Eco. Env. & Cons. 28 (February Suppl. Issue) : 2022; pp. (S459-S469)
Copyright@ EM International
ISSN 0971–765X

DOI No.: http://doi.org/10.53550/EEC.2022.v28i02s.070

Forecasting the Emission of Carbon-di-oxide Equivalent in Key Sectors of India Using ARIMA Model

Deepa S.[1] and Vijayasree G.[2]

Research Department of Statistics
Shrimathi Devkunvar Nanalal Bhatt Vaishnav College for Women, Chromepet, Chennai, Tamil Nadu 600 044, India

(Received 21 July, 2021; Accepted 20 August, 2021)

ABSTRACT

Climate Change is a global challenge and needs to be addressed immediately. The emission of Green House Gases in the atmosphere by anthropogenic factors is one of the major causes of Global Warming and Climate Change. India is working towards the control of global warming by focussing on controlling Green House Gas emissions. The emission of carbon-di-oxide in the atmosphere plays a predominant role in global warming. The Global Warming potential of the Green House Gases is measured in terms of carbon-di-oxide equivalent CO_2 (eq.). The annual CO_2 (eq.) emission data from 2005-2015 from four key sectors viz., Energy, Industry Process and Product Use (IPPU), Agriculture Forestry and Land Use (AFOLU), and the Waste sector were considered in the study. Classical Temporal disaggregation methods Denton, Denton Cholette, Chow - Lin, Fernandez, and Litterman methods were employed to disaggregate the low frequency (annual) data to high frequency (quarterly) data. The analysis revealed that the Chow - Lin method of disaggregation best suited to disaggregate the CO_2 (eq.) series for the three sectors except AFOLU with an Adjusted R square of 0.9 and the current Price GDP is the good indicator series for CO_2 (eq.). The disaggregated data is modelled using ARIMA modelling. The CO_2 (eq.) from 2021-Q1 to 2023-Q4 is forecasted using the fitted ARIMA model for each sector.

Key words: *Climate change, Greenhouse gas, Carbon-de-oxide equivalent CO_2 (eq.), Temporal disaggregation, Forecasting.*

Introduction and Review of Literature

Climate change faced by mankind for the last two decades, has been acknowledged as a major global issue at various international forums, which requires a concerted global response.

The emission of greenhouse gasses into the atmosphere has been identified as one of the major causes of climate change. The three major greenhouse gases (GHG) are Carbon dioxide (CO_2), Methane (CH_4) and Nitrous oxide (N_2O), of which CO_2 contributes a significant 58.8% of the GHGs responsible for climate change (Pao and Tsai, 2011). "The Paris Agreement" was adopted by 196 countries on December 12, 2015, to formulate a structure among various nations for controlling these emissions. The goal of the agreement is to limit global warming to well within 2 to 1.5 degrees Celsius.

India is the third largest emitter of GHGs in the world with roughly about 3 Giga tonnes of CO_2 (eq) and 7% of the global emissions and excessive usage of coal being the main source of these emissions. As per the Paris Climate Change Agreement 2015, India has committed to cut GHG emissions intensity of its GDP by 33-35 percent, increase non-fossil fuel power capacity to 40 % from 28%, add carbon sink

[1]Research Scholar, [2]Assistant Professor

of 2.5 -3 billion tonne of CO_2 per annum by increasing the forest cover by 2030. Therefore, it is important to employ appropriate statistical models for forecasting these emissions, so that necessary timely action can be taken to meet the goal.

The GHGs emission in India can be classified from four broad sectors namely the Energy Sector (ES), Industrial Process and Product-Use Sector (IPPUS), Agriculture, Forest and Land-use Sector (AFOLUS) and Waste Sector (WS).

The nature of data on GHGs is a Time-Series, which can be analysed and modelled using historical behaviour or pattern in the data for Forecasting. Time-series modelling can be done only when the data is of high frequency and Temporal Disaggregation techniques are employed in case of low-frequency data, to convert the series into required high frequency. Several approaches of disaggregation of time series like Sectoral-disaggregation Nain *et al.,* (2017), Temporal disaggregation Ajao *et al.* (2015), MIDAS regression (Guay and Maurin (2015) and Non-Parametric methods Ayodeji (2012) are available in literature, which have been applied on economic series. A literature review suggests that temporal disaggregation technique is a popular and widely used method for economic data. In order to apply Temporal disaggregation techniques to a given series, a related series is required to be identified. A review of literature on the relationship between economic growth and carbon emissions indicate that carbon emissions in the atmosphere have an impact on the nation's Gross Domestic Product (GDP), which is used as a proxy variable for economic activity. Refer Ghosh (2010); Pal and Kumar (2017); Alam *et al.* (2016) and Kanjilal and Ghosh (2013).

Several researchers Nyoni and Bonga (2019), Rahman and Hasan (2017); Sen *et al.* (2016) and Lotfalipour *et al.* (2013) have considered ARIMA models for forecasting CO_2 models. Lotfalipour *et al.* (2013), predicted CO_2 emissions in Iran based on grey system and Autoregressive Integrated Moving Average and compared the RMSE, MAE, and MAPE metrics. Pao and Tsai (2011) found that both Grey and ARIMA models exhibit high predicting performance for energy consumption and output in Brazil (from 1980-2007), with MAPEs of less than 3% in both cases.

Most of the studies have been made using CO_2 emissions only, while in the present study, we consider "Total CO_2 (eq.) of the three main GHGs" calculated according to the IPCC Assessment Report-II, because of its global warming potential. The nature of the data is a Low frequency Annual time series data from 2005 to 2015, which has to be disaggregated for times series modelling. This study attempts to disaggregate CO_2 (eq.) data from Low frequency (yearly) to High frequency (quarterly), by taking Current Price GDP as a related series for disaggregating Co_2(eq.). and forecast Co_2(eq.) emissions using the Auto Regressive Integrated Moving Average (ARIMA) model.

Objectives of the study

The aim of this study is to forecast the carbon-dioxide equivalent [CO_2 (eq.)] from various sectors in India, applying ARIMA model, by disaggregating the Low frequency data into high frequency data using Temporal disaggregation techniques by identifying a related indicator series for disaggregating CO_2 (eq.).

Material - Data Description

The data required for the study was extracted from www.ghgplatform-india.org. The "GHG Platform – India" is a civil society initiative that provides an unbiased estimate and analysis of India's greenhouse gas (GHG) emissions across important sectors like Energy, Waste, Industrial Process and Product Use (IPPU), Agriculture Forest and Other Land Use (AFOLU). National estimates of GHG emissions from 2005 to 2015 are presently available on the platform, which account for CO_2, CH_4, and N_2O gases, as well as their carbon dioxide equivalent CO_2 (eq). This study considers CO_2(eq) (in Giga Tonnes) from 2005-2015 and Current Price GDP (India) from (2005Q1-2020Q4) from the OECD (2010) database is taken as the indicator series for disaggregation of the CO_2 (eq.) data.

Methods

The various methods of Temporal disaggregation and ARIMA modelling which are already available in literature are discussed in this section.

Temporal Disaggregation Methods

Temporal Disaggregation methods used for disaggregating low frequency data to a higher frequency ensures that, the sum, the average, the first, or the last value of the resulting high frequency disaggregated time series is consistent with the low-frequency series. The five classical temporal disaggregation procedures viz., (a) Denton Process (Denton and Denton-Cholette method) (b) Chow-Lin method (c) Litterman method and (d) Fernandez method, are discussed below:

(a) Denton Process

Denton (1971) developed a simple disaggregation that tries to ensure consistency between the original and disaggregated series.

Let K represent intra-annual time periods, and the time series of interest has n=mk values and spans 'm' years. Let $y = (y_1 y_2 ......y_m)^t$ be a set of m annual totals to be disaggregated and $z = (z_1, z_2 ....z_n)^t$ be a set of original high-frequency values if available. The objective is to minimize the error between the new disaggregated series $x = (x_1, x_2 ....x_n)^t$ and the original vector z by a method which,

(1) Minimizes the distortion of the original series in some sense.
(2) Satisfies the condition that the 'k' values of the new series within each year sum to the given annual total for that year.

Mathematically, the objective is to choose 'x' so as to minimize a penalty function p(x,z) subject to the constraint, $\Sigma^{TK}_{(T-1)K+1} x_t = y_t$ for T = 1,2..., m

Let $p(x, z) = (x–z)'A(x–z)$ be a class of penalty functions, which is a quadratic form in the difference between original and adjusted time series values and let z be the $n \times n$ diagonal matrix with diagonal elements $z_1, z_2 ..., z_n$. The penalty function is minimized by distributing the discrepancy for each year in equal amounts among the k periods within the year. This leads to a discontinuity between the last period of one year and the first period of the next year. Therefore, a penalty function based on the difference of the original and adjusted series is introduced,

$p(x, z) = \Sigma^n_{t=1}(\Delta x_t – \Delta z_t)^2 = \Sigma^n_{t=1}[\Delta(x_t – z_t)^2]$ where $\Delta x_t = x_t – x_{t-1}$

Hence the penalty function includes the term,

$\Delta(x_1 – z_1) = (x_1 – z_1) – (x_0 – z_0)$

Where x_o & z_o refers to the period out of the range, then

$\Delta(x_1 – z_1) = x_0 – z_0$

The vector of backward first difference may be expressed as D(x – z), where D is an $n \times n$ matrix. Therefore, the quadratic form to be minimized is

$P(x, z) = (x – z)'D'D(x – z)$

although the difference $\Delta (x_1 – z_1) = x_1 – z_1$ affects the principle of movement preservation. Hence, Cholette (1984) resolved the issue by modifying the penalty function for the Additive first difference variant as $p(x,z) = \Sigma^n_{t=2}[\Delta(x_i – z_i)]^2$ and the method is referred to as Denton–Cholette process.

(b) Chow- Lin Method

A Best linear unbiased estimator approach was proposed by Chow and Lin (1971), to obtain the estimates of a monthly series by a regression model using related series

Let $y = X\beta + u$ ...(1)

be a multiple regression relationship with p related variables, $x_1, x_2 ..., x_p$, sampled during *4n* quarters, where y is 4n × 1 vector, *X is 4n* x *p* and *u* is a random vector with mean 0 and covariance matrix *V*. Let *C* be distribution matrix of order *n* x *4n*, that converts the *4n* quarterly observations into *n* annual observations C = [11110 ... 0 00 0: 0 ..... 1111].

Let $y. = Cy = CX\beta + Cu = X.\beta + u.$,

where y., X., u. denotes annual series, with $E(u.u.') = V. = CVC'$.

Now the objective is to estimate a vector z of m observations on the dependent variables, such that z is identical with y in the cases of interpolation and distribution and consists of observations outside the sample period in the case of extrapolation.

A linear unbiased estimator *z* of *z* satisfies, for some $m \times n$ matrix A,

$$\hat{z} = Ay_. = A(X_.\beta + u_.) \quad - (2)$$

and X_z *and* u_z denotes the variables in the regression model for z.

$$E(\hat{z} - z) = E[A(X_.\beta + u_.) - (X_z\beta + u_z)] = (AX_. - X_z)\beta = 0 \quad - (3)$$

From (2) and (3) we get,

$$AX_. - X_z = 0 \quad - (4)$$

$$\hat{z} - z = Au_. - u_z \quad - (5)$$

The covariance matrix of $(\hat{z} - z)$ is,

$$cov(\hat{z} - z) = AV.A' - AV_z - V_zA' + V_z \quad - (6)$$

with $V_z = Eu_.u_z'$, $V_{z.} = u_z u_.'$ and $V_z = u_z u_z'$

The best linear unbiased estimator (BLUE) of z is obtained by minimising the trace of cov($\hat{z}$–z) with respect to A, subject to the $m \times p$ matrix equations given in (4).

The resulting estimator is,

$$\hat{z} = Ay.=Xz\hat{\beta} + (V_{z.}V_{.}^{-1})\hat{u} \quad .. (7)$$

where

$$\hat{\beta} = (X_{.}'V_{.}^{-1}X_{.})^{-1}X_{.}'V_{.}^{-1}y_{.} \quad .. (8)$$

is the least square estimate of the regression using the n observations and

$$u = y. - X.\hat{\beta} \quad .. (9)$$

is the n × 1 vector of residuals in the regression using yearly data.

Estimation of covariance Matrix of the residual

The covariance matrix V of (7) has to estimate by assuming a structure to the residuals in (1). Chow and Lin (1971) proposed a method for estimation by assuming that regression residuals follow the first-order auto regression.

In the regression equation (1), $y = Xb + u$ it is assumed that, $u_t = \rho u_{t-1} + \varepsilon_t$ Where ε is WN(0,σ_ε) and |ρ| < 1. The resulting covariance matrix has the following form., V =

$$Euu' = \frac{\sigma_\varepsilon^2}{1-\rho^2}[1 \ \cdots \ \rho^{n-1} \ \vdots \ddots \vdots \ \rho^{n-1} \ \cdots \ 1]$$

$$=A \ \frac{\sigma^2}{1-\rho^2} \quad .. (10)$$

The AR1 parameter ρ needs to be estimated for the estimation of the covariance matrix V.

$$Vz.V.^{-1} = VC'(CVC')^{-1} = AC'(CAC')^{-1} \quad .. (11)$$

is required for interpolation and,

$Vz.V.^{-1} = (Eu_z u.')(CVC')^{-1}$ is required to be estimated in the case of extrapolation.

Since both (10) and (11) require the knowledge of ρ, a consistent estimator of ρ is required, which is obtained by observing the first-order autocorrelation coefficient of the annual residuals. The ratio of the second element to the first element on the first row of the matrix V.= CVC′ is Q=ρ^3, thus for interpolation C=C_1, ρ can be obtained from an initial guess of q. A set of regression residuals calculated from equations (10) and (9) are used for computing the first-order autocorrelation coefficient as the next guess of q.

Fernandez model

The quadratic loss function (QLF) approach by Denton (1971) and the best linear unbiased estimator (BLUE) approach by Chow and Lin (1971) yields the same results under the assumptions that high-frequency residuals are homoscedastic and serially independent when 54 is an identity matrix.

Fernandez (1981), proposed a new method by taking A =D′D Where D is the first difference transformation and considered the regression model, $X = z\beta + u$ where u_t follows the random walk model,

$u_t = u_{t-1} + \varepsilon_t$, ε_t bening a random variable, serially independent with zero mean and constant variance with $u_0 = 0$. Under these assumptions the residuals of the model is given by, $DX = DZb + Du$ *with* $E(Duu'D') = \sigma^2 I$.

Litterman Model

The random walk assumptions of the errors (u_t) defines a filter that removes all serial correlation in the residuals only when the model is correct. A generalization of Fernandez's approach suggested by Litterman (1983), helps in avoiding an adhoc search for an appropriate filter. This method uses a random walk Markov process i.e., AR (1) process for ε_t, in Fernandes's model $X = z\beta + u$, where ut follows the random walk model $u_t = u_{t-1} + \varepsilon_t$, and $\varepsilon_t = \varepsilon_{t-1} + \varepsilon_t$. Thus, both Fernandez and Litterman assume that the quarterly residuals follow a non-stationary process.

Auto-Regressive Integrated Moving Average (ARIMA) model

ARIMA model is a statistical time series model which is considered to outperform the econometric and regression models for time series data. It is one of the traditional methods applied for non-stationary time series analysis. ARIMA models were proposed by Box and Jenkins (1970). It aims at utilizing the autocorrelation present in the data, thus explained by its own lag values.

The autoregressive model of order (p) (denoted as AR (p)) can be written as,

$$y_t = c + \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \cdots + \varphi_p y_{t-p} + e_t$$

Where e_t the white noise term is $y_{t-1}, y_{t-2}, \ldots, y_{t-p}$ are the lag values of y_t

The moving average model of order (q) (denoted as MA (q)) can be written as,

$$yt = c + e_t + \theta_1 e_{t-1} + \theta_2 e_{t-2} + ... + \theta_q e_{t-q}$$

where e_t is the white noise term and $\theta_{t-1}, \theta_{t-2}, \dots \theta_{t-q}$ are the lag values of the previous forecasted errors.

A non–seasonal Auto-Regressive Integrated Moving Average (ARIMA) model is obtained by combining the differencing with Autoregressive Model and a Moving Average Model.

The full ARIMA model can be written as,

$$y'_t = c + \varphi_1 y'_{t-1} + \varphi_2 y'_{t-2} + ... + \varphi_p y'_{t-p} + \theta_1 e_{t-1} + \theta_2 e_{t-2} + ... + \theta_q e_{t-q} + e_t$$

where, y'_t is the differenced series $y_{t-1}, y_{t-2}, ...y_{t-p}$ are lagged values of y_t. The model is denoted as ARIMA (p, d, q) where, p, d and q are order of the autoregressive part, degree of differencing and order of the moving average respectively.

Hyndman and Khandakar (2008) proposed an automated algorithm for the selection of the model parameters that combines unit root tests, minimization of the AICc and MLE to obtain an ARIMA model. According to the algorithm considering the non-seasonal ARIMA (p, d, q) model, the differencing parameter is selected based on the KPSS unit root test and the other parameters (p and q) are selected using the following procedure:

Step 1

The parameter p and q will be selected by initially running the four following possible models, ARIMA (2, d,2), ARIMA (0, d,0), ARIMA (1, d,0), ARIMA (0, d,1)

If d =0 then the constant c is included, if d≥1 then the constant c is set to zero. A model with the lowest AICc value is set to be the current model. For ARIMA models corrected AICc can be written as

$$AICc = AIC + \frac{2(p+q+k+1)(p+q+k+2)}{T-p-q-k-2}, \text{ where } \quad AIC = -2\log(L) + 2(p+q+k+1),$$

L is the Likelihood of the data is the order of the autoregressive part, q is the order of the moving average part T is the length of the time series and

$$k = \begin{cases} 1 \text{ if } c \neq 0 \\ 0 \text{ if } c = 0 \end{cases}$$

Step 2

Variations in the current model are considered by

- Varying one of p or q by ± 1 in the current model
- Varying both p and q by ± 1 in the current model,
- Include/exclude constant c from the current model

A model with lower AICc will become the current model and the procedure is repeated until a model with no lower AICc can be detected in comparison with the current model.

Residual Analysis for ARIMA

ARIMA (p, d, q) model with appropriate parameters is fixed following the Hyndman and Khandakar algorithm and then the residuals of the model are checked for the Autocorrelation and Partial Autocorrelation. The Ljung-Box test is used to test the absence of serial correlation in the residuals of the fitted model up to a specified lag k. The Ljung-Box Q-Statistic is used for testing the null hypothesis "fitted model does not show a lack of fit´ against the alternative "fitted model shows a lack of fit".

The Ljung-Box Q-Statistic is defined as,

$$Q = n(n+2)\sum_{k=1}^{m}\frac{r_k^2}{n-k} \sim \chi^2 \; with \; (m-p)df$$

where n-the number of observations in the time series, k is the time lag, m is the number of time lags to be tested and n_k the sample autocorrelation function of the residuals lagged k time periods. Refer Hyndman and Athanasopoulos (2018) for a discussion on ARIMA models.

Statistical Analysis and Discussion

The CO_2 (eq.) emissions from four key sectors, namely ES, IPPUS, AFOLUS, WS have been disaggregated into quarterly data from 2005_Q1 to 2020_Q4 and then ARIMA model was fitted to the disaggregated data. The fitted model has been used for forecasting the CO_2 (eq.) emissions from 2021_Q1 to 2023_Q4.

Temporal Disaggregation of CO_2(eq.)

The estimates of regression coefficients along with the Adjusted R square and AR parameter are presented in Table 1. The estimates obtained by the three methods viz. Chow-Lin, Fernandez and

Litterman are compared using Adjusted R square values.

Comparing all the three models Chow - Lin model best predicts the quarterly emission with an approximate R^2 value of 0.9, for all the sectors except AFOLUS. The estimated regression coefficients of the three sectors ES, IPPUS and WS, shows a significant relationship between CO_2 (eq.) and GDP (with $p<0.05$). Thus CO_2 (eq.) for the above mentioned three sectors is significantly influenced by the current price GDP. Hence Current Price GDP is a good indicator of CO_2 (eq.) emission for the ES, IPPUS and WS. For disaggregating the AFOLUS Denton–Cholette method has been applied.

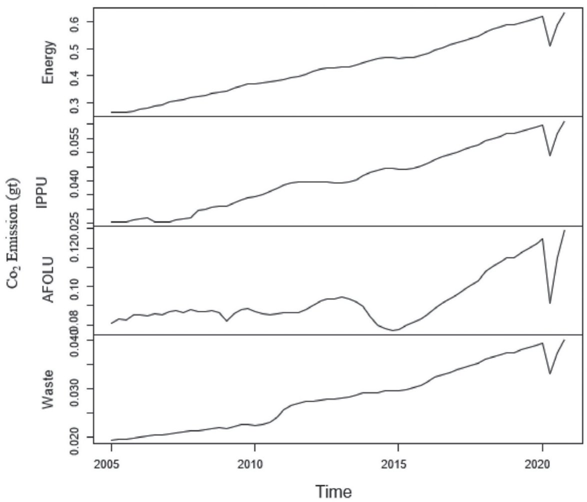


Fig. 1. Time series plot of Quarterly Emission for various sectors from 2005_Q1 to 2020_Q4

ARIMA Modelling and Forecasting CO_2 (eq.)

Model Identification

The disaggregated data for ARIMA modelling is obtained from the Chow-Lin method for the years 2005_Q1 to 2020_Q4 for ES, IPPUS and WS, while for AFOLUS the disaggregated data for the same period is obtained using the Denton-Cholette method.

The time plot for CO_2 (eq.) in Figure 1 for the four sectors, namely ES, IPPUS, AFOLUS and WS shows a progressive growth with no major atypical variations over time, indicating that data transformations is not required.

Table 1. Coefficients of Regression Estimates for various disaggregation methods

Sector		**Chow-lin**		Fernnadez		Litterman	
	Coefficients	(Intercept)	GDP	(Intercept)	GDP	(Intercept)	GDP
ENERGY	Estimate	0.2154	0.0079	0.2038	0.0069	0.2031	0.0070
	Std. Error	0.0197	0.0009	0.0160	0.0016	0.0160	0.0015
	t value	10.9330	9.0370	12.7500	4.3700	12.6670	4.6460
	Pr(>|t|)	0.0000**	0.0000*	0.6441	0.0018**	0.0000**	0.0012**
	Adjusted R-squared	0.8897		0.6441		0.6731	
	AR1-Parameter:	0.8951		0.0000		-0.602	
IPPU	Estimate	0.0202	0.0008	0.0197	0.0007	0.0196	0.0007
	Std. Error	0.0016	0.0001	0.0019	0.0002	0.0019	0.0002
	t value	12.9500	10.6000	10.2820	3.5740	10.2700	3.8610
	Pr(>|t|)	0.0000**	0.0000**	0.0000**	0.0060**	0.0000**	0.0038**
	Adjusted R-squared	**0.9176**		0.5407		0.5817	
	AR1-Parameter:	0.7984		0		-0.999	
AFOLU	Estimate	0.0871	-0.0001	0.0841	-0.0002	0.0840	-0.0002
	Std. Error	0.0043	0.0002	0.0071	0.0007	0.0071	0.0007
	t value	20.2120	-0.3380	11.9230	-0.2430	11.8800	-0.2200
	Pr(>|t|)	0.0000**	0.7430	0.0000**	0.8130	0.0000**	0.8300
	Adjusted R-squared	-0.09716		-0.1039		-0.1051	
	AR1-Parameter:	0.6943		0		0.1634	
WASTE	Estimate	0.0157	0.0005	0.0162	0.0004	0.0161	0.0004
	Std. Error	0.0007	0.0000	0.0014	0.0001	0.0014	0.0001
	t value	23.5000	14.4800	11.1550	2.6840	11.3180	2.9680
	Pr(>|t|)	0.0000**	0.00008*	0.0000**	0.0251*	0.0000**	0.0158*
	Adjusted R-squared	0.9543		0.3828		0.4385	
	AR1-Parameter:	0.4927		0		-0.999	

** Statistically Significant at 1% level, * Statistically significant at 5% level

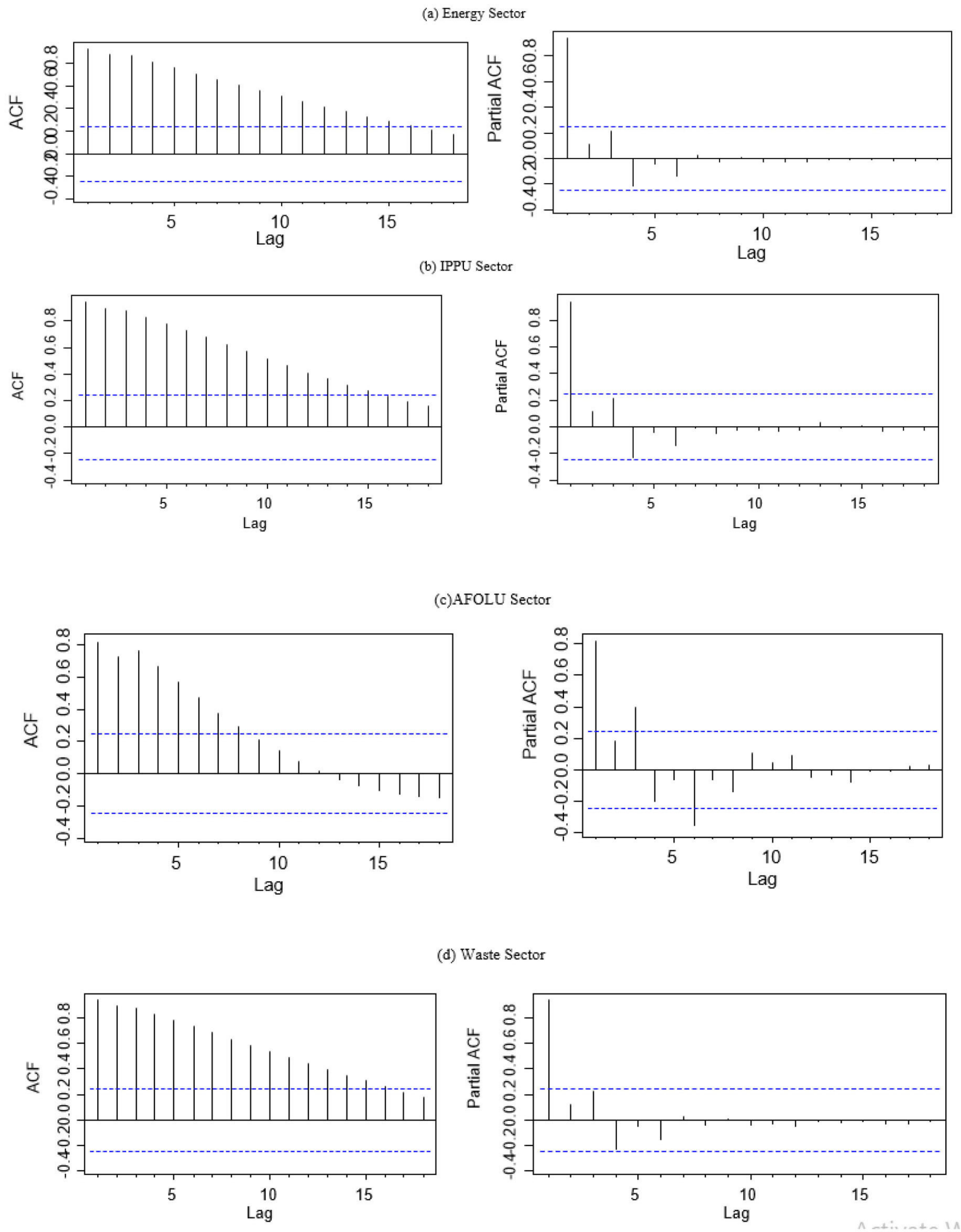


Fig. 2. ACF and PACF graphs for the CO_2 Emission from various Sectors

The presence of auto correlation in the series can be examined using ACF and PACF plots in Figure 2 and it is indicative of the presence of trend (i.e. non-stationary) which suggests that differencing is needed.

The disaggregated CO_2 (eq.) of ES, IPPUS, AFOLUS, WS are modelled using the automated algorithm and the models with minimum AICc, among the various fitted model using the combinations of the model parameters p, d and q, is summarized in Table 2.

The AICc value is minimum for ARIMA (2, 1, 0) model in all the sectors except AFOLU Sector, with AICc values -338.21, -620.59, -685.49 for Energy Sector, IPPU Sector and Waste sector respectively and the coefficients φ1 and φ2 of ARIMA (2,1,0) are statistically significant (p-value <0.01). The ARIMA (0, 1, 3) model has the minimum AICc value -476.72 for the AFOLU sector, among all the other models iterated with different combinations of the model parameters and the model coefficients θ_1, θ_2, θ_3 statistically significant (p-value<0.01).

Residual Analysis of the fitted model

The ARIMA model assumes that the residuals are normally distributed and not auto correlated. The graphs in Figure 3 provide a clear visual representation of the assumptions.

The absence of autocorrelation and partial autocorrelation among the residuals can be seen in the ACF and PACF graphs, which show no spikes in the first few lags for all four sectors. The pp-plot and Q-Q-plot show that the residuals are approximately normally distributed.

The Presence of White noise alone in the fitted ARIMA model, is evidenced by the Ljung Box Q Statistic in Table3 for the ES with p-value (0.9915), IPPUS (0.9799), AFOLUS (0.9896), and WS (0.9896) all (p value >0.05). The fitted model can be considered for prediction based on ACF, PACF plots, and the Ljung Box test.

Table 3. Residual Analysis

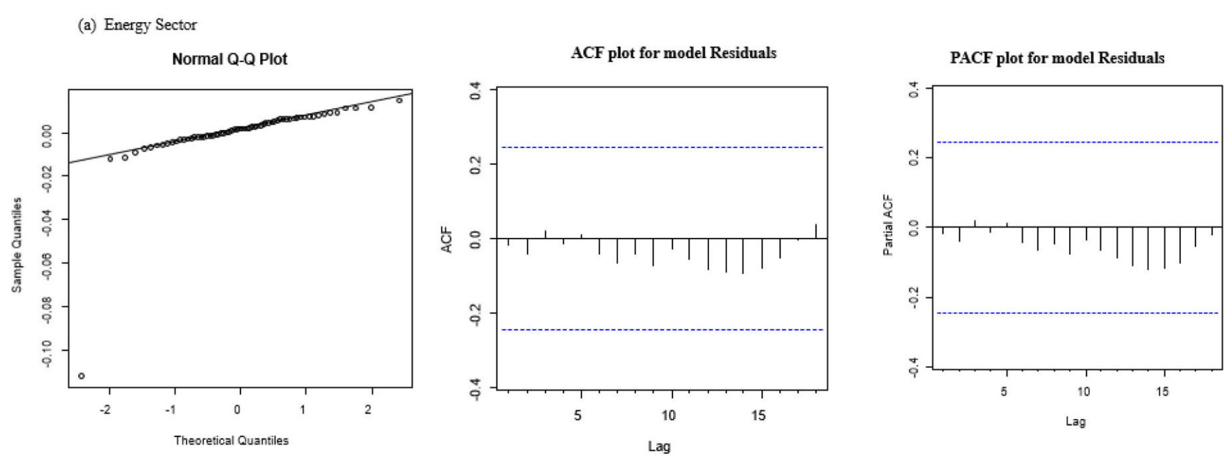
L-JUNG BOX TEST	Q-statistic	df	p-value
ENERGY	1.1756	7	0.9915
IPPU	1.567	7	0.9799
AFOLU	1.2551	7	0.9896
WASTE	1.2551	7	0.9896

Point Forecast using the fitted model

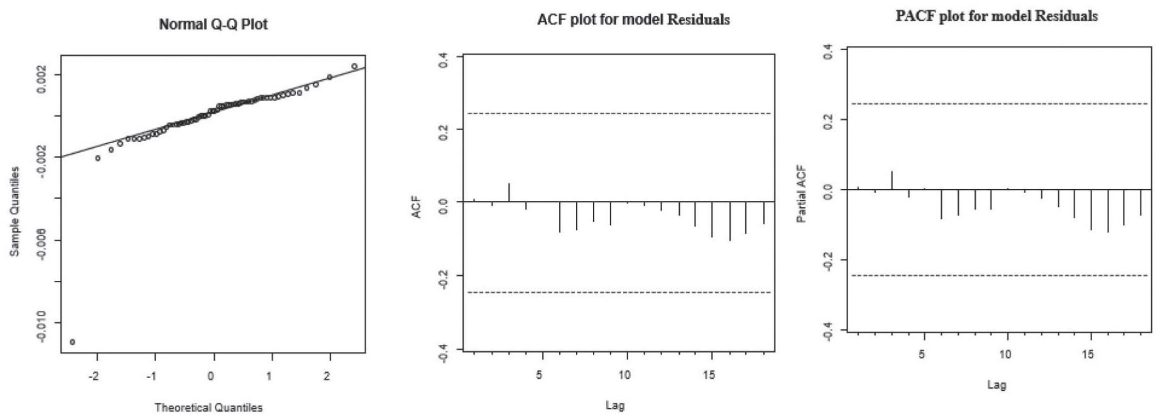
The CO_2 (eq.) for all four sectors have been forecast, using the sector wise fitted model for the period from 2021_Q1 to 2023_Q4, with a confidence interval of 80 percent to 95 percent and is presented in

Table 2. ARIMA model fitted to the series of different sectors

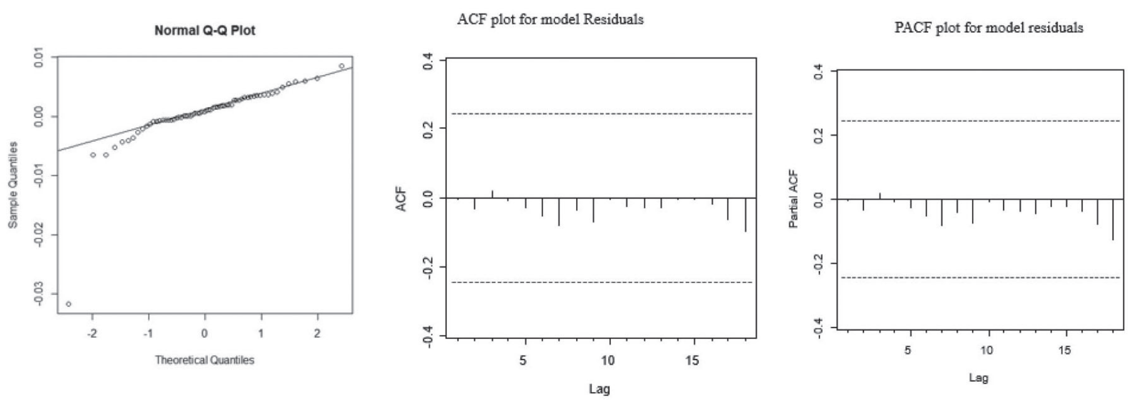
Sector	Best Model	Model Equation	AICc
Energy	ARIMA (2,1,0) with drift	$y_t=0.0053–0.5525y_{t-1} – 0.6445y_{t-2}+ \varepsilon_t$	-338.21
IPPU	ARIMA (2,1,0) with drift	$y_t=0.0005 – 0.4308y_{t-1} – 0.4986y_{t-2} + \varepsilon_t$	-620.59
AFLOU	ARIMA (0,1,3)	$y_t=0.657\varepsilon_{t-3} – 0.28775\varepsilon_{t-2} – 0.4479\varepsilon_{t-1}+\varepsilon_t$	-476.72
Waste	ARIMA (2,1,0) with drift	$y_t=0.0003 – 0.4181y_{t-1} – 0.4928y_{t-2} + \varepsilon_t$	-685.49



(b) IPPU Sector



(c) AFOLU Sector



(d) Waste Sector

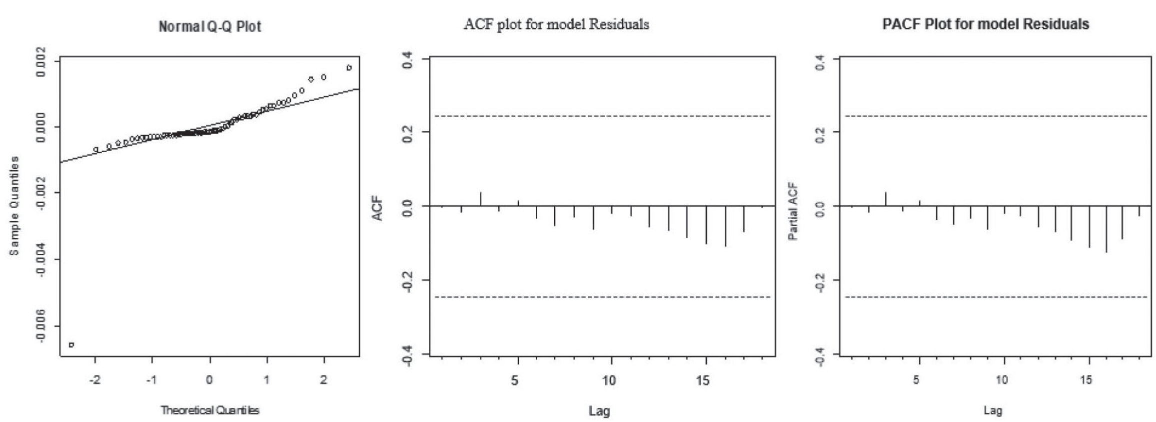


Fig.3. Normal Q-Q plot , ACF and PACF plots of the residual of the fitted model for various sectors

the following Figure 4.

The forecasted estimates for all the sectors except AFOLUS, reflect an increasing trend in emissions. The AFOLU sector's steady movement indicates that CO_2 absorption in the atmosphere is stable, however with a broad confidence interval.

The percentage of increase in emission for the predicted values shown in Figure 5. The emission from IPPUS and WS indicates almost a similar pattern and are modest, but the Emission from AFOLUS remains study. The Emission from ES shows a pronounced swing.

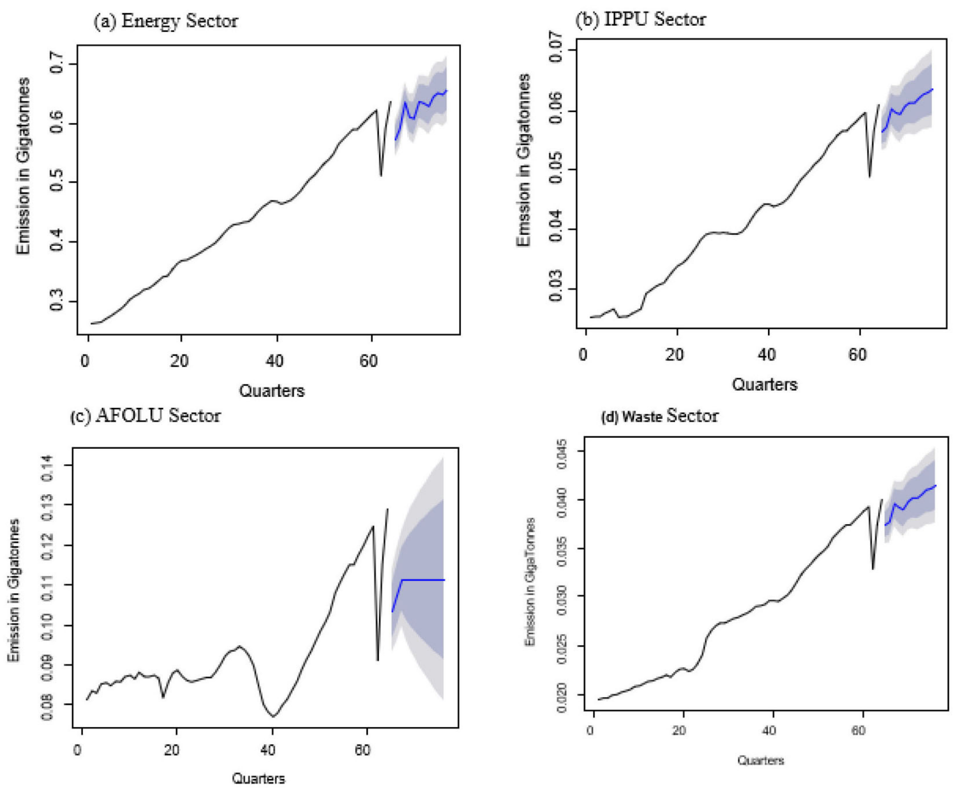
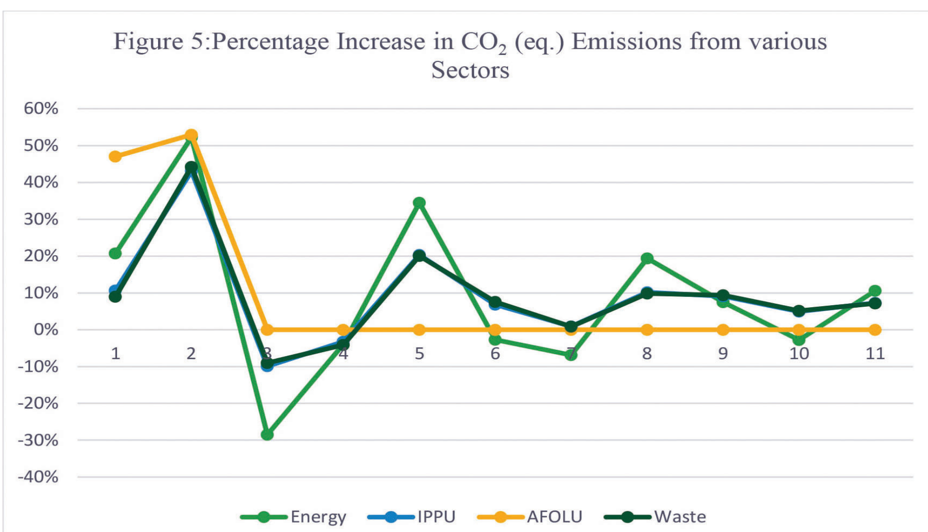


Fig.4. Emission forecast for various sector from 2021_Q1-2023_Q4



Conclusion and Recommendation

The Chow-Lin approach of disaggregating CO_2 (eq.) series using Current Price GDP as the indicator series was found to be best suited for the study. The sector-wise best fitted ARIMA model using the disaggregating CO_2 (eq.) were used for predicting CO_2 (eq.) for the period 2021_Q1 to 2023_Q4. The variation in percentage increase in predicted CO_2 (eq.) emissions for the Energy Sector indicates that, urgent action is needed to control the use of fossil fuels and promote clean energy sources for reduction of greenhouse gas emissions in the atmosphere. Suitable policies for increasing Carbon sink in particular, must be strengthened to keep up the Goal set by the Paris Agreement.

References

Ajao, I. O., Ayoola, F. J. and Iyaniwura, J. O. 2015. Temporal Disaggregation Methods in Flow Variables of Economic Data: Comparison Study. *International Journal of Statistics and Probability*. *5*(1) : 36. https://doi.org/10.5539/ijsp.v5n1p36

Alam, M., Murad, W., Hanifa, A. and Ozturk, I. 2016. Relationships among carbon emissions, economic growth, energy consumption, and population growth: Testing Environmental Kuznets Curve hypothesis for Brazil, China, India, and Indonesia. *Ecological Indicators*. 70 : 466–479. https://doi.org/10.1016/j.ecolind.2016.06.043

Ayodeji, S. D. I. 2012. Temporal Disaggregation of Time Series Data: A Non-Parametric Approach. *Mathematical Theory and Modeling*. *2*(5) : 7–14.

Chow, G. C. and Lin, A. 1971. Best Linear Unbiased Interpolation, Distribution and Extrapolation of Time Series by Related Series. *The Review of Economics and Statistics*. 53(4) : 372. https://doi.org/10.2307/1928739

Cholette, P.A. 1984. Adjusting sub-annual series to yearly benchmarks. *Survey Methodology*. 10(1).

Denton F.T. 1971. Adjustment of Monthly or Quarterly Series to Annual Totals: An Approach Based on Quadratic Minimization. *Journal of the American Statistical Association*. 66 : 333, 99-102, DOI: 10.1080/01621459.1971.10482227

Fernandez, R. B. 1981. A Methodological Note on the Estimation of Time Series. *The Review of Economics and Statistics*. 63(3) : 471. https://doi.org/10.2307/1924371

Ghosh, S. 2010. Examining carbon emissions economic growth nexus for India: A multivariate cointegration approach. *Energy Policy*. 38(6) : 3008–3014. https://doi.org/10.1016/j.enpol.2010.01.040

Guay, A. and Maurin, A. 2015. Disaggregation methods based on MIDAS regression. *Economic Modelling*, *50*, 123–129. https://doi.org/10.1016/j.econmod.2015.05.013

Hyndman, R. J. and Khandakar, Y. 2008. Automatic time series forecasting: The forecast package for R. *Journal of Statistical Software*. 27(3) : 1–22. https://doi.org/10.18637/jss.v027.i03

Hyndman J. and Athanasopoulos, G. 2018. Forecasting: Principles and Practice", 2nd edition, Otexts: Melbourne, Australia. Otexts.com/fpp2

Kanjilal, K. and Ghosh, S. 2013. Environmental Kuznets curve for India: Evidence from tests for cointegration with unknown structural breaks. *Energy Policy*. 56: 509–515. https://doi.org/10.1016/j.enpol.2013.01.015

Litterman, R. B. 1983. A Random Walk, Markov Model for the Distribution of Time Series. *Journal of Business & Economic Statistics*. 1(2) : 169–173. https://doi.org/10.1080/07350015.1983.10509336

Lotfalipour, M. R., Falahi, M. A. and Bastam, M. 2013. Prediction of CO_2 emissions in Iran using grey and ARIMA models. *International Journal of Energy Economics and Policy*. 3(3) : 229–237.

Nain, M. Z., Ahmad, W. and Kamaiah, B. 2017. Economic growth, energy consumption and CO_2 emissions in India: a disaggregated causal analysis. *International Journal of Sustainable Energy*. 36(8) : 807–824. https://doi.org/10.1080/14786451.2015.1109512

Nyoni, T. and Bonga, W. G. 2019. Prediction of CO2 Emissions in India Using ARIMA Models. *DRJ-Journal of Economics & Finance*. 4(2) : 1–10.

Pal, D. and Kumar, S. 2017. The environmental Kuznets curve for carbon dioxide in India and China: Growth and pollution at crossroad. *Journal of Policy Modeling*, 39(2) : 371–385. https://doi.org/10.1016/j.jpolmod.2017.03.005

Pao, H. T. and Tsai, C. M. 2011. Modeling and forecasting the CO2 emissions, energy consumption, and economic growth in Brazil. *Energy*. 36(5): 2450–2458. https://doi.org/10.1016/j.energy.2011.01.032

Rahman, A. and Hasan, M. M. 2017. Modeling and Forecasting of Carbon Dioxide Emissions in Bangladesh Using Autoregressive Integrated Moving Average (ARIMA) Models. *Open Journal of Statistics*. 07(04): 560–566. https://doi.org/10.4236/ojs.2017.74038

Sen, P., Roy, M. and Pal, P. 2016. Application of ARIMA for forecasting energy consumption and GHG emission: A case study of an Indian pig iron manufacturing organization. *Energy*. 116: 1031–1038. https://doi.org/10.1016/j.energy.2016.10.068

OECD. 2010. Main Economic Indicators- Complete database, Main Economic Indicator (database) http://dx.doi.org/10.1787/data-0005.