Modeling and analysis of CO2 emissions in million tons of sectoral greenhouse gases in Indonesia

Ruly Budiono¹, Hafizan Juahir², Mustafa Mamat¹, Sukono*, Sudradjat Supian⁴, Muhamad Nurzaman¹

¹Department of Biology, Faculty of Mathematics and Natural Sciences, Universitas Padjadjaran, Indonesia
²East Coast Environmental Research Institute (ESERI), Universiti Sultan Zainal Abidin, Malaysia
³Faculty of Informatics and Computing, Universiti Sultan Zainal Abidin, Malaysia
⁴Department of Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Padjadjaran, Indonesia

E-mail: ruly@unpad.ac.id; hafizanjuahir@unisza.edu.my; musmat567@gmail.com; sukono@unpad.ac.id: sudradjat@unpad.ac.id; m.nurzaman@unpad.ac.id

*Corresponding author: sukono@unpad.ac.id

Abstract. Economic sector development is an effort to increase national economic growth which is usually measured by increasing Gross Domestic Product (GDP). But economic sector development affects CO2 emissions in million tons of sectoral greenhouse gases. Therefore, this paper performs modeling and analysis of CO2 emissions in million tons of sectoral greenhouse gases in Indonesia. Sectors analysed include: industrial, commercial, household, transportation, power plant, and others. The objective is to obtain a model estimator and determine the level of sectoral CO2 emissions contribution to total CO2 emissions in Indonesia. Modeling is done by using Cobb-Douglas multiplicative production function approach. While to make estimation of model parameters is done by using Ordinary Least Square (OLS) method. The analysis results show that sectors: industrial, commercial, household, transportation, power plant, significantly contribute to total CO2 emissions. While the other sectors do not significantly affect the total CO2 emissions. The analysis results also show that the five-sector CO2 emission contribution model for total CO2 emissions significantly follows Cobb-Douglas multiplicative function, with correlation strength level which is 99.9% coefficient of determination. Thus, the model estimators obtained are excellent and can be used to determine the level of sectoral CO2 emissions contribution to total CO2 emissions in Indonesia.

1. Introduction

Economic sector is a framework of development of national monetary growth. Successful development of economic growth is usually measured by raising Gross Domestic Product (GDP). National economic development is a process of raising total and per capita income, which is accompanied by fundamental changes in the economic structure of a country and the equitable distribution of income for its people [13]. Through economical sector development, it is possible to change the agrarian economic structure into industrial economic structure. So that the economic activities undertaken by the state will be more diverse and dynamic [14; 15]. However, unplanned economic sector development may impact damage to the environment. Economic development through the industrial sector has indeed brought prosperity to its people, but has also given a negative impact on the ecological system of the world [3; 17]. It is now felt that the world community is in a great crisis situation. The world community has felt its own bad consequences of air pollution and global warming caused by the build-up of CO2 emissions in the
air [4; 10; 18]. CO2 emissions have the most important role in climate change because these gases continue to accumulate in the atmosphere in large numbers [5; 9; 22]. Therefore, it is emerging to study the effect of sectoral economic growth on total CO2 emission increase.

Filiz and Omer [7], has undertaken research on long-term and causal relationships between the industrial, cement and steel production sectors; power plants; oil consumption; and increased CO2 emissions in Turkey. The study was conducted using an auto-regression vector testing approach (VAR) in the period 1990 to 2010. The results showed that there are two way linkage of Granger causality between CO2 emissions and cement production and power plant. While the results of impulse-response, indicating that CO2 emissions largely influenced by the industrial production sector, followed by cement production sectors, power plants, oil consumption, and steel production. Aye and Edoja [2] examined the effect of economic sector growth on CO2 emission changes using panel data analysis consisting of 31 developing countries. The analysis results show that the growth of the economic sector has a negative effect on changes in CO2 emissions under low growth conditions, but has a positive effect on high growth conditions, with a marginal effect higher in high growth conditions. Coi and Abdullah [6], predicted changes in CO2 emissions using linear regression. CO2 emission change trends are related to the development of economic sectors and other variables, such as demand and supply sectors in the economy and energy consumption sectors. They say that the linear model is one of the common methods used to explain the correlation between CO2 emissions and related economic sector variables. Similar studies have also been conducted by Khobai & Roux [11], Mrabet et al. [16], and Shukla [19]. Kosztowniak [12] analysed the influence of foreign direct investment net entry on GDP in Poland between 1994-2012, using the Cobb-Douglas production function. In such an analysis defined conditions necessary for the positive influence of foreign direct investment in Poland. Also assumed the assumptions are the Cobb-Douglas production function and predicted changes in GDP value in Poland. In the analysis also identified factors that significantly affect economic growth in Poland. A similar study has also been conducted by Anghelache et al. [1], Gogtay et al. [8], Shukla [19], and Wang & Fu [21].

Referring to the above description, this paper intends to model and analyse CO2 emissions in million tons of sectoral greenhouse gases in Indonesia. The economic sectors analysed include: industrial, commercial, household, transportation, power plant, and others. The purpose of this study is to obtain a model estimator and determine the level of CO2 emissions contribution from economic sector development to the total CO2 emissions in Indonesia. The estimated model is expected to be used in the determination of CO2 emission level control policy from sectoral economic development activities to CO2 emissions.

2. Methodology

The research methodology here is a scientific process or way of getting the facts by a step or rule that can use to get an estimator of the correlation model between sectoral CO2 emissions with total CO2 emissions. After several observations, it is suspected that the correlation of sectoral CO2 emissions to total CO2 emissions follows the Cobb-Douglas production function model. Therefore, this section discusses the methods including: Cobb-Douglas production function, parameter estimation method, goodness of fit test, and forecasting.

2.1. Cobb-Douglas production function

The form of production function is the interaction between the input (input) and the output (output). There are several forms of production functions such as the Cobb-Douglas production function. Referring to Wang & Fu [21] and Kosztowniak [12], generally, the Cobb-Douglas is a multiplicative production function with more than two independent variables can be given as follows:

$$G = \phi z_1^{\beta_1} z_2^{\beta_2} z_3^{\beta_3} \cdots z_k^{\beta_k} e^c,$$

(1)
where \( G \) is the dependent variable (output); \( \phi \) is coefficient intercept; \( Z_1, Z_2, Z_3, \ldots, Z_k \) are independent variables (input); \( \beta_1, \beta_2, \beta_3, \ldots, \beta_k \) are coefficient of elasticity of independent variables; \( e = 2.7182818285 \) natural numbers; and \( \varepsilon \) is error (residual).

The sum of the elasticity coefficients can be used as a measurement of returns to scale. There are 3 characteristics of size return to scale, as follows:

- **Decreasing returns to scale**, when \( \sum_{i=1}^{k} \beta_i < 1 \)
  Indicates that the Cobb-Douglas production function has additional characteristics of diminishing returns on a production scale, where output increases with a proportion smaller than the proportion of inputs.
- **Constant returns to scale**, when \( \sum_{i=1}^{k} \beta_i = 1 \)
  Indicates that the Cobb-Douglas production function has a constant yield characteristic on the scale of production, when all inputs rise in certain proportions, then the produced output rises in exact proportion to the proportion of the input.
- **Increasing returns to scale**, when \( \sum_{i=1}^{k} \beta_i > 1 \)
  Indicates that the Cobb-Douglas production function has additional characteristics of increased yield on a production scale, where output increases with a proportion greater than the proportion of inputs.

When the left and right sides of equation (1) are transformed by natural logarithms, then equation (2) is obtained, and resulting the following linear regression equation:

\[
\ln G = \ln \phi + \beta_1 \ln Z_1 + \beta_2 \ln Z_2 + \beta_3 \ln Z_3 + \ldots + \beta_k \ln Z_k + \varepsilon
\]

Furthermore, if we give \( Y = \ln G, \ \beta_0 = \ln b, \ X_1 = \ln Z_1, \ X_2 = \ln Z_2, \ X_3 = \ln Z_3, \ldots, \ X_k = \ln Z_k \), then equation (2) is a multiple linear regression equation which can be expressed as follows [1]:

\[
Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \ldots + \beta_k X_k + \varepsilon.
\]

The estimation of the multiple linear regression equation (3) is expressed as follows:

\[
\hat{Y} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \ldots + \beta_k X_k.
\]

So that, the residual values can be determined as \( \varepsilon = Y - \hat{Y} \).

### 2.2. Method of parameter estimation

In this section we discuss the method of parameter estimation from multiple regression equation in general. Once the estimation is done by using the matrix equation approach, the multiple linear regression equation (3) can be given as follows [1]:

\[
Y = X \beta + \varepsilon.
\]

Where:

\[
Y = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}, \quad X = \begin{bmatrix} 1 & X_{12} & X_{13} & \cdots & X_{1k} \\ 1 & X_{22} & X_{23} & \cdots & X_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_{n2} & X_{n3} & \cdots & X_{nk} \end{bmatrix}, \quad \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix}, \quad \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix},
\]

where \( Y \) is matrix of \((n \times 1)\), \( X \) is matrix of \((n \times (k + 1))\), \( \beta \) is matrix of \(( (k + 1) \times 1)\), and \( \varepsilon \) is matrix of \((n \times 1)\).
Based on Anghelache et al. [1] and Gogtay et al. [8], to obtain the estimator values of matrix parameters, it can be determined by using ordinary least square method (OLS), i.e. by minimizing the residual squares amount which the equation is as follows:

\[
\text{Minimization } \sum \varepsilon_i^2 = \varepsilon^T \varepsilon = (Y - X\beta)^T (Y - X\beta).
\] (6)

Where \( \varepsilon^T = (Y - X\beta)^T \) is matrix transpose of \( \varepsilon \). Because \( \beta^T X^T Y \) is a scalar, so its value is equal to its transpose, i.e. \( Y^T XB \). In the process of minimizing, the equation (6) is obtained as follows:

\[
\frac{\partial \sum \varepsilon_i^2}{\partial \beta} = -2X^T Y + 2X^T X\beta = 0.
\] (7)

Further, solving equation (7) can be obtained the parameter estimator matrix as follows:

\[
\beta = (X^T X)^{-1} X^T Y.
\] (8)

Where the matrix \((X^T X)^{-1}\) is the inverse of the matrix \((X^T X)\).

Note that this approach method can only be used if the matrix \((X^T X)\) has an inverse.

2.3. Test of goodness of fit

After the parameter estimator values of multiple linear regression models are obtained, the parameter estimator need to be tested its goodness of fit at a defined level of significance. The goodness of fit test is to determine that the model is significantly able to describe the actual data. Goodness of fit test to model parameter estimator is done by using partial significance test, simultaneous significance test, assumption of residual normality test, and coefficient of determination test.

- Test of individual significance parameters

  The individual parameter estimator significance test is performed to test each parameter \( \beta_i \) (\( i = 0,1,2,\ldots,k \)) of equation (8), individually in affecting the dependent variable. To test parameter estimators \( \beta_i \), determine hypothesis \( H_0: \beta_i = 0 \) and alternate hypotheses \( H_1: \beta_i \neq 0 \). Testing is done by using statistic \( t_{\text{Stat}} \). The equation is as follows:

\[
t_{\text{Stat}} = \frac{\hat{\beta}_i}{SE(\hat{\beta}_i)}.
\] (9)

where \( SE(\beta_i) \) is the standard error of parameters \( \beta_i \) (\( i = 0,1,2,\ldots,k \)).

The test criterion is to reject the hypothesis \( H_0 \) and accept the hypothesis \( H_1 \), if statistic \(|t_{\text{Stat}}| \geq t_{(n-k, \frac{1}{2}\alpha)} \) or statistic \( \Pr[t_{\text{Stat}} < \alpha] \), where statistic \( t_{(n-k, \frac{1}{2}\alpha)} \) is the critical value of the distribution of \( t \) at the level of significance \( 100(1 - \alpha) \% \), \( k \) is the number of parameters, and \( n \) is the number of data. Instead, accept the hypothesis \( H_0 \) and reject the hypothesis \( H_1 \), if statistic \(|t_{\text{Stat}}| \leq |t_{(n-k, \frac{1}{2}\alpha)}| \) or statistic \( \Pr[|t_{\text{Stat}}| \geq \alpha] \), where statistic \( |t_{(n-k, \frac{1}{2}\alpha)}| \) is the critical value of the distribution of \( t \) at the level of significance \( 100(1 - \alpha) \% \), \( k \) is the number of parameters, and \( n \) is the number of data [1; 20].

- Test of the significance of parameters simultaneously
Test the significance of parameter estimators simultaneously, is conducted to simultaneously test parameter estimators $\beta_i \ (i = 0,1,2,\ldots,k)$ of equation (4), in affecting the dependent variable. In this test is determined the hypothesis $H_0$: $\beta_0 = \beta_1 = \beta_2 = \ldots = \beta_k = 0$ and alternate hypotheses $H_1$: $\exists \beta_0 \neq \beta_1 \neq \beta_2 \neq \ldots \neq \beta_k \neq 0$. Testing is done by using statistic $F$, the equations is as follows:

$$F_{stat} = \frac{MS_{Reg}}{MS_{Error}}, \quad (10)$$

where $MS_{Reg}$ mean square due to regression, and $MS_{Error}$ mean square due to residual variation.

The test criterion is to reject the hypothesis $H_0$ and accept the hypothesis $H_1$, if statistic $F_{stat} > F(k-1,n-k,1-\alpha)$ or statistic $Pr[F_{stat}] < \alpha$, where statistic $F(k-1,n-k,1-\alpha)$ is the critical value of the distribution of $F$ at the level of significance $100(1-\alpha)$%, $k$ is number of parameters, and $n$ is the number of data. Instead, accept the hypothesis $H_0$ and reject the hypothesis $H_1$, if statistic $F_{stat} \leq F(k-1,n-k,1-\alpha)$ or statistic $Pr[F_{stat}] \geq \alpha$, where statistic $F(k-1,n-k,1-\alpha)$ is the critical value of the distribution of $F$ at the level of significance $100(1-\alpha)$%, $k$ is number of parameters, and $n$ is the number of data [8; 20].

- Test of assumption of residual normality

The normality assumption test is performed to determine that the residual data is distributed following a normal distribution. Normality assumption test can be done using Kolmogorov-Smirnov (KS) statistic. In this test, the hypothesis is determined $H_0$: normally distributed data, and alternatives of hypotheses $H_1$: data is not normally distributed. Testing is done by determining residual standard deviation by using equation as follows:

$$S_{\varepsilon_i} = \sqrt{\frac{\sum_{i=1}^{n} (\varepsilon_i - \bar{\varepsilon})^2}{n-1}}. \quad (11)$$

Then, transformed residual value $\varepsilon_i$ become $z_i$ using equation $z_i = (\varepsilon_i - \bar{\varepsilon}) / S_{\varepsilon_i}$. The probability value $P(z_i)$ is determined using standard normal distribution tables. While the probability $S(z_i)$ is determined using the equation $S(z_i) = randl(z_i)/n$. Next, we calculate the values of absolute difference $|S(z_i) - P(z_i)|$. The statistical value of Kolmogorov-Smirnov $KS_{Stat}$ is determined using the following formula [20]:

$$KS_{Stat} = \max(|S(z_i) - P(z_i)|). \quad (12)$$

Next, the statistical value $KS_{Stat}$ is compared with statistical critical value of $KS(\alpha,n-1)$, at the level of significance $\alpha = 0.05$. The test criterion is to reject the hypothesis $H_0$ and accept the hypothesis $H_1$, if statistic $KS_{Stat} > KS(\alpha,n-1)$, where $n$ is the number of data. Instead, accept the hypothesis $H_0$ and reject the hypothesis $H_1$, if statistic $KS_{Stat} \leq KS(\alpha,n-1)$, where $n$ is the number of data [20].

- Coefficient of determination
Referring to Anghelache et al. [1] and Gogtay et al. [8], the coefficient of determination $R^2$ is used to measure how much the variability of independent variables to the dependent variable, based on the level of correlation power. So the coefficient of determination value is the ability or strength of the independent variable $X_i$ ($i = 1, 2, ..., k$) in affecting the dependent variable $Y$. Value of the coefficient of determination $R^2$ is determined using the following equation:

$$R^2 = \frac{\sum_{i=1}^{n}(\hat{Y}_i - \bar{Y})^2}{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}.$$  \hspace{1cm} (13)

The value of determination coefficient is ranged between 0 and 1. The value of determination is small close to 0 which means the variation of independent variable to the dependent variable is very weak. Conversely, a value close to 1 means that the variation of the independent variable to the dependent variable is very strong.

### 2.4. Forecasting (Prediction)

Forecasting (prediction) is very important in modeling. Forecasting is done where the complexity and uncertainty are faced by the model maker. So as to require the level of error as small as possible or the level of accuracy as big as possible, there are many methods that can be used to measure the error rate of a forecasting model. One of which is Mean Absolute Percentage Error (MAPE). Determining the MAPE value is determined by using the equation as follows:

$$\text{MAPE} = \left( \frac{1}{n} \sum_{i=1}^{n} \left| \frac{Y_i - \hat{Y}_i}{Y_i} \right| \right) \times 100\% \hspace{1cm} (14)$$

The smaller the MAPE value, the smaller the value of error, which also means the greater the accuracy of the model in predicting an event [20].

### 3. Result and Analysis

In this section we will discuss the result and analysis which include: data analysed, multiple linear regression model estimation, and estimation of Cobb-Douglas model, and forecasting (prediction).

#### 3.1. Data analysed

In this study the data analysed are CO2 emissions in million tons of sectoral greenhouse gases in Indonesia. It is a secondary data obtained from the Ministry of Energy and Mineral Resources of the Republic of Indonesia, in the period 2000 to 2015. Data are distributed from CO2 emissions emitted by each sector of the economy which includes: industrial, commercial, household, transportation, and power generation. The data can be shown as the graph given in Figure 1.
Figure 1. Graph of CO2 emissions of each sector and total

Taking into account to Figure 1, it appears that the total CO2 emissions emitted by the economic sector, from year to year, it tends to increase with a high trend. The contribution of total CO2 emissions from the largest economic sector is from the industry. The next contribution is from power generation, transportation, household, commercial and others. The contribution of each economic sector fluctuates with a relatively constant trend. Furthermore, based on the data in this study, it carried out modeling the contribution of CO2 emissions from each sectoral to total CO2 emissions from year to year, using Cobb-Douglas model. While for parameter estimation, it is done by using multiple linear regression equation, ordinary least square method (OLS).

3.2. Multiple linear regression modeling

In this section we estimate the multiple linear regression models of natural logarithm data from CO2 emissions emitted by sectors: industrial, commercial, household, transportation, power generation, and others, to the natural logarithm of total CO2 emissions. Estimation is done by using Ordinary Least Square (OLS) method approach which refers to equation (8). The estimation of multiple linear regression models is done by using Minitab 16 software, and the result of estimation process as given in Table 1.

Table 1. Results of estimation of multiple linear regression model of CO2 emission

| Predictor       | Coef | SE Coef |  T   |  P   |
|-----------------|------|---------|------|------|
| Constant        | 1.49100 | 0.30250 | 4.93 | 0.001 |
| ln(Industrial)  | 0.20228 | 0.01507 | 13.47 | 0.000 |
| ln(Commercial)  | -0.17962 | 0.09575 | -1.88 | 0.063 |
| ln(Household)   | 0.12674 | 0.02685 | 4.72 | 0.001 |
| ln(Transportation) | 0.31914 | 0.07033 | 4.54 | 0.001 |
| ln(Other)       | 0.10970 | 0.12420 | 0.88 | 0.380 |
| ln(Power Plant) | 0.25113 | 0.08329 | 3.02 | 0.003 |

S = 0.00629268 R-Sq = 99.9% R-Sq(adj) = 99.9%

Analysis of Variance

| Source         | DF | SS   | MS    | F     | P     |
|----------------|----|------|-------|-------|-------|
| Regression     | 6  | 0.65494 | 0.10916 | 2756.65 | 0.000 |
| Residual Error | 9  | 0.00036 | 0.00004 |       |       |
| Total          | 15 | 0.65530 |        |       |       |

By Noting the results of the estimation process in Table 1, If the significance level of the level is equal to \( \alpha = 0.05 \), then the estimator parameters: Constant, ln(Industrial), ln(Household), ln(Transportation), and ln(Power Plant), have a probability \( P < 0.05 \). Unless the estimator of the
parameters ln(Commercial) and ln(Other) has a probability $P \geq 0.05$, it means estimators of the parameters ln(Commercial) and ln(Other) individually significant have no effect on the dependent variable ln (Total CO2 Emissions). Therefore, gradually the non-significant parameter estimators are excluded from the multiple linear regression model estimators. We begin with the parameter estimator having the greatest probability, that is, the estimator of the coefficient parameter ln(Other), and then the re-estimation process is done. The re-estimation process is also performed using Minitab 16 software, and the estimation results are given in Table 2.

Based on the results of the re-estimation given in Table 2, it appears that the probability of the parameter estimator of all values is less than 0.05. This indicates that there is no longer an estimate to be issued, and thus no further re-estimation is required. Based on the values in Table 2, and with reference to equation (3) multiple linear regression models the estimation results can be expressed in terms of equations as follows:

$$Y = 1.7119 + 0.28587X_1 - 0.10505X_2 + 0.11783X_3 + 0.37434X_4 + 0.18695X_5 + \varepsilon. \quad (15)$$

Where $Y$ is ln(total CO2 emissions), $X_1$ is ln(Industrial), $X_2$ is ln(Commercial), $X_3$ is ln(Household), $X_4$ is ln(Transportation), $X_5$ is ln(Power Plant), and $\varepsilon$ is residual.

Furthermore, equation (15) needs to be tested its goodness of fit. First, the individual significance test is performed on the parameter estimators $\hat{\beta}_0 = 1.71190$, $\hat{\beta}_1 = 0.28587$, $\hat{\beta}_2 = -0.10505$, $\hat{\beta}_3 = 0.11783$, $\hat{\beta}_4 = 0.37434$, and $\hat{\beta}_5 = 0.18695$. The test is performed by referring to equation (9), using a significance level $\alpha = 0.05$. At the level of significance $\alpha = 0.05$ and with degrees of freedom $df = 16-6 = 10$, from the table of distribution of $t$, we obtained the critical value of statistic $t_{10;0.025} = 2.228$. Based on the estimation results presented in Table 2, the value of statistics for $t_{Stat} = 10.18$ so it is clear that statistic $t_{Stat} > t_{10;0.025}$. This shows that parameter estimator $\hat{\beta}_0$ is significant. The significance test also needs to be done for the parameter estimators $\hat{\beta}_1$, $\hat{\beta}_2$, $\hat{\beta}_3$, $\hat{\beta}_4$, and $\hat{\beta}_5$. Testing is also done in the same way, and each result also significantly affects the variable $Y$ (ln(total CO2 emissions)). Second, the simultaneous significance test of the parameter estimator $\hat{\beta}_0 = 1.71190$, $\hat{\beta}_1 = 0.28587$, $\hat{\beta}_2 = -0.10505$, $\hat{\beta}_3 = 0.11783$, $\hat{\beta}_4 = 0.37434$, and $\hat{\beta}_5 = 0.18695$. The test is performed by referring to equation (10), using a significance level $\alpha = 0.05$ and degrees of freedom $df = (6-1; 16-6)$. At the level of significance $\alpha = 0.05$.

| Table 2. Result of re-estimation process without independent variable ln(Other) |
|---------------------------------------------------------------|
| **The regression equation is** |
| ln(Total CO2 Emissions) = 1.71 + 0.285 ln(Industrial) - 0.105 ln(Commercial) + 0.118 ln(Household) + 0.374 ln(Transportation) + 0.187 ln(Power Plant) |
| **Predictor** | **Coeff** | **SE** | **T** | **P** |
|----------------|-----------|-------|-------|-------|
| Constant       | 1.71190   | 0.16810 | 10.18 | 0.000 |
| ln(Industrial) | 0.28587   | 0.01435 | 19.93 | 0.000 |
| ln(Commercial) | -0.10505  | 0.04489 | -2.36 | 0.040 |
| ln(Household)  | 0.11783   | 0.02450 | 4.79  | 0.001 |
| ln(Transportation) | 0.37434 | 0.05184 | 11.78 | 0.000 |
| ln(Power Plant) | 0.18695   | 0.04016 | 4.65  | 0.001 |
| $R^2 = 0.00522350$ | $R^2 = 99.9\%$ | $R^2 = 99.9\%$ |
| **Analysis of Variance** |
| **Source** | **DF** | **SS** | **MS** | **F** | **P** |
| Regression    | 5      | 0.65491 | 0.13090 | 3362.42 | 0.000 |
| Residual Error| 10     | 0.00039 | 0.00004 |        |        |
| Total         | 15     | 0.65530 |        |        |        |
0.05 and with degrees of freedom \( df = (5; 10) \), of the table of distribution of \( F \), then we obtained the critical value of statistic \( F_{(5;10;0.05)} = 4.10 \). Based on the estimation results presented in Table 2, the value of statistic \( F_{Stat} = 3382.42 \). So it is clear that statistic \( F_{Stat} > F_{(5;10;0.05)} \). This shows that parameter estimators \( \hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3, \hat{\beta}_4, \) and \( \hat{\beta}_5 \), simultaneously have significant effect on the variable \( Y \) (ln(total CO2 emissions)).

Third, the assumption of residual normality assumption, is intended to determine that residuals are normally distributed. Testing is done by referring to equation (12), also by using the level of significance \( \alpha = 0.05 \). Based on the results of testing using Minitab 16 software, it gives residual \( \varepsilon \sim N(0.0125, 3.617604) \). Fourth, the determination coefficient test is intended to control the level of strength of the relationship between independent variables with the dependent variable. Determination of coefficient value of determination is done by referring to equation (13). Based on the test results presented in Table 2, it is found that the value of determination \( R^2 = 99.9\% \). This shows that the relationship between the independent variable and the dependent variable is very strong.

After a goodness of fit test and all test results are significant, thereby obtaining an estimator of multiple linear regression models of CO2 emissions given in equation (15) can be expressed as equations:

\[
\hat{Y} = 1.7119 + 0.28587X_1 - 0.10505X_2 + 0.11783X_3 + 0.37434X_4 + 0.18695X_5. \tag{16}
\]

Furthermore, the model estimator of multiple linear regression equation (16), with reference to equation (1) can be obtained by estimating Cobb-Douglas production function model as follows:

\[
\hat{G} = e^{1.7119Z_1^{0.28587}Z_2^{-0.10505}Z_3^{0.11783}Z_4^{0.37434}Z_5^{0.18695}}, \tag{17}
\]

where \( \hat{G} \) is the output variable estimator (total CO2 emissions); \( e = 2.7182818285 \) is a natural number as a multiplier; \( Z_1 \) is a variable of industrial sector; \( Z_2 \) is variable of commercial sector; \( Z_3 \) is variable of household sector; \( Z_4 \) is a variable of transport sector; and \( Z_5 \) is a variable of power generation sector.

Furthermore, Cobb-Douglas production function estimator (17) is used for forecasting. First, forecasting is done by using in-sample data whose results are shown as forecasting graphs as shown in Figure 3. In Figure 3, it appears that the forecasting graph (prediction) coincides with the actual data (total CO2 emissions), and forecasting using the in-sample data gives a MAPE error rate of 0.00394169 or 0.394169%. This means that the model estimator has an accuracy level for forecasting of 99.605831%.
4. Second, forecasting based on out-sample data by estimating the contribution values of sectoral CO2 emissions. For example in certain years, it is estimated that the contribution values of sectoral CO2 emissions as given in Table 3. Using Cobb-Douglas production function estimator (17) obtained the forecast of total CO2 emissions in million tons as given in Table 3.

Table 3. Forecasting of sectoral and total CO2 emissions in million tons in Indonesia

| Years | Industrial | Commercial | Household | Transportation | Power Plant | Total CO2 Emissions |
|-------|------------|------------|-----------|----------------|-------------|---------------------|
| 2016  | 105.4      | 2.13       | 21.43     | 140.18         | 194.44      | 473.7564790         |
| 2017  | 117.58     | 1.78       | 21.85     | 142.42         | 213.26      | 510.9674851         |
| 2018  | 129.76     | 1.43       | 22.27     | 144.66         | 232.08      | 550.8038038         |

4. Conclusion

In this paper we have conducted modeling and analysis of CO2 emissions in million tons of sectoral greenhouse gases in Indonesia. The modeling is done by using Cobb-Douglas multiplicative production function, while for parameter estimation is done by using ordinary least square method (OLS). Based on the analysis results, it can be concluded that sectoral economic growth covering sectors: industrial, commercial, household, transportation, and power plant has resulted in increasing total CO2 emission in Indonesia. The estimated model obtained year by year significantly follows the Cobb-Douglas production function, with a deterministic coefficient of 99.9%. The estimated model of Cobb-Douglas production function model has an insurance rate for forecasting of 99.605831%. Using the estimated Cobb-Douglas production model estimator model, it can be used for forecasting total CO2 emissions in future periods, by estimating the values of sectoral economic variables. If in 2018 it is estimated that the contribution of industrial sector CO2 emission is 129.76; commercial 1.43; household 22.27; transportation 144.66; and 232.08 power plant, the total CO2 emissions will reach 550.8038038 million
tons. Therefore, the results of this study are expected to be taken into consideration in environmentally friendly economic development, to support efforts to decrease CO2 emissions in the air.

5. Acknowledgements

Acknowledgments are conveyed to the Rector, Director of Directorate of Research, Community Involvement and Innovation, and the Dean of Faculty of Mathematics and Natural Sciences, Universitas Padjadjaran, with whom the Internal Grant Program of Universitas Padjadjaran was made possible to, fund this research. The grant is a means of enhancing research and publication activities for researchers at Universitas Padjadjaran.

References

[1] Anghelache C, Anghel M G, and Popovic M 2015 Multiple Regressions Used in Analysis of Private Consumption and Public Final Consumption Evolution. *International Journal of Academic Research in Accounting, Finance and Management Sciences*, Vol. 5, No.4, October 2015, pp. 69–73.

[2] Aye G C and Edoja P E 2017 Effect of Economic Growth on CO2 Emission in Developing Countries: Evidence from a Dynamic Panel Threshold Model. *Cogent Economics & Finance* (2017), 5: 1379239, pp. 1-22. https://doi.org/10.1080/23322039.2017.1379239

[3] Asici A A 2011 Economic growth and its impact on environment: A panel data analysis. *MPRA* Paper No. 30238, posted 18. April 2011 12:50 UTC. Online at https://mpra.ub.uni-muenchen.de/30238/.

[4] Budiono R, Juahir H, Mamat M, Sukono and Nurzaman M 2018 Modelling Interaction of CO2 Concentration and the Biomass Algae Due to Reduction of Anthropogenic Carbon Based on Predator-Prey Model. *International Journal of Applied Environmental Sciences*, ISSN 0973-6077 Volume 13, Number 1 (2018), pp. 27-38. © Research India Publications. http://www.ripublication.com

[5] Cioca L I, Ivascu L, Rada E C, Torretta, V., and Ionescu, G. 2015. Sustainable Development and Technological Impact on CO2 Reducing Conditions in Romania. *Sustainability* 2015, 7, 1637-1650; doi:10.3390/su7021637, pp. 1-14.

[6] Choi C S and Abdullah L 2016 Prediction of Carbon Dioxide Emissions Using Two Linear Regression-based Models: A Comparative Analysis. *Journal of Applied Engineering (JOAE)*, 4(4), April-2016 (Volume-IV, Issue-IV), pp. 305-312.

[7] Filiz O and Omer O 2012 An Analysis of CO2 Emissions of Turkish Industries and Energy Sector. *Regional and Sectoral Economic Studies*, Vol.12-2 (2012), pp. 65-85.

[8] Gogtay N J, Deshpande S P and UM Thatte U M 2017 Principles of Regression Analysis. *Journal of The Association of Physicians of India*, Vol. 65, April 2017, pp. 48-52.

[9] Halushchak M, Bun R, Shpak N and Valakh M 2016 Modelling and spatial analysis of greenhouse gas emissions from fuel combustion in the industry sector in Poland. *Ecotechmod. An International Quarterly Journal*, 2016. Vol. 5. No. 1, pp. 19–26.

[10] Jin S H, Bai L, Kim J Y, Jeong S J and Kim K S 2017 Analysis of GHG Emission Reduction in South Korea Using a CO2 Transportation Network Optimization Model. *Energies* 2017, 10, 1027, pp. 1-18. doi:10.3390/en10071027.

[11] Khobai H and Roux P L 2017 The Relationship between Energy Consumption, Economic Growth and Carbon Dioxide Emission: The Case of South Africa. *International Journal of Energy Economics and Policy*, 2017, 7(3), pp. 102-109. available at http://www.econjournals.com.
[12] Kosztowniak A 2014 Analysis of the Cobb-Douglas Production Function as a Tool to Investigate the Impact of FDI Net Inflows on Gross Domestic Product Value in Poland in the Period 1994–2012. *Öconomia Copernicana*, 5(4), 169-190. https://doi.org/10.12775/OeC.2014.032.

[13] Knight K W and Schor J B 2014 Economic Growth and Climate Change: A Cross-National Analysis of Territorial and Consumption-Based Carbon Emissions in High-Income Countries. *Sustainability* 2014, 6, pp. 3722-3731; doi:10.3390/su6063722

[14] Lappalainen S and Kouvo P 2004 Evaluation of Greenhouse Gas Emissions from Landfills in the ST. Petersburg Area-Utilization of Methan in Energy Production, Metgas. *Publication* 10, 2004, pp. 1-56. Lappeenranta University of Technology, Northern Dimension Research Centre. URL: www.lut.fi/nordl.

[15] Mikayilov J I, Marzio G M and Hasanov F J 2018 The Impact Of Economic Growth On Co2 Emissions In Azerbaijan. *Working Paper*, No. 102, February 2018, pp. 1-33. Working Paper Series. ISSN 1973-0381

[16] Mrabet A, Achairi R and Ellouze A 2013 The Two-Way relationship between Economic Growth and CO2 Emissions. *International Conference on Business, Economics, Marketing & Management Research* (BEMM’13) Volume Book: Economics & Strategic Management of Business Process (ESMB), pp. 1-4. Copyright © IPCO 2013.

[17] Phimpanthavong H 2013 The Impacts of Economic Growth on Environmental Conditions in Laos. *Int.J.Buss.Mgt.Eco.Res.*, Vol 4(5),2013, pp. 766-774. www.ijbmer.com | ISSN: 2229-6247.

[18] Reinecke R and Casey N H 2017 A whole farm model for quantifying total greenhouse gas emissions on South African dairy farms. *South African Journal of Animal Science* 2017, 47 (No. 6), pp. 883-894.

[19] Shukla M 2007 Estimation of CO2 Emissions using Energy Input-Output (EIO) Tables for India. *V.R.F. Series*, No. 430, October 2007. Institute of Developing-Economies, Japan External Trade Organization.

[20] Sukono, Hidayat Y, Suhartono, Sutijo B, Bon A T B and Supian S 2016 Indonesian Financial Data Modeling and Forecasting By Using Econometrics Time Series and Neural Network. *Global Journal of Pure and Applied Mathematics*. ISSN 0973-1768 Volume 12, Number 4 (2016), pp. 3745-3757. © Research India Publications. http://www.ripublication.com/gjipam.htm.

[21] Wang X and Fu Y 2013 Some Characterizations of the Cobb-Douglas and CES Production Functions in Microeconomics. *Abstract and Applied Analysis*, pp. 1-6.

[22] Zushman J M 2016 Modelling greenhouse gas emissions from light duty vehicles. *30 ECTS thesis*, submitted in partial fulfillment of a *Magister Scientiarum* degree in Environment and Natural Resources, Faculty of Industrial Engineering, Mechanical Engineering and Computer Science University of Iceland.