

# Linear Relationship between CO<sub>2</sub> Emissions and Economic Variables: Evidence from a Developed Country and a Developing Country

Lazim Abdullah<sup>1</sup>

<sup>1</sup> School of Informatics and Applied Mathematics, University Malaysia Terengganu, Malaysia

Correspondence: Lazim Abdullah, School of Informatics and Applied Mathematics, University Malaysia Terengganu, 21030 Kuala Terengganu, Terengganu, Malaysia. Tel: 60-9-668-3335. E-mail: lazim\_m@umt.edu.my

Received: December 3, 2014 Accepted: December 19, 2014 Online Published: March 30, 2015

doi:10.5539/jsd.v8n2p66

URL: <http://dx.doi.org/10.5539/jsd.v8n2p66>

## Abstract

Multiple linear regressions (MLR) analyses have been used to explain various linear relationships between CO<sub>2</sub> emissions and economic variables. Previous research have suggested that predictors of CO<sub>2</sub> emissions can be vary depending on economic development of a country and also region where a country is located. This paper investigates the linear relationships between CO<sub>2</sub> emissions and its related economic variables using MLR analyses for the UK and Malaysia. Differently from the typical MLR analyses which directly identify the most prevalent predictors, these analysis includes F- test to check linearity property of the relationships, regression equations and also error analysis to validate the robustness of the multiple linear regressions as predictive tool. It is shown that the linear relationships of the UK data outperformed the linear relationship of Malaysian data where agriculture and transport are the most effective predictors for the UK and Malaysia data respectively. The most effective predictor from the linear relationship would provide valuable information for policy holders and environmental management authority on potential causes of CO<sub>2</sub> emissions.

**Keywords:** linear relationship, economic variable, development status, CO<sub>2</sub> emission, error analysis

## 1. Introduction

Over the past century there has been a dramatic increase in economic activities in all over the world. One of the side effects from the increasing of economic activities is escalating of CO<sub>2</sub> emissions. The trend of CO<sub>2</sub> emissions has been increased exponentially in the past few decades. Almost 30 billion tons of CO<sub>2</sub> enters the atmosphere as a result various human activities each year (Goodall, 2007). The effect of the higher concentrations of CO<sub>2</sub> to the world population could not be taken lightly. It has been reported that CO<sub>2</sub> emissions is held responsible for 58.8% of green house effects (World Bank, 2007). The effects may cause major environment threats. The increase in CO<sub>2</sub> emission would give disastrous environmental consequences such as droughts, storms, floods and other environmental calamities. According to a research, global sea level has increased by 10 -20 cm during the twentieth century (Mukhtar et al., 2004). Not only was the increment in sea level but temperature of sea also reported in risen trend. Spence (2005) reported that global CO<sub>2</sub> emissions have increased by 30 % and the world temperature has risen by 0.3-0.6 degree Celsius. These environmental instabilities are said to be caused by a combination of intangible and tangible variables such as population growth, economic growth, energy consumption, industrial activities and CO<sub>2</sub> emissions. Therefore, the intertwined relationships among all these variables with CO<sub>2</sub> emissions are a subject that attracted many concerned parties. Developments in these relationships have heightened the need for empirical research.

Many research have been conducted using various scientific methods across regions to investigate these relationships. Lizano and Gutierrez (2007), for example, employed non-parametric frontier approach to model the relationships among population, gross domestic product (GDP), energy consumption and CO<sub>2</sub> emissions. Moran and Gonzalez (2006) proposed a combination of sensitivity analysis and linear programming methods that could identify the main productive linkages between CO<sub>2</sub> emissions and human activities. The case study in Spain suggested that CO<sub>2</sub> emissions were related to productive relationships within economics activities. Freitas and Kaneko (2011) analyzed the decoupling of CO<sub>2</sub> emissions and economic growth in Brazil. They examined the occurrence of a decoupling between the growth rates in economic activity and CO<sub>2</sub> emissions from energy consumption in Brazil from 2004 to 2009. This decoupling was highlighted when economic activity and CO<sub>2</sub>

emissions moved in opposite directions in 2009. Ang (2009) explored the determinants of CO<sub>2</sub> emissions in China using aggregate data for more than half a century. The results indicated that CO<sub>2</sub> emissions in China are negatively related to research intensity, technology transfer and the absorptive capacity of the economy to assimilate foreign technology. Their findings also indicated that more energy use, higher income and greater trade openness tend to cause more CO<sub>2</sub> emissions.

The relationships between CO<sub>2</sub> emissions and economic variable were further investigated from comparative studies perspective. Other than research at a specific region, there were also handful of comparative research that have been embarked on different nations and different economic development status. Su et al. (2009), for example, analyzed CO<sub>2</sub> emissions embodied in trade. The effects of sector aggregation in trade were investigated to see the possible effects of sector aggregation. They conducted empirical studies using the data of China and Singapore where energy related CO<sub>2</sub> emissions embodied in their exports are estimated at different levels of sector aggregation. Martínez-Zarzoso and Maruotti (2011) investigated the impact of urbanization on CO<sub>2</sub> emissions. Evidence from developing countries was provided to analyze the impact of urbanization on CO<sub>2</sub> emissions in developing countries from 1975 to 2003. It corroborates the existing literature by examining the effect of urbanization. Dynamism and the presence of heterogeneity were taken into account in the sample of countries. The results showed an inverted-U shaped relationship between urbanization and CO<sub>2</sub> emissions was really exist. However, the relationships between CO<sub>2</sub> emissions and economic variables by embarking on cross tabulating analysis among developed and developing countries were given less attention. It was hypothesized that CO<sub>2</sub> emissions from developed countries are more efficient than developing countries. Rosa and Tolmasquim (1993), for example, proposed an analytical model to compare energy efficiency indices and CO<sub>2</sub> emissions in developed and developing countries. The index of CO<sub>2</sub> emissions was about ten times higher in Brazil than in the USA, Japan and Germany. This analysis shows that efficiency of CO<sub>2</sub> emissions between developed and developing countries seems differ significantly. Development status of a country is much depending on economic development. A developed country is known as a country which has a highly developed economy and advanced technological infrastructure relative to other less developed nations. There are no conclusive criteria for evaluating development status of a country. However, the three most typical criteria in evaluating economic development status are GDP, per capita income, and level of industrialization. A developing country contrarily is defined as a nation with a low living standard and undeveloped industrial base (Sullivan, et al., 2003). Clearly, a developing country and a developed country were segregated solely based on economic variables. It is imperative to investigate how the nations' segregation based on economic variables may affect the efficiency in CO<sub>2</sub> emissions.

This paper aims to investigate linear relationship between CO<sub>2</sub> emissions and economics variables for a developed country and a developing country. Data of CO<sub>2</sub> emissions and its related variables from Malaysia and the United Kingdom (UK) are employed to investigate the relationship. The rest of this paper is structured as follows. Section 2 describes a brief theoretical background of linear relationship between predictors and response variables. Section 3 presents analysis and results using data from developed and developing countries. Section 4 concludes.

## 2. Predictors and Response Variables in Linear Relationship

Linear relationship between two variables is typically explained by linear regression model. This model gives a straight-line relationship between two variables (Mann, 2007). Linear regression was the first type of regression analysis to be studied rigorously, and to be used extensively in many practical applications. The main reason behind this popularity is characterized by linear relationship between two variables. The linear relationship depends on their variables that easier to fit than models which are non-linearly related to their variables. The statistical properties of the resulting predictors are easier to determine. The function form which is most frequently used for expressing the linear relationship is given by

$$Y' = a + bX \quad (1)$$

where  $Y'$  is the predicted value of the  $Y$  variable for a selected  $X$  value. Constant  $a$  is the  $Y$ -intercept. It is the estimated value of  $Y$  when  $X = 0$ . Another way to put it is:  $a$  is the estimated value of  $Y$  where the regression line crosses the  $Y$ -axis when  $X$  is zero.  $b$  is the slope of the line, or the average change in  $Y'$  for each change of one unit (either increase or decrease) in the independent variable  $X$ . The equation (1) is function to describe the relationship between one response variable ( $Y'$ ) and one predictor. However, in many practical applications, there are more than one predictor that can be related to one response variable. The linear relationship between one response variable and several predictors is explained by multiple linear regressions.

Multiple regression analysis has been viewed only as a way to describe the relationship between a response variable and several predictors. In multiple linear regressions, additional predictors (denoted  $X_1, X_2, \dots$ , and so on) are used to help better explain or predict the response variable ( $Y$ ). The general descriptive form of a multiple linear equation is shown in Equation (2). The number of predictors is represented by  $k$ . So  $k$  can be any positive integer.

$$Y' = a + b_1X_1 + b_2X_2 + b_3X_3 + \dots + b_kX_k \quad (2)$$

where:

$a$  is the intercept, the value of  $Y$  when all the  $X$ 's are zero.

$b_k$  is the amount by which  $Y$  changes when that particular  $X_k$  increases by one unit with all other values held the same. The subscript  $j$  can assume values between 1 and  $k$ , which is the number of predictors.

Response variable and predictors are predominantly used in the next section for describing the relationship between CO<sub>2</sub> emissions and economic variables.

### 3. Analyses and Results

Data of CO<sub>2</sub> emissions and its associated economic variables from Malaysia and the UK were employed to be tested the relationships using multiple linear regression. This study collects data on CO<sub>2</sub> emissions for the period between 1990 to 2010 and 1981 to 2005 for the UK and Malaysia respectively. Different period of data were considered due to limitation in retrieving secondary data. Data of the response variable and predictors were retrieved from the official websites of World Bank (2012).

Predictors of the UK data are energy supply, business, transport, population, agriculture, industrial process and waste management. It is labeled as  $x_1, x_2, x_3, x_4, x_5, x_6$ , and  $x_7$  respectively. Predictors of Malaysia data are fuel mix, transport, GDP, and population. These predictors are labeled as  $x_{M1}, x_{M2}, x_{M3}$ , and  $x_{M4}$ . CO<sub>2</sub> emission is the response variable and denoted as  $y$ . The relationships between the predictors and the response variables are executed using the multiple linear regressions. Results are divided into four subsections as to accommodate the comprehensive analysis of linear relationships.

#### 3.1 Excluded Predictors

Decision on any of the predictors that can be excluded from the Malaysian model was made. Table 1 shows the predictor of fuel mix and population can be removed from the full model due to the significance level that greater 0.05. These two predictors are not significance when probability of rejection is defined at 0.05.

Table 1. Excluded predictors for the Malaysia data

Model		t-value	Sig
1	Fuel Mix	0.497	0.624
	GDP	2.183	0.040
	Population	1.875	0.074

The similar analysis was made to decide predictors that can be excluded from the UK model. With significant level at 0.05, the predictors of population, industrial process and waste management can be removed from the full model. The significance values, t value and predictors are shown in Table 2.

Table 2. Excluded predictors for the UK data

	Model	t-value	Sig
1	Energy supply	8.862	0.000
	Transport	1.1887	0.075
	Population	-0.923	0.368
	Agriculture	-2.135	0.047
	Industrial process	0.838	0.413
	Waste management	0.267	0.793

The not excluded predictors were used in the next three analyses. These predictors are used to determine the contributions of predictors towards CO<sub>2</sub> emission, to test the model linearity and to construct multiple regression equations.

### 3.2 Analysis of Variance (ANOVA) and Coefficient of Determination

The selected predictors were used to model the CO<sub>2</sub> emissions using multiple linear regressions. Table 3 shows summary results for the linear regression model for Malaysia data.

Table 3. Linearity test for Malaysia data

Regression statistics					
R <sup>2</sup>					0.477
Adjusted R <sup>2</sup>					0.430
SE					919.8784
	Df	SS	MS	F	Significance F
Analysis of variance (ANOVA)					
Regression	2	16986808	8493404.023	10.037	0.01
Residual	22	18615878	846176.260		
Total	24	35602686			

The coefficient of determination and the adjusted R squared are 0.477 and 0.430 respectively, which indicates that about 48 % of the variation in the CO<sub>2</sub> emissions is explained by transport and GDP. The analysis of variance indicates that the *p* value (probability of rejecting the null hypothesis) for the F test statistic is 0.01, which provide strong evidence against the null hypothesis. In other words, there is a linear relationship between the predictors and CO<sub>2</sub> emissions for Malaysia data.

The similar analysis was made to obtain the linear regression for the UK data. The regression statistics and F-test are shown in Table 4.

Table 4. Linearity test for the UK data

Regression statistics					
R <sup>2</sup>					0.991
Adjusted R <sup>2</sup>					0.989
SE					2.89989
	Df	SS	MS	F	Significance F
Analysis of variance (ANOVA)					
Regression	4	15596.156	3899.039	463.654	0.000
Residual	16	134.550	8.409		
Total	20	15730.705			

The coefficient of determination and the adjusted R squared are 0.991 and 0.989 respectively, which indicates that about 99 % of the variation in the CO<sub>2</sub> emissions is explained by business, energy supply, transport and agriculture. The analysis of variance indicates that the *p* value for the F test statistic is 0.000, which provide strong evidence against the null hypothesis. There is a linear relationship between the predictors and CO<sub>2</sub> emissions for UK data.

### 3.3 Multiple Regression Equations

Regression coefficients for the not excluded predictors for Malaysia data are shown in Table 5.

Table 5. Regression coefficients for Malaysia data

	Coefficient	SE	t Statistic	p- value	Lower 95%	Upper 95%
Intercept	249.527	372.788	0.699	0.510	-523.588	1022.642
Transport	0.792	0.210	3.767	0.001	0.536	1.227
GDP	0.685	0.314	2.183	0.040	0.034	1.337

The t test statistic shows that the p-value for the model intercept and the coefficient associated with the rejection region is less than 0.05. This result provides strong evidence against the null hypothesis. The p-value associated with transport and GDP predictors are less than 0.05. These results indicate that at least one of the predictors is useful for predicting CO<sub>2</sub> emissions. From the value of regression coefficients, the multiple regression equation can be written as

$$\hat{Y} = 249.527 + 0.792x_{M2} + 0.685x_{M3}$$

The regression equation shows that  $x_{M2}$  is the best predictor. Hence, transport is the most effective predictor in CO<sub>2</sub> emissions in Malaysia.

Regression coefficients for the UK data are computed with the similar fashion. The four not excluded predictors are energy supply, business, transport, and agriculture. The regression coefficients for the predictors are shown in Table 6.

Table 6. Regression coefficients for the UK data

	Coefficient	SE	t Statistic	p- value	Lower 95%	Upper 95%
Intercept	-7.287	30.804	-0.237	0.816	-72.589	58.015
Business	1.499	0.196	7.628	0.000	1.082	1.915
Energy supply	0.955	0.065	11.754	0.000	0.818	1.092
Transport	1.271	0.191	6.658	0.000	0.866	1.676
Agriculture	10.122	3.953	2.560	0.021	1.741	18.502

The relationship between economic predictors and CO<sub>2</sub> emissions can be translated into an equation. The regression equation is  $\hat{Y} = -7.287 + 0.955x_1 + 1.499x_2 + 1.271x_3 + 10.122x_5$

The regression equation shows that  $x_5$  is the most effective predictor for the UK data. Of the four not excluded predictors, agriculture is the most effective predictor.

### 3.4 Prediction Errors

Analysing the linear relationship seems incomplete without investigating the performance of predicted values against actual values. The predicted values of CO<sub>2</sub> emissions are computed using the regression equations and its deviations from the actual values are measured. The deviations of predicted values from actual values are called as prediction errors. The mean average percentage errors (MAPE) and root mean square errors (RMSE) are used to measure the prediction errors. These errors are computed using the following two equations.

$$MAPE = \frac{1}{n} \sum_{t=1}^n \left| \frac{A_t - F_t}{A_t} \right|$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (A_t - F_t)^2}$$

where  $A_t$  is the actual value,  $n$  is number of data and  $F_t$  is the predicted value.

Errors for the UK model and Malaysia model are summarized in Table 7

Table 7. Errors of the UK Model and Malaysia model

Model	MAPE	RMSE
UK data	0.0036272	2.53123
Malaysia data	2.610492	862.924

The errors show that the UK model has smaller errors than Malaysia model. It indicates that the predictors of the UK model performs better than predictors of Malaysia on CO<sub>2</sub> emission prediction.

#### 4. Conclusions

An important element in analyzing CO<sub>2</sub> emissions and its variables is a method which can take into account multiple variables and easily interpreted its relationships. The model should establish a decision to reflect the contribution of each accounted predictors toward CO<sub>2</sub> emissions. Apart from the identification of the best predictor, the method is also able to provide evidence on the efficiency of the predictors for two countries with different economic development profiles and regions. In this paper, the multiple linear regressions was utilized to capture the effective predictor of CO<sub>2</sub> emissions for the United Kingdom and Malaysia data. Furthermore, this paper also contributed to the identification of the better model for CO<sub>2</sub> emissions. The approach has successfully offered the prediction of CO<sub>2</sub> of the two countries from two different socioeconomic profiles. The CO<sub>2</sub> emissions prediction model of the United Kingdom outperformed the model of Malaysia. The results indicate that the chosen predictors of CO<sub>2</sub> emissions in the United Kingdom were the better predictors. One the other hand, the predictors of CO<sub>2</sub> emissions in Malaysia were not sufficient to be considered as good predictors. Therefore, it is suggested that several new predictors should be considered in predicting CO<sub>2</sub> emissions in Malaysia. Future research could be explored not only considering the number of predictors but also could be extended in searching the best model for prediction of CO<sub>2</sub> emissions.

#### Acknowledgments

The first author is deeply indebted to the undergraduate student, Chee San Choi for partly assisting in data collection and analysis.

#### References

- Ang, J. (2009). CO<sub>2</sub> emissions, research and technology transfer in China. *Ecological Economics*, 68(10), 2658-2665. <http://dx.doi.org/10.1016/j.ecolecon.2009.05.002>
- Freitas, L., & Kaneko, S. (2011). Decomposing the decoupling of CO<sub>2</sub> emissions and economic growth in Brazil. *Ecological Economics*, 70(8), 1459-1469. <http://dx.doi.org/10.1016/j.ecolecon.2011.02.011>
- Goodall, C. (2007). *How to live a low carbon life: the individual's guide to stopping climate change* (1st ed.). United Kingdom: Earthscan.
- Lizano, S., & Gutierrez, E. (2007). Non-parametric frontier approach to modelling the relationships among population, GDP, energy consumption and CO<sub>2</sub> emissions. *Ecological Economics*, 66(4), 687-699. <http://dx.doi.org/10.1016/j.ecolecon.2007.11.003>
- Mann, P. S. (2007). *Introductory Statistics*. John Wiley and Sons Inc: New Jersey.
- Martínez-Zarzoso, I., & Maruotti, A. (2011). The impact of urbanization on CO<sub>2</sub> emissions: Evidence from developing countries. FEEM Working Paper No. 50.2008 70(7), 1344-1353. <http://dx.doi.org/10.2139/ssrn.1151928>
- Moran, M., & Gonzalez, P. (2006). A combined input-output and sensitivity analysis approach to analyse sector linkages and CO<sub>2</sub> emissions. *Energy Economics*, 29(3), 578-597. <http://dx.doi.org/10.1016/j.eneco.2006.02.004>
- Mukhtar, H., Kamaruddin, P. N. F. M., & Radhakishnan, V. R. (2004). Carbon credit trading for CO<sub>2</sub> reduction: Opportunities for Malaysia. *Technology cluster: Oil and gas*, 4(2). Retrieved from <http://www.utp.edu.my/publications/platform/Platform%20v4n2.pdf>
- Rosa, L. P., & Tolmasquim, M. T. (1993). An analytical model to compare energy efficiency indices and CO<sub>2</sub> emissions in developed and developing countries. *Energy Policy*, 21(3), 276-283. [http://dx.doi.org/10.1016/0301-4215\(93\)90249-F](http://dx.doi.org/10.1016/0301-4215(93)90249-F)
- Spence, C. (2005). *Global warming: Personal solutions for a healthy planet* (1st ed.). New York: Palgrave

Macmillan.

Su, B., Huang, H., Ang, B., & Zhou, P. (2009). Input-output analysis of CO<sub>2</sub> emissions embodied in trade: The effects of sector aggregation. *Energy Economics*, 32(1), 166-175. <http://dx.doi.org/10.1016/j.enpol.2013.01.041>

Sullivan, A., & Sheffrin, S. M. (2003). *Economics: Principles in Action*. Upper Saddle River, New Jersey 07458: Pearson Prentice Hall. p. 471.

World Bank. (2007). *Growth and CO<sub>2</sub> emissions: how do different countries fare*. Environment Department, Washington, D.C.

World Bank. (2012). *World Development indicator*. Retrieved April 15, 2012, from <http://siteresources.worldbank.org/INTCC/214574-1192124923600/21511758/CO2DecompositionfinalOct2007.pdf>

### Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).