What to Expect with Cocoa Futures in Ghana?

Forecasting Cocoa Price: A Time Series Approach to Commodity Price Modeling

Emily Pan, Manjun Zhu, Edan Wong

April 4, 2025

1 Introduction

Cocoa is a major agricultural commodity in Ghana. The industry employs over 850,000 farmers and generates $2 billion in foreign exchange annually (Board 2024). Cocoa futures are traded primarily on the Intercontinental Exchange (ICE), which plays a significant role in global pricing, market expectations, and risk management. Although Ghana does not have a domestic cocoa futures market, its daily cocoa prices, export contracts, and farmgate prices are affected by international cocoa futures benchmarks (Nunley, Locopo, and Morganteen 2024). Daily cocoa prices show the current market value for immediate delivery, while futures prices indicate expected future values based on supply, demand, and market predictions influenced by macroeconomic factors and weather conditions.

Recent developments in the global cocoa market have highlighted the volatility and sensitivity of cocoa prices to both environmental shocks and market speculation. From 1994 to 2023, cocoa prices showed long-term growth, rising from approximately US$1,400 per ton to around US$4,000 per ton (Organization, n.d.). However, in 2024, cocoa prices experienced unprecedented fluctuations—peaking at US$12,072 per ton in February before sharply declining by September (Organization (n.d.). These fluctuations have been driven by supply constraints, rising demand, and adverse production conditions linked to climate change, aging trees, pests, and disease.

In response, the Ghana Cocoa Board raised the producer price to GH¢3,100 per 64kg bag to help farmers benefit from high international prices and cushion the effects of market volatility (Board 2024). The ongoing threats related to climate change highlight the urgent need for comprehensive and long-term strategies aimed at improving forecasting, stabilizing prices, and fostering sustainable practices in cocoa agriculture.

This study aims to improve forecasting accuracy for agricultural commodity prices, particularly cocoa, using time series models such as ARIMA, ARIMAX, GARCH, and VARMA, alongside machine learning approaches like XG-Boost. By integrating exogenous variables and modelling price levels and volatility, we seek to offer a comprehensive framework for understanding and anticipating market dynamics.

2 Literature Review

The ARIMA model is a foundational tool in time series forecasting but often struggles in volatile markets. For instance, Wang, Yue, Wei, and Lv (Wang et al. 2017) observed that ARIMA consistently underperformed in forecasting futures prices for wheat, corn, soybean, and sugar due to its assumptions of linearity and constant variance, resulting in the highest forecasting errors among tested models.

To better capture volatility, Katsiampa (Katsiampa 2017) applied GARCH-type models to Bitcoin and found the AR-CGARCH model most effective in modeling conditional variance by incorporating both short- and long-term volatility components. While the study focused on cryptocurrency, the findings are applicable to cocoa markets, which similarly exhibit high volatility and are influenced by speculative behavior and external shocks.

For multivariate relationships, models such as Vector Autoregression (VAR) and Vector Autoregressive Moving Average (VARMA) are used to assess the joint impact of exogenous factors like exchange rates and weather. Mayr and Ulbricht (Mayr and Ulbricht 2015), however, cautioned that log-transforming data in VAR frameworks does not consistently improve forecast accuracy and recommended testing for stationarity and variance instability before transformation—an important consideration when modeling cocoa prices.

Recent literature has increasingly turned to machine learning and deep learning for more robust forecasting. Menéndez-García, García-Nieto, García-Gonzalo, and Sánchez Lasheras (Menéndez-García et al. 2024) compared ARIMA and VARMA with algorithms such as Support Vector Machines (SVM), Multivariate Adaptive Regression Splines (MARS), and Multilayer Perceptrons (MLPs) for forecasting platinum prices. MLP, an artificial neural network (ANN), achieved the highest accuracy and lowest RMSE, while ARIMA again performed the worst, reinforcing its difficulty in capturing non-linear and noisy patterns.

Further supporting the shift toward advanced techniques, Nayak, Alam, Singh, and Sinha (Nayak et al. 2024) evaluated agricultural price forecasting models for tomatoes, onions, and potatoes in India. Their research found that deep learning models like TransformerX and NBEATSX, especially when incorporating exogenous variables such as weather, consistently outperformed traditional models including ARIMAX, multiple linear regression, and several ML algorithms (Random Forest, XGBoost, ANN, etc.) across various accuracy metrics.

Given the limitations of traditional models in volatile agricultural markets, hybrid approaches have gained traction. These approaches combine volatility modeling techniques such as GARCH with the flexibility and pattern-recognition strengths of ML and DL methods like XGBoost. As researchers increasingly address challenges like non-stationarity, seasonality, and noisy data, these integrated models provide a more robust framework for forecasting prices in markets like cocoa, where production uncertainty and speculative forces remain prominent.

3 Methodology

This section details the methodologies applied in forecasting cocoa prices using five different modeling approaches: ARIMA, ARIMAX, GARCH, VARMA and XGBoost. These models were chosen because of their capability in capturing both statistical and structural characteristics of the cocoa price series, including autocorrelation, volatility, multivariate influence and nonlinear patterns. For each fitted model, its performance will be evaluated through out-of-sample forecasts using metrics such as Mean Absolute Error (MAE), Root Mean Squared Error (RMSE) and Mean Absolute Percentage Error (MAPE).

Initially, the analysis was conducted using weekly data to closely match the frequency of international cocoa price fluctuations. However, this approach presented several limitations. A significant proportion of external datasets, including key variables like exchange rates, cocoa production figures, and macroeconomic indicators, were only available at a monthly or lower frequency.

The ARIMA model serves as a baseline univariate model and it will be first applied to forecasting the weekly cocoa price series after testing for stationarity using the ADF test because the original cocoa daily price excluded the weekends' data and . Where necessary, differencing or box cox transformation will be used to transform the series to stationary. The optimal parameters (p, d, q) were identified by examining the ACF and PACF of the transformed price series and AIC and BIC will be used to select the best ARIMA model after performing the residual diagnostic for each ARIMA models, where the diagnostic confirms the absence of autocorrelation in residuals.

3.1 ARIMA Models

The ARIMA model is used to create a baseline forecast of weekly cocoa prices. It relies only on past prices and patterns in the data, without using any external factors. This helps us understand how well the price can be predicted using just its past values.

$$Y_t = \phi_1 Y_{t-1} + \cdots + \phi_p Y_{t-p} + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q} + \varepsilon_t$$

where:

- Y_t is the log-transformed weekly cocoa price at time t,
- ε_t is a white noise error term,
- p, d, and q are the orders of autoregression, differencing, and moving average, respectively.

To incorporate environmental drivers of cocoa price, we extended the univariate ARIMA model into an ARIMAX framework by including exogenous variables such as weekly precipitation and weekly average temperature, which helps to test whether including these weather factors improves the accuracy of cocoa price forecasts compared to using price history alone. This approach allowed us to assess the linear effect of relevant predictors while maintaining the ARIMA structure to model autocorrelation in the residuals.

3.2 ARIMAX Model

The ARIMAX model builds on ARIMA by adding climate variables like precipitation and temperature. It helps us test whether including these weather factors improves the accuracy of cocoa price forecasts compared to using price history alone.

To incorporate climate-related external influences, the ARIMA model is extended to an ARIMAX(p, d, q) form:

$$Y_t = \phi_1 Y_{t-1} + \cdots + \phi_p Y_{t-p} + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q} + \beta_1 X_{1t} + \cdots + \beta_k X_{kt} + \varepsilon_t$$

where:

- $X_{1t}, \ldots, X_{kt}$ are the exogenous climate variables at time t (e.g., precipitation, average temperature),
- $\beta_1, \ldots, \beta_k$ are their corresponding coefficients.

In parallel, a separate GARCH model will be fitted to model time-varying volatility in weekly cocoa prices, which is a prominent feature in commodity markets. The GARCH model is capable of accounting time-varying conditional variance, making it well-suited for capturing market dynamics characterized by volatility clustering.

3.3 GARCH Model

The GARCH(1,1) model is used to capture volatility in weekly cocoa price returns. It helps estimate how uncertainty changes over time, offering insight into risk and price fluctuations beyond average trends.

$$\begin{aligned} r_t &= \mu + \varepsilon_t, \\ \varepsilon_t &= \sigma_t z_t, \\ \sigma_t^2 &= \omega + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \end{aligned}$$

where:

- r_t is the log return of weekly cocoa price at time t,
- σ_t^2 is the time-varying conditional variance (volatility),
- z_t is an i.i.d. innovation term, assumed to follow a Student's t-distribution,
- $\omega > 0$, $\alpha_1 \geq 0$, and $\beta_1 \geq 0$ are parameters to be estimated.

Model orders for ARIMA and ARIMAX were selected using AIC, BIC, and ACF/PACF diagnostics to balance fit and complexity. GARCH(1,1) was chosen for its simplicity and effectiveness in capturing volatility clustering, with diagnostics confirming model adequacy.

3.4 VARMA Model

$$\begin{aligned}\text{PriceDiff}_t = \beta_0 + \beta_1\text{PriceDiff}_{t-1} + \beta_2\text{ExRateDiff}_{t-1} + \beta_3\text{MonthlyAvgTemp}_{t-1} + \\ \beta_4\epsilon_{\text{PriceDiff},t-1} + \beta_5\epsilon_{\text{ExRateDiff},t-1} + \beta_6\epsilon_{\text{MonthlyAvgTemp},t-1} + \\ \beta_7\epsilon_{\text{PriceDiff},t-2} + \beta_8\epsilon_{\text{ExRateDiff},t-2} + \beta_9\epsilon_{\text{MonthlyAvgTemp},t-2} + \epsilon_t\end{aligned}$$

Where:

- `p = 1` means the model uses the lagged values from the previous time step (t-1) for PriceDiff, ExRateDiff, and MonthlyAvgTemp (AR part)
- `q = 2`: means the model uses lagged residuals (errors) from the previous two time steps (t-1 and t-2) for PriceDiff, ExRateDiff, and MonthlyAvgTemp (MA part).
- $\beta_1, \beta_2, ...$ are the coefficients for the lagged variables and error terms ϵ_t represents the residual (error) term at time t.

For modeling multivariate interdependencies, the VARMA model will be used to fit the monthly cocoa price, including the exchange rate as exogenous variables and changing the weekly precipitation and temperature into monthly frequency. The cocoa price series will be analyzed alongside relevant exogenous variables with each tested for stationarity using the ADF test, followed by differencing where necessary. The optimal lag structure for the VARMA model was identified using the VARselect function from the "var" package, where various VARMA models with the selected lags will be fitted and the final model will be determined by having the lowest AIC. The fitted VARMA model allowed for joint modeling of endogenous and exogenous variables, capturing the dynamic interrelationship among them.

3.5 XGBoost

To address nonlinearities and improve model robustness, XGBoost was also applied as a machine learning forecasting approach and it is selected for its ability to manage both linear and nonlinear relationships and its flexibility in dealing with non-stationary data without requiring transformation assumptions. Feature selection was conducted based on relative gain values, with variables above a 0.1 threshold retained to enhance interpretability and predictive accuracy. The XGBoost model will forecast the monthly cocoa price difference as a sum of multiple decision trees:

$$\hat{y}_t = f_1(x_t) + f_2(x_t) + \cdots + f_{500}(x_t)$$

where $\hat{y}_t$ is the predicted **cocoa price difference** at time t, x_t represents the **input features** and $f_m(x_t)$ is the **m-th decision tree** trained by XGBoost with total number of 500 boosting iterations.

4 Data

4.1 Weekly Data Description for ARIMA, ARIMAX, and GARCH Analysis

The data set used in this analysis consists of weekly observations on international cocoa prices and climate indicators from Ghana.

- **Cocoa Price Data**: Daily price series from the International Cocoa Organization (ICCO), provided in USD per metric tonne. This data set was aggregated to a weekly frequency by selecting the last available daily price for each week.
- **Ghana Climate Data**: Daily weather data for multiple Ghanaian weather stations obtained from the National Centers for Environmental Information (NCEI). Variables include: Total daily precipitation (PRCP), Average daily temperature (TAVG), Maximum temperature (TMAX), and Minimum temperature (TMIN)

These were cleaned, averaged across stations, and aggregated to weekly summaries (sum for precipitation, average for temperatures). Missing values were forward and backward filled using Last Observation Carried Forward (LOCF).

The final weekly data set spans from **January 2000 to December 2024**, containing the following variables:

Table 1: Summary Statistics for Weekly Cocoa Price and Climate Variables

Variable	Min	1st Qu.	Median	Mean	3rd Qu.	Max
Cocoa Price (USD/tonne)	774.1	1557.99	2192.96	2290.50	2715.88	10966.97
Total Weekly Precipitation (mm)	0.0	0.00	0.15	0.44	0.62	6.06
Average Weekly Temperature (°C)	74.0	78.99	81.32	81.05	83.00	88.14
Max Weekly Temperature (°C)	79.0	86.00	89.47	88.83	91.47	99.00
Min Weekly Temperature (°C)	66.6	72.85	74.00	74.09	75.34	80.29

4.2 Monthly Data Description for VARMA and XGBoost Analysis

The monthly dataset combines cocoa prices, climate indicators, exchange rates, and production volumes to support multivariate modeling of Ghana's cocoa sector.

- **Cocoa Price Data**: Sourced from the International Cocoa Organization (ICCO), daily prices in USD per tonne were aggregated to monthly frequency using the monthly average of daily values.
- **Ghana Climate Data**: Ghana Climate Data: Daily weather observations (PRCP, TAVG, TMAX, TMIN) from multiple stations were aggregated using the monthly mean across all available station values.
- **Exchange Rate Data**: Monthly average USD to GHS exchange rates were sourced from Investing.com ("USD/GHS - US Dollar Ghanaian Cedi Exchange Rate" 2024). These rates reflect the macroeconomic environment relevant to Ghana's cocoa exports.
- **Cocoa Production Data**: Annual cocoa production volumes for Ghana were obtained from Our World in Data (Data 2024). Production is reported by crop year (October–September), and annual totals were evenly distributed across months. This method is a practical simplification due to the lack of higher-frequency production data.

The final monthly dataset spans from **March 1995 to November 2024**, and includes variables for cocoa prices, climate conditions, exchange rates, and estimated production.

Table 2: Summary Statistics for Monthly Cocoa Price, Climate, Production, and Exchange Rate Variables

Variable	Min	1st Qu.	Median	Mean	3rd Qu.	Max
Cocoa Price (USD/tonne)	800.9000	1.5679e+03	2.2053e+03	2.29550e+03	2684.8000	9.876600e+03
Monthly Average Temperature (°C)	76.2700	7.8930e+01	8.1420e+01	8.10900e+01	82.8300	8.690000e+01
Monthly Average Precipitation (mm)	0.0000	1.5700e-02	5.7270e-02	7.43100e-02	0.1084	4.569400e-01
Cocoa Production (Tonnes)	288075.0000	4.3660e+05	7.0002e+05	6.62729e+05	835466.0000	1.047385e+06
Exchange Rate (USD/GHS)	0.0995	8.0470e-01	1.4318e+00	2.95170e+00	4.3690	1.625000e+01

We now consider the VARMA model due to the volatility of cocoa prices, while still incorporating parameters from the ARMA model. To ensure the variables are strong significant (p-value < 0.05), we first apply set difference on the variables that are not statistically significant (p-value > 0.05), such as `cocoa_price`, `price_lag`, and `currency_exchange_rate`. With these significant variables identified, we proceed to fit the VARMA model by selecting an appropriate lag order.

Table 3: Lag Order Selection

Criterion	Selected.Lag
AIC(n)	4
HQ(n)	2
SC(n)	1
FPE(n)	4

From the results, we will consider lag=1, 2, 4. We will use model selection to decide which one is better.

Table 4: AIC for Different Lag Orders

Model	AIC
VARMA(p=1, q=2)	10677.67
VARMA(p=2, q=2)	10713.37
VARMA(p=4, q=2)	10683.50

We built three VARMA models with different lag orders (p = 1, 2, 4, q = 2) to forecast price fluctuations. The difference between these models lies in the number of past observations (lags) they consider for predicting future values. To compare them, we used the AIC, which balances model fit and complexity. The results show that VARMA(p=1, q=2) has the lowest AIC, making it the best model among the three. We will continue the model check with VARMA(1, 2) model.

In XGBoost Model, the initial model will incorporate the following predictor variables (features):

- Monthly_Avg_Temp (Monthly Average Temperature)
- Monthly_Avg_Precip (Monthly Average Precipitation)
- Exchange_Rate_t (Exchange rate)
- $\text{Production\_in\_Ton}_t$ (Cocoa Production in tons)
- Price_diff_t (Price difference)
- Price_diff_{t-1} (Price difference with 1 lag)
- Price_diff_{t-2} (Price difference with 2 lag)
- Price_diff_{t-3} (Price difference with 3 lag)

The inclusion of 1-month lagged exchange rate follows Frankel and Hardouvelis' (Frankel and Hardouvelis 1985) observation supports that modeling exchange rate effects with 1-month delay account for market absorption time for currency fluctuations and reducing noise compared to immediate-rate models. Furthermore, three lagged price differences is grounded in both empirical evidence and commodity market dynamics, particular in agricultural commodities like cocoa, where short-term price momentum and cyclical patterns are exhibited frequently due to supply chain delays coming from production adjustments and inventory restocking typically manifest over quarterly horizons (Gilbert 2010).

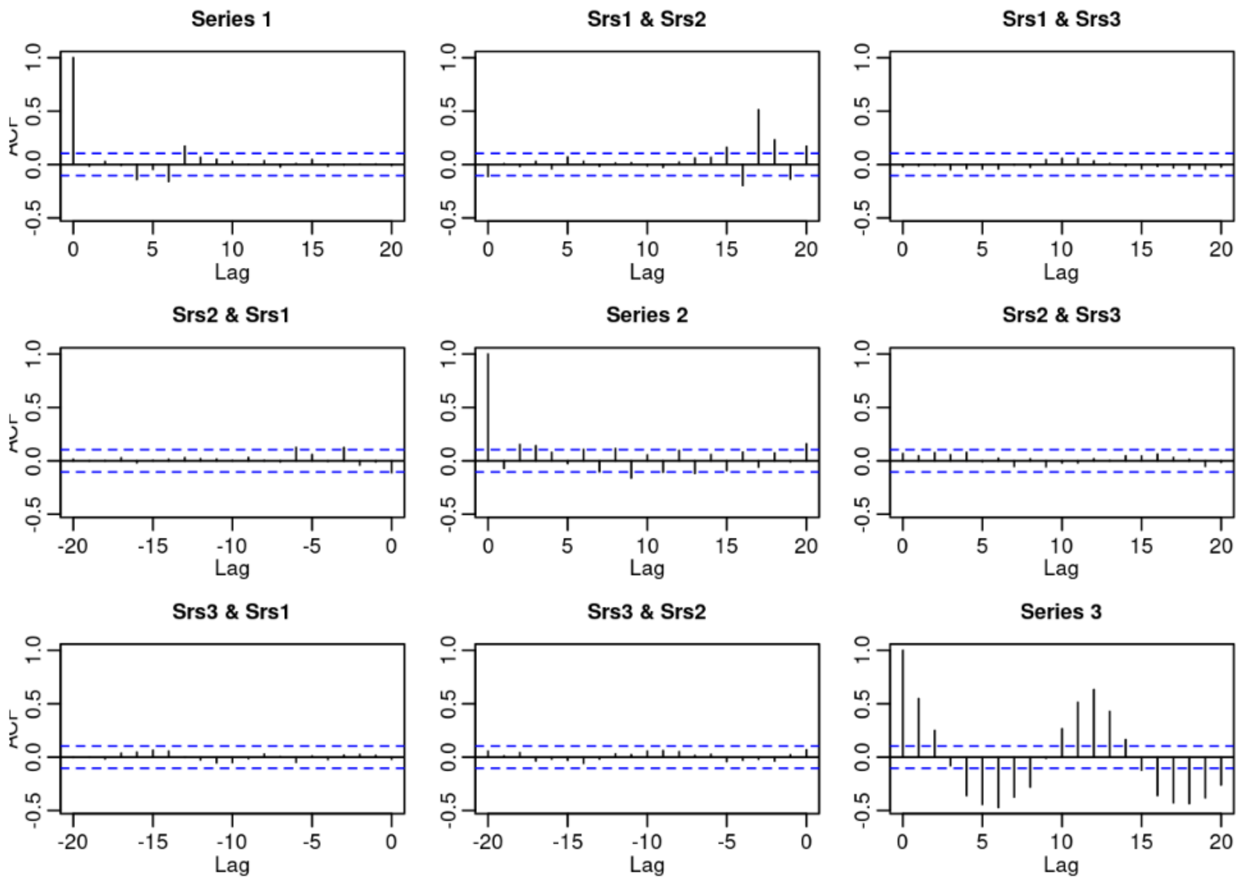

Figure 1: The residuals in the VARMA(1, 2) model shows that most of the bars, which represent autocorrelation at different lags, fall within the blue confidence interval bands, typically the default for ACF plots. This indicates that the residuals do not exhibit significant autocorrelation and are random. Therefore, the residuals can be considered white noise, suggesting that the model has effectively captured the underlying patterns. Based on this, we will forecast the price fluctuations over time using VARMA(1, 2) model.

5 Forecasting and Results

The performance of classical time series models like ARIMA and ARIMAX was assessed using weekly cocoa price data. Although these models are useful for understanding temporal dependencies and the impact of climate variables, their predictive performance was limited. Forecast accuracy metrics (MAE, RMSE, MAPE) indicated relatively high values for all ARIMA and ARIMAX models, with MAE ranging from 5009 to 5052, RMSE from 5429 to 5470, and MAPE from 175.41% to 176.77%, suggesting a poor predictive fit. This underperformance highlights the limitations of ARIMA in capturing cocoa price dynamics. In contrast, models like VARMA and XGBoost showed improved forecast accuracy, indicating a better ability to model the underlying data patterns. The VARMA model showed an RMSE of 1331.574 and an MAE of 1097.820, indicating a relatively higher magnitude of forecast errors compared to XGBoost. A GARCH(1,1) model was also applied to log returns to capture volatility in weekly cocoa prices. The model showed strong evidence of conditional heteroskedasticity and high volatility persistence (+ 0.99), consistent with commodity market behavior. Residual diagnostics confirmed the model's adequacy. While GARCH does not enhance point forecast accuracy, it effectively models the uncertainty and risk in price movements. In comparison, XGBoost, with MAE of 38.11, RMSE of 235.46, and MAPE of 41.14%, outperformed the other models, providing the most accurate forecast for cocoa prices.

Table 5: Forecast Error Metrics by Model. ARIMA and ARIMAX models are based on weekly data, while VARMA and XGBoost use monthly data. Lower values for MAE, RMSE, and MAPE indicate better forecast accuracy.

Model	MAE	RMSE	MAPE
ARIMAX(2,1,2)	5018.82	5438.84	175.73
ARIMAX(1,1,1)	5009.09	5429.23	175.41
ARIMA(0,1,1)	5034.85	5452.90	176.20
ARIMA(0,1,2)	5052.54	5470.55	176.77
VARMA	1097.82	1331.57	97.87
XGBoost	38.11	235.46	41.14

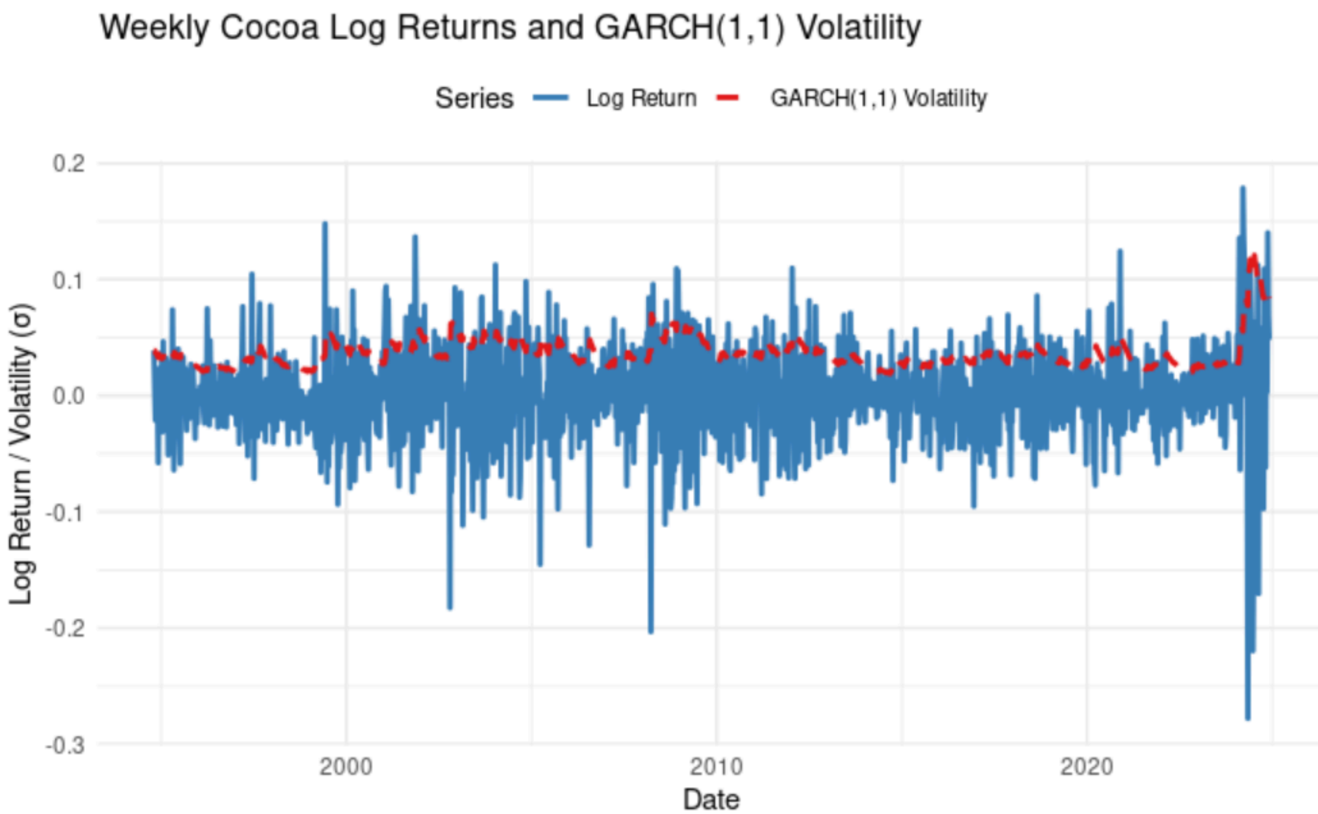

Figure 2: Weekly cocoa log returns with GARCH(1,1) estimated volatility. The dashed line shows changing volatility over time, highlighting periods of greater market risk.

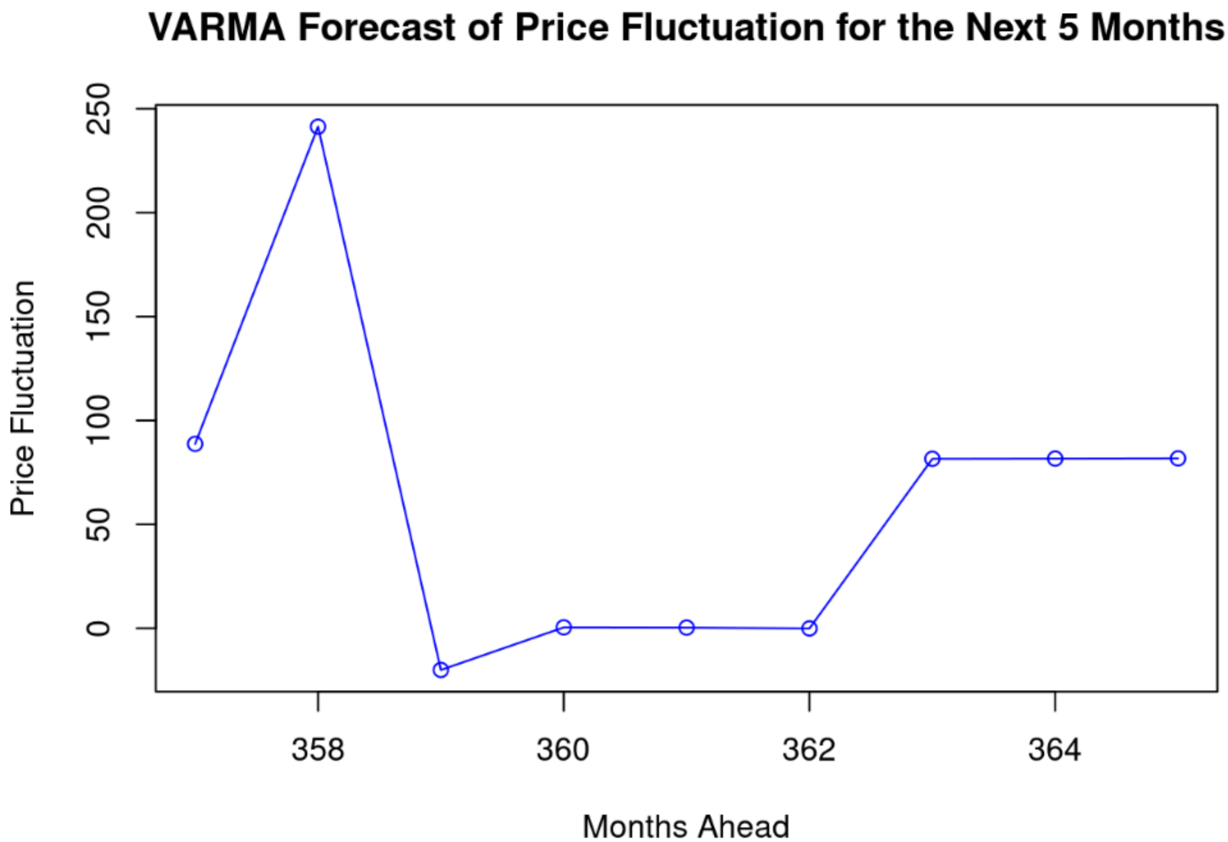

Figure 3: The forecast for price changes indicates significant short-term volatility, with an initial sharp increase followed by fluctuations before stabilizing around a moderate positive value. The first two steps show a large jump (+88.74 to +241.41), followed by a drop (-20.06), and then a gradual stabilization around 20–30. This suggests that prices may experience a sharp initial surge before settling into a more predictable upward trend. If applied to actual prices, this pattern implies a high-risk, high-reward scenario in the short term, followed by steady growth. For trading or investment decisions, this means navigating an initial period of uncertainty before benefiting from more stable price movements.

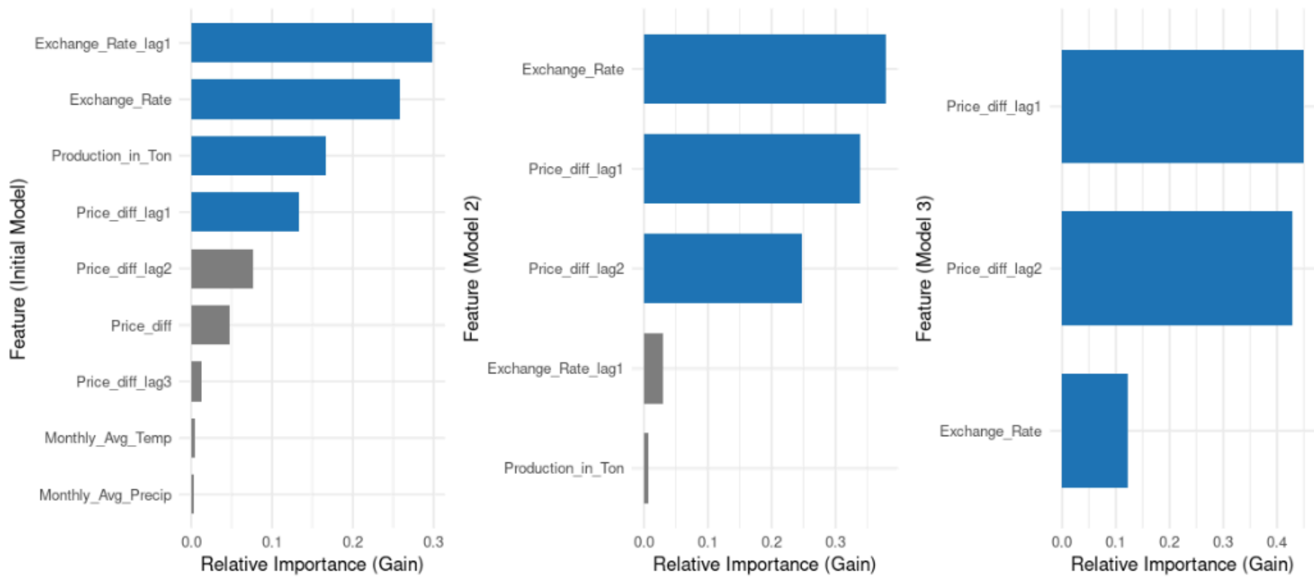

Figure 4: Feature Importance of Three Models: As XGBoost ranks feature by its importance and highlights which variables drive price fluctuations, focusing on substantial predictive importance features by dropping variables with importance below 0.1, the final XGBoost model is determines by the features Exchange rate, Price difference with 1 lag, Price difference with 2 lag

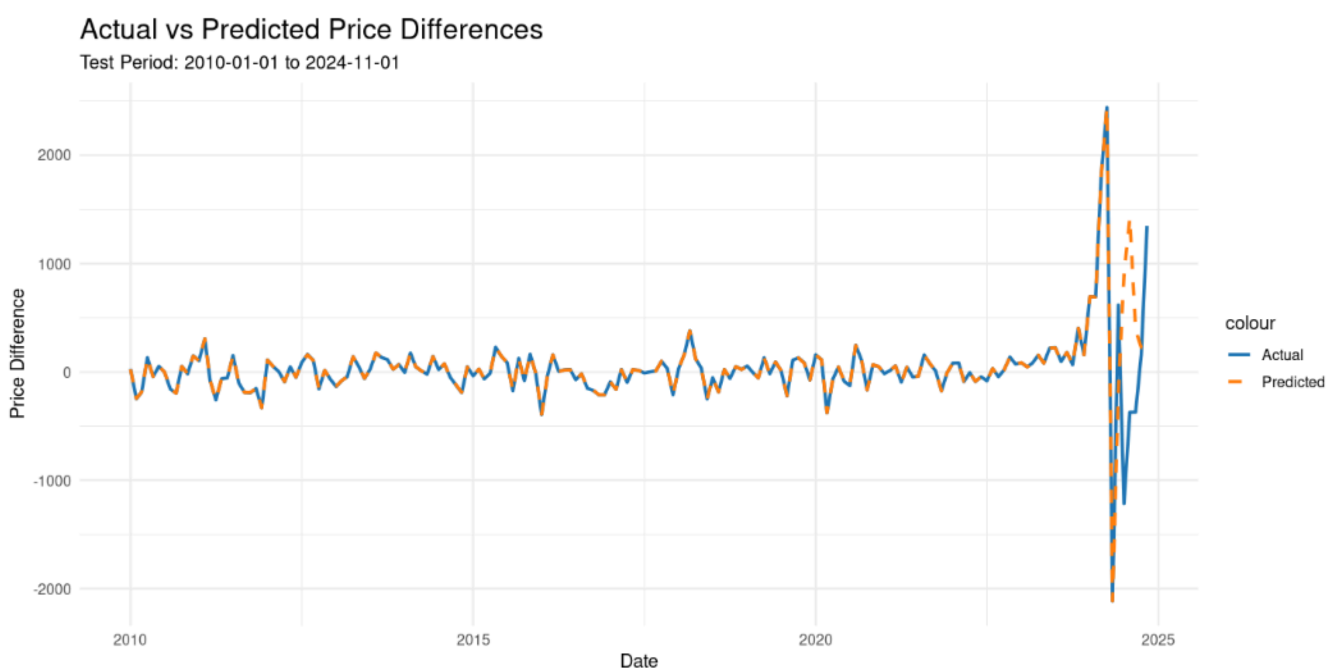

Figure 5: The discrepancies between the actual and predicted prices over the test period from January 1, 2010, to November 1, 2024 with the price differences fluctuate with notable deviations indicates the model significantly overestimated the actual prices. The trend over time shows variability with no clear consistent pattern, suggesting that the model's performance may be influenced by external factors or inherent limitations in capturing the underlying price dynamics.

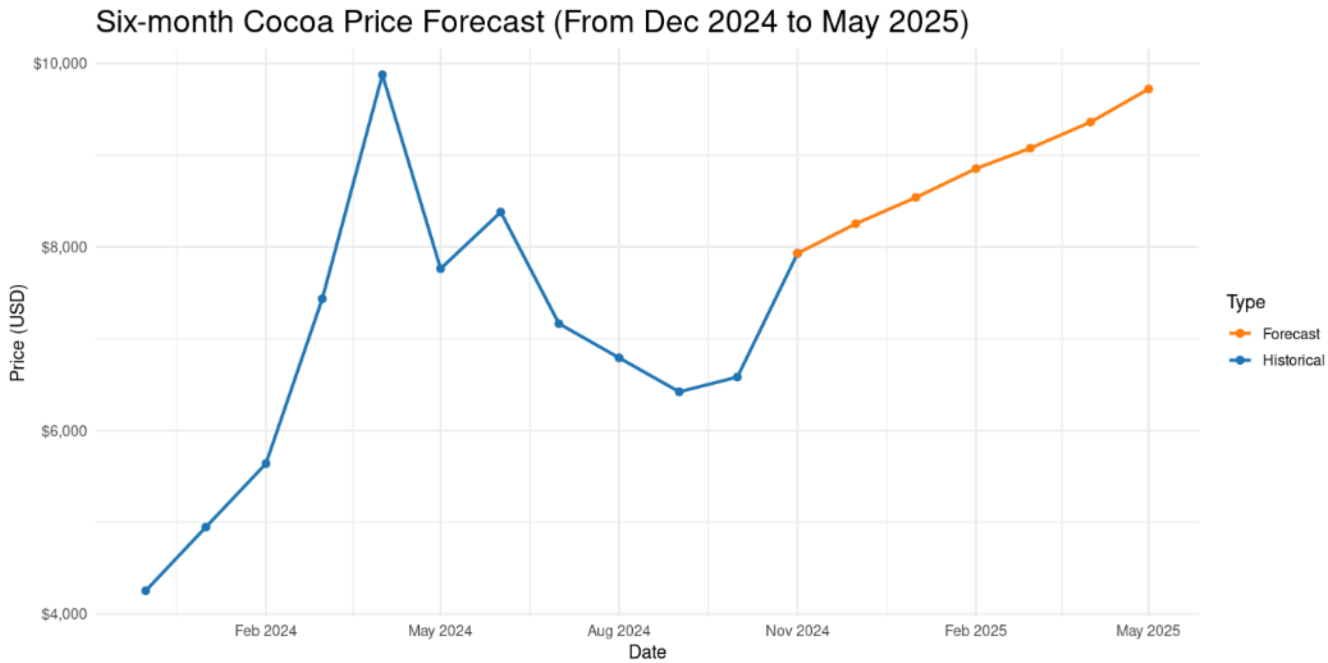

Figure 6: Since XGBoost predicts price differences $\hat{y}_t$, the actual cocoa price is calculated recursively by $\text{Price}_t = \text{Price}_{t-1} + \hat{y}_t$, where price at time t is derived by adding the predicted difference to the previous price. According to model 3, the forecasted monthly cocoa price during the period from December 2024 to May 2025 is increasing linearly approximately from \$8000 to \$9800

6 Discussion and Conclusion

6.1 Discussion

This analysis applied ARIMA, ARIMAX, VARMA, and XGBoost models to forecast future cocoa prices and assess volatility with GARCH. Traditional time series models provided a useful baseline and offered interpretability, but their forecasting performance was limited. The GARCH(1,1) model effectively captured volatility clustering, reflecting persistent risk in cocoa price movements. ARIMAX models included climate-related variables (such as precipitation and temperature), but these exogenous variables had limited influence on forecasting accuracy. Their coefficients were small and often not statistically significant, suggesting that while they may impact cocoa production over the long term, they did not contribute meaningfully to short-term price forecasts in this setting. XGBoost significantly outperformed all other models, including VARMA(1, 2), demonstrating a strong ability to capture nonlinear and complex relationships. This suggests that hybrid approaches may offer improved accuracy and robustness in forecasting commodity prices.

6.2 Conclusion

In our analysis of various forecasting models for cocoa prices, we evaluated several approaches, including ARIMA, ARIMAX, GARCH, VARMA, and XGBoost, using key performance metrics like MAE, RMSE, and MAPE. Among these models, XGBoost emerged as the top performer, delivering the most accurate predictions with the lowest error rates. The ARIMA and ARIMAX models, while useful in capturing price movements, demonstrated relatively higher forecast errors, especially in terms of MAE and RMSE. The GARCH model, although effective in modeling volatility and uncertainty, did not significantly improve the point forecast accuracy, highlighting its strength in risk management rather than direct price forecasting. VARMA, on the other hand, showed very poor performance, with notably high errors across all metrics, making it less suitable for this task.

After comparing the models, we chose XGBoost due to its superior ability to capture the underlying patterns in the cocoa price data, offering the most precise forecast while minimizing prediction error. As a result, the XGBoost model forecasts a steady increase in the monthly cocoa price from December 2024 to May 2025, rising from approximately $8000 to $9800. This forecast reflects a strong upward trend, likely driven by market factors such as supply-demand dynamics and broader economic influences.

6.3 Limitations

Forecasting Ghana's cocoa prices involved several limitations that potentially affected model accuracy and generalizability. A key challenge was data availability and quality. Crucial economic indicators, such as global demand and supply indices, were either inaccessible or available only upon request, limiting the inclusion of important explanatory variables. The dataset from Ghana's Cocoa Board also lacked full temporal coverage. Additionally, daily climate data contained substantial missing values and spatial inconsistencies across weather stations further reduced reliability. Methodologically, reliance on univariate time series models (e.g., ARIMA) constrained the analysis of external influences. Although multivariate models like VARMA offer a more comprehensive approach, their implementation was hindered by data sparsity and complexity. In the XGBoost model, several features had to be excluded due to excessive missingness, potentially reducing predictive power. Lastly, access to advanced tools and datasets was limited by paywalls and proprietary restrictions, impeding broader model exploration.

6.4 Further Suggestion

To enhance the robustness and accuracy of future forecasting efforts, several directions are worth considering. First, the integration of deep learning models, such as Long Short-Term Memory (LSTM) networks or Transformer-based architectures, may offer significant advantages in capturing complex, nonlinear patterns and long-range dependencies inherent in volatile cocoa markets. These models have shown promise in similar agricultural forecasting tasks and may outperform traditional and machine learning approaches, particularly when sufficient data is available.

Second, future research would benefit from incorporating higher-frequency data, particularly daily indicators of supply and demand dynamics, rather than relying solely on aggregate annual production figures. This would enable a more responsive modeling framework, better suited to capturing short-term market shifts driven by logistical, economic, or geopolitical events.

Finally, it is recommended that researchers explore the possibility of direct engagement with relevant institutions, such as governmental agencies, commodity boards, or international trade organizations. Formally requesting access to proprietary datasets may yield more granular and timely information, thus supporting the development of richer and more accurate forecasting models.

References

Board, Ghana Cocoa. 2024. "Review of the Producer Price of Cocoa for the 2024/2025 Cocoa Season." https://cocobod.gh/news/review-of-the-producer-price-of-cocoa-for-the-20242025-cocoa-season-wednesday-11th-september-2024.

Data, Our World in. 2024. "Cocoa Bean Production." https://ourworldindata.org/grapher/cocoa-bean-production?tab=chart&country=~GHA.

Frankel, Jeffrey A., and Georgios A. Hardouvelis. 1985. "Commodity Prices, Money Surprises and Fed Credibility." *Journal of Money, Credit and Banking* 17 (4): 425–38. https://doi.org/10.2307/1992439.

Gilbert, Christopher L. 2010. "Speculative Influences on Commodity Futures Prices 2006–2008." UNCTAD Discussion Papers 197. United Nations Conference on Trade; Development.

Katsiampa, Paraskevi. 2017. "Volatility Estimation for Bitcoin: A Comparison of GARCH Models." *Economics Letters* 158: 3–6. https://doi.org/10.1016/j.econlet.2017.06.023.

Mayr, Stefan, and Dirk Ulbricht. 2015. "Log Versus Level in VAR Forecasting: 42 Million Empirical Answers—Expect the Unexpected." *Economics Letters* 126: 40–42. https://doi.org/10.1016/j.econlet.2014.11.008.

Menéndez-García, Luis A., Pedro J. García-Nieto, Enrique García-Gonzalo, and Francisco Sánchez Lasheras. 2024. "Time Series Analysis for COMEX Platinum Spot Price Forecasting Using SVM, MARS, MLP, VARMA and ARIMA Models: A Case Study." *Resources Policy* 95: 105148. https://doi.org/10.1016/j.resourpol.2024.105148.

Nayak, G. H. H., M. W. Alam, K. N. Singh, and S. Sinha. 2024. "Exogenous Variable Driven Deep Learning Models for Improved Price Forecasting of TOP Crops in India." *Scientific Reports* 14: 17203. https://doi.org/10.1038/s41598-024-68040-3.

Nunley, Christopher (Producer & Editor), Claudia (Graphics) Locopo, and Jason (Supervising Producer) Morganteen. 2024. "How the Price of Chocolate's Key Ingredient Surged More Than 300% in 2024 [Video]." https://www.youtube.com/watch?v=tJwVoma8ZVk.

Organization, International Cocoa. n.d. "Statistics." https://www.icco.org/statistics/#price.

"USD/GHS - US Dollar Ghanaian Cedi Exchange Rate." 2024. https://ca.investing.com/currencies/usd-ghs-historical-data.

Wang, Dawei, Chao Yue, Shujie Wei, and Jin Lv. 2017. "Performance Analysis of Four Decomposition-Ensemble Models for One-Day-Ahead Agricultural Commodity Futures Price Forecasting." *Algorithms* 10 (3): 108. https://doi.org/10.3390/a10030108.