Modeling the impacts of constant price GDP and population on CO2 emissions using Cobb-Douglas model and ant colony optimization algorithm

Sukono1, Betty Subartini1, Priscila Thalia, Sudradjat Supian1, Eman Lesmana1, Ruly Budiono2, Hafizan Juahir3

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

Email: sukonon@unpad.ac.id; betty.subartini@unpad.ac.id; friscilathalia31@gmail.com; sudradjat@unpad.ac.id; man.msie@gmail.com; budiono.ruly@gmail.com; hafizanjuahir@unisza.edu.my

*corresponding author: sukonon@unpad.ac.id

Abstract. The per capita Gross Domestic Product (GDP) measures a country’s economic growth. Increasing GDP is a dream of all countries, but generally, GDP increases often have a negative impact with increasing CO2 emissions. This paper intends to model the impact of GDP growth based on constant prices and the population in increasing CO2 emissions in Indonesia. Modeling is done by using Cobb-Douglas model production function, where parameter estimation is done by using ant colony optimization algorithm. Furthermore, model estimators are used for forecasting CO2 emission concentrations. The results of the analysis show that the impact of GDP based on constant prices and population significantly follows the Cobb-Douglas model of production, with the coefficient of elasticity is 0.819405999 and 0.834930855, respectively. The value of determination was obtained at 97.4%, indicating that the correlation between GDP at constant prices and population with increasing CO2 emissions in air is very strong. Estimator model obtained has a level of accuracy for forecasting is 0.98478981 or 98.4798981%. Thus, the model estimator obtained is able to describe the actual data pattern.

Keywords: GDP, Population, CO2 emissions, Cobb-Douglas model, ant colony optimization.

1. Introduction

Economic development is an activity to promote economic growth for every country. One of the economic growths is measured by the amount of Gross Domestic Product (GDP) per capita. However, high GDP growth is often followed by declining environmental quality. GDP growth is not the only one to be a successful measure of development of a country [1; 2]. Keeping nature and the environment alive is also a factor of development success. GDP growth along with population growth also has a negative effect, and significantly affects the quality of the environment as indicated by increased CO2 emissions [7; 9]. Any 1 percent increase in GDP and population, will lead to a decrease in environmental quality by 9.11 percent. It shows that GDP growth and population growth will lead to a decrease in the level of environmental quality in the form of rising CO2 emissions [8; 10].

Thus, GDP growth and population growth have an impact on increasing CO2 emissions in the air [17; 19]. How patterns of GDP growth and population growth together contribute to increasing CO2 emissions, this becomes an important issue for a study. Previous researchers, among them, have done
many studies are: Laceheb et al. [6] reviewed the economic growth and CO$_2$ emissions in Algeria in the period 1971-2009. The study was conducted using an integrated autoregressive lag model, where the data were extracted from the World Bank Indicators. The results show that over the long term, it was revealed that income and population growth significantly affected CO$_2$ emissions. The results of the study also show that only the population showed increased CO$_2$ emissions from liquid fuel consumption. Yang et al. [15], investigating the interaction of economic growth and pollution in the industrialization process in Zhejiang, time series data from three types of pollution indices from 1981 to 2006. The investigation is done by using cointegration test and granger causality test. The test results show that GDP per capita results in increased emissions of industrial waste pollution except solid waste. Conversely, it is not true that industrial waste pollution increases can lead to economic growth. Conrad & Cassar [3], Sasana & Putri [13], and Zhao & Wang [18] have also conducted similar studies. Salami [11], in his study proposed a hybrid algorithm to find solutions to optimization problems using ant colony and genetic programming. An evolutionary process of ant colony optimization algorithms adapts genetic operations to improve ant motion to achieve optimal solutions. The proposed algorithm integrates with the optimal solution, by collecting the most effective sub-solutions. Samsami [12] conducted a study on the application of Ant Colony Optimization (ACO) to predict CO$_2$ emissions in Iran, based on socio-economic indicators. Forms of linear and non-linear equations were developed to predict CO$_2$ emissions using ACO. The results obtained from the study provide useful insights into energy systems and CO2 emission control modeling. Other studies that have done the same are Duan & Yong [4] and Yu & Wang [16].

Therefore, based on the exposure of previous researchers, in this paper the research on the modeling of the impact of GDP at constant price and population on CO$_2$ emission is done by using Cobb-Douglas model and particle swarm optimization algorithm. The research objects are GDP growth data and population, and CO$_2$ emissions in Indonesia in period 1967-2014. The goal is to derive a significant model estimate, and can be used for forecasting the development of CO$_2$ emissions with minimum errors or high accuracy.

2. Methodology

This methodology discusses some of the models and algorithms used in the study, which include Cobb-Douglas production function, model parameter estimation using ant colony optimization (ACO), model significance test, and forecasting (prediction).

2.1. Cobb-Douglas model

Referring to Felipe & Adams [5] and Omri [10], the Cobb-Douglas production function is an equation that contains two or more variables, a dependent variable and several independent variables. Cobb-Douglas multiplicative error term term production function is expressed by the following equation:

$$ G_t = \phi X_{1t}^{\theta_1} X_{2t}^{\theta_2} e^{\epsilon_t}, $$

where $G_t$ is to declare an output variable in the form of CO2 emissions; $X_{1t}$ is to state the input variable in the form of GDP at constant price per capita; $X_{2t}$ is to declare the input variable the number of population; $\phi$, $\theta_1$, and $\theta_2$ is to declare the Cobb-Douglas model parameters; and $\epsilon_t$ is exponential of residual.

Referring to Felipe and Adams [5], the coefficient of elasticity of production $\rho$ is to state the percentage change in output, divided by the percentage of input changes. The coefficient of production elasticity is the ratio of the relative change of output produced, to the relative changes in the number of inputs that affect. The coefficient of output elasticity of GDP at constant prices input, $\rho_1$ calculated using the equation as follows:
\[ \rho_1 = \frac{\% \Delta G_t}{\% \Delta X_{1t}} = \theta_1. \]

The output elasticity coefficient of GDP at constant price input can also be determined by using coefficient parameters \( \theta_1 \) of the Cobb-Douglas production function.

While the output elasticity coefficient of the input population, \( \rho_2 \) can be calculated using the equation as follows:

\[ \rho_2 = \frac{\% \Delta G_t}{\% \Delta X_{2t}} = \theta_2. \]

The output elasticity coefficient of the input population can also be calculated by using parameter coefficient \( \theta_2 \) of the Cobb-Douglas production function.

Referring to Felipe and Adams [5], the sum of the production elasticity coefficients \( \theta_1 + \theta_2 \) explain the size of the venture scale or often called return to scale. There are three forms of return to scale, as given below:

- When \( \theta_1 + \theta_2 = 1 \), the Cobb-Douglas production function shows a constant return to scale, meaning that the increase in input will be followed by an equivalent proportional output increase.
- When \( \theta_1 + \theta_2 < 1 \), Cobb-Douglas production function shows the scale with decreasing return (decreasing return to scale), meaning that the proportional increase of output is smaller than the proportional addition of input.
- When \( \theta_1 + \theta_2 > 1 \), Cobb-Douglas production function shows scale with increasing return (increasing return to scale), meaning that proportional addition of output greater than proportional addition of input.

Equation (1), if the left and right segments are taken natural logarithm, then obtained multiple linear regression equation as follows:

\[ \ln G_t = \ln \phi + \theta_1 \ln X_{1t} + \theta_2 \ln X_{2t} + \epsilon_t. \]

Suppose that \( Y_t = \ln G_t ; \theta_0 = \ln \phi ; Z_{1t} = \ln X_{1t} ; \) and \( Z_{2t} = \ln X_{2t} , \) then the equation can be expressed as follows:

\[ Y_t = \theta_0 + \theta_1 Z_{1t} + \theta_2 Z_{2t} + \epsilon_t. \] \hspace{1cm} (2)

The estimate of the multiple linear regression equation (2) is as follows:

\[ \hat{Y}_t = \theta_0 + \theta_1 Z_{1t} + \theta_2 Z_{2t}. \] \hspace{1cm} (3)

Equation (2) is a linear regression with parameters \( \theta_0 , \theta_1 , \) and \( \theta_2 \) as well as residuals (error) \( \epsilon_t \). Constants \( \theta_0 \) is intercept, \( \theta_1 \) and \( \theta_2 \) is a coefficient parameter of elasticity of Cobb-Douglas production function.

If the estimation is done by using ordinary least square method (OLS, then from multiple linear regression equation (2) it is necessary to establish the optimization model as follows:

\[ \text{Minimisasi } \sum \epsilon_t^2 = \sum (Y_t - \hat{Y}_t)^2 = \sum (Y_t - \theta_0 - \theta_1 Z_{1t} - \theta_2 Z_{2t})^2. \] \hspace{1cm} (4)
It means that, the estimation is aimed at determining the estimator values of $\theta_0$, $\theta_1$, and $\theta_2$ which can minimize the sum of squares of residuals $\sum e_i^2$. The process of minimizing the number of residual squares, in this research is done by using Ant Colony Optimization (ACO) algorithm.

2.2. Ant colony optimization algorithm

According to Duan & Yong [4], Salami [11], Samsami [12], and Yue & Wang [16], Ant Colony Optimization (ACO) is adopted from the treatment of ant colonies, called ant systems. Naturally, ant colonies are able to find the shortest route on the way from the nest to the source of food and back to the nest. At the time of ants traveling, ants always leave a piece of information called pheromone, where it travels and to mark the route. Pheromone works for communication between ants when creating routes. The road route the ant passes from the nest to where the food can be given as Figure 1. The ACO algorithm has been widely used to solve various optimization cases.

![Figure 1. The route of the ant path from the nest to the food place](image)

According to Duan & Yong [4], Salami [11], Samsami [12], and Yue & Wang [16], the workings of the Ant Colony algorithm are described as follows:

1. Initially, ant colonies stroll randomly.
2. When the ant colony finds a different path, such as finding the intersection, the ant colony determines the direction of the path randomly.
3. Some ant colonies choose to walk to the upper route, and some choose to walk down the route.
4. Whenever he finds some ant food back to his colony, while putting a sign with a pheromone trail.
5. Because the path taken down the shorter route, the ant colony passing down first arrives, which assume the velocity of all ant colonies to be the same.
6. The pheromone left by the ant colony on the shorter route smells stronger, than the pheromone on the longer path.
7. Other ant colonies are more interested in following the route down, because the pheromone aroma is stronger.

According to Duan & Yong [4], Salami [11], Samsami [12], and Yue & Wang [16], the ant colony algorithm requires several variables, and the steps to determine the shortest distance.

Stage 1:

a. The parameters required in the ant colony algorithm are as follows:

- Intemity of ant colony traces between places $r_{ij}$ and delta changes $\Delta r_{ij}$. 

Intensities $\tau_{ij}$ should be initialized before starting the iteration cycle. Change $\Delta\tau_{ij}$ initialized after the completion of one iteration cycle. Change $\Delta\tau_{ij}$ used to determine $\tau_{ij}$ on the next iteration cycle.

- The ants colony iteration cycle stays $Q$
  The iteration cycle $Q$ is a constant used in the equation used to determine the delta of change $\Delta\tau_{ij}$. The iteration cycle value $Q$ determined by the user.

- Ant colony trace intensity control constant $\alpha$
  Traffic intensity control constant $\alpha$ used in the probability equations of places visited, and serves as control of the intensity of ant colony traces. Value $\alpha$ determined by the user.

- Time visibility controller $\beta$
  Time visibility controller $\beta$ is used in the probability equation of locations visited by ant colonies, and serves as a visibility controller. Value $\beta$ determined by the user.

- Visibility between places $\eta_{ij}$
  Visibility between places $\eta_{ij}$ used in the probability equation of places visited by ant colonies. Value $\eta_{ij}$ is the result of $1/d_{ij}$ (the distance of the place).

- The number of ant colonies $m$
  The number of ant colonies $m$, is the number of ant colonies that the iteration cycle does in the ant colony algorithm. Value $m$ determined by the user.

- The vaporization constant of ant colony traces $\rho$
  The vaporization constant of ant colony traces $\rho$ used to determine for $\tau_{ij}$ on the next iteration cycle. Value $\rho$ determined by the user.

- Number of maximum iteration cycles $NC_{\text{max}}$
  Number of maximum iteration cycles $NC_{\text{max}}$, is the maximum number of shrimp cycles that will take place. The iteration cycle will stop according to the value $NC_{\text{max}}$ pre-defined by the user.

- Charging the coordinates of the place
  On the filling of the coordinates of the place can be input in accordance with the desired by the user. The more coordinates, the longer the ant colony travels.

b. Initialize first place of each ant colony.

After initialization $\tau_{ij}$ done, then $m$ ant colonies are placed in a particular random first place. For parameter values $\alpha$ should be rated between ranges $0 \leq \alpha \leq 1$, it is intended to avoid unlimited accumulation of pheromone on the route. Because the amount of pheromone left behind is unlikely to grow stronger, it gets weaker. For parameter values $\beta$ should not be given a value of 0, because if given a value of 0 then the results achieved are not optimum. Not optimum here means a condition where the length of the journey achieved is not minimum.

Stage 2:

Filling the first place into the taboo list. The initialization result of the first place of each ant in stage 1, must be filled as the first element in the taboo list. The result of this step is the taboo list for each ant colony with a specific spot index.

Stage 3:

Arrange the route of each ant colony's visit to a number of places. An ant colony that has spread to a number of places, began to travel from the first place of each as a place of origin, and to one of the other
places as the destination. Then from the second place, each ant colony continues its journey, choosing one of the places that is not on the taboo list, as the next destination. Ant colony trips continue continuously until all the places one by one visited. If \( s \) specifies the index of the order of visits, the place of origin is expressed as \( \text{tabu}_k(s) \), and other places are declared as \( \{N - \text{tabu}_k(s)\} \), then to determine the destination, used the probability equation where to visit as follows:

\[
P_{ij}^k(s) = \begin{cases} 
\frac{[\tau_{ij}(s)]^\alpha [\eta_{ij}]^\beta}{\sum_{k' \in \{N - \text{tabu}_k(s)\}} [\tau_{ij}(s)]^\alpha [\eta_{ij}]^\beta} & \text{if } j \in \{N - \text{tabu}_k(s)\}, \\
0 & \text{for other } j.
\end{cases}
\]

where \( i \) as the index of place of origin, and \( j \) as the destination index.

Stage 4:

a. Calculation of the route length of each ant colony

Long route calculation of closed or \( \text{L}_k \) each ant colony is performed after all ants complete one iteration cycle. Calculations are based \( \text{tabu}_k(s) \) each with the following equation:

\[
\text{L}_k = d_{\text{tabu}_k(n),\text{tabu}_k(1)} + \sum_{s=1}^{n-1} d_{\text{tabu}_k(s),\text{tabu}_k(s+1)},
\]

where \( d_{ij} \) is the distance between \( i \) and \( j \) which is calculated using following equation:

\[
d_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}.
\]

b. Shortest distance search

After \( \text{L}_k \) each ant colony is calculated, there will be a minimum price of the closed route length of each cycle, or \( \text{L}_{\min} \) and the minimum price of the overall closed line length is \( \text{L}_{\min} \).

c. Determination of the value of price change of the ant colony footprint intensity between places \( \Delta \tau_{ij} \)

An ant colony leaves footprints on the route between places it passes. The existence of evaporation and differences in the number of ant colonies that pass, causing the possibility of changes in the price of ant colony footprint intensity between places. The value of the price change in the intensity of the ant colony footprint is determined using the equation:

\[
\Delta \tau_{ij} = \frac{m}{\sum_{k=1}^{m} \Delta \tau_{ij}^k},
\]

where \( m \) is the number of ant colonies; and \( \tau_{ij} \) is the length of each ant colony's route. Where \( \Delta \tau_{ij}^k \) is the value of the price change of the ant colony footprint intensity between places, for each ant colony calculated using the equation:

\[
\Delta \tau_{ij}^k = \begin{cases} 
\frac{Q}{\text{L}_k} & \text{for } (i,j) \text{ is place of origin and destination in } \text{tabu}_k, \\
0 & \text{for other } (i,j)
\end{cases}
\]

where \( Q \) is the ant colony iteration cycle of the ants; and \( \text{L}_k \) is length close tour (lct).

Stage 5:

a. Calculation of intern colony footprint intensity values between places for subsequent iteration cycles. The intensity of the ant colony footprints between places on all routes is likely to change, as there is evaporation, and the differences and the number of ant colonies passing through. For the next iteration cycle, the ant colony passing through the route the intensity value has changed. The
intensity of the ant colony footprint between places for an iterative cycle is then calculated using the equation:

\[ \tau_{ij} = (1 - \rho) \times \tau_{ij} + \Delta \tau_{ij}. \]  

(10)

b. Rearranging the change in the intensity of ant colony footprints between places. For the next iteration cycle the change in the intensity value of ant colony traces between places needs to be reset to have a value equal to zero.

Stage 6:

Dismiss the taboo list and repeat step 2 if needed. Taboo lists need to be emptied to be filled again with a new place sequence in the next iteration cycle. If the maximum number of iteration cycles has not been reached, the algorithm is repeated from the taboos filling stage, with the ant colony footprint intensity parameter values between the updated places.

2.3. The significance test of model parameters

The significance test of model parameters is done by using simultaneous test (statistical test of \( F \)), test individually (statistical test of \( t \)), assumption of residual normality, and determination of efficiency. In this study, the significance test of model parameters was done using Minitab 16 assistance.

- **Statistic test of \( F \)**

This test is conducted with the aim to know that all parameter estimators simultaneously significantly affect the dependent variable. The parameter significance test simultaneously the hypothesis used is:

\[ H_0 : \theta_0 = \theta_1 = \theta_2 = 0, \]

meaning that free variables do not affect significantly to the dependent variable. While the alternative hypothesis is

\[ H_1 : \exists \theta_0 \neq \theta_1 \neq \theta_2 \neq 0, \]

meaning that at least there are independent variables that have a significant effect on the dependent variable. The value of count statistics \( F_{Stat} \) is determined using the equation [14; 5]:

\[ F_{Stat} = \frac{\text{MS}_{Reg}}{\text{MS}_{Error}} = \frac{\sum_{t=1}^{n} (Y_t - \bar{Y})^2}{k - 1} \times \frac{n - k}{\sum_{t=1}^{n} (Y_t - \hat{Y}_t)^2}. \]  

(11)

Next, the calculated statistical value \( F_{Stat} \) is compared with statistical critical values \( F(\alpha, df) \) with a significant level \( \alpha \) and degree of freedom \( df = (k - 1, n - k) \). Where \( n \) is number of samples, and \( k \) is the number of parameters. The test criterion is to reject the hypothesis \( H_0 \) and to accept the hypothesis \( H_1 \), if \( F_{Stat} > F(\alpha, df) \). This means that the independent variables affect the dependent variable. Instead, to accept the hypothesis \( H_0 \) and to reject the hypothesis \( H_1 \), if \( F_{Stat} \leq F(\alpha, df) \). Meaning that simultaneously independent variable does not affect the dependent variable.

- **Statistic test of \( t \)**

Individual significance test parameters, conducted with the aim to determine whether each parameter estimator significantly contributes to affect the dependent variable. The significance test of individual parameters is done with the hypothesis used is \( H_0 : \theta_i = 0, \) \( (i = 0, 1, 2), \) means that the independent variable does not significantly affect the dependent variable. While the alternative hypothesis is \( H_1 : \theta_i \neq 0, \) \( (i = 0, 1, 2), \) meaning that that the independent variables significantly influence the dependent variable. Value statistics \( t_{Stat} \) is determined using the equation [14; 5]:

...
\[ t_{\text{stat}} = \frac{\theta_i}{SE(\theta_i)}. \]  

Next, the calculated statistical value \( t_{\text{Stat}} \) is compared with statistical critical values \( t(\frac{\alpha}{2}, df) \) with a significant level \( \alpha \) and degree of freedom \( df = (n - k) \). Where \( n \) is number of samples, and \( k \) is the number of parameters. The testing criterion is to reject the hypothesis \( H_0 \) and to accept the hypothesis \( H_1 \), if \( |t_{\text{Stat}}| > |t(\frac{\alpha}{2}, df)| \). This means that each independent variable significantly affects the dependent variable. Instead accepting the hypothesis \( H_0 \) and rejecting the hypothesis \( H_1 \), if \( |t_{\text{Stat}}| \leq |t(\frac{\alpha}{2}, df)| \). This means that each independent variable has no significant effect on the dependent variable.

- Residual normality test
  The normality assumption test is performed to determine that the residual data follows the normal distribution. Normality assumption test in this paper is done by using Kolmogorov-Smirnov (KS) statistic. In this test, the hypothesis used is \( H_0 \): the data is normally distributed, with the alternative hypothesis being \( H_1 \): data is not normally distributed. Test assumption of residual normality is done by determining standard deviation by using equation:
  \[ S_{\varepsilon_i} = \sqrt{\frac{\sum_{i=1}^{n}(\varepsilon_i - \overline{\varepsilon})^2}{n-1}}. \]  

Then, transformed value \( \varepsilon_i \) into \( z_i \) using equations \( z_i = (\varepsilon_i - \overline{\varepsilon})/S_{\varepsilon_i} \). While the value of probability \( P(z_i) \) using standard normal distribution tables. For probability \( S(z_i) \) determined using the equation \( S(z_i) = \text{rand}l(z_i)/n \). Next, calculated the values of absolute difference \( |S(z_i) - P(z_i)| \). The statistical value of Kolmogorov-Smirnov \( KS_{\text{Stat}} \) determined using the expression [14]:
  \[ KS_{\text{Stat}} = \max\{|S(z_i) - P(z_i)|\}. \]  

Determine the critical value of statistics \( KS(\alpha, n-1) \), at a significant level \( \alpha = 0.05 \). The criteria of the test is to reject the hypothesis \( H_0 \) if statistics \( KS_{\text{Stat}} > KS(\alpha, n-1) \).

- Coefficient of determination value
  Determination coefficient value is used to measure correlation strength between all independent variables with non-free variable. Value of coefficient of determination \( R^2 \) can be determined by equation [14] as follows:
  \[ R^2 = \frac{SS_{\text{Reg}}}{SS_{\text{Corr}}} = \frac{\sum_{i=1}^{n}(\hat{Y}_i - \overline{Y})^2}{\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2}. \]  

If the value of coefficient of determination \( R^2 = 1 \), it means that the percentage of independent variable contribution to the dependent variable \( Y_i \) amount 100%. If the value of the coefficient of determination \( R^2 = 0 \), it means that the independent variable does not contribute to the dependent variable \( Y_i \).
2.4. Level of accuracy of forecasting

To measure the error rate of the model estimator in predicting a value, it can be measured based on the error generated from the forecast. According to Sukono et al. [14], to measure the level of forecast error is usually done by means of Mean Square Error (MSE). MSE is the average size of the quadratic value of forecasting errors. To determine the value of MSE can be done using the equation as follows:

\[
MSE = \frac{1}{n} \sum_{t=1}^{n} (Y_t - \hat{Y}_t)^2
\]

A model is said to be appropriate, if the error resulting from the prediction is very small close to zero or its accuracy rate approaches 100%.

3. Results and Anaysis

In this section, we discussed the result and analysis, which included data analyzed, model parameter estimation, model estimator significance test, deterministic coefficient determination, and measurement of model accuracy in forecasting.

3.1. Data analyzed

The data analyzed in this study is secondary data, i.e. GDP data based on constant prices, population, and CO\textsubscript{2} emissions 1967-2014. The data is obtained from the World Bank's official website (https://data.worldbank.org/). Descriptive statistics are presented with a view to providing an overview of the quantitative data used in the study. The presentation of descriptive statistical data in this study was conducted using Microsoft Excel 2013. The descriptive statistics can be seen in Table 1. Where \( Y_t \) is the CO\textsubscript{2} emissions, \( Z_{1t} \) is GDP at constant prices, and \( Z_{2t} \) is the total population.

| Statistic  | \( Y_t \)     | \( Z_{1t} \)  | \( Z_{2t} \)  |
|------------|---------------|---------------|---------------|
| Mean       | 1.02753743    | 1838.98859    | 181239285.4   |
| Median     | 0.89894000    | 1750.89300    | 1.8300000     |
| Maximum    | 2.55975023    | 3692.94288    | 2.5500000     |
| Minimum    | 0.23191548    | 656.746824    | 105907403     |
| Std.Dev    | 0.57736700    | 835.768600    | 446693252     |

Graphs of CO\textsubscript{2} emissions data, GDP at constant prices, and population numbers are presented in Figure 2.
3.2. Estimation of model parameters

GDP data at constant prices, population and CO$_2$ emissions obtained from World Bank official website are transformed into natural logarithmic values by reference equation (2). The transformed data is then modeled into the Cobb-Douglas model with reference to (1). Estimation of model parameters was performed using the ant colony optimization (ACO) algorithm developed in section 2.2.

Model parameter estimation process is done by using Matlab R2015a software, the stages are as follows:

- Click start → choose Matlab R2015a.
- On the File menu, select New Script.
- Then write the objective function along with the CO$_2$ emissions data as $Y_t$, GDP at constant prices as $Z_{1t}$, and the population as $Z_{2t}$ on the sheet script. Then save the script by name `obj1_function.m`. The objective function is as given in (4).
- Open Command Windows to call the script, which has been made in the previous stages.
- To get parameter values based on data already entered into function, run `obj1_function.m` how to click run then type `Y =particleswarm(@obj1_function,3)` on Command Windows and press enter, the execution results appear the estimated parameter values.

Based on the estimation process using ACO obtained values estimator model parameters are respectively

\[ \hat{\theta}_0 = -22.0640901, \quad \hat{\theta}_1 = 0.819405999, \quad \text{and} \quad \hat{\theta}_2 = 0.834930855. \]

Thus, by reference to equation (2), the multiple linear regression model of the estimate can be expressed as follows:

\[ Y_t = -22.0640901 + 0.819405999Z_{1t} + 0.834930855Z_{2t} + \xi_t. \]  

(17)

In order to ensure that the model of the estimation result using ACO is to significantly draw the pattern of the analyzed data, it is necessary to test the significance of the model estimator. The testing stages of significance of the model estimator are as follows.

3.3. Testing the significance of model estimators

The model significance significance test performed in this section includes: simultaneous parameter estimator significance test, individual parameter estimator significance test, and assumption of residual normality test.

- Test of simultaneous significance

In the simultaneous test of significant parameters, the hypothesis used is $H_0: \hat{\theta}_0 = \hat{\theta}_1 = \hat{\theta}_2 = 0$, which means the independent variables together do not have a significant effect on the dependent variable. While the alternative hypothesis used is $H_1: \exists \hat{\theta}_0 \neq \hat{\theta}_1 \neq \hat{\theta}_2 \neq 0$, which means there are independent variables that influence significantly to the dependent variable.
Statistic value of $F_{Stat}$ is calculated by using equation (11), and the statistic value is obtained $F_{Stat} = 853.2750831$. Meanwhile, using the level of significance $\alpha = 0.05$, and degrees of freedom $df = (k - 1; n - k) = (2; 45)$, of the distribution table of $F$ can be obtained the critical value of statistics $F_{(0.05;2;45)} = 3.205$. Thus, it can clearly be shown that $F_{Stat} > F_{(0.05;2;45)}$, therefore it rejects the hypothesis $H_0$ and accepts the hypothesis $H_1$. This means that at a significant 5% or 95% confidence level, GDP at constant prices and population totals, or at least a parameter estimator $\hat{\theta}_0 = -22.0640901$, $\hat{\theta}_1 = 0.819405999$, and $\hat{\theta}_2 = 0.834930855$ affect the CO$_2$ emission-free variable in Indonesia.

- Test of individual significance
In testing the significance of parameter estimators individually, each parameter estimator one by one is tested. Starting with the test for the parameter estimator $\hat{\theta}_0$, where the hypothesis used is $H_0: \hat{\theta}_0 = 0$, it means that a constant estimator $\hat{\theta}_0$ has no partial effect on the dependent variable. While the alternative hypothesis is $H_1: \hat{\theta}_0 \neq 0$, meaning that a constant estimator $\hat{\theta}_0$ has a partial effect on the dependent variable.

Value of statistics $t_{Stat}$ is determined using equation (12), and the statistical value is obtained $t_{Stat} = 3.103705989$. Meanwhile, by using a significant level $\alpha = 0.05$, and degree of freedom $df = n - k = 45$, from the distribution table of $t$ can be obtained critical value statistic $t_{(0.025;45)} = 2.01410$. Therefore, it can be shown that $|t_{Stat}| > |t_{(0.025;45)}|$, rejects the hypothesis $H_0$ and therefore accepts the hypothesis $H_1$. This means that with a significant level of 5% or 95% confidence level, the parameter estimator is a constant $\hat{\theta}_0 = -22.0640901$ individually has an effect on the CO2 emission-free variable.

Tests of individual significance also need to be made against the estimator coefficient parameters $\hat{\theta}_1$ and $\hat{\theta}_2$. Using the same way, the test results show that the parameter estimator coefficients $\hat{\theta}_1 = 0.819405999$, and $\hat{\theta}_2 = 0.834930855$, significantly each also has an effect on the CO$_2$ emission-free variable.

- Test of the assumption of residual normality
Testing assumption of residual normality in this paper, conducted by using statistical test Kolmogorov-Smirnov. In testing the assumption of this residual normality, the hypothesis used is $H_0$: the residual data follows the normal distribution. While the alternative hypothesis is $H_1$: the residual data does not follow the normal distribution.

Determination of statistical value of Kolmogorov-Smirnov $D_{Stat}$ can be done by using equation (13), and obtained statistical value $D_{Stat} = 0.07004871$. Meanwhile, using the level of significance $\alpha = 0.05$ and the number of data $n = 48$, from the Kolmogorov-Smirnov table obtained the critical value of statistics $D_{(0.05;48)} = 0.198$. Therefore, it can be shown that $D_{Stat} > D_{(0.05;48)}$, thus rejecting the hypothesis $H_0$ and therefore accepting hypothesis $H_1$. This means that the residual data follows the normal distribution. From the process of testing the residual normality also obtained value estimator mean $\hat{\mu}_e = 0.005327$, and variance estimator $\hat{\sigma}_e^2 = 0.015567$. Thus, it can be shown that $\varepsilon_1 \sim N(0.005327, 0.015567)$.

3.4. Determination of the value of the coefficient of determination
In this section, calculating the coefficient of determination can be used to measure the strength of the correlation between independent variables with the dependent variable. Value of coefficient of determination \( R^2 \) using equation (15), and the result is \( R^2 = 0.980211005 \). This means that the large contribution of GDP at constant prices and population to the variation (ups and downs) of CO\(_2\) emissions by 98%, while the remaining 2% is caused by other factors outside the GDP at current prices and the population of Indonesia.

3.5. Measurement of the accuracy of forecasting

Measuring the level of accuracy of forecasting is very important, to measure the accuracy of the model estimator when used for forecasting. It is known that the model obtained from the estimation process using ACO is given as in (17). If it refers to equation (3), then equation (17) means having the model estimator as follows:

\[
\hat{Y}_t = -22.0640901 + 0.819405999Z_{1t} + 0.834930855Z_{2t}
\]

While it is known that \( Y_t = \ln G_t \); \( \theta_0 = \ln \phi \); \( Z_{1t} = \ln X_{1t} \); and \( Z_{2t} = \ln X_{2t} \), so this last equation can be expressed as follows:

\[
\hat{G}_t = e^{-22.0640901}X_{1t}^{0.819405999}X_{2t}^{0.834930855}
\]

In equation (18), the amount of elasticity \( \hat{\theta}_1 + \hat{\theta}_2 = 1.654336854 \), show that \( \hat{\theta}_1 + \hat{\theta}_2 > 1 \). Therefore, equation (18) is a Cobb-Douglas production function that has a characteristic increasing return to scale. This means that there is an increase in the values of the proportion of input units \( X_{1t} \) and \( X_{2t} \), will lead to an increase in the proportion of output of the CO\(_2\) emission-free variable doubled from the proportion of the input unit \( X_{1t} \) and \( X_{2t} \). Furthermore, equation (18) is used to forecast the data out sample. The objective is to measure the accuracy of model estimators when used in forecasting.

Measurements are performed starting with the determination of the error rate, and then used to calculate the accuracy of the model estimator. Determination of error value can be done by using data in sample. The error value is calculated by using the MSE as given in equation (16), and the value of MSE is obtained 0.01521019 or 1.521019%. Therefore, it can be interpreted that the accuracy of the model estimator of GDP at constant prices and the population on CO\(_2\) emissions, using the ACO algorithm is 0.98478981 or 98.478981%. The graph of actual data and forecasting results as given in Figure 3, as follows:
Furthermore, if we estimate the approximate values $X_1$ and $X_2$ as given in Table 2, then forecasting out sample values of CO$_2$ emissions by 2015, 2016, 2017, and 2018 can be done. Incorporates the values of each variable that increase from the previous year, and forecasting is done using equation (18) the result as given in Table 2, column $\hat{G}_t$.

**Table 2.** Forecasting out sample value of emissions CO$_2$

| Years ($t$) | $X_1$ | $X_2$ | $\hat{G}_t$ |
|------------|-------|-------|-------------|
| 2015       | 3825.77917 | 258229969 | 2.381406066 |
| 2016       | 3958.615465 | 261328822 | 2.473462628 |
| 2017       | 4091.45176  | 264427675 | 2.566405774 |
| 2018       | 4224.288055 | 267526528 | 2.660236089 |

From Table 2, it appears that there is an increase of input unit in the form of GDP based on constant prices and population, causing an increase in CO$_2$ emissions that multiply in Indonesia.

### 4. Conclusion

In this paper, we have been done in estimating the model of the impact of GDP growth at constant prices and population on CO$_2$ emissions, using the Cobb-Douglas model of the parameters estimated using ant colony optimization (ACO). Based on the analysis, it can be concluded that the impact of the growth of GDP at constant prices and population on CO$_2$ emissions is significantly following the Cobb-Douglas production function with elasticity estimator $\hat{\theta}_1 = 0.819405999$, and $\hat{\theta}_2 = 0.834930855$. Because $\hat{\theta}_1 + \hat{\theta}_2 = 1.654336854$, show that $\hat{\theta}_1 + \hat{\theta}_2 > 1$, the obtained production function estimator has the characteristics of increasing return to scale. This indicates that the proportion of output additions doubled from the proportion of input additions. Furthermore, forecasting based on data in sample, error rate measured using MSE obtained value of 0.01521019 or 1.521019%. This means that the estimator accuracy rate of the model estimator of the impact of GDP at constant prices and the population on CO$_2$ emissions, using the PSO algorithm is 0.98478981 or 98.478981%. This suggests that the Cobb-Douglas production function estimator obtained is relatively accurate to be used for increasing CO$_2$ emissions, as a result the growth of GDP at constant prices and population. Taking into account such conditions of high CO$_2$ emissions, such modeling is expected to be taken into consideration for the authorities in policy making in controlling the increase of CO$_2$ emissions in Indonesia.

### 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] Alam J 2014 On the Relationship Between Economic Growth and CO2 Emissions: The Bangladesh Experience. IOSR Journal of Economics and Finance (IOSR-JEF). Volume 5, Issue 6, pp-36-41. E-ISSN: 2321-5933.

[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] Conrad E and Cassar L F 2014 Decoupling Economic Growth and Environmental Degradation: Reviewing Progress to Date in the Small Island State of Malta. Sustainability 2014, 6, pp. 6729-6750; doi:10.3390/su6106729.

[4] Duan P and Yong A I 2016 Research on an Improved Ant Colony Optimization Algorithm and its Application. International Journal of Hybrid Information Technology. Vol.9, No.4 (2016), pp. 223-234. http://dx.doi.org/10.14257/ijhit.2016.9.4.20.

[5] Felipe J and Adams F G 2005 The Estimation of the Cobb-Douglas Function: A Retrospective View. Eastern Economic Journal, Vol. 31, No. 3, Summer 2005, pp. 427-445.

[6] Lacheheb M, Rahim A S A and Sirag A 2015 Economic Growth and Carbon Dioxide emissions: Investigating the Environmental Kuznets Curve Hypothesis in Algeria. International Journal of Energy Economics and Policy, 2015, 5(4), 1125-1132.

[7] Mehrizi M A, Atashi S M and Elahi M 2012 The effect of variables population growth, urbanization and economic growth on CO2 Emissions in Iran. African Journal of Business Management, Vol.6 (28), pp. 8414-8419, 18 July, 2012, pp. 8414-8418. Available online at http://www.academicjournals.org/AJBM. DOI: 10.5897/AJBM11.2020. ISSN 1993-8233 ©2012 Academic Journals.

[8] Mitic P, Ivanovic O M and Zdravkovic A 2017 A Cointegration Analysis of Real GDP and CO2 Emissions in Transition Countries. Sustainability2017, 9, 568, pp. 1-18. doi:10.3390/su9040568.

[9] Obradovic S and Lojanica N 2017 Energy Use, CO2 Emissions and Economic Growth-Causality on a Sample of SEE Countries. Economic RexEurEch-Ekonomskais Ražīvania, 2017, voL. 30, no. 1, 511–526. http://dx.doi.org/10.1080/1331677X.2017.1305785.

[10] Omri A 2013 CO2 Emissions, Energy Consumption and Economic Growth Nexus in MENA Countries: Evidence from Simultaneous Equations Models. Energy Economics 40 (2013) 657–664. http://dx.doi.org/10.1016/j.eneco.2013.09.003.

[11] Salami, Al N M A 2009 Ant Colony Optimization Algorithm. UbiCC Journal, Volume 4, Number 3, August 2009, pp. 823-826.

[12] Samsami R 2013 Application of Ant Colony Optimization (ACO) to forecast CO2 Emission in Iran. Bulletin of Environment, Pharmacology and Life Sciences. Bull. Env. Pharmacol. Life Sci., Vol 2 (6) May 2013: 95-99. ©2013 Academy for Environment and Life Sciences, India. Online ISSN 2277-1808. Journal’s URL:http://www.bepls.com.

[13] Sasana H and Purti AE 2018 The Increase of Energy Consumption and Carbon Dioxide (CO2) Emission in Indonesia. E3S Web of Conferences31, 01008 (2018), ICENIS 2017, pp. 1-5. https://doi.org/10.1051/e3sconf/20183101008.

[14] Sukono, Hidayat Y, Suhartono Sutijjo B BonA.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.rippublication.com/gjpm.htm.
[15] Yang L, Yuan S and Sun L 2011 The Relationships between Economic Growth and Environmental Pollution Based on Time Series Data: An Empirical Study of Zhejiang Province. *Journal of Cambridge Studies*, Volume 7, No. 1, pp. 33-42.

[16] Yue and Wang X 2015 An Improved Ant Colony Optimization Algorithm for Solving TSP. *International Journal of Multimedia and Ubiquitous Engineering*. Vol.10, No.12 (2015), pp.153-164. http://dx.doi.org/10.14257/ijmue.2015.10.12.16.

[17] Zarzoso I. M, Morancho A B and Lage R M 2006 The Impact of Population on CO2 Emissions: Evidence from European Countries. *Paper. NOTA DI LAVORO 98.2006*. web site: www.feem.it.

[18] Zhao Y and Wang S 2015 The Relationship between Urbanization, Economic Growth and Energy consumption in China: An Econometric Perspective Analysis. *Sustainability* 2015, 7, pp. 5609-5627; doi:10.3390/su7055609.

[19] Zhu Q and Peng X 2012 The Impacts of Population Change on Carbon Emissions in China During 1978–2008. *Environmental Impact Assessment Review* 36 (2012) 1–8. journal homepage: www.elsevier.com/locate/eiar.