iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 8 Issue: VIII Month of publication: August 2020

DOI: https://doi.org/10.22214/ijraset.2020.31160

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 8 Issue VIII Aug 2020- Available at www.ijraset.com

Time Series Forecasting of Malaysia Producer Price Index using ARIMA and Grey Models

Azme Bin Khamis[1], Tay Li Xin[2]
[1, 2]Department of Mathematics and Statistics, Faculty of Applied Sciences and Technology, Universiti Tun Hussein Onn Malaysia, Johor, Malaysia

Abstract: The continuous increment inflation rate was seen as a popular topic in Malaysia. Producer Price Index (PPI) is an index to measure the average change in prices of goods and services sold by producers. Therefore, PPI is used to estimate the percentage change or inflationary pressure in producer price in this study. The aim of this study is to build the forecasting models of Malaysia PPI by using ARIMA model and GM (1,1) model, identify the most appropriate model to forecast one year ahead PPI and estimate the percentage change in producer price by using predicted PPI as a measure. Monthly Malaysia PPI with the base year 2010 from January 2013 to July 2019 is retrieved from Bank Negara Malaysia (BNM) and used in this study. Two forecasting methods have been applied in this study which are ARIMA model and GM (1,1). The results show that ARIMA (1,2,0) is outperformed compared with GM_5 (1,1) model based on MAPE. Hence, this model is chosen to forecast the volume index of producer price for the coming year. The negative values in the percentage change of producer price for the local production reflected the deflationary pressure that might be faced by Malaysia in the coming year.
Keywords: Malaysia Producer Price Index, Forecasting, Inflation rate, ARIMA, Grey model

I. INTRODUCTION

The continuous increment inflation rate in Malaysia and over the world was seen as a popular topic in recent years. The Producer Price Index (PPI) is a key indicator of the economic stability of a country as it measures the average change in selling prices of goods and services received by producers and manufacturers during a given period [1]. According to Department of Statistics Malaysia, the overall PPI index for local production in Malaysia was contributed by the sectors Agriculture, forestry & fishing, Manufacturing, Mining, Water supply and Electricity & gas supply [2]. Malaysia economy experienced high episodes in the years 1973 to 1974 and 1980 to 1981 and low inflation rate in the year 1985 to 1987. Malaysia able to maintain a low and stable inflation rate during the high economic growth period caused Malaysia has a unique experience involve in inflation [3]. Meanwhile, [4] stated that CPI is one of the most common communication methods used by the researchers to measure inflation rate nationwide such as discussed by researcher [5], [6] and [7]. However, the theory of the production chain argued that the consumer prices are calculated based on the prices of production. Hence, the inflation through the price indices in the production could be transmitted to inflation through consumer price index. Furthermore, [8] has the same viewpoint with [4] which stated that the PPI as a short-term leading indicator of final price changes at the consumer level and subsequently act as a leading inflation index.

High production of raw materials and semi-finished products prices will increase the overall price level and eventually causes inflation. Hence, PPI can serve as a short-term leading indicator of the ultimate price changes at consumer level and subsequently as a leading inflation index [8]. PPI can used as an analytic tool for tracking economic trends and act as an information to the future inflation by overcoming the challenges to control inflation in Malaysia [9]. Forecasting PPI can assist economists to predict the future CPI movement and eventually provides indicators to economist researchers as well as government towards the inflation in Malaysia as Malaysia has a unique experience involved inflation [10]. The low inflation rate is beneficial to motivate consumer spending, corporate profits and finally simulated the stock market [8].

Forecasting uncertainty phenomenon are being great interest for many researchers. Many researchers have involved in socio-economic research or financial time series forecasting such as [11], [12], [13], [14], [15] and [16]. A comparison studied of the ARIMA model and Fourier Series model on time series forecasting for Consumer Price Index (CPI) in Nigeria [6]. ARIMA models and Holt's exponential smoothing method are implemented by [7] to analyse the movement of CPI in Zambia and subsequently used to predict the inflation rate of Zambia. Research done by [17] conducted comparative study on the accuracy of the forecasting model of gold price including original GM (1,1), GM (1,1) model using Fourier series, (FGM(1,1)) and ARIMA model, meanwhile [18] point out that there are two reasons ARIMA model is chosen for the comparison with grey model.

The aim of the study is to develop the forecasting models of Malaysia Producer Price Index (PPI) by using ARIMA model and grey prediction model also known as GM (1,1). Besides, to identify the most appropriate forecasting model by comparing the accuracy of several forecasting methods in order to obtain the most precise data prediction. Furthermore, to determine Malaysia Producer Price inflationary pressure or percentage change by using forecasted PPI as a measure. This is because lack of robust research concerning to used PPI as a relevant indicator for assessing inflationary pressure.

The data used in this study is the monthly PPI in Malaysia from January 2013 to July 2019 with a base year of 2010 as presented in Figure 1. The dataset consists of total 79 observation. For ARIMA model, the data set was split in a ratio of 75:25 which the first 60 data points, from January 2013 to December 2017 for training purpose; whereas the data from March 2017 to December 2017 (4 to 10 sample size) are used to develop GM (1,1) models with different sample sizes. The remaining 19 data points is used for the testing purpose from January 2018 to July 2019 for both of the methods.

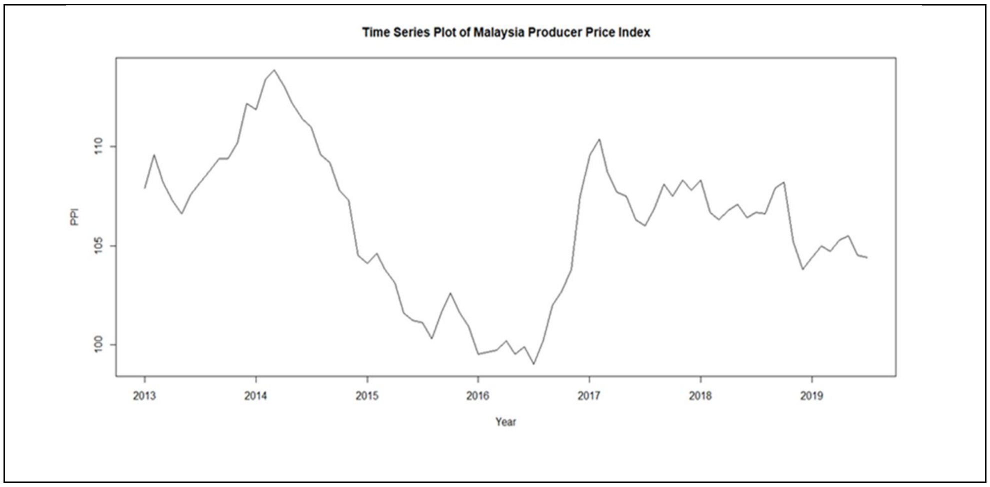


Figure 1: Time Series Plot of Malaysia Monthly Producer Price Index from January 2013 to July 2019

II. METHODOLOGY

Two forecasting approaches: ARIMA model and grey model are used in this study. The reasons that these two methods chosen is due to the ARIMA model as a conventional forecasting model which produces more reliable and accurate forecasts. Secondly, ARIMA model and grey model can be directly compared on same base.

A. *Autoregressive Integrated Moving Average (ARIMA) Model*

According to [19], a time series ARIMA model is produced by Box and Jenkins for the series is non-stationary and differencing is recommended to achieve stationary with the combination of AR and MA processes. The AR components represent weighted sum of the past values, MA components for the weighted sum of pass errors and I stands for integrated. The full ARIMA (*p*,*d*,*q*) model can be written as follows:

$$y'_t = c + \phi_1 y'_{t-1} + \cdots + \phi_p y'_{t-p} + \theta_1 e_{t-1} + \cdots + \theta_q e_{t-q} + e_t \qquad (1)$$

where y'_t is the differenced series,

ϕ is the parameter estimate AR components,

θ is the parameter estimate MA components, and

e are independent and identically distributed normal error with zero mean

The Box-Jenkins approach referring to three steps iterative method: model identification, parameter estimation, and diagnostic checking is used to build the ARIMA model [20]. For model identification, the appropriate order of the tentative models is identified through visual inspection of both ACF and PACF plot. The statistical Augmented Dickey Fuller (ADF) test and Kwiatkowski-Philips-Schmidt-Shin (KPSS) test are used to test the stationarity of the PPI series.

The parameters ϕ and θ of the selected model are estimated using maximum likelihood techniques in second step. The null hypothesis is let any parameters in ARIMA model equal to zero. If the p-value is less than significance level of 0.05, the null hypothesis is rejected and conclude that the parameter is significant in the ARIMA model. The final step in Box-Jenkins approach is the diagnostic checking where the residuals of the selected model is tested against the model assumptions once a model has been fitted. The Ljung-Box test is a diagnostic tool used to determine the lack of fit of the time series model. The null hypothesis represent that the model is adequate which the model does not exhibit lack of fit. If the p-value of Ljung Box statistics is larger than significance level of 0.05, the null hypothesis is not rejected thus the model is adequate.

B. *Grey Forecasting Model*

Grey system theory was invented by Deng in 1985 to catch the system development tendency in order to overcome the limitation of information and small sample size problem by estimating the interrelation of the data [21]. Since only as few as 4 sample size are needed to build a GM (1,1) model, the optimal sample size for grey model development is determined. The most appropriate GM (1,1) model is chosen by comparing the MAPE of all the GM (1,1) with different sample sizes. The grey modelling algorithm can be summarized in a few of steps:

1) *Step 1:* The original raw data sequence, $x^{(0)}$ is assumed with *n* samples:

$$x^{(0)} = \left[x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n)\right] = [x^{(0)}(k)] \quad (2)$$

where *n* is the optimal sample size of the data.
$x^{(0)}$ is the non-negative original historical time series data

2) *Step 2:* First-order accumulated generation operator (1-AGO) is defined as follows which used to transform the original sequence $x^{(0)}$ in the grey system into a new sequence $x^{(1)}$ for the purpose of identifying the potential pattern and trend of the original data sequence [10].

$$x^{(1)}(k) = \sum_{m=1}^{k} x^{(0)}(m), k = 1,2,3, \dots, n \quad (3)$$

$$\frac{dx^{(1)}}{dt} + ax^{(1)} = b \quad (4)$$

3) *Step 3:* The first-order differential equation that shown in Equation (4) represents the approximation data:

where $\frac{dx^{(1)}}{dt}$ is the derivative of the function x
parameter *a* is the development coefficient
parameter *b* is the grey input

4) *Step 4:* The equation is reduced to Equation (5) by discretization Equation (4) for the purpose of solving the model parameters *a* and *b*.

$$\frac{dx^{(1)}}{dt} = x^{(1)}(k+1) - x^{(1)}(k) = x^{(0)}(k+1) \quad (5)$$

The generated sequence of the consecutive neighbour of predicted value $x^{(1)}$ is defined as Equation (6) by setting the generation coefficient of 0.5 in the original model.

$$z^{(1)}(k) = 0.5x^{(1)}(k) - 0.5x^{(1)}(k-1) \quad (6)$$

A grey differential model is then obtained and shown in Equation (7).

$$x^{(0)}(k) + az^{(1)}(k) = b, \qquad k = 1,2, \dots, n, \dots \quad (7)$$

5) *Step 5:* Least square method is used to solve parameters *a* and *b* which stated in Equation (7).

$$\hat{a} = \begin{bmatrix} a \\ b \end{bmatrix} = (B^T B)^{-1} B^T Y_N \quad (8)$$

6)

where $Y_N = \begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \vdots \\ x^{(0)}(n) \end{bmatrix}$, $B = \text{data matrix} = \begin{bmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ \vdots & \vdots \\ -z^{(1)}(n) & 1 \end{bmatrix}$

7) *Step 6:* The particular solution is given approximately by

$$\hat{x}^{(1)}(k+1) = \left(x^{(0)}(1) - \frac{b}{a}\right) e^{-ak} + \frac{b}{a}, k = 1,2, \dots, n \quad (9)$$

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 8 Issue VIII Aug 2020- Available at www.ijraset.com

where $\hat{x}^{(1)}(k+1)$ is the forecasted value of $x^{(1)}(k+1)$ at time $(k+1)$

8) *Step 7:* The forecasted value $\hat{x}^{(0)}(k+1)$ can be obtained by using the inverse accumulated generation operator (I-1AGO) as follow:

$$\hat{x}^{(0)}(k+1) = \hat{x}^{(1)}(k+1) - \hat{x}^{(1)}(k) \qquad (10)$$

C. *Performance of Forecast*

In this study, the results of mean absolute percentage error (MAPE) is used to compare the forecast accuracy of the models. The model that gained the smallest values in MAPE error magnitude measurement will be chosen as most appropriate forecasting model in this study.

D. *Inflationary Pressure Determination*

The inflationary pressure was subsequently calculated by using forecasted PPI as follows:

$$I_t = \frac{PPI_t - PPI_{t-1}}{PPI_{t-1}} \times 100 \qquad (11)$$

where PPI_t is forecast PPI, PPI_{t-1} is actual PPI a year before and I_t is inflation rate at time t

III. RESULTS AND DISCUSSIONS

A. *ARIMA Model*

The incidence data of Malaysia PPI from January 2013 to December 2017 shows a non-stationary trend with time. The Box-Cox transformation and differencing are applied in order to ensure the dataset follow the stationary assumption. The obtained lambda value of Box-Cox equal to -0.50 indicates an inverse of square root transformation is needed to stabilize the variance of the time series. Double differencing is done to stabilize the mean of the PPI time series since trend and level changes exist in times series. ARIMA models are build based on ACF and PACF plots.

According to Table 1, ARIMA (1,2,0) and ARIMA (0,2,1) have significant estimated parameters as the p-values in these two ARIMA models are less than significance level of 0.05. Based on the Ljung-Box statistics, the p-value of these two ARIMA models greater than 0.05 indicates that there is no autocorrelation in the residuals and the models are adequate. Based on the obtained MAPE, ARIMA (1,2,0) model is the most appropriate ARIMA model as it obtained the lowest values in error measurements. Hence, ARIMA (1,2,0) model is then compared with the optimal GM (1,1) model.

Table 1: Summary of parameters estimation and diagnostic checking of ARIMA models.

No.	Box-Jenkins Model ARIMA(p,d,q)	Parameters Estimation		p-value of Ljung-Box Statistics	MAPE in testing set
		Coefficient	p-value		
1.	ARIMA (1,2,0)	$\widehat{\phi_1}$ = -0.37298	0.003612	0.3136	0.3985454
2.	ARIMA (0,2,2)	$\widehat{\theta_1}$ = -0.63288 $\widehat{\theta_2}$ = -0.22874	0.000003 0.1249	0.6623	1.1055840
3.	ARIMA (0,2,1)	$\widehat{\theta_1}$ = -0.70328	0.000006	0.3185	1.1602904

B. *GM (1,1) Model*

Grey prediction method is employed in Malaysia PPI data from March 2017 to December 2017 to develop the GM (1,1) model with different sample size. Equation (2) to Equation (8) are used to determine the parameters, development coefficient, a and grey variable, b for different GM_n (1,1) model. The forecast values of Malaysia PPI at time $k + 1$ for grey model are computed using Equation (9) and (10). Table 2 shows the computed MAPE for training set and test set in order to evaluate the forecasting model performance. Based on the obtained results in Table 2, the optimal sample size of 5, GM_5 (1,1) produced the lowest MAPE (1.82%).

C. Comparative Study

The ARIMA (1,2,0) model is then compared with GM_5 (1,1) model in Malaysia PPI prediction. Figure 2 shows the fitting and forecasting curves of these two models. ARIMA (1,2,0) model has outperformed than GM_5 (1,1) model as ARIMA (1,2,0) model has the lower value of MAPE (0.3985%) compared with GM_5 (1,1) model (1.82%). Therefore, ARIMA (1,2,0) model is used to forecast one year ahead of Malaysia PPI from and subsequently used the forecasted values to determine the percentage change in producer price or inflationary pressure.

The negative values of PPI for the local production based on year-on-year measurement reflected the deflationary pressure that might be faced by Malaysia on the next coming year as compared to the same month of the preceding year as shown in Figure 3

Table 2: Fitted and forecast value and MAPE of $GM_n(1,1)$ with different sample sizes

Sample size	$GM_{10}(1,1)$	$GM_9(1,1)$	$GM_8(1,1)$	$GM_7(1,1)$	$GM_6(1,1)$	$GM_5(1,1)$	$GM_4(1,1)$
Fitted Values in Training Set (March 2017 – December 2017)							
MAPE	0.53691	0.97499	0.88955	0.36631	0.35649	**0.25023**	0.26759
Forecast Values in Test Set (January 2018 – December 2018)							
MAPE	2.95998	4.01119	5.58910	5.73464	3.94904	**1.81776**	3.37586

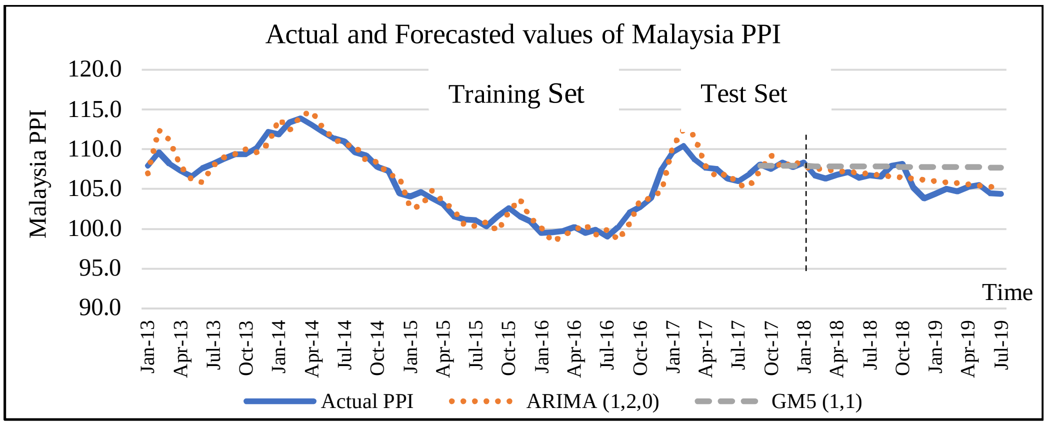


Figure 2: The observed, fitting and forecasting PPI by ARIMA and GM (1,1)

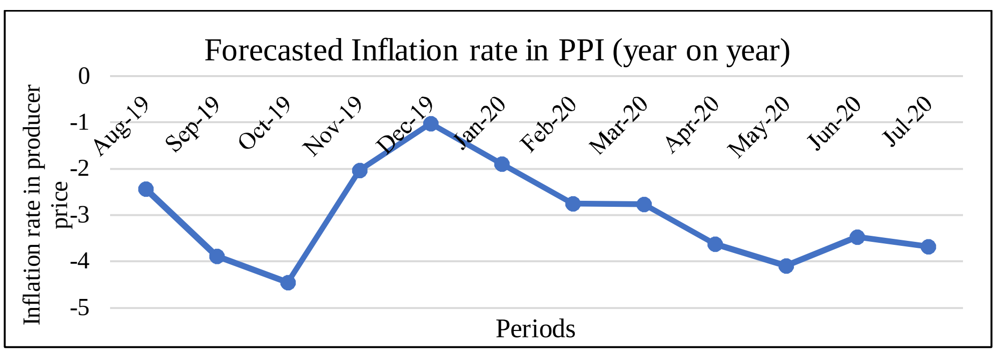


Figure 3: Forecasted inflation rate (year on year) based on forecasted PPI variation

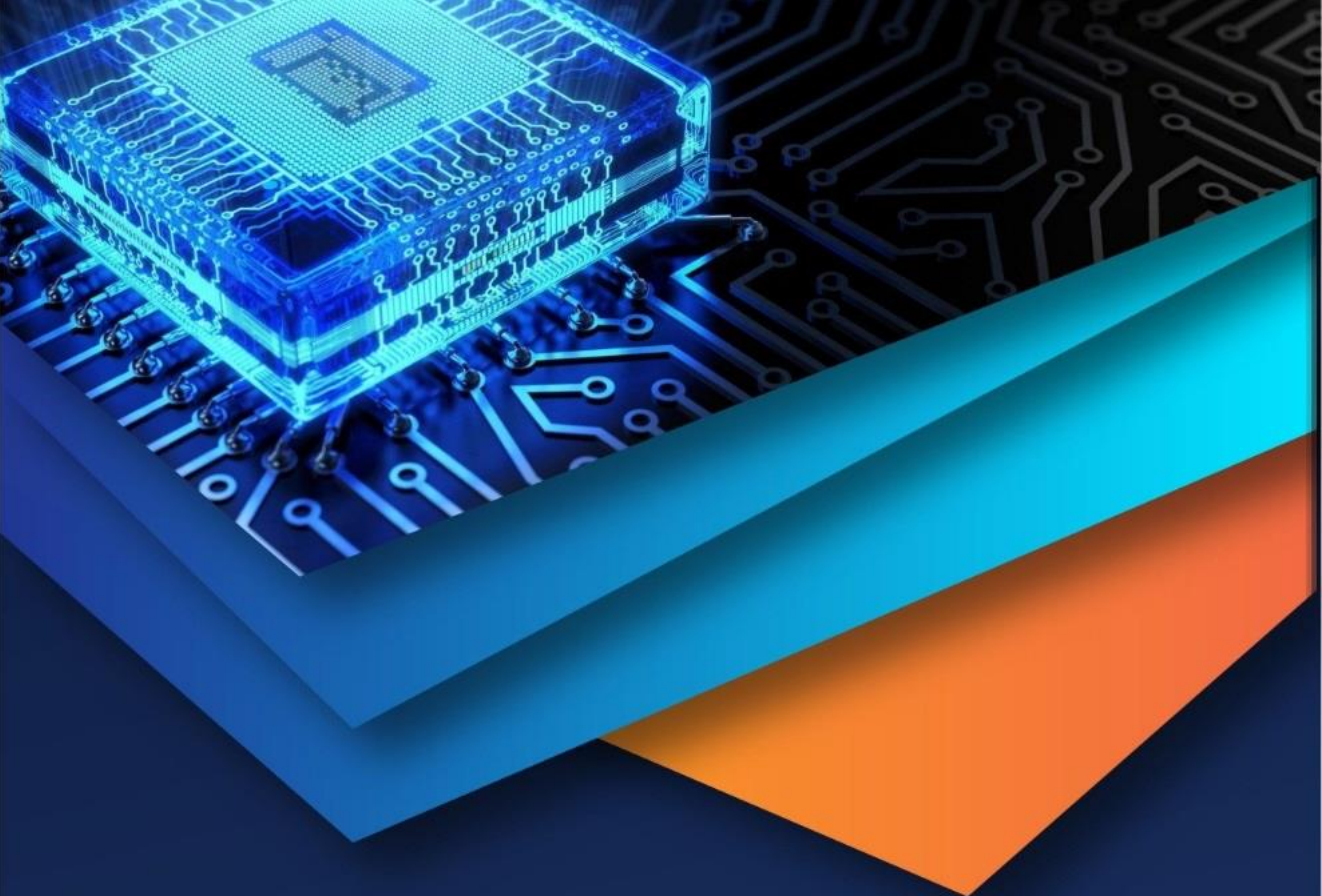
IV. CONCLUSION

Maintain a low and stable inflation rate has become one of the challenges in the macroeconomic management of Malaysia. PPI is an index that used to measure the average change in prices of goods and services sold by manufacturers in the wholesale market during a given period. In this study, a comparative study between ARIMA (1,2,0) model and GM_5 (1,1) model is done. ARIMA (1,2,0) model provided a more accurate result than GM_5 (1,1) model since it obtained the lower MAPE value (0.3985%). Hence, ARIMA (1,2,0) model is used to forecast one year ahead Malaysia PPI and the forecasted value is subsequently used as a measure for determining the percentage changes or inflationary pressure in PPI.

The negative values show in percentage changes indicates that the deflationary pressure that might be faced by Malaysia on the next coming year. Multivariate time series analysis is suggested to be done to determine the inflation rate in Malaysia if other information is provided as this study involved only the univariate time series analysis – Malaysia PPI. Besides, machine learning techniques can be used for further study if the information of the dataset is large enough.

REFERENCES

[1] S. Neda. Multivariate Time Series Analysis of Inflation: The Case of Ethiopia. Master Thesis. Addis Ababa University; 2011
[2] Producer Price Index Malaysia September 2019. Retrieved on October 31, 2019, from https://www.dosm.gov.my/v1/index.php?r=column/cthemeByCat&cat=107&bul_id=anI5NXFtYmdlOHBmTERzUDltQ1VUZz09&menu_id=bThzTHQxN1ZqMVF6a2I4RkZoNDFkQT09
[3] R.Murdipi, , & S. H. Law. Dynamic Linkages between Price Indices and Inflation in Malaysia. Jurnal Ekonomi Malaysia, 50(1), 41-52, 2016.
[4] R. C. Vilcu. Inflation by Producer Price Index-predictive factor for inflation by Consumer Price Index? The case of Romania. Romanian Statistical Review, Supplements, (22), 2015.
[5] R. F. Zafar, A. Qayyum & S.P.Ghouri. Forecasting Inflation using Functional Time Series Analysis. MPRA Munich Personal RePEc Archive. 2015.
[6] E. E. Umanah. Time Series Modeling of All Items Consumer Price Index in Nigeria. Master Thesis. University of Nigeria. 2010.
[7] S. Jere & M. Siyanga. Forecasting Inflation Rate of Zambia Using Holt's Exponential Smoothing. Open Journal of Statistics, 06(02), 363–372, 2016.
[8] L.W. M. Then. The Relationship Between Consumer Price Index (CPI) and Producer Price Index (PPI) in Malaysia Ph.D. Thesis Universiti Malaysia Sarawak; 2011.
[9] Okyere, F., & Kyei, L. Temporal Modelling of Producer Price Inflation Rates of Ghana. IOSR Journal of Mathematics. 10(3); 70–77, 2014.
[10] J.Y.M.Wong, Y.T. Lim, H.Y. Loke & J.J. Tai. Effect of Macroeconomic Variables toward Inflation in Malaysia's Economy. Degree Thesis. University Tunku Abdul Rahman; 2017.
[11] L. Zhang & J.Li. Inflation forecasting using support vector regression. Proceedings of the 2012 4th International Symposium on Information Science and Engineering, ISISE 2012, pp.136–140, 2012.
[12] T. Nyoni. ARIMA modeling and forecasting of Consumer Price Index (CPI) in Germany. https://mpra.ub.uni-muenchen.de/92414/1/MPRA_paper_92414.pdf, 2019.
[13] A. Phatchakorn, S. Tomonobu, T. Hirofumi, & Y. Atsushi. A Hybrid ARIMA and Neural Network model for short term Price Forecasting in Deregulated Market. IEEE Trans. on Power Systems.2010.
[14] S. J. Rani, V. V. Haragopal, & M. K. Reddy. Forecasting Inflation Rate of India using Neural Networks. International Journal of Computer Applications, 158, 5. 2017.
[15] G. Grudnitski & L. Osburn. Forecasting S&P and gold futures prices: An application of neural networks. Journal of Futures Markets, 13(6), 631–643, 1993.
[16] S. Telli & M. Coşkun. Forecasting the BIST 100 Index Using Artificial Neural Networks with Consideration of the Economic Calendar. International Review of Economics and Management, 4(3), 26–46, 2016.
[17] M. Askari & H. Askari. Time series grey system prediction-based models: Gold price forecasting. Trends in Applied Sciences Research, 6(11), 1287-1292, 2011.
[18] P. H. Ho. Comparison of the grey model and the Box–Jenkins model in forecasting manpower in the UK construction industry. In Proceedings of the 28th Annual ARCOM Conference, 369-379, 2012.
[19] R. J. Hyndman & G. Athanasopoulos. *Forecasting: principles and practice*, 2nd edition, OTexts: Melbourne, Australia. OTexts.com/fpp2, 2018.
[20] N. Habimana, A. Wanjoya & A.Waititu. Modeling and Forecasting Consumer Price Index: Case of Rwanda. American Journal of Theoretical and Applied Statistics. 5(3), 101-107, 2016.
[21] J. L. Deng. Introduction to grey system theory. The Journal of grey system. 1(1), 1-24, 1989.



10.22214/IJRASET

INDEX COPERNICUS 45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY