WIND SPEED FORECASTING IN THE PRESENCE OF SOME METEOROLOGICAL VARIABLES USING VECTOR AUTOREGRESSIVE AND MULTINOMIAL LOGISTIC REGRESSION MODEL

By

Gabriel Onuche ODEKINA

(P15SCMT8057)

DEPARTMENT OF STATISTICS,
AHMADU BELLO UNIVERSITY,
ZARIA, NIGERIA

DECEMBER, 2017

WIND SPEED FORECASTINGUSING IN THE PRESENCE OF SOME METEOROLOGICAL VARIABLES USING VECTOR AUTOREGRESSIVE AND MULTINOMIAL LOGISTIC REGRESSION MODEL

By

Gabriel Onuche ODEKINA

B.Sc. Statistics (University of Ilorin, 2014)
P15SCMT8057

A DISSERTATION SUBMITTED TO THE SCHOOL OF POSTGRADUATE STUDIES, AHMADU BELLO UNIVERSITY, ZARIA IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE AWARD OF MASTER DEGREE IN STATISTICS

DEPARTMENT OF STATISTICS,
FACULTY OF PHYSICAL SCIENCE,
AHMADU BELLO UNIVERSITY,
ZARIA, NIGERIA

DECEMBER, 2017

Declaration

I declare that the work in this Dissertation entitled "Wind Speed Forecasting Using Vector Autoregressive and Multinomial Logistic Regression Model" has been carried out by me in the Department of Statistics. The information derived from the literature has been duly acknowledged in the text and a list of references provided. No part of this dissertation was previously presented for another degree or diploma at this or any other Institution.

Odekina Onuche Gabriel _____ _____

(Name of Student) Signature Date

Certification

This dissertation entitled "Wind Speed Forecasting in the presence of some meteorological variables Using Vector Autoregressive and Multinomial Logistic Regression Model" by Gabriel Onuche ODEKINA meets the regulations governing the award of the degree of Masters of Science in Statistics of the Ahmadu Bello University, Zaria and is approved for its contribution to knowledge and literary presentation.

Dr. H. G. Dikko Chairman, Supervisory Committee	_____ Signature	_____ Date
Dr.T. Dahiru Member, Supervisory Committee	_____ Signature	_____ Date
Dr. H.G. Dikko Head of Department	_____ Signature	_____ Date
_____ External Examiner	_____ Signature	_____ Date
Prof. S. Z. Abubakar Dean, School of Postgraduate Studies	_____ Signature	_____ Date

Dedication

This research work is dedicated to my parents, Mr. Alilu Ibrahim Odekina and Mrs Veronica Odekina.

Acknowledgements

My profound gratitude goes to God Almighty for making this research work a success, and for turning every milestone met during the course of this work into a huge success.

I also want to thank my supervisors; Dr. H. G. Dikko, Dr. T. Dahiru and other lecturers of the department of Statistics, Ahmadu Bello University, Zaria for their intellectual guidance towards the success of this research work.

I want to thank my parents: Mr Alilu Ibrahim Odekina and Mrs Veronica Odekina, my brothers, Moses Odekina, David Odekina, and sister, Esther Odekina for the support rendered to me, both financially and morally. I also thank them for the love and words of encouragement showered on me.

My gratitude is not complete without mentioning the support rendered to me by my friend-Mr Mathew Aina, and other friends who have in one way or the other contributed to the success of this research work. I say a big thank you to everyone and God bless.

Abstract

Wind speed forecasting is pertinent in order to have a power system that is reliable and secured. The Vector Autoregressive model and the Multinomial logistic regression model were applied on a 30 year daily data obtained from the department of Soil science, Ahmadu Bello University, Zaria. The Vector Autoregressive model shows that non of the meteorological variables namely; rainfall, air temperature, earth temperature, sunshine, humidity, pressure and evaporation significantly affect wind speed. It was also found that each of the meteorological variables have varying effects on Wind speed over a future time horizon as depicted by the impulse response function. The Wind direction was also examined using the Multinomial logistic regression where it was found that there is no Statistical significant difference between Wind speed and the Northern and Western direction. To this end, this study posits that wind speed is not significantly being influenced by the meteorological variables and the northern and Western directions do not exert any effect on Wind speed.

TABLE OF CONTENTS

Title page……..……..ii
Declaration iii
Certification..... iv
Dedication v
Acknowledgements vi
Abstract vii
TABLE OF CONTENTS viii
CHAPTER ONE 1
INTRODUCTION 1
1.1 Background to the Study 1
1.2 Statement of the Problem 3
1.3 Aim and Objectives of the Study 4
1.4 Significance of the Study 4
1.5 Scope of the Study 4
1.6 Definition of Terms 5
CHAPTER TWO 7
LITERATURE REVIEW 7
2.1 Introduction 7
2.2 Empirical Review 7
CHAPTER THREE 12
METHODOLOGY 12
3.1 Introduction 12
3.3 Vector Autoregressive (VAR) Model 13
3.4 Information Criteria 13
3.5 The Impulse Response Function (IRF) 14
3.6 Lagrange Multiplier Statistic 15
3.7 Residual Autocorrelation 15
3.8 Normality Test 16
3.9 Multinomial Logistic Regression 16

CHAPTER FOUR 20
RESULTS AND DISCUSSION 20
4.1 Introduction 20
4.2 Time Series Plot 20
4.3 Descriptive Statistics 22
4.4 Test for Stationarity 23
4.5 Lagrange Multiplier Test for Conditional Heteroscedasticity (ARCH) and Autocorrelation 25
4.6 Fitting of the Vector Autoregressive Model 26
4.7 The Impulse Response Function 29
4.8 Wind speed forecast 31
4.9 Multinomial Logistic Regression Model 32
CHAPTER FIVE 34
SUMMARY, CONCLUSIONS AND RECOMMENDATIONS 34
5.1 Summary 34
5.2 Conclusion 34
5.3 Recommendations 35
5.4 Contribution to Knowledge 35
REFERENCE 37
APPENDIX 41

CHAPTER ONE

INTRODUCTION

1.1 Background to the Study

With the advent of science and technology, the demand for electrical energy becomes inevitable. Nigeria is a country endowed with abundant energy resources like coal, solar, water (Dam), wind and so on which can be used as a form of electricity generation. Despite the abundance of these energy resources, there is inconsistent supply of electricity, which may be due to underutilization of the potentials.

Wind energy is the fastest growing renewable source of energy. With respect to this, the need for wind speed modelling and forecasting becomes paramount. Energy generation by wind is of great advantage because wind turbines do not produce any form of pollution when sited strategically. Moreover, it blends with the natural landscape. The utilization of wind energy will ensure the growth of socio-economic development and improvement in the quality of life of the citizens. The demand for more sustainable energy sources are on the increase in order to address the growing needs of humans. It is also in line with taking care of the environment and the minimal use of natural resources, which means there is an urgent need for developing renewable energy.

Wind energy is now becoming the current trend in renewable energy as it addresses rising energy demand while being nature friendly at the same time. In the long run, electricity generated from the wind turbines cost less than the conventional power plants since it does not consume fossil fuel. Researches on the potentials of wind energy in some major cities in

Nigeria show high wind speed in Lagos, Maiduguri, Enugu, Jos, Kano, Funtua and Sokoto (Idris *et al.*, 2012).

The gathering of wind data is important for the wind farm beginning from its feasibility to its actual operation. Prior to the construction of a wind farm, at least one year of meteorological study is necessary and a detailed verification of the specific on-site wind conditions are necessary. Meteorological values, more specifically wind speed and wind direction are necessary for the calculation of the wind farm's yearly electrical generation profile. The harnessing of Kinetic energy through the wind has been used for centuries, be it in form of powering sail boats, wind mills, or furnaces (Aliyu and Mohammed 2014). But, it was not until 1979 that the modern wind power industry began in earnest with the production of wind turbines. The use of wind energy as a form of renewable energy gained momentum in the 80s and 90s and there are now thousands of wind turbines operating all over the world (Minh *et al.*, 2011). The modern and most commonly used wind turbine has a horizontal axis with two or more aerodynamic blades mounted on the shaft. These blades can travel at over several times the wind speed, generating electricity which is captured by a medium voltage power collection system and fed through the power transmission network (Garba and Al-Amin, 2014).

Wind energy is unarguably the most economic renewable energy apart from hydropower, Its usage, versatility and ability to use it as a decentralized energy form make its applications possible in rural areas where it is technically and economically feasible in the country. Winds are caused by the uneven heating of the atmosphere by the sun, irregularities of the earth's surface, and rotation of the earth. Wind flow patterns are modified by the earth's terrain, water bodies and vegetative cover (Reddy *et al.*, 2015). The major challenge of using Wind

as a source of energy is that winds are intermitted and it is not available always when electricity is needed.

Wind speed forecasting is essential for a secured and reliable power system for a particular site. This research forecasts wind speed in Samaru, Zaria using the Vector Auto Regressive (VAR) model. The explanatory variables used are rainfall, relative humidity, air temperature, earth temperature, sunshine and pressure. In addition, the wind directions were also studied in relation to wind speed using the Multinomial Logistic Regression model.

1.2 Statement of the Problem

The growing demands of energy supply by mankind has triggered the potentials of using wind energy due to its constant availability and nature-friendly environment. The importance of wind energy cannot be over emphasized. Wind energy can be seen as the energy of the future for Nigerians if properly harnessed and utilized. For example, the 10MW Katsina wind farm project which is owned by the Federal Ministry of Power is a pioneer project in Nigeria aiming to generate 10MW of power via wind turbine with the Federal Government's desire to boost electricity generation and have constant power supply. Wind speed modelling and forecasting are necessary for a reliable and secured power system. And it is therefore of paramount importance to know the impact of meteorological variables on wind speed. Cherie *et al.,* (2014) used the Vector Autoregressive Model in modelling wind speed in the presence of some meteorological variables which are humidity, temperature and pressure. However, we applied the Vector Autoregressive Model and Multinomial Logistic Regression in studying wind speed and its direction respectively with some selected meteorological variables which are rainfall, air temperature, earth temperature, humidity, evaporation, sunshine and pressure.

1.3 Aim and Objectives of the Study

The aim of this study is to forecast wind speed using the Vector Autoregressive and the Multinomial Logistic Regression Model. This aim will be achieved through the following objectives;

(i) determination of the co-integrating variables and order of co-integrating VAR model.

(ii) fitting of the Cointegrated Vector Autoregressive Model

(iii) model adequacy check and forecasting.

(iv) fitting of the Multinomial Logistic Regression Model

(v) to make recommendation or possible implication of the result

1.4 Significance of the Study

This study will serve as a contribution to the existing literatures on modelling of wind speed in the scientific world. In addition, the result of this study will be of paramount importance to the Kaduna State Government, the Federal Government and also to the Nigerian Meteorological Agency as it creates awareness on the importance of wind speed, can provide valuable information on the expected daily and seasonal load for a project. The model which will be adopted can be used as a tool by the government in making forecasts on wind speed in order to facilitate proactive planning.

1.5 Scope of the Study

The data used for this study is a 30 year daily data from 1984 to 2015 obtained from the Department of Soil Science, Ahmadu Bello University, Zaria. The data consist of the daily amount of rainfall, relative humidity, pressure, sunshine, air temperature and earth temperature in Zaria, Kaduna state.

1.6 Definition of Terms

i. **Wind:** This refers to the movement of air.

ii. **Air temperature:** This is the degree of hotness or coldness of the air.

iii. **Earth temperature:** This is the degree of hotness or coldness of the earth's surface.

iv. **Humidity:** This is the presence of water vapor in the atmosphere.

v. **Pressure:** This refers to the force exerted onto a surface by the weight of the air.

vi. **Evaporation:** This is the process in which a substance in its liquid state changes into gaseous state.

vii. **Rainfall:** This refers to the amount of precipitation falling in a given area at a particular period of time.

viii. **Sunshine:** This refers to the intensity of the light and heat coming from the sun.

ix. **Time series:** This is a collection of data on a particular phenomenon or variable over a period of time. The time period can be hourly, daily, weekly or yearly.

x. **Stationary series:** These are series with constant mean and variance. These series have a mean that is defined in which it can fluctuate around with a constant variance.

xi. **Nonstationary series:** These are series with no constant mean and variance over a period of time. These series have different mean at different time period and with variance increasing with respect to the sample size.

xii. **Cointegration:** Two or more time series variables are said to be cointegrated if the linear combination of the series results into a stationary series. Cointegration is therefore a linear combination of nonstationary time series variables whose linear combination gives a stationary series.

xiii. **Autocorrelation:** This is the correlation of a variable with itself over a successive time interval.

xiv. **Heteroscedasticity**: This is a situation in which the variance of the errors is not constant.

xv. **Cointegration rank:** This is referred to as the number of independent linearly cointegrating vectors (Johansen, 2000)

CHAPTER TWO

LITERATURE REVIEW

2.1 Introduction

Researches have been carried out in the past on wind speed modelling. It is therefore important to review some of these studies. This chapter presents works related to this study. The empirical review brings some of the scientific researches and their findings on wind speed modelling

2.2 Empirical Review

Wind speed forecasting using spatio-temporal indicators by Ohashi and Torgo (2012) described a new methodology for short-term wind speed prediction with extreme relevance for electricity markets and Wind power generation. Van *et al.*, (2005) also worked on wind speed modelling where a comparison of the Weibull model with measured wind speed distributions for stochastic wind generation was studied. Stochastic trends have been proposed as one technique to generate scenarios of future climate change. Temperature and rainfall are among the main indicators in climate study. Babazadeh and Shamsnia (2014) carried out a simulation and modelling of monthly precipitation and the mean of monthly temperature using stochastic methods based on Autoregressive Integrated Moving Average (ARIMA) model.

Fadare (2008) modelled wind speed variation at Ibadan using the Weibul distribution function where it was found that with a yearly wind speed of 2.748m/s, Ibadan is in a region with low wind speed. Odo *et al.,* (2013) also applied the Weibul distribution in modeling wind data for Enugu, Onitsha and Owerri whose result shows that the model can be used in predicting wind energy with a precision that is acceptable. Evaluation of wind power density in Gombe, Yola

and Maiduguri was carried out by Dikko and Yahaya (2012) where they compared the Weibul Model and Rayleigh models and discovered that the Weibul model is the best model that describes the best wind speed data at 10m heights

Given more than one locally stationary time series, Cardinali and Nason (2010) described a method to discover time-varying linear combinations of the locally stationary series that are stationary. Systems for which this can occur are called stationary and the associated time-varying linear combinations are called co-stationary vectors. Ampaw *et al.,* (2013) developed a time series model for rainfall pattern and its prediction for the new Juaben Municipality of the Eastern region of Ghana. Rainfall monthly data on New Juaben Municipality was used by applying the Box-Jenkins approach to develop a SARIMA (0, 0, 1), $(1,1,1)_{12}$model to forecast the rainfall statistics of the New Juaben Municipality for a year.

Ayanlowo (2014) studied climatic variables and its effect on climatic conditions using square regression, two stage least square, co-integration, granger causality test and Vector Autoregressive (VAR) model. The temperature, humidity and radiation were stationary at 5% level, the Vector Autoregressive models revealed a dynamic relationship and the granger causality test indicated both unidirectional and bidirectional causal relationship between the climatic variables at 5% level of significance.

Brown *et al.,* (1984) applied time series models to simulate and forecast wind speed and wind power. In their work, a general approach for wind speed modelling was studied because wind energy is a function of wind speed. Values of wind energy were estimated by applying appropriate transformation to values of wind speed. The wind speed modelling approach they used takes into account some basic features of wind speed data, including autocorrelation and

non-Gaussian distribution. Ibrahim *et al.,* (2014) applied the spatial wind speed modeling which described a spatial wind distribution. Spatial Wind data were interpolated by kriging. The spatial wind modelling extrapolated the spatial wind data to its altitude from the elevation by power law approach and it was found that in Southern parts, wind speed was higher compared to other parts throughout the year.

In the forecasting of wind speed using a univariate ARIMA model and a multivariate neural network by Cadenas *et al.,* (2016), a two step ahead wind speed forecasting models were compared. The impact of the various meteorological variables (barometric pressure, air temperature, Solar radiation, wind direction and humidity) were compared on the performance of a multivariate model of wind speed forecast with respect to high performing univariate linear model. Sloughter *et al.,* (2010) introduced the Bayesian model averaging for probabilistic wind speed forecasting using ensembles and Bayesian model averaging where the weather forecast paradigm is deterministic, based on models. Their method yields a calibrated and sharp probabilistic forecast.

Extreme wind speed modeling by Tilman (2006) shows that a common problem with wind data is the problem of masking, where only the largest wind speed across all directions is recorded while speeds in other directions are not recorded. In order to overcome this challenge, they suggested a model that takes into account the masking problem, where a simulation was carried out to analyse the behavior of this model under different conditions. Oluseyi *et al.,* (2013) studied wind profile characteristics and turbine performance analysis using 21 years monthly wind speed data. They subjected the data to different statistical test

and compared the two-parameter Weibull probability density function where they found that the average monthly wind power ranges between 6.6 to 9.5 m/s.

Minano *et al.,* (2013) proposed the continuous wind speed models based on stochastic differential equations. The models were constructed using parameters obtained from the statistical properties of real-world measurement data. The developed models were able to provide wind speed trajectories with similar marginal distribution and autocorrelation function to those observed in wind speed data. Gnana (2011) studied the computing models for wind speed prediction in renewable energy systems in which different techniques for wind speed forecast were described and the model based on neural network produced the wind forecast with high accuracy. The short-term forecasting of wind speed and direction exploiting data non-stationarity by Alice *et al.,* (2015) gives the exploration of how the precision of short-term prediction of wind speed and direction can be enhanced by considering the daily variation of wind. The wind speed and direction were modelled as the magnitude and phase of a complex- valued time series. The prediction was performed using the spatio-temporal correlation between measurements at different geographical locations.

Wind speed estimation involving the incorporation of seasonal data using Markov chain models by Selin *et al.,* (2013) presented a method for accurately modelling and predicting wind speed in cases where incomplete data sets are available. They applied seasonal simulation for the synthetic generation of wind speed data which was achieved through the Markov Chain Monte Carlo technique. Petroula *et al.,* (2008) applied the Kalman filtering in improving wind speed forecasts for wind power prediction purposes. Two atmospheric models were employed with different options of horizontal resolution to provide wind speed

forecast and the application of the Kalman filter led to the elimination of possible systematic errors .

Frandsen *et al.*, (2006) studied the analytical modelling of wind speed in large offshore farms. They proposed a model for wind speed deficit in wind farm which is analytical and encompasses both small wind farms and wind farms extending over large areas. The model handles a regular array geometry with straight rows of wind turbines and equidistant spacing between each rows.

Cherie *et al.*, (2014) investigated wind speed with some selected weather variables using the Vector Autoregressive (VAR) model where they found that wind speed, temperature and humidity are important determinants of the wind speed of turbines. This work is the motivation behind our work as we include more variables namely sunshine, evaporation, air temperature, earth temperature and wind direction using the Vector Autoregressive model and Multinomial Logistic Regression.

CHAPTER THREE

METHODOLOGY

3.1 Introduction

This chapter presents the methods and materials used in this research work. The Vector Autoregressive model was used on the time series data and the Augmented Dickey Fuller (ADF) test was used in determining the stationarity of the variables. In addition, the variables were tested for heteroscedasticity and autocorrelation using the Lagrange multiplier and Box-Pierce statistic respectively. In addition, the impulse response function was used in determining the possible shocks expected on wind speed in the presence of other meteorological variables and finally, the multinomial logistic regression model was applied on the wind direction considering the categorical nature of wind direction.

3.2 Testing for Stationarity

The Dickey Fuller (DF) test is used for testing if the given series is stationary or not. The approach involves testing the null hypothesis that a series is not stationary against the alternative of stationarity. The Dickey Fuller test which is a regression model that nests a mean, a lagged term and a deterministic trend term in one model is given below;

$$y_t = \alpha + y_{t-1} + \beta t + e_t \tag{3.1}$$

The Augmented Dickey Fuller test involves adding an unknown number of lagged first differences of the dependent variable to capture autocorrelated omitted variables that would otherwise, by default, enter the error term ε_t. Its regression equation is given below;

$$y_t = \alpha_0 + \rho_1 y_{t-1} + \sum_{j=2}^{\rho-1} \beta_j \nabla y_{t-j} + e_t \tag{3.2}$$

3.3 Vector Autoregressive (VAR) Model

A multivariate time series y_t is a VAR process of order 2, or VAR (2) if it follows the model,

$$y_t = \Phi_1 y_{t-1} + \Phi_2 y_{t-2} + a_t \tag{3.3}$$

where $a_t \sim N(0, \sum)$, Φ_1 and Φ_2 are k×k matrices of unknown coefficients, and k is the total number of variables under consideration.

3.4 Information Criteria

There are standard information criteria for selecting a vector autoregressive model of order (p). These information criteria are Akaike Information Criterion (AIC), Bayesian Information Crietrion (BIC) and the Hannan Quinn Information Criterion (HQ). Any one of the information criteria mentioned can be used in for lag order selection, but it is advisable to go with one. According to Kuha (2004), relying on only one information criterion for model selection would produce more conclusive result than using multiple ones, as each criterion is designed to identify the candidate estimated to be the best. For this study, we used the Akaike Information Criterion.

The available information criteria for selecting a vector autoregressive model of order (p) are given below.

$$AIC = \ln\left(\left|\hat{\Omega}_i\right|\right) + \frac{2k^2 i}{T} \tag{3.4}$$

$$BIC = \ln\left(\left|\hat{\Omega}_i\right|\right) + \frac{k_i^2 \ln(T)}{T} \tag{3.5}$$

$$HQ = \ln\left(\left|\hat{\Omega}_i\right|\right) + \frac{2k_i^2 \ln[\ln(T)]}{T} \tag{3.6}$$

Where T is the sample size, k is the number of parameters,

$$\hat{\Omega} = \frac{1}{T}\sum_{t=i+1}^{T} \hat{a_t^i}(\hat{a_t^i})' \tag{3.7}$$

a_t stands as the error term.

3.5 The Impulse Response Function (IRF)

This technique gives the simulation of how a variable responds to a sudden change by structural effect from other variables over a future time horizon. These structural effects can be referred to as shocks or innovations in a variable (Harvey, 1994). This innovation is commonly referred to as impulse.

Let y_t be a k dimensional vector series generated by

$$y_t = F_1 y_{t-1} + F_2 y_{t-2} + ... + u_t,$$
$$u_t = error$$

$$= \phi(B)u_t = \sum_{i=0}^{\infty} \phi_i u_{t-i}$$
$$\phi_i = F_1\phi_{i-1} + F_2\phi_{i-2} + ...F_p\phi_{i-p} \tag{3.8}$$

where $\mathrm{E}(u_t u_t') = \sum$ According to Cholesky decomposition of $\sum$, let P be a lower triangular matrix such that

$\sum = PP'$. Then,

$$y_t = \sum_{i=0}^{\infty} (F_i P)\left(P^{-1}u_{t-1}\right)$$

The impulse response function is given by

$$\Psi^{0}_{j(n)} = \phi_n P e_j , \qquad \text{n=0, 1, 2…}$$

Where e_j is an m× 1 selection vector with unity as its jth element and 0 else where.

3.6 Lagrange Multiplier Statistic

The Autoregressive Conditional Heteroscedastic (ARCH) Lagrange Multiplier (LM) is used to test the hypotheses of homoscedasticity. This involves regressing the squared residuals on the conditional mean equation which may be an autoregressive, or moving average model. For example, considering an ARMA (1,1) process,

$$r_t = \phi r_{t-1} + \phi e_{t-1} + e_t ,$$

$$e_t^2 = \alpha_0 + \alpha_1 e_{t-1}^2 + \alpha_2 e_{t-2}^2 + ... + \alpha_q e_{t-q}^2 \quad (3.9)$$

$$H_0 : \alpha_1 = \alpha_2 = \alpha_3 = ...\alpha_q = 0 \text{ vs } H_1 : \alpha_i > 0$$

$$LM_E = TR^2 \quad (3.10)$$

where T= Sample size, R^2 = R squared

3.7 Residual Autocorrelation

The Box-Pierce statistic which was proposed by Box and pierce (1970) will be used to test the autocorrelation in the residuals.

H_0: There is no autocorrelation up to order k Vs H_1: There is autocorrelation up to order k. The test statistic is given as

$$Q = n\sum_{j=1}^{k} r_j^2 \tag{3.11}$$

where n= Sample size, r= Autocorrelation at lag j

3.8 Normality Test

The Jarque-Berra test will be used to determine whether the error correction model is normally distributed. This test measures the difference in Kurtosis and Skewness of a variable compared to those of the normal distributions.

H_0: The variable is normally distributed Vs H_1: The variable is not normally distributed.

$$JB = \frac{N-p}{6}\left[S^2 + \frac{(k-3)^2}{4}\right] \tag{3.12}$$

N= number of observations

p=Number of estimated parameter

S= Skewness

K=Kurtosis of the variable.

We reject the null hypotheses if the p-value $\leq$ level of significance.

3.9 Multinomial Logistic Regression

Czepiel (2002) presented the maximum likelihood estimation of the multinomial logistic Regression as follows.

The Logit model having a baseline category with a predictor x is given by

$$\log\left(\frac{\pi_{ij}}{\pi_{iJ}}\right)=\log\left(\frac{\pi_{ij}}{1-\sum_{j=1}^{J-1}\pi_{ij}}\right)=\sum_{k}^{K}x_{ik}\beta_{kj} \tag{3.13}$$

i =1, 2,…N

j=1,2,…,J−1

$$\text{where } \pi_{ij}=\frac{e^{\sum_{k}^{K}x_{ik}\beta_{kj}}}{1+\sum_{j=1}^{J-1}e^{\sum_{k}^{K}x_{ik}\beta_{kj}}}, \qquad j< J \tag{3.14}$$

and

$$\pi_{iJ}=\frac{1}{1+\sum_{j=1}^{J-1}e^{\sum_{k=0}^{K}x_{ik}\beta_{kj}}} \tag{3.15}$$

The joint probability density function is

$$f\left(y/\beta\right)=\prod_{i=1}^{N}\left[\frac{n_i!}{\prod_{j=1}^{J}y_{ij}}\times\prod_{j=1}^{J}\pi_{ij}^{y_{ij}}\right] \tag{3.16}$$

The log likelihood function for multinomial logistic regression model is

$$L\left(\beta/y\right)\cong\prod_{i=1}^{N}\prod_{j=1}^{J}\pi_{ij}^{y_{ij}} \tag{3.17}$$

Replacing the Jth terms, equation (3.17) becomes

$$L\left(\beta/y\right)=\prod_{i=1}^{N}\prod_{j=1}^{J-1}\pi_{ij}^{y_{ij}}\times\pi_{ij}^{n_i-\sum_{j=1}^{J-1}y_{ij}}$$

$$L\left(\beta/y\right)=\prod_{i=1}^{N}\prod_{j=1}^{J-1}\pi_{ij}^{y_{ij}}\times\frac{\pi_{ij}^{n_i}}{\pi_{ij}^{\sum_{j=1}^{J-1}y_{ij}}}$$

$$L(\beta / y) = \prod_{i=1}^{N} \prod_{j=1}^{J-1} \pi_{ij}^{y_{ij}} \times \frac{\pi_{ij}^{n_i}}{\prod_{j=1}^{J-1} \pi_{ij}^{y_{ij}}} \tag{3.18}$$

$$L(\beta / y) = \prod_{i=1}^{N} \prod_{j=1}^{J-1} \left(\frac{\pi_{ij}}{\pi_{iJ}} \right)^{y_{ij}} \times \pi_{iJ}^{n_i} \tag{3.19}$$

Substituting for π_{ij} and π_{iJ} using equation (3.14) and (3.15), we have

$$L(\beta / y) \prod_{i=1}^{N} \prod_{j=1}^{J-1} \left(e^{\sum_{k=0}^{K} x_{ik} \beta_{kj}} \right)^{y_{ij}} \times \left(\frac{1}{1 + \sum_{j=1}^{J-1} e^{\sum_{k=0}^{K} x_{ik} \beta_{kj}}} \right)^{n_i}$$

$$L(\beta / y) = \prod_{i=1}^{N} \prod_{j=1}^{J-1} e^{y_{ij} \sum_{k=0}^{K} x_{ik} \beta_{kj}} \times \left(1 + \sum_{j=1}^{J-1} e^{\sum_{k=0}^{K} x_{ik} \beta_{kj}} \right)^{-n_i} \tag{3.20}$$

By taking the log of equation (3.20), we have

$$l(\beta) = \sum_{i=1}^{N} \sum_{j=1}^{J-1} \left(y_{ij} \sum_{k=0}^{K} x_{ik} \beta_{kj} \right) - n_i \log \left(1 + \sum_{j=1}^{J-1} e^{\sum_{k=0}^{K} x_{ik} \beta_{kj}} \right) \tag{3.21}$$

which is the log likelihood function for the multinomial logistic regression. The values for β which maximize equation (3.21) can be found by differentiating equation (3.21) with respect to β.

$$\frac{d(\beta)}{d(\beta_{kj})} = \sum_{i=1}^{N} y_{ij} x_{ik} - n_i \times \frac{1}{1 + \sum_{j=1}^{J-1} e^{\sum_{k=0}^{K} x_{ik} \beta_{kj}}} \times \frac{d}{d\left(\beta_{kj}\right)} \left(1 + \sum_{j=1}^{J-1} e^{\sum_{k=0}^{K} x_{ik} \beta_{kj}} \right)$$

$$\frac{d(\beta)}{d(\beta_{kj})} = \sum_{i=1}^{N} y_{ij} x_{ik} - n_i \times \frac{1}{1 + \sum_{j=1}^{J-1} e^{\sum_{k=0}^{K} x_{ik} \beta_{kj}}} \times e^{\sum_{k=0}^{K} x_{ik} \beta_{kj}} \times \frac{d}{d\beta_{kj}} \sum_{k=0}^{K} x_{ik} \beta_{kj}$$

$$\frac{d(\beta)}{d(\beta_{kj})} = \sum_{i=1}^{N} y_{ij} x_{ik} - n_i \times \frac{1}{1+\sum_{j=1}^{J-1} e^{\sum_{k=0}^{K} x_{ik}\beta_{kj}}} \times e^{\sum_{k=0}^{K} x_{ik}\beta_{kj}} \times x_{ik}$$

$$\frac{d(\beta)}{d(\beta_{kj})} = \sum_{i=1}^{N} y_{ij} x_{ik} - n_i \pi_{ij} x_{ik} \tag{3.22}$$

Equation (3.22) consist of (J−1) (K+1) equations to be set to zero and solve for each β_{kj}.

$$\frac{d^2(\beta)}{d\beta_{kj} d\beta_{k^1 j^1}} = \frac{d(\beta)}{d\beta_{k^1 j^1}} \sum_{i=1}^{N} y_{ij} x_{ik} - n_i \pi_{ij} x_{ik}$$

$$\frac{d^2(\beta)}{d\beta_{kj} d\beta_{k' j'}} = \frac{d(\beta)}{d\beta_{k' j'}} \sum_{i=1}^{N} -n_i x_{ik} \pi_{ij}$$

$$\frac{d^2(\beta)}{d\beta_{kj} d\beta_{k' j'}} = -\sum_{i=1}^{N} n_i x_{ik} \frac{d}{d\beta k' j'} \left(\frac{e^{\sum_{k=0}^{K} x_{ik}\beta_{kj}}}{1+\sum_{j=1}^{J-1} e^{\sum_{k=0}^{K} x_{ik}\beta_{kj}}} \right) \tag{3.23}$$

When $j' = j$, equation (3.23) becomes

$$\frac{d^2(\beta)}{d\beta_{kj} d\beta_{k' j'}} = \pi_{ij} x_{ik^1} \left(1 - \pi_{ij}\right) \tag{3.24}$$

When $j' \neq j$, equation (3.23) becomes

$$\frac{d^2(\beta)}{d\beta_{kj} d\beta_{k' j'}} = -\pi_{ij} x_{ik'} \pi_{ij'} \tag{3.25}$$

CHAPTER FOUR

RESULTS AND DISCUSSION

4.1 Introduction

This chapter presents an overview of the analysis of the study and its findings. The meteorological variables (wind speed, rainfall, humidity, sunshine, air temperature, earth temperature and pressure) will be tested for stationarity using the Augmented Dickey Fuller test. The order of the model will also be determined using the Akaike Information Criterion (AIC). The residuals of the variables will be assessed for heteroscedasticity and the variables for autocorrelation. Moreover, the model will be fitted and its normality will be tested using the Jarque-Berra test. In addition, a 30 day forecast will be made using the fitted model. Finally, wind direction and wind speed will be modelled using the multinomial logistic regression.

4.2 Time series plot

The first stage in the analysis of any time series data is to plot the data. This gives the time series plot. This plot enables us to visually see the nature of the data and its behavior over time.

From the plots given below, consisting of sunshine, wind speed, evaporation, rainfall, humidity, pressure, air temperature, earth temperature, it can be seen that each of these meteorological variables appear to be stationary over time. A stationarity test was further conducted to certify this claim.

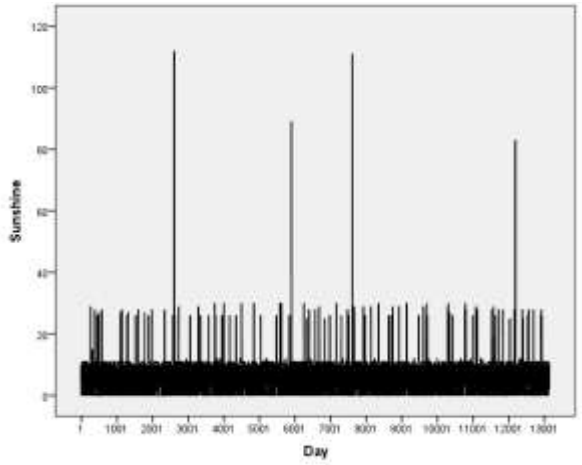


Figure 1: Sunshine

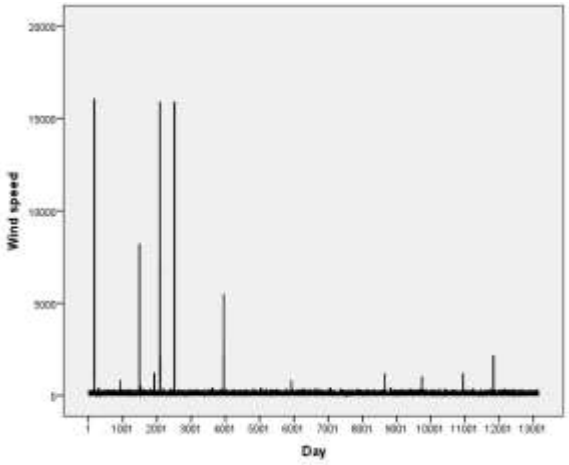


Figure 2: Wind speed

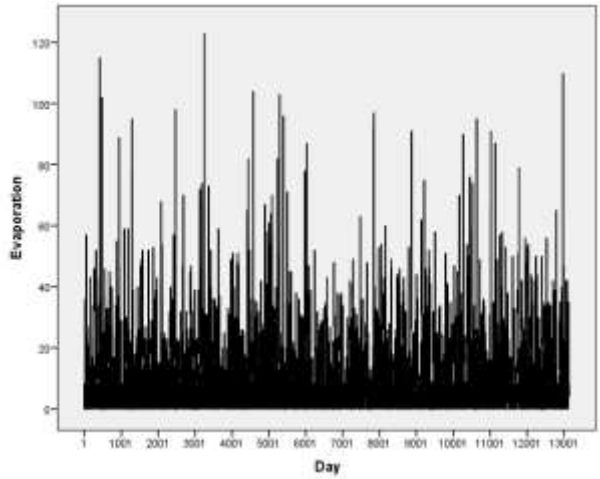


Figure 3: Evaporation

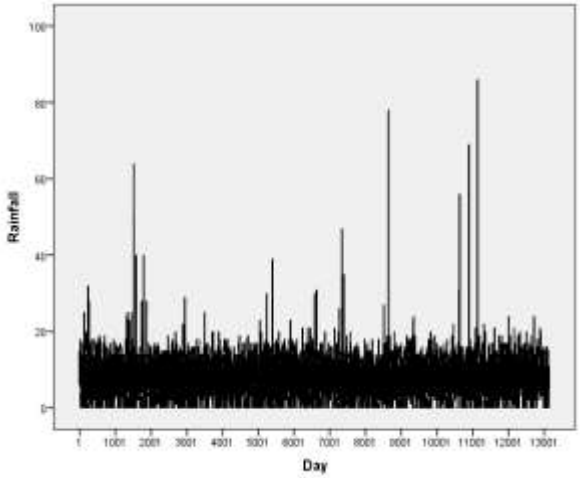


Figure 4: Rainfall

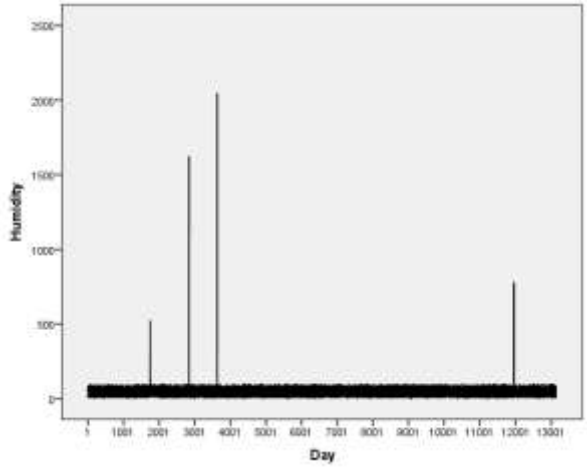


Figure 5: Humidity

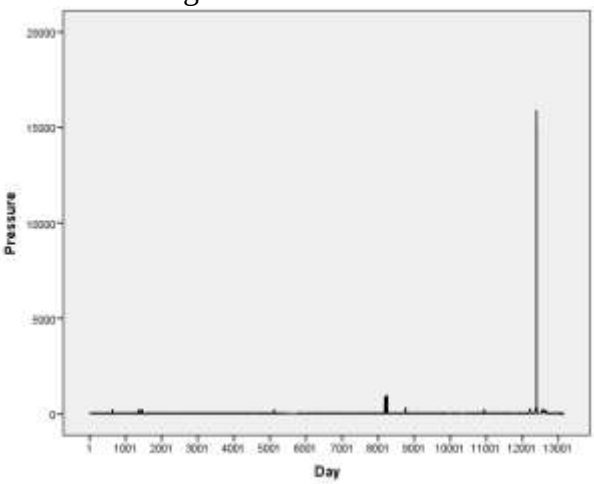


Figure 6: Pressure

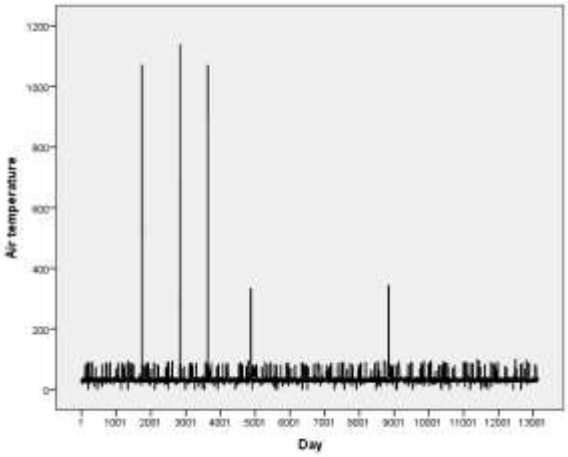


Figure 7: Air temperature

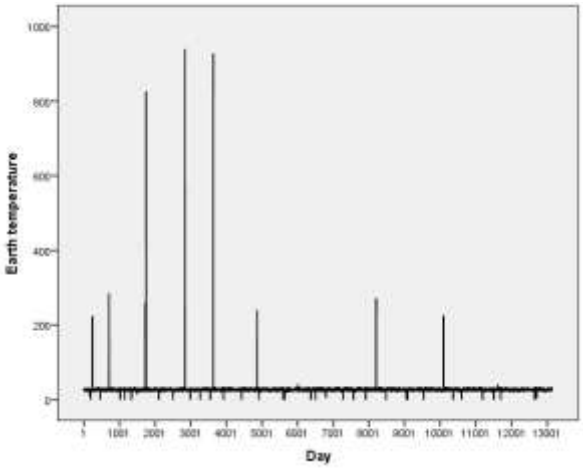


Figure 8: Earth temperature

4.3 Descriptive Statistics

Table 4.1 reports the descriptive statistics of the variables. The standard deviation of wind speed and pressure reflect a wider spread of values within the data from their true mean compared to other variables. The standard error across the variables is minimal due to the large sample size of 13,140. Though, the variables show departure from symmetry as indicated by the skewness and kurtosis, all the meteorological variables are positively skewed and leptokurtic.

Table 4.1 Descriptive Statistics

Variable	N	Mean	Standard deviation	Standard error	Skewness	Kurtosis
Wind speed	13140	127.78	266.410	2.324089	50.780	2925.328
Air temperature	13140	32.95	18.096	.157864	47.196	2725.491
Earth temperature	13140	26.84	14.305	.1247944	53.713	3216.729

Rainfall	13140	7.88	7.803	.068071	4.869	39.879
Evaporation	13140	7.88	3.744	.0326579	2.558	38.399
Humidity	13140	51.19	37.237	.3248451	18.055	879.434
Pressure	13140	30.71	140.188	1.22296	110.624	12519.690
Sunshine	13140	6.91	3.841	.0335046	4.891	115.479

4.4 Test for Stationarity

From Table 4.2 to Table 4.8 given below, it can be seen that the test statistic for all the meteorological variables is larger than the critical values across different level of significance. We therefore reject the null hypothesis of non stationarity and conclude that all the meteorological variables are stationary. With this, we proceed to use the Vector Auto regressive model.

Table 4.2 Air temperature

	Test statistic	1% critical value	5%critical value	10% critical value
Z (t)	-114.778	-3.430	-2.860	-2.570

Table 4.3 Earth temperature

	Test statistic	1% critical value	5%critical value	10% critical value
Z (t)	-114.530	-3.430	-2.860	-2.570

Table 4.4 Sunshine

	Test statistic	1% critical value	5%critical value	10% critical value
Z (t)	-113.268	-3.430	-2.860	-2.570

Table 4.5 Wind speed

	Test statistic	1% critical value	5%critical value	10% critical value
Z (t)	-114.838	-3.430	-2.860	-2.570

Table 4.6 Evaporation

	Test statistic	1% critical value	5%critical value	10% critical value
Z (t)	-113.757	-3.430	-2.860	-2.570

Table 4.7 Pressure

	Test statistic	1% critical value	5%critical value	10% critical value
Z (t)	-110.966	-3.430	-2.860	-2.570

Table 4.8 Humidity

	Test statistic	1% critical value	5%critical value	10% critical value
Z (t)	-114.632	-3.430	-2.860	-2.570

Table 4.9 Lag order selection

Lag order	AIC
1	28.2955
2	28.2956
3	28.2963
4	28.2959
5	28.2943
6	28.2955
7	28.2944
8	28.2944
9	28.2946
10	28.295
11	28.2954
12	28.2958
13	28.2959
14	28.2962
15	28.2962
16	28.2969
17	28.2968

18	28.2974
19	28.2984
20	28.2992
21	28.3
22	28.3009
23	28.3018
24	28.3028
25	28.3037
26	28.3049
27	28.3059
28	28.3069
29	28.3077
30	28.3085

The Akaike Information Criterion selects a lag of order 5 which has the minimum AIC for the Vector Autoregressive model.

4.5 Lagrange Multiplier Test for Conditional Heteroscedasticity (ARCH) and Autocorrelation

Table 4.10 below reports the heteroscedasticity test of the meteorological variables (Wind speed, Air temperature, Earth temperature, Humidity, Sunshine, and Pressure). The test result shows that we fail to reject the null hypothesis of no ARCH effect. That is, there is no ARCH effect. Table 4.11 gives the Box-Pierce statistic for autocorrelation test of the 5 lags which shows no autocorrelation is depicted in lag one and lag two..

Table 4.10 Testing for ARCH effect

Lags (p)	Chi2	Prob> Chi2
5	0.298	0.997

H_0: $no\ ARCH\ effects\ Vs\ H_1$: $ARCH\ (p)\ disturbance$

Table 4.11 Box-Pierce statistic test of autocorrelation

Lag	Chi2	Prob>Chi2
1	73.8460	0.1870
2	83.4712	0.05158
3	83.7472	0.04945
4	84.6913	0.04274
5	95.4457	0.00659

H_0: *no autocorrelation at lag order*

4.6 Fitting of the Vector Autoregressive Model

After the preliminary analyses were carried out, the stationarity test indicates that all the variables (Wind speed, Air temperature, Earth temperature, Humidity, Rainfall, Sunshine and Evaporation) are stationary. This stationarity condition satisfies the Vector Autoregressive model. From the Wind speed equation, there exists a short-run causality running from each of the variables (Wind speed, Air temperature, Earth temperature, Humidity, Rainfall, Sunshine and Evaporation) across the five different lags to Wind speed. A relationship exists between Wind speed and the meteorological variables. There is a linear dependence between Wind speed and Pressure in the presence of other meteorological variables which are Rainfall, Air temperature, Earth Temperature, Humidity and Sunshine. However, the result of the fitted model shows that each of these variables is not significant on Wind speed. These variables have been shown not to have any influence on Wind speed, having P-values greater than 0.05 which has made the null hypothesis not to be rejected as seen in table 4.13. For a one unit increase in wind speed, there is a 0.0579573 %, 0.0461428% and 0.0441661% increase in air temperature at lag1, lag 2 and lag 5 respectively which are not significant. Sunshine shows a positive increase of 0.2605161%, 0.9499953%, and 0.632326% at lag 3, lag 4 and lag 5 for a one unit increase on wind speed which are not significant. For a one unit increase in wind speed, rainfall shows 0.21722%, and 0.1987497% increase at lag 1 and lag 2 respectively which are not significant. Evaporation also shows 0.109271%, 1.098779%, 0.1357885%,

0.4432814% increase at lag 1, lag 3, lag 4 and lag 5 respectively which are not significant. For a one unit increase in wind speed, there are 0.0152288%, 0.0740397%, 0.1452261 and 0.0526453% increase on humidity at lag 1, lag 3, lag 4 and lag 5which are not significant. Pressure shows a 0.0091202%, and 0.0002462% increase at lag 2 and lag 5 respectively which are not significant as well. The non significance of meteorological variables on wind speed can be attributed to the large sample size. Lindley and Scott (1984) postulated that "the interpretation to be placed on significant at 5% depends on the sample size. It is more indicative of the falsity of the null hypothesis with a small sample than with a larger one". This postulate therefore confirms the reason behind the meteorological variables not being significant on wind speed.

Table 4.12 gives the normality test of the disturbances as not being normally distributed. With reasonably large sample sizes, the violation of the normality assumption should not cause a major problem (Ghasemi and Zahediasl , 2012).

Table 4.12 Jarque-Berra Test

Equation	**Chi2**	**Prob>Chi2**
Wind speed	2.5e09	0.0000

Table 4.13 Vector Autoregressive Model

Wind speed	**Coefficient**	**P> |Z|**
Wind speed		
L1	−0.0020484	0.815
L2	−0.0038457	0.660
L3	−0.002135	0.807

L4	−0.0055139	0.528
L5	0.036552	0.000
Air temperature		
L1	0.0579573	0.808
L2	0.0461428	0.847
L3	−0.0330338	0.890
L4	−0.0706185	0.769
L5	0.0441661	0.853
Earth temperature		
L1	−0.1719705	0.589
L2	−0.0401178	0.900
L3	−0.2015501	0.526
L4	−0.1454581	0.647
L5	−0.0866181	0.785
Sunshine		
L1	−0.4121061	0.506
L2	−0.0692198	0.911
L3	0.2605161	0.674
L4	0.9499953	0.125
L5	0.6323261	0.308
Rainfall		
L1	0.21722	0.480
L2	0.1987497	0.518
L3	−0.1827793	0.552
L4	−0.1392096	0.651
L5	−0.3023057	0.326
Evaporation		
L1	0.109271	0.870
L2	−0.6315361	0.345
L3	1.098779	0.101
L4	0.1357885	0.839
L5	0.4432814	0.508
Humidity		
L1	0.0152288	0.854
L2	−0.0701318	0.395
L3	0.0740397	0.369
L4	0.1452261	0.078
L5	0.0526453	0.523
Pressure		
L1	−0.0083662	0.615
L2	0.0091202	0.583
L3	−0.0009319	0.955
L4	−0.0030476	0.855
L5	0.0002462	0.988
Constant	111.834	0.000

4.7 The Impulse Response Function

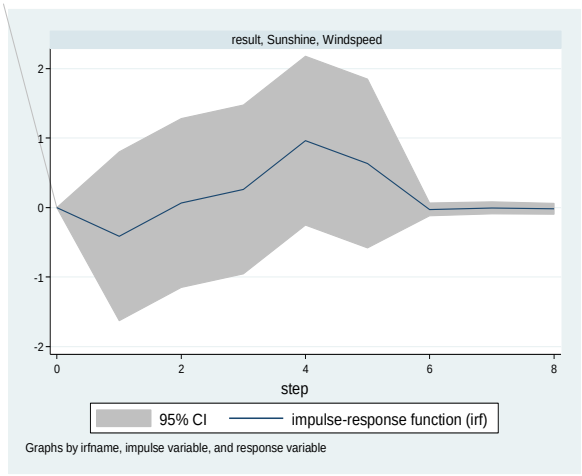


Figure (a): Sunshine and wind speed

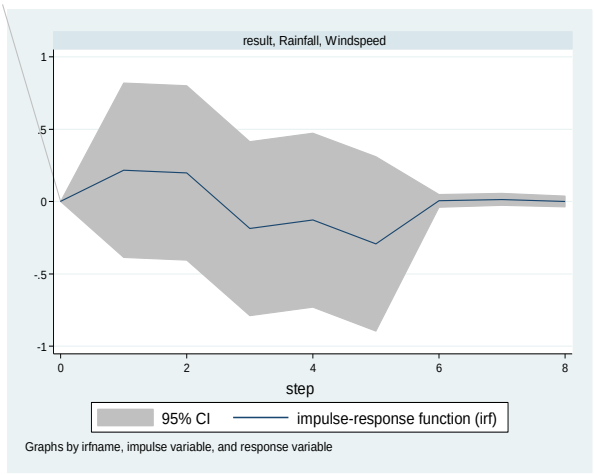


Figure (b): Rainfall and Wind speed

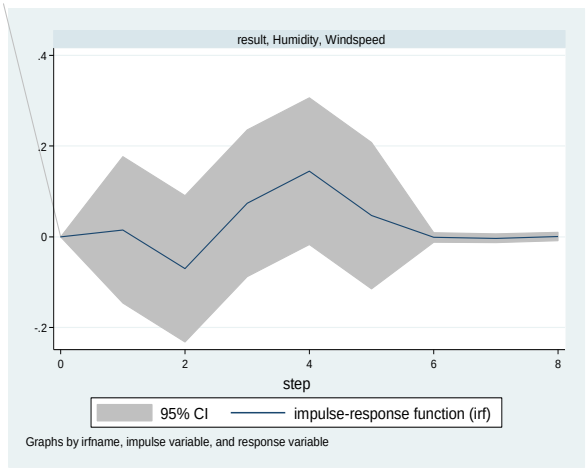


Figure (c): Humidity and Wind speed

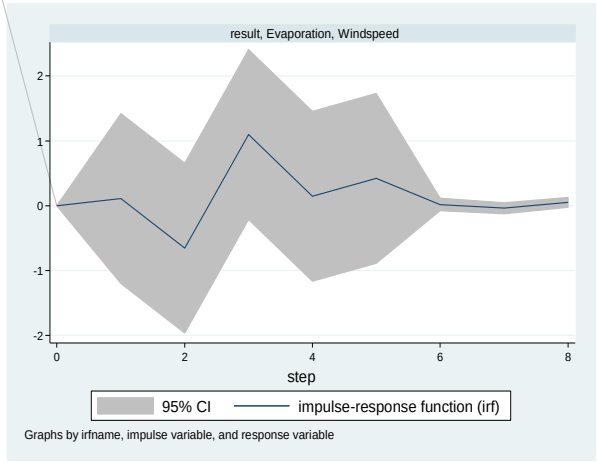


Figure (d): Evaporation and wind speed

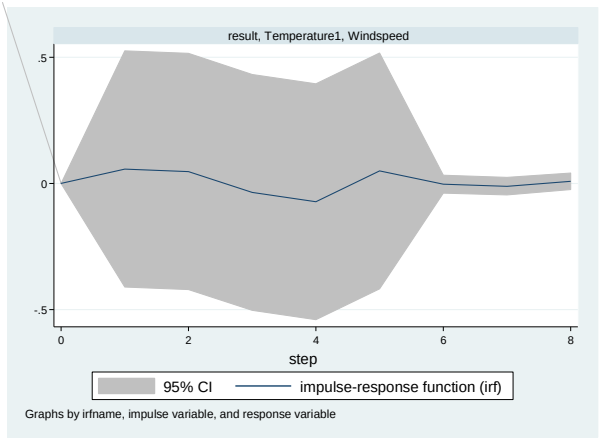


Figure (e): Air temperature and wind speed

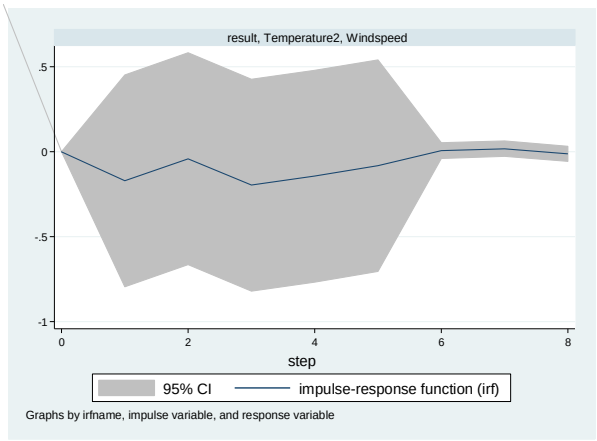


Figure (f): Earth temperature and wind speed

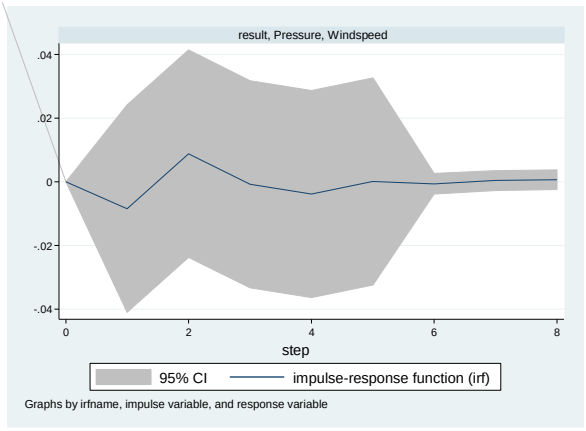


Figure (g): Pressure and Wind speed

Figure (a) is a graph of the Impulse-response function of Sunshine and Wind speed, where Sunshine is the Impulse variable and wind speed the response variable. The plot shows that Sunshine has an effect on Wind speed, with some degree of fluctuations around the zero axis. Majority of the effect exerted on Wind speed by Sunshine are positive.

Figure (b) shows the plot of the impulse-response function of Rainfall and Wind speed, where Rainfall is the impulse variable and wind speed is the response variable. From the plot, there exists rapid fluctuation around the zero axis. With Rainfall being the impulse variable, Wind speed depicts a negative effect from Rainfall. Though some days between 0 and 2 indicate a positive effect, but majority of the effects are negative as wind speed tends to drop on days after 2.

Figure (c) shows the plot of Humidity and Wind speed, where Humidity is the impulse variable and Wind speed is the response variable. A shock from Humidity has both positive and negative effect on Wind speed, with the majority being positive. The response of Wind speed to Humidity however dropped below 0 on days between 0 and 2 but increased on day 2, with its highest speed on day 4.

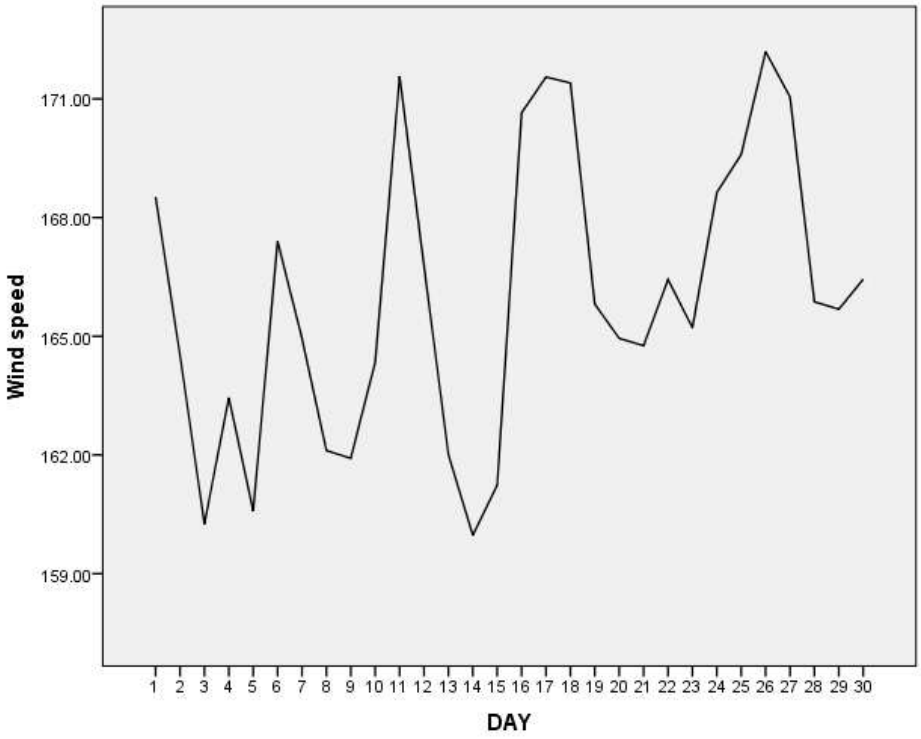
Figure (d) shows the impulse-response function of Evaporation and Wind speed, where Evaporation is the impulse variable and Wind speed is the response variable. Almost all the responses of Wind speed are positive. But there exists a sharp negative effect on day 2. The effect of Evaporation on Wind speed fluctuates rapidly above the 0 axis, with the highest positive effect on days 2 and 4.

Figure (e) shows the Impulse-response function of Temperature 1 (Air temperature) and Wind speed. As can be seen, there exists smooth fluctuations of Wind speed around the 0 axis, but with a negative effect on day 4.

Figure (f) is a graph of the impulse-response function of Temperature 2 (Earth temperature) and Wind speed. Wind speed depicts some degree of fluctuations on days between 0 and 3, with the highest positive effect on day 2. Figure (g) shows the impulse-response function of Pressure and Wind speed. With Pressure as the impulse variable, Wind speed fluctuates on days between 0 and 2, but becomes fairly smooth on days after 2.

4.8 Wind speed forecast

The values of Wind speed forecast for 30 days are presented in the appendix. This forecast shows some degree of fluctuations across the 30 days with an upward trend.



4.9 Multinomial Logistic Regression Model

This part presents the relationship that exist between Wind speed and Wind directions which are North, East, West, South, North east, North west, South east and South west . Considering the fact that there are more than two categories of the response variables which is the Wind direction, we applied the Multinomial Logistic Regression due to the categorical nature of Wind direction, where East is used as the reference category. The Multinomial Logistic Regression gives $K-1$ equations, where K is the total number of categories in the response variable.

Table 4.14 gives the parameter estimate of the Multinomial Logistic Regression. It can be seen from the table that there is no statistical significance between Wind speed and any of the wind directions. This is evident from the probability values (Sig column) which are all greater than the level of significance of 0.05. Looking at the model of Wind directions (North, West)

and Wind speed, the coefficient β is reported to be 0. This means there is no association between wind direction of North and Wind speed. The parameter β is estimated based on the Chi-square test that tests the null hypothesis that the estimate equals 0. It is also obvious from its odds ratio of exponentiated β which is given to be 1.00. North east, North West, South east, South west, South and west show a positive association with wind speed. It can also be seen from their respective odd ratio of exponential β which are greater than 1. Only Wind speed towards the South direction depicts a negative association, having a coefficient of 0.002. The negative association is also reflected in its odds ratio which is less than1.

Table 4.14 Multinomial Logistic Regression Model

Wind direction	$\boldsymbol{\beta}$	**Wald**	**Sig**	**Exp($\boldsymbol{\beta}$)**
Intercept	4.535	5.983	0.14	
North				
Wind speed	0.000	0.000	0.988	1.00
North east				
Wind speed	0.01	0.11	0.915	1.001
North west				
Wind speed	0.001	0.010	0.919	1.001
South				
Wind speed	-0.002	0.015	0.902	0.998
South east				
Wind speed	0.001	0.012	0.913	1.001
South west				
Wind speed	0.002	0.013	0.908	1.002
West				
Wind speed	0.000	0.001	0.977	1.000

CHAPTER FIVE

SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

5.1 Summary

In this study, Wind speed was modelled in the presence of some selected meteorological variables which are rainfall, air temperature, earth temperature, humidity, sunshine, pressure and evaporation. Before we went into the actual modelling, some preliminary works were carried out. The Augmented Dickey Fuller (ADF) test shows that all the variables are stationary, which prompted us to use the Vector Autoregressive model. The AIC was used to select a lag order of 5 for the Vector Autoregressive model. The variables were also tested for ARCH effect and autocorrelation of its residuals. There exist no heteroscedasticity from the test result but there exist no autocorrelation at the first two lags. The Vector Autoregressive model was fitted and it was found that there is a linear dependence between Wind speed and pressure in the presence of other meteorological variables. Though the result of this model shows that each of these variables have no significant influence on Wind speed. The Jarque-Berra test was also used to check the normality of the fitted model, which shows it is not normally distributed. This can be ignored considering the fact that the sample size is reasonably large (>30). The multinomial logistic regression model was also applied in modelling Wind speed and wind direction.

5.2 Conclusion

This research work modelled wind speed with some selected meteorological variables (Rainfall, air temperature, earth temperature, humidity, sunshine and evaporation). The result of the stationarity test shows that all the meteorological variables are stationary. Using the Vector Autoregressive model of order 5, it was discovered there exist a linear dependence between wind speed and pressure in the presence of other meteorological variables. But none

of these variables significantly affect wind speed, though there exist some degree of relationship among them. This can be seen visually from the impulse response function. Pressure and evaporation depict more fluctuation from the impulse response function on wind speed compared to other meteorological variables. A 30 day forecast was also generated using the Vector Autoregressive model which is presented in the appendix.

Wind speed however does not vary significantly across its wind directions, but shows some degree of dependence with its directions. This was discovered from the Multinomial logistic regression model of wind speed and its directions, with East as the baseline category. The directions of North and West however do not show dependence. This was also supported by the result of its odds ratio.

5.3 Recommendations

From the results of this research work, we make the following recommendations

- **i.** Wind speed of the turbines should be increased by increment of pressure and evaporation rate. This will in turn lead to greater output in wind energy .
- **ii.** Direction of wind speed should not be focused towards the Northern and Western directions as it does not have any effect on its speed
- **iii.** Wind turbines should be sited in places with high level of temperature or sunshine.

5.4 Contribution to Knowledge

Wind speed modeling with some selected meteorological variables including Sunshine and Wind direction were carried out using the vector autoregressive model and the multinomial logistic regression. The inclusion of Sunshine, air temperature, earth temperature and wind directions which was modeled through the application of the multinomial logistic regression

due to the categorical nature of wind direction forms the most important contribution of this research work because such phenomenon has not been previously investigated.

REFERENCE

Alice, M., Jethro, D., Stephan, W., and David, I., (2015). Electrical and Electronic Engineering, University of Strathclyde, Glasgow, Scotland

Ampaw, E. M., Akuffo, B., Larbi, S. O., and Lartey, S. (2013). Time Series Modelling of Rainfall in New Juaben Municipality of the Eastern Region of Ghana. *International Journal of Business and Social Science*, 4(8), 116-129.

Ayanlowo, E.A. (2014). A Time series modeling of climatic variables and its effect on climatic conditions. *Journal of Advances in Scientific Research and Applications.* 67-84.

Babazadeh, H., and Shamsnia, S. A. (2014). Modeling climate variables using time series analysis in arid and semi arid regions. *African Journal of Agricultural Research, 9*(26), 2018-2027.

Box, G. E. P. and Pierce, D. A. (1970), Distribution of residual correlations in autoregressive-integrated moving average time series models. *Journal of the American Statistical Association*, 65, 1509–1526.

Brown, B. G., Katz, R. W., and Murphy, A. H. (1984). Time series models to simulate and forecast wind speed and wind power. *Journal of Climate and Applied Meteorology*, 23(8), 1184-1195.

Cadenas, E., Rivera, W., Campos-Amezcua, R., and Heard, C. (2016). Wind speed prediction using a univariate ARIMA model and a multivariate NARX model. *Energies*, 9(2), 1-15.

Cardinali, A., and Nason, G. P. (2010). Costationarity of locally stationary time series. *Journal of Time Series Econometrics*, *2*(2). 1-33

Cherie, B. O., Dennis, S. M., Julius, C. O. (2014). Time series analysis using vector autoregressive model of wind speed in Bangui Bay and selected weather variables in Laoagcity, philipines. MPRA paper No. 54448. Retrieved from http://mpra.ub.unimuenchen.de/54448/

Czepiel, S. A. (2002). Maximum likelihood estimation of logistic regression models: theory and implementation. *Available at czep.net/stat/mlelr.pdf.*

Dikko, I. and Yahaya, D.B. (2012). Evaluation of wind power density in Gombe, Yola and Maiduguri, Northeastern Nigeria. *Journal of Research in Peace, Gender and Development*, 2(5), 115-122.

Fadare, D. A. (2008). A statistical analysis of wind energy potential in Ibadan, Nigeria, based on Weibull distribution function. *The Pacific Journal of Science and Technology*, 9(1), 110-119.

Frandsen, S., Barthelmie, R., Pryor, S., Rathmann, O., Larsen, S., Højstrup, J., and Thøgersen, M. (2006). Analytical modelling of wind speed deficit in large offshore wind farms. *Wind energy*, 9(1-2), 39-53.

Garba, A. D., and Al-Amin, M. (2014). Assessment of Wind Energy Alternative in Nigeria from the Lessons of the Katsina Wind Farm. *Civil and Environmental Research*. *6*(4), 91-94.

Ghasemi. A. and Zahediasl. S. (2012). Normality tests for statistical analysis: a guide for non-statisticians. *International Journal of Endocrinology and Metabolism*, *10*(2), 486-489.

Gnana, S.K. (2011). Computing models for wind speed prediction in renewable energy systems. *IJCA Special Issue on "Computational Science, New Dimensions and Perspectives NCCSE*.

Harvey, P.D. (1994). The impact of condom prices on sales in social marketing programs. *Studies in family planning*, 52-58

Ibrahim, M. Z., Hwang, Y. K., Ismail, M., and Albani, A. (2014). Wind Speed Modeling for Malaysia. *International Journal of Renewable Energy Research (IJRER)*, *4*(4), 912-923.

Idris N.A, Lamin H. S, Ladan M. J and Yusuf B. H. (2012). Nigeria's wind energy potentials; the path to a diversified electricity generation-mix. *International Journal of Modern Engineering Research*,2(4),2434-2437

Johansen, S. (1988). Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control*, 12(2-3), 231-254.

Johansen, S. (2000). Modelling of cointegration in the vector autoregressive model. *Economic Modelling*, 17(3), 359-373.

Katarina, J. (2007). *Cointegration analysis of climate change. An exposition.* Department of Economics, university of Copenhagen.

Kuha, J. (2004). AIC and BIC: Comparisons of Assumptions and Performance. *Sociological methods and research*, 33(2), 188-229

Lindley, D.V., and Scott W.F. (1984). *New Cambridge Statistical Tables*. Cambridge University press

Minano, Z. R., Anghel, M., and Milano, F. (2013). Continuous wind speed models based on stochastic differential equations. Applied Energy, 104, 42-49

Minh, H.Q., Frederic,N., Najib, E, and Abdelaziz, H. (2011). Power management of a variable speed wind turbine for stand-alone systems using Fuzzy logic. *International conference* on (1404-1410).

Odo, F. C., Ugbor, D. O., and Ugwuoke, P. E. (2013). Twenty-five years of wind data in south-eastern locations of Nigeria: Modeling and prediction of wind potential. *International Journal of Physical Sciences*, 8(2), 83-88.

Ohashi, O., and Torgo,L.(2012). Wind speed forecasting using spatio-temporal indicators. *Proceedings of the 20th European Conference on Artificial intelligence*,pp.975-980.IOS Press

Oluseyi. A., Fagbenle, R.O., Katende, J., Aasa, S.A., and Okeniyi, J.O. (2013). Wind profile characteristics and turbine performance analysis in Kano, north-western Nigeria. *International Journal of Energy and Environmental Engineering*. 4(1),1-15

Petroula. L., Galanis, G., Siebert, N., Kariniotakis, G., Katsafados, P., Pytharoulis, I., and Kallos, G. (2008). Improvements in wind speed forecasts for wind power prediction purposes using Kalman filtering. *Journal of Wind Engineering and Industrial Aerodynamics*, 96 (12), 2348-2362

Rao, C. R. (2002). Linear Statistical Inference and Its Applications. John Wiley and sons, inc, second edition.

Reddy, G. K. K., Reddy, S. V., Sarojamma, B., and Ramkumar, T. K. (2015). Statistical Analysis for Wind Energy Estimation over Gadanki, India. *Research Journal of Engineering and Technology*, *1*(2), 30-40.

Selin, K., and Kenneth W. Corscadden.(2013). Wind speed estimation:Incorporating seasonal data using Markov chain.Department of Business Administration, Faculty of Economics and Administrative Sciences, Yalova University,Yalova, Turkey

Sloughter, J. M., Gneiting, T., and Raftery, A. E. (2010). Probabilistic wind speed forecasting using ensembles and Bayesian model averaging. *Journal of the American Statistical Association*, 105(489), 25-35.

Tilman, P. (2006). Department of Mathematics, informatics and Statistics, University of Munchen

Tsay, R. S. (2005). *Analysis of Financial Time series*. John Wiley and Sons.

Turasie, A. A. (2012). Cointegration modelling of climatic time series.

Van, D., S.J., Wagner, L.E., Skidmore, E.L., and Tatarko,J. (2005). Comparison of the Weibull model with measured wind speed distributions for stochastic wind generation. *Transactions of the ASAE*,48(2),503-510.

APPENDIX

30 days Wind speed forecast values

Predicted wind speed	Day
168.52	1
164.52	2
160.25	3
163.45	4
160.58	5
167.41	6
164.94	7
162.11	8
161.92	9
164.35	10
171.57	11
166.79	12
162.01	13
159.97	14
161.24	15
170.65	16
171.55	17
171.40	18
165.81	19
164.95	20
164.76	21
166.44	22
165.22	23
168.63	24
169.60	25
172.20	26
171.05	27
165.87	28
165.68	29
166.44	30