Tanzania Journal of Science
Volume 51(4), 2025
journal homepage: https://commons.udsm.ac.tz/tjs/

Modeling and forecasting wholesale gasoline prices in Tanzania using ARIMA and neural network autoregressive models

Thadei D Sagamiko[1*] and Edward K Ngailo[1, 2]
[1]*Department of Physics, Mathematics and Informatics, Dar es Salaam University College of Education, University of Dar es Salaam, P. O. Box 2329, Tanzania.*
[2]*Department of Computer and Information Science, Division of Statistics and Machine Learning, Linköping University, P. O. Box S-58 183 Linköping, Sweden.*
E-mail co-author: edward.ngailo@udsm.ac.tz

Keywords

Gasoline Prices;
Forecasting;
ARIMA;
Neural Networks

Abstract

Fluctuations in wholesale gasoline prices significantly impact the economy of import-dependent countries like Tanzania, making accurate forecasting essential for policymakers and market participants. This study forecasts monthly wholesale gasoline prices in Dar es Salaam by comparing the Autoregressive Integrated Moving Average (ARIMA) model with a Neural Network Autoregressive (NNAR) model. Using data from January 2015 to December 2024 from the Energy and Water Utilities Regulatory Authority (EWURA), the models were evaluated. The ARIMA(1,1,1) model was identified as the best-fitting linear model, but the NNAR(11,7) model demonstrated superior accuracy. On the holdout test set, the NNAR model achieved a lower Mean Absolute Percentage Error (MAPE) of 2.3447% and a Root Mean Square Error (RMSE) of 76.6797, compared to the ARIMA model's MAPE of 2.049% and RMSE of 95.8016. The findings indicate that the NNAR model is more effective at capturing the complex and non-linear patterns in gasoline prices.

*Corresponding author email: tsagamiko@gmail.com
17 Dec 2024; Revised 16 Nov 2025; Accepted 22 December 2025; Published 29 December 2025
https://doi.org/10.65085/2507-7961.1116

ISSN 0856-1761, e-ISSN 2507-7961

Introduction

The global energy market is characterized by its complexity and volatility, with gasoline prices serving as a key indicator of economic growth and consumer behaviour. Fluctuations in gasoline prices can have implications affecting everything from transportation costs to inflation rates (Kafuku 2025). In Africa, the energy landscape is particularly dynamic, influenced by a combination of local production capabilities, international oil prices, and geopolitical factors. As countries across the continent strive for economic growth, understanding the factors that drive gasoline prices becomes increasingly important for policymakers, businesses, and consumers alike. In Tanzania, the situation is not different as the country heavily relies on imported petroleum products, making it susceptible to global price changes and supply chain disruptions (Kojima et al. 2010). Dar es Salaam, as the largest city and economic hub, experiences the most effects of these fluctuations. The Energy and Water Utility Regulatory Authority (EWURA) plays a crucial role in monitoring and publishing prices of petroleum products, providing essential data that reflects the local market conditions. This data is important for stakeholders who need to navigate the complexities of fuel pricing in a rapidly changing economic environment.

For more than half a century, Box and Jenkins (1970, 1974, 1976), also known as Auto-Regressive Integrated Moving Average (ARIMA) time series models, have been predominant in many areas of time series forecasting. ARIMA models assume that the future values of a variable are linear function of several past observations and random errors (Brockwell 2016, Shumway and Stoffer 2017). The ARIMA model is based on the premise that the series is generated from a linear process (Khashei and Bijari 2010). Several researchers have applied ARIMA models for forecasting crude oil prices; for example, Gasper and Mbwambo (2023) applied ARIMA model for forecasting crude oil prices, Ntare (2023) used ARIMA models for the prediction of petrol and diesel prices in Dar es Salaam, Gasper and Mbwambo (2023) fitted ARIMA models and compared with Holt's method in forecasting petroleum prices in Tanzania. However, the linearity assumption of ARIMA model may be inappropriate when the underlying mechanism is nonlinear. In fact, real-world applications are often nonlinear (Zhang et al. 1998). To capture the nonlinear behaviour within variables, more sophisticated tools and methods may be necessary.

Artificial Neural Networks (ANNs) provide a technique capable of modeling complex nonlinear relationships among input regressors and response variables (Thoplan 2014). One of the most significant advantages of ANN models over other nonlinear statistical models is their ability to act as universal approximators, enabling them to approximate a wide range of functions with a high degree of accuracy (Zhang and Qi 2005). Feed-forward ANNs, combined with the back-propagation algorithm, have been widely used in various applications. For instance, Rostami et al. (2015) applied this approach to predict air quality indices, while EkmiŞ et al. (2017) developed a feed-forward neural network in combination with the ARIMA model for forecasting daily revenue of a fashion retail chain in Istanbul. Recently, Junita and Kartikasari (2024) employed this method to forecast the value of oil and gas exports in Indonesia, Vijayalakshmi et al. (2023) used ANN to predict rice production in India, and Daniyal et al. (2022) compared conventional modeling techniques with the neural network autoregressive model using COVID-19 data.

The Neural Network Autoregressive (NNAR) model is a type of ANN in which the lagged values of a time series are used as input predictors, and the output is the predicted values of the series (Wah 2023). One of the primary differences between NNAR and ARIMA models is that NNAR does not impose any restrictions on its parameters to ensure stationarity (Thoplan 2014).

Unlike many previous studies that have focused on retail fuel prices, which are often shaped by local politics, taxation policies, and domestic market interventions, our analysis emphasizes wholesale prices. Wholesale prices are more directly linked to international oil markets, making them sensitive to global shocks such as the COVID-19 pandemic, geopolitical tensions, and supply chain disruptions (Sharif et al. 2020, Yang and Fu 2025). This distinction is particularly important in Tanzania, an import dependent country where wholesale prices better reflect global energy dynamics than retail prices. In this study, we compare Box-Jenkins ARIMA and NNAR models to predict monthly wholesale gasoline prices in Dar es Salaam, Tanzania. The comparison is based on the forecasting performance of the best-performing ARIMA and NNAR models using common accuracy metrics such as the root mean squared error (RMSE), mean absolute percentage error (MAPE), the Akaike information criterion (AIC) and Bayesian information criterion (BIC).

The novelty of this study lies less in the methodological comparison itself, which has been examined in other contexts, and more in the application to wholesale gasoline prices in an import-reliant African economy, under conditions shaped by recent global disruptions. By focusing on this unique dataset, our research provides new insights into how international price dynamics are transmitted to local markets in Tanzania. The findings have direct implications for policymakers, energy regulators, and businesses that rely on accurate forecasts for decision-making.

Materials and Methods

Study area

This study focuses on the wholesale gasoline market in Tanzania, with a specific emphasis on Dar es Salaam. Tanzania is an East African nation whose economy is highly dependent on imported petroleum products, making it particularly sensitive to fluctuations in the global oil market (Kojima et al. 2010). Dar es Salaam serves as the country's primary economic hub and the main port of entry for over 95% of its petroleum imports. Consequently, the wholesale gasoline prices in Dar es Salaam are a critical benchmark for the entire national market. The choice to analyze wholesale prices, as opposed to retail prices, is strategic; they are more directly influenced by international cost factors and less distorted by local taxes, subsidies, and regulatory interventions, providing a clearer signal of fundamental market dynamics.

Data sources

This study utilized a time-series dataset of monthly wholesale gasoline prices in Tanzania. The data were sourced from the Energy and Water Utilities Regulatory Authority (EWURA), the official regulatory body, and retrieved from their publicly available webpage (EWURA, n.d.). The dataset spans a period of ten years, from January 2015 to December 2024. All statistical modeling and analysis were conducted using the R programming language (R Core Team 2024), leveraging the following key packages: the *auto.arima()* and *forecast* package (Hyndman and Khandakar 2008) for ARIMA modeling and the *nnet* package (Venables and Ripley 2002) for neural network implementation.

The ARIMA model

Box–Jenkins method is a classical statistical modeling technique for analyzing and forecasting time series data (Shumway and Stoffer 2017). The method is named after the statisticians George Box and Gwilym Jenkins who developed the autoregressive integrated moving average (ARIMA) in the early 1970 (Box and Jenkins 1970), and since then they have been applied to a wide variety of time series prediction applications (Box et al. 2015, Brockwell 2016).

The ARIMA model is composed of three main components: AutoRegressive (AR) terms, Integrated (I) terms, and Moving Average (MA) terms. An AR(p) component refers to the use of past values in the regression equation for the series z_t. The autoregressive parameter p specifies the number of lags used in the model. For example, an AR(2) model, which is equivalently represented as ARIMA (2,0,0), can be written as;

$$z_t = \mu + \phi_1 z_{t-1} + \phi_2 z_{t-2} + \varepsilon_t$$

where ϕ_1 and ϕ_2 are the partial autocorrelation parameters for the $AR(2)$.

The general form of the $AR(p)$ model is given by;

$$z_t = \mu + \sum_{i=1}^{p} \phi_i z_{t-i} + \varepsilon_t,$$

where p is the order of the AR model and $\phi_1, \phi_2, .., \phi_p$ are the partial autocorrelation parameters for the $AR(p)$ model.

The parameter d represents the degree of differencing in the integrated $I(d)$ component. Differencing a series involves subtracting its current value from its previous value d times.

A moving average $MA(q)$ component represents the error of the model as a linear combination of previous error terms ε_t. The order q signifies the order of the moving average component. The moving average is a linear combination of past forecast errors.

The general $MA(q)$ model is expressed as;

$$z_t = \mu + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \ldots + \theta_q \varepsilon_{t-q} + \varepsilon_t$$

$$= \mu + \sum_{i=1}^{q} \theta_i \varepsilon_{t-1} + \varepsilon_t \ ,$$

where $\theta_1, \theta_2, \ldots, \theta_q$ are the parameters of the model. The terms $\varepsilon_t, \varepsilon_{t-1}, \ldots, \varepsilon_{t-2}$ are the white noise error terms.

Combining autoregressive, differencing, and moving average components make up a non-seasonal ARIMA model, which can be written as a linear equation;

$$z_t = \mu + \sum_{i=1}^{p} \phi_i z_{t-i} + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i} + \varepsilon_t \ , \qquad (1)$$

where z_t is the observed value at time t, μ is a constant mean, ϕ_i are the autoregressive parameters, and θ_i are the moving average parameters.

The methodology starts by identifying models through the analysis of the autocorrelation function (ACF) and partial autocorrelation function (PACF) plots. Once a model is defined, its parameters are estimated, and the model's adequacy is assessed through diagnostic testing, typically using the Ljung-Box statistic test. If the model is deemed adequate, it can then be utilized for forecasting purposes. The Box and Jenkins methodology involves three model construction stages and a forecasting step: (i) Identification: This step ensures stationarity in time series by analyzing ACF and PACF to determine model orders. (ii) Estimation: This determines optimal and orders in ARIMA models by comparing sum of square errors through F-tests and estimating parameters using the least-squares method. (iii) Diagnostic checking: Evaluates model fit using diagnostic tests and plots, adjusting or selecting new models as needed for adequacy. (iv) Forecasting: In this step, the model is selected with the lowest MSE for forecasting and compares its performance with an NNAR model through cross-validation.

Neural network auto-regressive (NNAR) model

Neural networks, particularly those applied to time series forecasting, have gained considerable attention due to their ability to capture complex nonlinear relationships in data (Maleki et al. 1993). Unlike traditional linear models such as ARIMA model given in Equation (1), neural networks can model and learn from patterns in the data without explicitly defining the underlying relationship between variables. The most basic neural network consists of an input layer, one or more hidden layers, and an output layer as shown in Figure 1. There are three types of neural network architectures (Aggarwal 2018, Ghatak 2019): feed forward neural networks, convolutional neural networks, and recurrent neural networks. Feed forward network is the most widely used neural network model for time series modeling and forecasting (Hyndman and George 2018). The feed forward artificial neural network has found extensive application in various environmental contexts, including the prediction of wastewater treatment plant performance (Khodadadi et al. 2016), assessing pollutant removal efficiency (Mansoorian et al. 2017), and forecasting air quality indices (Rostami et al. 2015). The NNAR model is a three-layer feed forward neural network, comprising a linear combination function and an activation function. In the following is Figure 1 which provides the conceptual basis for the neural network models.

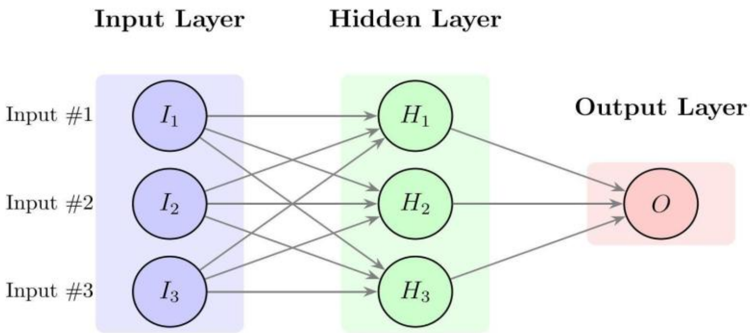


Figure 1: Example architecture of a feed forward neural network. This schematic illustrates the fundamental components (input (I), hidden (H), output layers (O)) that form the basis of the NNAR models used in this study.

The relationship between the model output z_t and the inputs $z_{t-1},\ldots,z_{t-p}$ can mathematically be expressed as (Ghatak 2019);

$$z_t = w_0 + \sum_{j=1}^{h} w_j . g\left(w_{0,j} + \sum_{i=1}^{n} w_{i,j} . z_{t-i} \right) + \varepsilon_t , \quad (2)$$

where, $w_{i,j} (i = 0,1,2,\ldots,n,\ j = 1,\ 2,\ \ldots,h),\ w_j \left(j = 0,\ 1,\ 2,\ \ldots.,\ h \right)$ represent the connection weights, n denotes the number of input nodes, and h signifies the number of hidden nodes.

The logistic sigmoid function, defined in Equation (3) below, is utilized as the activation function in the hidden layer. For the output layer, the most commonly used activation function in machine learning is a linear function (Ghatak 2019)

$$g(z) = \frac{\exp(-z)}{1+\exp(-z)} = \frac{1}{1+\exp(-z)} \ , \quad (3)$$

where $\exp(.)$ represents exponent.

Lagged values of the time series can be used as inputs to a neural network $NNAR(p,k)$, lagged inputs and k nodes in the single hidden layer.The $NNAR(p,0)$ model is equivalent to an $ARIMA(p,0,0)$ model but without stationarity restrictions. In Figure 2, is an illustration of hidden layer's role in learning complex patterns.

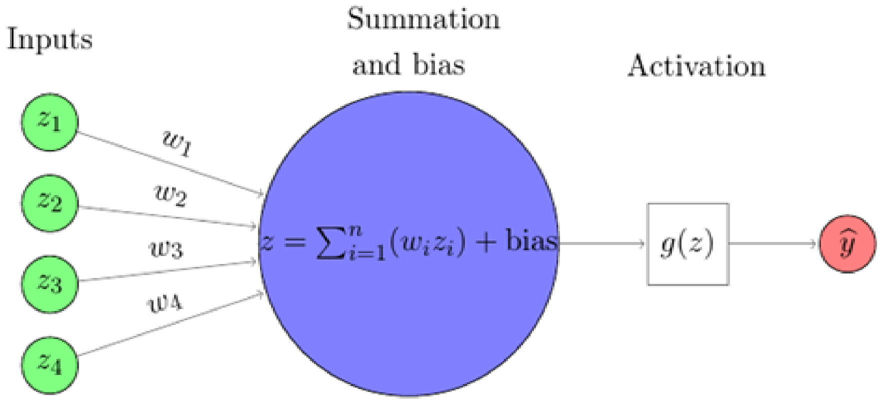


Figure 2: Conceptual diagram of the non-linear transformation performed by a hidden layer. The layer maps input data into a higher-dimensional space (a geometric transformation), allowing the network to model complex, non-linear relationships in the time series data.

The optimal order for the NNAR model is determined by analyzing the PACF plot of the training data to identify significant lags, which suggest an initial input parameter (p). The number of neurons in the hidden layer (k) is then empirically tested based on the common heuristic $k=(p+1)/2$ evaluating configurations and neurons to find the model with the best performance on the testing set, thus avoiding overfitting (Wah et al. 2023).

Forcasting accuracy metrics

To validate the forecasting models, the dataset is to be partitioned into a training set and a testing (holdout) set. The training set to comprise approximately 80% of the data (from January 2015 to December 2023), used for model fitting and parameter estimation. The remaining 20% (January 2024 to December 2024) to be used as an out-of-sample test set. The forecasting accuracy metrics (MAPE, RMSE, MAE) for the testing phase are calculated by comparing the model forecasts for this period against the actual observed values, ensuring an unbiased assessment of predictive performance.

To assess the effectiveness of each forecasting model in predicting wholesale gasoline prices, the RMSE, MAPE, MAE, AIC and BIC are employed.

i. The Akaike information criterion (AIC): is utilized to determine the best model based on its performance. If k represents the number of model parameters and L denotes the estimated value of the maximum likelihood, the AIC equation can be expressed as follows (Cavanaugh and Neath 2019);

$$AIC = 2k - 2\ln(L).$$

ii. The mean absolute percentage error (MAPE): is the average of absolute errors over a given period, multiplied by 100% to express the results as percentages. The formula for calculating the MAPE is as follows (Desiyanti et al. 2022);

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{z_i - \hat{z}_i}{z_i}\right| \times 100\% \ ,$$

where z_t is the actual data in period t, $\hat{z}_t$ is the forecasted data in the period t, and n is the number of data points.

iii. The root mean square error (RMSE): is a measure of the differences between the actual and forecasted values. It gives a relatively high weight to large errors due to the squaring of differences, making it sensitive to outliers. The RMSE formula (Hyndman and George 2018) is given by

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(z_i - \hat{z}_i)^2},$$

where z_t is the actual data in period t, $\hat{z}_t$ is the forecasted data in period t, and n is the number of data periods.

iv. The mean absolute error (MAE): measures the average magnitude of the errors in a set of forecasts, without considering their direction. It is less sensitive to outliers than RMSE because it does not square the errors. The MAE (Hyndman and George 2018) is calculated as

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|z_i - \hat{z}_i|,$$

where z_t is the actual data in period t, $\hat{z}_t$ is the forecasted data in period t, and n is the number of data periods.

Results

This section presents the results of the ARIMA and NNAR models, followed by a comparative analysis of their performance.

Overview of price dynamics and exogenous shocks

Figure 3 shows the gasoline wholesale price in Dar es Salaam city, Tanzania, between January 2015 and December 2024. It can be observed that the price of gasoline generally increased, with some fluctuations, during the period from 2015 to 2024. For example in 2020, there was a sharp decline in gasoline prices, which can be attributed to the impact of the Corona virus (COVID-19) pandemic on global fuel demand (WHO 2020). A significant price jump in 2022 is most likely influenced by the Russia-Ukraine war, which disrupted global energy markets and led to a sharp increase in oil prices (Yagi and Managi 2023).

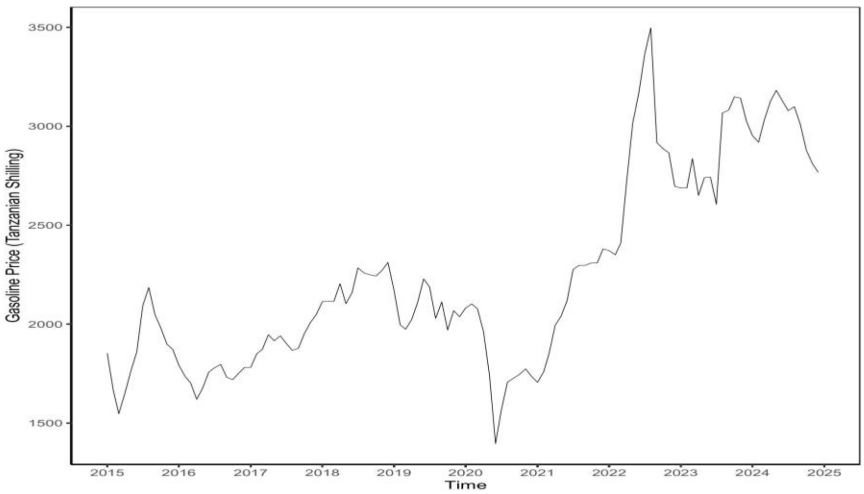


Figure 3: Monthly wholesale gasoline prices in Dar es Salaam, Tanzania (January 2015 – December 2024), showing a general upward trend with significant volatility attributable to global events.

ARIMA model

The ACF in Figure 4 shows that there is a significant autocorrelation at lag 1, and that the autocorrelation decreases as the lag increases. This suggests a strong dependence between the current price of gasoline and the price in the previous period. On the other hand, the PACF in Figure 4 indicates a significant partial autocorrelation at lag 1, suggesting a direct relationship between the current price of gasoline and the price in the previous period, independent of the prices in other periods.

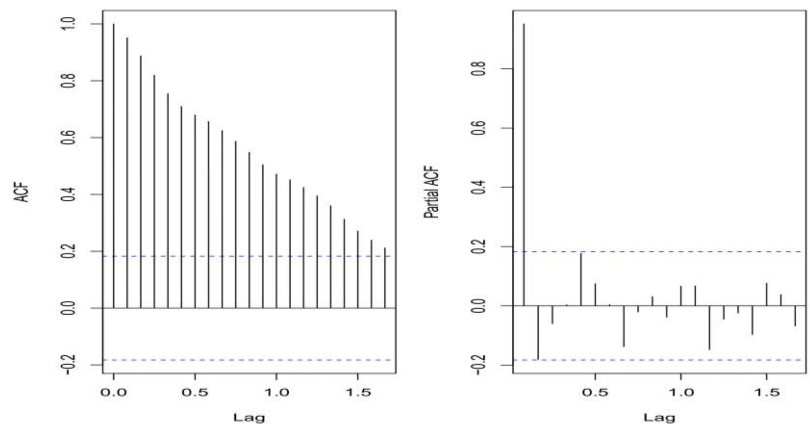


Figure 4: ACF and PACF plots of the original gasoline price series, indicating non-stationarity. The slow decay in the ACF and the significant spike at lag 1 in the PACF suggest the need for differencing and inform the initial ARIMA(p,d,q) order selection.

Table 1: Stationarity assessment using the Augmented Dickey-Fuller (ADF) test. The null hypothesis of a unit root (non-stationarity) cannot be rejected for the original series (p > 0.05) but is rejected for the first differenced series ($p < 0.01$) confirming stationarity after differencing.

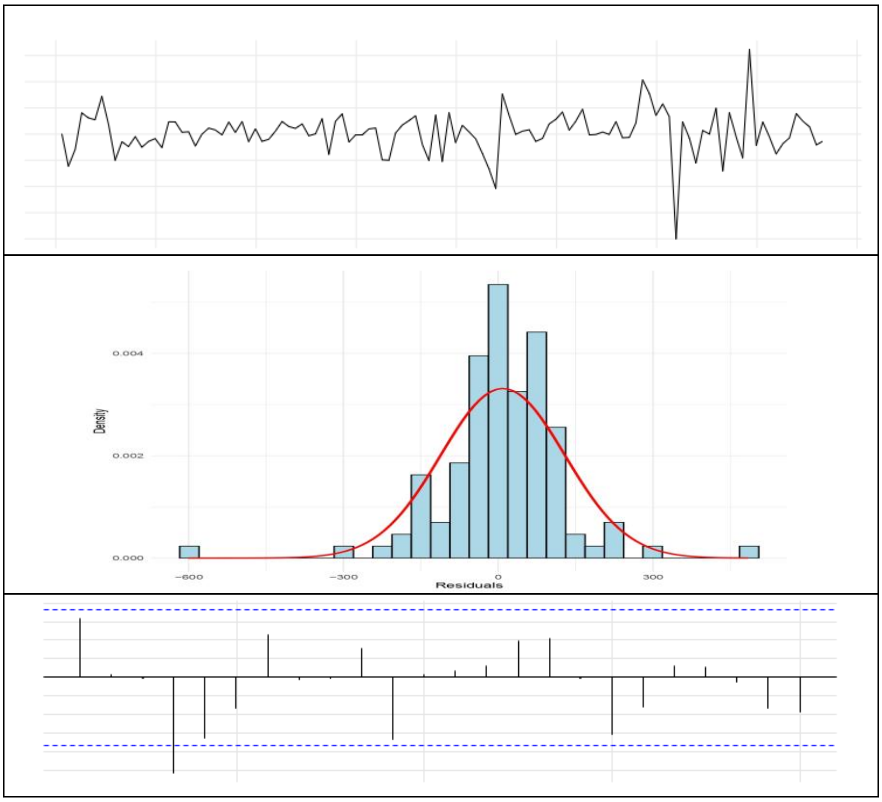
	Dickey-Fuller	p-value
Before differencing	-2.0269	0.5654
After differencing	-5.9839	0.01

The results in Table 1 indicate that the original time series is not stationary. This is evident from the Augmented Dickey-Fuller (ADF) test, where the p-value is larger than the 5% significance level. However, the ADF test performed on the differenced data, as shown in Table 1, reveals that the p-value is significantly smaller than the 5% significance level. This suggests that the differenced series is stationary, and first order differencing (with a differencing order of 1, i.e., $d=1$) may be suitable. The ACF plot in Figure 4 and Table 1 indicate that ARIMA (1,1,1) is a suitable model. However, we compare it with three other ARIMA models which are ARIMA (0,1,0), ARIMA (0,1,1), and ARIMA (1,1,0). According to Table 2, the ARIMA (1,1,1) model appears to outperform the other models in forecasting petroleum prices in Tanzania, as evidenced by its superior log likelihood AIC, and BIC value. The AIC and BIC for the ARIMA (1,1,1) model are the lowest among the models considered. The results of all four ARIMA models are summarized in Table 2.

Table 2: Comparison of ARIMA model candidates based on Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC). The ARIMA(1,1,1) model demonstrates the best fit to the historical price data, as indicated by its minimized information criteria.

	ARIMA (0, 1, 0)	**ARIMA (1, 1, 1)**	**ARIMA (1, 1, 0)**	**ARIMA (1, 0, 1)**
Log likelihood	-709.8534	-704.4324	-721.8125	-714.9262
AIC	1421.707	1414.865	1447.625	1437.852
BIC	1424.443	1423.047	1453.0800	1448.832

A diagnostic test is conducted to evaluate how well the selected model fits the data. It can be observed in Figure 5 that the residuals are normally distributed and that there is no serial correlation. To assess the normality of the residuals, histograms and QQ plots are used. In Figure 5, the histogram fitted with density plot have bell-shaped curve, indicating normal distribution. Also, the points on QQ plot follow straight line. Additionally, the Box-Ljung test and the ACF of the residuals are examined and displayed in Figure 5 to determine whether any residual correlations exist.



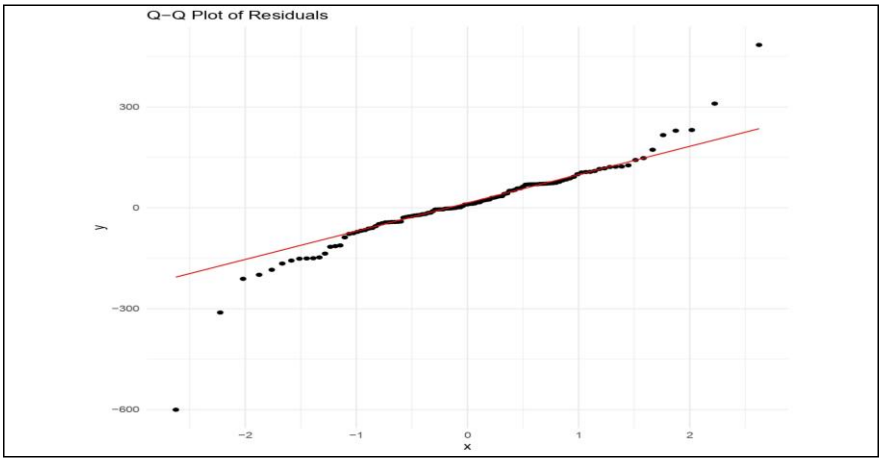


Figure 5: Diagnostic checks for the residuals of the selected ARIMA(1,1,1) model: First plot: Residuals plotted over time; Second plot: Histogram of the residuals with a normal distribution curve; Third plot: Normal Q-Q plot for assessing normality of residuals; Fourth plot: Autocorrelation function plot of the residuals.

Neural Network Autoregressive model input

The PACF plot in Figure 6, suggests that the NNAR model is capturing previous relationship at lag 11 and has an AIC value of 1259.7942.

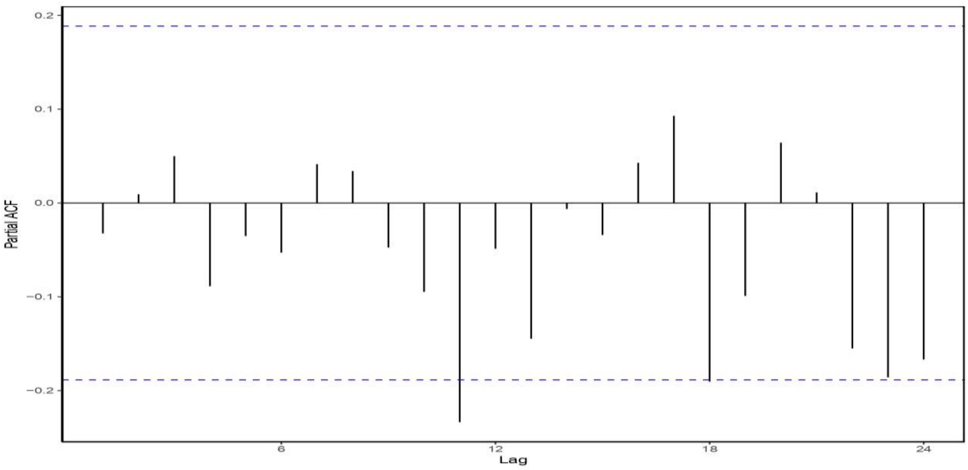


Figure 6: Partial Autocorrelation Function for determining NNAR model inputs. The PACF plot of the training data identifies significant lags, leading to the selection of 11 lagged values (p=11) as inputs for the neural network architecture.

From PACF plot in Figure 6, the training data model and testing data were determined. The analysis shows that the smallest AIC value was obtained at lag 11, indicating that the input layer should have 11 neurons. To determine the number of neurons in the hidden layer, the following formula was used, $k = \frac{(p+1)}{2} = \frac{(11+1)}{2} = 6,$ where p is the number of input neurons. Based on this calculation, the number of neurons in the hidden layer was set to 5, 6, and 7 to investigate the potential for overfitting.

Forecasting using Neural Network Autoregressive models

The forecasting of the wholesale gasoline price in Dar es Salaam, Tanzania will be conducted using three different neural network autoregressive models. The specific models to be used are; NNAR(11,5), NNAR(11,6), and NNAR(11,7). The forecasting will cover the period from January 1 2025 to December 31, 2026. Table 3 compares the performance of three different NNAR models, each with varying number of neurons (5, 6, 7), in terms of MAPE and RMSE for both training and testing. NNAR(11,5) shows the highest MAPE(1.0806% training, 3.4492% testing) and RMSE (28.9468 training, 120.6338 testing), indicating it performs the least well out of the three models. NNAR(11,6) improves on the performance with lower MAPE in training (0.8233%) but slightly higher MAPE in testing (3.5500%). The RMSE values also show a reduction in training (23.3025) but an increase in testing (126.7358), implying that while the model fits better to the training data, it has average generalization to unseen data. NNAR(11,7) exhibits the best performance overall, with the lowest MAPE (0.6543% training, 2.3447% testing) and RMSE (19.2312 training, 76.6797 testing), suggesting this configuration provides the best balance between training accuracy and generalization.

Table 3: Comparison of forecasting accuracy (MAPE and RMSE) for NNAR models with varying hidden layer sizes (NNAR(11,5), NNAR(11,6), and NNAR(11,7)). The NNAR(11,7) model demonstrates the best generalization performance on the testing set.

Model	MAPE (%)		RMSE	
	Training	Testing	Training	Testing
NNAR (11,5)	1.0806	3.4492	28.9468	120.633[illegible]
NNAR (11,6)	0.8233	3.5500	23.3025	126.735[illegible]
NNAR (11,7)	0.6543	2.3447	19.2312	76.679[illegible]

Neural Network Autoregressive Model Architectures

To visually represent the candidate Neural Network Autoregressive (NNAR) architectures evaluated in this study, architectural diagrams for the three finalists are presented below. The structure of a $NNAR(p,k)$ model is defined by two key parameters: p, the number of lagged inputs (neurons in the input layer), and k, the number of neurons in the single hidden layer. As detailed in the methodology, the input parameter was determined to be $p=11$ based on the analysis of the Partial Autocorrelation Function (PACF). Subsequently, the optimal number of hidden neurons k was empirically tested. The following figures illustrate the architectures of the three competing models NNAR(11,5), NNAR(11,6), and NNAR(11,7) which were trained and compared to identify the configuration with the best predictive performance on the holdout sample.

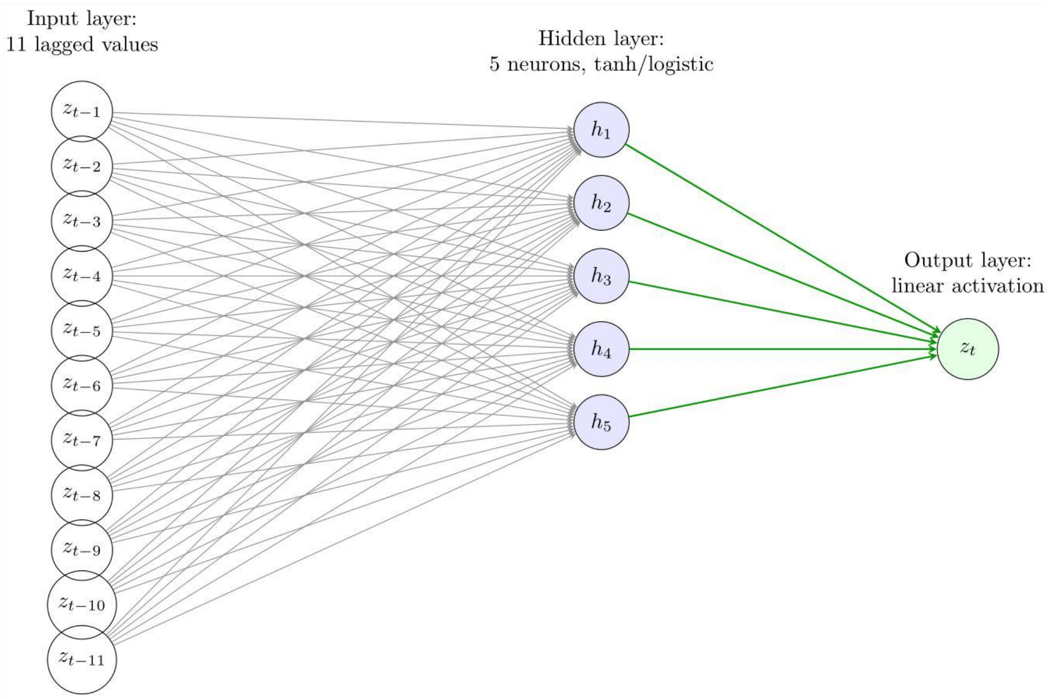


Figure 7: Architectural diagram of the NNAR(11,5) model. The network uses the previous 11 lagged values of the gasoline price series as inputs (input layer) to learn complex patterns through a hidden layer with 5 neurons, producing a one-step-ahead forecast (output layer).

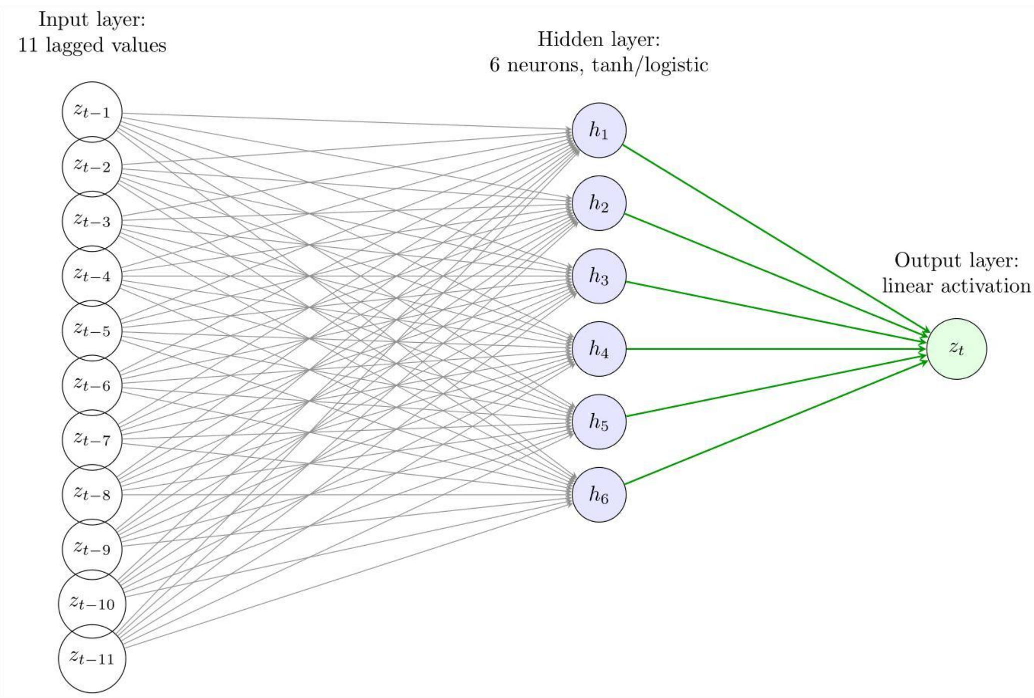


Figure 8: Architectural diagram of the NNAR(11,6) model. The model structure features 11 input nodes (lagged values), a single hidden layer containing 6 neurons where non-linear transformations occur, and a single output node for the forecasted value.

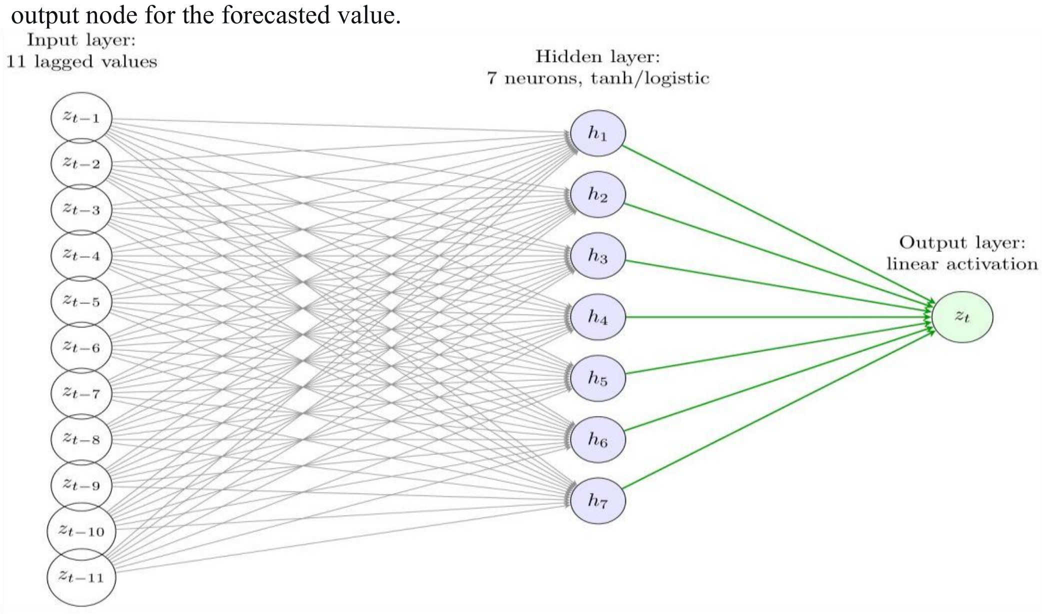


Figure 9: Architectural diagram of the optimal NNAR(11,7) model. This configuration, with 7 neurons in the hidden layer, demonstrated the best forecasting performance by effectively capturing the non-linear patterns in the gasoline price time series.

Further, Figure 10 shows that the NNAR(11,7) model is able to capture the general trend of the gasoline price data, and fits the data well.

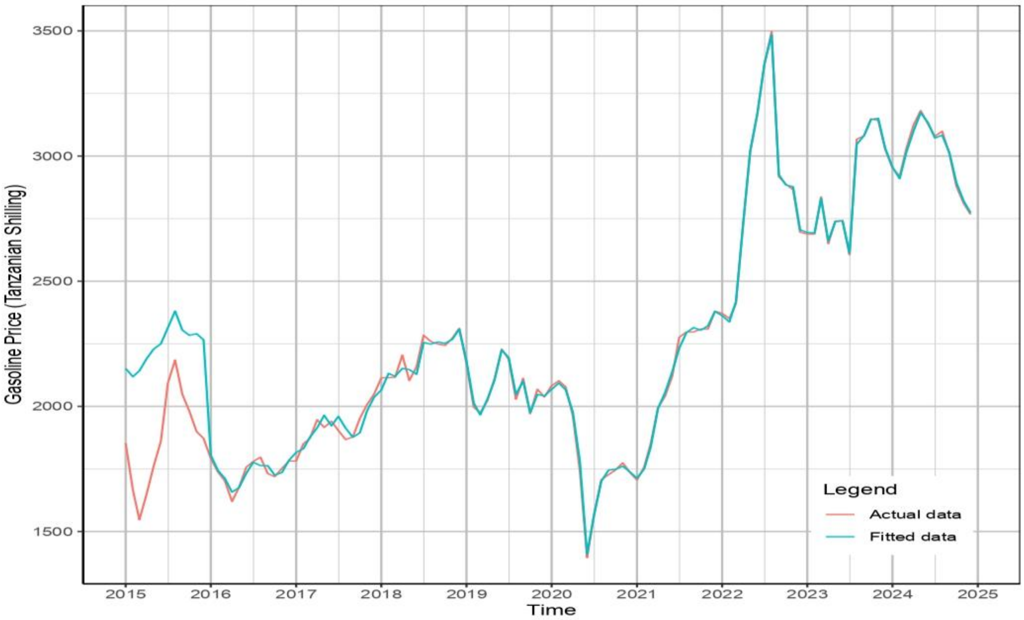


Figure 10: In-sample fit of the NNAR(11,7) model to the historical gasoline price data. The close alignment between the fitted values (red line) and the actual data (blue line) demonstrates the model's strong capability in capturing the underlying non-linear trends and volatilities.

Comparing ARIMA(1,1,1) and NNAR(11,7) Models for Gasoline Price Forecasting

Table 4 compares the performance of two models, ARIMA(1,1,1) and NNAR(11,7), in forecasting wholesale gasoline prices, based on MAPE and RMSE for both training and testing phases. ARIMA(1,1,1) shows higher errors compared to NNAR in both MAPE and RMSE. Its training MAPE is 3.7538%, while the testing MAPE is lower at 2.3447%. However, the RMSE values are high, with 115.5199 for training and 95.8016 for testing. NNAR(11,7) demonstrates better performance overall, with a much lower training MAPE of 0.6543% and a testing MAPE of 2.049%, which is moderately better than ARIMA in the testing phase. The RMSE values for NNAR(11,7) are significantly lower, with 19.2312 in training and 76.6797 in testing, indicating a better fit and generalization.

The performance of the ARIMA(1,1,1) and NNAR(11,7) models was evaluated based on their forecasting accuracy on a holdout test set. The initial 80% of the data (January 2015 to December 2023) was used for model training, while the remaining 20% (January 2024 to December 2024) was retained as an out of sample test set to validate the models. The metrics reported in Table 4 for the testing phase therefore compare the model forecasts against the actual known values within this holdout period (2024), providing a robust evaluation of their true predictive performance on unseen data.

Table 4: Performance comparison of ARIMA(1,1,1) and NNAR(11,7) models. The NNAR model demonstrates superior forecasting accuracy for wholesale gasoline prices, as evidenced by lower MAPE and RMSE values across both training and testing phase

.Model	MAPE(%)		RMSE	
	Training	Testing	Training	Testing
ARIMA(1,1,1)	3.7538	2.049	115.5199	95.8016
NNAR(11,7)	0.6543	2.3447	19.2312	76.6797

Figure 10 shows the ARIMA(1,1,1) model's forecast for gasoline prices over the next 24 months, with actual (blue line), fitted (red dotted line), and forecasted values (red dashed line). The shaded areas indicate 80% and 95% prediction intervals, showing the uncertainty of the forecast. The forecast remains relatively stable with small fluctuations, and the intervals widen as the forecast horizon increases, reflecting increasing uncertainty in predictions over time.

In Table A1, the ARIMA(1,1,1) model forecasts gasoline prices starting from January 2025, showing a gradual increase from approximately 2716.78 TZS in January to about 2856.50 TZS by December 2026. The 80% confidence intervals indicate a range between approximately 1792.60 TZS and 4133.93 TZS by the end of the forecast period, suggesting a moderate level of uncertainty. The 95% confidence intervals widen the range further, reflecting greater uncertainty in the long-term forecast.

In Table A2, the NNAR(11,7) model shows a slightly higher forecast for gasoline prices, starting at approximately 2756.06 Tanzanian Shillings (TZS) in January 2025 and reaching around 2737.84 TZS by December 2026. This indicates a more stable forecast compared to the ARIMA model, the results which has been reflected in Figure 11 for NNAR(11,7). Similar to the ARIMA results, the NNAR model also provides 80% and 95% confidence intervals. The 80% intervals range from 2593.26 TZS to 2892.25 TZS, while the 95% intervals suggest a similar level of uncertainty in the long-term predictions.

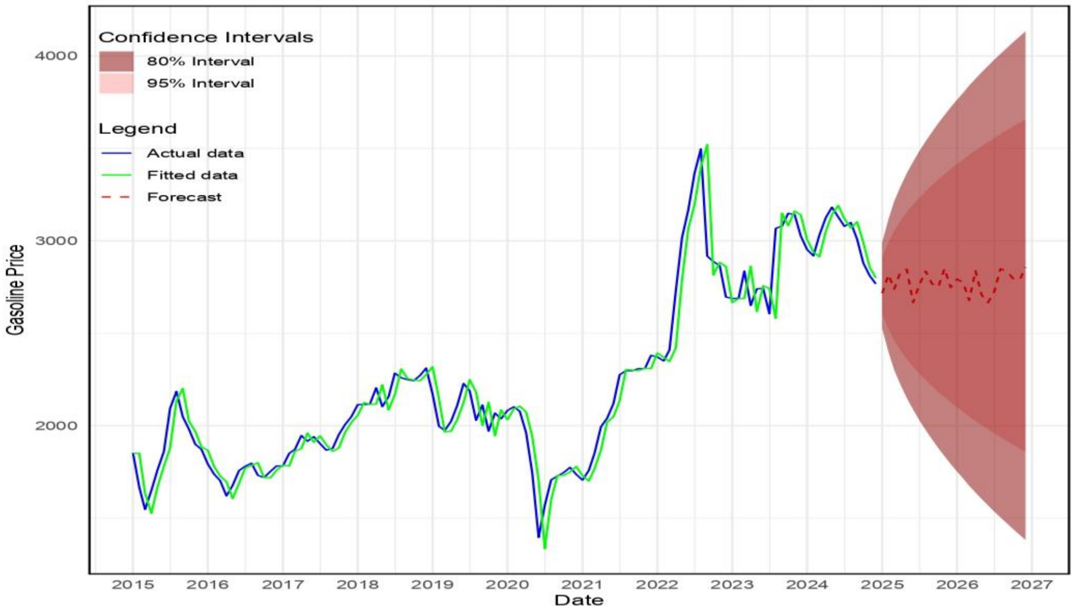


Figure 10: Twenty-four-month forecast of wholesale gasoline prices generated by the ARIMA(1,1,1) model. The forecast exhibits a gradual upward trend with fluctuations, while the widening prediction intervals indicate growing uncertainty over the forecast horizon.

Figure 11: Twenty-four-month forecast of wholesale gasoline prices using the optimal NNAR(11,7) model. The forecast displays a more stable and consistent trajectory compared to the ARIMA model, reflecting the NNAR's superior ability to capture underlying complex patterns.

Discussion

This study set out to model and forecast wholesale gasoline prices in Tanzania using two key time-series approaches: the linear ARIMA framework and the nonlinear Neural Network Autoregressive model. The findings reveal several important insights concerning the dynamics of gasoline prices in an import-dependent economy and the comparative suitability of linear versus nonlinear models when forecasting highly volatile economic series.

The analysis of monthly wholesale gasoline prices from 2015 to 2024 shows that the Tanzanian market is heavily influenced by global economic conditions. The sharp decline in 2020 corresponds with the COVID-19 pandemic, which dramatically reduced global fuel demand, while the pronounced surge in 2022 coincides with the Russia–Ukraine war that destabilized energy markets worldwide (Yagi and Managi 2023). These disruptions created structural breaks that contributed to the pronounced non-linearity in the price series, underscoring the vulnerability of Tanzania's petroleum supply chain to international shocks. The observed volatility confirms that wholesale prices more than retail prices directly reflect global cost movements, validating the study's focus on wholesale dynamics.

The linear ARIMA models were first evaluated, with ARIMA(1,1,1) emerging as the best-fitting candidate based on AIC and BIC values. Diagnostic checks further confirmed that the model's residuals were approximately normal and free from serial correlation, indicating an adequate linear representation of the differenced series. However, despite its statistical adequacy, the ARIMA model exhibited limitations in capturing sharp fluctuations and nonlinear patterns, as reflected in higher RMSE and MAPE values during out-of-sample testing. This behaviour aligns with the inherent constraints of ARIMA models, which assume linearity and are less effective in environments where abrupt structural shocks occur (Zhang 2003).

In comparison, the NNAR models demonstrated stronger predictive performance, with the NNAR(11,7) configuration outperforming both the ARIMA model and the other neural network configurations tested. Its superior accuracy particularly the lower RMSE of 76.6797 on the testing set indicates that neural networks are better suited to recognize complex, nonlinear, and volatile trends characteristic of petroleum price series. The NNAR model also produced more stable long-term forecasts, with smoother trajectories and comparatively narrower confidence intervals than ARIMA. This suggests that NNAR's flexibility in approximating nonlinear relationships allowed it to capture hidden patterns associated with global disruptions, making it a more robust forecasting tool for Tanzania's fuel market.

The results from this study are consistent with previous research in Tanzania and internationally. Studies such as Gasper and Mbwambo (2023) and Ntare (2023) found that ARIMA models can adequately forecast Tanzanian petroleum prices in stable periods but struggle to account for abrupt market shocks similar to the limitations observed in the ARIMA(1,1,1) model in this study. Conversely, research by Kafuku (2025), Vijayalakshmi et al. (2023), and Daniyal et al. (2022) revealed that ANN and NNAR models consistently outperform ARIMA in forecasting non-linear and volatile series, particularly in energy and agricultural contexts. The superior performance of the NNAR(11,7) model in this study corroborates these findings and reinforces the argument that neural-network-based approaches are more effective than traditional linear models in markets subject to external geopolitical and economic disturbances (Sharif et al. 2020).

The practical implications of these results are significant for policymakers, importers, and regulatory bodies such as EWURA. The demonstrated superiority of NNAR models suggests that regulators should integrate advanced machine-learning forecasting tools into their decision-making processes to improve the accuracy of price projections and reduce uncertainty in import planning. Accurate forecasts are crucial for budgeting, fuel import scheduling, stabilization policy, and minimizing price shocks that may affect transport, manufacturing, and consumer welfare. Furthermore, the widening confidence intervals observed in multi-step forecasts highlight the need for constant real-time updates to maintain accuracy, especially in an unpredictable global energy environment.

While the NNAR(11,7) model achieved strong performance, it is important to acknowledge that both the ARIMA and NNAR models used in this study relied solely on historical price data. The absence of exogenous variables such as global oil benchmarks, exchange rates, inflation, freight charges, and geopolitical risk indicators limits the models' ability to anticipate structural breaks before they occur. Future studies should consider ARIMAX or NNARX models incorporating these variables to enhance predictive power. Additionally, exploring deep learning architectures such as Long Short-Term Memory (LSTM) or Gated Recurrent Unit (GRU) networks may yield even stronger long-term forecasting accuracy.

Conclusion

This study demonstrates the superiority of the NNAR(11,7) model over the ARIMA(1,1,1) model in fitting historical trends and generating accurate short-term, out-of-sample forecasts, as rigorously validated on the 2024 holdout sample. This robust performance is largely due to the model's ability to capture the complex, non-linear patterns observed in the data. While the model projects a potential price trajectory through 2026 (see Figure 11, Table A2), long-term forecasting remains inherently uncertain due to unpredictable global economic and geopolitical shocks. These projections should be interpreted as a plausible scenario based on current trends and patterns rather than a fixed outlook; they are highly susceptible to unforeseen structural breaks in the global energy market.

While the NNAR(11,7) model is designed for one-step-ahead prediction, multi-step forecasts in this study (24 months ahead) are generated iteratively. In this approach, each subsequent forecast uses the model's previous predictions as inputs, allowing the model to project further into the future. It is important to note that this iterative method leads to widening prediction intervals over the forecast horizon, reflecting the accumulation of uncertainty as can be seen for Figure 11 and Table A2. This means that the model is superior in short term forecast.

There are several potential research problems related to the present paper. First, the models could be extended by incorporating exogenous variables such as global crude oil prices, the exchange rate, and domestic inflation indices to create ARIMAX or NNARX frameworks, which would better quantify the impact of external economic shocks on local prices. Second, the integration of datasets capturing geopolitical risk metrics and climate conditions could further enhance the model's ability to anticipate volatility triggered by supply chain disruptions and environmental factors. Finally, beyond the current NNAR architecture, more sophisticated deep learning techniques, such as LSTM networks, could be investigated to capture even more intricate long-term temporal dependencies and non-linear patterns within the data, potentially leading to a significant improvement in forecasting precision.

Data availability

The data that support the findings of this study are publicly available from the Energy and Water Utilities Regulatory Authority (EWURA) of Tanzania. The monthly wholesale gasoline price data for Dar es Salaam from January 2015 to December 2024 can be accessed directly from their official fuel prices portal at: https://www.ewura.go.tz/fuel-prices/.

Conflict of Interests

No conflicts of interest declared by the authors.

Source of Fund

Edward Ngailo acknowledges financial support of Swedish International Development Cooperation Agency (SIDA) through TZ-Sweden Bilateral Program for Research.

Acknowledgements

Authors would acknowledge the Energy and Water Utilities Regulatory Authority (EWURA) for the availability of data

References

Aggarwal CC 2018 Neural Networks and Deep Learning: A Textbook, 1st ed, Springer International Publishing, Cham, Switzerland. https://doi.org/10.1007/978-3-031-29642-0

Box GEP and Jenkins GM 1970 Time Series Analysis, Forecasting and Control, Holden Day.

Box GEP and Jenkins GM 1974 Time Series Analysis, Forecasting and Control, Holden Day.

Box GEP and Jenkins GM 1976 Time Series Analysis: Forecasting and Control, Revised Edition, Holden Day.

Box GE, Jenkins GM, Reinsel GC and Ljung GM 2015 *Time series analysis: forecasting and control*. Wiley, Wisconcin-Madison.

Brockwell PJ and Davis RA 2016 *Introduction to time series and forecasting, 3rd ed,* Springer, New York. https://doi.org/10.1007/978-3-319-29854-2

Cavanaugh JE and Neath AA 2019 The Akaike information criterion: Background, derivation, properties, application, interpretation, and refinements. *Wiley Interdiscip. Comput. Stat.* 11(3): e1460. https://doi.org/10.1002/wics.1460

Daniyal M, Tawiah K, Muhammadullah S and Opoku-Ameyaw K 2022 Comparison of Conventional Modeling Techniques with the Neural Network Autoregressive Model (NNAR): Application to COVID-19 Data. *J. Healthc. Eng.* 2022(1): 1–9

Desiyanti V, Rahayu YD and Umilasari R 2022 Comparisonal Analysis of DMA and DES (HOLT) Methods in Forecasting GKP Prices At Farmers Level. *J. Smart Tekn.* 3(5): 552–559

EkmiŞ MA, HekiMoğlu M and Bülbül BA 2017 Revenue forecasting using a feed-forward neural network and Arima model. *Sigma J. Eng. Nat. Sci.* 8(2): 129–134

Energy and Water Utilities Regulatory Authority (EWURA) n.d. Cap prices. Retrieved December 5, 2024, from https://www.ewura.go.tz/fuel-prices/

Hyndman RJ and Khandakar Y 2008 Automatic time series forecasting: The forecast package for R. *J. Stat. Softw.* 27(3): 1-22.

Gasper L and Andwilile AG 2023 Modelling Petroleum Prices in Tanzania. A comparative analysis between ARIMA and Holt's Method. *Int. J. Manag. Account. Econ*. 10(9).

Gasper L and Mbwambo H 2023 Forecasting Crude Oil Prices By Using ARIMA Model: Evidence From Tanzania. *J. Account. Finance Audit. Stud*. 158–175

Ghatak A 2019 Deep Learning with R, 1st ed, Springer Singapore, Singapore.

Mansoorian HJ, Karimaee M, Hadi M, Porazmey EJ, Barati F, Baziar M 2017 Feed Forward Artificial Neural Network Model to Estimate the TPH Removal Efficiency in Soil Washing Process. *Arch. Hyg. Sci*. 6(1): 96–104.

Junita TP and Kartikasari MD 2024 Application of the Neural network Autoregressive (NNAR) method for forecasting the value of oil and gas exports in Indonesia. *BAREKENG: J. Ilm. Mat. Terap*. 18(1): 0341–0348.

Kafuku JM 2025 Application of Artificial Neural Network Models for Predicting Diesel and Petrol Prices in the Geographically Sparsed Regions in Tanzania. *Tanz. J. Engrg. Technol. 44*(1), 56-71. https://doi.org/10.52339/tjet.v44i1.1082

Khashei M and Bijari M 2010 An artificial neural network (p,d,q) model for time series forecasting. *Expert Syst. Appl*. 37(1): 479–489.

Khodadadi M, Mesdaghinia A, Nasseri S, Ghaneian MT, Ehrampoush MH, Hadi M 2016 Prediction of the waste stabilization pond performance using linear multiple regression and multi-layer perceptron neural network: a case study of Birjand, Iran. *Env Health Eng Manage*. 3(2): 81–89.

Kojima M, Matthews K and Sexsmith F 2010 Petroleum Markets in Sub-Saharan Africa. Working paper, The World Bank.

Lo KL and Wu YK 2003 Risk assessment due to local demand forecast uncertainty in the competitive supply industry. *IEE Proc. Gener. Transm. Distrib*. 150(5): 573–582.

Maleki A, Nasseri S, Aminabad MS, Hadi M 2018 Comparison of ARIMA and NNAR Models for Forecasting Water Treatment Plant's Influent Characteristics. *KSCE J. Civ. Eng*. 22(9): 3233–3245.

Ntare PC 2023 Prediction of Petrol and Diesel Prices in Dar Es Salaam using ARIMA models. *Afr. J. Appl. Res.* 9(2): 82–89.

R Core Team 2024 *R: A language and environment for statistical computing.* R Foundation for Statistical Computing.

Rostami FZ, Miesdaghinia AR, Nadafi K, Nabizadeh NR, Mahvi AH and Hadi M 2015 Forecasting the air quality index based on meteorological variables and autocorrelation terms using artificial neural network. *Razi J Med Sci.* 22(137): 31-43.

Sharif A, Aloui C and Yarovaya L 2020 COVID-19 pandemic, oil prices, stock market, geopolitical risk and policy uncertainty nexus in the US economy: Fresh evidence from the wavelet-based approach. *Int Rev. Finance Anal.* 70: 101496. doi: 10.1016/j.irfa.2020.101496Get rights and content

Shumway RH and Stoffer DS 2017 Time Series Analysis and Its Applications: With R Examples, 4th ed, Springer

International Publishing, Cham.

Thoplan R 2014 Simple and Sophisticated Methods of Forecasting for Mauritius Monthly Tourist Arrival Data. *Int. J. Stat. Appl*. 4(5): 217–223.

Venables WN and Ripley BD 2002 Modern Applied Statistics with S. Fourth Edition. Springer.

Vijayalakshmi G, Pushpanjali K and Babu AM 2023 A comparison of ARIMA & NNAR models for production of rice in the state of Andhra Pradesh. *Int. J. Stat. Appl. Math*. 8(3): 251–257.

Wah YB, Berry MW, Mohamed A and Al-Jumeily D 2023 Data Science and Emerging Technologies.*Springer Nature*.

World Health Organization (WHO) 2020 Coronavirus disease 2019 (COVID-19). Situation Report-55. World Health Organization https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports Accessed on October. 16, 2024.

Yagi M and Managi S 2023 The spillover effects of rising energy prices following 2022 Russian invasion of Ukraine. *Econom. Anal Policy*. 77: 680-695. doi: 10.1016/j.eap.2022.12.025

Yang S and Fu Y 2025 Interconnectedness among supply chain disruptions, energy crisis, and oil market volatility on economic resilience. *Energy Econ.* 143: 108290. doi: 10.1016/j.eneco.2025.108290

Zhang G, Patuwo EB and Hu MY M 1998 Forecasting with artificial neural networks: The state of the art. *Int. J. Forecast*. 14(1): 35–62.

Zhang GP 2003 Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing* 50: 159-175. doi:10.1016/S0925-2312(01)00702-0

Zhang GP and Qi M 2005 Neural network forecasting for seasonal and trend time series. *Eur. J. Oper. Res*. 160(2): 501–514.