Parameter: Journal of Statistics
Available at https://bestjournal.untad.ac.id/index.php/parameter
Vol. 4 No. 1 2024, 30 - 36
DOI: *https://doi.org/10.22487/27765660.2024.v4.i1.17120*
eISSN: 2776-5660

APPLICATION OF THE ARIMA METHOD IN FORECASTING THE PRICE RED CAYENNE PEPPER IN MAKASSAR CITY

Arwini Arisandi[1*], Ismail Gaffar[2], Andi Ridwan Makkulawu[3]

[1,2,3]Politeknik Pertanian Negeri Pangkajene Kepulauan

***e-mail**: [1]arwini.arisandi@polipangkep.ac.id

ABSTRACT

Red chili is one of the commodities with very tall cost changes. The cost variance of red chili can be caused by a huge amount of supply and request. The higher the amount of supply, the lower the cost, and the lower the amount of supply, the higher the cost. This study aims to implement the ARIMA method in forecasting red cayenne pepper prices in Makassar City. Data analysis to forecast red cayenne pepper prices used the ARIMA method with the results show that the price range of chili is from IDR 13,000 to IDR 80,000, with a mean value of IDR 38,218. The model with the minimum SSE and MSE value is ARIMA(1,1,1), so this model be used in time series data modeling for forecasting. The results of forecasting using the best model obtained a MAPE value of 15.90%, which is in the range of 10-20%, so it can be concluded that the ability of the ARIMA(1,1,1) model in forecasting the price of red cayenne pepper includes the good category.

Keywords: *ARIMA, forecast, MAPE, red cayenne pepper.*

Cite: Arisandi, A., Gaffar, I., & Makkulawu, R. A. (2024). *Application of the Arima Method in Forecasting the Price Red Cayenne Pepper in Makassar City*. *Parameter: Journal of Statistics*, *4*(1), 30-36, https://doi.org/10.22487/27765660.2024.v4.i1.17120.

[*]Received: Jun 2024; Reviewed: Jun 2024; Published: Jun 2024

INTRODUCTION

Cayenne pepper (*Capsicum frutescens*) is a fruit and a plant from the genus *Capsicum* whose fruit grows upwards. The fruits are size small green when not yet ripe and turn dark red when ripe. This type of chili is widely cultivated and used in dishes in various countries, including Indonesia. This chili is known for its strong spicy taste and diverse nutritional content. Chili is one of the leading national vegetable commodities with high adaptability and economic value. Chili is an agricultural strategic commodity that receives serious attention from the government and business actors due to its contribution to national economics (Wulandari, 2020). Red chili is one of the commodities with quite high price fluctuations. The price fluctuation of red chili can be caused by a large quantity of supply and demand. The higher the quantity of supply, the lower the price, and the lower the quantity of supply, the higher the price (Sukmawati et al., 2016).

Data from the Food Security Service of Makassar City show that from January to May 2024, the average price of red cayenne pepper in Makassar City is between IDR 37,000 and 47,000/kg. Data from 2023 show that the lowest price of red cayenne pepper in Makassar City occurs in September at IDR 15,000 and the highest price in December at IDR 74,000. This fluctuating price has led to several methods for forecasting chili prices being applied. Price changes will produce a data distribution that fluctuates over time, so this called a time series. Data that is not stationary over time and fluctuates constantly can be forecast through time series (Angelo et al., 2023).

Several methods for forecasting red cayenne pepper prices include Holt-Winters exponential smoothing, which was conducted by (Siregar et al., 2021) and obtained a MAPE forecasting result of 31%. Komaria et al. (2020) used the Fuzzy Time Series Model Lee with a MAPE result of 13.09% (Komaria et al., 2023). The forecast of red cayenne pepper used the Kalman Filter Algorithm with MAPE results of 34.3% in the Legi Market and 31.76% in the Sumoroto Market in forecasting the red cayenne pepper prices (Rohmah, 2024). Besides those methods, the integrated autoregressive moving average (ARIMA) method which provided a MAPE result of 3.84% with the ARIMA model (1,0,1) (Putra et al., 2021). The ARIMA model is effective in forecasting with minor forecasting error (Yasmin & Moniruzzaman, 2024). The ARIMA model is effectively used to forecast historical data and is popularly used to forecast time series data (Bagwan, 2024).

This study aims to implement the ARIMA method in forecasting red cayenne pepper prices in Makassar City. ARIMA is combination of several function autoregression (AR), integration (I), and moving averages (MA) which are able to forecast future values based on historical data (Abu et al., 2024). It used an iterative approach to identify the most appropriate model of all possible models, which can use all types of data (Buchori & Sukmono, 2018).

MATERIALS AND METHODS

This study used daily data on red cayenne pepper prices in Makassar City from July 14, 2022, to May 14, 2024. There were 671 observation data obtained from the Food Security Service of Makassar City. Data analysis to forecast red cayenne pepper prices used the ARIMA method with the assistance of Minitab software version 16. Analysis began by conducting descriptive statistics to determine the distribution of research data. Data processing was divided into two data sets: data for modeling (training data) of 640 data and testing data of 31 data (31 days) in the last one month. The testing data used to compare actual and predict value to get the ability forecast. This modeling data was used to establish the ARIMA model with the following analysis steps.

a. Time series model

This method is a non-stationary time series modeling. The stationary test was conducted by checking the stationarity of the variance and mean. Stationarity of variance can be found through the Box-Cox transformation. Non-stationarity data of mean can be subjected to a differencing process. A time series data $\{Y_t\}$ is stated to follow an autoregressive integrated moving average model if the d-th difference $Z_t = \nabla^d Y_t$ is a stationary ARMA process. If $\{Z_t\}$ follows an ARMA(p,q) model, we say that $\{Y_t\}$ is in the ARIMA(p,d,q) process with d=1 or at most 2. Stationary data then can be modeled with ARIMA (p,d,q) by the checking cut-off after lag in Autocorrelations Function (ACF) and Partial Autocorrelations Function (PACF) values with the conditions as in Table 1 below (Raftery, 2008).

Table 1. Characteristic of the ACF and PACF

	AR(p)	**MA(q)**	**ARMA(p,q), p>0 and q>0**
ACF	Tails off	Cuts off after lag q	Tails off
PACF	Cuts off after lag p	Tails off	Tails off

ARIMA(p,d,q) model will form several tentative models according to the cut-off values in ACF and PACF.

b. The models were selected as best based on the minimum criteria of Sum Square Error (SSE) and Mean Square Error (MSE) values. SSE and MSE values can be calculated through equation (1) and (2) (El-Azab et al., 2024).

$$\mathrm{SSE}=\sum_{t=1}^{n}(y_t-\hat{y}_t)^2 \tag{1}$$

$$\mathrm{MSE}=\frac{1}{n}\sum_{t=1}^{n}(y_t-\hat{y}_t)^2 \tag{2}$$

with y_t is the t-th actual data, $\hat{y}_t$ is the t-th forecast value (t=1,2,3,…,n).

The best ARIMA(p,d,q) model was then calculated with the forecast value $\hat{y}_t$, then compared with the actual value and testing data through the Mean Absolute Percentage Error (MAPE) value in equation (3) below.

$$\mathrm{MAPE}=\frac{1}{n}\sum_{t=1}^{n}\left|\frac{y_t-\hat{y}_t}{y_t}\right| \tag{3}$$

The forecasting ability can be measured through the range of MAPE values with MAPE ≤20% has good forecast ability (Arisandi & Hafid, 2024).

RESULTS AND DISCUSSION

The description of the research data begins with analyzing time series data distribution through time series plot. The distribution of research data is shown in Figure 1.

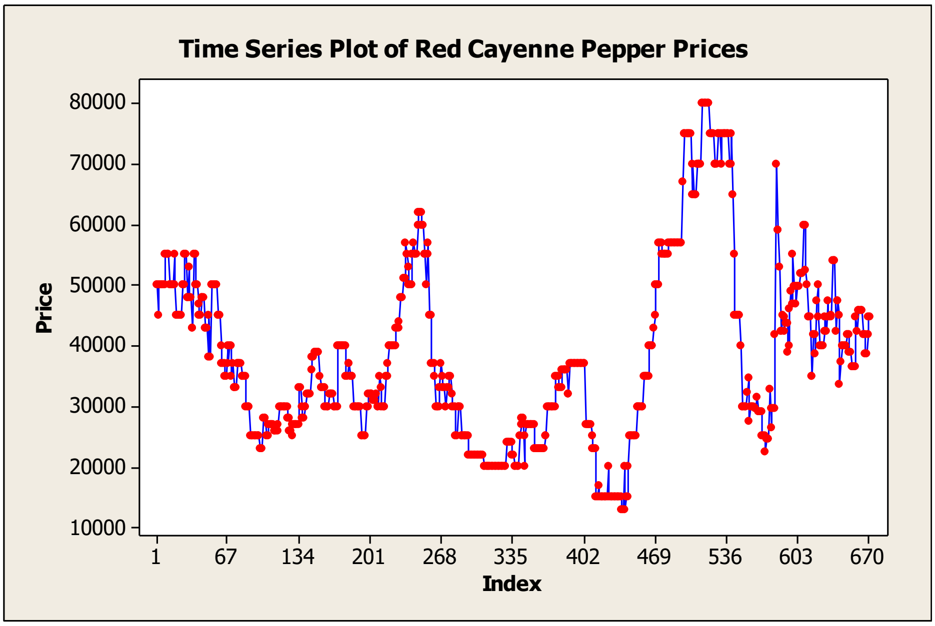


Figure 1. Description of red cayenne pepper prices

Figure 1 is a distribution of red cayenne pepper price data from 671 observations. The results show that the price range of chili is from IDR 13,000 to IDR 80,000, with a mean value of IDR 38,218. Furthermore, this data is separated by training and testing data. The ARIMA model was built with training data from 640 data (Figure 2) and testing data from 31 data to test the model.

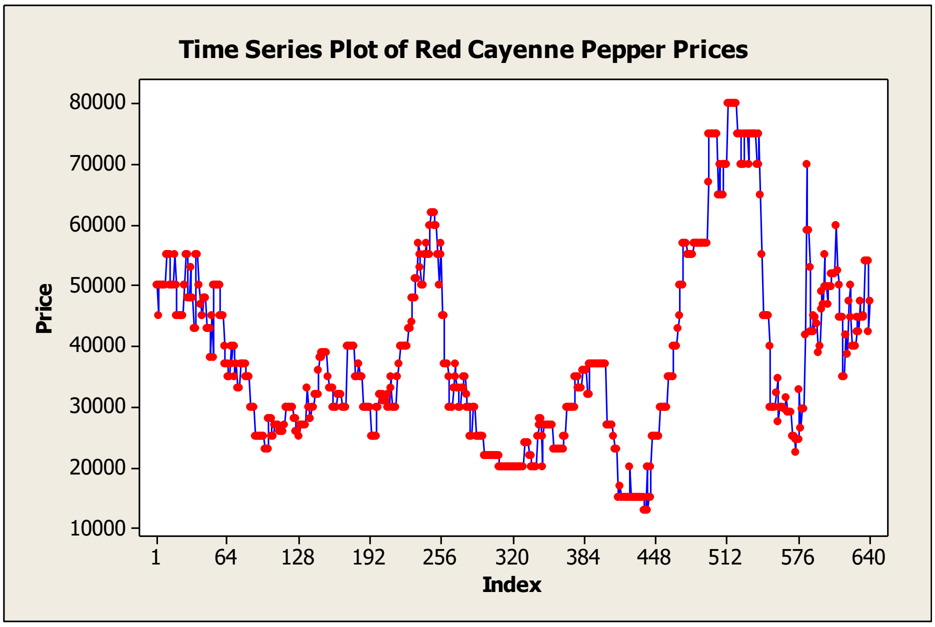
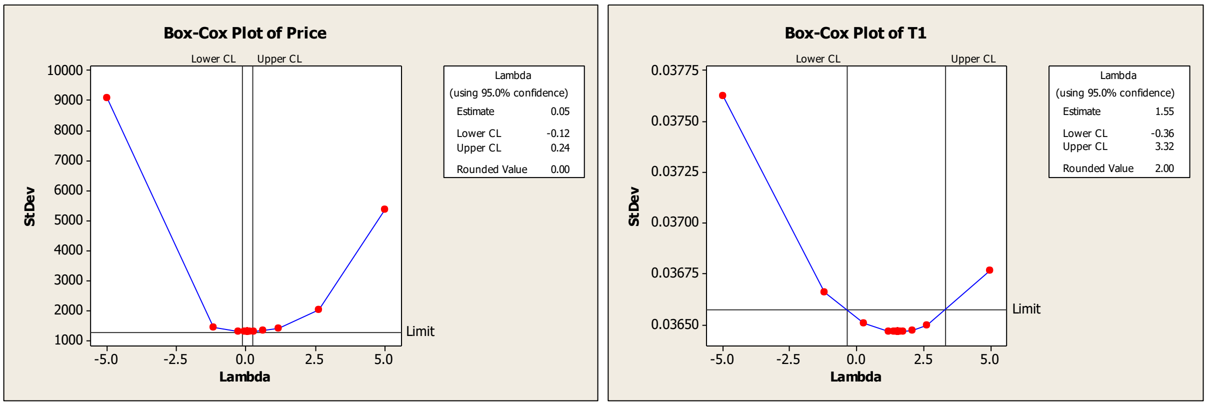


Figure 2. Time series plot of training data

The concept of stationarity can be practically described if a time series data plot has no changes in mean value over time (stationary to mean) and does not show a clear change of variance over time. Based on the time series plot in Figure 2, it can be found that the data is not yet stationary toward variance and mean, so the data requires Box-Cox transformation to overcome non-stationary toward variance and a differencing process to overcome non-stationary toward the mean.



(a) (b)

Figure 3. Box-Cox of price (a), Box-Cox of first transformation of price

Figure 3 is the Box-Cox transformation to overcome non-stationary variance. Assessment of stationary to variance is known through a rounded value of 1. The modeling data obtained a rounded value of 0.00 (Figure 3a), so the Box-Cox transformation still requires to be carried out. The results showed that the rounded value obtained was 2 in the first transformation (Figure 3b), and then the transformation was carried out again until the rounded value of 1 (Figure 4).

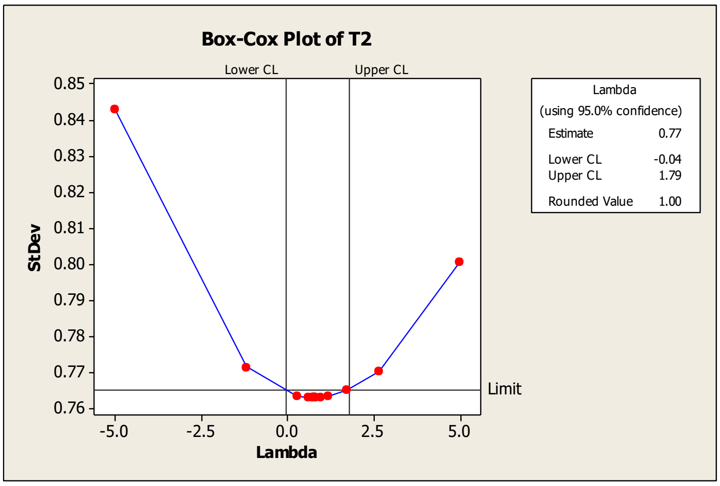


Figure 4. Box-Cox of second transformation

Figure 4 shows a rounding value of 1, which means the data is stationary regarding the variance so that a differentiation process is carried out if the data is not stationary regarding the mean. The results showed that the data were stationary, which means that the observation time did not affect the mean and variance of the data. This means that time series data fluctuates around a constant mean and variance, and it can be stated that time series data is stationary in mean and variance (Figure 5).

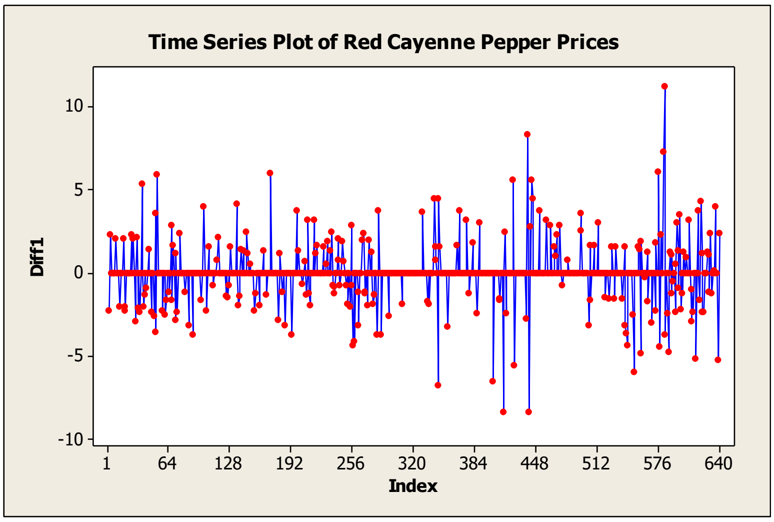


Figure 5. Differencing process

Stationary data was then modeled with ARIMA(p,d,q) modeling with a value of d=1 (1st differencing), so it will form the ARIMA(p,1,q) model. The determination of the p-th order in AR and q-th order in MA was based on the cut-off on the p-th lag of the PACF plot for the AR(p) model and the cut-off on the q-th lag of the ACF plot for the MA(q) model.

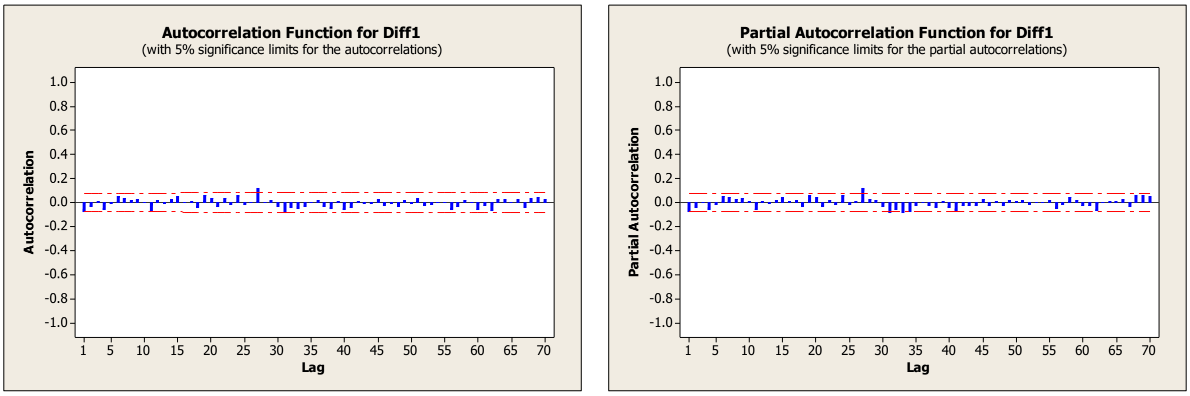


Figure 6. ACF and PACF plot

According to Figure 6, cut off after 1st lag occurs, so that the MA(1) model was obtained. In the PACF plot, it can also be seen that cut off after 1st lag occurs, so it can be concluded that the hat AR(1) model was obtained. The candidates for the ARIMA model that will be identified are ARIMA(1,1,0), ARIMA(0,1,1), and ARIMA(1,1,1) models. These model candidates will be identified and selected as the best model with the minimum SSE and MSE values.

Table 2. Tentative ARIMA Model for red pepper price

Model	SSE	MSE
ARIMA(1,1,0)	6368773027	9998074
ARIMA(0,1,1)	6368288240	9997313
ARIMA(1,1,1)	6354527269	9991395

Table 2 shows the model accuracy measured through SSE and MSE values. According to the three models, the model with the minimum SSE and MSE value is ARIMA(1,1,1), so the model formed is ARIMA(1,1,1), which will then be used in time series data modeling for forecasting for the next 31 periods.

Table 3. Forecast ARIMA model for red cayenne pepper prices

Period	Actual	Forecast	MAPE
641	47430	47466.1	0.08%
642	45000	47487.9	5.53%
643	33540	47500.6	41.62%
644	37420	47507.4	26.96%
645	40000	47510.5	18.78%
646	40000	47511.2	18.78%
647	40000	47510.3	18.78%
648	40000	47508.4	18.77%
649	40000	47506	18.77%
⋮	⋮	⋮	⋮
663	45830	47457.5	3.55%
664	45830	47453.9	3.54%
665	41830	47450.3	13.44%
666	41830	47446.7	13.43%
667	38730	47443.1	22.50%
668	38730	47439.5	22.49%
669	41830	47435.9	13.40%
670	44720	47432.3	6.07%
671	44720	47428.7	6.06%
	Mean		15.90%

ARIMA(1,1,1) model is used to forecast red cayenne pepper prices for periods 641 to 671 and compare the forecast values with actual values (Table 3). The results of the comparison are conducted MAPE calculation to measure forecasting accuracy. According to Table 3, a MAPE value of 15.90% was obtained. This is in the range of 10-20%, so it can be concluded that the ability of the ARIMA(1,1,1) model to forecast the price of red cayenne pepper includes a good category.

CONCLUSION

Forecasting the price of red cayenne pepper in Makassar City can be used using the ARIMA model. The best model is ARIMA(1,1,1) with a minimum SSE value of 6354527269 and minimum MSE value of 9991395 if compared with ARIMA(0,1,1) model and ARIMA(1,1,0) model. The results of forecasting using the best model obtained a MAPE value of 15.90%, which is in the range of ≤20%, so the researchers concluded that the ability of the ARIMA(1,1,1) model in forecasting the price of red cayenne pepper includes the good category.

REFERENCES

Abu, M., Siddique, B., Mahalder, B., Haque, M., Hossain, M., Chandra, J., Akhtar, S., & Ahammad, A. K. S. (2024). Heliyon Forecasting of tilapia (Oreochromis niloticus) production in Bangladesh using ARIMA model. *Heliyon*, *10*(5), e27111. https://doi.org/10.1016/j.heliyon.2024.e27111

Angelo, M. D., Fadhiilrahman, I., & Purnama, Y. (2023). Comparative Analysis of ARIMA and Prophet Algorithms in Bitcoin Price Forecasting. *Procedia Computer Science*, *227*, 490–499. https://doi.org/10.1016/j.procs.2023.10.550

Arisandi, A., & Hafid, H. (2024). Akurasi Model Prediksi Menggunakan Metode Automatic Clustering Fuzzy Time Series pada Indeks Harga Konsumen di Kota Makassar. *Journal of Mathematics: Theory and Applications*, *6*(1), 97–103. https://doi.org/10.31605/jomta.v6i1.3621

Bagwan, W. A. (2024). Electronic waste (E-waste) generation and management scenario of India , and ARIMA forecasting of E-waste processing capacity of Maharashtra state. *Waste Management Bulletin*, *1*(4), 41–51. https://doi.org/10.1016/j.wmb.2023.08.002

Buchori, M., & Sukmono, T. (2018). Peramalan Produksi Menggunakan Metode Autoregressive Integrated Moving Average (ARIMA) di PT. XYZ. *PROZIMA (Productivity, Optimization and Manufacturing System Engineering)*, *2*(1), 27–33. https://doi.org/10.21070/prozima.v2i1.1290

El-Azab, H. A. I., Swief, R. A., El-Amary, N. H., & Temraz, H. K. (2024). Machine and deep learning approaches for forecasting electricity price and energy load assessment on real datasets. *Ain Shams Engineering Journal*, *15*(4), 102613. https://doi.org/10.1016/J.ASEJ.2023.102613

Komaria, V., Maidah, N. El, & Furqon, M. A. (2023). Prediksi Harga Cabai Rawit di Provinsi Jawa Timur Menggunakan Metode Fuzzy Time Series Model Lee. *Komputika : Jurnal Sistem Komputer*, *12*(2), 37–47. https://doi.org/10.34010/komputika.v12i2.10644

Putra, J. E., Jamhari, & Perwitasari, H. (2021). *Pemrakiraan Harga Cabai Rawit di Kota Surabaya*.

Raftery, A. E. (2008). Time series analysis with R. In *European Journal of Operational Research* (Vol. 20, Issue 2). https://doi.org/10.1016/0377-2217(85)90052-9

Rohmah, N. A. (2024). *Prediksi harga cabai di kabupaten ponorogo menggunakan algoritma kalman filter*. *01*(01).

Siregar, Y. J., Hartono, R., Hardana, A. E., Agribisnis, P. S., Brawijaya, U., & Smoothing, E. (2021). *Peramalan harga cabai rawit di kota malang dengan metode holt-winters exponential smoothing*. *6*, 99–110.

Sukmawati, D., Sulistyowati, L., Karmana, maman H., & Wikarta, E. K. (2016). Fluktuasi Harga Cabai Merah Keriting (Capsicum annum L) DI SENTRA PRODUKSI DAN PASAR INDUK. *Mimbar Agribisnis*, 25–26.

Wulandari, S. A. (2020). Fluktuasi Harga Cabai Merah Di Masa Pandemi Covid 19 Di Kota Jambi. *Jurnal MeA (Media Agribisnis)*, *5*(2), 112. https://doi.org/10.33087/mea.v5i2.82

Yasmin, S., & Moniruzzaman, M. (2024). Forecasting of area, production, and yield of jute in Bangladesh using Box-Jenkins ARIMA model. *Journal of Agriculture and Food Research*, *16*, 101203. https://doi.org/10.1016/J.JAFR.2024.101203