J.Agromet 23(1):1-10,2009

CLIMATE PROJECTION OVER INDONESIA BASED ON THE TOTAL FOSSIL FUEL CO_2 EMISSION PREDICTION USING THE BOX-JENKINS ARIMA MODEL

(Proyeksi Iklim Wilayah Indonesia Berdasarkan Prakiraan Emisi CO2 dari Penggunaan Bahan Bakar Fosil Menggunakan Model Arima Box-Jenkins)

Eddy Hermawan

Center for Application of Atmospheric Sciences and Climate of
National Institute of Aeronautics and Space (LAPAN)
Jalan Dr. Djundjunan 133, Bandung 40173, Indonesia
E-mail : eddy@bdg.lapan.go.id

ABSTRACT

This paper mainly discusses about the development of estimation models raising the rate of gas emissions of carbon dioxide (CO_2) as the main parameters of global warming in Indonesia. This is important to remember not many comprehensive scientific study which shows that the impact of global warming has actually experienced by Indonesia. Using Box-Jenkins method and the stage of identification, assessment, and testing, then the best prediction model obtained for the above data, the model of ARIMA (8,1,3). This means that the predicted value for the next year depending on the data before and 8 years 3 years earlier error. In the validation data with predicted results, the MAD (Mean Absolute Deviation) is relatively high. However, the pattern of results followed the pattern predicted almost the original data with a correlation value of 99%. Based on this result, we can estimate the climate projection over Indonesia, especially during 2012-2014.

Keywords : ARIMA model (8,1,3), Box-Jenkins methods, climate projection, total fossil fuel emission CO_2, Indonesia

INTRODUCTION

During the pre-industrial era the atmospheric carbon dioxide concentration has been stable (IPCC 1996). This stability is due to the equilibrium situation when the global carbon dioxide absorption rate of about 220 GtC/a carbon to cold ocean water and growing biomass is balanced by an emission of 220 GtC/a from warm ocean water and decomposing biomass. When the global mean temperature has been high, the equilibrium has changed towards a slightly higher atmospheric carbon dioxide concentration, probably because of decreased solubility of carbon dioxide in the warmer ocean water (Ahlbeck 2000).

When carbon dioxide is emitted from fossil fuels, cement production, or deforestation, the increased partial pressure of carbon dioxide in the atmosphere will force an increase of the absorption rate and thus a net sink flow of carbon to the backmixed surface layer of the oceans and to the biosphere.As we know in 1992, the Intergovernmental Panel on Climate Change (IPCC), presented a group of emission scenarios for different greenhouse gases. A "mid-range" emission scenario was called IS92a.

Penyerahan naskah : 02 Januari 2009
Diterima untuk diterbitkan : 12 Februari 2009

However, due to limited fossil fuel reserves, IS92a seems exaggerated when looking 100 years into the future. Numerous new emission scenarios, higher and lower than IS92a, have been created recently. In order to predict future atmospheric carbon dioxide concentrations, emission scenarios may be inserted into computerized global carbon dioxide models. For IS92a, the IPCC claims that the atmospheric concentration would increase from today's value of 369 ppm (ppm=parts per million by volume) to 705 ppm in the year 2100. This is possible only if the rate of atmospheric carbon dioxide increase would very soon begin to increase from today's value of 1.5 ppm/year up to 4 ppm/year.

In reality, we can see that the increase rate of atmospheric carbon dioxide has, despite the substantial increase of carbon dioxide emissions, remained on a very stable level during the recent 30 years. In fact, the airborne fraction, or the portion of the yearly emissions that stays in the atmosphere, has decreased from 52% in the year 1970 to 39% today. The IPCC model using IS92a implies however a nearly constant future airborne fraction.

Although, is not included in the list of countries as the largest contributor to global warming, but with the forest fires which occurred almost throughout the year, especially in the dry season length (as in 1982 and 1997), estimated there were about 2.5 billion tons of CO_2 that we contribute to global warming.In this paper, we mainly concern on the projection of the total fossil fuel of carbon dioxide (CO_2) emission over Indonesia based on the Box Jenkins ARIMA model analysis. The steps analysis to get that the best model prediction of that data will be discussed in this paper.

MATERIALS AND METHODS

The main data used in this study is the CO_2 emission taken from Indonesian territory that are downloaded from the web-side http://cdiac.ornl.gov/ftp/trends/emissions/ido.dat. From this web-site address, then the set of numbers obtained as follows (Table 1). The data is then in-plot in the form of time-series to be investigated the variations with time. The Complete data were calculated from 1889 to 2004 (about 115 years observation). Since that data is relatively long to be shown (Table 1).

Tabel 1 The increasing of CO_2 emission over Indonesia since 1889 to 2004

Year	Total Fossil Fuel CO2 Emissions	CO2 Emissions from Gas Fuels	CO2 Emissions from Liquid Fuels
1889	1	0	0
1890	4	0	0
1891	6	0	0
1892	49	0	3
1893	110	0	62
1894	131	0	61
1895	210	0	116
-------	---------	---------	---------
2000	99728	22237	58348
2001	98331	15821	57194
2002	113285	21410	60159
2003	111345	22216	63969
2004	103170	17363	68378

Source: http://cdiac.ornl.gov/ftp/trends/emissions/ido.dat

Please note here, we applied the Box-Jenkins method with the following steps, namely: identification, assessment and testing before the application of the model itself.

Identification of Model

The first step that we need to do is we need to check if the data is stationery or no. If the data used are not stationary, we need do distinction get a stationary time series. A non stationer time series data can be transformed into stationary by transforming the values of the time series. If the time series does not have seasonal variation, the transformation into a stationary form is often used the first difference transformation of the values from time series. If the distinction first had to produce a stationary time series, it would require a more complex distinction anymore. In the identification model, the first thing to do is :

a. Make a plot of data (time plots) are useful to see whether the data visible stationary or not.
b. Checking autocorrelation plot of the function (ACF) and partial autocorrelation function (PACF) to see the model from data.

If ACF is significant at lag (lead time) q and PACF decreased exponentially, so the data can be modeled with a moving average model of degree q (MA (q)) and if it falls exponentially ACF and PACF lag is significant at p, then the data can be modeled by p degrees autoregressive model (AR (p)). If these two things are not obtained, there is the possibility of a joint process model is the AR and MA or ARMA (p, q).

So to determine the order of the AR process is to look at PACF. Another case with MA model to determine the order of the model used ACF. But both ACF and PACF of each model must be considered because it could have obtained the model was ARMA model. Therefore, to identify the time series model is better to use both the ACF and PACF. Here is the behavior of ACF and PACF for the model AR (p), MA (q), and ARMA (p, q):

Tabel 2 Identification model for time series data AR(p), MA(q), and ARMA (p,q)

	AR (p)	**MA (q)**	**ARMA (p,q)**
ACF	Eksponential decrease	*Cut – off* at lag to- q	Eksponential decrease with start lag to - p
PACF	*Cut – of f* pada lag ke – p	Eksponential decrease	Eksponensial decrease with start lag to - q

Suspect of Model Parameters

To help choose the type of tentative (temporary), using the results of the analysis and partial autocorrelation with a certain lag length. After the model the time series had been identified, the next step is to suspect the model parameters are based on least square criteria. There are two basic ways to obtain these parameters:

a. By way of experimentation (trial and error) that is testing several different values and selecting a value (or set of values, if there are more than one parameter to be estimated) that minimizes the sum of squares residual value / value of the error (sum of squared residuals).
b. Iterative improvement of selecting initial estimates and then let the computer programs are watched by iterative forecasting (Makridakis, 1999).

Validation Model

After the ARIMA model is determined, the next step is to conduct diagnostic tests to test the feasibility of the model and suggest improvements if necessary. One way that can be done is by analyzing the error (residual). In other words, examining the difference (difference) between observation data and model output. Error value (error) that remains after matching is ARIMA model, expected only a random disturbance. Therefore, if the plot function and autocorrelation partial of error values have been obtained, is expected to:

a. There was no significant autocorrelation.
b. There was no significant partial autocorrelation.

The second is to study the statistical sampling of the optimum solution to see whether the model can still be simplified. Statistical assumptions underlying the general model of ARIMA that gave some statistics that should be calculated after the values measured optimum coefficients. For example, for each coefficient / parameter values that are obtained will be calculated so that the error sum of squares error value. Coefficient value is selected that has the smallest squared error values. Error values can be obtained from (Makridakis, 1999).

Forecasting Model

The next step is to forecast (forecasting) if the model is suitable. The next step is to forecast (forecasting) if the model is suitable.

RESULTS AND DISCUSSIONS

Identification of Data

The data used to make this prediction model is data on the CO2 emissions of Indonesia since 1889 to 2004. In this study analysis, we applied the ARIMA (Autoregressive Integrated Moving Average), because it involves time series data, thus obtained a model that describes the time series data.

Stationery test needs to be done before the creation of models for forecasting in time series data requires that data must be stationary. The number of time series data distinction will become the order of d values in the model used ARIMA. A stationary data when said average value and variance are constant over time. Is not stationary data need to be modified (made the distinction) to generate stationary data. Here is a plot autocorrelation function (ACF), and partial autocorrelation function (PACF) as shown in Figure 1 and 2 below. We present also for the PACF and the first distinction at Figure 3 and 4, respectively.

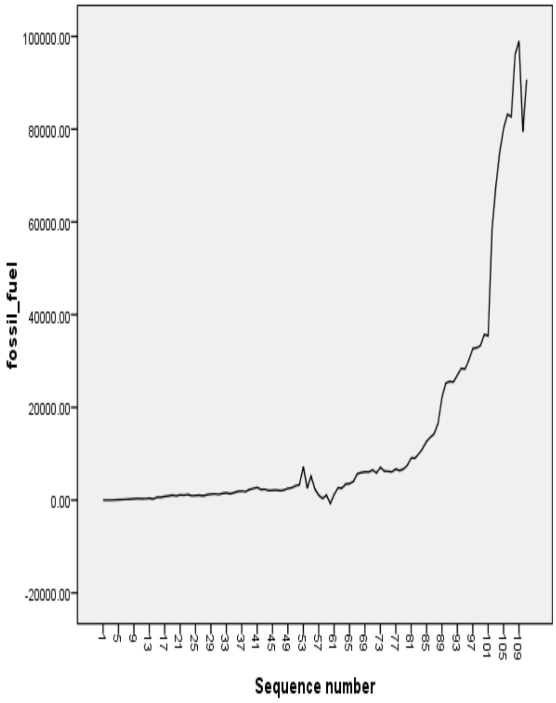

Figure 1 The time-series of the Total Fossil Fuel CO_2 Emissions (in 1000 metric tons of carbon) since 1889 to 2004

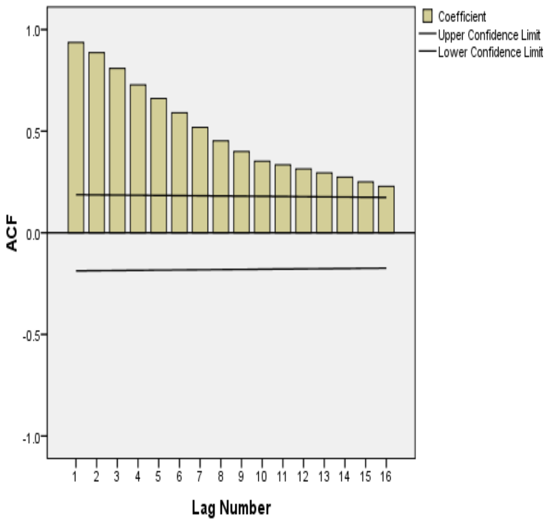

Figure 2 The Autocorrelation Function (ACF) of the Total Fossil Fuel CO_2 Emission since 1989 to 1999

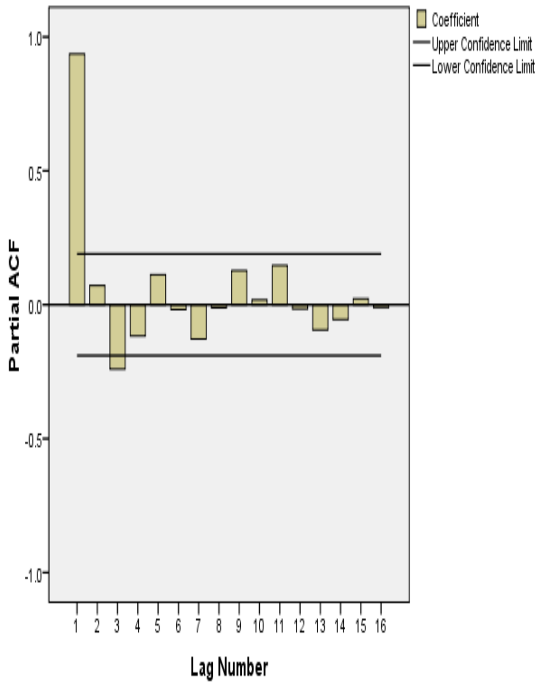

Figure 3 The time-series of the Total Fossil Fuel CO_2 Emissions (in 1000 metric tons of carbon) since 1889 to 2004, for Partial Autocorrelation Function (PACF)

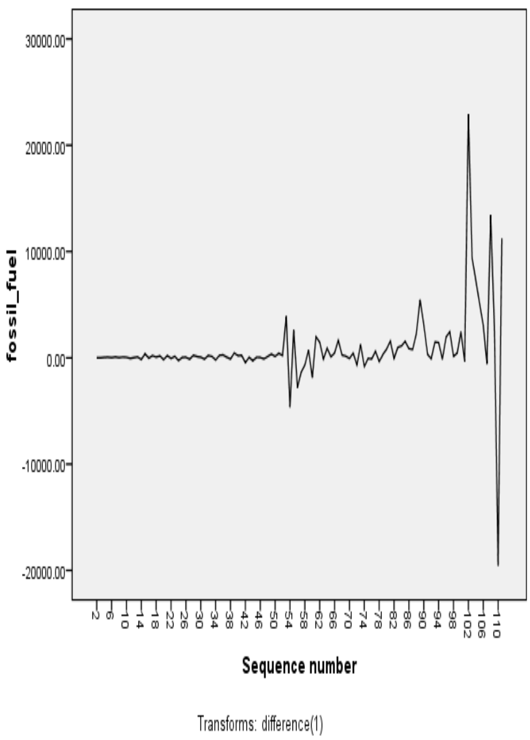

Figure 3 The time-series of the Total Fossil Fuel CO_2 Emissions (in 1000 metric tons of carbon) since 1889 to 2004but after we do the first distinction

The Estimated and Validation Model

Through the ACF and PACF plot of the original data is performed first distinction, while the model is determined CO2 emissions data period 1889 to 2004. From the ACF plot (Figure 3-2) and PACF (Figure 3-3) obtained information that the CO_2 emissions ACF lag signnifikan at 1,2,3,4,5. While CO_2 significant PACF at lag 1 and 2. Thus while the model of the data plot is a mixture of CO_2 emissions from autoregressive, the first distinction, and moving averages or ARIMA model (p, 1, q). With the p-value is 1 and 2 while the value of q selected 1, 2, 3, 4, and 5. Next is an estimate of the lag-lag is to get the best model. After establishing the identification of the model temporarily, then the parameters AR and MA should be established.

Table 3 *Mean Absolute Deviation* (MAD) for ARIMA model of the *Total Fossil Fuel* CO_2 *Emission for period of* 1989 to 1999

Model ARIMA	**MAD (*Mean Absolute Deviation*)**
(3,1,3)	4208.099106
(3,1,6)	2502.466386
(3,1,8)	3407.307572
(3,1,9)	22945.05449
(8,1,3)	2093.597265
(8,1,6)	14197.47792
(8,1,8)	3320.764213
(8,1,9)	2728.053095

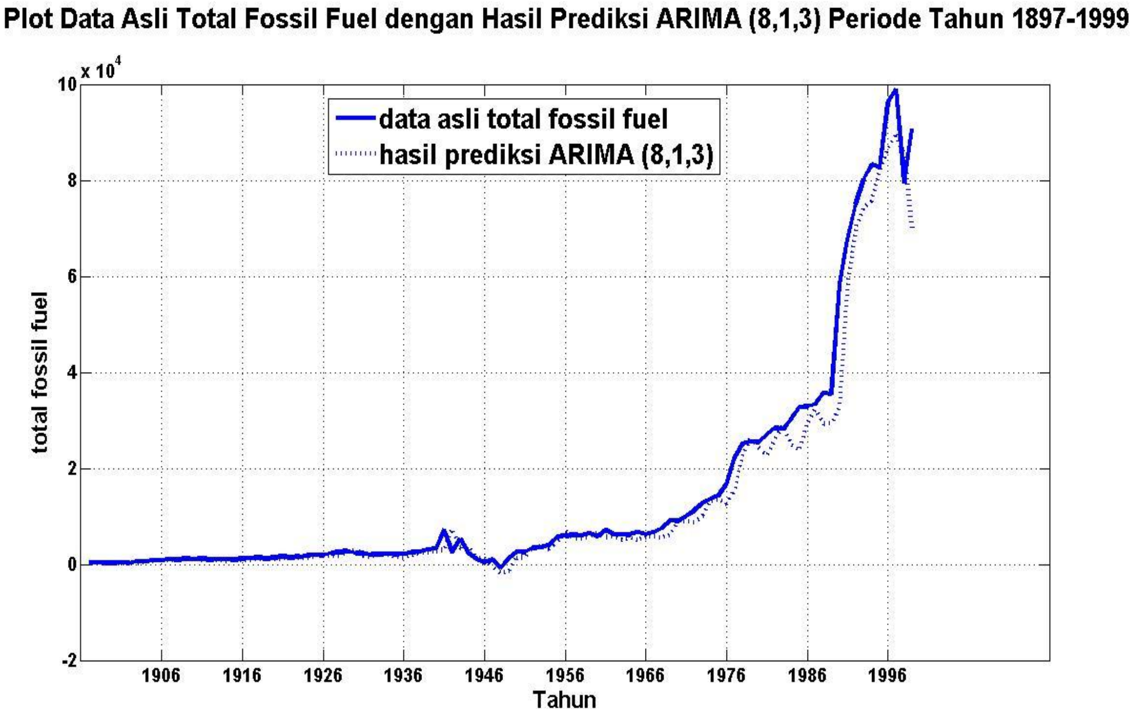

Figure 4 Total fossil fuel observed and predicted with ARIMA (8,1,3) 1897 - 1999

The Applied Model

Using Box-Jenkins method and the stage of identification, assessment, and testing, the best prediction model obtained for data Total Fossil Fuel CO2 Emissions, with the model prediction ARIMA (8,1,3)

0298 zt = zt-1 - 0, .423 ZT 0221-2 + 3 + ZT-0203 ZT-0059 ZT 4 + 5 + 0239-ZT-6-0368 ZT-7 - 0:55 + ZT-8-0779 at 1 + 0.7 at-2 +-0473 at 3 (harus dikoreksi)

where: ZT = predictive value on day t and at = error (the difference between the original values and results of prediction) on day-t.

Cross-Checking between Model and Observed Data

Tabel 4 Output model ARIMA and obsereved data

Year	Original Data	Predition ARIMA (8,1,3)	Galat/Error
2000	99728	73048.6105	26679.39
2001	98331	89378.3311	8952.669
2002	113285	96043.1392	17241.86
2003	111345	95640.649	15704.35
2004	103170	86672.1384	16497.86

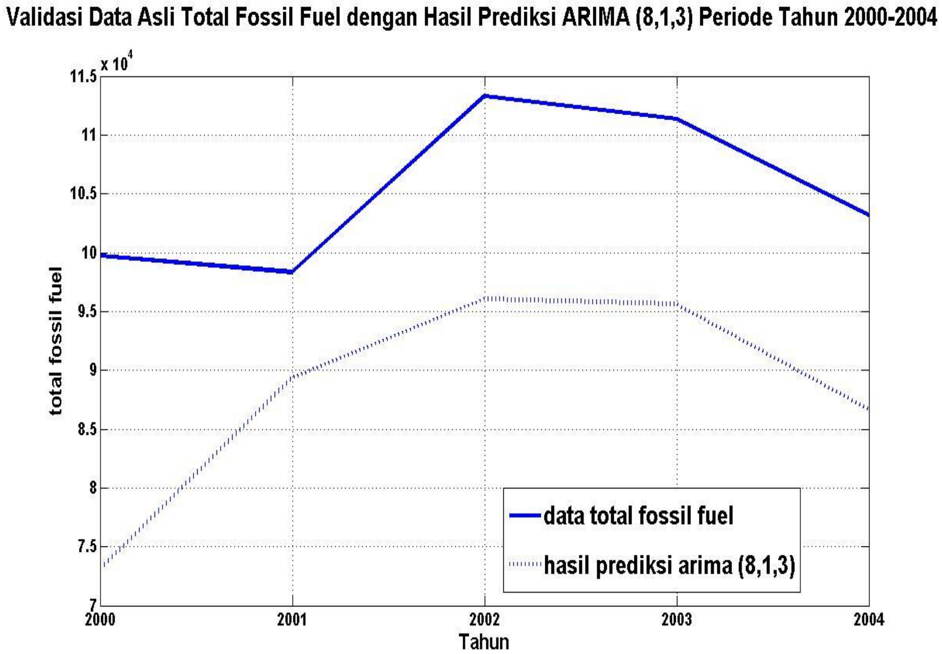

Figure 5 Total fossil fuel observed and predicted with ARIMA (8,1,3) period of 2000-2004

Tabel 5 Prediction Result's of the *Total Fossil Fuel CO_2 Emission* using ARIMA (8,1,3) model for period of 2005-2014

Year	The Total Fossil Fuel CO_2 Emissions (in thousand metric tons of carbon) based on the ARIMA (8,1,3) model prediction
2005	118467
2006	129816
2007	119756
2008	123018
2009	126634
2010	122146
2011	133614
2012	137314
2013	123583
2014	128041

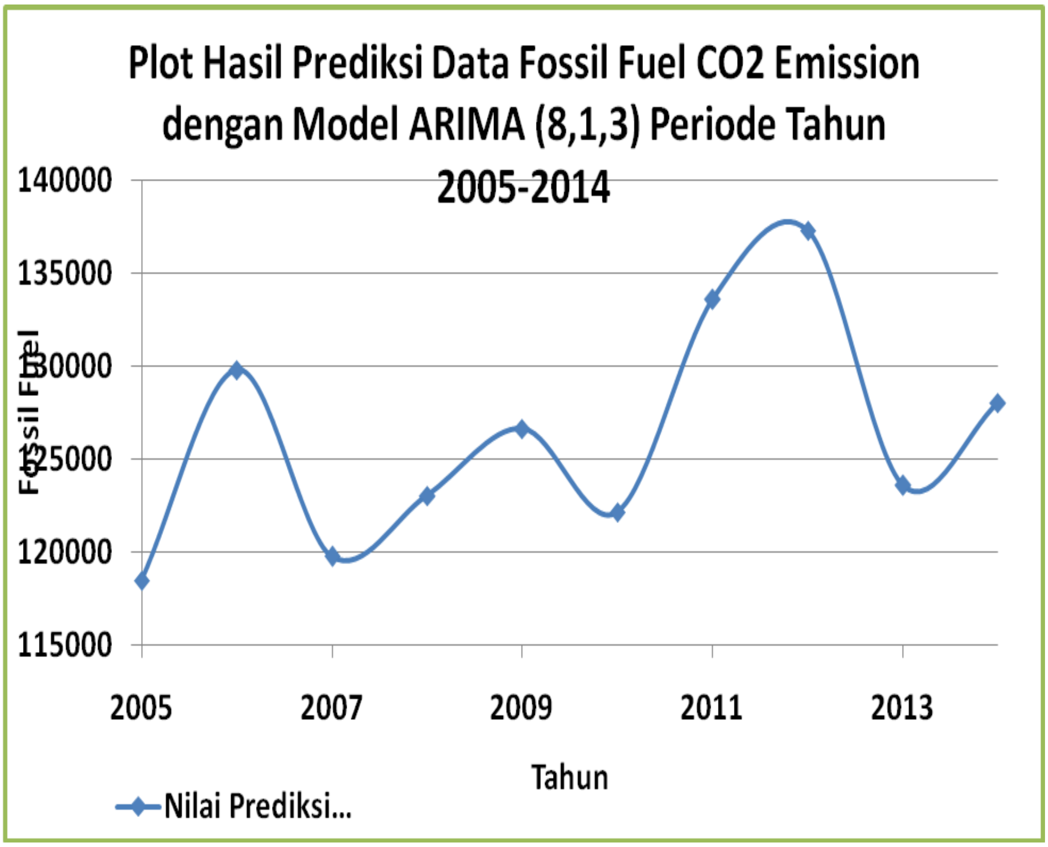

Figure 6 The total fossil fuel CO_2 emissions (in 1000 metric tons of carbon) based on the ARIMA (8,1,3) model prediction periods of 2005 - 2014

SUMMARY

Based on the above results it can be concluded that the best predictor model for the Total Fossil Fuel CO_2 Emissions over Indonesia is ARIMA (8,1,3). This means that the predicted value for the next year depending on the data before and 8 years 3 years earlier error. In the validation data with predicted results, the MAD (Mean Absolute Deviation) is relatively high. However, the pattern of results followed the pattern predicted almost the original data with a correlation value of 99%. Based on this result, we can estimate the climate projection over Indonesia, especially during 2012-2014.

REFERENCES

Ahlbeck J. 2000: *The Carbon Dioxide Thermometer*. Energy & Environment, 11, 3, 355-358.

Makridakis *et al*. 1999: *Metode dan Aplikasi Peramalan*. Jilid Satu Ed ke-2. Untung SA dan Abdul B, penerjemah. Jakarta: Erlangga. Terjemahan dari: Forecasting, 2nd Edition.

http://cdiac.ornl.gov/ftp/trends/emissions/ido.dat

sumber pustaka kurang