T. R.
VAN YÜZÜNCÜ YIL UNIVERSITY
INSTITUTE OF NATURAL AND APPLIED SCIENCES
DEPARTMENT OF STATISTICS

IMPACT OF CLIMATE CHANGE ON RESIDENTIAL ELECTRICITY CONSUMPTION IN ERBIL: AN ECONOMETRIC ANALYSIS

M. Sc. THESIS

PREPARED BY : Zewar Omar ISMAEL
SUPERVISOR : Asst. Prof. Dr. Muhammed Hanifi VAN

VAN-2021

T. R.
VAN YÜZÜNCÜ YIL UNIVERSITY
INSTITUTE OF NATURAL AND APPLIED SCIENCES
DEPARTMENT OF STATISTICS

IMPACT OF CLIMATE CHANGE ON RESIDENTIAL ELECTRICITY CONSUMPTION IN ERBIL: AN ECONOMETRIC ANALYSIS

M. Sc. THESIS

PREPARED BY: Zewar Omar ISMAEL

VAN-2021

ACCEPTANCE and APPROVAL PAGE

This thesis entitled "…………………………………………………….............. ...." presented by .... under supervision of .... in the department of .... has been accepted as a M. Sc. / Ph.D. thesis according to Legislations of Graduate Higher Education on ….... /......./.......... with unanimity / majority of votes members of jury.

Chair: .... Signature:

Member: .... Signature:

Member: .... Signature:

Member: .... Signature:

Member: .... Signature:

This thesis has been approved by the committee of The Institute of Natural and Applied Science on ………/........../….…….. with decision number ....

Signature
……………..………..
Director of Institute

THESIS STATEMENT

All information presented in the thesis obtained in the frame of ethical behavior and academic rules. In addition, all kinds of information that does not belong to me have been cited appropriately in the thesis prepared by the thesis writing rules.

Signature
Zewar Omar ISMAEL

ABSTRACT

IMPACT OF CLIMATE CHANGE ON RESIDENTIAL ELECTRICITY CONSUMPTION IN ERBIL: AN ECONOMETRIC ANALYSIS

ISMAEL, Zewar Omar
M.Sc. Thesis, Department of Statistics
Supervisor: Asst. Prof. Dr. Muhammed Hanifi VAN
December 2021, 83 pages

Identifying changes in energy consumption can be affected by climate change. Rising intensity of climate change, especially global warming, may have a negative impact on energy consumers and power companies. It is neccessary to identify the challenges that determine how changes in the economy, environment, and technology affect the future of energy system because of the climate changes. The objective of thesis is to analyze the long-term link between electricity usage and identifying climatic factors such as air temperature, humidity, and wind speed in Erbil city of Iraq, using monthly data from 2009 to 2019. Data were obtained from two sources. First one, the electricity load data wass obtained from the Erbil directorate of electricity. The latter, wind speed, temperature, and rate of humidity, were obtained from the general directorate of agriculture and weather. In this study, the dependent variable is electrical load, and the independent variables are humidity, temperature, and wind speed. The graphs of the series were examined to see the general trends of the study variables. Then, ACF and PACF graphs were examined to understand whether there was seasonality or not. Then, the stationarity of the series was examined by conventional unit root tests. The cointegration relationship between the electricity load and wind speed series is tested using the Jonahansen cointegration method. The cointegration relationship between the electrical load, wind speed, and air temperature and humidity is also tested using the ARDL method. According to the results obtained, it has been determined that the series of electric load and wind speed move together in the long run. According to results of the ARDL model, it is found that there is a statistically significant long-term relationship between electric charge and humidity, wind speed and air temperature. These results show that climate change will increase energy use.

Keywords: Climate change, Cointegration analysis, Electricity consumption.

ÖZET

İKLİM DEĞİŞİKLİĞİNİN ERBİL'DE KONUT ELEKTRİK TÜKETİMİNE ETKİSİ: EKONOMETRİK BİR ANALİZ

ISMAEL, Zewar Omar
Yüksek Lisans Tezi, İstatistik Anabilim Dalı
Tez Danışmanı : Dr. Öğr. Üyesi Muhammed Hanifi VAN
Aralık 2021, 83 sayfa

Enerji tüketimindeki değişikliklerin belirlenmesi iklim değişikliğinden etkilenebilir. Küresel ısınma başta olmak üzere iklim değişikliğinin artan yoğunluğu, enerji tüketicileri ve enerji şirketlerini ekonomik olarak olumsuz etkileyecektir. Sonuç olarak, iklim değişikliğinin bir sonucu olarak ekonomide, çevrede ve teknolojide meydana gelen değişikliklerin enerji sisteminin geleceğini nasıl etkileyeceğini belirleyecek temel zorlukların belirlenmesi hayati önem taşımaktadır. Bu tezin amacı, Irak'ın Erbil kentinde elektrik tüketimi ile hava sıcaklığı, nem ve rüzgar hızı gibi iklim faktörleri arasındaki uzun dönemli ilişkiyi 2009-2019 arasındaki aylık verileri kullanarak araştırmaktır. Verileri iki kaynaktan elde edilmiştir. Birincisi, elektrik yükü verileri Erbil elektrik müdürlüğünden elde edilmiştir. İkincisi, rüzgar hızı, sıcaklık ve nem oranı, tarım ve hava durumu genel müdürlüğünden alınmıştır. Bu çalışmada bağımlı değişken elektrik yükü, bağımsız değişkenler ise nem, sıcaklık ve rüzgar hızıdır. Bu amaçla çalışmada ilk olarak serilerin genel seyrini görmek için serilerin grafikleri incelenmiştir. Ardından mevsimselliğin olup olmadığını anlamak için ACF ve PACF grafikleri incelenmiştir. Daha sonra serilerin durağanlıkları geleneksel birim kök testleri ile incelenmiştir. Birinci farkalarında durağan olan electricity load and wind speed serileri arasında eşbütünleşme illişkisi Jonahansen eşbütünleşme medodu ile test edilmiştir. Farklı düzeyde durağan olan elektrik yükü, rüzgar hızı, ve hava sıcaklığı ve nem oranı serileri arasındaki eşbütünleşme ilişki ARDL metodu kullanılarak test edilmiştir. Elde edilen sonuçlara göre, elektrik yükü ve rüzgar hızı serilerin uzun dönemde birlikte hareket ettiği belirlenmiştir. ARDL modeli tahmin sonuçlarına göre ise elektrik yükü ile nem oranı ve rüzgar hızı Bu sonuçlar ilkim değişikliğini enerji kullanımını artıracağını göstermektedir.

Anahtar kelimeler: Elektrik tüketimi, Eşbütünleşme analizi, İklim değişikliği.

ACKNOWLEDGMENT

First and foremost, I would like to thank Allah for helping me in my successful study and I would like to thank my supervisors Asst. Prof. Dr. Muhammed Hanifi van and for their supports, guidance, encouragement, and patience. Then, my best regards also to all other members of the Statistic Department, especially to Prof. Dr. H. Eray Çelik, Dr. K. Dler Hussein and Assistant Lecturer M. Yaseen Omar I am also grateful to everyone whom I have learned many forms. And the great characteristics of my parents who they always help me and support me in every field in my life, so words can't show gratitude to their supports. Furthermore, my evergreen gratefulness goes to my wife for helping me and encouraging me in achieving my best goals. I am deeply indebted to Assistant Lecturer Y. Rebaz Othman and all friends for their support. I would like to express my deepest appreciation to other my friends and all. Those who provided me with assistance to complete this thesis. Finally, thanks for everyone who has helped me even by a word.

2021

Zewar ISMAEL

TABLE OF CONTENTS

Page

Page

LIST OF TABLES

Table **Page**

LIST OF FIGURES

Figure **Page**

SYMBOLS AND ABBREVIATIONS

Some symbols and abbreviations used in this study are presented below, along with descriptions.

Symbols	Description
$\boldsymbol{a}$	Eigenvector
$\boldsymbol{\beta}$	Constant term
$\mathbf{C_t}$	Cycle
$\boldsymbol{E}$	Expectation value
$\boldsymbol{\Sigma}$	Covariance matrix
$\mathbf{\Sigma}$	Covariance matrix
$\mathbf{I_t}$	Noise
$\mathbf{p}$	Correlation
$\mathbf{r_k}$	Autocorrelation
$\mathbf{S_t}$	Seasonal
$\mathbf{T_t}$	Trend
$\mathbf{U_t}$	Error term
$\mathbf{W(r)}$	Standard Brownian motion
$\bar{\boldsymbol{y}}$	Mean of variable
$\mathbf{z_t}$	The value of a variable at time period t
%	Percent
$\emptyset$	Constant time series
t	Time
$\boldsymbol{\mu}$	Mean of variables
$\boldsymbol{\sigma}$	Standard deviation
$\boldsymbol{\sigma 2}$	Variance
$\boldsymbol{\Omega}$	Omega
$\boldsymbol{\lambda}$	Eigenvalue

Abbreviations **Description**

ACF	Auto Correlation Function
ADF	Augment Dickey Fuller
AR	Autoregressive
ARDL	Auto-regressive Distributed Lag
ARIMA	Autoregressive integrated moving average
ARMA	Autoregressive moving average
COV	Covariance
DW	Durbin-Watson
EKC	Environmental Kuznets curve
i.i.d	Independent identically distributed
KPSS	Kwiatkowski, Phillips, Schmidt and Shin
MA	Moving average
OLS	Ordinary list square
PACF	Partial Auto Correlation Function
PP	Phillip Peron
SMC	Squared Multiple Correlation
V	Variable
VAR	Variance
VAR	Vector Autoregressive
VEC	Vector error correction
VECM	Vector error correction model

1. INTRODUCTION

Climate change can have an impact on determining changes in energy consumption. The increasing intensity of climate change, the increase in carbon dioxide emissions leads to global warming. Thus, there is a need to identify key challenges that how changes in economic, environment and technology under climate change will affect the future of power system. The connection between power system, climate change and economic has many aspects and it is complicated Queiroz et al. (2016). A major effect of climate change on energy usage is how much it raises or lowers the temperature. There are a number of empirical studies that suggest temperature is the most important weather component when it comes to determining how much energy people consume. Buildings need to be heated in cold weather, and cooled in hot weather (Silhvola et al., 1997).

Iraq has performed relatively few researches on the coincidental link between climate change and energy consumption. Electricity is heavily utilized for heating in the winter and cooling in the summer, particularly in Erbil. However, the felt air temperature, humidity and wind speed can increase or decrease the electrical load.

Cooler or warmer weather introduces a high demand for temperature conditioning in tropical and subtropical countries. The main source of indoor temperature conditioning is electricity. The fluctuation of weather can have a major impact on electricity via different aspects such as socially and economically.

Erbil city has suffered a huge damage of electricity supply as the result of Gulf was in 1991 and civil war between 1992-1996. Most of the stations, transmission lines and distributional stations were destroyed. The main source of electricity for was from the hydropower stations of Dokan and Derbandikhan in Sulaymaniya province (Kadir, 2020).

The system power load is equal to the total of all users' electrical loads. The purpose of predicting system-level power load is to forecast future system-level electricity load. A thorough grasp of the system's features enables the development of realistic forecasting models and the selection of relevant models for various scenarios. Erbil city is very cold in winter and very hot in the winter. Most people are relying on electricity for the main sources of cooling and heating.

1.1. The Aim of the Thesis

The purpose of this thesis is to examine the link between climate change and power usage in Iraq using time series data from 2009 to 2019. To avoid omitted variables that occur from bivariate models this thesis will apply a multivariate model by including several variables, which are humidity, temperature, and wind speed, in order to examine cointegration and casual relationship between climate change and energy consumption.

Research Aim 1: To review the available literature on the relationship between climate change and energy consumption

Research Aim 2: To assess the long run and short run relationship between electricity load and other factors such as temperature, wind speed and rate of humidity in Erbil

1.2. Overview of Thesis

This thesis includes the following chapters. The next chapter is a literature review in which the overview of cointegration analysis, unit root test, causality analysis and the relationship between electricity consumption and climate factors are extensively reviewed and presented.

The chapter three provides an overview of extensive methodologies that we will apply in our study. It highlights the basic definition of time series including trend, seasonality, autocorrelation and partial autocorrelation and decomposition. It provides the methodologies for AR, MA, ARMA, and ARIMA. An extensive background of unit root test has been discussed such as Dickey-Fuller, ADF, the Phillip Peron approach and KPSS test. Adonal, we have provided an overview of structural breaks, vector autocorrelation model, cointegration analysis, causality analysis and cointegration with structural beak.

The chapter four provides the main results of the study. It provides graphical presentation to detect stationarity of time series, explanatory of data analysis, unit root test including ADF and Phillip perron approaches. We have provided the results for cointegration analysis using Johansen, VEC and ARDL approaches. Additionally, chapter five provides the relevant discussion of our results obtained. Our results are compared with other studies for the consistency and reliability of our results. Finally, an appropriate conclusions and recommendations are given in chapter six.

2. LITERATURE REVIEW

2.1. Background

This chapter highlights the revision of various studies about the cointegration and cointegration analysis. It reviews different methods regarding to cointegration analysis including VEM, Johansen test, the Auto-regressive Distributed Lag cointegration (ARDL) method, VECM Granger causality framework, linear and non-linear causality test. It also highlights methodologies for testing unit root in relation to the electricity consumption and climate change variables.

2.2. Relative Works

Belloumi (2009) used cointegration and causality analysis to examine Tunisia's energy usage and GDP. The author employed the Johansen cointegration technique to examine the causal link between per capita GDP (PCGDP) and per capita energy consumption (PECE) using the Vector Error Correction (VEC) model. Prior to implementing the Johansen cointegration method, multivariate cointegration between variables is investigated. Between 1971 and 2004, data were collected. Rather than using a vector autoregressive model, the Granger causality test was applied using a vector error correction model (VECM) (VAR). The author discovered that the PECE and PCGDP are linked in Tunisia using a cointegration vector. Between energy use and GDP, a long term bi-conditional casual connection was discovered. Similarly, a short-run unidirectional causal relationship between energy use and GDP was discovered. His findings highlighted a critical implication: energy consumption should be viewed as a constraint on Tunisia's GDP development. The author concluded by emphasizing the presence of energy inside the cointegration space and a long run by conditional link between the two time series variables (Belloumi, 2009).

Boukhelkhal and Bengana (2018) studied the link between economic development, electricity consumption, and carbon dioxide emissions in relation to additional factors such as average temperature and openness of trade. The research examined data from north African nations between 1971 and 2014. The author employed the Auto-regressive Distributed Lag cointegration (ARDL) technique, sometimes referred to as the Bounds approach. The major goal of this study was to determine the existence of cointegration between these variables that are used to assess long-term and casual relationships using the Bounds testing approach. Their findings suggest that in Egypt and Algeria, there is a long run link between economic development and CO2 emissions, whereas in Morocco, there is a long run relationship between power consumption and CO2. The data indicate that energy consumption is driving economic development in the aforementioned nations. Additionally, temperature fluctuation is a key element impacting Egypt's energy usage. There are a large number of causal links, as shown by the Yamamoto and Toda Granger causality test. Their findings revealed that trade liberalization and economic expansion are responsible for CO2 emissions. It is worth noting that the growth theory was verified in the long and short run in Morocco and Egypt, but not in the long run in Algeria. This indicates that the critical component for economic development in all three countries is the ability to sustainably increase energy output. Additionally, the author stated that rising temperatures resulted in an increase in energy consumption, which may be a critical element in economic development and environmental sustainability (Boukhelkhal and Bengana, 2018).

Rahman and AbulKashem (2017) evaluated industrial growth, carbon emissions, and energy consumption in Bangladesh using empirical data from ARDL cointegration and Granger causality analysis. Their study examined the empirical cointegration, short- and long-run connections, and casual dynamics of industrial growth, carbon emissions, and energy consumption in Bangladesh between 1977 and 2011. The author applied the Granger causality test and the ARDL Bounds Testing technique on the AVR. The evaluation of ARDL limitations and subsequent cross-checking investigations indicated that industrial production, carbon emissions, and energy consumption are all long- and short-term cointegrated. Energy consumption and industrial expansion have a significant favorable influence on carbon emissions in the long and short run. According to the causality test, there is a unidirectional causal relationship between industrial growth and carbon emissions, energy consumption and carbon emissions, and industrial growth and energy consumption (Rahman and AbulKashem, 2017).

Hatzigeorgiou et al in 2011 examined the causal relationships between Carbon emission, GDP and energy intensity in Greece between the time period of 1977–2007. They applied Granger-causality tests based on VECM for this purpose. Apart from statistical independent test and causality test, tests Johansen multivariate cointegration rank and unit root were applied. They have found a bidirectional and unidirectional causalities between variables. A Variance Decomposition Analysis model based on Choleski approach was performed and comparisons of their results with previous research were provided. They obtained stationarity in the first differences in the selected variables, and they were cointegrated. The author found a uni-directional causality between Carbon emission, GDP and energy intensity and bi-directional relationship was found between Carbon emission, GDP and energy intensity (Hatzigeorgiou et al., 2011).

Zhang et al (2019) studied Climate effects such as temperature and electricity consumption. The objective of the study was to obtain parameter estimation of electricity because of temperature at the country scale in China. The temperature data collecting between 2006 and 2015 from the National Climatic Data Center (NCDC) of the National Oceanic and Atmospheric Administration (NOAA). Fixed effect panel model has been applied to the data and they found that increasing daily temperature influences the electricity consumption in the rural area (Zhang et al., 2019).

In a research published in Environ Dev Sustain, the cointegration connection between GDP and energy use in Italy was examined. He examined the relationship between per capita GDP and per capita energy use in Italy between 1970 and 2009, using yearly data. The time series characteristics KPSS, PP, ERS, and ADF were evaluated using four-unit root tests. Additionally, XMR and ZA tests were used to determine the presence of potential structural fractures. The author demonstrated that both series were non-stationary and that the cointegration relationship between the two variables was strong. The short run causality test established a causal link between energy use and GDP. Additionally, there was a long-run bidirectional causal relationship between energy use and GDP. The log transformation was used to create the two-time series. The inter-quality range method was used to identify many outliers in the samples (Zhang et al., 2019).

Nazlioglu et al. (2014) used linear and nonlinear causation and cointegration analysis to explore the causative connection between economic development and energy consumption. The data were collected between 1967 and 2007 in Turkey. Three analytical techniques were utilized in this study: linear Granger causality testing, nonlinear Granger causality testing, and bound testing cointegration. Their findings revealed that electricity consumption and economic growth had a long run cointegration connection. Economic growth and power consumption had a long-run cointegration connection. The investigation of linear Granger causality using an error correction model revealed a bidirectional causal relationship between economic development and electricity consumption in Turkey in the short and long term. The authors established the series' non-linearity by employing a Dechert and Scheinkman test (Nazlioglu et al., 2014).

Another research published in 2013 by Solarin and Shahbaz evaluated the causal relationship between urbanization, economic growth, and electricity consumption in Angola from 1971 to 2009. The unit root tests of Lee and Strazicich, 2003, and Lee and Strazicich, 2004 were used to evaluate the time series' stationarity characteristics. They examined long run relationships using the ARDL bounds test in conjunction with the Gregory–Hansen structural break cointegration method (Solarin and Shahbaz, 2013).

The Granger causality test was used to determine the direction of causation between urbanization, electricity consumption, and economic growth using the vector error correction model. Long range correlations were discovered, as well as bidirectional causation between economic development and power usage. Additionally, power usage and urbanization are causally related. The author found that Angola's power consumption has a beneficial effect on economic growth in the long run, as indicated by the long run coefficients of the ARDL, the FMOLS, and the DOLS. During the civil war, energy consumption had a less beneficial effect on economic growth (Solarin and Shahbaz, 2013).

Johansen at (1991) has tested and hypothesized Vectors of Cointegration in Gaussian Vector Autoregressive (VAR) Models. He presented a likelihood approach to analyze cointegration in VAR models using seasonal dummies, constant terms and

Gaussian error. A likelihood ratio test of cointegration was discussed in order to obtain asymptotic distribution of the test. the maximum likelihood estimator of the cointegrating associations were characterized and structural hypothesis of these associations was formulated. The author showed asymptotic distribution of likelihood estimator was mixed Gaussian. The author stated that when a specific eigenvalue issue is tackled, the eigenvalues and eigenvectors can be calculated. Chi-square distribution can be used to inference on the cointegration rank (Johansen, 1991).

Rafindadi and Ozturk in 2017 studied an effect of renewable energy consumption on economic growth in German using cointegration test combination. A quarterly time data was collected between 1971Q1 and 2013QIV. A Clemente- Montanes Reyes detrended structural break test was employed, the Bayer-Hanck combined cointegration test and the ARDL bounds testing methods to cointegration. further, the vector error correction model was used when causality analysis was observed. Their results illustrated that cointegration was existed between variables of interest. They showed that 1 % increase renewable energy boosted economic growth in Germany by 0.2194%. Further, increasing capital led to increase in economic growth by 1.1320% while 1% rise in labor productivity boosted 0.5125% rise in economic growth. The causality relationship was existed between economic growth and renewable energy consumption. Whilst a bi-directional relationship was found between renewable energy consumption and capital (Rafindadi and Ozturk, 2017).

Another research published in 2012 by Shahbaz et al. examined the link between electricity consumption, capital accumulation, and economic growth in Romania using cointegration and causality methods. They examined the relationship between electricity use and capital per capita consumption and economic development. The ARDL bounds testing approach was used to examine the long term relationship between capital per capita usage, economic growth, and electricity use from 1980 to 2011. (Muhammad Shahbaz et al, 2012). Additionally, to examine the direction of causality between capital per capita use, economic growth, and electricity consumption, the Toda and Yamamoto (1995) non-causality test was used to determine whether a bidirectional causal relationship existed between capital per capita use, economic growth, and electricity consumption. Meanwhile, a unidirectional causal relationship between capital usage and electricity consumption was also found. They argued that energy conservation regulations might potentially stifle economic growth by lowering power usage (Toda and Yamamoto, 1995).

Kumar et al. (2013) examined the link between the environmental Kuznets curve and the role of coal use in India using cointegration and causality analysis. In the case of India, they were interested in coal use, economic development, trade openness, and carbon emissions. The sequence of integration of the variables of interest was determined using a structural break unit test. The long-run relationship between the variables was examined using Pesaran et alARDL.'s bounds testing approach for cointegration. Economic growth, coal consumption, CO2 emissions, and trade openness were all cointegrated over the long run. Additionally, a long run and a short run environmental Kuznets curve (EKC) existed. Carbon emissions were exacerbated by coal usage and trade liberalization. Causality analysis revealed a feedback hypothesis between CO2 and economic growth, and the same conclusion was reached for CO2 and coal use (Kumar et al., 2013).

Salahuddin et al in 2014 explored the association between economic growth, reduction in poverty and financial development in Bangladesh using cointegration and causality analysis. The quarterly data were obtained between 1975 and 2011. A structural break unit root test was applied to examine the order of integration of the variables. They examined long run association among variables by using ARDL bounds testing while utilizing dummy to accommodate for a structural break stemming in the series. The authors found a long run association between economic growth, reduction in poverty and financial development and financial growth assist in poverty reduction but it was not linear (Salahuddin et al., 2014).

In 2015, Baker et al. used cointegration and causality tests to determine the link between regional aviation and Australia's economic development. We offered empirical evidence for a long- and short-run causal connection between regional aviation and economic growth. The authors used data from 1985 to 2010 to study 88 regional airports in Australia in order to determine the influence of regional air aviation on regional economic growth. The airport activity level and real aggregate taxable income were the factors of interest. Between variables, there existed a bidirectional connection. They discovered that regional airports have an effect on regional economic growth and that regional air aviation is directly impacted by the economy. Additionally, a significant connection was discovered between the two series. Correlation coefficient of 0.789 was determined between the two series. Correlations were comparable for regional airports (0.795) and distant airports (0.690), (Baker et al., 2015).

Chontanawat conducted another study in which he examined the link between carbon emissions, energy consumption, and economic growth in ASEAN using cointegration and causality analysis. Between 1971 and 2015, the primary objective of this study was to examine the connection between carbon emissions, energy consumption, and economic production in ASEAN (Chontanawat and Jaruwan et al., 2006). The author found that there was a long run relationship and there was causality relationship between the interested covariates this means that energy consumption and output were cointegrated to carbon emissions. The results from cointegration illustrated a long run association between variables, meaning that energy consumption and economic activities were cointegrated to carbon emissions. The results causality analysis showed the existence of the association between these variables (Jaruwan et al., 2006).

Ramos investigated the Granger-causality link between imports, exports, and economic development in Portugal between 1865 and 1998 in 2001 using data from 1865 to 1998. The stationarity characteristics of the data were determined by performing DF, SDF, and PP tests on the single- and second-unit root. Their empirical findings indicated that unidirectional causation did not exist between the variables of concern. There was a feedback effect between the rise of exports and imports. They discovered no evidence of a causal relationship between import and export growth (Francisco and Ramos, 2001).

Kassouri and Altıntaş in 2020 examined nonlinearity threshold cointegration, and frequency domain causality relationship among Turkish Lira and stock price. They analyzed the reaction of Modelling complex asymmetric impacts and non-linear time series methods to find association between exchange rate and stock prices using data between 2003 and 2018. According to the authors, several kinds of economic methodologies have been employed in order for exchange rate and Turkish stock market movements to be characterized as dynamically intertwined. The evidence of asymmetric threshold cointegration indicated that the Turkish financial market can be disguised by linear time series methods. Cointegration was found between Turkey's stock and currency markets, showing that one market may be anticipated from the other. This contravened the hypothesis's claims about its usefulness. He and his coworker Halil Alttaş (Kassouri, 2020) showed that they could accurately forecast the USD-Turkish lira exchange rate as well as the money supply and interest rates when stock prices fluctuated at different frequencies.

A study by Owyong at el in 2015 was conducted under title of "Cointegration and causality among the onshore and offshore markets for China's currency". Granger causality was tested using data from the onshore and offshore markets for the renminbi between 2010 and 2015. They examined how the causality associations were varied by the changes in policy and discuss the significance of their results. They have forecasted one-month-ahead using the estimated causality algorithm and showed that upon occasion they were consistent with the in estimating sample. The authors discovered a greater correlation between spot offshore and spot onshore rates than vice versa. The evidence of bi-directional nonlinear and linear causality was discovered, implying that foreign impulses influenced the domestic market. As a consequence of their sub-period study, they discovered that exchange rate fluctuations were most predictable during Sub-period 1, less predictable during Sub-period 2, and basically random during Sub-period 3, while the links between offshore and onshore markets were highest during Sub-period 3, (Owyong et al., 2015).

Esso (2010) used 1970–2007 data to examine the threshold cointegration relationship between economic growth and energy in seven Sub-Saharan countries. The threshold cointegration methods of Gregory and Hansen (1996 a, b) and the Granger-causality approach of Toda and Yamamoto (1995) were utilized. The African Development Bank's statistics and the World Bank's 2008 world development indices. In the presence of a structural break, they discovered a cointegration connection between economic development and energy consumption in South Africa, Cameroon, Cote d'Ivoire, Ghana, and Nigeria. Additionally, the computation of error-correction models and threshold cointegration tests indicated that economic growth had a positive long-run influence on energy consumption in both nations before to 1988, but this effect was reversed for Ghana and South Africa after the breakpoint, 1988. Furthermore, the Granger causality test indicated a bidirectional causal link between energy consumption and real GDP in Cote d'Ivoire, and a unidirectional causal relationship between real GDP and energy consumption in Ghana and Congo (Esso, 2010).

Bekiros and Diks (2008) used cointegration, nonlinear and linear causality methods to analyze the connection between crude futures prices and oil spot. The data was collected between 1991 and 1999 and 1999 to 2007. After adjusting for cointegration, they used a novel nonparametric test for nonlinear causation and the traditional linear Granger test. Additionally, causality was established after adjusting for the effects of other variables. They analyzed nonlinear causal relationships between VECM filtered residuals to ascertain when the measured causation was nonlinear. After correcting for conditional heteroscedasticity, a GARCH-BEKK model was used to test the hypothesis of nonlinear non-causality for the data. After filtering VECM cointegration, linear causal linkages disappeared. Even after GARCH filtering, non-linear causal connections were identified in specific circumstances over both time periods. Futures and spot prices may exhibit asymmetric GARCH effects and statistically significant higher order conditional moments as a result. They contended that when nonlinear factors are included, neither market consistently leads or lags the other, and that the pattern of leads and lags fluctuates over time (Diks, 2008).

Esso and Keho (2016) conducted another study in which they evaluated the long-run and causative relationships between energy consumption, CO2 emissions, and economic growth in 12 Sub-Saharan African nations. Annual data from 1971 to 2010 were subjected to the Bounds test for Granger causality and cointegration. In the Democratic Republic of Congo, Ghana, Benin, Nigeria, and Senegal, Granger causality studies found evidence of economic expansion affecting carbon emissions in the short term, suggesting that economic development cannot be achieved without influencing the environment. For Gabon, Nigeria, and Togo, reverse causation was shown between CO2 emissions and economic growth, implying that environmental initiatives aimed at reducing air pollution may have a detrimental effect on economic growth. Additionally, a bidirectional causal relationship between economic growth and carbon emissions was discovered in the short run for Nigeria and in the long run for Gabon and Congo. In the long run, economic expansion and energy consumption in Benin, Togo, Côte d'Ivoire, Senegal, Nigeria, and South Africa result in carbon emissions (Esso and Keho, 2016).

Shahbaz et al. (2012) examined the relationship between economic growth, carbon (CO2) emissions, energy consumption, and Pakistan's trade openness. The statistics were collected between 1971 and 2009. The empirical investigation used the Bounds test for cointegration and the Granger causality technique. Granger causality testing indicated a one-way causal relationship between income and carbon emissions. Energy usage increased both long-term and short-term CO2 emissions. Trade liberalization reduced carbon emissions in the long run but had a negligible effect in the short term. The presence of EKC demonstrated the country's efforts to reduce carbon emissions. This demonstrates the government's success in reversing environmental deterioration in Pakistan from the start of the NEP in 2005. On the other hand, conclusions based on aggregated data cannot adequately depict the trend of four provinces separately. The functioning of NEP alone was insufficient (Shahbaz et al., 2012).

Gebre-Mariam (2011) conducted a study to determine the unit root, causation, cointegration, and efficiency of the northwest US natural gas market. They examined the natural gas market's unit roots, causation, cointegration, and efficiency using the Northwest US natural gas markets as a case study. The research conducted a variety of statistical analysis on data from the spot and futures markets. The results revealed that after the first differencing, natural gas market prices remained stable and that spot and futures natural gas prices move in the same way. The influence of the spot price on futures prices was substantial for contracts with a maturity date of about six months (Mariam, 2011).

Alvarado et al. (2019) used causality and cointegration analysis to examine the causative link between non-sustainable energy, real per capita production and the growth rates of sustainable energy consumption in Latin America. Pedroni (1999) and Westerlund (2007) utilized cointegration methods to determine the link between the variables, whereas Dumitrescu and Hurlin (2012) used causality tests. We utilized a panel dynamic ordinary least squares model for grouped nations and a dynamic ordinary least squares model for individual countries to determine the intensity of the cointegration vector. They identified a short- and long-term equilibrium connection between renewable energy consumption growth rates, non-renewable energy consumption growth rates, and real per capita output growth. The vector of cointegration between renewable energy and production was more powerful in countries with a medium-low or medium-high income. The cointegration vector between production and non-renewable energy was higher in high-income countries. The causality test found a bidirectional relationship between renewable energy and real per capita productivity in low-middle income countries (Alvarado et al., 2019).

Bélad and Abderrahmani (2013) investigated the link between economic development and power consumption in Algeria using a multivariate causality analysis in the context of structural change. The objective of this study was to investigate and evaluate the causal link between power usage, Brent oil prices, and Algeria's GDP between 1971 and 2010. We used a multivariate cointegration approach based on recent advances in time series analysis to explore short-run, long-run, and joint causation relationships. There was a bidirectional relationship between energy use and GDP in both the long and short run. Additionally, the study discovered no clear correlation between electricity use and the price of Brent crude oil (Bélad and Abderrahmani, 2013).

Johansen's (1993) recursive cointegration technique was used to explore the consequences of the time varying nature of both the connection between five markets. They observed a substantial cointegration between the three metal prices, oil prices, and the US dollar exchange rate during the period following 1995. Additionally, the European sovereign debt crisis between 2010 and 2012 had an effect on the connection between these five markets. There was a positive long run correlation with gold price for the duration of the study period, however structural breakdowns in the two correlations occurred at various periods. After 2003, the commodities growth of ETFs erodes silver and copper prices' effect on gold prices. For the majority of the period following 2009, the link between silver and gold prices was negligible, while the association between copper and gold prices was negligible due to the influence of China's sharp decline in copper consumption following 2012. Additionally, the authors observed that the long-term link between oil and gold prices was negative before to 2003 and became murky during the Iraq War. The US dollar has traditionally decreased gold prices, and the correlation was especially significant following the January 1999 introduction of the Euro. They discovered that gold and oil prices do not have a long-run causal connection. While there was a long-run causal link between gold, silver, and copper prices, the long-run causal relationship between gold and copper prices was negligible during the 2008–2009 global financial crisis (Mei-Se et al, 2008).

3. MATERIALS AND METHODS

This chapter provides an overview of extensive methodologies that we will apply in our study. It highlights the basic definition of time series including trend, seasonality, autocorrelation and partial autocorrelation and decomposition. It provides the methodologies for AR, MA, ARMA, and ARIMA. An extensive background of unit root test has been discussed such as Dickey-Fuller, ADF, the Phillip Peron approach and KPPS test. We have provided an overview of vector autocorrelation model, cointegration analysis and ARDL method.

3.1. Time Series

A time series is a collection of observations made over time, in which the time dimension is critical for determining annual, monthly, weekly, and daily activity. Economic, social, environmental, commercial/business, medical, and biological areas all have a plethora of time series. There are two theoretically distinct methods for collecting time series data. The first method is to observe the data for a certain timestamp, which may occur infrequently or periodically. This may be a time series for a district. The second method is to constantly record data at the time interval. Numerous techniques are utilized to conduct time series analysis in order to make inferences about the data's characteristics (Falk, 2012).

The objective of analyzing time series data can be as follows:

1- To describe a data value to make time series, thus summary of data can be beneficial. Descriptions could be based on the idea of using visuals and they might be expressed through a mathematical model.
2- To understand that could be connected to the uncovering of structure – for example, identification and quantification of regular cyclic behavior.

3- To forecast future values as the basis for decisions and quantifying forecast errors.

4- Time series data can be categorized as either stationary or non-stationary.

Stationarity refers to the fact that the mean and variance remain constant across time. When the characteristics of a time series remain constant throughout time, even though the series' values change, the potential of predicting future behavior becomes possible. Stationarity may be described as the probabilistic model underlying the observed series non-stationarity: this refers to the fact that the mean and variance of time series data do not remain constant across time. This might be statistics relating to a trend or seasonality (Falk, 2012).

3.1.1. Trend

Trent is a general smooth tendency to rise or fall, ignoring short-term fluctuations. The increase and decrease movement can be described by trend of the time series which can be non-linear. The seasonal fluctuation could be year-to-year (Hyndman and Athanasopoulos, 2018).

3.1.2. Seasonality

Seasonality can be defined as a pattern of variation related with time of year, repeated from year to year. Seasonality pattern can be repeated daily, weekly, monthly or yearly. To identify seasonality in the time series data, a simple method can be used which is a graphical inspection using different scale. There are many ways to incorporate seasonality once it is detected in the model to have better forecasting value. This can be done by adding seasonality term in the model (Hyndman and Athanasopoulos, 2018).

3.1.3. Autocorrelation and partial autocorrelation

In time series, auto-correlation calculates the linear association among lagged values of a time series data. This means that time series is linearly associated with lagged a value which measures the degree of similarity among lags and time series. The reason that autocorrelation is important is that forecasting time series using some methods requires the assumption of no autocorrelation in the residuals (Hyndman and Athanasopoulos, 2018).

The mathematical equation of autocorrelation can be written as follow:

$$r_k = \frac{\sum_{t=k+1}^{T}(y_t-\bar{y})(y_{t-k}-\bar{y})}{\sum_{t=1}^{T}(y_t-\bar{y})^2} \tag{3.1}$$

T is the length of time series. A Correlogram which is known as ACF plot can be used to investigate the lags that have great correlation, detecting patterns, understanding properties and model time series by using this information. The ACF graph looks like the following.

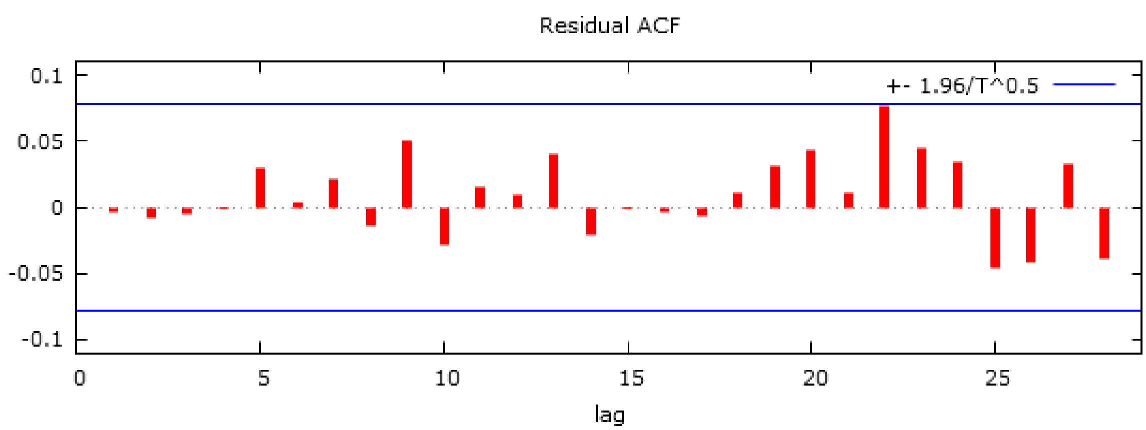

Figure 3.1. Residual of auto correlation function

In figure 3.1 the ACF graph can be used to assess the stationarity of time series data. it can reveal seasonality and trend patterns. Each bar in ACF plot represents direction and size of the correlation (Hyndman and Athanasopoulos, 2018).

On the other hand, the correlation among lagged versions of itself and a time series is measured by the partial correlation function (PACF). The PACF is used to determine the correlation between data points at time t and data at time t2. Additionally, graphic approaches can be used to analyze the PACF. The graph can be as follow:

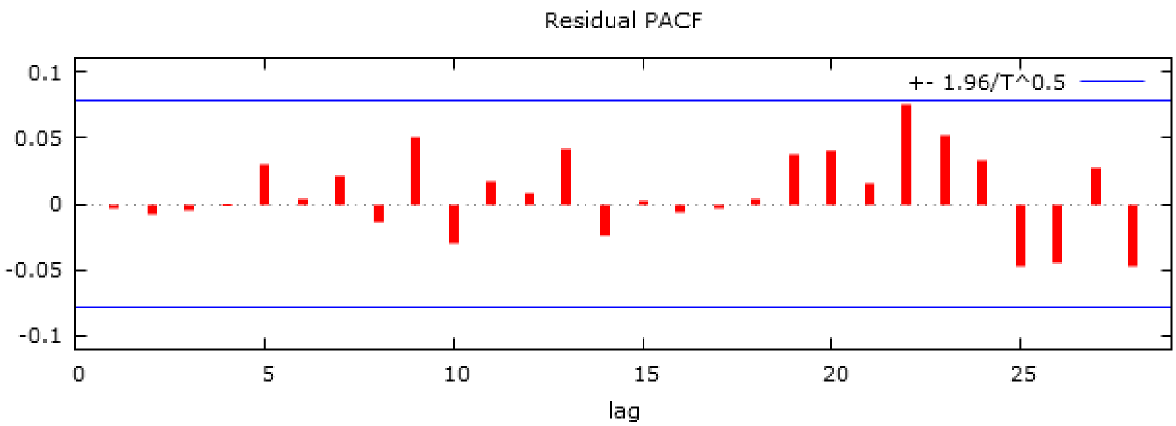

Figure 3.2. Residual of partial auto correlation function.

In figure 3.2 the PACF plots partial autocorrelation coefficients at lag h versus time to examine whether the AR model is suitable for the data and to identifying orders in AR model. The most used approach to calculate partial autocorrelation coefficients is by using the equation the following equation.

$$\frac{cov(y_i, y_{i-h} | y_{i-1}, \ldots, y_{i-h+1})}{\sqrt{var(y_i | y_{i-1}, \ldots, y_{i-h+1}).var(y_{i-h} | y_{i-1}, \ldots, y_{i-h+1})}} \tag{3.2}$$

The above equation is for the h^{th} order partial autocorrelation Hamilton, (1994). The PACF is usually used to identify the order of an autoregressive integrated moving average (ARIMA) model (Hyndman and Athanasopoulos, 2018).

3.1.4. Decomposition

Decomposition means estimation of the components of I_t, S_t , $C_t$and T_t where I_t, S_t , $C_t$and T_t are noise, seasonal, cycle and trend components respectively. Decomposition can be useful due to the components that may be of interest. It can be used to examine seasonal variation. Second, it becomes seasonal adjustment, which is the process of subtracting predicted seasonal influences from the original series in order to expose any underlying changes more clearly. Monthly economic and unemployment statistics are typically de-seasonalized in this way to facilitate month-to-month comparisons. Finally, if the model performs poorly as a result of the trend or seasonal component, we must address the issue in the specific component, not the complete model, which is generally more sophisticated in comparison to the component (Hyndman, R.J. and Athanasopoulos, G, 2018).

3.1.4.1. Additive decomposition

If additive decomposition was assumed, the model can be as follow:

$$\mathcal{Y}_t = T_t + C_t + S_t + I_t$$

This is addictive model which is more suitable when the seasonal fluctuation does not differ from time series level (Hyndman and Athanasopoulos, 2018).

3.1.4.2. Multiplicative decomposition

Multiplicative decomposition is widely used in case of having economic time series data. It is suitable in case of existence of variation around trend or seasonal as a proportional to the time series level. The model can be as follow:

$$\mathcal{Y}_t = T_t * C_t * S_t * I_t$$

The alternative approach of addictive and multiplicative decomposition can be used which is transforming data (Hyndman and Athanasopoulos, 2018).

3.2. ARMA Model

ARMA is a forecasting model in which both auto-regression (AR) analysis and moving average (MA) are used to well-behaved time series data. In ARMA, it is assumed that the time series is stationary and that when it varies, it does so evenly around a particular instant of time (Gordon et al., 2020).

Auto regressive (AR) model uses the current value is influenced by previous values to forecast next step ahead.

$$\boldsymbol{z_t = \emptyset_1 z_{t-1} + \emptyset_2 z_{t-2} + \emptyset_3 z_{t-3} + \cdots + \emptyset_p z_{t-p} + u_t} \quad (3.3)$$

$$\boldsymbol{U_t \sim N\ (0, \sigma^2)}$$

A simple way to provide AR order p and more complex models is through the backward shift operator B. this can be defined as a functional who allows us to express long formulae in compact form. In the context of a stochastic process, we define B so that.

$$\boldsymbol{z_t = \emptyset_1 z_{t-1} + \emptyset_2 z_{t-2} + \emptyset_3 z_{t-3} + \cdots + \emptyset_p z_{t-p} + u_t}$$

$$\boldsymbol{z_t - \emptyset_1 z_{t-1} - \emptyset_2 z_{t-2} - \emptyset_3 z_{t-3} - \cdots - \emptyset_p z_{t-p} = u_t}$$

$$\boldsymbol{z_t\left(1 - \emptyset_1 \beta - \emptyset_2 \beta^2 - \emptyset_3 \beta^3 - \cdots - \emptyset_p \beta^p\right) = u_t}$$

$$\text{Let}\boldsymbol{\left(1 - \emptyset_1 \beta - \emptyset_2 \beta^2 - \emptyset_3 \beta^3 - \cdots - \emptyset_p \beta^p\right) = \ \emptyset\beta}$$

$$\boldsymbol{\emptyset(\beta) z_t = u_t} \quad (3.4)$$

Where $\emptyset(B) = (1 - \alpha B)$; here $\emptyset$ (B) is a polynomial of degree one of B.

3.2.1. Moving average model (MA)

MA is a method used to model the time series data with trends. This is very useful to forecast trends especially when the trend is long term. MA model is using errors from past forecast in a regression model. The MA model of order q can be defined as follow:

$$\boldsymbol{z_t = e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \theta_3 e_{t-3} - \cdots - \theta_q e_{t-q}}$$

MA(1)

$$\boldsymbol{z_t = e_t - \theta_1 e_{t-1}} \quad (3.5)$$

$$\boldsymbol{e_t} \sim \boldsymbol{N(0, \sigma^2)}$$

Where $\boldsymbol{e_t}$is the white noise? $\boldsymbol{z_t}$ is a weighted moving average of the past forecast errors (Hyndman and Athanasopoulos, 2018).

3.2.2. ARMA

ARMA is an autoregressive moving average model in which the value is stated as a linear mix of white noise and historical data. The ARMA model of order p and q consists of AR (p) and MA(q) that can be as follow:

$$\boldsymbol{z_t} = \emptyset_1 \boldsymbol{z_{t-1}} + \emptyset_2 \boldsymbol{z_{t-2}} + \cdots + \emptyset_p \boldsymbol{z_{t-p}} + \boldsymbol{e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q}} \quad (3.6)$$

The parameters of the above equations $\emptyset_i$ can be estimated using maximum likelihood methods and mean least squares methods. The ARMA model is very vital in the time series analysis description and widely used in regard with AR models (Gregory et al., 2015). The backshift operator can be written in the model as follow:

$$\boldsymbol{z_t} - \emptyset_1 \boldsymbol{z_{t-1}} - \emptyset_2 \boldsymbol{z_{t-2}} - \cdots - \emptyset_p \boldsymbol{z_{t-p}} = \boldsymbol{e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q}}$$

$$\boldsymbol{z_t(1-\emptyset_1\beta-\emptyset_2\beta^2-\cdots-\emptyset_p\beta^p)=e_t(1-\theta_1\beta-\theta_2\beta^2-\cdots-\theta_p\beta^p)}$$

$$\boldsymbol{\emptyset\beta=(1-\emptyset_1\beta-\emptyset_2\beta^2-\cdots-\emptyset_p\beta^p)}$$

$$\boldsymbol{\theta\beta=(1-\theta_1\beta-\theta_2\beta^2-\cdots-\theta_p\beta^p)}$$

$$\boldsymbol{z_t\emptyset\beta=e_t\theta\beta} \tag{3.7}$$

3.2.3. ARIMA

Autoregressive integrated moving average (ARIMA) is indicated by Box and Jenkins as an autoregressive integrated moving average model (1976). The constant mean and variance cannot be always obtained in the time series data. This leads to non-stationarity time series data. ARIMA model is widely used to forecast future values. The model allows a non-zero auto-correlation function. When they are being utilized for stationary time series data, the ARIMA models must be employed. Differencing method should be utilized if the data isn't stationary (Hyndman and Athanasopoulos, 2018). The ARIMA model can be written as follows:

$$\boldsymbol{\emptyset\beta z_t=\theta\beta e_t}$$

$$\boldsymbol{\emptyset\beta w_t=\theta\beta e_t}$$

$$\boldsymbol{w_t=\nabla^d z_t}$$

$$\boldsymbol{w_t=\emptyset_1 w_{t-1}+\emptyset_2 w_{t-2}+\cdots+\emptyset_p w_{t-p}+\ e_t-\theta_1 e_{t-1}-\theta_2 e_{t-2}-\cdots-\theta_q e_{t-q}}$$

$$\boldsymbol{w_t=\nabla^d z_t=\ z_t-z_{t-1}}$$

$$\boldsymbol{z_t-z_{t-1}=\emptyset_1(z_{t-1}-z_{t-2})+\emptyset_2(z_{t-2}-z_{t-3})+\cdots+\emptyset_p(z_{t-p}-z_{t-p})+\ e_t-\theta_1 e_{t-1} -\theta_2 e_{t-2}-\cdots-\theta_q e_{t-q}}$$

$$\boldsymbol{z_t-z_{t-1}=\emptyset_1 z_{t-1}-\emptyset_1 z_{t-2}+\emptyset_2 z_{t-2}-\emptyset_2 z_{t-3}+\cdots+\emptyset_p z_{t-p}-\emptyset_p z_{t-p})+\ e_t -\theta_1 e_{t-1}-\theta_2 e_{t-2}-\cdots-\theta_q e_{t-q}}$$

$$z_t = (1 + \emptyset_1)z_{t-1} + (\emptyset_2 - \emptyset_1)z_{t-2} + (\emptyset_3 - \emptyset_2)z_{t-3} + \cdots + \left(\emptyset_{p-1} - \emptyset_{p-2}\right)z_{t-p-1} + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q} \quad (3.8)$$

3.3. Unit Root

A unit root is defined as a sequence which includes one or more characteristic roots. In any time series, a unit root can be referred as a stochastic trend. For example, in case of having AR (1) as follow:

$$z_t = \emptyset_1 z_{t-1} + u_t \quad (3.9)$$

u_t is a white noise, If the $\emptyset_1 = 1$, then we can say that z_t has a unit root. The process of unit root is not stationary in which it known as random walk model. The time series data can be stationary using differencing. This means the process has unit root which the recent value of z_t equals to last value of $z_t - 1$ in addition to error terms. It is very important for the time series to be stationary, because correlation can make the series non-stationary even in case of having large sample size. Thus, we can differencing the time series data to achieve stationarity (Wei, 2006)

Different methods are available in order to test unit root including Dickey-fuller, Augment Dickey-fuller, the Phillip Peron and Durbin-Watson (DW). These approaches are most commonly used among others.

3.3.1. Dickey Fuller

In the test of Dickey Fuller, the null hypothesis of AR model has a unit root is tested. The parameter estimation of the test can be obtained using OLS and t-test. The test is developed by Dickey and Fuller in 1979 as they examined the existence of unit root in the AR in the first order. Three versions of tests of AR processes were found as follows:

1. Unit root test

$$\Delta z_t = \emptyset z_{t-1} + u_t \tag{3.11}$$

2. Unit root test with constant which is called drift:

$$\Delta z_t = a_0 + \emptyset z_{t-1} + u_t \tag{3.12}$$

3. Unit root test with drift and deterministic time trend:

$$\Delta z_t = a_0 + a_1 t + \emptyset z_{t-1} + u_t \tag{3.13}$$

The assumptions are $z_t = 0$ and u_t is independent identically distributed, i.i.d (0, σ^2). The test hypothesis can be written as follows:

H_0: $\emptyset = 1$

H_1: $|\emptyset| < 1$

When $\alpha = 1$, it means that there is unit root. The time series is stationary with zero mean in case p ≥ 0. The assumption of using Dickey-fuller is that the error term U_t is uncorrelated. When U_t is correlated, he Augmented Dickey-Fuller (ADF) can be used. The limitation of Dickey fuller is that it does not take auto correlation into consideration in the process of error term U_t. In order to tackle this issue, the ADF can be applied by utilizing the difference lagged dependent covariate as independent covariates (Fuller, 1979).

3.3.2. Augment Dickey Fuller (ADF)

The ADF is an extended augmented version of Dickey Fuller test in which serial correlation is accommodated. It is also set for more and large complex models. Unlike DF test the ADF test can be used for the higher order of AR model. The test can be used to test the stationarity of time series. The procedure of the test is similar to the Dickey Fuller except AR (p) can be considered here as follows:

The equation of the ADF test can be as follow to test the unit root:

$$\Delta \boldsymbol{z} = \boldsymbol{\beta_0} + \emptyset^* \boldsymbol{z_{t-1}} + \sum_{\boldsymbol{i=1}}^{\boldsymbol{p-1}} \emptyset_{\boldsymbol{i}} \boldsymbol{z_{t-1}} + \boldsymbol{u_t} \tag{3.14}$$

Where $\boldsymbol{z_t}$ is the value of a variable at time period t, $\Delta \boldsymbol{z_t} = \boldsymbol{z_t} - \boldsymbol{z_{t-1}}$, $\boldsymbol{\beta_0}$ denotes a constant term; t is a linear time trend and $\boldsymbol{u_t}$ is an error term. The null hypothesis of the ADF test is that the parameters are equal to zero in which the time series data is not stationary. The alternative hypothesis which is the parameters are not zero.

The DF test can be compared with appropriate critical value. Because the test is asymmetrical, negative values are our concern. If the computed test statistic is more negative (lees) than critical value, we can reject null hypothesis and the time series data has no unit root (Said and Dickey, 1984).

3.3.3. The Phillip Peron

The test of many unit roots is developed by Phillips and Perron (1988). Nowadays, the test is widely used in the time series analysis especially in case of having financial data. The PP test differs from the ADF test especially the way the PP test deals with heteroskedasticity and serial correlation in the errors. Particularly, a parametric auto-regression can be used to approximate ARMA structure of the errors. In the test regression, a serial correlation can be ignored in the PP test. The test can be as follow:

$$\Delta \boldsymbol{Z_t} = \boldsymbol{\beta' D_t} + \boldsymbol{\pi z_{t-1}} + \boldsymbol{u_t} \tag{3.15}$$

Here $\boldsymbol{u_t}$ is I (0) and can be heteroskedastic. The heteroskedasticity and serial correlation in errors of $\boldsymbol{u_t}$ can be corrected by the PP test. The test statistics can be modified to $\boldsymbol{t_\pi = 0}$ and T$\boldsymbol{\hat{\pi}}$. The modification can be given by $\boldsymbol{Y_t}$ and $\boldsymbol{Y_\pi}$, are given by

$$\boldsymbol{Y_t} = \left(\frac{\boldsymbol{\hat{\sigma}^2}}{\boldsymbol{\hat{\lambda}^2}}\right)^{1/2} . \boldsymbol{t_\pi} = \boldsymbol{0} - \frac{\boldsymbol{1}}{\boldsymbol{2}}\left(\frac{\boldsymbol{\hat{\lambda}^2 - \hat{\sigma}^2}}{\boldsymbol{\hat{\lambda}^2}}\right) . \left(\frac{\boldsymbol{T.SE(\hat{\pi})}}{\boldsymbol{\hat{\sigma}^2}}\right)$$

$$\boldsymbol{Y_\pi} = \boldsymbol{T\hat{\pi}} - \frac{\boldsymbol{1}}{\boldsymbol{2}} \frac{\boldsymbol{T^2.SE(\hat{\pi})}}{\boldsymbol{\hat{\sigma}^2}} \left(\boldsymbol{\hat{\lambda}^2 - \hat{\sigma}^2}\right) \tag{3.16}$$

Here the variance of parameters of $\hat{\boldsymbol{\sigma}}^{2}$ and $\hat{\boldsymbol{\lambda}}^{2}$ are consistent

$$\boldsymbol{\sigma}^{2} = \mathbf{lim}_{T\rightarrow\infty}\, \boldsymbol{T}^{-1}\sum_{t=1}^{T}\boldsymbol{E}[\boldsymbol{u}_{t}^{2}] \quad (3.17)$$

$$\boldsymbol{\lambda}^{2} = \mathbf{lim}_{T\rightarrow\infty}\sum_{t=1}^{T}\boldsymbol{E}[\boldsymbol{T}^{-1}\boldsymbol{S}_{t}^{2}] \quad (3.18)$$

Again $\boldsymbol{S}_{t}\sum_{t=1}^{T}\boldsymbol{u}_{t}$. We have consistent least square residuals $\hat{\boldsymbol{u}}_{t}$ of sample variance. Additionally, the Newey-West long-run estimation of variance of $\boldsymbol{u}_{t}$ applying $\hat{\boldsymbol{u}}_{t}$ is a consistent estimate of $\boldsymbol{\lambda}^{2}$. The same asymptotic distribution as the ADF can be obtained for the PP $\boldsymbol{Y}_{t}$ and $\boldsymbol{Y}_{\pi}$ under the null hypothesis which is $\boldsymbol{\pi} = \mathbf{0}$. One of the advantages of the PP test is that it can be robust to the general form of heteroskedasticity in the error term $\boldsymbol{u}_{t}$ unlike the ADF test. Also, it is not necessary to specify lag length for the test (Phillips and Perron, 1988).

3.3.4. KPSS test

The unit root tests of PP and ADF are for the $\boldsymbol{H}_{0}$ which a time series $\boldsymbol{z}_{t}$ is I(1). On the other hand, stationarity tests are for the $\boldsymbol{H}_{0}$ which $\boldsymbol{z}_{t}$ is I (0). Thus, KPSS is the most widely used test which is proposed by Kwiatkowski, Phillips, Schmidt and Shin (1992). They started to test the following mode:

$$\boldsymbol{z}_{t} = \emptyset'\boldsymbol{D}_{t} + \boldsymbol{\mu}_{t} + \boldsymbol{u}_{t} \quad (3.19)$$

$$\boldsymbol{\mu}_{t} = \boldsymbol{\mu}_{t-1} + \boldsymbol{\varepsilon}_{t},\ \boldsymbol{\varepsilon}_{t} \sim \boldsymbol{WN}(\mathbf{0}, \boldsymbol{\sigma}_{\varepsilon}^{2})$$

The deterministic components are included in $\boldsymbol{D}_{t}$ (constant with time trend or constant), $\boldsymbol{u}_{t}$ is I (0) and heteroskedastic may be present. The pure random walk is donated by $\boldsymbol{\mu}_{t}$ with $\boldsymbol{\sigma}_{\varepsilon}^{2}$ which is innovation variance of that random walk. The hypothesis is as follows:

$\boldsymbol{H_0}: \boldsymbol{\sigma_\varepsilon^2 = 0}$

$\boldsymbol{H_1}: \boldsymbol{\sigma_\varepsilon^2 \neq 0}$

The null hypothesis displays a constant $\boldsymbol{\mu_t}$. The $\boldsymbol{H_0}$ also displays a unit MA root in the ARMA characterization of $\Delta\boldsymbol{z_t}$.the score statistic to test $\boldsymbol{\sigma_\varepsilon^2 = 0}$ against the alternative which $\boldsymbol{\sigma_\varepsilon^2 > 0}$ can be given as follow:

$$\boldsymbol{KPSS = \left(T^{-1} \sum_{t=1}^{T} \widehat{S}_t^2\right) / \hat{\lambda}^2} \tag{3.20}$$

Here, $\boldsymbol{\sum_{j=1}^{t} \widehat{u}_j, \widehat{u}_t = \widehat{S}_t}$ which represent the regression residuals of $\boldsymbol{z_t}$ on $\boldsymbol{D_t}$ and $\boldsymbol{\hat{\lambda}^2}$ represents a consistent estimate of the long-run variance of $\boldsymbol{u_t}$ utilizing $\boldsymbol{\widehat{u}_t}$. The authors showed that $\boldsymbol{z_t}$ is I (0) under $\boldsymbol{H_0}$ any they illustrated that the convergence of KPSS can be obtained to standard Brownian motion function which relays on the form of the deterministic terms $\boldsymbol{D_t}$ but not the values of parameters of $\boldsymbol{\beta}$. If $\boldsymbol{D_t = 1}$ then

$$\boldsymbol{KPSS \xrightarrow{d} \int_0^1 V_1 (r) dr} \tag{3.21}$$

Where $\boldsymbol{V_1(r) = W(r) - rW(1)}$ and $\boldsymbol{W(r)}$ is a standard Brownian motion for $\boldsymbol{r} \in [\boldsymbol{0, 1}]$ if $\boldsymbol{D_t = (1, t)'}$ then

$$\boldsymbol{KPSS \xrightarrow{d} \int_0^1 V_2 (r) dr} \tag{3.22}$$

Where $\boldsymbol{V_2(r) = W(r) - r(2 - 3r)W(1) + 6r(r^2 - 1) \int_0^1 W_s \, ds}$.

We shall obtain critical values from the asymptotic distributions (3.21) and (3.22) using simulation approaches. The test for stationarity is a one-sided right-tailed test. The null of stationarity can be rejected at the 100 · α% level if the KPSS test statistic (3.20) is more than the 100, (1 −α) % quantile from the appropriate asymptotic distribution (3.21) or (3.22) (Kwiatkowski and Shin, 1992).

3.4. VAR Model

Sims proposed a Vector auto-regressive (VAR) models which can eb used to capture the interdependency and dynamics of multivariate time series. AR models are only considering one time series $\boldsymbol{Z_t}$, while VAR models multiple time series. It is known as a generalization of univariate AR models or a mix among univariate time series models and parallel equations models. The main concenters are on the parameter estimation and model specification. Let is assume a k-dimensional multivariate time series $\boldsymbol{Z_1} \ldots \boldsymbol{Z_t}$ with $\boldsymbol{Z_t}$ = ($\boldsymbol{Z_{1t}}$, …, $\boldsymbol{Z_{kt}}$). This can be generated by VAR model. The basic Var model with p lags can be given by:

$$\boldsymbol{Z_t = v + \emptyset_1 Z_{t-1} + \emptyset_2 Z_{t-2} + \cdots + \emptyset_p Z_{t-p} + \mathrm{u_t}} \qquad (3.38)$$
$$\boldsymbol{t = 1, 2, 3, \ldots, T}$$

Here $\boldsymbol{Z_t}$= ($\boldsymbol{Z_{1t}}$, $\boldsymbol{Z_{2t}}$, … , $\boldsymbol{Z_{kt}}$) ′ which is donated by a (k × 1) vector of time series covariates. v is a (k × 1) vector of intercepts. $\boldsymbol{\emptyset_m}$ Are (k × k) coefficient matrices, and $\boldsymbol{u_t}$ is a (k × 1) error term with zero-mean white noise vector process.

The evaluation of k sets of variables over time period can be described by a VAR model. Each time period can be numbered as t = 1, ..., T. three assumptions should be made by the error terms as follows:

Errors have zero mean, $\boldsymbol{E(u_t) = 0}$

1-$\boldsymbol{E(u_t u_t') = \Omega}$. The error terms is a k × k positive-semi definite matrix denoted Ω

2-$\mathbf{E}(\boldsymbol{u_t u_{t-k}'}) = \mathbf{0}$. For any non-zero k. There is no association between times.

Selecting maximum lag in the model needs special caution. This is because inference is dependent on the accuracy of chosen lag order (Brooks, 2002). There are three assumptions with the VAR model. The first assumption is that the intervals among measurements have same length. The second assumption is that the observations are normally distributed with mean vector and covariance matrix **Σ**. We can tackle the issue of violating second assumption using mixture distribution. It is assumed in VAR that the time series is stationary (Lütkepohl, 2005).

3.5. Co-integration Analysis

Cointegration analysis concentrates on testing whether long run relationship is stationary between several time series. This long run can be measured and tested utilizing the idea of cointegration. Cointegration can occur in case several time series datasets have long run equilibrium, move together and sharing common stochastic trend. To test cointegration, it can be considered whether there are multiple cointegrating vectors, whether the cointegration vector is known and are there any structural breaks in the cointegration relationships. The cointegration test can be used to identify the degree of sensitivity of several datasets in a period of time (Erica, 2020).

According Marseet and Khadija (2015), several approaches can be used to test long run cointegration between two or more set of data. Long run association between variables can be assessed using cointegration analysis. Cointegration between variables indicates that a long run or equilibrium association exists between variables. If a time series is not stationary, but they may more together through time and their difference can be stationary. The long run linkage between variables means the system coverages through time and we can interpret the error term as the disequilibrium over time. A great method has been emerged by cointegration to examine multivariate time series trends and providing a technique for long run and short run relationship between variables (Marseet and Khadija, 2015).

Many approaches are available to investigate the fact that integrated variables of order one may have a cointegration. When each covariate in the set of variables is integrated of the equal order, and when at least one linear combination exists among those variables, this can be called as stationary. Testing for cointegration means that we are testing the presence of long run relationship. These methods need to some order of integration for investigating long-run connection between the two sets of data (Marseet and Khadija, 2015).

3.5.1. Johansen (VECM)

Johansen test based on VECM consists of two tests which are the maximal Eigen value and the trace test. These tests can be used for long run relationship based on the VECM. VECN models can be estimated using different assumption such as the model with or without constant or trend and with different number of cointegration vectors. Then models can be compared using likelihood ratio test. The trace test can be obtained using likelihood ration test of a restricted VECM versus unrestricted VECM with $\boldsymbol{k}$ vectors.

The basic VECM cab is as follow for $\boldsymbol{k}$ lags:

$$\Delta \boldsymbol{y_t} = \boldsymbol{\Pi y_{t-k}} + \boldsymbol{\Gamma_1} \Delta \boldsymbol{y_{t-1}} + \boldsymbol{\Gamma_2} \Delta \boldsymbol{y_{t-2}} + \cdots + \boldsymbol{\Gamma_{k-1}} \Delta \boldsymbol{y_{t-(k-1)}} + \boldsymbol{u_t} \tag{3.43}$$

Where $\boldsymbol{\Pi} = \sum_{i=1}^{k} \boldsymbol{\beta_i}) - \boldsymbol{I_g}$

And $\boldsymbol{\Gamma_i} = \left(\sum_{j=1}^{i} \boldsymbol{\beta_i}\right) - \boldsymbol{I_g}$

The $\boldsymbol{\Gamma}$ is the parameters which can be referred as short run matrices. Nevertheless, the concentrate here on the Johansen's cointegration $\boldsymbol{\Pi}$ which referred as long run parameter matrix. The rank of $\boldsymbol{\Pi}$ can be investigated to define the cointegration between covariates using $\boldsymbol{\Pi}$ matrix (Franses, 1998). The formula for the test statistics can be as follow:

$$\boldsymbol{\lambda_{trace}(r)} = -\boldsymbol{T} \sum_{i=r+1}^{g} \ln\left(\mathbf{1} - \boldsymbol{\hat{\lambda}_i}\right) \tag{3.44}$$

$$\boldsymbol{\lambda_{max}} = -\boldsymbol{T} \ln\left(\mathbf{1} - \boldsymbol{\hat{\lambda}_i} + \mathbf{1}\right) \tag{3.45}$$

It can be noticed that $\boldsymbol{\hat{\lambda}}$ is the obtained values for the ordered Eigen values from the Π matrix and $\boldsymbol{r}$ cointegrating vectors with $\boldsymbol{H_0}$. Also, T denotes the total number of observations. The null hypothesis of the trace test is that there are $\boldsymbol{r}$ cointegration vectors, whereas the alternative hypothesis is that there are more than $\boldsymbol{r}$ cointegration vectors. The null hypothesis for the maximum test is that there are $\boldsymbol{r}$ cointegration vectors, as opposed to the alternative hypothesis that there are more than $\boldsymbol{r+1}$ cointegration vectors.

VECM can be used for investigating the short-term association among variables. It has some advantages over VAR. this is because VAR has the issue of possibility of misspecification in the first differences (Franses, 1998).

3.5.2. Autoregressive distributed lag (ARDL)

Testing cointegration using Johansen procedure needs that all integrated variables to be in the same order. These may not occur every time. Thus, an alternative approach has been developed by Pesaran et al in 2001. This is known based testing that is auto-regressive distributed lag model. The assumption that variables should be integrated in the same sequence is not required. Unlike Johnsen, the method is more suitable in case of small sample. Implementing and interpreting are quite straightforward because it has one equation.

The method of ARDL has many benefits in comparison with historical cointegration test as follows:

- It is resilient and analyzing I (0) and/or I(1) is allowed.
- Interpretation and computing ARDL method are easy because of one equation.
- Various number of lag lengths can be used in the model with various covariates.
- It is appropriate in case of having small sample.
- It is possible to acquire an unbiased estimate of the long term connection.
- The issues of endogeneity and autocorrelation can be addressed (Jalil and Ma, 2008).

The ARDL (p, q) model as:

$$\mathbf{y_t} = \mathbf{a_0} + \mathbf{a_1} + \sum_{\mathbf{i=1}}^{\mathbf{p}} \emptyset_{\mathbf{i}} \mathbf{y_{t-i}} + \boldsymbol{\beta}' \mathbf{x_t} + \sum_{\mathbf{i=0}}^{\mathbf{q-1}} \boldsymbol{\beta}_{\mathbf{i}}^{*\prime} \Delta \mathbf{x_{t-i}} + \boldsymbol{u}_{\mathbf{t}} \quad (3.46)$$

$$\Delta \mathbf{x_t} = \mathbf{P_1} \Delta \mathbf{x_{t-1}} + \mathbf{P_2} \Delta \mathbf{x_{t-2}} + \cdots + \mathbf{P_s} \Delta \mathbf{x_{t-s}} + \boldsymbol{u}_\mathbf{t} \tag{3.47}$$

When $\boldsymbol{x_t}$ contaions k-dimensional I (1) non-cointegrated covariates, $\boldsymbol{P_i}$ is the parameter matrix that generates a stable autoregressive process, and $\boldsymbol{u_t}$ represents white noise error term (Peseran et al., 2001). The bound testing technique may be used to conduct a cointegration test for the ARDL approach. Additionally, the test has two sets of asymptotic values in which all variables are assumed to be I (1) in one set and I(0) in the other. For both the I (1) and I(0) data sets, the two sets can provide critical value limits for cointegration. Three steps are needed to apply the ARDL procedure:

1- Bounds testing procedure should be applied to detect cointegration ranks among covariates.

2- Long run relationship parameter can be estimated with respect to cointegration relations estimated in first step.

3- Short run dynamic parameters can be estimated through vector error (Peseran et al., 2001).

The ARDL cointegration approach is more appropriate when covariates are integrated in the order I (0), I (1), or a mixture of both. If a single long run link exists between variables with a small sample size, the test is robust. There are some requirements for the application of ARDL as follows:

First, the ARDL technique may be utilized regardless of whether the variables are I (0), I (1), or a combination of both. This enables us avoiding pre-testing issues related with cointegration analysis that classifying variables into I (0), I (1) is required. This means that we do not need to pre-test variables is not needed for the bound cointegration testing technique.

Second, when the test statistics indicate the presence of a single long run connection and the sample size is small or finite, the ARDL's error correction representation can be more efficient.

Third, when the test statistics reveals the existence of a multiple long run relationship, we cannot use the ARDL method. Thus, we can apply an alternative method such as Johansen and Juselius (1990). This means that if there are different single equation of variables as dependent variable, a multivariate techniques can be used.

Fourth, If the maximum eigenvalue, the F-statistics, or the generates a single long-run connection, the ARDL technique may be employed instead of the Johansen and Juselius approach (Peseran et al., 2001).

3.6. Cointegration with Structural Break

The power of cointegration can be affected by presence of structural breaks. Spurious unit root behavior can be introduced by a structural break. Thus, it is difficult to reject null hypothesis of no cointegration. Neglecting the problem of structural break leads to inaccurate statistical results of cointegration analysis. It is important to make suitable adjustment when the structural breaks are known. To analysis the data, it can be assumed that the break date in unknown. This is because tests can be applied to find the most appropriate structural breaks in the data set. There are many approaches available for cointegration analysis with structural breaks. However, here the methods of Gregory and Hansen (1996) and Hatemi-J (2008) can be discussed and applied in study our study.

3.6.1. Gregory and Hansen (1996)

Gregory and Hansen (1996) developed a method as an extension of residual-based cointegration testing that takes into account the potential of unforeseen regime changes in either the coefficient or the intercept vector. The null hypothesis is that no cointegration occurs when a structural break occurs, whereas the alternative hypothesis is that cointegration occurs when a structural break occurs. Using ADF test for cointegration analysis without time regime shift could lead to false conclusion that the long run relationship is not exist between dependent covariate and its determinants. The probability of rejecting the null hypothesis is higher with the Gregory and Hansen test. Different assumptions about structural break are mentioned in the proposed method by Gregory and Hansen as follows: The first assumption is the level shift:

$$\boldsymbol{y_t + \alpha_1 + \alpha_2 D_t(\tau) + \delta_1 X_t + e_t} \tag{3.48}$$

The second assumption is the level shift with trend:

$$\boldsymbol{y_t + \alpha_1 + \alpha_2 D_t(\tau) + \gamma_t + \delta_1 X_t + e_t} \tag{3.49}$$

The third assumption is the regime shift:

$$\boldsymbol{y_t + \alpha_1 + \alpha_2 D_t(\tau) + \delta_1 X_t + \delta_2 X_t D_t(\tau) + e_t} \tag{3.50}$$

The categorical variable $\boldsymbol{DU_{tb}}$ in which takes 0 when $\boldsymbol{t \leq b}$ and one when $\boldsymbol{t > b}$. Here, $\boldsymbol{t}$ is time trend and $\boldsymbol{b}$ is the time that the structural break happens. It has been shown that the length lag, $\boldsymbol{k}$ can be selected based on a test statistic following the same procedure of (Perron and Vogelsang, 1992).

The GH test assumes unknown structural break time; thus the structural break is endogenously determined. Three test statistics: $\mathbf{ADF^* = inf_{\tau \in T}\, ADF}$, $\mathbf{Z_t^* = inf_{\tau \in T}\, Z_t}$, $\mathbf{Z_a^* = inf_{\tau \in T}\, Z_a}$ are proposed, that is the modified version of the Engel and Granger (1987) cointegration test. $\mathbf{Z_t^*} = \mathbf{\inf\limits_{\tau \in T}\ Z_t\ (t)}$ and $\mathbf{Z_a^* = inf_{\tau \in T}\, Z_a\ (t)}$ which can be considered as a modification version of Phillips and Quliaris (1990). The lowest value of the three test statistics can be considered as the break point. The GH utilizes a modified Mackinnon (1991) critical value for the cointegration test. These critical values vary from the critical values utilized by the Engle and Granger method (Gregory and Hansen, 1996).

3.6.2. Hatemi-J (2008)

Three residual based tests are extended by (Hatemi-J, 2008) for cointegration which takes to consideration two shifts. There is unknown timing of each shift, and it is determined. There are no standard distributions of the tests. Critical values can be calculated using simulation approach. Hatemi-J extended the tests for cointegration by taking the possibility of two structural shifts into consideration. To consider the impact of two structural breaks on both the slop and intercept, the author generalized following equations:

$$y_t = \alpha_0 + \alpha_1 D_{1t} + \alpha_2 D_{2t} + \beta'_0 x_t + \beta'_1 D_{1t} x_t + \beta'_2 D_{2t} x_t + u_t \tag{3.51}$$

Where D_{1t} *and* D_{2t} are dummy variables defined as

$$D_{1t} = \begin{cases} 0 & if\ t \leq [n\tau_1] \\ 1 & if\ t > [n\tau_1] \end{cases}$$

And

$$D_{2t} = \begin{cases} 0 & if\ t \leq [n\tau_2] \\ 1 & if\ t > [n\tau_2] \end{cases}$$

τ_1 and $n\tau_2$ ∈ (0, 1) are unknown parameters. This signifying integer part and relative timing of the break change point (Hatemi, 2008).

ADF test can be computed to test null hypothesis by corresponding t-test for the slop of $\hat{u}_t$ on $\hat{u}_{t-k}$. The null hypothesis is that there is no cointegration. Here $\hat{u}_t$ is a calculated error term from the above equation. The Z_t and Z_a tests can be calculated using the bias-corrected first-order serial correlation coefficient estimate $\hat{p}^*$, as follow:

$$\hat{p}^* = \frac{\sum_{t=1}^{n-1}(\hat{u}_t \hat{u}_{t-1} - \sum_{j=1}^{B} w\left(\frac{j}{B}\right)\hat{\gamma}(j))}{\sum_{t=1}^{n-1} \hat{u}_t^2} \tag{3.52}$$

The function, $w(.)$, provides kernel weights to meet standard conditions for a density estimator, B is the number satisfying the conditions of B → ∞ and B/n5 = O (1), and $\hat{\gamma}(j)$ donates an autocovariance function. The autocovariance function can be given by:

$$\hat{\gamma}(j) = \frac{1}{n}\sum_{t=j+1}^{T}(\hat{u}_{t-j} - \hat{\rho}\hat{u}_{t-j-1})(\hat{u}_t - \hat{\rho}\hat{u}_{t-1}) \tag{3.53}$$

Where $\hat{\rho}$ denotes parameter estimate of the effect with no intercept of $\hat{u}_{t-1}$ on $\hat{u}_t$. The Z_a and Z_t test statistics can be as follow:

$$\boldsymbol{Z_t} = \frac{(\widehat{\rho}^{8} - 1)}{(\widehat{\gamma}(0) + 2\sum_{j=1}^{B} w(j/B)\widehat{\gamma}(j))/\sum_{1}^{n-1}\widehat{u}_t^2} \tag{3.54}$$

Where $\widehat{\gamma}(0) + 2\sum_{j=1}^{B} w(j/B)\widehat{\gamma}(j)$ can be defined as the long-run variance estimate of the residuals of $\widehat{u}_t$on $\widehat{u}_{t-1}$. These tests have no distribution and the asymptotic distribution of ADF is identical to the distribution of Z_t distribution.

The test statistics are the lowest values of these three tests among values for τ_1 and τ_2, with $\tau_1 \in T_1$ = (0.15, 0.70) and $\tau_2 \in T_2$ = (0.15+τ_1, 0.85). The reason of selecting the lowest value for each test statistic is that the lowest value shows the empirical evidence against the null hypothesis. These test statistics are defined as

$$ADF^* = \inf_{(\tau_1,\tau_2)\in \mathrm{T}} ADF(\tau_1, \tau_2), \tag{3.55}$$

$$Z_t^* = \inf_{(\tau_1,\tau_2)\in \mathrm{T}} Z_t(\tau_1, \tau_2), \tag{3.56}$$

$$Z_a^* = \inf_{(\tau_1,\tau_2)\in \mathrm{T}} Z_a(\tau_1, \tau_2), \tag{3.57}$$

Here, T = (0.15n, 0.85n). The concept to truncate the data by 15% on each side follows the idea of Gregory and Hansen (1996). Depending on the same logic, distance between the two regimes shifts be at least 15% can be applied (Hatemi, 2008).

4. RESULTS

This chapter provides the main results of the study. It provides graphical presentation to detect stationarity of time series, explanatory of data analysis, unit root test including ADF and Phillip Perron approaches. We have provided the results for cointegration analysis using Johansen, VEC and ARDL approaches.

4.1. Explanatory Data Analysis

We have provided some descriptive statistics for the data in order to know the characteristic of the data set. The descriptive statistics for the data including mean and standard deviation have been calculated. We have collected the data between 2009 and 2019.

Table 4.1. Descriptive statistics of the obtained data

Variable	N	Mean	Std. Dev.	Min.	Max.
Total electricity load	132	23451.36	8645.417	53880.58	53880.58
Rate of humidity	132	38.79517	7.273610	13.90189	62.58136
Tempreture	132	21.96987	1.466530	16.69333	27.20889
Wind speed	132	6.165348	3.595883	1.375000	17.54000

Table 4.1 shows that the average electricity load is 23,451.36 KW in 10 years while the average of humidity rate is 38.8. The average of temperature in Erbil is below 22 for 10 years. Additionally, the average of maximum wind speed is 6.2 () for 10 years in Erbil.

4.1.1. Data

We have obtained data from two sources for the period of 2009 and 2019. Several variables were collected including wind speed, temperature, electricity load, rate of humidity and time. The electricity load data is obtained from Erbil directorate of electricity.

Wind speed, temperature and rate of humidity are obtained from general directorate of agricultural and weather.

4.1.2. Variable explanation

All the variables will be explained in detail in terms of their unit measurement. The variables are electricity load, humidity, temperature, wind speed and time.

4.1.2.1. Load

The electricity load is the sum of total consumer. We have collected this data because we want to do forecasting the future values of electricity load in Erbil. The data is numerical, and it is collected from 2009 to 2019.the data is hourly electricity load which covers the period starting on (Jun 1, 2009) and ending on (Dec 31, 2019). Understanding the characteristics of electricity load can help us to forecast accurately and better model selection. There are many factors affecting electricity load which is defined as time and weather.

4.1.2.2. Humidity rate

The level of humidity in atmosphere can be influenced by weather variables such as air conditioning that leads to cooling loads in summer. The impact of humidity is greatly noticeable especially if there is high temperature. Humidity can affect power frequency as the power will be lower when humidity arises (Yang et al., 2018). Seasonal discomfort or heat stress equivalent is measured by the temperature humidity index, which depends on both surface air temperature and relative humidity. It typically has the highest association with season load and has an effect on electricity consumption only above a set cutoff temperature (Panjwani and Narejo, 2014).

4.1.2.3. Temperature degree

It's impossible to forecast the weather without considering temperature. It generates the majority of the load that is weather dependent. When temperatures deviate from normal, it might affect how much electricity people use. Heat storage in buildings delays the onset of the changes. While a drop in temperature below freezing increases heating energy consumption, a rise in temperature above freezing during the summer increases air conditioning energy consumption (increasing the cooling electricity load). The electrical load is often treated as a function of the effective temperature or temperature fluctuation when temperature impacts are predicted (Mawlood and Yahya, 2018).

4.1.2.4. Time

Time variables affecting the power load at hourly, daily, and seasonal intervals. The load curves demonstrate that there are definite laws governing the fluctuation of the electrical load with the hour or day. That's because seasonal, trend, and cycle power loads already take into account the overall influences of base temperature and hence have different levels of fluctuation in their electrical loads (Fahad and Arbab, 2014; Taylor and Buizza 2003).

4.1.2.5. Wind speed

This is a key factor that is related to weather-dependent power load. Wind speed has an impact on the weather throughout the winter and can be a direct result of the wind's cooling effect. Speed and temperature both have an impact on wind's cooling ability (Taylor and Buizza, 2003).

4.1.3. Visual explanation

In order to investigate the feature of the data, it is possible to use visualization approaches. The common application of weather data is that it involves seasonal cycles in many time series applications.

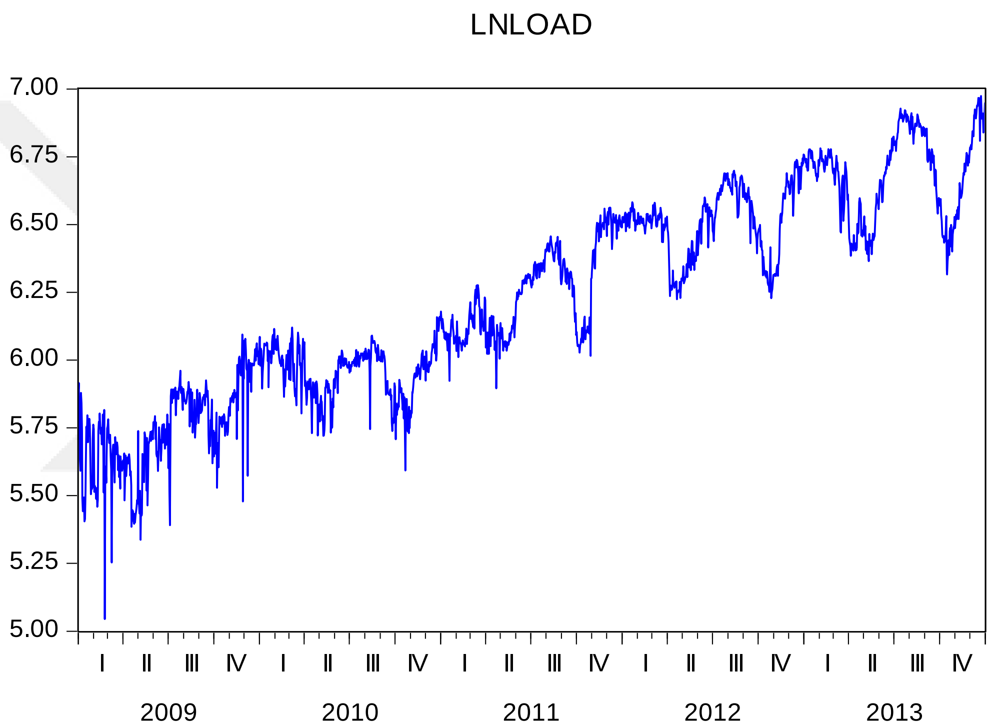

Figure 4.1. Load of electricity from 2009 to 2019.

Figure 4.1 shows log of electricity load between 2009 and 2013. The graph shows a clear trend with the present of seasonal changes. It reveals that the demand of electricity has been increased throughout the year. Thus, more loads have been consumed.

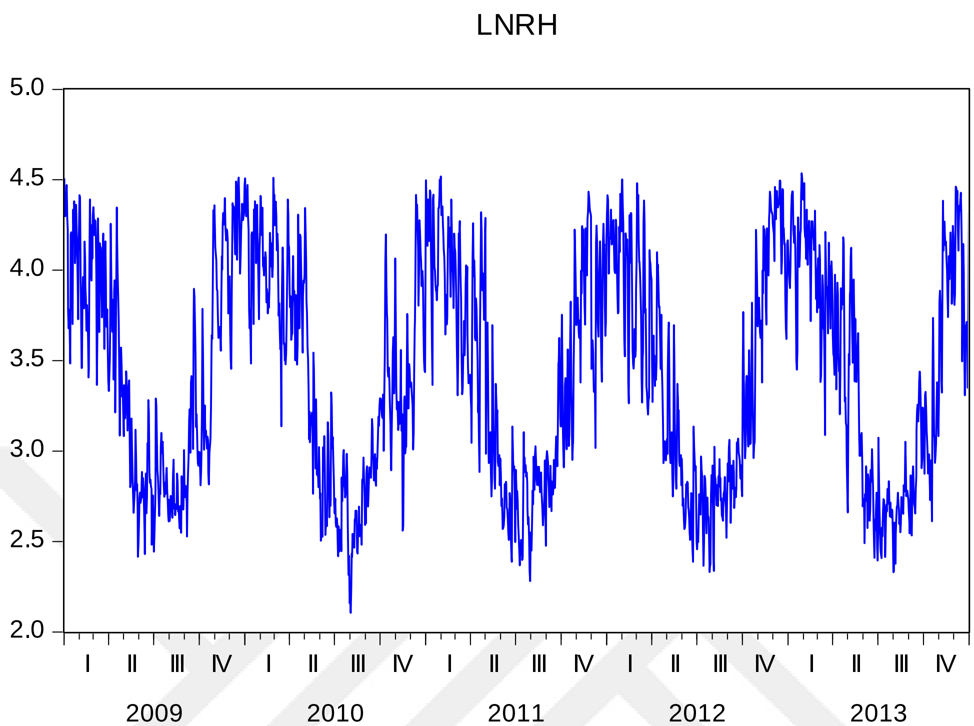

Figure 4.2. Relative humidity measure between 2009 and 2013.

In figure 4.2, the plot corresponding to log of the humidity rate illustrates that there is a clear seasonal pattern without trend. It shows that humidity rates change yearly. It can be noticed that humidity rate in Erbil-Iraq is between 2 and 5.

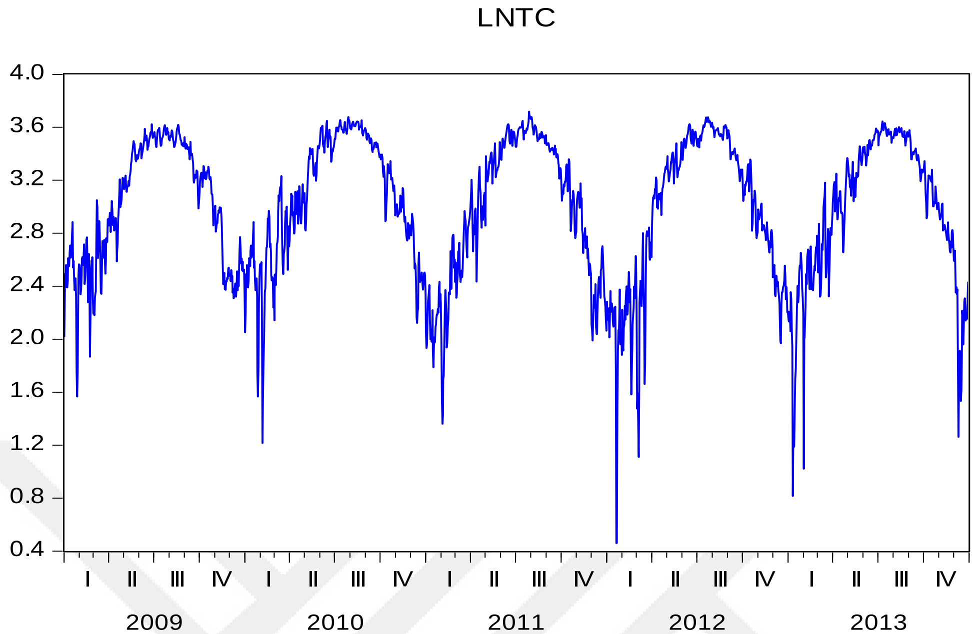

Figure 4.3. Shows the temperature degree of Erbil city throughout year 2009-2013

Figure 4.3 corresponding to log of the maximum temperature in Erbil shows that there is a clear seasonal pattern. There are some extreme values appear especially in fourth quarter of 2011 and fourth quarter of 2012.

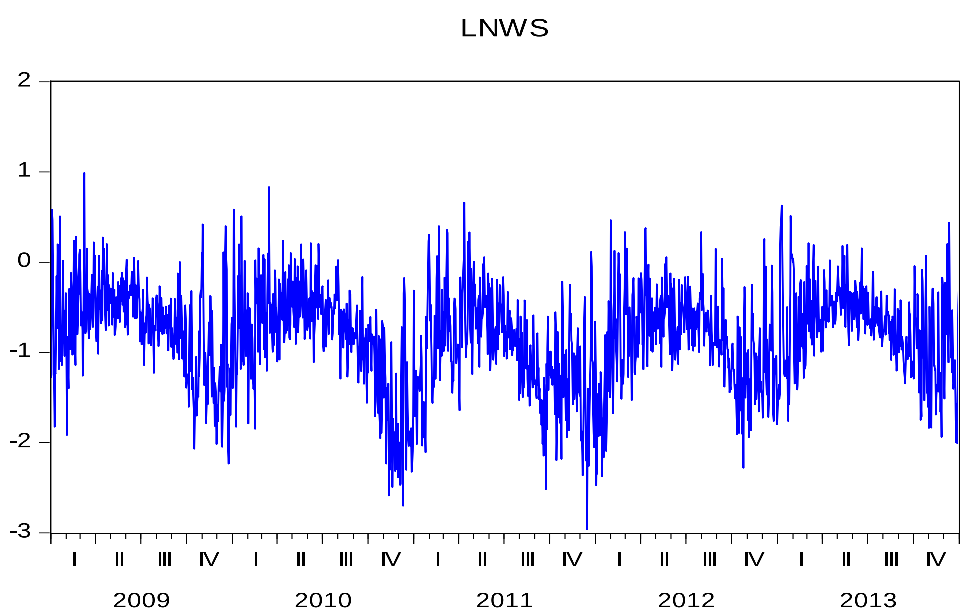

Figure 4.4. Shows the maximum wind speed in Erbil city throughout year 2009-2013

Figure 4.4 the plot corresponding to log the maximum wind speed in Erbil shows the clear seasonal pattern in the time series data. The seasonal pattern of five years can be seen and there is an extreme value in the fourth quarter in 2011.

Table 4.2. Load autocorrelation and partial correlation graph

Corr.	P. Corr.	AC	PAC	Q-Stat	Probability
|******	|*****	.990	.990	1791.9	0.000
|******	|* |	.984	.203	3563.2	0.000
|******	|* |	.979	.092	5317.6	0.000
|******	| |	.975	.065	7057.6	0.000
|******	| |	.970	.021	8782.8	0.000
|******	| |	.966	.046	10496.	0.000
|******	| |	.963	.031	12197.	0.000
|******	| |	.958	-.060	13882.	0.000
|******	| |	.954	.002	15553.	0.000
|******	| |	.950	.032	17211.	0.000
|******	| |	.946	.003	18856.	0.000
|******	| |	.941	-.032	20487.	0.000
|******	| |	.936	-.052	22100.	0.000
|******	| |	.931	.004	23698.	0.000
|******	*| |	.925	-.079	25276.	0.000
*******	| |	.920	-.001	26837.	0.000
|*******	| |	.915	.001	28381.	0.000
|*******	| |	.909	-.023	29907.	0.000
|******	| |	.904	-.005	31415.	0.000

Table 4.2. Load autocorrelation and partial correlation graph (Continued)

Corr.	P. Corr.	AC	PAC	Q-Stat	Probability
|******|	| |	.893	.006	34381.	0.000
|******|	| |	.887	-.021	35838.	0.000
|******|	| |	.882	-.011	37277.	0.000
|******|	| |	.877	.051	38702.	0.000
|******|	| |	.872	-.006	40110.	0.000
|******|	| |	.866	-.027	41501.	0.000
|******|	| |	.860	-.005	42875.	0.000
|******|	| |	.856	0.040	44234.	0.000
|******|	| |	.851	0.005	45578.	0.000
|******|	| |	.845	-0.014	46906.	0.000
|******|	| |	.840	-0.019	48218.	0.000
|******|	| |	.835	0.017	49515.	0.000
|******|	| |	.830	0.001	50796.	0.000
|******|	| |	.824	-0.012	52061.	0.000
|******|	| |	.819	0.019	53313.	0.000
|******|	| |	.814	-0.051	54547.	0.000

The table 4.2 shows ACF and PACF for log of electricity load. It can be notices that there is no seasonality in the series. Thus, the series is stationary.

Table 4.3. Wind spead autocorrelation and partial correlation graph

Corr.	P. Corr.	AC	PAC	Q-Stat	Probability
.|******|	.|******|	.795	0.795	85.418	0.000
.|***** |	.|** |	.723	0.247	156.62	0.000
.|***** |	.|** |	.709	0.229	225.48	0.000
.|***** |	.|* |	.714	0.210	296.01	0.000
.|***** |	.|. |	.670	0.017	358.53	0.000
.|***** |	.|* |	.680	0.164	423.51	0.000
.|***** |	.|. |	.644	-0.027	482.14	0.000
.|**** |	*|. |	.576	-0.134	529.42	0.000
.|**** |	.|. |	.567	0.049	575.72	0.000
.|**** |	.|. |	.580	0.059	624.46	0.000
.|**** |	.|* |	.585	0.107	674.47	0.000
.|**** |	.|* |	.588	0.115	725.45	0.000
.|**** |	.|. |	.565	-0.019	772.87	0.000
.|**** |	.|. |	.552	0.045	818.53	0.000
.|**** |	.|. |	.528	-0.048	860.71	0.000
.|*** |	*|. |	.481	-0.178	896.02	0.000
.|**** |	.|. |	.489	0.046	932.87	0.000
.|*** |	.|. |	.476	-0.048	968.00	0.000
.|*** |	.|. |	.468	0.062	1002.2	0.000
.|*** |	.|. |	.414	-0.062	1029.3	0.000
.|*** |	*|. |	.376	-0.115	1051.8	0.000
.|*** |	.|. |	.371	0.070	1073.9	0.000

Table 4.3. Wind spead autocorrelation and partial correlation graph (Continued)

<table>
<tr><th>Corr.</th><th>P. Corr.</th><th>AC</th><th>PAC</th><th>Q-Stat</th><th>Probability</th></tr>
<tr><td>.|*** |</td><td>.|. |</td><td>.371</td><td>-0.004</td><td>1096.3</td><td>0.000</td></tr>
<tr><td>.|*** |</td><td>.|* |</td><td>.391</td><td>0.092</td><td>1121.3</td><td>0.000</td></tr>
<tr><td>.|** |</td><td>*|. |</td><td>.348</td><td>-0.086</td><td>1141.3</td><td>0.000</td></tr>
<tr><td>.|** |</td><td>*|. |</td><td>.308</td><td>-0.105</td><td>1157.1</td><td>0.000</td></tr>
<tr><td>.|** |</td><td>.|. |</td><td>.275</td><td>-0.028</td><td>1169.9</td><td>0.000</td></tr>
<tr><td>.|** |</td><td>.|* |</td><td>.310</td><td>0.121</td><td>1186.2</td><td>0.000</td></tr>
<tr><td>.|** |</td><td>.|. |</td><td>.310</td><td>0.016</td><td>1202.7</td><td>0.000</td></tr>
<tr><td>.|** |</td><td>*|. |</td><td>.255</td><td>-0.134</td><td>1214.0</td><td>0.000</td></tr>
<tr><td>.|** |</td><td>.|. |</td><td>.229</td><td>-0.035</td><td>1223.2</td><td>0.000</td></tr>
<tr><td>.|** |</td><td>.|. |</td><td>.225</td><td>0.049</td><td>1232.2</td><td>0.000</td></tr>
<tr><td>.|** |</td><td>.|. |</td><td>.215</td><td>0.061</td><td>1240.4</td><td>0.000</td></tr>
<tr><td>.|* |</td><td>.|. |</td><td>.202</td><td>-0.032</td><td>1247.8</td><td>0.000</td></tr>
<tr><td>.|* |</td><td>.|. |</td><td>.200</td><td>-0.010</td><td>1255.0</td><td>0.000</td></tr>
<tr><td>.|* |</td><td>.|. |</td><td>.198</td><td>0.051</td><td>1262.3</td><td>0.000</td></tr>
</table>

We can see in table 4.3 that the series is not stationary using Correlogram which do not show stationarity of electricity load. It is clear from Correlogram that the non-decaying behavior of the sample ACF is because of lack of stationarity. It can be noticed that ACF suffered from linear decline but the PACF decaying sharply.

Table 4.4. Rate of hummidity autocorrelation and partial correlation graph

A. Corr.	P. Corr.	AC	PAC	Q-Stat	Probability
.|******|	.|******|	.778	0.778	81.745	0.000
.|*** |	***|. |	.444	-0.409	108.59	0.000
.|. |	***|. |	.029	-0.425	108.71	0.000
***|. |	***|. |	-.392	-0.424	130.00	0.000
*****|. |	***|. |	-.721	-0.393	202.35	0.000
******|. |	*|. |	-.840	-0.180	301.35	0.000
*****|. |	.|. |	-.727	0.002	376.08	0.000
***|. |	.|* |	-.417	0.144	400.93	0.000
.|. |	.|* |	-.019	0.076	400.99	0.000
.|*** |	.|. |	.374	0.032	421.31	0.000
.|***** |	.|. |	.673	0.057	487.51	0.000
.|******|	.|. |	.774	-0.036	575.85	0.000
.|***** |	.|* |	.689	0.104	646.40	0.000
.|*** |	*|. |	.394	-0.102	669.72	0.000
.|. |	*|. |	.008	-0.066	669.73	0.000
***|. |	.|. |	-.350	0.071	688.41	0.000
*****|. |	.|. |	-.628	-0.044	749.00	0.000
*****|. |	.|. |	-.740	-0.030	834.01	0.000
*****|. |	.|. |	-.637	0.032	897.60	0.000
***|. |	*|. |	-.385	-0.076	920.95	0.000
.|. |	.|. |	-.031	-0.006	921.10	0.000
.|** |	.|* |	.342	0.122	939.95	0.000

Table 4.4. Rate of hummidity autocorrelation and partial correlation graph (Continued)

<table>
<tr><th>A. Corr.</th><th>P. Corr.</th><th>AC</th><th>PAC</th><th>Q-Stat</th><th>Probability</th></tr>
<tr><td>.|**** |</td><td>.|. |</td><td>.601</td><td>0.003</td><td>998.48</td><td>0.000</td></tr>
<tr><td>.|***** |</td><td>.|. |</td><td>.713</td><td>0.068</td><td>1081.7</td><td>0.000</td></tr>
<tr><td>.|**** |</td><td>*|. |</td><td>.607</td><td>-0.126</td><td>1142.7</td><td>0.000</td></tr>
<tr><td>.|** |</td><td>*|. |</td><td>.348</td><td>-0.099</td><td>1162.9</td><td>0.000</td></tr>
<tr><td>.|. |</td><td>.|. |</td><td>.012</td><td>0.026</td><td>1162.9</td><td>0.000</td></tr>
<tr><td>**|. |</td><td>.|* |</td><td>-.319</td><td>0.108</td><td>1180.2</td><td>0.000</td></tr>
<tr><td>****|. |</td><td>*|. |</td><td>-.589</td><td>-0.102</td><td>1239.7</td><td>0.000</td></tr>
<tr><td>*****|. |</td><td>.|. |</td><td>-.684</td><td>-0.043</td><td>1320.7</td><td>0.000</td></tr>
<tr><td>****|. |</td><td>.|. |</td><td>-.592</td><td>-0.002</td><td>1382.2</td><td>0.000</td></tr>
<tr><td>***|. |</td><td>.|. |</td><td>-.346</td><td>-0.027</td><td>1403.3</td><td>0.000</td></tr>
<tr><td>.|. |</td><td>.|. |</td><td>-.022</td><td>0.010</td><td>1403.4</td><td>0.000</td></tr>
<tr><td>.|** |</td><td>*|. |</td><td>.288</td><td>-0.084</td><td>1418.4</td><td>0.000</td></tr>
<tr><td>.|**** |</td><td>.|. |</td><td>.552</td><td>0.041</td><td>1473.9</td><td>0.000</td></tr>
<tr><td>.|***** |</td><td>.|* |</td><td>.660</td><td>0.098</td><td>1554.2</td><td>0.000</td></tr>
</table>

Table 4.4 the table above indicates that rate of humidity has a seasonality pattern. o rh has seasonality. It can be noticed that ACF suffered from linear decline and fluctuation but the PACF decaying sharply. Thus, the series is not stationary as we should take differencing to adjust seasonality.

Table 4.5. Temperature’s autocorrelation and partial correlation graph

A. Corr.	P. Corr	AC	PAC	Q-Stat	Probability
.|******|	.|******|	0.845	0.845	96.433	0.000
.|*** |	******|. |	0.476	-0.833	127.28	0.000
.|. |	****|. |	-0.010	-0.501	127.29	0.000
****|. |	***|. |	-0.486	-0.358	159.87	0.000
******|. |	*|. |	-0.821	-0.190	253.67	0.000
*******|. |	**|. |	-0.936	-0.238	376.72	0.000
******|. |	*|. |	-0.799	-0.067	467.07	0.000
***|. |	*|. |	-0.455	-0.122	496.62	0.000
.|. |	*|. |	0.001	-0.092	496.62	0.000
.|*** |	.|. |	0.454	0.053	526.50	0.000
.|******|	.|. |	0.781	0.041	615.72	0.000
.|******|	.|. |	0.895	-0.043	733.75	0.000
.|******|	.|. |	0.769	-0.011	821.66	0.000
.|*** |	.|. |	0.440	-0.013	850.68	0.000
.|. |	.|. |	0.000	0.017	850.68	0.000
***|. |	.|. |	-0.433	0.005	879.21	0.000
*****|. |	.|. |	-0.741	0.060	963.74	0.000
******|. |	.|. |	-0.849	-0.025	1075.7	0.000
*****|. |	*|. |	-0.736	-0.093	1160.4	0.000
***|. |	.|. |	-0.427	0.040	1189.3	0.000

Table 4.5. Temperature's autocorrelation and partial correlation graph (Continued)

A. Corr.	P. Corr	AC	PAC	Q-Stat	Probability
.|. |	.|. |	-0.009	0.045	1189.3	0.000
.|*** |	.|. |	0.409	0.046	1216.2	0.000
.|***** |	.|. |	0.710	-0.033	1298.0	0.000
.|******|	*|. |	0.815	-0.072	1406.8	0.000
.|***** |	.|. |	0.702	-0.022	1488.2	0.000
.|*** |	*|. |	0.399	-0.073	1514.8	0.000
.|. |	.|* |	-0.001	0.077	1514.8	0.000
***|. |	.|. |	-0.391	0.012	1540.8	0.000
*****|. |	.|. |	-0.667	0.065	1617.2	0.000
******|. |	.|. |	-0.762	-0.027	1717.8	0.000
*****|. |	.|. |	-0.654	0.009	1792.6	0.000
***|. |	.|. |	-0.374	0.019	1817.4	0.000
.|. |	.|. |	-0.004	-0.052	1817.4	0.000
.|*** |	.|* |	0.365	0.098	1841.4	0.000
.|***** |	.|. |	0.629	-0.018	1913.6	0.000
.|***** |	.|. |	0.719	-0.039	2008.9	0.000

Table 4.5 the table above indicates that tempreture has also a seasonality pattern. Has seasonality. It can be noticed that ACF suffered from linear decline and fluctuation but the PACF decaying sharply. Thus, the series is not stationary as we should take differencing to adjust seasonality.

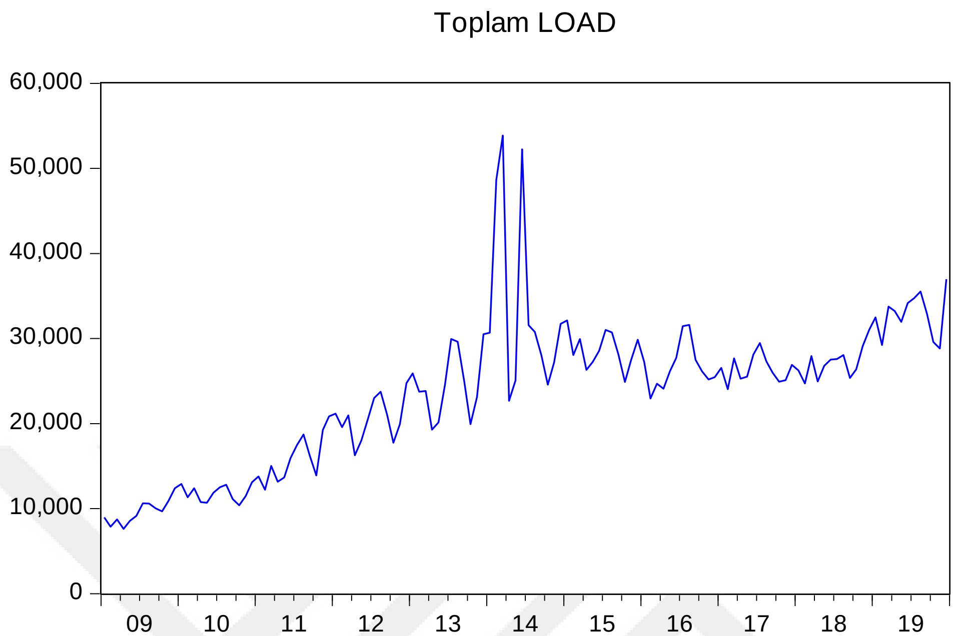

Figure 4.5. Sum of electricity load from 2009 and 2019

In figure 5.5 we have taken the sum of load yearly to see the pattern of the data. It can be seen from figure 6 that there is still trend and seasonal patterns in the data. The graph shows clearly that the load data is not stationary.

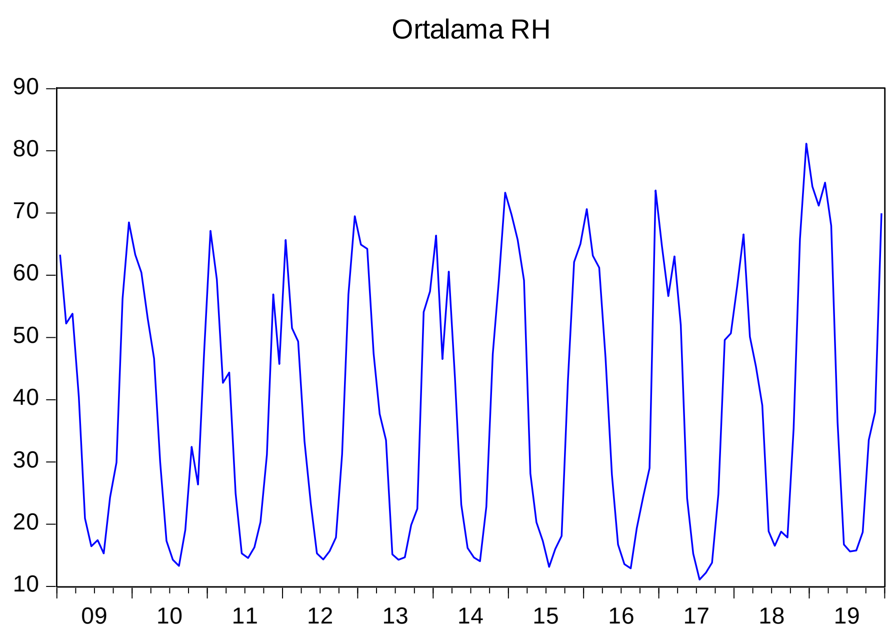

Figure 4.6. Seasonality adjustments to the data using MA

In figure 5.6 we have taken a monthly average of the rate of humidity. We have adjusted seasonality to the data by using moving average approach. It can be seen the data is stationary.

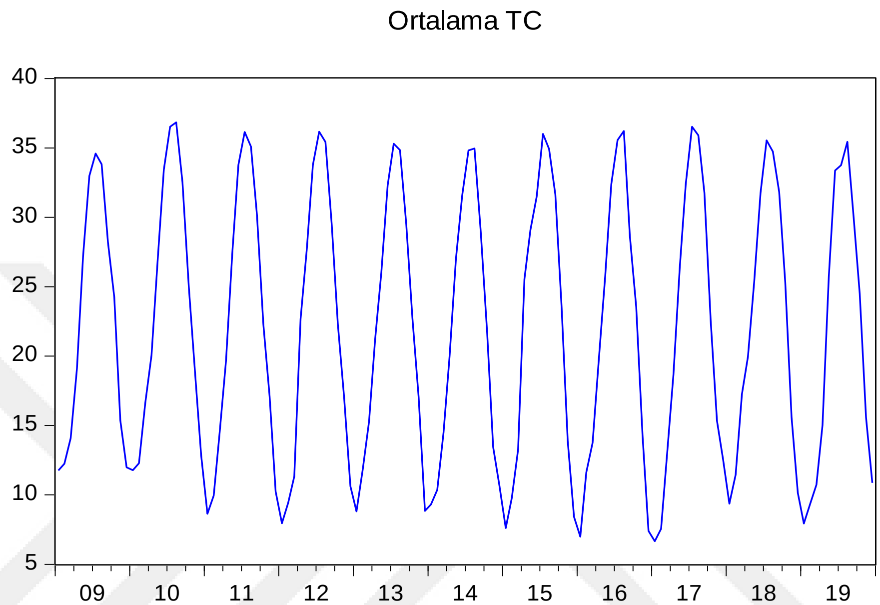

Figure 4.7. Stationarity of temperature data from 2009 to 2019

In figure 5.7 we have taken a monthly average of the temperature degree have adjusted seasonality to the data by using moving average approach. It can be seen the data is stationary.

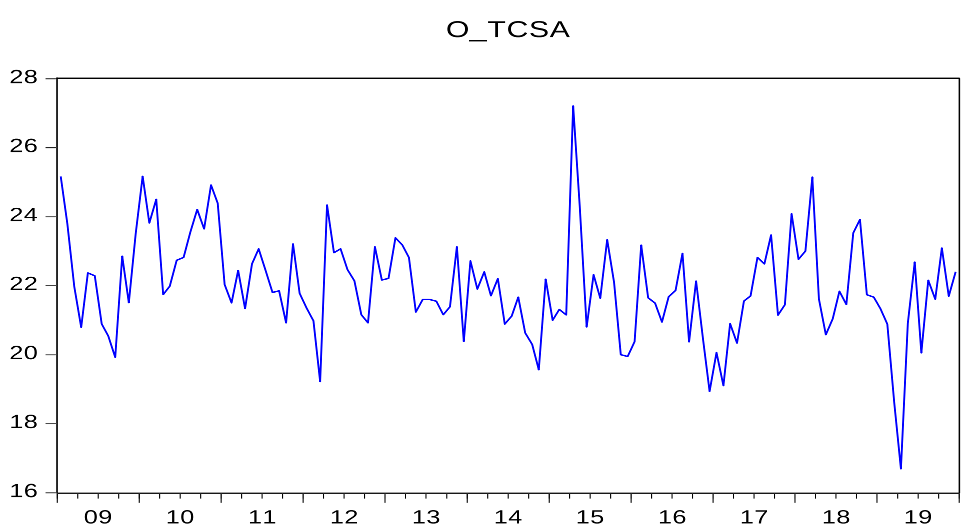

Figure 4.8. Average of seasonality adjusted temperature between 2009 and 2019.

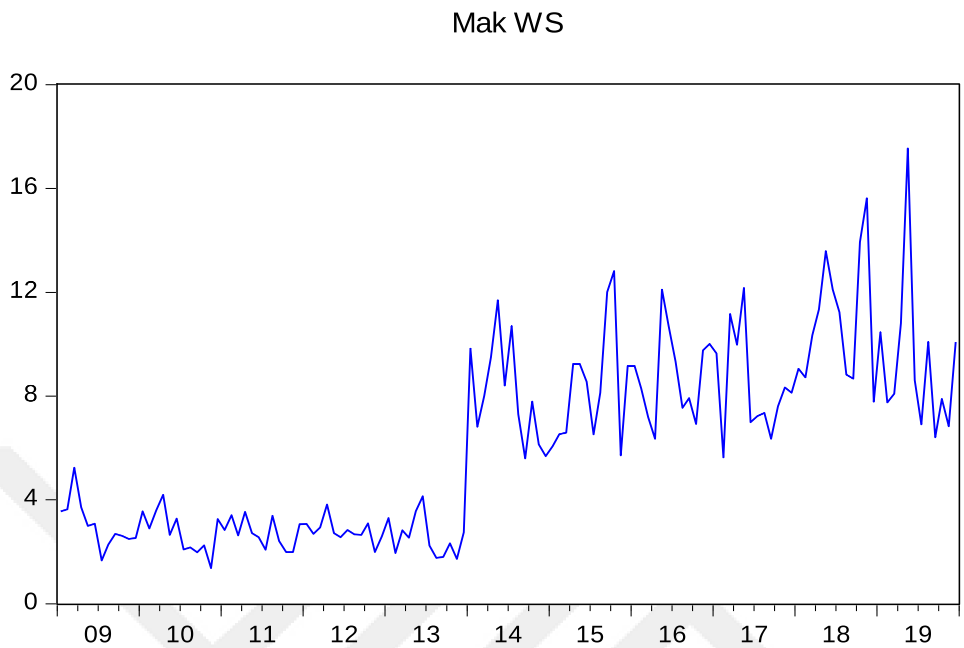

Figure 4.9. Distribution of average wind speed data from 2009 to 2019.

4.2. Unit root test

We tested for unit root to see if the time series data was stationary. The unit root has been tested using Augmented Dickey-Fuller (ADF) and Phillip Perron (PP) methods.

Table 4.6. Augmented Dickey-Fuller (ADF) results

Variable	Intercept		None	
	Level	1st Difference	Level	1st Difference
T_Loadsa	-1,66	-9,95***	0,75	-9,83***
O_Rhsa	-7.93***	-13,76***	-0,43	-13,81***
o_Tcsa	-7,61***	-11,58***	-0,28	-11,63***
m_Ws	-1,37	-10,69***	-0,15	-10,71***

Note: Statistical significance is shown by ***, **, and * at percentages of 1, 5, and 10%.

Table 4.6 shows the result of an Augmented Dickey Fuller test for the unit root. Variables are tested including wind speed, electricity load, temperature, and rate of humidity. We have used Augmented Dickey Fuller approach to test unit root for all variables separately. The null hypothesis is that each variable has unit root. The tests for unit root in levels with no intercept illustrates all the variables are non-stationary. After taking fist deference, the variables became stationary. However, it can be noticed that the ADF test results with allowing intercept reveal that only variables of rate of humidity and temperature are stationary. After taking first difference with allowing intercept, all the variables exhibit unit root process.

It is worth mentioning that the ADF tests do not allow for the presence of structural brecks. Thus, the results may be biased (Glynn et al, 2009). We can conclude that not all the four series are stationary, so we have taken difference and all the series become stationary with or without allowing intercept.

We know from the unit root test that the three series aren't stationary, so we make them stationary by taking the first difference between each of them.

Table 4.7. Philip Perron test results

Variable	Intercept		None	
	Level	1st Difference	Level	1st Difference
T_Loadsa	-2,59	-18,47***	0,33	-17,74***
O_Rhsa	-3,06*	-9,07***	-0,28	-9,09***
o_Tcsa	-7,60***	-33,18***	-0,61	-33,07***
m_Ws	-1,64	-10,34***	-0,41	-10,37***

Note: Statistical significance is shown by ***, **, and * at percentages of 1, 5, and 10%.

Table 4.7 shows the result of a Phillip Perron test for the unit root for all variables separately. The same asymptotic distribution as the ADF can be obtained for the PP $\boldsymbol{Y_t}$ and $\boldsymbol{Y_\pi}$ under the null hypothesis which is $\boldsymbol{\pi = 0}$. One of the advantages of the PP test is that it can be robust to the general form of heteroskedasticity in the error term $\boldsymbol{u_t}$ unlike the ADF test. Also, it is not necessary to specify lag length for the test. The tests for unit root in levels without intercept illustrates all the variables are non-stationary. After taking first difference, the variables became stationary. Additionally, it can be noticed that the PP test results with allowing intercept reveal that only variables of rate of humidity and temperature are stationary. After taking first difference with allowing intercept, all the variables exhibit unit root process.

4.3. Co-integration

Since wind speed and load are stationary at the same level, the cointegration relationship between them has been investigated.

Table 4.8. Johansen cointegration results

Null Hypothesis	Trace Statistic	0.05 Critical Value	Prob.
No Cointegration	21.47371	20.26184	0.0339
At most 1	4.342999	9.164546	0.3634
Null Hypothesis	Trace Statistic	0.05 Critical Value	Prob.
No Cointegration	17.13071	15.89210	0.0318
At most 1	4.342999	9.164546	0.3634

Table 4.8 the long-run relationship between the covariates estimated applying Johansen cointegration method. There results of trace statistics reveals that there at least one cointegration equation. We test the null hypothesis of no cointegration in which there is no cointegration among variables. Since the trace statistic value is greater than 5% critical value, we reject null hypothesis. At most 1 reveals that there is at least one cointegration equation and it is statistically significant (trace statistics is lower than 5% critical value). Thus, from the cointegration test results given in Table 4; from the trace statistics and the maximum eigenvalues, it is seen that there is at most one cointegration between the series.

Table 4.9. VECM estimation results

Long run equation		
	t_load_t	m_ws_t
Constant	-9.022635***	15.23428
t_load_t	-	-1.688452 (0.30931)
m_ws_t	-0.592259 *** (4.12085)	
Short run equation		
	Δt_load_t	Δm_ws_{t-1}
$VECT_t$	-0.119075*** [-2.72307]	0.269706*** [3.25243]
Δt_load_{t-1}	-0.067209 [-0.74496]	0.057340 *** [0.33515]
Δm_ws_{t-1}	-0.033461 [-0.75598]	-0.248338*** [-2.95864]

Note: Statistical significance is shown by ***, **, and * at percentages of 1, 5, and 10%.

Table 4.9 according to the Weak Erogeneity tests, the externality hypothesis was rejected for all series. Therefore, long and short term (Error correction model) VECM (1) estimation results between electricity city load and wind speed values are given in Table 4.5

The vector correction coefficient of the model estimated according to Table 4.5 is negative and significant. This indicates that the error correction mechanism is working. Deviations from equilibrium come to equilibrium in the long run. A 1% increase in wind speed in the long term reduces the total load by 0.59%. In the short run equation, wind speed has no effect on the total load.

Table 4.10. Lagrange multiplier (LM) autocorrelation test

Number of lags	LM-Stat	Prob
1	15.74558	0.2034
2	35.94271	0.1922
3	9.412644	0.0616
4	7.504750	0.1115
5	4.596099	0.3313
6	21.83374	0.0002
7	2.579093	0.6305
8	5.350340	0.2532
9	3.043923	0.5505
10	4.609793	0.3297
11	0.761617	0.9435
12	25.99336	0.7512

The LM test was performed to find out if there is autocorrelation between the error terms in the VAR model.

Table 4.10 according to the results obtained from the LM test, it was determined that there was no autocorrelation problem at the 5% significance level according to the 12th lag. Then, whether the variance of the error terms is constant for all observations was examined with the White Variance Test.

Table 4.11. VEC residual heteroskedasticity tests

White heteroskedaticity tests	
χ^2 Stat	Prob.
38.73644	0.1017

Table 4.11 according to the results obtained in Table 4.7, it was seen that there was no heteroskedasticity problem at the 5% significance level in the model whose χ2 value was estimated. In other words, it was concluded that the variance of the error term did not change for the whole sample. Thus, it was decided that the 12th lag was appropriate for the

Tado-Yamamoto causality test due to the absence of autocorrelation and varying variance. Considering the unit root test results in the study, it was determined as k+d_max=12+1=13 lagged since the variables had the greatest degree of integration (d_max) of 1 (i.e. the variables were I (1) at the most).

4.4. Auto-Regressive Distributed Lag (ARDL)

We have used ARDL model because of the fact that not all variables are stationary at the same level. The maximum lag length for the ARDL(p, q1,q2) model, with total load series dependent variables temperature and rate of humidity as the independent variables, was taken as 4 and the ARDL(3,0,0) model was chosen by deciding the best one among 500 models out of Model 3 according to the AIC criterion. ARDL (3,0,0) model results are given in Table 4.12.

Table 4.12. ARDL (3,0,0) test results

Dependent variable: t_load	Coefficientt	Std. Error	t	Prob.
LNT_LOOAD(-1)	0.713124	0.081969	8.699910	0.0000
LNT_LOOAD(-2)	-0.227634	0.102391	-2.223182	0.0280
LNT_LOOAD(-3)	0.442318	0.078449	5.638266	0.0000
LNO_TC	-0.196111	0.063102	-3.107841	0.0023
LNO_RH	-0.152374	0.053450	-2.850756	0.0051
C	1.854697	0.490324	3.782594	0.0002
$\bar{R}^2$=0,88 F=188.89 (P=0.000), DW=1.88				
Diagnostic Tests:				
Serial Corelation (Breush-Godfrey): F= 0.750562 (P= 0.3863)				
Specification:(Ramsey-RESET): F=3.065375 (P= 0.1825)				
Normality: (Jarque-Bera): JB=0.38 (P=0.5544)				
Heteroscedasticity: /Breush-Pagan-Godfrey): F= 8.763309 (P= 0.1189)				

Table 4.12 illustrates the results of ARDL bounds testing. The result of F-test reveals that the cointegration exists between variables of interest (p-value=0.000).

According to the ARDL (3,0,0) model estimation results given in Table 4.12, all coefficients were found to be statistically significant, and as a result of the diagnostic tests carried out, it was determined that the model did not have any serial correlation, changing variance, specification and normality problems. In Table 4.13, the results of the boundary test to investigate the existence of a cointegration relationship between the series are given.

Table 4.13. ARDL bounds test

F-Bounds Test				
H_0: Null Hypothesis: No levels relationship				
Test Statistic	Value	Signif.	I(0)	I(1)
F-statistic	4.790	10%	3.17	4.14
K	2	5%	3.79	4.85
		1%	5.15	6.36
t-Bounds Test				
Test Statistic	Value	Signif.	I(0)	I(1)
t-statistic	-2.344	10%	-2.57	-3.21
		5%	-2.86	-3.53
		1%	-3.43	-4.1

Table 4.13 it was calculated as F=4.78 for the F limit test seen in Table 13. Since this value is greater than the upper critical values for only 10% (F > I (1)), the null hypothesis of "no cointegration" will be rejected (for 0.01 the upper critical value is 6.36, for 0.05 the upper critical value = 4.7 and for 10% the upper critical value = 3.97).vAccording to the F bounds test, the series are cointegrated according to 10%. However, it is necessary to test whether this cointegration is a valid cointegration, since model 3 is used as the conditional error correction model. For this reason, the t bounds test is given in the last section of Table 4.13. For the t-bonds test, t=2.34 was calculated and since this statistic is small in absolute value (e.g., 3.34<3.78 for 5% error level), cointegration between the series is not.

5. DISCUSSION AND CONCLUSION

The objective of this thesis was generally to assess the relationship between climate change and electricity consumption in Iraq using the cointegration analysis for the time series data between 2009 and 2019. We obtained the data from the two sources and the obtained variables were electricity load, temperature, humidity rate and Wind speed. We have provided descriptive statistics to know the mean and SD of the variables and we have provided various graphs to see whether the stationarity is presented for all variables. We have found using graphic inspection that variables of rate of humidity, wind speed and temperature were not stationarity. We adjusted seasonality for the variables of rate of humidity and wind speed.

Additionally, we have used Augmented Dickey-Fuller (ADF) and Phillip Perron (PP) approaches to test unit root. The tests for unit root in levels with no intercept using ADF revealed that all the variables were non-stationary. After taking fist deference, the variables became stationary. The ADF test results with allowing intercept revealed that only variables of rate of humidity and temperature were stationary. After taking first difference with allowing intercept, all the variables exhibited unit root process. There are some studies that used ADF to test stationarity of weather data using ADF approach. For example, Tran et al (2020) have used ADF method to test their temperature data.

The long-run relationship between the covariates estimated applying Johansen cointegration method. There results of trace statistics revealed that there at least one cointegration equation. We looked at the null hypothesis of no cointegration, which says that no variables are linked together. Null hypothesis was ruled out since the trace statistic value exceeded the crucial value by 5 percent. We thus rejected it. Wind speed and electrical load have a long-term connection, which means that when wind speed increases, so can the power load.

Similar results have been obtained by Zhang et al., (2019) as studied Climate effects such as temperature and electricity consumption as they found that increasing daily temperature influences the electricity consumption in the rural area. Moreover, A. Ngah

Nasaruddin et al showed with a positive relationship between temperature-electricity consumption, and rainfall-electricity consumption and Hernández, L et al have found a correlation between weather and electric power demand, as they designed a new ANN-based architectural models for electric load forecasting.

According to the ARDL (3,0,0) model estimation results, all coefficients were found to be statistically significant. There was a relationship between load and RH and temperature. The result of the diagnostic tests was also indicated that the model did not have any serial correlation, changing variance, specification, and normality problems. The results of the boundary test to investigate the existence of a cointegration relationship between the series are given.

RECOMMENDATION AND FURTHER RESEARCH

Based on the results and studied variables, we will recommend the followings:

1- Weather variables such as temperature, wind speed and rate humidity can be considered when the electricity system is set up
2- Because of the weather change, separate models can be applied for seasonal data.
3- Granger causality can be also applied to see bi-directional relationship between electricity load and weather variables.
4- The ministry of electricity can use these methodologies and results obtained in the future to estimate hourly and daily load demand.
5- The cointegration analysis is flexible tool for analyzing dynamic phenomenon; thus, it can be applied in application in many different fields such GDP, economic and marketing.
6- It is important for government to enhance and establish renewable energy.
7- Further research is needed using different variables such as GDP and environmental variables to find long run correlation using different methods such as forecasting and Artificial Neural Network.

REFERENCES

Abdallah, K., Belloumi, M., De Wolf, D., 2013. Indicators for sustainable energy development: A multivariate cointegration and causality analysis from Tunisian road transport sector. Renewable and Sustainable, ***Energy Reviews***, **25**: 34-43.

Yong, A. H., Christopher, G., Sirimon, T., 2004. Co-integration and causality in the Asian and emerging foreign exchange markets: Evidence from the 1990s financial crises. ***International Review of Financial Analysis,*** **13** (4): 479-515.

Boukhelkhal, A., Bengana, I., 2018. Cointegration and causality among electricity consumption, economic, climatic and environmental factors. ***Energy***, **163:** 1193-1206.

Brooks, C., 2002. ***Introductory Econometrics for Finance.*** United Kingdom, Cambridge University Press, xxv+701 pp.

Belloumi, M., 2009. Energy consumption and GDP in Tunisia: cointegration and causality analysis. ***Energy policy,*** **37** (7): 2745-2753.

Bélaïd, F., Abderrahmani, F., 2013. Electricity consumption and economic growth in Algeria: A multivariate causality analysis in the presence of structural change. ***Energy Policy, Elsevier,*** **55** (C): 286-295.

Chien, M., Shu-Jung, C. L., Lee, C. C., 2018. Time-varying co-movement of the prices of three metals and oil: Evidence from recursive cointegration. **Resources Policy, 57**: 186-195.

Caporale, G., Gil-Alana, L. A., 2019. Long-term interest rates in Europe: A fractional cointegration analysis***. International Review of Economics & Finance,*** **61***:* 170-178.

DBaker, D., Md.Kamruzzaman, R. M., 2015. Regional aviation and economic growth: cointegration and causality analysis in Australia. ***Transport Geography,*** **43**: 140-150.

Dickey, D. A., Fuller, W. A., 1981. Likelihood ratio statistics for autoregressive time series with a unit root. ***Econometric Society*****, 49**: 1057–1072***.***

Dickey, D. A., Fuller, W. A., 1979. Distribution of the estimators for autoregressive time series with a unit root. ***American statistical association*****, 74**: 427–431.

Esso, L. J., 2010. Threshold cointegration and causality relationship between energy use and growth in seven African countries. Energy Economics, 32 (6): 1383-1391.

Esso, L. J., Keho, Y., 2016. Energy consumption, economic growth and carbon emissions: Cointegration and causality evidence from selected African countries. ***Energy,*** **114** (1): 492-497.

Erica., 2020. A Guide to Conducting Cointegration Tests. Aptech, *https://www.aptech.com/blog/a-guide-to-conducting-cointegration-tests/.* Published January 28, 2020 · Updated May 12, 2021.

Emmanouil, H., Heracles, P., Haralambopoulos, D., 2011. Co2 emission, GDP and energy intensity: A Multivariate cointegration and causality analysis for Greece, 1977–2007. ***Energy,*** **88** (1): 1377-1385.

Francisco, F., Ramos, R., 2001. Exports, imports, and economic growth in Portugal: evidence from causality and cointegration analysis. ***Economic Modelling,*** **18** (4): 613-623.

Falk, M., 2012. ***A First Course on Time Series Analysis — Examples with SAS***. Chair of Statistics. Wurzburg, Germany, University of Wurzburg.

Franses, Philip Hans, 1998. ***Time series models for business and economic forecasting.*** United Kingdom. The press syndicate of the University of Cambridge.

Salahuddin, G., Shahbaz, M., Mohamed, A., and Teulon, F., 2014. Financial development and poverty reduction nexus: A cointegration and causality analysis in Bangladesh. ***Economic Modelling,*** **36**: 405-412

Gujarati, D. N., Porter, D. C., 2009. ***Basic Econometrics***. McGraw-Hill/Irwin, a business unit of The McGraw-Hill Companies, Inc., 1221 Avenue of the Americas, New York, NY, 10020.

Granger, C. W. J., 1981. Some properties of time series data and their use in econometric model specification. ***Econometrics,*** **16** (1): 121-130

Gregory C. R., Ljung, G. M., George, E. P., Box Jenkins, G. M., 2015. ***Time Series Analysis: Forecasting and Control***. United states. Wiley, fifth edition. John Wiley and Sons Inc., Hoboken, New Jersey, pp. 712. ISBN: 978-1-118-67502-1.

Gebre-Mariam, Y. K., 2011. Testing for unit roots, causality, co-integration, and efficiency: The case of the northwest US natural gas market. ***Energy,*** **36** (5): 3489-3500.

Huang, B. N., Yang, C. W., Shan Hu, J. W., 2000. Causality and co-integration of stock markets among the United States, Japan and the South China Growth Triangle. ***International Review of Financial Analysis,*** **3**: 281-297**.**

Hyndman, R. J., Athanasopoulos, G., 2018. ***Forecasting: principles and practice,*** 2nd edition, OTexts. Melbourne, Australia. Monash University, Accessed on Jun 2021.

Hatemi-J, A., 2012b. Asymmetric causality tests with an application. ***Empirical Economics,*** **43**: 447–456.

Hamilton, J. D., 1994. ***Time Series Analysis,*** Princeton University Press, New Jersey. United States.

Hernández, L., Baladrón, C., Aguiar, J. M., Calavia, L., Carro, B., Sánchez-Esguevillas, A., Cook, D. J., Chinarro, D., & Gómez, J., 2012, A Study of the Relationship between Weather Variables and Electric Power Demand inside a Smart Grid/Smart World Framework. ***Sensors (Basel, Switzerland),*** **12** (9): 11571–11591.

Kadir, D., 2020. Time Series Modeling to Forecast on Consuming Electricity: A case study Analysis of electrical consumption in Erbil City from 2014 to 2018. ***Al Rafidain,*** **46**: 473-485.

Khadija M., 2015. ***Estimation and Forecasting of the Dynamic Relationship between Stock Prices and Exchange Rates: A Comparative Study.*** Doctoral Thesis, University of Huddersfield. United Kingdom.

Kassouri, Y., Altıntaş, H., 2020. Threshold co-integration, nonlinearity and frequency domain causality relationship among Turkish Lira and stock price. ***Research in International Business and Finance,*** **52**: 101097.

Lütkepohl, H., 2005. ***New introduction to multiple time series analysis.*** Berlin, Germany, Springer.

Mawlood, K. I., Yahya, R. O., 2018. Using Dynamic Linear Models and Kalman Filter for modeling and forecasting electricity load in Erbil city. ***Humanity Sciences,*** **22** (4).

Muhammad, S., Mihai, M., Aviral Kumar, T., 2012. Revisiting the relationship between electricity consumption, capital and economic growth: cointegration and causality analysis in Romania***. Economic forecasting,*** **3**: 97-120.

Maddala, G. S., & Kim, I. M., 1998. ***Unit roots, cointegration, and structural change.*** Cambridge University Press, 1998 - Business & Economics.

Nasaruddin, A. N., Tee, B. T., Tahir, M. M., Jasman, E. S., 2021. Data Assessment on the relationship between typical weather data and electricity consumption of academic building in Melaka. ***Data in brief,*** **35:** 106797.

Nazlioglu, S. K., Adiguzel, U., 2014. Electricity Consumption and Economic Growth in Turkey: Cointegration, Linear and Nonlinear Granger Causality. ***Energy sources,*** **9** (4): 315-324.

Owyong, D., Wong, W. K., Horowitz, I., 2015. Cointegration and causality among the onshore and offshore markets for China's currency. ***Asian Economics,*** **41**: 20-38.

Perron, P., 1989. The great crash, the oil price shock and the unit root hypothesis. ***Econometrica,*** **57**: 1361-1401.

Perron, P., Vogelsang, T. J., 1992. Non-stationarity and level shifts with an application to purchasing power parity. ***J Bus and Econ Statist***, **10**: 301-20.

Pilli-Sihvola, K., Aatola, P., Ollikainen, M., & Tuomenvirta, H., 2010. Climate change and electricity consumption—Witnessing increasing or decreasing use and costs?. Energy Policy, 38 (5), 2409-2419.

Rafindadi, A., Ozturk, I., 2017. Impacts of renewable energy consumption on the German economic growth: Evidence from combined cointegration test. ***Renewable and Sustainable Energy Reviews*, 75:** 1130-1141.

Rahman, M. M., AbulKashem, M., 2017. Carbon emissions, energy consumption and industrial growth in Bangladesh: Empirical evidence from ARDL cointegration and Granger causality analysis. ***Energy Policy,*** **110**: 600-608.

Runsheng Yin, R., Jintao, X., 2003. Identifying the inter-market relationships of forest products in the Pacific Northwest with cointegration and causality tests. ***Forest policy and Economics,*** **5** (3): 305-315.

Shahbaz, M., Lean, H. H., Shabbir, M. S., 2012. Environmental Kuznets Curve hypothesis in Pakistan: Cointegration and Granger causality. ***Renewable and Sustainable Energy Reviews,*** **16** (5): 2947-2953.

Said, S. E., Dickey, D. A., 1984. Testing for unit roots in autoregressive-moving average models of unknown order. ***Biometrika,*** **71** (3): 599–607.

Sailor, D. J., & Muñoz, J. R., 1997. Sensitivity of electricity and natural gas consumption to climate in the USA—Methodology and results for eight states. Energy, 22 (10), 987-998.

Stelios, D. B., Cees, G. H. D., 2008. The relationship between crude oil spot and futures prices: Cointegration, linear and nonlinear causality. ***Energy Economics,*** **30** (5): 2673-2685.

Solarin, S. A., Shahbaz, M., 2013. Trivariate causality between economic growth, urbanization and electricity consumption in Angola: Cointegration and causality analysis. ***Energy Policy,*** **60**: 876-884

Tran, T. T. K., Lee, T., Shin, J. Y., Kim, J. S., Kamruzzaman, M., 2020. Deep Learning-Based Maximum Temperature Forecasting Assisted with Meta-Learning for Hyper parameter Optimization. ***Atmosphere,*** **11** (5): 487-495.

Tiwari, A. K., Shahbazb, M., Adnan Hye, Q. M., 2013. The environmental Kuznets curve and the role of coal consumption in India: Cointegration and causality analysis in an open economy. ***Renewable and Sustainable Energy Reviews,*** **18:** 519-527.

Wei W. S., 2006. ***Time series analysis: univariate and multivariate***. Boston: Pearson.

Yang, C., Zhang, W., Zhou, N., Tang, L., Li, T., 2018. Influence of temperature and humidity on power frequency electric field intensity. ***Integrated Ferroelectrics,*** **191** (1): 104–110.

Zivot, E., Andrews, D., 1992. Further evidence of great crash, the oil price shock and unit root hypothesis, ***Business and Economic Statistics,*** **10**: 251-270.

EXTENDED TURKISH SUMMARY
(GENİŞLETİLMİŞ TÜRKÇE ÖZET)

İKLİM DEĞİŞİKLİĞİNİN ERBİL'DE KONUT ELEKTRİK TÜKETİMİNE ETKİSİ: EKONOMETRİK BİR ANALİZ

ISMAEL, Zewar Omar
Yüksek Lisans Tezi, İstatistik Anabilim Dalı
Tez Danışmanı : Dr. Öğr. Üyesi. Muhammed Hanifi VAN
Aralık 2021, 83 sayfa

ÖZET

İklim değişikliğinin enerji tüketimindeki değişiklikleri belirlemede etkisi olabilir. İklim değişikliğinin yoğunluğunun artması, karbondioksit emisyonlarının artması küresel ısınmaya yol açmaktadır. Bu nedenle, iklim değişikliği altında ekonomik, çevre ve teknolojideki değişikliklerin enerji sisteminin geleceğini nasıl etkileyeceği konusundaki temel zorlukların belirlenmesine ihtiyaç vardır. Güç sistemi, iklim değişikliği ve ekonomi arasındaki bağlantının birçok yönü vardır ve karmaşıktır Queiroz ve diğerleri (2016). İklim değişikliğinin enerji kullanımı üzerindeki önemli bir etkisi, sıcaklığı ne kadar yükselttiği veya düşürdüğüdür. İnsanların ne kadar enerji tükettiğini belirlemeye gelince, sıcaklığın en önemli hava bileşeni olduğunu öne süren bir dizi ampirik çalışma var. Binaların soğuk havalarda ısıtılması ve sıcak havalarda soğutulması gerekir (Silhvola ve diğerleri; Sailor ve Muñoz, 1997).

Irak, iklim değişikliği ve enerji tüketimi arasındaki ilişki ile ilgili nispeten az araştırma yapılmıştır. Erbil'de elektrik, kışın ısınma ve yazın soğutma için yoğun olarak kullanılmaktadır. Ancak hissedilen hava sıcaklığı nem ve rüzgar hızıda ellektirik yükünü artırıp ya da azaltabilmektedir.

Daha soğuk veya daha sıcak hava, tropik ve subtropikal ülkelerde sıcaklık koşullandırma için yüksek bir talep getirir. İç ortam sıcaklık koşullandırmasının ana kaynağı elektrik olmaktadır. Havanın dalgalanması, sosyal ve ekonomik gibi farklı yönlerden elektrik üzerinde büyük bir etkiye sahip olabilmektedir.

Bu tezin amacı, Irak'ın Erbil kentinde elektrik tüketimi ile hava sıcaklığı, nem ve rüzgar hızı gibi iklim faktörleri arasındaki uzun dönemli ilişkiyi 2009-2019 arasındaki aylık verileri kullanarak araştırmaktır.

Verileri iki kaynaktan elde edilmiştir. Birincisi, elektrik yükü verileri Erbil elektrik müdürlüğünden elde edilmiştir. İkincisi, rüzgar hızı, sıcaklık ve nem oranı, tarım ve hava durumu genel müdürlüğünden alınmıştır. Bu çalışmada bağımlı değişken elektrik yükü, bağımsız değişkenler ise nem, sıcaklık ve rüzgar hızıdır.

Bu amaçla çalışmada ilk olarak serilerin genel seyrini görmek için serilerin grafikleri incelenmiştir. Ardından mevsimselliğin olup olmadığını anlamak için ACF ve PACF grafikleri incelenmiştir. Daha sonra serilerin durağanlıkları geleneksel birim kök testleri ile incelenmiştir. Birinci farkalarında durağan olan electricity load and wind speed serileri arasında eşbütünleşme illişkisi Jonahansen eşbütünleşme medodu ile test edilmiştir. Farklı düzeyde durağan olan elektrik yükü, rüzgar hızı, ve hava sıcaklığı ve nem oranı serileri arasındaki eşbütünleşme ilişki ARDL metodu kullanılarak test edilmiştir.

Elde edilen sonuçlara göre, elektrik yükü ve rüzgar hızı serilerin uzun dönemde birlikte hareket ettiği belirlenmiştir.

RDL modeli tahmin sonuçlarına göre ise elektrik yükü ile nem oranı ve rüzgar hızı ve hava sıcaklığı arasında uzun dönemde istatiki olarak anlamlı ilişki olduğu bulunmuştur. Bu sonuçlar ilkim değişikliğini enerji kullanımını artıracağını göstermektedir.

Anahtar kelimeler: Elektrik tüketimi, Eşbütünleşme analizi, İklim değişikliği.

1. GİRİŞ

Enerji tüketimindeki değişikliklerin belirlenmesi iklim değişikliklerinden etkilenebilmektedir. İklim değişikliğinin artan yoğunluğu, özellikle küresel ısınma, enerji tükettimi ve enerji şirketleri üzerinde ekonomik olarak olumsuz etki yapacaktır. Bu nedenle, iklim değişikliği altında ekonomik, çevre ve teknolojideki değişikliklerin enerji sisteminin geleceğini nasıl etkileyeceği konusundaki temel zorlukların belirlenmesine ihtiyaç bulunmaktadır.

Erbil şehri, 1991 yılındaki Körfez savaşında ve 1992-1996 yılları arasında yaşanan iç savaştan dolayı elektrik arzı büyük zarar görmüştür. İstasyonların, iletim hatlarının ve dağıtım istasyonlarının çoğu imha edilmiştir. Bu süreçte ana elektrik kaynağı Süleymaniye ilindeki Dokan ve Derbandikhan hidroelektrik santrallerinden sağlanmaktaydı (Kadir, 2020).

Elektrik sistem yükü, tüm tüketicilerin, aynı zamanda elektrik yükünün toplamını ifa etmektedir. Sistem özelliklerinin iyi anlaşılması, makul tahmin modelleri tasarlamaya ve farklı durumlarda çalışan uygun modelleri seçmeye yardımcı olur. Erbil şehri kışın çok soğuk, kışın çok sıcaktır. Çoğu insan, ana soğutma ve ısıtma kaynakları için elektriğe güveniyor.

Bu tezin amacı, genel olarak 2009 ve 2019 yılları arasındaki zaman serisi verileri için eşbütünleşme analizini kullanarak Irak'ta iklim değişikliği ve elektrik tüketimi arasındaki ilişkiyi değerlendirmektir.

2. METODOLOJİ

Bu bölüm, çalışmamızda uygulayacağımız kapsamlı metodolojilere genel bir bakış sunmaktadır. yapısal kırılmalar, vektör otokorelasyon modeli, eşbütünleşme analizi ve ARDL yöntemine genel bir bakış sağladık.

2.1. Zaman Serisi

Zaman serisi, zaman boyutunun yıllık, aylık, haftalık ve günlük aktiviteyi belirlemek için, zaman içinde yapılan gözlemlerin bir koleksiyonudur. Ekonomik, sosyal, çevresel, ticari/iş, tıbbi ve biyolojik alanların hepsinin çok sayıda zaman serisi kullanılmaktadır. Zaman serisi verilerini toplamak için teorik olarak farklı iki yöntem vardır. İlk yöntem, seyrek veya periyodik olarak meydana gelebilecek belirli bir zaman süreci için verileri gözlemlemektir. Bu bir bölge için bir zaman serisi de olabilmektedir. İkinci yöntem, zaman aralığında verileri sürekli olarak kaydetmektir. Verilerin özellikleri hakkında çıkarımlarda bulunmak için zaman serisi analizi yapmak için çok sayıda teknik kullanılmaktadır (Falk, 2012).

2.1.1. Trent

Trent, kısa vadeli dalgalanmaları göz ardı ederek, yükselmeye veya düşmeye yönelik genel olarak yumuşak bir eğilimdir. Artış ve azalış hareketi, doğrusal olmayan zaman serisinin eğilimi ile tanımlanabilir. Mevsimsel dalgalanma yıldan yıla olabilir (Hyndman ve Athanasopoulos, 2018).

2.1.2. Mevsimsellik

Mevsimsellik, yıldan yıla tekrarlanan, yılın belli bir zamanı ile ilgili bir varyasyon modeli olarak tanımlanabilir. Mevsimsellik deseni günlük, haftalık, aylık veya yıllık olarak tekrarlanabilir. Zaman serisi verilerinde mevsimselliği belirlemek için farklı ölçekler kullanılarak grafiksel inceleme olan basit bir yöntem kullanılabilir. Modelde daha iyi tahmin değerine sahip olduğu tespit edildiğinde mevsimselliği dahil etmenin birçok yolu vardır. Bu, modele mevsimsellik terimi eklenerek yapılabilir (Hyndman ve Athanasopoulos, 2018).

2.1.3. Otokorelasyonun ve kısmi otokorelasyonun

Zaman serilerinde, otokorelasyon fonksiyonu bir zaman serisi verilerinin gecikmeli değerleri arasındaki doğrusal ilişkiyi hesaplar. Bu, zaman serisinin, gecikmeler ve zaman serileri arasındaki benzerlik derecesini ölçen gecikmeli bir değerle doğrusal olarak ilişkili olduğu anlamına gelir. Otokorelasyonun önemli olmasının nedeni, bazı yöntemler kullanılarak zaman serilerinin tahmin edilmesinin artıklarda otokorelasyon olmadığı varsayımını gerektirmesidir (Hyndman ve Athanasopoulos, 2018).

Otokorelasyonun matematiksel denklemi aşağıdaki gibi yazılabilir:

$$r_k = \frac{\sum_{t=k+1}^{T}(y_t-\bar{y})(y_{t-k}-\bar{y})}{\sum_{t=1}^{T}(y_t-\bar{y})^2} \tag{3.1}$$

Birim kök

Birim kök, bir veya daha fazla karakteristik kök içeren bir dizi olarak tanımlanır. Herhangi bir zaman serisinde, bir birim kök, stokastik bir eğilim olarak ifade edilebilir. Örneğin AR (1) süreci aşağıdaki gibi;

$$z_t = \emptyset_1 z_{t-1} + u_t$$

$\emptyset_1 = 1$, ise, $\mathbf{u_t}$ ne zaman beyaz gürültü olur. $\mathbf{z_t}$'nin birim kökü olduğunu söyleyebiliriz. Rassal yürüyüş modeli olarak bilinen birim kök süreci durağan değildir. Zaman serisi verileri, fark alma kullanılarak durağanlaştırılabilir.

2.1.4. Dickey Fuller

Dickey Fuller'ın testinde birim köke sahip AR modelinin sıfır hipotezi test edilir. Testin parametre tahmini, t-testi kullanılarak elde edilebilir. Test, Dickey ve Fuller tarafından ilk olarak AR'de birim kökün varlığını inceledikleri için 1979 yılında geliştirilmiştir.

2.1.5. Artırma Dickey Fuller (ADF)

ADF, Dickey Fuller testinin seri korelasyonun yer aldığı genişletilmiş genişletilmiş bir versiyonudur. Ayrıca daha fazla ve büyük karmaşık modeller için ayarlanmıştır. DF testinden farklı olarak ADF testi, daha yüksek AR modeli için kullanılabilir. Test, zaman serilerinin durağanlığını test etmek için kullanılabilir. Testin prosedürü Dickey Fuller'a benzerdir (Said ve Dickey, 1984).

2.1.6. Philip Peron

Birçok birim kök testi Phillips ve Perron (1988) tarafından geliştirilmiştir. Günümüzde test, özellikle finansal verilere sahip olunması durumunda zaman serisi analizinde yaygın olarak kullanılmaktadır. PP testi, ADF testinden özellikle PP testinin hatalardaki değişen varyans ve seri korelasyonu ele alma biçiminden farklıdır. Özellikle, hataların ARMA yapısını tahmin etmek için parametrik bir otomatik regresyon kullanılabilir.

2.2. VAR Modeli

Sims, çok değişkenli zaman serilerinin karşılıklı bağımlılığını ve dinamiklerini yakalamak için kullanılabilecek bir Vektör oto-regresif (VAR) modelleri önermiştir. AR modelleri yalnızca bir Z_t zaman serisini dikkate alırken, VAR birden çok zaman serisini modeller. Tek değişkenli AR modellerinin genelleştirilmesi veya tek değişkenli zaman serisi modelleri ile paralel denklem modelleri arasında bir karışım olarak bilinir. Ana odak noktaları parametre tahmini ve model spesifikasyonu üzerindedir K boyutlu çok değişkenli bir zaman serisi ele alındığında $\boldsymbol{Z_1} \ldots \boldsymbol{Z_t}$ with $\boldsymbol{Z_t}$ = ($\boldsymbol{Z_{1t}}$, …, $\boldsymbol{Z_{kt}}$). Bu VAR modeli genelleştirilebilir. P gecikmeli VAR modeli:

$$\boldsymbol{Z_t = v + \emptyset_1 Z_{t-1} + \emptyset_2 Z_{t-2} + \cdots + \emptyset_p Z_{t-p} + \mathbf{u}_t} \quad (3.38)$$
$$\boldsymbol{t = 1, 2, 3, \ldots, T}$$

Olarak ele alınmaktadır. Burada $\boldsymbol{Z_t}$= $(\boldsymbol{Z_{1t}}, \boldsymbol{Z_{2t}}, \ldots, \boldsymbol{Z_{kt}})'$ $(k \times 1)$ boyutlu zaman serisi değişkeni olarak ele alınmaktadır. . v, $(k \times 1)$ boyutlu vektörün sabitidir (Lütkepohl, 2005).

2.3. Eşbütünleşme Analizi

Eşbütünleşme analizi, birkaç zaman serisi arasında uzun dönemli ilişkinin durağan olup olmadığını test etmeye odaklanır. Bu uzun dönem, eşbütünleşme fikri kullanılarak ölçülebilir ve test edilebilir. Birkaç zaman serisi veri setinin uzun dönem dengesine sahip olması, birlikte hareket etmesi ve ortak stokastik eğilimi paylaşması durumunda eşbütünleşme meydana gelebilir. Eşbütünleşmeyi test etmek için birden fazla eşbütünleşme vektörünün olup olmadığı, eşbütünleşme vektörünün bilinip bilinmediği ve eşbütünleşme ilişkilerinde yapısal kırılma olup olmadığı dikkate alınabilir. Eşbütünleşme testi, belirli bir zaman diliminde birkaç veri setinin duyarlılık derecesini belirlemek için kullanılabilir (Erica, 2020).

Marseet ve Khadija'ya (2015), iki veya daha fazla seri arasındaki uzun dönemli eşbütünleşmeyi test etmek için çeşitli yaklaşımlar kullanılabilir. Değişkenler arasındaki uzun dönemli ilişki, eşbütünleşme analizi kullanılarak değerlendirilebilir. Değişkenler arası eşbütünleşme, değişkenler arasında uzun dönemli veya denge ilişkisinin var olduğunu gösterir. Bir zaman serisi durağan değilse, ancak zamanla birlikte hareket edebilir ve farkları durağan olabilir. Değişkenler arasındaki uzun dönemli bağlantı, sistemin zaman içinde kapsadığı anlamına gelir ve hata teriminin zaman içindeki dengesizliği olarak yorumlanabilir. Eşbütünleşme ile çok değişkenli zaman serisi eğilimlerini incelemek ve değişkenler arasında uzun dönem ve kısa dönem ilişkileri için bir teknik sağlamak için ortaya çıkmıştır (Marseet ve Khadija, 2015).

2.3.1. Johansen (VECM)

VECM'ye dayalı Johansen testi, maksimum Eigen değeri ve izl testi olmak üzere iki testten oluşur. Bu testler VECM'ye dayalı uzun dönemli ilişkiler için kullanılabilir. VECN modelleri, sabit veya trendli veya sabit olmayan model ve farklı sayıda eşbütünleşme vektörü gibi farklı varsayımlar kullanılarak tahmin edilebilir. Daha sonra modeller olabilirlik oranı testi kullanılarak karşılaştırılabilir. İz testi, k vektörlü kısıtlanmamış VECM'ye karşı kısıtlı bir VECM'nin olabilirlik oranı testi kullanılarak elde edilebilir (Frans, 1998).

2.3.2. Otoregresif dağıtılmış gecikme (ARDL)

Johansen prosedürünü kullanarak eşbütünleşmeyi test etmek, tüm bütünleşik değişkenlerin aynı düzeyde durağan olmasını gerektirir. Bu değişkenler her seferinde aynı düzeyde durağan olmayabilir. Bu nedenle, 2001 yılında Pesaran ve diğerleri tarafından alternatif bir yaklaşım geliştirilmiştir. Bu yöntemde değişkenlerin aynı düzeyde entegre edilmesi gerektiği varsayımı gerekli değildir. Johnsen'den farklı olarak, yöntem küçük örneklem olması durumunda daha uygundur. Tek bir denklemi olduğu için bu yöntemi uygulamak ve yorumlamak oldukça basittir (Peseran ve diğerleri, 2001).

ARDL (p, q) modeli:

$$\mathbf{y_t = a_0 + a_1 + \sum_{i=1}^{p} \emptyset_i y_{t-i} + \beta' x_t + \sum_{i=0}^{q-1} \beta_i^{*'} \Delta x_{t-i}} + \boldsymbol{u}_\mathbf{t} \tag{3.46}$$

$$\mathbf{\Delta x_t = P_1 \Delta x_{t-1} + P_2 \Delta x_{t-2} + \cdots + P_s \Delta x_{t-s}} + \boldsymbol{u}_\mathbf{t} \tag{3.47}$$

olarak gösterilebilir.

Burada $\boldsymbol{x_t}$ k boyutlu I (1) eşbütünleşik olmayan bir değişken, $\boldsymbol{P_i}$ genelleştirilmiş otoregresif bir süreci gösteren parametredir.

2.4. Sonuçlar

Veri setinin özelliklerini belirlemek için veriler için bazı tanımlayıcı istatistikler tablo 1'de verilmiştir.

Tablo 4.1. Elde edilen verilerin tanımlayıcı istatistikleri

Değişken	Açıklama	N	Ortalama	Std. Dev.	Min.	Maks.
Toplam elektrik yükü		132	23451.36	8645.417	53880.58	53880.58
Nem oranı		132	38.79517	7.273610	13.90189	62.58136
Sıcaklık		132	21.96987	1.466530	16.69333	27.20889
Rüzgar hızı		132	6.165348	3.595883	1.375000	17.54000

Tablo 4.1, 10 yılda ortalama elektrik yükünün 23.451 KW olduğunu, ortalama nem oranının ise 38.8 olduğunu göstermektedir. Erbil'de sıcaklık ortalaması 10 yıldır 22'nin altında.

2.4. Birim Kök Testi

Birim kökü test ederek zaman serisi verilerinin durağanlığını araştırılmıştır. Birim kökü test etmek için Artırılmış Dickey-Fuller (ADF) ve Phillip Perron (PP) test sonuçları sırasıyla tablo 4.6 ve tablo 4.7 de verilmiştir.

Tablo 4.6. Artırılmış Dickey-Fuller (ADF) birim kök testi sonuçları

Değişken	Sabitli		Yok	
	Level	1st Difference	Level	1st Difference
T_Loadsa	-1,66	-9,95***	0,75	-9,83***
O_Rhsa	-7.93***	-13,76***	-0,43	-13,81***
o_Tcsa	-7,61***	-11,58***	-0,28	-11,63***
m_Ws	-1,37	-10,69***	-0,15	-10,71***

Not: ***, ** ve * Yüzde %1, %5 ve %10 düzeylerinde istatistiksel olarak anlamlıdır.

Table 4.7: Philip Perron test results

Variable	Intercept		None	
	Level	1st Difference	Level	1st Difference
T_Loadsa	-2,59	-18,47***	0,33	-17,74***
O_Rhsa	-3,06*	-9,07***	-0,28	-9,09***
o_Tcsa	-7,60***	-33,18***	-0,61	-33,07***
m_Ws	-1,64	-10,34***	-0,41	-10,37***

Not: İstatistiksel anlamlılık ***, ** ve * ile %1, 5 ve %10 oranlarında gösterilmiştir.

Rüzgar hızı, elektrik yükü, sıcaklık ve nem oranı gibi değişkenlerin durağanlığı araştırılmıştır. Tüm değişkenler için birim kökü ayrı ayrı test etmek için Augmented Dickey Fuller yaklaşımını kullandık. Temel hipotez, her değişkenin birim köke sahip olduğudur. Sabitli modelde birim kök testleri, tüm değişkenlerin durağan olmadığını göstermektedir. Birim kök testinden iki serinin durağan olmadığı sonucuna varılmıştır, bu yüzden durağan olmayan bu iki seriyi birinci farkını alarak durağan hale getirilmiştir.

Tüm değişkenler için ayrı ayrı birim kök için yapılan Phillip Perron testinin sonucu sunulmuştur. PP için ADF ile aynı asimptotik dağılım elde edilebilir. Birim kök testleri, tüm değişkenlerin durağan olmadığını göstermektedir. Birinci farkı alındıktan sonra değişkenler durağan hale gelmiştir. Ek olarak, kesmeye izin veren PP test sonuçlarının, yalnızca nem oranı ve sıcaklık değişkenlerinin durağan olduğunu ortaya koyduğu anlaşılmaktadır. Birinci farkı alındıktan sonra, tüm değişkenler birim kök süreci sergilemektedir.

2.5. Eşbütünleşme

Rüzgar hızı ve yükü aynı seviyede durağan olduğundan aralarındaki eşbütünleşme ilişkisi tablo 4.8 dearaştırılmıştır.

Tablo 4.8. Johansen eşbütünleşme testi sonuçları

Temel Hipotez	İz İstatistik	0.05 Kritik Değer	Prob.
No Cointegration	21.47371	20.26184	0.0339
en fazla 1	4.342999	9.164546	0.3634
Boş Hipotez	**İz İstatistik**	**0.05 Kritik Değer**	**Prob.**
No Cointegration	17.13071	15.89210	0.0318
en fazla 1	4.342999	9.164546	0.3634

Johansen eşbütünleşme yöntemi uygulanarak tahmin edilen ortak değişkenler arasındaki uzun dönemli ilişki iz istatistiklerinin sonuçları, en az bir eşbütünleşme denkleminin olduğunu ortaya koymaktadır. İz istatistik değeri %5 kritik değerden büyük olduğu için temel hipotezi red edilmektedir. En fazla 1, en az bir eşbütünleşme denklemi olduğunu ve istatistiksel olarak anlamlı olduğunu ortaya koymaktadır (iz istatistikleri %5 kritik değerden düşüktür). Böylece Tablo 4.8'te verilen eşbütünleşme testi sonuçlarından; iz istatistiklerinden ve maksimum öz değerlerden seriler arasında en fazla bir eşbütünleşme olduğu görülmektedir.

2.6. ARDL

Zayıf Erojenite testlerine göre tüm seriler için dışsallık hipotezi reddedilmiştir. Bu nedenle, elektrik şehir yükü ve rüzgar hızı değerleri arasındaki uzun ve kısa vadeli (Hata düzeltme modeli) VECM (1) tahmin sonuçları verilmiştir. LM testinden elde edilen sonuçlara göre 12. gecikmeye göre %5 anlamlılık düzeyinde otokorelasyon sorunu olmadığı tespit edilmiştir. Daha sonra tüm gözlemler için hata terimlerinin varyansının sabit olup olmadığı Beyaz Varyans Testi ile incelenmiştir.

ARDL sınır testi sonuçları 1sunulmaktadır. F-testi sonucu, ilgilenilen değişkenler arasında eşbütünleşmenin var olduğunu ortaya koymaktadır (p-değeri=0.000). Tablo 4.13'te verilen ARDL (3,0,0) model tahmin sonuçlarına göre tüm katsayılar istatistiksel olarak anlamlı bulunmuş ve yapılan tanısal testler sonucunda modelin herhangi bir seriye sahip olmadığı belirlenmiştir. Korelasyon, değişen varyans, spesifikasyon ve normallik problemleri. Seriler arasında eşbütünleşme ilişkisinin varlığını araştırmak için yapılan sınır testi sonuçları verilmiştir.

Tablo 4.13. ARDL sınır testi

F-Sınır Testi				
		H_0: Boş Hipotez: Düzey ilişkisi yok		
Test istatistiği	Değer	anlam.	I(0)	I(1)
F- istatistiği	4.790	10%	3.17	4.14
K	2	5%	3.79	4.85
		1%	5.15	6.36
t- Sınır Testi				
Test istatistiği	Değer	anlam.	I(0)	I(1)
t- istatistiği	-2.344	10%	-2.57	-3.21
		5%	-2.86	-3.53
		1%	-3.43	-4.1

Tablo 9'da görülen F limit testi için F=4.78 olarak hesaplanmıştır. Bu değer sadece %10 (F > I (1)) için üst kritik değerlerden büyük olduğundan, "eşbütünleşme yok" sıfır hipotezi olacaktır. reddedildi (0,01 için üst kritik değer 6,36, 0,05 için üst kritik değer = 4,7 ve %10 için üst kritik değer = 3,97).

F sınır testine göre seriler %10'a göre eşbütünleşiktir. Ancak koşullu hata düzeltme modeli olarak model 3 kullanıldığından bu eşbütünleşmenin geçerli bir eşbütünleşme olup olmadığının test edilmesi gerekmektedir. Bu nedenle Tablo 9'un son bölümünde t-sınır testi verilmiştir), seriler arasında eşbütünleşme geçerli değildir.

3. SONUÇ

Bu tezin amacı genel olarak Irak'ta iklim değişikliği ile elektrik tüketimi arasındaki ilişkiyi 2009 ve 2019 yılları arasındaki zaman serisi verileri için eşbütünleşme analizi kullanarak değerlendirmektir. Verileri iki kaynaktan elde edilmiştir ve elde edilen değişkenler elektrik yükü, sıcaklık ve sıcaklıktır. , nem oranı ve Rüzgar hızıdır. Birim kökü test etmek için Artırılmış Dickey-Fuller (ADF) ve Phillip Perron (PP) yaklaşımlarını kullanılarak verilerin durağanlığı araştırılmıştır. ADF kullanılarak sabitsiz modelde birim kök testleri, tüm değişkenlerin durağan olmadığını sonucuna varılmıştır. Birinci fark alındıktan sonradeğişkenler durağan hale gelmiştir. ADF test sonuçları, yalnızca nem oranı ve sıcaklık değişkenlerinin durağan olduğunu ortaya koymuştur.

Birinci farkalarında durağan olan elektrik yükü and rüzgar hızı serileri arasında eşbütünleşme illişkisi Jonahansen eşbütünleşme medodu ile test edilmiştir. Farklı düzeyde durağan olan elektrik yükü, rüzgar hızı, ve hava sıcaklığı ve nem oranı serileri arasındaki eşbütünleşme ilişki ARDL metodu kullanılarak test edilmiştir. Bu sonuçlar, A. Ngah Nasaruddin ve arkadaşlarının çalışmaları ile parellik göstermiştir. Elde edilen sonuçlara göre, elektrik yükü ve rüzgar hızı serilerin uzun dönemde birlikte hareket ettiği belirlenmiştir. ARDL modeli tahmin sonuçlarına göre ise elektrik yükü ile nem oranı ve rüzgar hızı ve hava sıcaklığı arasında uzun dönemde istatiki olarak anlamlı ilişki olduğu bulunmuştur. Bu sonuçlar ilkim değişikliğini enerji kullanımını artıracağını göstermektedir.

CURRICULUM VITAE

He lives in Erbil province, Iraq. He completed primary, secondary and high school in ERBIL city. I graduated from Statistic Department taken second degree in more than one hundred student in my class in Salahaddin University - College of Administration and Economic in 2015. He started his Postgraduate study at the department of Statıstıc, Institute of Natural Applied Sciences at Van Yüzüncü Yil University in Turkey on October 2019.

VAN YUZUNCU YIL UNIVERSITY
THE INSTITUTE OF NATURAL AND APPLIED SCIENCES
THESIS ORIGINALITY REPORT

Date: 19/10/2021

Thesis Title: **IMPACT OF CLIMATE CHANGE ON RESIDENTIAL ELECTRICITY CONSUMPTION IN ERBIL: AN ECONOMETRIC ANALYSIS**

The title of the mentioned thesis, above having total 84 pages with cover page, introduction, main parts and conclusion, has been checked for originality by Turnitin computer program on the date of 19/10/2021 and its detected similar rate was 15 % according to the following specified filtering

Originality report rules:

- Excluding the Cover page,
- Excluding the Thanks,
- Excluding the Contents,
- Excluding the Symbols and Abbreviations,
- Excluding the Materials and Methods
- Excluding the Bibliography,
- Excluding the Citations,
- Excluding the publications obtained from the thesis,
- Excluding the text parts less than 7 words (Limit match size to 7 words)

I read the Thesis Originality Report Guidelines of Van Yuzuncu Yil University for Obtaining and Using Similarity Rate for the thesis, and I declare the accuracy of the information I have given above and my thesis does not contain any plagiarism; otherwise I accept legal responsibility for any dispute arising in situations which are likely to be detected.

Sincerely yours,

19/10/2021

Name and Surname: Zewar Omar Ismael

Student ID#: 19910001260

Department: Statistics

Program: M. Sc. **X** Ph.D. ☐

APPROVAL OF SUPERVISOR
SUITABLE

APPROVAL OF THE INSTITUTE
SUITABLE

Asst. Prof. Dr. Muhammed Hanifi VAN