Time Series Modeling and Forecasting on Carbon dioxide Emission in Sri Lanka

K.P.Hansika[1]* and N.R.Abeynayake[2]
[1,2]Postgraduate Institute of Agriculture, University of Peradeniya
pavanih92@gmail.com

Abstract

The continuous rise of anthropogenic carbon dioxide (CO2) atmospheric emissions is a major cause of global warming with adverse environmental effects. Most developed countries have a higher share of annual CO2 emissions in global emissions. As a developing country, Sri Lanka's annual CO2 emission levels are lowest, but the total annual CO2 emission has increased at an annual growth rate of 5.06%. As a signatory of the Kyoto Protocol and the Paris Agreement, and being highly vulnerable to climate change, Sri Lanka commits to reduce its CO2 emissions. Valid database analysis of CO2 emission modeling and forecasting in Sri Lanka will help to policy makers in reducing CO2 emissions in Sri Lanka. In the present study, different Autoregressive Integrated Moving Average (ARIMA) models were developed to model the CO2 emission by using time series data from 1950-2019. The performance of these developed models was assessed with different selection measure criteria, and the model having the minimum value of these criteria was considered as the best forecasting model. Based on findings, ARIMA (0, 1, 1) is the best fitted model in predicting the emission of CO2 in Sri Lanka. Vector Auto Regressive with exogenous variable (VARX) model was used to assess the impact of energy consumption, GDP and urban population on CO2 emission in Sri Lanka with the time series secondary data from 1965 to 2019. Based on the results, the VARX (1,3) model was the best model for the relationship among these variables.

Keywords: *Carbon Dioxide, Modeling, Forecasting, Time Series, ARIMA, VARX*

1. Introduction

Atmospheric CO2 is the primary carbon source for life on earth. All living creatures produce this CO2 through respiration, decaying of organic material, combustion of fossil fuels and fermentation processes. Generally, the emission of CO2 to the atmosphere is balanced by the same amount being removed to the atmosphere by plant photosynthesis and by the oceans. Human activity has disturbed this equilibrium significantly by generating increased levels of CO2 from excessive consumption of fossil fuels, combustion and by deforestation. These imbalances have a greater impact on the enhanced greenhouse effect (Razmjoo and Davarpanah, 2019). Greenhouse effect is a natural process that warms earth's surface by downward radiation of radioactive greenhouse gases which includes carbon dioxide (CO2), methane (CH4), nitrous oxide (N2O), hydrofluoro- carbons (HFCs), perfluorocarbons (PFCs) and sulphur hexafluoride (SF6) (EPA, 2014). This process of absorbing part of sun's energy and re-radiation, maintain the earth temperature around 33oC and it may around 14oC without this greenhouse effect. Increasing the concentration of greenhouse gases due to anthropogenic emission has caused the enhanced greenhouse effect, referred to as global warming. Global warming is one of the biggest threats the world has faced in the recent centuries. In 2007, the Intergovernmental Panel on Climate Change (IPCC) reported that there would be an estimated rise in the average global temperature between 1.1oC and 6.4oC within the next 100 years. This cause adverse effects such as melting of polar ice cap, sea level rise and ocean acidification which will hamper many eco systems.

Among the greenhouse gases CO2 is the most dominant which accounted 77% of the total global greenhouse gas emission (Parry, 2007). According to environmental research, in 1960 CO2 level was about 270 (ppm) but it reached at 405 (ppm) in 2017 (Samreen et al, 2019). Increase in population, industrial revolution, urbanization aggravates the continuous rise in CO2 emission in most of the developed countries and have a much higher share in global emissions than the developing ones (Nebojsa, 1994). China is the Asia's and the world's largest emitter. It emits nearly ten billion tons each year, more than one-quarter of global emissions. North America is the second largest regional emitter at 18% of global emissions. It's followed closely by Europe with 17% (Hannah et al, 2020). Sri Lanka as a developing country is a low carbon emitting country with a global share of less than 0.1%. A recent analysis showed that Sri Lanka has achieved both high human development and managed to keep CO2 emissions below the long-term average needed to contain global warming targets of the Paris Agreement (Ministry of Environment, 2021). However, Sri Lanka's total annual CO2 emission was increased from 3.1 million tons in 1970 to 21.11 million tons 2020, growing at an annual growth rate of 5.06% (Knoema, 2021). Due to this trending increment of CO2 emissions and highly vulnerable to climate change, Sri Lanka commits to reducing its CO2 emissions. Sri Lanka as signatory of the United Nations Framework Convention on Climate Change (UNFCCC) and Kyoto Protocol has made a significant contribution towards strengthening of national policy, legal and institutional capabilities to reduce the greenhouse gas emissions. To realize these policies and to implement action plans at the right time in future; valid database analysis is essential to identify the Sri Lanka's CO2 emission path and make reliable prediction of its future emissions. The objective of this study is to model and forecast the CO2 emission in Sri Lanka by developing an analytical statistical time series model. Therefore in this study, the time series secondary data on total annual CO2 emission from 1950 to 2019 were analyzed with ARIMA models to select the most appropriate model for prediction. Further, Vector Auto Regressive with exogenous variable VARX model was constructed to assess the impact of energy consumption, GDP and urban population on CO2 emission in Sri Lanka with the time series secondary data from 1965 to 2019. Forecasted values of CO2 emission in Sri Lanka were obtained from the most appropriate model. The results of the study can be used in policy making to implement necessary actions to minimize the effect of CO2 emissions by emphasizing environment friendly systems at the right time in future.

2. Materials and Methods

Materials

In the present study, time series secondary data on total annual CO2 emissions in Sri Lanka from 1950 to 2019 were considered for modeling and forecasting CO2 emissions in Sri Lanka. Time series secondary data on energy consumption, GDP, and urban population in Sri Lanka from 1965 to 2019 were also considered for the construction of the VARX model. All the data was sourced from the World Bank online database.

Methods

ARIMA Model

Box-Jenkins ARIMA, which is one of the commonly used methods of non-stationary time series analysis of environmental, financial, energy and engineering data (Bowden and Payne, 2008 and Pappas, et al., 2008), was used in the analysis of total CO_2 emission in Sri Lanka. The ARIMA model explains time series by its past or lagged values and stochastic error terms. ARIMA models use a combination of autoregressive (AR), integration (I), refers to the reverse process of differencing to produce the forecast and moving average (MA) operations (Box and Jenkins, 1970).

Auto Regressive (AR) models forecast the variable of interest using a linear combination of past values of the variable. AR model of order "p" denoted by AR (p) and is written as,

$$Y_t = C + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \phi_3 Y_{t-3} + \ldots + \phi_p Y_{t-p} + \varepsilon_t$$

where, Y_t is time series data, C is a constant, ε_t is white noise, assumed to be independently distributed across time with mean 0 and variance σ^2 and $\phi_1, \phi_2, \ldots, \phi_p$ are autoregressive coefficients.

Moving Average (MA) model uses past forecast errors in a regression like model. MA model of order q, MA (q) can be written as,

$$Y_t = C - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \theta_3 \varepsilon_{t-3} - \ldots - \theta_q \varepsilon_{t-q} + \varepsilon_t$$

where, C is a constant, ε_t is white noise and $\theta_1, \theta_2, \ldots, \theta_q$ are moving average (MA) coefficients.

An ARMA model is the sum of an Autoregressive and Moving Average processes proposed by Box and Jenkins. Mathematically an ARMA model is defined by the following equation,

$$Y_t = C + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \ldots + \phi_p Y_{t-p} + \varepsilon - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \ldots - \theta_q \varepsilon_{t-q}$$

One extension to the ARMA (p, q) model which greatly enhances the value as empirical descriptors of non-stationary time series is the class of autoregressive–integrated-moving average (ARIMA) models. The number of times 'd' that the integrated process must be differentiated to make stationary is said to be the order of the integrated process and is called non seasonal ARIMA (p, d, q) model which can be written as,

$$\grave{Y}_t = C + \phi_1 \grave{Y}_{t-1} + \phi_2 \grave{Y}_{t-2} + \ldots + \phi_p \grave{Y}_{t-p} + \varepsilon - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \ldots - \theta_q \varepsilon_{t-q}$$

where, $\grave{Y}_t$ is the differenced series.

The steps followed to define an ARIMA model, as stated by Box & Jenkins, are determining the stationarity of the model, model identification, parameter estimation, diagnostic checking, and forecasting. Identification of non-stationarity of data series was done by the Augmented Dickey-Fuller (ADF) unit root test, where null hypothesis(H_0): there is a unit root, which means data is not stationary and alternative hypothesis(H1): there is no unit root, which means data is stationary. Identification of specific number and type of parameters to be estimated using series plots, correlograms of autocorrelation (ACF) and partial autocorrelation (PACF). Mean Squared Error (MSE), Mean Absolute Percentage Error (MAPE) and Akaike Information Criteria (AIC) were used as model selection criteria to compare different parametric combinations of ARIMA (p,d,q) on the basis of minimum value.

MSE = $\frac{1}{n}\sum_{t=1}^{n} |e_t^2|$ where, e_t is forecast error

MAPE = $\frac{1}{n}\sum_{t=1}^{n} |PE_t|$

Where, $PE_t = \frac{Yt-Ft}{Yt} * 100$

PE_t =Percentage Error at t time

Y_t= Observed value at t time

F_t =Forecasted value at t time

AIC = -2 log L + 2m where, L= Likelihood and m= number of parameters excepted

After estimating a tentative model, diagnostic checking was done by generating the set of residuals and testing whether they satisfy the characteristics of a white noise process. A run chart and normal probability plot were carried out to check the randomness and normality of the residuals of selected models.

If the characteristics of the white noise process are not satisfied, model re-selection and repetition from model identification, which is the second stage in the Box-Jenkins ARIMA procedure, are followed until the model's efficiency is achieved.

VARX Model

The VARX model is one of the statistical analyses frequently used in many studies involving time series data, used to model several endogenous variables interconnected and influenced by the previous time lags and between exogenous variables that affect the endogenous variable. Based on literature, energy consumption, GDP and the urban population were selected as the CO_2 emission impact factors in Sri Lanka to develop VARX model. Total annual CO_2 emission as an endogenous variable and energy consumption, as well as GDP and urban population (people living in urban areas as defined by national statistical offices) as exogenous variables, were considered in constructing the VARX model.

VARX modeling is done by simultaneously estimating the VAR model and exogenous variables that influence the model. The general form of VARX model is as follows:

$$Y_t = c + \varphi_1 Y_{t-1} +\ldots+ \varphi_p Y_{t-p} + \beta_1 EC_t +\ldots+ \beta_p EC_{t-p} + \theta_1 GD_t +\ldots+ \theta_p GD_{t-p} + \delta_1 UP_t +\ldots+ \delta_p UP_{t-p} + \mu_t$$

Where Y_t represents current value of total CO_2 emission, EC represents energy consumption, GD represents GDP, UP represents urban population and φ, β, θ, δ are the relevant coefficients.

The optimal VARX model was selected using AIC values, and its stationarity was tested with an AR unit root test.

The VARX model's diagnostic test is done by the multivariate extensions of the Jarque-Bera residual normality test, which compares the residuals' third and fourth moments to those from the normal distribution. The null hypothesis of the test is that the residuals are multivariate normal, and the orthogonalization method selected was Cholesky of covariance (Lutkepohl).

To test whether there is a correlation between residuals, the Portmanteau test was carried out. The null hypothesis is that there are no residual autocorrelations up to lag h. The null hypothesis is accepted if the p-value is greater than 0.05; thus, the residuals are not correlated.

Forecasting of the estimated VARX model is done by solving the model in a deterministic simulation where all equations in the model are solved so that they hold without error during the solution sample, all coefficients are held fixed at their point estimates and all exogenous variables are held constant at their historical values in the sample. The VARX model results are evaluated by calculating the MAPE value for the forecast results of testing data.

3. Result and Discussion

Table 1 Descriptive statistics of total annual CO_2 emission in Sri Lanka from 1950-2019

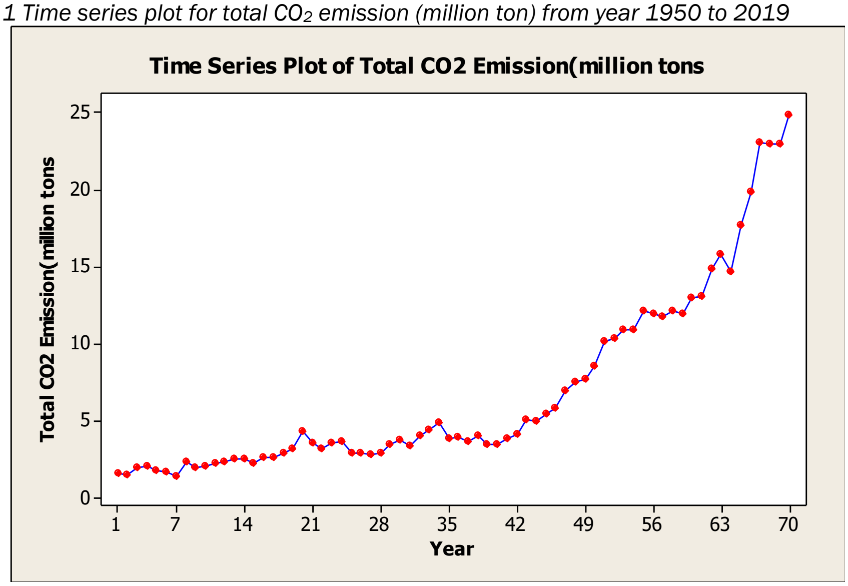
Description	Statistic
Mean	6.902571
Median	3.885000
Maximum	24.84000
Minimum	1.420000
Standard Deviation	6.056319
Skewness	1.432886
Kurtosis	4.151132

Table 1 presents descriptive statistics of CO2 emissions in Sri Lanka. The wide gap between the minimum CO2 emission (1.42 million tons in 1956) and the maximum CO2 emission (24.84 million tons in 2019) implies that the CO2 emission series is sharply trending upwards. The skewness is 1.432886, where it is positively skewed and non-symmetric. Kurtosis is 4.151132, indicating that the CO_2 emission data series is not normally distributed.

Modeling and forecasting CO_2 emission in Sri Lanka

To identify the CO_2 emission pattern in Sri Lanka, a time series plot was created for the data obtained from 1950 to 2019 (Figure 1). Time series plot of total CO_2 emission shows an increasing trend.

Figure 1 Time series plot for total CO_2 emission (million ton) from year 1950 to 2019



ARIMA Model

A common assumption in ARIMA modeling and forecasting is that the data are stationary. Time series plot of total CO_2 emission revealed that the series was non-stationary as it displays a trend or random walk behavior, and the variance was not constant over time. Log transformation was applied to stabilize the variance (Figure 2). Results of the Augmented Dickey-Fuller unit root test confirm that the transformed series of CO2 emission was not stationary, so that the series was differenced once and the null hypothesis (data is non-stationary) of the test was rejected at first difference for the transformed series (Table 2) which confirms that the series is stationary.

Figure 2 Time series plot of transformed data of CO_2 emission from year 1950 to 2019

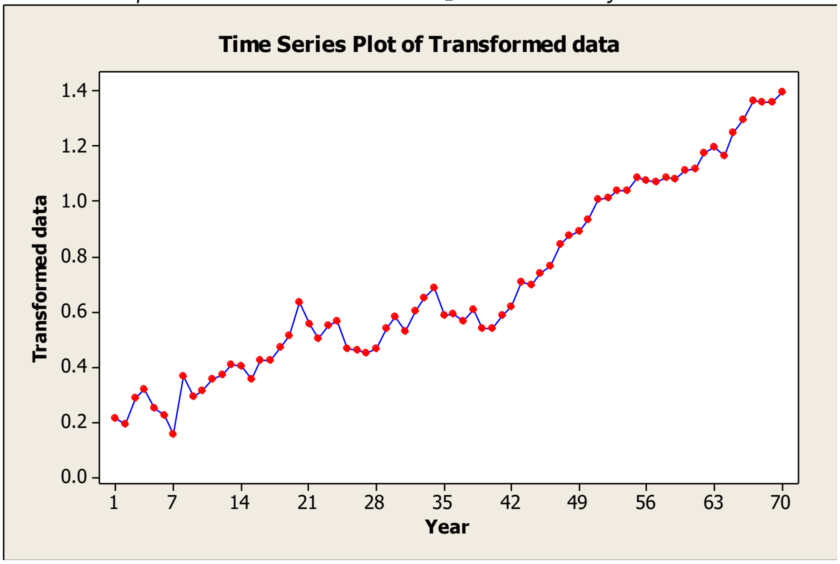


Table 2 Augmented Dickey Fuller Unit Root test for data of total annual CO_2 emission in Sri Lanka before and after differentiation

Variable	Type	P value of level data	P value after 1st differencing
Total annual CO_2 emission	Intercept	0.9711	0.0001
	Trend & intercept	0.4812	0.0000
	None	0.9967	0.0000

After making the series stationary, different parametric combinations of the ARIMA (p, d, q) model were tried to analyze the seventy-year data (1950 to 2019) of CO2 emission. The best-fitted model was accepted based on the minimum value of the selection criteria as per the methods. The results of the performance of the developed ARIMA (p, d, q) models with minimum MS, MAPE, and AIC values were presented in Table 3.

Table 3 Performance of developed ARIMA (p, d, q) models

Model	Model Selection Criteria		
	MS	MAPE	AIC
ARIMA (0,1,1)	0.002716	8.50962	-3.042003
ARIMA (0,1,2)	0.002749	8.51219	-2.993164

ARIMA (1,1,0)	0.002740	8.51973	-3.025765
ARIMA (1,1,1)	0.002749	8.50421	-3.005615
ARIMA (1,1,2)	0.002791	8.51204	-3.00721
ARIMA (2,1,0)	0.002756	8.49669	-3.005290
ARIMA (2,1,1)	0.002790	8.52686	-3.021753
ARIMA (2,1,2)	0.002748	8.44277	-3.110268
ARIMA (3,1,0)	0.002764	8.60359	-2.992831
ARIMA (3,1,1)	0.002797	8.31800	-3.007294
ARIMA (3,1,2)	0.002255	7.24933	-2.963012

According to Table 3, the lowest MAPE value was reported on ARIMA (3,1,2), but the parameters were insignificant in that model. (p value 0.566 > 0.05 for AR 3) ARIMA (2,1,2) and ARIMA (0,1,1) are the following two models having the minimum value of all selection criteria. Out of which all the parameters were significant in the model ARIMA (0,1,1) as in ARIMA (2,1,2) the p value of AR 2 component was not significant (0.191). Therefore, it was concluded that the appropriate model for forecasting the CO_2 emission in Sri Lanka is ARIMA (0, 1, 1), which has a constant minimum selection criteria value and significant parameters compared to other models. The final estimates of parameters of developed ARIMA (0, 1, 1) model were presented in Table 4. P values at lags 12, 24, 36 and 48 were above 0.05 in Modified Box-Pierce (Ljung-Box) Chi-Square statistic revealed no auto correlation in residuals (Table 5).

Table 4 Final estimates of parameters of ARIMA (0 1 1) model

Type	Coef.	SE Coef.	T	P value
MA1	0.2449	0.1185	2.07	0.043
Constant	0.017172	0.004739	3.62	0.001

Table 5 Modified Box-Pierce (Ljung-Box) Chi-square statistic of ARIMA (0 1 1) model

Lag	12	24	36	48
Chi-square	12.2	14.2	25.5	35.3
DF	10	22	34	46
P value	0.272	0.896	0.852	0.875

Diagnostic checking

The residuals of the model ARIMA (0,1,1) were tested for normality and randomness. The normal probability plot and Run Chart proved that the residuals were random (Cluster probability is 0.729) and normal (Anderson Darling p = 0.256). Figures 3 and 4 represent the relevant graphs.

Figure 3 Run chart for residuals

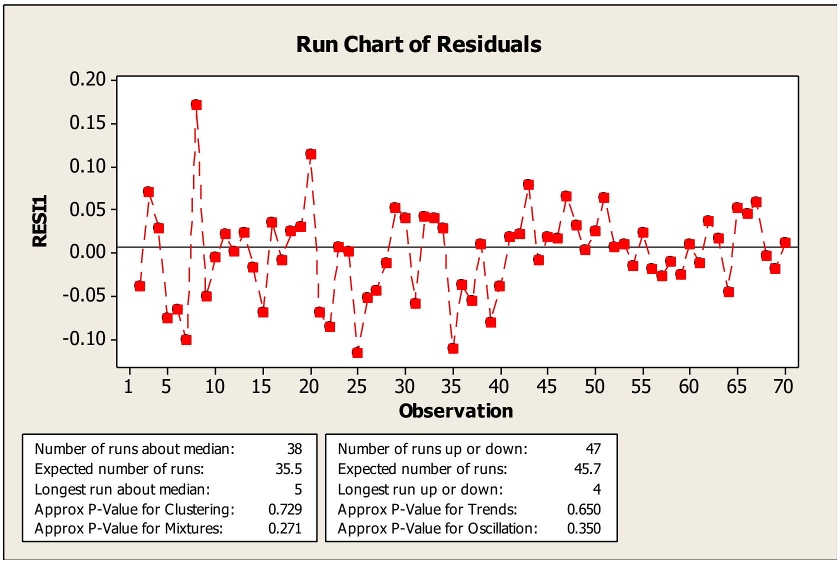
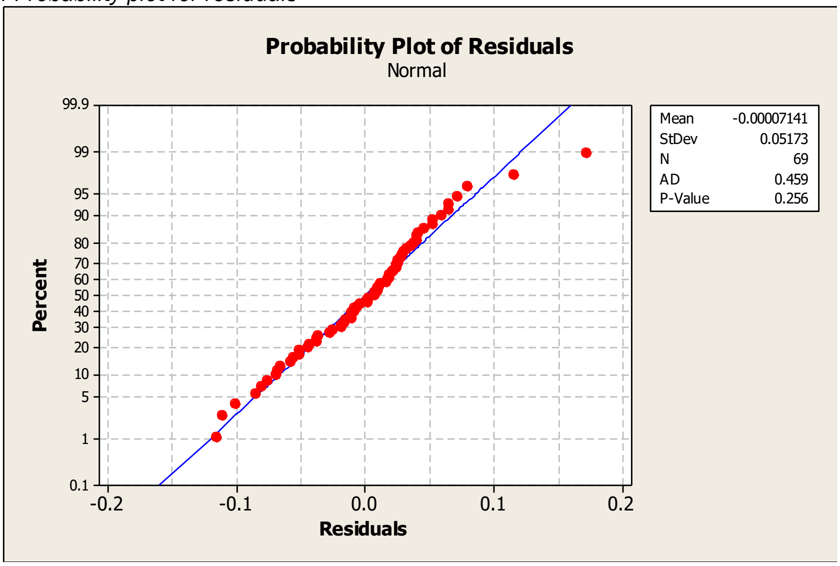


Figure 4 Probability plot for residuals



ARIMA (0,1,1) was the appropriate model amongst all models tested, for predicting the CO_2 emission in Sri Lanka. The ARIMA (0,1,1) model with constant has the prediction equation of,

$Y_t = C + Y_{t-1} - \theta_1 e_{t-1}$
$Y_t = 0.017172 + Y_{t-1} - 0.2449 e_{t-1}$

Y_t is forecasted value, C is constant, θ_1 is MA co-efficient and e_{t-1} is MA component.

The model performed well in explaining variability in the data series and in its predicting ability. The forecasted values of total CO2 emissions (million tons) in Sri Lanka from 2020 to 2025 were obtained using the best-fitted model ARIMA (0,1,1) (Table 6).

Table 6 Forecast of CO_2 emission (million tons) in Sri Lanka using ARIMA (0,1,1) during 2020 to 2025 (95% confidence interval)*

Year	Forecast	Lower*	Upper*
2020	25.6626	20.2829	32.4691
2021	26.6975	19.8811	35.8509
2022	27.7741	19.6866	39.1850
2023	28.8948	19.6155	42.5628
2024	30.0601	19.6304	46.0310
2025	31.2723	19.7088	49.6204

In 2037, total CO_2 emission in Sri Lanka will be 50 million tons according to the prediction equation. This implies that Sri Lanka will continue to face challenges of global warming and be more susceptible to climate change shortly. By using this prediction equation, Sri Lanka can be better prepared and aware when controlling the emissions in future. However, total carbon emissions of most countries worldwide, including Sri Lanka, were reduced in 2020 compared to 2019 due to lockdowns in the first half of the year during the COVID-19 pandemic. These kinds of situations may affect the accuracy of the model. However, according to Bhanumati, 2022 global CO2 emissions will be back above the pre-pandemic levels again in 2021, which suggests many deviations from the prediction models.

Construction and Testing of VARX model

Correlograms of ACF and PACF and the results of the Augmented Dickey-Fuller (ADF) unit root test shows that the three variables, energy consumption, GDP and urban population were not stationary, so that they were differenced, then viewed the ACF and PACF plots again and the formal results of the Augmented Dickey-Fuller unit roots. Results of the ADF test are presented in Table 7.

Table 7 Augmented Dickey Fuller Unit Root test for energy consumption data, GDP and urban population before and after differentiation.

Variable	Type	P value of level data	P value after 1st differencing	P value after 2nd differencing
Energy consumption	Intercept	0.9999	0.0000	0.0000
	Trend & intercept	0.9785	0.0000	0.0000
	None	1.0000	0.0000	0.0000
GDP	Intercept	1.0000	0.5247	0.0000
	Trend & intercept		0.0242	0.0000
	None	1.0000	0.1881	0.0000
		1.0000		
		1.0000		
Urban population	Intercept	0.8669	0.5264	0.0000
	Trend & intercept	0.2147	0.9211	0.0000
	None	0.9284	0.3824	0.0000

The optimal lag order of the VARX model was determined before constructing the VARX model. Lag order refers to the number of previous observations in time series used as predictors in the VAR model. Models with different lag orders with CO_2 emission as endogenous variable and energy consumption, GDP and urban population as exogenous variables were constructed through Eviews8 software. Selection of an appropriate lag is critical to build a parsimonious model as using too few lags can result in autocorrelated errors whereas using too many lags results in over-fitting, causing an increase in mean square forecast errors of the VAR model. The lag length for the VAR (*p*) model was determined using model selection criteria by fitting VAR(*p*) models with orders $p = 0, 1, \ldots, p_{max}$ and choosing the value of p which minimises the AIC, which is one of the commonly used model selection criteria. It was obtained that the smallest AIC value from the VAR analysis using the CO_2 emission in the lag 4 amounted to -3.240481 and lag five amounted to -3.241634. Lag 4 was selected to include in the model as there is not much more difference in the AIC value, and using a longer lag will result in more forecast errors. Then after over fitting from VARX (1,1) to VARX (4,3), the AIC values of each model are obtained as follows (Table 8).

Table 8 Overfitting VARX models

VARX model	AIC value
VARX (1,0)	-3.165220
VARX (1,1)	-3.153763
VARX (1,2)	-3.146833
VARX (1,3)	-3.295259
VARX (2,0)	-3.123166
VARX (2,1)	-3.132057
VARX (2,2)	-3.115098
VARX (2,3)	-3.278617
VARX (3,0)	-3.085118
VARX (3,1)	-3.087767
VARX (3,2)	-3.082664
VARX (3,3)	-3.246259
VARX (4,0)	-3.240481
VARX (4,1)	-3.250343
VARX (4,2)	-3.238149
VARX (4,3)	-3.274274

Based on the AIC values of the VARX models estimated in Table 8 above, the smallest AIC value was obtained in VARX (1,3), which was selected as the optimal model. To test whether the selected VARX model is effective, the AR unit root test was used to test its stability. As shown in the graph of AR root test results of the VARX model (Figure 5), the characteristic roots were all in the unit circle, indicating that the VARX model is stable.

Diagnostic Checking of selected VARX model

The residual normality test results (orthogonalization: Cholesky of covariance) revealed that the residuals are multivariate normal as the probability values are greater than 0.05 and the null hypothesis is not rejected (Table 9).

Based on the results of the Portmanteau test, almost until the 12th lag had a p-value greater than 0.05, which indicates that the residuals are not correlated (Table 10).

Figure 5 Inverse roots of AR characteristic polynomial of VARX (1,3)

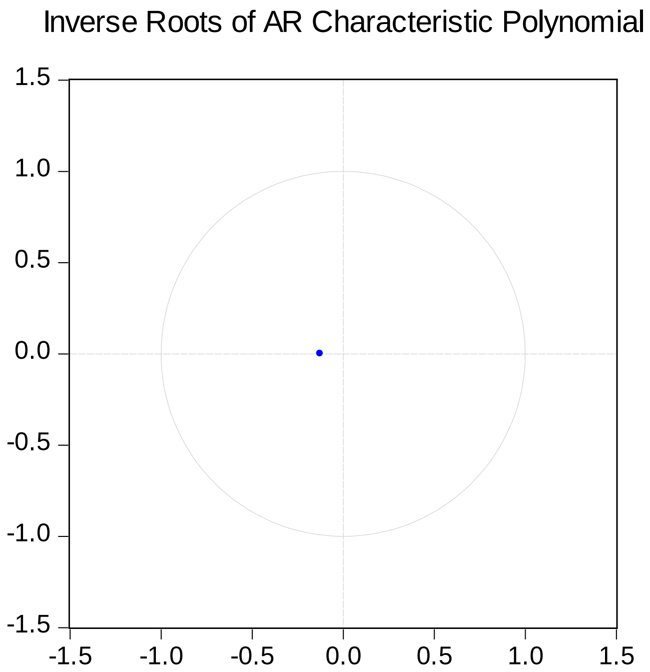


Table 9 Results of the residual normality test (orthogonalization: Cholesky of covariance)

Component	Skewness	Chi-sq	df	Prob.
1	-0.312557	0.814100	1	0.3669
Joint		0.814100	1	0.3669

Component	Kurtosis	Chi-sq	df	Prob.
1	2.702164	0.184805	1	0.6673
Joint		0.184805	1	0.6673

Component	Jarque-Bera	df	Prob.
1	0.998904	2	0.6069
Joint	0.998904	2	0.6069

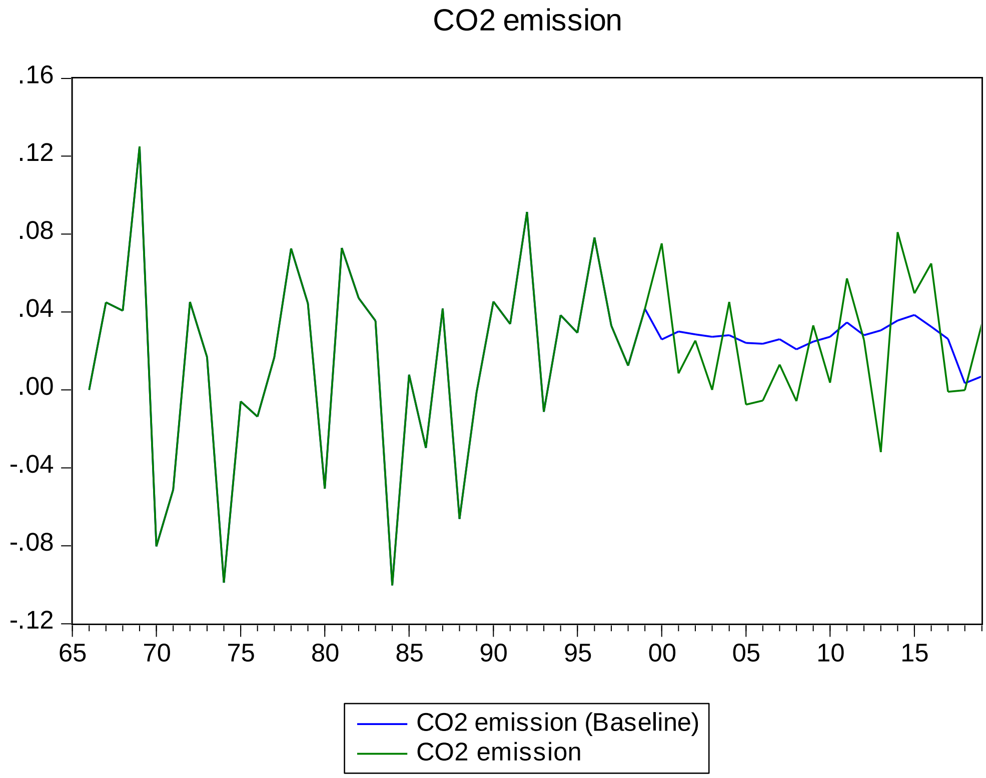
Table 10 Results of the Portmanteau test

Lags	Q-Stat	Prob.	Adj Q-Stat	Prob.	df
1	0.268997	NA*	0.274487	NA*	NA*
2	0.419910	0.5170	0.431687	0.5112	1
3	0.565531	0.7537	0.586604	0.7458	2
4	3.231374	0.3573	3.484259	0.3228	3
5	3.352448	0.5007	3.618786	0.4600	4
6	4.973448	0.4191	5.460831	0.3623	5
7	10.14918	0.1185	11.47913	0.0747	6
8	13.24583	0.0663	15.16561	0.0339	7
9	13.26382	0.1031	15.18756	0.0556	8
10	13.41126	0.1449	15.37185	0.0812	9
11	15.06955	0.1295	17.49786	0.0640	10
12	16.32148	0.1296	19.14514	0.0585	11

Forecasting and evaluation of VARX model

Solving the model over the forecast period achieved a forecast from the VARX. Upon solving the model, forecasted values were created for the endogenous variable: total annual CO_2 emission (Figure 6).

Figure 6 Forecasted values of CO_2 emission from VARX (1,3)



The MAPE value for CO_2 emission was 5.11325, which is less than 10%, and the VARX (1,3) model has excellent forecasting capabilities. So, the VARX (1,3) model can be used to forecast total annual CO2 emissions in the future.

Conditional forecasts with alternative scenarios were generated by defining new scenarios and specifying the time path of the exogenous variables under that scenario. There are three scenarios to examine what would happen if those values are higher than in the baseline forecast, assuming that the energy consumption, GDP and urban population values are 10% higher since 2000. This is done by overriding the values of exogenous variables during the forecast horizon and solving the model under new scenarios. The resulting forecasts under the baseline (unconditional forecast) and alternative (conditional) forecast for CO_2 emission were obtained under scenario 1, scenario two and scenario three, which represents the override forecast values of energy consumption, GDP and urban population respectively (Figure 7,8,9).

Figure 7 Conditional forecasts for CO_2 emission under alternative scenario (scenario1): energy consumption values are 10% higher than the actual values since 2000

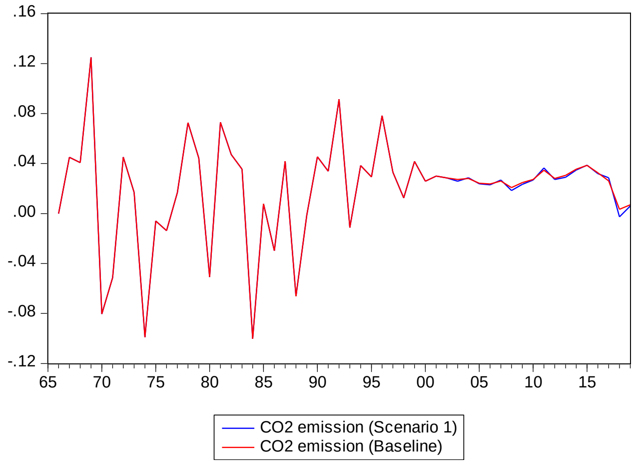


Figure 8 Conditional forecasts for CO_2 emission under alternative scenario (scenario2): GDP values are 10% higher than the actual values since 2000

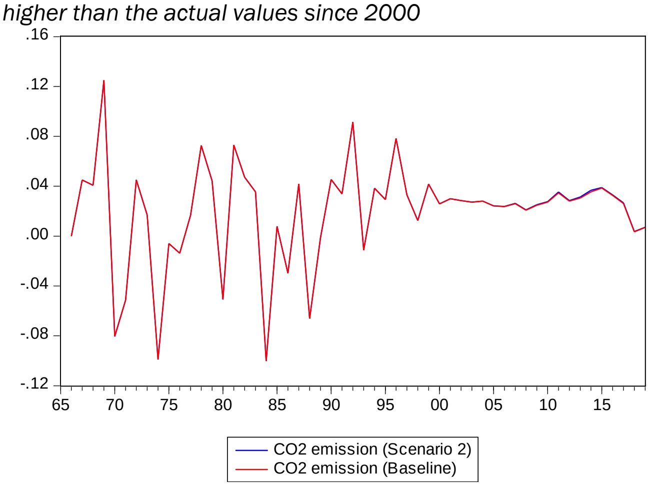
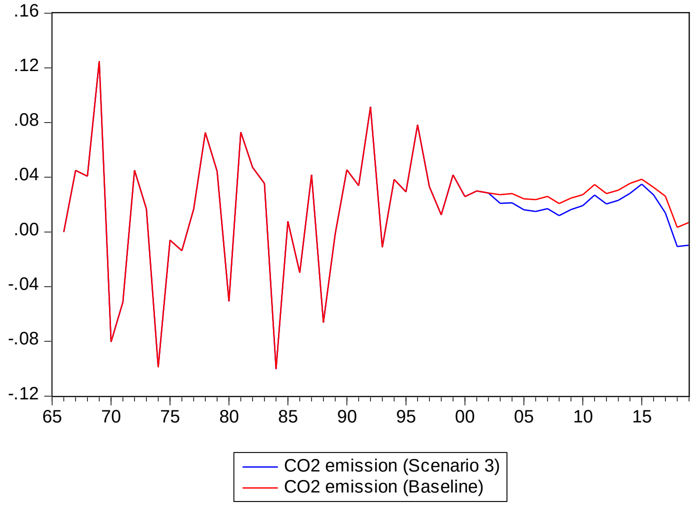


Figure 9 Conditional forecasts for CO_2 emission under alternative scenario (scenario3): urban population values are 10% higher than the actual values since 2000



There is no deviation in CO2 emission forecasts for scenario 1 (increase of energy consumption) and scenario 2 (increase of GDP) from baseline forecasts, but the increase in urban population has led scenario 3's forecasts for CO_2 emission to shift upward compared to baseline.

4. Conclusion

Based on the results obtained, it is concluded that the ARIMA (0, 1, 1) model, which had the minimum value of all the selection criteria, was the most appropriate model for forecasting the CO2 emission in Sri Lanka. Using the model ARIMA (0, 1, 1) forecasted CO2 emission in Sri Lanka were 27.7741 million tons, 28.8948 million tons, 30.0601 million tons and 31.2723 million tons in the years 2022, 2023, 2024 and 2025 respectively. Based on the results of the analysis of the relationship between the endogenous (annual CO2 emission) and exogenous variables (energy consumption, GDP and urban population), the VARX (1,3) model was found to be the best model for the relationship among these variables. This study will help the government of Sri Lanka, especially when it comes to short-term and long-term planning, and for policy makers to shape up the policies to take necessary actions to reduce the CO2 emission in Sri Lanka. The study results may be applied to the design and implementation of the energy audit concept, energy management and the energy conservation practices.

References

Allen, M. R., Frame, D. J., Huntingford, C., Jones, C. D., Lowe, J. A., Meinshausen, M., & Meinshausen, N. (2009). Warming caused by cumulative carbon emissions towards the trillionth tonne. *Nature*, *458*(7242), 1163-1166.

Baba, N. (2010). Sinking the pearl of the Indian Ocean: Climate change in Sri Lanka. *Global Majority E-Journal*, *1*(1), 4-16.

Bhanumati, P. (2022) Data for a greener world: A guide for practitioners and policy makers. International Monetary Fund.

Basnayake, B.R.S.B., Abjhyasinghe, K.R., Sumathipala, W.L., Punyawardene, B.V. R.P., Perera, N., and Joseph, P.G. (2007). Climate Change in Sri Lanka: Impacts, Adaptation & Mitigation. *Proceedings of the National Conference on Climate Change, Centre for Climate Change Studies, and Kobe-Asia-Pacific Network for Global Change Research, Department of Meteorology, Colombo, Sri Lanka.*

Bowden, N., & Payne, J. E. (2008). Short term forecasting of electricity prices for MISO hubs: Evidence from ARIMA-EGARCH models. *Energy economics*, *30*(6), 3186-3197.

Box, G.E.P. and Jenkins, G.M. (1970). Time Series Analysis: Forecasting and Control, *Holden-Day*, San Francisco.

Box, G. E., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). *Time series analysis: forecasting and control*. John Wiley & Sons.

EPA (2014). Climate Change Indicators in the United States, Global Greenhouse Gas Emissions. Available from: www.epa.gov/climatechange/indicators [22nd February 2020].

Fan, W., Luo, X., Yu, J., & Dai, Y. (2021). An empirical study of carbon emission impact factors based on the vector autoregression model. *Energies*, *14*(22), 7797.

Halicioglu, F. (2009). An econometric study of CO2 emissions, energy consumption, income and foreign trade in Turkey. *Energy policy*, *37*(3), 1156-1164.

Hannah, R., Max, R., & Pablo, R. (2020). CO_2 and greenhouse gas emissions. *Our world in data*. Available from: <'https://ourworldindata.org/co2-and-other-greenhouse-gas-emissions'>. [24 March 2020].

Hosseini, S. M., Saifoddin, A., Shirmohammadi, R., & Aslani, A. (2019). Forecasting of CO2 emissions in Iran based on time series and regression analysis. *Energy Reports*, *5*, 619-631.

Knoema (2021) Available at: https://knoema.com/atlas/SriLanka/CO2-emission [28th July 2021].

Kumari, D.D.N.D.S., and Abeynayake, N.R. (2017). Modeling of carbon dioxide emission pattern in Sri Lanka. *Proceedings of 16th Agricultural Research Symposium (2017) 75-78. Wayamba University of Sri Lanka.*

Lotfalipour, M. R., Falahi, M. A., & 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.

Ministry of Environment (2021). Updated Nationally Determined contributions. URL: http://www.climatechange.lk/CCS2021/UpdatedNDCsSriLanka2021.pdf

Muth, J. F. (1960). Optimal properties of exponentially weighted forecasts. *Journal of the american statistical association*, *55*(290), 299-306.

Naranpanawa, A. (2011). Does trade openness promote carbon emissions? Empirical evidence from Sri Lanka. *The Empirical Economics Letters*, *10*(10), 973-986.

Nebojsa, N. (1994). Energy Strategies for Mitigating Global Change. Technology Responses to Global Environmental Challenges.

Nyoni, T., & Bonga, W. G. (2019). Prediction of co2 emissions in india using arima models. *DRJ-Journal of Economics & Finance*, *4*(2), 01-10.

Pappas, S. S., Ekonomou, L., Karamousantas, D. C., Chatzarakis, G. E., Katsikas, S. K., & Liatsis, P. (2008). Electricity demand loads modeling using AutoRegressive Moving Average (ARMA) models. *Energy*, *33*(9), 1353-1360.

Parry, M. L. (Ed.). (2007). *Climate change 2007-impacts, adaptation and vulnerability: Working group II contribution to the fourth assessment report of the IPCC* (Vol. 4). Cambridge University Press.

Rahman, A., & Hasan, M. M. (2017). Modeling and forecasting of carbon dioxide emissions in Bangladesh using Autoregressive Integrated Moving Average (ARIMA) models. *Open Journal of Statistics*, *7*(4), 560-566.

Razmjoo, A., & Davarpanah, A. (2019). Developing various hybrid energy systems for residential application as an appropriate and reliable way to achieve energy sustainability. *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects*, *41*(10), 1180-1193.

Samreen, F., Sayed, S.A., Syeda, S.Z., Ehtesham, H., Tayyab, R.F. and Mehwish, S.K. (2019). Forecasting Carbon Dioxide Emission of Asian Countries Using ARIMA and Simple Exponential Smoothing Models. *Econ. Environ. Geol*. 10 (1), 64-69.

Septiani, A., Sumertajaya, I. M., & Aidi, M. N. (2019). Vector autoregressive X (VARX) modeling for Indonesian macroeconomic indicators and handling different time variations with cubic spline interpolation. *Repositories-Dept. of Statistics, IPB University*, 175-180.

Shih, S. H., & Tsokos, C. P. (2008). Prediction models for carbon dioxide emissions and the atmosphere. *Neural, Parallel and Scientific Computations*, *16*(1), 165.

Sutthichaimethee, P., & Ariyasajjakorn, D. (2018). Forecast of carbon dioxide emissions from energy consumption in industry sectors in Thailand. *Environmental and Climate Technologies*, *22*(1), 107-117.

Xu, B., & Lin, B. (2015). Carbon dioxide emissions reduction in China's transport sector: a dynamic VAR (vector autoregression) approach. *Energy*, *83*, 486-495.