Time Series Modelling of DAP Fertilizer Demand in Kenya

James Mwiti Mutegi, Moses Mahugu Muraya and Elizabeth W. Njoroge
Department of Physical Sciences, Chuka University, P.O. Box 109-60400 Chuka Kenya
***Correspondence Email:* mwitiwamutegi@gmail.com**

ABSTRACT
Agriculture for many years has remained to be the backbone of the Kenya's economy and has been a source of livelihood of the population in rural areas. The report "Agricultural Sector Development Strategy (2022) points that agriculture is the mainstay of the Kenyan economy contributing 26% of the GDP and another 25% indirectly. However, agricultural productivity has stagnated in recent years due to various constraints including poor agronomic practices such as fertilizer application. Fertilizers are an important component in increasing agricultural output. One of the major constraints is understanding the demand for fertilizer across different agroecological zones and timely delivering of fertilizer to the farming communities. Hence, Kenya struggles to make this agricultural input adequate because the country does not manufacture it but relies on imported fertilizer. This study investigated the trend and patterns of Diammonium Phosphate (DAP) fertilizer demand in Kenya within a span of 13 years to understand seasonal trends and underlying cycles and develop a SARIMA (p, d, q) (P, D, Q) s. The secondary monthly data from 2010-2023 that was obtained from the Ministry of Agriculture and Livestock Development headquarters, Nairobi, Kenya. R-software was then be utilized to analyse data for descriptive statistics, fertiliser demand variability and model fitting. The Box-Jenkins method (model identification, model estimation, model validation) was used to determine the best models for the data. The findings of the study pointed out that the demand for DAP fertiliser can be predicted by SARIMA (0,0,0) (2,0,0) [12] w/ mean. These findings are crucial as will play a significant role in promoting sustainable farming practices, influencing policy decisions, and contributing to discussions in agricultural economics.

Keywords: SARIMA Model, Autoregressive, Normality, Trend, Seasonality, Fertilizer Demand

INTRODUCTION

Agriculture for many years has remained to be the backbone of the Kenya's economy and has been a source of livelihood of the population in rural areas. The report "Agricultural Sector Development Strategy (2022) points that agriculture is the mainstay of the Kenyan economy contributing 26% of the GDP and another 25% indirectly. The Agricultural Sector Development Strategy (2022) outlines key objectives for the sector, with a primary focus on achieving an average growth rate of 7 percent per year over the next five years. One of the core strategies to achieve this growth is increasing productivity, commercialization, and the competitiveness of agricultural commodities. One of the main ways to improve productivity is through the effective application of fertilizers (Ritchie *et al.,* 2022). Fertilizers play a vital role in enhancing soil fertility, boosting crop yields, and ensuring that farmers can produce more with the same land resources. By improving the availability of fertilizers, Kenya can significantly raise the productivity of its agricultural sector, ensuring sustainable growth and competitiveness in the global market. Thus, fertilizer application is a key factor in realizing the sector's strategic mission of creating an innovative, commercially oriented, and modern agricultural framework. Different types of fertilisers are used in crop production in Kenya, which include calcium ammonium nitrate (CAN), diammonium phosphate (DAP), calcium nitrate, Muriate of potash, NPK, and Urea, among others (Ritchie *et al.,* 2022).

Predicting fertilizer demand presents several challenges, as different studies have shown varying levels of success depending on the methodology used. Fertilizer demand is a complex, multi-layered issue influenced by various factors such as weather patterns, economic conditions, government policies, and shifts in agricultural practices. These complexities make accurate forecasting difficult, and several studies highlight the limitations of current methods. Tenkorang *et al.* (2008) used simple linear regression to forecast global fertilizer demand in Asia. The study emphasized that accurate demand forecasting is crucial for ensuring long-term global food security and the profitability of the fertilizer industry. However, one of the shortcomings of linear regression is its assumption of a constant relationship between variables over time. This method does not account for trends, seasonality, or autocorrelation, which are often present in time series data. As McQuarrie and Tsai (1998) point out, this can lead to inaccurate forecasts since the method overlooks the sequential dependence inherent in time series data. Borkar (2023) addressed these challenges by using the ARIMA model to forecast fertilizer consumption in India, using data from 1950-2021. While ARIMA models are better suited for handling trends and autocorrelation, the study revealed that even with the optimal model challenges remain in capturing the complex dynamics of fertilizer demand. Bezerra *et al.* (2013) used both ARIMA and logistic function models to forecast fertilizer demand in Brazil for the 20 years. Although time series models like ARIMA can handle trends, they may still face challenges in adapting to sudden external shocks, such as changes in global markets or weather-related disruptions. In Thailand, Pisuttinusart *et al.* (2022) used the SARIMA model to forecast the demand for imported fertilizer because it accounts for seasonality and is more advanced than linear regression. In Kenya, studies by Sheahan *et al.* (2016) and Okello (2023) explored hybrid models, such as Seasonal ARIMA-GARCH, to forecast fertilizer prices and consumption. Despite the advancements in modelling techniques, it is evident from these studies that are highly reliable model for predicting fertilizer demand has not yet been realized. Challenges such as changing market conditions, external shocks, and the inherent variability in agricultural systems continue to pose obstacles to accurate forecasting. Accurately determining the amount of fertilizer to import will help address the issue of shortages. This can be achieved by forecasting the monthly fertilizer requirements. Models like regression, ARIMA, and SARIMA have been applied to demand forecasting. Although various forecasting methods are available, Kenya's agricultural production is seasonal, and demand fluctuates over time, making SARIMA model better for forecasting seasonal demand (Filder *et al*., 2019).

Seasonal ARIMA models are an extension of traditional ARIMA models, specifically designed to manage time series data with seasonal trends. They incorporate autoregressive (AR), integrated (I), and moving average (MA) elements along with seasonal components. SARIMA models are particularly useful for forecasting data that follows recurring time-based patterns. The model is represented as SARIMA (p, d, q) (P, D, Q), where the lowercase (p, d, q) refers to the non-seasonal components, and the uppercase (P, D, Q) _s represents the seasonal components (Shumway *et al.,* 2017). The general equation for SARIMA model is

$$Y_t = C + \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \cdots + \varphi_p Y_{t-p} - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_q \varepsilon_{t-q} + \varepsilon_t - \Phi_1 Y_{t-s} + \Phi_2 Y_{t-2s} + \cdots + \Phi_P Y_{t-Ps} - \Theta_1 \varepsilon_{t-s} - \Theta_2 \varepsilon_{t-2s} - \cdots - \Theta_Q \varepsilon_{t-Qs} + \varepsilon_t$$

where Y_t is the value of the time series at time t,

C is a constant or drift term (only included if the time series has a trend component).

p represents the order of the autoregressive (AR) component, and φ_i are the autoregressive coefficients.

d is the order of differencing, indicating the number of times differencing applied to make the time series data stationary.

q is the order of the moving average (MA) component, and θ_i are the moving average coefficients.

P and *Q* are the orders of the seasonal AR and MA components with corresponding coefficients. Φ_i and Θ_i.

One of the best-known methods for modelling SARIMA is the Box-Jenkins methodology, a systematic approach to identifying, estimating, and diagnosing time series models, particularly ARIMA and SARIMA models. This methodology was introduced by Box and Jenkins (1976) and is widely used due to its structured process, making it effective for time series modelling. The Box-Jenkins methodology is crucial for the SARIMA model, as it provides a structured framework to identify the seasonal and non-seasonal components of the time series data. For instance, in forecasting fertilizer demand in Kenya, seasonality plays a major role due to the agricultural cycles, making SARIMA an ideal model. The Box-Jenkins approach ensures that both seasonal (P, D, Q) and non-seasonal (p, d, q) parameters are identified correctly and that the model can adequately forecast future demand.

The Box and Jenkins (1976) entails four main steps in modelling:

Identification: The first step in the Box-Jenkins methodology is analysis of the time series to determine whether it is stationary. If the data is not stationary, differencing is applied, either non-seasonal, seasonal, or both, to achieve stationarity. Once stationarity is achieved, the autocorrelation function (ACF) and partial autocorrelation function (PACF) are used to identify the orders of the autoregressive (AR), moving average (MA), and seasonal components. For the SARIMA model, this process leads to determining the orders of (p), (d), (q) for the non-seasonal components, and (P), (D), (Q) for the seasonal components.

Estimation: After identifying the appropriate model, the next step is to estimate the parameters for the AR and MA components. Techniques such as maximum likelihood estimation (MLE) are commonly used for this purpose. The SARIMA model equation at this stage is as follows:

$$\Phi_P(B^S)\phi_p(B)(1-B^S)^D(1-B)^d y_t = \Theta_Q(B^S)\theta_q(B)\epsilon_t$$

Where $\Phi_p(B)$ is the non-seasonal AR component, $\theta_q(B)$ is the non-seasonal MA component, $\Phi_P(B^S)$ is the seasonal AR component, $\Theta_Q(B^S)$ is the seasonal MA component, B is the backshift operator and ϵ_t is the error term (white noise).

Diagnostic Checking: Once the parameters are estimated, diagnostic checks are conducted by analysing the residuals. The residuals should behave like white noise, indicating that the model has successfully captured the underlying data patterns. The residuals' ACF is checked to ensure no significant correlations remain, which would suggest a well-fitted model.

Forecasting: After successfully identifying and estimating the model, it can then be used to forecast future values. SARIMA models are especially well-suited for forecasting time series data that display both non-seasonal and seasonal trends.

In this study, SARIMA modelling was carried out using the Box-Jenkins methodology, with the primary aim of investigating trends in fertilizer demand in Kenya through time series analysis techniques. The study focused on analysing and detecting trends and seasonal patterns in the monthly fertilizer demand, specifically for Diammonium Phosphate (DAP). A key objective was to develop SARIMA model tailored to DAP fertilizer demand in Kenya and subsequently apply this model to forecast future demand, providing a robust framework for understanding and predicting fertilizer needs in the country.

METHODOLOGY

The study was conducted using data on fertilizer demand in Kenya, as recorded by the Ministry of Agriculture, covering the period from January 2010 to December 2023. Data inspection and analysis were carried out at Chuka University. A longitudinal observational research design was employed to investigate the patterns of fertilizer demand over the 13-year period. This approach allowed for the examination of changes and trends in DAP monthly fertilizer demand, providing a comprehensive understanding of the variations across different seasons and years.

For data collection, the study utilized secondary data from the Ministry of Agriculture and Livestock Development in Kenya. The use of official data ensured accuracy and representativeness, covering 168 months of DAP fertilizer consumption. This extensive dataset was ideal for the application of the Box-Jenkins methodology, which requires a minimum of 50 observations to develop reliable models (Box & Jenkins, 1976). The time series data was visualized using time series plots to highlight trends, seasonality, and irregularities in the fertilizer demand. The Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test was used to assess the stationarity of the data. Model identification followed, focusing on identifying the appropriate SARIMA components (p, d, q for non-seasonal and P, D, Q for seasonal) using ACF and PACF plots. The R program was employed to perform differencing and model fitting, including training and testing the model to optimize parameters and ensure accuracy. Once a satisfactory SARIMA model was established, it was used to forecast future fertilizer demand in Kenya.

RESULTS AND DISCUSSION

Descriptive statistics

Descriptive statistics For DAP Fertiliser demand (2010-2023)

Variable	n	Mean	Median	Min	Max	Skewness
DAP	168	17906.31	10962	0	1005 98	1.4

The descriptive statistics for Diammonium Phosphate (DAP) in metric tonnes show considerable variability in the data. With 168 observations, the mean value is 17,906.31 metric tonnes, indicating a high average demand. The median, at 10,962 metric tonnes, is lower than the mean, suggesting a right-skewed distribution. The dataset ranges from a minimum of 0 to a maximum of 100,598 metric tonnes, reflecting significant variability in demand. The positive skewness of 1.4 further supports this, indicating that while most values are clustered towards the lower end, there are a few very high values extending the distribution's tail to the right.

Visualisation of Diammonium Phosphate Time Series

Figure 1 is the time series plot of Diammonium Phosphate (DAP) fertiliser from January 2010 to 2023. The graph shows a cyclic pattern with recurring peaks and troughs, indicating varying levels of fertiliser demand. A notable drop in demand in 2010 and 2014 suggests periods of decreased demand. From 2015 to 2019, there is significant demand variability, with highly pronounced peaks indicating high demand for DAP in these years. However, in 2021, there was a significant drop in demand for DAP, which may have been attributed to government policies and the COVID-19 effects (Mutegi *et al*., 2024). This suggests that demand for DAP fertiliser is highly variable and influenced by seasonal requirements, market conditions, or external events.

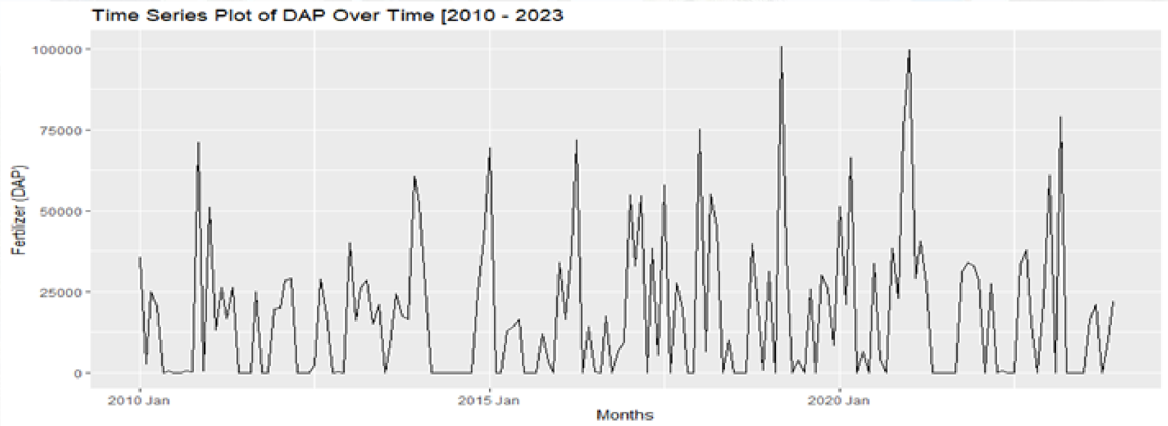

Figure 1: Diammonium phosphate time series plot 2010 – 2023

Normality Test of Diammonium Phosphate Time Series

A Jarque-Bera test was conducted to assess the normality of the Diammonium Phosphate data. The test yielded a test statistic of 78.566 with two degrees of freedom and a p-value less than $2.2e^{-16}$. Given the extremely small p-value, less than the critical $\alpha = 0.05$, we reject the null hypothesis that the data follows a normal distribution. This result indicates that the distribution of Diammonium Phosphate data deviates significantly from normality.

TABLE 2: JARQUE-BERA TEST

Statistic	Value
Test Statistic (X-squared)	78.566
Degrees of Freedom (d.f)	2
p-value	< 2.2e-16
Conclusion	Reject H_0: Data does not follow a normal distribution

In order to better understand the seasonality of the data, a plot of seasonal variation in calcium nitrate fertiliser was done.

Figure 2 is a GG-Season plot that overlays the DAP demand data for each year using differently coloured lines (See appendix 4). This plot highlights clear monthly trends, with noticeable peaks in January, March, and June and smaller peaks in October and December. The variability between years is evident, as each line represents a different annual pattern, suggesting significant shifts in demand. This variability implies that external factors such as weather, market conditions, or policy changes significantly impact DAP demand.

Seasonal decomposition of DAP time series

The graph in Figure 4 demonstrates the seasonal variation in Diammonium Phosphate (DAP) usage over 2010-2023, with each line representing a different year and distinguished by colours. The data shows a clear seasonal pattern, with two noticeable peaks in January and March, indicating higher usage during these months. Usage drops significantly from April to August, reflecting a trough during the summer months, followed by a gradual increase towards the year's end. The variations among the years are relatively consistent, with some years exhibiting slightly higher or lower usage during peak months.

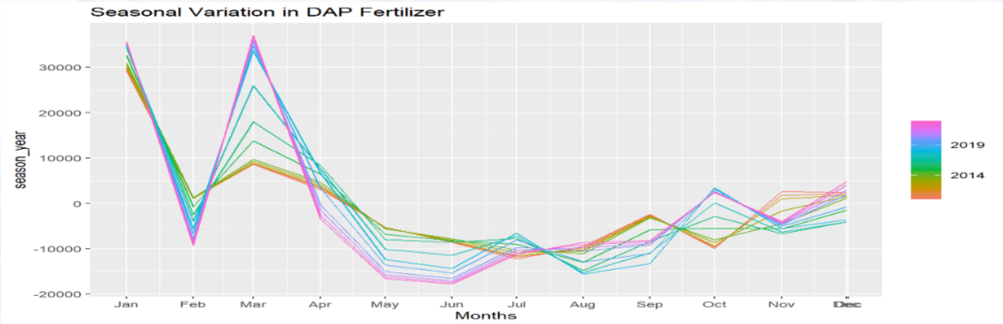

Figure 4: Seasonal variation in diammonium phosphate

The characteristics of the demand patterns are attributed to the pre-planting period when farmers prepare the soil for the growing season. A significant drop in usage is observed from April to August, corresponding to the summer months when fertiliser application typically decreases. The demand for DAP rises towards the end of the year due to preparations for the next planting season and replenishing soil nutrients. While the seasonal pattern remains consistent across the years, individual years have slight variations in peak consumption because of the changes in agricultural practices or climatic conditions that influence fertiliser application.

Figure 5 divides the plot into twelve subplots, each corresponding to a different month. The y-axis represents DAP usage within each subplot, while the x-axis covers 2010 to 2023. The black lines depict the monthly usage trends over the years, and the blue lines show the average usage for each month. This plot emphasizes the monthly seasonality of DAP usage, with notable spikes in January and March, aligning with the findings of the graph in Figure 4. The trend lines indicate a cyclical pattern with high demand in the early months, a decline in the middle of the year, and a rise towards the year's end.

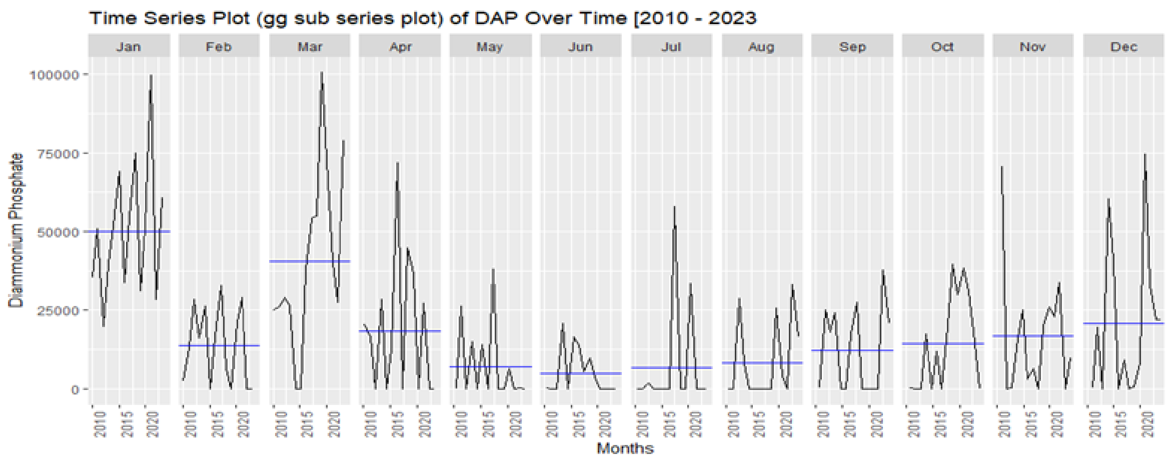

Figure 5: Subplots of diammonium phosphate

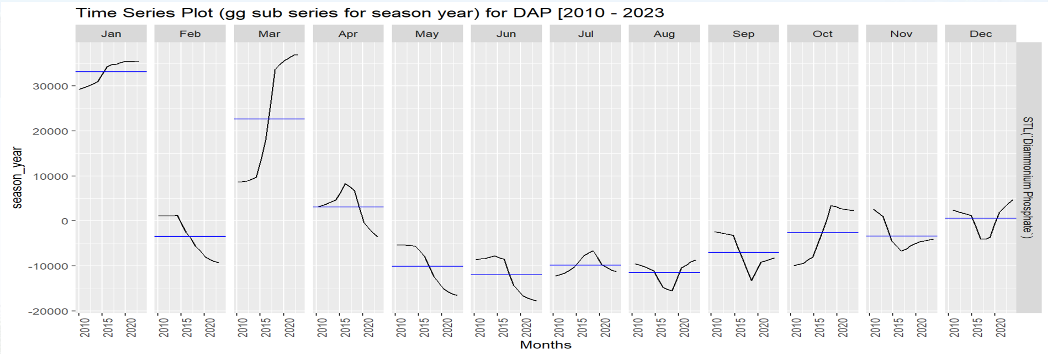

Figure 5: Subplots of diammonium phosphate

The findings are that the demand for DAP is highest in January and March, which confirms the findings in figure 4. Figure 6, on the other hand, displays a seasonal trend decomposition using LOESS (STL) of the DAP time series. This decomposition divides the data into four components: the observed data, the trend, the seasonal component, and the remainder (residuals). The first panel shows the raw data, which includes evident periodic spikes. The second panel illustrates the trend component, showing a general increase in DAP usage up to around 2017, followed by a slight decline. The third panel highlights the seasonal component, which reveals a consistent pattern of high usage in the early months and lower usage in the middle of the year across the entire timeframe. The final panel represents the remainder component, capturing the irregularities or noise not accounted for by the trend or seasonality.

Figure 6: STL Decomposition diammonium phosphate fertiliser time series

The STL (Seasonal and Trend decomposition using Loess) analysis of the Diammonium Phosphate (DAP) time series from 2010 to 2023 reveals both a trend and seasonality in the data. The analysis shows an upward trend, indicating a gradual rise in DAP demand over the years. This suggests that, even with seasonal fluctuations, the overall demand for fertiliser has been increasing over the years. The seasonal component displays recurring higher consumption patterns during certain months, corresponding with planting cycles. These findings highlight the seasonal variations and long-term increase in DAP demand.

The three graphs reveal that the Diammonium Phosphate (DAP) time series exhibits a strong seasonal pattern, with high usage in January and March and low usage during summer (June to August). The trend component indicates a general increase in DAP usage over the years, reaching a peak around 2017 before a slight decline. Figures 4 and 5 show that the consistent seasonal pattern suggests a predictable cyclic behaviour in fertiliser usage. The STL decomposition further explains this cyclic behaviour by separating it from long-term trends and random fluctuations, providing a clearer understanding of the factors driving DAP usage. The data demonstrates clear seasonality, which may be influenced by agricultural cycles, potential policy changes, and market dynamics that affect fertiliser application practices.

Testing for stationarity Diammonium Phosphate time series

The KPSS (Kwiatkowski-Phillips-Schmidt-Shin) test was conducted to check the stationarity of the time series data. The test returned a KPSS statistic of 0.1985524 and a p-value (KPSS p-value) of 0.1. The null hypothesis of the KPSS test is that the data is stationary, while the alternative hypothesis is that it is non-stationary. Since the p-value exceeds the common significance level (e.g., 0.05), we fail to reject the null hypothesis, indicating that the time series data is likely stationary. The test on the number of differencing needed was done using are, and the test results indicated that no differencing was needed. The ACF and PACF plots below indicate a minimal correlation between the time series and its lags and the partial correlation coefficients between the series and its lags.

The Autocorrelation Function (ACF) plot for DAP demand in Figure 7 shows the correlation of the time series with its lagged values over different time lags. The plot reveals a significant positive correlation at lag one and another significant spike at lag 12. The positive spike at lag 1 indicates that the current demand for DAP is positively correlated with the previous month's demand, indicating some persistence in the demand pattern. The spike at lag 12, on the other hand, indicates a seasonal pattern, suggesting that demand one year ago (12 months prior) strongly influences the current demand.

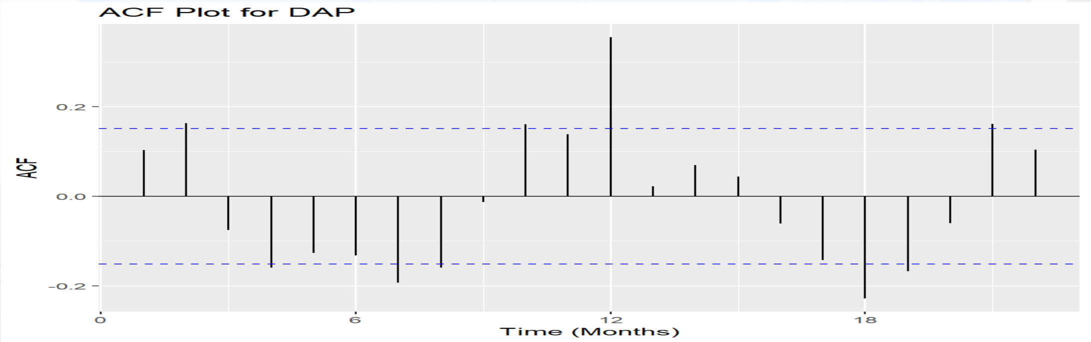

Figure 7: ACF diammonium phosphate

The Partial Autocorrelation Function (PACF) plot in Figure 8 is for DAP demand. It presents the extent of correlation between the time series and its lagged values after removing the effects of intermediate lags. The PACF plot indicates a significant positive partial autocorrelation at lag 1, which implies that the previous month's demand directly influences the current month's demand, with no significant influence from the intermediate months. The spike at lag 12 in the PACF plot suggests that after controlling for the effects of lags 1 through 11, the demand from 12 months prior still directly influences the current demand, reinforcing the seasonal pattern observed in the ACF plot in Figure 7.

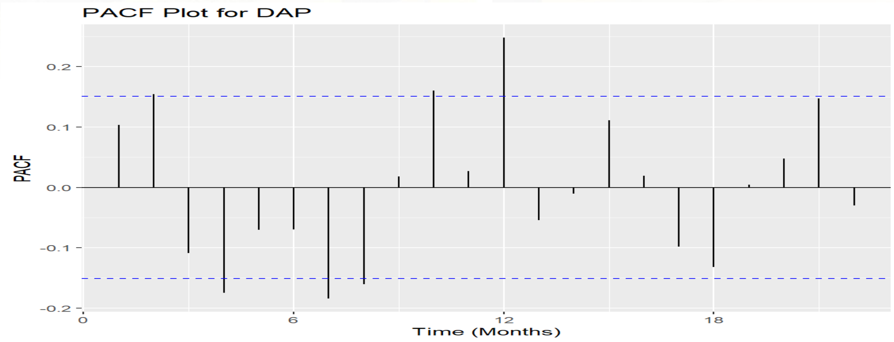

Figure 8: PACF diammonium phosphate

The ACF and PACF plots in Figures 8 and 9 suggest that the DAP demand series exhibits trend and seasonality. The trend is evident from the significant autocorrelation at lag 1, indicating that past demand values have a lasting effect on future demand. The spikes at lag 12 in both plots indicate the seasonality, pointing to a yearly cyclical pattern in demand. These characteristics suggest that a seasonal ARIMA model is appropriate for modelling this time series, incorporating terms that account for both the autoregressive behaviour and the seasonal effects. The significant lags observed in both ACF and PACF plots guide the selection of the order of the SARIMA model.

The seasonal patterns observed in DAP demand align with Kenya's bimodal agricultural calendar, where the long rains (March to May) and short rains (October to December) dictate planting seasons. DAP is used during the planting phase, which is why there is higher demand during these periods. The periods of zero demand likely correspond to off-seasons when top dressing is unnecessary. The National Fertiliser Subsidy Program, particularly the subsidy allocation changes, caused farmer preference shifts. The demand for DAP (Diammonium Phosphate) decreased in 2010 and 2014, likely due to shifts in government policies or cropping patterns. For instance, shifting from an e-voucher system to direct distribution through the National Cereals and Produce Board (NCPB) may have altered accessibility, affecting demand (Njagi et al. 2024).

DAP is primarily used during planting, especially for cereals like maize and rice (Yaseen et al., 2023). The high demand observed in January and March corresponds with the short rainy season, essential for crop establishment in Kenya. The consistent demand throughout the rest of the year suggests a balanced application across various crops to maintain soil fertility. Kenya's agricultural practices, which vary significantly by region and crop type, also explain the variability in fertiliser demand. DAP is essential for maize and other staple crops. For example, the high demand for DAP in January and March corresponds with the preparation for the long rains, a critical period for maize planting.

Model building

Diammonium Phosphate; SARIMA (0,0,0) (2,0,0) [12] w/ mean

This model includes two seasonal autoregressive terms: SAR1, estimated at 0.309194, and SAR2, estimated at 0.205886. Both terms are statistically significant, with SAR1 having a p-value of 9.74E-05 and SAR2 having a p-value of 9.75E-03 (Table 8). The constant term is estimated at 8432.586 and is highly significant, with a p-value of 7.22E-09. The variance of the residuals is estimated at 395,584,183. The model's log-likelihood is -1901.23, with an Akaike Information Criterion (AIC) of 3810.46, a corrected AIC (AICc) of 3810.7, and a Bayesian Information Criterion (BIC) of 3822.95. These results show that the seasonal autoregressive and constant terms play a significant role in the model. In addition, the model includes a mean term, which implies that the model accounts for a constant average level in the time series.

TABLE 3: DIAMMONIUM PHOSPHATE SARIMA MODEL

Model	Term	Estimate	Std. error	t-statistic	p-value
ARIMA	sar1	0.309194	7.74E-02	3.992806	9.74E-05
ARIMA	sar2	0.205886	7.88E-02	2.614315	9.75E-03
ARIMA	constant	8432.586	1.38E+03	6.09617	7.22E-09

sigma^2 estimated as 395584183: log likelihood=-1901.23, AIC=3810.46 AICc=3810.7 BIC=3822.95: This Model order is autogenerated from the auto. Arima () function in R based on the minimization of AIC and BIC

$$Y_t = 8432.586 + 0.309194 y_{t-12} + 0.205886\, y_{t-24}$$

where y_{t-12} and y_{t-24} represents the value of the time series 12 periods (months) and the value of the time series 24 months before the current time t.

The SARIMA (0,0,0) (2,0,0) [12] model with a mean for Diammonium Phosphate incorporates two significant seasonal autoregressive terms (SAR1 and SAR2) and a constant term, which indicates a stable average level in the time series. The significance of SAR1 (p-value = 9.74E-05) and SAR2 (p-value = 9.75E-03) suggests that the demand for Diammonium Phosphate is influenced by patterns that recur yearly, with lagged effects from 12 and 24 months prior playing a critical role. The constant term (8432.586, p-value = 7.22E-09) supports the steady demand level, while the model's residual variance (395,584,183) and log-likelihood (-1901.23) implies a relatively higher error variability compared to the previous models.

Testing model accuracy

To evaluate SARIMA model accuracy, residuals are analysed to ensure they resemble white noise, indicating the model has captured underlying patterns. This involves plotting residuals, checking autocorrelation with ACF and PACF, and assessing normality using statistical tests and Q-Q plots. Forecasting accuracy is evaluated through comparison with a hold-out sample or cross-validation, and by comparing forecast plots to actual values.

Diammonium phosphate (DAP) Demand Forecast

Figure 9 presents the forecasted demand for DAP fertiliser, with Table 4 displaying the 24-month forecast points. The forecasted values are all within the required bounds, indicating reliability in the predictions. The deep blue shaded area represents the 80% confidence level, while the light blue shaded area corresponds to the 95% confidence level. Notably, all forecasted points fall within the 95% confidence bounds, demonstrating that the predictions are within an acceptable range and, therefore, reliable.

Model building

Diammonium Phosphate; SARIMA (0,0,0) (2,0,0) [12] w/ mean

This model includes two seasonal autoregressive terms: SAR1, estimated at 0.309194, and SAR2, estimated at 0.205886. Both terms are statistically significant, with SAR1 having a p-value of 9.74E-05 and SAR2 having a p-value of 9.75E-03 (Table 8). The constant term is estimated at 8432.586 and is highly significant, with a p-value of 7.22E-09. The variance of the residuals is estimated at 395,584,183. The model's log-likelihood is -1901.23, with an Akaike Information Criterion (AIC) of 3810.46, a corrected AIC (AICc) of 3810.7, and a Bayesian Information Criterion (BIC) of 3822.95. These results show that the seasonal autoregressive and constant terms play a significant role in the model. In addition, the model includes a mean term, which implies that the model accounts for a constant average level in the time series.

TABLE 3: DIAMMONIUM PHOSPHATE SARIMA MODEL

Model	Term	Estimate	Std. error	t-statistic	p-value
ARIMA	sar1	0.309194	7.74E-02	3.992806	9.74E-05
ARIMA	sar2	0.205886	7.88E-02	2.614315	9.75E-03
ARIMA	constant	8432.586	1.38E+03	6.09617	7.22E-09

sigma^2 estimated as 395584183: log likelihood=-1901.23, AIC=3810.46 AICc=3810.7 BIC=3822.95: This Model order is autogenerated from the auto. Arima () function in R based on the minimization of AIC and BIC

$$Y_t = 8432.586 + 0.309194 y_{t-12} + 0.205886\, y_{t-24}$$

where y_{t-12} and y_{t-24} represents the value of the time series 12 periods (months) and the value of the time series 24 months before the current time t.

The SARIMA (0,0,0) (2,0,0) [12] model with a mean for Diammonium Phosphate incorporates two significant seasonal autoregressive terms (SAR1 and SAR2) and a constant term, which indicates a stable average level in the time series. The significance of SAR1 (p-value = 9.74E-05) and SAR2 (p-value = 9.75E-03) suggests that the demand for Diammonium Phosphate is influenced by patterns that recur yearly, with lagged effects from 12 and 24 months prior playing a critical role. The constant term (8432.586, p-value = 7.22E-09) supports the steady demand level, while the model's residual variance (395,584,183) and log-likelihood (-1901.23) implies a relatively higher error variability compared to the previous models.

Testing model accuracy

To evaluate SARIMA model accuracy, residuals are analysed to ensure they resemble white noise, indicating the model has captured underlying patterns. This involves plotting residuals, checking autocorrelation with ACF and PACF, and assessing normality using statistical tests and Q-Q plots. Forecasting accuracy is evaluated through comparison with a hold-out sample or cross-validation, and by comparing forecast plots to actual values.

Diammonium phosphate (DAP) Demand Forecast

Figure 9 presents the forecasted demand for DAP fertiliser, with Table 4 displaying the 24-month forecast points. The forecasted values are all within the required bounds, indicating reliability in the predictions. The deep blue shaded area represents the 80% confidence level, while the light blue shaded area corresponds to the 95% confidence level. Notably, all forecasted points fall within the 95% confidence bounds, demonstrating that the predictions are within an acceptable range and, therefore, reliable.

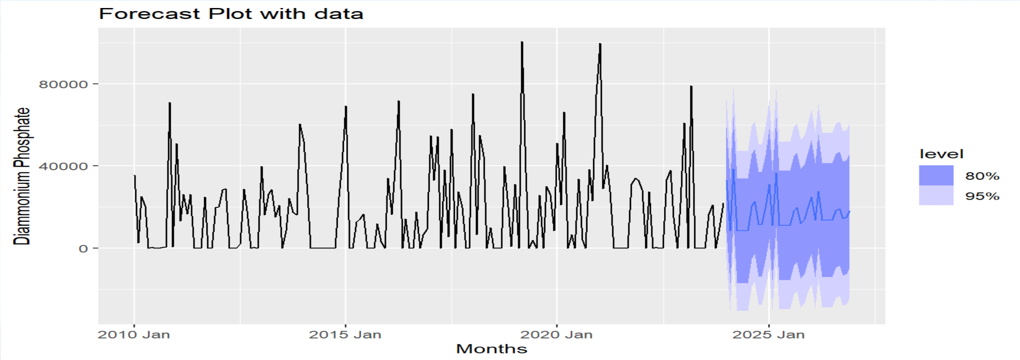

Figure 9: Forecasts plot for diammonium phosphate fertiliser

Table 4: Forecasts of diammonium phosphate fertiliser

Month	Qty (in metric tons)	Month	Qty (in metric tons)
2024 Aug	N (20405, 4e+08)	2025 Aug	N (18139, 4.3e+08)
2024 Sep	N (22718, 4e+08)	2025 Sep	N (19780, 4.3e+08)
2024 Oct	N (11500, 4e+08)	2025 Oct	N (11988, 4.3e+08)
2024 Nov	N (11518, 4e+08)	2025 Nov	N (14048, 4.3e+08)
2024 Dec	N (19764, 4e+08)	2025 Dec	N (19073, 4.3e+08)
2025 Jan	N (31203, 4.3e+08)	2026 Jan	N (24896, 4.7e+08)
2025 Feb	N (11040, 4.3e+08)	2026 Feb	N (13582, 4.7e+08)
2025 Mar	N (36638, 4.3e+08)	2026 Mar	N (27698, 4.7e+08)
2025 Apr	N (11040, 4.3e+08)	2026 Apr	N (13582, 4.7e+08)
2025 May	N (11068, 4.3e+08)	2026 May	N (13610, 4.7e+08)
2025 Jun	N (11040, 4.3e+08)	2026 Jun	N (13582, 4.7e+08)
2025 Jul	N (11040, 4.3e+08)	2026 Jul	N (13582, 4.7e+08)

The plots to test the normality of the forecasts indicated that the forecasts are normally distributed (as shown in Figure 10). A plot to check the deviation between the fitted and actual data also indicates slight deviations, confirming that the model was a good fit for forecasting. The Shapiro-Wilk normality test yielded W = 0.90764 and a p-value = 8.636e-09. W is close to 1, indicating that the residues closely approached normal distribution. The graph in Figure 13 shows standardized deviations of residuals from normality

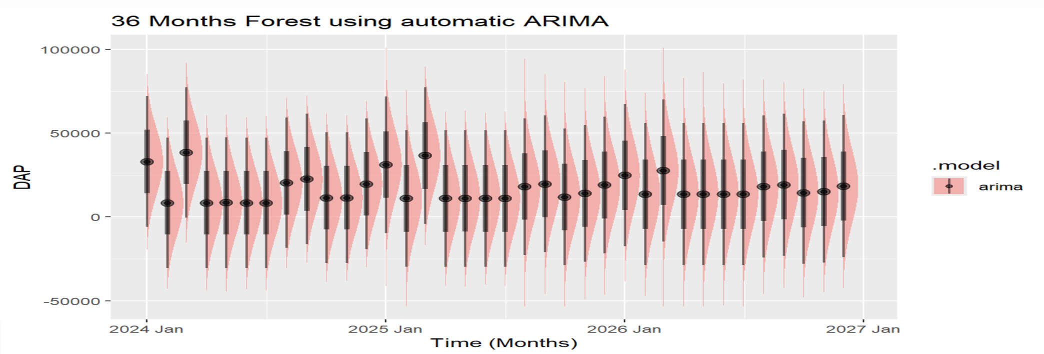

Figure 10: Distribution of the forecast for diammonium phosphate.

Figure 10 presents the actual and fitted values of the Diammonium Phosphate demand from January 2010 to Dec 2023. The red line indicates the actual data, while the blue line represents the fitted model. The graph indicates pronounced variability, with frequent and sharp spikes in demand. Although the fitted values follow the general pattern of the actual data, there are noticeable deviations, particularly during periods of high demand. Despite these discrepancies, the fitted model still captures the data's overall trend and seasonal fluctuations. The model's ability to generally track the actual demand, even with some deviations during peak periods, suggests that the predictions remain reliable. A plot to check the normality of the residues (in figure 12) indicated that the residues slightly follow normality, a confirmation of the model's accuracy.

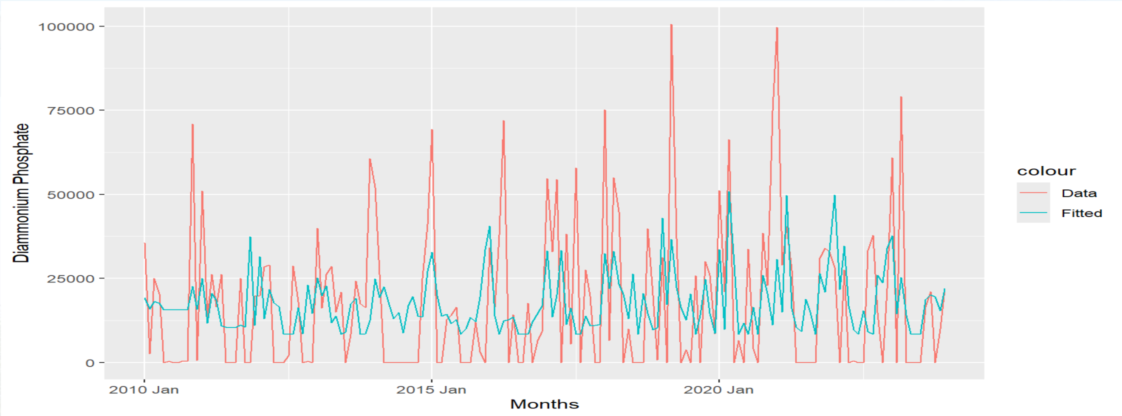

Figure 11: Fitted vs. actual time series for diammonium phosphate

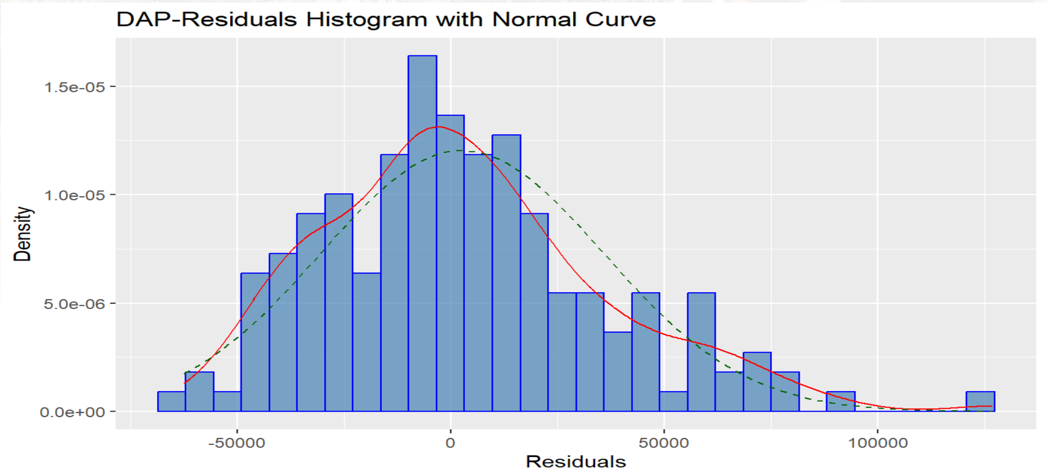

Figure 12: Residuals vs. normal for diammonium phosphate

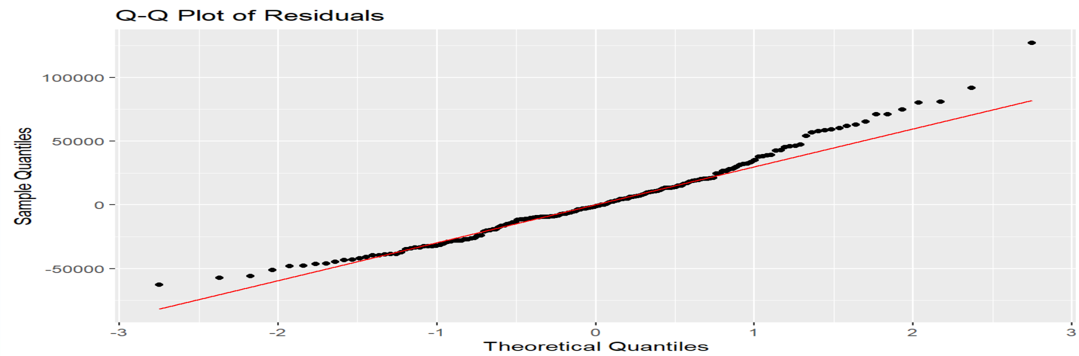

Figure 13: Standardised residuals plot for diammonium phosphate

All the test done to evaluate the accuracy of the model aligns with various test done. Shapiro-Wilk normality test applied to DAP fertiliser model residuals assessed the normality of residuals for the SARIMA model, following the approach of Harris (2001). Harris (2001) demonstrates the efficacy of the Shapiro-Wilk test in validating model residuals, ensuring that the assumptions of normality are met for reliable forecasts. Similarly, the application of this test confirmed that the residuals for DAP fertiliser from the fertilizer demand model adhered to normal distribution assumptions, thereby reinforcing the accuracy and validity of forecasting results. Q-Q residual plots utilized in this analysis visually assess the distribution of residuals from the model, echoing the methodology employed by Harris (2001). As indicated by Harris (2001) Q-Q plots effectively reveal deviations from normality by comparing the quantiles of residuals to a theoretical normal distribution. By applying this technique, verification that the residuals from DAP fertilizer demand forecasts approximately followed a normal distribution was ascertained, thereby supporting the robustness of our models and aligning with established practices in time series forecasting.

Conclusion and recommendation

This study addressed the urgent need to forecast DAP fertiliser demand in Kenya, a country where agriculture plays a pivotal role in the economy and sustenance of the population. Despite the increasing population, agricultural productivity has stagnated due to factors like suboptimal fertiliser application and distribution. This study's analysed trends and seasonal patterns in Kenya's DAP fertiliser demand and developed appropriate SARIMA models that can be used to forecast future demand for this fertiliser type. Using time series analysis, the study successfully identified patterns in fertiliser consumption and developed SARIMA model for DAP fertiliser using the Box Jankins approach. This model was then applied to forecast future fertiliser demand, providing a valuable tool for stakeholders in the agricultural sector to make informed decisions and improve planning processes in Kenya.

The study's key findings reveal that DAP fertiliser demand in Kenya exhibits distinct seasonal patterns, with higher demand during January, March and April due to the country’s agricultural practices. Notably, there is an increasing trend in DAP fertiliser demand, indicating that future needs will grow as the population and agricultural activity expand. The seasonal nature of fertiliser demand is particularly pronounced during the long and short rainy seasons when agricultural activities peak and the need for fertiliser increases. The finding of this study indicated that the DAP fertiliser demand can be represented by SARIMA (0,0,0) (2,0,0) [12]. This study recommended incorporation of SARIMAX and Hybrid Models for Comprehensive Forecasting. Therefore, future research should consider employing SARIMAX (Seasonal Autoregressive Integrated Moving Average with Exogenous Regressors) models alongside hybrid forecasting models. SARIMAX models can account for seasonal and non-seasonal components and external factors influencing fertiliser demand. Exploring hybrid models that combine SARIMAX with machine learning techniques, such as ensemble methods or neural networks, could offer enhanced predictive accuracy by leveraging the strengths of both statistical and data-driven approaches.

Acknowledgement

We sincerely thank everyone who contributed to the completion of this research. This work was made possible through the combined efforts of our colleagues and mentors who generously shared their knowledge and expertise.

REFERENCES

Bezerra, E, Ogasawara, E.S., Oliveira, D.D., Junior, F.P., Castañeda, R., Amorim, M., Mauro, R.C., Soares, J.D., Quadros, J.R., &. (2013). A Forecasting Method for Fertilisers Consumption in Brazil. *Int. J. Agric. Environ. Inf. Syst., 4*, 23-36.

Borkar, P. (2023). Statistical Modelling for Forecasting Fertiliser Consumption in India. https://orcid.org/0000-0002-5437-7001

Box, G. E. P., & Jenkins, G. M. (1976). *Time Series Analysis: Forecasting and Control*. San Francisco: Holden-Day.

Box, G. E. P., Jenkins, G. M., & Reinsel, G. C. (2015). Time Series Analysis: Forecasting and Control (5th ed.). Wiley.

Brockwell, P. J., & Davis, R. A. (2016). Introduction to Time Series and Forecasting (3rd ed.). Springer.

Filder, T.N., Muraya, M.M., & Mutwiri, R.M. (2019). Application of seasonal autoregressive moving average models to analysis and forecasting of time series monthly rainfall patterns in Embu County, Kenya, *Asian Journal of Probability and Statistics*, 4 (4): 1-15

Hannah Ritchie, Max Roser and Pablo Rosado (2022) - "Fertilisers". Published online at OurWorldInData.org. Retrieved from: https://ourworldindata.org/fertilisers

Harris, D. (2001). Statistical Modelling in S-Plus. Springer.

McQuarrie, A. D., & Tsai, C. L. (1998). *Regression and time series model selection*. World Scientific.

Mutegi, J., Adolwa, I., Kiwia, A., Njoroge, S., Gitonga, A., Muthamia, J., & Kansiime, M. (2024). Agricultural production and food security implications of Covid-19 disruption on small-scale farmer households: lessons from Kenya. World Development, 173, 106405.

Okello, E. A. (2023). Application of Hybrid seasonal ARIMA-GARCH Model in modelling and forecasting fertiliser prices in Kenya [Strathmore University]. http://hdl.handle.net/11071/15390

Pisuttinusart, C., Jatuporn, C., Suvanvihok, V., & Seerasarn, N. (2022). Forecasting The import Demand for chemical fertiliser in Thailand. *The EUrASEANs: journal on global socio-economic dynamics*, (3 (34)), 61-70.

Sheahan, M., Ariga, J., & Jayne, T. S. (2016). Modelling the effects of input market reforms on fertiliser demand and maize production: A case study from Kenya. *Journal of Agricultural Economics*, *67*(2), 420-447.

Shumway, R. H., & Stoffer, D. S. (2017). Time Series Analysis and Its Applications: With R Examples (4th Edition). Springer.

Shumway, R. H., Stoffer, D. S., & Stoffer, D. S. (2000). *Time series analysis and its applications* (Vol. 3). New York: springer.

Shumway, R. H., Stoffer, D. S., Shumway, R. H., & Stoffer, D. S. (2017). ARIMA models. Time series analysis and its applications: with R examples, 75-163.

Shumway, R.H., & Stoffer, D.S. (2017). Time Series Analysis and Its Applications: With R Examples. Springer.

Tenkorang, Frank & Lowenberg-DeBoer, James. (2008). Forecasting Long-Term Global Fertiliser Demand. Nutrient Cycling in Agroecosystems. 83. 233-247. 10.1007/s10705-008-9214-y.

Wanjuki, T. M., Wagala, A. and Muriithi, D. K. (2021). Sarima models: review and its application to Kenyan's commodity price index of food and beverage. In*: Isutsa, D. K. (Ed.). Proceedings of the 7th International Research Conference held in Chuka University from 3rd to 4th December 2020, Chuka, Kenya, p. 574-586*

Yaseen, M., Ahmad, A., Younas, N., Naveed, M., Ali, M. A., Shah, S. S. H., ... & Mustafa, A. (2023). Value-Added Fertilisers Enhanced Growth, Yield and Nutrient Use Efficiency through Reduced Ammonia Volatilization Losses under Maize–Rice Cropping Cultivation. Sustainability, 15(3), 2021.