Asian Journal of Research in Education and Social Sciences
e-ISSN: 2682-8502 | Vol. 7, No. 2, 257-265, 2025
https://asianscholarsnetwork.com/asnet-journals
SPECIAL ISSUE: International Conference on Future and Sustainable Education (ICFSE) 2024

Modelling of RON97 Fuel Prices in Malaysia Using Time Series Analysis

Suhaila Bahrom[1,2*], Noratikah Abu[2]

[1] Department of Mathematics, Centre for Foundation Studies, International Islamic University Malaysia (IIUM), Malaysia
[2] Centre for Mathematical Sciences, University Malaysia Pahang Al-Sultan Abdullah), Malaysia

*Corresponding Author: suhaila_b@iium.edu.my

Received: 15 July 2024 | Accepted: 1 December 2024 | Published: 1 March 2025

DOI: https://doi.org/10.55057/ajress.2025.7.2.25

Abstract: *Forecasting the prices of RON97 fuel is crucial for economic planning and policy-making due to its impact on transportation costs and inflation. This study aims to analyse the RON97 fuel prices and develop a forecasting model using Box-Jenkins modelling. The study employs a comprehensive historical weekly dataset of RON97 fuel prices in Malaysia from 30th March 2017 to 18th April 2024 obtained from Malaysia's Official Open Data Portal. This research focuses on Stage I and Stage II of the Box-Jenkins approach. Stage I involves model identification, including a preliminary assessment of the data's stationarity through differencing and unit root tests. In Stage II, model estimation is applied to identify the most significant Box-Jenkins model to forecast the fuel price. This stage also includes identifying potential Box-Jenkins models based on autocorrelation functions (ACF) and partial autocorrelation functions (PACF). In this study, the analysis is done with the aid of R software, where the potential of this software in forecasting weekly RON97 fuel prices time series data is explored. The result from the analysis revealed that ARIMA (0,1,2) is the best model for forecasting RON97 fuel prices in Malaysia. Box-Jenkins modelling effectively captures the underlying patterns and trends in RON97 fuel prices, highlighting its applicability in economic and financial time series forecasting.*

Keywords: ARIMA, RON97, Box-Jenkins, Time Series Analysis

---

1. Introduction

Research Octane Number 95 (RON95) and Research Octane Number 97 (RON97) are the two types of petrol in Malaysia. RON95 is a widely used type of petrol in Malaysia, known for its affordability and popularity among private car owners. On the other hand, RON97 is considered a premium-grade fuel, offering higher octane and potentially better performance for certain vehicles (Mohamad & How, 2014). The increase in fuel prices in Malaysia is causing concern as it indicates higher grocery and living costs, with minorities feeling the most burdened (Rohani & Pahazri, 2018)Fluctuations in fuel prices significantly impact various sectors, including transportation, manufacturing, healthcare, entertainment, agriculture, construction, business, and the overall economy (Akhmad et al., 2019). Expenditure for private car owners will increase as fuel prices increase. The fluctuation in fuel prices in Malaysia is influenced by global crude oil prices and foreign exchange volatility, resulting in unpredictable price changes (Majuca, 2020) Crude oil prices are influenced by various factors, including geopolitical tensions, supply-demand imbalances, natural disasters, and economic fluctuations

(Zhang et al., 2024). The instability in fuel prices since the switch to a managed float system (MFS) in 2016 is primarily due to fluctuations in international crude oil prices, foreign exchange rates, and the reduction of subsidies (Sarpong-Streetor et al., 2021). As the price of crude oil increases, the expense of refining and manufacturing RON97 also increases.

Conversely, when crude oil prices drop, production costs usually decrease. While RON97 fuel is not as commonly used as RON95 fuel in the country, forecasting price changes for RON97 fuel is crucial as it is often used by owners of higher-end vehicles and luxury cars (Yunus et al., 2019). The ability to predict future fuel price patterns can impact vehicle decision (Santini et al., 2000). People considering purchasing a new car are more likely to choose a more fuel-efficient vehicle if they expect fuel prices to stay high when they plan to own the new vehicle. It also plays a role in decisions about long-term maintenance and usage plans. Understanding fuel price trends is essential for market analysis and strategic planning, especially for business owners and fleet managers. Assisting in making well-informed choices related to logistics, expenses for transportation, and general financial strategising. Moreover, the rise in oil prices will directly affect consumer goods and services, posing unpredictability in the future (Ahmad et al., 2022).

Previously, the Malaysian government used this mechanism to balance global market fluctuations and local economic conditions, ensuring that fuel prices remain manageable for consumers while maintaining market integrity (Tan, 2009). Forecasting fuel prices in Malaysia has been challenging due to the structural model's reliance on the inaccessible input variable, Mean of Platts Singapore (MOPS). The shift from the Automatic Price Mechanism (APM) to MFS in 2016 did not stabilise prices, leading to uncertainty in predictions (Sokkalingam et al., 2021). The need for an alternative forecasting method arises due to the constraints of the current structural model. Hence, this research aims to analyse the RON97 fuel prices in Malaysia and develop a forecasting model using the Box-Jenkins modelling. This study focuses on model identification and parameter estimation, which are Stage I and Stage II of the Box Jenkins methodology.

Autoregressive integrated moving average (ARIMA) models are among the most common univariate time series models used when the time series data is non-stationary. The autocorrelation function (ACF) and partial autocorrelation function (PACF) are used to identify the parameters of the autoregressive and moving average parameters (Snyder et al., 2017). Once the parameter has been identified, a possible model can be developed to find a significant model according to the parsimony principle using the Box-Jenkins methodology. In the next section, we will discuss methodology, describing the fuel price data used for analysis. The Box-Jenkins modelling approach is employed, starting with the model identification phase, where the Box-Jenkins model is selected as the framework for analysis. A stationary test follows this to ensure the time series data meets the necessary assumptions for modelling. The model is further refined through the parameter estimation stage. The results and discussion section presents the findings from the analysis, the list of possible models from ACF and PACF plots, and the equation for the best model. Finally, the conclusion summarises the key points of the study and suggests areas for future research in modelling RON97 fuel prices in Malaysia.

2. Methodology

Data Description

This research employs a comprehensive historical weekly dataset of RON97 fuel prices in Malaysia, obtained from Malaysia's Official Open Data Portal website. The data consists of

Asian Journal of Research in Education and Social Sciences
e-ISSN: 2682-8502 | Vol. 7, No. 2, 257-265, 2025
https://asianscholarsnetwork.com/asnet-journals
SPECIAL ISSUE: International Conference on Future and Sustainable Education (ICFSE) 2024

352 weekly observations from 30 March 2017 to 18 April 2024. The price of the RON97 fuel in the dataset is in Ringgit Malaysia (RM). R software was utilized to analyse the data as a time series. The possible models identified from the ACF and PACF are then compared with the main modeling automation algorithm, autoarima (Hyndman & Khandakar, 2008).

Box-Jenkins Modelling

The Box-Jenkins model, introduced by Box and Jenkins in 1968, is one of the most powerful forecasting methods available for time series analysis in research practice (Box & Jenkins, 1968). The methodology consists of four stages: model identification, parameter estimation, diagnostic checking, and forecasting. Figure 1 illustrates a summary of the stages in Box-Jenkins. Stage I of Box-Jenkins model identification consists of data screening and model identification, while Stage II involves model estimation. The time series data must be stationary to proceed with the second stage, known as parameter estimation. Stationarity in time series can be classified into stationary in variance and stationary in mean. The Box-Cox transformation is used when time series data is not stationary in variance. Table 1 presents commonly used values and their associated transformations (Box & Cox, 1964). The differencing, difference is often used to make series stationary in mean. In model identification, the autocorrelation function (ACF) and partial autocorrelation function (PACF) plots were used to check the stationarity of the data. In addition to the ACF and PACF plot, the unit root test, known as the Augmented Dickey-Fuller test (ADF-test), was used to confirm stationarity. Stage I and II Box-Jenkins modelling typically provide a robust model that fits the data well and captures the essential dynamics of the time series. By stopping at Stage II, the research avoids the potential overfitting and complexity that might arise in Stage III (diagnostic checking) and Stage IV (forecasting). The models developed in Stages I and II were suitable for the intended purpose, ensuring simplicity and interpretability without compromising forecast accuracy.

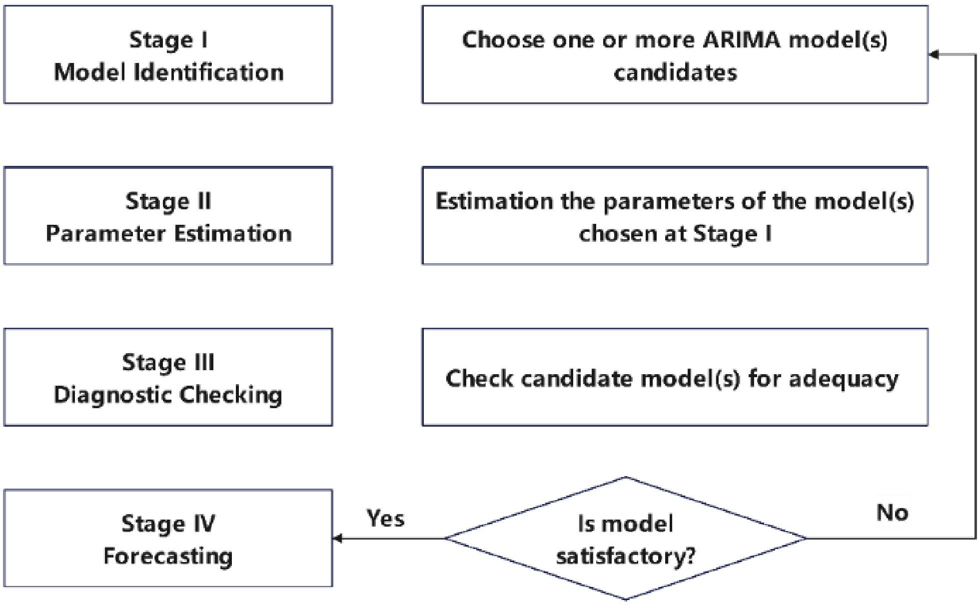

Figure 1: General Box-Jenkin's Framework

Table 1: Box-Cox Transformation

Values of λ	**-1.0**	**-0.5**	**0**	**0.5**	**1.0**
Transformation	$\frac{1}{y_t}$	$\frac{1}{\sqrt{y_t}}$	$\log_e y_t$	$\sqrt{y_t}$	y_t

Adopted from (Box & Cox, 1964)

Asian Journal of Research in Education and Social Sciences
e-ISSN: 2682-8502 | Vol. 7, No. 2, 257-265, 2025
https://asianscholarsnetwork.com/asnet-journals
SPECIAL ISSUE: International Conference on Future and Sustainable Education (ICFSE) 2024

Model Identification: Box-Jenkins Model

The Box-Jenkins model consists of five models based on linear stationary and nonstationary assumptions. Three models under linear stationary assumptions are Autoregressive (AR), Moving Average (MA), and Autoregressive Moving Average (ARMA). The AR, MA, and ARMA models are based on the stationary assumption; the series has constant variance and constant mean, and autocovariance depends only on the time lag. Two models under linear nonstationary are ARIMA (Autoregressive Integrated Moving Average) and Seasonal ARIMA (SARIMA). In order to identify the model, the ACF and PACF were used to analyse the RON97 fuel price over time and determine if the data was stationary. A very slow linear decay pattern can be corrected by first-degree differentiation order. The Autoregressive Integrated Moving Average (ARIMA) is a Box-Jenkins method widely used in time series modelling and forecasting. The three components encompassing the general term for ARIMA $(p, d, q)$are autoregressive (AR), integrated (Differencing), and moving average (MA), which are used in the respective order of p, d, q. When the original time series is stationary in mean, differencing is not required. In this case, the ARIMA model simplifies to an ARMA model. The general form of the ARIMA model, using a backshift operator and difference form, is provided in Equation 1.

$$\varphi_P(B)\nabla^d y_t = c + \theta_q(B)a_t \tag{1}$$

p:	the order of AR model
q:	the order of MA model
d:	the order of differencing
B:	backshift operator, $B^k y_t = y_{t-k}$
∇:	difference operator, $\nabla^d = (1-B)^d$

ARIMA modeling methods were used in this study based on a common method available for modeling and forecasting time series data. If the time series data pass all the preliminary tests in Stage I of Box-Jenkins modelling, then Stage II, which is parameter estimation, can be employed.

Stationarity Test

In Box-Jenkins modelling, the time series data must be stationary. If a time series is non-stationary, it means that the statistical properties of the data, such as the mean and variance, change over time. In order to apply the Box-Jenkins modelling, it is compulsory to make the time series stationary. One common method to accomplish this is by differencing the time series data, which entails calculating the variance between consecutive observations. This process aids in eliminating trends and seasonality, resulting in more consistent data that is suitable for analysis using the Box-Jenkins approach. There are several methods to detect non-stationarity, with common tests such as unit root and trend tests. In this study, the Augmented Dickey-Fuller (ADF) test was used to test for the stationarity of the dataset.

Parameter Estimation

Model or parameter estimation is applied to identify the best significant Box-Jenkins model. There are 2 approaches which are Maximum Likelihood Estimation (MLE) and Least Squares Estimates (OLS). This study uses Maximum Likelihood Estimation (MLE) to estimate the model's parameters. After identifying the possible model, the next step involved finding precise estimates of the model parameters. This study employs two different methods for parameter estimation: ACF and PACF analysis, as well as using the auto.arima function in R software.

Asian Journal of Research in Education and Social Sciences
e-ISSN: 2682-8502 | Vol. 7, No. 2, 257-265, 2025
https://asianscholarsnetwork.com/asnet-journals
SPECIAL ISSUE: International Conference on Future and Sustainable Education (ICFSE) 2024

The main modelling automation algorithm has been implemented in the auto.arima function. The summary of Stage I and Stage II of the Box-Jenkins procedure for forecasting the RON97 fuel prices is depicted in Figure 2.

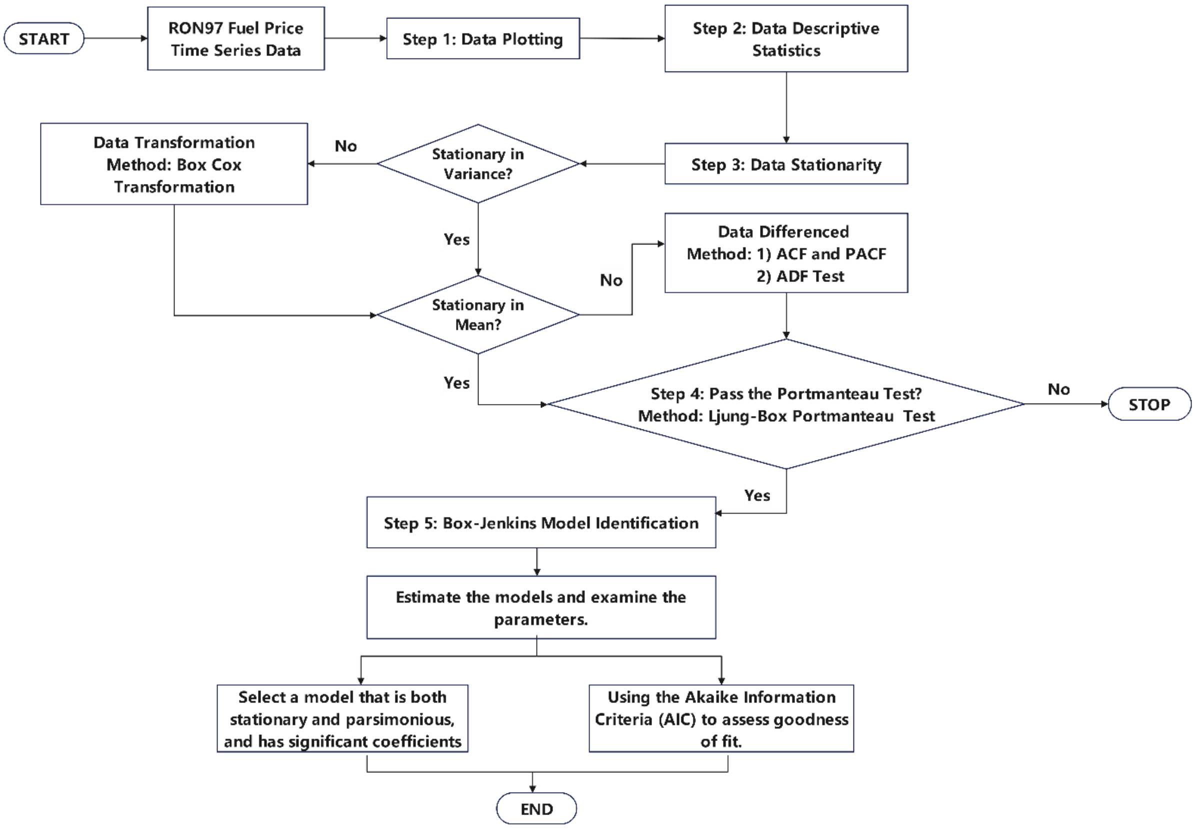

Figure 2: Stage I and Stage II Box-Jenkins Modelling

3. Results and Discussion

Malaysia's Official Open Data Portal recorded 352 weekly RON97 fuel prices from 30 March 2017 until 18 April 2024. Table 2 presents the descriptive statistics of the time series data. The mean price of RON97 fuel over this period is RM2.85, with the prices ranging from a minimum of 1.55 to a maximum of 4.84. The standard deviation, which measures the variability or dispersion of the prices, is 0.69, indicating moderate variation around the mean. The skewness value of 0.72 suggests a positive skew, while the kurtosis value of 0.23 indicates that the distribution is slightly platykurtic. The histogram in Figure 3 represents the distribution of the data.

Table 2: Descriptive Statistics

Mean	Maximum	Minimum	Standard deviation	Skewness	Kurtosis	Observation
2.85	4.84	1.55	0.69	0.72	0.23	352

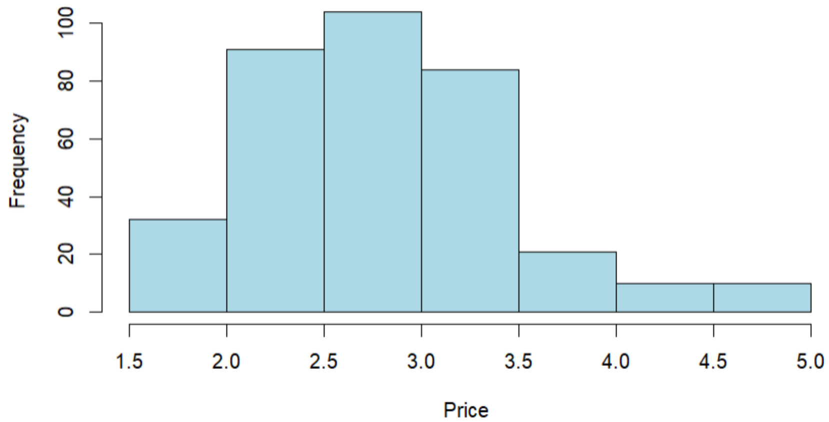

Figure 3: Histogram of RON97 Weekly Prices from 30th March 2017 to 18th April 2024

The data are split into train-test datasets based on a ratio of 90:10. Training data are used to estimate the model and test data are used to forecast RON97 fuel price purposes. The first part of the sample period varies from 30th April 2017 to 10th August 2023. This data will be used to estimate the models. Meanwhile, the second part, which is called an out-of-sample period, varies from 17th August 2023 to 18th April 2024. Figure 4 illustrates the data plot for RON97 weekly prices in Malaysia from March 30, 2017, to April 18, 2024. The data plot in Figure 4 reveals fluctuations in RON97 fuel prices from 2017 to 2020. There is an increasing trend from 2020 to 2022. The data price is slowly consistent from 2023 until the first quarter of 2024. This indicates that the data is not stationary. A stationary time series is defined as its means and variance are constant over time. The auto-correlation Function (ACF) and partial correlation function (PACF) plots are used to identify initial autoregressive (AR) and moving average (MA) model parameters. Using autocorrelation functions in a linear time series model can capture the linear dynamics of the data (Tsay, n.d.). The model coefficients are estimated recursively using maximum likelihood estimation (MLE), where the number of parameters is regularised using the Akaike Information Criterion (AIC). The ACF plot in Figures 5(a) and 5(b) shows that the value of the auto-correlation coefficient is large, and the ACF plot or ACF value decreases slowly. This plot supports the initial finding from data plotting. The correlogram reporting the ACF and PACF showed that seasonality does not influence the prevalence and incidence of RON97 fuel prices. Using the raw training data from March 30, 2017, to August 10, 2023, the data was transformed to stabilise the variance using Box-Cox transformation, and a first-order difference $(d = 1)$ was applied to make the data stationary in mean. The augmented Dickey-Fuller (ADF) test was used to determine whether the sequence was stationary, and the result supported that the data was a stationary time series $\left(p-value=0.01\right)$.

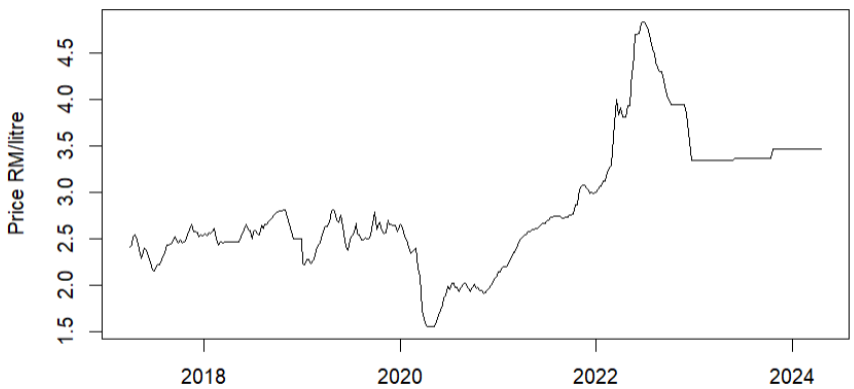

Figure 4: RON97 weekly prices in Malaysia from 30th March 2017 to 18th April 2024

Asian Journal of Research in Education and Social Sciences
e-ISSN: 2682-8502 | Vol. 7, No. 2, 257-265, 2025
https://asianscholarsnetwork.com/asnet-journals
SPECIAL ISSUE: International Conference on Future and Sustainable Education (ICFSE) 2024

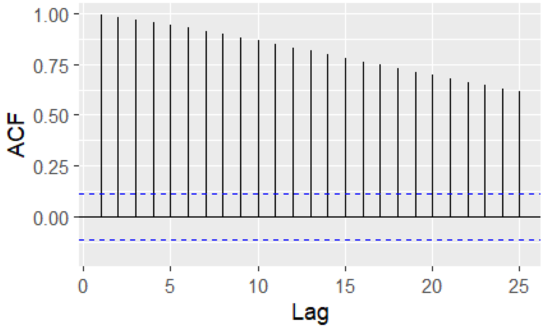

Figure 5(a): ACF plot of the in-sample data

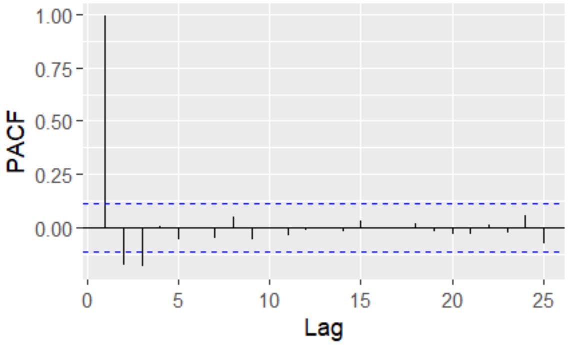

Figure 5(b): PACF plot for in-sample data

The ACF and PACF plots for the stationary in-sample data are presented in Figures 6(a) and 6(b). The significant spikes drop to near zero after lag 2, indicating a rapid decay characteristic of a moving average process of order 2, as shown in Figure 6(a). Hence the possible values for the MA component are 0,1 and 2 $(q = 0,1,2)$. In the PACF plot depicted in Figure 6(b), there are significant spikes at lag 1 and lag 2, with the rest of the lags dropping to near zero, suggesting an autoregressive process of order 2. Therefore, the possible AR components are 0,1 and 2 $(p = 0,1,2)$. Table 2 lists possible models derived from the ACF and PACF plots of stationary in-sample data or training datasets.

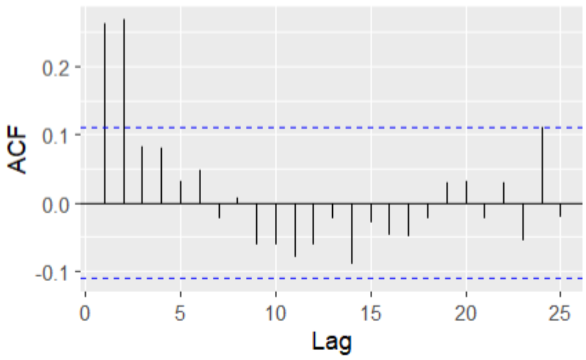

Figure 6(a): ACF plot of stationary data

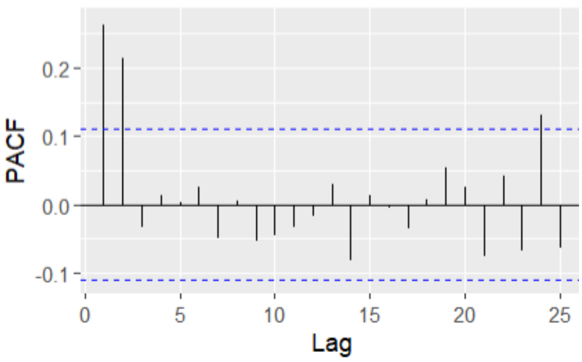

Figure 6(b): PACF plot of stationary data

Table 3: List of Possible Models

No	Model	AIC	Model Significance	No. of Parameter(s)
1	ARIMA $(0,1,0)$ with drift	2103.06	Significant	0
2	ARIMA $(0,1,1)$ with drift	-2116.33	Significant	1
3	ARIMA $(0,1,2)$ with drift	-2134.67	Significant	2
4	ARIMA $(1,1,0)$ with drift	-2123.69	Significant	1
5	ARIMA $(1,1,1)$ with drift	-2131.50	Significant	2
6	ARIMA $(1,1,2)$ with drift	-	-	3
7	ARIMA $(2,1,0)$	-2136.53	Significant	2
8	ARIMA $(2,1,1)$	-	-	3
9	ARIMA $(2,1,2)$ with drift	-	-	4

Based on the results of the goodness-of-fit test statistics presented in Table 4, ARIMA$(0,1,2)$ model with drift was identified as the optimal model due to its lowest AIC value. This model also passed the Ljung–Box Q Test $(z = 25.607, p = 0.060)$, and all the parameter estimates were significant. The ARIMA model is implemented using a forecast package from R studio. The equation for the ARIMA$(0,1,2)$ is represented as Equation 2.

$$y_t = 0.2055a_{t-1} + 0.2448a_{t-2} + a_t \quad (2)$$

4. Conclusion

The historical weekly RON97 fuel prices have been analysed to identify trends and patterns. By utilising the ACF and PACF, we can systematically identify and evaluate potential ARIMA models based on the observed ACF and PACF patterns. Our findings indicate that the ARIMA model is suitable for forecasting RON97 fuel prices due to its nonstationary and non-seasonal patterns, and this paper proposes ARIMA as the best model. Future research should proceed with Stage III and Stage IV of the Box-Jenkins methodology. In Stage III, diagnostic checks are conducted to ensure the model's adequacy and identify potential inadequacies or improvement areas. By addressing any residual issues or model misspecifications, the accuracy and reliability of the model can be further enhanced. Stage IV focuses on generating and evaluating forecasts, allowing for a comprehensive assessment of the model's predictive performance. The continuation will confirm the initial findings and establish a robust model for accurate predictions of RON97 fuel prices. This can be valuable for policymakers, energy analysts, market participants, stakeholders, and consumers. Consumers can utilise these predictions to prepare for future fuel expenses, manage their budgets effectively, and make well-informed transportation and household spending decisions.

Acknowledgement
The authors express their gratitude to the Malaysia Official Open Data Portal (data.gov.my) for providing access to the dataset employed in this study, which was essential for the research's successful completion.

References

Ahmad, I., Iqbal, S., Khan, S., Han, H., Vega-Muñoz, A., & Ariza-Montes, A. (2022). Macroeconomic effects of crude oil shocks: Evidence from South Asian countries. *Frontiers in Psychology*, *13*. https://doi.org/10.3389/fpsyg.2022.967643

Akhmad, Romadhoni, B., Karim, K., Tajibu, M. J., & Syukur, M. (2019). The impact of fuel oil price fluctuations on indonesia's macro economic condition. *International Journal of Energy Economics and Policy*, *9*(2), 277–282. https://doi.org/10.32479/ijeep.7470

Box, G. E. P., & Cox, D. R. (1964). An Analysis of Transformations. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *26*(2), 211–243. https://doi.org/10.1111/j.2517-6161.1964.tb00553.x

Box, G. E. P., & Jenkins, G. M. (1968). Some Recent Advances in Forecasting and Control. In *Source: Journal of the Royal Statistical Society. Series C (Applied Statistics)* (Vol. 17, Issue 2). https://about.jstor.org/terms

Hyndman, R. J., & Khandakar, Y. (2008). *Journal of Statistical Software Automatic Time Series Forecasting: The forecast Package for R* (Vol. 27). http://www.jstatsoft.org/

Majuca, R. (2020). *Assessing the Impact of Oil Prices on the Malaysian Economy*. https://www.amro-asia.org/wp-content/uploads/2020/04/Impact-of-Oil-Prices-on-Malaysia_final.pdf

Mohamad, T. I., & How, H. G. (2014). Part-load performance and emissions of a spark ignition engine fueled with RON95 and RON97 gasoline: Technical viewpoint on Malaysia's fuel price debate. *Energy Conversion and Management*, *88*, 928–935. https://doi.org/10.1016/j.enconman.2014.09.008

Rohani, M. M., & Pahazri, N. (2018). Survey on how fluctuating petrol prices are affecting Malaysian large city dwellers in changing their trip patterns. *IOP Conference Series: Earth and Environmental Science*, *140*(1). https://doi.org/10.1088/1755-1315/140/1/012085

Santini, D. J., Patterson, P. D., & Vyas, A. D. (2000). Importance of Vehicle Costs, Fuel Prices, and Fuel Efficiency in Hybrid Electric Vehicle Market Success. *Transportation Research Record: Journal of the Transportation Research Board*, *1738*(1), 11–19. https://doi.org/10.3141/1738-02

Sarpong-Streetor, R. M. N. Y., Sokkalingam, R., Othman, M., Daud, H., & Owusu, D. A. (2021). *ARIMAX Modelling of Ron97 Price with Crude Oil Price as an Exogenous Variable in Malaysian* (pp. 679–691). https://doi.org/10.1007/978-981-16-4513-6_59

Snyder, R. D., Ord, J. K., Koehler, A. B., McLaren, K. R., & Beaumont, A. N. (2017). Forecasting compositional time series: A state space approach. *International Journal of Forecasting*, *33*(2), 502–512. https://doi.org/10.1016/j.ijforecast.2016.11.008

Sokkalingam, R., Sarpong-Streetor, R. M. N. Y., Othman, M., Daud, H., & Owusu, D. A. (2021). Forecasting Petroleum Fuel Price in Malaysia by ARIMA Model. *Springer Proceedings in Complexity*, 671–678. https://doi.org/10.1007/978-981-16-4513-6_58

Tan, P. (2009). APM: How fuel prices are calculated in Malaysia. *Malaysian Fuel Prices*.

Tsay, R. S. (n.d.). *Analysis of Financial Time Series, Third Edition (Wiley Series in Probability and Statistics)*.

Yunus, S. M., Hodin, S., Zulkifli, F. H., Mustaffa, N., & Osman, S. A. (2019). Fuel, Mixture Formation and Combustion Process The Comparative Analysis between Gasolines RON95, RON97 and RON100 on Engine Performance. *Fuel*, *1*(2), 1–7. www.fazpublishing.com/fmc

Zhang, Q., Hu, Y., Jiao, J., & Wang, S. (2024). The impact of Russia–Ukraine war on crude oil prices: an EMC framework. *Humanities and Social Sciences Communications*, *11*(1), 8. https://doi.org/10.1057/s41599-023-02526-9