Performance Evaluation of ARIMA and GRU Models for Forecasting Chili Price in East Java

Windi Pangesti[1*], Nabila Syukri[2], Khairil Anwar Notodiputro[3], Yenni Angraini[4], Laily Nissa Atul Mualifah[5]

[1,2,3,4,5]*Statistics and Data Science, IPB University, Indonesia*

*corr-author: Windi_pangesti@apps.ipb.ac.id

Abstract – Time series forecasting plays a crucial role in predicting future conditions based on historical data, particularly in the food sector, which is highly susceptible to price fluctuations. This study compares two approaches: the conventional ARIMA method and the deep learning method GRU, to forecast the price of red chillies in East Java. East Java was chosen because it is the largest national producer of chilies, thus the stability of its prices has a broad impact. The research results indicate that the GRU model outperforms the ARIMA model with a MAPE value of 19.80% compared to a MAPE of 27.63% for the latter. The benefit of this research is to contribute to the literature on developing agricultural commodity price forecasting models as a basis for enhancing food security policies and stabilizing commodity prices, particularly in East Java Province, Indonesia.

Keywords: ARIMA, GRU, Time Series, Chili Price

I. INTRODUCTION

Time series data is a sequence of observations recorded over a specific period such as hourly, daily, weekly, monthly, or yearly [1]. In the analysis of time series data, forecasting is one of the research focuses and has always been a concern. In addition to gaining insights from data, predicting future events based on historical data patterns and trends is no less important [2]. In general, there are two approaches to forecasting, namely conventional methods and machine learning-based methods. Conventional methods require the fulfillment of several statistical assumptions before the model is used for prediction. Meanwhile, the machine learning method applies algorithms to recognize patterns within data automatically by studying historical data, without the need to rely on certain assumptions [3].

One of the conventional methods that is commonly used in time series analysis is ARIMA (Autoregressive Integrated Moving Average). ARIMA is a statistical model that combines three main components, namely Autoregressive (AR), differentiating (I), and Moving Average (MA) to capture patterns in univariate time series data. The ARIMA model is quite effective in forecasting and is able to explain the structure of the model explicitly and interpretively [4]. However, the weakness of the ARIMA model in its use is that it is not able to handle non-linear or non-stationary properties, making it less effective in prediction, especially for the accurate long-term [5].

Therefore, machine learning-based approaches, especially deep learning, are considered to be able to overcome the complexity of time series data patterns that cannot be handled well by conventional models such as ARIMA [6]. Deep learning works by automatically recognizing patterns in data through learning from historical data, without the need for explicit modeling [7]. Deep learning is a new approach that can derive representations of original inputs by combining simple but non-linear models. Each neural network module converts one representation level into a more abstract level [8]. One of the deep learning models is GRU (Gated Recurrent Unit) which is a modified version of the Recurrent Neural Network (RNN) architecture that can retain both long-term and short-term information from data. The structure of GRU is less complex compared to Long Short-Term Memory (LSTM), but it is still able to retain important information over the long term [9].

Several studies have compared the performance of conventional methods in predictive accuracy levels such as those conducted by [10] on the bitcoin price data which concludes that the ARIMA model offers better predictions than the GRU model, and the GRU is better than the LSTM. Ref [11] compared the ARIMA and GRU models on minute-by-minute stock price data on HIMBARA banks by concluding that the GRU outperformed the ARIMA model for all types of banks analyzed. Furthermore, research conducted by [12] in predicting bitcoin price movements with the results showing that GRU attains greater accuracy in long-term forecasting, while ARIMA is optimal for short-term predictions. Based on the finding of several studies, model performance is highly sensitive by the characteristics of the analyzed data. Consequently, the choice of forecasting method should be tailored to the specific nature of the dataset and the objectives of the analysis.

Accurate forecasting results are needed in decision-making and future planning, especially in the management of the agricultural sector and the stability of food commodity prices [13]. The price of red chili is one of the strategic commodities in Indonesia and is a contributor to inflation due to high fluctuations. Several factors cause high volatility such as weather, the month of Ramadan or other significant days making the price of chili difficult to predict with a simple pattern [14]. By using the right forecasting method, policymakers are expected to be able to take more responsive policies such as crop distribution, ideal planting time, and price intervention when there is a surge.

Based on data from the Central Statistics Agency (BPS), there are 7 provinces that serve as chili production centers, collectively contributing 86.37% to the total national chili production. East Java Province is the largest producer, accounting for 37.35% of Indonesia's total chili production in 2023 [15]. So the stability of chili prices in this region is very important to analyze because of the influence on national market conditions. Therefore, a forecasting approach is needed that can capture fluctuating price patterns based on available historical data. In this study, the ARIMA and GRU methods were chosen to represent conventional and machine learning approaches, aiming to determine which model offers better performance in forecasting chili prices in East Java Province.

II. METHOD

A. Data

The data utilized in this research is secondary data, specifically red chili price data in East Java obtained from the Information System on Availability and Price Development of Basic Materials in East Java (SISKAPERBAPO). It consists of monthly observations for January 2014 to December 2024 which amounts to 132 observed data. The data is divided into training and test data The training data is used to construct the forecasting model, while the test data is used to generate forecasts and assess the model's accuracy by comparing the predicted values to actual observations.

B. Analysis Procedure

- Exploring the data descriptively by plotting the data to learn its characteristics.
- Dividing the data into two parts, namely training data and test data. Data division is done by considering the data's historical pattern.
- ARIMA Model Procedure
  - Checking data stationarity visually using the Autocorrelation Function (ACF) plot and confirming it with the Augmented Dickey-Fuller (ADF) test. Stationarity is generally classified into two types: stationarity in mean and stationarity in variance. on-stationary mean can be addressed through differencing [2], while non-stationary variance can be handled using a Box-Cox transformation.
  - Determine the order p,d,q for the ARIMA model seen based on the ACF and PACF plots of the stationary chili price data. Then determine several tentative models with the appropriate combination of AR and MA orders. Table I is used to determine the initial estimate of parameters based on the ACF and PACF plots [16].
  - Perform parameter estimation and significance testing on the tentative ARIMA model obtained at the model identification stage.
  - Perform diagnostic tests using the Ljung-Box test to check white noise, Breusch Pagan test to test homogeneity, and Shapiro-Wilk test to check the assumption of normality.
  - Determine the best ARIMA model using the Akaike Information Criterion (AIC) of several tentative models. Models with smaller AIC values are considered better since they keep a good balance between model fit and complexity. The AIC formula can be seen in (1).

$$AIC = 2k - 2\ln(L) \tag{1}$$

  Where k is the number of parameters, and L refers to the maximum value of the likelihood function [17].
  - Predict the test data using the best ARIMA model.
  - Evaluate the goodness of the ARIMA model by calculating the MAPE, MAE and RMSE, as defined in (2)-(4). These metrics are calculated using the following formulas:

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_t-\hat{y}_t}{y_t}\right| \times 100\% \tag{2}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_t - \hat{y}_t| \tag{3}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_t - \hat{y}_t)^2} \tag{4}$$

  The smaller value of MAPE, MAE, or RMSE indicates better predictive accuracy and model performance [18].
- GRU Model Procedure

- Changing the data structure to supervised learning, where y_{t-1} data is the predictor variable, and y_t is the target variable.
- Data transformation by rescaling the data to a scale of 0 to 1 using MinMaxScaler from scikitlearn. This normalization process is defined by the formula in (5):

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)} \quad (5)$$

Normalization is necessary because the sigmoid activation function used in the GRU network is highly sensitive to input scaling [11]. Without normalization, the model may have convergence difficulties or get stuck on a very small gradient area, thus slowing down the training process.

- Changing the data dimension from two-dimensional (number of observations and number of variables) to three-dimensional (number of observations, timestep, and number of variables) is adjusted to the needs of the GRU algorithm which uses three-dimensional input.
- Hyperparameter tuning is done to find the best hyperparameter combination that produces optimal model performance with the number of combinations tested is 36. Units are the number of units (neurons) in the GRU layer. Epochs refers to the number of times the entire dataset is traversed by the model during the training process. Dropout is the percentage of neurons that are randomly turned off during the training process to prevent overfitting. Learning rate is a measure of how much updating is done to the model weights after each training step, the smaller the more stable but slower. Table II shows the combination of hyperparameter GRU:

To obtain the best hyperparameter combination, a cross-validation process was performed by dividing the data into five groups using the Walk Forward Validation method, which is presented in Table III. In each group, the data is divided into two subsets: a training set (48 months) and a test set (12 months).

- After the model is trained, the prediction process is performed on the test data using the best hyperparameter combination to estimate future chili prices. The prediction results and actual data, previously normalized, are transformed back to their original scale using an inverse transform, allowing the predictions to be interpreted directly in actual price units.
- Model evaluation with the MAPE, MAE and RMSE metric, as defined in (2)-(4).
  Evaluate the performance of the ARIMA and GRU models by interpreting visual comparisons between actual and predicted data, and by comparing their MAPE, MAE, and RMSE values. A flowchart of the research workflow is provided in Fig. 1.

TABLE I
MODEL ESTIMATION BASED ON ACF AND PACF PLOTS

Model	**ACF**	**PACF**
AR(p)	Tails off	Cut off after-lag p
MA(q)	Cut off after-lag q	Tails off
ARMA(p, q)	Tails off	Tails off

TABEL II
HYPERPARAMETER GRU

Hyperparameter	Value
Number of layers	1
Dropout	0.15
Number of neurons	[10, 16, 32, 50, 64, 128]
Learning rate	[0.01, 0.001]
Epochs	[100, 200, 300]

TABEL III
SPLITTING DATA FOR CROSS VALIDATION

Group	**Period**	**Splitting data**
1	January 2014 – December 2017	Data Training
	January 2018 – December 2018	Data Testing
2	January 2015 – December 2018	Data Training
	January 2019 – December 2019	Data Testing
3	January 2016 – December 2019	Data Training
	January 2020 – December 2020	Data Testing
4	January 2017 – December 2020	Data Training
	January 2021 – December 2021	Data Testing
5	January 2018 – December 2021	Data Training
	January 2012 – December 2022	Data Testing

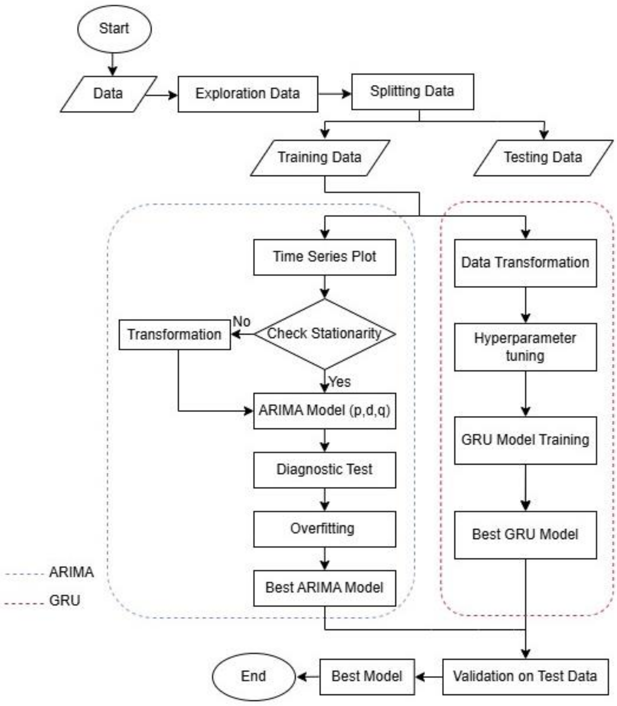


Fig. 1 Flowchart of the research stages

III. RESULT AND DISCUSSION

A. Data exploration

Data exploration was carried out to analyze and identify the characteristics and patterns of the red chili prices data in East Java from January 2014 to December 2024. A visualization of the monthly chili price data is presented in Fig. 2.

Fig. 2 shows that the highest chili price occurred in February 2017, at Rp129,068, while the lowest chili price occurred in June 2014, at Rp11,657. Based on the pattern in the plot, it can be seen that the data does not show any seasonal patterns or consistent upward or downward trends. In general, chili prices tend to fluctuate around a stable mean throughout the observation period, suggesting that the data is stationary in the mean. Furthermore, for analysis purposes the data is divided into two parts, the data division is done by considering the historical pattern of the chili price data shown in Fig. 3.

The training data includes 108 monthly observations from January 2014 to December 2022 to build the model, while the test data consists of 24 monthly observations from January 2023 to December 2024. This split allows for evaluating the performance and generalization ability of the model in predicting chili prices in East Java.

B. ARIMA Model

Stationarity is a fundamental requirement in time series analysis, as most time series models, including ARIMA, assume a stable mean and variance over time. The monthly red chili price time series plot in Fig. 2 suggests that the data is not stationary in the variance, indicating the need for variance-stabilizing transformation before proceeding with ARIMA modelling. Before performing the transformation, it is necessary to determine the optimal lambda value using a Box-Cox plot. The corresponding Box-Cox plot used to assess the appropriate lambda shown in Fig. 4.

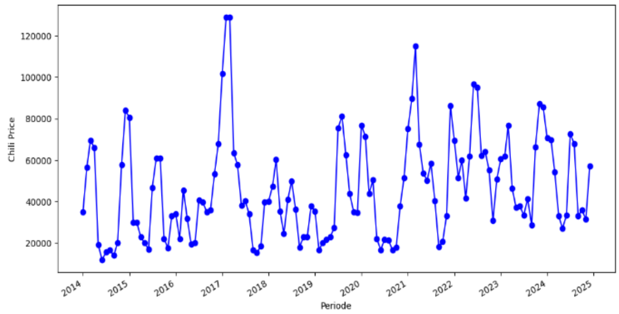


Fig. 2 Monthly trend of red chili price

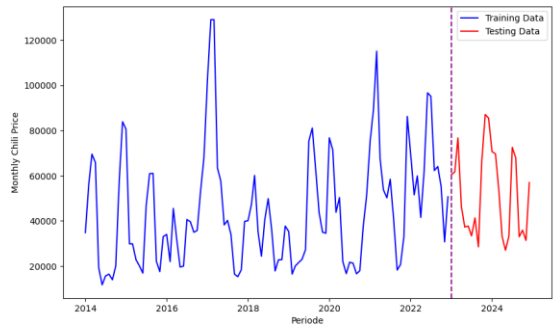
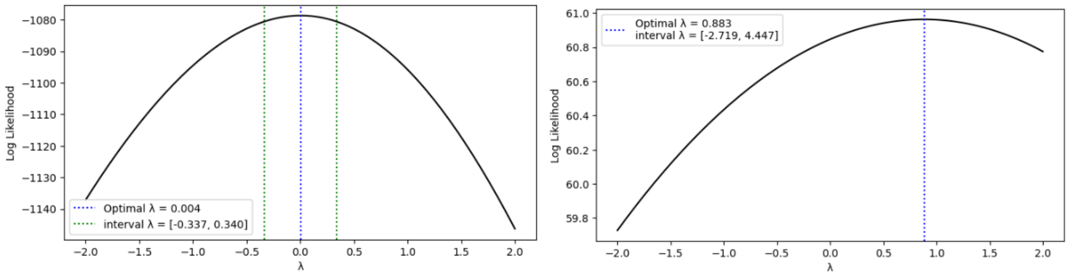


Fig. 3 Time series plot of training data and test data

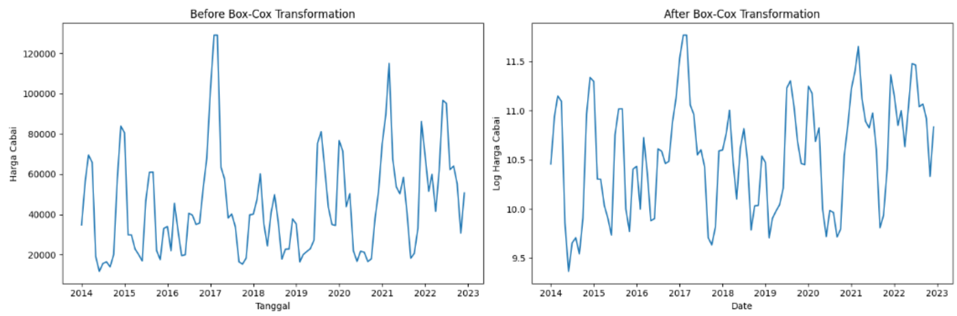


(a) (b)

Fig. 4 (a) Box-Cox plot before transformation; (b) Box-Cox plot after transformation

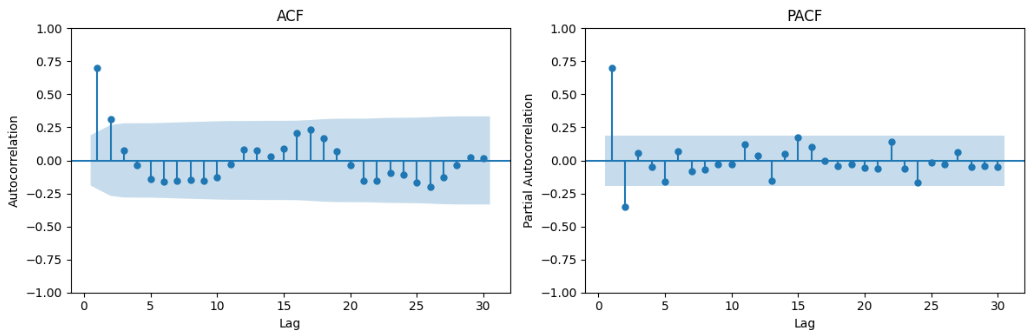
Fig. 4(a) shows the Box-Cox plot before transformation, which yields an optimal lambda value of 0.004, or approximately zero, suggesting that the data remains non-stationary in variance and that a logarithmic transformation is appropriate for stabilizing it. Therefore, a log transformation was applied to the monthly chili price training data. Fig. 4(b) shows the Box-Cox plot after this transformation, where the estimated lambda value shifts to 0.883, close to 1.0, suggesting that variance stationarity was achieved after a single logarithmic transformation. The time series plots before and after the Box-Cox transformation are shown in Fig. 5.

Fig. 5(b) shows that the time series plot after applying the Box-Cox transformation, is stationary in both variance and mean as the data fluctuates around a relatively constant in mean. This is evidenced by the ADF test results which show a p-value of 0.0000001, well below the significance level α = 0.05. Therefore, it can be inferred that the data is stationary in the mean. After verifying that the data is stationary in both mean and variance, no differencing was required ($d = 0$). The next step is to determine the appropriate order of the ARIMA model. This is done by analyzing the ACF and PACF plots generated from the stationary data, which help identify the potential values for the AR and MA order.



(a) (b)

Fig. 5 (a) Time series plot before transformation; (b) Time series plot after transformation



(a) (b)

Fig. 6 (a) Plot ACF; (b) Plot PACF

TABEL IV
MODEL PARAMETER ESTIMATION

Model	Parameter	Coefficient	P-value	AIC
ARIMA (0,0,2)	Const*	10.5565	0.000	107.821
	MA(1)*	0.9660	0.000	
	MA(2)*	0.4583	0.000	
ARIMA (2,0,0)	Const*	10.5531	0.000	109.232
	AR(1)*	0.9258	0.000	
	AR(2)*	-0.3363	0.004	
ARIMA (2,0,2)	Const*	10.5499	0.000	112.402
	AR(1)*	1.3331	0.000	
	AR(2)*	-0.5013	0.003	
	MA(1)	-0.4079	0.259	
	MA(2)	0.1471	0.284	

* Parameters are significant at 5% significance level

Based on Fig. 6, it can be observed that both the ACF and PACF plots cut off at lag 2. This pattern indicates that the tentative models that may be suitable are ARIMA (0,0,2), ARIMA (2,0,0), and ARIMA (2,0,2). Furthermore, parameter estimation of the three models was conducted to determine the most optimal model. The parameter estimation results of each candidate model are presented in Table IV.

Table IV shows the estimation results and significance test for the parameters derived from the ARIMA modeling. Among the models evaluated, ARIMA (0,0,2) and ARIMA (2,0,0) have all parameters statistically significant at the 5% level. Comparing the models using the AIC value, the ARIMA (0,0,2) model has the lowest AIC value, suggesting it is the most suitable among the three options evaluated. Following the selection of the optimal model, diagnostic tests are conducted to assess its adequacy, with results summarized in Table V.

Furthermore, the overfitting model was checked by trying the ARIMA (1,0,2) and ARIMA (0,0,3) models. However, the AR (1) parameter in the ARIMA (1,0,2) model is not significant, with an AIC value = 109.780, then the MA (3) parameter in the ARIMA (0,0,3) model is not significant, with an AIC value = 109.802. Both model AIC values are still greater than the AIC value of the ARIMA (0,0,2) model. Thus, it is concluded that the ARIMA (0,0,2) is the best model. The ARIMA (0,0,2) model equation can be written as follows in (6):

$$\hat{Y}_t = 10.5565 + 0.9660\varepsilon_{t-1} + 0.4583\varepsilon_{t-2} + \varepsilon_t \quad (6)$$

Based on (6), the constant value of 10.5565 represents the long-term mean value of the predicted time series. The MA coefficient at lag 1 of 0.9660 and lag 2 of 0.4583 indicates that price fluctuations in the previous one and two months have a positive influence on the current price, with the influence of the previous one month being greater. In the absence of autoregressive and differencing components, this prediction model forecasts chili prices based solely on the residual pattern of the previous period.

C. GRU Model

GRU modeling does not require data stationarity assumptions like ARIMA models. However, before GRU modeling is performed, the data is normalized in the range (0,1) with minmaxscaler. By performing data normalization, the data is easier to train the model and can improve accuracy. Fig. 7 is a plot of chili price data after normalization.

TABEL V
DIAGNOSTIC TEST RESULTS

Statistic Test	P-value	Decision
Ljung-Box	0.8299	Residual *white noise*
Breush-Pagan	0.5618	Constant residual variance
Shapiro-Wilk	0.3949	Residual normal

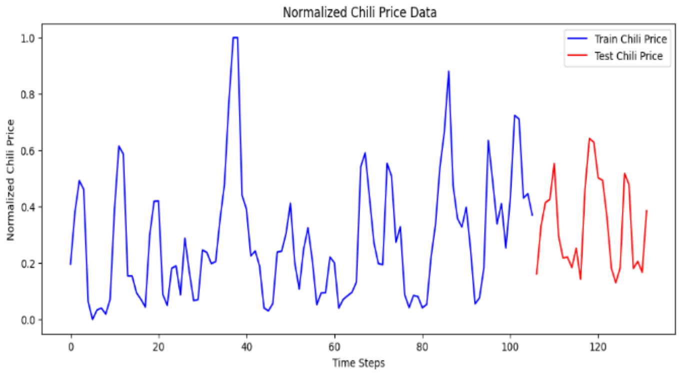


Fig. 7 Normalization plot of chili price

Furthermore, the data is converted into a three-dimensional form, namely the sample size, timestep and number of features to match the input required by the GRU model. In this study, timestep = 1 is used, meaning the model will learn patterns based on the previous 1 month. There is one feature used, namely the price of chili in the previous month $y_{(t-1)}$. Therefore, the form of the training data becomes (107, 1, 1) which means there are 107 samples, each with 1 timestep and 1 feature. While the test data has dimensions (24, 1, 1).

GRU modeling is built using several combinations of hyperparameters that have gone through a tuning process. The hyperparameter tuning process is important in deep learning modeling because it affects the performance of the model in prediction. In this study, several Hyperparameters were adjusted to find the combination that produces optimal results. Hyperparameters include the number of units (neurons), epochs, dropouts and learning rates. During model training, the Adam optimizer was used with a fixed learning rate selected through tuning to achieve fast and stable convergence. The optimal hyperparameter combination was determined based on the lowest MSE value, which served as the primary performance metric. The best combination of hyperparam eters obtained through this tuning process using cross-validation is shown in Table VI.

All hyperparameter combinations listed in Table II were tuned to identify the best GRU model. The best GRU model selected has 128 neurons, meaning the model built has 128 neurons or units in the hidden layer, epochs of 200 mean the dataset is trained 200 times during the training process, a dropout rate of 0.15 means 15% of neurons will be randomly deactivated at each iteration during the training process, this aims to avoid overfitting, and a learning rate of 0.001 means the model updates the weights with a step of 0.001 based on how big the prediction error is.

To prevent overfitting and ensure effective generalization of the GRU model, a validation process is implemented during training. The loss value of training and validation shows in Fig. 8.

Fig. 8 shows that the loss values on the training and validation data together decrease consistently until around the 50th epoch. After that, the loss values tend to be more stable and do not show any gap between the training and validation losses. Validation loss is lower than training loss indicating that the model has good generalization performance to the testing data and there is no indication of overfitting.

TABEL VI
THE BEST HYPERPARAMETER COMBINATION OF THE GRU MODEL

Hyperparameter	Value
Number of layers	1
Dropout	0.15
Number of neurons	128
Learning rate	0.001
Epochs	200

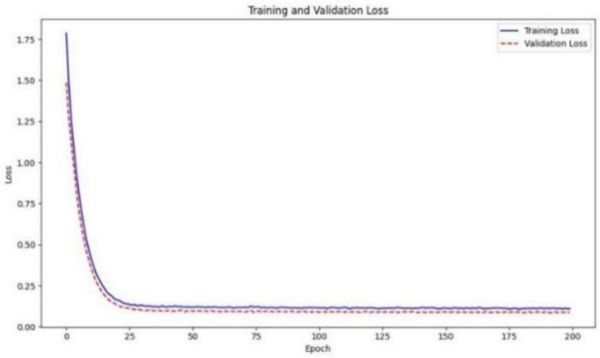


Fig. 8 Training vs validation loss plot

D. Evaluation of ARIMA and GRU Models

The predicted values of the ARIMA and GRU models on the testing data are evaluated against the actual data visualized with a comparison plot of predictions from both models in Fig. 9. The blue line shows the actual values, while the green and red lines demonstrate the prediction results from the ARIMA and GRU models, respectively. From this visualization, it is evident that the GRU model is better able to follow chili price fluctuations than the ARIMA model, especially when there are sharp price spikes and drops. However, the GRU model is still underestimated against the actual data. Meanwhile, the ARIMA model tends to produce smoother predictions and is less responsive to extreme changes. This is because ARIMA is unable to capture long-term patterns so it tends to respond slowly to sharp spikes or drops in actual data.

Model evaluation is conducted using several metrics, including MAPE, MAE and RMSE. MAPE is used to determine the level of prediction error by expressing the difference between predicted and actual values in percentage form to make it easier to understand in interpretation and compare performance between models. Meanwhile, MAE and RMSE provide complementary insights by showing the magnitude of average errors in absolute terms. Table VII presents the metrics value of each model.

Based on the evaluation results in Table VII, the GRU model demonstrates consistently better performance than the ARIMA model for all evaluated metrics. The GRU model achieves a lower MAPE value (19.80%) compared to the ARIMA model (27.63%), along with lower MAE (12152.31) and RMSE (15867.25) values, indicating greater predictive accuracy and smaller forecast errors. These results confirm that the GRU model is more effective at capturing the identifying patterns in chili price data and provides superior forecasting performance compared to the ARIMA model.

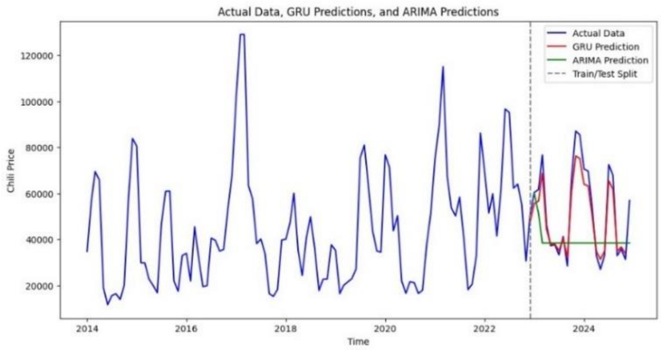


Fig. 9 Test data prediction vs actual plot

TABEL VII
MODEL EVALUATION

Model	MAPE	MAE	RMSE
ARIMA (0,0,2)	27.63%	16612.06	22385.03
GRU (*unit* 128, *epoch* 200, *dropout* 0.15, *learning rate* 0.001)	19.80%	12152.31	15867.25

IV. CONCLUSION

Based on the analysis results, the GRU model outperforms the ARIMA model in forecasting chili prices in East Java Province. This is supported by the lower MAPE value of 19.80% for the GRU model, compared to 27.63% for the ARIMA model, as well as lower MAE and RMSE values. The best GRU model has a hyperparameter configuration of 128 units (neurons), 200 epochs, a dropout rate of 0.15, and learning rate of 0.001. The ability of GRU to capture nonlinear patterns in data allows this model to provide higher accuracy, especially in long-term forecasting. In contrast, the ARIMA model tends to be more optimal for short-term forecasting. Therefore, the GRU model is recommended over the ARIMA model for forecasting red chili prices in East Java Province. As a suggestion for future research, it would be valuable to explore other deep learning models, including LSTM or Transformer, and incorporate additional external variables to enhance the accuracy of the forecasting.

REFERENCES

[1] G. Ghaly Ghiffary, E. Dicky Darmawan Yanuari, K. Anwar Notodiputro, Y. Angraini, and L. Nissa Atul Mualifah, "COMPARATIVE PERFORMANCE OF SARIMAX AND LSTM MODEL IN PREDICTING IMPORT QUANTITIES OF MILK, BUTTER, AND EGGS "COMPARATIVE PERFORMANCE OF SARIMAX AND LSTM MODEL IN PREDICTING IMPORT QUANTITIES OF MILK COMPARATIVE PERFORMANCE OF SARIMAX AND LSTM MODEL IN PREDICTING…," *BAREKENG: J. Math. & App*, vol. 19, no. 1, pp. 407–0418, 2025, doi: 10.30598/barekengvol19iss1pp0407-0418.

[2] K. A. Notodiputro, Y. Angraini, and L. N. A. Mualifah, *Analisis Data Deret Waktu dengan Python* . Bogor: IPB Press, 2025.

[3] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*, 2nd ed. New York, 2009.

[4] D. Peter and P. Silvia, "ARIMA vs ARIMAX-which approach is better to analyze and forecast macroeconomic time series?," in *Proceedings of 3oth International Conference Mathematical Methods in Economic*, Sep. 2012, pp. 136–140.

[5] W. A. Pratiwi, A. F. Rizki, K. A. Notodiputro, Y. Angraini, and L. N. Mualifah, "THE COMPARISON OF ARIMA AND RNN FOR FORECASTING GOLD FUTURES CLOSING PRICES," *BAREKENG: J. Math. & App*, vol. 19, no. 1, pp. 397–0406, 2025, doi: 10.30598/barekengvol19iss1pp0397-0406.

[6] X. Li, X. Ma, F. Xiao, C. Xiao, F. Wang, and S. Zhang, "Time-series production forecasting method based on the integration of Bidirectional Gated Recurrent Unit (Bi-GRU) network and Sparrow Search Algorithm (SSA)," *J Pet Sci Eng*, vol. 208, no. A, pp. 1–13, Jan. 2022, doi: 10.1016/j.petrol.2021.109309.

[7] P. L. Seabe, C. R. B. Moutsinga, and E. Pindza, "Forecasting Cryptocurrency Prices Using LSTM, GRU, and Bi-Directional LSTM: A Deep Learning Approach," *Fractal and Fractional*, vol. 7, no. 2, Feb. 2023, doi: 10.3390/fractalfract7020203.

[8] Z. Shen, Y. Zhang, J. Lu, J. Xu, and G. Xiao, "A novel time series forecasting model with deep learning," *Neurocomputing*, vol. 396, pp. 302–313, Jul. 2020, doi: https://doi.org/10.1016/j.neucom.2018.12.084.

[9] A. P. Meriani and A. Rahmatulloh, "PERBANDINGAN GATED RECURRENT UNIT (GRU) DAN ALGORITMA LONG SHORT TERM MEMORY (LSTM) LINEAR REGRESSION DALAM PREDIKSI HARGA EMAS MENGGUNAKAN MODEL TIME SERIES," *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 12, no. 1, Jan. 2024, doi: 10.23960/jitet.v12i1.3808.

[10] P. T. Yamak, L. Yujian, and P. K. Gadosey, "A Comparison between ARIMA, LSTM, and GRU for Time Series Forecasting," in *ACAI '19: Proceedings of the 2019 2nd International Conference on Algorithms, Computing and Artificial Intelligence*, Feb. 2020, pp. 49–55. doi: https://doi.org/10.1145/3377713.3377722.

[11] M. Ridwan, K. Sadik, and M. Afendi, "Comparison of ARIMA and GRU Models for High-Frequency Time Series Forecasting," *Scientific Journal of Informatics*, vol. 10, no. 3, p. 389, 2023, doi: 10.15294/sji.v10i3.45965.

[12] I. W. R. Pinastawa, M. G. Pradana, D. S. Setiawan, and A. Izzety, "Comparison of ARIMA and GRU Methods in Predicting Cryptocurrency Price Movements," *sinkron*, vol. 9, no. 1, pp. 96–105, Jan. 2025, doi: 10.33395/sinkron.v9i1.14235.

[13] A. M. Windhy and A. Syariful Jamil, "Peramalan Harga Cabai Merah Indonesia: Pendekatan ARIMA Forecasting Indonesian Red Chilli Prices: The ARIMA Approach," 2021.

[14] R. Akademi, M. Cirebon, A. Badan, and P. Statistik, "PENGARUH HARGA CABAI TERHADAP TINGKAT INFLASI DI INDONESIA TAHUN 2016," 2017.

[15] Kementerian Pertanian, *ANALISIS KINERJA PERDAGANGAN CABAI MERAH Pusat Data dan Sistem Informasi Pertanian Sekretariat Jenderal Kementerian Pertanian 2024*, vol. 14. Pusat Data dan Sistem Informasi Pertanian, 2024.

[16] J. D. Cryer and K.-S. Chan, *Time Series Analysis*, 2nd ed. Springer, 2008.

[17] W. W. S. Wei, *Time Series Analysis: Univariate and Multivariate Methods*, 2nd ed. New York: Greg Tobin, 2006.

[18] Siswanto, J. Sembiring, I. Setiawan, and A. Setyawan, "Number of Cyber Attacks Predicted With Deep Learning Based LSTM Model," 2024.