HANDLING ARCH EFFECTS IN WIND SPEED DATA USING STATE SPACE APPROACH MODEL

‘AAISHAH RADZIAH BINTI JAMALUDIN

UNIVERSITI TEKNOLOGI MALAYSIA

HANDLING ARCH EFFECTS IN WIND SPEED DATA USING STATE SPACE APPROACH MODEL

‘AAISHAH RADZIAH BINTI JAMALUDIN

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

Faculty of Science
Universiti Teknologi Malaysia

FEBRUARY 2017

To my beloved husband, *Islahuddin Md Sahat*,
my charming prince, *Umar Affan* and
my sweet little munchkin, *Naurah Hannan*.

ACKNOWLEGEMENT

In preparing this thesis I was in contact with many people, researchers, academicians, and practitioners. They have contributed towards my understanding and thoughts. In particular, I wish to express my sincere appreciation to my main thesis supervisor Prof Madya Dr. Fadhilah Yusof , for encouragement, guidance, critics and friendship. I am also very thankful to my co-supervisor, Dr Ibrahim Lawal Kane for his guidance, advices and motivation. Without their continued support and interest, this thesis would not have been the same as presented here.

I am also indebted to Universiti Teknologi Malaysia (UTM) for funding my Master study. Librarians at UTM also deserve special thanks for their assistance in supplying relevant literatures.

My beloved husband should also be recognised for his constant support and motivation throughout the journey. My sincere appreciation also extends to all my colleagues and others who have provided assistance at various occasions. Their views and tips are useful indeed. Unfortunately, it is not possible to list all of them in this limited space. I am grateful to all my family members.

ABSTRACT

In general, Malaysia experiences low wind speed, but some areas in this country experience strong wind in certain periods of time within a year. In line with the necessity to enhance the utilization of indigenous renewable energy resources in order to contribute towards national electricity supply, the study on the potential of the wind speed as a new source of renewable energy is significantly crucial. For that reason, this study aims to model and forecast wind speed data in 10 stations all over Peninsular Malaysia by using three different methods; Autoregressive Moving Average (ARMA), hybrid model (ARMA with Generalized Autoregressive Conditional Heteroscedasticity (ARMA-GARCH)) and Dynamic Linear models (DLM). ARMA was used as the benchmark in identifying an adequate linear model. The Autoregressive Conditional Heteroscedasticity (ARCH) effect in the residuals data from the developed conventional model was determined. The presence of ARCH shows that the model is not appropriate to be treated as a linear model. Therefore, to overcome this problem, ARMA model was hybridized with GARCH model. However, there is still some remaining ARCH exists in the residuals data for several datasets. Thus, a new approach namely DLM was introduced in order to treat the shortcoming. At the end of the research, a comparative study was made. It was discovered that in most cases, DLM outperforms than other models. DLM is found to be flexible in treating the dynamical fluctuation of the data and superior in terms of predictive accuracy with just a small error when compared with other methods.

ABSTRAK

Secara umum, Malaysia mengalami kelajuan angin yang rendah, tetapi di sesetengah kawasan dalam negara ini mengalami tiupan angin yang kuat dalam tempoh masa tertentu di sepanjang tahun. Selari dengan keperluan untuk menggalakkan penggunaan sumber tenaga diperbaharui dalam menyumbang kepada bekalan tenaga elektrik nasional, kajian ke atas potensi kelajuan angin sebagai satu sumber baru bagi tenaga diperbaharui adalah sangat penting. Justeru itu, kajian ini bertujuan untuk menghasilkan model dan meramal data kelajuan angin bagi 10 stesen di Semenanjung Malaysia dengan menggunakan tiga kaedah berbeza; Autoregresif Purata Bergerak (ARMA), model hybrid (ARMA dan model Teritlak Autoregresif Bersyarat Heteroskedastisiti (GARCH)) dan Model Linear Dinamik (DLM). ARMA telah digunakan sebagai penanda aras dalam mengenalpasti model linear yang sesuai. Kesan Autoregresif Bersyarat Heteroskedastisiti (ARCH) dalam baki data daripada model konvensional yang dibina telah ditentukan. Kehadiran kesan ARCH menunjukkan model tersebut tidak sesuai dibina sebagai model linear. Oleh itu, untuk mengatasi masalah ini, model ARMA telah digabungkan dengan model GARCH. Walaubagaimanapun, masih terdapat baki kewujudan ARCH dalam baki data pada sesetengah set data. Justeru itu, suatu pendekatan baru dinamakan DLM diperkenalkan bagi mengatasi masalah ini. Pada akhir kajian, suatu kajian perbandingan telah dilakukan. Ia telah ditemui dalam kebanyakan kes, DLM lebih baik berbanding model lain. DLM didapati fleksibel dalam merawat perubahan dinamik data dan lebih baik dari segi ketepatan ramalan dengan ralat yang kecil dibandingkan dengan kaedah lain.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ACF	-	autocorrelation function
ADF	-	Augmented Dicker-Fuller
AIC	-	Akaike's Information Criterion
ANN	-	Artificial Neural Networks
ARCH	-	Autoregressive Conditional Heteroscedasticity
ARFIMA	-	Autoregressive Fractional Integrated Moving Average
ARIMA	-	Autoregressive Integrated Moving Average
ARMA	-	Autoregressive Moving Average
BPNN	-	Back Propagation Neural Network
BSO	-	brainstorm optimization
CC	-	cross correlation
CCM	-	Cumulative Climate Memory
CEC	-	Control Energy Centre
Coeff.	-	Coefficient
CVE	-	cross-validation estimators
DEVS	-	Discrete Event System Specification
DFA	-	De-trended fluctuation analysis
DFM	-	dynamic factor model
DIR	-	Directional-bias forecasts
DLM	-	Dynamic Linear Model
DRL	-	diurnal cycle forecast correction
E	-	East
e.g	-	example
EMD	-	empirical mode decomposition

ESNs	-	echo state networks
ET0	-	evapotranspiration
EWMA	-	European Wind Energy Association
FISM	-	Fractional Integral Statistical Model
FVMF	-	finite von Mises–Fisher
G7	-	The Group of Seven
GARCH	-	Generalised Autoregressive Conditional Heteroscedasticity
GARMA	-	Gegenbauer autoregressive moving average
GDP	-	Gross Domestic Product
GLS	-	Gaussian linear state-space
GP	-	Genetic Programming
GPOK	-	Genetic-Programming-Based-Ordinary-Kriging
GPP	-	Gaussian predictive processes
KAL	-	Kalman Filter Forecasts
LARIMA	-	limited Autoregressive Integrated Moving Average
LB	-	Ljung Box
LLS	-	Linear Least Square corrected forecasts
LM	-	Langrage Multiplier
MAPE	-	Mean Absolute Percentage Error
MAV	-	Mean and Variance corrected forecast
MK	-	Mann Kendall
MKD	-	Mann Kendall Data
MKDD	-	Mann Kendall Detrended Data
MKRD	-	Mann Kendall Rank Detrended
ML	-	Maximum Likelihood
MLCV	-	Maximum Likelihood Cross Validation
ML-DFA	-	Maximum Likelihood - Detrended Fluctuation Analysis
MMD	-	Malaysian Meteorological Department
MSE	-	Mean Square Error
MW	-	MegaWatt
N	-	North
NZWEA	-	New Zealand Wind Energy Association
OK	-	Ordinary Kriging

OLS	-	Ordinary Least Square
PACF	-	Partial autocorrelation function
PSO	-	particle swarm optimization
RAMS	-	Regional Atmospheric Modeling System
RK	-	regression kriging
RKNNRK	-	Regression Kriging and neural network residual kriging
RMSE	-	root mean square error
RP	-	recurrence plot
SC	-	spectral clustering
Seda	-	Sustainable Energy Development Authority Malaysia
SK	-	Simple Kriging
SOC	-	surface ozone concentrations
SSA	-	singular spectrum analysis
STB	-	Short-term bias-correction forecasts
Std. errror	-	Standard error
STEO	-	Short-term Energy Outlook
STT	-	Sen trend test
sVAR	-	Sparse Vector Autoregression
SVM	-	Support Vector Machine
SVR	-	support vector regression
TDD	-	trigonometric direction diurnal
TES	-	Triple Exponential Smoothing
TF	-	trachomatous inflammation–follicular
TI	-	trachomatous inflammation-intense
TK	-	Taylor Kriging
UK	-	universal kriging
UKF	-	unscented Kalman filter
US	-	United States
VARTA	-	Vector-Autoregressive-To-Anything
WP	-	water parameter
WSE	-	Weather-Scale Excitation

LIST OF SYMBOLS

$g(h)$	-	Exponential model
Δy_t	-	The differenced series
y_{t-1}	-	Immediate previous observation
x_t	-	Optional exogenous regressor
$\beta_1,....,\beta_p$	-	Coefficients of the lagged difference term up to lag p
ε_t	-	Error term
α, β	-	Parameter of GARCH model
$C_R(h)$	-	Covariance function of the residual component of the variable
Y_t	-	Observation equation
r_k	-	Sample autocorrelation function
r_k	-	Sample partial autocorrelation function
s_0	-	Prediction location
s^2 @ σ^2	-	Variance
$\bar{x}$ @ μ	-	Mean of the sample
y_t	-	A sequence of random variables indexed by some time subscript *t* known as time series
$\hat{\varepsilon}(k)$	-	McLeod-Li test
λ_i	-	Unknown weight for the measured value at the *i*-th location
μ_t	-	State equation
$\psi_\tau(0,1)$	-	The probability density function of the residuals with 0 mean and variance 1
∞	-	Infinity

a	-	Practical range
α and δ	-	Parameters to be estimated
c	-	Sill
C_0	-	Variance of the prior distribution
D	-	Fractal dimension
f	-	Frequency
FF	-	Covariates
GG	-	Evolution
h	-	Lag distance
H	-	Hurst exponent
h	-	Distance
H_0, H_1	-	Probability level
L	-	Number of autocorrelation lags included in the statistic
$L(k)$	-	Gradually varying function at infinity
$L(t)$	-	Slow-varying function
m	-	Length of boxes
M_0	-	Mean of the prior distribution
n	-	Sample size
N	-	Number of measured values
N	-	Sample size
p,q	-	Parameter estimation
s	-	Standard deviation
s_i	-	Recognized points
t	-	Time
T_j and T_i	-	Daily values in days j and i
V	-	Variance of observation
W	-	Variance of evolution
x	-	Observed values
x_i	-	Original series
$y(k)$	-	Integrated series
z	-	Contiguous data points
Z_s	-	Standard test statistic
α	-	The degree of long memory

$\gamma(k)$	-	Autocovariance function
L	-	Likelihood function
Q	-	Ljung Box statistics test
$g(f)$	-	Spectral density of a long-memory process
r	-	Signal-to-noise ratio
s	-	Mann-Kendall *S* Statistic
$z(s_i)$	-	Measured value in the *i*-th location
$\gamma(h)$	-	Semivariance
τ	-	The distributional parameter that define the shape of the distribution

LIST OF APPENDICES

The shortest distance between a problem and its solution

is the distance between your knees and the floor.

Be somebody in the eyes of ALLAH,

even if you are nobody in the eyes of people.

CHAPTER 1

INTRODUCTION

1.0 Introduction

The development of wind energy technology shows the fastest growing in the world. Wind energy offers many advantages. Currently, the wind energy usage was emphasized by the researchers all around the world. Wind as a fuel source free, does not contaminate the air like other power plants that certainly depends on burning of fossil fuels such as coal or natural gas. The installation of wind turbines helps our ecosystem in reducing the greenhouse effect as wind does not produce acid rain or greenhouse gases. United States Energy Information Administration (2016) reported that in the last 10 years, wind power capacity has grown with an increment of 30% which represents 28% of worldwide capacity.

This sustainable source actually originated from the solar energy. The heat of the atmosphere which comes from the sun produced the wind. Another two factors that contribute to the production of wind are the earth's rotation and the geographical earth's surfaces. The energy could be assembled as long as the sun shines and the wind blows. Therefore, the energy generated from the wind turbines can send the power across the grid along the times. In economical context, wind power is quite profitable. Among the renewable energy technologies today, wind is the lowest-priced with the cost estimation of four to six cents per kilowatt-hour. However, it relies on the wind resource and its particular financial projects. Moreover, referring to the Wind Vision Report (US Department of Energy), wind has a bright prospective to support more than 600,000 jobs in manufacturing, installation, maintenance and supporting services by 2050.

Recent statistics show that by the year 2050, the total population of Malaysian is estimated to be 40.7 million people (Bujang *et al*., 2015). Energy has become as fundamental needs of the development in Malaysia. Hence, the demand of energy is also projected to increase. Practically, energy is essential in all aspect of everyday life including agricultural, drinking water, lighting health care, telecommunication, and industrial activities. Wind can be seen as a suitable resource of energy since Malaysia experience continuous wind and it also produce minor impacts on the environment. As the rising bids of wind energy, it is essential for power utilities to plan the assimilation of wind power. Wind with the intermittency and random character makes the modeling and the forecasting of wind speed more valuable and it can be a support tool for the operators of the Control Energy Centre (CEC) of the power utility (Cadenas and Rivera, 2010). Hence, an accurate measurement of wind speed prediction is preferred for providing an overview on how the behavior and trend of historical pattern and future projected pattern of wind speed could be, which the power energy could be estimated from. Wind power is extensively exploited in the countries like Germany, Denmark and Spain. As an example, it was reported that more than 4% of electricity in Spain comes from the wind source (Sanchez, 2006). However in Malaysia the installation of wind turbine is still in the research stage.

As reported by the star online 28th October 2013, the Sustainable Energy Development Authority Malaysia (Seda) was considered to include wind as an alternative source in the renewable energy scheme. Hence, a comprehensive study and thorough analysis on the availability of wind energy in Malaysia is immediately essential to decide whether wind should be included as another renewable resource in the scheme. Thus, this study provides a dynamical model capable in describing and forecasting the wind speed data sets of 10 stations across peninsular Malaysia. The assessments of the wind models can help to express the dynamic on how much wind energy can be produced at a particular area.

Therefore, the characteristic of Peninsular Malaysia daily wind speed will be explored intensively and comprehensively in this study and both criteria will be considered; spatial and temporal properties. This will enhance the understanding of the process for the effective development and execution of wind energy installation as one of the renewable energy resources.

1.1 Background of the study

In Malaysia, the highest maximum wind speed was recorded at Kuching, Sarawak on 15 September 1992 with 41.7 m/s whereas the highest mean daily wind speed with 3.8 m/s was recorded at Mersing, Johor. As a whole, Malaysia experiences low wind speed. However certain areas in this country recorded high readings. Therefore it is essential to study the Malaysia's wind speed behavior and identify the potential area to be developed with the wind energy generation system. At Malaysian meteorological department, the wind speed data is measured by using anemometers (Figure 1.1), the most common instrument of which the cups are mounted symmetrically from the right angle to the vertical shaft. Basically the operation of this instrument is based on the difference between the wind pressures from one side of the cup to the other which causes the cup to spin about the shaft. The rate of the rotation is directly proportional to the speed of wind.

Figure 1.1 Anemometers

(Picture courtesy from *http://www.met.gov.my*)

Hafezatul (2010) found that from 22 weather stations studied across Peninsular Malaysia, Sabah and Sarawak, there are only two types of wind; weak wind (discomfort thermal) and strong wind (comfort thermal). In terms of building and construction, for different regional areas different types of buildings need to be considered based on the wind speed flow. For southern part of peninsular Malaysia the construction of high-rise building is quite appropriate since this location was recorded with the weak wind condition. The rest of the areas are considered with the comfort thermal condition which are suitable for terrace housing and low-rise building.

Nowadays, researchers all around the world have focused on an alternative in the production of clean and renewable energy that comes from resources which are naturally replenished on a human timescale. Wan *et al*., (2011) explored the wind energy potential at east coast of peninsular Malaysia and found that the wind power of study site (Terengganu, 4^{o}13.6'N and 103^{o}26.1' E) is lowest during the south west monsoon season, while it is highest during the north east monsoon season with the value of 84.60 W/m^2. Therefore, small wind turbines could be used in order to deliver the energy power which is particularly generated during the north east monsoon which starts from late November and ends in early March. Nurhayati (2010) and Siti *et al*. (2011) discovered that Mersing is the most potential site for installing wind turbine due to peak mean wind speed during the northeast monsoon with approximately of 62 W/m^2. This reveals that there is a high potential on applying the small scale wind turbine system at Mersing for power generation purposes.

Furthermore, wind energy is an ideal alternative energy source which is rapidly developed all around the world. As an example in China, it was reported that the total current capacity of wind farms is 25805.3 MW, with a progress rate of 114% in 2009. As the applications of the wind energy grows, it is essential for power utilities to plan the assimilation of wind power and other traditional powers. Hence, an accurate measurement of wind speed prediction is preferred (Hui *et al*. 2010).

The research flow of this study can be illustrated in Figure 1.2. Firstly, the wind speed data will go through the pre-processing analysis starting with spatial analysis known as Kriging. It is then followed by the data explanatory analysis comprises of descriptive statistics, checking of long memory and Mann Kendall trend test. After all the pre-processing analysis is done, the modelling process begins. All wind speed datasets will be modelled using conventional models; ARMA. The error component will be tested with Mc-Leod Li test (ML test) to check the presence of ARCH. If ARCH does not exist, the model is adequate. But if ARCH exists, the model is inadequate. Therefore, the ARCH effects need to be treated. In this study, GARCH model will be hybridized into the conventional ARMA.

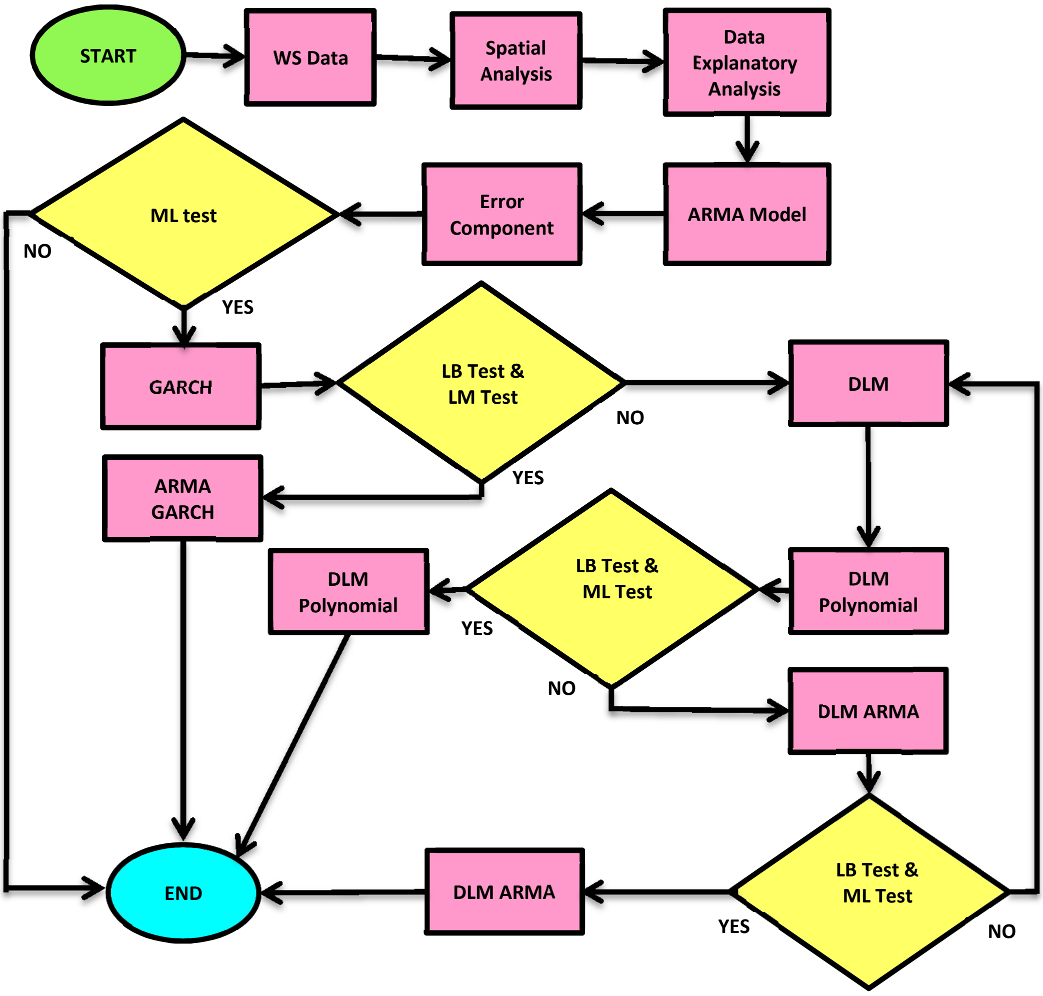

Figure 1.2 Flowchart summarizing the methodological framework

The adequacy of ARMA-GARCH will be checked by Ljung-Box (LB) test and Lagrange-Multiplier (LM) test. If the model is adequate, the end result will be ARMA-GARCH model. However, if there is still excessive ARCH in ARMA-GARCH, then Dynamic Linear Model (DLM) will be introduced. The first class of DLM introduced is DLM Polynomial. If the model is adequate, the end result is DLM Polynomial. If not, we introduced the second class of DLM; DLM-ARMA. The adequacy checking by LB test and ML test will be carried out. If the model is adequate, the end result is DLM-ARMA model.

1.2 Problem Statement

Currently, the production of electrical energy sources in Malaysia was entirely fabricated from water basis. Even though water is a ceaseless source, other renewable energy alternatives are still needed in order to support the production of energy for occupying people's necessities. In line with the save our earth campaign, one alternative that can be seen helpful and effective is through the energy generated from wind speed since wind is natural, continuous and clean source.

However, researches in this area are still limited, not much information can be acquired in Malaysia. Modeling wind speed data may become a very thoughtful task due to an erratic of natural phenomena. As a substitutable, time series models are used. Time series models normally have a comparatively simple structure that integrates the component of uncertainty in the outcome. This uncertainty signifies the portion of the process that cannot be explained deterministically. Time series models are aimed to duplicate the imperative patterns of evident in the data set based on the present information of the physical processes.

Wind with the volatility characteristics may become a challenging task in the data analysis process. Therefore a good analysis of the data is vital to apprehend the entire behaviour of the wind speed data which lead to an accurate modeling and forecasting. Moreover, when it comes to the wind energy calculation, a precise and lesser error method is preferred. Therefore in this study, an attempt has been made in finding the best method to represent the wind speed model. The occurrence of Autoregressive Conditional Heteroscedasticity (ARCH) effect that was found in the residuals data is another problem that needs to be treated. An alternative has been studied critically in order to remove the ARCH effect. Hence, this study will explore the most suitable method that is able to overcome the problems of excessive ARCH effect in the ARMA-GARCH model.

Another aspect that needs to be considered is in terms of forecasting ability. Forecasting wind speed data commonly involves short-term forecast. Therefore a dynamic time series model that is capable to forecast in short term duration is essential. A model that is flexible in capturing any pattern of the data series was examined comprehensively. Furthermore, the analysis of wind speed time series with conventional models (e.g Ahmed *et al*, 2010) requires at least a preliminary transformation of the data to get stationarity; but we might feel more natural to have models which allow analyzing more directly data which show instability in the mean level and in the variance, without transformation. So this study hopefully will fulfill the researcher's essential and contributes a piece of knowledge in wind speed energy generated field in terms of modeling and forecasting. As a conclusion; three main problems was identified to be handled in this study. Firstly is in term of modeling Malaysia's wind speed using time series approach, secondly is in term of treating the excessive ARCH effect in the ARMA-GARCH model and thirdly is to select the best model that capable to do short-term forecasting.

1.3 Objectives of the study

The objectives of the research are:

(i) To investigate the wind speed trend in Peninsular Malaysia.
(ii) To propose the hybrid time series linear model
(iii) To propose state space approach known as Dynamic Linear Model (DLM)
(iv) To compare the performance of conventional and hybrid times series models with the state space approach model.
(v) To use the best model to forecast the wind speed data

1.4 Scope and limitations

The data used in this study involves daily wind speed data (secondary data) dated from January 1st 1985 to December 31st 2009 covering ten stations in Peninsular Malaysia only due to unavailability of complete wind speed stations. The unit of wind speed data is meter per second (m/s). However, the data contains missing values. Therefore the Mean Imputation method was used to fill up the missing values. The pre-processing method includes descriptive statistics, Ordinary Kriging, Long memory and Mann Kendall trend test. The tools that were applied throughout the analysis process comprises of R-software and Gretl software. This study focuses on short-term forecasting model only due to necessity of wind energy calculation. Wind farm managers and workers require immediate forecasting tools for the instantaneous plan of maintenance. It also can optimize the whole electricity supply, stabilize the load and supply of electricity and also improve the cost effectiveness of energy supply. The analysis of the wind speed data include the conventional method based on Box-Jenkins approach, hybrid with volatility model; Generalised Autoregressive Conditional Heteroscedasticity (GARCH) model and a proposed model; DLM that allows the short term forecasting method.

1.5 Significance of the research

Selecting the best model is important to forecast the wind speed data. An accurate forecasting model will reduce the planning errors and increase the reliability of the electric power grid (Masseran, 2016). This thesis aims to provide a short term forecasting model of the daily wind speed data. The prediction of wind speed from the best fitted model will provide an accurate view of future pattern and the calculation of energy generated can be estimated instantly for future usage. Therefore, several contributions will be highlighted here.

Firstly, this thesis examined the spatial variability of Peninsular Malaysia wind speed and areas with high, moderate and low wind speed event was recognized. The structure of wind speed data that was identified is important in deriving the wind speed patterns, measuring wind speed amounts in experimental locations and at the same time helping in classifying locations of high, moderate and low wind speed events. A thorough study is effectively beneficial for planners and other users. The aims of this research are to study the properties of wind speed time series through conventional and volatility models. Hence, through the details analysis the problem solving of excessive persistence can be identified. At the end of the study, a good forecasting model could be produced and able to describe the dynamics of wind speed data of the study areas.

The wind speed data even though performed well with the linear model but the checking of heteroscedasticity presence is necessary to reconfirm either the model constructed is adequate or not. If this effect exists, it indicates that the linear model is not suitable to fit the behavior of the data. So, in this research the dealing of the heteroscedasticity effect will be highlighted since this effect plays a major concern in regression analysis. The presence of heteroscedasticity effect can invalidate the statistical tests of significance.

Furthermore, a novel model approach was developed in this thesis that captures all the characteristics observed in wind speed time series. The thesis first considers the class of conventional Autoregressive Moving Average (ARMA) model with hybrid GARCH. However, from the analysis we found that the hybrid model ARMA-GARCH fails to capture the entire volatility of wind speed data as the remaining of ARCH effect still exists in the residuals data. Therefore, an alternative to overcome the excessive of ARCH effect in ARMA-GARCH model was presented by introducing Dynamic Linear Model (DLM) based on State Space approach that is capable to capture the volatility of the wind speed data without going through any transformation of the data. This approach was found to be easier and simpler yet flexible in any kind of condition. The dynamic properties of this method allow the model to follow closely the behavior of the wind speed data. Thus, a more accurate model was created in presenting daily wind speed data in Peninsular Malaysia.

Modeling Malaysian wind speed data could be a vital contribution since it provides an overview on how the trend and nature of the data. The trend of historical pattern will be exploited to project the short term future pattern which the instantaneous power energy could be estimated from. Hence, it will help the authorities in planning the total power grid and the effective cost in particular location for public usage. In other words, this paper aims to develop a model that signifies a prediction of the wind speed data and an attempt has been made in order to execute a formula that is capable to capture the future pattern of the data.

1.6 Organization of the thesis

The rest of the thesis is organized as follows: Chapter 2 presents a review of the related literature. Chapter 3 gives the methodology used in this thesis. The main contributions of this research start in Chapter 4. Chapter 4 explores the basic characteristics of the daily wind speed data sets and the spatial variability representing locations with high, moderate and low wind speed event. Chapter 5 presents a conventional wind speed model that describes the basic characteristics found in the wind speed data exploration and chapter 6 proposes a new approach that captures the limitations of the previous models. Chapter 7 closes the thesis with conclusions and recommendations for future research.

REFERENCES

Adhikary, S. K., Muttil, N., and Yilmaz, A. G. (2016). Genetic Programming-Based Ordinary Kriging for Spatial Interpolation of Rainfall. *Journal of Hydrologic Engineering*, 21(2), 1–14. American Society of Civil Engineers.

Agdas, D., Webster, G. D., and Masters, F. J. (2012). Wind Speed Perception and Risk. *Journal of Plos One*, 7(11), 1-8. PLOS.

Ahmed, M. A., Akhter, M. W., Hussain, G., and Ahmed, F. (2013). *Analysis and Forecasting Wind Speed At Six Locations of Sindh , Pakistan for Power Generation Using Box-Jenkins ARMA models.* Unpublished note, University of Karachi, Pakistan.

Amath, E. (2011). *Time Series Concepts*. Unpublished note, University of Washington.

Ambach, D. (2015). Short-Term Wind Speed Forecasting in Germany. *Journal of Applied Statistics*, 43(2), 351-369. Taylor & Francis Group.

Arandia, E., Ba, A., Eck, B., Asce, M., McKenna, S. (2015). Tailoring Seasonal Time Series Models to Forecast Short-term Water Demand. *Journal of Water Resource and Planning Management*, 142(3), 1-10. American Society of Civil Engineers.

Ashley, R. A., and Patterson, D. M. (2010). Apparent Long Memory in Time Series As an Artifact of a Time-Varying Mean: Considering Alternatives To the Fractionally Integrated Model Macroeconomic Dynamics. *Journal of Macroeconomics Dynamics*, 14(1), 59–87. Cambridge University Press.

Bakar, K. S. (2012). *Bayesian analysis of daily maximum ozone levels*. PhD Thesis Page 33, University of Southampton, England.

Balluff, S., Bendfeld, J., and Krauter, S. (2015). Short term wind and energy prediction for offshore wind farms using neural networks. *4th International Conference on Renewable Energy.Research and Applications*. 22-25 November 2015. Palermo, Itally, 379-382.

Bessac, J., Ailliot, P., and Monbet, V. (2015). Gaussian linear state-space model for wind fields in the North-East Atlantic. *Journal of Environmetrics*, 26(1) , 29–38. John Wiley & Sons.

Blain, G.C. (2013). The modified Mann-Kendall test: on the performance of three variance correction approaches. *Journal of Agrometeorology*,72(4), 416–425.

SciELO.

Bohling. G., (2005). *Introduction to Geostatistics and Variogram Analysis*. Unpublished note, Kansas Geological Survey.

Botcharova, M., Farmer, S. F., and Berthouze, L. (2013). *A maximum likelihood based technique for validating detrended fluctuation analysis (ML-DFA).* Unpublished note, Cornell University Library.

Bouaddi, S., Ihlal, A., and Fernández-García, A. (2015). Soiled CSP solar reflectors modeling using dynamic linear models. *Journal of Solar Energy*, 122, 847–863. Elsevier.

Brodsky, B., and Darkhovsky, B. (2012). Structural changes and unit roots in non-stationary time series. *Journal of Statistical Planning and Inference*, 142(1), 327–335. Elsevier.

Bujang, A.S., Bern, C.J. and Brumm, T.J.(2015). Summary of energy demand and renewable energy policies in Malaysia. *Renewable and Sustainable Energy Reviews*, 53, 1459-1467. Elsevier.

Cadenas, E. and Rivera. W. (2010). Wind speed forecasting in three different regions of Mexico, using a hybrid ARIMA-ANN model. *Journal of Renewable Energy*, 35, 2732–2738. Elsevier.

Candelieri, A., Soldi, D. and Archetti, F. (2015). Short-term Forecasting of Hourly Water Consumption by Using Automatic Metering Readers Data. *Journal of Procedia Engineering*, 119, 844-853. Elsevier.

Cellura, M., Cirrincione, G., Marvuglia, A. and Miraoui, A. (2008). Wind speed spatial estimation for energy planning in Sicily: Introduction and Statistical Analysis. *Journal of Renewable Energy*, 33(2008), 1237–1250. Elsevier.

Chan, N.H. and Palma,W.(1998). State Space Modeling of Long-Memory Processes. The Annals of Statistics, 26(2): 719–740. IMS.

Chen, H., Wan, Q., Li, F., and Wang, Y. (2013). GARCH in mean type models for wind power forecasting. *Power and Energy Society General Meeting*. 21-25 July 2013. Nanjing, China, 1-5.

Chen, K., and Yu, J. (2014). Short-term wind speed prediction using an unscented Kalman filter based state-space support vector regression approach. *Journal of Applied Energy*, 113, 690–705. Elsevier.

Chen, P., Pedersen, T., Bak-jensen, B. , and Chen, Z. (2010). ARIMA-Based Time Series Model of Stochastic Wind Power Generation. *IEEE Transactions on*

Power Systems .25(2), 667–676.

Chen, X., Jin, S., Qin, S., and Li, L. (2009). Short-Term Wind Speed Forecasting Study and Its Application Using a Hybrid Model Optimized by Cuckoo Search. *Journal of Mathematical Problems in Engineering*, *2015*,1-18. Hindawi Publishing Corporation.

Conor, P. S., Peter, L., and Paul, L. (2013). Reducing errors of Wind Speed forecasts by an optimal combination of post-processing methods. Journal of Meterological Applications, 20, 32-40. John Wiley & Sons.

Corbetta, M., Sbarufatti, C., Manes, A., and Giglio, M. (2014). On Dynamic State-Space models for fatigue-induced structural degradation. *International Journal of Fatigue*, 61, 202–219. Elsevier.

Cortez, P., Rocha, M., and Neves, J. (2001). Evolving time series forecasting neural network models. 1–25. Kluwer Academic Publishers.

Danneberg, J. (2012). Changes in runoff time series in Thuringia, Germany-Mann-Kendall trend test and extreme value analysis. *Journal of Advances in Geosciences*, 31, 49–56. Copernicus Publications.

Dissanayake, G. S., Peiris, M. S., and Proietti, T. (2014). State space modeling of Gegenbauer processes with long memory. *Journal of Computational Statistics & Data Analysis*, 100, 115-130. ScienceDirect.

Dowell, J., and Pinson, P. (2015). Very-Short-Term Probabilistic Wind Power Forecasts by Sparse Vector Autoregression. *IEEE Transactions on Smart Grid*. United Kingdom, 1–8.

Erdem, E., and Shi, J. (2011). ARMA based approaches for forecasting the tuple of wind speed and direction. *Journal of Applied Energy*, 88(4): 1405–1414. Elsevier.

EWEA. (2016). *Wind in power*. European Statistics. 1–14.

Fant, C., Schlosser, C. A., and Strzepek, K. (2016). The impact of climate change on wind and solar resources in southern Africa. *Journal of Applied Energy*, 161, 556–564. Elsevier.

Gil-Alana, L. A. (2003). An application of fractional integration to a long temperature series. *International Journal of Climatology*. 23(14): 1699–1710. John Wiley & Sons.

Giovanni Petris, Petrone, S., and Campagnioli, P. (2009). Dynamic Linear Models with R. *Applied Spatial Data Analysis with R*. London: Springer.

Gomes, P., and Castro, R. (2012). Wind Speed and Wind Power Forecasting using Statistical Models: AutoRegressive Moving Average (ARMA) and Artificial Neural Networks (ANN). *Journal of Sustainable Energy*, 36–45. Infonomics Society.

Gomez-Biscarri, J., and Hualde, J. (2015). A residual-based ADF test for stationary cointegration in i (2) settings. *Journal of Econometrics*, 184(2), 280–294. ScienceDirect.

Guo, H., Xu, M., and Hu, Q. (2011*).* Changes in near-surface wind speed in China: 1969-2005. *International Journal of Climatology*, 31,349–358. John Wiley & Sons.

Guo, Z., Dong, Y., Wang, J., and Lu, H. (2010). The forecasting procedure for long-term wind speed in the Zhangye area. *Journal of Mathematical Problems in Engineering*, (2010), 1-17. Hindawi Publishing Corporation.

Hamed, K. H. (2008). Trend detection in hydrologic data: The Mann-Kendall trend test under the scaling hypothesis. *Journal of Hydrology*, 349(3-4), 350–363. Elsevier.

Hamilton, J. (1994). *State-space models*. Unpublished note, University of Washington.

Hodge, B., Florita, A., Sharp, J., Margulis, M., Mcreavy, D., and Martin, L. (2015). *The Value of Improved Short- Term Wind Power Forecasting*. Technical Report. U.S.Department of Energy.

Hussain, M. a. (2012). Urban Wind Speed Analysis in Global Climate Change Perspective: Karachi as a Case Study. *International Journal of Geosciences*, 03(25), 1000–1009. Scientific Research An Academic Publisher.

Ibrahim, M. Z., Yong, K. H., Ismail, M., and Albani, A. (2014). Wind Speed Modeling for Malaysia. *International Journal of Renewable Enegry Research*, 4(4), 1-12. Thomson Reuters.

Jiang, P., Qin, S., Wu, J., and Sun, B. (2015). Time Series Analysis and Forecasting for Wind Speeds Using Support Vector Regression Coupled with Artificial Intelligent Algorithms. Journal of Mathematical Problems in Engineering, 2015, 1-14. Hindawi Publishing Corporation.

Jing, S., Xiuli, Q., and Song T, Z,. (2011). Short-term Wind Power Generation Forecasting: Direct versus Indirect ARIMA-Based Approaches. *International Journal of Green Energy*, 8 ,100-112. Taylor & Francis.

Kane, I. L., and Yusof, F. (2013). *Assessment of Risk of Rainfall Events with a Hybrid of ARFIMA-GARCH*. Modern Applied Scienc, *7*(12), 78–89. Canadian Center of Science and Education.

Kariniotakis, G. N., Marti, I., Nielsen, T. S., Giebel, G., Tambke, J., Waldl, I., and Frayssinet, P. (2002). Advanced Short-term Forecasting of Wind Generation. *IEEE Trans.. On Power Syste*m. Greece, 1-9.

Karmeshu, N. (2012). *Trend Detection in Annual Temperature & Precipitation using the Mann Kendall Test – A Case Study to Assess Climate Change on Select States in the Northeastern United States Trend Detection in Annual Temperature & Precipitation using the Mann*. Master Thesis, University of Pennsylvania.

Khairul, M., Mohd, B., Yusof, F., Daud, Z. M., Yusop, Z., and Afif, M. (2016). Optimal Design of Rain Gauge Network in Johor By Using Geostatistics and Particle Swarm Optimization. International Journal of Geomate, 11(25), 2422–2428. Hindawi Publishing Corporation.

Kim, J. Y., Kim, D. Y., and Oh, J. H. (2013). Projected changes in wind speed over the Republic of Korea under A1B climate change scenario. *International Journal of Climatology*, 1356, 1346–1356. John Wiley & Sons.

Kisi, O., and Ay, M. (2014). Comparison of Mann-Kendall and innovative trend method for water quality parameters of the Kizilirmak River, Turkey. *Journal of Hydrology*, 513, 362–375. ScienceDirect.

Koivisto, M., Seppänen, J., Mellin, I., Ekström, J., Millar, J., Mammarella, I., and Lehtonen, M. (2016). Electrical Power and Energy Systems Wind speed modeling using a vector autoregressive process with a time-dependent intercept term. *International Journal of Electrical Power and Energy Systems*, 77, 91–99. ScienceDirect.

Kwapien, J., Oswiecimka, P., and Drozdz, S. (2015). *Detrended fluctuation analysis made flexible to detect range of cross-correlated fluctuations*. 1–13. Cornell University Library.

Lillo, F. and Farmer, J.D. (2004). *The Long Memory of the Efficient Market . Studies in Nonlinear Dynamics & Econometrics (1-33).* United States :The Berkeley Electronics Press.

Liu, D., Wang, J., and Wang, H. (2015). Short-term wind speed forecasting based on spectral clustering and optimised echo state networks. *Journal of Renewable*

Energy, 78, 599–608. Elsevier.

Liu, F., Porco, T.C., Amza, A., Kadri, B., Nassirou, B., West, S.K., Bailey, R.L., Keenan, J.D. and Lietmann, T.M. (2015). Short-term Forecasting of the Prevalence of Clinical Trachoma: Utility of Including Delayed Recovery and Tests for Infection. *Journal of Parasites & Vectors*, 8(535), 1-8. Biomed Central.

Liu, H., Shi, J., and Erdem, E. (2010). Prediction of wind speed time series using modified Taylor Kriging method. *Journal of Energy*, 35(12), 4870–4879. Elsevier.

Liu, H., Tian, H. qi, and Li, Y. fei. (2012). Comparison of two new ARIMA-ANN and ARIMA-Kalman hybrid methods for wind speed prediction. *Journal of Applied Energy*, 9,415–424. ScienceDirect.

Loha, E., and Lindtjørn, B. (2010). Model variations in predicting incidence of Plasmodium falciparum malaria using 1998-2007 morbidity and meteorological data from south Ethiopia. *Malaria Journal*, 9(166), 1-12. National Center for Biotechnology Information U.S.

Ludlow, J., and Enders, W. (2000). Estimating non-linear ARMA models using Fourier coefficients. *International Journal of Forecasting*, 16(3), 333–347. ScienceDirect.

Makridakis, S. and Hibon, M. (1998). *ARMA models and The Box-Jenkins methodology*. Unpublished note, INSEAD, France.

Masseran, N. (2015a). Evaluating wind power density models and their statistical properties. *Journal of Energy*, 84, 533–541. Elsevier.

Masseran, N. (2015b). Markov Chain model for the stochastic behaviors of wind-direction data. *Journal of Energy Conversion and Management*, 92, 266–274. Elsevier.

Masseran, N. (2016). Modeling the fluctuations of wind speed data by considering their mean and volatility effects. *Journal of Renewable and Sustainable Energy Reviews*, 54, 777–784. Elsevier.

Masseran, N., and Razali, A. M. (2016). Modeling the wind direction behaviors during the monsoon seasons in Peninsular Malaysia. *Journal of Renewable and Sustainable Energy Reviews*, 56, 1419–1430. Elsevier.

Mcinnes, K. L., Erwin, T. A., & Bathols, J. M. (2011). Global Climate Model projected changes in 10 m wind speed and direction due to anthropogenic climate change. *Journal of Atmospheric Science*, 12, 325–333. John Wiley & Sons.

Ministry of Business, Innovation and Employment, New Zealands. (2016). *Short-term Employment Forecasts: 2016-2019*.

Moura, G. V., & Turatti, D. E. (2014). Efficient estimation of conditionally linear and Gaussian state space models. Journal of *Economics Letters*. 124(3): 494–499. SienceDirect.

Mudelsee, M. (2007). Long memory of rivers from spatial aggregation. *Journal of Water Resources Research*, 43(1), 1–6. John Wiley & Sons.

New Zealand Wind Energy Association (NZWEA). (2004). *Introduction to Wind Energy*.

Nurhayati, R. (2010). *Current Situtation of Wind Energy in Malaysia (2010)*. Unpublished note, Universiti Putra Malaysia.

Of, & Loa, M. (2010). *Detrending time series*, Unpublished note. University of Arizona.

Ooms, M. (2011). Statistical Software for State Space Methods. *Journal of Statistical Software*, 41(1), 1–18. The Foundation for Open Access Statistics.

Pandey, R.M. (2016). Commonly Used *t*-test in Medical Research. *Journal of the Practice of Cardiovascular Sciences*, 1(2), 185-188. Wolters Kluwer-MedKnow.

Pardo-Igúzquiza, E., and Dowd, P. A. (2013). Comparison of inference methods for estimating semivariogram model parameters and their uncertainty: The case of small data sets. *Journal of Computers & Geosciences*, 50, 154–164. Elsevier.

Pereira, E. B., Martins, F. R., Pes, M. P., da Cruz Segundo, E. I., and Lyra, A. D. a. (2013). The impacts of global climate changes on the wind power density in Brazil. *Journal of Renewable Energy*, 49, 107–110. Elsevier.

Petris, G., Petrone, S., and Campagnoli, P. (2009). *Dynamic linear models*. (1st ed.). U.S: Springer International Publishing.

Porshakov, A. and Deryugina, E. (2015). *Nowcasting and Short-term Forecasting of Russian GDP with a Dynamic Factor Model*. Unpublished note, Bank of Finland (BOFIT).

Prema, V., and Rao, K. U. (2015). Time series decomposition model for accurate wind speed forecast. *Journal of Renewables: Wind, Water, and Solar*, 2(18), 1-11. Springer.

Pryor, S. C., and Barthelmie, R. J. (2011). Assessing climate change impacts on the near-term stability of the wind energy resource over the United States. *Proceedings of the National Academy of Sciences of the United States of America*. India, 8167–8171.

Racine, R. (2011). *Estimating the Hurst Exponent*. Bachelor Thesis, Mosaic Group.

Rea, W., Reale, M., and Brown, J. (2011). *Long memory in temperature reconstructions*. Journal of Climatic Change, 107(3), 247–265. Springer.

Rosmann, T., Domínguez, E., and Chavarro, J. (2016*).* Comparing trends in hydrometeorological average and extreme data sets around the world at different time scales. *Journal of Hydrology: Regional Studies*, 5, 200–212. Elsevier.

Saini, N., and Mittal, A. K. (2014). Forecasting Volatility in Indian Stock Market using State Space Models. *Journal of Statistical and Econometris Methods*, 3(1), 115–136. Elsevier.

Sánchez, I. (2006). *Short-term prediction of wind energy production. International Journal of Forecasting*, 22(1), 43–56. Elsevier.

Sang, Y. F., Wang, Z., and Liu, C. (2014). Comparison of the MK test and EMD method for trend identification in hydrological time series. *Journal of Hydrology*, 510, 293–298. Elsevier.

Schmitt, P., Mandel, J., and Guedj,M. (2015). A Comparison of Six Methods for Missing Data Imputation. *Journal of Biometrics & Biostatistics*, 6(1),1-6. OMICS International.

Seo, Y., Kim, S., and Singh, V. P. (2015). Estimating Spatial Precipitation Using Regression Kriging and Artificial Neural Network Residual Kriging (RKNNRK) Hybrid Approach. *Journal of Water Resources Management*, 2189–2204. Elsevier.

Shadmani, M., Marofi, S., and Roknian, M. (2012). Trend Analysis in Reference Evapotranspiration Using Mann-Kendall and Spearman's Rho Tests in Arid Regions of Iran. *Journal Water Resources Management*, 26(1), 211–224. ScienceDirect.

Sharif, M., Archer, D. R., Fowler, H. J., and Forsythe, N. (2013). Trends in timing and magnitude of flow in the Upper Indus Basin. Journal of Hydrology and Earth System Sciences, 17(4), 1503–1516. ScienceDirect.

Shen, Z., and Ritter, M. (2015). *Forecasting volatility of wind power production*.

Discussion Paper 2015-2016. Berlin, 1-32.

Sirdaş, S. (2005). Daily wind speed harmonic analysis for Marmara region in Turkey. *Journal of Energy Conversion and Management*, 46(7-8), 1267–1277. ScienceDirect.

Siti, M. R., Norizah, M., and Syafrudin, M. (2011). The Evaluation of Wind Energy Potential in Peninsular Malaysia. *Journal of International Journal of Chemical & Environmental Engineering,* 2(4), 284-299. Academic Journal Database.

Soleimanzadeh, M., Wisniewski, R., and Brand, A. (2014). State-space representation of the wind flow model. *Journal of Wind Energy*, 17, 627–639. John Wiley & Sons.

Stringhi, A., and Figini, S. (2016). *How to improve accuracy for DFA technique*. Unpublished note, University of Pavia, Italy.

Svensson, M. (2015). *Networks Solar Power Forecasting Using Artificial Neural Networks*. Master Thesis , Royal Institute of Technology KTH.

Tatinati, S., and Veluvolu, K. C. (2013). A hybrid approach for short-term forecasting of wind speed. *The Scientific World Journal*, 2013, 1-8. Hindawi Publishing Corporation.

The European Wind Energy Association (EWEA). (2016). *Wind in Power*. European Statistics. February 2016.

Tian, Y. E., Xu, Y., Booij, M. J., Zhang, Q., and Lin, S. (2011). Trends in precipitation extremes and long-term memory of runoff records in Zhejiang , East China. Proceedings J-HO2. July 2011. Australia,227–232.

Tiriolo, L., Torcasio, R.C., Montesanti, S., Sempreviva, A.M., Calidonna, C.R., Transerici, C., and Federico, S. (2015). Forecasting wind power production from a wind farm. *Journal of Advances in Science & Research*, 12, 37–44. Copernicus Publications.

Troldborg, N., and Sørensen, J. (2014). A simple atmospheric boundary layer model applied to large eddy simulations of wind turbine wakes. *Journal of Wind Energy*, 657–669. John Wiley & Sons.

United States Energy Information Administration. (2016). *Short-term Energy Outlook (STEO)*. Independent Statistics & Analysis July 2016.

Valipour, M. (2012). Ability of Box-Jenkins Models to Estimate of Reference Potential Evapotranspiration (A Case Study: Mehrabad Synoptic Station, Tehran, Iran). *IOSR Journal of Agriculture and Veterinary Science*, 1(5), 01–11.

International Organzation of Scientific Research.

Van Ackere, S., Van Eetvelde, G., Schillebeeckx, D., Papa, E., Van Wyngene, K., and Vandevelde, L. (2015). Wind resource mapping using landscape roughness and spatial interpolation methods. *Journal of Energies*, 8(8), 8682–8703.MDPI.

Varotsos, C., Tzanis, C., Efstathiou, M., and Deligiorgi, D. (2015). Tempting long-memory in the historic surface ozone concentrations. *Journal of Atmospheric Pollution Research*, 6(6), 1055–1057. ScienceDirect.

Wan, N. W.B., Ahmad, M.F., Ibrahim, M.Z., Samo, K.B., and Muzhatik, M.Z. (2011). Wind energy potential at East Coast of Peninsular Malaysia. *International Journal of Applied Engineering Research*, 2(2), 360–366. Integrated Publishing Association.

Wang, J., Zhou, Q., Jiang, H., and Hou, R. (2015). Short-Term Wind Speed Forecasting Using Support Vector Regression Optimized by Cuckoo Optimization Algorithm. *Journal of Mathematical Problems in Engineering*, 1–13. Hindawi Publishing Corporation.

Wang, W., Gelder, P. H. A. J. M. Van, Vrijling, J. K., and Ma, J. (2005). *Nonlinear Processes in Geophysics Testing and modelling autoregressive conditional heteroskedasticity of streamflow processes*. Journal of Non-linear Processes in Geophysics, 12, 55–66. European Geoscience Union.

Wang, Y., Wang, J., and Wei, X. (2015). A hybrid wind speed forecasting model based on phase space reconstruction theory and Markov model : A case study of wind farms in northwest China. *Journal of Energy*, 91, 556–572. ScienceDirect.

Wu, T., and Li, Y. (2013). Spatial interpolation of temperature in the United States using residual kriging. *Journal of Applied Geography*, 44, 112–120. Elsevier.

Ye, W. (2013). *Spatial Variation and Interpolation of Wind Speed Statistics and Its Implication in Design Wind Load*. PhD Thesis, The University of Western Ontario.

Yuan, N., Fu, Z., & Liu, S. (2014). Extracting climate memory using Fractional Integrated Statistical Model : A new perspective on Climate Prediction. *Scientific Reports*, 1–10. McMillan Publishers Limited.

Yusof, F., & Kane, I. L. (2013). Volatility modeling of rainfall time series. *Journal of Theoretical and Applied Climatology*, 113(2), 247–258. Springer.

Yusof, F., Kane, I. L., & Yusop, Z. (2013). Structural break or long memory: An empirical survey on daily rainfall data sets across Malaysia. *Journal of*

Hydrology and Earth System Sciences, 17(4), 1311–1318. Copernicus Publications.

Zhang,W., Su, Z., Zhang, H., Zho, Y.,Zhao, Z. (2014). Hybrid Wind Speed Forecasting Model Study Based on SSA and Intelligent Optimized Algorithm. *Journal of Abstract and Applied Analysis*, 2014, 1-14. Hindawi Publishing Corporation.

Zhu, Q., & Lin, H. S. (2010). Comparing ordinary kriging and regression kriging for soil properties in contrasting landscapes. *Journal of Pedosphere*, 20(5), 594–606. Elsevier.

Zhu, X., Kenneth P.B., & Genton, M.G. (2014). Incorporating Geostrophic Wind Information for Improved Space-time Short-Term Wind Speed Forecasting. *The Annals of Applied Statistics,* 3, 1782-1799. Institute of Mathematical Statistics.