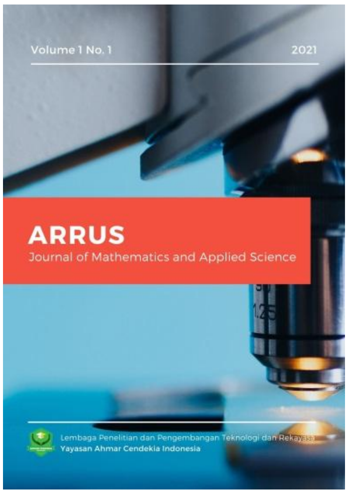
OPEN ACCESS

ISSN : 2776-7922 (Print) / 2807-3037 (Online)



*Corresponding author: Nur Ikhwana, Statistics Department, Faculty of Mathematics and Natural Science, Universitas Negeri Makassar, Makassar, Indonesia

E-mail: nurikhwana@unm.ac.id

RESEARCH ARTICLE

Forecasting Monthly Red Chili Prices in South Sulawesi Using Prophet Model with Time Series Cross-Validation*

Nur Ikhwana*, Agung Tri Utomo, Annisa Syalsabila,& Lalu Ramzy Rahmanda

Statistics Department, Faculty of Mathematics and Natural Science, Universitas Negeri Makassar, Makassar, Indonesia

Abstract: This study examines monthly red chili price movements in South Sulawesi using the Prophet forecasting model. Daily price data from the National Strategic Food Price Information Center (PIHPS) covering January 2020 to May 2026 were aggregated into 77 monthly observations. Missing values were handled using linear interpolation and Last Observation Carried Forward (LOCF) before modeling. The Prophet model using for forecasting and time series cross-validation was used as a validation method. The model performance was evaluated by Root Mean Square Error (RMSE), Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE). The results indicate that the model produced an average RMSE = 11,024.68 IDR/kg, MAE = 9,139.04 IDR/kg and MAPE = 23.92%, suggests an acceptable forecasting performance for a highly volatile agricultural commodity. Results pattern demonstrate an annual seasonality on red chili prices, usually lowest price occurs in September and highest one in March of next year. The 12 months projections also foresee a maximum price of 72,597 IDR/kg in March 2027. These results show that the Prophet model can capture trend and seasonality, especially in predicting red chili prices of South Sulawesi.

Keywords: Prophet, Red Chili Prices, Forecasting, Time Series, Cross-Validation

1. Introduction

Red chili is a core food commodity in Indonesia, and its price fluctuations are deeply tied to the country's food inflation level and food security status. Reports from Food and Agriculture Organization, 2023, and the World Bank (2024) both note that price volatility of fresh agricultural products globally has generated widespread socioeconomic impacts. This study sets its core research site as South Sulawesi, Indonesia's main production and consumption center for red chili. Local red chili prices are primarily driven by five core factors: weather, harvest cycles, transportation costs, supply chain disruptions, and seasonal demand (Fadillah et al., 2025). Sharp price fluctuations have brought widespread uncertainty to all value-chain stakeholders including producers, traders, consumers, and policymakers, which highlights the practical necessity of building an accurate price forecasting model. In previous research on agricultural product price forecasting, traditional time series models represented by exponential smoothing and ARIMA have achieved sound application outcomes. A study by Hyndman & Athanasopoulos (2021) clearly points out that agricultural price forecasting inherently faces challenges stemming from strong seasonality, structural shifts, and irregular fluctuations, while the Prophet model proposed by (Taylor & Letham, 2018) can specifically

https://doi.org/10.35877/mathscience4893

ISSN : 2776-7922 (Print) / 2807-3037 (Online)

adapt to these complex fluctuation scenarios. At present, there are two distinct research gaps in the field. First, applied research using the Prophet model for red chili in South Sulawesi is extremely scarce. Second, very few studies adopt the rigorous rolling-origin time series cross-validation method to conduct model evaluation. To address these gaps, this study uses monthly Indonesian red chili price data from January 2020 to May 2026 provided by PIHPS, adopts RMSE, MAE, and MAPE as core evaluation metrics, and assesses the Prophet model's performance via time series cross-validation. The study finds that the Prophet model can accurately capture the annual seasonal characteristics of this product's prices, a conclusion that supplements the empirical gap in research on the Prophet model's applicability to price forecasting for highly volatile agricultural products.

2. Literature Review

2.1. Time Series Forecasting

Time series forecasting is the process of predicting future values based on patterns observed in time-sequenced historical data (Hyndman & Koehler, 2006). The ARIMA method, introduced by Box & Jenkins (1976), is the most widely used classical approach, but it requires data stationarity and does not explicitly handle complex seasonal components. The Exponential Smoothing (ETS) method offers more flexibility in handling trends and seasonality, but remains limited to relatively simple patterns.

Forecast accuracy evaluations commonly use metrics such as MAE, RMSE, and MAPE. (Hyndman & Koehler, 2006) emphasize the importance of selecting an appropriate evaluation metric according to the data characteristics and forecasting objectives. MAPE is particularly popular because it produces relative values in percentages that are easy to interpret practically.

2.2. Prophet Model

Prophet, an open-source time series forecasting program led by the Meta (formerly Facebook) team and proposed by Taylor and Letham in 2018 (Taylor & Letham, 2018), is suitable for fitting time series with strong seasonal patterns, trend changepoints, and holiday effects. Its core framework adopts the additive decomposition formula:

$$y(t) = g(t) + s(t) + h(t) + \varepsilon(t)$$

where each component corresponds in order to the trend $g(t)$, seasonal term $s(t)$, holiday term $h(t)$, and error term $\varepsilon(t)$. It fits the seasonal component using Fourier series and constructs its trend via piecewise linear or logistic growth. Its core advantages include automatic changepoint detection and simple, accessible configuration (Taylor & Letham, 2018)

2.3. STL Decomposition

STL (the Seasonal and Trend decomposition method based on Loess) was first proposed by Cleveland et al. in 1990. It can decompose a time series into three components: trend, seasonal, and residual. The additive decomposition framework can be expressed as:

$$Y_t = T_t + S_t + R_t$$

Where (Y_t) represents the observed series, (T_t) represents the trend component, (S_t) represents the seasonal component, (R_t) represents the remainder or irregular component.

Relying on locally weighted regression, STL does not require the assumption that the trend is linear, can process outliers, and outperforms classical decomposition methods. In the data exploration stage before Prophet modeling in this study, STL is used to verify the existence of the time series seasonal pattern.

2.4. Time Series Cross-Validation

The evaluation of time series forecasting models has unique, dedicated requirements. Research by (Hyndman & Koehler, 2006) notes that standard k-fold cross-validation can introduce look-ahead bias, making it unsuitable for time series data. Time series cross-validation (TSCV) adopts a rolling origin evaluation method to guarantee that training data always chronologically precedes test data. The cross validation function of the Prophet tool, proposed by (Taylor & Letham, 2018), can implement this method via the initial, period, and horizon parameters. MAPE evaluation criterion is widely used to assess commodity price forecasting models (Lewis et al., 1983).

2.5. Red Chili Price Volatility and Food Inflation

Red chili is one of the food commodities with the most dramatic price fluctuations in Indonesia, and a common trigger of regional inflation volatility. Its price is affected by five core factors; combined with its own short production cycle and limited storage capacity, it is extremely prone to sharp rises and falls in price over a short period. Accurate price forecasting can not only support the implementation of macro policies such as

food price monitoring and inflation control, but also help all actors across the entire supply chain—including producers and distributors—anticipate market conditions, optimize supply, stabilize price fluctuations, and improve the operational efficiency of the market.

3. Research Method and Materials

3.1. Data Source

The core dataset used in this study for red chili price forecasting is drawn from the official market monitoring records of the National Strategic Food Price Information Center (PIHPS), covering daily price data from South Sulawesi Province, Indonesia, spanning January 2020 to May 2026. As the original dataset contains missing values, we first applied linear interpolation to fill these gaps to preserve the continuity of the time series. We then aggregated the daily data into monthly average prices to mitigate short-term fluctuations, producing a stable time series for use in forecasting.

3.2. Missing Value Handling

Daily data contains a number of missing values marked with a "-" character in the original source. Handling is carried out in stages using the zoo package (Zeileis & Grothendieck, 2005): first, linear interpolation with na.approx() to fill in missing values between two available observations; second, LOCF (Last Observation Carried Forward) with na.locf() for missing values at the end of the series; and third, NOCB (Next Observation Carried Backward) with na.locf (fromLast = TRUE) for missing values at the beginning of the series. A final verification confirms that there are no remaining missing values (sum(is.na) = 0).

3.3. Data Aggregation

Aggregation of daily data to monthly level was performed using the dplyr (Wickham et al., 2023) and lubridate (Grolemund & Wickham, 2011) packages. Daily red chili prices were aggregated into monthly average prices using the mean aggregation approach. This process reduced short-term fluctuations and produced 77 monthly observations for forecasting analysis.

3.4. Exploratory Time Series Analysis

This study takes forecasting the monthly price of red chili in South Sulawesi Province as its core goal. Before constructing formal models, it conducts exploratory time series analysis: first, it visualizes the monthly price series to identify its long-term trend, fluctuation patterns, and potential seasonal patterns; then it adopts the Loess-based Seasonal-Trend Decomposition using Loess (STL) to decompose the series into three components: trend, seasonality, and residual. It sorts out these data characteristics to support subsequent model development.

https://doi.org/10.35877/mathscience4893

ISSN : 2776-7922 (Print) / 2807-3037 (Online)

3.5. Prophet Forecasting Model

This study uses the Prophet prediction model to conduct monthly price forecasting for red chili peppers. First, complete data preprocessing and exploratory analysis, converting the data into the date + response variable format required by the model. Since this study uses monthly observational data, only enable annual seasonality and disable weekly and daily seasonality. The model fit with the full set of data ranging from January 2020 to May 2026, and generate forecast values for the following 12 months.

3.6. Model Evaluation

This study adopts the time series cross-validation method to evaluate the predictive performance of the Prophet model. This method preserves the chronological order of observed data and conducts evaluations using a rolling forecast origin. The evaluation procedure employed an initial training window of 1,460 days, a forecasting horizon of 365 days, and a cutoff interval of 90 days, resulting in six evaluation folds. Forecast accuracy was assessed using Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE). These metrics were used to quantify forecasting errors and evaluate the predictive performance of the Prophet model.

4. Results and Discussion

4.1. Descriptive Statistics

Descriptive statistics were calculated to provide an overview of red chili price behavior in South Sulawesi. After data preprocessing and monthly aggregation, a total of 77 monthly observations were obtained covering the period from January 2020 to May 2026.

Table 1. Descriptive Statistics

Statistics	Minimum	Q1	Median	Mean	Q3	Maximum
Value (IDR/kg)	15075	27633	37420	39847	49638	83040

Monthly red chili price data in South Sulawesi (Table 1) shows very high variability. The price range between IDR 15,075/kg and IDR 83,040/kg reflects the extreme volatility common to horticultural commodities. The average price of IDR 39,847/kg is above the median (IDR 37,420/kg), indicating a slightly right-skewed distribution, meaning several extremely high price events pushed the average upward.

The last six months of data (December 2025–May 2026) recorded the highest price in March 2026 at IDR 64,330/kg, reflecting supply pressure due to seasonal conditions. This fluctuation pattern is consistent with the BPS South Sulawesi (2026) report on regional red chili price dynamics.

4.2. Time Series Visualization

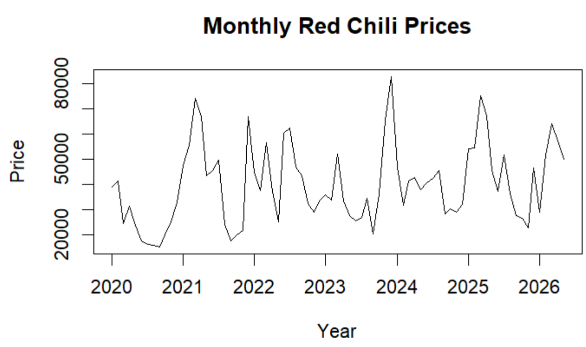


Figure 1. Monthly Red Chili Prices in South Sulawesi

Based on Figure 1. During the observation period, the price of red chili peppers fluctuated roughly between 15,000 IDR/kg and over 80,000 IDR/kg. Four price peaks exceeding

70,000 IDR/kg were recorded successively, occurring in early 2021, late 2021, early 2024, and early 2025 respectively. In addition, three price troughs that fell below 25,000 IDR/kg emerged in mid-2020, mid-2021, and late 2025. This price series has no strong long-term upward trend. The intervals between its peaks and troughs follow a regular pattern, and it shows stable annual seasonality, which aligns with the common characteristics of bulk agricultural commodities affected by climate and harvesting plans. The core causes of these fluctuations are seasonal factors, production cycles, and supply-demand imbalances. Follow-up studies will use the STL decomposition method and the Prophet forecasting model, which can capture both trend and seasonal patterns simultaneously, to conduct quantitative analysis.

4.3. STL Decomposition Analysis

To clarify the underlying structure of the data, this study used the STL decomposition method proposed by Cleveland et al. in 1990 before modeling. This method splits time series data into three components: trend, seasonal, and residual.

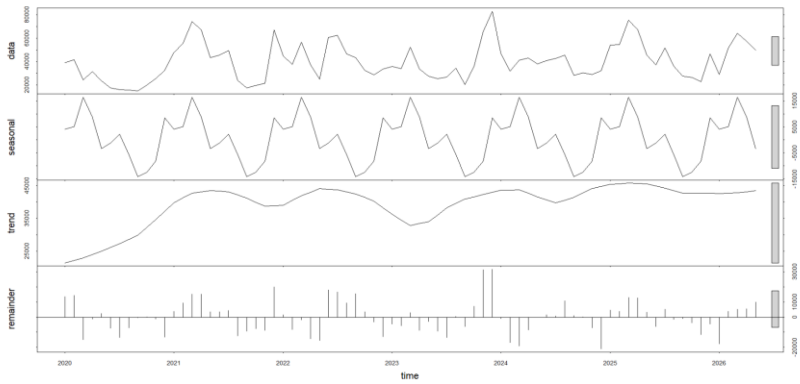


Figure 2. STL Decomposition

Based on figure 2, The seasonal component derived from the decomposition shows a stable annual cycle: positive seasonal effects drive up prices during specific periods, while negative effects pull prices down during other periods. Its core driving factors are the natural harvest cycle of red chili peppers, regional climatic conditions, and seasonal consumer demand. The trend component indicates that prices rose from roughly 23,000 IDR/kg in early 2020 to 44,000–45,000 IDR/kg at the end of the study period, marking a sustained long-term increase overall, with only a temporary decline recorded in 2023. Driving factors include rising production costs, inflationary pressure, and changes to regional market conditions. The remainder (residual) component fluctuates around the zero value with no fixed systematic pattern, which means the trend and seasonal components explain the vast majority of price fluctuations. Large residual values only emerged in a small number of periods, which correspond to irregular shocks such as sudden supply disruptions, extreme weather events, and abrupt demand shifts. The clear time series structure obtained from this STL decomposition provides solid support for follow-up research that uses the Prophet model, which is suited to capture trend and seasonal characteristics.

4.4. Prophet Forecasting Results

The Prophet model was successfully trained using 77 monthly observations. The resulting trend component shows a long-term upward price trend with several changepoints automatically detected by the algorithm (Taylor & Letham, 2018), reflecting structural changes in the dynamics of the red chili market. The annual seasonal component, modeled using Fourier series, captures a periodic pattern consistent with the findings of the STL decomposition (Cleveland et al., 1990)

ISSN : 2776-7922 (Print) / 2807-3037 (Online)

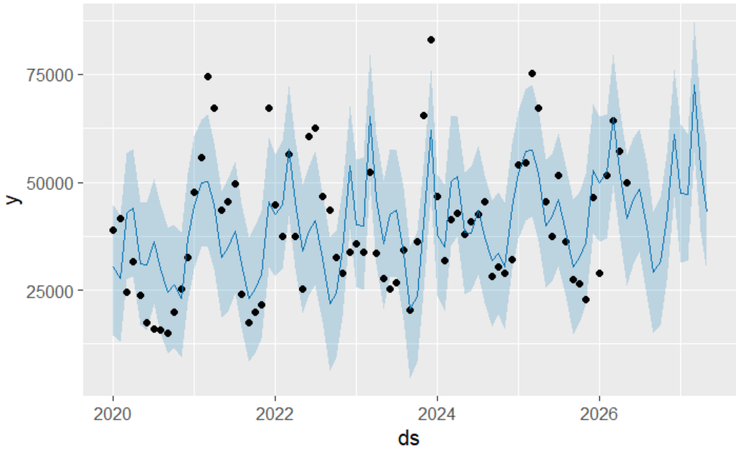


Figure 3. Prophet Forecasting Results

Figure 3 sequentially marks three core elements, namely the actual monthly observed prices of red chili peppers, the model's fitted values, and the 95% prediction interval, which align with the foundation of the sequence characteristic analysis completed in the early stage of this study using STL decomposition. It then evaluates the model's performance from three dimensions: the first dimension verifies fitting capacity, as the model can accurately capture the overall price trend and annual seasonal fluctuations; the second dimension verifies the reliability of the prediction interval, as the vast majority of actual observed values fall within the prediction interval output by the model, with a reasonable representation of uncertainty; the third dimension verifies the consistency of long-term patterns, as the core cyclical characteristics output by the model fully match the sequence rules extracted via STL decomposition. This study also finds that the model cannot fit the extreme price spikes of red chili peppers. This limitation conforms to the established domain logic that agricultural product prices are affected by three types of sudden shocks: abnormal weather, supply disruptions, and sudden demand shifts. Finally, this study concludes that the model is only suitable for capturing the long-term seasonal patterns of this price sequence, and is not applicable to forecasting short-term market shocks.

4.5. Forecast Accuracy Evaluation

Model evaluation was performed using Time Series Cross-Validation with 6 iterations. Table 2 presents the results of the evaluation metrics per prediction horizon:

Table 2. Prophet Model Cross-Validation Evaluation Results (6 Cutoffs)

Horizon	MSE	RMSE (IDR/kg)	MAE (IDR/kg)	MAPE (%)	Coverage
31 days	160006251	12649.36	10208.97	24.66	71.43%
55 days	171985787	13114.34	11045.34	26.91	71.43%
56 days	162190997	12735.42	10151.37	24.72	71.43%
58 days	168634995	12.985.95	11041.27	27.70	71.43%
60 days	148721873	12195.16	10169.76	22.15	71.43%
61 days	86687657	9310.62	7756.05	16.60	85.71%

This study adopts the time series cross-validation method, and sets 6 rolling prediction starting points to complete performance verification. The final quantitative evaluation results based on table 2 are as follows: RMSE ranges from 9310.62 to 13114.34 Indonesian rupiah per kilogram; MAE ranges from 7756.05 to 11045.34 Indonesian rupiah per kilogram; MAPE ranges from 16.60% to 27.70%; and the prediction interval coverage rate falls between 71.43% and 85.71%. Among all test intervals, the 61-day prediction interval has the lowest error, while the 55-day prediction interval has the highest error. The general industry accuracy benchmark defines MAPE <20% as good, and a MAPE of 20%-50% as acceptable. Measured against this standard, the MAPE of most prediction intervals in this study falls within the acceptable range. Combined with the conclusion drawn from the earlier STL decomposition conducted in this study, which found that the time series has strong annual seasonality and a stable long-term trend, the Prophet model can capture the core trend and seasonality of the price series. Errors only rise during periods of extreme price fluctuations, so the model's overall performance meets the research requirements.

Table 3. Average Forecast Accuracy Metrics

Metrics	Average	Unit
RMSE	11024.68	IDR/kg
MAE	9139.04	IDR/kg
MAPE	23.92%	%

The average RMSE value of IDR 11,024.68/kg and MAE of IDR 9,139.04/kg indicate an average prediction deviation of around IDR 9000–11000 from the actual price. The MAPE value of 23.92% is in the reasonable forecasting category based on Lewis's (1982) criteria. The high MAPE for red chili is a common phenomenon in the literature on horticultural commodity price forecasting, considering that prices are influenced by factors that are difficult to model quantitatively, such as pest attacks, natural disasters, and import policies (Hyndman & Koehler, 2006).

4.6. Forecast of Future Red Chili Prices

The final Prophet model was used to forecast monthly red chili prices for the subsequent twelve months.

Table 4. Forecasted Monthly Red Chili Prices

Month	Predicted (IDR/kg)	Lower Limit 80%	Upper Limit 80%
June 2026	46053.80	30955.60	60243.26
July 2026	48501.49	34016.72	62415.99
August 2026	39767.45	23894.91	54496.12
September 2026	29194.49	15135.12	43159.28
October 2026	31648.30	17034.16	46490.48
November 2026	41827.77	27137.83	56547.25
December 2026	61164.70	46689.73	76270.82
January 2027	47499.70	31382.32	63492.42
February 2027	47217.35	31758.37	61007.07
March 2027	72597.47	57061.64	87084.73
April 2027	53615.49	38641.69	68246.48
May 2027	43055.28	29054.59	57713.64

The forecast results (Table 4) show a fluctuation pattern consistent with the seasonal component identified by the Prophet model (Taylor & Letham, 2018). Prices are projected to reach their lowest point in September 2026 (Rp 29,194/kg), coinciding with the peak harvest period, which typically increases market supply (BPS South Sulawesi, 2026). A significant increase will occur in December 2026 (Rp 61,165/kg) due to year-end seasonality, and the highest price peak is predicted in March 2027 (Rp 72,597/kg).

The relatively wide 80% confidence interval (average ±14,000 IDR/kg) reflects the inherent uncertainty in forecasting highly volatile horticultural commodities (Hyndman & Koehler, 2006). This information is important for policymakers to prepare price stabilization mechanisms, such as market operations or distribution interventions, especially ahead of peak price periods.

5. Conclusion

This study successfully developed a monthly red chili price forecasting model in South Sulawesi using the Prophet algorithm with time series cross-validation evaluation. The following conclusions can be drawn:

(1). Monthly red chili price data in South Sulawesi for the 2020–2026 period shows high volatility (Rp 15,075–83,040/kg) with an average of Rp 39,847/kg.
(2). The Prophet model, configured with an annual seasonal component, successfully captured trend and seasonal patterns. Time series cross-validation yielded an average

OPEN ACCESS

ISSN : 2776-7922 (Print) / 2807-3037 (Online)

RMSE of IDR 11,024.68/kg, MAE of IDR 9,139.04/kg, and MAPE of 23.92%, which are considered reasonable forecasts based on Lewis's (1982) criteria.

(3). The June 2026–May 2027 price forecast estimates a lowest price in September 2026 (Rp 29,194/kg) and a peak in March 2027 (Rp 72,597/kg), consistent with the historical seasonal pattern of the red chili market in South Sulawesi.

(4). To improve accuracy, it is recommended to explore log-transformation and add external regressors (rainfall and CPI) as additional regressors in the Prophet model.

References

Box, G. E. P., & Jenkins, G. M. (1976). Time Series Analysis Forecasting and Control. Holden-Day.

Cleveland, R. B., Cleveland, W. S., McRae, J. E., & Terpenning, I. (1990). STL: A Seasonal-Trend Decomposition Procedure Based on Loess. Journal of Official Statistics, 6(1), 3–73.

Fadillah, M. R., Fahlapi, R., Solehah, A., Ningrum, N. W., Apriliani, S., & Marwati, N. (2025). Analisis Fluktuasi Harga Cabai Merah Keriting Periode Januari-Desember 2024 di Pasar Kramat Jati. Jurnal Perencanaan Dan Pengembangan Ekonomi, 8(2). https://journal.umgo.ac.id/index.php/JPPE

Food and Agriculture Organization. (2023). Food Outlook: Biannual Report on Global Food Markets.

Grolemund, G., & Wickham, H. (2011). Dates and Times Made Easy with lubridate. Journal of Statistical Software, 40(3). https://doi.org/10.18637/jss.v040.i03

Hyndman, R. J., & Athanasopoulos, G. (2021). Forecasting: Principles and Practice (3rd ed.). OTexts.

Hyndman, R. J., & Koehler, A. B. (2006). Another look at measures of forecast accuracy. International Journal of Forecasting, 22(4), 679–688. https://doi.org/10.1016/j.ijforecast.2006.03.001

Lewis, C. D., Green, B., & Sevenoaks. (1983). Industrial and business forecasting methods. Journal of Forecasting, 2(2), 194–196. https://doi.org/10.1002/for.3980020210

Taylor, S. J., & Letham, B. (2018). Forecasting at Scale. The American Statistician, 72(1), 37–45. https://doi.org/10.1080/00031305.2017.1380080

Wickham, H., François, R., Henry, L., Müller, K., & Vaughan, D. (2023). dplyr: A Grammar of Data Manipulation.

World Bank. (2024). Commodity Markets Outlook.

Zeileis, A., & Grothendieck, G. (2005). Zoo: S3 Infrastructure for Regular and Irregular Time Series. Journal of Statistical Software, 14(6). https://doi.org/10.18637/jss.v014.i06