functions and the inputs of the nodes in the present layer are obtained from previous layers (Tavana, Fallahpour, Di Caprio, & Santos-Arteaga, 2016). The consecutive layers of ANFIS structure are as follows: layer 1 is fuzzification (if-part), layer 2 is production part, layer 3 is normalization part, layer 4 is defuzzification (then-part), and eventually layer 5 is total output generation part. Figure 3.4 shows the structure of ANFIS with two independent variables *x* (prices) and *y* (demands) as input and one dependent variable price or f_{out} as an output.

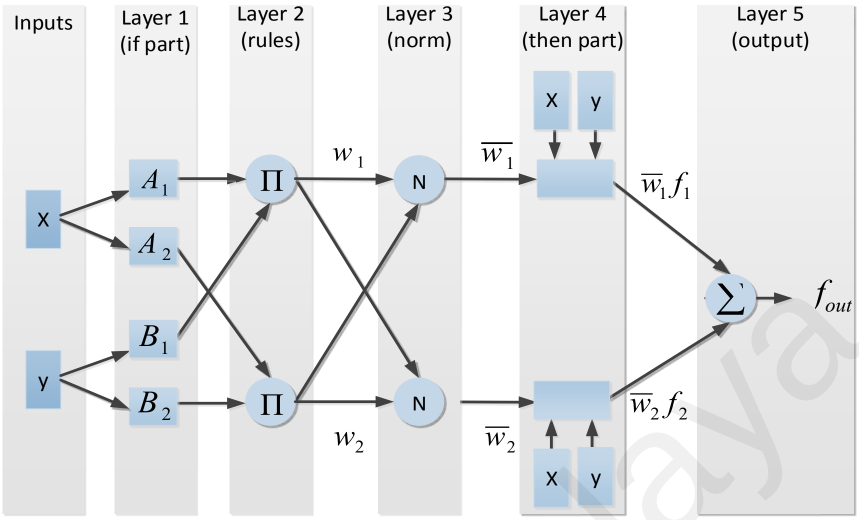

Figure 3.4: The general structure of ANFIS

For fuzzy inference systems, difference in the consequence of the set of fuzzy rules (if-then) and defuzzification procedures leads to two different types of fuzzy inference systems known as Mamdani type FIS and Sugeno type FIS.

In many aspects, Mamdani type FIS is similar to Sugeno method. The fuzzifying the inputs data and executing the fuzzy operators are similar fuzzy inference process in both types. The main difference between Sugeno type FIS and Mamdani type FIS is the way the fuzzy inputs are converted to the crisp output. In Mamdani type FIS for computing the crisp output the defuzzification technique of a fuzzy output is used while in Sugeno type FIS the weighted average method is used. As the consequents of the rules are not fuzzy in the Sugeno method, the interpretability and expressive power of Mamdani output are eliminated in this method. In comparison to Mamdani type FIS, Sugeno has faster processing time since instead of the time consuming defuzzification process the weighted average method is applied. Particularly, in decision support applications, Mamdani method is widely applied, due to intuitive nature and the interpretable of the rule base provided in this method. Moreover, another difference between Sugeno and

Mamdani type FIS is that Sugeno has no output membership functions whereas Mamdani FIS has output membership. Hence, Sugeno method provides an output that is either linear (weighted) mathematical expression or a constant. Instead, Mamdani method provides an output that is a fuzzy set. ANFIS architectures representing both the Mamdani and Sugeno methods. In comparison to Mamdani type FIS, Sugeno has more flexibility in system design as latter can be integrated with ANFIS tool to model the systems more precisely (Svalina, Galzina, Lujić, & Šimunović, 2013).

Considering ANFIS with Sugeno type FIS, the rule base of ANFIS contains fuzzy IF-THEN rules of a first order Sugeno type FIS are stated as:

Rule 1: *If x is A_1 and y is B_1 then z is* $f_1(x, y; p_1, q_1, r_1) = x\,p_1 + y\,q_1 + r_1$

Rule 2: *If x is A_2 and y is B_2 then z is* $f_2(x, y; p_2, q_2, r_2) = x\,p_2 + y\,q_2 + r_2$

where $f_i(x, y; p_i, q_i, r_i)$ is a first order polynomial function which represents the outputs of the Sugeno type FIS, A_i and B_i are the fuzzy sets, and x and y are two different input and z is an output of ANFIS model.

In the ANFIS structure, different layers consists of different node function. As shown in Figure 3.4, adaptive nodes which represent the adjustable parameter sets are denoted by squares whereas fixed nodes, which represent the fixed parameter sets in the system are denoted by circles.

- **Layer 1**

Every node in this layer is an adaptive node with a node function as follow:

$$Q_{1,i} = \mu_{Ai}(x), \qquad i = 1,2 \tag{3.17}$$

$$Q_{1,i} = \mu_{Bi-2}(y), \quad i = 3,4 \tag{3.18}$$

where x and y are the inputs to node i , A_i and B_i are linguistic labels, μ_{Ai} and μ_{Bi} are the membership functions for Ai and Bi fuzzy sets, respectively and $Q_{I,i}$ is the membership grade of a fuzzy set and considered as the output of node i in the first layer which specifies the degree to the given input (x or y) satisfies the quantifies.

Typically, in ANFIS, the MF (membership function) for a fuzzy set can be any parameterized membership function, such as generalized Bell shaped function, Gaussian, trapezoidal or triangular.

A generalized Bell shaped MF (bell MF) is specified as follows:

$$\mu_A(x; a, b, c) = \frac{1}{1 + (\frac{x - c}{a})^{2b}} \tag{3.19}$$

A Gaussian MF is specified as follows:

$$\mu_A(x; c, \sigma) = e^{-0.5(\frac{x-c}{\sigma})^2} \tag{3.20}$$

where σ and c determine the width and center of Guassian MF, respectively.

A trapezoidal MF is specified as follows:

$$\mu_A(x; a, b, c, d) = max(min\left(\frac{x - a}{b - a}, 1, \frac{d - x}{d - c}\right), 0) \tag{3.21}$$

The parameters with $a < b \leq c < d$ specify the x cordinates of the four corners for the underlying trapezoidal MF.

A triangular MF is specified as follows:

$$\mu_A(x; a, b, c) = max(min\left(\frac{x - a}{b - a}, \frac{c - x}{c - b}\right), 0) \tag{3.22}$$

The parameters with $a < b < c$ specify the x cordinates of the three corners for the underlying triangular MF. In this layer, the parameters a,b,c,d and σ are the antecedent parameters.

- **Layer 2**

Every node in this layer is a fixed node whose output is the product of all the incoming signals. In this layer, through multiplication of input signals, the firing strength of every rule is determined,

$$Q_{2,i} = w_i = \mu_{Ai}(x)\mu_{Bi}(y), \qquad i = 1,2 \tag{3.23}$$

where w_i is the output signal, which represents the firing strength of a rule.

- **Layer 3**

Every node in this layer is a fixed node. In this layer, the firing strength provided in previous layer is normalized by computing the ratio of the ith rule's firing strength to the sum of all rules' firing strengths,

$$Q_{3,i} = \overline{w_i} = \frac{w_i}{w_1 + w_2} \qquad i = 1,2 \tag{3.24}$$

where $\overline{w}$ is output signal which represents the normalized firing strength of a rule.

- **Layer 4**

In this layer every node i is adaptive with a node function,

$$Q_{4,i} = \overline{w_i} f_i \qquad i = 1,2 \tag{3.25}$$

where f_1 and f_2 are the fuzzy IF-THEN rules as follows:

Rule1: *If x is A_1 and y is B_1 then $z=f_1(x, y; p_1, q_1, r_1)$*

Rule2: *If x is* A_2 *and y is* B_2 *then* $z=f_2(x, y; p_2, q_2, r_2)$

where r_i, q_i and p_i are the parameter set, referred to as the linear consequent parameters.

- **Layer 5**

This layer has only one fixed node that computes the overall output of ANFIS by summation of all incoming signals,

$$Q_{5,i} = f_{out} = \sum_i \overline{w_i} f_i = \frac{\sum_i w_i f_i}{\sum_i w_i} = \ overal\ output \qquad i = 1,2 \tag{3.26}$$

The overall output is linear combination of the consequent parameters. Thus, the final output of ANFIS is expressed as

$$\begin{aligned} f_{out} &= \overline{w_1} f_1 + \overline{w_2} f_2 = \frac{w_1}{w_1 + w_2} f_1 + \frac{w_2}{w_1 + w_2} f_2 \\ &= (\overline{w_1} x) p_1 + (\overline{w_2} x) p_2 + (\overline{w_1} y) q_1 + (\overline{w_2} y) q_2 + (\overline{w_1}) r_1 + (\overline{w_2}) r_2 \end{aligned} \tag{3.27}$$

Eventually, ANFIS applies a hybrid learning algorithm for parameters tuning. It utilizes the back propagation algorithm and the least squared method for updating the input MF parameters (antecedent parameters) in layer 1, and training the consequent parameters, respectively.

3.4 The proposed (MI+ (MOBBSA+ANFIS)) hybrid feature selection method for short-term electricity price forecasting

In this stage, the most influential prices and demands are selected for forecasting electricity price. The main idea of the proposed feature selection technique is to choose a subset of candidate inputs by eliminating irrelevant and redundant candidates through mutual information in first stage.

In the second stage, (MOBBSA+ANFIS) is used in this study to select prices and demands subsets, which have substantial impact on forecasting of electricity price. In this stage, hybrid feature selection method has been used to choose dissimilar and most relevance features with minimum number of features and maximum accuracy among the previously selected candidates from first stage, which for the process of forecasting have been used as input. The description of the implemented feature selection method is presented in Figure 3.5.

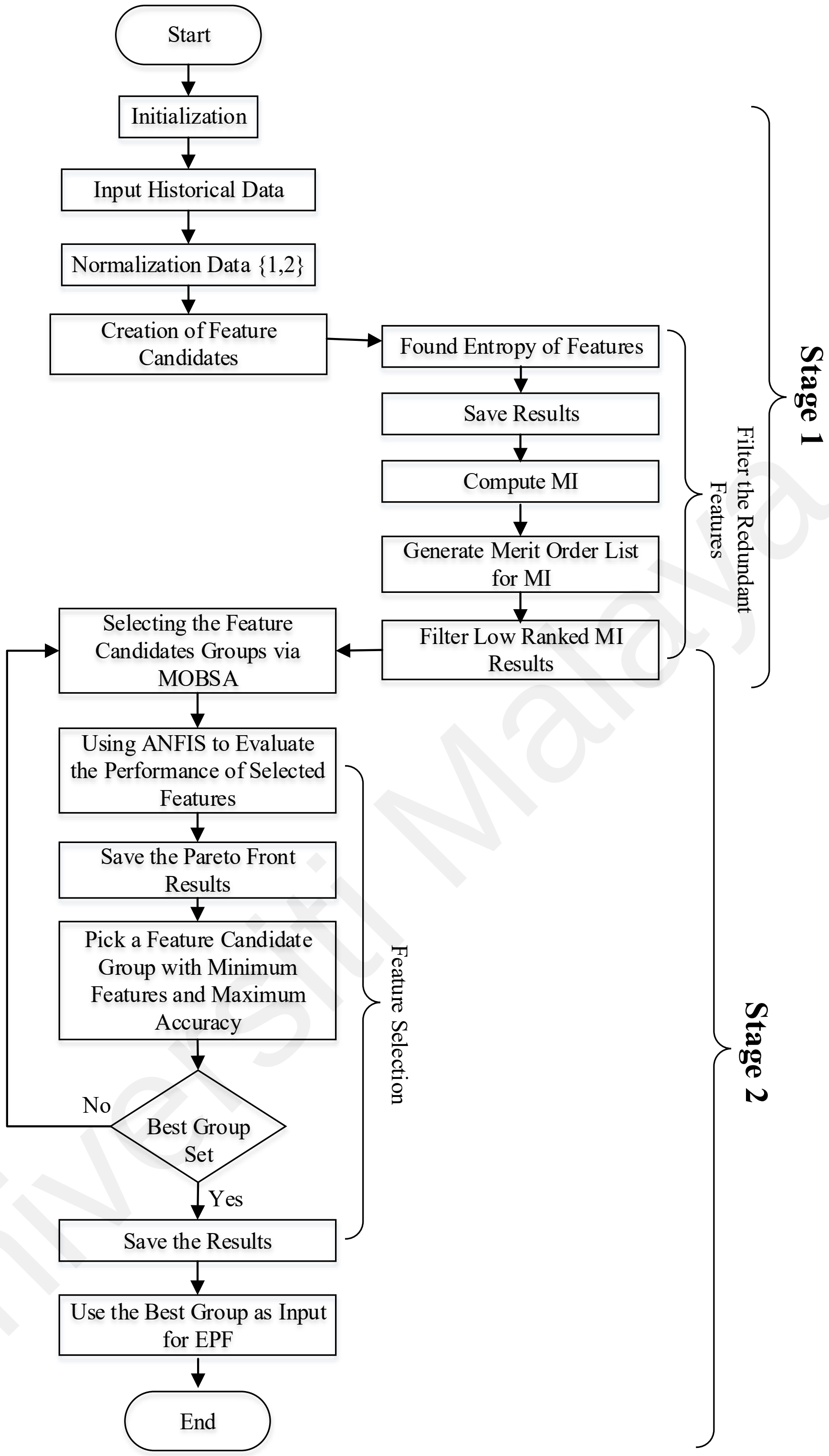

Figure 3.5: The procedure of feature selection development for short-term electricity price forecasting (EPF)

3.5 The proposed (ANFIS+BSA) hybrid technique for short-term electricity price forecasting

The procedure of the proposed methodology for short-term electricity price forecasting (EPF) of Canadian and Australian de-regulated electricity market is illustrated in Figure 3.6. In general, a similar procedure is followed to find the empirical models by the soft computing techniques. Thus, the following sequential steps are carried out for all applied methods in this study to obtain the optimal AI-based models for short-term EPF:

a) The electricity price is considered as a time series with hourly intervals that is a function of electricity price and demand in preceding hours. The particular data of Ontario and Queensland deregulated electricity markets in 2017 (HOEP and HOED) and 2018 (EP and ED) are selected by the proposed feature selection method and are taken as independent variables. HOEP on that period is considered as a dependent variable.

b) To predict the electricity price (EP) precisely, metaheuristic methods is implemented for seeking the optimal coefficients of ANFIS by minimizing the cost function as follows:

$$F = \sum_{t=1}^{N} \left| \left(EP\,(t)_{actual} - EP(t)_{estimated} \right) \right| \tag{3.28}$$

where $EP\ (t)_{actual}$ and $EP(t)_{estimated}$ are the actual and predicted electricity price respectively and N represents the number of observation.

The coefficients of ANFIS models (w) are determined by:

w=argmin F (3.29)

c) As the pattern of electricity demand seasonally changes, one month from each season (i.e., February, May, August and November) is considered to assess the effectiveness of applied methods for forecasting the electricity price in different seasons. Both independent and dependent variables are divided into two subsets, where the hourly data of first three weeks is used for training of design phase and that last week of each month is utilized for testing phase of obtained models.

d) The learning process occurs in the training phase. The computer programs that connect the independent variables to the dependent variable are derived through learning process. To speed up the learning process, both independent and dependent variables are normalized according to Eq. (4.2) as stated in Chapter 4.

e) Although the testing phase does not have any role in developing the models, it is utilized to evaluate the performance of the models obtained by AI-based methods. To quantify the forecasting performance of the obtained models, various evaluation criteria are carried out such as MAPE (mean absolute percentage error), RMSE (root mean square error) and Thiel's inequality coefficient (U-statistic). The mathematical equations of these criteria are:

$$MAPE\% = \frac{1}{N}\sum_{t=1}^{N}\frac{\left|\left(EP(t)_{actual} - EP(t)_{estimated}\right)\right|}{EP(t)_{actual}} \times 100 \tag{3.30}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{t=1}^{N}\left(EP(t)_{actual} - EP(t)_{estimated}\right)^2} \tag{3.31}$$

$$U = \frac{RMSE}{\sqrt{\frac{1}{N}\sum_{t=1}^{N}\left(EP(t)_{actual}\right)^2} + \sqrt{\frac{1}{N}\sum_{t=1}^{N}\left(EP(t)_{estimated}\right)^2}} \tag{3.32}$$

U-statistic provides a measure of how well fitted a time series of forecasted values to a corresponding time series of actual data. The U-statistic is always laid between [0,1] while the value closer to zero indicating greater forecasting accuracy with a perfect fit and the value closer to one means the estimation is no better than a naive guess.

f) Whiteness test (Durbin-Watson test) is used to ensure that the obtained models adequately describe a given data series (AlRashidi & El-Naggar, 2010). This test is obtained through a confirmatory analysis. The purpose of confirmatory analysis is to ensure the whiteness of estimated residuals (*e(t)*). The whiteness of estimated residuals implies that they are uncorrelated. Residuals autocorrelation function (RACF) is utilized to analyze the correlation of whiteness of estimated residuals by:

$$RACF = \frac{\sum_{t=2}^{N}(e(t)e(t-1))}{\sum_{t=1}^{N}(e(t))^2} \tag{3.33}$$

The RACF values come into the range of [0, 1], if the RACF value is significantly different from zero, it will fall outside a confidence level. This specifies that the residuals are not white (correlated) and an important independent variable has been omitted from the tested model.

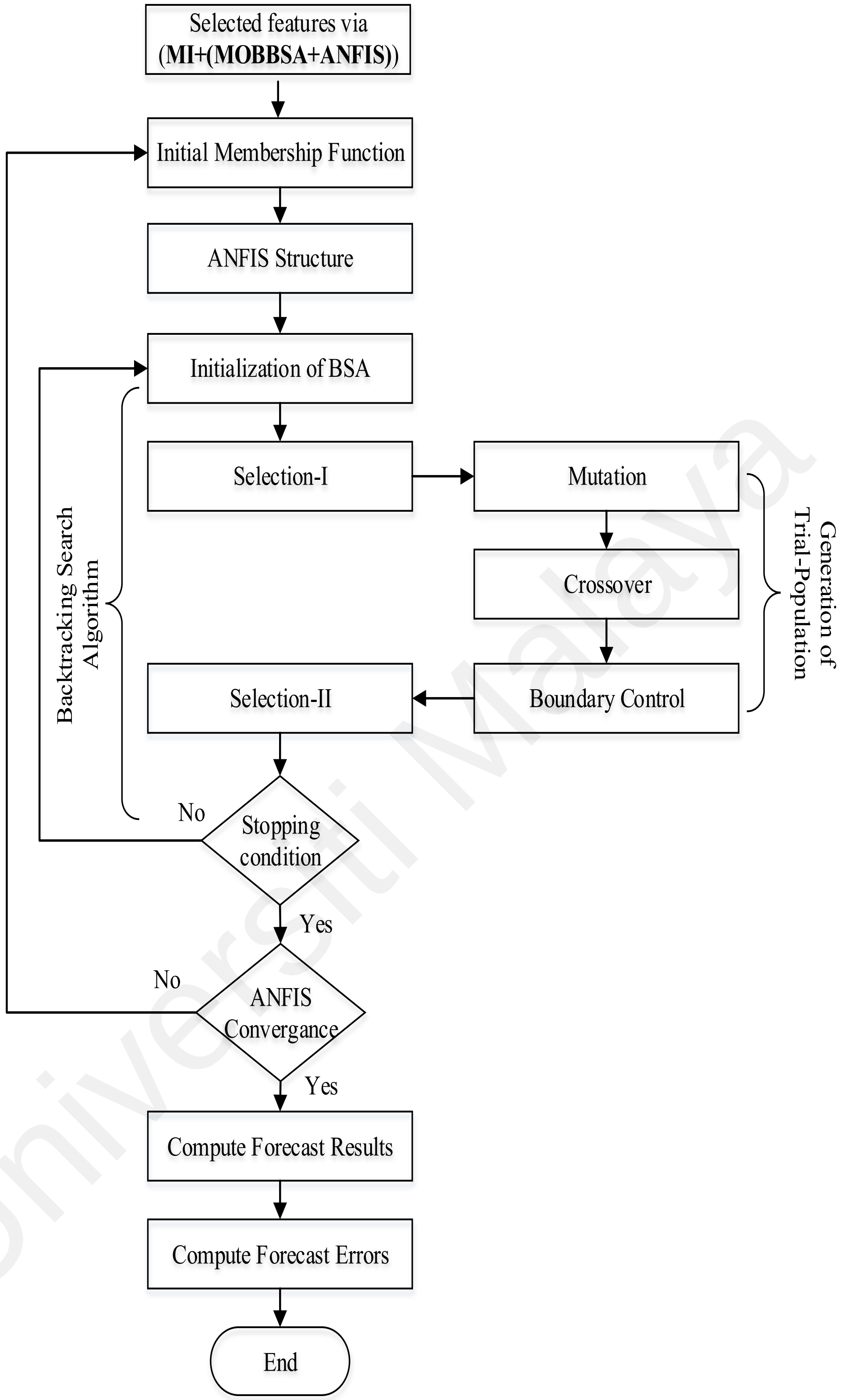

Figure 3.6: The procedure of ANFIS-BSA for short-term electricity price forecasting (EPF)

3.6 Summary

In this work, a new forecast strategy has been proposed for day-ahead price forecast of competitive electricity markets. The proposed method consists of two-stage feature selection method comprising of multi-objective binary-valued backtracking search algorithm (MOBBSA) and optimized adaptive neuro-fuzzy inference system (ANFIS) technique as a forecasting engine. A new multi-objective backtracking search algorithm (MOBBSA) is developed and its procedure based on minimize the number of feature and maximize the accuracy is presented in this chapter.

Firstly, two stages feature selection technique are developed based on combination of mutual information to extract redundant features (prices and demands) in first stage and hybrid (MOBBSA+ANFIS) to select the best feature subsets of input variables with maximum accuracy and minimum features in second stage. Secondly, the best feature subsets will be fed into the ANFIS again to forecast the price. Here, the optimization techniques are implemented to exploit the solution structure and explore its appropriate weighting factors of the learning process of ANFIS and ANN in order to increase the forecast accuracy. The performance of the proposed technique is verified and tested on Ontario and Queensland competitive electricity markets for different seasons in chapter 4. The simulation results confirm the ability of the proposed technique to improve the forecasting accuracy of existing artificial based methods for short-term electricity price.

CHAPTER 4: RESULTS AND DISCUSSIONS

4.1 Introduction

This chapter presents the results of the proposed feature selection methods and the proposed forecasting method for short-term electricity price of volatile systems. An efficient feature selection technique in the applied method has the ability to provide the highly accurate and efficient day-ahead price forecasting. The presented method consists of two-stage feature selection method comprising of multi-objective binary-valued backtracking search algorithm (MOBBSA) and optimized adaptive neuro-fuzzy inference system (ANFIS) technique as a forecasting engine. At the first stage of feature selection, mutual information is used to extract redundant features. Then, MOBBSA will search within different combinations of input variables before being evaluated by ANFIS to select the best feature subsets of input variables with maximum relevancy and minimum redundancy in second stage of feature selection. In the second stage of forecasting process, the best feature subsets will be fed into the ANFIS again to forecast the price. Several actual data sets are collected over the different weather seasons of Ontario electricity market of year 2017 and Queensland electricity market of year 2018 for the verification of simulation result.

4.2 Selected features for short-term electricity price forecasting

It can be considered that in electricity price forecasting process, electricity price in time *t* depends not only on the electricity demand at time *t* but also on their past values and even the past values of the electricity demand which is expressed as follows:

$$EP(t) = f\begin{pmatrix} EP(t-1), EP(t-2), EP(t-3),\ \ldots\ , EP(t-NL_{EP}), \\ ED(t), ED(t-1), ED(t-2), ED(t-3),\ \ldots\ , ED(t-NL_{ED}) \end{pmatrix} \tag{4.1}$$

where $EP(t)$ and $ED(t)$ represent the electricity price and demand at time t while they are assumed as a time series with t interval, NL_{EP} presents the number of lag order for the electricity price and similarly NL_{ED} denotes electricity demand lag order.

Accurate short-term forecasting is driving the expansion of de-regulated electricity market paving the way to sustainable development as the global population continues to grow at an unprecedented rate. As part of this, Ontario being second largest province in Canada with respect to surface area succeeding Quebec is the most densely populated. Ontario's electricity market is interconnected directly or indirectly to several other electricity markets i.e. direct connection with New York and Midwest electricity markets, and indirect connection with New England and PJM markets. Furthermore, Ontario trades electricity with regulated utilities in Quebec and Manitoba which further have connections with several other utilities in US. The hybrid ANFIS-BSA is applied in this study to forecast hourly Ontario electricity price (HOEP) according to two types of input historical data. The input historical data sets of HOEP and hourly Ontario electricity demand (HOED) in 2017 are adapted from ((IESO), 2018-11-17). Since the range of historical data varies widely, both dependent and independent variables are normalized according to Eq. (4.2),

$$\overline{Z}(t) = \frac{Z(t) - min(Z)}{max(Z) - min(Z)} + 1 \tag{4.2}$$

where $\overline{Z}$ is the normalized value, Z is the value to be normalized, and t is hourly interval. The main feature of data normalization is adjusting raw data observed on different scales to a notionally common scale, often prior to data processing.

Assuming only one week with hourly lagged, values of exogenous variables (NL_{EP} = NL_{ED} = 168) are used to forecast electricity price, which constitutes 336 lagged values of exogenous variables. If all aforementioned exogenous variables are applied in the forecasting process, this will not only slow down the learning process but will also give rise to a poor performance and overfitting the training data. In other words, although these factors are of vital importance for electricity price forecasting process, only those features exhibiting significant influence on the output should be picked. Although all subsets of variables have causal relation with electricity price, it is neither efficient nor feasible to employ all of them as inputs and assess their performances. Instead, an efficient feature selection is applied to select the most effective subset of variables in model construction.

Wrapper methods often provide the most relevant features for particular type of model but they need a systematic searching algorithm in their evolutionary training process. Since the exhaustive search is generally impractical, the sequential search as a heuristic technique to feature selection has been proposed to add or remove features sequentially until there is no improvement in the accuracy of model. The sequential search has a tendency to become stagnated in local optima. Thus, the randomized search algorithm known as metaheuristic algorithm is proposed to explore the space of different feature subsets. Metaheuristic algorithms incorporating randomness into search procedure of feature selection to escape local optima. GA, SA, ACO and PSO are different metaheuristic algorithms applied in the context of feature selection (Unler & Murat, 2010).

These metaheuristic algorithms deal with feature selection as a single objective optimization problem so that the number of relevant features should be predefined and always find the subset of features with fixed number of features. Generally, feature selection has two main conflicting objectives, which are minimizing simultaneously both the estimation error and the number of features. Therefore, feature selection problem can be expressed as a multi-objective problem that has two main objectives, maximizing the accuracy of model and minimizing the number of features and the decision is a tradeoff between these two objectives. Treating feature selection as a multi-objective problem leads to a set of non-dominated feature subsets to meet different requirements in real-world applications.

Multi-objective particle swarm optimization (MOPSO) and NSGA-II have been investigated in (Xue, Zhang, & Browne, 2013, 2014) to generate a Pareto front of feature subsets, but still more efficient searching strategy is essential to better address feature selection problems (C. Zhang, Zhou, Li, Fu, & Peng, 2017). Existing multi-objective feature selection algorithms suffer from the problems of high computational cost, too many control parameters, and over sensitivity to initial value of these parameters, BSA with only one control parameter is argued computationally less expensive than other metaheuristic algorithms (Modiri-Delshad & Rahim, 2016). A binary-valued BSA (BBSA) proposed in (Ahmed et al., 2017) to solve the optimization of discrete parameters. In BBSA, the individuals in the population are encoded as a binary vector, and the population value is converted to zero or one according to Eq. (4.3),

$$
\begin{aligned}
S_{i,j} &= \frac{1}{1+e^{-w}} \\
BP_{i,j} &= \begin{cases} 0 & if \;\; s_{i,j} < 0.5 \\ 1 & if \;\; s_{i,j} \geq 0.5 \end{cases}
\end{aligned}
\tag{4.3}
$$

where *S* is the sigmoid function, *PB* is its binary value and *w* is the population value.

However, BBSA has been exploited as an efficient searching algorithm for feature selection in (C. Zhang et al., 2017). It treats the task as a single objective problem and it could not directly be used to address multi-objective feature selection problems. Different versions of multi-objective BSA (MOBSA) have been developed in (Modiri-

Delshad & Rahim, 2016; Zou, Chen, Li, Lu, & Lin, 2017). Statistical analyses in (Modiri-Delshad & Rahim, 2016) confirm that MOBSA is a promising optimization method for solving high dimensional multi-objective problems over different well-known multi-objective evolutionary algorithms (e.g. MOPSO and NSGA-II). Therefore, in this study BBSA-based multi-objective feature selection algorithm is developed as a promising technique to generate a Pareto front of non-dominated feature subsets.

To assess the performance of every candidate feature subsets, any learning algorithm (e.g. ANN, SVM, ANFIS) can be used in the evolutionary training process of BSA-based multi-objective feature selection algorithm. ANFIS is considered as a universal estimator due to its fast learning capability to approximate nonlinear functions (Tavana et al., 2016). Hence, it is adopted as an evaluation metric method in the proposed multi-objective wrapper-based feature selection method. In particular, ANFIS employs an efficient hybrid learning method that combines the least squares method and gradient descent. The least square method is the main factor for quick training (Alizadeh, Jolai, Aminnayeri, & Rada, 2012). Thus, after only few epoch of training, ANFIS is able to construct the predictive model. Since the least square method is computationally efficient, the models are constructed for various combinations of features selected by multi-objective BBSA (MOBBSA) and train them with single/few running of the least-squares method. Then, a non-dominated subset of features with the best performance is chosen for constructing the model.

Before applying the feature selection to extract the most influential subsets of input variables with maximum relevancy and minimum redundancy for short-term EPF, both dependent and independent variables are randomly divided into two sets: 70% as the training set and 30% as the test set. The training set is used to construct the ANFIS models with different subsets of input variables while the test set is used to access the strength and utility of generated models.

In the developed multi-objective feature selection method, MOBBSA is used to search within different combinations of input variables and selects the non-dominated feature subsets, while ANFIS is applied as evaluation metric to determine the performance of every feature subset. During the training process of the applied learning method (ANFIS), every individual of MOBBSA represents one input variable. The developed feature selection strategy uses an elitist external archive to store non-dominated feature subsets, which simultaneously minimizes both the root mean square error (RMSE) on the test set and the number of input variables as the optimal solution set according to the concept of Pareto dominance. MOBBSA, as an extension of BSA, has only one control parameter named “mixrate,” which controls the number of individuals to be engaged in the crossover process.

The population of individuals is set to 100 and the maximum value of mixrate (i.e. 100% of the population size) is considered in the developed feature selection strategy to engage all the individuals in the crossover process. To form the structure of ANFIS for feature selection, the Sugeno-type FIS is used as a promising alternative to Mamdani-type since the Sugeno-type is well suited for modeling nonlinear systems by interpolating between multiple linear models. Based on the work conducted in (J.-S. Jang, 1996), the scatter partitioning can be used to facilitate the training process of ANFIS in feature selection. Since subtractive clustering is a realization of scatter partitioning, it is used to set up the ANFIS for feature selection. Here, firstly, the principal of information theory based on mutual information and entropy to eliminate redundant features for electricity price forecasting is explained.

In order to reduce the running time of feature selection, two stages feature selection is proposed in this work. The stages are presented in Fig 3.5. In the first stage of the proposed feature selection technique, the mutual information between every individual input variable and output feature is computed according to equation (3.1). As the value goes high, it indicates that each input variable and output is very dependent on each other. To filter the redundant features, the relevancy threshold is considered as TH = 0.46. After the filtering process, the most relevant attributes are selected as 60 features. In the 2nd stage, MOBBSA is used to search non-dominated feature subsets, while ANFIS is applied as evaluation metric to determine the performance of each feature subset. In this stage, among 60 selected candidates, most relevance and dissimilar are 23 features, which are selected by hybrid MOBBSA and ANFIS and later used as inputs for the next forecasting process.

Additionally, to calculate the efficiency of the suggested multi-objective feature selection method, comparison is done with 5 different hybrid well-known optimization techniques namely MOPSO+ANFIS, NSGA-II+ANFIS, MOBSA+ANN, MOPSO+ANN and NSGA-II+ANN. The optimal subsets of the input variables selected by studied multi-objective feature selection methods and their corresponding performances in terms of RMSE value are tabulated in Table 4.1 and Table 4.2. The *EP (t)* and *ED (t)* in both tables are EP and ED at t-time as optimal subsets of input variables selected by different multi-objective feature selection techniques. According to the obtained results, the suggested selection technique is very much efficient than other during the same test as it provides less estimation error and number of features. The first column of table 4.1 represents the features that are selected for forecasting process.

Table 4.1: The optimal subsets of input variables selected in Ontario by different multi-objective feature selection techniques and their corresponding performances in terms of RMSE value

MOBBSA+ ANFIS	MOPSO+ ANFIS	NSGAII+ ANFIS
EP(t-1)	EP(t-1)	EP(t-1)
EP(t-2)	EP(t-24)	EP(t-2)
EP(t-23)	EP(t-25)	EP(t-3)
EP(t-24)	EP(t-47)	EP(t-24)
EP(t-47)	EP(t-72)	EP(t-48)
EP(t-71)	EP(t-95)	EP(t-72)
EP(t-95)	EP(t-120)	EP(t-95)
EP(t-119)	EP(t-144)	EP(t-119)
EP(t-143)	EP(t-168)	EP(t-121)
EP(t-167)	EP(t-169)	EP(t-143)
EP(t-191)	EP(t-191)	EP(t-144)
EP(t-335)	EP(t-335)	EP(t-167)
EP(t-504)	EP(t-336)	EP(t-168)
ED(t)	EP(t-503)	EP(t-192)
ED(t-1)	ED(t)	EP(t-193)
ED(t-23)	ED(t-1)	EP(t-335)
ED(t-71)	ED(t-23)	EP(t-336)
ED(t-96)	ED(t-24)	EP(t-503)
ED(t-120)	ED(t-96)	EP(t-504)
ED(t-144)	ED(t-120)	ED(t)
ED(t-168)	ED(t-144)	ED(t-1)
ED(t-192)	ED(t-168)	ED(t-2)
ED(t-335)	ED(t-191)	ED(t-24)
	ED(t-192)	ED(t-48)
	ED(t-336)	ED(t-72)
	EP(t-504)	ED(t-96)
		ED(t-120)
		ED(t-144)
		ED(t-336)
RMSE		
12.23	13.05	13.87

Table 4.2: The optimal subsets of input variables selected in Ontario by different multi-objective feature selection techniques and their corresponding performances in terms of RMSE value

MOBBSA+ ANN	MOPSO+ ANN	NSGAII+ ANN
EP(t-1)	EP(t-1)	EP(t-1)
EP(t-2)	EP(t-2)	EP(t-2)
EP(t-24)	EP(t-3)	EP(t-3)
EP(t-25)	EP(t-24)	EP(t-24)
EP(t-48)	EP(t-25)	EP(t-25)
EP(t-49)	EP(t-48)	EP(t-48)
EP(t-72)	EP(t-49)	EP(t-49)
EP(t-73)	EP(t-72)	EP(t-72)
EP(t-96)	EP(t-73)	EP(t-73)
EP(t-97)	EP(t-96)	EP(t-96)
EP(t-120)	EP(t-97)	EP(t-97)
EP(t-121)	EP(t-120)	EP(t-120)
EP(t-144)	EP(t-121)	EP(t-121)
EP(t-145)	EP(t-144)	EP(t-144)
EP(t-168)	EP(t-145)	EP(t-145)
EP(t-169)	EP(t-168)	EP(t-168)
EP(t-192)	EP(t-169)	EP(t-169)
EP(t-193)	EP(t-192)	EP(t-192)
ED(t)	EP(t-193)	EP(t-193)
ED(t-1)	ED(t)	ED(t)
ED(t-24)	ED(t-1)	ED(t-1)
ED(t-72)	ED(t-3)	ED(t-24)
ED(t-96)	ED(t-24)	ED(t-25)
ED(t-120)	ED(t-72)	ED(t-48)
ED(t-144)	ED(t-96)	ED(t-49)
ED(t-168)	ED(t-120)	ED(t-72)
ED(t-336)	ED(t-144)	ED(t-96)
	ED(t-168)	ED(t-120)
		ED(t-144)
		ED(t-168)
		ED(t-504)
RMSE		
13.47	13.85	14.74

4.3 Simulation Results and Discussion

In this study, ANFIS-BSA is employed to enhance accuracy of electricity price forecasting (EPF) of HOEP, as its electricity market is recognized as the one of the unstable market due to its single settlement nature. The most effective input variables are selected by proposed feature selection method. Thus, for electricity price forecasting, only the value of the selected inputs should be predefined as they are input subset of forecasting engine. Furthermore, to assess the effectiveness of ANFIS-BSA for short-term EPF, its estimates are compared with those obtained from the following techniques: ANN, ANFIS, ANN-GA, ANN-PSO, ANN-BSA, ANFIS-GA, and ANFIS-PSO.

As there is no consensus about the optimum values of the AI-based methods parameters setting, the control parameters of applied methods are set according to the similar methodologies have been successfully applied in the literature for energy price/demand forecasting. All parameter settings of applied methods are summarized in Table 4.3.

The back propagation MLP as a type of ANN using feed-forward architecture trained with BP (back propagation) learning algorithm is used. This network is known as universal estimator due to its simple solution network and faster computational procedure which allows supervised training over large input data sets. There are some factors such as the network structure, transfer function and learning algorithm that affect the effectiveness of the developed models by ANN. For short-term EPF, two hidden layers, logarithmic sigmoid transfer function and Levenberg-Marquardt PB learning are used to form a MLP model as recommended in (Ekonomou, 2010).

Instead of Mamdani-type FIS (fuzzy inference system), the Sugeno-type FIS is used to form the ANFIS structure. The Sugeno-type is effectively suitable for modeling nonlinear systems by interpolating between multiple linear models. The membership function type is Gaussian as recommended in (Ali Azadeh, Saberi, Gitiforouz, & Saberi, 2009) and subtractive clustering (radii = 0.8) methods is used to set up the ANFIS.

The GA is executed by genetic algorithm optimization toolbox (GAOT) with Gaussian mutation and scattered crossover as considered in Ref. (Ünler, 2008). All the control parameters, e.g., mutation rate and crossover rate, are set to be default as recommended in Ref. (A Azadeh, Saberi, & Seraj, 2010) and the population of the GA is set to 100.

The standard form of PSO algorithm is executed in this study. The population is set to 100 in this algorithm. The acceleration factors c_1 and c_2 are both 2.0, a decaying inertia weight ω starting at 0.9 and ending at 0.4 with run time increasing is used as specified in (Askarzadeh, 2014).

The number of individuals, which are going to participate in cross over process, are set by the control parameter of BSA. Generally, this parameter varies from 0% to 100% of population size. However, the performance of BSA is not affected by variation of this parameter. Therefore, to achieve the best solutions the maximum value of individuals is considered. The population of individuals is set to 100.

Table 4.3: Parameter setting of applied methods

Methods		Parameters	Value
ANN	MLP	Hidden layer	2
		Transfer function	logarithmic sigmoid
		Learning algorithm	Levenberg-Marquardt PB
ANFIS	SC (Subtractive clustering)	Cluster radius	0.8
		FIS structure	Sugeno-type
		Membership function	Gaussian
Metaheuristic optimization	GA	Population	Size:100,Type: double vector, Creation function: uniform
		Selection function	Stochastic uniform
		Elite count	5.0
		Crossover	Fraction: 0.8, Function: scattered
		Mutation	Function: Gaussian, Scale:1.0, Shrink:1.0
		Migration	Direction: forward, Fraction:0.2, Interval:20
	PSO	Swarm population	100
		w	[0.4, 0.9]
		c1=c2	2
	BSA	Number of individuals	100
		Control parameter rate (P)	100%

4.3.1 Validation of the Model Using Machine Learning Methods in Ontario market

Table 4.4 summarizes the performances of applied machine learning methods for Ontario EPF in February 2017. It is observed that the calculated RACF values validate the whiteness of estimated residuals for all developed models in an affirmed confidence range. According to Table 4.4, the forecasting accuracy of methods on Ontario electricity market according to multi-criteria decisions using the mean rank of the methods for each indicator (absolute error, RMSE, U-statistic and MAPE) is ranked as ANFIS-BSA > ANFIS-PSO > ANFIS-GA > MLP-BSA > MLP-PSO > MLP-GA > ANFIS > MLP. The enhanced results in Table 4.4 show that the optimized ANFIS methodologies provide better-fit estimation than other studied methods. Moreover, it is determined that BSA is the most efficient optimization algorithm for training the ANFS, while ANFIS-BSA achieved to MAPE = 2.79%, U-statistic = 0.03, RMSE = 0.07 and absolute error = 21.5, that is less than the values achieved by the other optimization algorithms. Figure 4.1 presents the performance test of ANFIS-BSA when the training is executed for design phase and testing phase.

Table 4.4: Comparison between forecasting accuracy of studied methods for EPF of Ontario in February 2017

Performance Indexes		Methods							
		MLP	MLP-GA	MLP-PSO	MLP-BSA	ANFIS	ANFIS-GA	ANFIS-PSO	ANFIS-BSA
RACF	Training	0.0005	0.0021	0.0017	0.0002	0.0012	0.0145	0.0045	0.0055
	Testing	0.0007	0.0009	0.0014	0.0005	0.0023	0.0034	0.0103	0.0037
	Whole set	0.0003	0.0012	0.0001	0.0019	0.0029	0.0006	0.0025	0.0143
Absolute error	Training	20.0043	19.0045	18.1357	17.7653	19.5678	17.6546	16.9876	**16.7089**
	Testing	7.0543	6.0654	5.9876	5.5673	6.7864	5.1328	4.8987	**4.7936**
	Whole set	27.0586	25.0699	24.1233	23.3326	26.3542	22.7874	21.8863	**21.5026**
RMSE	Training	0.0875	0.0853	0.0823	0.0789	0.0815	0.7654	0.7001	**0.0683**
	Testing	0.0845	0.0812	0.0802	0.0775	0.0798	0.0734	0.0678	**0.0628**
	Whole set	0.0867	0.0834	0.0821	0.0781	0.0801	0.0745	0.0699	**0.0670**
U-statistic	Training	0.0449	0.0413	0.0404	0.0367	0.0409	0.0353	0.0321	**0.0317**
	Testing	0.0411	0.0396	0.0384	0.0346	0.0386	0.0332	0.0309	**0.0297**
	Whole set	0.0438	0.0409	0.0391	0.0355	0.0395	0.0345	0.0327	**0.0312**
MAPE (%)	Training	5.0345	4.8965	4.5615	4.0018	4.8763	3.5683	3.0054	**2.8777**
	Testing	4.8865	4.4445	4.2347	3.7654	4.7899	3.0643	2.7647	**2.5237**
	Whole set	4.9818	4.5679	4.3781	3.8967	4.8011	3.1268	2.8964	**2.7892**

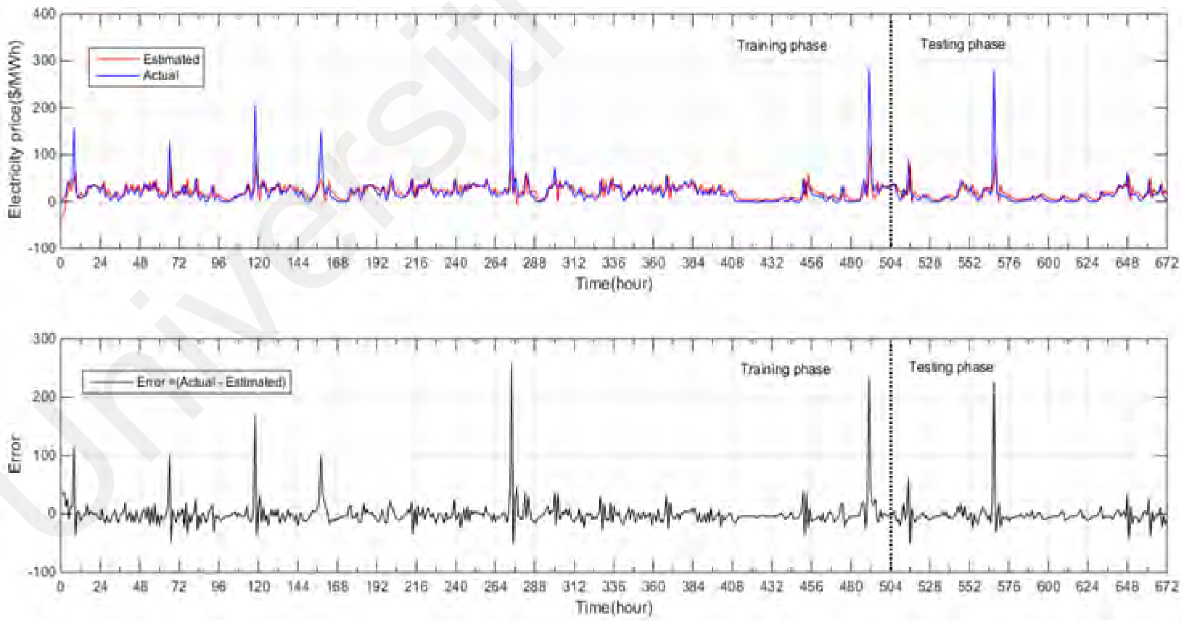

Figure 4.1: The performance of ANFIS-BSA during training of design phase and testing phase and its corresponding error for EPF of Ontario in February 2017

The performances of the applied machine learning methods for EPF of Ontario in May 2017 are tabulated in Table 4.5. The RACF values reported in this table confirm that the estimated residuals of all obtained models are white at a confidence interval level. The forecasting accuracy of applied methods for EPF according to multi-criteria decisions using the mean rank of the methods for each indicator (absolute error, RMSE, U-statistic and MAPE) in whole set is ranked as ANFIS-BSA > ANFIS-GA > MLP-BSA > ANFIS-PSO > MLP-PSO > MLP-GA > ANFIS > MLP. The comparison reveals that ANFIS-BSA is the most efficient model, while the best reported values in Table 4.5 in term of MAPE (0.87%), U-statistic (0.01), RMSE (0.02) and absolute error (6.98) belong to this model. The performances of ANFIS–BSA method for Ontario EPF in May is illustrated in Figure 4.2. The statistical factors of the ANFIS-BSA model for forecasting the Ontario's electricity price in this month are computed.

Table 4.5: Comparison between forecasting accuracy of studied methods for EPF of Ontario in May 2017

Performance Indexes		Methods							
		MLP	MLP-GA	MLP-PSO	MLP-BSA	ANFIS	ANFIS-GA	ANFIS-PSO	ANFIS-BSA
RACF	Training	0.0002	0.0033	0.0021	0.0005	0.0030	0.0019	0.0014	0.0004
	Testing	0.0011	0.0016	0.0009	0.0003	0.0004	0.0008	0.0019	0.0010
	Whole set	0.0004	0.0005	0.0024	0.0015	0.0013	0.0026	0.0005	0.0016
Absolute error	Training	6.4578	6.2346	6.0045	5.8941	6.2354	5.7654	6.0007	**5.5994**
	Testing	2.5436	2.2549	2.1643	1.9745	2.4543	1.8435	2.1003	**1.3796**
	Whole set	9.0014	8.4895	8.1688	7.8686	8.6897	7.6089	8.1010	**6.9790**
RMSE	Training	0.0321	0.0301	0.0288	0.0267	0.0293	0.0251	0.0279	**0.0247**
	Testing	0.0245	0.0244	0.0229	0.0194	0.0251	0.0178	0.0201	**0.0133**
	Whole set	0.0311	0.0298	0.0262	0.0243	0.0271	0.0239	0.0256	**0.0226**
U-statistic	Training	0.0194	0.0184	0.0155	0.0145	0.0164	0.0121	0.0129	**0.0119**
	Testing	0.0107	0.0092	0.0082	0.0077	0.0089	0.0069	0.0071	**0.0065**
	Whole set	0.0191	0.0178	0.0147	0.0141	0.0155	0.0118	0.0124	**0.0110**
MAPE (%)	Training	1.3223	1.2132	1.2001	1.0059	1.2267	0.9278	0.9976	**0.8987**
	Testing	1.1397	1.1014	1.0009	0.9879	1.1165	0.8963	0.9181	**0.7957**
	Whole set	1.2466	1.1887	1.0679	0.9977	1.1999	0.9145	0.9459	**0.8754**

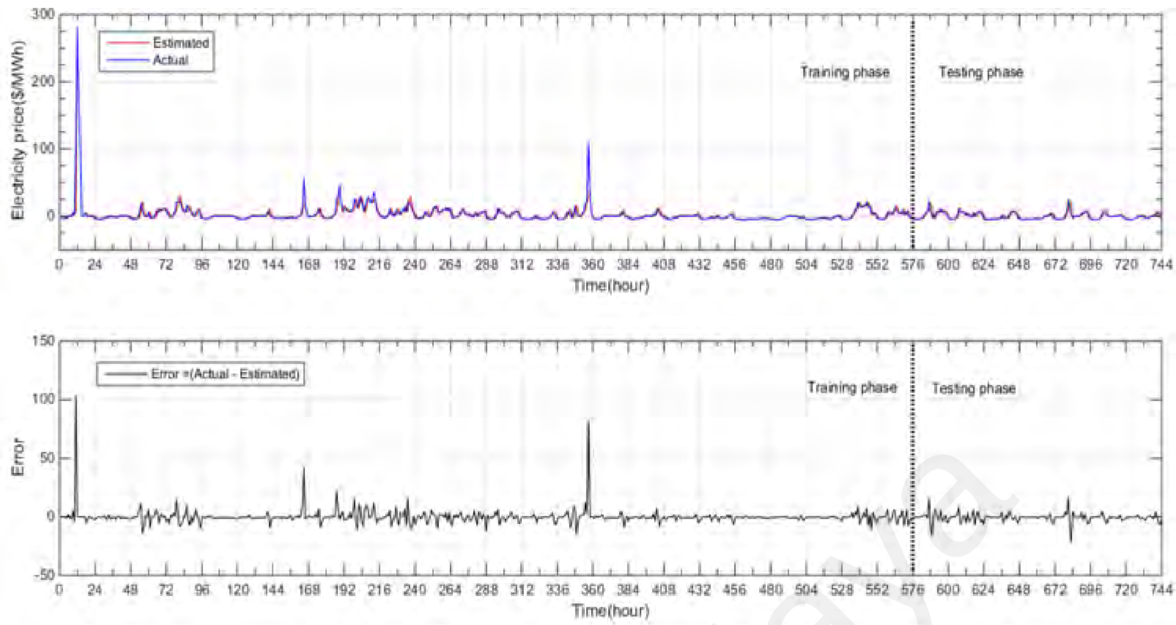

Figure 4.2: The performance of ANFIS-BSA during training of design phase and testing phase and its corresponding error for EPF of Ontario in May 2017

The performances of the applied methods for EPF of Ontario in August 2017 are quantified in Table 4.6. The RACF values in this table confirm the whiteness of estimated residuals at a confidence interval level for all obtained models. According to Table 4.6, the forecasting accuracy of studied methods in terms of multi-criteria decisions using the mean rank of the methods for each indicator (absolute error, RMSE, U-statistic and MAPE) in the whole set is ranked as ANFIS-BSA > ANFIS-PSO > ANFIS-GA > ANFIS > MLP-BSA > MLP-GA > MLP-PSO > MLP. The comparison between forecasting accuracy of studied methods on Ontario electricity market in August reveals that optimized ANFIS approaches outperform the other studied methods. Whereas the superior MAPE (1.7%), U-statistic (0.02), RMSE (0.05) and absolute error (14.98) values reported in Table 4.6 belong to ANFIS-BSA. Therefore, it is concluded that the most efficient optimization algorithm for training ANFIS is BSA. Figure 4.3 depicts the ANFIS-BSA performance for forecasting Ontario electricity price in August during training of design phase and testing phase. For further verification, the statistical factors of the ANFIS-BSA model for forecasting the Ontario's electricity price in this season are computed.

Table 4.6: Comparison between forecasting accuracy of studied methods for EPF of Ontario in August 2017

Performance Indexes		Methods							
		MLP	MLP-GA	MLP-PSO	MLP-BSA	ANFIS	ANFIS -GA	ANFIS -PSO	ANFIS -BSA
RACF	Training	0.0011	0.0014	0.0012	0.0001	0.0010	0.0006	4.4E-5	0.0003
	Testing	0.0002	0.0018	0.0009	5.6E-6	0.0007	0.0003	0.0015	0.0005
	Whole set	0.0004	0.0003	0.0008	0.0013	0.0002	0.0004	0.0004	0.0005
Absolute error	Training	16.6789	16.5432	14.5431	15.3456	16.3458	13.0432	12.9807	**12.4135**
	Testing	3.9875	3.8976	3.3247	3.4568	3.4567	3.0543	2.7896	**2.5635**
	Whole set	15.7646	15.5432	15.2123	15.3002	15.3246	15.0004	14.9996	**14.9770**
RMSE	Training	0.0682	0.0646	0.0603	0.0612	0.0653	0.0601	0.0598	**0.0581**
	Testing	0.0357	0.0345	0.0312	0.0323	0.0337	0.0278	0.0254	**0.0223**
	Whole set	0.0643	0.0637	0.5987	0.0601	0.0632	0.5432	0.5362	**0.0522**
U-statistic	Training	0.0348	0.3246	0.3001	0.0314	0.0335	0.2798	0.2781	**0.0264**
	Testing	0.0198	0.0188	0.0178	0.0185	0.0189	0.0132	0.0115	**0.0102**
	Whole set	0.0289	0.0278	0.0261	0.0269	0.0267	0.0256	0.0241	**0.0237**
MAPE (%)	Training	2.7897	2.6543	2.4321	2.5638	2.4567	2.2271	1.9861	**1.7909**
	Testing	2.6543	2.5328	2.3456	2.4538	2.2456	1.9643	1.7632	**1.4117**
	Whole set	2.7065	2.5895	2.3953	2.5431	2.3689	2.0048	1.8732	**1.7053**

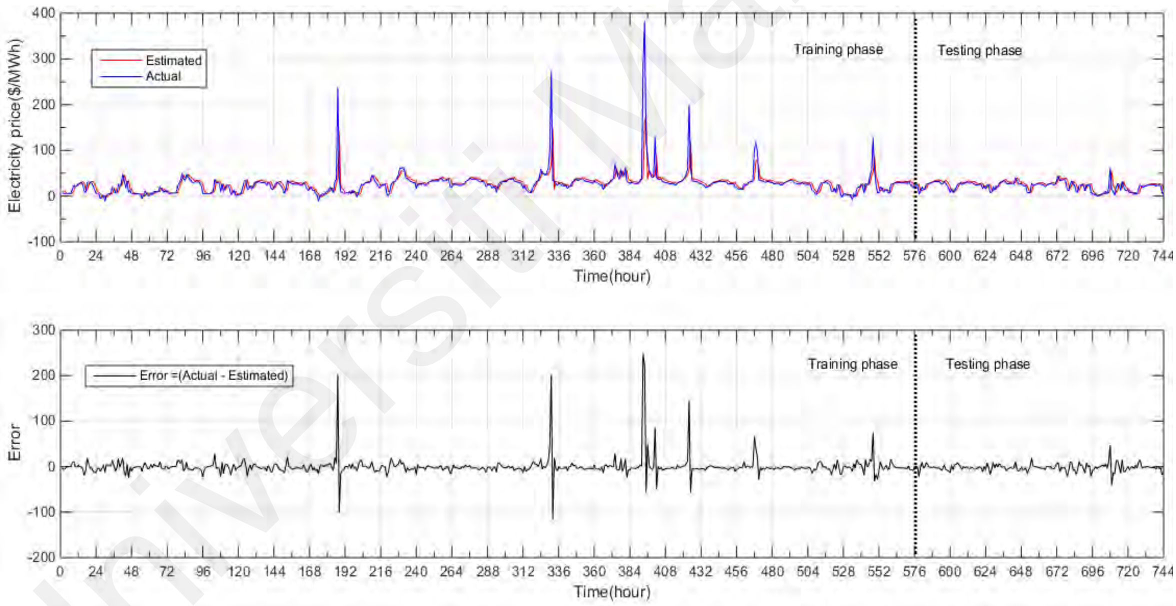

Figure 4.3: The performance of ANFIS-BSA during training of design phase and testing phase and its corresponding error for EPF of Ontario in August 2017

For further examination of solution methodology, the performance of optimized ANFIS methods for EPF of Ontario in November 2017 are compared with those from other soft computing approaches as shown in Table 4.7. The calculated RACF values in this table indicate that the estimated residuals of all models are uncorrelated and obtained models adequately describe the given set of data. According to Table 4.7, the forecasting accuracy of applied methods in terms of multi-criteria decisions using the mean rank of the methods for each indicator (absolute error, RMSE, U-statistic and MAPE) in whole set is ranked as ANFIS-BSA > ANFIS-PSO > ANFIS-GA > MLP-BSA > MLP-PSO > ANFIS > MLP-GA > MLP. The comparison between forecasting accuracy of studied methods on Ontario's electricity market in November reveals that optimized ANFIS approaches outperform the other studied methods. Whereas the superior MAPE (3.17%), U-statistic (0.03), RMSE (0.07) and absolute error (27.12) values reported in Table 4.7 belong to ANFIS-BSA model. Therefore, it is concluded that the most efficient optimization algorithm for training ANFIS. Figure 4.4 demonstrates the performance of ANFIS-BSA method for forecasting Ontario's electricity price in November during training of design phase and testing phase. The statistical factors of the ANFIS-BSA model for forecasting the Ontario electricity price in this month are computed.

Table 4.7: Comparison between forecasting accuracy of studied methods for EPF of Ontario in November 2017

Performance Indexes		Methods							
		MLP	MLP-GA	MLP-PSO	MLP-BSA	ANFIS	ANFIS-GA	ANFIS-PSO	ANFIS-BSA
RACF	Training	0.0012	0.0020	0.0018	0.0022	0.0011	0.0003	0.0012	0.0015
	Testing	0.0017	0.0023	0.0011	0.0013	0.0020	0.0008	0.0028	0.0019
	Whole set	0.0002	0.0016	0.0007	0.0004	0.0016	0.0010	0.0006	0.0009
Absolute error	Training	28.0031	27.9832	27.5421	26.3454	27.6067	27.0003	23.0045	**22.6287**
	Testing	8.2345	8.0046	7.8743	7.2451	7.9812	6.0321	5.0001	**4.4944**
	Whole set	36.2376	35.9878	35.4164	33.5905	35.5879	33.0324	28.0046	**27.1231**
RMSE	Training	0.0826	0.0819	0.0811	0.0802	0.0814	0.0804	0.0799	**0.0798**
	Testing	0.0794	0.0743	0.0720	0.0701	0.0785	0.0676	0.0503	**0.0431**
	Whole set	0.0803	0.0786	0.0754	0.0704	0.0793	0.0743	0.0737	**0.0729**
U-statistic	Training	0.0302	0.0298	0.0287	0.0278	0.0283	0.0276	0.0241	**0.0202**
	Testing	0.0497	0.0478	0.0463	0.0456	0.0482	0.0451	0.0389	**0.0358**
	Whole set	0.0392	0.0378	0.0364	0.0356	0.0375	0.0355	0.0341	**0.0331**
MAPE (%)	Training	4.6429	4.5329	4.4723	4.4011	4.4963	4.1212	3.7591	**3.4031**
	Testing	4.0428	3.9797	3.8346	3.7543	3.8203	3.3451	2.7521	**2.4180**
	Whole set	4.5465	4.3458	4.0768	3.9732	4.3103	3.7543	3.3541	**3.1732**

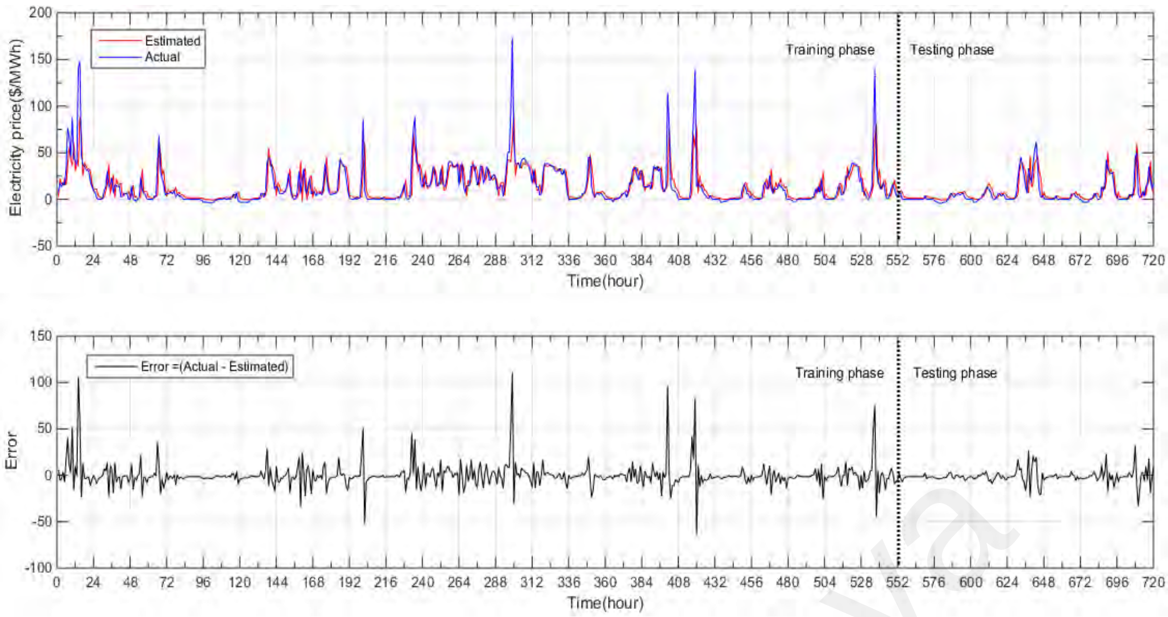

Figure 4.4: The performance of ANFIS-BSA during training of design phase and testing phase and its corresponding error for EPF of Ontario in November 2017

From the demand side management perspective, the negative price can be controlled in whole months of seasons by providing incentive for using the electricity in a particular time to curtail the consumption accordingly when the demand is low in order to stabilize the frequency and voltage of grid.

4.3.2 Validation of the Model Using Statistical Methods in Ontario market

Different statistical methods are also applied as external validation to verify the validity of models developed by ANFIS-BSA. To evaluate the performance of the obtained model the following attributes were recommended (Mostafavi, Mousavi, & Hosseinpour, 2014):

i- If a model gives $|R| > 0.8$, a strong correlation exists between the predicted and observed values.

ii- If a model gives $0.2 < |R| < 0.8$, a correlation exists between the predicted and observed values.

iii- If a model gives $|R| < 0.2$, a weak correlation exists between the predicted and observed values.

In addition, new factors suggested by (Golbraikh & Tropsha, 2002) are checked for external validation of the obtained models on the testing phase. It is recommended that at least one slope of the regression lines (k or k') through the origin should be close to one. It should be noted that k and k' are the slopes of the regression lines between the regressions of actual output (h_i) against predicted output (t_i) or t_i against h_i through the origin, i.e. h_i=k t_i and $t_i = k'h_i$, respectively.

In addition, the performance indexes of *m* and *n* should be less than 0.1 (*m and n* are the two factors for evaluating the model performance). Recently, Roy and Roy (Roy & Roy, 2008) presented a confirmed indicator (R_m) for external predictability of models. For $R_m > 0.5$, the condition is satisfied. Either the squared correlation coefficient (through the origin) between predicted and experimental values (${R_o}^2$), or the squared correlation coefficient between experimental and predicted values (${R_{o'}}^2$) should be close to R^2 and to one (Alavi, Aminian, Gandomi, & Esmaeili, 2011). In item one, *R* should be greater than 0.8. In the second item, *k* should be between 0.85 and 1.15. In the third item, k' should be between 0.85 and 1.15. According to items four and five, *m* and *n* values should be smaller than 0.1. Finally, R_m should be greater than 0.5.

Table 4.8 tabulates the statistical factors of the ANFI-BSA model for EPF of Ontario in February, May, August and November. As shown in this table, the developed models satisfy all the requisite conditions. The validation phase ensures that ANFIS-BSA provides precise models, which is strongly applicable for short term electricity price forecasting of Ontario.

Table 4.8: Statistical factors of the ANFIS-BSA model for EPF of Ontario in February (F), May (M), August (A) and November (N) 2017

Item	Formula	Condition	F	M	A	N
1	R	$0.8<R_0$	0.9992	0.9977	0.9966	0.9992
2	$K=\frac{\sum_{i=1}^{n}(h_i\times t_i)}{\sum_{i=1}^{n}h_i^2}$	0.85<k<1.15	0.9989	0.9991	0.9959	0.9983
3	$K'=\frac{\sum_{i=1}^{n}(h_i\times t_i)}{\sum_{i=1}^{n}t_i^2}$	0.85<k'<1.15	1.0010	1.0006	1.0038	1.0016
4	$m=\frac{R^2-R_o^2}{R^2}$	$\mid m \mid<0.1$	-0.0015	-0.0045	-0.0056	-0.0016
5	$n=\frac{R^2-R_{o'}^2}{R^2}$	$\mid n \mid<0.1$	-0.0015	-0.0045	-0.0057	-0.0016
6	$R_m=R^2\times\left(1-\sqrt{\lvert R^2-R_o^2\rvert}\right)$	$0.5<R_m$	0.9970	0.9910	0.9881	0.9967
Where	$R_o^2=1-\frac{\sum_{i=1}^{n}(t_i-h_i^o)^2}{\sum_{i=1}^{n}\left(t_i-\bar{t}_i\right)^2},h_i^o=k\times t_i$	$0.8<R_0^2<1$	1.0000	1.0000	0.9996	0.9999
	$R_{o'}^2=1-\frac{\sum_{i=1}^{n}(h_i-t_i^o)^2}{\sum_{i=1}^{n}\left(h_i-\bar{h}_i\right)^2},t_i^o=K'\times h_i$	$0.8<R_{0'}^2<1$	1.0000	1.0000	0.9997	0.9999

4.4 Australia day-a head electricity market

The Australian National Electricity Market (NEM) includes all states and territories other than Western Australia and the Northern Territory. Its centerpiece is a set of regional gross-pool spot energy and ancillary services markets that solve a security-constrained dispatch every 5 min. The Australian Energy Market Operator (AEMO) is the wholesale market operator and transmission system operator (TSO) for the entire system. Regions are currently located at all borders between states within the NEM. All generating plants of greater than 30-MW capacity (except intermittent generation including wind) are required to participate as scheduled generators and submit offers to sell or bids to buy energy (and/or ancillary services) in the NEM dispatch process. The pre dispatch processes forecasts up to 40 hours ahead of real time and provides public forecasts of energy and ancillary service prices and (privately to each dis patchable participant) dispatch levels based on participant bids and offers, the demand forecasts and the estimated effects of dispatch constraints.

Demand is permitted to participate directly in the wholesale market; however, nearly all end-users interface with the market through an electricity retailer. There are eight Frequency Control Ancillary Services (FCAS) markets to provide load following (raise and lower) and three contingency responses of different speed (raise and lower) between the 5-min energy dispatches. Market dispatch co-optimizes energy and FCAS bids and offers to establish regional prices for both energy and FCAS for every 5-min period. Commercial trading is based on these prices averaged over 30 min. Locational pricing within regions is achieved using averaged loss factors. Importantly, all generators are permitted to change their offers (rebid) just prior to every 5-min dispatch. Furthermore, the only commercially significant prices in the NEM are these averaged 30-min prices (the pre dispatch prices are advisory only). Also, note that NEM is an energy-only market and participants are required to manage their own unit commitment and other intertemporal scheduling challenges (within a range of technical dispatch constraints).

From the relation between electricity price and demand situations, it exhibits a time series spread over hourly intervals, which can be seen in a competitive electricity market. On the contrary, the price of electricity is a function of demand of electricity. The price of electricity depends on its present value t as well as electricity price's and demand's past values. It can be expressed as:

$$EP(t) = f\begin{pmatrix} EP(t-1), EP(t-2), EP(t-3), \ \dots \ , EP(t-NL_{EP}), \\ ED(t), ED(t-1), ED(t-2), ED(t-3), \ \dots \ , ED(t-NL_{ED}) \end{pmatrix} \tag{4.4}$$

where NL_{ED} is the electricity demand lag order, $EP\ (t)$ and $ED\ (t)$ are the price and demand of electricity at instantaneous time t assuming them as a time series and NL_{EP} is the lag order number for the price of electricity.

The hybrid ANFIS-BSA method is implemented for Queensland electricity price forecasting. From ((AEMO), 2019-04-06), the input ED and EP past data sets for the year of 2018 have been acquired. Historical data has variable range and Eq. (4.2) is used to normalize the independent and dependent variables. Normalizing specific data entails calibrating the data collected on distinct scales to an estimate common scale, usually applied before data processing.

For the purpose of electricity price forecasting in this study, only one week of exogenous variables ($NL_{EP} = NL_{ED} = 168$) with hourly lagged values are assumed, where, total 336 exogenous variables lagged values are available. Overburdening the machine learning algorithms with excess amount of features results in a sluggish learning process, rendering a deplorable performance and overfitting the training data. Therefore, only features significantly affecting the output (for electricity price forecasting process) should be assigned to machine learning algorithm.

Feature selection technique is of utmost concern for selecting the important input variables. By selecting a feature, the contributions of the final predictor variables (ED and EP in preceding hours) in the best hybrid (MOBBSA-ANFIS) model were evaluated. After developing and controlling several models with different combinations of input variables, these variables were identified for QLD market. A hybrid feature selection is applied in order to reduce the running time. In the first stage of hybrid feature selection, relevancy threshold of TH = 0.46 has been chosen for filtering the redundant features, and after filtering 69 relevant features are selected. In the second stage, hybrid MOBBSA-ANFIS is used in this study to select the input variables subsets, which have substantial impact on forecasting of electricity price of Australia market. In this stage, hybrid technique has been used to choose 27 dissimilar and most relevance features among the previously selected 69 candidates, which for the process of forecasting have been used as input. The input variables subsets chosen by two stages feature selection techniques are computed in table 4.9 and 4.10. The features are selected by (MI+ (MOBBSA-ANFIS)) for forecasting process is bolded in table 4.9.

Table 4.9: The optimal subsets of input variables selected for QLD by different multi-objective feature selection techniques and their corresponding performances in terms of RMSE value

MOBBSA+ ANFIS	MOPSO+ ANFIS	NSGAII+ ANFIS
EP(t-1)	EP(t-1)	EP(t-1)
EP(t-2)	EP(t-2)	EP(t-2)
EP(t-3)	EP(t-3)	EP(t-3)
EP(t-23)	EP(t-23)	EP(t-23)
EP(t-24)	EP(t-25)	EP(t-24)
EP(t-25)	EP(t-47)	EP(t-48)
EP(t-47)	EP(t-48)	EP(t-49)
EP(t-48)	EP(t-49)	EP(t-72)
EP(t-72)	EP(t-71)	EP(t-73)
EP(t-95)	EP(t-72)	EP(t-96)
EP(t-120)	EP(t-96)	EP(t-97)
EP(t-167)	EP(t-119)	EP(t-120)
EP(t-168)	EP(t-168)	EP(t-121)
EP(t-169)	EP(t-169)	EP(t-144)
EP(t-191)	EP(t-192)	EP(t-145)
EP(t-336)	EP(t-334)	EP(t-168)
EP(t-504)	EP(t-335)	EP(t-169)
ED(t)	EP(t-336)	EP(t-192)
ED(t-1)	ED(t)	EP(t-193)
ED(t-2)	ED(t-1)	ED(t)
ED(t-23)	ED(t-2)	ED(t-1)
ED(t-24)	ED(t-24)	ED(t-2)
ED(t-25)	ED(t-25)	ED(t-24)
ED(t-167)	ED(t-167)	ED(t-25)
ED(t-168)	ED(t-168)	ED(t-72)
ED(t-169)	ED(t-169)	ED(t-96)
ED(t-335)	ED(t-335)	ED(t-120)
	EP(t-336)	ED(t-144)
		ED(t-168)
		ED(t-336)
RMSE		
17.35	17.59	17.96

Table 4.10: The optimal subsets of input variables selected for QLD by different multi-objective feature selection techniques and their corresponding performances in terms of RMSE value

MOBBSA+ ANN	MOPSO+ ANN	NSGAII+ ANN
EP(t-1)	EP(t-1)	EP(t-1)
EP(t-2)	EP(t-2)	EP(t-2)
EP(t-3)	EP(t-3)	EP(t-6)
EP(t-23)	EP(t-71)	EP(t-23)
EP(t-24)	EP(t-73)	EP(t-24)
EP(t-25)	EP(t-96)	EP(t-48)
EP(t-48)	EP(t-97)	EP(t-120)
EP(t-49)	EP(t-120)	EP(t-144)
EP(t-73)	EP(t-121)	EP(t-168)
EP(t-94)	EP(t-144)	EP(t-169)
EP(t-120)	EP(t-145)	EP(t-192)
EP(t-167)	EP(t-167)	EP(t-193)
EP(t-168)	EP(t-168)	EP(t-335)
EP(t-169)	EP(t-169)	EP(t-337)
EP(t-191)	EP(t-335)	EP(t-503)
EP(t-192)	EP(t-336)	EP(t-504)
EP(t-336)	EP(t-337)	ED(t)
EP(t-504)	EP(t-504)	ED(t-1)
ED(t)	EP(t-505)	ED(t-4)
ED(t-1)	ED(t)	ED(t-12)
ED(t-2)	ED(t-1)	ED(t-24)
ED(t-24)	ED(t-3)	ED(t-25)
ED(t-25)	ED(t-24)	ED(t-48)
ED(t-72)	ED(t-72)	ED(t-72)
ED(t-96)	ED(t-96)	ED(t-96)
ED(t-120)	ED(t-120)	ED(t-120)
ED(t-144)	ED(t-144)	ED(t-144)
ED(t-168)	ED(t-168)	ED(t-168)
ED(t-336)	EP(t-192)	ED(t-335)
	ED(t-335)	ED(t-504)
RMSE		
18.70	18.87	18.94

4.5 Numerical results

In this chapter, the electricity price forecasting (EPF) accuracy of Queensland mainland is forecasted by employing ANFIS-BSA, which is known as the one of the most volatile electricity market. Significant features are determined by developing a feature selection (MI+ (MOBBSA-ANFIS)) as input for forecasting analysis in this section. Moreover, to evaluate the usefulness of ANFIS-BSA for short-term EPF accuracy, the proposed method is compared with well-known AI- techniques that include ANN, ANFIS, ANN-GA, ANN-PSO, ANN-BSA, ANFIS-GA, and ANFIS-PSO.

4.5.1 Validation of the Model Using Machine Learning Methods in QLD market

The performance results of different machine learning methods are tabulated in Table 4.11 for EPF for the same area in summer (February 2018). The RACF values, like in summer season, declare the whiteness of estimated residuals at a confidence interval level for all obtained models. The forecasting accuracy of these methods is ranked as ANFIS-BSA > ANFIS- PSO > ANFIS- GA > MLP-BSA> ANFIS> MLP-GA > MLP-PSO >MLP in terms of multi-criteria decisions, which the mean rank of the methods for each indicator (U-statistic, MAPE, RMSE, and absolute error) is used to extract this ranking. To be more precise, U-statistic, MAPE, RMSE and absolute error are 0.04, 3.43%, 0.08 and 22.99 respectively for ANFIS-BSA. Additionally, ANFIS-BSA can be concluded as the most promising and advanced model than any other applied methods used for Queensland's short-term electricity price forecasting. The performance analysis of ANFIS-BSA for EPF in summer 2018 is depicted in Figure 4.5 during both training and testing design phases for Queensland region.

Table 4.11: Comparison between forecasting accuracy of studied methods for EPF of Queensland in February2018

Performance Indexes		Methods							
		MLP	MLP-GA	MLP-PSO	MLP-BSA	ANFIS	ANFIS-GA	ANFIS-PSO	ANFIS-BSA
RACF	Training	0.0012	0.0005	0.0025	0.0020	0.0027	0.0006	0.0009	0.0002
	Testing	0.0014	0.0017	0.0006	0.0005	0.0014	0.0018	0.0022	0.0004
	Whole set	0.0016	0.0010	0.0034	0.0040	0.0020	0.0007	0.0003	0.0013
Absolute error	Training	19.2058	18.9636	18.9975	18.6235	18.8524	18.5326	18.1478	**17.7108**
	Testing	7.1034	6.5628	6.6024	6.2468	6.3832	6.2280	5.4871	**5.2836**
	Whole set	26.3092	25.5264	25.5999	24.8703	25.2356	24.7606	23.6349	**22.9944**
RMSE	Training	0.1245	0.0925	0.0973	0.0861	0.0882	0.0846	0.0798	**0.0772**
	Testing	0.0982	0.0838	0.0844	0.0803	0.0817	0.0789	0.0742	**0.0724**
	Whole set	0.1106	0.0873	0.0882	0.0838	0.0846	0.0827	0.0780	**0.0752**
U-statistic	Training	0.0525	0.0498	0.0510	0.0462	0.0477	0.0443	0.0397	**0.0367**
	Testing	0.0489	0.0442	0.0453	0.0417	0.0415	0.0402	0.0357	**0.0321**
	Whole set	0.0518	0.0481	0.0492	0.0443	0.0458	0.0406	0.0381	**0.0353**
MAPE (%)	Training	5.0563	4.8523	4.9322	4.5721	4.6237	4.4172	3.9823	**3.8542**
	Testing	4.5624	4.3411	4.4503	3.9846	4.0798	3.7671	3.2207	**3.0785**
	Whole set	4.9247	4.6408	4.7126	4.3587	4.4424	4.1031	3.6412	**3.4309**

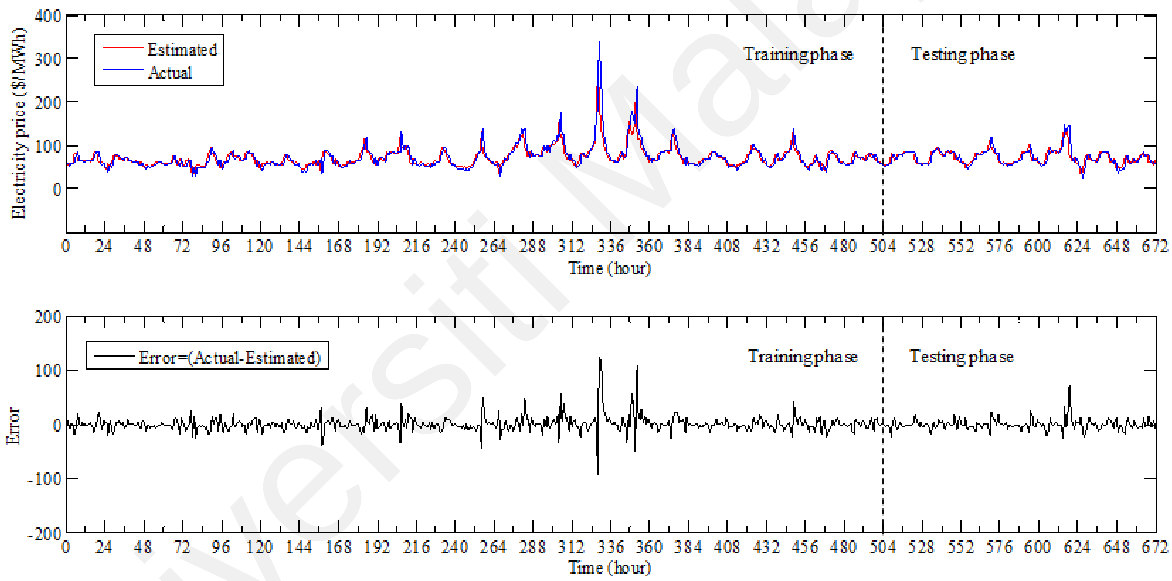

Figure 4.5: The performance of ANFIS-BSA during training of design phase and testing phase and its corresponding error for EPF of Queensland in February2018

The ANFIS-BSA performance for EPF of Queensland region in autumn season of May 2018 is compared with other AI- methods for the purpose of further examination of solution methodology as shown in Table 4.12. From the values obtained for RACF, it can be shown that the estimated residuals of all models are uncorrelated and the obtained models sufficiently describe the given set of data. Based on Table 4.12, the forecasting accuracy of the applied methods is ranked as ANFIS-BSA> ANFIS- GA > ANFIS- PSO > MLP-BSA > ANFIS > MLP-PSO > MLP-GA > MLP. By using the mean rank of multi-criteria decisions methods, the ranking for each indicator (absolute error, RMSE, U-statistic and MAPE) in the whole set has been performed. According to the findings, from the comparison of studied methods for electricity market in Queensland region in autumn season, ANFIS-BSA approach performs much better than other methods in terms of electricity forecasting since it has higher precision.

The values of U-statistic (0.04), MAPE (4.36%), RMSE (0.08), and absolute error (29.85) are presented in Table 4.12. Figure 4.6 depicts the ANFIS-BSA performance for forecasting the electricity price in autumn season of 2018 during the training and testing phase for Queensland region.

Table 4.12: Comparison between forecasting accuracy of studied methods for EPF of Queensland in May2018

Performance Indexes		Methods							
		MLP	MLP-GA	MLP-PSO	MLP-BSA	ANFIS	ANFIS-GA	ANFIS-PSO	ANFIS-BSA
RACF	Training	0.0011	0.0013	0.0028	0.0146	0.0046	0.0112	0.0045	0.0014
	Testing	0.0142	0.0002	0.0016	0.0018	0.0026	0.0035	0.0003	0.0038
	Whole set	0.0002	0.0006	0.0008	0.0022	0.0005	0.0006	0.0008	0.0020
Absolute error	Training	21.8950	21.7156	21.6982	21.6532	21.4425	20.9273	21.0214	**20.5412**
	Testing	12.2236	12.2134	12.2033	11.8245	12.1437	11.6578	12.0056	**11.3045**
	Whole set	34.1186	33.9290	33.9015	33.4777	33.5862	32.5851	33.0270	**29.8457**
RMSE	Training	0.1512	0.1314	0.1304	0.1233	0.1178	0.0918	0.0923	**0.0863**
	Testing	0.0965	0.0947	0.0935	0.0885	0.0921	0.0874	0.0898	**0.0847**
	Whole set	0.0991	0.0963	0.0956	0.0923	0.0942	0.0896	0.0907	**0.0852**
U-statistic	Training	0.0602	0.0568	0.0556	0.0531	0.0526	0.0487	0.0494	**0.0462**
	Testing	0.0563	0.0514	0.0503	0.0476	0.0491	0.0466	0.0465	**0.0437**
	Whole set	0.0599	0.0521	0.0517	0.0487	0.0509	0.0469	0.0472	**0.0443**
MAPE (%)	Training	4.9784	4.8804	4.8657	4.8563	4.8245	4.6281	4.7841	**4.5721**
	Testing	4.5324	4.4612	4.4569	4.2832	4.4423	4.0578	4.2730	**3.9846**
	Whole set	4.9562	4.8235	4.8052	4.5656	4.7136	4.4855	4.5637	**4.3587**

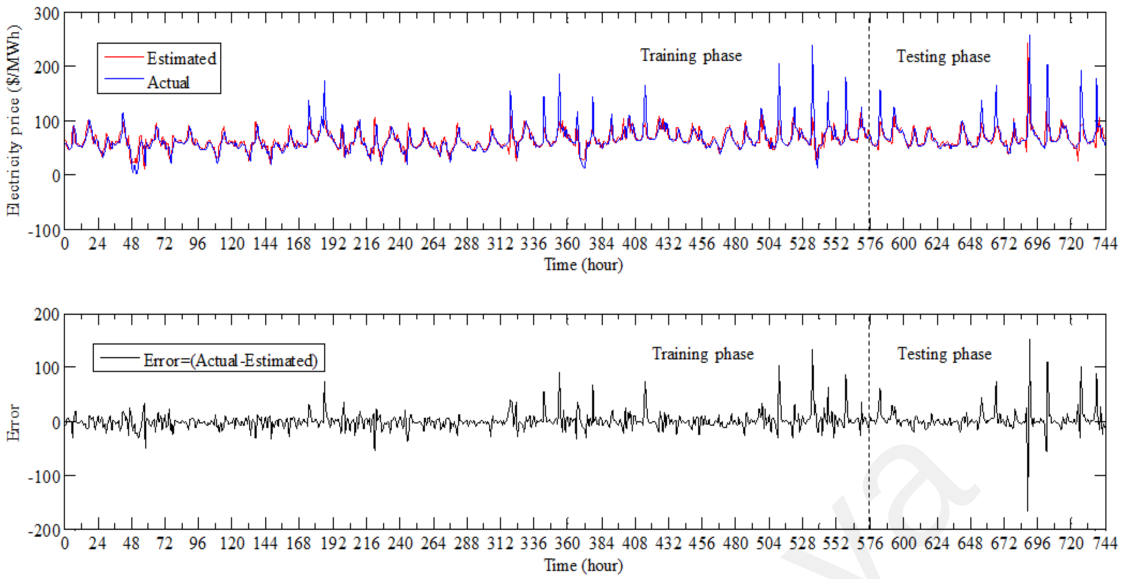

Figure 4.6: The performance of ANFIS-BSA during training of design phase and testing phase and its corresponding error for EPF of Queensland in May2018

The machine learning methods performances in winter season (August) of 2018 for Queensland EPF is presented in Table 4.13. From the results tabulated in Table 4.13, it can be said that the whiteness of the estimated residuals for all developed models has been validated by the calculated *RACF* values, which are in an affirmed confidence range. Moreover, all developed models are able to describe the given set of data sufficiently. The analysis that has been done on Queensland electricity market regarding the evaluation of forecasting accuracy of methods concludes that according to multi-criteria decisions using the mean rank of the methods, each indicator (absolute error, *RMSE*, *U*-statistic and *MAPE*) is ranked as ANFIS-BSA > ANFIS- PSO > ANFIS- GA > MLP-BSA > MLP-GA > MLP-PSO > ANFIS > MLP. The comparison of the developed models with the existing similar models concludes that ANFIS-BSA approach performs exceptionally better. Table 4.13 also shows that the ANFIS-BSA based method is superior in terms of *MAPE* = 5.41 %, *U*-statistic = 0.07, *RMSE* = 0.17 and absolute error = 40.58. Figure 4.7 presented the performance test of ANFIS-BSA method in winter season of 2018 when the training is executed for both testing and design phase.

Table 4.13: Comparison between forecasting accuracy of studied methods for EPF of Queensland in August 2018

Performance Indexes		MLP	MLP-GA	MLP-PSO	MLP-BSA	ANFIS	ANFIS-GA	ANFIS-PSO	ANFIS-BSA
		Methods							
RACF	Training	0.0015	0.0024	0.0019	0.0007	0.0004	0.0009	0.0029	0.0001
	Testing	0.0016	0.0005	0.0014	0.0025	0.0010	0.0011	0.0027	0.0014
	Whole set	0.0012	0.0008	0.0015	0.0020	0.0024	0.0010	0.0018	0.0013
Absolute error	Training	25.7567	25.4287	25.5600	25.3459	25.6895	25.3214	24.2596	**24.0278**
	Testing	18.2863	17.2110	18.2365	17.1237	17.2466	17.0003	16.6582	**16.5542**
	Whole set	44.0430	42.6397	43.7965	42.4696	42.9361	42.3217	40.9178	**40.5820**
RMSE	Training	0.2550	0.2138	0.2246	0.2001	0.2489	0.1952	0.1826	**0.1789**
	Testing	0.1948	0.1824	0.1934	0.1763	0.1871	0.1715	0.1604	**0.1563**
	Whole set	0.2134	0.1937	0.1976	0.1856	0.1994	0.1833	0.1765	**0.1684**
***U*-statistic**	Training	0.0935	0.0876	0.0902	0.0849	0.0914	0.0824	0.0798	**0.0785**
	Testing	0.0902	0.0800	0.0883	0.0784	0.0856	0.0767	0.0723	**0.0701**
	Whole set	0.0912	0.0834	0.0895	0.0810	0.0905	0.0789	0.0756	**0.0742**
MAPE (%)	Training	6.2376	5.9456	6.1014	5.9127	6.1878	5.8455	5.7745	**5.6542**
	Testing	6.0024	5.7524	5.9896	5.6524	5.8233	5.6047	5.4268	**5.3125**
	Whole set	6.2077	5.8330	6.0001	5.7951	6.1268	5.6493	5.5096	**5.4108**

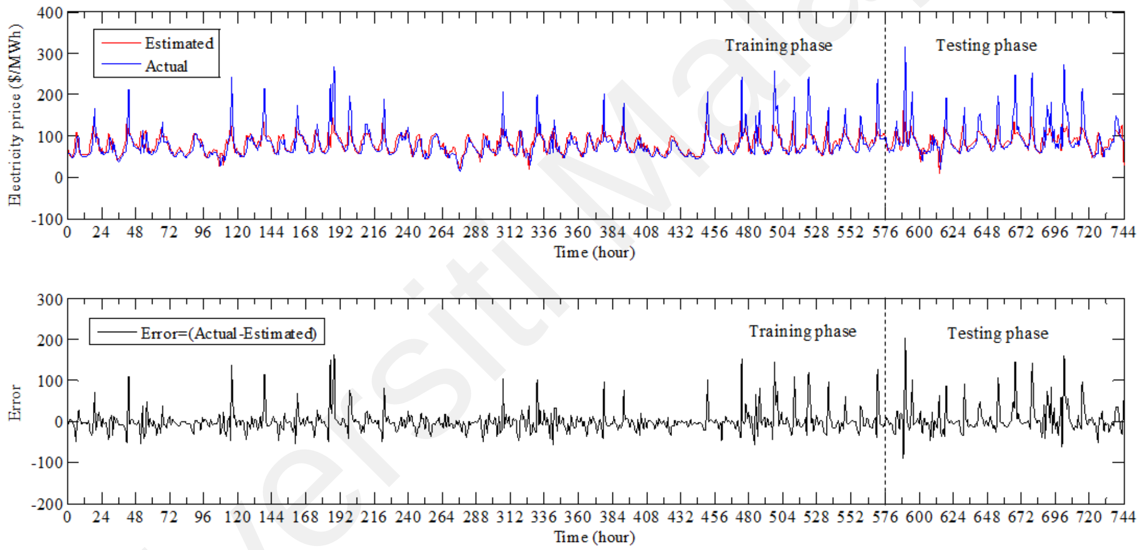

Figure 4.7: The performance of ANFIS-BSA during training of design phase and testing phase and its corresponding error for EPF of Queensland in August 2018

Table 4.14shows a tabulated performance of the applied machine learning techniques in Queensland region for EPF in spring season (November) of 2018. According to the *RACF* values, the conclusion can be drawn that the estimated residuals of all obtained models are white at a confidence interval level. The accuracy of the methods is ranked as ANFIS-BSA > ANFIS- PSO > ANFIS- GA > MLP-BSA > MLP-GA > MLP-PSO > ANFIS > MLP, which is based on multi-criteria decisions adopting the mean rank of the methods for each indicator (*U*-statistic (0.05), *MAPE* (4.46%), *RMSE* (0.09), and absolute error (32.12)). Hence, the most efficient model is ANFIS-BSA. Figure 4.8 illustrates the performance of ANFIS-BSA in spring season for the same region.

Table 4.14: Comparison between forecasting accuracy of studied methods for EPF of Queensland in November 2018

Performance Indexes		Methods							
		MLP	**MLP-GA**	**MLP-PSO**	**MLP-BSA**	**ANFIS**	**ANFIS-GA**	**ANFIS-PSO**	**ANFIS-BSA**
RACF	Training	0.0007	0.0008	0.0014	2.50E-05	0.0002	0.0015	0.0007	0.0019
	Testing	0.0004	3.20E-06	0.0012	0.0006	0.0005	0.0014	0.0011	0.0005
	Whole set	0.0009	0.0009	1.70E-06	0.0004	0.0001	0.0003	0.0006	0.0010
Absolute error	Training	26.6740	26.0021	26.2140	25.3145	26.5412	25.2364	25.3011	**22.1123**
	Testing	12.7713	12.1456	12.3214	12.0138	12.5677	11.9823	10.5673	**10.0032**
	Whole set	39.4453	38.1477	38.5354	37.3283	39.1089	37.2187	35.8684	**32.1155**
RMSE	Training	0.1409	0.1021	0.1143	0.0933	0.1325	0.0912	0.0924	**0.0876**
	Testing	0.1217	0.0985	0.1006	0.0928	0.1105	0.0886	0.0857	**0.0825**
	Whole set	0.1324	0.0965	0.1024	0.0925	0.1243	0.0897	0.0884	**0.0861**
***U*-statistic**	Training	0.0678	0.0621	0.0643	0.0556	0.0654	0.0523	0.0537	**0.0476**
	Testing	0.5823	0.0503	0.0525	0.0484	0.0538	0.0477	0.0449	**0.0415**
	Whole set	0.0651	0.0600	0.0617	0.0521	0.0622	0.0518	0.0482	**0.0450**
MAPE (%)	Training	5.2347	5.0031	5.1314	4.7810	5.1884	4.6734	4.7126	**4.5862**
	Testing	4.7624	4.4575	4.5710	4.0036	4.6217	3.9620	3.8755	**3.8023**
	Whole set	5.1975	4.9562	5.0078	4.6123	5.1059	4.6021	4.5237	**4.4654**

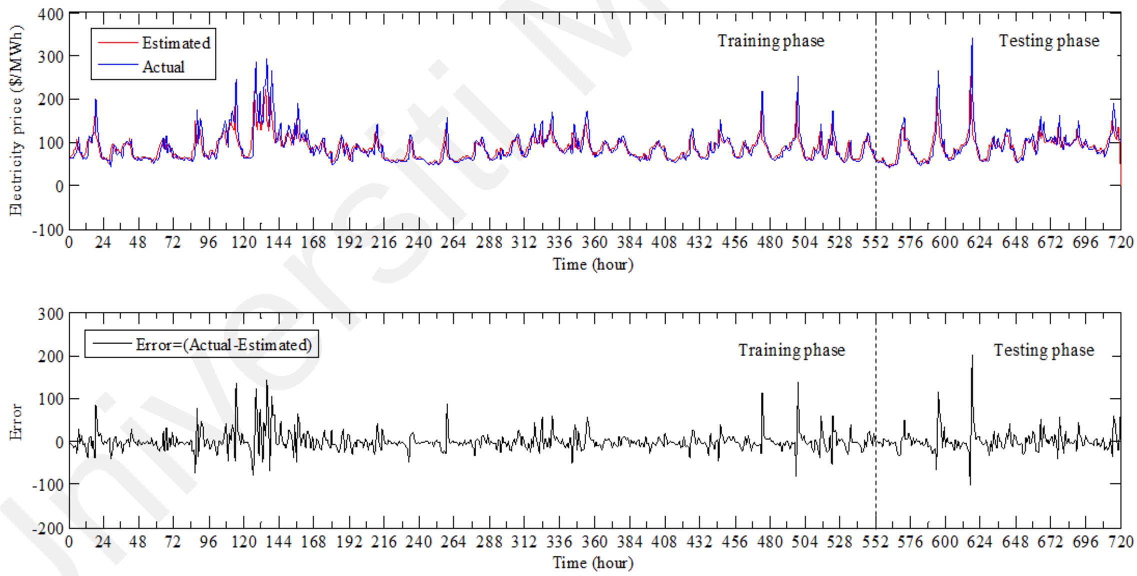

Figure 4.8: The performance of ANFIS-BSA during training of design phase and testing phase and its corresponding error for EPF of Queensland in November 2018

Since the utilities provide incentives in their EP, the presence of negative price is certain. However, the incentives are provided to adopt the frequent change of power market with the power station and to create new business opportunity.

4.5.2 Validation of the Model Using Statistical Methods in QLD market

Different statistical methods are also applied as external validation to verify the validity of models developed by ANFIS-BSA. To evaluate the performance of the obtained model the following attributes were recommended (Mostafavi et al., 2014):

iv- If a model generates $|R| > 0.8$, a strong correlation exists between the predicted and observed values.

v- If a model generates $0.2 < |R| < 0.8$, a correlation exists between the predicted and observed values.

vi- If a model generates $|R| < 0.2$, a weak correlation exists between the predicted and observed values.

Table 4.15 tabulate the statistical factors of the ANFI-BSA model for EPF of Queensland in February, May, August and November of different seasons. As shown in this table, the developed models satisfy all the requisite conditions. The validation phase ensures that ANFIS-BSA provides precise models, which is strongly applicable for short term electricity price forecasting of Queensland.

Table 4.15: Statistical factors of the ANFIS-BSA model for EPF of QLD in February (F), May (M), August (A) and November (N) 2018

Item	Formula	Condition	F	M	A	N
1	R	$0.8<R_0$	0.9934	0.9970	0.9960	0.9962
2	$K=\frac{\sum_{i=1}^{n}(h_i\times t_i)}{\sum_{i=1}^{n}h_i^2}$	$0.85<k<1.15$	0.9980	1.0014	0.9935	0.9954
3	$K'=\frac{\sum_{i=1}^{n}(h_i\times t_i)}{\sum_{i=1}^{n}t_i^2}$	$0.85<k'<1.15$	1.0015	0.9983	1.0060	1.0042
4	$m=\frac{R^2-R_o^2}{R^2}$	$\lvert m\rvert<0.1$	-0.0119	-0.0056	-0.0053	-0.0061
5	$n=\frac{R^2-R_{o'}^2}{R^2}$	$\lvert n\rvert<0.1$	-0.0121	-0.0056	-0.0054	-0.0063
6	$R_m=R^2\times\left(1-\sqrt{\lvert R^2-R_o^2\rvert}\right)$	$0.5<R_m$	0.9764	0.9892	0.9873	0.9867
Where	$R_o^2=1-\frac{\sum_{i=1}^{n}(t_i-h_i^o)^2}{\sum_{i=1}^{n}\left(t_i-\bar{t}_i\right)^2},\ h_i^o=k\times t_i$	$0.8<R_0^2<1$	1.0000	0.9999	0.9992	0.9997
	$R_{o'}^2=1-\frac{\sum_{i=1}^{n}(h_i-t_i^o)^2}{\sum_{i=1}^{n}\left(h_i-\bar{h}_i\right)^2},\ t_i^o=K'\times h_i$	$0.8<R_{0'}^2<1$	1.0000	0.9999	0.9993	0.9998

4.6 Comparison of Different Forecasting Methods based on MAPE

Figure 4.9 and figure 4.10 illustrate the performance of AI based methods for EPF of Ontario and Queensland respectively based on MAPE (%).The proposed hybrid ANFIS-BSA forecasting approach provided a higher forecasting accuracy of EPF of Ontario in figure 4.9 in terms of, MAPE = 2.79%, 0.87%, 1.7% and 3.17% in February, May, August and November respectively compared to different artificial intelligence (AI) models. In addition, the robustness of the developed ANFIS-BSA model in Queensland electricity market compare to other AI methods in terms of MAPE 3.43%, 4.36%, 5.41% and 4.46% in February, May, August and November respectively is shown in figure 4.10.

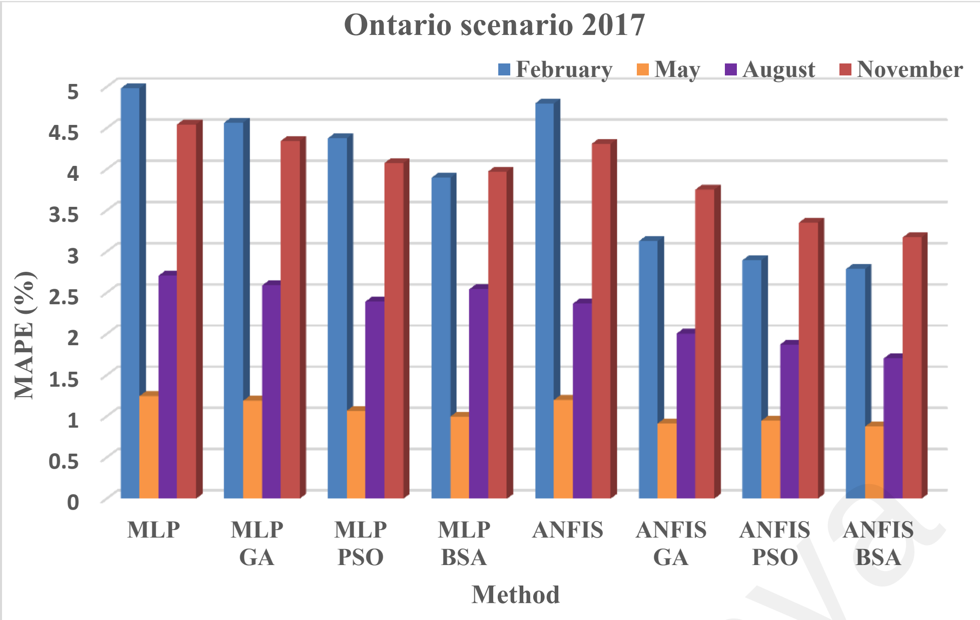

Figure 4.9: The performance of AI –based method for EPF of Ontario based on MAPE (%) in 2017

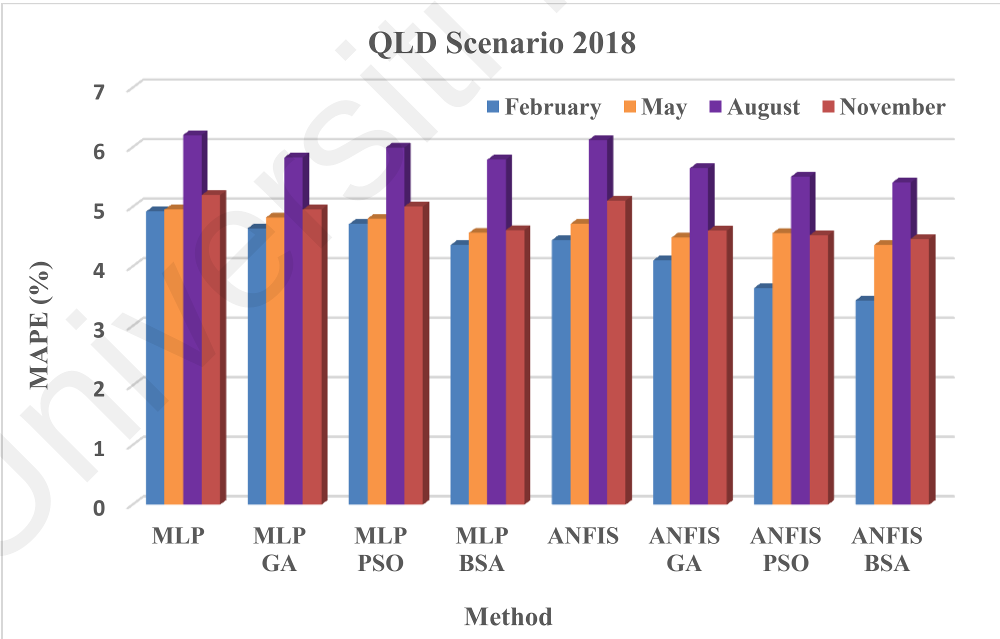

Figure 4.10: The performance of AI-based method for EPF of Queensland based on MAPE (%) in 2018

4.7 Summary

Companies that are trading in electricity markets make extensive use of price forecast techniques either to bid or hedge against volatility. However, price prediction has its own complexities. At the same time, utility companies usually have limited and uncertain information for price prediction. Thus, there is essential need to an efficient and robust forecasting method for the companies that trade in the electricity markets. In this chapter, the proposed hybrid forecast and feature selection strategy has been examined for day-ahead price forecast of Canada and Australia deregulated electricity markets. Our comprehensive experiments, based on MAPE and error variance, not only reveal the forecast capability of the proposed strategy but also demonstrate that the forecast accuracy can be significantly improved based on the proposed feature selection algorithm. The application of the proposed approach in electricity prices forecasting on the Ontario and Queensland markets is both novel and effective.

From the obtained results, it has been seen that the number of selected features for Ontario mainland is 23 and RMSE value is 12.23 for the proposed MOBBSA feature selection method which are less than the features number (26, 29, 26, 29 & 32) and RMSE values (13.05, 13.87, 13.47, 13.85 & 14.74) obtained by using other multi-objective approaches namely MOPSO + ANFIS, NSGA-II + ANFIS, MOBAS + AN, MOPSO + ANN and NSGA-II + ANN respectively. The proposed hybrid BSA and ANFIS forecasting approach provided a higher forecasting accuracy with the least complexity for electricity price forecasting in term of, MAPE = 2.79%, 0.87%, 1.7% and 3.17% in February, May, August and November respectively compared to other artificial intelligence (AI) models for Canadian market.

The results have shown the robustness of the developed ANFIS-BSA model in Queensland electricity market. In the case of electricity price forecasting, it provides a higher forecasting precision and simplicity compared to other AI methods in terms of MAPE 3.43%, 4.36%, 5.41% and 4.46%in February, May, August and November respectively. Besides, in this market, by developing the feature selection technique for the selection of optimum subset of features within a pool of 69 features, only 27 features to be employed as input for direct prediction method. The robustness of the proposed technique is evident through efficient selection of the most suitable features by removal irrelevant and redundant attributes. In developing the sustainable smart grid in forthcoming days, the importance of the proposed approach is inevitable to EP forecasting. Therefore, the presented synthesis can be a profound contrivance to develop energy strategies for the electricity participants in bidding as well as for the EP forecasting researchers plainly.

CHAPTER 5: CONCLUSIONS AND FUTURE WORKS

5.1 Conclusions

This chapter fills the gaps and supplements existing research in the literature, concerning the evolving issue of the day ahead electricity price forecasting. Thus, a hybrid price forecast strategy composed of a multi-objective feature selection technique and hybrid forecast engine has been proposed for day ahead prediction of prices in electricity markets.

To simplify the learning process of a forecasting model and the interpretation of a dataset, feature selection is an essential stage in day ahead electricity price forecasting. To address the feature selection problem, the multi-objective feature selection technique has been successfully proposed. The proposed feature selection approach is composed of a multi-objective binary-valued backtracking search algorithm (MOBBSA) and adaptive neuro-fuzzy inference system (ANFIS) approach. In order to form a reduced input subset, firstly the mutual information of input features is measured and the irrelevant and redundant features are filtered. Then, the proposed feature selection technique is applied to the reduced subset to find a small set of features with high forecasting accuracy.

The proposed feature selection methodology not only evaluates the non-linear dependencies of the price signal on its input variables better than the other well-known multi-objective optimization methods e.g. NSGAII and MOPSO, but also selects those relevant features that are mutually dissimilar by removing redundant features. The proposed MOBBSA-ANFIS feature selection approach can identify less input variables and provides higher forecasting accuracy. The proposed strategy has an information theoretic feature selection technique with the ability of selecting a minimum subset of the most relevant and non-redundant features among a large set of candidate inputs.

Thus, the finally selected candidate inputs have minimum redundancy with maximum relevance with the output. For example, number of selected features (Ontario) is 23 and RMSE value is 12.23 for proposed MOBBSA feature selection method which are less than features number (26 & 29) and RMSE values (13.05 & 13.87) obtained by using other multi-objective approaches namely MOPSO and NSGAII respectively.

Finally, to provide more accurate forecast engine for electricity price forecasting, backtracking search algorithm (BSA) as an efficient optimization algorithm has been successfully implemented in the learning process of ANFIS approach. To assess the applicability and accuracy of the proposed method for electricity price forecasting, its estimates are compared with those obtained from artificial neural network (ANN), ANFIS, ANN and ANFIS models optimized by particle swarm optimization (PSO) and genetic algorithm (GA). The simulation results are validated by actual data sets observed from Ontario electricity market; which is reported as among the most volatile electricity market worldwide. The results confirm the higher accuracy and reliability of the proposed ANFIS-BSA method as compared with other artificial intelligence (AI) based models. On the basis of the favorable results obtained, it can be concluded that recent enhancements in AI-based approaches, as in this study, could result higher accuracy with the least complexity for electricity price forecasting. For instance ANFIS+BSA method provided better performance for short-term electricity price forecasting of Ontario in term of, MAPE = 2.79%, 0.87%, 1.7% and 3.17% in February, May, August and November respectively compared to other AI models.

The AFNIS approach optimized by BSA provides evidence that it could be considered as robust and useful forecast engine to the actual needs of electricity market participants, including the self-producers and traditional generation companies, suppliers/retailers and aggregators. These contributions may help market participants to bid effectively, maintaining efficient daily operation, and eventually increasing company's profit.

5.2 Future Works

The following tasks can be carried out as the future works.

(a) Considering spike forecasting in EPF

The proposed method provides significant improvement in both normal price and price spike prediction accuracy compared to other forecast techniques applied for case studies of energy markets. However, electricity price is a complex volatile signal owning many spikes. Most of the electricity price forecast techniques focus on the normal price prediction while price spike forecast is a different prediction process. For future work, the proposed method can be extended to predict price spike occurrence.

(b) Adding different features in EPF

In this work, the analysis uses only based on the data that is publicly available. Therefore, the proposed EPF considered only demands and prices in different time interval. Other features like total market demand (TMD), imports and exports can also be considered in the future work of the EPF. As well as short-term electricity price forecasting technologies considering feed-in tariffs, carbon pricing and other policy drives to accelerate renewable energy development can also be conducted.

REFERENCES

(IESO), T. I. E. S. O. (2018-11-17). http://www.ieso.ca/.

(AEMO), N. E. M. (2019-04-06). https://www.aemo.com.au/Electricity/National-Electricity-Market-NEM/Data-dashboard#aggregated-data

Abedinia, O., & Amjady, N. (2016). Short‐term load forecast of electrical power system by radial basis function neural network and new stochastic search algorithm. *International transactions on electrical energy systems, 26*(7), 1511-1525.

Abedinia, O., Amjady, N., Shafie-Khah, M., & Catalão, J. (2015). Electricity price forecast using combinatorial neural network trained by a new stochastic search method. *Energy Conversion and Management, 105*, 642-654.

Aggarwal, S. K., Saini, L. M., & Kumar, A. (2008). Electricity price forecasting in Ontario electricity market using wavelet transform in artificial neural network based model. *International Journal of Control, Automation, and Systems, 6*(5), 639-650.

Aggarwal, S. K., Saini, L. M., & Kumar, A. (2009). Electricity price forecasting in deregulated markets: A review and evaluation. *International Journal of Electrical Power & Energy Systems, 31*(1), 13-22.

Ahmed, T., Mekhilef, S., Shah, R., Mithulananthan, N., Seyedmahmoudian, M., & Horan, B. (2017). ASEAN power grid: A secure transmission infrastructure for clean and sustainable energy for South-East Asia. *Renewable and Sustainable Energy Reviews, 67*, 1420-1435.

Alamaniotis, M., Gatsis, N., & Tsoukalas, L. H. (2018). Virtual Budget: Integration of electricity load and price anticipation for load morphing in price-directed energy utilization. *Electric power systems research, 158*, 284-296.

Alavi, A. H., Aminian, P., Gandomi, A. H., & Esmaeili, M. A. (2011). Genetic-based modeling of uplift capacity of suction caissons. *Expert Systems with Applications, 38*(10), 12608-12618.

Alizadeh, M., Jolai, F., Aminnayeri, M., & Rada, R. (2012). Comparison of different input selection algorithms in neuro-fuzzy modeling. *Expert Systems with Applications, 39*(1), 1536-1544.

AlRashidi, M., & El-Naggar, K. (2010). Long term electric load forecasting based on particle swarm optimization. *Applied Energy, 87*(1), 320-326.

Amjady, N., & Daraeepour, A. (2009a). Design of input vector for day-ahead price forecasting of electricity markets. *Expert Systems with Applications, 36*(10), 12281-12294.

Amjady, N., & Daraeepour, A. (2009b). Mixed price and load forecasting of electricity markets by a new iterative prediction method. *Electric power systems research, 79*(9), 1329-1336.

Amjady, N., Daraeepour, A., & Keynia, F. (2010). Day-ahead electricity price forecasting by modified relief algorithm and hybrid neural network. *IET generation, transmission & distribution, 4*(3), 432-444.

Amjady, N., & Hemmati, M. (2006). Energy price forecasting-problems and proposals for such predictions. *IEEE Power and Energy Magazine, 4*(2), 20-29.

Amjady, N., & Hemmati, M. (2009). Day‐ahead price forecasting of electricity markets by a hybrid intelligent system. *International Transactions on Electrical Energy Systems, 19*(1), 89-102.

Amjady, N., & Keynia, F. (2008). Day ahead price forecasting of electricity markets by a mixed data model and hybrid forecast method. *International Journal of Electrical Power & Energy Systems, 30*(9), 533-546.

Amjady, N., & Keynia, F. (2009a). Day-ahead price forecasting of electricity markets by a new feature selection algorithm and cascaded neural network technique. *Energy Conversion and Management, 50*(12), 2976-2982.

Amjady, N., & Keynia, F. (2009b). Day-ahead price forecasting of electricity markets by mutual information technique and cascaded neuro-evolutionary algorithm. *IEEE Transactions on Power Systems, 24*(1), 306-318.

Amjady, N., & Keynia, F. (2010). Application of a new hybrid neuro-evolutionary system for day-ahead price forecasting of electricity markets. *Applied Soft Computing, 10*(3), 784-792.

Anbazhagan, S., & Kumarappan, N. (2012). Day-ahead deregulated electricity market price classification using neural network input featured by DCT. *International Journal of Electrical Power & Energy Systems, 37*(1), 103-109.

Anbazhagan, S., & Kumarappan, N. (2013). *Classification of Day-Ahead Deregulated Electricity Market Prices Using DCT-CFNN.* Paper presented at the International Conference on Swarm, Evolutionary, and Memetic Computing.

Anbazhagan, S., & Kumarappan, N. (2014). Day-ahead deregulated electricity market price forecasting using neural network input featured by DCT. *Energy Conversion and Management, 78*, 711-719.

Askarzadeh, A. (2014). Comparison of particle swarm optimization and other metaheuristics on electricity demand estimation: a case study of Iran. *Energy, 72*, 484-491.

Azadeh, A., Saberi, M., Gitiforouz, A., & Saberi, Z. (2009). A hybrid simulation-adaptive network based fuzzy inference system for improvement of electricity consumption estimation. *Expert Systems with Applications, 36*(8), 11108-11117.

Azadeh, A., Saberi, M., & Seraj, O. (2010). An integrated fuzzy regression algorithm for energy consumption estimation with non-stationary data: a case study of Iran. *Energy, 35*(6), 2351-2366.

Bask, M., & Widerberg, A. (2009). Market structure and the stability and volatility of electricity prices. *Energy Economics, 31*(2), 278-288.

Baziar, A., & Kavousi-Fard, A. (2015). Short term load forecasting using a hybrid model based on support vector regression. *International journal of scientific & technology research, 4*(5), 189-195.

Benth, F. E., Kiesel, R., & Nazarova, A. (2012). A critical empirical study of three electricity spot price models. *Energy Economics, 34*(5), 1589-1616.

Bian, X.-Q., Han, B., Du, Z.-M., Jaubert, J.-N., & Li, M.-J. (2016). Integrating support vector regression with genetic algorithm for CO2-oil minimum miscibility pressure (MMP) in pure and impure CO2 streams. *Fuel, 182*, 550-557.

Bigdeli, N., Afshar, K., & Amjady, N. (2009). Market data analysis and short-term price forecasting in the Iran electricity market with pay-as-bid payment mechanism. *Electric power systems research, 79*(6), 888-898.

Catalão, J., Mariano, S., Mendes, V., & Ferreira, L. (2007). *An artificial neural network approach for short-term electricity prices forecasting.* Paper presented at the Intelligent Systems Applications to Power Systems, 2007. ISAP 2007. International Conference on.

Catalão, J., Pousinho, H., & Mendes, V. (2009). *Neural networks and wavelet transform for short-term electricity prices forecasting.* Paper presented at the Intelligent System Applications to Power Systems, 2009. ISAP'09. 15th International Conference on.

Catalão, J. P. d. S., Mariano, S. J. P. S., Mendes, V., & Ferreira, L. (2007). Short-term electricity prices forecasting in a competitive market: A neural network approach. *Electric power systems research, 77*(10), 1297-1304.

Catalão, J. P. d. S., Pousinho, H. M. I., & Mendes, V. M. F. (2011). Hybrid wavelet-PSO-ANFIS approach for short-term electricity prices forecasting. *IEEE Transactions on Power Systems, 26*(1), 137-144.

Cerjan, M., Krželj, I., Vidak, M., & Delimar, M. (2013). *A literature review with statistical analysis of electricity price forecasting methods.* Paper presented at the EUROCON, 2013 IEEE.

Chang, W.-Y. (2015). Short-term load forecasting using radial basis function neural network. *Journal of Computer and Communications, 3*(11), 40.

Chogumaira, E. N., & Hiyama, T. (2011). Short-term electricity price forecasting using a combination of neural networks and fuzzy inference. *Energy and Power Engineering, 3*(01), 9.

Civicioglu, P. (2013). Backtracking Search Optimization Algorithm for numerical optimization problems. *Applied Mathematics and Computation, 219*(15), 8121-8144.

Coello, C. C. (2006). Evolutionary multi-objective optimization: a historical view of the field. *IEEE computational intelligence magazine, 1*(1), 28-36.

Conejo, A. J., Contreras, J., Espinola, R., & Plazas, M. A. (2005). Forecasting electricity prices for a day-ahead pool-based electric energy market. *International journal of forecasting, 21*(3), 435-462.

Cover, T. M., & Thomas, J. A. (2006). Elements of information theory 2nd edition. *Willey-Interscience: NJ.*

Cuaresma, J. C., Hlouskova, J., Kossmeier, S., & Obersteiner, M. (2004). Forecasting electricity spot-prices using linear univariate time-series models. *Applied Energy, 77*(1), 87-106.

Deb, K., Pratap, A., Agarwal, S., & Meyarivan, T. (2002). A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE transactions on evolutionary computation, 6*(2), 182-197.

Diongue, A. K., Guegan, D., & Vignal, B. (2009). Forecasting electricity spot market prices with a k-factor GIGARCH process. *Applied Energy, 86*(4), 505-510.

dos Santos Coelho, L., & Santos, A. A. (2011). A RBF neural network model with GARCH errors: application to electricity price forecasting. *Electric power systems research, 81*(1), 74-83.

Ebrahimian, H., Barmayoon, S., Mohammadi, M., & Ghadimi, N. (2018). The price prediction for the energy market based on a new method. *Economic research-Ekonomska istraživanja, 31*(1), 313-337.

Ekonomou, L. (2010). Greek long-term energy consumption prediction using artificial neural networks. *Energy, 35*(2), 512-517.

Eydeland, A., & Wolyniec, K. (2003). *Energy and power risk management: New developments in modeling, pricing, and hedging* (Vol. 206): John Wiley & Sons.

Gao, W., Darvishan, A., Toghani, M., Mohammadi, M., Abedinia, O., & Ghadimi, N. (2019). Different states of multi-block based forecast engine for price and load prediction. *International Journal of Electrical Power & Energy Systems, 104*, 423-435.

Golbraikh, A., & Tropsha, A. (2002). Beware of q 2! *Journal of molecular graphics and modelling, 20*(4), 269-276.

Gong, D.-s., Che, J.-x., Wang, J.-z., & Liang, J.-z. (2008). *Short-term electricity price forecasting based on novel SVM using artificial fish swarm algorithm under deregulated power.* Paper presented at the 2008 Second International Symposium on Intelligent Information Technology Application.

Hahn, H., Meyer-Nieberg, S., & Pickl, S. (2009). Electric load forecasting methods: Tools for decision making. *European journal of operational research, 199*(3), 902-907.

Hamzaçebi, C. (2008). Improving artificial neural networks' performance in seasonal time series forecasting. *Information Sciences, 178*(23), 4550-4559.

Hernandez, L., Baladron, C., Aguiar, J. M., Carro, B., Sanchez-Esguevillas, A. J., Lloret, J., & Massana, J. (2014). A survey on electric power demand forecasting: future trends in smart grids, microgrids and smart buildings. *IEEE Communications Surveys & Tutorials, 16*(3), 1460-1495.

Holland, J., & Goldberg, D. (1989). Genetic Algorithms in Search, Optimization and Machine Learning. *ed: Addison-Wesley, Reading, MA*.

Hong, T. (2015). Crystal ball lessons in predictive analytics. *EnergyBiz Mag, 12*(2), 35-37.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning* (Vol. 112): Springer.

Jang, J.-S. (1996). *Input selection for ANFIS learning.* Paper presented at the Fuzzy Systems, 1996., Proceedings of the Fifth IEEE International Conference on.

Jang, J. S. R. (1993). ANFIS: adaptive-network-based fuzzy inference system. *IEEE Transactions on Systems, Man, and Cybernetics, 23*(3), 665-685.

Jin, C. H., Pok, G., Paik, I., & Ryu, K. H. (2015). Short‐term electricity load and price forecasting based on clustering and next symbol prediction. *IEEJ Transactions on Electrical and Electronic Engineering, 10*(2), 175-180.

Jónsson, T., Pinson, P., Nielsen, H. A., Madsen, H., & Nielsen, T. S. (2013). Forecasting electricity spot prices accounting for wind power predictions. *IEEE Transactions on Sustainable Energy, 4*(1), 210-218.

Kankal, M., Akpınar, A., Kömürcü, M. İ., & Özşahin, T. Ş. (2011). Modeling and forecasting of Turkey's energy consumption using socio-economic and demographic variables. *Applied Energy, 88*(5), 1927-1939.

Karakatsani, N. V., & Bunn, D. W. (2008). Forecasting electricity prices: The impact of fundamentals and time-varying coefficients. *International journal of forecasting, 24*(4), 764-785.

Keles, D., Scelle, J., Paraschiv, F., & Fichtner, W. (2016). Extended forecast methods for day-ahead electricity spot prices applying artificial neural networks. *Applied energy, 162*, 218-230.

Keynia, F. (2012). A new feature selection algorithm and composite neural network for electricity price forecasting. *Engineering Applications of Artificial Intelligence, 25*(8), 1687-1697.

Koopman, S. J., Ooms, M., & Carnero, M. A. (2007). Periodic seasonal Reg-ARFIMA–GARCH models for daily electricity spot prices. *Journal of the American Statistical Association, 102*(477), 16-27.

Lago, J., De Ridder, F., & De Schutter, B. (2018). Forecasting spot electricity prices: Deep learning approaches and empirical comparison of traditional algorithms. *Applied Energy, 221*, 386-405.

Landeras, G., López, J. J., Kisi, O., & Shiri, J. (2012). Comparison of Gene Expression Programming with neuro-fuzzy and neural network computing techniques in estimating daily incoming solar radiation in the Basque Country (Northern Spain). *Energy Conversion and Management, 62*, 1-13.

Li-xiao, N., Zheng-fang, W., Chuan-zhi, Z., Wen-li, S., & Sheng-shan, Z. (2014). Hybrid model based on wavelet transform and ARIMA for short-term electricity price forecasting. *Application Research of Computers/Jisuanji Yingyong Yanjiu, 31*(3), 688.

Lin, W.-M., Gow, H.-J., & Tsai, M.-T. (2010). Electricity price forecasting using enhanced probability neural network. *Energy Conversion and Management, 51*(12), 2707-2714.

Maca M., & Lhotská, L. (2013). *Wrapper feature selection significantly improves nonlinear prediction of electricity spot prices.* Paper presented at the Systems, Man, and Cybernetics (SMC), 2013 IEEE International Conference on.

McClymont, K., & Keedwell, E. (2012). Deductive sort and climbing sort: New methods for non-dominated sorting. *Evolutionary computation, 20*(1), 1-26.

Modiri-Delshad, M., & Rahim, N. A. (2016). Multi-objective backtracking search algorithm for economic emission dispatch problem. *Applied Soft Computing, 40*, 479-494.

Mostafavi, E. S., Mousavi, S. M., & Hosseinpour, F. (2014). Gene Expression Programming as a Basis for New Generation of Electricity Demand Prediction Models. *Computers & Industrial Engineering*.

Niu, D.-X., Wang, Q., & Li, J.-C. (2006). Short term load forecasting model based on support vector machine *Advances in Machine Learning and Cybernetics* (pp. 880-888): Springer.

Niu, D., Liu, D., & Wu, D. D. (2010). A soft computing system for day-ahead electricity price forecasting. *Applied Soft Computing, 10*(3), 868-875.

Nowotarski, J., & Weron, R. (2016). On the importance of the long-term seasonal component in day-ahead electricity price forecasting. *Energy Economics, 57*, 228-235.

Panapakidis, I. P., & Dagoumas, A. S. (2016). Day-ahead electricity price forecasting via the application of artificial neural network based models. *Applied Energy, 172*, 132-151.

Pousinho, H., Mendes, V., & Catalão, J. (2010). Hybrid PSO-ANFIS Approach for Short-Term Electricity Prices Prediction.

Pousinho, H. M. I., Mendes, V. M. F., & Catalão, J. P. d. S. (2012). Short-term electricity prices forecasting in a competitive market by a hybrid PSO–ANFIS approach. *International Journal of Electrical Power & Energy Systems, 39*(1), 29-35.

Rafal, W. (2006). Modeling and forecasting electricity loads and prices: John Wiley & Sons.

Raza, M. Q., & Khosravi, A. (2015). A review on artificial intelligence based load demand forecasting techniques for smart grid and buildings. *Renewable and Sustainable Energy Reviews, 50*, 1352-1372.

Razak, I. A. W. A., Abidin, I. Z., Yap, K. S., Abidin, A. A. Z., Rahman, T. K. A., & Nasir, M. N. M. (2016). *A novel hybrid method of LSSVM-GA with multiple stage optimization for electricity price forecasting.* Paper presented at the Power and Energy (PECon), 2016 IEEE International Conference on.

Renani, E. T., Elias, M. F. M., & Rahim, N. A. (2016). Using data-driven approach for wind power prediction: A comparative study. *Energy Conversion and Management, 118*, 193-203.

Rodriguez, C. P., & Anders, G. J. (2004). Energy price forecasting in the Ontario competitive power system market. *IEEE Transactions on power systems, 19*(1), 366-374.

Roy, P. P., & Roy, K. (2008). On some aspects of variable selection for partial least squares regression models. *QSAR & Combinatorial Science, 27*(3), 302-313.

Ruibal, C. M., & Mazumdar, M. (2008). Forecasting the mean and the variance of electricity prices in deregulated markets. *IEEE Transactions on power systems, 23*(1), 25-32.

Shafie-Khah, M., Moghaddam, M. P., & Sheikh-El-Eslami, M. (2011). Price forecasting of day-ahead electricity markets using a hybrid forecast method. *Energy Conversion and Management, 52*(5), 2165-2169.

Sharma, V., & Srinivasan, D. (2013). A hybrid intelligent model based on recurrent neural networks and excitable dynamics for price prediction in deregulated electricity market. *Engineering Applications of Artificial Intelligence, 26*(5-6), 1562-1574.

Shayeghi, H., & Ghasemi, A. (2013). Day-ahead electricity prices forecasting by a modified CGSA technique and hybrid WT in LSSVM based scheme. *Energy Conversion and Management, 74*, 482-491.

Shayeghi, H., Ghasemi, A., Moradzadeh, M., & Nooshyar, M. (2015). Simultaneous day-ahead forecasting of electricity price and load in smart grids. *Energy Conversion and Management, 95*, 371-384.

Shiri, A., Afshar, M., Rahimi-Kian, A., & Maham, B. (2015). *Electricity price forecasting using Support Vector Machines by considering oil and natural gas price impacts.* Paper presented at the Smart Energy Grid Engineering (SEGE), 2015 IEEE International Conference on.

Shrivastava, N. A., & Panigrahi, B. K. (2014). A hybrid wavelet-ELM based short term price forecasting for electricity markets. *International Journal of Electrical Power & Energy Systems, 55*, 41-50.

Singh, N., Mohanty, S. R., & Shukla, R. D. (2017). Short term electricity price forecast based on environmentally adapted generalized neuron. *Energy, 125*, 127-139.

Singhal, D., & Swarup, K. (2011). Electricity price forecasting using artificial neural networks. *International Journal of Electrical Power & Energy Systems, 33*(3), 550-555.

Srinivas, N., & Deb, K. (1994). Muiltiobjective optimization using nondominated sorting in genetic algorithms. *Evolutionary computation, 2*(3), 221-248.

Svalina, I., Galzina, V., Lujić, R., & Šimunović, G. (2013). An adaptive network-based fuzzy inference system (ANFIS) for the forecasting: The case of close price indices. *Expert Systems with Applications, 40*(15), 6055-6063.

Tavana, M., Fallahpour, A., Di Caprio, D., & Santos-Arteaga, F. J. (2016). A hybrid intelligent fuzzy predictive model with simulation for supplier evaluation and selection. *Expert Systems with Applications, 61*, 129-144.

Uniejewski, B., Nowotarski, J., & Weron, R. (2016). Automated variable selection and shrinkage for day-ahead electricity price forecasting. *Energies, 9*(8), 621.

Ünler, A. (2008). Improvement of energy demand forecasts using swarm intelligence: The case of Turkey with projections to 2025. *Energy Policy, 36*(6), 1937-1944.

Unler, A., & Murat, A. (2010). A discrete particle swarm optimization method for feature selection in binary classification problems. *European journal of operational research, 206*(3), 528-539.

Vahidinasab, V., Jadid, S., & Kazemi, A. (2008). Day-ahead price forecasting in restructured power systems using artificial neural networks. *Electric power systems research, 78*(8), 1332-1342.

Ventosa, M., Baıllo, A., Ramos, A., & Rivier, M. (2005). Electricity market modeling trends. *Energy policy, 33*(7), 897-913.

Wan, C., Xu, Z., Wang, Y., Dong, Z. Y., & Wong, K. P. (2013). A hybrid approach for probabilistic forecasting of electricity price. *IEEE Transactions on Smart Grid, 5*(1), 463-470.

Weron, R. (2014). Electricity price forecasting: A review of the state-of-the-art with a look into the future. *International journal of forecasting, 30*(4), 1030-1081.

Weron, R., & Zator, M. (2015). A note on using the Hodrick–Prescott filter in electricity markets. *Energy Economics, 48*, 1-6.

Wu, L., & Shahidehpour, M. (2014). A hybrid model for integrated day-ahead electricity price and load forecasting in smart grid. *IET Generation, Transmission & Distribution, 8*(12), 1937-1950.

Xue, B., Zhang, M., & Browne, W. N. (2013). Particle swarm optimization for feature selection in classification: A multi-objective approach. *IEEE transactions on cybernetics, 43*(6), 1656-1671.

Xue, B., Zhang, M., & Browne, W. N. (2014). Particle swarm optimisation for feature selection in classification: Novel initialisation and updating mechanisms. *Applied Soft Computing, 18*, 261-276.

Yan, X., & Chowdhury, N. A. (2013). Mid-term electricity market clearing price forecasting: A hybrid LSSVM and ARMAX approach. *International Journal of Electrical Power & Energy Systems, 53*, 20-26.

Yang, Z., Ce, L., & Lian, L. (2017). Electricity price forecasting by a hybrid model, combining wavelet transform, ARMA and kernel-based extreme learning machine methods. *Applied Energy, 190*, 291-305.

Yuan, C., Liu, S., & Fang, Z. (2016). Comparison of China's primary energy consumption forecasting by using ARIMA (the autoregressive integrated moving average) model and GM(1,1) model. *Energy, 100*, 384-390.

Zareipour, H., Bhattacharya, K., & Cañizares, C. A. (2007). Electricity market price volatility: The case of Ontario. *Energy Policy, 35*(9), 4739-4748.

Zareipour, H., Cañizares, C. A., Bhattacharya, K., & Thomson, J. (2006). Application of public-domain market information to forecast Ontario's wholesale electricity prices. *IEEE Transactions on Power Systems, 21*(4), 1707-1717.

Zhang, C., Zhou, J., Li, C., Fu, W., & Peng, T. (2017). A compound structure of ELM based on feature selection and parameter optimization using hybrid backtracking search algorithm for wind speed forecasting. *Energy Conversion and Management, 143*, 360-376.

Zhang, L., Luh, P. B., & Kasiviswanathan, K. (2003). Energy clearing price prediction and confidence interval estimation with cascaded neural networks. *IEEE Transactions on Power Systems, 18*(1), 99-105.

Zhang, X., Tian, Y., Cheng, R., & Jin, Y. (2015). An efficient approach to nondominated sorting for evolutionary multiobjective optimization. *IEEE transactions on evolutionary computation, 19*(2), 201-213.

Ziel, F., Steinert, R., & Husmann, S. (2015). Efficient modeling and forecasting of electricity spot prices. *Energy Economics, 47*, 98-111.

Zitzler, E., & Thiele, L. (1999). Multiobjective evolutionary algorithms: a comparative case study and the strength Pareto approach. *IEEE transactions on evolutionary computation, 3*(4), 257-271.

Zou, F., Chen, D., Li, S., Lu, R., & Lin, M. (2017). Community Detection in Complex Networks: Multi-objective Discrete Backtracking Search Optimization Algorithm with Decomposition. *Applied Soft Computing*.