Scientific Digest
An International publication
Journal of Applied Engineering
ISSN: 2348-4802

Prediction of Carbon Dioxide Emissions Using Two Linear Regression-based Models: A Comparative Analysis

Chee San Choi[1], Lazim Abdullah[2]

[1]School of Informatics and Applied Mathematics, Universiti Malaysia Terengganu, Malaysia
[2]School of Informatics and Applied Mathematics, Universiti Malaysia Terengganu, Malaysia

Abstract: Carbon dioxide (CO2) emissions have been continuously escalating in recent years. The escalating trend is consistent with the current economic activities and other uncertain variables such as demand and supply in businesses and energy needs. Linear model is one of the most commonly used methods to explain the relationship between CO2 emissions and the related economic variables. The conventional linear regression model has a disadvantage in describing the relationships due to the variables' uncertainty and vague information. To address this problem, the fuzzy linear regression model has been proposed for explaining the relationships. However, the performance of the two linear models for predicting CO2 emissions is not immediately known. This paper presents a comparative study of conventional linear regression model and linear regression with fuzzy numbers model for predicting CO2 emissions in Malaysia. Twenty five years data from 1981 to 2005 of CO2 emissions, fuel mix, transportation, gross domestic product, and population have been used to develop the model of possibilistic fuzzy linear regression (PFLR) and multiple linear regression (MLR). The criteria of performance evaluation are calculated for estimating and comparing the performances of PFLR and MLR models. The performance comparison of PFLR and MLR models due to mean absolute percentage errors, root mean squared error criteria; indicate that MLR performed better on CO2 emissions prediction. A considerable further work needs to be done to determine the flexibility of fuzzy numbers in enhancing the performance of PFLR against the MLR.

Keywords:Fuzzy linear regression, multiple linear regression, Error analysis, Predictive model, CO2 emissions

I. INTRODUCTION

Carbon dioxide (CO2) emission is one of the important components in conserving the stability of climate system, and plays a key role in greenhouse effect. In the past few decades, CO2 emissions have been increasing exponentially. It has been reported that almost 30 billion tons of CO2 enters the atmosphere as a result various human activities each year [1]. The effect of the higher concentrations of CO2 to people could not be taken lightly. CO2 is held responsible for 58.8% of greenhouse effect. The effect may cause major environment pollutions and climate instability. The increase in CO2 emission would give disastrous environmental consequences such as droughts, storms, floods and other environmental calamities. As a result of large volume of CO2 in atmosphere, it has been reported that global sea level has increased by 10-20 cm during the twentieth century [2]. Not only the increment in sea level, but temperature of sea also reported in risen trends. Spence [3] reported that global CO2 emissions have increased by 30% and temperature has risen by 0.3-0.6 degree Celsius. These are among the many feared examples of environmental instability as a result of uncontrolled CO2 emissions. The risk of environmental catastrophe is therefore considered great enough to justify the importance of research in CO2 emissions.

A considerable amount of literature about CO2 emissions has been published especially on causal relationships between CO2 emissions and its contributing factors. Hwang and Yoo [4], for example, analyzed the short and long-run causality issues between energy consumption, CO2 emissions, and economic growth in Indonesia using time-series techniques. Chang [5], and Zhang and Cheng [6] investigated the relationships between CO2 emissions with Gross Domestic Product (GDP) and energy consumption in China. The similar studies were also conducted in South Africa [7], and Turkey [8]. However, the causal relationships among variables could also be extended to prediction studies. The relationships between CO2 emissions and its related variables have been widely investigated in projections, and predictions based research. Nor Sharfiza et al. [9], for example, projected the CO2 emissions in Malaysia. They used the model of long range energy alternative planning system to make the CO2 emissions projection until the year 2020. Electricity generation, transportation, industrial and residential were used as the causal factors. However, this is a projection research where the accuracy or errors of the model were rarely measured. Apart from projection research, there were many other related prediction researches in CO2 emissions have been undertaken. Most of the prediction models, not only provide the predicted values, but performance measures as well. The methods based on artificial intelligence, traditional linear regression, computer based simulation, optimal growth model are among the popular approaches [10]. Prediction of CO2 emissions has become an important research as it would provide clues and awareness to achieve environmental stability. Emissions of gaseous elements such as CO2 become a worldwide concern as the greenhouse gas proved to contribute most impact on environmental problems. Nevertheless, choosing the right methods in predicting CO2 emissions depend on a wide range of factors which involved both qualitative and quantitative variables.

II. Related research

One of the popular approaches in predicting CO2 is conventional linear models. Bronfield et al. [11], for example, applied a linear model of the linear regression downscaling to model on-road CO2 emissions in Boston, Massachusetts and tested the approach with surface-level CO2 observations. The accuracy of the estimated model was measured using coefficient of determination (R2). Murad et al. [12] used ordinary least squares methods to estimate parameters in three linear regression models. The findings of the research revealed three important observations for Malaysia: the link between agricultural growth rate and climate change score was proven to be negative, but insignificant; the link between per capita CO2 emissions and agricultural production index was found to be direct and highly significant; and the link between per capita agricultural production index and per capita CO2 emissions was proven to be positive and highly significant. Also, an increasing level of per capita CO2 emissions in the country was proven to have both detrimental and beneficial effects on its agricultural growth. Kone and Buke [13] employed regression analyses to predict energy-related CO2 emissions. Trend analysis, which depends on linear approach was also used for its modelling. Trends in CO2 emissions for the top-25 countries and the world total CO2 emissions during 1971 to 2007 were identified. These data were regressed against the year using a least squares technique. The CO2 emissions for eleven countries and world total CO2 emissions correlation versus the year was obtained from modelling with the fit coefficients and correlation coefficient values for each fitting. The results obtained from the analyses showed that the models in those countries can be used for CO2 emission projections into the future planning.

Besides linear models, nonlinear models were also being carried out for predicting CO2. Ming and Niu [14], for example, used nonlinear logistic regression to improve the goodness of fit of the CO2 emissions prediction model. Mean absolute percentage errors (MAPE) were measured to check the efficiency of estimated values. The empirical analysis in China shows that the logistic regression method was better than the linear model in terms of goodness of fit and simulation risk. This is one of the examples of research in CO2 emissions that have been conducted using linear and nonlinear logistic models. Nevertheless, with the advent of intelligence based research and the uncertainty in the causes and consequences of CO2 emissions, most of the recent predictive models tend to rely on nonlinear intelligence models. Zhou et al. [15], for example, studied the relationships among the significant parameters impacting CO2 production. The adaptive neural network fuzzy inference system technique was trained with historical data and generated the membership functions and rules which best interpret the input-output relationships in the process. The model validation process showed that modeling accuracies of these fuzzy inference systems are within acceptable limits. Martinez-Lopez [16] applied a fuzzy controller in order to obtain a CO2 emissions path which leads to a temperature increase of 2 degrees when it is used to drive a simple, linear climate model. In another intelligence-based research, Azadeh et al. [17] introduced an integrated fuzzy regression and data envelopment analysis algorithm for oil consumption estimation and optimization with uncertain and ambiguous data. Three types of relative errors, root mean squared error (RMSE), mean absolute error and MAPE were considered to validate the applicability and superiority of the proposed algorithm.

Apart from linear models, intelligence-based models and integrated models were also carried out in comparative analyses of CO2 emissions and its variables. In a comparison study, Ionescu and Candau [18] performed multiple linear regression (MLR) to predict CO2 and NO2 released in the process of reheating furnace in the iron and steel industry. Furthermore, they also built artificial neural network (ANN) for the same purposes. RMSE values were calculated to find the accuracy of the modelling. It appeared that CO2 can be satisfactorily estimated by a linear regression. Meanwhile the NO2 appeared to have problem with this model. Hence, NO2 emission modelling required a non-linear model. It is found that the ANN modelling can be considered as reliable method for NO2 and CO2. Moreover, in the case of CO2, a simple linear model gave less efficient results than the ANN (5.6%), but was still comparable to the measurement error. Another comparison study was carried out by Pao et al. [19]. They presented an improved grey model called nonlinear grey Bernoulli model (NGBM) in their research. The model was developed to predict three indicators which were: carbon emissions, energy consumption and real output in China. This study collected annual data on energy consumption, CO2 emissions, energy intensity, carbon intensity and the real GDP. It considered the data for the period between 1980 and 2009. The performance between NGBM and ARIMA were compared. For the purpose of evaluating the out-of-sample prediction capability, the prediction accuracy was examined by calculating three different error measures: the RMSE, the mean absolute error (MAE) and the MAPE. The predicting ability of NGBM with optimal parameter model, namely NGBM-OP has remarkably improved because it obtained robust results in terms of MAPE, RMSE and MAE.

In another comparative study, Pau and Tsai [20] applied GM on three variables of CO2 emissions, energy consumption and real GDP to investigate the dynamic relationships between the variables. Data of Brazil for the period between 1980 and 2007 were used in this investigation. The finding of the inverted U-shaped relationships of both emissions-income and energy consumption-income imply that both environmental damage and energy consumption firstly increase with income, and then stabilize. ARIMA model was also built in order to compare the forecasting ability of GM model. Both of the models have shown strong forecasting performance with MAPEs of less than 3%. MohdPauzi and Abdullah [21] compared the performance of fuzzy inference system (FIS) and adaptive neuro fuzzy inference system (ANFIS) models in predicting CO2 emissions. The inputs to the models were simulated using the Malaysian data for the period from 1980 to 2009. The prediction performances were measured using RMSE, MAE and MAPE. The performance of the two models against the CO2 emission clearly shows that the ANFIS outperforms the FIS model.
So far, however, little research has been carried out in the area of comparison between linear model with fuzzy parameters and conventional linear models specifically in CO2 emissions. Considering the issue of prediction performance in particular, the literature is almost absent on the detailed comparison between the two linear models with different nature of coefficients. In spite of the widespread use of multiple linear regression methods in many day life activities, there exists uncertainty in variables used. As to handle the uncertainty, fuzzy linear regression was introduced. Conventional linear regression cannot handle visual inspection results that are inherently non-crisp or linguistic. On the other hand, fuzzy linear regression provides an effective means for coping with such fuzzy data or linguistic data. One of the earliest fuzzy linear regression is known as possibilistic regression since fuzzy data can be regarded as distribution of possibility [22]. Against all this background, the aim of this paper is to compare the prediction performance of possibilistic fuzzy linear regression (PFLR) against multiple linear regression (MLR) in CO2 emissions. The CO2 emissions data of Malaysia together with the variables of fuel mix, transport, GDP and population were employed in this comparative study. Differently from the typical linear model which directly used crisp numbers, this paper introduces fuzzy number based on multiple linear regression. The rest of this paper is structured as follows. Section III elucidates the brief theoretical review of PFLR and MLR. Section IV presents a comparison of the CO2 emissions predictive models between PFLR and MLR using Malaysian data. This paper is finally concluded in Section V.

III. SHORT REVIEW OF MLR AND PFLR

This section elucidates some basics theoretical background of MLR and PFLR.

A. Multiple Linear Regression

Linear regression was the first type of regression analysis to be studied rigorously, and to be used extensively in practical applications. This is because models which depend linearly on their unknown parameters are easier to fit than models which are non-linearly related to their parameters and because the statistical properties of the resulting estimators are easier to determine. The function form which is omost frequently used for expressing the relationship is the linear form:

$$Y' = a + bX \tag{1}$$

where:

Y' is the predicted value of the Y variable for a selected X value.

a is the Y -intercept. It is the estimated value of Y when X =0. In other words, a is the estimated value of Y where the regression line crosses the Y -axis.

b is the slope of the line, or the average change in Y' for each change of one unit (either increase or decrease) in the independent variable X .

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

$$Y' = a + b_1X_1 + b_2X_2 + b_3X_3 + \ldots + b_kX_k \tag{2}$$

where a is the intercept, the value of Y when all the X's are zero and b_k is the amount by which Y changes when that particular X_k increases by one unit with all other values held the same. The subscript j can assume values between 1 and k, which is the number of independent variables.

B. Possibilistic Fuzzy Linear Regression

In this sub-section, formulations for fuzzy linear regression estimation are presented. The inputs and outputs of the model are non-fuzzy observations. The base model is assumed to be a fuzzy linear function as below:

$$\tilde{y} = f(x,\tilde{A}) = \tilde{A}_0 + \tilde{A}_1 x_1 + \tilde{A}_2 x_2 + \ldots + \tilde{A}_n x_n \quad (3)$$

where A_i(i= 1,2,…,n) are the fuzzy coefficients in the form of (p_i, c_i, c_i). The real number p_i is the middle and c_i is the spread. The spread value denotes the fuzziness of the function.

The membership functions for each type of A_i are assumed a triangular membership. So it can be expressed by this equation,

$$\tilde{A}_i(a) = \begin{cases} 1 - \dfrac{|a - p_i|}{c_i}, & p_i - c_i \le a \le p_i + c_i \\ 0, & otherwise \end{cases} \quad (4)$$

Equation (3) can be written as:

$$\tilde{y} = (p_0, c_0) + (p_1, c_1)x_1 + (p_2, c_2)x_2 + \ldots + (p_n, c_n)x_n \quad (5)$$

By applying the Extension Principle [23], it implies that the membership function of fuzzy number $\tilde{y}$ is given by:

$$\tilde{y}(y) = \begin{cases} max(min_i\{\tilde{A}_i(a_i)\}, \{a_i | y = f(x,a_i) \neq \varphi\} \\ 0, \quad otherwise \end{cases} \quad (6)$$

From equation (5) and (6), membership function of $\tilde{y}$ is

$$\tilde{y}(y) = \begin{cases} 1 - \dfrac{\left|y - \sum_{i=1}^{n} p_i x_i\right|}{\sum_{i=1}^{n} c_i |x_i|}, & x_i \neq 0 \\ 1, & x_i = 0, y_i = 0 \\ 0, & x_i = 0, y_i \neq 0 \end{cases} \quad (7)$$

The spread of $\tilde{y}$ is $\sum_{i=1}^{n} c_i |x_i|$ and the middle of $\tilde{y}$ is $\sum_{i=1}^{n} p_i x_i$.

Equation (5) can be re-written as:

$$\tilde{y}_j = (p_0, c_0) + (p_1, c_1)x_{1j} + (p_2, x_2)x_{2j} + \ldots + (p_n, c_n)x_{nj}$$
$$j = 1,2,3,\ldots,m \quad (8)$$

where m is the number of observation.

We try to find the coefficient $\tilde{A}_i = (p_i, c_i)$ that minimize the spread of the fuzzy output for all data sets. From Montgomery and Peck [24], the objective function is given as

$$Min \sum_{j=1}^{m} \sum_{i=1}^{n} (c_0 + \sum_{i=1}^{n} c_i |x_{ij}|) \quad (9)$$

The constraints require that each observation y_j has at least h degree of belonging to $\tilde{y}(y)$ [25]. So,

$$\tilde{y}(y_j) \ge h, \quad j = 1,2,\ldots,m \quad (10)$$

The degree h is specified by the user.

By substituting equation (7) into equation (10), we obtain:

$$y_j \ge p_0 + \sum_{i=1}^{n} p_i x_{ij} - (1-h)(c_0 + \sum_{i=1}^{n} c_i |x_{ij}|), \quad j = 1,2,\ldots,m$$
$$y_j \le p_0 + \sum_{i=1}^{n} p_i x_{ij} + (1-h)(c_0 + \sum_{i=1}^{n} c_i |x_{ij}|), \quad j = 1,2,\ldots,m \quad (11)$$

The aforementioned analysis leads to the following linear programming problem [24].

$$Min \sum_{j=1}^{m} \sum_{i=1}^{n} (c_0 + \sum_{i=1}^{n} c_i |x_{ij}|)$$

s.t.

(12)

$$y_j \geq p_0 + \sum_{i=1}^{n} p_i x_{ij} - (1-h)(c_0 + \sum_{i=1}^{n} c_i |x_{ij}|, \qquad j = 1,2,\dots m$$

$$y_j \leq p_0 + \sum_{i=1}^{n} p_i x_{ij} + (1-h)(c_0 + \sum_{i=1}^{n} c_i |x_{ij}|, \qquad j = 1,2,\dots m$$

$$c_i \geq 0, \qquad p_i \geq 0$$

Solution of the linear programming problem would provide the fuzzy coefficients of possibilistic fuzzy linear regression.

IV. IMPLEMENTATION

In general, one may presume that the performances of the two predictive models in CO2 emissions differ due to different theoretical backgrounds. As to confirm this hypothesis, the annual data of CO2 emissions and its associated variables from Malaysia were employed to be tested using PFLR and MLR. Input variables or predictors of Malaysia data are fuel mix, transportation, GDP, and population. These variables are labeledas x_1, x_2, x_3, and x_4. CO2 emissions is the response variable and denoted as $\tilde{y}$. Summary of the variables are shown in Table 1.

TABLE 1
Input and response variables for the models

Input variables	
x_1	Fuel mix
x_2	Transportation
x_3	Gross Domestic Product
x_4	Population
Response variable	
$\tilde{y}$	CO2 emissions

The prediction abilities of the PFLR model are compared with MLR model using the actual data of CO2 emissions in Malaysia over the period between 1981 to 2005. This sample data is used to build the models and also used to evaluate the prediction accuracy using RMSE and MAPE.

A. Possibilistic Fuzzy Linear Regression Model

The optimization software LINGO successfully yielded the fuzzy coefficients of the PFLR model. Spreads and centre values of fuzzy coefficients for Malaysia are shown in Table 2.

TABLE 2
Fuzzy coefficients of the PFLR

	p	*c*
0	163.9972	0
1	0	0.3322688
2	0.2284970	0.8459350
3	0.9931189	0.2905979
4	0.1645203	0.0859338

The PFLR model for CO2 emissions data can be written as

$$\tilde{y} = (163.9972,0) + (0,0.3322688)x_1 + (0.2284970,0.8459350)x_2 + (0.9931189,0.2905979)x_3 + (0.1645203,0.0859338)x_4$$

The FLR model identifies GDP as the most effective predictor of CO2 emissions in Malaysia [26].

B. Multiple Linear Regression Model

The same variables were used to model the CO2 emissions using multiple linear regression. The multiple linear regression model for the estimation of the CO2 emissions were examined. Table 3 summarizes the results for the MLR model.

TABLE 3
Linear regression analysis outputs

Regression statistics					
R^2					0.691
Adjusted R^2					0.430
SE					919.8784
	Df	SS	MS	F	Significance F

Analysis of variance (ANOVA)						
Regression		2	16986808	8493404.023	10.037	0.01
Residual		22	18615878	846176.260		
Total		24	35602686			
	Coefficient	SE	t Statistic	P value	Lower 95%	Upper 95%
Intercept	249.527	372.788	0.699	0.510	-523.588	1022.642
Transportation	0.792	0.210	3.767	0.001	0.536	1.227
GDP	0.685	0.314	2.183	0.040	0.034	1.337

The coefficient of determination, R^2 and the adjusted R^2 are 0.691 and 0.477 respectively, which indicates that about 69.1% of the variation in the CO2 emissions is explained by transport and GDP. The analysis of variance indicates that the *p*-value (probability of rejecting the null hypothesis) for the F test statistic is 0.01, which provide strong evidence against the null hypothesis. The t-test statistic shows that the *p*-value for the model intercept and the coefficient associated with the rejection region is less than 0.05, which again provide statistical evidence of rejecting the null hypothesis. The significant level associated with transportation and GDP variables is more than 0.05. These results indicate that least one of the predictors are useful for predicting CO2.

From the output, the multiple regression equation can be written as

$$\hat{Y} = 249.527 + 0.792x_2 + 0.685x_3$$

The regression equation shows that x_2 is the best predictor. Hence, the variable of transport is the most effective variable in CO2 emissions followed by GDP.

C. Comparative Analysis

The results of predicted values from the two models and the actual values were also examined. Figure 1 shows the results produced by the two models against the actual values.

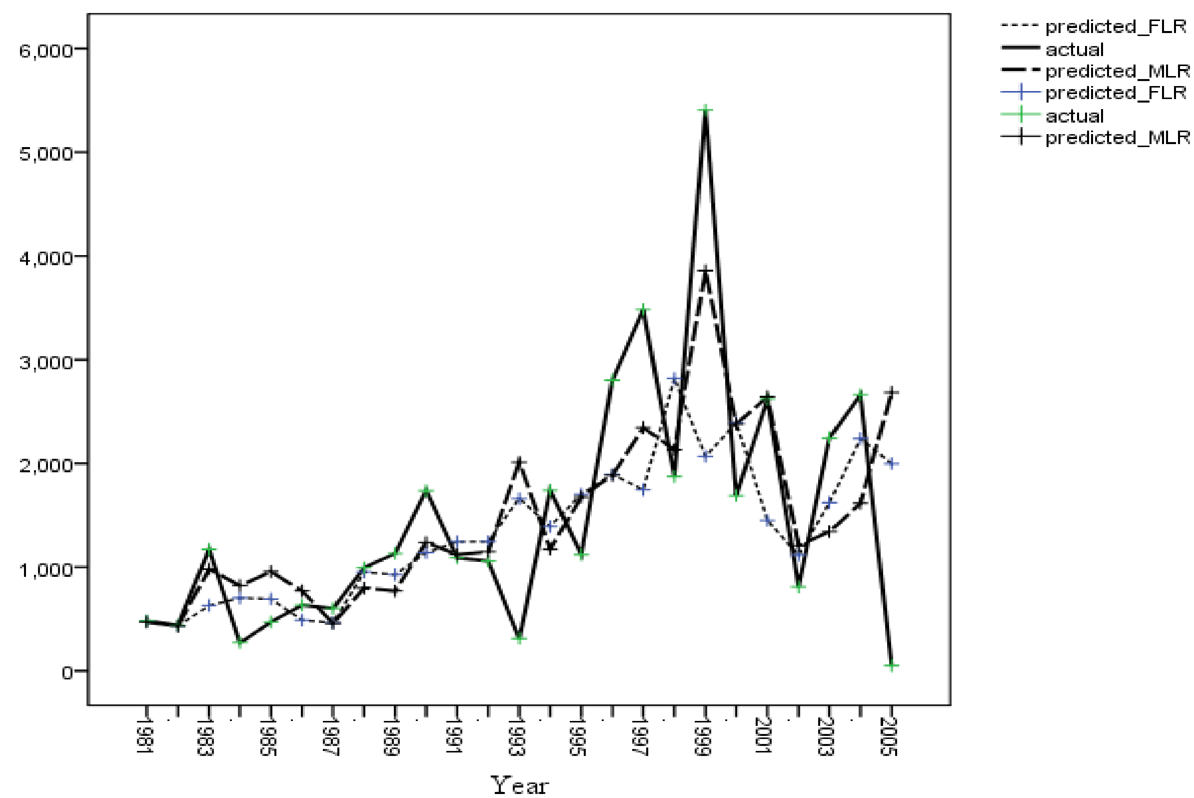

Fig. 1 Actual values against the predicted values using MLR and PFLR

As to check the performance of the PFLR and the MLR models, MAPE and RMSE are calculated using the equations (13) and equations (14).

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{A_t - F_t}{A_t}\right| \tag{13}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(A_t - F_t)^2} \tag{14}$$

where A_t is the actual value and F_t is the predicted value. The real number *n* represents number of data used in the analysis.

The performance comparison can be shown in Table 4.

TABLE 4

Comparison or errors between MLR and PFLR models

Model	MAPE	RMSE
Forecasts of CO2 emissions (million metric tons)		
MLR	2.010492	862.964
PFLR	2.028522	1010.117

The errors show that MLR model has smaller errors compared to PFLR model. It indicates that the MLR is the better model for CO2 emissions prediction. It is good to mention that the PFLR model with degree of belonging *h*= 0.5 was used in this study. Perhaps, theoretical background of fuzzy linear regression can explain the results. Theoretically, when the value of *h* is one, the PFLR coefficient has zero width of fuzzy number (see Eq 11). Consequently, this particular PFLR coefficient is no longer a fuzzy number and the fuzzy coefficient is now crisp number and equivalent to the multiple linear regression coefficients. Therefore, the efficiency of PLFR model is much depending on degree of belonging, *h*

V. CONCLUSION

Conventional multiple linear regression is one of the commonly used models in prediction and relationship analysis. The fuzzy linear regression model is another linear model with the addition of fuzzy numbers that normally used to deal with uncertainty and vagueness of data. However, the performance of these two models specifically for the CO2 emissions case has not immediately known. This paper has provided a comparative study of multiple linear regression and possibilistic linear regression models for the case of CO2 emissions prediction in Malaysia. Twenty-year data of four predictors and CO2 emissions were employed. The performance of multiple linear regression and possibilistic linear regression was compared using mean average percentage errors and root mean squared errors. The results indicate that the multiple linear regression model performed better than possibilistic fuzzy linear regression model. It should be noted that the results may no longer favour to multiple linear regression if different degree of belonging of fuzzy linear regression were considered. This hypothetical statement could be left for further research.

REFERENCES

[1] C. Goodall, "How to live a low carbon life:the individual's guide to stopping climate change" (1st Edition), United Kingdom: Earthscan, 2007.

[2] H. Mukhtar, P.N.F.MKamaruddin and V.R. Radhakishnan, "Carbon credit trading for CO2 reduction: Opportunities for Malaysia", Technology cluster: Oil and gas, (Volume 4, Number 2) Available at: http://www.utp.edu.my/publications/platform/Platform%20v4n2.pdf. 2004.

[3] C. Spence, "Global warming: Personal solutions for a healthy planet" (1st Edition), New York: Palgrave Macmillan. 2005.

[4] J.H.Hwang, and S.H. Yoo. "Energy consumption, CO2 emissions, and economic growth: evidence from Indonesia", Qualitative and Quantitative, DOI 10.1007/s11135-012-9749-5, 2012.

[5] C.C. Chang, "A multivariate causality test of carbon dioxide emissions, energy consumption and economic growth in China", Applied Energy, 87, pp. 3533–3537, 2010.

[6] X.P. Zhang and X.M. Cheng, "Energy consumption, carbon emissions and economic growth in China", Ecology Economics, 68, pp. 2706–2712, 2009.

[7] K.Menyah, and Y. Wolde-Rufael, "Energy consumption, pollutant emissions and economic growth in South Africa", Energy Economic, 32, pp. 1374–1382, 2010.

[8] A. Ozturk, and A. Acaravci. "CO2 emissions, energy consumption and economic growth in Turkey", Renewable and Sustainable Energy Reviews. 14(9), pp. 3220–3225, 2010.

[9] Nor Sharliza, ZaimnuraZainon Noor, U. Zaini, and T. Juhaizah, "Projection of CO2 emission in Malaysia", Wiley Online Library. 2010.

[10] L.Abdullah, andH.MohdPouzi, "Methods in forecasting carbon dioxide emissions: a decade review", JurnalTeknologi, 75(1), pp. 67-82, 2015.

[11] M.N.Brondfield, L.R. Hutyra, C.K., Gately, S.M., Raciti, and S.A. Peterson. "Modeling and validation of on-road CO2 emissions inventories at the urban regional scale", Environmental Pollution, 70, pp.123-133, 2012.

[12] Murad, M.W. Molla, R.I. Mokhtar, and Raquib, M.A. 2010. "Climate change and agricultural growth: an examination of the

link in Malaysia", International Journal of Climate Change Strategies and Management, 2(4), pp.403 – 417. .

[13] A.C.Köne, and T. Büke, "Forecasting of CO2 emissions from fuel combustion using trend analysis", Renewable and Sustainable Energy Reviews, 14(9), pp.2906-2915, 2010.

[14] M. Ming, and D. Niu, "Modeling CO2 emissions from fossil fuel combustion using the logistic equation", Energy, 36(5), pp. 3355-3359, 2011.

[15] Q. Zhou, C.W. Chan and P. Tontiwachwuthikul, "An application of neuro-fuzzy technology for analysis of the CO2 capture process", Fuzzy Sets and Systems, 161(19), pp. 2597-2611, 2010.

[16] B. Martinez-Lopez, "Fuzzy control of CO2 emissions". Eighth International Conference on Fuzzy Systems and Knowledge Discovery (FSKD), 1, pp. 626-629, 2011.

[17] A. Azadeh, O.Seraj, S.M Asadzadeh, and M. Saberi, An integrated fuzzy regression-data envelopment analysis algorithm for optimum oil consumption estimation with ambiguous data. Applied Soft Computing, 12 (8), pp. 2614-2630, 2012.

[18] A. Ionescu and Y. Candau, "Air Pollutant Emissions Prediction by Process Modelling – Application in the Iron and Steel Industry in the Case of a Re-heating Furnace", Environmental Modelling & Software, 22(9), pp. 1362-1371, 2007.

[19] H.T. Pao, H.C. Fu and C.L. Tseng, "Forecasting of CO2 emissions, energy consumption and economic growth in China using an improved grey model", Energy, 40(1), pp.400-409, 2012.

[20] H. Pao and C.M. Tsai, "Modeling and forecasting the CO2 emissions, energy consumption, and economic growth in Brazil", Energy, 36(5). pp. 2450-2458, 2011.

[21] H. MohdFauzi, and Abdullah, L. "A comparative analysis of two fuzzy based models in predicting CO2", Proceedings of the first international conference on advanced data and information engineering", 285, pp. 203-211, 2014.

[22] H.Tanaka, Uejima, K.SandAsai, "Fuzzy linear regression model". IEEE Transactions on System, Man and Cybernetics, 10, 1 pp. 2933-2938, 1982.

[23] LA. Zadeh, "The concept of a linguistic variable and its application to approximate reasoning-I", Information Sciences, 8, pp. 199-249, 1975.

[24] D. Montgomery, and E.A. Peck, "Introduction to Linear Regression Analysis", Wiley, New York. 1982.

[25] D. T . Redden, and W. H. Woodall, "Further examination of fuzzy linear regression", Fuzzy Sets and Systems, 79, pp. 203-211, 1996.

[26] L. Abdullah and N. Khalid , " Prediction of CO2 emissions using fuzzy linear regression model: A case of developed and developing countries," Journal of Sustainability Science and Management, 9(1), 69-77, 2014.