SHORT TERM ELECTRICITY PRICE FORECASTING USING BIOLOGICALLY INSPIRED SUPPORT VECTOR MACHINE

MOHAMAD SHAFIQ BIN ROSLAN

BACHELOR OF ELECTRICAL ENGINEERING WITH HONOURS

(INDUSTRIAL POWER)

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

SUPERVISOR DECLARATION

"I hereby declare that I have read through this report entitled "Short Term Electricity Price Forecasting using Biologically Inspired Support Vector Machine" and found that it complies the partial fulfillment for awarding the degree of Bachelor of Electrical Engineering"

Signature	:	…………………………………….
Supervisor's Name	:	DR. INTAN AZMIRA BINTI WAN ABDUL RAZAK
Date	:	6 JUNE 2018

SHORT TERM ELECTRICITY PRICE FORECASTING USING BIOLOGICALLY INSPIRED SUPPORT VECTOR MACHINE

MOHAMAD SHAFIQ BIN ROSLAN

A report submitted in partial fulfillment of the requirement for the degree of

Bachelor of Electrical Engineering (Power Industrial)

Faculty of Electrical Engineering

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

2018

I declare that this report entitled Short Term Electricity Price Forecasting using Biologically Inspired Support Vector Machine" is the result of my own research except as cited in the references. The report has not been accepted for any degree and is not concurrently submitted in candidature of any other degree.

Signature	:	…………………………………….
Name	:	MOHAMAD SHAFIQ BIN ROSLAN
Date	:	6 JUNE 2018

To my beloved mother and father

ACKNOWLEDGEMENT

In preparing this report, I used to be in contact with some people, academicians, and practitioners. They have contributed towards my understanding and thought. In particular, I wish to express my sincere appreciation to my main project supervisor, Dr. Intan Azmira Binti Wan Abdul Razak, for guidance, encouragement, critics, and friendship. I am also very thankful to my friends who are in about the same case study for the help and give some suggestions. Without their continued support and interest, this project would not have been same as presented here.

I would like to express my gratitude to Universiti Teknikal Malaysia Melaka (UTeM) for giving me the chance to pursue my Bachelor in Electrical Engineering. My fellow postgraduate students should also be recognized for their support. My honest appreciation also extends to all my colleagues and others who have provided assistance on various occasions. Their views and tips are beneficial certainly. Unluckily, it is not viable to listing all of them in this limited space.

L ast but not least, I am grateful to all my family members for all their moral support. I also would like to thank all those involved directly or indirectly in carrying out this project. Without their aide, I might not be able to complete this project at any

ABSTRACT

In deregulated electricity markets, consumers have the ability to choose their electricity supplier. Power producers and customers use short term price forecasts to manage and plan for bidding approaches, and hence increase the utility's profit and energy efficiency. Thus, it allows for greater price flexibility and increased competition between electric providers. This project proposes a prediction model for short-term electricity price forecasting using Support Vector Machine (SVM) and Bacterial Foraging Optimization Algorithm (BFOA). Least Square Support Vector Machine (LSSVM) which is an algorithm that is improved from the SVM is used in this project. Bacterial Foraging Optimization Algorithm (BFOA) is used to optimize the parameters of the LSSVM model which is gamma (γ) and sigma (σ). Furthermore, BFOA optimizes number of features to be fed into LSSVM. The parameters of BFOA are varied to find the best LSSVM-BFOA configuration. The result showed that LSSVM-BFOA able to predict electricity price with good Mean Absolute Percentage Error (MAPE). The model is examined in the Ontario, electricity market in Canada.

ABSTRAK

Dalam pasaran elektrik yang tidak dapat dikawal, pengguna mempunyai keupayaan untuk memilih pembekal elektrik mereka. Pengeluar dan pelanggan kuasa menggunakan ramalan harga jangka pendek untuk mengurus dan merancang pendekatan bidaan, dan dengan itu meningkatkan keuntungan dan kecekapan tenaga utility. Oleh itu, ia membolehkan fleksibiliti harga yang lebih tinggi dan peningkatan persaingan antara pembekal elektrik. Projek ini mencadangkan model ramalan untuk ramalan harga elektrik jangka pendek menggunakan Mesin Vektor Sokongan Biologi (SVM) dan Algoritma Pengoptimuman Pengekalan Bacterial (BFOA). Mesin Vector Sokongan Sisi Kecil (LSSVM) yang merupakan algoritma yang diperbaiki daripada SVM digunakan dalam projek ini. Algoritma Pengoptimuman Pengekalan Bacterial (BFOA) digunakan untuk mengoptimumkan parameter model LSSVM iaitu gamma (γ) dan sigma (σ). Selain itu, BFOA mengoptimumkan jumlah ciri yang akan dimasukkan ke dalam LSSVM. Parameter BFOA adalah berbeza untuk mencari konfigurasi LSSVM-BFOA terbaik. Hasilnya menunjukkan bahawa LSSVM-BFOA dapat meramalkan harga elektrik dengan Ralat Peratusan Maksimum Mutlak yang baik (MAPE). Model ini diperiksa di Ontario, pasaran elektrik di Kanada.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

MCP	-	Market Clearing Price
LSSVM	-	Least Square Support Vector Machine
BFOA	-	Bacterial Foraging Optimization Algorithm
NN	-	Neural Network
MAPE	-	Mean Absolute Percentage Error
ARIMA	-	Autoregressive Integrated Moving Average
ANN	-	Artificial Neural Network
SVM	-	Support Vector Machine
KKT	-	Karush- Kuhn-Tucker
QP	-	Quadratic Programming
SSE	-	Sum Square Error
GA	-	Genetic Algorithm
PSO	-	Particle Swarm Optimization
HOEP	-	Hourly Ontario Electricity Price
S	-	Number of bacteria
Nc	-	Number of Chemotactic
Ns	-	Number of steps during swim
Nre	-	Number of reproduction steps
Ned	-	Number of elimination-dispersal
Ped	-	Number of Probability elimination-dispersal

MAE	-	Mean absolute error
R	-	Regression
PDP	-	Pre-Dispatch Prices
MARS	-	Multivariate adaptive regression splines

CHAPTER 1

INTRODUCTION

1.1 Project Introduction

The world population is increasing day by day and somehow affected the usage of electricity. Furthermore, with current technologies, electricity becomes the source to deliver the power. Electricity needed to run everything in daily lives. When the usage is increasing, the price of electricity for sure will increase as well. With this problem, electricity price forecasting is performed. Electricity Price Forecasting (EPF) plays an important role in the wholesale electricity market. Forecasting electricity price essential for better bidding which is important for both consumer and suppliers. The consumer can maximize their usage and minimize their expense. Energy supplier aim is sell all their generated power and maximize their income [1].

A power market company can forecast the fluctuation of wholesale prices with an affordable level of accuracy. This may regulate its bidding strategy and its own particular generation or utilization plan in order to maximize the profits in day-ahead mercantilism cut back the risk. This also helps the customer to control their usage of the electricity in their daily activities.

Somehow, as compared to predicting the load or demand, forecasting electricity price is more challenging. This is because of fluctuation of price where unexpected spikes might occur for any point of series. There are others factors that affecting the unpredictability in value such as the weather condition, an imbalance between supply and demand and unexpected disturbance at generation and transmission sites. Time-Series (TS), Neural Network (NN) and Support Vector Machine (SVM) are some of the methods that the previous researchers used to forecast electricity price.

1.2 Motivation

Nowadays, fluctuation of electricity price is very common due to various factors. Demand is the most important factor that affects the electricity price. Demand increase because of consumer behavior toward electricity, holidays and peak hours during weekdays. Another important factor is fuel cost that affected by gas prices or coal prices. Besides, weather condition, power plant, transmission, and distribution system for maintenance as known as the factor affecting electricity price forecasting. Both of the suppliers and consumers have the same goal which is to gain the profit. The consumers want to schedule their energy consumption for the next days and maximize their usage, while the suppliers aim to sell all generated energy and maximize their profit [1].

1.3 Problem Statement

There are a lot of methods used for electricity price forecasting. This method can be divided into two categories which are simpler approaches and complex approaches. From previous researches, most of the methods can predict electricity well during normal condition. However, when the spike occurrences happened, the forecast error become large. Some researchers reported least Square Support Vector Machine (LSSVM) can deal with the spikes occurrences [2]. Hence, Least Square Support Vector Machine (LSSVM) is chosen as forecast model.

One of the factors of that cause a bad result of forecast price is the improper selection of features. Due to that, it leads to low accuracy and efficiency of price forecast. Too many feature selection cause the time taken to carry out the analysis become longer. While, insufficient features may also lead to high forecast error. So, the selecting on the features based on the correlation result should be chosen correctly.

Some researched reported that stand-alone forecast model cannot produce a good accuracy. Therefore, this project proposed to combine the optimizer to the forecast model. This optimizer works to select the significant inputs or optimize parameter of forecast model [3]. The selection of parameter is important for optimization to improve the accuracy of optimization.

1.4 Objectives

The objectives of this project areas below:

i. To analyze the correlation between forecast input and target price by using correlation analysis.
ii. To develop Least Square Support Vector Machine (LSSVM) as main forecast engine and optimize LSSVM parameter and features using Bacterial Foraging Optimization Algorithm (BFOA).

1.5 Scope

Scopes of this project are:

i. This project uses LSSVM as main forecast engine.
ii. BFOA is proposed to optimize LSSVM parameter and features of forecasting.
iii. Programming language of MATLAB is written.
iv. The data are taken from Ontario, Canada.
v. Objective function of Mean Absolute Percentage Error (MAPE) is used to observe the forecast error.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

Electricity price forecasting has become one of the serious tasks in the operation of the electrical power system.

This participant consists of consumers and suppliers. Both suppliers and consumers need an accurate electricity price prediction to make their profit. The consumer's wish is to maximize power usage and minimize their price while the suppliers desire to sell all their produced power and boost their income.

Based on the previous forecast researchers, electricity price is hard to forecast due to some factors. Some of the factors affected this forecast are load behavior, unstable climate and fuel cost [4].

Many methods have been used to forecast electricity price during normal condition as it works well. However, during the occurrences of a spike, the forecast error becomes high. In this project, Least Square Support Vector Machine (LSSVM) was chosen as a main method and Bacterial Foreign Optimization Algorithm (BFOA) was chosen as an optimization technique.

2.2 Methods of Electricity Price Forecasting

Some of the popular methods of forecast are Time-Series, Neural Network, and Support Vector Machine. The theory and basic principle of each of the method will be explained in the next section.

2.2. 1 Time-Series

A time series is a consecutive arrangement of information data, measured typically over amount of times. A time series can be constant or discrete. In a constant time series information are measured at each example of time, while a discrete time series contains perceptions measured at discrete purposes of time [5].

Strong explanatory model which is Time-Series has been used for forecasting by simulating the variation of historical price [6]. There are many techniques that have been used for time series and one of those techniques is Autoregressive Integrated Moving Average (ARIMA). From previous researches, times series forecasting by ARIMA has become one of the favorite methods [7].

ARIMA models are essentially created to estimate the comparing subordinate variable. This ARIMA model can be partitioned into two kinds of figure which are post sample period forecast and sample period forecast. The sample period forecast is utilized to create certainty interim in the model and the post sample period figure is utilized to produce honest to goodness gauges for arranging and different purpose [8].

Rafał Weron & Misiorek (2008): The authors examine two markets using a wide range of advanced Time Series models: For the Nordpool market, they report an MAPE of only 3.2%. The other market they examine is the Californian market where a MAPE of 12.96% is the best result for a model in which spikes have been pre-processed in a way that they are dampened [20]. Dawit Hailu Mazengia (2008): In a recent study of the Ontario Electricity market in 2007, the authors report a MAPE is 17.85%. The forecasting was done by suing multiple linear regression (MLR) and included exogenous variables for technologies, market power and network congestion using a rolling window approach [21].

2.2.2 Neural Network

Neural Network is a mathematical model that emulates the functional architecture of the human brain [9]. NN works like a human brain. Thus, NN will learn and identify the design of input data. After that, NN will generate the result by applying the knowledge that has been stored in their brain [5].

Neural Network (ANN) is a frequent forecasting technique of electricity price. However, an artificial neural network has intrinsic restitution, and for instance, when the data is out of training sample, the error is extremely large. Thus, it has limited generalization capability and uncontrolled convergence [10].

In the application of electricity price forecasting, NN is analyzed in two cases. In the primary case, we consider just the cost of electric vitality as information and the system preparing is done exclusively in light of valuable information. At that point in the second case, both electric vitality utilization and costs are considered as sources of inputs [11].

Hamdireza Zareipour developed Artificial Neural Network (ANN) for forecasting short-term electricity market price and its application to operation planning of demand-side Bulk Electricity Market Customer (BECMs) in Ontario market [24]. The authors reported that the MAPE from the model is 18.30% based on electricity price in 2004 which come out better that the previous researches.

2.2.3 Support Vector Machine

SVM is a model that uses to analyze data and recognize the pattern for estimation and classification [12]. There are many advantages of SVM compared to the other methods.

SVM can deal with many problems such as local minima, high structural input data, and over-fitting. However, SVM has a disadvantage which is bind with the burden of large set input of data. In order to overcome this problem, Least Square Support Vector Machine (LSSVM) is suggested to lower the burden of SVM [4].

LSSVM is the least squares formulation of a standard SVM. [13]. LSSVM solves a system of linear equations that namely as Karush- Kuhn-Tucker (KKT), instead of a quadratic programming (QP) [14].

In other words, KKT is more easy to use than QP. Even though SVM is improved, LSSVM still retains the concept of old SVM which great in observation capability. LSSVM also reduce the sum square errors (SSEs) of practical statistics sets while synchronously lowering error.

2.3 Optimization Algorithm

In this project, optimization is used to be hybrid with the forecast engine. Optimization can be defined as a minimization or maximization problem. From previous researchers, parameters of forecast engine can be optimized or important inputs can be chosen by optimization algorithm [3]. There are many methods of optimization algorithm such as Genetic Algorithm (GA) and Particle Swarm Optimization (PSO).

2.3.1 Genetic Algorithm

GA is an algorithm based on the survival of the fittest and the natural evolution process via reproduction. The objective function is often referred to as fitness function. Three main operations in GA are selection, crossover and mutation.

There are four core elements that influencing the performance of GA which are population size, number of generations, crossover rate and mutation rate. Furthermore, the algorithm is usually terminated when the generation reaches its maximum value or an acceptable fitness value is obtained.

2.3.2 Particle Swarm Optimization

PSO was introduced by James Kennedy and Rusell Eberhart. PSO mimics the behavior of a group of migrating birds or fish trying to reach unknown destination in the search space changing its velocity. Each individually in a group will randomly move around to find food and announce the sources of the food to its neighbors so the neighbors will approach the same location [2].

A lot of researchers are interested in using PSO since then. The field of swarm intelligence is now become an exciting and constantly evolving the subject of research. Almost every area in design, optimization or scheduling applications, and computational intelligence has used PSO.

2.3.3 Bacterial Foraging Optimization Algorithm (BFOA)

The BFOA was found to search food quicker than other optimization methods [16]. Moreover, the BFOA shows a great performance rather than other meta-heuristic optimization approaches [17]. The BFOA was found to perform better in terms of fast convergence, simplicity in programming, accuracy, and flexibility.

In this project, BFOA is modelled to be hybrid with the LSSVM engine. BFOA is a foraging activity that ingesting and locating nutrient or food that present in human intestines based on bacteria named E.coli.

2.4 Summary and Discussion of the Review

Least Square Support Vector Machine (LSSVM) shows better MAPE than other methods at Ontario, Canada. LSSVM produce 13.0871% MAPE compared to Time-Series and Neural-Network which are 17.63% and 18.80% respectively. Table 2.1 below shows the result of previous MAPE.

Table 2.1: Previous Result of MAPE

References	**Method**	**Test data**	**MAPE (%)**
[2]	Time Series (ARIMA)	2006	17.6300
[2]	Neural Network	2006	18.8000
[2]	LSSVM	2010	13.0871

Based on the result, it can be concluded that LSSVM is a good method for conducting the electricity price forecasting. Each of the methods has advantages and disadvantages that affect the result of MAPE. LSSVM is a technique that not requires big set of data to attain the connection between input and output. LSSVM is good to deal with immense dimensional input and also can step down the over-fitting. Moreover, LSSVM minimizes the sum square errors (SSEs) of training data set while simultaneously lowering margin error [4].

Compare to LSSVM, Time-Series method more suitable for linear problem while price sequence is a non-linear pattern. Meanwhile, NN has problem with over-fitting and under-fitting. This can cause generalization problem as the model cannot figure out the connection between input and output. Besides, training process of NN require extra time.

2.5 Conclusion

Based on the previous research, Least Square Support Vector Machine (LSSVM) shows a better result on performance and accuracy than the other two methods which are Time-Series and Neural Network. So, LSSVM is chosen to be the main forecast engine for this project. In addition, this project proposed to add the optimizer Bacterial Foraging Optimization Algorithm (BFOA) is selected as the optimization technique to improve the performance and accuracy of Support Vector Machine.