# REGRESSION ANALYSIS TO FORECAST MALAYSIA'S IMPORTS OF CRUDE MATERIAL

**[1]MOHAMED A.HMILAD, [2]ROSE IRNAWATY IBRAHIM, [3]SAMIAPPANMARAPPAN**

[1,2,3]Universiti Sains Islam Malaysia (USIM) ,Mathematical Finance, Faculty of Science & Technology, Bandar BaruNilai, 71800 Nilai, Negeri Sembilan, Malaysia.
E-mail: agga2425@yahoo.com, rose_irnawaty@yahoo.co.uk

**Abstract-** In statistics, there are six kinds of linear regression analyses which can be done are namely :simple linear regression, multiple linear regression, ordinal regression, logistic regression, decriminant analysis and multinominal regression [1]. The purpose of conducting the linear regression analysis is to analyze and model relationships that exist between a dependent variable and one or more independent variables using various techniques. The present research used a stepwise multiple regression which is a combination of forward selection and backward elimination method. This stepwise regression method was utilised to select suitable controlled variables which were used to forecast the Malaysia's imports of crude material. The findings obtained from the stepwise regression method revealed that only four controlled variables namely;Predict price index of (CM), Gross Domestic Products (GDP), the value of exports and the average of sales tax of (CM) were selected. This proved that only these variables influenced the value of imports.

**Keywords-** Forecast; Malaysia 'Imports; Regression Analysis;. Stepwise Multiple Regression

## I. INTRODUCTION

Worldwide, the imports have occupy a big place. It has been proven that there is a relationship between imports and economic growth of any country. Previous research has proved an existing relationship between the imports and the economic growth of any country. Results obtained by previous researchers (Liu. 2001, Chen 2009, Azgun & Servinc 2010, Alam et al. 2009, Wong 2004 ,and Kogid, et al. 2011 showed a similar conclusion of the effect of imports on the economic growth in the Malaysian context, and therefore, suggested that future studies should further investigate the effect of such imports. Based on such useful recommendations, the present study reported in this paper aimed to develop a statistical model on prediction of the value of imports in the Malaysian context. A decision support system that could forecast the Malaysia' imports through computing the forecasting values for production of imports in Malaysia for the period from the first quarter 1991 to the fourth quarter 2013 was developed for the purpose of assisting Malaysia to draw imports policies in the future. In achieving this aim of the study, a regression analysis was used that could draw valuable interpretation of the data collected. In other words, such analysis is valued for its objectivity and systematic nature especially when it is used for the purpose of modelling and forecasting discrete time series. Based on the six identified types of linear regression analysis reported by Ghani & Ahamad (2010)[1], the current study employed a multiple linear regression to developing a model of Malaysia's imports. It also used some other methods categorized under the stepwise type or stepwise regression. This was to choose the most relevant controlled variables that could explain and predict the Malaysia's imports. According to the report of the Department of Statistic in Malaysia, classifying the imports is usually different based on their objectives. Therefore, the current study used the (SITC) classification created by the United Nations and which divides imports of commodities into 10 parts as in Figure (1) below. However, the focus of investigation in the present study is exclusive to the imports of Crude Material.

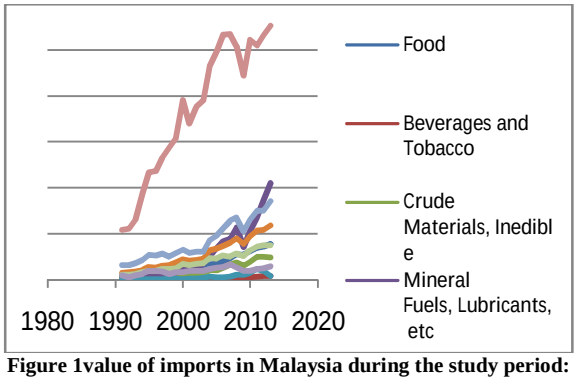



**Figure 1value of imports in Malaysia during the study period: 1991 to 2013.**

## II. NOTATION AND ABBREVIATION

The following notation and abbreviation are used.
CM: Crude Material, SITC: Standard International Trade Classification, D.W: Durbin Watson, VIF: Variance inflation factor, $R^2$: coefficient of determination, PPI: Producer Price Index, GDP: Gross Domestic Products, MLR: Multile linear regression.

## III. LITERATURE REVIEW:

There are several previous studies which used such type of regression analysis for a similar purpose. For instance, Abu Bakar (2000) [2] in analyzing the long-run relationship of Malaysia's import demand function used a multiple linear regression concept including the dynamic OLS method and the Johansen

Maximum Likelihood. The researcher used two variables: real GDP and price and the function form was log-log. His results obtained from the most preferred regression equation revealed that the R 2 = 0.48, S.E of Regression = 0.18, DW= 2.14, and F-stat = 4.25. Moreover, Tang & Nair (2002) [3], in empirically analyzing the aggregated import demand behaviour for Malaysia, used a regression method particularly for analyzing the long-run elasticises of import demand in relation to the income and relative prices. The regression formula was a log-log regression, and based on the results gained from the most preferred regression equation, it was found the followings: the R –Squared = 0.782, DW= 2.407, F-stat = 4,25 ,and Sum Squared Error = 0.116. Another study by Arshad et al. (2011) [4] applied a multiple linear regression method to selection of the suitable controlled variables of Malaysia 'imports of palm oil. The researchers used three different variables namely; G.D.P, Real world price of palm oil, stock, and real world price of soybean oil. The estimated structural equation was linear-linear, and R2 = 0.5607, F-stat = 10.25, DW = 2.1839. The relevant DW-Statistic and F- stat revealed that the autocorrelation is not series. All the variables showed signs and significant coefficients as expected and the only exception was that the stock of palm oil showed a correct sign but an insignificant value. Prambudia & Nakano (2012)[5] in utilizing a multiple linear regression concept for analyzing the relationship of Malaysia's oil import demand function for the period from 1985 to 2009 identified two main factors affecting the imports: oil production and consumption. Moreover, the results revealed that the function form was linear-linear and coefficient of determination equal to 0.92. (Milad, el at.2014)[6] Considered variables such as exchange rate, tariff, sales tax, producer price index, and value of export in previous years and used the regression method to forecast future values of imports in Malaysia.

## IV. METHODOLOGY

Data: The data of this study was collected from Bank Negara Malaysia, Department of Statistics Malaysia and Royal Malaysian Customs Department. It covered a period of 23 years starting from the first quarter of 1991 until the third quarter in 2013. The main focus of the current research was on the imports of Crude Material.

Table (1) shows the data definitions and sources used in the present study.

Table1: Data definition and sources:

| Variable | Definition | Source |
|---|---|---|
| The value of imports of (CM). | The value of imports of (CM) of Malaysia, unit (Million RM) | Department of Statistics Malaysia. |
| The value of exports of (CM). | The value of export of (CM) of Malaysia, unit (Million RM). | Department of Statistics Malaysia. |
| PPI of (CM) | PPI (producer price index) constitutes local production and import price indices of (CM), 2005=100. | Department of Statistics Malaysia. |
| G.D.P | Gross domestic product, 2005=100. | Department of Statistics Malaysia. |
| Exchange Rate | Exchange Rate, against the USD. | Bank Negara Malaysia |
| Tariff tax | The average of tariff tax of imports of (CM). | Royal Malaysian customs department. |
| Sales tax | The average sales tax of imports of (CM). | Royal Malaysian customs department. |

Research method: Using a multiple regression analysis, the relation between the value of imports of (CM) and its influence factors such as Exchange rate, Producer price index of (CM), G.D.P, the value of exports of (CM), the average of tariff tax of (CM), and the average sales tax of (CM) was investigated and the model was revised and completed. There were five general steps in building a forecasting model of imports of (CM) in multiple linear regression. The general steps followed in developing the model in the present study were checking assumptions, developing equitation of multiple linear regression, regression model and hypothesis, diagnostic measures, and predictive ability.

Step1: Checking assumption:

Table2. PARTAL CORRELATION COEFICIENT

| Second Variables | First Variables | | | | |
|---|---|---|---|---|---|
| | Exchange rate $x_1$ | PPI $x_2$ | G.D.P. $x_3$ | Exports $x_4$ | Tariff tax $x_5$ |
| Price index $x_2$ | -0.273* | | | | |
| Sig. | 0.010 | | | | |
| G.D.P. $x_3$ | 0.283 | 0.641* | | | |
| Sig. | 0.008* | 0.000 | | | |
| The value of | -0.132 | 0.607* | 0.122 | | |

| | | | | | |
|---|---|---|---|---|---|
| exports $x_4$ | | | | | |
| Sig. | 0.221 | 0.000 | 0.262 | | |
| The average of tariff tax $x_5$ | -0.455* | 0.252 | -0.172 | -0.025 | |
| Sig. | 0.000 | 0.019* | 0.114 | 0.821 | |
| The average sales tax $x_6$ | 0.307* | 0.281 | -0.143 | 0.002 | -0.593* |
| Sig. | 0.004 | 0.009* | 0.189 | 0.982 | 0.000 |

There is a weak relationship between the average of Exchange rate $x_1$ and The value of exports $x_4$ ; also there is a weak relationship between the G.D.P $x_3$ and the value of exports $x_4$ ,The average of tariff tax $x_5$ ,The average sales tax $x_6$ , also there is a weak relationship between The value of exports $x_4$ and The average of tariff tax $x_5$ The average sales tax $x_6$ ,While there is significance partial correlation between other variables.

Previous research pointed at the importance of checking four assumptions: normality, linearity, heteroscedasticity and multicollinearity (Ghani et al.2010)[1].

The dependent variable must be a normal distribution which can be assessed using a chi-squared test (when $n > 30$), histogram graph, plot P-P, plot Q-Q, kurtosis and skewness.

If the distribution of data is not normal, so we need to use transformation. From the result of the chi-square test, normality assumption was acceptable (p>0.01)

Table3. Chi-Squared Goodness of fit test, Chi-Square = 14.03796, df = 1, p = 0.00018.

| | Observed Frequency | Percent Observed | Expected Frequency | Percent Expected | Observed-Expected |
|---|---|---|---|---|---|
| < 1100 | 17 | 18.6813187 | 18.3823265 | 20.2003588 | -1.38232651 |
| 1100-2700 | 41 | 45.0549451 | 28.7466804 | 31.5897587 | 12.2533196 |
| 2700-4300 | 13 | 14.2857143 | 27.702508 | 30.4423165 | -14.702508 |
| 4300-5900 | 12 | 13.1868132 | 12.9248346 | 14.203115 | -0.924834634 |
| > 5900 | 8 | 8.79120879 | 3.24365047 | 3.56445106 | 4.75634953 |

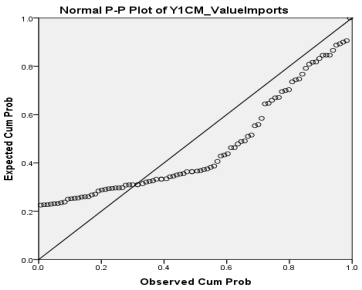



**Figure (1): Normal P-P plot of Malaysya'imports of Crude Material.**

Figure1: The P-P plot suggested that the time data plotted is not normally distributed as there is a considerable departure from the straight line. At the lower values, there is far more data points expected than observed, middle we have less than expected. This may be also seen in the histogram of the data. A transformation of the data may be necessary. Then, it is important for a multiple linear regression to exhibit a linear relationship between the response variable and controlled variables. In regression the model, we fit it using a linear mode. In this regard, Andy (2005) [7] indicated that a linear model is known as a model which is usually based on a straight line. In the present study, a Pearson linear correlation was used for checking whether there was a relationship between two variables. In other words, a simple correlation analysis can be used for measuring the linear relationship between dependent and independent variables.

Table 4. Correlations between dependent variable and independent variables.

| Dependent Variable | Independent Variables | | | | | |
|---|---|---|---|---|---|---|
| | Exchange | PPI | G.D.P | Exports | The average of tariff tax $x_5$ | the average |

| | rate $x_1$ | $x_2$ | $x_3$ | $x_4$ | | sales tax $x_6$ |
|---|---|---|---|---|---|---|
| Correlation | 0.158 | 0.953 | 0.974 | 0.887 | -0.578 | 0.537 |
| Sig. | 0.135 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |

Using a 0.01 particular significance level and 0.05 significance level, results showed a significance positive linear correlation between the value of imports $y$. (Crude Materials) and the producer price index of (Crude Materials), G.D.P $x_{3t}$, The value of exports $x_{4t}$ (Crude Materials) , there is a significance negative linear correlation between the value of imports $y$ and the average of tariff tax $x_{5t}$ (Crude Materials), and the average sales tax $x_{6t}$ (Crude Materials). Secondly, there no significance linear correlation was found between the value of imports $y$ (Crude Materials) and influence factors Exchange rate $x_{1t}$ .Next, in MLR, it is assumed that any data should be free from heteroscedasticity which usually occurs in a case when an interruption in the model is not fulfilled. This means that it occurs when there are missing important variables in the model. Thus, any accurate developed model is capable of checking whether there is heteroscedasticity or not based on Spearman's rank correlation test. Or based on $F$ – test.

Step 2 : Developing Equation of Multiple Linear Regression:

In this research, the hypotheses that used:

$H_0 : \beta_1 = \beta_2 = ..... = \beta_6$

$H_1$ : At least one of the $\beta_1, \beta_2, .....and \beta_6$, does not equal to 0

Which says that?

$H_0$ : None of the controlled variables $x_1, x_2, x_3, x_4, x_5, and x_6$ is significantly related to $y$.

$H_1$ : At least one of the controlled variables $x_1, x_2, x_3, x_4, x_5, and x_6$ is significantly related to $y$.

The model of multiple linear regression can be represent as:

$$y_t = \beta_0 + \beta_1 x_{1t} + \beta_2 x_{2t} + .....+ \beta_6 x_{6t} + et$$

Where

$y_t$ =Dependent variable (The value of imports), at period $t$.

$\beta_0$ = Constant variable.

$\beta_1$ = Coefficient of first control variable, $x_1$

$\beta_2$ = Coefficient of second control variable, $x_2$

$\beta_3$ = Coefficient of third control variable, $x_3$

$\beta_4$ = Coefficient of fourth control variable, $x_4$

$\beta_5$ = Coefficient of fifth control variable, $x_5$

$\beta_6$ = Coefficient of sixth control variable, $x_6$

$x_{1t}, x_{2t}, x_{3t}, x_{4t}, x_{5t}, and x_{6t}$ are called predictor variables or independent variables at period t.

$e_t$ = The random Error.

Table 5. Model Summary

| Model | $R^2$ | adjusted $R^2$ | D.W. | F Test | Sig. |
|---|---|---|---|---|---|
| (I)- Linear-Linear | 97.21 | 97.08 | 1.63659 | 748.46 | 0.000 |
| (II)- Log-Log | 97.89 | 97.82 | 1.31245 | 1345.29 | 0.000 |
| (III)- log-linear | 97.73 | 97.63 | 1.27732 | 927.41 | 0.000 |
| (IV)- Linear-log | 94.95 | 94.72 | 1.21357 | 404.64 | 0.000 |

Table 5 provides a summary of the model.

Model I: The following represents the model of multiple linear-linear regression:

$$y_t = \beta_0 + \beta_1 x_{1t} + \beta_2 x_{2t} + ..... + \beta_6 x_{6t} + e_t$$

Based on Table 5, the value of $R^2$ is .9721 (Model I) indicates that there are 97.21% changes in the response variable y which happened due to the changes in combination of four controlled variables which are $x_{2t}, x_{3t}, x_{4t}, and x_{6t}$.

Model II: The following represents the model of multiple log-log regression:

$$Lny = \beta_0 + \beta_1 Lnx_{1t} + \beta_2 Lnx_{2t} + .....+ \beta_6 Lnx_{6t} + e_t$$

From Table 5, it is evident that the value of $R^2$ is .9789 (Model II) reveals 97.89% changes in the response variable $y$ which are attributed to because the changes in combination of three controlled variables: $x_{2t}, x_{3t}, and x_{4t}$.

Model III: The following equation represents the model of multiple log-linear:

$$Lny = \beta_0 + \beta_1 x_{1t} + \beta_2 x_{2t} + ..... + \beta_6 x_{6t} + e_t$$

As seen in Table 5, the value of $R^2$ is .9773 (Model III) indicates that there 97.73% changes in the response variable $y$ due to the changes in combination of four controlled variables namely; $x_1, x_3, x_4, and x_5$.

Model IV: This equation represents the model of multiple linear-log regression:

$$y_t = \beta_0 + \beta_1 Lnx_{1t} + \beta_2 Lnx_{2t} + ..... + \beta_6 Lnx_{6t} + e_t$$

From Table 5, the value of $R^2$ is 94.95 (Model IV) reveals the existence of 94.95% changes in the response variable $y$ which are caused by because the changes in combination of four controlled variables: $x_2, x_3, x_4, and x_6$.

In comparing the above four models, Models I, II and III provide better fits to the data than Model IV. This is due to the evidence that the higher the value or $R^2$ and adjusted $R^2$ (Model I = .9721; Model II = .9789; Model III =0.9773 and Model IV=0.9495).

Table 6. Stepwise regression results

| Models | | Coefficients | Standard Error | $t$-value | Significance value | *VIF* |
|---|---|---|---|---|---|---|
| I | Constant | -1047.5 | 150.9 | -6.94 | 0.000 | |
| | $x_{2t}$ | 6.947 | 3.708 | 1.87 | 0.064 | 18.759 |
| | $x_{3t}$ | 0.021487 | 0.001534 | 14.00 | 0.000 | 8.916 |
| | $x_{4t}$ | 0.25422 | 0.08303 | 3.06 | 0.003 | 8.157 |
| | $x_{6t}$ | -65763 | 20032 | -3.28 | 0.001 | 1.666 |
| II | Constant | -6.6382 | 0.5543 | -11.98 | 0.000 | |
| | $x_{2t}$ | 0.2946 | 0.1389 | 1.47 | 0.144 | 18.186 |
| | $x_{3t}$ | 1.00879 | 0.05079 | 19.86 | 0.000 | 7.178 |
| | $x_{4t}$ | 0.21129 | 0.09034 | 2.34 | 0.022 | 6.859 |
| III | Constant | 6.2658 | 0.1453 | 43.14 | 0.000 | |
| | $x_{1t}$ | 0.07321 | 0.03433 | 2.13 | 0.036 | 2.46 |
| | $x_{3t}$ | 0.00000731 | 0.00000053 | 13.87 | 0.000 | 8.745 |
| | $x_{4t}$ | 0.00010773 | 0.00002474 | 4.35 | 0.000 | 6.017 |
| | $x_{5t}$ | -0.39424 | 0.05599 | -7.04 | 0.000 | 3.880 |
| IV | Constant | -39925 | 2862 | -13.95 | 0.000 | |
| | $x_{2t}$ | 2659.8 | 562.4 | 4.73 | 0.000 | 18.299 |
| | $x_{3t}$ | 1588.6 | 220.0 | 7.22 | 0.000 | 8.261 |
| | $x_{4t}$ | 820.8 | 368.5 | 2.23 | 0.000 | 7.002 |
| | $x_{6t}$ | -1079.9 | 216.2 | -5.00 | 0.000 | 2.190 |

As shown in Table 6, the coefficients indicate that it is possible to develop four models using MLR equation such as below:

Model I: $y_t = -1047.5 + 6.947 x_{2t} + 0.021487 x_{3t} + 0.25422 x_{4t} - 65763 x_{6t}$

Model II: $Lny_t = -6.638 + 0.295 Lnx_{2t} + 1.009 Lnx_{3t} + 0.211 Lnx4t$

Model III: $Lny_t = 6.266 + 0.0732 x_{1t} + 0.0000073 x_{3t} + 0.00011 x_{5t}$ Model IV:

$y_t = -39925 + 2660 Lnx_{2t} + 1589 Lnx_{3t} + 821 Lnx_{4t} - 1079.9 Lnx_{6t}$

Based on the Model I, the coefficients for $x_{2t}$ ($\beta_2 = 6.947, p < 0.05$), $x_{3t}$ ($\beta_3 = 0.0214, p < 0.05$), $x_{4t}$ ($\beta_4 = 0.254, p < 0.05$) and $x_{6t}$ ($\beta_6 = -65763, p < 0.05$) are significant. This explained that $x_{2t}, x_{3t}, x_{4t}, and x_{6t}$ are factors to the y.

In Model II, the coefficients for the coefficients for $x_{2t}$ ($\beta_2 = 0.20, p < 0.05$), $x_{3t}$ ($\beta_3 = 1.009, p < 0.05$), and $x_{4t}$ ($\beta_4 = -0.211, p < 0.05$) are significant. This explained that $x_{2t}, x_{3t}, and x_{4t}$ are factors to the $y$.

In Model III, the coefficients for the coefficients for $x_{1t}$ ($\beta_1 = 0.073, p < 0.05$), $x_{3t}$ ($\beta_3 = 0.00001, p < 0.05$), $x_{4t}$ ($\beta_4 = -0.00011, p < 0.05$) and $x_{5t}$ ($\beta_5 = -0.394, p < 0.05$) and and are significant. This explained that $x_{1t}, x_{3t}, x_{4t}, and x_{5t}$ are factors to the $y$.

While in Model IV, the coefficients for $x_{2t}$ ($\beta_2 = 2660, p < 0.05$), $x_{3t}$ ($\beta_3 = 1589, p < 0.05$), $x_4$ ($\beta_4 = 821, p < 0.05$), and $x_{6t}$ ($\beta_6 = -1080, p < 0.05$) are significant. This explained that $x_{2t} x_{3t}, x_{4t}, and x_{6t}$ are factors to the $y$.

Step 3: Regression Model and Hypothesis Testing:
Investigation was started off by using a multiple linear regression analysis and obtaining the regression model between the value of imports of (CM) and exchange rate, producer price index of (CM), Gross Domestic Products (GDP), the value of exports of (CM), average sales tax and the average of tariff tax of (CM).

Variance inflation factor VIF: Any analysis is capable of detecting whether multicollinearity is present or not just by checking the value of the variation inflation factor (VIF). It is not series when the value of VIF is less than 5 whereas if the VIF is more than 5, then multicollinearity is substantial. It is also expected that multicollinearity can be more serious whenever the value of VIF is more than 10.

Decision: The values of VIF indicate that multicollineairty is substantial in all models.
Watson statistic (D.W): We will need to compare the displayed statistic with lower and upper bounds in a table. If D.W > upper bound, no correlation exists; if D.W < lower bound, positive correlation exists; if D.W is in between the two bounds, the test is inconclusive.

Decision: At the 5% level of significance with n=91 and there is evidence of positive autocorrelation for the second, third and fourth model, but in first model is inconclusive, thus model 1 is better than others models, but it still suffers from the presence of heterosedasticity. Then, in tackling the problem heterosedasticity, the results were as follows.

The regression equation is:

$$\frac{\hat{y}_t}{x_{2t}} = 10.176 - \frac{968.9}{x_{2t}} + 0.0205\frac{x_{3t}}{x_{2t}} + 0.153\frac{x_{4t}}{x_{2t}} - 61405\frac{x_{6t}}{x_{2t}}$$

Table 7. Model Summary

| $R^2$ | adjusted $R^2$ | $R^2$(Pred) | PRESS | S | D.W. |
|---|---|---|---|---|---|
| 91.1 % | 90.6 | 89.77% | 681.950 | 2.63266 | 1.419 |

Table 8. Analysis of variance

| Source | *DF* | SS | MS | F | P |
|---|---|---|---|---|---|
| Regression | 4 | 6071.6 | 1517.9 | 219.01 | 0.000 |
| Residual Error | 86 | 596.1 | 6.9 | | |
| Total | 90 | 6667.7 | | | |

The model is good in terms of (VIF), but it is not good in terms of D.W, which refers to problem of autocorrelation in residuals. Then, in tackling the problem of autocorrelation in residuals, the results were as follows.

The regression equation is

$$\frac{y_t^*}{x_{2t}} = 5.61 - 955\frac{1^*}{x_{2t}} + 0.0212\frac{x_{3t}^*}{x_{2t}} + 0.209\frac{x_{4t}^*}{x_{2t}} - 56271\frac{x_{6t}^*}{x_{2t}}$$

Table 9. Analysis of Variance Table

| Source | DF | SS | MS | F | P |
|---|---|---|---|---|---|
| Regression | 4 | 3795.09 | 948.77 | 219.01 | 0.000 |
| Residual Error | 86 | 552.63 | 6.43 | | |
| Total | 90 | 4347.72 | | | |

Table 10. Model Summary

| $R^2$ | adjusted $R^2$ | S | D.W. |
|---|---|---|---|
| 87.3% | 86.7% | 2.535 | 1.69 |

Step 4: predictive ability: In this step, a predicted $R^2$ to was used to determine how well the model predicts responses for new observations. Thus, a larger value of predicted $R^2$ indicates that the model is of a greater predictive ability. In this study, the

Predicted $R^2$ ranged from 0 and 100% and was calculated from the PRESS statistic.

Through the value of $R^2$ (Pred), it can be concluded that the model has a good predictive ability.

$$\text{PRESS} = \sum_{i=1}^{n} \left(\frac{e_i}{1-h_i}\right)^2$$

$$R^2\,(\text{Pred}) = 1 - \frac{\sum_{i=1}^{n}\left(\frac{e_i}{1-h_i}\right)^2}{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}$$

PRESS = 631.731 $R^2$ (Pred) = 85.47%

Through the value of $R^2$ (Pred), conclude that the model has good predictive ability.

Step 5: Diagnostic measures:

Residual: The difference between an observed value y and its corresponding fitted value $\hat{y}$.

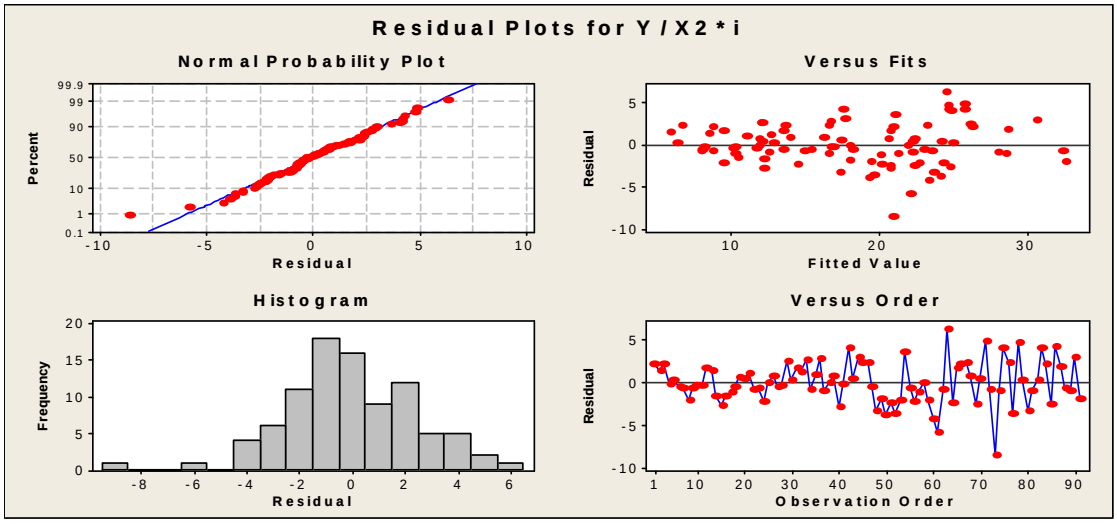



**Figure (2) for residual plots for $\frac{y_t^*}{x_{2t}}$ Model:**

Table 11. Illustrates the diagnostic measures to make sure that the model suitable to predict.

| Residual | Standard residual | Studentized Residual | Leverage | Cook 'distance | DFIT | Fitted values | I |
|---|---|---|---|---|---|---|---|
| 2.25513 | 0.96898 | 0.96863 | 0.157096 | 0.034998 | 0.41817 | 6.6016 | 1 |
| 1.44402 | 0.58618 | 0.58393 | 0.055638 | 0.004049 | 0.14174 | 5.8783 | 2 |
| 2.19744 | 0.90468 | 0.90371 | 0.081862 | 0.014595 | 0.26985 | 8.7197 | 3 |
| -0.25234 | -0.10250 | -0.10191 | 0.056800 | 0.000127 | -0.02501 | 8.2530 | 4 |
| 0.28083 | 0.11408 | 0.11342 | 0.056941 | 0.000157 | 0.02787 | 6.3887 | 5 |
| -0.52491 | -0.21312 | -0.21193 | 0.055961 | 0.000538 | -0.05160 | 8.0953 | 6 |
| -0.70074 | -0.28270 | -0.28118 | 0.043862 | 0.000733 | -0.06022 | 8.7059 | 7 |
| -2.10379 | -0.85602 | -0.85468 | 0.060066 | 0.009365 | -0.21606 | 9.4681 | 8 |
| -0.67287 | -0.27155 | -0.27008 | 0.044527 | 0.000687 | -0.05830 | 7.8748 | 9 |
| -0.28905 | -0.11763 | -0.11696 | 0.060384 | 0.000178 | -0.02965 | 10.2138 | 10 |
| -0.36306 | -0.14610 | -0.14527 | 0.039028 | 0.000173 | -0.02928 | 10.1529 | 11 |
| 1.71311 | 0.68884 | 0.68672 | 0.037505 | 0.003698 | 0.13556 | 9.4992 | 12 |
| 1.39610 | 0.56647 | 0.56422 | 0.054776 | 0.003719 | 0.13582 | 8.3873 | 13 |
| -1.53319 | -0.61687 | -0.61464 | 0.038691 | 0.003063 | -0.12331 | 10.4666 | 14 |
| -2.73738 | -1.10418 | -1.10560 | 0.043570 | 0.011108 | -0.23598 | 12.2687 | 15 |
| - | -0.66459 | -0.66241 | 0.038475 | 0.003535 | - | 12.1827 | 16 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 1.65196 | | | | | 0.13251 | | |
| -1.12712 | -0.45566 | -0.45355 | 0.047796 | 0.002084 | -0.10161 | 10.3387 | 17 |
| -0.45114 | -0.18128 | -0.18026 | 0.036205 | 0.000247 | -0.03494 | 11.6764 | 18 |
| 0.62691 | 0.25041 | 0.24904 | 0.024645 | 0.000317 | 0.03959 | 11.8997 | 19 |
| 0.45068 | 0.17979 | 0.17877 | 0.022097 | 0.000146 | 0.02687 | 12.1084 | 20 |
| 1.06545 | 0.42952 | 0.42748 | 0.042472 | 0.001637 | 0.09003 | 11.0799 | 21 |
| -0.89373 | -0.35696 | -0.35514 | 0.024496 | 0.000640 | -0.05628 | 12.4716 | 22 |
| -0.65327 | -0.25992 | -0.25851 | 0.016992 | 0.000234 | -0.03399 | 13.5031 | 23 |
| -2.29439 | -0.91395 | -0.91307 | 0.019271 | 0.003283 | -0.12799 | 14.4117 | 24 |
| -0.02175 | -0.00881 | -0.00875 | 0.050952 | 0.000001 | -0.00203 | 11.8528 | 25 |
| 0.81365 | 0.32396 | 0.32226 | 0.018338 | 0.000392 | 0.04405 | 13.9154 | 26 |
| -0.58279 | -0.23288 | -0.23159 | 0.025391 | 0.000283 | -0.03738 | 15.3757 | 27 |
| -0.28481 | -0.11396 | -0.11330 | 0.027922 | 0.000075 | -0.01920 | 16.8713 | 28 |
| 2.55427 | 1.03105 | 1.03144 | 0.044941 | 0.010005 | 0.22374 | 12.1662 | 29 |
| 0.24762 | 0.09921 | 0.09863 | 0.030446 | 0.000062 | 0.01748 | 12.8838 | 30 |
| 1.72522 | 0.69093 | 0.68881 | 0.029745 | 0.002927 | 0.12060 | 13.4746 | 31 |
| 1.15871 | 0.46669 | 0.46456 | 0.040696 | 0.001848 | 0.09568 | 12.6190 | 32 |
| 2.68626 | 1.08439 | 1.08552 | 0.045050 | 0.011095 | 0.23577 | 11.9262 | 33 |
| -0.76296 | -0.30467 | -0.30305 | 0.024093 | 0.000458 | -0.04762 | 14.9180 | 34 |
| 0.90411 | 0.36243 | 0.36059 | 0.031613 | 0.000858 | 0.06515 | 16.2600 | 35 |
| 2.74980 | 1.09700 | 1.09831 | 0.022202 | 0.005465 | 0.16550 | 16.6726 | 36 |
| -0.97978 | -0.39623 | -0.39428 | 0.048454 | 0.001599 | -0.08897 | 16.4882 | 37 |
| -0.07895 | -0.03159 | -0.03140 | 0.027765 | 0.000006 | -0.00531 | 18.0278 | 38 |
| 0.71961 | 0.29536 | 0.29379 | 0.076281 | 0.001441 | 0.08443 | 20.5270 | 39 |
| -2.81972 | -1.17405 | -1.17667 | 0.102371 | 0.031440 | -0.39737 | 20.6643 | 40 |
| -0.19853 | -0.08103 | -0.08056 | 0.065918 | 0.000093 | -0.02140 | 16.6571 | 41 |
| 4.14435 | 1.71585 | 1.73582 | 0.092150 | 0.059768 | 0.55303 | 17.6281 | 42 |
| 0.47255 | 0.19434 | 0.19325 | 0.079911 | 0.000656 | 0.05695 | 17.4660 | 43 |
| 3.03297 | 1.26433 | 1.26881 | 0.104486 | 0.037302 | 0.43340 | 17.6444 | 44 |
| 2.31500 | 1.00744 | 1.00753 | 0.178287 | 0.044042 | 0.46931 | 13.6374 | 45 |
| 2.36584 | 0.97314 | 0.97283 | 0.080217 | 0.016518 | 0.28730 | 16.5320 | 46 |
| -0.51821 | -0.21096 | -0.20978 | 0.060961 | 0.000578 | -0.05345 | 18.1945 | 47 |
| -3.33290 | -1.35628 | -1.36303 | 0.060267 | 0.023594 | -0.34518 | 17.2534 | 48 |
| -1.88227 | -0.76390 | -0.76204 | 0.055171 | 0.006815 | -0.18414 | 17.9839 | 49 |
| -3.86980 | -1.54520 | -1.55797 | 0.023956 | 0.011721 | -0.24408 | 19.2446 | 50 |
| -2.39850 | -0.96091 | -0.96048 | 0.030435 | 0.005797 | -0.17017 | 20.2013 | 51 |
| -3.67640 | -1.46354 | -1.47347 | 0.018027 | 0.007864 | -0.19964 | 19.6393 | 52 |
| -2.07629 | -0.82671 | -0.82518 | 0.018422 | 0.002565 | -0.11305 | 19.3719 | 53 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 3.57862 | 1.42372 | 1.43240 | 0.016803 | 0.006928 | 0.18725 | 21.0335 | 54 |
| -0.64571 | -0.25904 | -0.25763 | 0.033052 | 0.000459 | -0.04763 | 23.0284 | 55 |
| -2.18856 | -0.87553 | -0.87433 | 0.027633 | 0.004357 | -0.14739 | 22.6611 | 56 |
| -1.21191 | -0.49486 | -0.49267 | 0.066656 | 0.003498 | -0.13166 | 20.0610 | 57 |
| -0.02857 | -0.01156 | -0.01149 | 0.049290 | 0.000001 | -0.00262 | 21.8320 | 58 |
| -2.16226 | -0.89239 | -0.89132 | 0.086372 | 0.015057 | -0.27405 | 24.3084 | 59 |
| -4.23786 | -1.71756 | -1.73761 | 0.052611 | 0.032765 | -0.40948 | 23.2748 | 60 |
| -5.80152 | -2.38784 | -2.45676 | 0.081387 | 0.101034 | -0.73126 | 22.0979 | 61 |
| -0.79432 | -0.32831 | -0.32660 | 0.089088 | 0.002108 | -0.10214 | 23.5268 | 62 |
| 6.33882 | 2.64279 | 2.74104 | 0.104734 | 0.163415 | 0.93753 | 24.5315 | 63 |
| -2.46973 | -0.99784 | -0.99782 | 0.046687 | 0.009753 | -0.22082 | 22.3504 | 64 |
| 1.64044 | 0.66052 | 0.65834 | 0.040132 | 0.003648 | 0.13461 | 20.7094 | 65 |
| 2.19308 | 0.88543 | 0.88431 | 0.045318 | 0.007443 | 0.19267 | 20.9313 | 66 |
| 2.35238 | 0.94632 | 0.94574 | 0.038380 | 0.007148 | 0.18894 | 23.1948 | 67 |
| 0.76360 | 0.30832 | 0.30669 | 0.045446 | 0.000905 | 0.06692 | 22.4550 | 68 |
| -2.55864 | -1.04656 | -1.04714 | 0.069850 | 0.016450 | -0.28695 | 20.7707 | 69 |
| 0.44552 | 0.18099 | 0.17997 | 0.057079 | 0.000397 | 0.04428 | 22.2846 | 70 |
| 4.90155 | 1.97433 | 2.00887 | 0.040848 | 0.033201 | 0.41457 | 25.7520 | 71 |
| -0.85604 | -0.35374 | -0.35194 | 0.088691 | 0.002436 | -0.10979 | 28.0700 | 72 |
| -8.61852 | -3.54692 | -3.81642 | 0.081198 | 0.222359 | -1.13453 | 20.9571 | 73 |
| -1.04439 | -0.42507 | -0.42303 | 0.060567 | 0.002330 | -0.10741 | 21.2676 | 74 |
| 4.06840 | 1.63438 | 1.65068 | 0.035718 | 0.019789 | 0.31769 | 24.8295 | 75 |
| 2.37968 | 0.95652 | 0.95604 | 0.036814 | 0.006994 | 0.18691 | 26.2273 | 76 |
| -3.73142 | -1.49740 | -1.50847 | 0.033656 | 0.015618 | -0.28151 | 24.0383 | 77 |
| 4.76455 | 1.90786 | 1.93820 | 0.029467 | 0.022103 | 0.33773 | 24.6060 | 78 |
| 0.21127 | 0.08470 | 0.08421 | 0.031848 | 0.000047 | 0.01527 | 25.0642 | 79 |
| -3.32035 | -1.34725 | -1.35375 | 0.054781 | 0.021039 | -0.32590 | 23.6764 | 80 |
| -0.95260 | -0.39295 | -0.39101 | 0.085461 | 0.002886 | -0.11953 | 22.3037 | 81 |
| 0.32230 | 0.13137 | 0.13062 | 0.063341 | 0.000233 | 0.03397 | 24.1903 | 82 |
| 4.13390 | 1.67613 | 1.69426 | 0.053397 | 0.031695 | 0.40240 | 24.5542 | 83 |
| 2.19516 | 0.88927 | 0.88818 | 0.051746 | 0.008631 | 0.20748 | 26.3214 | 84 |
| -2.63310 | -1.09090 | -1.09212 | 0.093384 | 0.024516 | -0.35051 | 24.7884 | 85 |
| 4.28432 | 1.75494 | 1.77681 | 0.072525 | 0.048165 | 0.49686 | 25.8408 | 86 |
| 1.88989 | 0.77260 | 0.77078 | 0.068853 | 0.008828 | 0.20960 | 28.6093 | 87 |
| -0.69763 | -0.29307 | -0.29151 | 0.118212 | 0.002303 | -0.10673 | 32.4081 | 88 |
| -1.01086 | -0.41309 | -0.41109 | 0.068137 | 0.002495 | -0.11116 | 28.5683 | 89 |
| 2.90845 | 1.20145 | 1.20460 | 0.088057 | 0.027877 | 0.37432 | 30.5721 | 90 |
| -1.99496 | -0.83852 | -0.83706 | 0.119156 | 0.019023 | -0.30787 | 32.5367 | 91 |

## CONCLUSIONS

The current study aimed to build a forecasting model for $y$ by using the multiple regressions (MR). The statically analyzed and obtained output from SPSS 21.0 showed four models for building this. The four multiple regression models were estimated and were found to be all good in terms the coefficient of determination R2.

It was found that the model 1(linear Model) is the best model among the four models, but it still suffers from problem of multicollinearity . However, after treating the model and solving the problem of multicollinearity , there was still a problem of autocorrelation in the model which was later resolved.

We obtained an effective model which does not suffer from this problem and is capable of prediction, and consistent with the economic theory in terms of signal parameters. It was also found that the first model is good in terms of its predictive ability. Diagnostic measures showed that the model is very suitable for predicting this Model 1(linear model) shows that there are only three controlled variables by using stepwise regression and produced a regression model such as follows:

$$\frac{y_t^*}{x_{2t}} = 5.61 - 955\frac{1^*}{x_{2t}} + 0.0212\frac{x_{3t}^*}{x_{2t}} + 0.209\frac{x_{4t}^*}{x_{2t}} - 56271\frac{x_{6t}^*}{x_{2t}}$$

## REFERENCES


[1] Ghani, I. M. M., & Ahmad, S. (2010). Stepwise multiple regression method to forecast fish landing. Procedia-Social and Behavioral Sciences, 8, 549-554.

[2] Abu Bakar, Nor'Aznin (2000) Time series estimation of Malaysia's export and import demand :a dynamic old method. Analisis, 7 (1&2). pp. 61-77, UUM.

[3] Tang, T. C., & Nair, M. (2002). A cointegration analysis of Malaysian import demand function: reassessment from the bounds test. Applied Economics Letters, 9(5), 293-296.

[4] Arshad, F. M., Shamsudin, M. N., & Hameed, A. A. A. (2011). An econometric analysis of the link between biodiesel demand and malaysian palm oil market.International Journal of Business and Management, 6(2), p35.

[5] Prambudia, Y., & Nakano, M. (2012). Exploring Malaysia's transformation to net oil importer and oil import dependence. Energies, 5(8), 2989-3018.

[6] Milad, M. A., Rose, I. B. I., & Marappan, S. (2014). Modeling and Forecasting the Volumes of Malaysia's Import. In International Conference on Global Trends in Academic Research, Bali, Indonesia Global Illuminators, Kuala Lumpur, Malaysia.

[7] Andy, P. (2005). Discovering Statistics Using SPSS:(and Sex and Drugs and Rock'n'Roll). Sage publications.


★★★