Far East Journal of Mathematical Sciences (FJMS)

http://www.pphmj.com
http://dx.doi.org/10.17654/MS101122731
Volume 101, Number 12, 2017, Pages 2731-2744 **ISSN: 0972-0871**

COMPOSITE MODEL APPROACH FOR FORECASTING MALAYSIAN IMPORTS OF MACHINERY AND TRANSPORT EQUIPMENTS: A COMPARATIVE STUDY

Mohamed A. H. Milad, Rose Irnawaty Ibrahim and Samiappan Marappan

Faculty of Science and Technology
Universiti Sains Islam Malaysia (USIM)
Bandar Baru Nilai, 71800 Nilai
Negeri Sembilan, Malaysia

Abstract

Imports are regarded as an important factor in contributing to the economic growth. Specifically, this study proposes a composite model for predicting the future imports of machinery and transport equipments in the Malaysian context. In this study, a proposed composite model (with regression processing of autocorrelation) was employed for extracting information that increases the accurate prediction of the size of future imports as well as enhances forecasting methods in Malaysia and compares with commonly used methods including regression method, ARIMA models, composite model (without regression processing), and simple seasonal exponential smoothing model. The forecasting results of this study reveal that the proposed composite model (with regression processing of autocorrelation) offers more probabilistic information that improves

Received: January 13, 2017; Revised: March 2, 2017; Accepted: March 28, 2017

2010 Mathematics Subject Classification: 62P05.

Keywords and phrases: Malaysian imports, forecasting, composite model.

the forecasting of Malaysian imports of machinery and transport equipment than other models. The proposed model offers a good solution to the problem of autocorrelation in residuals. The emerging problem cannot be processed in the regression model; it can be processed only through the composite model (without regression processing).

Abbreviations

MTE refers to machinery and transport equipments. SSES refers to seasonal exponential smoothing. RM-WOP is simple for regression model (without processing). RM-WPA refers to regression model (with processing of autocorrelation). CM-WRP is a composite model (with regression processing). CM-WORP refers to composite model (without regression processing). CM-WRPA is a composite model (with regression processing of autocorrelation). CM-WRPH refers to composite model (with regression processing of heteroscedasticity). CM-WRPHA is simple for composite model (with regression processing of heteroscedasticity and autocorrelation).

Introduction

Previous studies have defined the relationship between imports and economic growth. Imports seriously affect and significantly facilitate economic growth and development.

In the Malaysian context, previous studies intended to develop a model for predicting Malaysian imports. These studies applied various statistical methods [1-5]. Most of such studies used an ordinary least square regression method as basis in which the response variable is the value of imports and the explanatory variables are the GDP and prices. Osman [6] determined the best-fitted model among exponential smoothing methods, while Shabri et al. [7] developed a model to predict the yields of rice imports in Malaysia using two methods of artificial neural network (ANN) predictability, namely, the statistical autoregressive integrated moving average and double exponential smoothing.

The composite model (combining regression and ARIMA) is a confirmed valuable for predicting a better forecast than using the two techniques separately. Another previous study [8] proposed two composite models, namely, the CM-WRPH and CM-WRPHA, to forecast the value of imports in Malaysia.

The overall purpose of this study is to propose a composite model to forecasting the values of imports in Malaysia and the proposed approach is compared empirically with other used methods in terms of the measurement criteria on the forecasting performance.

Material and Methods

Material. This section describes the case study which is regarded as an effective research approach to investigating and comparing the proposed model and other relevant models. This case study is adopted based on the steps discussed below.

Data collection. The efficiency and reliability of the proposed forecasting model, which used data on the imports of MTE in Malaysia covering period of 23 years have been tested and demonstrated. In particular, the period covered by the data is from the first quarter of 1991 to the fourth quarter of 2013. Figure 1 and Table 1 illustrate the period of data collection in conjunction with the definitions and sources.

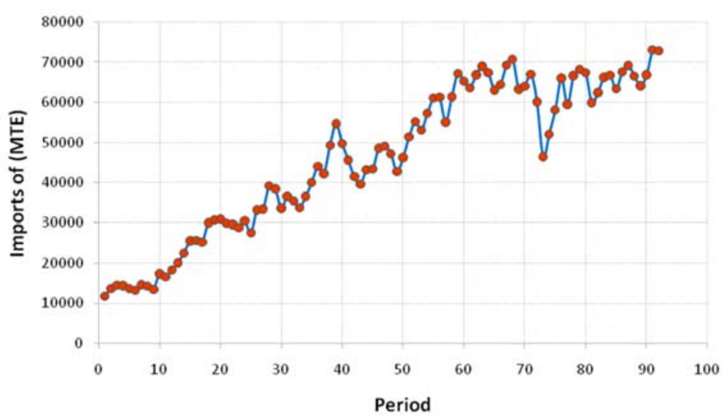

Figure 1. Time series of imports of MTE (Q1: 1991-Q4: 2013).

Table 1. Data, definition and sources

Variable	Definition	Source
The value of imports of MTE	The value of imports of MTE from Malaysia, unit (Million RM)	Department of Statistics of Malaysia
The value of exports of MTE	The value of exports of MTE from Malaysia, unit (Million RM)	Department of Statistics of Malaysia
PPI for the domestic economy	The PPI (producer price index) for the domestic economy is a composite index based on the price data derived from that of local production and import price indices of MTE, 2005=100	Department of Statistics of Malaysia
GDP	Gross domestic product, 2005=100	Department of Statistics of Malaysia
Exchange rate	Exchange rate, against the USD	Bank Negara Malaysia
Tariff tax	The average of tariff tax on imports of MTE	Royal Malaysian Customs Department
Sales tax	The average sales tax of imports of MTE	Royal Malaysian Customs Department

Evaluation of the forecasting performance indices. This study employed test significance of parameters and two statistical indices to measure and evaluate the proposed approach and in particular, its forecasting accuracy. Theil's inequality coefficients were used as the first statistical indices. The small values indicate that the forecast performance is high. This coefficient has a value between 0 and 1, which indicate a perfect prediction and a perfect inequality, respectively. The following equation illustrates this coefficient:

$$U = \sqrt{\frac{1}{n}\sum (y_t - \hat{y}_t)^2} \Big/ \sqrt{\frac{1}{n}\sum \hat{y}_t^2} + \sqrt{\frac{1}{n}\sum y_t^2},$$

where y_t is the real data for period t and $\hat{y}_t$ represents the predicted value and point at the same time, while n refers to the number of periods [9].

This study also used another type of statistical index, that is, the predicted R^2. This index was used to test the capability of the proposed model in forecasting the responses to new observations.

In using the predicted R^2, its higher value indicates the high capability of the developed model for forecasting such responses. Thus, its value is between 0 and 1. For its calculation, if $R^2(Pred) = 0$, then the predictive performance is regarded as poor. On the contrary, if $R^2(Pred) = 1$, then the model has a perfect fit [5]. This value can be expressed in the following formula:

$$R^2(Predicted) = 1 - \left[PRESS / \sum_{i=1}^{n} (y_i - \bar{y})^2 \right],$$

where $PRESS = \sum_{i=1}^{n} e_i^2$.

Methods. This section describes the method proposed in this study along with other methods, including the multiple linear regression method, ARIMA models, exponential smoothing method, and CM-WORP. This section also describes the operation of the proposed CM-WRPA. As indicated in Figure 2, the red square represents the proposed model, whereas the white squares represent other methods compared with the proposed model in the case study of MTE.

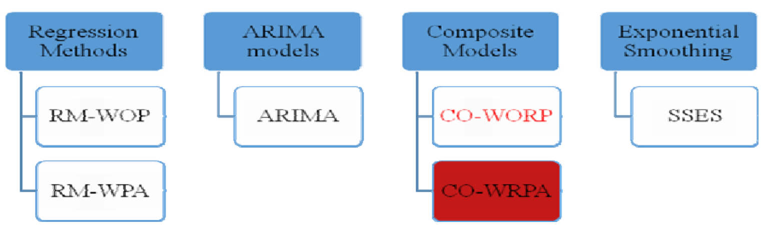

Figure 2. Overall forecasting models.

Multiple linear regression method. Regression analysis is defined as a statistical methodology that uses the relation between two or more quantitative variables. In using this analysis, the variable which is a response or an outcome can be predicted from other variables. The use of the linear regression analysis has been documented or reported in research in various areas, including social science, business, behavioral science, and biological sciences as well as many other disciplines [10]. The following equation illustrates the multiple linear regression model:

$$Y = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + \cdots + \beta_k X_{kt} + E_t,$$

where Y denotes the development trend of profession, which is generally expressed by demand, and X_{1t}, X_{2t}, ..., X_{kt} represent the influential factors of the profession development trend, which are the explanatory variables. In addition, β_0, β_1, β_2, ..., β_k refer to the regression coefficients, and E_t is the random error with its zero mean and constant variance σ. E_t is also assumed to have no correlation [11].

ARIMA model. The ARIMA model analyzes and forecasts equally spaced univariate time series data. An ARIMA model predicts a value in a response time series as a linear combination of its own past values. The ARIMA approach was first popularized by Box and Jenkins, and thus, ARIMA models are often referred to as Box-Jenkins models. In this study, the analysis performed by ARIMA is divided into three stages: identifying the model, estimating the parameter, and diagnosing and forecasting [12].

Composite model. The composite model refers to a combination of forecasts from both the regression and ARIMA models. The advantage of the composite model is that, in most cases, it outperforms any of the individual forecasts. The composite model (combining regression and ARIMA models) has been well documented in previous researches [13]. This model can also be explained as follows: In case y_t is forecasted, it would include the independent variables:

$$Y_t = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_p X_p + \varepsilon_t.$$

This equation includes an additive error term that accounts for unexplained variance in y_t, that is, it accounts for that part of the variance of y_t which is not explained by x_1, x_2, ..., x_p. The above equation can be estimated using a regression analysis, as one source of forecast error would come from the additive noise term whose future values cannot be predicted. An effective application of time series analysis is the construction of an ARIMA model for the residual series ε_t of this regression. This step is followed by substituting the ARIMA model for the implicit error term in the original regression equation. This application also enables the forecasting of the error term ε_t using the ARIMA model. The ARIMA model provides information about the likely future values of ε_t. It helps to "explain" the unexplained variance in the regression equation. The combined regression-time series model is given as follows:

$$Y_t = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_p X_p + \varnothing^{-1}(B)\theta(B)\eta_t,$$

where Y_t is the dependent variable, $X_1, X_2, ..., X_p$ are the independent variables, $\beta_0, \beta_1, \beta_2, ..., \beta_p$ are the regression parameters, $\varnothing$ and θ are AR and MA parameters, and η_t is the error random variable.

SSES method. This method is suitable for examining data series with no trend and a seasonal effect. It suits data that are constant over time. Thus, level and season are the smoothing parameters of this model. The following equations can illustrate this model:

$$L(t) = \alpha(Y(t) - S(t-s)) + (1-\alpha)L(t-1),$$

$$S(t) = \delta(Y(t) - L(t)) + (1-\delta)S(t-s),$$

$$\hat{Y}_t(k) = L(t) + S(t+k-s).$$

Based on the abovementioned equations that considered the last available estimated level state, $\hat{y}$ is forecasted, followed by the last available smoothing seasonal factor. In addition, for $S(t+k-s)$, matching the month of the forecast horizon is added. Hence, the ARIMA model equivalent to the seasonal exponential smoothing model is the ARIMA $(0, 1, p+1)(0, 1, 0)p$, where p stands for the number of periods in a seasonal interval (for monthly data, $p = 12)$ with restrictions [14, 15].

Results and Discussion

Results. The study reported in this paper compares all abovementioned models following various steps to test the significance of the estimated parameters and measures the forecasting error.

First step - Testing the significance of the estimated parameters. This step tested the significance of the estimated coefficients of Malaysia's imports of MTE models.

Based on the results, particularly the *p*-values for the estimated coefficients of all Malaysia's imports of MTE in Tables 2 to 7, the estimated coefficients were significant at different levels.

Table 2. Results of testing the significance of the estimated parameters of the RM-WOP

Model	Parameters	Estimate	Standard error	*t*-value	Significance value
RM-WOP	β_0	14058	2174.000	6.47	0.000
	β_3	0.099	0.011	9.39	0.000
	β_4	0.474	0.037	12.91	0.000
	β_5	–18646.000	4937.000	–3.78	0.000

Source: Own data calculations

Table 3. Results of testing the significance of the estimated parameters of the RM-WPA

Model	Parameters	Estimate	Standard error	*t*-value	Significance value
RM-WPA	β_0	5536	1040	5.321	0.000
	β_3	0.168	0.018	9.190	0.000
	β_4	0.246	0.042	5.870	0.000
	β_5	–14567	6828	–2.133	0.036

Source: Own data calculations

Table 4. Results of testing the significance of the estimated parameters of the ARIMA model

Model	Parameters	Estimate	Standard error	t-value	Significance value
ARIMA model $(0, 1, 0)(1, 1, 1)$	Seasonal $\varnothing_1$	–0.326	0.113	–2.893	0.005
	Seasonal θ_1	0828	0086	9.636	0000

Table 5. Results of testing the significance of the estimated parameters of the CM-WORP

Model	Parameters	Estimate	Standard error	t-value	Significance value
CM-WORP	β_0	14058	2174.000	6.47	0.000
	β_3	0.099	0.011	9.39	0.000
	β_4	0.474	0.037	12.91	0.000
	β_5	–18646.000	4937.000	–3.78	0.000
	AR	0.412	0.095	4.324	0.000

Source: Own data calculations

Table 6. Results of testing the significance of the estimated parameters of the CM-WRPA

Model	Parameters	Estimate	Standard error	t-value	Significance value
CM-WRPA	β_0	5536	1040.000	5.321	0.000
	β_3	0.168	0.018	9.190	0.000
	β_4	0.246	0.042	5.870	0.000
	β_5	–14567	6828.000	–2.133	0.036
	MA	0.847	0.071	11.982	0.000
	MA, seasonal	0.980	0.394	2.491	0.015

Source: Own data calculations

Table 7. Results of testing the significance of the estimated parameters of the SSES

Model	Parameters	Estimate	Standard error	t-value	Significance value
SSES	Alpha α (level)	0.900	0.099	9.119	0.000
	Delta δ (season)	1.000	1.040	0.961	0.339

Source: Own data calculations

Tables 2 to 7 present the p-values for the estimated coefficients of all Malaysia's imports of MTE. The p-values for the estimated coefficients of the RM-WOP, ARIMA model, and the CM-WORP are less than 0, indicating that they are highly significant. The p-values for β_5 in the RM-WPA and the CM-WRPA are equal to 0.036. Such results also indicate that the estimated coefficients are highly significant. However, the p-value for at least one of the estimated coefficient of the SSES model is greater than 0.1. This finding implies that the value is insignificant at the 0.05 level. Thus, the authors drop it and do not consider it in the next step.

Second step - Measuring the forecast error. As shown in Table 8, the five models, namely, the RM-WOP, ARIMA model, RM-WPA, CM-WORP and CM-WRPA are compared. The comparison among these models focused on various measures of error. The results of the forecasting performance of these five models are summarized in Table 8 below:

Table 8. Statistical measures of forecast error for the Malaysia's imports of MTE

Models	R^2 (prediction)	U
ARIMA model	0.96829	0.00173
RM-WOP	0.96170	0.00179
CM-WORP	0.96835	0.00148
RM-WPA	0.82270	0.01400
CM-WRPA	0.82110	0.98610

Source: Own data calculations

The results of the comparison among the five abovementioned models in terms of their forecasting performance in Table 8 include the U test and predicted R^2. Figure 3 also illustrates the results generated from such robustness assessment of the different methods. The forecasting performance of these different models upon further comparison is indicated by each bar that shows the number of best forecasts obtained by its corresponding model in terms of a specified accuracy measure.

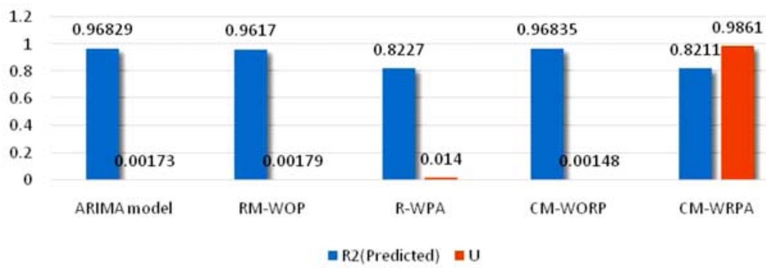

Figure 3. Results of the comparison of the forecasting performance among the different models.

The authors interpreted and discussed the relevant issues according to the results illustrated in Figure 3.

Discussion. This section provides a detailed discussion of the abovementioned results concerning the forecasting performance of the proposed approach. Initially, it discusses the performance as indicated by the abovementioned results obtained from the previous experiments, and then it provides additional insights into this performance by highlighting that of the composite model.

Performance comparison and analysis. The previously stated experimental design and methodologies employed in the current study aimed to show the experimental forecasts of the imports of MTE. Subsequently, the forecasting performance of the models was evaluated by testing the significance of the estimated parameters and two main measurement criteria.

The results presented in Table 8 revealed that the U and predicted R^2 values of CM-WORP are 0.00148 and 0.96835, respectively, for the time series of the imports of MTE. Such results clearly indicate that whereas the U value is lower than those of other methods, the predicted R^2 value is higher than those of other methods. A further comparison of all models used in the present study shows that the CM-WORP achieved the best performance among all models because its fit was the best. Such a result corroborates the one reported in a previous study showing that this model performed better [16]. The result of this earlier study was obtained by comparing the CM-WORP and the CM-WRP. However, the study found that the CM-WORP was capable of solving the problem of autocorrelation in residual. Moreover, it can increase the level of accuracy in forecasting. Such result is also consistent with other earlier studies [8, 17-21]. All these previous studies provided evidence of the contribution of this CM-WORP to the forecasting accuracy because CM-WORP increases forecasting accuracy.

Additional insights into using the composite model. The results listed in Tables 2 to 8 present a remarkable conclusion. The most remarkable result obtained in the current study is that using the CM-WORP can solve the emergent problem of autocorrelation, which is not solved by the regression model.

Conclusions

The present study proposed and evaluated a model for forecasting the imports of MTE in Malaysia. The proposed model, the CM-WRPA, was evaluated by comparing it to other methods that were made based on the time series of MTE in Malaysia. This study has a useful contribution to the literature because it represents the first empirical study that applied the CM-WRPA in this research area. The results provided evidence of the importance and value of such composite model as a robust forecasting method that improves or increases the accurate prediction of the value of imports of MTE and enhances forecasting methods in the Malaysian context. As observed from the results, when the problem of autocorrelation emerged, it was solved or processed only with the CM-WORP.

Acknowledgements

The authors thank the Faculty of Science and Technology, University Science Islam Malaysia (USIM) for the facilities they provided during this work. The authors also wish to extend their gratefulness and appreciation to the reviewers for their valuable suggestions.

References

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

[2] S. D. A. Applanaidu, F. M. Arshad, M. N. Shamsudin and A. A. A. Hameed, An econometric analysis of the link between biodiesel demand and Malaysian palm oil market, International Journal of Business and Management 6(2) (2011), 35-45.

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

[4] J. Mohamad, The impact of tariff reductions on real imports in Malaysia from 1980-2010: an empirical study, Journal of the Graduate School of Asia - Pacific Studies (2012), 181-199.

[5] M. Milad, R. Ibrahim and S. Marappan, Regression analysis to forecast Malaysia's imports of crude material, Int. J. Manage. Appl. Sci. 1 (2015), 121-130.

[6] L. I. B. Osman, Best Fitted Model to Forecast the Trade Balance of Malaysia, Faculty of Computer Science and Mathematics, UiTM, 40450 Shah Alam, Selangor, 2012.

[7] A. Shabri, R. Samsudin and Z. Ismail, Forecasting of the rice yields time series forecasting using artificial neural network and statistical model, Journal of Applied Sciences 9(23) (2009), 4168-4173.

[8] M. A. Milad and R. I. Ibrahim, A robust composite model approach for forecasting Malaysian imports: a comparative study, Journal of Applied Sciences 16(6) (2016), 279-285.

[9] R. M. Leuthold, On the use of Theil's inequality coefficients, American Journal of Agricultural Economics 57(2) (1975), 344-346.

[10] O. S. Ardekani and S. R. Shadizadeh, Development of drilling trip time model for southern Iranian oil fields: using artificial neural networks and multiple linear regression approaches, Journal of Petroleum Exploration and Production Technology 3(4) (2013), 287-295.

[11] G. K. Uyanık and N. Güler, A study on multiple linear regression analysis, Procedia-Social and Behavioral Sciences 106 (2013), 234-240.

[12] R. Tripathi, A. K. Nayak, R. Raja, M. Shahid, A. Kumar, S. Mohanty, B. B. Panda, B. Lal and P. Gautam, Forecasting rice productivity and production of Odisha, India, using autoregressive integrated moving average models, Advances in Agriculture 2014 (2014), Article ID 621313, 9 pp.

[13] R. S. Pindyck and D. L. Rubinfeld, Econometric Models and Economic Forecasts, Irwin/McGraw-Hill, 1998.

[14] I. Corporation, IBM SPSS statistics for Windows, version 21.0, IBM Corporation Armonk, New York, 2012.

[15] T. B. Fomby, Exponential smoothing models, 2008. Retrieved from Southern Methodist University website: http://www.google.com/url.

[16] M. A. Milad, R. I. Ibrahim and S. Marappan, A comparison among two composite models (without regression processing) and (with regression processing), applied on Malaysian imports, Appl. Math. Sci. 9 (2015), 5757-5767.

[17] R. S. Pindyck and D. L. Rubinfeld, Econometric mod Analysis of Cointegration Vectors Using the GMM Approach els and Economic Forecasts, Irwin/McGraw-Hill, Boston, Vol. 4, 1998.

[18] M. N. Shamsudin and F. Mohd Arshad, Composite models for short term forecasting for natural rubber prices, Pertanika 13(2) (1990), 283-288.

[19] A. Islam, Explaining and forecasting investment expenditure in Canada: combined structural and time series approaches, 1961-2000, Applied Econometrics and International Development 7(1) (2007), 76-89.

[20] A. A. Khin, F. C. Eddie Chiew, M. N. Shamsudin and Z. A. Mohamed, Natural rubber price forecasting in the world market, Agricultural Sustainability through Participative Global Extension (AGREX 2008), Kuala Lumpur, Universiti Putra Malaysia, 2008.

[21] A. K. Aye, Z. Mohamed and M. N. Shamsudin, Comparative forecasting models accuracy of short-term natural rubber prices, Trends in Agricultural Economics 4(1) (2011), 1-17.