African Journal of Mathematics and Statistics Studies
ISSN: 2689-5323
Volume 9, Issue 1, 2026 (pp. 13-26)
www.abjournals.org

TIME SERIES MODELING AND FORECASTING OF PETROL SALES AT A FUEL STATION IN ZIMBABWE

Henry Samambgwa*, Simbarashe Chinyere, and Thomas Musora.

Department of Mathematics and Statistics, School of Natural Sciences and Mathematics, Chinhoyi University of Technology.

*Corresponding Author's Email: henrysamambgwa@gmail.com

Cite this article:

H., Samambgwa, Simbarashe, C., T., Musora (2026), Time Series Modeling and Forecasting of Petrol Sales at a Fuel Station in Zimbabwe. African Journal of Mathematics and Statistics Studies 9(1), 13-26. DOI: 10.52589/AJMSS-HSZ2HI0H

Manuscript History

Received: 12 Nov 2025

Accepted: 15 Dec 2025

Published: 8 Jan 2026

ABSTRACT: *Petrol is an important commodity in the Zimbabwean economy. Its availability, shortage and pricing have a significant impact on key economic indicators and the livelihoods of the general population in the country. This study focuses on petrol sales volumes at a fuel station in Zimbabwe. There are costs and negative implications that emanate from overstocking or understocking of petrol at the fuel station. This makes it necessary to forecast future demand, so as to plan and maintain adequate volumes of petrol. This study sought to fit a SARIMA model using monthly sales data from January 2015 to December 2024. Minimal information criteria were used to compare and select from candidate models. Parameter p-values were then used to identify and eliminate likely insignificant parameters. This was because insignificant parameters contribute to model overfitting. The best fitting model was then used to forecast petrol sales volumes for the months from January to December 2025.*

KEYWORDS: Time series, seasonal autoregressive integrated moving average (SARIMA).

Article DOI: 10.52589/AJMSS-HSZ2HI0H
DOI URL: https://doi.org/10.52589/AJMSS-HSZ2HI0H

African Journal of Mathematics and Statistics Studies
ISSN: 2689-5323
Volume 9, Issue 1, 2026 (pp. 13-26)

INTRODUCTION

Petrol is a critical commodity for the wellbeing of the Zimbabwean economy (Chayita & Kaseke, 2021). Times of inadequate fuel supply are characterised by economic turmoil, unfavorable inflation rates, and depreciated standards of living for the general population (Toriro & Mugadza, 2020). Petrol sales constitute an important indicator of economic activity and energy consumption trends. Predicting petrol sales volumes allows for effective distribution and logistics planning (Zema et al., 2024). For fuel companies, distributors, and policymakers, understanding and predicting sales patterns is essential for effective planning, inventory management, pricing strategies, and infrastructure investment.

Forecasting petrol sales is important for managing distribution in fuel companies (Sun et al, 2018). Effective management of fuel sales is crucial for businesses in the energy sector, as it directly impacts revenue generation, customer satisfaction, and operational efficiency. Predicting sales helps managers in strategic planning to prepare for fluctuations in demand (Alwadi, 2025). Predictive analytics are essential in the management and planning of business operations (Karonski, 2024). Petrol sales are closely linked to economic activity, consumer behaviour, and transportation patterns (Chadha et al., 2023). Analysing fuel sales data provides useful insights on changes in demand, market trends, and competitive strategy in supply chain management and market strategies. Overall, petrol sales are an important component of the energy sector, having a significant impact on businesses, consumers, and the broader economy. This study analysed and forecasted petrol sales volumes of a fuel company in Zimbabwe. The fuel company had several service stations and distribution centers across the country.

Stock management strategies envisage efficient supply of petrol to satisfy demand (Seroney et al., 2019). Accurate prediction of monthly fuel demand can improve supply chain management, strategic decision-making, and financial planning for businesses (Krause et al., 2024). Redundancy can be avoided, losses prevented and profits increased if petrol distributions across fuel stations is optimised. For strategic planning purposes, it is necessary for the company to forecast petrol sales volumes at its different fuel stations.

STATISTICAL METHODS

Time Series Analysis and Modeling Approach

Time series modeling is a statistical technique for developing mathematical models of sequential data occurring in time. It includes building models to predict future sales based on historical data (Ezeliora, 2013). Predictive analytics are used for strategic planning and economic management (Al-Chalabi et al., 2018). In these models, values of variables are obtained from their predecessors, such that using the model, and given historical data, present and future values can be predicted as forecasts.

ARIMA Modeling

The autoregressive integrated moving average (ARIMA) model is made up of autoregressive and moving average components. Among the multitude of methods available for predictive analytics, the ARIMA model stands out for its simplicity and effectiveness (Rizvi, 2024).

Article DOI: 10.52589/AJMSS-HSZ2HI0H
DOI URL: https://doi.org/10.52589/AJMSS-HSZ2HI0H

African Journal of Mathematics and Statistics Studies
ISSN: 2689-5323
Volume 9, Issue 1, 2026 (pp. 13-26)

Autoregressive Model

In an autoregressive model, subsequent values are obtained from past values, a constant term and the present error (Equation (1)):

$$Y_t = \lambda_0 + \lambda_1 Y_{t-1} + \lambda_2 Y_{t-2} + \ldots + \lambda_p Y_{t-p} + \epsilon_t, \quad (1)$$

where:

Y_t = value of variable Y at time

$\lambda_0, \lambda_1, \lambda_2, \ldots, \lambda_p$ are constants

ϵ_t = error at time t

Parameter p is equal to the number of past Y values used in the calculation.

Moving Average Model

In the moving average model, subsequent values are obtained from past errors plus a constant term and the present error (Equation (2)):

$$Y_t = \mu_0 + \mu_1 \epsilon_{t-1} + \mu_2 \epsilon_{t-2} + \ldots + \mu_q \epsilon_{t-q} + \epsilon_{t,} \quad (2)$$

where:

Y_t = value of variable Y at time t

$\mu_0, \mu_1, \mu_2, \ldots, \mu_q$ are constants

ϵ_t = error at time t

Parameter q is equal to the number of past error values used in the calculation.

Autoregressive Integrated Moving Average Process (ARIMA)

When the autoregressive (AR) and moving average (MA) models are combined, the result is an ARMA model. If differencing is also used to make the data stationary, the result is an ARIMA model (Musora et al., 2023).

Seasonal Autoregressive Integrated Moving Average Process (SARIMA)

When seasonality is incorporated into an ARIMA model, it becomes a SARIMA model. Seasonality is the repetitive patterns that occur at fixed intervals (Arumugam & Natarajan, 2023).

Musora et al. (2022) formally state the SARIMA model in Equation (3):

$$\lambda_p(B)\Phi_P(B^s)\nabla^d \nabla_s^D Y_t = \mu_q(B)\Theta_Q(B^s)X_t, \quad (3)$$

where:

Y_t is an $ARIMA(p, d, q)$ process

Article DOI: 10.52589/AJMSS-HSZ2HI0H
DOI URL: https://doi.org/10.52589/AJMSS-HSZ2HI0H

African Journal of Mathematics and Statistics Studies
ISSN: 2689-5323
Volume 9, Issue 1, 2026 (pp. 13-26)
www.abjournals.org

X_t is a random process with mean zero and constant variance

s is the length of the season

$\nabla_s^D X_t = \sum_{j=0}^{D} \quad (D\ j\ ) Y_{t-js}$, and $\lambda_p(B)$ and $\mu_q(B)$ are polynomials in B of order p and q respectively (Equations (4) and (5)):

$$\mu_q(B) = \left(1 - \mu_1 B - \mu_2 B^2 - \ldots - \mu_q B^q\right), \tag{4}$$

$$\lambda_q(B) = \left(1 - \lambda_1 B - \lambda_2 B^2 - \ldots - \lambda_p B^p\right). \tag{5}$$

Model Validation

The model is evaluated for validity by examining residual plots of the autocorrelation function and the partial autocorrelation function (ACF and PACF), histogram, normal probability plot, plot of residuals against fitted values, and the plot of residuals against order (sequential position in the time series). The model is considered valid if no further relationship is detected between residuals and fitted or observed values. The residuals have to have white noise characteristics. Stationarity of data can be evaluated from the autocorrelation function (ACF) plot (Salman & Kanigoro, 2021). The Augmented Dickey-Fuller test can also be used to confirm stationarity (Shubam et al, 2024).

ANALYSIS RESULTS AND DISCUSSION

Analysis Software Environment

Minitab 22 software environment was used in this study for analysis, generating results, tables and plots.

Data Source

Petrol sales data was obtained from the service station. The data contained monthly sales volumes in litres of petrol for a period of 120 months (10 years), from January 2015 to December 2024.

The data was checked for stationarity using an Augmented Dickey-Fuller (ADF) test (**Table 1**).

Table 1: Augmented Dickey-Fuller Test

Null hypothesis:		Data are non-stationary
Alternative hypothesis:		Data are stationary
Test Statistic	**P-Value**	**Recommendation**
-3.18836	0.021	Test statistic <= critical value of -2.88636. Significance level = 0.05 Reject null hypothesis. Data appears to be stationary, not supporting differencing.

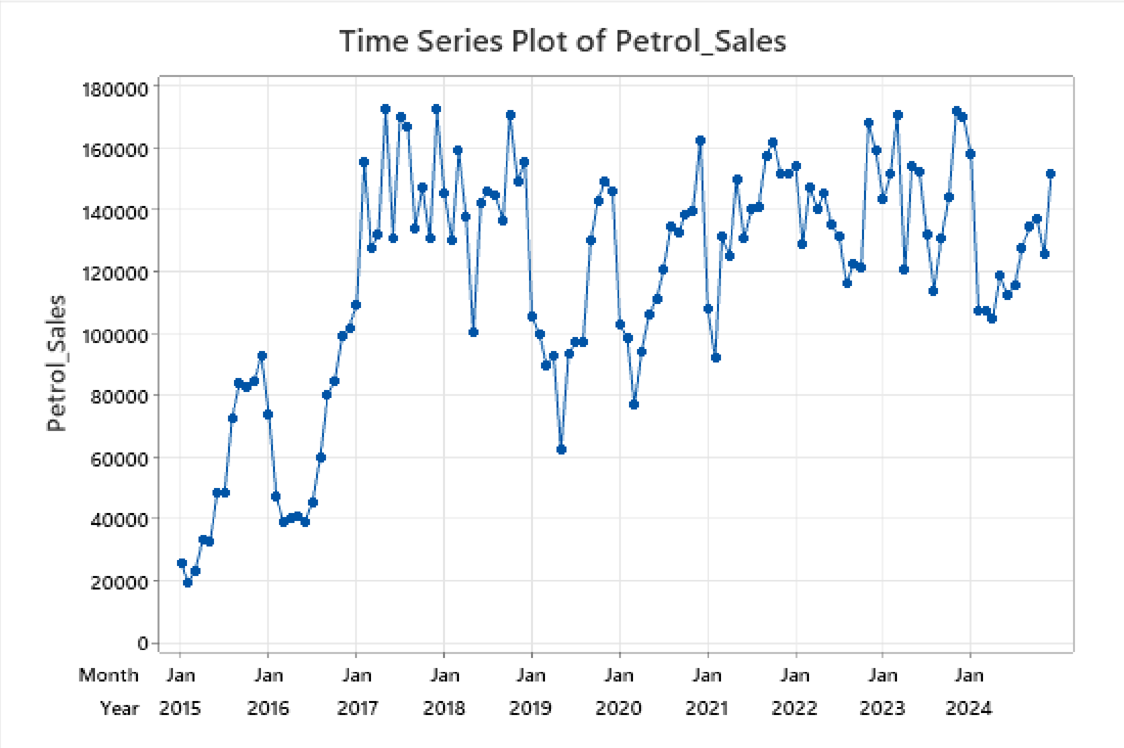
Article DOI: 10.52589/AJMSS-HSZ2HI0H
DOI URL: https://doi.org/10.52589/AJMSS-HSZ2HI0H

The ADF test confirmed that the data was stationary. A time series plot (**Figure 1**) was generated to confirm stationarity and detect any seasonal patterns.

Time Series Plot

The time series plot was generated showing the month-on-month petrol sales volumes in litres for the 120 months from January 2015 to December 2024 (**Figure 1**).

Figure 1: Time series plot of petrol sales

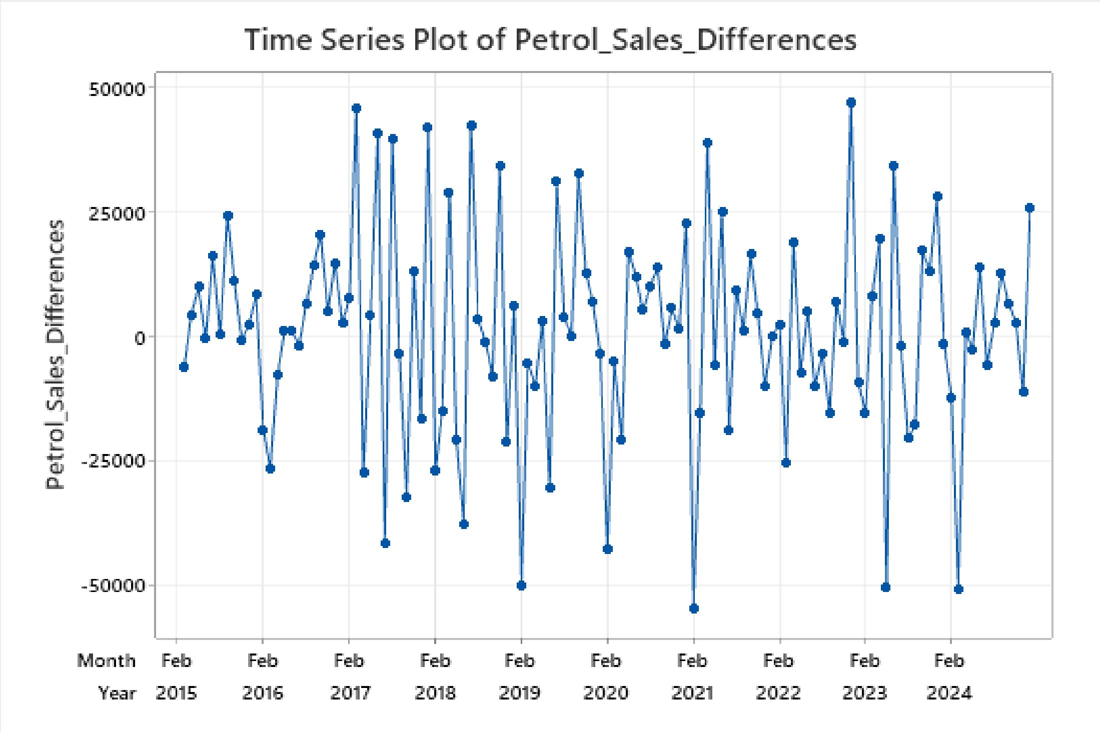


Although the data had passed the ADF test for stationarity, visual inspection of the plot revealed the likely presence of a trend as well as a seasonal pattern. It was thus necessary to difference the data to make it stationary.

Article DOI: 10.52589/AJMSS-HSZ2HI0H
DOI URL: https://doi.org/10.52589/AJMSS-HSZ2HI0H

African Journal of Mathematics and Statistics Studies
ISSN: 2689-5323
Volume 9, Issue 1, 2026 (pp. 13-26)

Figure 2: Time series plot of petrol sales after differencing



The plot of first differences (**Figure 2**) was observed to achieve stationarity. An ADF test (**Table 2**) was done, once more, to confirm the stationarity of the data.

Table 2: Augmented Dickey-Fuller Test

Null hypothesis:	Data are non-stationary
Alternative hypothesis:	Data are stationary

Test Statistic	**P-Value**	**Recommendation**
-5.82116	0.000	Test statistic <= critical value of -2.88820. Significance level = 0.05 Reject null hypothesis. Data appears to be stationary, not supporting differencing.

Thus, a differencing parameter of 1 was determined.

Article DOI: 10.52589/AJMSS-HSZ2HI0H
DOI URL: https://doi.org/10.52589/AJMSS-HSZ2HI0H

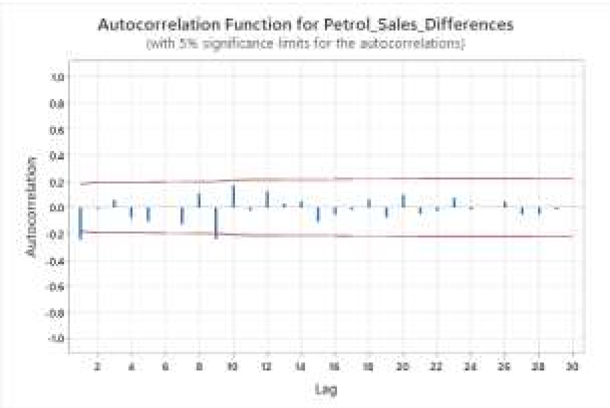
African Journal of Mathematics and Statistics Studies
ISSN: 2689-5323
Volume 9, Issue 1, 2026 (pp. 13-26)

Autocorrelation Function and Partial Autocorrelation Function Plots of Differenced Data

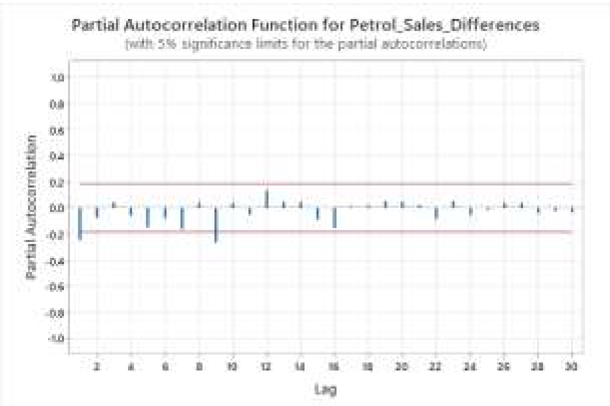
Autocorrelation function (ACF) (**Figure 3(a)**) and partial autocorrelation function (PACF) (**Figure 3(b)**) plots were generated for the differenced data.

Figure 3: (a) Autocorrelation and (b) partial autocorrelation, plots of difference petrol sales data

(a)



(b)



It was observed that the ACF decayed soon after the significant spike at the first lag, and the PACF also decayed soon after a significant spike at the first lag. This suggested the suitability of an ARIMA (1,1,1) model.

Seasonality had been observed in the data from the time series plot of petrol sales. The Minimal Information Criteria (MINIC) selection technique was then used to find the best fitting parameters for the seasonal part of the model. Information criterion automatic selection methods generate reliable results for time series models (Billah et al., 2006).

Article DOI: 10.52589/AJMSS-HSZ2HI0H
DOI URL: https://doi.org/10.52589/AJMSS-HSZ2HI0H

Model Selection Using Minimal Information Criteria

The ARIMA(1,1,1) model was deduced from the ACF and PACF plots after differencing once. From the observed seasonality in the data, the model was proposed as follows:

$$ARIMA(1,1,1)(p,d,q)_{12}$$

MINIC selection was used in the following ranges:

$0 \leq p \leq 3, 0 \leq d \leq 1$, and $0 \leq q \leq 3$.

The possible models were estimated using Minitab 22. **Table 3** shows the resulting Akaike Information Criteria (AIC) and Bayesian Information Criteria (BIC) values. Inestimable models were excluded from the results.

Table 3: Model Comparison and Selection

p	d	q	AIC	BIC
0	0	0	2701.96	2710.3
0	0	1	2698.99	2710.11
0	0	2	2700.69	2714.59
0	0	3	2714.03	2730.7
0	1	0	2486.07	2494.08
0	1	1	2478.94	2489.63
0	1	2	2466.94	2480.3
0	1	3	2475.95	2491.99
1	0	0	2700.43	2711.55
1	0	1	2719.06	2732.96
1	0	2	2718.22	2734.89
1	1	0	2472.65	2483.34
1	1	1	2476.41	2489.77
1	1	2	2458.3	2474.34
1	1	3	2470.85	2489.56
2	0	0	2702.33	2716.23
2	1	0	2464.96	2478.33
2	1	1	2470.43	2486.46
2	1	2	2530.17	2548.88
2	1	3	2565.5	2586.88
3	0	0	2703.57	2720.24
3	1	0	2506.35	2522.39
3	1	1	2474.92	2493.63
3	1	2	2475.59	2496.97
3	1	3	2498.47	2522.53

The best model is as simple as possible and minimizes information criteria (Fattah et al., 2018). By choosing the lowest Bayesian Information Criteria (BIC), the following model was selected:

$$ARIMA(1,1,1)(1,1,2)_{12}$$

Article DOI: 10.52589/AJMSS-HSZ2HI0H
DOI URL: https://doi.org/10.52589/AJMSS-HSZ2HI0H

Table 4 shows the final parameter estimates for the model.

Table 4: Final Estimates of Parameters

Type	Coef	SE Coef	T-Value	P-Value
AR 1	-0.074	0.281	-0.26	0.792
SAR 12	-0.647	0.500	-1.29	0.199
MA 1	0.301	0.267	1.13	0.262
SMA 12	0.120	0.475	0.25	0.802
SMA 24	0.674	0.381	1.77	0.080

Model Refinement

Table 4 was interpreted to determine the statistical significance of estimated parameters. It was noted that the parameter p-values were far above the threshold level of significance (0.05). The SMA(24) parameter had a p-value of 0.08, marginally close to the threshold of 0.05, suggesting that a seasonal moving average parameter may be necessary. It was concluded that the autoregressive, seasonal autoregressive and moving average parameters were likely statistically insignificant. Insignificant parameters must be excluded from the model as they contribute to the model overfitting. A new model was sought using MINIC selection (**Table 5**), with parameter ranges that eliminate the autoregressive (p), seasonal autoregressive (P) and moving average (q) parameters from the $SARIMA(p,d,q)(P,D,Q)_{12}$ model. The seasonal moving average parameters (Q)were estimated from zero to five. The parameter estimation ranges were thus: $0 \le p \le 0, 1 \le d \le 1, 0 \le q \le 0, 0 \le P \le 0, 1 \le D \le 1,$ and $0 \le Q \le 5$.

Table 5: Model Comparison and Selection

Model (d = 1, D = 1)	**AIC**	**BIC**
p = 0, q = 0, P = 0, Q = 1	2452.76	2458.10
p = 0, q = 0, P = 0, Q = 2	2455.74	2463.76
p = 0, q = 0, P = 0, Q = 3	2457.70	2468.39
p = 0, q = 0, P = 0, Q = 4	2458.47	2471.83
p = 0, q = 0, P = 0, Q = 5	2459.57	2475.60

The $SARIMA(0,1,0)(0,1,1)_{12}$ was found to minimize the BIC. The final parameter estimates and p-values were analysed in **Table 6**.

Table 6: Final Estimates of Parameters

Type	Coef	SE Coef	T-Value	P-Value
SMA 12	0.8579	0.0829	10.34	0.000

The SMA(12) parameter was noted to be statistically significant.

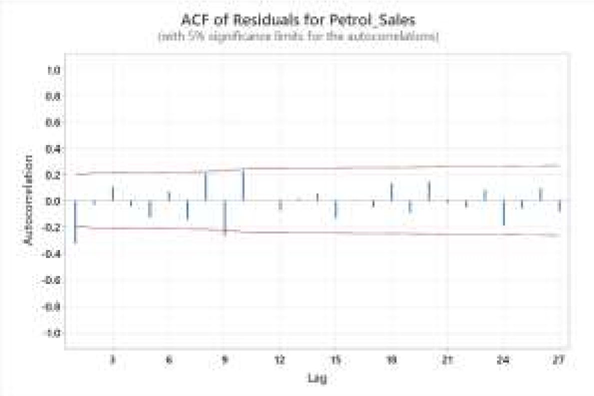
Article DOI: 10.52589/AJMSS-HSZ2HI0H
DOI URL: https://doi.org/10.52589/AJMSS-HSZ2HI0H

Model Validation

Six residual plots were used to evaluate the performance of the model, namely, ACF (**Figure 4(a)**), PACF (**Figure 4(b)**), Normal Probability Plot (**Figure 5(a)**), Residuals vs Fitted Values ((**Figure 5(c)**), and Residuals vs Order Plots (**Figure 5(d)**).

Figure 4: (a) Autocorrelation and (b) partial autocorrelation, plots of residuals

(a)



(b)

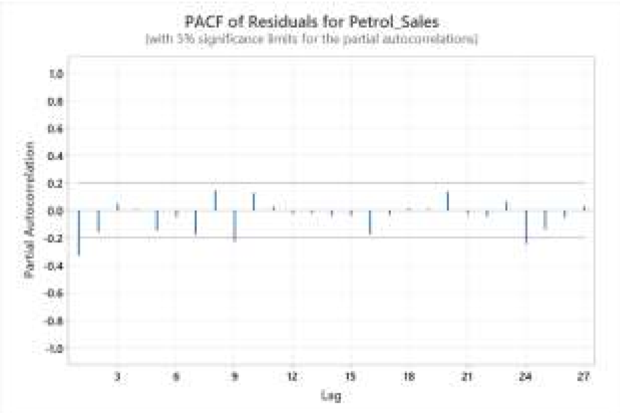
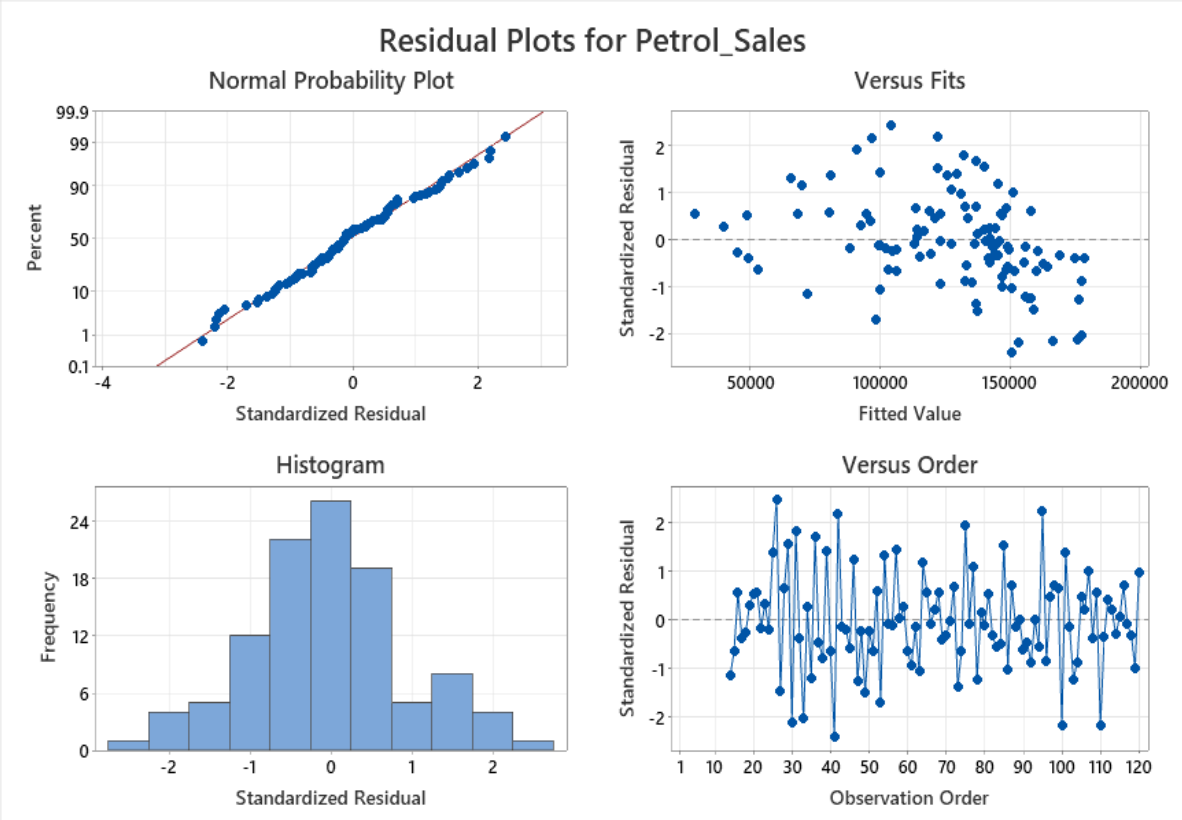


Figure 5: (a) Normal probability, (b) residual vs fitted values, (c) residuals histogram and (d)residuals vs order plots

Article DOI: 10.52589/AJMSS-HSZ2HI0H
DOI URL: https://doi.org/10.52589/AJMSS-HSZ2HI0H



Visual inspection of the residual ACF and PACF plots (**Figure 4(a) and 4(b)**) revealed no significant spikes, inferring that the selected model parameters were sufficient. In the normal probability plot (**Figure 5(a)**), the points were very close to the straight line, as expected for normally distributed residuals. The plot of Residuals vs Fitted values (**Figure 5(b)**) were randomly distributed, indicating no further relationship between residuals and fitted values. Observation of the histogram of residuals (**Figure 5(c)**) confirmed that the residuals were normally distributed. It was thus concluded that the final model was a sufficiently good fit, and could then be used for forecasting.

Forecasting Using the Chosen Model

The $ARIMA(1,1,1)(1,1,2)_{12}$ model was used to forecast petrol sales volumes for each month from January 2025 to December 2025 (**Table 6**).

Table 6: Petrol Sales Forecasts (Jan 2025 – Dec 2025)

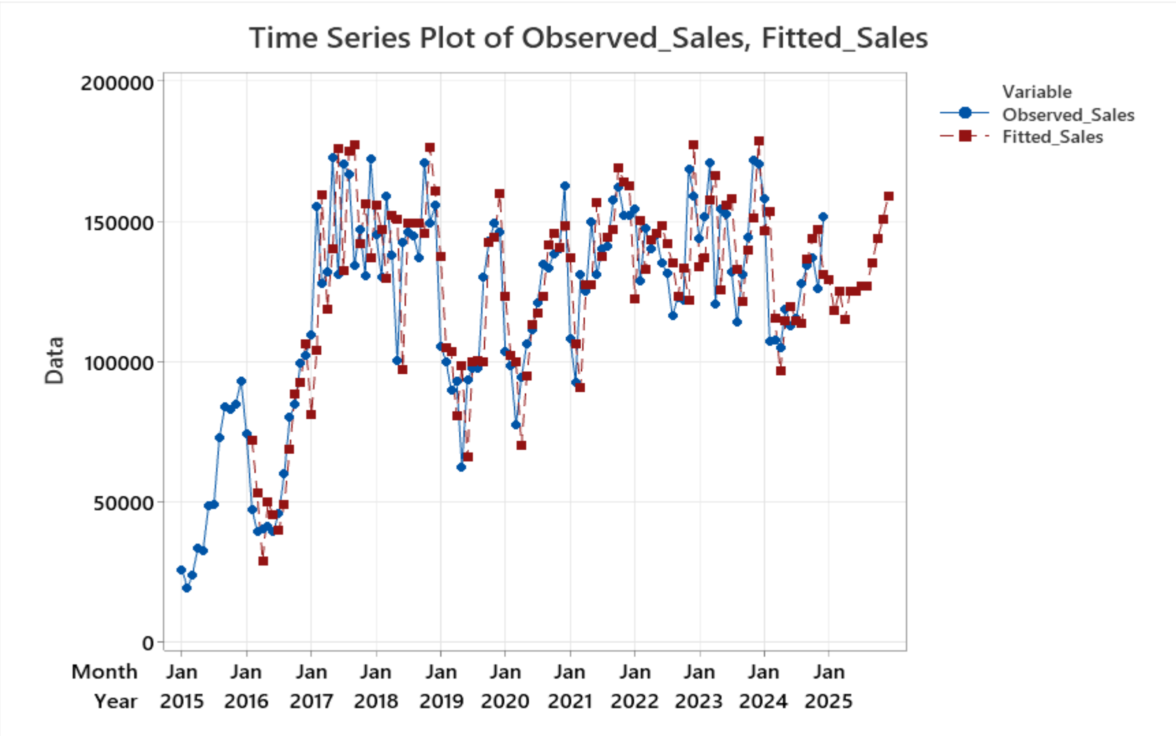
Month	Forecasted Sales (Litres)	Observed Sales (Litres)
Jan-24	146480	158052
Feb-24	153229	107132
Mar-24	115234	107662
Apr-24	96481	104935
May-24	114274	118557
Jun-24	119425	112731
Jul-24	114389	115386

Article DOI: 10.52589/AJMSS-HSZ2HI0H
DOI URL: https://doi.org/10.52589/AJMSS-HSZ2HI0H

Aug-24	113629	127920
Sep-24	136473	134373
Oct-24	143750	136993
Nov-24	146945	125819
Dec-24	131159	151471
Jan-25	129288	
Feb-25	117916	
Mar-25	124942	
Apr-25	114962	
May-25	124909	
Jun-25	124826	
Jul-25	126626	
Aug-25	126900	
Sep-25	135155	
Oct-25	143572	
Nov-25	150523	
Dec-25	158749	

Figure 6 illustrates the observed petrol sales values from January 2015 to December 2025, plotted against the fitted values, as well as the forecast values from January 2025 to December 2025.

Figure 6: Times series plots of observed petrol sales values and fitted petrol sales values with forecasts



Article DOI: 10.52589/AJMSS-HSZ2HI0H
DOI URL: https://doi.org/10.52589/AJMSS-HSZ2HI0H

Visual inspection of this time series plot revealed a close-fitting relationship between the petrol sales values generated by the model and the observed values. This further endorsed the suitability of the $SARIMA(0,1,0)(0,1,1)_{12}$ model.

DISCUSSION

Strategic planning is critical in fast moving consumer goods (FMCG) industries, particularly in cases involving high sales volumes with very high gross revenue implications. Commodity supply needs to be maintained within optimal levels. Failure to maintain adequate supply of goods could lead to customer dissatisfaction, inadequate revenue generation, and a wide array of emanating negative consequences, including significant financial losses. Delivering more goods than needed can lead to storage logistics challenges, increased storage maintenance costs, opportunity costs, and a wide array of emanating negative consequences, including significant financial losses. It is thus necessary to strategise adequately and manage the supply goods within optimal limits.

In such situations, it is critical to make reliable forecasts of service and goods demand. A reliable forecasting model is a splendid instrument in the strategic management of such high sales volume and high gross revenue markets. In this study, a model was formulated that satisfies fundamental validation criteria. The forecasted petrol sales volumes for the months from January 2025 to December 2025 were generated within 95% confidence limits. They are thus a suitable and critically necessary data sheet for strategic planning for the year 2025.

REFERENCES

Al-Chalabi, H., Al-Douri, Y. K., Lundberg, J. (2018). Time Series Forecasting using ARIMA Model. The Twelfth International Conference on Advanced Engineering Computing and Applications in Sciences.

Alwadi, M. A. (2025). Fuel Sales Price Forecasting using Time Series, Machine Learning, and Deep Learning Models. Engineering Technology and Applied Science Research.

Arumugam, V., Natarajan, V. (2023). Time Series Modelling and Forecasting using Autoregressive Integrated Moving Average and Seasonal Autoregressive Integrated Moving Average Models. Instrument Mesure Metrologie.

Billah, B., King, M.L., Snyder, R.D. and Koehler, A.B. (2006). Exponential smoothing model selection for forecasting. International journal of forecasting.

Chadha, R. S., Jugesh, Parveen, S., Singh, J. (2023). Fuel Sales Forecasting with SARIMA-GARCH and Rolling Window. Journal of Soft Computing Paradigm.

Chayita, R., Kaseke, N. (2021). Survival Business Strategies in the Petroleum Industry in Zimbabwe. South East Asia Journal of Contemporary Business, Economics and Law.

Ezeliora, C. D., Chinwuko, E. C., Ogunoh, A. V., Nwosu, M. C. (2013). Regression and Time Series Analysis of Petroleum Product Sales in Masters Energy oil and Gas. International Journal of Science, Engineering and Technology Research.

Fattah, J., Ezzine, L., Aman, Z., Moussami, H. E., Lachhab, A. (2018). Forecasting of demand using ARIMA model. International Journal of Engineering Business Management.

Article DOI: 10.52589/AJMSS-HSZ2HI0H
DOI URL: https://doi.org/10.52589/AJMSS-HSZ2HI0H

Karonski, A., Hernes, M., Walaszczyk, E., Rot, A., 2024. Forecasting Sales at Fuels Stations using a Multilayer Perceptron. Communications in Computer and Information Science. Springer.

Krause, J., Beiruth, A. C. A., Barddal, J. P., Britto, A. S., Souza V. M. A. (2024). Fuels Demand Forecasting: Identifying Leading Feature Sets, Prediction Strategy, and Regressors. International Conference on Machine Learning and Applications.

Musora, T., Chazuka, Z., Jaison, A., Mapurisa, J., Kamusha, J. (2023). Demand Forecasting of a Perishable Dairy Drink: An ARIMA Approach.

Musora, T., Chazuka, Z., Matarise, F. (2022). Foreign Direct Investment Inflow Modelling and Forecasting; A Case Study of Zimbabwe. CARI 2022.

Rizvi, M. F., Sahu, S., Rana, S. (2024). ARIMA Model Time Series Forecasting. Ijraset Journal for Research in Applied Science and Engineering Technology.

Salman, A. G., Kanigoro, B. (2021). Visibility Forecasting using Autoregressive Integrated Moving Average (ARIMA) Models. Elsevier.

Seroney, J. K., Wanyoike, D. M., Langat, E. K. (2019). Influence of Demand Forecasting on Supply Chain Performance of Petroleum Marketing Companies in Nakuru County, Kenya. The International Journal of Business Management and Technology.

Shubam, S., Nirmalkar, R., Wadhwani, T., Shivani, V., Pandey, A. (2024). Time Series Forecasting and Anomaly Detection with ARIMA and SARIMAX. International Journal of Research Publication and Reviews.

Sun, L., Xing, X., Zhou, Y., Hu, X. (2018). Demand Forecasting for Petrol Products in Gas Stations using Clustering and Decision Tree. Journal of Advanced Computational Intelligence and Intelligent Informatics.

Toriro, P., Mugadza, T. (2020). Socio-spatial Constructions Surrounding Queuing for Fuel in Harare. Journal of Urban Systems and Innovations for Resilience in Zimbabwe.

Zema, T., Wójcik, F., Sulich, A., Hernes, M. (2024). Forecasting Fuel Retail Sales Volume Using Machine Learning for Sustainable Decision-Making. Emerging Challenges in Intelligent Management Information Systems.

Article DOI: 10.52589/AJMSS-HSZ2HI0H
DOI URL: https://doi.org/10.52589/AJMSS-HSZ2HI0H