J. Stat. Appl. Pro. **12**, No. 3, 941-959 (2023)

Journal of Statistics Applications & Probability
An International Journal

http://dx.doi.org/10.18576/jsap/120306

Forecasting Emissions of Carbon Dioxide (CO_2), Methane (CH_4) and Energy Consumption in Egypt Using VECM and ARIMAX Models

Hanaa H. A. Aboul Ela

Department of Statistics, Mathematics and Insurance, Faculty of Business, Ain Shams University, Cairo, Egypt

Received: 6 Nov. 2022, Revised: 22 Dec. 2022, Accepted: 7 Feb. 2023
Published online: 1 Sep. 2023

Abstract: Greenhouse gas emissions are one of the important environmental problems in Egypt that do not harm only humans, but also contribute to climate changes all over the world. The emissions of carbon dioxide (CO_2) and methane (CH_4) are the most important of these emissions. The decision makers seek to use renewable energies to reduce greenhouse gas emissions. Therefore, this paper aims to measure the factors affecting CO_2 and CH_4 emissions in Egypt during the period from 1980 to 2019 and to predict of these emissions and energy sources from 2020 to 2030. The study applied the Vector Error Correction Model (VECM) and Autoregressive Integrated Moving Average with Exogenous variables (ARIMAX) models. The study results found that the most influential variables on CO_2 gas emissions are energy consumption, gross domestic product, and international trade. It was also found that livestock production, energy consumption and agricultural fertilizers are the most influential variables on CH_4 emissions. It was also found that the predictability of VECM is better than the ARIMAX model, so we can use it to predict emissions of CO_2 and CH_4.

Keywords: ARIMAX Model, Climate Changes, Energy, Greenhouse Emissions, VECM.

1 Introduction

The sustainable development is the main goal of all countries of the world, and the energy consumption is the main factor for all economic sectors because it helps in achieving economic development. Most of the world's countries use traditional energy based on fossil sources such as oil, natural gas and coal, which pollute the environment and cause emissions of harmful gases to society. For these reasons, international organizations have held many conferences aimed at the need for governments to commit to implementing their promises in achieving sustainable development. Especially the non-renewable energies began to dry up and it was necessary to look for renewable sources of energy such as solar energy, wind energy and hydroelectric energy that preserve the environment and achieve sustainable development. The importance of this study is to choose the best model to predict greenhouse gas emissions in Egypt in the future. This helps in taking the necessary procedures to mitigate climate change and achieve the visions of the Sustainable Development Strategy 2030. Therefore, the current study aims to measure the impact of energy consumption sources and other economic factors on CO_2 and CH_4 emissions in Egypt using the VECM in the short and long term. In addition, to analysis of the causal relationship and analysis of the components of variation for greenhouse gas emissions, then predict these emissions after comparing the VECM & ARIMAX models. To achieve the objectives of the study, it seeks to test the following hypotheses: There is a significant a statistical relationship between CO_2 emissions and sources of energy consumption, gross domestic product, foreign direct investment. There is also an insignificant statistical relationship with commodity trade and agricultural production. There is significant a statistical relationship between CH_4 emissions and sources of energy consumption, livestock production, and consumption of agricultural fertilizers. There is also an insignificant statistical correlation with Gross Domestic Product (GDP). There is a two-way causal relationship between the CO_2 gas emissions variable and the following variables (sources of energy consumption, gross domestic product and agricultural production index), as well as there is a one-way causal relationship between the CH_4 gas emissions variable and the following variables (sources of energy consumption and Livestock Production Index, agricultural fertilizer

* Corresponding author e-mail: dr.hanaaaboualaela@bus.asu.edu.eg

consumption). The predictability of the VECM model is expected to be better than that of the ARIMAX model when predicting emissions of CO_2&CH_4.

2 Literature Review

Bayar et al. [4] studied the impact of economic growth, financial development, and energy consumption on CO_2 emissions. This study was in European Union countries during the period (1995 - 2017) using the Co-integration model and Granger causality analysis. The study found a two-way relationship between economic growth and energy consumption. It was also found a two-way relationship between economic growth and CO_2 emissions. In addition, it was found that the development of the financial sector and energy consumption have a positively affect carbon dioxide emissions in the long run. Bekun et al. [5], Bilgili et al. [6], and Jabil et al. [14] found that the renewable energy reduces CO_2 gas emissions in seventeenth Organisation for Economic Co-operation and Development (OECD) countries over the period 1977 to 2010 using the Full Modified Ordinary Least Squares (FMOLS) model and the Dynamic Ordinary Least Squares (DOLS) model, As well as the same models were used in twenty five Organisation for Economic Co-operation and Development (OECD) countries over the period1980-2010. The Panel Autoregressive Distributive Lag (PARDL) was also used with the pooled mean group (PMG) estimator in sixteenth European Union (EU) countries over the period 1996 to 2014. Garip and Oktay [12] introduced a study aimed to predict the CO_2 Emissions in Turkey by using machine learning methods. The study also, showed that the support vector machine model is better than the random forest model to predict the CO_2 Emission in Turkey. Nyoni et al. [21] used the Box -Jenkins methodology to predict CO_2 emissions in China during the period 1960 to 2017. The study also showed that the autoregressive integrated moving average (ARIMA $(1,2,1)$) model is the most appropriate model to predict the total CO_2 emissions in China over the next ten years. The CO_2 emissions in China in 2024 will be expected to be ten million kilo tonnes. These emissions are also expected to increase, further climate changes and global warming. Khobai [18] studied the impact of renewable energy consumption on economic growth expressed in Gross Domestic Product (GDP), CO_2 emissions, employment and fixed capital composition. This study was applied in Indonesia over the period 1990 to 2014 using the Auto Regressive Distributed Lagged (ARDL) model and Vector Error Correction Model (VECM). The study showed a direct relationship between economic growth and renewable energy consumption in the short and long term. Anwar [3] applied the panel data model in East Asian countries during the period 1980 to 2017 and found that the best model is the fixed effect model. The study found that urbanization, economic growth and trade openness are the most influential on carbon dioxide emissions. The study recommended adopting policies to reduce CO_2 emissions and encourage sustainable urbanization because it helps in the economic growth. Alam and Alarjani [2] applied the Artificial Neural Networks (ANNs), Holt and Winters Exponential Smoothing (HWES), and ARIMA models. The ARIMA $(2,1,2)$ model was found to be suitable for predicting the CO_2 emissions in Saudi Arabia. Rahman and Hasan [24] have developed different autoregressive integrated moving average (ARIMA) models to predict the carbon dioxide emissions by using forty-four-year time series data from 1972 to 2015 . Based on the results it was found that the ARIMA $(0,2,1)$ model is the best suitable model for forecasting carbon dioxide emissions in Bangladesh. The literature review has focused on studying the impact of energy consumption and some economic variables on CO_2 gas emissions, and few studies have been interested in studying the impact of these variables on CH_4 gas emissions as one of the most important greenhouse gas emissions affecting global warming and climate changes in Egypt. Also, we rarely find a study that compared linear models such as VECM and ARIMAX, choosing the best of them to predict emissions of CO_2 gas and CH_4 gas emissions.

3 Materials and Methods

This paper is based on the published data on the amount of petroleum consumption (PC), the amount of Natural Gas Consumption (NGC), the amount of Coal Consumption (CC), Foreign Direct Investment (FDI), Gross Domestic Product (GDP), Commodity Trade (CT) volume, Agricultural Production Index (API), Livestock Production Index (LPI), Agricultural Fertilizer Consumption (AFC), Carbon Dioxide (CO_2), Methane Emissions (CH_4) over the period 1980 to 2019. The data was obtained from the World Bank (WB) Group [25], United States (US) Energy Information Administration (EIA) [26] and Food and Agricultural Organization (FAO) [11]. The statistical analysis will be done using EVIEWS12 software after using normal logarithm for original data. The following equations represent the CO_2 and CH_4 emissions models as follows:

$$CO_{2t} = \alpha_0 + \alpha_1 PC_t + \alpha_2 NGC_t + \alpha_3 CC_t + \alpha_4 GDP_t + \alpha_5 FDI_t + \alpha_6 CT_t + +\alpha_7 API_t + \varepsilon_{1t}$$
$$CH_{4t} = \beta_0 + \beta_1 PC_t + \beta_2 NGC_t + \beta_3 CC_t + \beta_4 GDP_t + \beta_5 LPI_t + \beta_6 AF_t + \varepsilon_{2t}$$

where α_i&β_j are constants, ε_{1t}&ε_{2t} are errors.

The study will use the multivariate linear models such as the VECM and ARIMAX models as follows:

3.1 Vector Error Correction Model (VECM)

The first step in this model is using the Augmented Dickey Fuller (ADF) tests Dickey and Fuller [8] and Phillips-Perron (PP) [23] tests to know whether the study variables have the unit root that causes spurious regression, or they are stationary. Adkins and Hill [1].

3.1.1 Unit Root Tests

–Augmented Dickey- Fuller (ADF) Test

The ADF test depends on the following three models by least squares method as follows:

$$\triangle X_t = pX_{t-1} - \sum_{j=2}^{P} \phi_j \triangle X_{t-j+1} + \varepsilon_t$$

$$\triangle X_t = pX_{t-1} - \sum_{j=2}^{P} \phi_j \Delta X_{t-j+1} + c + \varepsilon_t$$

$$\triangle X_t = pX_{t-1} - \sum_{j=2}^{P} \phi_j \Delta X_{t-j+1} + c + bt + \varepsilon_t$$

where ΔX_t reflects the first differentiation of variable X in year t. P is the number of lags to delete the autocorrelation of the random error ε_t. The ADF test depends on two hypotheses $H_0 : \phi - 1 = 0$ and $H_1 : \phi - 1 \neq 0$. The ADF test statistic is calculated from the following formula. Dickey and Fuller [9].

$$T_{cal} = \frac{\hat{\phi}_1 - 1}{SE\left(\hat{\phi}_1\right)}$$

where $\hat{\phi}_1$ is the least square estimator. $SE\left(\hat{\phi}_1\right)$ is the standard error estimate. If $T_{\text{cal}} \geq T_{\text{tab}}$ it means rejecting the null hypothesis H_0 and we accept the alternative hypothesis H_1 that means the time series is stable. If $T_{\text{cal}} < T_{\text{tab}}$ it means accept the null hypothesis H_0 and we reject the alternative hypothesis H_1 that means the time series is unstable.

–Phillips - Perron (PP) Test.

The Phillips - Perron test includes the conditional variance of errors. It allows for the elimination of biases that result from random volatility. Phillips-Perron statistic calculates from the following equation. Philip and Dick [22].

$$t_h^* = \sqrt{k} \times \frac{\left(\hat{\phi}_1 - 1\right)}{\hat{\sigma}_{\hat{\phi}_1}} + \frac{n(k-1)\hat{\sigma}_{\hat{\phi}_1}}{\sqrt{k}}$$

where $k = \frac{\sigma^2}{{s_t}^2}$ and k $= 1$ if the series of residuals (e_t) constitute white noise, then this statistic will compare with critical values in (Mackinnon) table.

3.1.2 Johansen Co-integration Tests

The co-integration test is used if the time series data is stable, and its degree of integration is one. There are many tests for co-integration but the Johansen test. Johansen [15] is broader than the methodology applied in the Engle and Granger [10] test because it allows to determine the number of long-term balance relationships between several integrated variables of the same degree. Its steps will be shown as follows:

–The residuals $\hat{v}_t \& \hat{u}_t$ are calculated through the following two models:

$$\nabla Y_t = \hat{A}_0 + \hat{A}_1 \nabla Y_{t-1} + \hat{A}_2 \nabla Y_{t-2} + \ldots . + \hat{A}_p \nabla Y_{t-p} + \hat{v}_t$$

$$\nabla Y_{t-1} = \hat{A}_0^v + \hat{A}_1^v \nabla Y_{t-1} + \hat{A}_2' \nabla Y_{t-2} + \ldots . + \hat{A}_p \nabla Y_{t-p} + \hat{u}_t$$

$$\text{where} \quad \boldsymbol{Y}_t = \begin{bmatrix} \boldsymbol{Y}_{1,t} \\ \boldsymbol{Y}_{2,t} \\ \cdots \\ \cdots \\ \boldsymbol{Y}_{\boldsymbol{k},t} \end{bmatrix}$$

$\hat{v}_t \& \hat{u}_t$ are the residuals matrices with of rank(k,n), where k is the number of variables, and n is the number of observations.

–Calculate the covariance from the following four matrices which its rank (k,k).

$$\hat{\sum_{vv}} = \left(\frac{1}{n}\right)\sum_{t=1}^{T} v_t v_t', \quad \sum_{uv} = \left(\frac{1}{n}\right)\sum_{t=1}^{T} u_t v_t', \quad \hat{\sum_{uu}} = \left(\frac{1}{n}\right)\sum_{t=1}^{T} u_t u_t', \quad \sum_{vu} = \left(\frac{1}{n}\right)\sum_{t=1}^{T} v_t u_t'$$

K is calculated as an eigenvalue of the following matrix M of rank (k,k) as follows :

$$M = \hat{\sum}_{vv}^{-1}\hat{\sum}_{vu}\hat{\sum}_{uu}^{-1}\hat{\sum}_{uv}$$

Johansen and Juselius [16] proposed two tests to determine the number of co-integration vectors, the first test is the trace test which test the hypothesis that say there are at most q co-integration vectors versus the general unrestricted model r = q. The probability ratio statistic for this test is calculated from the following relationship:

$$\lambda_{\text{trace}} = -n\sum_{i=r+1}^{P} \ln\left(1-\hat{\lambda}_i\right)$$

where $\lambda_{r+1,..........}\lambda_p$ are the smallest values of the eigenvectors $p-r$, there are a number of cointegration vectors less than or equal to r. The second test is the Max-Eigen test, which calculates its statistic from the following relationship:

$$\lambda_{\max}(r, r+1) = -T\ln(1-\lambda_{r+1})$$

The null hypothesis assumes that there is r of co-integration vectors versus the alternative hypothesis assumes that there is of $r+1$ of the co-integration vectors.

3.1.3 Vector Error Correction Model (VECM) Estimation

The vector error correction model is used to adapt the behaviour of the variables in the short term and the long term. The general equation of a VECM (p) where p is the lag of endogenous variables with the rank of co-integration r $\leq$ k is as follows. Lütkepohl [20].

$$\Delta y_t = \Pi y_{t-1} + \sum_{i-1}^{p-1} \Gamma_i \Delta y_{t-1} + D_t + \varepsilon_t$$

where Δ is the differences factor, $\Delta y_t = y_t - y_{t-1}, y_{t-1}$ is the vector of variable endogenous with one lag, ε_t is the vector of residuals with rank $\mathrm{k}\times 1, D_t$ is the vector of constant with rank $\mathrm{k}\times 1, \Pi$ is the matrix of coefficient of co-integration where $\Pi = \alpha\beta \mathrm{t}, \alpha$ represents the modification vector by rank $(\mathrm{k}\times \mathrm{r}), \beta$ represents the co-integration matrix by rank $(\mathrm{k}\times\mathrm{r})$, also it represents the long term parameter , and Γ_i is the coefficient matrix of the endogenous variables with rank $\mathrm{k}\times\mathrm{k}$.

3.1.4 Causality Granger Test

The economic theory assumes that there is a causal relationship between the independent variables and the dependent variable , but there are economic relationships in which the theory has not clarified the direction of causality, and therefore the Granger test is used to ensure that there is an interrelationship between the economic variables to be studied, and we consider the autoregressive model which is consisting of the two variables Y_{1t}, Y_{2t} and the degree of lag is (p) Granger [13] .

$$\begin{pmatrix} Y_{1t} \\ Y_{2t} \end{pmatrix} = \begin{pmatrix} a_0 \\ b_0 \end{pmatrix} + \begin{pmatrix} a_1^1 b_1^1 \\ a_1^2 b_1^2 \end{pmatrix}\begin{pmatrix} Y_{1t-1} \\ Y_{2t-1} \end{pmatrix} + \ldots\ldots + \begin{pmatrix} a_p^1 b_p^1 \\ a_p^2 b_p^2 \end{pmatrix}\begin{pmatrix} Y_{1t-p} \\ Y_{2t-p} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix}$$

where $a_0, b_0, a_1^1, a_1^2, b_1^1, b_1^2$ are the parameters of the model, $\varepsilon_{1t} \& \varepsilon_{2t}$ represent an errors. The variable Y_{2t} does not cause changes in the variable Y_{1t} if the following hypothesis accepts $H_0 : b_1^1 = b_2^1 \ldots\ldots = b_p^1 = 0$, and the variable Y_{1t} does not cause changes in the variable Y_{2t} if the following hypothesis accepts $H_0 : a_1^2 = a_2^2 \ldots\ldots = a_p^2 = 0$ but if the two hypotheses are accepted together, the variable Y_{1t} causes changes in the variable Y_{2t} and at the same time Y_{2t} causes changes in the variable Y_{1t}, that means there is a two-way causal relationship which is called feedback effect between $Y_{1t} \& Y_{2t}$.

3.1.5 Variance Decompositions Analysis

This analysis aims at calculating the contribution of the error variation of each variable for a certain period in the prediction error variation. To know the ratio of each variance, we will divide this variance to the total variance of the prediction error. After the shocks become normal, the response analysis will be occurred.

3.2 Autoregressive Integrated Moving Average with Exogenous Variables (ARIMAX) Models

The Box-Jenkins methodology is a method for finding the appropriate model for estimating time series values using the Autocorrelation Function (ACF) and the Partial Autocorrelation Function (PACF). The model consists of three main parts which are the Auto Regressive AR (p) rank, the Integration I(d) rank and the Moving average MA (q) rank. The Autoregressive Integrated Moving Average (ARIMA) model is used for unstable time series then we prepare a stable Auto Regressive Moving Average (ARMA (p, q)) model with finding the difference for the time series to develop the ARIMA (p, d, q) model which we can express about it as follows:

$$y_t = \sum_{i=1}^{p} \phi_i \beta^d y_{t-i} + \sum_{j=0}^{q} \theta_j w_{t-j}$$

The ARIMAX model is adopted by including an X variable with the ARIMA model, which is an exogenous variable to improve the accuracy of forecasting, that is more applicable for time series with sudden changes in trends. The process of ARIMA (p, d, q) model with the presence of the previous values (m) for an external variable x_i we develop the ranks of the ARIMAX model to become ARIMAX (p, d, q, m) and is represented by the following equation. Kaur and Rakshit [17].

$$y_t = \sum_{i=1}^{p} \phi_i \beta^d y_{t-i} + \sum_{j=0}^{q} \theta_j w_{t-j} + \sum_{k=1}^{m} \lambda_k e_{t-k}$$

where w_j represents the white noise. ϕ_i, θ_j and λ_k represent the coefficients of the autoregressive, moving average and exogenous variables, respectively. In the following points we can explain stages of time series analysis according to the methodology of Box et al. [7] as follows:

–Identification. This stage is one of the most important stages of the forecasting process according to the Box-Jenkins methodology, where the ranks (p, d, q) of the ARIMA model are determined using the Autocorrelation Function (ACF) and the Partial Autocorrelation Function (PACF).

–Estimation. After the identification stage. There are several methods that it can be used to estimate the parameters of the Auto Regressive AR(P) model such as Ordinary Least Squares (OLS), Maximum Likelihood (ML) and Walker -Yule equations method.

–Diagnostic Testing. At this stage, it is necessary to make sure that the assumptions of the ARIMA model are available, including that the residuals represent independent random changes with an average of zero and constant variance, and the Ljung-Box test [19]. can be used to know whether this assumption is correct or not, as in the following equation:

$$LB = n(n+2) \sum_{k=1}^{m} \frac{1}{n-k} \rho_k^2$$

–Forecasting. At this stage, the expected values of the dependent variable will be known. There are many criteria for choosing the optimal forecasting model. These criteria are root mean square errors (RMSE), Theil's Inequality Coefficient (TIC), Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE).

4 Results and Discussion

4.1 Vector Error Correction Model (VECM)

The figure 1 shows a summary of steps for vector error correction Model as follows:

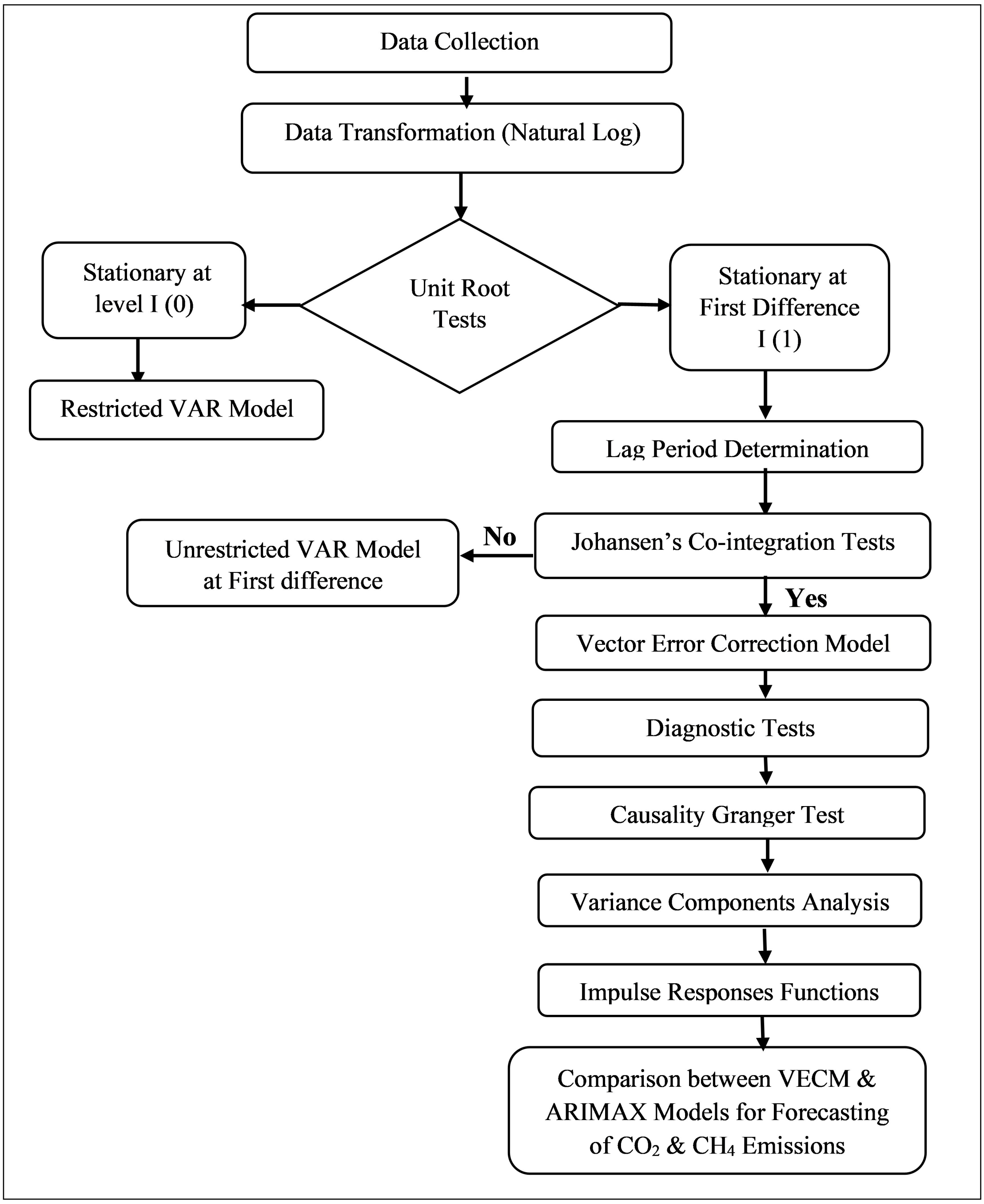

Fig. 1: Steps of VECM Analysis for Forecasting of CO_2&CH_4 Emissions

4.1.1 Unit Roots Tests

The table 1 shows the results of stability tests of the time series for the following variables (CO_2, CH_4, PC, NGC, CC, GDP, FDI, TC, AGPI, LPI, AF), which are not stable at the Level, but they are stable at the 1^{st} difference. This means that they suitable to estimate VECM and ARIMAX models.

Table 1: Results of unit root tests (ADF and PP)

		ADF				PP			
		Level		1st diff.		Level		1st diff.	
		Cons.	Cons. & Tre.	Cons.	Cons. & Tre.	Cons.	Cons. & Tre.	Cons.	Cons. & Tre.
CO_2	t-Stat.	-4.09	-4.11	-4.32	-4.72	-2.99	-3.75	-4.32	-4.72
	Prob.	0.00	0.01	0.00	0.00	0.04	0.03	0.00	0.00
		(***)	(**)	(***)	(**)	(**)	(**)	(***)	(***)
CH_4	t-Stat.	-2.22	0.46	-4.41	-5.02	-1.73	-0.14	-4.41	-5.02
	Prob.	0.20	1.00	0.00	0.00	0.41	0.99	0.00	0.00
		(no)	(no)	(***)	(***)	(no)	(no)	(***)	(***)
PC	t-Stat.	-3.50	-3.07	-4.57	-4.71	-2.95	-3.46	-4.57	-4.71
	Prob.	0.01	0.13	0.00	0.00	0.05	0.06	0.00	0.00
		(**)	(no)	(***)	(***)	(**)	(*)	(***)	(***)
NGC	t-Stat.	-6.24	-5.88	-4.15	-4.28	-4.81	-4.94	-4.15	-4.28
	Prob.	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.01
		(***)	(***)	(***)	(***)	(***)	(***)	(***)	(***)
CC	t-Stat.	-2.08	-4.08	-4.47	-4.48	-1.51	-2.06	-4.47	-4.48
	Prob.	0.25	0.01	0.00	0.01	0.52	0.55	0.00	0.01
		(no)	(**)	(***)	(***)	(no)	(no)	(***)	(***)
GDP	t-Stat.	-0.90	-3.57	-3.04	-3.04	-1.89	-3.26	-3.70	-3.81
	Prob.	0.78	0.05	0.04	0.14	0.33	0.09	0.01	0.03
		(no)	(**)	(**)	(no)	(no)	(*)	(***)	(**)
FDI	t-Stat.	-3.17	-3.15	-7.37	-7.33	-3.21	-3.18	-7.37	-7.33
	Prob.	0.03	0.11	0.00	0.00	0.03	0.10	0.00	0.00
		(**)	(no)	(***)	(***)	(**)	(no)	(***)	(***)
AF	t-Stat.	-2.08	-3.17	-7.67	-7.61	-2.08	-3.12	-7.67	-7.61
	Prob.	0.25	0.11	0.00	0.00	0.25	0.12	0.00	0.00
		(no)	(no)	(***)	(***)	(no)	(no)	(***)	(***)
API	t-Stat.	-2.52	0.91	-5.81	-7.22	-2.19	0.37	-5.81	-7.22
	Prob.	0.12	1.00	0.00	0.00	0.21	1.00	0.00	0.00
		(no)	(no)	(***)	(***)	(no)	(no)	(***)	(***)
LPI	t-Stat.	-10.60	-0.69	-4.93	-5.51	-2.31	-0.61	-4.93	-5.51
	Prob.	0.00	0.97	0.00	0.00	0.18	0.97	0.00	0.00
		(***)	(no)	(***)	(***)	(no)	(no)	(***)	(***)
CT	t-Stat.	-2.22	-2.23	-5.88	-6.10	-2.26	-2.35	-5.88	-6.10
	Prob.	0.20	0.46	0.00	0.00	0.19	0.40	0.00	0.00
		(no)	(no)	(***)	(***)	(no)	(no)	(***)	(***)

Notes: (*) Significant at the 10%; (**) Significant at the 5%; (***) Significant at the 1%; and (no) Not Significant.
- MacKinnon (1996) one-sided p-values.

4.1.2 The Number of Optimal Lags Periods

The table 2 shows the number of optimal lags periods for the CO2 and CH4 emissions models, which reduce the value of the following criteria such as (Sequential modified Likelihood Ratio (LR) test statistic, Final Prediction Error (FBE) Criterion, Akaike Information Criterion (AIC), Schwarz Information Criterion (SIC) and Hannan-Quinn Information (HQ) Criterion as follows:

Table 2: VAR Lag Order Selection Criteria for CO_2 and CH_4 Models.

Endogenous variables: CO_2 PC NGC CC CT FDI GDP API						
Lag	LogL	LR	FPE	AIC	SIC	HQ
0	405.62	NA	0.00	-20.93	-20.58	-20.80
1	707.14	460.22	0.00	-33.43	-30.33*	-32.32
2	803.67	106.68*	1.30e-25*	-35.140*	-29.28	-33.055*
Endogenous variables: CH_4 PC NGC CC AF LPI GDP						
Lag	LogL	LR	FPE	AIC	SC	HQ
0	375.50	NA	0.00	-19.39	-19.09	-19.29
1	690.06	496.68*	7.93e-24*	-33.37	-30.96*	-32.51*
2	739.74	60.15	0.00	-33.41*	-28.88	-31.80

* Indicates lag order selected by the criterion

It is clear from the table 2 that there are two lags' periods in the CO_2 gas model, as explained by most of the criteria except the Schwartz Information Criterion (SIC), and it also shows that there is one lag period in CH_4 model, as explained by most of the criteria except the Akaike Information criterion (AIC), and by determining these lag periods we can determine the best CO_2 and CH_4 emissions models.

4.1.3 Johansen's Co-integration Tests

The table 3 shows Johansson's cointegration tests for the two models as follows:

Table 3: Johansen's Co-integration tests Results for CO_2&CH_4 Emissions Models

Series: CO_2 PC NGC CC CT FDI GDP API							
Hypothesized	Trace Test 0.05				Max-Eigen Test 0.05		
No. of CE(s)	Eigen value	Stat.	Crit. Value	Prob.**	Stat.	Crit. Value	Prob.**
None *	0.98	412.96	159.53	0.00	151.88	52.36	0.00
At most 1 *	0.89	261.08	125.62	0.00	82.02	46.23	0.00
At most 2 *	0.80	179.06	95.75	0.00	60.33	40.08	0.00
At most 3 *	0.77	118.73	69.82	0.00	53.85	33.88	0.00
At most 4 *	0.55	64.88	47.86	0.00	29.76	27.58	0.03
At most 5 *	0.45	35.12	29.80	0.01	21.92	21.13	0.04
At most 6	0.26	13.20	15.49	0.11	10.99	14.26	0.15
At most 7	0.06	2.21	3.84	0.14	2.21	3.84	0.14
Series: CH_4 PC NGC CC AF LPI GDP							
None *	0.758	171.60	125.62	0.00	53.88	46.23	0.01
At most 1 *	0.690	117.72	95.75	0.00	44.48	40.08	0.02
At most 2 *	0.611	73.24	69.82	0.03	35.88	33.88	0.03
At most 3 *	0.410	37.35	47.86	0.33	20.02	27.58	0.34
At most 4 *	0.281	17.33	29.80	0.62	12.53	21.13	0.50
At most 5 *	0.116	4.80	15.49	0.83	4.70	14.26	0.78
At most 6	0.003	0.11	3.84	0.74	0.11	3.84	0.74

* Denotes rejection of the hypothesis at the 0.05 level.
**MacKinnon-Haug - Michelis (1999) p-values.

It is clear from the table 3 that in the long run there are six co-integration relationships in the CO_2 emissions model . There are also six co-integration relationships in the CH_4 emissions model according to The Trace and Maximum Eigenvalue tests at a significant level of 0.05. Therefore, the VECM model can be used to estimate CO_2 and CH_4 emissions.

4.1.4 Vector error correction model estimates

The table 4 shows that there is a significant long-term statistical relationship between CO2 emissions, Gross Domestic Product (GDP), Agricultural Production Index (API), Commodity Trade (CT) and Foreign Direct Investment (FDI), Natural Gas Consumption (NGC), Coal Consumption. There is also an insignificant statistical relationship between CO_2 emissions and Petroleum Consumption (PC).

Table 4: Vector Error Correction Model Estimation for CO_2&CH_4 Emissions

CO_2					CH_4				
	Coeff.	Std. Error	t-Stat.	Prob.		Coeff.	Std. Error	t-Stat.	Prob.
GDP (-1)	-0.84	-0.02	-52.71	0.00	PC (-1)	2.04	-0.53	3.87	0.00
API (-1)	0.25	-0.01	23.77	0.00	NGC (-1)	0.10	-0.16	0.62	0.54
CT (-1)	-0.04	0.01	-10.01	0.00	CC (-1)	0.29	-0.10	3.03	0.01
FDI (-1)	0.01	0.00	3.74	0.00	LPI (-1)	-1.11	-0.17	-6.53	0.00
NGC (-1)	-0.14	0.00	-34.32	0.00	GDP (-1)	-0.55	-0.36	-1.53	0.14
PC (-1)	0.01	-0.02	0.44	0.66	AF (-1)	0.70	-0.26	2.75	0.01
CC (-1)	0.12	0.00	39.47	0.00	C	2.49			
C	7.92								
CointEq1	-1.15	-0.57	-2.03	0.04	CointEq1	-0.06	0.02	-2.55	0.01
D (CO_2 (-1))	0.47	-0.50	0.94	0.35	D(CH4(-1))	0.08	0.20	0.41	0.68
D (CO_2 (-2))	0.21	-0.34	0.63	0.53	D (PC (-1))	-0.04	0.10	-0.37	0.71
D (GDP (-1))	-0.20	-0.47	-0.42	0.67	D (NGC (-1))	-0.10	0.04	-2.18	0.03
D (GDP (-2))	0.49	-0.47	1.05	0.30	D (CC (-1))	0.02	0.02	0.72	0.47
D (API (-1))	0.36	-0.23	1.53	0.13	D (LPI (-1))	0.11	0.06	1.74	0.08
D (API (-2))	0.18	-0.23	0.82	0.42	D (GDP (-1))	0.19	0.28	0.68	0.50
D (CT (-1))	0.05	-0.03	1.67	0.10	D (AF (-1))	0.07	0.05	1.41	0.16
D (CT (-2))	0.03	-0.03	0.89	0.38	C	0.01	0.01	1.55	0.12
D (DIF (-1))	0.01	-0.01	1.10	0.27					
D (DIF (-2))	0.01	-0.01	0.80	0.43					
D (NGC (-1))	-0.12	-0.17	-0.70	0.48					
D (NGC (-2))	-0.07	-0.12	-0.60	0.55					
D (PC (-1))	0.13	-0.18	0.72	0.47					
D (PC (-2))	0.20	-0.19	1.02	0.31					
D (CC (-1))	0.04	-0.05	1.83	0.44					
D (CC (-2))	0.12	-0.06	1.83	0.07					
C	-0.01	-0.02	0.43	0.66					
R-squared		0.68			R-squared		0.57		
F-statistic		2.32			F-statistic		4.75		
Prob.(F-Statistic)		0.01			Prob.(F-Statistic)		0.01		
Durbin-Watson stat		2.32			Durbin-Watson stat		2.37		
Akaike AIC		-5.62			Akaike AIC		-6.17		
Schwarz SC		-4.84			Schwarz SC		-5.78		

The table 4 shows that a CO_2 emissions model is significant using the F-test. The explanatory ability of this model was good where the coefficient of determination equals 0.68. As well as the value of the Durbin-Watson test equals 2.32. It means that there is no autocorrelation between the residuals. The results also showed that the error correction coefficient (CoinEq1) was significant and negative where equals (-1.15). This means that the adaptation speed between the short and long run. As well as it can be corrected any imbalance in CO_2 emission rates and returning it to a stable and balance state. It is clear from the table 4 that in the long run there is a statistically significant relationship between CH_4 emissions and Petroleum Consumption (PC), Coal Consumption (CC), Agricultural Fertilizer (AF) and Livestock Production Index (LPI). Also, it was found a statistical non-significant relationship between CH_4 emissions and Natural Gas Consumption (NGC) and Gross Domestic Product (GDP). The table 4. Also, shows that a CH_4 gas emissions model is significant using the F-test. The explanatory ability of this model was good where the coefficient of determination equals 0.57. As well as the value of the Durbin-Watson test equals 2.37. It means that there is no autocorrelation between the residuals. The results also showed that the error correction coefficient (CoinEq1) was significant and negative where equals (-0.06). This means that the adaptation speed between the short and long run. As well as it can be corrected any imbalance in CH_4 emission rates and returning it to the stability and balance.

4.1.5 Diagnostic Tests

The table 5 shows that the significant level of the Vector Error Correction (VEC) residual serial correlation Lagrange Multiplier (LM) test for both models (CO_2&CH_4) was greater than 0.05, therefore we will accept the null hypothesis that means the residuals do not suffer from the autocorrelation problem.The VEC residual normality test is presented, which shows that the residuals of most variables follow the normal distribution for both models $(CO_2\&CH_4)$ using the Jarque-Bera test, where their probabilities were greater than 0.05. It is also clear from the table 5 that the probability result of the chi -square test is equal to 0.24 in the CO_2 emissions model, as well as in the CH_4 emissions model, its value was 0.59, which is greater than 0.05, according to the VEC resident Heteroscedasticity Test. This means that there is no problem of Heteroscedasticity.

Table 5: Diagnostic Tests for CO_2 and CH_4 Emissions Models.

VEC Residual Serial Correlation LM Tests

Lag	LRE* stat	df	Prob.	Rao F-stat	df	Prob.
			CO_2			
1	64.58	64	0.46	0.88	(64, 29.6)	0.68
2	83.47	64	0.05	1.37	(64, 29.6)	0.17
			CH_4			
1	44.53	49	0.65	0.88	(49, 85.7)	0.68
2	31.39	49	0.98	0.58	(49, 85.7)	0.98

VEC Residual Normality Tests

	CO_2			CH_4		
Component	Jarque - Bera	df	Prob.	Jarque - Bera	df	Prob.
1	4.1	2.0	0.1	0.48	2	0.79
2	0.3	2.0	0.9	0.95	2	0.62
3	8.4	2.0	0.0	2.89	2	0.24
4	42.8	2.0	0.0	9.39	2	0.01
5	4.2	2.0	0.1	265.34	2	0.00
6	0.3	2.0	0.8	0.62	2	0.73
7	0.2	2.0	0.9	0.36	2	0.83
8	0.8	2.0	0.7			
Joint	16.15	16	0.00	280.04	14	0.00

VEC Residual Heteroscedasticity Tests (Levels and Squares)

CO_2			CH_4		
Chi-sq.	df	Prob.	Chi-sq.	df	Prob.
1258	1224	0.24	441	448	0.59

* Edgeworth expansion corrected likelihood ratio statistic.

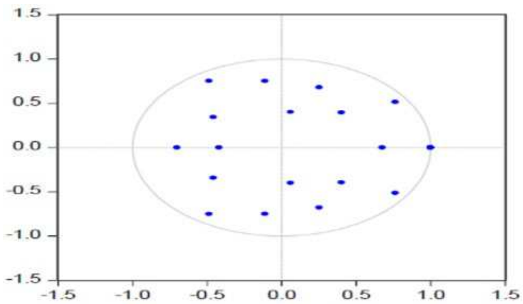

Fig. 2: An Inverse Roots of AR Characteristic Polynomial for CO_2 Emissions Model

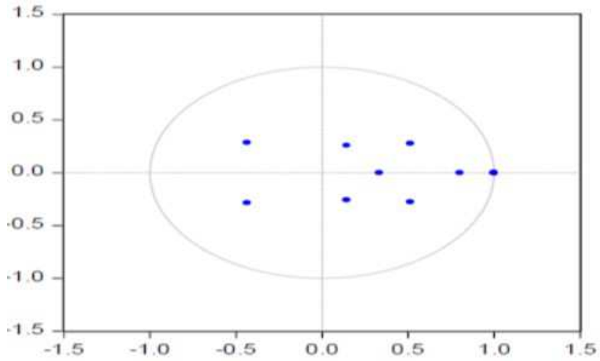

Fig. 3: An Inverse Roots of AR Characteristic Polynomial for CH_4 Emissions Model

As for the stability test for the two models, the following figures 2 and 3 show that the inverse roots of Autoregressive characteristic polynomial for two models. They are located inside and on the boundaries of the circle, which indicates that the stability condition for the two models has been achieved.

4.1.6 Causality Granger Test

it is clear from table 6 that in the third slowdown period there is a one-way causal relationship between the GDP variable, the NGC consumption variable, and the Agricultural Production Index (API) to Carbon Dioxide (CO_2) emissions. There is also a one-way causal relationship from the variable of carbon dioxide CO_2 emissions to Petroleum Consumption (PC), and there is no causal relationship between CO_2 emissions and (Coal Consumption (CC), Foreign Direct Investment (FDI), and Commodity Trade (CT)). It is also clear that in the eighth lag period there is a one-way causal relationship between both the Petroleum Consumption (PC) variable and the Livestock Production Index (LPI) to methane (CH_4) emissions .There was also a one-way causal relationship from the methane emissions (CH_4) variable to both the consumption variable of Natural Gas (NGC) and Agricultural Fertilizers (AF), and there is no causal relationship between methane emissions (CH_4) and both (Coal Consumption (CC), and Gross Domestic Product(GDP)).

Table 6: VEC causality Granger test for CO_2 and CH_4 Emissions Models.

Pairwise Granger Causality Tests (CO_2) – Lag:3			
Null Hypothesis:	Obs.	F-Statistic	Prob.
GDP does not Granger Cause CO_2	37	3.704	0.022
CO_2 does not Granger Cause GDP		0.536	0.661
PC does not Granger Cause CO_2	37	0.659	0.584
CO_2 does not Granger Cause PC		9.581	0.000
NGC does not Granger Cause CO_2	37	3.691	0.023
CO_2 does not Granger Cause NGC		0.114	0.951
CC does not Granger Cause CO_2	37	0.574	0.637
CO_2 does not Granger Cause CC		1.899	0.151
API does not Granger Cause CO_2	37	4.745	0.008
CO_2 does not Granger Cause API		1.011	0.401
FDI does not Granger Cause CO_2	37	0.207	0.891
CO_2 does not Granger Cause FDI		2.253	0.103
CT does not Granger Cause CO_2	37	0.546	0.654
CO_2 does not Granger Cause CT		1.140	0.349
Pairwise Granger Causality Tests (CH_4) – Lag:8			
PC does not Granger Cause CH_4	32	2.843	0.039
CH_4 does not Granger Cause PC		1.373	0.284
NGC does not Granger Cause CH_4	32	0.407	0.900
CH_4 does not Granger Cause NGC		3.522	0.017
CC does not Granger Cause CH_4	32	0.341	0.936
CH_4 does not Granger Cause CC		0.735	0.661
GDP does not Granger Cause CH_4	32	1.749	0.167
CH_4 does not Granger Cause GDP		1.334	0.300
AF does not Granger Cause CH_4	32	0.452	0.871
CH_4 does not Granger Cause AF		2.853	0.038
LPI does not Granger Cause CH_4	32	3.574	0.016
CH_4 does not Granger Cause LPI		0.354	0.929

4.1.7 Variance Components Analysis

The table 7 shows the analysis of the variance components of the prediction error in the CO_2 emissions variable and the role of each shock of independent variables in explaining the fluctuations that occur in CO_2 emissions during ten future periods. We note that there were no changes in CO_2 gas emissions in the first period, where it was 100%, and then decreases in the second period to 85% and so on until it reaches 55.7% in the long term in the tenth period, and the remaining 44.3% of the explanation for the variation of the prediction error is due to changes in Natural Gas Consumption (NGC), Gross Domestic Product(GDP), Coal Consumption(CC), Petroleum Consumption (PC), Commodity Trade(CT), Agricultural Production Index (API), and Foreign Direct Investment (FDI) respectively.

Table 7: Variance Components Analysis for CO_2&CH_4 Emissions Models.

Period	S.E.	CO2	PC	NGC	CC	CT	FDI	GDP	API
1	0.01	100.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2	0.02	85.13	0.93	4.92	6.72	1.05	1.21	0.03	0.02
3	0.02	76.84	0.79	10.68	7.66	1.46	1.03	0.77	0.77
4	0.03	73.70	0.96	12.93	7.53	1.70	0.86	1.78	0.54
5	0.03	68.00	2.46	13.65	9.60	1.83	0.67	3.34	0.45
6	0.03	64.27	3.86	13.40	9.27	1.87	0.70	6.17	0.47
7	0.04	60.94	4.96	14.08	8.21	2.01	0.77	8.25	0.78
8	0.04	58.42	5.19	14.91	8.14	2.15	0.80	9.61	0.79
9	0.04	57.15	5.38	15.21	7.61	2.18	0.88	10.80	0.80
10	0.05	55.69	6.12	15.57	7.01	2.25	1.07	11.21	1.08

Period	S.E.	CH4	PC	NGC	CC	AF	LPI	GDP
1	0.01	100.00	0.00	0.00	0.00	0.00	0.00	0.00
2	0.02	75.84	0.06	3.78	0.59	0.03	19.39	0.32
3	0.02	59.99	0.03	4.87	2.20	1.45	31.31	0.15
4	0.03	48.31	0.19	4.33	4.77	1.81	40.48	0.11
5	0.04	40.74	0.38	3.55	6.50	2.34	46.40	0.11
6	0.05	35.87	0.56	2.86	7.62	2.59	50.38	0.12
7	0.05	32.51	0.71	2.35	8.33	2.75	53.23	0.12
8	0.06	30.19	0.81	1.97	8.79	2.87	55.26	0.12
9	0.07	28.48	0.88	1.69	9.10	2.94	56.78	0.12
10	0.07	27.19	0.94	1.48	9.32	3.01	57.94	0.12

It is clear also from the table 7 the analysis of the variance components of the prediction error in the methane (CH_4) gas model, we note that there were no changes in CH_4 gas emissions in the first period, where they were 100%, then they decrease in the second period to reach 76% and so on until they reach in the long term to 27% in the tenth period, and the remaining percentage is approximately 73% due to changes in the livestock Production Index (LPI), Coal Consumption (CC), Agricultural Fertilizers (AF), Natural Gas Consumption (NGC), Petroleum Consumption (PC), and Gross Domestic Product (GDP) respectively .

4.1.8 Impulse Responses Functions

The figure 4 shows the response of CO_2 emissions to the shocks of the following independent variables in the future for ten years as follows:

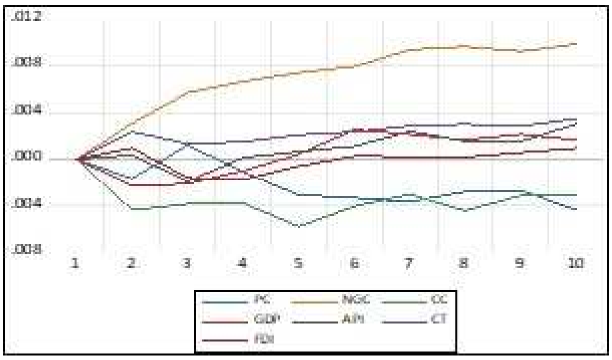

Fig. 4: Response of CO_2 to Innovations using Cholesky (d.f. adjusted) Factors

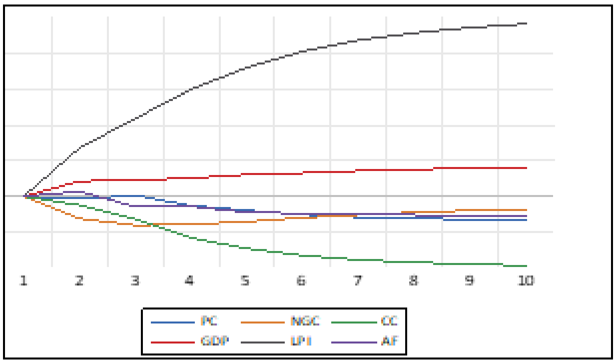

Fig. 5: Response of CH_4 to Innovations using Cholesky (d.f. adjusted) Factors

- –The shocks of Petroleum Consumption (PC) by one standard deviation is negative during the first period, then it tend to the positive in the second period and continue to negative until the tenth period.
- –The shocks of Natural Gas Consumption (NGC) and Commodity Trade (CT) are positive from the first period until the tenth period.
- –The shocks of Coal Consumption (CC) will be negative in the first period until the tenth period.
- –The shocks of GDP will be negative in the first period, then changes to positive until the tenth period. The shocks of Foreign Direct Investment (FDI) and Agricultural Production Index (API) will be positive during the first period, then decreases in the second period and change to positive until the tenth period.

The figure 5 shows also the response of CH_4 emissions to the shocks of the following independent variables in the future for ten years as follows. The shocks of livestock Production Index (LPI) and GDP will be positive from the first period until the tenth period. The shocks of Agricultural Fertilizer (AF) will be positive during the first period, then decrease in the second period and continue until the tenth period. The shocks of Petroleum Consumption (PC) will be negative during the first period, then become positive in the second period and change to negative until the tenth period. The shocks of Coal Consumption (CC) will be negative in the first period until the tenth period. The shocks of Natural Gas Consumption (NGC) will be negative in the first and second periods then increase until the tenth period.

4.2 Autoregressive Integrated Moving Average with Exogenous Variable (ARIMAX)

The following figures from no.6 to no.9 show that the time series of emissions of CO_2 and CH_4 is not stable at the level and will be stable at the first difference. We note that the same result achieved as shown in Table 1 when we applied the ADF and PP stability tests.

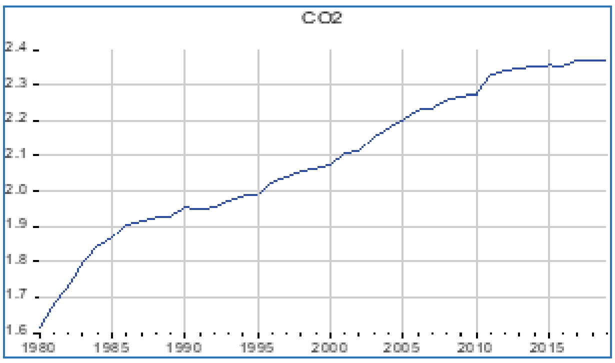

Fig. 6: CO_2 at level

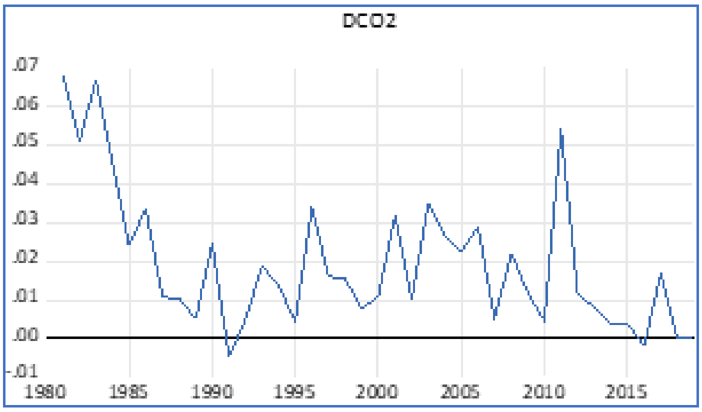

Fig. 7: CO_2 at 1^{st} difference

Fig. 8: CH_4 at level

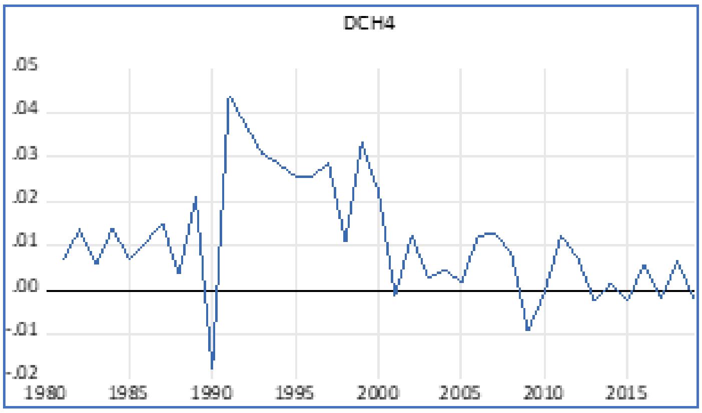

Fig. 9: CH_4 at 1^{st} difference

The following figures 10 and 11 will determine the ranks of all possible ARIMAX models using the autocorrelation function (ACF) and the partial autocorrelation function (PACF) of the D (CO_2) and D(CH_4) series.

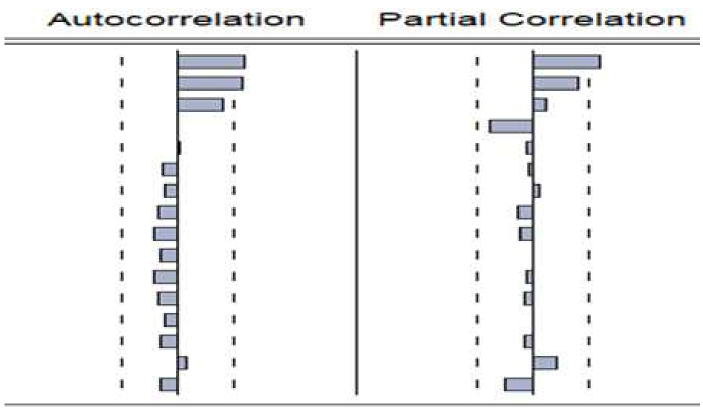

Fig. 10: ACF and PACF plot of D (CO2)

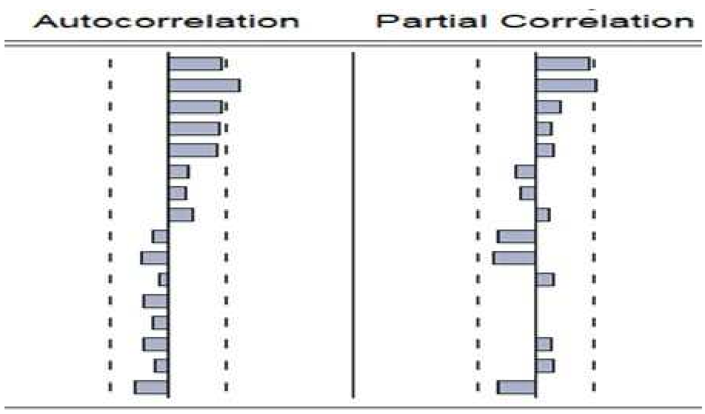

Fig. 11: ACF and PACF plot for D (CH4)

It is clear also from the figures 10 and 11 that according to ACF and PACF there are several models of CO_2 and CH_4 emissions series whose values fall outside the confidence limits. Based on the estimation of these models which appears in table 8 we can choose the best model which has the lowest value according to the (AIC) criterion. It is clear also from table 8 that the best model of CO_2 emissions is ARIMAX $(2,1,2)$ and the best model of CH_4 emissions is ARIMAX $(2,1,0)$

Table 8: Comparison between ARIMAX models according to D (CO_2) and (CH_4) series

Dependent Variable: D (CO₂)				
Model	LogL	AIC	BIC	HQC
ARIMAX (2,1,2)	131.439272	-6.073809	-5.519288	-5.874852
ARIMAX (0,1,1)	128.209329	-6.062017	-5.635463	-5.908973
ARIMAX (2,1,0)	130.064428	-6.054586	-5.542721	-5.870933
ARIMAX (0,1,2)	128.579703	-6.029728	-5.560519	-5.861380
ARIMAX (1,1,1)	128.286413	-6.014688	-5.545478	-5.846339
ARIMAX (1,1,2)	128.649958	-5.982049	-5.470184	-5.798396
ARIMAX (2,1,0)	125.730633	-5.883622	-5.414412	-5.715274
ARIMAX (1,1,0)	124.680335	-5.881043	-5.454489	-5.727999
ARIMAX (0,1,0)	122.661481	-5.828794	-5.444895	-5.691054
Dependent Variable: D(CH₄)				
Model	Log L	AIC	BIC	HQC
ARIMAX (2,1,0)	125.652740	-5.930910	-5.504355	-5.777866
ARIMAX (1,1,1)	125.042792	-5.899630	-5.473076	-5.746586
ARIMAX (1,1,2)	125.824388	-5.888430	-5.419220	-5.720082
ARIMAX (2,1,1)	125.697454	-5.881921	-5.412711	-5.713572
ARIMAX (2,1,2)	125.819649	-5.836905	-5.325040	-5.653252
ARIMAX (1,1,0)	122.467230	-5.818832	-5.434933	-5.681093
ARIMAX (0,1,2)	122.319261	-5.759962	-5.333408	-5.606918
ARIMAX (0,1,1)	118.600750	-5.620551	-5.236652	-5.482812
ARIMAX (0,1,0)	116.005670	-5.538752	-5.197509	-5.416317

The table 9 shows the estimation of ARIMAX $(2,1,2)$ model for and ARIMAX $(2,1,0)$ model for CH_4 emissions. It is clear from the diagnostic stage for both models ARIMAX (2,1,2) and ARIMAX $(2,1,0)$ that the values of Durbin-Watson statistic are $(1.93\&2)$ respectively, which means that there is no autocorrelation between the residuals.

Table 9: Estimation of D (CO_2)-ARIMAX $(2,1,2)$ and D(CH_4)-ARIMAX $(2,1,0)$ Models

D (CO₂)-ARIMAX (2,1,2)					D(CH₄)-ARIMAX (2,1,0)				
Variable	Coeff.	Std. Err.	t-Stat.	Prob.	Variable	Coeff.	Std. Err.	t-Stat.	Prob.
C	0.012	0.003	3.604	0.001	C	0.018	0.010	1.756	0.090
D(PC)	0.390	0.157	2.485	0.020	D(PC)	-0.220	0.156	-1.413	0.168
D(NGC)	0.212	0.025	8.489	0.000	D(NGC)	0.016	0.052	0.311	0.758
D(CC)	-0.051	0.021	-2.450	0.021	D(CC)	-0.008	0.036	-0.212	0.833
D(GDP)	-0.130	0.239	-0.543	0.592	D(GDP)	-0.138	0.278	-0.496	0.623
D(FDI)	0.004	0.009	0.443	0.661	D(LPI)	-0.091	0.084	-1.085	0.287
D(CT)	0.045	0.018	2.462	0.021	D(AF)	-0.080	0.069	-1.160	0.255
D(API)	-0.209	0.051	-4.062	0.000	AR (1)	0.335	0.299	1.120	0.272
AR (1)	0.548	0.453	1.210	0.237	AR (2)	0.430	0.270	1.591	0.123
AR (2)	-0.191	0.343	-0.560	0.581	SIGMASQ	0.000	0.000	4.054	0.000
MA (1)	-1.993	2421.14	-0.000	0.999					
MA (2)	0.910	2430.13	0.000	0.999					
SIGMASQ	5.4E-05	0.065	0.001	0.999					
R-squared	0.834	Mean dep. var.		0.019	R-squared	0.469	Mean dep. var.		0.011
Adj. R-sq.	0.757	S.D. dep. var.		0.018	Adj. R-sq.	0.304	S.D. dep. var.		0.013
S.E. of reg.	0.009	Akaike info. crit.		-6.074	S.E. of reg.	0.011	Akaike info. crit.		-5.931
Sum squ.res.	0.002	Schwarz crit.		-5.519	Sum.squ.res.	0.004	Schwarz crit.		-5.504
Log like.	131.44	Han. - Quinn crit.		-5.875	Log like.	125.7	Han. - Quinn crit.		-5.778
F-statistic	10.854	Durb. -Wats. stat.		1.929	F-statistic	2.842	Durb. -Wats. Stat.		2.004
Prob.(F-stat.)	0.000				Prob.(F-stat.)	0.016			

The table 10 shows that vector error correction models for CO_2 and CH_4 emissions are better than ARIMAX $(2,1,2)$ and ARIMAX $(2,1,0)$ models because they have the lowest values according to the following predictive ability criteria (AIC, SIC, RMSE, MAE, Theil Inequality). This means that we can use them to predict CO_2 and CH_4 emissions.

Table 10: The Predictive Ability of ARIMAX & VECM models for CO_2&CH_4 emissions

	Model	AIC	SIC	RMSE	MAE	MAPE	Theil Ine.
D (CO_2)	ARIMAX (2,1,2)	-6.08	-5.52	2.09	2.08	99.02	0.98
	VECM	-5.62	-4.84	0.03	0.03	1.34	0.01
D (CH_4)	ARIMAX (2,1,0)	-5.93	-5.50	4.589	4.586	99.76	0.99
	VECM	-6.17	-5.785	0.09	0.08	1.67	0.01

The table 11 shows that carbon dioxide emissions will increase in the future during the period 2020 to 2030 due to an increase in consumption of oil, natural gas consumption, coal consumption, GDP and agricultural production index. We also note that methane emissions will also increase during the same period due to an increase in consumption of oil and natural gas consumption, coal consumption, GDP, livestock production index and agricultural fertilizers consumption.

Table 11: the forecast values of CO_2&CH_4 emissions and the factors affecting them during the period (2020-2030) using VECM

Year	CO2_F	PC_F	NGC_F	CC_F	GDP_F	FDI_F	CT_F	API_F
2020	2.400	0.220	0.400	-1.042	12.601	-0.008	1.464	2.046
2021	2.403	0.215	0.403	-0.974	12.617	0.301	1.481	2.048
2022	2.409	0.211	0.439	-1.014	12.635	0.179	1.504	2.046
2023	2.432	0.227	0.476	-0.995	12.649	0.013	1.474	2.079
2024	2.443	0.230	0.489	-0.910	12.666	0.184	1.442	2.086
2025	2.453	0.239	0.523	-0.930	12.684	0.247	1.454	2.095
2026	2.472	0.246	0.552	-0.886	12.701	0.180	1.479	2.120
2027	2.480	0.247	0.569	-0.815	12.722	0.314	1.491	2.125
2028	2.491	0.256	0.606	-0.833	12.742	0.312	1.489	2.137
2029	2.511	0.265	0.637	-0.807	12.761	0.209	1.475	2.160
2030	2.520	0.267	0.659	-0.766	12.780	0.248	1.471	2.167
Year	CH4_F	PC_F	NGC_F	CC_F	GDP_F	LPI_F	AF_F	
2020	4.711	0.166	0.160	-1.042	12.588	1.975	2.736	
2021	4.720	0.171	0.183	-1.026	12.607	1.991	2.740	
2022	4.728	0.176	0.206	-1.010	12.625	2.007	2.744	
2023	4.737	0.180	0.229	-0.994	12.643	2.024	2.748	
2024	4.745	0.185	0.253	-0.979	12.662	2.040	2.752	
2025	4.754	0.190	0.276	-0.963	12.680	2.056	2.756	
2026	4.763	0.195	0.299	-0.947	12.698	2.072	2.760	
2027	4.771	0.200	0.322	-0.931	12.716	2.089	2.764	
2028	4.780	0.204	0.346	-0.916	12.735	2.105	2.768	
2029	4.788	0.209	0.369	-0.900	12.753	2.121	2.772	
2030	4.797	0.214	0.392	-0.884	12.771	2.137	2.776	

5 Conclusion

The study aimed to measure the impact of energy sources consumption on CO_2 and CH_4 emissions in Egypt during the period 1980 to 2019 in the short and long term, and to predict these emissions during the period 2020 to 2030 . When we applied the analysis of the variance components of the prediction error of the CO_2 and CH_4 emissions models, it was already found that an energy consumption and GDP were the most effected on CO_2 emissions. Also, an energy consumption, livestock production and agricultural fertilizers were the most influential on CH_4 emissions. The VECM was found to be better than ARIMAX models for predicting emissions of CO_2 and CH_4. It is expected that the emissions of CO_2 and CH_4 will rise, as well as the consumption of petroleum, natural gas consumption, and coal consumption will rise, which means that we are facing a big problem in the future, which leads to more climate instability and change in ecosystems in Egypt. This also negatively affects human security and may lead to internal or external migration. Therefore, based on the previous results, this study recommends the following:

– It is clear that the agricultural production and fertilizers consumption variables are the main source of methane emissions from human activities. Therefore, the Ministry of Agriculture should continuously raise awareness among farmers about the production of agricultural varieties that bear negative environmental changes that affect crop productivity.
– It is necessary look for agricultural varieties that require little nitrogen fertilizers that reduce methane and nitrous oxide emissions, and plant more high productivity varieties with a higher photosynthetic capacity to absorb more carbon dioxide. In addition, stop the burning of agricultural waste in the fields by researching smart practices that contribute to reducing methane emissions from rice fields.
– It must take care of the health of livestock because this will reduce the presence of dead animals that decompose organically and thus reduce methane emissions that negatively affect the environment and increase climate changes.
– Focusing on investments in renewable and clean energy and all modern technological investments that have a positive impact on the environment. Also, it must leave the investments that pollute the environment and increase greenhouse gas emissions.
– Increasing the dependence of the renewable energy sector on low-cost natural energy sources that are inexhaustible (wind, solar, water, biofuels, etc.), reducing dependence on traditional fossil energy sources (petroleum, natural gas, coal) that have the high carbon intensity, and providing practical solutions capable of reducing harmful emissions and negative effects resulting from traditional industries.
– Encourage importers to import environmentally friendly and low-fuel production methods from technologically advanced countries and that protect the environment.
– It must activate the role of the Ministry of Environment Continuously by the development of policies and procedures which protect the environment. because the increasing of economic development and energy consumption will affect negatively on greenhouse gas emissions.

Acknowledgement

I would like to thank Prof. Hayam Wahba, Vice Dean for Postgraduate Studies and Prof. of Finance and Investment, Faculty of Business, Ain Shams University for her advice and continuous support, and I would also like to thank the editor and referees for their suggestions and corrections to improve the quality of this paper.

References

[1] L. Adkins and R. Hill. Using STATA for principles of Econometrics. 5th edition, Wiley, 391, (2018).

[2] T. Alam and A. Alarjani. Forecasting CO_2 emissions in Saudi Arabia using artificial neural network, holt-winters exponential smoothing and autoregressive Integrated moving average models. 2021 International Conference on Technology and Policy in Energy and Electric Power (ICT-PEP), (2021).

[3] A. Anwar, M. Younis and I. Ullah. Impact of urbanization and economic growth on CO_2 emissions: a Case of Far East Asian Countries. International Journal of Environmental Research and Public Health, 17, 2531, (2020).

[4] Y. Bayar, L. Diaconu and A. Maxim. Financial development and CO_2 emissions in post- transition European Union Countries. Sustainability,12, 2640, (2020).

[5] F. Bekun, A. Alola and S. Sarkodie. Toward a sustainable environment: nexus between CO_2 emissions, resource rent, renewable and non-renewable energy in 16-EU countries. The Science of The Total Environment, 657, 1023-1029, (2019).

[6] F. Bilgili, E. Kocak and U. Bulut.The dynamic impact of renewable energy consumption on CO_2 emissions: a revisited Environmental Kuznets Curve approach. Renewable and Sustainable Energy Reviews,54, 838-845, (2016) .

[7] G. Box, G. Jenkins, G. Reinsel, and , G. Ljung. Time series analysis: forecasting and control. 5th edition, John Wiley & Sons Inc., Hoboken, New Jersey,712, (2016).

[8] D. Dickey. and W. Fuller. Distribution of the estimators for autoregressive time series with a unit root. Journal of the American Statistical Association, 74,(366a), 427-431,(1979).

[9] D. Dickey and W. Fuller. Likelihood ratio statistics for autoregressive time series with a unit root. Econometrica, 49, 1057-1072,(1981).

[10] R. Engle and C. Granger. Co-integration and error correction: representation, estimation and Testing. Econometrica, 55, 251-276, (1987).

[11] Food and Agriculture Organization (FAO), https://www.fao.org/home/ar.

[12] E. Garip, and A. Oktay. Forecasting CO_2 emission with machine learning Methods. International Conference on Artificial Intelligence and Data Processing (IDAP),Malatya, Turkey, (2018).

[13] G. Granger. Co-integrated variables and error-correcting Models. Unpublished discussion paper, University of California, San Diego, 83-13, (1983).

[14] M. Jebli, S. Youssef, and I. Ozturk. Testing environmental Kuznets curve hypothesis: the role of renewable and non-renewable energy consumption and trade in OECD countries. Ecological Indicators, 60,824-831, (2016).

[15] S. Johansen. Statistical analysis of co-integrating vectors. Journal of Economic Dynamics and Control, 12, 231-254. http://dx.doi.org/10.1016/01651889 90041-3, (1988).

[16] S. Johansen and K. Juselius. Maximum likelihood estimation and inference on co-integration with applications to demand for money.Oxford Bulletin of Economics and Statistics, 52,169-210,(1990).

[17] S. Kaur, and M. Rakshit. Performance comparison of forecasting on solar plant generation data. International Journal of Computer Sciences and Engineering (IJCSE). Open Access, 6, 10, 827-834, (2018).

[18] H. Khobai. Renewable energy consumption and economic growth in Indonesia : Evidence from the ARDL bounds testing approach. MPRA Paper, 85081, University Library of Munich, Germany, https://mpra.ub.uni-muenchen.de/85081. (2018).

[19] G. Ljung and G. Box. On a measure of a lack of fit in time series models. Biometrika , 65, 2, 297-303, , (1978).

[20] H. Lutkepohl. New introduction to multiple time Series analysis. Berlin, Springer Verlag, (2005).

[21] T. Nyoni, H. Hapanyengwi, and C. Mutongi. Modelling and forecasting carbon Dioxide emissions in China using autoregressive integrated moving average (ARIMA) Models. EPRA International Journal of Multidisciplinary Research (IJMR), (5)(4) 215- 224, (2019).

[22] H. Philip and V. Dick. Nonlinear time series models in empirical finance. Cambridge University Press, $136-137$, (2003).

[23] P. Phillips and P. Perron. Testing for a unit root in time series regression. Biometrika,75, 335-346, (1988).

[24] A. Rahman and M. Hasan. Modeling and forecasting of carbon dioxide emissions in Bangladesh using autoregressive integrated moving average (ARIMA) Models. Open Journal of Statistics, 7, 560-566, (2017).

[25] The World Bank Group, https:/www.worldbank.org.

[26] United States (US) Energy Information Administration (EIA), https://www.eia.gov/international/data/country/EGY/other-statistics/more\OT1\textendashother\OT1\textendashstatistics-data?pd=40&p=&.