31

Statistical Modeling of Non-stationary Carbondioxide Emissions versus Income Data

N. Tharshanna[1], K. Perera[2] and R. Shanthini[3]

[1]*Department of Mathematics and Computer Science, Eastern University*

[2]*Department of Engineering Mathematics,*

[3]*Department of Chemical & Process Engineering, Faculty of Engineering, University of Peradeniya*

Abstract

This paper explores the relationship between per capita carbon dioxide emissions and per capita income for Sri Lanka, Japan and United States. Unit root tests carried out showed that these variables are non-stationary. Since non-stationary data outperform regression models developed with them, regression models were developed in this study with stationarized variables. For the best fitted models for Sri Lanka, Japan and United States, it was found that the per capita emissions were driven mainly by its own autoregressive term, GDP per capita and its autoregressive term. For the best fitted models, statistical characteristics such as Mean Square Error, Akaike Information Criterion and Bayesian Information Criterion took the least values among the models studied. Also, the residuals of the best fitted models were found to posses the characteristics of white noise and they were normally distributed.

Keywords: CO_2 emissions; economic growth; GDP per capita; stationarity.

1 Introduction

Increase in national income is often seen as having strong links with increase in burning of fossil fuels which results in a considerable amount of carbon dioxide (CO_2) emissions (Shafik, 1994), which is the major source of global warming (IPCC, 2001). The relationship between income, measured by gross domestic product per capita and CO_2 emissions per capita is extensively researched and regression models have been developed (see, for example, Heil and Seldon, 2001).

A linear relationship for CO_2 emissions per capita and GDP per capita was confirmed in early studies (Shafik and Bandyopadhyay, 1992; Shafik, 1994). An inverted U-shaped function, popularly known as Environmental Kuznets Curve, has also been identified (De Bruyn et al., 1998; Heli and Selden, 2001; Holtz-Eakin and Selden, 1995; Moomau and Unruh, 1997) and also an N-shaped (cubic) specifications (see, for example, Galeotti and Lanza, 1999).

However, the abovementioned regressions models could be outperformed by the data if they are non-stationary. But, the stationarity of the data has seldom been tested before developing regression models (Friedl and Getzner, 2003). In this paper, regression models have been developed accounting for the non-stationarity of emissions and income data for Sri Lanka, United States (US) and Japan.

2 Data and Methods

2.1 Data Used

The data used in this study are obtained from World Development Indicators Online (World Bank, 2006) for the years 1960 to 2002. The annual carbon dioxide emissions data are those stemming from the burning of fossil fuels, during the consumption of solid, liquid and gas fuels and gas flaring and the manufacture of cement. The annual carbon dioxide emissions divided by midyear population provides the annual carbon dioxide emissions per capita (denoted by CO2) and it is in tonnes (= 1000 kg) of CO_2. The income per capita of a country is represented by its gross domestic product per capita (denoted by GDP) which is the annual gross domestic product divided by midyear population and it is in constant 2000 US$.

2.2 Methods

To test the stationarity of the variables, CO_2 and GDP, *Phillips–Perron* unit root test was applied. Time series regression models were developed using the AUTOREG procedure available in SAS statistical package. When time series data are used in regression analysis, often the error term is not independent through time. Instead, the errors are serially correlated or auto correlated. The autoregressive error model used in the AUTOREG procedure corrects for serial correlation.

To test for the best fitted model, residuals are tested for white noise characteristics. That is, the residuals are zero-mean, homoscedasticity and serially uncorrelated random variables. The *Phillips–Perron* stationarity test was used for this purpose. *Kolmogorov–Smirnov* test was used to test the normality of the residuals.

In the best fitted models for the countries studied, the CO_2 emissions per capita was mainly driven by its own autoregressive term, the GDP per capita and its autoregressive term. For the purpose of predicting future emissions, the autoregressive term of the CO_2 emissions per capita was obtained from the model itself. The GDP per capita term and its autoregressive term for predicting future CO_2 emissions were generated assuming per capita GDP growth rate at the preferred value of 10% (Case 1), at the world average of 3% (Case 2) and at the worst case of 0%, which is no growth (Case 3). That is, for the future, $GDP_{t+1} = (1 + x)\, GDP_t$, where x = 0.1, 0.03 or 0.

3 Results and Discussions

3.1 Stationarity Test

Results of the *Phillips–Perron* unit root test carried out to test the stationarity of the variables concerned are shown in Table 1. It exhibits that the *Phillips–Perron* test statistics rho and tau P-values (denoted by Pr on the table) for all cases studied are greater than 0.05, Therefore, we concluded that the null hypothesis (Ho: unit root exists) cannot be rejected at a reasonable level of significance. This means that the variables testes are non-stationary. From a qualitative viewpoint, it is concluded that the time paths of CO_2 emissions per capita and income per capita for all countries studied are not independent of time, i.e., per capita CO_2 emissions and per capita income grew during the relevant period without a significant mean-reverting trend.

Table 1: *Phillips–Perron* unit root test results

Countries	Variable	Rho	Pr < Rho	Tau	Pr < Tau
Sri Lanka	$CO_{2(t)}$	-2.30	0.96	0.67	0.97
	$GDP_{(t)}$	-2.52	0.95	-1.59	0.78
United States	$CO_{2(t)}$	-6.14	0.71	-2.05	0.56
	$GDP_{(t)}$	-10.27	0.36	-2.24	0.45
Japan	$CO_{2(t)}$	-4.51	0.84	-1.93	0.62
	$GDP_{(t)}$	-3.06	0.93	-0.82	0.95

3.2 Model Development

This section describes the two models analysed in this study.

Model 1 was developed by fitting a simple linear regression model with lagged versions. The AUTOREG procedure in SAS with the option NLAG was used. First, NLAG was set to 1 and the coefficients of the regression equation,

$$CO\,2_t - \rho_1 CO2_{t-1} = a + b(GDP_t - \rho_1 GDP_{t-1}) + \varepsilon_t,$$

including ρ_1, were estimated. In case of obtaining significant values for the coefficients, the variables ($CO\,2_t - \rho_1 CO2_{t-1}$) and ($GDP_t - \rho_1 GDP_{t-1}$) were tested for stationarity. If they were not stationary, then NLAG was to be set at 2 and the coefficients of the regression equation,

$$CO\,2_t - \rho_1 CO2_{t-1} - \rho_2 CO2_{t-2} = a + b(GDP_t - \rho_1 GDP_{t-1} - \rho_2 GDP_{t-2}) + \varepsilon_t,$$

were to be estimated. But this step was not necessary for the three countries considered in this study.

Model 2, shown below, was developed using the differenced versions of the variables concerned, which are found to be stationary by the results of the *Phillips–Perron* unit root test.

$$DIF(CO2_t) = a + b[DIF(GDP_t)] + \varepsilon_t,$$

where $DIF(CO2_t) = CO2_t - CO2_{t-1}$ and

$DIF(GDP_t) = GDP_t - GDP_{t-1}$

In case of one or more coefficients of the above regression equation being insignificant, a sub model of Model 2, named Model 2a, was developed using the AUTOREG procedure in SAS with NLAG = 1 option. Model 2a is as shown below:

$$DIF(CO2_t) - \rho_1 DIF(CO2_{t-1}) = a + b[DIF(GDP_t) - \rho_1 DIF(GDP_{t-1})] + \varepsilon_t,$$

This was necessary for the case of United States.

3.3 Selecting the Best Model

Model selection is best done based on the statistics, Mean Square Error (MSE), Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC), which are reported in Table 2.

For Sri Lanka, the AIC, BIC and MSE values of Model 1 are smaller than those of Model 2. Because of that and because of the high R^2 value, Model 1 is chosen as the best model for Sri Lanka. Also, Table 3 shows that the Model 1 parameter estimates for Sri Lanka are significant.

For Japan, the AIC, BIC and MSE values of Model 2 are smaller than those of Model 1. Therefore, Model 2 is chosen as the best model for Japan despite its low R^2 value. (Difference models are known to give low R^2 values.) Table 4 shows that the Model 2 parameter estimates are significant.

For United States, AIC, BIC and MSE values of Model 2a are smaller than those of Model 1. Therefore, Model 2a is chosen as the best model for United States. Also, Table 4 shows that the Model 2a parameter estimates are significant.

Table 2: Statistical characteristics of the models studied

Country	Model	AIC	BIC	MSE	Total R^2 (%)
Sri Lanka	Model 1	-167.33	-163.81	0.0011	98.82
	Model 2	-162.82	-159.35	0.0012	12.3
Japan	Model 1	11.73	15.25	0.069	99.88
	Model 2	7.65	11.12	0.067	54.45
United States	Model 1	66.18	69.7	0.23	99.94
	Model 2a	53.099	56.57	0.197	40.29

Table 3: Parameter estimates for Model 1

Country	Parameter estimates	
	b	ρ_1
Sri Lanka	0.000575 (P<0.0001)	0.8902 (P<0.0001)
Japan	0.000278 (P<0.0001)	0.978 (P<0.0001)
United States	0.000494 (P<0.0001)	0.9976 (P<0.0001)

Table 4: Parameter estimates for Model 2 and Model 2a

Country	Model	Parameter estimates	
		b	ρ_1
Japan	Model 2	0.000314 (P<0.0001)	
United States	Model 2a	0.000586 (P<0.0001)	0.4050 (P=0.0093)

The best models chosen in this study are summarized below:

For Sri Lanka,

$CO2_t = 0.8902CO2_{t-1}+0.000575(GDP_t-0.8902GDP_{t-1})+\varepsilon_t$

For Japan,

$$CO2_t = CO2_{t-1} + 0.000314(GDP_t - GDP_{t-1}) + \varepsilon_t$$

For United States,

$$CO2_t = CO2_{t-1} + 0.405(CO2_{t-1} - CO2_{t-2}) + \varepsilon_t$$
$$+0.000586(GDP_t - GDP_{t-1}) - 0.405(GDP_{t-1} - GDP_{t-2}) + \varepsilon_t$$

3.4 Diagnostic Checking of the Fitted Model Residuals

In the time series regression modeling, we often assume that the regression errors are zero- mean, homoscedasticity and serially uncorrelated random variables, i.e., the errors are white noise. Table 5 shows that the P-values for the white noise tests of all chosen models are greater than 0.05. Therefore the null hypothesis (Ho: white noise of the variable (residual)) cannot be rejected at a reasonable level of significance. The above result shows that the residuals of the models chosen are white noise. The third column of Table 5 shows that the P-values for the normality test of the residuals are greater than 0.15, which means the residuals are normally distributed. Therefore, we conclude that the residuals of the models chosen in this study are mean zero, constant variance, serially uncorrelated and normally distributed variables.

Table 5: White noise and normality test for the residuals of the models chosen

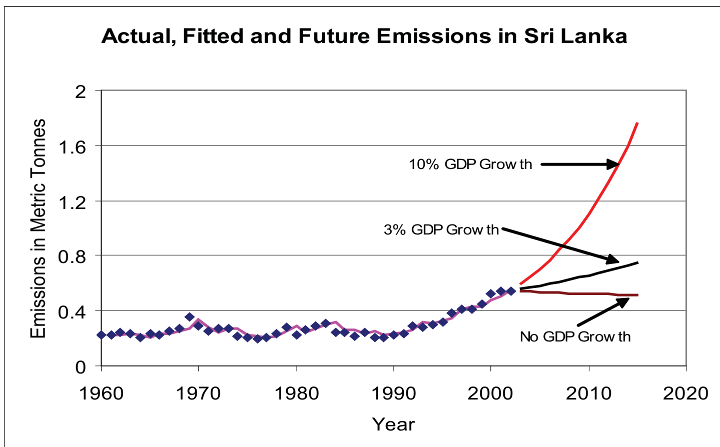
Country	P-value for	
	White Noise Test	Kolmogorov-Smimov Normality Test
Sri Lanka	0.3129	> 0.15
Japan	0.2952	> 0.15
United States	0.2082	> 0.15

3.5 Predicting the Carbon Dioxide per Capita Emissions

Figure 1 and 2 shows the actual, fitted and future values of the per capita carbon dioxide emissions for the countries studied. As long as there is a growth in the per capita income, the figure shows, the per capita emissions also increase. Therefore, an upward trend in per capita carbon dioxide emissions seems to be the case for the future of the countries studied.

4 Conclusion

Carbon dioxide emissions per capita and the gross domestic product per capita data are non-stationary for Sri Lanka, Japan and United States. The regression models were fitted for stationarized variables. For the best fitted models of all three countries studied, CO_2 emissions per capita is driven by its own autoregressive term, the GDP per capita and its autoregressive term.



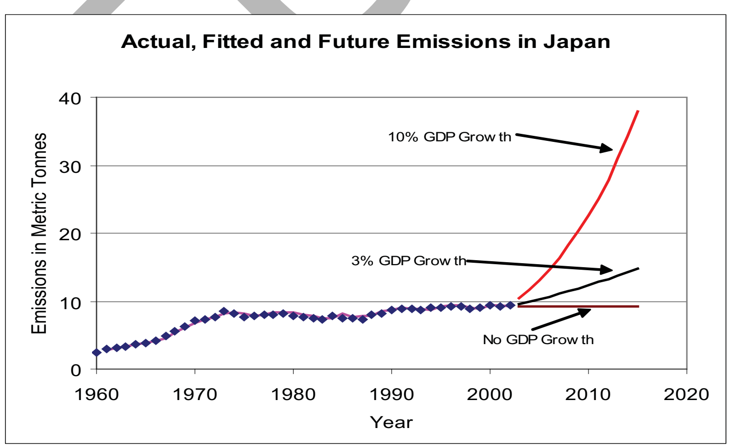


Figure 1: Actual, Fitted and Future CO_2 Emissions for Sri Lanka and Japan

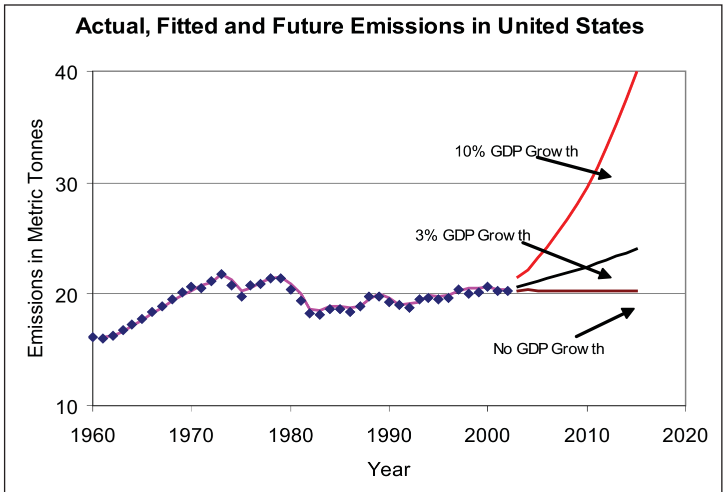


Figure 2: Actual, Fitted and Future CO_2 Emissions for United States

References

Friedl, B. and Getzner, M., 2003, "Determinants of C02 emissions in a small open economy", Ecol. Econ, 45:133-48.

Heil, M.T. and Seldon, T.M., 2001, "Carbon Emissions and Economic Development: future trajectories based on historical experience", Environ, Dev. Econ., 6(1):63-8.

IPCC, 2001, "Climate Change 2001: The Scientific Basis", Cambridge University Press, Cambridge.

Shafik, N., 1994, "Economic Development and Environmental Quality: An Econometric Analysis", 0xford Economic Papers, 46:757-73.

World Bank, 2006, "World Development 1ndicators 0nline", World Bank Publication, http://publications.worldbank.org/susbcriptions/WDI/

Shafik and Bandyopadhyay, 1992, "A linear relationship for CO_2 emissions per capita and GDP per capita"

De Bruyn et al., 1998; Heli and Selden, 2001, "An inverted U-shaped relationship for CO_2 emissions per capita and GDP per capita"

Holtz-Eakin and Selden, 1995, Moomau and Unruh, 1997, Galeotti and Lanza, 1999. "N-shaped (cubic) relationship for CO_2 emissions per capita and GDP per capita".

http://www.duke.edu/~rnau/411l696.htm. "Decision forecasting 411 Fitting time series regression models".

http://www.duke.edu/~rnau/simpreg.htm. "Not-so-simple regression models Decission forecating 411".

http://swan.econ.ohio-state.edu/econ840/note1.pdf. "Stationary Time series".

http://www.tcd.ie/Economics/staff/frainj/main/2005_06_MSc/~Session20/UNIVAR4.pdf. "Time series".

Ross Ihaka. Time Series Analysis, Statistics Department University of Auckland.

Dan Vlad Metes. Visual, Unit Root and Stationarity Tests and Their Power and Accuracy, University of Alberta Edmonton.

http://faculty.washington.edu/ezivot/econ584/584notes.htm. "Unit Root Tests".

http://Vosvrdaweb.utia.cas.cz/derivaty/LukasVanchastationarity\%20and\%20Unit\%20Root\20Testing.pdf. "Stationarity and Unit Root Testing".

http://www2.stat.unibo.it/ManualiSas/ets/chap7.pdf. "SAS ARIMA Procedure".

http://faculty.washington.edu/ezivot/econ584/584notes.htm. "Unit Root Tests".

http://www2.stat.unibo.it/ManualiSas/ets/chap8.pdf. "SAS AUTOREG Procedure".

Footnotes

1 **Altering/controlling the parameters** of the system can be done through the strict enforcement of legislation.

2 This is part of the EMS (environmental Management System) and is necessary for finding out the sensitive points of origins of chaotic situations and devising appropriate controlling/eradicating strategies.