11th MALAYSIA STATISTICS CONFERENCE 2024
Data and Artificial Intelligence: Empowering the Future
Sasana Kijang, Bank Negara
19th September 2024

Predictive Analysis of Malaysia's Food Price Index: An ARIMA Model Analysis

Wan Ameerah Fatini Wan Razali[1]; Syamimi Izzati Saifol Bakhri [2]; Noreha Mohamed Yusof[3]

[1, 2, 3] College of Computing, Informatics and Mathematics Universiti Teknologi MARA Negeri Sembilan Seremban Campus, 70300 Negeri Sembilan
[*] noreh144@uitm.edu.my

Abstract:
Inflation, a critical economic indicator, is often measured by the Consumer Price Index (CPI). It is known that the CPI of FNAB in Malaysia has consistently remained over 100 since the year 2010 which indicates ongoing food price inflation, which is the primary cause of Malaysia's overall inflation rate due to its significant weightage in the headline CPI. The objective of this study is to forecast the CPI of FNAB in Malaysia for the upcoming 24 months by using ARIMA modelling. The data from the Department of Statistics Malaysia (DOSM) spanning January 2010 to August 2023 was used. The research identifies the ARIMA (1, 1, 2) as the most accurate for predicting the future values of CPI of FNAB 2 years ahead. Moreover, the CPI values are predicted to continue to increase for the next 2 years with values between 151.7325 and 166.5813 with a 95% confidence interval.

Keywords:
CPI of FNAB (Consumer Price Index of food and non-alcoholic beverages); ARIMA; inflation

1. Introduction:
CPI is a great measurement to reflects the inflation of one's country. According to Salwati & Wessel (2021), the Bureau of Labor Statistics (BLS) produces the CPI, which is the most generally used measure of inflation. The primary CPI is designed to measure price changes faced by urban consumers. According to DOSM (2024), the CPI of FNAB has consistently remained over 100 since June 2010, which indicates increasing food price inflation in Malaysia. By having the heaviest weight in the headline CPI (29.5%), FNAB will greatly affect Malaysia's inflation trend (DOSM, 2023). As a result, food price increases in Malaysia have been the primary cause of the country's inflation rate.

This study aims to generate the prediction ARIMA models of CPI of FNAB in Malaysia. Other than that, the purpose of this study is to determine the future values of the CPI of FNAB in Malaysia 2 years ahead. The objectives are achieved by utilising the monthly

CPI of FNAB in Malaysia from January 2010 until August 2023 that can be retrieved from DOSM.

A study done by Radzi et al. (2022), utilized ARIMA and Vector Autoregression (VAR) models to predict Malaysia's GDP and CPI, selecting ARIMA (0,1,1) for predicting the country's GDP and utilized ARIMA (1,1,0) to forecast the CPI for upcoming 5 years, both showing upward trends. Similarly, Mohamed (2020) compared ARIMA models and regression with ARIMA errors, finding that ARIMA (0,1,3) most suitable to forecast CPI values of Somaliland, and it also predicted that the CPI values will continue to rise in upcoming years. In summary, these studies show that ARIMA models are useful for predicting CPI values and trends in across different countries. From these studies, it is shown that ARIMA models is effective for forecasting CPI trends in Malaysia. Hence, to achieve the objectives, ARIMA modelling is applied in this study.

2. Methodology:

2.1. Box-Jenkins Methodology

The Box-Jenkins forecasting model assumes a stationary time series if it has a constant mean, constant variance, and constant autocorrelation. Additionally, a unit root test namely the Augmented Dickey-Fuller (ADF) test should be conducted to further prove that the CPI of FNAB is a stationary series.

In this study, the monthly CPI data for FNAB in Malaysia is utilized. The dataset spans from January 2010 to August 2023, containing 164 data points.Hence, this dataset went through the ADF test to check if it is a stationary time series.

The time plot of Malaysia's CPI of FNAB series autocorrelation function (ACF) and partial autocorrelation function (PACF) will be visually reviewed during the identification step. The ACF of the series will not decrease or exhibit any indication of decay if the series is non-stationary. Should this be the case, then the series of CPI of FNAB must undergo several differencing degrees, *d* for it to be a stationary series. The differencing operation is the "I" in ARIMA, which stands for Integrated. Differencing a series *d* times transforms an I(*d*) non-stationary series into a stationary I(1) series. The first order of differencing can be written as

$$\Delta y_t = \Delta y_t - \Delta y_{t-1}$$
$$w_t = (1 - B)y_t$$

where Δy_t and w_t is the first difference, and *B* is the number of backward steps.

The next step is to determine the ARIMA model's *p* and *q* orders. An autoregressive time series, AR(*p*) model should have the general form of

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + \varepsilon_t$$

with *p* denotes the order of the autoregressive process and thus the number of lagged dependent variables in the model.

On the other hand, for a moving average, MA(*q*) process, follows the general form of

$$Y_t = \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_p \varepsilon_{t-p} + \varepsilon_t$$

An MA(*q*) model is an average of *q* stationary white-noise processes; thus, it follows that every MA(*q*) model is stationary if *q* is finite. Each of the chosen models was estimated and the various coefficients were checked during the estimation stage. The estimated models of Malaysia's CPI of FNAB series will be compared in this stage using the criterion for model selection, the Bayesian Information Criterion (BIC).

Furthermore, it is possible to predict the CPI values of FNAB for the following two years using best ARIMA(*p, d, q*). The forecast value and forecast line of the model can be shown either as a single value or within a confidence interval. The confidence interval gives a measure of the level of certainty and uncertainty underlying the predicted values.

3. Result:

3.1. Model Identification

The data series of CPI of FNAB were initially plotted to examine the overall trend. Figure 1 shows the model of the CPI of FNAB series.

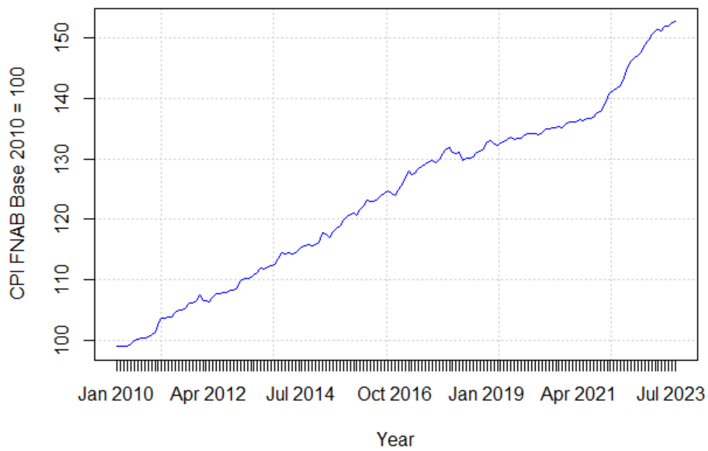

Figure 1: CPI of FNAB from January 2010 to July 2023

Figure 1 exhibits a significant upward trend, indicating the existence of a trend component in the series, and no seasonal pattern was observed in the data. Hence, the data series is non-stationary.

Additionally, the stationarity of the data series can be examined by analysing the patterns in the ACF and PACF. Figure 2 shows the ACF and PACF of the data series.

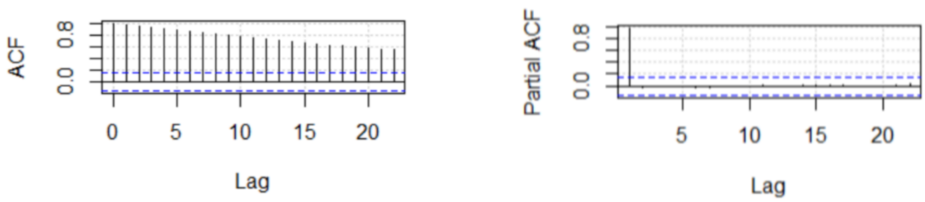

Figure 2: ACF and PACF series of CPI of FNAB

Figure 2 shows an ACF series of CPI of FNAB with a slow decay and a PACF series with a significant spike at lag 1. Thus, it also proved that the series is non-stationary.

An ADF test was conducted to determine if the series has a unit root. The ADF test produced a *p*-value of 0.9904, more than the critical value of α=0.05, thus indicating that the series has a unit root. Since the series contains a unit root, it is further proven non-stationary.

Given that the data series satisfies all the assumptions of non-stationary data, the CPI series of FNAB must undergo differencing until the ACF and PACF plot shows no sign of decay to achieve stationarity. Once stationary, the data series could be applied for forecasting.

ACF and PACF plots after first-order differencing were used to determine whether the series satisfied the stationarity assumption. Figure 3 shows the ACF and PACF after first-order differencing.

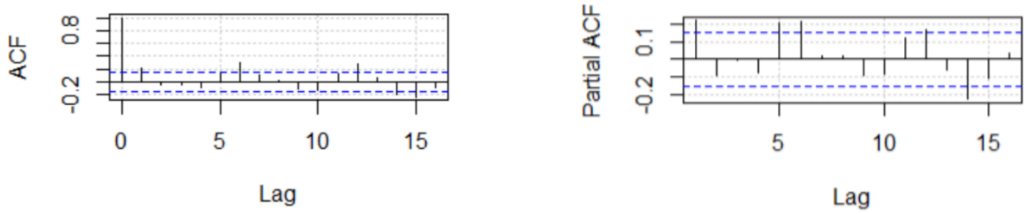

Figure 3: ACF and PACF series of CPI of FNAB after first-order differencing

Figure 3 shows that after first-order differencing, the ACF and PACF plots show no signs of decay. Thus, the series has reached stationarity since the order differencing has been done where d=1.

In addition, an ADF test was used to assess whether the series has a unit root. The ADF test produced a p-value of 8.223962×10^{-15}, lower than $\alpha=0.05$, indicating that the series had no unit root thus it has further proved that the series has reached stationary. The model identification stage could then be executed using ARIMA(p, 1, q).

The lag that exceeded the standard error (SE) band in both the ACF and PACF plots in Figure 3 must be observed to determine the appropriate ARIMA model. The ACF plot shows significant lags at lags 1 and 6, which exceed the SE band. This indicates that the time series contains an MA component of order 2, MA(2). Thus, the model identification would have q=2, which would produce ARIMA(p,1,2).

Meanwhile, the PACF plot in Figure 3 showed 3 significant lags: 1, 5, and 6 thus the time series has an AR component of order 3 AR(3). Hence, the model would have p=3, so it was identified that the appropriate ARIMA model would be ARIMA(3, 1, 2).

3.2. Model Estimation

After identifying the ARIMA(3, 1, 2), four additional ARIMA models were suggested to determine the best model to forecast the CPI of FNAB in Malaysia for the next two years. The four additional ARIMA models were picked in descending order from ARIMA(3, 1, 2).

Following are the suggested ARIMA models:

- ARIMA(3, 1, 2)
  $$w_t = 0.2667w_{t-1} + 0.9203w_{t-2} - 0.1973w_{t-3} - 0.0067\varepsilon_{t-1} - 0.9677\varepsilon_{t-2} + \varepsilon_t$$
- ARIMA(3, 1, 1)
  $$w_t = 0.196w_{t-1} - 0.3318w_{t-2} + 0.1319w_{t-3} - 0.927\varepsilon_{t-1} + \varepsilon_t$$
- ARIMA(2,1, 2)
  $$w_t = 0.6723w_{t-1} + 0.3242w_{t-2} - 0.4013\varepsilon_{t-1} - 0.5181\varepsilon_{t-2} + \varepsilon_t$$
- ARIMA(2, 1, 1)
  $$w_t = 1.2777w_{t-1} - 0.2279w_{t-2} - 0.9899\varepsilon_{t-1} + \varepsilon_t$$

- ARIMA(1, 1, 2)

$$w_t = 0.9999w_{t-1} - 0.7258\varepsilon_{t-1} - 0.2619\varepsilon_{t-2} + \varepsilon_t$$

3.3. Model Evaluation

Five suggested models have been evaluated based on information criterion values, particularly BIC. Table 1 shows the BIC values for the suggested models.

Table 1: BIC values for suggested ARIMA models

ARIMA Model	BIC
ARIMA(3, 1, 2)	229.4446
ARIMA(3, 1, 1)	224.2210
ARIMA(2, 1, 2)	224.1910
ARIMA(2, 1, 1)	220.9796
ARIMA(1, 1, 2)	219.7365

The model with the lowest information criterion would be selected to get the best fit. Table 1 shows that the ARIMA(1, 1, 2) has BIC values of 219.7365. The remaining competing models show slightly different findings in BIC values. Furthermore, when examined alongside ARIMA(2, 1, 1), ARIMA(1, 1, 2) has only three parameters, making it one of the simplest models and thus obeying the parsimony principle. Despite also having the least parameters, ARIMA(2, 1, 1) was rejected due to its larger BIC. Hence, the most fit model for predicting the CPI of FNAB would be ARIMA(1, 1, 2).

3.4. Forecast of CPI of FNAB in Malaysia

The forecast values and model were generated based on ARIMA(1, 1, 2) since it was determined to be the best model to predict the CPI of FNAB for the next two years. Figure 4 shows the CPI forecast model for FNAB for the next two years using ARIMA(1, 1, 2).

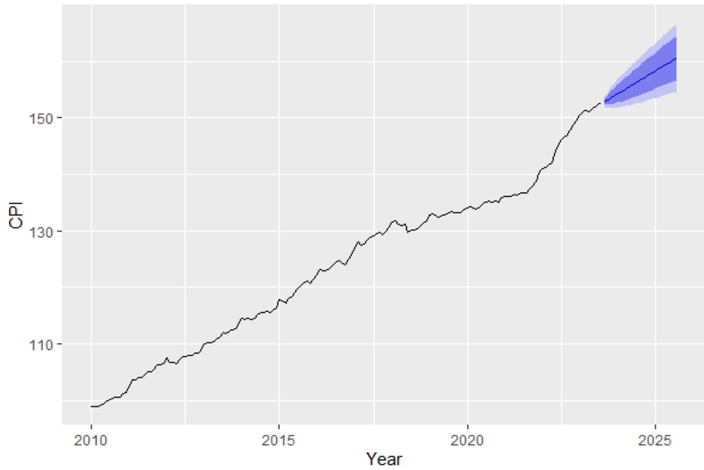

Figure 4: Forecast of CPI of FNAB for 2 years by ARIMA(1, 1, 2)

Figure 4 shows a line graph indicating the time series of the actual CPI values for FNAB from January 2010 to August 2023. The line then meets with the predicted values from September 2023 to August 2025. The forecasted CPI of FNAB in September 2023 is estimated to be 152.8501. At an 80% confidence interval, the lowest forecasted value of CPI of FNAB is 152.2700 while the highest would be at 164.5163. The 95% confidence interval for the forecasted CPI of FNAB is from 151.7325 to 166.5813. The CPI of FNAB is expected to progressively rise and reach 160.6154 by August 2025. Thus, the model has shown a consistently increasing trend in its prediction.

4. Discussion and Conclusion:

The pattern of the CPI of FNAB analysis shows the food inflation in Malaysia, thus knowledging it is fundamental. The rise in the CPI of FNAB will have a significant impact on the main CPI, as FNAB has the highest weightage of 29.54% (DOSM, 2023).

The objectives of this study is succesfully achieved since this study able to identify ARIMA (1,1,2) as the most suitable ARIMA model to forecast the future values of CPI of FNAB in Malaysia. The ARIMA (1, 1, 2) is selected since it has the lowest AIC and BIC values for all possible ARIMA models. There was no sudden increase in the projected values, but there was a consistent and gradual growth in the projected values. The forecast model's pattern shows this consistent growth pattern without a sudden spike.

This aligns with a study by Radzi et al. (2022), where they utilised the ARIMA model too specifically, ARIMA(0, 1, 1) model to predict the CPI in Malaysia for the next five years. The results indicate that the CPI will increase in accordance with the data. In addition, the finding obtained by Mohamed (2020) further supports the projected model of CPI of FNAB which suggests the model with the lowest AIC and BIC values is better than other competing models.

Moreover, this study able to forecast the future values of CPI of FNAB in Malaysia with the values from 151.7325 in September 2023 to 166.5813 in August 2025. These findings indicate that the prices of goods in Malaysia are expected to increase over the next 24 months, emphasizing the need for proactive measures by the government. Besides, this study provides knowledge about the CPI of FNAB that is relevant to researchers, enterprises, Malaysians, and the government.

From the findings, it is important to highlight there are very few studies that focus on forecasting the CPI specifically the FNAB subgroup. Moreover, there are also very limited number of studies about this topic in Malaysia even for the main CPI. It is crucial to address these gaps in the research to better understand and improve the accuracy of CPI forecasts in these contexts.

References:

DOSM. (2023). *Selected Agricultural Indicators 2023*.
DOSM. (2024). *Consumer Prices*. OpenDOSM. https://open.dosm.gov.my/dashboard/consumer-prices
Mohamed, J. (2020). Time Series Modeling and Forecasting of Somaliland Consumer Price Index: A Comparison of ARIMA and Regression with ARIMA Errors. *American Journal of Theoretical and Applied Statistics*, *9*(4), 143. https://doi.org/10.11648/j.ajtas.20200904.18
Radzi, H. M., Sufian, A., & Bakar, A. (2022). Forecasting GDP and CPI of the Malaysian Economy Using ARIMA Models. *8th Annual ECOFI Symposium 2022*. https://www.researchgate.net/publication/364358216
Salwati, N., & Wessel, D. (2021). How does the government measure inflation? *Brookings*. https://www.brookings.edu/articles/how-does-the-government-measure-inflation/