International Journal of Business and Economics
Vol. 10, No. 1, 2025, pp. 104-124

https://ielas.org/ijbe/index.php/ijbe
ISSN (online) 2545-4137

GLOBAL MONTHLY COTTON PRICE TRENDS: INSIGHTS FROM ARIMA TIME-SERIES FORECASTING

Rumita Limbu Sanwa[1]* and Yeong Nain Chi[1]
[1]Department of Agriculture, Food, and Resource Sciences, University of Maryland Eastern Shore, Princess Anne, MD 21853, USA

*Corresponding author: rlimbusanwa@umes.edu

Received: 2025-05-12 Accepted: 2025-10-02 Published online: 2025-10-11

Abstract

Cotton is a crucial cash crop in the global economy, yet its prices often change due to market trends and seasonal variations. Accurate price forecasting can help farmers, traders, and policymakers make better, timely decisions and promote sustainable production. This study uses a Seasonal Autoregressive Integrated Moving Average (SARIMA) model to analyze and forecast global monthly cotton prices using data from January 1990 to January 2025 obtained from the FRED database. Model selection was performed with the auto.arima function from the forecast package in R, guided by the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC). Following the Box–Jenkins methodology, the final model, ARIMA (4,1,1) (0,0,2) [12], satisfied key assumptions of stationarity, normality, and homoscedasticity, confirmed through the Augmented Dickey–Fuller test, Ljung–Box test, and residual ACF/PACF diagnostics. Forecast accuracy was evaluated using Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), Mean Percentage Error (MPE), Mean Absolute Percentage Error (MAPE), and Mean Error (ME), showing modest out-of-sample errors consistent with prior studies on commodity price forecasting. A 30-month forecast indicates a steady increase in cotton prices, from 76.60 U.S. cents per pound in February 2025 to 91.66 cents by January 2027, with minor monthly fluctuations. These findings highlight SARIMA as a practical tool for anticipating cotton price trends and supporting informed decision-making. However, forecasts should be interpreted as guidance rather than exact predictors of future prices. Future studies may benefit from integrating SARIMA with hybrid models or machine learning approaches to better capture complex, non-linear relationships in commodity price forecasting.

Keywords: Cotton, Box–Jenkins methodology, price forecasting, ARIMA.

1. Introduction

Cotton is a major cash crop and serves as a vital raw material for the global textile industry. It contributes significantly to the world economy, with an estimated annual economic impact of over $600 billion (Khan et al., 2020). However, cotton prices are subject to high volatility due to various factors, including global economic fluctuations. This price instability poses challenges for stakeholders across the supply chain, particularly farmers and policymakers. Reliable cotton price forecasts are essential for informed decision-making among producers and market participants (Darekar & Reddy, 2017).

Accurate forecasting enables farmers to plan agricultural activities more effectively, such as choosing optimal selling periods and identifying markets that offer higher prices. Since cotton prices often vary by season and geographic location, timely market information can enhance farm-level income stability and risk management (Kumar et al., 2024). Moreover, promoting sustainability in cotton production is critical, as the sector plays a significant role in employment generation, economic development, and the growth of related industries (Ashraf et al., 2024). From a macroeconomic perspective, price stability is a fundamental objective of monetary policy, ensuring fair outcomes for both producers and consumers. In this context, the development and implementation of robust forecasting models are necessary to support economic planning and contribute to the stabilization of commodity markets (Wanjuki et al., 2021).

Predicting prices in agricultural markets, especially for global commodities like cotton, is important because prices are affected by changing weather, international trade, and government policies (Weng et al., 2019). To understand and forecast these prices, researchers use different types of time series models to predict future values and identify trends in data, such as non-seasonal, univariate seasonal, and multiple seasonal models. One popular method is the ARIMA model, which is often used to predict trends in univariate time series (Shumway & Stoffer, 2025). Numerous studies have employed ARIMA models to forecast agricultural commodity prices. Iqbal et al. (2005) modeled wheat prices, and Darekar and Reddy (2017) utilized ARIMA models to forecast cotton prices, demonstrating the model's strength in capturing linear trends and its reliance on past data for accurate predictions. However, traditional ARIMA models may underperform when strong seasonal components are present in the data. To address this, the Seasonal ARIMA (SARIMA) model extends ARIMA by explicitly modeling seasonal patterns.

Several studies have demonstrated the effectiveness of SARIMA models in forecasting agricultural prices. For example, Divisekara et al. (2021) used a SARIMA (2,1,2) (0,1,1) [52] model to forecast red lentil prices in Canada using weekly data and found it effective for generating accurate short-term forecasts that support better decision-making and price risk management. (Adanacioglu and Yercan, 2012) applied a SARIMA (1,0,0) (1,1,1) [12] model to predict tomato prices in Turkey, successfully capturing seasonal trends to aid in market planning. Similarly, Khadka and Chi (2024) employed a SARIMA (0,1,1) (0,0,1) [12] model to forecast global corn prices; likewise, Ojha and Karki (2025) used the same model structure for forecasting global wheat prices, further confirming the suitability of SARIMA models for analyzing seasonality and trends in agricultural commodities.

More recent work has reinforced the practical value of ARIMA and SARIMA models in cotton price forecasting. Kumar et al. (2024) studied major cotton markets in Haryana, India, and found that an ARIMA (1,1,1) model yielded highly accurate forecasts, supporting its applicability at the regional market level. Bhopala et al. (2023) compared ARIMA and SARIMA models across six major Indian cotton markets and concluded that model performance varied by region, with SARIMA outperforming in some markets while ARIMA excelled in others. Their findings highlight the importance of selecting appropriate models based on the unique seasonal patterns and statistical characteristics of each dataset, rather than relying on a one-size-fits-all approach. Biswal and Sahoo (2020) used weekly price data to evaluate forecasting methods for agricultural products and found that a SARIMA (1,1,1) (1,0,1) [4] model provided the best fit for accurate forecasting. Their analysis aims to support farmers in making informed pricing decisions and adjusting crop schedules in response to expected market trends. Jadhav et al. (2017) also emphasized the reliability of the ARIMA model for agricultural price prediction, illustrating its potential to assist in strategic planning within the farming sector.

These studies collectively support the view that time series models, particularly ARIMA and SARIMA, are effective tools for forecasting agricultural commodity prices. They not only enhance market predictability but also provide essential insights that can help stakeholders navigate price volatility and improve economic outcomes.

2. Materials and Methods

The forecasting of global monthly cotton prices was conducted using the ARIMA (Autoregressive Integrated Moving Average), following the structured Box-Jenkins methodology (Box & Jenkins, 1970; Montgomery et al., 2015). This process includes three main stages: model identification, parameter estimation, and diagnostic checking. To begin, the time series was assessed for stationarity using the Augmented Dickey-Fuller (ADF) test (Dickey & Fuller, 1981). Since the original series was non-stationary, both first-order and seasonal differencing were applied to stabilize the mean and address seasonal patterns. Visual diagnostics, such as ACF and PACF plots, alongside seasonal decomposition, helped determine the appropriate model structure. The best-fitting ARIMA model was selected using the auto.arima () function from the *forecast* package in R, with model selection based on AIC, BIC, and residual diagnostics. The selected model, ARIMA (4,1,1) (0,0,2) [12], met key diagnostic criteria, including the Ljung-Box test for residual independence, confirming its adequacy. Parameters were estimated using Maximum Likelihood Estimation (MLE), and model performance was evaluated using forecast accuracy metrics such as Mean Error (ME), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and Mean Absolute Scaled Error (MASE) (Hyndman & Koehler, 2006). Forecasts were then produced for a 30-month period, including 80% and 95% confidence intervals to capture uncertainty. Reproducibility was ensured by setting a fixed random seed during the forecasting process. This approach provided a reliable way to project future cotton price trends using patterns observed in historical data.

2.1 ARIMA Model

The Autoregressive Integrated Moving Average (ARIMA) model, originally proposed by Box and Jenkins (1970), is a foundational statistical method for time series forecasting, widely used and especially suited for data that exhibits trends and autocorrelation. The ARIMA model is specified by three parameters: p, d , and q, where:

- p represents the order of the autoregressive (AR) component,
- d denotes the number of differencing required to make the series stationary,
- q represents the order of the moving average (MA) component.

The general ARIMA model can be mathematically represented using the backshift operator B as:

$$\phi_p(B)(1 - B)^d y_t = \theta_q(B)\epsilon_t \quad (1)$$

Where,

y_t is the observed time series value at time *t,*

$\phi_p(B)$ is the autoregressive polynomial,

$\theta_q(B)$ is the moving average polynomial,

$(1 - B)^d$ represents the differencing operator to induce stationarity,

ϵ_t is a white noise error term.

2.1.1 Autoregressive (AR) Model

The Autoregressive (AR) model is one of the foundational time series models in which the current value of a variable is expressed as a linear function of its past values and a stochastic error term (Box & Jenkins, 1970). It is particularly useful for data with autocorrelation, where past values influence future outcomes. By assuming that historical patterns can inform future outcomes, the AR model is well-suited for stationary time series (Shumway & Stoffer, 2025). The AR (p) model predicts the current value of a time series using a linear combination of its past values p values and a random error. Formally introduced by Yule (1927), the AR model gained broader popularity with the development of the ARMA framework by Box and Jenkins (1970). The order of the model, represented as AR(p), refers to the number of umber of lagged terms used and is typically determined using information criteria like the Akaike Information Criterion (AIC) (Akaike, 1974) or the Bayesian Information Criterion (BIC) (Schwarz, 1978).

International Journal of Business and Economics
Vol. 10, No. 1, 2025, pp. 104-124
https://ielas.org/ijbe/index.php/ijbe
ISSN (online) 2545-4137

Graphical tools such as the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) play a crucial role in identifying the appropriate model structure in the time series analysis. For a pure autoregressive process of order p (AR(p)), the PACF sharply declines after lag pp, signifying direct dependence only up to that point. In contrast, the ACF exhibits a gradual decay, capturing the indirect effects of previous observations (Shumway & Stoffer, 2017).

The AR model of order p, denoted as AR (p), is mathematically represented as:

$$y_t = \sum_{i=1}^{p} \phi_i \; y_{t-i} + \epsilon_t \qquad (2)$$

Where,

y_t is the value of the time series at time t,

p: the order of the autoregressive process (i.e., how many past values the model depends on),

ϕ_i: the coefficients that determine the influence of each past value y_{t-i},

ϵ_t: a white noise error term, typically assumed to be independently and identically distributed with mean 0 and constant variance

2.1.2 Moving Average (MA) model

It is a basic time series tool that explains the current value based on past errors or shocks, rather than past values. It works well for short-term patterns in data that are stable over time. This makes it useful for capturing quick, temporary effects in a time series (Box et al., 2015).

The MA model of order q, denoted as MA(q), is mathematically represented as:

$$y_t = \mu + \sum_{i=1}^{q} \theta_i \epsilon_{t-i} + \epsilon_t \qquad (3)$$

Where,

y_t is the value of the time series at time t,

μ is the mean of the series (maybe zero in differenced form),

θ_i is the MA coefficient,

ϵ_t represents a white noise error term,

q isthe order of the MA process

The MA model is commonly checked using the Autocorrelation Function (ACF). If the ACF shows a sudden drop to zero after a certain lag (q), it suggests that an MA(q) model might be a good fit for the data.

Vol. 10, No. 1, 2025, pp. 104-124
https://ielas.org/ijbe/index.php/ijbe
ISSN (online) 2545-4137

2.2 The Seasonal Autoregressive Integrated Moving Average (SARIMA) model

It builds on the ARIMA framework by incorporating both seasonal and non-seasonal patterns, making it better suited for capturing complex time series behavior (Box et al., 2015). Using the backshift operator B, the general form of the SARIMA model can be expressed as:

$$\Phi_P(B^s)\phi_p(B)(1-B)^d(1-B^s)^D\,y_t = C + \Theta_Q(B^s)\theta_q(B)\epsilon_t \quad (4)$$

Where,

$\phi_p(B)$: non-seasonal autoregressive polynomial,
$\theta_q(B)$: non-seasonal moving average polynomial,
$\Phi_P(B^s)$: seasonal autoregressive polynomial,
$\Theta_Q(B^s)$: seasonal moving average polynomial,
$(1-B)^d$: non-seasonal differencing,
$(1-B^s)^D$:seasonal differencing,
C : constant (intercept)
ϵ_t: white noise error term

3. Results

To forecast the global monthly price of cotton, a Seasonal ARIMA (SARIMA) model was applied. The analysis was conducted using R version 4.4.3 on Windows, with data from January 1990 to January 2025 obtained from the FRED database. This data was then converted into a time series format for modeling and forecasting.

3.1 Identification of the Seasonal ARIMA Model

Table 1: *Summary Statistics of Global Monthly Cotton Prices (1990–2025)*

Statistic	Value (U.S. Cents per Pound)
Min.	37.22
1st Qu.	60.44
Median	75.91
Mean	77.86
3rd Qu.	87.04
Max.	229.67
Variance	616.2144
Standard Deviation	24.82367

The summary statistics (Table 1) of the cotton price series show considerable variability over the study period. Prices ranged from a minimum of 37.22 to a maximum of 229.67 U.S. cents per pound. The mean price was approximately 77.86, with a median of 75.91, indicating a slightly right-skewed distribution. The standard

International Journal of Business and Economics
Vol. 10, No. 1, 2025, pp. 104-124

https://ielas.org/ijbe/index.php/ijbe
ISSN (online) 2545-4137

deviation of 24.82 and variance of 616.21 reflect moderate to high price volatility across the dataset.

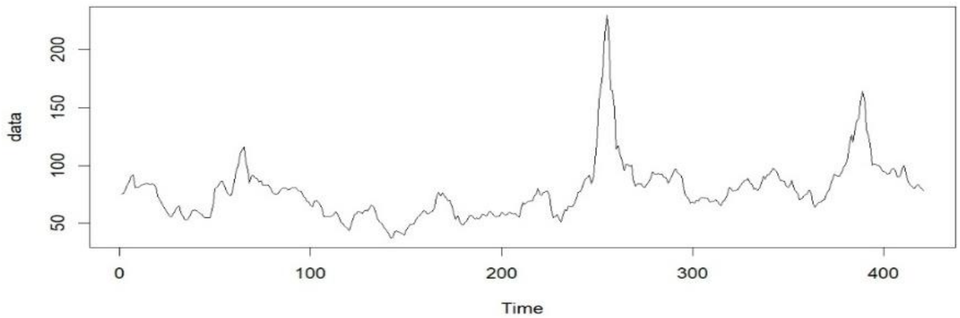

Figure 1: Time series plot of the monthly global price of Cotton (January 1990 ~ January 2025).

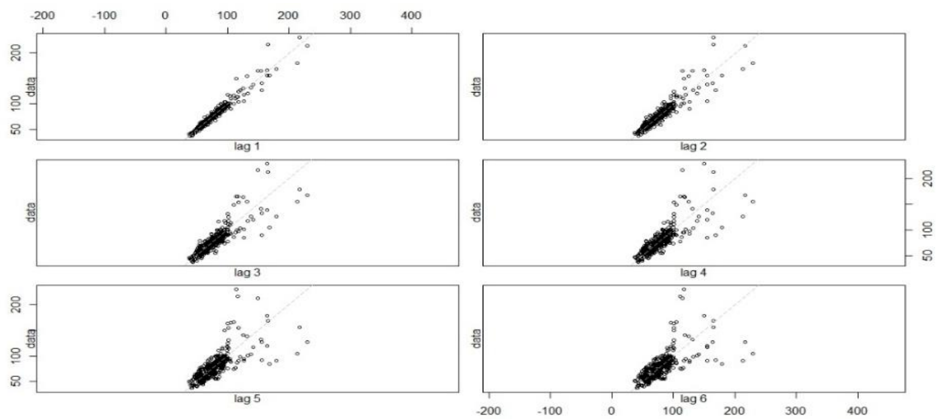

Figure 2: Lagged plots of the monthly global price of cotton from January 1990 to January 2025.

The Augmented Dickey-Fuller (ADF) test was conducted on the cotton price data to check if it is stationary. The test gave a result of -3.8534 with a p-value of 0.01649. Since the p-value is below 0.05, we can confidently say the data is stationary, meaning it doesn’t have a trend or a random drift. It is suitable for modeling with ARIMA or SARIMA, which assume stationarity in the data.

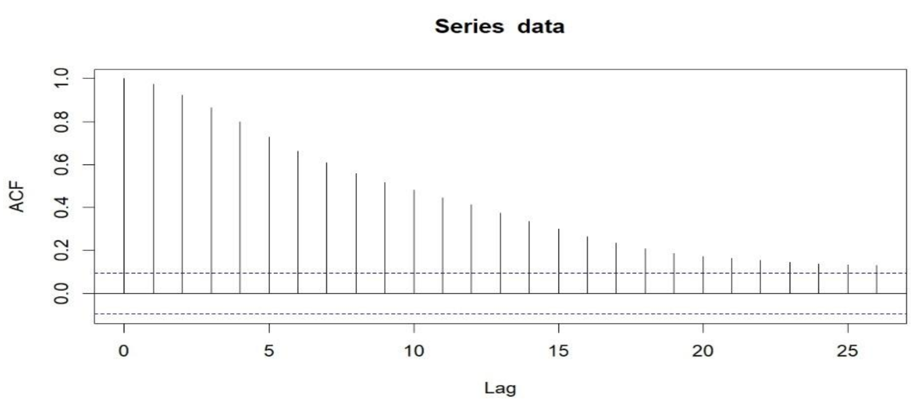

Figure 3: ACF plot of the global price of cotton before differencing.

The autocorrelation function (ACF) is a helpful tool for assessing stationarity in a time series. In stationary data, the ACF values drop quickly to zero. In contrast, a slow decline in the ACF suggests non-stationarity, where patterns persist over longer periods. This visual pattern helps us understand the time series behavior and guides appropriate model selection (Chi, 2022).

Figure 3 displays the autocorrelation function (ACF) plot for the time series data, showing correlations between the series and its lagged values. The ACF exhibits slow and gradual decline, characteristic of a non-stationary time series. High autocorrelation values at initial lags, which decrease slowly, suggest a strong pattern or trend in the data. The dashed blue lines represent the 95% confidence intervals; since many bars exceed these lines at early lags, the autocorrelations are statistically significant. This pattern implies that the series requires differencing to achieve stationarity before fitting a suitable ARIMA model. Identifying and transforming non-stationary time series is important for reliable forecasting, especially when working with economic or agricultural datasets such as cotton prices.

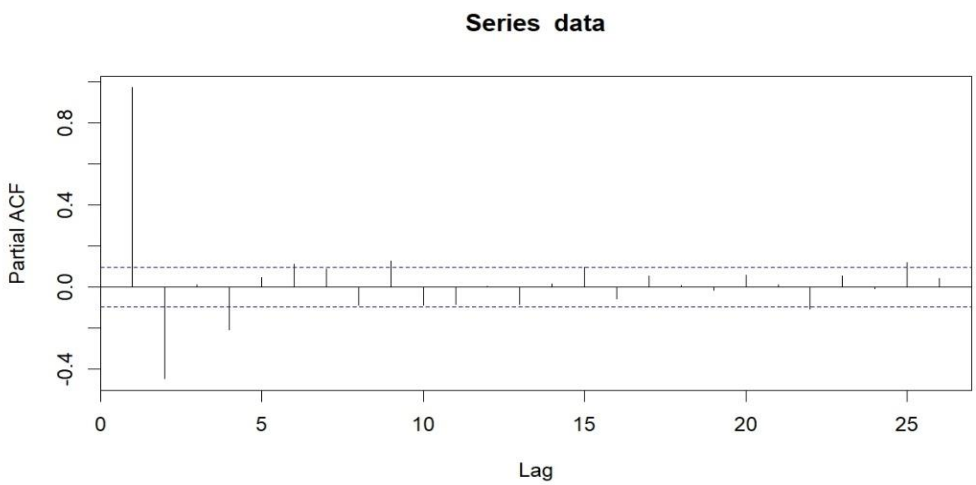

Figure 4: PACF plot of the global price of cotton before differencing.

Figure 4 presents the Partial Autocorrelation Function (PACF) of the time series. It shows the direct relationship between a value and its previous values, excluding the influence of the values in between. The significant spike at lag 1 means that the current value is mainly influenced by the previous value. After that, the bars stay close to zero within the shaded area, meaning there’s little to no additional influence from earlier points. This suggests that most of the correlation in the series can be explained by just the immediate past value, and further lags do not add much predictive power. In practical terms, this pattern supports using an ARIMA model with an autoregressive AR (1).

International Journal of Business and Economics
Vol. 10, No. 1, 2025, pp. 104-124
https://ielas.org/ijbe/index.php/ijbe
ISSN (online) 2545-4137

Figure 5 displays the cotton price time series decomposition using an additive model. The top panel shows the original monthly data from 1990 to 2025, highlighting fluctuations in observed prices. The second panel shows the overall trend, which rises over time with some noticeable peaks, especially around 2011 and 2022. The third panel highlights a repeating seasonal pattern that comes up every year, confirming the presence of annual seasonality. The bottom plot shows random changes in the data that don't follow a clear pattern. This decomposition allows for a clearer understanding of the long-term price trends, seasonal cycles, and random variations in the data.

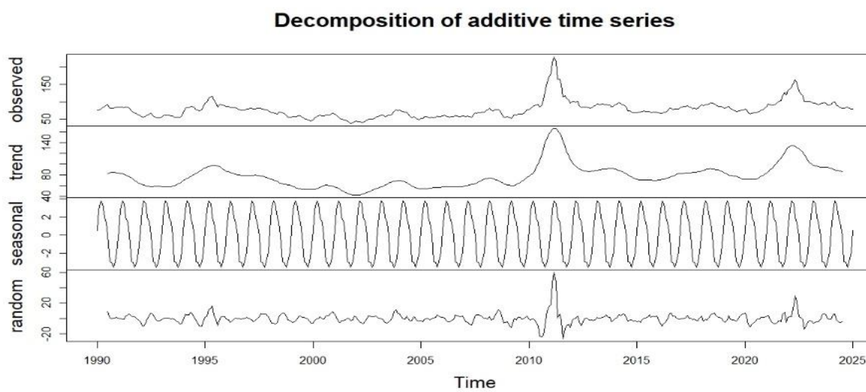

Figure 5: Decomposition of the monthly global price of cotton.

The purpose of seasonal adjustment is to identify and remove regular seasonal patterns from a time series to better understand the real trends and unexpected changes. This is done by subtracting the estimated seasonal part from the original data, so the adjusted data gives a clearer view of how things are changing over time without the usual ups and downs that happen every year. Once the seasonal variation is removed, the seasonally adjusted time series highlights longer-term trends and unexpected changes more clearly (Chi, 2022).

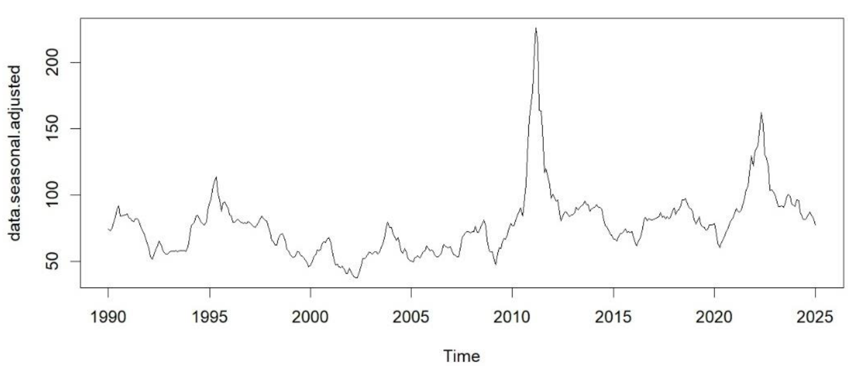

Figure 6: Time series plot of seasonally adjusted monthly global price of Cotton.

International Journal of Business and Economics
Vol. 10, No. 1, 2025, pp. 104-124
https://ielas.org/ijbe/index.php/ijbe
ISSN (online) 2545-4137

The Augmented Dickey-Fuller (ADF) test was applied to the seasonally adjusted cotton price series to evaluate its stationarity. The test yielded a Dickey-Fuller statistic of -4.0393 with a lag order of 7, and the p-value of the test was 0.01. This result led to the rejection of the null hypothesis of non-stationarity. It indicated that the seasonally adjusted time series is stationary, making it appropriate for further modeling and forecasting.

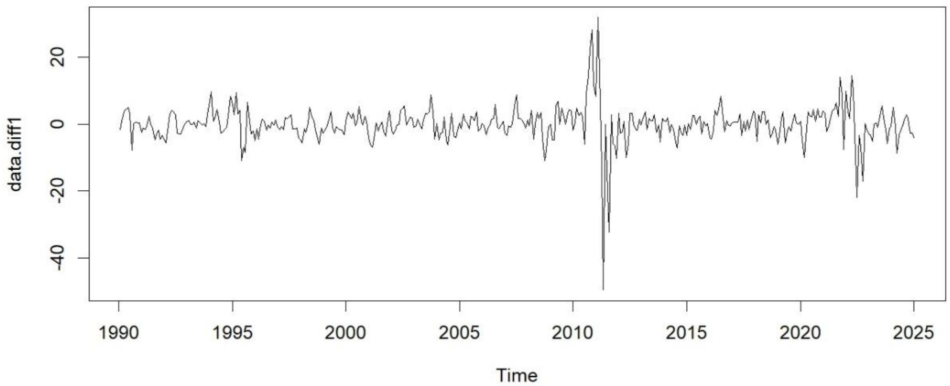

Figure 7: *Time* series plot of the first difference of the seasonally adjusted monthly global price of Cotton.

Figure 8 displays an autocorrelation function (ACF) plot of the first-order differenced series (data. diff1), which shows a significant spike at lag 1, with subsequent lags falling within the 95% confidence bounds. This indicates that differencing has effectively reduced autocorrelation, suggesting the series is now approximately stationary and suitable for further time series modeling.

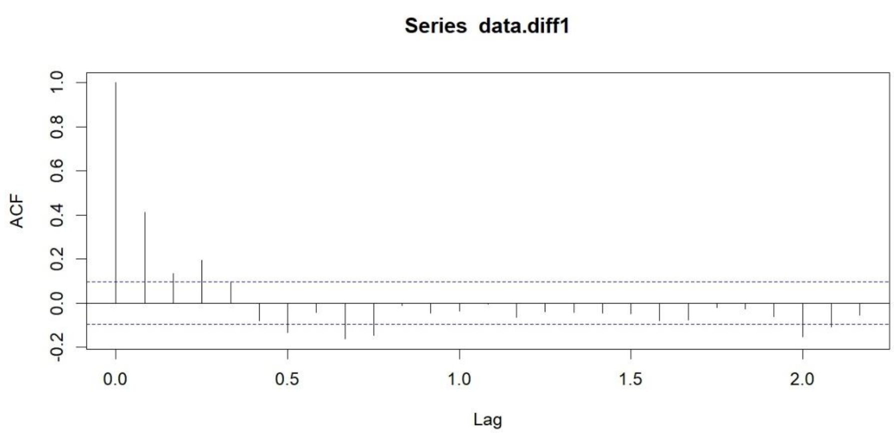

Figure 8: ACF plot of the global price of cotton after differencing (d=1).

This PACF plot of the first-order differenced series (data. diff1) shows a strongspike at lag 1, while the rest stay within the confidence limits. This suggests that only the

International Journal of Business and Economics
Vol. 10, No. 1, 2025, pp. 104-124

https://ielas.org/ijbe/index.php/ijbe
ISSN (online) 2545-4137

first lag has a meaningful partial correlation, pointing to an AR (1) structure. It also confirms that the series is likely stationary and ready for time series modeling.

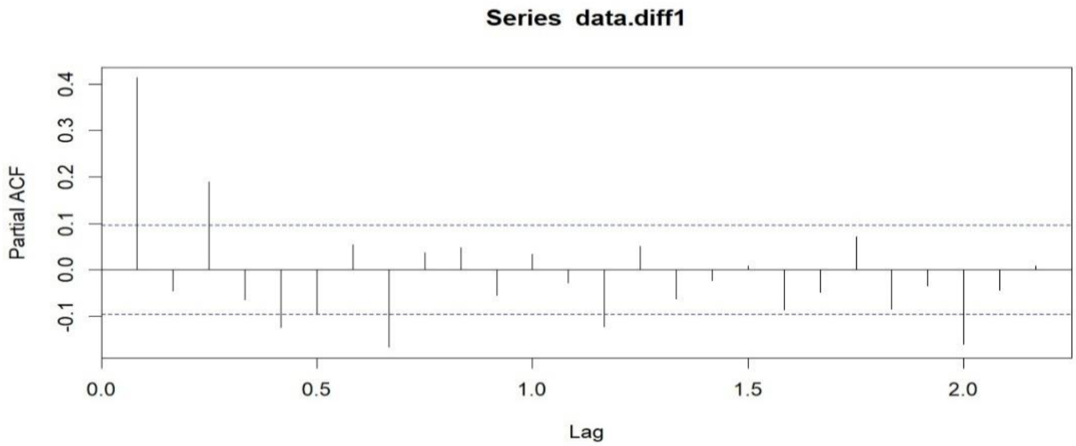

Figure 9: PACF plot of the global price of cotton after differencing (d=1)

Seasonal differencing is a technique where each data point is compared to the value from one full season earlier. This helps remove repeating seasonal patterns and makes the data more stable for modeling (Erasmus Kabu Aduteye et al., 2023). The seasonal differencing of the seasonally adjusted cotton price data was performed using a lag of 12 months to address any remaining annual seasonality. Dickey-Fuller statistics were -36.203 with a lag order of 7 and a p-value of 0.01. This support rejecting the null hypothesis of non-stationarity, confirming that the differenced series is stationary in Figure 10.

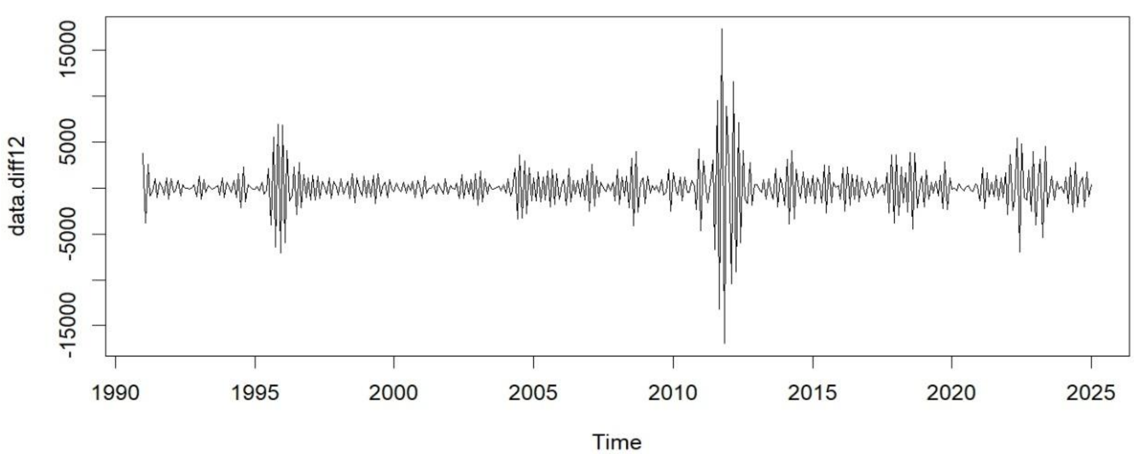

Figure 10: Time Series Plot of the 12th Difference of the Seasonal Adjusted Monthly Global Price of Cotton.

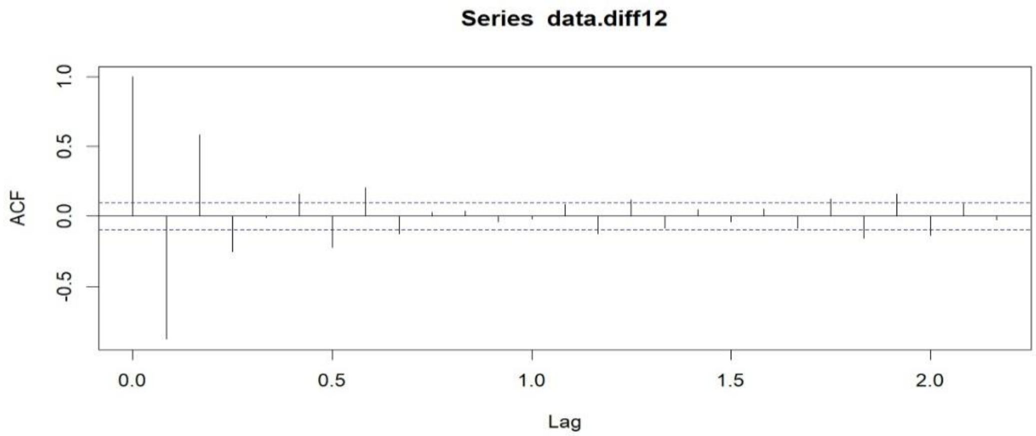

Figure 11: ACF plot of the global price of cotton after seasonal differencing.

The ACF plot of the seasonally differenced series shows a strong positive spike at lag 1, followed by values that stay within the confidence bounds and fluctuate around zero. This suggests short-term autocorrelation remains. This indicates that adding an autoregressive (AR) term may help improve the model after seasonal differencing. This PACF plot of the seasonally differenced series shows several significant negative spikes at early lags, especially at lag 1, followed by values near zero. This pattern suggests a strong short-term autoregressive (AR) effect, supporting the inclusion of an AR term in the model.

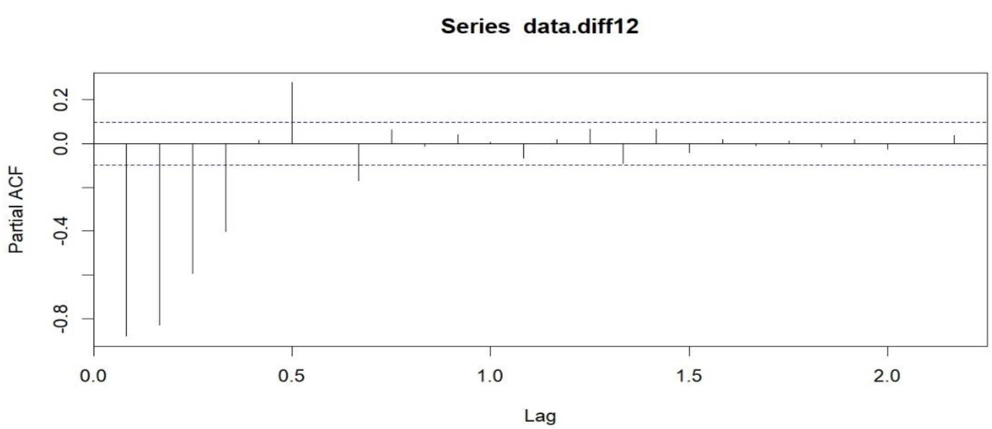

Figure 12: PACF plot of the global price of cotton after seasonal differencing.

3.2 Estimation of the Seasonal ARIMA Model Parameter

The auto.arima () function from the forecast package in R was used to determine the best-fitting seasonal ARIMA model. This function automatically determines AIC, AICc, or BIC, selecting the best-fitting model. The ARIMA model is defined by three components: AR (how much the model uses past values), I (how many times the

International Journal of Business and Economics
Vol. 10, No. 1, 2025, pp. 104-124
https://ielas.org/ijbe/index.php/ijbe
ISSN (online) 2545-4137

data is differenced to make it stable), and MA (how it accounts for past errors). The selected ARIMA (4,1,1) (0,0,2) [12] model indicates that the time series required one differencing step to achieve stability, meaning the original data showed a trend that needed to be removed. The non-seasonal AR (4) and MA (1) components suggest the current value is influenced by combining the last four observations and the most recent forecast error. Additionally, the seasonal MA (2) component, with a seasonal period of 12, captures the repeating yearly patterns in the data. Together, these elements reflect that the series is shaped by both short-term fluctuations and systematic seasonal effects, making this model well-suited for capturing the complex dynamics of the data over time.

Table 2: *Parameters of the ARIMA (4,1,1) (0,0,2)12 Model*

Parameters	Coefficients	Standard Error
Differencing	1	
ar1	1.3818	0.0520
ar2	-0.5391	0.0812
ar3	0.3397	0.0812
ar4	-0.2402	-0.2402
ma1	-0.9792	-0.9792
sma1	0.0271	0.0271
sma2	-0.1358	0.0519
$Sigma^2$ = 23.68: log likelihood = -1257.82		
AIC=2531.64 AICc=2531.99 BIC=2563.96		

ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
0.116421	4.819376	2.9887	-0.08646286	3.626474	0.1688978	0.003726883

Note: AR= Auto-Regressive, MA = Moving Average, SMA = Seasonal Moving Average, S.E. = Standard Error, AIC = Akaike Information Criterion, AICc = Corrected Akaike Information Criterion, BIC = Bayesian Information Criterion, ME = Mean Error, RMSE = Root Mean Squared Error, MAE = Mean Absolute Error, MPE = Mean Percentage Error, MAPE = Mean Absolute Percentage Error, MASE = Mean Absolute Scaled Error, and ACF1 = Autocorrelation of residuals at lag 1.

An ARIMA model with first-order differencing and seasonal components was fitted to the cotton price series, incorporating AR (1~4), MA (1), and seasonal MA terms (SMA1 and SMA2). Key coefficients such as AR (1) = 1.3818 and MA (1) = -0.9792 were statistically significant, while SMA1 showed limited influence. The model showed a good fit, with a log-likelihood of -1257.82, AIC = 2531.64, AICc = 2531.99, and BIC = 2563.96, indicating a strong balance between accuracy and complexity. Performance metrics such as RMSE (4.82), MAE (2.99), and MAPE (3.63%) demonstrate high predictive accuracy and minimal bias (ME = 0.12, MPE = -0.09%), with low autocorrelation in residuals (ACF1 = 0.0037), affirming the model's reliability for forecasting purposes.

Vol. 10, No. 1, 2025, pp. 104-124

https://ielas.org/ijbe/index.php/ijbe
ISSN (online) 2545-4137

The selected ARIMA (4,1,1) (0,0,2) [12] model will generate the following forecasting equation:

$$\phi(B)(1 - B)y_t = C + \Theta B^{\{12\}}\theta(B)\varepsilon_t \tag{5}$$

With the result presented above, the ARIMA (4,1,1) (0,0,2) [12] model can be written as:

$$(1 - 1.3818B + 0.5391B^2 - 0.3397B^3 + 0.2402B^4)(1 - B)y_t = \left(1 - 0.0271B^{\{12\}} + 0.1358B^{\{24\}}\right)(1 + 0.9792B)\varepsilon_t \tag{6}$$

1. Non-seasonal Autoregressive part (AR): $\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \phi_3 B^3 - \phi_4 B^4$
2. Non-seasonal Moving average part (MA): $\theta(B) = 1 + \theta_1 B$
3. Seasonal MA (Q = 2): $\Theta\left(B^{\{12\}}\right) = 1 + \Theta_1 B^{\{12\}} + \Theta_2 B^{\{24\}}$

3.3 Diagnostic Checking of the Seasonal ARIMA Model

The suitability of the model was also evaluated using the Ljung–Box Q-test. It is a statistical test used to check whether the residuals of a time series model are independently distributed (i.e., no autocorrelation). It helps validate the adequacy of a fitted time series model (like ARIMA) by testing the null hypothesis that all autocorrelations up to a certain lag are zero (Ljung & Box, 1978).

The diagnostic plots for the ARIMA (4,1,1) (0,0,2) [12] model show that the residuals are centered around zero with no clear pattern over time, and the autocorrelation function (ACF) plot indicates that most residual autocorrelations lie within the confidence bounds, suggesting no significant autocorrelation meaning zero mean and constant variation. The histogram shows an approximately normal distribution of residuals. These diagnostics support the idea that the model provides a good fit and that the residuals behave like white noise.

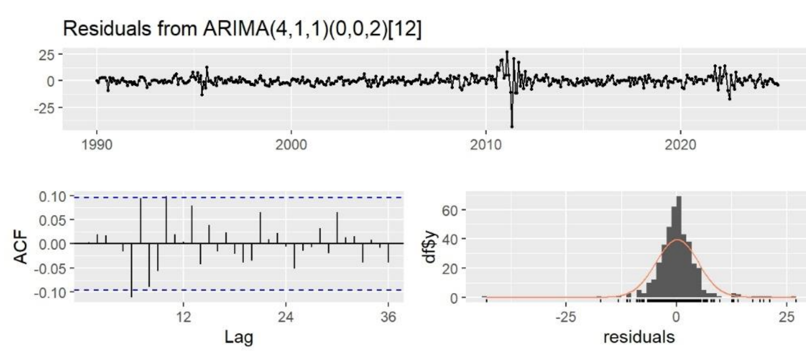

Figure 13: Ljung-Box test of the seasonally adjusted monthly global price of Cotton.

In this study, the Ljung–Box Q-test gave a statistic of 26.682 with 17 degrees of freedom and a p-value of 0.0629 (model df: 7, total lags used: 24). Since the p-value is above 0.05, it suggests that the residuals are likely random, meaning the model fits the data well. Based on this result, the ARIMA (4,1,1) (0,0,2) [12] model appears to capture the time series patterns effectively.

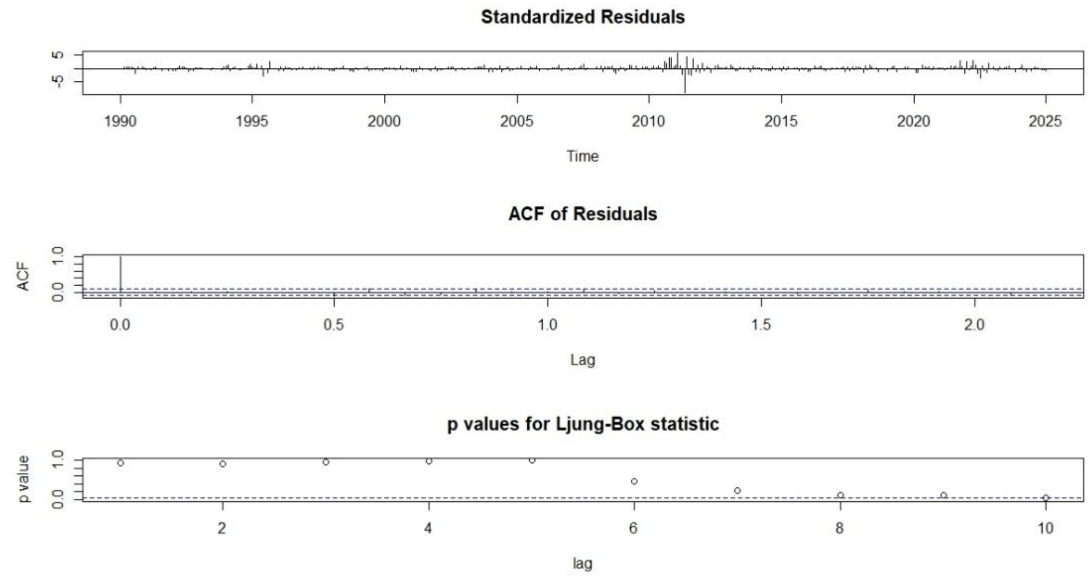

Figure 14: Residual plots of the seasonally adjusted monthly global price of Cotton.

Figure 14 shows that the model's residuals are stable, with no significant autocorrelation. The Ljung–Box test p-values are above 0.05, indicating the residuals behave like white noise and the model fits the data well.

3.4. Using the Seasonal ARIMA Model in Forecasting Process

The forecasted values of monthly global cotton prices are shown in the figure using the ARIMA (4,1,1) (0,0,2) [12] model. The black line represents historical data, while the blue line shows the forecasted values extending into the future. The shaded areas around the forecast line indicate confidence intervals, with darker shading for higher confidence. The model suggests a slight upward trend in the short term, and the forecast appears stable, indicating a reasonably good model fit for short-term prediction.

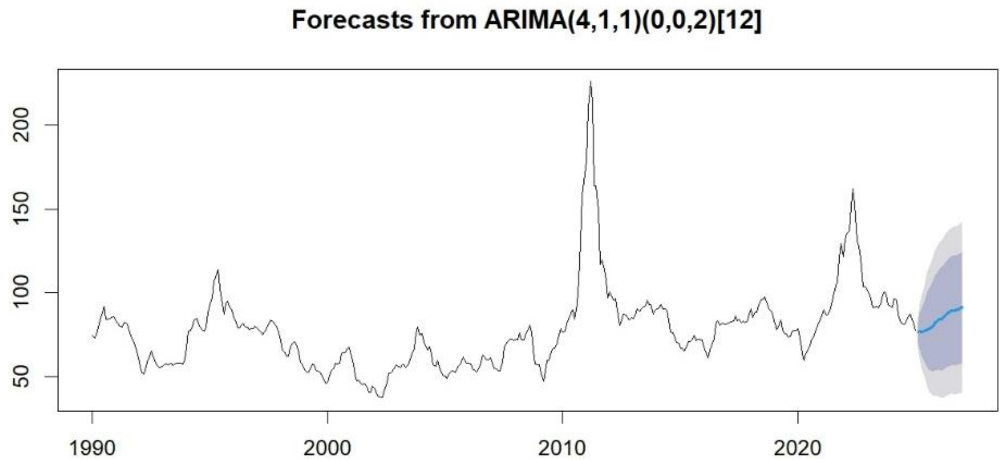

Figure 15: Forecasted monthly global price of cotton from February 2025 to January 2027.

In Figure 15, the time series reveals two notable price spikes, occurring around 2010 and 2021. The first spike corresponds to the early 2011 surge in global cotton prices, which exceeded $2.00 per pound, the highest level in over a century. This increase was driven by a combination of global cotton shortages, poor harvests, and extreme weather events, like severe droughts that disrupted supply chains (Agarwal & Narala, 2020; Anyamba et al., 2014). Prices stabilized between 90–100 cents per pound from 2012 to 2014, restoring competitiveness in the global cotton market (Agarwal & Narala, 2020). The second spike, between 2020 and 2022, reflects volatility caused by the COVID-19 pandemic and the Russia–Ukraine conflict. The pandemic severely disrupted global supply chains, particularly in the textile sector, reducing demand and rising production costs. Initial lockdowns and economic uncertainty led to a steep decline in textile consumption, which initially depressed cotton prices before contributing to their eventual rise (Emran & Schmitz, 2023). These unusual events underscore the importance of robust market monitoring and responsive policy mechanisms to manage price volatility in global commodity markets.

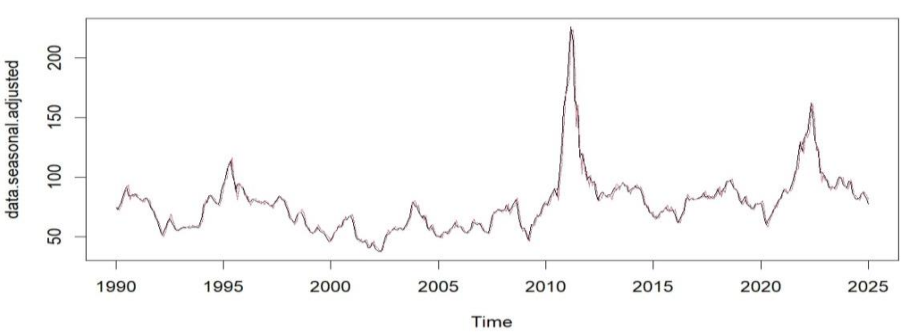

Figure 16: Observed and forecasted monthly global price of Cotton.

International Journal of Business and Economics
Vol. 10, No. 1, 2025, pp. 104-124
https://ielas.org/ijbe/index.php/ijbe
ISSN (online) 2545-4137

Table 3: *The forecasted monthly global price of cotton (February 2025 - January 2027) (U.S. Cents per Pound)*

Month	Point Forecast	Confidence Interval	
		Lo 95	Hi 95
Feb 2025	76.60242	67.06557	86.13927
Mar 2025	77.12543	60.69777	93.55308
Apr 2025	76.77418	55.48632	98.06204
May 2025	77.02752	51.02389	103.03115
Jun 2025	77.84928	47.42923	108.26933
Jul 2025	78.29492	44.22402	112.36582
Aug 2025	78.53431	41.42377	115.64484
Sep 2025	79.43717	39.76504	119.1093
Oct 2025	80.48139	38.71308	122.2497
Nov 2025	82.17247	38.71434	125.63059
Dec 2025	83.43126	38.61096	128.25156
Jan 2026	84.16978	38.25853	130.08104
Feb 2026	84.24976	37.42009	131.07942
Mar 2026	85.08351	37.5005	132.66652
Apr 2026	86.65525	38.46502	134.84548
May 2026	87.45612	38.76892	136.14331
Jun 2026	88.40024	39.30365	137.49684
Jul 2026	89.28107	39.84665	138.71549
Aug 2026	89.52291	39.80634	139.23949
Sep 2026	89.57093	39.61507	139.52679
Oct 2026	89.89362	39.7319	140.05533
Nov 2026	90.42896	40.08718	140.77074
Dec 2026	90.94829	40.44607	141.4505
Jan 2027	91.65768	41.00993	142.30542

The ARIMA (4,1,1) (0,0,2) [12] model predicts a steady increase in global monthly cotton prices, starting at 76.60 U.S. cents per pound in February 2025 and reaching 91.66 U.S. cents per pound by January 2027. The forecast indicates a consistent upward trend, with small fluctuations of about 1-2 cents per month. This suggests gradual growth in cotton prices over the forecast period, with the highest values occurring toward the end of the timeline.

4. Discussion

The ARIMA model has proven to be a practical tool for short-term forecasting of monthly global cotton prices, supporting the foundational work of Box and Jenkins (1976), who emphasized the model's ability to capture both trends and seasonality. Recent applications of ARIMA in forecasting agricultural commodity prices, such as sugarcane (Paswan et al., 2022), tomato (Adanacioglu & Yercan, 2012), and cucumber (Luo et al., 2013), further validate its effectiveness across various commodities. These studies demonstrate how ARIMA models can successfully address seasonal fluctuations and trends typical in agricultural data, providing reliable and actionable forecasts. By identifying and modeling these patterns, ARIMA facilitates better-informed policy and operational decisions, enabling stakeholders to anticipate price changes and reduce risks associated with market volatility.

Agricultural commodity prices, however, often change unpredictably due to weather variability and changing market demand, making accurate forecasting challenging. Manogna et al. (2025) suggest that hybrid models, which incorporate external factors like weather data, can improve forecasting accuracy. Likewise, Guindani et al. (2024) highlighted that external events can significantly affect agricultural commodity prices and suggested that forecasts should incorporate hybrid methods and combine them with relevant variables. Sun et al. (2023) also emphasized that combining models is an effective strategy for improving forecast reliability.

For cotton producers and market intermediaries, accurate and reliable price forecasts are critical to minimizing the risks of costly decision-making errors. Given the dynamic and fluctuating nature of the global cotton market, it is essential to carefully consider the factors influencing price forecasts to improve the reliability of future predictions (Agarwal & Narala, 2020). Even though SARIMA is a conventional approach, its application in forecasting cotton prices remains relevant. They are particularly effective when seasonal patterns are present, and recent advancements suggest that combining ARIMA with machine learning techniques can further enhance forecast accuracy by leveraging the strengths of both approaches (Ly et al., 2021). Stempień and Ślepaczuk (2025) showed that hybrid models, combining ARIMA with machine learning and deep learning methods such as Support Vector Machines and Long Short-Term Memory networks, can perform better than single models. These approaches work well because ARIMA handles linear patterns, while the machine learning components capture nonlinear dynamics in time-series data.

5. Conclusions

The Box-Jenkins methodology was successfully applied to forecast global monthly cotton prices using an ARIMA (4,1,1) (0,0,2) [12] model, highlighting a clear seasonal pattern in the data. The results demonstrate the effectiveness of time series models in providing reasonably accurate price forecasts, which can support farmers in making informed market decisions and optimizing resource allocation. Additionally, these forecasts can guide policymakers in developing price support strategies and managing risks in the supply chain. However, the ARIMA model's linearity limits its ability to capture non-linear patterns or external factors such as policy changes, climate events, and geopolitical shifts. As a univariate model, it also overlooks key exogenous variables like crude oil prices, global GDP, and exchange rates, which may limit its explanatory power. The model's fixed seasonal structure may not fully adapt to the evolving dynamics of agricultural markets, and its reliance on historical data makes it less effective in predicting unexpected events. While SARIMA provides valuable insights, it should be considered as guidance for informed decision-making rather than a definitive predictor of future prices. Future research could explore hybrid approaches, such as combining ARIMA with machine learning methods, to capture complex patterns in price behavior and improve forecasting accuracy. Applying training and testing techniques to data would also help increase the reliability of results. In addition, incorporating external factors such as geopolitical events or economic shocks could make the models more robust and reflective of real-world market conditions.

Acknowledgments

The corresponding author is deeply grateful to Dr. Yeong Nain Chi for his unwavering support, thoughtful guidance, and encouragement throughout the course of this work.

Conflict of interest

The authors declare that they have no conflicts of interest related to this publication.

References

1. Adanacioglu, H., & Yercan, M. (2012). An analysis of tomato prices at the wholesale level in Turkey: an application of the SARIMA model. Custos e@ gronegócio on line, 8(4), 52–75.
2. Agarwal, I., & Narala, A. (2020). Comparing predictive accuracy through price forecasting models in cotton.
3. Akaike, H. (1974). A New Look at the Statistical Model Identification. IEEE Transactions on Automatic Control, 19, 716–723. https://doi.org/10.1109/TAC.1974.1100705
4. Ashraf, S., Ch, K. M., Ashraf, I., & Akbar, N. (2024). A phenomenological inquiry into farmers' experiences growing cotton in Punjab, Pakistan. Scientific Reports, 14. https://doi.org/10.1038/s41598-024-62950-y
5. Bhopala, U. D., Dhandhalya, M. G., Kumar, M., & Suthar, B. (2023). Forecasting cotton prices in major domestic markets of India: An analytical approach. International Journal of Mathematics Trends and Technology, 8(6B), 172–176. https://doi.org/10.22271/maths.2023.v8.i6b.1480
6. Biswal, Dr. S. K., & Sahoo, Mrs. A. (2020). Agricultural Product Price Forecasting using ARIMA Model. International Journal of Recent Technology and Engineering (IJRTE), 8, 5203–5207. https://doi.org/10.35940/ijrte.d7606.018520
7. Box, G. E., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). Time series analysis: forecasting and control. John Wiley & Sons.
8. Box, G. E. P., & Jenkins, G. M. (1970). Time series analysis: Forecasting and control. San Francisco: Holden-Day.
9. Chi, Y. N. (2022). Time Series Modeling and Forecasting of Monthly Mean Sea Level (1978 – 2020): SARIMA and Multilayer Perceptron Neural Network. International Journal of Data Science, 3, 45–61. https://doi.org/10.18517/ijods.3.1.45-61.2022

10. Darekar, A., & Reddy, A. A. (2017). Cotton price forecasting in major producing states. Economic Affairs, 62(3), 373–378. https://doi.org/10.5958/0976-4666.2017.00047.X
11. Dickey, D. A., & Fuller, W. A. (1981). Likelihood ratio statistics for autoregressive time series with a unit root. Econometrica: journal of the Econometric Society, 1057-1072. https://doi.org/10.2307/1912517
12. Divisekara, R. W., Jayasinghe, G. J. M. S. R., & Kumari, K. W. S. N. (2021). Forecasting the red lentils commodity market price using SARIMA models. SN Business & Economics, 1(1). https://doi.org/10.1007/s43546-020-00020-x
13. Erasmus Kabu Aduteye, Tseganesh Tegegne Sete, & Yeong Nain Chi. (2023). Time series analysis of global prices of coffee: Insights into a complex market. International Journal of Business & Economics (IJBE), 8, 138–151. https://doi.org/10.58885/ijbe.v08i2.138.ea
14. Guindani, L. G., Oliveirai, G. A., Ribeiro, M. H. D. M., Gonzalez, G. V., & de Lima, J. D. (2024). Exploring current trends in agricultural commodities forecasting methods through text mining: Developments in statistical and artificial intelligence methods. Heliyon, 10(23), e40568. https://doi.org/10.1016/J.HELIYON.2024.E40568
15. Hyndman, R. J., & Koehler, A. B. (2006). Another look at measures of forecast accuracy. International journal of forecasting, 22(4), 679-688.https://doi.org/10.1016/j.ijforecast.2006.03.001.
16. Iqbal, N., Bakhsh, K. H. U. D. A., Maqbool, A. S. I. F., & Ahmad, A. S. (2005). Use of the ARIMA model for forecasting wheat area and production in Pakistan. Journal of Agriculture and Social Sciences, 1(2), 120–122.
17. Jadhav, V., Chinnappa, R. B., & Gaddi, G. M. (2017). Application of ARIMA model for forecasting agricultural prices.
18. Khadka, R., & Chi, Y. N. (2024). Forecasting the Global Price of Corn: Unveiling Insights with SARIMA Modelling Amidst Geopolitical Events and Market Dynamics. Am. J. Appl. Stat. Econ, 3, 124-135. https://doi.org/10.54536/ajase.v3i1.2776
19. Khan, M. A., Wahid, A., Ahmad, M., Tahir, M. T., Ahmed, M., Ahmad, S., & Hasanuzzaman, M. (2020). World cotton production and consumption: An overview. In Cotton Production and Uses: Agronomy, Crop Protection, and Postharvest Technologies (pp. 1–7). Springer Singapore. https://doi.org/10.1007/978-981-15-1472-2_1
20. KUMAR, A., KUMAR, V., CHETNA, GHALAWAT, S., KAUR, J., KUMARI, K., SAHARAN, H., CHHABRA, S., & RAI, B. (2024). Forecasting cotton (Gossypium spp.) prices in major Haryana markets: A time series and ARIMA approach. The Indian Journal of Agricultural Sciences, 94, 1013–1018. https://doi.org/10.56093/ijas.v94i9.150524
21. Ljung, G. M., & Box, G. E. P. (1978). On a measure of lack of fit in time series models. Biometrika, 65(2), 297–303. https://doi.org/10.1093/biomet/65.2.297
22. Luo CS, Zhou LY, Wei QF (2013) Application of SARIMA modeling in cucumber price forecast. Appl Mech Mater (Trans Tech Publications, Switzerland). 373–375:1686–1690. https://doi.org/10.4028/ www.scientific.net/AMM.373-375.1686

23. Ly, R., Traore, F., & Dia, K. (2021). Forecasting commodity prices using long-short-term memory neural networks. https://doi.org/10.2499/p15738coll2.134265
24. Manogna, R. L., Dharmaji, V., & Sarang, S. (2025). Enhancing agricultural commodity price forecasting with deep learning. Scientific Reports, 15(1), 1–24. https://doi.org/10.1038/S41598-025-05103-Z;SUBJMETA=4066,4068,704,843,844;KWRD=ENERGY+ECONOMICS,ENVIRONMENTAL+ECONOMICS
25. Montgomery, D. C., Jennings, C. L., & Kulahci, M. (2015). Introduction to time series analysis and forecasting (2nd ed.). Wiley.
26. Ojha, S., & Karki, L. B. (2025). Forecasting Global Wheat Price in the Context of Changing Climate and Market Dynamics: An Application of SARIMA Modeling Technique. American Journal of Applied Statistics and Economics, 4(1), 1–11. https://doi.org/10.54536/ajase.v4i1.4116
27. Paswan, S., Paul, A., Paul, A., & Noel, A. S. (2022). Time series prediction for sugarcane production in Bihar using ARIMA & ANN model. The Pharma Innovation Journal, 11(4), 1947-1956.
28. Schwarz, G. (1978). Estimating the dimension of a model. The Annals of Statistics, 461–464.
29. Shumway, R. H., & Stoffer, D. S. (2025). Time Series Analysis and Its Applications. Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-70584-7
30. Stempie´n, D. S., & Robert´slepaczuk, R. R. (2025). Hybrid Models for Financial Forecasting: Combining Econometric, Machine Learning, and Deep Learning Models. https://arxiv.org/pdf/2505.19617v1
31. Sun, F., Meng, X., Zhang, Y., Wang, Y., Jiang, H., & Liu, P. (2023). Agricultural Product Price Forecasting Methods: A Review. Agriculture 2023, Vol. 13, Page 1671, 13(9), 1671. https://doi.org/10.3390/AGRICULTURE13091671
32. Wanjuki, T. M., Wagala, A., & Muriithi, D. K. (2021). Forecasting Commodity Price Index of Food and Beverages in Kenya Using Seasonal Autoregressive Integrated Moving Average (SARIMA) Models. European Journal of Mathematics and Statistics, 2(6), 50–63. https://doi.org/10.24018/ejmath.2021.2.6.80
33. Weng, Y., Wang, X., Hua, J., Wang, H., Kang, M., & Wang, F. Y. (2019). Forecasting Horticultural Products Price Using ARIMA Model and Neural Network Based on a Large-Scale Data Set Collected by Web Crawler. IEEE Transactions on Computational Social Systems, 6, 547–553. https://doi.org/10.1109/TCSS.2019.2914499
34. Yule, G. U. (1971). On a method of investigating periodicities in disturbed series with special reference to Wolfer's sunspot numbers. Statistical Papers of George Udny Yule, 226, 389–420.