CRUDE OIL PRICE FORECASTING BASED ON THE RECONSTRUCTION OF IMFS OF DECOMPOSITION ENSEMBLE MODEL WITH ARIMA AND FFNN MODELS

MUHAMMAD AAMIR

UNIVERSITI TEKNOLOGI MALAYSIA

CRUDE OIL PRICE FORECASTING BASED ON THE RECONSTRUCTION OF IMFS OF DECOMPOSITION ENSEMBLE MODEL WITH ARIMA AND FFNN MODELS

MUHAMMAD AAMIR

A thesis submitted in the fulfilment of the
requirements for the award of the degree of
Doctor of Philosophy

Faculty of Science
Universiti Teknologi Malaysia

AUGUST 2018

"Dedicated to my beloved family!"

ACKNOWLEDGEMENT

In the name of Allah, the most beneficent and merciful, who made all things possible, gave me the strength and power to complete this thesis successfully. I wish to express my sincere gratitude to my supervisor, Dr. Ani Bin Shabri. During the last few years, he was more than just a formal advisor, he helped me to understood new concepts and working hard to produce this work. Thank you very much, Dr. Ani, for showing me the importance of researching without having numbers as parameter. For sure, our collaboration and friendship will never end.

Well, I was impatient to start talking about my wife and my sons. Saba, you are my strength, my motivation, and my inspiration. Thank you very much for your patience and love. You are present in every page of this thesis. Thank you very much for making my days and trips funnier. Wajdaan and Sanan, my two sons, you came up a few years ago, but I cannot even remember my life before you. You suddenly became the main reason for everything. Your tender smile moves me and makes me forget any tough situation.

I am deeply grateful to my father Wakeel Dad, my mother, my brothers Dr. Muhammad Ishaq, Hazrat Ali, Wajid Ali, and my sisters. You are the greatest models for me. I am pretty sure you are always by my side, regardless the situation I know I can count on you. I love you so much.

I am also grateful to the staff members of the Department of Mathematical Sciences, Universiti Teknologi Malaysia (UTM) for their help and support. The UTM library deserves special thanks for supplying the relevant resources.

Finally, I also thank all my friends and colleagues from AWKUM Pakistan (Prof. Dr. Zahoor ul Haq, Dr. Muhammad Tariq, Dr. Shehzad Khan, Sayim Sohail, Arsh ur Rehman and Hassan Khan), from UTM (Prof. Dr. Nor Haniza Sarmin, Muhammad Faizan Malik, Abrar Ullah, Aliyu Isah Chikaji (Nigeria), Ghani ur Rehman, Shoukat Ali), from Pakistan (Dr. Muhammad Yasir, Dr. Imran Ullah, Dr. Arshad Hussain), from Malaysia (Dr. Muhammad Imran Qureshi (UniKL)) and last but not the least Dr. Muhammad Adil Khattak my travelling partner and my motivational mentor, your dedication to work inspires me a lot, you were my teacher and now I am sure our friendship goes on for many years. For those whom I could not mention your names, please know that I cherish and appreciate all your support and prayers.

This research was supported by Abdul Wali Khan University Mardan, KPK, Pakistan, under the faculty development program (FDP) project sponsored by KPK Government.

ABSTRACT

The development of economic and industry depend upon how well the accuracy of crude oil price forecasting is managed. The study aims to reduce computation complexity and enhance forecasting accuracy of decomposition ensemble model. The propose model comprises four steps which are (i) decomposing the complex data into several IMFs using ensemble empirical mode decomposition (EEMD) method, (ii) reconstructing the decomposed IMFs through autocorrelation into stochastic and deterministic components, (iii) forecasting every reconstructed component, and (iv) ensemble all forecasted components for the final output. IMFs in the stochastic component are analysed separately. The findings confirm that the stochastic component contributed more variation as compared to deterministic component. For verification and illustration, Brent, West Texas Intermediate (WTI) daily, weekly, monthly and yearly, and Pakistan monthly spot crude oil prices were used as sample study. The empirical results indicated that the proposed model statistically outperformed all the considered benchmark models including the most popular auto-regressive integrated moving average (ARIMA) model, feed forward neural network (FFNN) model, decomposition ensemble model (EEMD-ARIMA and EEMD-FFNN), reconstruction decomposition ensemble model with stochastic and deterministic components (EEMD-(S+D)-ARIMA and EEMD-(S+D)-FFNN) and Rios and De Mello (RD) reconstruction decomposition ensemble model with stochastic and deterministic components (EEMD-RD-ARIMA and EEMD-RD-FFNN). To determine the performance, two descriptive statistical measures were applied, including the root mean square error (RMSE) and mean absolute percentage error (MAPE). The MAPE of the proposed EEMD-individual stochastic and deterministic (ISD)-FFNN model for daily and weekly data of Brent and WTI are <1%, however, for monthly Brent, WTI and Pakistan data are <5% shows a good fit produce by EEMD-ISD-FFNN. The MAPE of the model EEMD-ISD-FFNN for yearly Brent data is <30% indicate a reasonable fit and for WTI <20% implies a good fit. Whereas the MAPE of the EEMD-(S+D)-FFNN model for Brent yearly data <20% display a good fit and for WTI data <10% indicate excellent fit. In nutshell, the recommended model for yearly data is EEMD-(S+D)-FFNN. In conclusion, the proposed method of reconstruction of IMFs based on autocorrelation enhanced the forecasting accuracy of the EEMD model.

ABSTRAK

Perkembangan ekonomi dan industri bergantung kepada sejauh mana ketepatan ramalan harga minyak mentah diuruskan. Kajian ini bertujuan untuk mengurangkan kerumitan perhitungan dan meningkatkan ketepatan ramalan menggunakan model pengabunggan. Model pengabungan ini terdiri daripada empat langkah iaitu (i) menguraikan data kompleks ke dalam beberapa IMF menggunakan kaedah penguraian mod empirikal (EEMD), (ii) mengkelaskan IMF yang diuraikan melalui autokorelasi ke dalam komponen stokastik dan deterministik, (iii) meramalkan setiap komponen yang dibina semula, dan (iv) mengabungkan semua ramalan untuk dijadikan sebagai hasil ramalan akhir. IMF dalam komponen stokastik dianalisis secara berasingan. Penemuan mengesahkan bahawa komponen stokastik menyumbang lebih banyak variasi berbanding komponen deterministik. Untuk pengesahan dan ilustrasi, harga minyak mentah Brent, West Texas Intermediate (WTI) harian, mingguan, bulanan dan tahunan, dan harga minyak mentah bulanan Pakistan digunakan sebagai sampel kajian. Keputusan empirikal menunjukkan bahawa model yang dicadangkan secara statistik mengatasi semua model penanda aras yang dipertimbangkan termasuk model purata gerak bersepadu auto-regresif yang paling popular (ARIMA), model rangkaian neural suap ke hadapan (FFNN), model pengabungan penguraian (EEMD-ARIMA dan EEMD-FFNN), model pengabungan penguraian pembinaan semula dengan komponen stokastik dan deterministik (EEMD-(S+D)-ARIMA dan EEMD-(S+D)-FFNN) dan Rios dan De Mello (RD) model gabungan penguraian pembinaan dengan komponen stokastik dan deterministik (EEMD-RD-ARIMA dan EEMD-RD-FFNN). Untuk menentukan prestasi model yang dicadangkan, dua statistik deskriptif telah digunakan, termasuk punca min ralat kuasa dua (RMSE) dan min ralat mutlak (MAPE). MAPE bagi model stokastik EEMD-individu dan deterministik (ISD)-FFNN yang dicadangkan untuk setiap data harian dan mingguan untuk Brent dan WTI adalah <1%, bagaimanapun, untuk data bulanan bagi Brent, WTI dan Pakistan adalah <5%, menunjukkan pemadanan yang begitu baik dihasilkan oleh EEMD-ISD-FFNN. MAPE bagi model EEMD-ISD-FFNN untuk data tahunan Brent adalah <30% menandakan pemadanan yang munasabah dan untuk WTI adalah <20% menunjukkan pemadanan yang baik. Sedangkan MAPE bagi model EEMD-(S+D)-FFNN untuk data tahunan bagi Brent adalah <20% memaparkan pemadanan yang baik dan untuk data WTI adalah <10% menandakan pemadanan yang begitu baik. Secara ringkasnya, model yang dicadangkan untuk data tahunan adalah model EEMD-(S+D)-FFNN. Kesimpulannya, kaedah cadangan pembinaan semula IMF berdasarkan autokorelasi meningkatkan ketepatan ramalan model EEMD.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ACF	-	Autocorrelation Function
ADF	-	Augmented Dickey-Fuller
AIC	-	Akaike Information Criterion
ANN	-	Artificial Neural Network
ARIMA	-	Autoregressive Integrated Moving Average
Corr.	-	Correlation
DC	-	Deterministic Component
EEMD	-	Ensemble Empirical Mode Decomposition
EMD	-	Empirical Mode Decomposition
GARCH	-	Generalized Autoregressive Conditional Heteroscedasticity
IMF	-	Intrinsic Mode Function
LB	-	Ljung-Box test
MAE	-	Mean Absolute Error
MAPE	-	Mean Absolute Percentage Error
MLE	-	Maximum Likelihood Estimator
PACF	-	Partial Autocorrelation Function
R	-	R- software
RD	-	Rios and DeMello
RMSE	-	Root Mean Square Error
SC	-	Stochastic Component
TDS	-	True Deterministic Series
TSS	-	True Stochastic Series
t-stat	-	t-statistic
WTI	-	West Texas Intermediate
Z-stat	-	Z-statistic

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

1.1 Introduction

This chapter provides an overview of the research undertaken in this research project. First background of this study will briefly be discussed followed by problem statement which describes the problem. Besides, the chapter presents the research questions which this study attempts to answer. Moreover, the chapter puts forward the research objectives followed by the significance of the study. The scope of this research study is also explained in this chapter. The second and last section of this study dwells upon the contribution of the study and the origination of the thesis respectively.

1.2 Background of the Study

Crude oil is a very important commodity in the world because of its unique nature as it affects the life of every individual in many ways. According to International Energy Agency (IEA), as of early 2017, the world currently consumes 97.9 million barrels of oil and liquid fuels daily. In global markets it is the most active and heavily traded commodity. Due to high demand of crude oil in every field of life, it needs more attention as compared to other commodities. Oil is a non-renewable commodity, but the world consumes it in different ways thus it is a challenge for mathematician, statistician and econometrician to develop a better strategy for understanding the price changing aspect of crude oil**.** Crude oil is treated as a special commodity among different world players such as oil companies, oil producing nations, oil importing nations, speculators and individual refineries. Due to the uneven nature of geo political and global socio-economic events the prices of crude oil are sensitive**.** Fluctuation in the price of the crude oil involves several factors like supply, demand, inventory and consumption but irrespective of these factors the oil prices are also influenced by irregular and unpredictable elements which are random in nature. These characteristics of the crude oil price make fluctuation in the market. Due to the irregular and stochastic nature of oil prices it is a very complex and challenging task for researchers to develop appropriate models for forecasting the crude oil prices. The complex and compound nature of crude oil price makes this area widely opened for researchers to develop different procedures for forecasting the crude oil prices in a good manner. As far as the world economy is concerned, crude oil plays an increasingly important role as two third of the world's energy demand is met by crude oil (Alvarez-Ramirez *et al.*, 2003). Like other commodities, the price of crude oil is also measured from the demand and supply. Experimental evaluations propose that these factors are accountable for latest drop in crude oil prices. Since both supply and demand are related factors underlying the recent decline in crude oil prices are likely to continue over the near- to medium-term, crude oil prices are expected to remain soft but unstable, with a steady regaining over the next decade (Baffes *et al.*, 2015).

Additionally, the whole aggregate economic activity can be disturbed by a sharp movement of crude oil price which may fluctuate the nation's economy significantly. Furthermore, the impacts can be reflected in two ways on nation's economy. Firstly, the oil importing countries' economic growth is adversely influenced increasing inflation with sharp rise in crude oil prices. Secondly, the oil exporting countries face a serious budgetary shortfall problems by a small fall in crude oil prices (Abosedra and Baghestani, 2004). The topic of forecasting crude oil price is very important although it is difficult to foretell it due to its high volatility and inherent difficulties (Wang, 2005).

In the last two decades, typical statistical tools and econometric methods were used for forecasting crude oil prices, such as generalized autoregressive conditional heteroscedasticity (GARCH) models, autoregressive moving average (ARIMA) models, linear regression, naïve random walk, error correction model (ECM), vector autoregressive (VAR) models, and also computational approaches such as artificial neural networks (ANN), decomposition-ensemble techniques of wavelet decomposition, empirical mode decomposition (EMD) and ensemble empirical mode decomposition (EEMD) have been used. Nowadays the decomposition procedures are more popular in the financial markets due to the reason that these techniques are more flexible and produce high level of accuracy. Recently, some modified decomposition-ensemble technique is used for forecasting the crude oil prices such as (Yu *et al.*, 2015) suggested an EEMD based technique to enhance forecasting accuracy and reduce complexity in computation as well as in time.

Furthermore, for forecasting crude oil prices (Yu *et al.*, 2016) proposed a new EEMD paradigm and extended extreme learning machine (EELM) based on the principle of "decomposition and ensemble". Moreover, (Zhu *et al.*, 2016) proposed a new multi-scale paradigm through a kernel function incorporating EEMD, particle swarm optimization (PSO) and least-square support vector machines (LSSVM). The EEMD method was used to decompose the original time series into intrinsic mode functions (IMFs) and a trend function through sifting processes. However, an important issue could arise regarding the model computational time, cost and complexity. Because, the decomposition–ensemble models break down the time series into many IMFs and a trend function.

To solve the above issue different authors introduced the reconstruction of IMFs and trend function into some components aiming to reduce the computational time and model selection complexity (Rios and De Mello, 2013; Shu-ping *et al.*, 2014; Yu *et al.*, 2015; Zhu *et al.*, 2016). Rios and De Mello (2013) divided the IMFs and trend function obtained from EMD into two groups called stochastic and deterministic components. They used the recurrence plot (RP) and average mutual information (AMI) for dividing the IMFs and trend function into stochastic and deterministic components. The empirical study assured that the reconstruction of IMFs and trend function into stochastic and deterministic components enhanced the forecasting accuracy. Shu-ping *et al.* (2014) also divided the IMFs and trend function of EMD into high, medium, and low frequencies and in a trend sequence through run-length judgment method and outperformed the decomposition ensemble results using all IMFs and trend function for forecasting the crude oil prices.

EEMD paradigm IMFs and trend function was also reconstructed into components by (Yu *et al.*, 2015) using the "data-characteristic-driven" approach. The main data characteristics were data complexity (i.e. high and low) and pattern characteristics (i.e. cyclicity, mutability, and tendency), the empirical results showed that the reconstruction of IMFs and trend function into components enhanced the forecasting accuracy. Zhu *et al.* (2016) also divided the EEMD IMFs and trend function into groups i.e. high frequency's (HFs), low frequency's (LFs) and trend. Later, the HFs was forecasted using ARIMA model, while LFs and trend were forecasted by the PSO-LSSVM methods and for the final prediction the whole forecasted results are simply combined. Most of the above studies have used two or more criteria for reconstruction of IMFs and trend function such as recurrence plot, average mutual information, high frequency, low frequency, trend, data complexity and data pattern characteristics. The use of these different criteria motivates this thesis to develop a single criterion for reconstruction of EEMD IMFs and trend function and to enhance the forecasting accuracy as compared to the other procedures of reconstruction of IMFs and trend function.

1.3 Problem Statement

The complex data of crude oil prices was effectively handled by the decomposition ensemble models EEMD-ARIMA and EEMD-FFNN as compared to single ARIMA and FFNN models. EEMD divided the data into several IMFs and in a trend function to simplify the task of forecasting. However, an important issue arose regarding the model computational time, cost and complexity and most probably sometimes leading to a poor result because the estimation errors of all models be accumulated in the last ensemble step of forecasting (Yu *et al.*, 2015).

To overcome the above problem different studies have been executed (Rios and De Mello, 2013; Shu-ping *et al.*, 2014; Yu *et al.*, 2015; Zhu *et al.*, 2016) which reconstructed the IMFs and trend function into some meaningful components using the different data characteristics, data patterns, high frequencies, medium frequencies, low frequencies, trend, run-length judgement method, recurrence plot, and average mutual information. All studies of reconstruction of IMFs and trend function obtained from EEMD exploited different procedures to make new components and most of them used more than one procedure for reconstructing of IMFs and trend function.

The above problem of reconstruction of IMFs and trend function motivated the researcher to develop a single procedure for reconstruction of IMFs and trend function. Thus, in this research project IMFs and trend function obtained from EEMD are divided into two components i.e. stochastic and deterministic based on autocorrelation. A threshold value of autocorrelation will be fixed (see chapter 4) for dividing the IMFs and trend function into stochastic and deterministic components by simulating different time series with different number of observations and threshold values. All IMFs and trend function with determinism rate greater than a threshold are added to form the deterministic component and the remaining ones are summed to form the stochastic component. After getting the stochastic and deterministic components, appropriate model selection for crude oil price forecasting is attempted. To check the effect of the reconstruction of IMFs and trend function into stochastic and deterministic components, the two models are used (see chapter 5). i.e. EEMD-Stochastic(S)+Deterministic(D)-ARIMA or EEMD-(S+D)-ARIMA and EEMD-(S+D)-FFNN.

Furthermore, in this thesis the IMFs being part of the stochastic component is studied in more details because from the analysis it proves that it significantly influences the overall results. Hence, the IMFs being part of the stochastic component will be modelled individually for possible increase in forecasting accuracy. Thus, the models EEMD-Individual Stochastic (IS) Deterministic (D)-ARIMA or EEMD-ISD-ARIMA and EEMD-ISD-FFNN are fitted for crude oil prices (see chapter 6). Therefore, the focus of this study is to establish the best optimal forecasting model for crude oil price. By doing so, it is expected to further improve the forecasting accuracy of the crude oil price.

1.4 Research Questions

The research questions of this study are as follows:

i) Would decomposition techniques improve forecasting accuracy as compared to single ARIMA and FFNN models?
ii) Could the IMFs obtained from decomposition ensemble models be separated further into stochastic and deterministic components?
iii) Could the reconstructed decomposition ensemble models outperform the ARIMA, FFNN, EEMD-ARIMA and EEMD-FFNN models?
iv) Could the reconstructed decomposition ensemble models perform well for daily, weekly, monthly and yearly data or its use is specific to certain data sets?

The above all research questions will be answered through empirical analysis all over the study

1.5 Research Objectives

The sole aim of this thesis is the reconstruction of IMFs and trend function obtained from EEMD aiming to enhance forecasting accuracy and reducing the computational time and model selection complexity. The objectives of this thesis are as follows:

i) To determine if the decomposition-ensemble models can effectively modelled the complex data of crude oil price as compared to single ARIMA and FFNN models.

ii) To determine a threshold value which divided the IMFs and trend function obtained from EEMD into stochastic and deterministic components.

iii) To determine the best model for reconstructed components for forecasting the crude oil price by using the ARIMA and FFNN models.

iv) To compare the IMFs and trend function obtained from EEMD and reconstructed components by using the ARIMA and FFNN models.

1.6 Significance of the Study

The objective of this research is to develop a new procedure for reconstruction of IMFs and trend function obtained from EEMD and designs an appropriate model for world crude oil price forecasting with reconstructed stochastic and deterministic components. The advantages of the new proposed method of reconstruction of IMFs and trend function are as listed as (i) the proposed procedure of reconstruction of IMFs and trend function based on only one criterion which is autocorrelation of all IMFs, (ii) the proposed procedure of reconstruction of IMFs and trend function takes less time in model selection for every reconstructed component because the number of IMFs reduces from the original number, (iii) the forecasting accuracy also increases with the use of new reconstructed components which is essential for new models, and (iv) the new proposed procedure of reconstruction of IMFs and trend function could also be implemented in automatic function of some software like R and Matlab from which the system will automatically reconstruct the new components because the new procedure does not require human monitoring. The models which proposed the reconstructed components are EEMD-(S+D)-ARIMA, EEMD-(S+D)-FFNN, EEMD-ISD-ARIMA and EEMD-ISD-FFNN in this thesis.

The new proposed models EEMD-ISD-ARIMA and EEMD-ISD-FFNN significantly improved the forecasting accuracy of the crude oil price as compared to single ARIMA and FFNN models and EEMD-ARIMA and EEMD-FFNN which used all IMFS and trend function and EEMD-(S+D)-ARIMA, EEMD-(S+D)-FFNN which used only stochastic and deterministic components. As a test case different crude oil price series were used and forecasted including daily, weekly, monthly and yearly data to check its usefulness, generalizability and robustness of the proposed method. The empirical results assured the importance of the proposed reconstruction technique and significantly improved the forecasting accuracy of the crude oil price and recommended for the forecasting of crude oil price.

In general, the new procedure which is designed in this study for forecasting crude oil price would be useful for investors, suppliers, government agencies for planning their activities within the available resources and the statisticians, econometrician and researchers in particular to grasp the crude oil price understanding and will produce more up-to-date and better forecasts for future crude oil price.

1.7 Scope of the Study

This study only focused on the crude oil price forecasting. The well-known linear and non-linear single models ARIMA and FFNN are applied to check the effect of reconstruction of IMFs and trend function obtained from EEMD. For decomposition of time series, only the EEMD method is exploited. For reconstruction of IMFs and trend function only two methods are employed; the average mutual information (AMI) and the proposed method of autocorrelation. For fixing the threshold value of determinism rate of autocorrelation, the simulation of different time series was carried out using four different scenarios. In this study only one step ahead forecast is performed. Three different crude oil price series Brent, WTI and Pakistan were utilized as a test case. The Brent and WTI crude oil price series consist of daily, weekly, monthly and yearly while Pakistan has only monthly data. For analysis all crude oil price time series were distributed into two different groups such as training and testing. The training set consists of the first 80 percent of the total observations while the last 20 percent was used as a testing set for model evaluation.

1.8 Contribution of the Thesis

The contributions of this research study are listed as below:

i) Simulations were carried out to fix the threshold value of autocorrelation to divide the IMFs into stochastic and deterministic components using different number of observations with four different scenarios (Chapter 4 page 97).

ii) The new procedure of reconstruction of IMFs and trend function obtained from EEMD is developed (Chapter 5 page 123).

iii) The new models EEMD-ISD-ARIMA and EEMD-ISD-FFNN are developed for forecasting the crude oil price which empirically assured the improvement in accuracy (Chapter 6 page 171).

iv) In this study different crude oil price series is used including daily, weekly, monthly and yearly data of Brent, WTI and Pakistan to check the usefulness, generalizability and robustness of the proposed models (Chapter 6 page 172).

1.9 Organization of the Thesis

This research study is organized in seven chapters. The contents of each chapter are outlined as follows:

Chapter one is the introduction. In first section the chapter is introduced followed by the background of the study which shows an overview on time series analysis and the methods commonly considered to decompose the time series. Next, is the problem statement which outlined the problem to be solved in this thesis followed by the research questions which are to be answered followed by the objectives of this study to achieve. The next section presents the significance of this thesis followed by the scope of the study and then summary of the contribution of this thesis. Finally, the organization of the thesis is presented which outlined the whole chapters in detail.

Chapter two consists of the literature review. The chapter starts with the introduction which describes the chapter accordingly followed by different approaches used for forecasting the crude oil price including cause and effect and a univariate approach. The chapter further presents that this study is focused on the time series analysis so next are the reviews of different procedures used for modelling the time series including ARIMA, GARCH and FFNN models. Next section highlights the work on decomposition ensemble techniques including the EMD and EEMD followed by the reconstructed of IMFs and trend function obtained from EMD and EEMD. The data sets used in different studies are also presented with complete details followed by the critical review to identify the research gap of the study. The last section puts forward of the chapter summary.

Chapter three consists of the research methodology. The chapter starts with the introduction and describes the chapter accordingly followed by the mathematical formulation and steps involved in every technique and model used in this thesis like ARIMA, GARCH, FFNN and EEMD. Next, are the methodologies for reconstruction of IMFs and trend function, simulations. The chapter further explains the chosen tests and proposed models. The real-world crude oil price data are also described in this chapter followed by the forecasting accuracy measures which are RMSE and MAPE. The last section presents the summary of the chapter.

Chapter four consists of the threshold value determination. The first section introduces the chapter contents followed by the experimental setup which describes the different scenarios used for data generation for simulations. Next, is discussed the computation and validation of the threshold values using four different scenarios. The last section is the summary of the chapter.

Chapter five consists of the real-world application using stochastic and deterministic components. The first section introduces the chapter contents followed by the data used in this research. Next, are the decomposition of daily, weekly, monthly and yearly data by EEMD. After the decomposition the reconstruction of IMFs and trend function outlined in detail and divide the IMFs into stochastic and deterministic components. The chapter further explains and analyse the stochastic and deterministic componetnts followed by the forecasting accuracy measures which are RMSE and MAPE. The last section presents the summary of the chapter.

Chapter six consists of the analysis. The chapter starts with the introduction and describes the chapter accordingly followed by the ARIMA modelling approach consists of identification, diagnostic checking, estimation, forecasting and evaluation of different models. Next is the ANN approach describe the steps involves in FFNN modelling including the transfer function and number of hidden nodes. The chapter further consists of analysis of all data followed by the forecasting accuracy measures which are RMSE and MAPE. The last section presents the summary of the chapter.

Chapter Seven concludes the thesis by summarizing the results, discussing the conclusion, limitations and future work.

REFERENCES

Abosedra, S., & Baghestani, H. (2004). On the predictive accuracy of crude oil futures prices. *Energy Policy, 32*(12), 1389-1393.

Agnolucci, P. (2009). Volatility in crude oil futures: a comparison of the predictive ability of GARCH and implied volatility models. *Energy Economics, 31*(2), 316-321.

Ahmad, M. (2012). Modelling and forecasting Oman crude oil prices using Box–Jenkins techniques. *International Journal of Trade and Global Markets, 5*(1), 24-30.

Ahmed, R. A., & Shabri, A. B. (2013). Fitting GARCH models to crude oil spot price data. *Lıfe Science Journal, 10*(4).

Ahmed, R. A., & Shabri, A. B. (2014). Daily crude oil price forecasting model using arima, generalized autoregressive conditional heteroscedastic and support vector machines. *American Journal of Applied Sciences, 11*(3), 425-432.

Alquist, R., & Kilian, L. (2010). What do we learn from the price of crude oil futures? *Journal of Applied Econometrics, 25*(4), 539-573.

Alvarez-Ramirez, J., Soriano, A., Cisneros, M., & Suarez, R. (2003). Symmetry/anti-symmetry phase transitions in crude oil markets. *Physica A: Statistical Mechanics and its Applications, 322*, 583-596.

An, X., Jiang, D., Zhao, M., & Liu, C. (2012). Short-term prediction of wind power using EMD and chaotic theory. *Communications in Nonlinear Science and Numerical Simulation, 17*(2), 1036-1042.

Ariño, M. A., & Marmol, F. (2004). A permanent-transitory decomposition for ARFIMA processes. *Journal of Statistical Planning and Inference, 124*(1), 87-97.

Azadeh, A., Moghaddam, M., Khakzad, M., & Ebrahimipour, V. (2012). A flexible neural network-fuzzy mathematical programming algorithm for improvement of oil price estimation and forecasting. *Computers & Industrial Engineering, 62*(2), 421-430.

Baffes, J., Kose, M. A., Ohnsorge, F., & Stocker, M. (2015). The great plunge in oil prices: Causes, consequences, and policy responses. *Consequences, and Policy Responses (June 2015)*.

Beale, M. H., Hagan, M. T., & Demuth, H. B. (2012). *Neural network toolbox™ user's guide*. Paper presented at the R2012a, The MathWorks, Inc., 3 Apple Hill Drive Natick, MA 01760-2098,, www. mathworks. com.

Bloomfield, P. (2004). *Fourier analysis of time series: an introduction*: John Wiley & Sons.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics, 31*(3), 307-327.

Box, G. E., & Jenkins, G. M. (1976). *Time series analysis: forecasting and control, revised ed*: Holden-Day.

Box, G. E., Jenkins, G. M., & Reinsel, G. C. (2011). *Time series analysis: forecasting and control* (Vol. 734): John Wiley & Sons.

Box, G. E., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). *Time series analysis: forecasting and control*: John Wiley & Sons.

Brockwell, P. J., & Davis, R. A. (2013). *Time series: theory and methods*: Springer Science & Business Media.

Bui, D. T., Pradhan, B., Lofman, O., Revhaug, I., & Dick, O. B. (2012). Landslide susceptibility assessment in the Hoa Binh province of Vietnam: a comparison of the Levenberg–Marquardt and Bayesian regularized neural networks. *Geomorphology, 171*, 12-29.

Caballero, J., & Fernández, M. (2008). Artificial neural networks from MATLAB® in medicinal chemistry. Bayesian-regularized genetic neural networks (BRGNN): Application to the prediction of the antagonistic activity against human platelet thrombin receptor (PAR-1). *Current topics in medicinal chemistry, 8*(18), 1580-1605.

Chang, C.-L., McAleer, M., & Tansuchat, R. (2010). Analyzing and forecasting volatility spillovers, asymmetries and hedging in major oil markets. *Energy Economics, 32*(6), 1445-1455.

Chatfield, C. (1996). Model uncertainty and forecast accuracy. *Journal of Forecasting, 15*(7), 495-508.

Cheong, C. W. (2009). Modeling and forecasting crude oil markets using ARCH-type models. *Energy Policy, 37*(6), 2346-2355.

Chinn, M. D., LeBlanc, M., & Coibion, O. (2005). *The predictive content of energy futures: an update on petroleum, natural gas, heating oil and gasoline*. Retrieved from http://www.nber.org/papers/w11033

Chiroma, H., Abdulkareem, S., & Herawan, T. (2015). Evolutionary Neural Network model for West Texas Intermediate crude oil price prediction. *Applied Energy, 142*, 266-273.

Darbellay, G. A., & Tichavsky, P. (2000). *Independent component analysis through direct estimation of the mutual information.* Paper presented at the ICA.

Dawson, C., & Wilby, R. (2001). Hydrological modelling using artificial neural networks. *Progress in physical Geography, 25*(1), 80-108.

Dées, S., Gasteuil, A., Kaufmann, R., & Mann, M. (2008). Assessing the factors behind oil price changes. *European Central Bank Working Paper*, 4-36.

Eckmann, J.-P., Kamphorst, S. O., & Ruelle, D. (1987). Recurrence plots of dynamical systems. *EPL (Europhysics Letters), 4*(9), 973.

Elliott, G., Rothenberg, T. J., & Stock, J. H. (1996). Efficient tests for an autoregressive unit root. *Econometrica, 64*(4), 813-836.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica: Journal of the Econometric Society*, 987-1007.

Flandrin, P., & Goncalves, P. (2004). Empirical mode decompositions as data-driven wavelet-like expansions. *International Journal of Wavelets, Multiresolution and Information Processing, 2*(04), 477-496.

Flandrin, P., Rilling, G., & Goncalves, P. (2004). Empirical mode decomposition as a filter bank. *IEEE signal processing letters, 11*(2), 112-114.

Goswami, J., Chan, A., & Chui, C. (1995). On a spline-based fast integral wavelet transform algorithm *Ultra-Wideband, Short-Pulse Electromagnetics 2* (pp. 455-463): Springer.

Gunay, S. (2015). Markov Regime Switching GARCH Model and Volatility Modeling for Oil Returns. *International Journal of Energy Economics and Policy, 5*(4).

Guo, Z., Zhao, W., Lu, H., & Wang, J. (2012). Multi-step forecasting for wind speed using a modified EMD-based artificial neural network model. *Renewable Energy, 37*(1), 241-249.

Hamilton, J. D. (1994). *Time series analysis* (Vol. 2): Princeton university press Princeton.

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

Harvey, A. C. (1990). *Forecasting, structural time series models and the Kalman filter*: Cambridge university press.

He, C., & Ma, C. (2010). A smoothing self-adaptive Levenberg–Marquardt algorithm for solving system of nonlinear inequalities. *Applied Mathematics and Computation, 216*(10), 3056-3063.

He, K., Yu, L., & Lai, K. K. (2012). Crude oil price analysis and forecasting using wavelet decomposed ensemble model. *Energy, 46*(1), 564-574.

Herrera, A. M., Hu, L., & Pastor, D. (2014). Forecasting crude oil price volatility.

Hirschen, K., & Schäfer, M. (2006). Bayesian regularization neural networks for optimizing fluid flow processes. *Computer methods in applied mechanics and engineering, 195*(7-8), 481-500.

Hou, A., & Suardi, S. (2012). A nonparametric GARCH model of crude oil price return volatility. *Energy Economics, 34*(2), 618-626.

Hsu, L.-C., & Wang, C.-H. (2009). Forecasting integrated circuit output using multivariate grey model and grey relational analysis. *Expert Systems with Applications, 36*(2), 1403-1409.

Huang, N. E., Shen, Z., & Long, S. R. (1999). A new view of nonlinear water waves: The Hilbert Spectrum 1. *Annual review of fluid mechanics, 31*(1), 417-457.

Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., . . . Liu, H. H. (1998). *The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis.* Paper presented at the Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences.

Huang, N. E., & Wu, Z. (2008). A review on Hilbert-Huang transform: Method and its applications to geophysical studies. *Reviews of Geophysics, 46*(2).

Hyndman, R. J., & Khandakar, Y. (2007). *Automatic time series for forecasting: the forecast package for R*: Monash University, Department of Econometrics and Business Statistics.

Jammazi, R., & Aloui, C. (2012). Crude oil price forecasting: Experimental evidence from wavelet decomposition and neural network modeling. *Energy Economics, 34*(3), 828-841.

Jiang, W., Zhang, L., & Wang, P. (2008). Nonlinear time series forecasting of time-delay neural network embedded with Bayesian regularization. *Applied Mathematics and Computation, 205*(1), 123-132.

Kaboudan, M. (2001). *Compumetric forecasting of crude oil prices.* Paper presented at the Evolutionary Computation, 2001. Proceedings of the 2001 Congress on.

Kang, S. H., Kang, S.-M., & Yoon, S.-M. (2009). Forecasting volatility of crude oil markets. *Energy Economics, 31*(1), 119-125.

Kermani, B. G., Schiffman, S. S., & Nagle, H. T. (2005). Performance of the Levenberg–Marquardt neural network training method in electronic nose applications. *Sensors and Actuators B: Chemical, 110*(1), 13-22.

Khan, N., Gaurav, D., & Kandl, T. (2013). Performance evaluation of Levenberg-Marquardt technique in error reduction for diabetes condition classification. *Procedia Computer Science, 18*, 2629-2637.

Kraskov, A., Stögbauer, H., & Grassberger, P. (2004). Estimating mutual information. *Physical review E, 69*(6), 066138.

Kulkar, S., & Haidar, I. (2009). Forecasting model for crude oil price using artificial neural networks and commodity future prices. *International Journal of Computer Science and Information Security, 2*(1), 81-88.

Kumar, P., Merchant, S., & Desai, U. B. (2004). Improving performance in pulse radar detection using Bayesian regularization for neural network training. *Digital signal processing, 14*(5), 438-448.

Lewis, C. D. (1982). *Industrial and business forecasting methods: A practical guide to exponential smoothing and curve fitting*: Butterworth-Heinemann.

Li, X., Yu, L., Tang, L., & Dai, W. (2013). *Coupling firefly algorithm and least squares support vector regression for crude oil price forecasting.* Paper presented at the Business Intelligence and Financial Engineering (BIFE), on Sixth International Conference

Lin, A. (2009). *Prediction of international crude oil futures price based on GM (1, 1)*. Paper presented at the Grey Systems and Intelligent Services, 2009. GSIS in IEEE International Conference

Liu, L. M. (1991). Dynamic relationship analysis of US gasoline and crude oil prices. *Journal of Forecasting, 10*(5), 521-547.

Ljung, G. M., & Box, G. E. (1978). On a measure of lack of fit in time series models. *Biometrika, 65*(2), 297-303.

Luukko, P., Helske, J., & Räsänen, E. (2016). Introducing libeemd: A program package for performing the ensemble empirical mode decomposition. *Computational Statistics, 31*(2), 545-557.

Marwan, N., Romano, M. C., Thiel, M., & Kurths, J. (2007). Recurrence plots for the analysis of complex systems. *Physics reports, 438*(5), 237-329.

Marzo, M., & Zagaglia, P. (2010). Volatility forecasting for crude oil futures. *Applied Economics Letters, 17*(16), 1587-1599.

Mirmirani, S., & Li, H. C. (2004). A comparison of VAR and neural networks with genetic algorithm in forecasting price of oil. *Advances in Econometrics, 19*, 203-223.

Montgomery, D. C., Jennings, C. L., & Kulahci, M. (2015). *Introduction to time series analysis and forecasting*: John Wiley & Sons.

Moshiri, S., & Foroutan, F. (2006). Forecasting nonlinear crude oil futures prices. *The Energy Journal*, 81-95.

Movagharnejad, K., Mehdizadeh, B., Banihashemi, M., & Kordkheili, M. S. (2011). Forecasting the differences between various commercial oil prices in the Persian Gulf region by neural network. *Energy, 36*(7), 3979-3984.

Mukherjee, I., & Routroy, S. (2012). Comparing the performance of neural networks developed by using Levenberg–Marquardt and Quasi-Newton with the gradient descent algorithm for modelling a multiple response grinding process. *Expert Systems with Applications, 39*(3), 2397-2407.

Ng, H., & Lam, K.-P. (2006, 2006). *How Does Sample Size Affect GARCH Models?*

Pan, H., Haidar, I., & Kulkarni, S. (2009). Daily prediction of short-term trends of crude oil prices using neural networks exploiting multimarket dynamics. *Frontiers of Computer Science in China, 3*(2), 177-191.

Papana, A., & Kugiumtzis, D. (2008). Evaluation of mutual information estimators on nonlinear dynamic systems. *arXiv preprint arXiv:0809.2149*.

Piotrowski, A. P., & Napiorkowski, J. J. (2011). Optimizing neural networks for river flow forecasting–Evolutionary Computation methods versus the Levenberg–Marquardt approach. *Journal of hydrology, 407*(1-4), 12-27.

Rast, M. (2001). *Fuzzy neural networks for modelling commodity markets.* Paper presented at the IFSA World Congress and 20th NAFIPS International Conference, 2001. Joint 9th.

Rios, R. A., & De Mello, R. F. (2013). Improving time series modeling by decomposing and analyzing stochastic and deterministic influences. *Signal Processing, 93*(11), 3001-3013.

Sadorsky, P. (2006). Modeling and forecasting petroleum futures volatility. *Energy Economics, 28*(4), 467-488.

Schlotthauer, G., Torres, M. E., & Rufiner, H. L. (2009). *A new algorithm for instantaneous F 0 speech extraction based on ensemble empirical mode decomposition.* Paper presented at the Signal Processing Conference, 2009 17th European.

Schwarz, G. (1978). Estimating the dimension of a model. *The annals of statistics, 6*(2), 461-464.

Shannon, C. E. (2001). A mathematical theory of communication. *ACM SIGMOBILE Mobile Computing and Communications Review, 5*(1), 3-55.

Sheela, K. G., & Deepa, S. N. (2013). Review on methods to fix number of hidden neurons in neural networks. *Mathematical Problems in Engineering, 2013*.

Shibata, R. (1976). Selection of the order of an autoregressive model by Akaike's information criterion. *Biometrika*, 117-126.

Shrestha, R. R., Theobald, S., & Nestmann, F. (2005). Simulation of flood flow in a river system using artificial neural networks. *Hydrology and Earth System Sciences Discussions, 9*(4), 313-321.

Shu-ping, W., Ai-mei, H., Zhen-xin, W., Ya-qing, L., & Xiao-wei, B. (2014). Multiscale combined model based on run-length-judgment method and its application in oil price forecasting. *Mathematical Problems in Engineering, 2014*.

Shukur, O. B., & Lee, M. H. (2015). Daily wind speed forecasting through hybrid KF-ANN model based on ARIMA. *Renewable Energy, 76*, 637-647. doi:http://dx.doi.org/10.1016/j.renene.2014.11.084

Singh, N., Singh, A., & Tripathy, M. (2012). Selection of hidden layer neurons and best training method for ffnn in application of long term load forecasting. *Journal of Electrical Engineering, 63*(3), 153-161.

Tang, B., Dong, S., & Song, T. (2012). Method for eliminating mode mixing of empirical mode decomposition based on the revised blind source separation. *Signal Processing, 92*(1), 248-258.

Tang, L., Yu, L., Wang, S., Li, J., & Wang, S. (2012). A novel hybrid ensemble learning paradigm for nuclear energy consumption forecasting. *Applied Energy, 93*, 432-443.

Ticknor, J. L. (2013). A Bayesian regularized artificial neural network for stock market forecasting. *Expert Systems with Applications, 40*(14), 5501-5506.

Wang, J., Xu, W., Zhang, X., Bao, Y., Pang, Y., & Wang, S. (2010). Data Mining Methods for Crude Oil Market Analysis and Forecast. *Data Mining in Public and Private Sectors: Organizational and Government Applications: Organizational and Government Applications*, 184.

Wang, S. Y., L.A. Yu and K.K. Lai,. (2005). Crude oil price forecasting with TEI@ I methodology. *Journal of Systems Science and Complexity, 18*(2), 145-166.

Weiqi, L., Linwei, M., Yaping, D., & Pei, L. (2011). *An econometric modeling approach to short-term crude oil price forecasting.* Paper presented at the Control Conference (CCC), 2011 30th Chinese.

Wu, Z., Feng, J., Qiao, F., & Tan, Z.-M. (2016). Fast multidimensional ensemble empirical mode decomposition for the analysis of big spatio-temporal datasets. *Phil. Trans. R. Soc. A, 374*(2065), 20150197.

Wu, Z., & Huang, N. E. (2004). *A study of the characteristics of white noise using the empirical mode decomposition method.* Paper presented at the Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences.

Wu, Z., & Huang, N. E. (2009). Ensemble empirical mode decomposition: a noise-assisted data analysis method. *Advances in Adaptive Data Analysis, 1*(01), 1-41.

Wu, Z., & Huang, N. E. (2010). On the filtering properties of the empirical mode decomposition. *Advances in Adaptive Data Analysis, 2*(04), 397-414.

Wu, Z., Huang, N. E., Wallace, J. M., Smoliak, B. V., & Chen, X. (2011). On the time-varying trend in global-mean surface temperature. *Climate Dynamics, 37*(3-4), 759-773.

Xie, W., Yu, L., Xu, S., & Wang, S. (2006). A new method for crude oil price forecasting based on support vector machines *Computational Science–ICCS 2006* (pp. 444-451): Springer.

Xiong, T., Bao, Y., & Hu, Z. (2013). Beyond one-step-ahead forecasting: evaluation of alternative multi-step-ahead forecasting models for crude oil prices. *Energy Economics, 40*, 405-415.

Yan, Q., Wang, S., & Li, B. (2014). Forecasting uranium resource price prediction by extreme learning machine with empirical mode decomposition and phase space reconstruction. *Discrete Dynamics in Nature and Society, 2014*.

Ye, M., Zyren, J., & Shore, J. (2005). A monthly crude oil spot price forecasting model using relative inventories. *International Journal of Forecasting, 21*(3), 491-501.

Yi, Y., & Qin, N. (2009). *Oil price forecasting based on self-organizing data mining.* Paper presented at the Grey Systems and Intelligent Services, 2009. GSIS 2009. IEEE International Conference on.

Yonaba, H., Anctil, F., & Fortin, V. (2010). Comparing sigmoid transfer functions for neural network multistep ahead streamflow forecasting. *Journal of Hydrologic Engineering, 15*(4), 275-283.

Yu, L., Dai, W., & Tang, L. (2016). A novel decomposition ensemble model with extended extreme learning machine for crude oil price forecasting. *Engineering Applications of Artificial Intelligence, 47*, 110-121.

Yu, L., Wang, S., & Lai, K. K. (2008). Forecasting crude oil price with an EMD-based neural network ensemble learning paradigm. *Energy Economics, 30*(5), 2623-2635.

Yu, L., Wang, Z., & Tang, L. (2015). A decomposition–ensemble model with data-characteristic-driven reconstruction for crude oil price forecasting. *Applied Energy, 156*, 251-267.

Yu, L., Zhao, Y., & Tang, L. (2016). Ensemble Forecasting for Complex Time Series Using Sparse Representation and Neural Networks. *Journal of Forecasting.*

Zhang, G., Patuwo, B. E., & Hu, M. Y. (1998). Forecasting with artificial neural networks:: The state of the art. *International Journal of Forecasting, 14*(1), 35-62.

Zhang, G., Wu, Y., & Liu, Y. (2014). An advanced wind speed multi-step ahead forecasting approach with characteristic component analysis. *Journal of Renewable and Sustainable Energy, 6*(5), 053139.

Zhang, G. P. (2003). Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing, 50*, 159-175.

Zhang, G. P., & Qi, M. (2005). Neural network forecasting for seasonal and trend time series. *European journal of operational research, 160*(2), 501-514.

Zhang, J., Yan, R., Gao, R. X., & Feng, Z. (2010). Performance enhancement of ensemble empirical mode decomposition. *Mechanical Systems and Signal Processing, 24*(7), 2104-2123.

Zhu, B., Shi, X., Chevallier, J., Wang, P., & Wei, Y. M. (2016). An Adaptive Multiscale Ensemble Learning Paradigm for Nonstationary and Nonlinear Energy Price Time Series Forecasting. *Journal of Forecasting*.