Table of Contents

1 Abstract

This paper forecasts monthly cocoa prices using data from the International Cocoa Organization and climate records from Ghana. After standardized the data, a series of time series models were applied to forecast the prices, including Exponential Smoothing (ETS), ARIMA, SARIMA, linear regression with climate covariates, and GARCH models. what was found, and why this matters

2 Introduction

Forecasting commodity prices is a challenge in economics and statistical modelling because of the multi-factorial drivers of price behaviour. However, forecasting commodity prices is important for producers, traders and policymakers to have more understanding about the market. Cocoa is a globally traded commodity with significant economic relevance, particularly in regions where it is both produced and consumed at scale such as Ghana. For stakeholders such as producers, traders, and policymakers, accurate forecasting is vital to design procurement strategies, manage supply chain risks, and stabilize income.

Some real-world examples also shows the importance of forecasting the Cocoa price. In 2016–2017, global cocoa prices declined by over 30% (International Cocoa Organization 2016), leading to significant income losses for smallholder farmers in Ghana. However, cocoa exports constitute a major share of national revenue in Chana. This forced some farmers to abandon cocoa cultivation or turn to alternative livelihoods, including environmentally damaging activities such as illegal mining (Bryant and Mitchell 2021). In order to stabilize the price of cocoa, Ghana and Côte d'Ivoire jointly introduced the Living Income Differential (LID) in 2019, establishing a $400-per-ton premium on cocoa exports to support farmer incomes (Squicciarini, Vandeplas, and Barreiro-Hurle 2021). This real-world example shows how price instability can widely influence the economic and social consequences, and highlights the importance of forecasting models.

This paper aims to develop an accurate model for predicting cocoa prices. The paper investigates the monthly behaviour of cocoa prices by using two key datasets, including daily cocoa futures prices from the International Cocoa Organization and daily climate data from Ghana, the largest cocoa-producing country in the world. The analysis focuses on modelling the monthly change in log-transformed cocoa prices. The differencing method was use to address non-stationarity. A series of forecasting models was evaluated, including Exponential Smoothing (ETS), ARIMA, SARIMA, linear regression with climate covariates, and GARCH models. Each model was trained on a 70% subsample and assessed using the remaining 30% sample, which is a 70/30 train-test split. Forecast accuracy was assessed with root mean square error (RMSE), AIC, and BIC, with all predictions back-transformed to the original price scale.

3 Methodology

###

previous study cocoa

```
### Load and Aggregate Cocoa Price Data to Monthly
prices <- read.csv(here::here("00-data/price.csv"))
prices$Date <- as.Date(prices$Date, format='%d/%m/%Y')
prices$Price <- as.numeric(gsub(",", "", prices$ICCO.daily.price..US..tonne.))
prices <- prices %>%
  mutate(YearMonth = floor_date(Date, "month")) %>%
  group_by(YearMonth) %>%
  summarise(Price = mean(Price, na.rm = TRUE)) %>%
  ungroup()

### Load and Aggregate Ghana Weather Data to Monthly
climate <- read.csv(here::here("00-data/climate.csv"))
climate$DATE <- as.Date(climate$DATE)
climate <- climate %>%
  mutate(YearMonth = floor_date(DATE, "month")) %>%
  group_by(YearMonth) %>%
  summarise(across(c(PRCP, TAVG, TMAX, TMIN), mean, na.rm = TRUE))
```

```
Warning: There was 1 warning in `summarise()`.
i In argument: `across(c(PRCP, TAVG, TMAX, TMIN), mean, na.rm = TRUE)`.
i In group 1: `YearMonth = 1990-01-01`.
Caused by warning:
! The `...` argument of `across()` is deprecated as of dplyr 1.1.0.
```

```
Supply arguments directly to `.fns` through an anonymous function instead.

  # Previously
  across(a:b, mean, na.rm = TRUE)

  # Now
  across(a:b, \(x) mean(x, na.rm = TRUE))
```

```r
### Merge and Clean Monthly Data
cocoa_data <- left_join(prices, climate, by = "YearMonth")
```

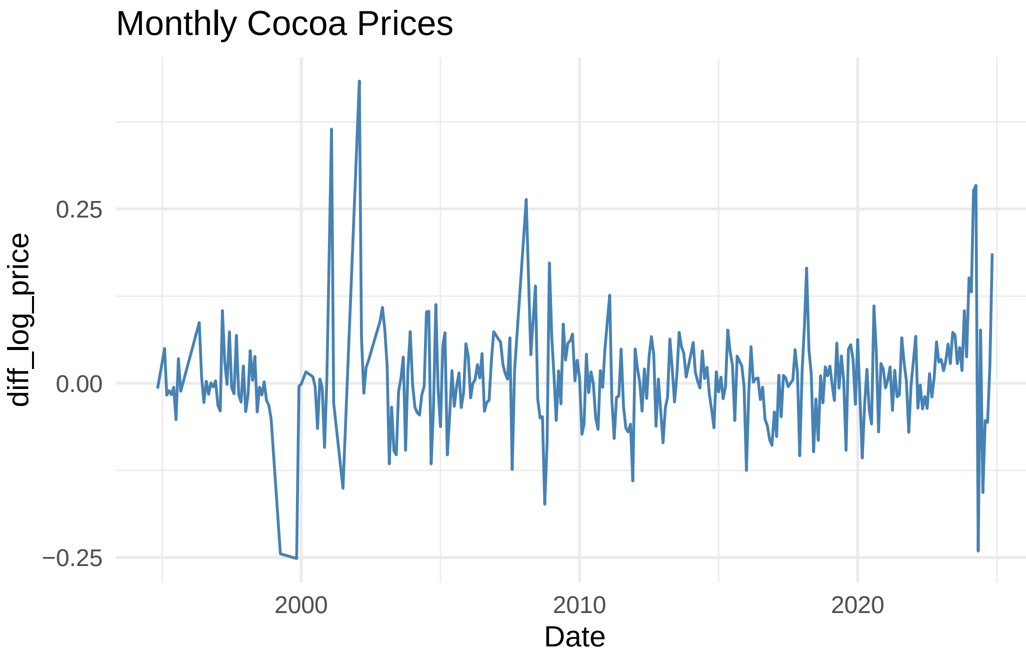
```r
### Plot Monthly Time Series
ggplot(cocoa_data, aes(x = YearMonth)) +
  geom_line(aes(y = Price), color = "steelblue") +
  labs(title = "Monthly Cocoa Prices", y = "Price", x = "Date") +
  theme_minimal()
```

response variable price log difference stationary

```r
cocoa_data <- cocoa_data %>%
  mutate(log_price = log(Price)) %>%
  drop_na()
```

```r
### Plot Monthly Time Series
ggplot(cocoa_data, aes(x = YearMonth)) +
  geom_line(aes(y = log_price), color = "steelblue") +
  labs(title = "Monthly Cocoa Prices", y = "log_price", x = "Date") +
  theme_minimal()
```


```r
### Merge and Clean Monthly Data
cocoa_data <- cocoa_data %>%
  mutate(diff_log_price = c(NA, diff(log_price))) %>%
  drop_na()
### Plot Monthly Time Series
ggplot(cocoa_data, aes(x = YearMonth)) +
  geom_line(aes(y = diff_log_price), color = "steelblue") +
  labs(title = "Monthly Cocoa Prices", y = "diff_log_price", x = "Date") +
  theme_minimal()
```



7:3 7 3

ETS ARIMA SARIMA Linear Regression Model ARCH/GARCH

transform transform predictions back to origin log exponential

sum square of erro, AIC BIC

4 Data

Table 1: Summary Statistics of Important Variables

Date	Price	Daily Perception	Average Temperature	Maximum Temperature	Minimum Temperature
Min. :1994-10-12	Min. : 778.4	Min. : 0.00000	Min. :73.60	Min. : 76.50	Min. :61.00
1st Qu.:2005-02-01	1st Qu.: 1689.4	1st Qu.: 0.00000	1st Qu.:78.56	1st Qu.: 85.50	1st Qu.:72.57
Median :2014-10-16	Median : 2330.7	Median : 0.07775	Median :80.50	Median : 88.67	Median :73.67
Mean :2012-09-06	Mean : 2589.3	Mean : 0.24756	Mean :80.62	Mean : 88.44	Mean :73.85

Table 1: Summary Statistics of Important Variables

Date	Price	Daily Perception	Average Temperature	Maximum Temperature	Minimum Temperature
3rd Qu.:2020-09-16	3rd Qu.: 2931.2	3rd Qu.: 0.30042	3rd Qu.:82.60	3rd Qu.: 91.25	3rd Qu.:75.00
Max. :2024-11-28	Max. :10690.7	Max. :10.28000	Max. :88.00	Max. :101.00	Max. :82.00

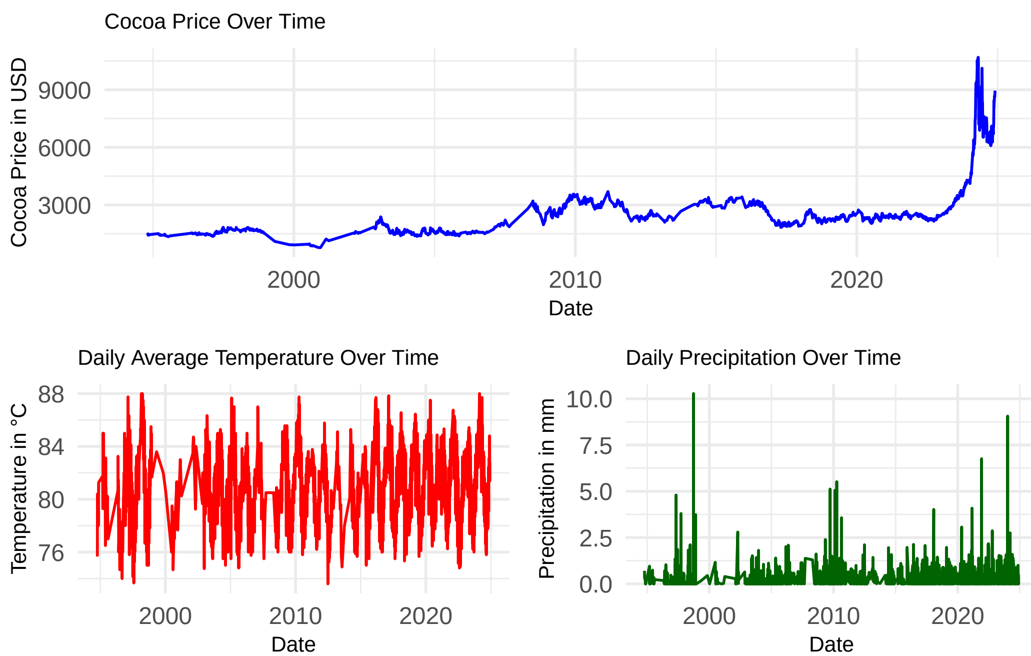


Figure 1: Time Series of Cocoa Price, Local Average Temperature, and Precipitation

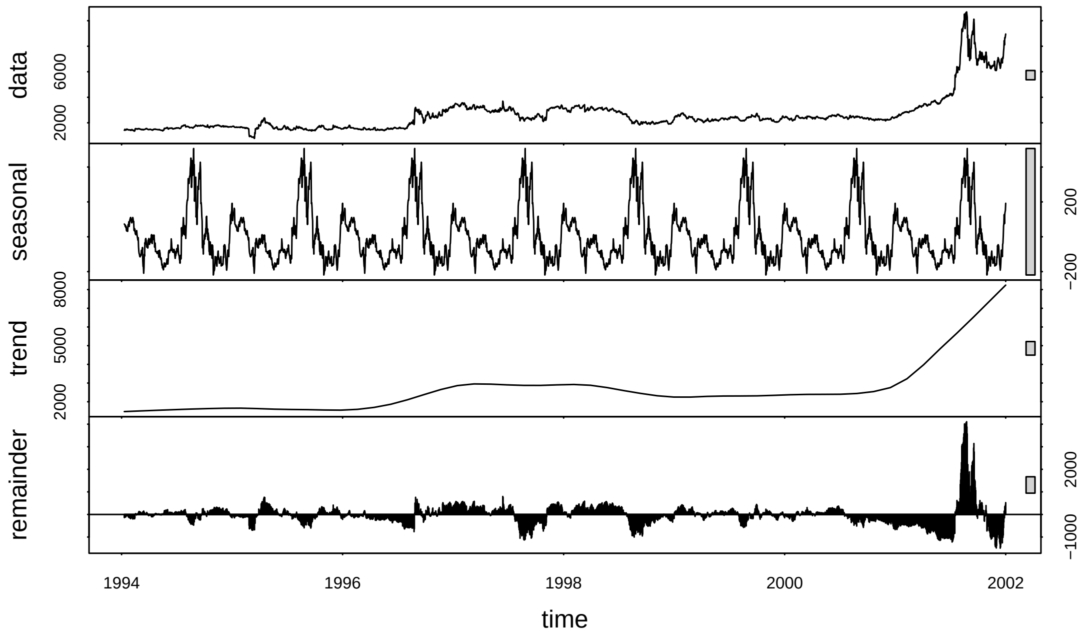


Figure 2: Decompose of Series of Prices

Bryant, Chris, and Matthew I Mitchell. 2021. "The Political Ecology of Cocoa in Ghana: Past, Present and Future Challenges." In *Natural Resources Forum*, 45:350–65. 4. Wiley Online Library.

International Cocoa Organization. 2016. "Monthly Cocoa Market Review - November 2016." International Cocoa Organization. https://www.icco.org/wp-content/uploads/2019/07/ICCO-Monthly-Cocoa-Market-Review-November-2016.pdf.

Squicciarini, Mara P., Anneleen Vandeplas, and Jesús Barreiro-Hurle. 2021. "Living Income Differential in the Cocoa Sector: Theory and Impact." JRC125754. European Commission, Joint Research Centre. https://publications.jrc.ec.europa.eu/repository/bitstream/JRC125754/lid_paper_sfpr_final.pdf.