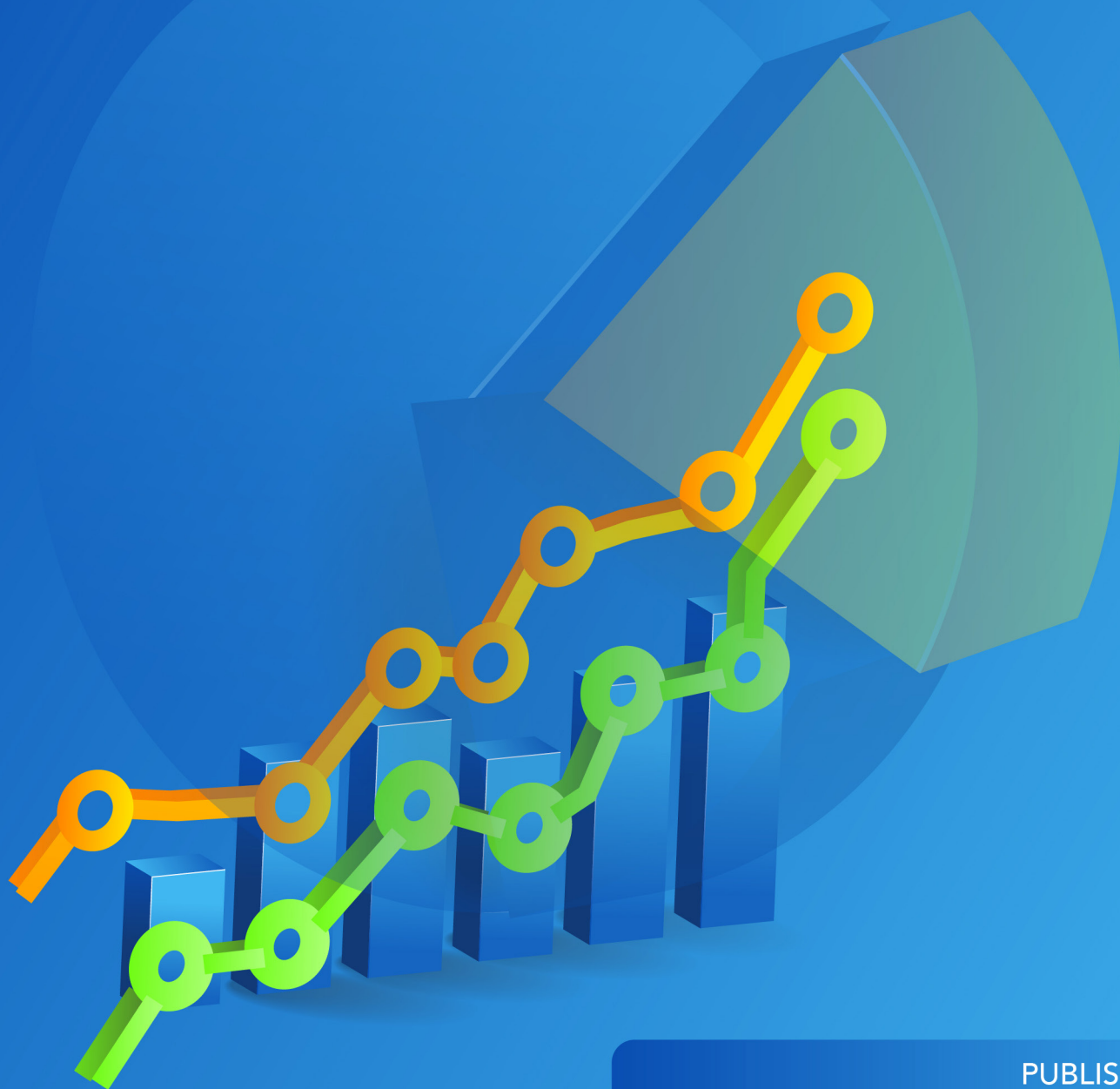


AMERICAN JOURNAL OF APPLIED STATISTICS AND ECONOMICS (AJASE)

ISSN: 2992-927X (ONLINE)

VOLUME 3 ISSUE 1 (2024)



PUBLISHED BY
E-PALLI PUBLISHERS, DELAWARE, USA

Forecasting the Global Price of Corn: Unveiling Insights with SARIMA Modelling Amidst Geopolitical Events and Market Dynamics

Raksha Khadka^{1*}, Yeong Nain Chi¹

Article Information

Received: July 05, 2024

Accepted: August 07, 2024

Published: August 10, 2024

Keywords

Corn, Price, Time Series, Forecasting, SARIMA

ABSTRACT

Corn is pivotal in global agriculture, serving diverse purposes in the food, feed, and biofuel sectors. Despite its economic significance, corn price volatility, influenced by supply-demand dynamics, climate variations, and geopolitical tensions, poses challenges in decision-making processes. This necessitates accurate price forecasting of corn for producers and government alike to formulate effective policies that uphold stability and enhance efficiency within the corn market. Using long-term records of the monthly global price of corn spanning from January 2014 to December 2023, this study employs SARIMA modeling techniques to forecast the global price of corn. To find a solution, the `auto.arima()` function from the “forecast” package in R 4.3.2 for Windows was employed to identify both the structure of the series (stationary or not) and type (seasonal or not) and sets the model’s parameters, which takes into account the AIC, AICc or BIC values generated to determine the best fitting seasonal ARIMA model. Following the Box–Jenkins methodology, the best-fitting SARIMA (0,1,1) (0,0,1) [12] model was identified, supported by the lowest AIC value. The Ljung–Box Q–test further validated the model’s adequacy in capturing the data’s behavior, with a non-significant p-value of 0.7013. This analysis uncovered valuable insights into the fluctuations of corn prices, providing a comprehensive understanding of the interplay between economic factors and external influences. This study underscores the practical utility of SARIMA modeling for farmers and other relevant stakeholders in anticipating market fluctuations and devising adaptive strategies in response to evolving corn market dynamics.

INTRODUCTION

Corn stands as one of the paramount grain crops globally, holding the prestigious rank of third, trailing only behind wheat and rice. Its cultivation sprawls across more than 100 countries, with the United States spearheading production, contributing approximately 40% of the world’s total output. Alongside the US, other key corn-producing nations encompass China, Brazil, Mexico, Indonesia, India, France, and Argentina (Darekar & Reddy, 2017). Beyond its sheer volume of production, corn plays a pivotal role in various sectors, including both food and industrial domains. Notably, corn finds its way into the production lines of diverse products, prominently featuring in the creation of starch (Yu & Moon, 2021). Corn’s multifaceted utility extends to its applications in livestock feed production (Fauziah *et al.*, 2023), corn oil extraction (Wheals *et al.*, 1999), and the production of ethanol, a renewable and widely used biofuel. Ethanol production from corn boasts an impressive conversion rate, generating 2.7 gallons of ethanol per bushel of corn (Baker & Zahniser, 2006). Given its indispensable role in various industries, any fluctuations in the price of corn reverberate across sectors, impacting stakeholders at various levels.

Price volatility in commodity markets stems from a myriad of factors, encompassing intricate interplays of supply-demand dynamics, crop yield variations, geopolitical tensions, economic downturns, and even

global health crises. Corn prices are inherently susceptible to these forces and have witnessed significant oscillations over time. In this context, harnessing the power of time series analysis emerges as a potent tool for navigating the complexities of pricing decisions.

Time series analysis entails the systematic examination of data points recorded over sequential time intervals. These data, arranged chronologically, offer insights into the temporal evolution of a specific variable. Represented mathematically as $y(t)$, where ‘y’ denotes the variable of interest and ‘t’ denotes time, time series data enable analysts to discern patterns, trends, and anomalies, thereby empowering informed decision-making (Montgomery *et al.*, 2015). By leveraging historical trends, forecasting techniques embedded within time series analysis equip market participants with valuable foresight, facilitating proactive adjustments to pricing strategies in anticipation of future market dynamics.

Moreover, the significance of corn transcends mere economic considerations. It is deeply intertwined with agricultural practices, environmental sustainability, and food security on a global scale. As populations burgeon and climates fluctuate, the resilience and adaptability of corn cultivation become increasingly pivotal. Understanding the intricate dynamics governing corn prices not only informs commercial decisions but also holds implications for broader socio-economic and environmental contexts. Furthermore, the evolution of technology, particularly

¹ Department of Agriculture, Food, & Resource Sciences, School of Agricultural and Natural Sciences, University of Maryland Eastern Shore (UMES), Princess Anne, MD 21853, USA

* Corresponding author’s e-mail: rkhadka@umes.edu

in data analytics and computational methods, has revolutionized the landscape of market analysis. Advanced statistical models and machine learning algorithms offer unprecedented capabilities in extracting actionable insights from voluminous datasets. Integrating these technological advancements with traditional economic principles enhances the efficacy of pricing strategies, positioning market participants to navigate the intricacies of the corn market with greater precision and agility.

In light of these considerations, this paper embarks on a comprehensive exploration of corn price dynamics, employing a multifaceted approach that melds economic theory with cutting-edge analytical methodologies. Through a nuanced examination of historical trends, statistical modeling, and forecast projections, this study endeavors to shed light on the underlying drivers of corn price fluctuations, thereby empowering stakeholders with actionable intelligence to optimize pricing strategies and mitigate risks in the volatile landscape of commodity markets.

LITERATURE REVIEW

The fluctuations in global corn prices over the years have been influenced by a myriad of interconnected factors, resulting in a complex and dynamic market landscape. These fluctuations can be attributed to shifts in global demand, disruptions in supply chains, geopolitical events, disease outbreaks, sudden climate changes, increased demand for corn-based products (such as biofuels), and speculation in commodity markets. The COVID-19 pandemic in 2020 exemplified how external shocks can significantly impact corn prices and production, particularly in the United States. Disruptions in ethanol, gasoline, and oil markets led to a notable decrease in corn prices, as highlighted by Schmitz *et al.* (2020). Beghin & Timalsina (2020) observed a decrease in corn prices from \$3.74 per bushel in December 2019 to \$2.94 per bushel in

May 2020, while Liu *et al.* (2024) noted that the pandemic influenced subsequent increases in global corn prices.

Following the economic slowdown as a result of COVID-19, biofuel prices experienced a significant decline in 2020, followed by their main feedstocks, maize, and oilseeds (Elleby *et al.*, 2020). Additionally, the Russia-Ukraine conflict, which began in 2022, exacerbated the situation by causing an energy crisis and disrupting food production and commodity markets, including corn prices, as both countries are major exporters of staple crops (Avalos & Huang, 2022). The conflict also highlighted the paradoxical potential for increased biofuel usage to moderate rising oil prices, consequently boosting demand and prices for corn, a crucial feedstock for ethanol production. Furthermore, fluctuations in both supply and demand for corn, driven by factors such as natural conditions, imports, changing needs in animal feed, food production, and alternative energy sources, have contributed to consumer-level price volatility (Baladina *et al.*, 2021).

The seasonal fluctuations in crop production further exacerbate the challenge of market strategy and investment planning due to price uncertainty (Brandt & Bessler, 1983). Consequently, accurate price forecasting becomes essential to assist farmers, stakeholders, consumers, policymakers, and investors in making informed decisions. Numerous studies have focused on forecasting commodity prices, including corn, utilizing various time series forecasting models.

Table 1 presents some of the identified models used for forecasting corn prices, along with the evaluation tools and performance metrics employed by different researchers. Similarly, Table 2 illustrates the diversity of applications of time series forecasting models in value forecasting across different fields, showcasing the range of models utilized and the performance metrics evaluated by various authors.

Table 1: SARIMA models identified by authors for commodity price forecasting

Commodities	Models Identified	Evaluation Tools/Performance Metrics	Contributors
Corn	$(2,1,0)(3,1,1)_{12}$	MSE, RMSE, Ljung-Box Q	(Lv & Wu, 2022)
Soybean	$(0,1,3)(0,0,2)_{12}$	MSE	(Chi, 2021)
Red Lentil	$(2,1,2)(0,1,1)_{52}$	AIC, Ljung-Box Q	(Divisekara <i>et al.</i> , 2020)
Potato	$(1,1,1)(1,0,0)_{12}$	SBC, AIC	(Chandran & Pandey, 2007)
Tomato	$(2,1,1)(1,0,1)_{12}$	AIC	(Mutwiri, 2019)
Bajra	$(0,1,1)(0,1,1)_{12}$	AIC, SBC, MAD, MAPE, MSE	(Sharma & Burark, 2015)

Table 2: SARIMA models identified by various authors for value forecasting

Forecasts	Models Identified	Evaluation Tools/Performance Metrics	Contributors
Monthly mean surface air temperature	$(2,1,1)(1,1,2)_{12}$	ME, RMSE, MAE, MPE, MAPE, MASE, BIC	(Asamoah-Boaheng, 2014)
Frequency of monthly rainfall	$(1,0,1)(1,1,1)_{12}$	S.E, Variance, AIC	(Adams <i>et al.</i> , 2019)
Temperatures	$(1,1,1)(1,0,1)_{12}$	AIC	(Chen <i>et al.</i> , 2018)

Corn plays a crucial role in ethanol production, a biofuel often blended with gasoline. Over the last two decades, ethanol production has been the only use of corn in the United States that has seen a notable increase, consuming about 40% of the U.S. harvest on average over the past five years (Avalos & Huang, 2022). Higher oil prices incentivize gasoline blenders to increase the ethanol content in their products, potentially mitigating oil price spikes but simultaneously boosting corn demand and prices (Avalos & Huang, 2022). This inflation was fueled by post-pandemic economic adjustments since mid-2020, and persistent constraints on aggregate supply due to disruptions in global supply chains (Goryunov *et al.*, 2023).

Economic challenges, including accelerating inflation since 2021, further compounded the situation, impacting countries worldwide. In the USA, inflation reached close to 10%, while it was even higher in the euro area (Goryunov *et al.*, 2023). Although inflation rates began to decrease by 2023, they remained elevated, reflecting the enduring impact of the factors influencing corn prices on the global economy. Dohlman *et al.* (2024) forecasts suggest that there will likely be a decrease or stability in crop prices from 2024 to 2033. The United States Department of Agriculture (USDA) also anticipates a decline in corn prices to \$4.50 per bushel, followed by a period of stabilization around the 2025/26 timeframe (Dohlman *et al.*, 2024). Due to the abundant supply of corn in the United States, it is anticipated that corn prices will experience a decline throughout 2024 (UGA Cooperative Extension, 2024).

In addition to understanding the factors influencing price fluctuations, forecasting models play a crucial role in providing valuable insights for decision-making processes across various sectors, ranging from agriculture to climate science. By utilizing sophisticated modeling techniques and evaluating performance metrics, researchers strive to enhance the accuracy and reliability of forecasting models to better navigate the complexities of global markets and environmental systems.

MATERIALS AND METHODS

The main purpose of this study was to demonstrate the role of the time series model in predicting processes and to pursue the analysis of time series data using long-term records of the monthly global price of corn from January 2014 to December 2023. The monthly global price of corn, (units: U.S. dollars per metric ton, monthly, not seasonally adjusted) from January 2014 to December 2023, is available to the public from International Monetary Fund, Global price of Corn [PMAIZMTUSD], retrieved from FRED, Federal Reserve Bank of St. Louis; <https://fred.stlouisfed.org/series/PMAIZMTUSD>, March 10, 2024. The average monthly global price of corn from January 2014 to December 2023 was \$200.6 U.S. dollars per metric ton with a standard deviation of \$84.35 (Minimum: \$144.0, Maximum: \$348.5, and Median: \$171.9).

Time series analysis is based on the underlying assumption that the data is stationary. Thus, it is crucial to identify whether time series data is stationary or non-stationary. Data is considered stationary if its mean and variance do not change over time. Conversely, non-stationary data exhibit a long-term increase or decrease over time, along with periodic fluctuations and changes in variance. A series is strictly stationary if the marginal distribution of Y at time t [$P(Y_t)$] is the same as at any other point in time. $P(Y_t) = P(Y_{t+k})$ and $P(Y_t, Y_{t+k})$ does not depend on t ($t \geq 1$ and k is any integer). This implies that the mean, variance, and covariance of the series Y_t are time-invariant. However, a series is said to be weakly stationary if the following conditions are met:

$$E(Y_1) = E(Y_2) = \dots = E(Y_t) = \mu \quad (1)$$

$$\text{Var}(Y_1) = \text{Var}(Y_2) = \dots = \text{Var}(Y_t) = \gamma_0 \text{ (a constant)} \quad (2)$$

$$\text{Cov}(Y_t, Y_{t+k}) = \text{Cov}(Y_2, Y_{2+k}) = \dots = \text{Cov}(Y_t, Y_{t+k}) = \gamma_k \text{ (depends only on lag } k) \quad (3)$$

Because the statistical properties of time series data change over time in non-stationary data, it is difficult to make predictions and draw conclusions. We cannot determine the appropriate model. Non-stationary data may sometimes result in false regressions, meaning the regression equation shows a significant relationship between two variables when there wasn't any.

The Dickey-Fuller Test was developed by David Dickey and Wayne Fuller in 1979. It tests for the presence of trends and seasonality in data. This test tests the presence of unit roots (Phillips & Perron, 1988) in an autoregressive model. If a unit root is present, it means the data is non-stationary suggesting that it is difficult to model and forecast the future values. Consider the following hypothesis:

Null Hypothesis (H_0) = The model is non-stationary (the time series has a unit root)

Alternate Hypothesis (H_1) = The model is stationary (the time series does not have a unit root).

If the test statistic produced by the Dickey-Fuller test is at the significance level of 0.05, the null hypothesis is rejected and the data set suggests stationarity. Again, if the test statistics are greater than the significance level i.e., >0.05 , the time series suggests the non-stationarity of data which implies there is a need for differencing to make the series stationary before applying time series models such as ARIMA.

It is a method of transforming a non-stationary time series into a stationary time series. It is used in removing the trend in the time series (McGonigle *et al.*, 2022). This is an important step in preparing data to be used in an ARIMA model. When the difference between the current period and the previous period is made, it is called first-order differencing and can be denoted as $I(1)$. It can be expressed as:

$$\Delta y_t = y_t - y_{(t-1)} \quad (4)$$

Where y_t represents time series data at time t and the current value we are trying to model. If the values exhibit non-stationarity properties, the process is repeated twice and thus called second-order differencing and can be

denoted as I(2). It can be expressed as:

$$\Delta^2 y_t = (y_t - y_{(t-1)}) - (y_{(t-1)} - y_{(t-2)}) \quad (5)$$

Where y_t represents time series data at time t and the current value we are trying to model. This process is continued until the values show stationary properties (constant mean and variance). A series that is stationary after being differentiated d times is said to be integrated of order d , denoted by I(d). However, when a series is stationary without differencing is said to be I(0).

Autoregressive (AR) Model

AR models are used to forecast future values only based on their previous values, typically called lags. Thus, the forecasted value ‘ Y ’ and time ‘ t ’ in AR is the function of its past values $Y_{t-1}, Y_{t-2}, Y_{t-3}, Y_{t-4}, \dots$. Thus,

$$Y_t = f(Y_{(t-1)}, Y_{(t-2)}, Y_{(t-3)}, Y_{(t-4)}, \dots, \epsilon_t) \quad (6)$$

The model can be mathematically expressed as:

$$y_t = \beta_0 + \beta_1 y_{(t-1)} + \epsilon_t \quad (7)$$

where: y_t represents the value of the variable at time t , β_0 is the constant term, β_1 is the coefficient of the lagged variable, y_{t-n} represents the effect of the n th period’s value on the current period, ϵ_t is the error term at time t and represents the deviation of the actual value from the predicted value based on the model.

AR models that depend only on one lag in the past are called First-order Auto-Regressive Model or AR(1) models and can be expressed as:

$$y_t = \beta_0 + \beta_1 y_{(t-1)} + \epsilon_t \quad (8)$$

AR models that depend only on two lags in the past are called Second-order Auto-Regressive Model or AR(2) models and can be expressed as:

$$y_t = \beta_0 + \beta_1 y_{(t-1)} + \beta_2 y_{(t-2)} + \epsilon_t \quad (9)$$

The variables of interest in the AR model are forecasted using the linear combination of the past values of the variable and the term ‘Autoregression’ indicates that it is a regression of the variable against itself. Hence, the AR model of order p can be represented as:

$$y_t = \beta_0 + \beta_1 y_{(t-1)} + \beta_2 y_{(t-2)} + \beta_3 y_{(t-3)} + \beta_4 y_{(t-4)} + \beta_5 y_{(t-5)} + \beta_6 y_{(t-6)} + \dots + \beta_p y_{(t-p)} + \epsilon_t \quad (10)$$

The AR models are also called long-memory models as they can take into account a large range of past observations to predict the current value.

Moving Average (MA) Model

Models used to forecast the series based solely on the past errors in the series i.e., error lags are called Moving Average (MA) models. In MA, the forecasted value ‘ Y ’ at a time ‘ t ’ is the function of the lags of its error value at the time ‘ t ’,

$$Y_t = f(\epsilon_t, \epsilon_{(t-1)}, \epsilon_{(t-2)}, \epsilon_{(t-3)}, \dots) \quad (11)$$

The model MA(q) can be mathematically expressed as:

$$y_t = \mu + \epsilon_t + \theta_1 * \epsilon_{(t-1)} + \theta_2 * \epsilon_{(t-2)} + \theta_3 * \epsilon_{(t-3)} + \theta_4 * \epsilon_{(t-4)} + \dots + \theta_q * \epsilon_{(t-q)} \quad (12)$$

where: y_t represents time series data at time t and the current value we are trying to model, μ is the mean of the time series data and the expected value of y_t when all other terms in the equation are zero, ϵ_t is the error term at time t and represents the deviation of the actual

value from the predicted value based on the model, and θ represents the weight assigned to the lagged error terms in the model.

MA models that depend upon only one error lag are called first-order MA models, denoted by

$$MA(1): y_t = \mu + \epsilon_t + \theta_1 * \epsilon_{(t-1)} \quad (13)$$

The second-order MA model, denoted by

$$MA(2) is: y_t = \mu + \epsilon_t + \theta_1 * \epsilon_{(t-1)} + \theta_2 * \epsilon_{(t-2)} \quad (14)$$

Moving average models are the short memory models since the errors in them don’t last long into the future.

Auto-Regressive Moving Average (ARMA) Model

ARMA is the combination of AR and MA models and is used to describe the behavior of the time series and to forecast future values based on historical data. The ARMA(p,q) model can be mathematically represented as:

$$y_t = c + \beta_1 y_{(t-1)} + \beta_2 y_{(t-2)} + \beta_3 y_{(t-3)} + \dots + \beta_p y_{(t-p)} + \theta_1 * \epsilon_{(t-1)} + \theta_2 * \epsilon_{(t-2)} + \theta_3 * \epsilon_{(t-3)} + \dots + \theta_q * \epsilon_{(t-q)} + \epsilon_t \quad (15)$$

where: y_t represents time series data at time t and the current value we are trying to model, c is a constant term, $\beta_1, \beta_2, \beta_3, \dots, \beta_p$ are the coefficients of the autoregressive part, $\theta_1, \theta_2, \theta_3, \dots, \theta_q$ are the coefficients of the moving average part, and ϵ_t is the error term at time t . ARMA model is purely stationary without difference and a blend of AR and MA models.

Auto-Regressive Integrated Moving Average (ARIMA) Model

ARIMA models, also known as Box-Jenkins models (Montgomery *et al.*, 2015), are a statistical method useful for forecasting data based on their temporal structures. This model is useful for analyzing and forecasting time series exhibiting trends and seasonal patterns. It is, in simple terms, an integration of the Auto Regression and Moving Average model with differencing made to make the time series stationary. Its parameters can be represented as (p, d, q) and can be expressed as ARIMA (p, d, q), where, p = number of lag observations or the order for the autoregression, d = order of differencing made to make the data stationary for the non-stationary series, and q = order of the moving average.

Seasonal Auto-Regressive Integrated Moving Average (SARIMA) Model

A widely used variation of the ARIMA model called the SARIMA model (Chen *et al.*, 2018) was proposed by Box and Jenkins. This model shows the significance of seasonality (Vaswani *et al.*, 2023). It integrates both the seasonal and non-seasonal components (Adams *et al.*, 2019) and can be represented as: ARIMA (p, d, q) (P, D, Q)(s), where, p = non-seasonal Auto Regression order, d = non-seasonal differencing order, q = non-seasonal Moving Average order. Similarly, P = seasonal Auto Regression order, D = seasonal differencing order, Q = seasonal Moving Average order, and s = period in a seasonal pattern.

Box-Jenkins Approach

The Box-Jenkins approach to modeling was developed by George Box and Gwilym Jenkins in the early 1970s. This methodology consists of three major stages: Identification, Estimation, and Diagnostic Checking. At the identification stage, stationarity is checked. The process advances to model estimation if the data is stationary; otherwise, the data is transformed to make it stationary through decomposition and differencing methods. The Autocorrelation Function (ACF) and Partial Auto-Correlation functions (PACFs) are visualized in this stage.

After identifying the model structure, parameters for the model are estimated by testing least squares estimation, such as Mean Error (ME), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Percentage Error (MPE), Mean Absolute Percentage Error (MAPE), and Mean Absolute Scale Error (MASE). ME represents the average error, whereas RMSE is the square root of the average of the squared errors. Similarly, MAE represents the average of the absolute errors, MPE is the average of the percentage errors, and MAPE is the average of the absolute percentage of errors. Likewise, MASE measures prediction accuracy, and ACF1 is the first-order autocorrelation of the residuals. Lower values indicate a better-fitting model. Maximum Likelihood Estimation is also one of the techniques used for estimating models, such as Akaike’s Information Criterion (AIC), AICs, Bayesian Information Criterion (BIC), etc.

To determine the representativeness of the model regarding the dataset, diagnostic tests are conducted after model estimation (Young, 1977). The diagnostic test ensures whether or not the model adequately captures the underlying patterns and features of the data. A famous test called the Box-Ljung test (Ljung & Box, 1978), used in time series analysis, determines the presence/absence

of autocorrelations among the residuals. The residuals are examined for randomness and autocorrelation, i.e., the graphs, test statistics, ACFs, and PACFs of the residuals are used for model verification. The steps of Identification, Estimation, and Diagnostics are repeated if the model is not verified as the best model. After completing the above-mentioned stages, the model is ready to forecast the future values of any time series data.

RESULTS AND DISCUSSION

A forecast of the global price of corn was attempted using the SARIMA (Seasonal Auto-Regressive Integrated Moving Average) model. R 4.3.2 for Windows was used to model and forecast the monthly global price of corn from January 2014 to December 2023, retrieved from FRED and converted into a time series object. The time series plot (Figure 1) and its lag plot (Figure 2) of global corn

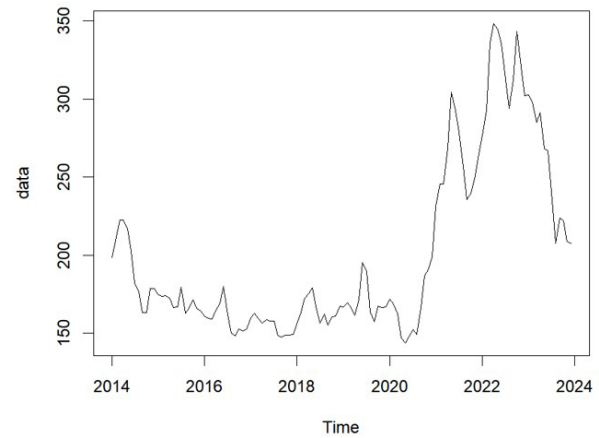


Figure 1: Time series plot of the monthly global price of Corn (January 2014 ~ December 2023)

Source: Own computation based on FRED (January 2014 ~ December 2023) data

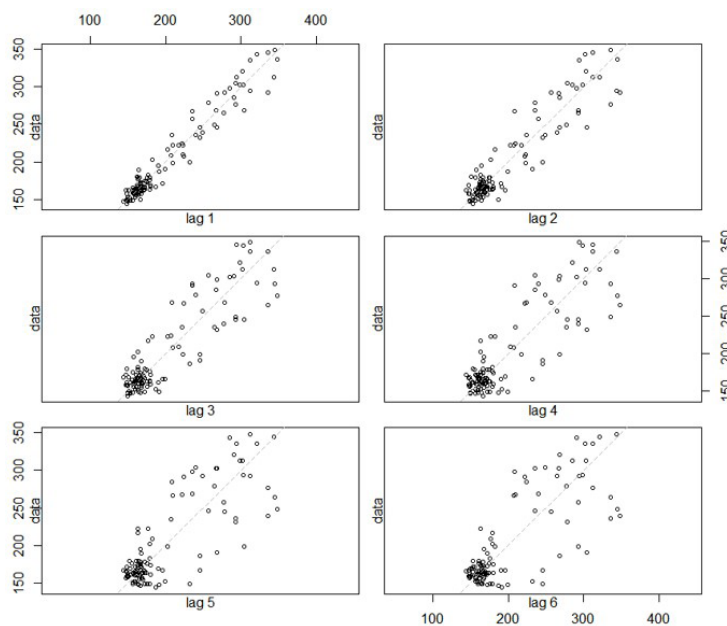


Figure 2: Lag plot of the monthly global price of Corn (January 2014 ~ December 2023)

Source: Own computation based on FRED (January 2014 ~ December 2023) data

prices recorded from January 2014 to December 2023 show fluctuation over time but show some trends and patterns. Seasonal variations are evidenced in the prices which can be potentially influenced by weather events, harvest periods, demand, etc. There are fluctuations within each year. There appear to be periods of both increases and decreases in the prices over the years. The price was lowest in 2020 and was a record high in 2022. Its stationarity was checked using the Augmented Dickey-Fuller (ADF) test. The value of ADF statistic was -2.3498 which reflects that the data is weakly stationary. The p-value suggests the level of significance of any observation. The high p-value (typically $> \alpha = 0.05$) suggests that we fail to reject the null hypothesis of non-stationarity. Here, the p-value (0.4313) obtained from the ADF test for the original data suggested rejecting the null hypothesis of non-stationarity. Autocorrelation (Figure 3)

and partial autocorrelation (Figure 4) were also visualized. In the ACF plot (Figure 3) it can be seen that there is a gradual decay of spikes but never cut off to zero meaning the data needs to be made stationary for further testing. Using decomposition, the monthly global price of corn time series was decomposed into three components - trend, seasonal, and random - and each component was visualized (Figure 5) which shows the decomposed corn price for the various years recorded. It also can be observed that the existence of the seasonal variation is constant over time. The random effect also seems to be constant over time. However, the trend of the series seems to be constant till 2020 and then gradually rising upwards peaking at 2022 and gradually sliding downwards after that. The seasonal component was removed to make the data stationary (Figure 6), and the stationarity was again tested using the ADF test.

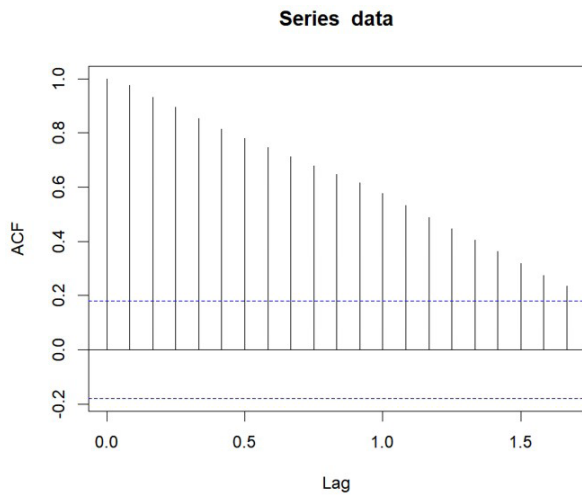


Figure 3: ACF plot of the monthly global price of Corn (January 2014 ~ December 2023)
 Source: Own computation based on FRED (January 2014 ~ December 2023) data

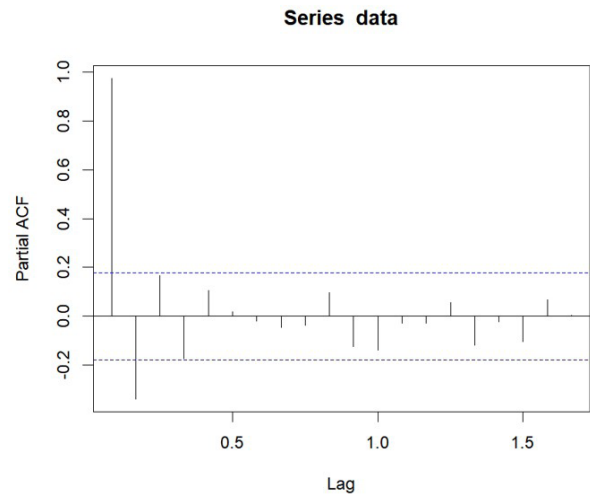


Figure 4: PACF plot of the monthly global price of Corn (January 2014 ~ December 2023)
 Source: Own computation based on FRED (January 2014 ~ December 2023) data

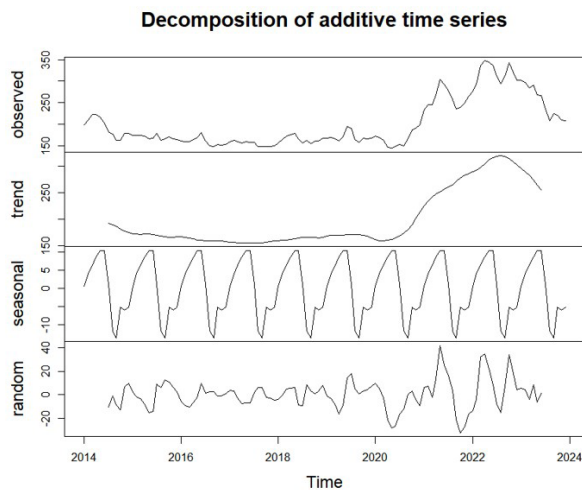


Figure 5: Decomposition of the monthly global price of Corn (January 2014 ~ December 2023)
 Source: Own computation based on FRED (January 2014 ~ December 2023) data

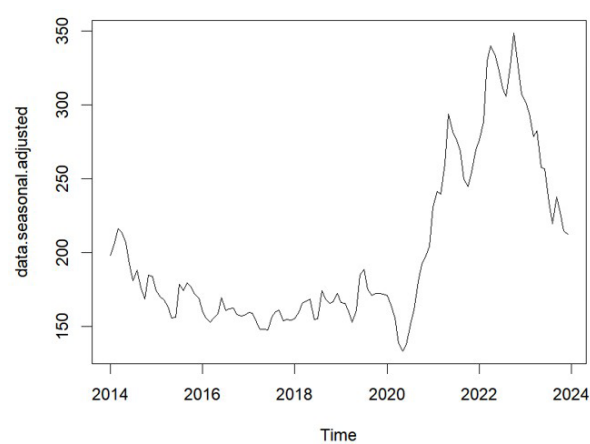


Figure 6: Time series plot of seasonally adjusted monthly global price of Corn (January 2014 ~ December 2023)
 Source: Own computation based on FRED (January 2014 ~ December 2023) data

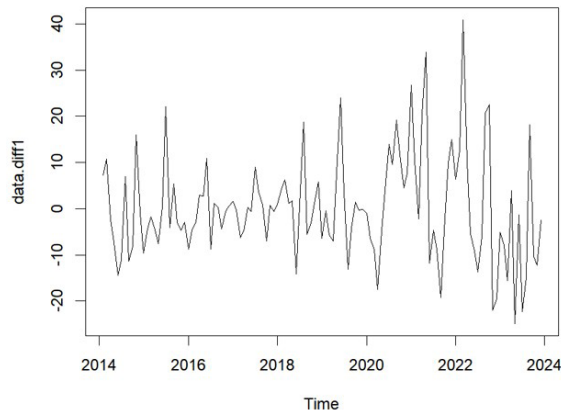


Figure 7: Time series plot of the first difference of the seasonally adjusted monthly global price of Corn (January 2014 ~ December 2023)
 Source: Own computation based on FRED (January 2014 ~ December 2023) data

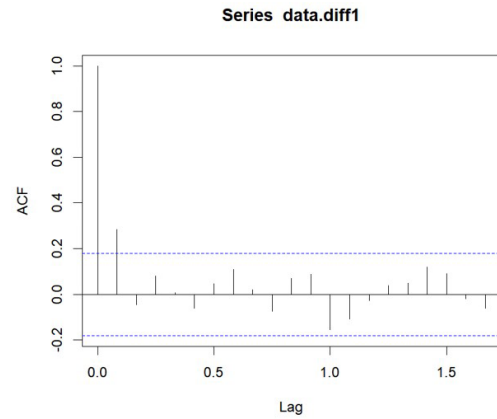


Figure 8: ACF plot of the first difference of the seasonally adjusted monthly global price of Corn (January 2014 ~ December 2023)
 Source: Own computation based on FRED (January 2014 ~ December 2023) data

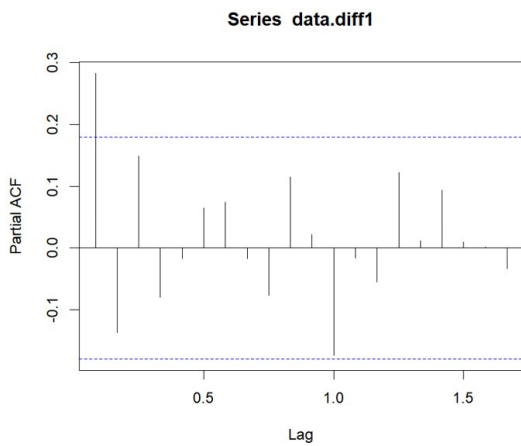


Figure 9: PACF plot of the first difference of the seasonally adjusted monthly global price of Corn (January 2014 ~ December 2023)
 Source: Own computation based on FRED (January 2014 ~ December 2023) data

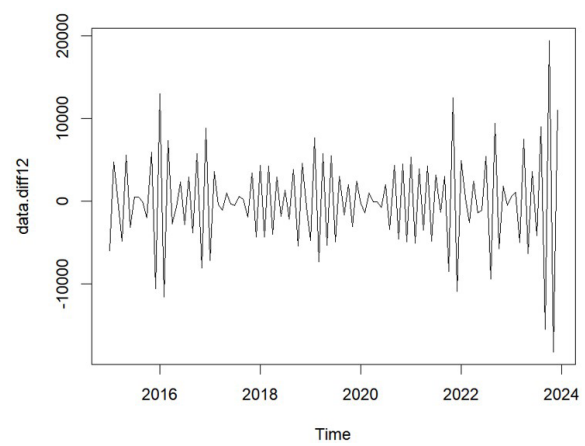


Figure 10: Time series plot of the twelfth difference of the seasonally adjusted monthly global price of Corn (January 2014 ~ December 2023)
 Source: Own computation based on FRED (January 2014 ~ December 2023) data

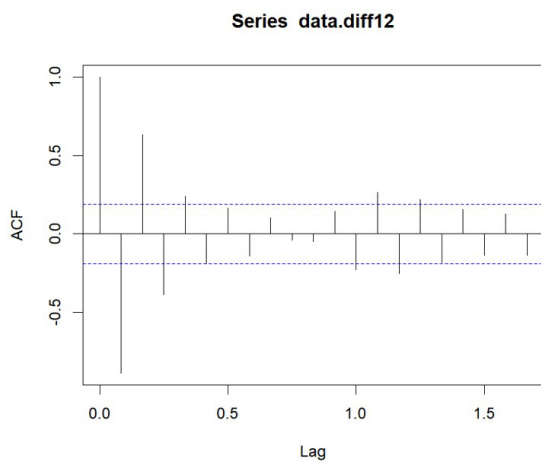


Figure 11: ACF plot of the twelfth difference of the seasonally adjusted monthly global price of Corn (January 2014 ~ December 2023)
 Source: Own computation based on FRED (January 2014 ~ December 2023) data

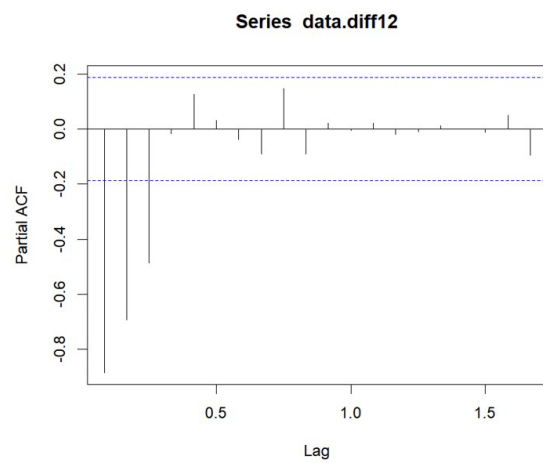


Figure 12: PACF plot of the twelfth difference of the seasonally adjusted monthly global price of Corn (January 2014 ~ December 2023)
 Source: Own computation based on FRED (January 2014 ~ December 2023) data

Given the p-value of 0.5211, which is quite high, we again failed to reject the null hypothesis and the data is still non-stationary. So, differencing (Figures 7 & 10) was attempted. The test statistics thus obtained were negative and significant

The p-value was 0.01 in both cases which is $< \alpha = 0.05$ suggesting the stationarity of data. The differenced series removed the trends and seasonality, making the data more stationary and ready for modeling. ACF (Figures 8 & 11) and PACF (Figures 9 & 12) from both cases were also visualized. These plots show the ACF and PACF of the global corn price series with 95% confidence limits. It can be seen from the ACF plot that the spike is significant till the first lag and there are no significant spikes in the PACF plot.

In terms of identifying and fitting the model, the best-fitting seasonal ARIMA model was determined by using `auto.arima()` function in R which automatically took into account the AIC, AICc, or BIC values. The notation in the

ARIMA model represents the parameters of the ARIMA model. It indicates the order of the Autoregressive (AR), differencing (I), and moving average (MA) components of the model. In this case, an ARIMA (0,1,1)(0,0,1) [12] model was obtained suggesting the series has a seasonal pattern and a first-order non-seasonal moving average term and a first-order seasonal moving average component has been influencing the series.

Thus, it is ARIMA (0,1,1) for the non-seasonal part and ARIMA (0,0,1) for the seasonal part, with a seasonal period of 12 months. This simply also suggested a moving average term and a seasonal moving average term. The coefficient for the MA term was 0.4628, and for the seasonal MA was -0.3777 and the estimated variance of the residuals was 105.9 (Table 3). The log-likelihood of the model is the measure of how well the model fits the data and it was -446.29 in this case. The model was selected based on AIC and was evaluated using its summary statistics.

Table 3: Parameters of the ARIMA(0,1,1)(0,0,1)₁₂ Model

Parameter	Estimate	Standard Error
Difference	1	
MA1	0.4268	0.0919
SMA1	-0.3777	0.1312
Sigma ² = 105.9; Log Likelihood = -446.29		
AIC = 898.57, AICc = 898.78, BIC = 906.91		
RMSE = 10.16116, MAE = 7.525733, MAPE = 3.665245		

Source: Own computation based on FRED (January 2014 ~ December 2023) data

The residual analysis was performed for the model verification. It can be seen from the Standardized Residuals (Figure 13) that the residuals of the model have zero mean and a constant variance concentrated around 2 and -2. It can be visualized from an ACF plot of residuals (the second panel of Figure 13) that there is no autocorrelation among the residuals; the autocorrelation is zero. This implies zero mean and constant variance among the residuals. Hence, the residuals follow a white noise process. The adequacy of the model was also judged by the Ljung-Box test and the Box-Ljung test which were used to assess the autocorrelation of the residuals from the model where autocorrelation means that the values of the series at different points in time are correlated with each other.

In the Ljung-Box test, the test statistics Q^* are compared against a chi-squared distribution with degrees of freedom equal to the number of lags used in the test minus the model degrees of freedom. The p-value (0.9303) obtained suggested no significant autocorrelation in the residuals. In the Box-Ljung test

also, the X-squared statistic is compared against a chi-squared distribution. The p-value (0.7013) in this test also indicated no significant autocorrelation in the residuals. This can also be visualized from the p-value of the residuals (the third panel of Figure 13) that its value exceeds the 5% confidence interval showing no significant departures from the white noise of the residuals. Also, from the normality plot of residuals (Figure 14), the residuals follow the normal distribution process i.e., there is a symmetric distribution of the residuals around the mean residual. Thus, SARIMA (0,1,1) (0,0,1) [12] was selected to be the best-fitting model for this time series.

Forecasts were generated using the fitted ARIMA model and the forecasted values were visualized (Figure 15 and Figure 16). The forecast values – January to December 2025 – seem to be relatively stable over the period, with a slight upward trend. The forecast values for each month in the first year (2024) are generally lower than those in the second year (2025) (Table 4).

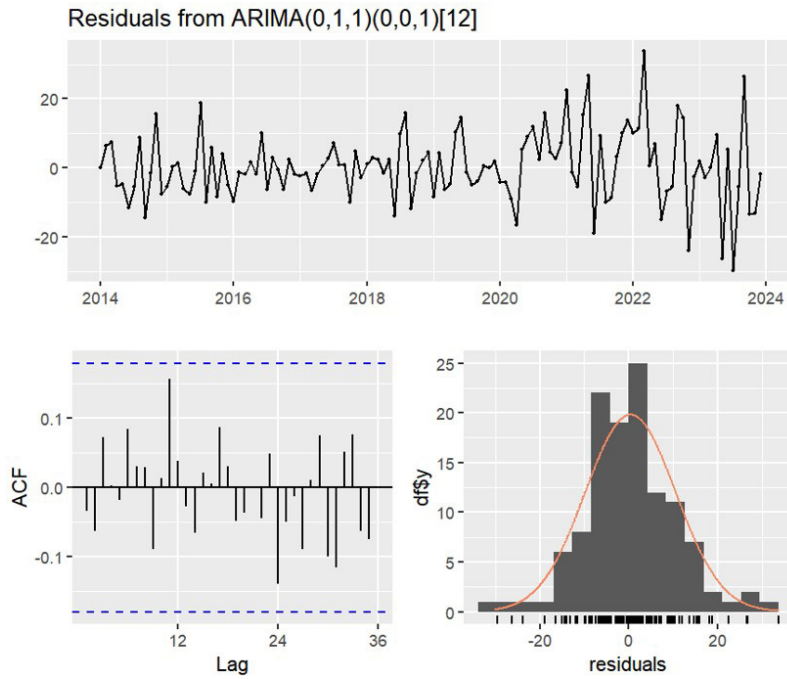


Figure 13: Ljung-Box test of the seasonally adjusted monthly global price of Corn (January 2014 ~ December 2023)
 Source: Own computation based on FRED (January 2014 ~ December 2023) data

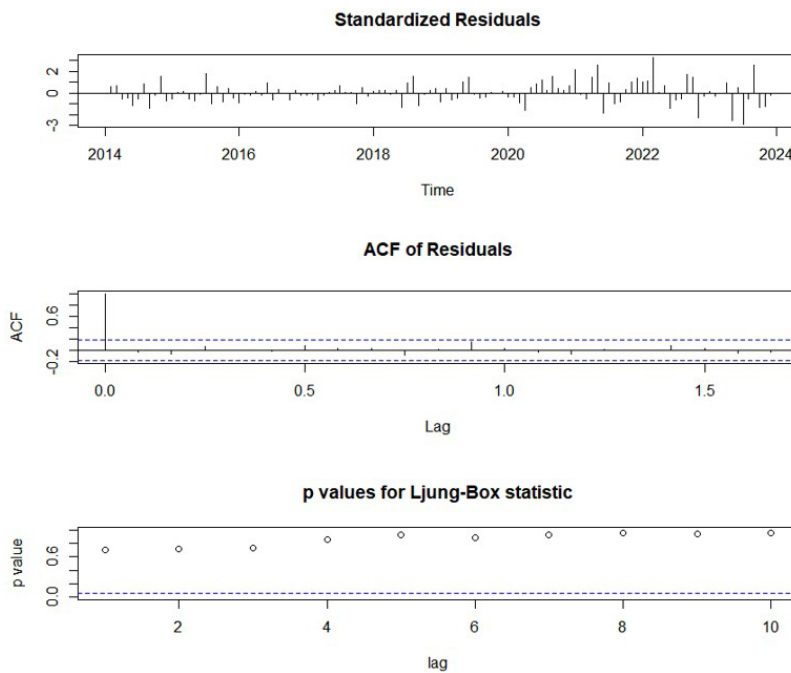


Figure 14: Residual plots of the seasonally adjusted monthly global price of Corn (January 2014 ~ December 2023)
 Source: Own computation based on FRED (January 2014 ~ December 2023) data

Table 4: The forecasted monthly global price of Corn for 2024 and 2025

Month	Point Forecast		Lower 95% Confidence Limit		Upper 95% Confidence Limit	
	2024	2025	2024	2025	2024	2025
January	211.4124	238.2565	191.2432	138.4384	231.5816	338.0746
February	212.1315	238.2565	176.9903	136.8447	247.2727	339.6682
March	212.5306	238.2565	167.1102	135.2757	257.9509	341.2373
April	208.9084	238.2565	155.1393	133.7303	262.6775	342.7827

May	217.2596	238.2565	156.2742	132.2074	278.2450	344.3056
June	219.4492	238.2565	152.0154	130.7060	286.8831	345.8070
July	229.7617	238.2565	156.4444	129.2253	303.0790	347.2877
August	236.5466	238.2565	157.7841	127.7644	315.3091	348.7486
September	227.4025	238.2565	143.5477	126.3226	311.2574	350.1904
October	228.1253	238.2565	139.4701	124.8992	316.7805	351.6138
November	235.2223	238.2565	142.0137	123.4934	328.4309	353.0196
December	237.9829	238.2565	140.4332	122.1046	335.5326	354.4084

Source: Own computation based on FRED (January 2014 ~ December 2023) data

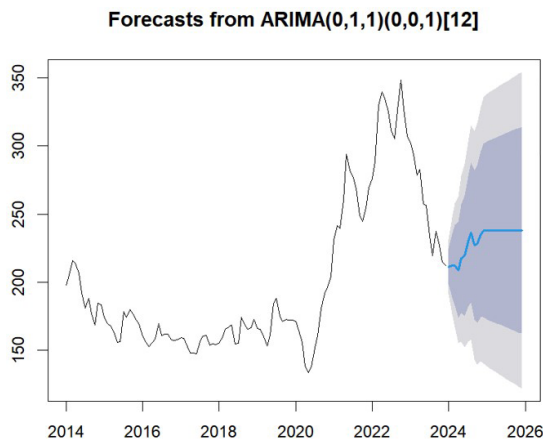


Figure 15: Forecasted monthly global price of Corn under upper and lower confidence levels

Source: Own computation based on FRED (January 2014 ~ December 2023) data

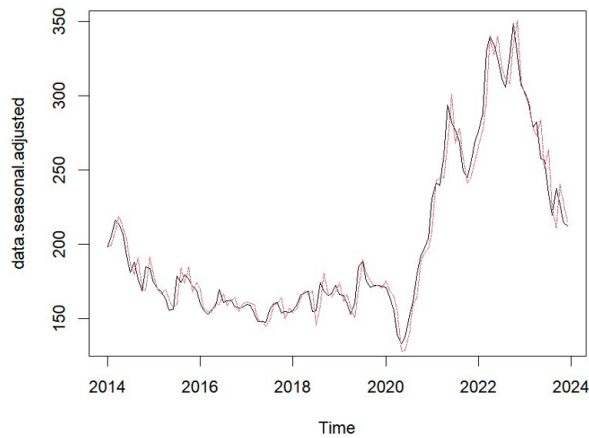


Figure 16: Observed and the forecasted monthly global price of Corn

Source: Own computation based on FRED (January 2014 ~ December 2023) data

The fluctuations in global corn prices over the years reflect the intricate interplay of multifaceted factors, necessitating a comprehensive understanding of the market dynamics. Global demand for corn is influenced by diverse factors, including population growth, dietary preferences, industrial uses, and government policies related to biofuel production and food security. Shifts

in these demand drivers can lead to significant price fluctuations, affecting stakeholders across the corn supply chain, from producers to consumers.

Disruptions in supply chains, whether due to natural disasters, trade conflicts, or logistical challenges, can have profound effects on corn prices. For instance, extreme weather events, such as droughts or floods, can disrupt corn production, leading to reduced yields and increased prices. Similarly, trade tensions between major corn-producing countries can impact market access and trade flows, affecting price dynamics on a global scale.

Geopolitical events and policy decisions also play a crucial role in shaping corn prices. Changes in trade agreements, tariffs, and subsidies can influence market conditions, creating uncertainties for market participants. Moreover, government policies promoting or restricting the use of corn-based ethanol as a renewable fuel source can impact both demand and prices in the corn market.

Disease outbreaks, such as the spread of corn diseases or pests, pose additional challenges to corn production and prices. Crop diseases can devastate yields, leading to supply shortages and price spikes. Furthermore, the emergence of new pests or pathogens can necessitate costly control measures, adding to production costs and potentially driving up prices for consumers.

Sudden climate changes, including shifts in temperature and precipitation patterns, pose significant risks to corn production worldwide. Climate variability and extreme weather events, exacerbated by climate change, can disrupt planting schedules, reduce yields, and increase the likelihood of crop failures. These climate-related risks underscore the importance of adaptive strategies and resilience-building efforts within the agricultural sector to mitigate the impacts of climate change on corn prices and food security.

In recent years, the growing demand for corn-based products, such as ethanol and animal feed, has contributed to increased competition for corn resources. The expansion of biofuel production, driven by renewable energy policies and environmental concerns, has led to a surge in corn consumption for ethanol production. This heightened demand for corn in non-food sectors has implications for food prices and market dynamics, underscoring the need for integrated approaches to balance competing demands for corn resources.

Moreover, speculation in commodity markets can

exacerbate price volatility, amplifying the effects of supply and demand shocks on corn prices. Financial market participants, including investors and hedge funds, engage in trading corn futures and derivatives, seeking to profit from price movements. However, speculative activities can introduce additional uncertainties and distortions into the corn market, leading to heightened price volatility and market inefficiencies.

In light of these complex dynamics, forecasting corn prices becomes imperative for stakeholders across the corn supply chain. Accurate price forecasts enable farmers to make informed decisions regarding planting, input use, and marketing strategies. Similarly, policymakers rely on price forecasts to formulate agricultural policies, manage food security risks, and mitigate market disruptions. Additionally, consumers and food industry stakeholders use price forecasts to anticipate changes in food prices and adjust consumption patterns accordingly.

Given the importance of price forecasting in navigating the uncertainties of the corn market, researchers have developed and applied various time series forecasting models to predict corn prices. These models utilize historical price data, along with relevant explanatory variables, to generate forecasts of future price movements. By evaluating the performance of different forecasting models and refining their methodologies, researchers aim to enhance the accuracy and reliability of corn price forecasts, thereby empowering stakeholders to make informed decisions in an increasingly complex and dynamic market environment.

CONCLUSION

The price of corn has exhibited fluctuations over the years, starting from a relatively low point in early 2014 and steadily increasing, peaking around March of that year. These elevated prices persisted throughout 2014 until mid-March of 2015, after which they began to decline, reaching a nadir in mid-2016. Subsequently, from late 2016 to 2017, global corn prices displayed some volatility but maintained a moderate and stable trajectory compared to the preceding years of 2014 and 2015. From 2018 to early 2019, corn's price gradually increased, reaching a relatively high point by early 2019. During the period from mid-2019 to 2020, there were some fluctuations in price, but overall stability within a moderate range was observed. However, from mid-2020 onwards, a significant surge in corn prices ensued, culminating in a record high by early 2022, followed by a gradual decline to lower levels by early 2024. The application of the Box-Jenkins methodology in time series forecasting effectively identified a seasonal pattern in global corn prices, resulting in an ARIMA (0,1,1)(0,0,1)[12] model. Furthermore, the model pinpointed the influence of a first-order non-seasonal moving average term and a first-order seasonal moving average component on the series. The validation of the model's accuracy was conducted through the analysis of the Ljung-Box Q-test and AIC value. Such forecasting endeavors offer substantial benefits to farmers,

stakeholders, policymakers, and investors alike. These forecasts provide farmers and stakeholders with valuable insights for price adjustments, while policymakers can utilize this information to formulate well-informed marketing strategies. Additionally, investors can leverage these forecasts to make prudent decisions regarding their investment choices. This comprehensive approach to price forecasting enhances decision-making processes across various sectors, facilitating more effective resource allocation and risk management strategies. Therefore, by furnishing insights for price adjustments, informing marketing strategies, and aiding investment decisions, these forecasts contribute to more informed decision-making processes across sectors, ultimately improving resource allocation and risk management strategies.

REFERENCES

- Adams, S. O., Mustapha, B., & Alumbugu, A. I. (2019). Seasonal autoregressive integrated moving average (SARIMA) model for the analysis of frequency of monthly rainfall in Osun state, Nigeria. *Physical Science International Journal*. <https://doi.org/10.9734/psij/2019/v22i430139>
- Asamoah-Boaheng, M. (2014). Using SARIMA to forecast monthly mean surface air temperature in the Ashanti Region of Ghana. *International Journal of Statistics and Applications*, 4(6), 292–298. <https://doi.org/10.5923/j.statistics.20140406.06>
- Avalos, F., & Huang, W. (2022). Commodity markets: shocks and spillovers. *BIS Quarterly Review*, 19, 15–29. https://www.bis.org/publ/qtrpdf/r_qt2209b.pdf
- Baker, A., & Zahniser, S. (2006). Ethanol reshapes the corn market. *Amber Waves: The economics of food, farming, natural resources and rural America*. <https://www.ers.usda.gov/amber-waves/2006/april/ethanol-reshapes-the-corn-market/>
- Baladina, N., Sugiharto, A. N., Anindita, R., & Laili, F. (2021). Price volatility of maize and animal protein commodities in Indonesia during the Covid-19 season. *IOP Conference Series: Earth and Environmental Science*, 803(1), 012060. <https://doi.org/10.1088/1755-1315/803/1/012060>
- Beghin, J. C., & Timalisina, S. (2020). The impact of the COVID-19 Crisis on Nebraska's ethanol industry. *Cornhusker Economics*, 1056. <https://agecon.unl.edu/cornhuskereconomics>
- Brandt, J. A., & Bessler, D. A. (1983). Price forecasting and evaluation: An application in agriculture. *Journal of Forecasting*, 2(3), 237–248. <https://doi.org/10.1002/for.3980020306>
- Chandran, K. P., & Pandey, N. K. (2007). Potato price forecasting using seasonal ARIMA approach. *Potato Journal*, 34(1–2). <https://epubs.icar.org.in/index.php/PotatoJ/article/view/33141>
- Chen, P., Niu, A., Liu, D., Jiang, W., & Ma, B. (2018). Time series forecasting of temperatures using SARIMA: An example from Nanjing. *IOP Conference Series: Materials Science and Engineering*, 394, 052024. <https://>

- doi.org/10.1088/1757-899X/394/5/052024
- Chi, Y. N. (2021). Time series forecasting of global price of soybeans using a hybrid SARIMA and NARNN model: Time series forecasting of global price of soybeans. *Data Science: Journal of Computing and Applied Informatics*, 5(2), 85–101. <https://doi.org/10.32734/jocai.v5.i2-5674>
- Darekar, A., & Reddy, A. A. (2017). Price forecasting of maize in major states. *Maize Journal*, 6(1), 1–5. <https://www.researchgate.net/publication/325755970>
- Divisekara, R. W., Jayasinghe, G., & Kumari, K. (2020). Forecasting the red lentils commodity market price using SARIMA models. *SN Business & Economics*, 1(1), 20. <https://doi.org/10.1007/s43546-020-00020-x>
- Dohlman, E., Hansen, J., & Chambers, W. (2024). USDA Agricultural Projections to 2033 Interagency Agricultural Projections Committee Acknowledgments and Contacts. <https://www.ers.usda.gov/webdocs/outlooks/108567/oce-2024-01.pdf?v=9694.8>
- Elleby, C., Domínguez, I. P., Adenauer, M., & Genovese, G. (2020). Impacts of the COVID-19 pandemic on the global agricultural markets. *Environmental and Resource Economics*, 76(4), 1067–1079. <https://doi.org/10.1007/s10640-020-00473-6>
- Fauziah, L., Anggraeni, L., Latifah, E., Istiqomah, N., & Khamidah, A. (2023). Increased yield and quality of corn by inorganic fertilizers and utilization of corn as food to support food security. *IOP Conference Series: Earth and Environmental Science*, 1253(1), 012129. <https://doi.org/10.1088/1755-1315/1253/1/012129>
- Goryunov, E. L., Drobyshevsky, S. M., Kudrin, A. L., & Trunin, P. V. (2023). Factors of global inflation in 2021–2022. *Russian Journal of Economics*, 9(3), 219–244. <https://doi.org/10.32609/j.ruje.9.111967>
- Liu, S., Liu, D., & Ge, S. (2024). Impact of external shocks on international corn price fluctuations. *Agricultural Economics*, 70(1), 1–11. <https://doi.org/10.17221/318/2023-AGRICECON>
- 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>
- Lv, J., & Wu, X. (2022). Corn price prediction in China's futures market during COVID-19. *Acade J Business Manag*, 4(11), 7–12. <https://doi.org/10.25236/AJBM.2022.041102>
- McGonigle, E. T., Killick, R., & Nunes, M. A. (2022). Modelling time-varying first and second-order structure of time series via wavelets and differencing. *Electronic Journal of Statistics*, 16(2), 4398–4448. <https://doi.org/10.1214/22-EJS2044>
- Montgomery, D. C., Jennings, C. L., & Kulahci, M. (2015). Introduction to time series analysis and forecasting. John Wiley & Sons.
- Mutwiri, R. M. (2019). Forecasting of tomatoes wholesale prices of Nairobi in Kenya: time series analysis using SARIMA model. <https://doi.org/10.11648/j.ijstd.20190503.11>
- Phillips, P. C. B., & Perron, P. (1988). Testing for a unit root in time series regression. *Biometrika*, 75(2), 335–346. <https://doi.org/10.1093/biomet/75.2.335>
- Schmitz, A., Moss, C. B., & Schmitz, T. G. (2020). The economic effects of COVID-19 on the producers of ethanol, corn, gasoline, and oil. *Journal of Agricultural & Food Industrial Organization*, 18(2), 20200025. <https://doi.org/10.1515/jafio-2020-0025>
- Sharma, H., & Burark, S. S. (2015). Bajra price forecasting in Chomu market of Jaipur district: An application of SARIMA model. *Agricultural Situation in India*, 71(11), 7–12.
- UGA Cooperative Extension. (2024, January 22). 2024 Corn, Soybean, and Wheat Outlook. UGA Cooperative Extension. <https://extension.uga.edu/publications/detail.html?number=AP130-2-06>
- Vaswani, A., Prasad, P. C., & Padhi, P. K. (2023). Time series analysis: An application of SARIMA godel in General trade to forecast sales. *Journal of University of Shanghai for Science and Technology*, 25(3). <https://jusst.org/wp-content/uploads/2023/03/Time-Series-Analysis-An-Application-of-SARIMA.pdf>
- Wheals, A. E., Basso, L. C., Alves, D. M. G., & Amorim, H. V. (1999). Fuel ethanol after 25 years. *Trends in Biotechnology*, 17(12), 482–487. [https://doi.org/10.1016/S0167-7799\(99\)01384-0](https://doi.org/10.1016/S0167-7799(99)01384-0)
- Young, W. L. (1977). The Box-Jenkins approach to time series analysis and forecasting: principles and applications. *RAIRO-Operations Research-Recherche Opérationnelle*, 11(2), 129–143. <https://doi.org/10.1051/ro/1977110201291>
- Yu, J.-K., & Moon, Y.-S. (2021). Corn starch: quality and quantity improvement for industrial uses. *Plants*, 11(1), 92. <https://doi.org/10.3390/plants11010092>