Evaluation of Global Corn Price Forecasting Models  
Callum Barr  
*School of Accounting, Finance and Economics, University of Waikato*

# Motivation and research design

Corn is an important agricultural commodity traded worldwide as it is used in a range of industries. In the United States it has three primary uses, 38.7% is allocated to animal feed, 26.6% in ethanol fuel production, and 17.5% is exported. The remainder is used in products for human consumption (World Economic Forum, 2021). As the leading producer, the United States generated 390 million metric tons of corn over the 2023/2024 season, which accounts for 32% of global production. China follows at 24% and Brazil contributes 10% (USDA, 2021). China spends $9 billion annually on corn imports which leads the world and are trailed by Mexico and Japan (Shahbandeh, 2024).

Forecasting corn prices is crucial for risk management among suppliers and consumers. This is especially important for farmers without supply contracts in place, who rely on future market prices due to the 8-month plant-to-harvest cycle of corn. An accurate forecast, 8 months ahead, would assist growers with making informed decisions about how much corn to plant, aid with contract negotiation and determining whether to hold out for better market rates. Buyers on other side of the trade also benefit from insights to price movements, improving their procurement strategies.

Since corn prices are non-stationary, a random walk with drift will be used as the base model to measure performance against. This approach is a common benchmark in economic and financial literature due to its simplicity and ability to capture long-term trends. If more complex models fail to outperform this, they are unnecessarily complex and do not add value. Previous studies have shown that seasonal ARIMA models worked well in corn price applications. For instance, a SARIMA (0,1,1) (0,0,1) [12] was identified as the best fit in a study on global corn prices from January 2014 to December 2023 (Raksha & Yeong, 2024). Similarly, another study found that a SARIMA (2,1,0) (3,1,1) [12] applied to China’s futures market can accurately predict settlement price (Junxue & Xi, 2022). Therefore, both ARIMA and SARIMA models will be evaluated in this analysis, expanding on it to identify if a model consistently performs best across various datasets, rather than relying on a single sample trail. An ARIMA uses past prices and moving averages (autoregressive patterns) to predict future prices. This typically performs well in short-term forecasts, as past shocks are incorporated into predictions. SARIMA, an extension to ARIMA, also accounts for seasonal patterns, making it particularly suitable for corn, which is grown and harvested annually. The seasonality may result in patterns of over or under supply patterns throughout the year which would affect the price.

Explanatory variables related to corn production and consumption were added to an ARIMA in an attempt to improve its performance. These variables help identify key drivers behind price movements which provides useful insights to market participants as it will reveal which factors focus on and what to expect under conditions such as elevated geopolitical risk. Corn price is known to experience volatility clustering, periods of heighten or diminished volatility often triggered by geopolitical or weather events. These periods are best modelled using GARCH models, which account for volatility varying over time. To incorporate this an ARIMA model was used as the mean model, with GARCH applied to factor in past volatility and assessed how it fared against the other models.

This study aims to identify the most accurate model for predicting monthly global corn prices (FRED, 2024) eight months in advance by comparing the Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) of different models. The comparison will be conducted across four different samples of training and test data from the period 1993 to 2024. The RMSE and MAE will be averaged to determine which model consistently performs the best across the four tests. By identifying the best-performing model, this study aims to contribute valuable insights into price risk management, benefiting both corn producers and consumers in navigating market conditions.

# Data and econometric methods

## 2.1. Data

The monthly global corn price, sourced from Federal Reserve Economic Data (FRED) and covering the period 1993 to 2024 (FRED, 2024), was tested for stationarity using the Augmented Dickey-Fuller (ADF) test. The test confirmed corn price was non-stationary, and was subsequentially differenced using the auto.arima function in R when ARIMA and SARIMA models were used. To improve the improve forecasting performance additional explanatory variables were added to an ARIMA. These were tested for stationarity, with only nitrogen price found to be non-stationary. It was transformed by calculating its monthly return in nitrogen price.

The first variable considered was geopolitical risk, which constructs a measure of adverse geopolitical events based on a tally of newspaper articles covering geopolitical tensions (Caldara & Iacoviello, 2024). Higher geopolitical risk is known to foreshadow lower investment and down side risk to the economy (Caldara & Iacoviello, 2024). This consideration is relevant in light of recent events like the war in Ukraine, a majority exporter of corn, and how such conflicts affect global corn prices. Additionally, trade wars may impact corn price through reduced imports or exports between countries.

The second variable examined was measure of China’s GDP, the country who spends the most on corn imports. While GDP data was only available quarterly, China’s total export-to-import ratio, sourced from the OECD was used as a substitute (OECD, 2024). This measure helps capture China’s demand for corn; for example, when they generally import more does global corn price increase.

Fertilizer prices were also included as they represent a significant input cost in growing corn. Nitrogen is the highest rate nutrient for corn production so by incorporating its prices we can access how rising costs impacts growers’ usage (Bayer, 2022). If farmers use less nitrogen due to rising prices, it may lead to a reduction in area of corn grown or lower nitrogen application rates, ultimately reducing supply and driving up corn prices.

Weather is a crucial factor affecting corn growth, as both excessive and insufficient rainfall can include yields and supply. To account for this, average monthly rainfall and temperature data were gathered from counties (Des Moines and Normal) in the two top corn producing states in US, Iowa and Illinois. The data was transformed into deviations from the monthly average to identify the impact of abnormally large or small amounts of rainfall or extreme temperatures on corn prices (NWS, 2024). Missing values were replaced with the historical average for that month.

Aside from the weather variables, all variables had four lags to capture any delayed effects that may arise.

## 2.2. Econometric methods

The random walk with drift was used as the base model, implemented using the rwf function in R. It is expressed as shown in equation 1.

[1]

Where CP is corn price and µ is the drift term.

The optimal ARIMA model was identified using the auto.arima function in R. It tests different p and q values through a variation of the Hyndman- Khandakar algorithm until the minimum AICc is achieved (Hyndman & Athanasopoulos, 2024). The ARIMA model is represented by equation 2.

ARIMA (p, d, q):

[2]

The Seasonal ARIMA (SARIMA) determined similarly but also includes an ARIMA component from the previous season as shown in equation 3.

SARIMA (p, d, q) (P, D, Q) [m]:

[3]

After fitting an ARIMA model we check the autocorrelation function (ACF) plot to ensure residuals of the lags are within the thresholds, particularly in the early lags. Additionally, a Ljung-Box Test is completed with the following hypotheses:

* H0: Residuals have no autocorrelation.
* HA: Residuals have autocorrelation.

The test ensures the model is valid, if autocorrelation is present, the model is not capturing all underlying patterns which may affect forecast accuracy.

To find the optimal ARIMAX model (Equation 4), all explanatory variables were included. Then the backwards elimination method was applied until the model only contained statistically significant coefficients (95% confidence interval) and had lower AIC, RMSE and MAE than the ARIMA model. The aim was to keep the number of explanatory variables as low as possible since these variables will also need to be forecast, which could introduce additional errors into corn price forecast.

ARIMAX (p, d, q):

[4]

The GARCH model is made of two equations: a level and variance equation. In this study the mean model was an ARIMA, found using auto.arima. The variance equation was a GARCH(1,1) expressed in equation 5.

GARCH(1,1):

[5]

To assess the model’s effectiveness two tests were completed. Test 1 checks for auto correlation in the residuals, where:

* H0: Residuals have no autocorrelation.
* HA: Residuals have autocorrelation.

Test 2 is similar but checks the squared residuals:

* H0: Squared residuals have no autocorrelation.
* HA: Squared residuals have autocorrelation.

Models were evaluated using RMSE and MAE, with AIC also considered in the ARIMA models. RMSE squares the errors before averaging and taking the square root. This means outliers are more heavily weighted resulting in the error being more sensitive to them. MAE is the average of the absolute value of the errors. In this method all errors are equally weighted, both were included as they treat errors slightly differently. AIC is a goodness-of-fit measure used to compare a set of statistical models, revealing the best model while considering its fit and complexity (Activeloop, 2024). Minimising all three measures unveils the best model.

Four tests were conducted within the period of 1993 to 2024. R was used to generate four random dates within this period, marking the forecast start dates. The first date was 01/09/2003, so the model was fitted to the training data from 01/01/1993 to 01/09/2003, and the out of sample test was conducted over the following 8 months. The other dates were 01/01/2014, 01/10/2021 and 01/08/2023, resulting in over lapping test periods. The average error was taken across the four out of sample tests to determine which model performs best across the entire data set.

Thresholds were established for the 10% lowest and highest monthly returns of corn price. Any monthly return greater than the upper threshold indicates that the volatility is in the highest 10% of the data, while returns below the lower threshold indicate the lowest 10% volatility (largest decreases in price change). A dummy variable was created, receiving a value of 1 if it exceeded either threshold, with high volatility periods shaded on the forecast plots.

# Results and interpretation

## 3.1. Forecast 1 summary

The random walk model performed relatively poorly, with an in-sample RMSE of $6.60 and an out-of-sample RMSE of $18.63. The ARIMA(2,1,3) model provided improvement with an in-sample RMSE of $5.61 and passing the Ljung-Box test (p-value 0.83), showing no autocorrelation in residuals. Despite higher out-of-sample errors, the ARIMA still outperformed the random walk. The seasonal ARIMA found no seasonal effects to improve upon the ARIMA. Introducing China's export-to-import ratio through an ARIMAX(0,1,1) model further lowered in-sample errors but struggled out-of-sample. Finally, the ARIMA(0,0,1)-GARCH(1,1) model had the lowest out-of-sample RMSE ($17.76), showing strong predictive ability of price volatility, with past volatility having a lasting impact (squared residuals coefficient of 0.998). All forecast models under predicted the corn price (Figure 1).

Table . Forecast 1 results.



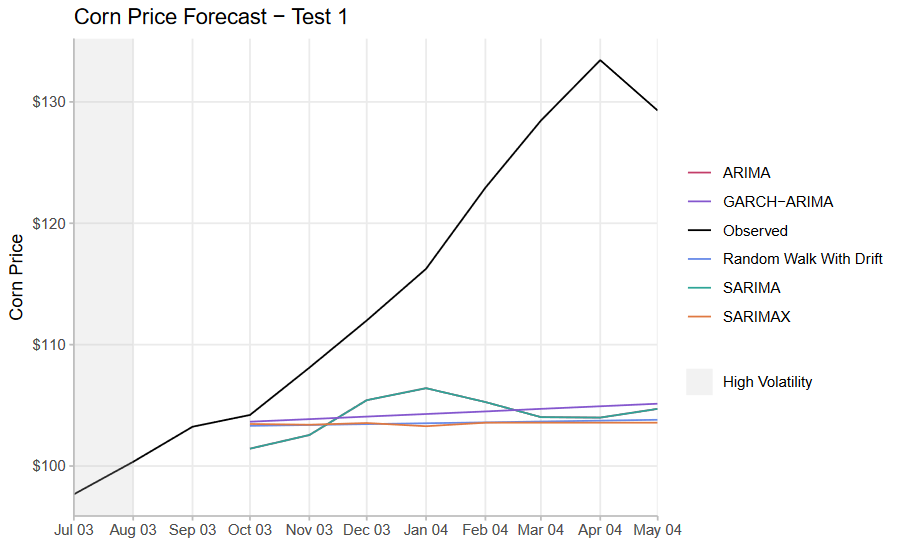


Figure . Test 1, corn price forecast from October 2003 to May 2004.

## 3.2. Forecast 2 summary

The random walk with drift had a in sample RMSE of $11.33 which increased to $22.12 out of sample. The ARIMA(2,1,3) improved on this baseline in both tests, however it failed the Ljung-Box test (p-value 0.025) indicating autocorrelation was present in the residuals. The ACF plot reveals it exceeds the thresholds at lags 14, 16 and 23. The SARIMA(1,1,0)(0,0,2)[12] further improved on the accuracy seen in the other two models. Out of sample it had an RMSE of $19, the lowest among the models. With a p-value of 0.195, it passed the Ljung-Box test, indicating a good model fit. Adding nitrogen price return and two lags were statistically significant when applied in a ARIMA(3,1,2). The residuals exhibited no autocorrelation (p-value 0.115), and it produced the lowest in sample error with and RMSE of $10.5. This accuracy improvement did not translate to the out of sample test. The ARIMA(3,0,2)-GARCH(1,1) did not maintain its performance from the first test, it performed the worst of all the models, with an out-of-sample RMSE of $24.46. The squared residuals coefficient did remain high, at 0.965. The models all had periods of under and over prediction (Figure 2).

Table . Forecast 2 results.



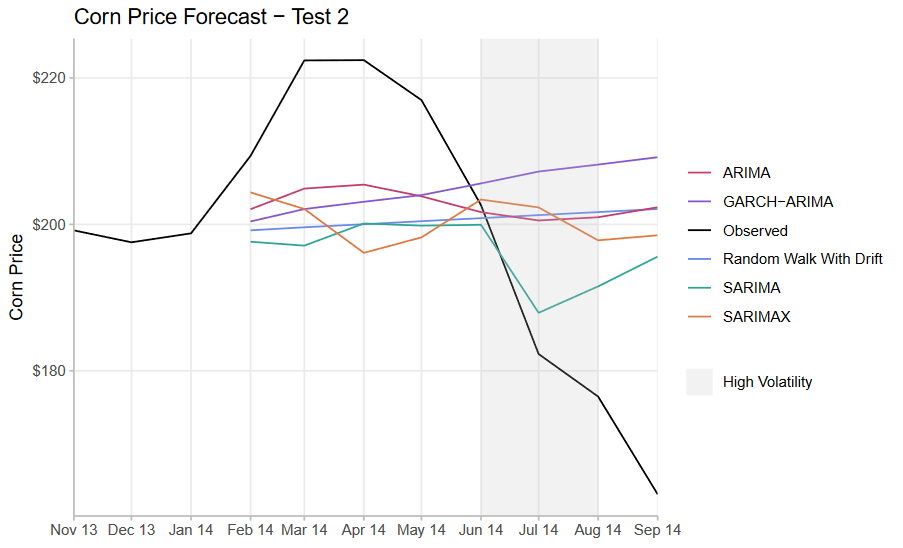


Figure . Test 2, corn price forecast from February 2014 to September 2014.

## 3.3. Forecast 3 summary

In the third test all models struggled with accuracy. Over the training period the random walk performed similarly to before with a RMSE of $11.14, however out of sample this increased to $74.04. Using an ARIMA (2,1,3) improved the in-sample error but had an even larger out of sample error of $75.27. This discrepancy was not due to autocorrelation in residuals as it passed the Ljung-Box test. The SARIMA did not fare any better. The SARIMA(1,1,0)(1,0,2)[12] had the same in sample error as the ARIMA but had the largest out of sample RMSE at $75.97. Auto correlation was detected in its residuals (p-value 0.0396) and thus failed the Ljung-Box test. No statistically significant variables were found to apply in an ARIMAX. However, the ARIMA(3,0,2)-GARCH(1,1) out performed all models with and out of sample RMSE of $71.60. Its performance possibly being aided by capturing the two highly volatile periods prior to the forecast (Figure 3). The poor performance across all the models is likely due to a trend reversal in corn price. It had been trending downward before increasing just prior to the forecast date, resulting in all models underpredicting the corn price (Figure 3).

Table . Forecast 3 results.



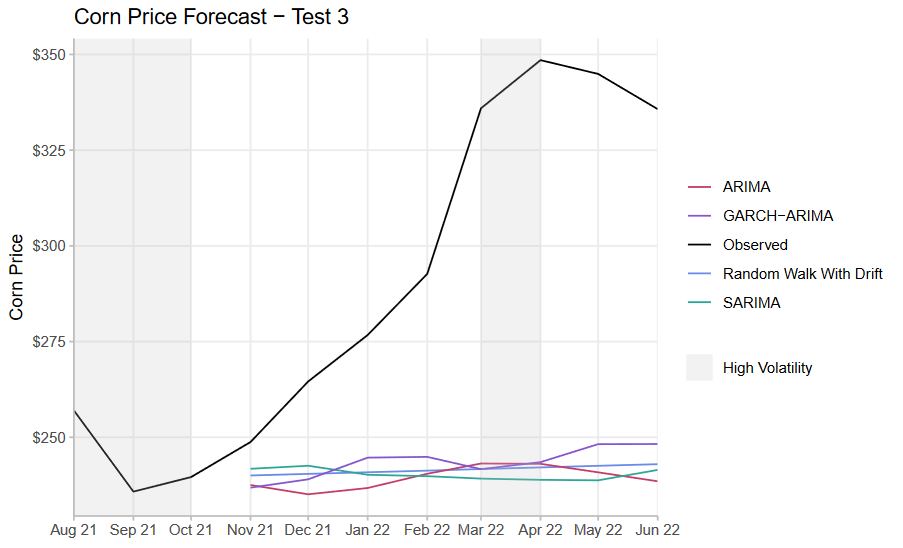


Figure . Test 3, corn price forecast from November 2021 to June 2022.

## 3.4. Forecast 4 summary

The results in the fourth test align more closely to outcomes observed in the first two tests. The random walk model performed slightly worse out of sample with RMSE of $14.70 compared to $11.82 in sample. Applying the ARIMA(1,1,0) improved performance with an out of sample RMSE of $13.89, with no autocorrelation in the residuals. The SARIMA also had no correlation in its residuals when modelled by ARIMA(1,1,0)(2,0,1)[12]. This model had the lowest error in both tests, with an RMSE of $12.76 out of sample. Applying a ARIMA(3,0,2)-GARCH(1,1) produced the second-lowest out-of-sample error of $13.46. It effectively captured the volatility passing both autocorrelation tests and maintained a high squared residuals coefficient of 0.95. No statistically significant variables were found to apply in ARIMAX. The models initially under precited the price and then began over predicting in the latter periods (Figure 4).

Table . Forecast 4 results.



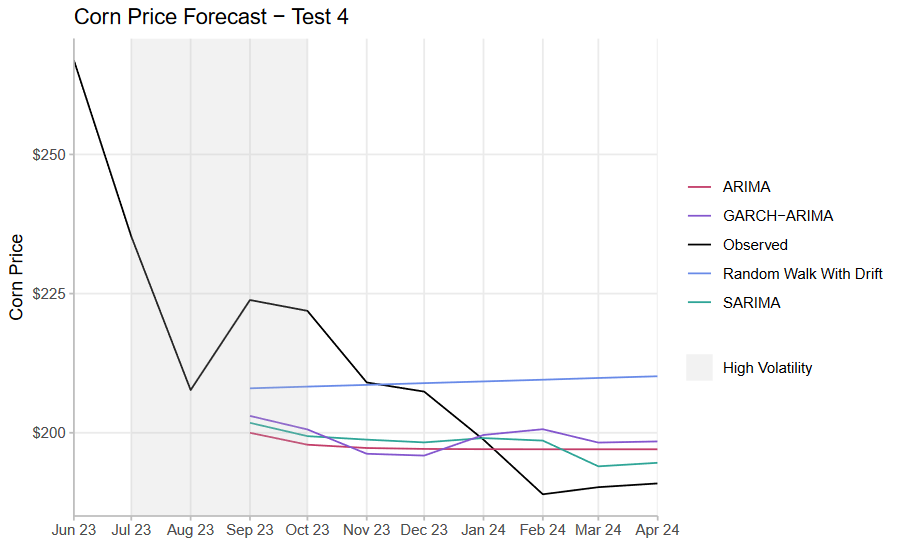


Figure . Test 4, corn price forecast from September 2023 to April 2024.

## 3.5. Summary of forecasts

Averaging the errors across the four tests reveals that the random walk with drift has the highest forecast error over an 8-month period, with an average out of sample RMSE of $32.37. It was closely followed by ARIMAX at $32.27. In tests three and four, where no explanatory variables were identified for the ARIMAX, the error of the ARIMA model was used, thereby lowering its average error which in reality should be worse. This had to be done otherwise the outlier test error would not have been in its average. The best performing model was the SARIMA which achieved an average out of sample RMSE of $31.41. The ARIMA-GARCH was marginally better than the ARIMA and thus the second-best model. It had and RMSE of $31.81 compared to $31.83 for the ARIMA model (Table 5).

Table . Average model errors.



All models experienced outlier errors in the third forecast. This anomaly can be attributed to a trend reversal in corn prices just before the forecast period. In the months leading up to the forecast, corn prices exhibited a sharp downward trend, followed by a rapid increase just prior to the forecast. Given that ARIMA models rely on autoregressive patterns, they are expected to perform poorly, as evidenced by the largest errors observed in the test. The outlier errors on this test skewed overall averages, with forecast errors in the range of $65 - $75, compared to the other tests where ranges of $10 – $22 were seen. Coefficients on the squared residuals in GARCH models were high on all forecasts, which indicates that past volatility has a large impact on current volatility and takes a long time to return to normal levels. This suggests the presence of volatility clustering.

# Conclusion

The results indicate that, on average, a seasonal ARIMA is the best model for forecasting corn prices, followed by ARIMA-GARCH and ARIMA models. However, none of the models demonstrated exceptional performance with the best average RMSE for an 8-month forecast being $31.48. At this stage the models offer limited practical value for stakeholders given their performance, though there is potential for improvement through analysis of other models and variables.

No variables were found that consistently had a relationship with corn prices, but adding them showed potential. During in sample tests when using statically significant variables in an ARIMAX they achieved the lowest in sample error. However, there effectiveness was restricted as they performed poorly out of sample, partly due to compounding errors. To forecast of corn prices, the explanatory variables themselves need to be forecast, meaning that the RMSE of the corn price forecast is influenced by the RMSE of the variable forecast.

For future research, it is recommended to prioritise identifying relationships with lagged variables. This approach reduces the reliance on forecasting of the explanatory variables, potentially yielding better results. Additionally, focusing on macro-scale variables may enhance the likelihood of finding statistically significant relationships. Companies such as LSEG offer agricultural commodities data packages that contain high quality and relevant data, though a subscription is required (LSEG, 2024). These packages include weather and crop forecasting data which result in better models as they capture a larger group than in this study. Given corns many uses it is a good candidate for a ridge machine learning model. This method could effectively handle a large number of correlated variables, which may lead to more accurate forecasts.

# References

Activeloop. (2024, 10 16). *Akaike Information Criterion (AIC)*. Retrieved from Activeloop: https://www.activeloop.ai/resources/glossary/akaike-information-criterion-aic/

Bayer. (2022, 03 08). *Balancing Fertilizer Needs for the Corn Crop*. Retrieved from Bayer: https://www.cropscience.bayer.us/articles/bayer/balancing-fertilizer-needs-corn

Caldara, D., & Iacoviello, M. (2024, 10 16). *Overview and Paper*. Retrieved from matteoiacoviello: https://www.matteoiacoviello.com/gpr.htm

FRED. (2024, 10 16). *Global price of Corn*. Retrieved from FRED: https://fred.stlouisfed.org/series/PMAIZMTUSDM

Hyndman, R. J., & Athanasopoulos, G. (2024, 10 19). *ARIMA modelling in R*. Retrieved from Forecasting: Principles and Practice: https://otexts.com/fpp2/arima-r.html

Junxue, L., & Xi, W. (2022). Corn Price Prediction in China's Futures Market. *Academic Journal of Business & Management*.

LSEG. (2024, 10 21). *LSEG Data & Analytics*. Retrieved from Commodities - Agriculture: https://www.lseg.com/en/data-analytics/financial-data/commodities-data/agricultural-commodities-data#overview

NWS. (2024). *National Weather Service*. Retrieved from National Weather Service: https://www.weather.gov/wrh/Climate?wfo=dmx

OECD. (2024). *International Trade: Ratio: Exports to Imports: Total for China*. Retrieved from OECD: https://fred.stlouisfed.org/series/XTEITT01CNM156N

Raksha, K., & Yeong, N. C. (2024). Forecasting the Global Price of Corn: Unveiling Insights with SARIMA Modelling. *American Journal of Applied*, 124-135.

Shahbandeh, M. (2024, 8 20). *Import value of corn worldwide in 2023, by leading country*. Retrieved from Statista: https://www.statista.com/statistics/1167839/leading-global-corn-importers/

USDA. (2021, 10 16). *Production - Corn*. Retrieved from USDA: https://fas.usda.gov/data/production/commodity/0440000

World Economic Forum. (2021, 6 2). *From feed to fuel: This is how corn is used around the world*. Retrieved from World Economic Forum: https://www.weforum.org/agenda/2021/06/corn-industries-sustainability-food-prices/

# Appendix

R code and data set were submitted to Moodle.