Available online www.worldscientificnews.com

World Scientific News

An International Scientific Journal

WSN 217 (2026) 24-36 EISSN 2392-2192

Evaluating the Accuracy of ARIMA and Least Square Spline Methods in Forecasting Daily Gold Futures Prices

Veronica Clasrissa Putrie[1], Zahira Najmatul Hayyah[2], Sri Purwani[3]

[1,2] Master Program in Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Padjadjaran, Jatinangor, Sumedang 45363, West Java, Indonesia

[3] Department of Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Padjadjaran, Jatinangor, Sumedang 45363, West Java, Indonesia

E-mail address: veronica21003@mail.unpad.ac.id

https://doi.org/10.65770/LIKR1073

ABSTRACT

Gold futures prices are dynamic financial instruments that are often influenced by economic uncertainty, market changes, and investment sentiment. This study aims to compare the accuracy of Autoregressive Integrated Moving Average (ARIMA) and Least Square Spline in forecasting daily gold futures prices. The data used in this study consist of daily Gold Futures prices from 4 January 2021 to 30 December 2025, totaling 1063 observations, which were divided into 850 training data and 213 testing data. The ARIMA model was developed through stationarity testing, differencing, ACF-PACF identification, model selection based on AIC, and residual diagnostics. The Least Square Spline model was developed by selecting the spline order and the optimal number of knots based on Generalized Cross Validation (GCV). Accuracy was evaluated using RMSE, MAE, and MAPE. The results show that the best ARIMA model used was ARIMA(1,1,0), while the best Least Square Spline model was a first-order spline with 15 knot points. Based on the evaluation results on the testing data, ARIMA(1,1,0) produced an RMSE of 909.5807, an MAE of 761.1346, and a MAPE of 20.59894%. Meanwhile, Least Square Spline produced an RMSE of 355.8918, an MAE of 267.4223, and a MAPE of 7.11823%. Therefore, Least Square Spline provides better prediction accuracy than ARIMA in forecasting daily gold futures prices.

Keywords: ARIMA, gold futures, Least Square Spline, MAPE, forecasting.

(Received 8 May 2026; Accepted 15 June 2026; Date of Publication 11 July 2026)

1. INTRODUCTION

Gold is one of the important commodities in economic and investment activities because it is often used as a hedging asset when markets experience uncertainty. The fluctuating movement of gold prices makes forecasting important for investors and market participants in determining buying, selling, and risk management strategies. Several studies have shown that daily gold prices can be analyzed using a time series approach because historical patterns still provide information about subsequent movements [1], [2].

Gold prices are not only influenced by past patterns but are also related to economic indicators such as crude oil prices, interest rates, consumer price indices, exchange rates, and stock market indices [2]. In addition, market sentiment can also increase the complexity of gold price movements [3]. This shows that gold price forecasting is not a simple process because gold prices are affected by both internal time series patterns and external economic conditions. Therefore, the selection of an appropriate forecasting method must consider the fluctuating, nonlinear, and dynamic characteristics of gold price data.

Several forecasting methods have been applied in previous studies to predict gold prices. ARIMA and ARIMA-GARCH models have been used to model gold price movements because these methods can capture historical patterns and volatility in time series data [1], [4], [5]. This shows that gold price data often require forecasting methods that are not only able to model trend patterns but also respond to volatility changes. Other studies have also compared ARIMA with machine learning-based methods such as SVM, LSTM, Bi-LSTM, GRU, and Least Square Support Vector Machine to improve prediction accuracy in financial and commodity markets [2], [6], [7], [8]. These studies indicate that no single method is always superior for all data conditions because model performance depends on the pattern, volatility, and nonlinearity of the data used.

One widely used time series method is Autoregressive Integrated Moving Average (ARIMA). ARIMA has a clear parametric structure and is able to handle non-stationary data through the differencing process [6], [9]. However, in financial data with sharp trend changes or high volatility, ARIMA can produce predictions that are too conservative. This is one reason why many studies have developed or compared ARIMA with other methods such as ARIMA-GARCH, SVM, LSTM, Least Square Support Vector Machine, and other machine learning-based models [1], [2], [4], [10], [7]. Although ARIMA is useful for short-term forecasting, its linear structure may limit its ability to follow sudden changes in gold futures prices.

On the other hand, nonparametric approaches such as Least Square Spline can serve as an alternative because they do not require the functional form to be strictly specified from the beginning. Spline uses knot points to form curve segments, making it more flexible in following changes in local patterns [11], [12]. In financial data, Least Square Spline has also been reported to produce good prediction accuracy when data patterns are not entirely linear [13], [14]. The ability of spline models to adjust curve segments through knot points makes this method suitable for data with changing trends and fluctuating movements.

Based on the discussion above, this study compares the accuracy of ARIMA and Least Square Spline in forecasting daily gold futures prices. ARIMA is used as a representation of a parametric time series method, while Least Square Spline is used as a nonparametric approach. The accuracy of the two methods is compared using RMSE, MAE, and MAPE, with MAPE as the main measure because it is easy to interpret in percentage form. This comparison is expected to provide insight into which method is more suitable for forecasting daily gold futures prices when the data show sharp trend changes and fluctuating patterns.

2. MATERIALS AND METHODS

2.1. Materials

This study uses secondary data in the form of daily Gold Futures prices from 4 January 2021 to 30 December 2025. Gold Futures price data were selected because gold is one of the financial commodities with fluctuating movements and is widely used as an investment instrument. The data consist of date and closing price information, which were then arranged based on chronological order to form a time series dataset.

Before modeling, preprocessing was carried out to ensure that the data were ready for analysis. The preprocessing steps included arranging the data by date, checking missing values, checking duplicate dates, converting price data into numerical form, and forming a time index. After preprocessing, 1063 observations were obtained. The time index was used as the predictor variable in the Least Square Spline model, while the Gold Futures price was used as the response variable to be forecasted.

The data were divided chronologically into training data and testing data. This division was not carried out randomly because random splitting can disrupt the temporal structure of time series data. The training data were used to build the ARIMA and Least Square Spline models, while the testing data were used to evaluate forecasting performance. In this study, 850 observations were used as training data and 213 observations were used as testing data.

Table 1. Training and Testing Data Split.

Data type	Number of observations	Period
Training data	850	4 January 2021 - 4 December 2024
Testing data	213	5 December 2024 - 30 December 2025
Total data	1063	4 January 2021 - 30 December 2025

Source: Processed data, 2026.

The forecasting error is defined as:

$$e_t = y_t - \hat{y}_t$$

where y_t is the actual value and $\hat{y}_t$ is the predicted value.

2.2. ARIMA Model

ARIMA is a parametric time series model that uses past values and past errors to forecast future values. This model is widely used in financial and gold price forecasting because it can handle non-stationary data through differencing [1], [2], [9].

The general form of ARIMA (p, d, q) is:

$$\phi_p(B)(1-B)^d y_t = \theta_q(B)\varepsilon_t$$

where p is the autoregressive order, d is the differencing order, and q is the moving average order. The first-order differencing is written as:

$$\nabla y_t = y_t - y_{t-1}$$

In this study, stationarity was tested using the Augmented Dickey-Fuller test. ACF and PACF plots were used to identify candidate models. The best ARIMA model was selected based on the smallest Akaike Information Criterion (AIC):

$$AIC = -2\ln(L) + 2k$$

where L is the likelihood value and k is the number of parameters.

2.3. Least Square Spline Model

Least Square Spline is a nonparametric regression method that can model fluctuating data patterns by using knot points. This method is more flexible because it does not require a fixed functional form from the beginning [11], [13], [14].

The nonparametric regression model is:

$$y_t = g(x_t) + \varepsilon_t$$

The spline function with order $p$and knot points $\tau_1, \tau_2, \dots, \tau_m$is:

$$g(x_t) = \sum_{j=0}^{p} \beta_j x_t^j + \sum_{k=1}^{m} \beta_{p+k} (x_t - \tau_k)_+^p$$

with:

$$\left(x_t - \tau_k\right)_+^p = \begin{cases} (x_t - \tau_k)^p, & x_t > \tau_k \\ 0, & x_t \le \tau_k \end{cases}$$

The parameter estimation is obtained using least squares:

$$\hat{\beta} = (X_\lambda^T X_\lambda)^{-1} X_\lambda^T y$$

The best spline model was selected using the smallest Generalized Cross Validation (GCV) value:

$$GCV(\lambda) = \frac{MSE(\lambda)}{\left[\frac{1}{n} trace(I - H(\lambda))\right]^2}$$

2.4. Forecast Accuracy

The accuracy of ARIMA and Least Square Spline was evaluated using RMSE, MAE, and MAPE. These measures are commonly used in forecasting studies, especially for gold price and financial data [1], [2], [9], [11].

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

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

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

MAPE was used as the main criterion because it is expressed in percentage form and is easier to interpret. The model with the smallest MAPE value was selected as the best forecasting model.

Table 2. Interpretation of MAPE Value.

MAPE value	Interpretation
$MAPE \leq 10\%$	Highly accurate
$10\% < MAPE \leq 20\%$	Good
$20\% < MAPE \leq 50\%$	Fair
$MAPE > 50\%$	Inaccurate

3. RESULT AND DISCUSSION

3.1. Data Description

Table 3. Descriptive Statistics of Daily Gold Futures Prices.

Statistics	Value
Number of observations	1063
Mean	2253.472502
Standard deviation	667.940988
Minimum	1623.6
Median	1936.4
Maximum	4527.5

Source: Processed data, 2026.

Table 3 shows that gold futures prices had a mean of 2253.472502 with a standard deviation of 667.940988. The large difference between the minimum and maximum values indicates significant price changes during the research period.

Figure 1. Daily Gold Futures Price Plot.

Figure 1 shows that gold futures prices moved fluctuatively. From 2021 to 2023, prices moved within a relatively narrower range. After entering 2024 and 2025, prices showed a stronger upward trend. This condition poses a challenge for forecasting models because the testing data are in a phase of sharp increase.

Figure 2. Training and Testing Data Split.

3.2. ARIMA Modeling Results

The ADF test on the initial training data produced an ADF statistic of -1.093541 with a p-value of 0.9236281, indicating that the data were not yet stationary. After first-order differencing, the ADF statistic became -9.972863 with a p-value of 0.01. Therefore, the differencing value used in the ARIMA model was d = 1.

Table 4. ADF Test Results.

Testing stage	ADF statistic	p-value	Conclusion
Initial training data	-1.093541	0.9236281	Non-stationary
First differencing	-9.972863	0.01	Stationary

Source: Processed data, 2026.

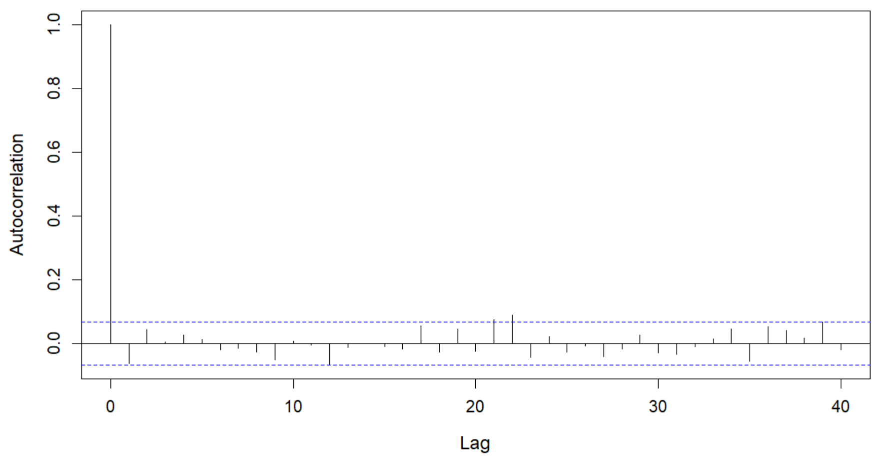

Figure 3. ACF Plot After First Differencing.

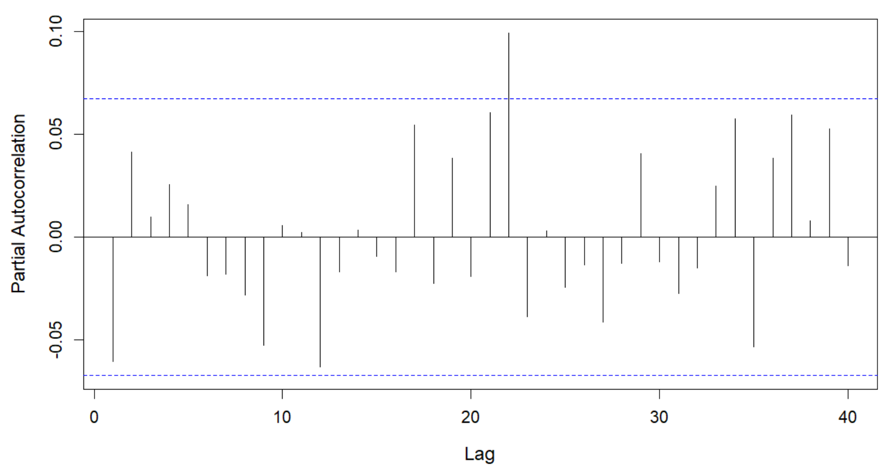

Figure 4. PACF Plot After First Differencing.

Based on the ACF and PACF plots after first-order differencing, the data had met the stationarity condition, so the differencing value used was d=1. The ACF plot did not show a strong cutoff at the early lags, while the PACF plot showed an effect at the early lag. Therefore, several ARIMA candidate models in the form ARIMA(p,1,q) were further tested. Based on the comparison results of the candidate models shown in Table 4, the best ARIMA model used in this study was ARIMA(1,1,0). ARIMA(1,1,0) was selected because it had the most suitable model evaluation results compared with other ARIMA candidates in the model selection process.

Table 5. Best ARIMA Candidate Models Based on Aic.

Rank	ARIMA model	AIC	BIC
1	ARIMA(1,1,0)	7485.627	7495.115
2	ARIMA(0,1,1)	7485.871	7495.359
3	ARIMA(4,1,2)	7485.945	7519.154
4	ARIMA(2,1,0)	7486.024	7500.256
5	ARIMA(0,1,2)	7486.038	7500.270

Source: Processed data, 2026.

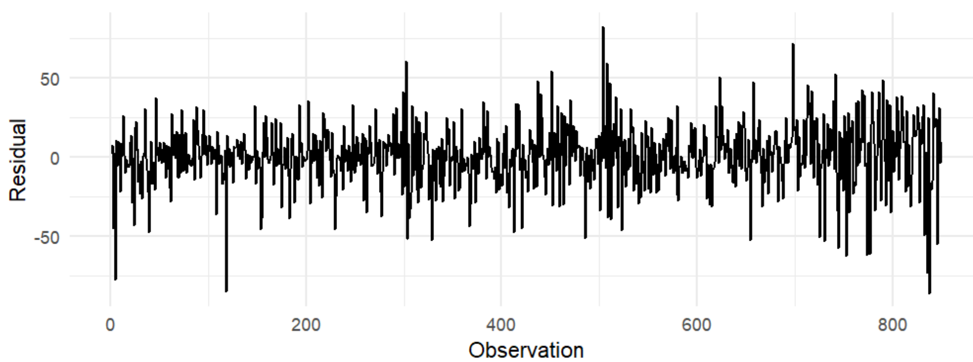

Figure 5. Residual Plot of ARIMA(1,1,0).

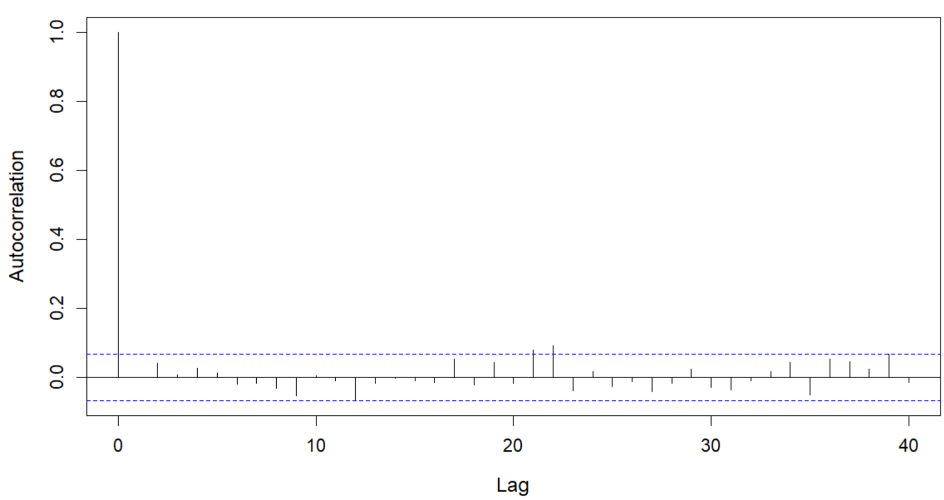

Figure 6. ACF Plot of ARIMA Residuals.

Residual diagnostic checking was carried out on the ARIMA(1,1,0) model. Based on the residual plot, the model residuals generally moved around zero, although some fluctuations were still present in certain periods. Furthermore, the residual ACF plot showed that most lags were within the confidence interval limits. This indicates that the residuals of the ARIMA(1,1,0) model did not show strong autocorrelation.

Table 6. Ljung-Box Test Results for Arima Residuals.

Lag	Ljung-Box statistic	p-value
10	5.943178	0.8200144
20	14.871062	0.7837373

Source: Processed data, 2026.

The Ljung-Box test results show that the p-values at lag 10 and lag 20 were greater than 0.05, indicating that the ARIMA residuals did not have significant autocorrelation. However, the model summary indicated signs of non-normal residuals and heteroscedasticity. This condition is commonly found in highly fluctuating financial data and shows that ARIMA still has limitations in capturing changes in volatility [4], [10].

3.3. Least Square Spline Modeling Results

The Least Square Spline model was selected based on the smallest GCV value. The selection results showed that the best model was a first-order spline with 15 knot points, with a GCV of 1796.553 and an MSE of 1725.410. The selection of 15 knots indicates that the gold futures price data had several pattern changes during the training period, so the model required flexibility to follow local patterns.

Table 7. Best Least Square Spline Candidate Models Based On Gcv.

Rank	Spline order	Number of knots	GCV	MSE
1	1	15	1796.553	1725.410
2	2	14	2043.473	1962.551

Source: Processed data, 2026.

Methodologically, this result supports the characteristics of spline as a nonparametric approach that is able to form curve segments based on knot points. This flexibility becomes important when price patterns are not entirely linear and experience relatively sharp trend changes [11], [13], [14].

3.4. Forecast Accuracy and Prediction Results

Table 8. Forecasting Accuracy Comparison.

Model	RMSE	MAE	MAPE	Interpretation
ARIMA(1,1,0)	909.5807	761.1346	20.59894%	Fair
Least Square Spline	355.8918	267.4223	7.11823%	Highly accurate

Source: Processed data, 2026.

Based on Table 8, the Least Square Spline model produced lower error values than ARIMA(1,1,0) across all evaluation measures. The ARIMA(1,1,0) model produced an RMSE of 909.5807, an MAE of 761.1346, and a MAPE of 20.59894%. Meanwhile, Least Square Spline produced an RMSE of 355.8918, an MAE of 267.4223, and a MAPE of 7.11823%. The MAPE value of Least Square Spline was below 10%, so it is included in the highly accurate category. In contrast, ARIMA(1,1,0) was in the fair category because its MAPE value was slightly above 20%. Therefore, the best model based on RMSE, MAE, and MAPE values was Least Square Spline.

Figure 7. Comparison of Actual and Predicted Gold Futures Prices.

Figure 7 shows the comparison between actual gold futures prices and the prediction results of ARIMA(1,1,0) and Least Square Spline on the testing data. Based on the graph, the ARIMA(1,1,0) predictions tended to be flatter and were less able to follow the sharp increase in gold prices during the testing period. In contrast, Least Square Spline produced a prediction pattern that increased gradually and was closer to the actual values. This shows that Least Square Spline is more adaptive in capturing changes in gold futures price patterns than ARIMA.

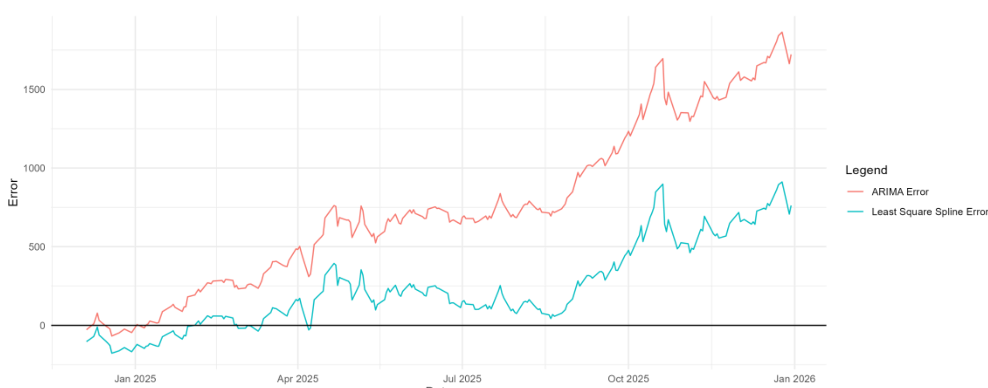

Figure 8. Forecasting Error of ARIMA and Least Square Spline.

Figure 8 shows that the prediction error of ARIMA(1,1,0) tended to increase when the actual gold prices experienced a sharp rise. This occurred because the ARIMA predictions were relatively constant, while the actual prices moved upward quite strongly in the testing data.

In the Least Square Spline model, errors also appeared in several periods, but their magnitude was lower than that of ARIMA. This condition strengthens the numerical evaluation results showing that Least Square Spline has better prediction capability than ARIMA(1,1,0).

Table 9. Selected Future Forecasting Results.

Date	ARIMA forecast	Least Square Spline forecast
2025-12-31	2663.808	3628.284
2026-01-01	2663.837	3632.455
2026-01-02	2663.835	3636.626
2026-01-05	2663.835	3640.968
2026-01-06	2663.835	3644.968

Source: Processed data, 2026.

Figure 9. Future Forecasting of Gold Futures Prices.

Based on Table 9 and Figure 9, the forecasting results for the following period show different patterns between ARIMA(1,1,0) and Least Square Spline. The ARIMA(1,1,0) forecast tended to be constant at around 2663.8, while the Least Square Spline forecast showed a gradual increasing tendency. This difference indicates that ARIMA produced more conservative predictions, while Least Square Spline more closely followed the increasing trend tendency at the end of the actual data.

Overall, the research results show that Least Square Spline is more suitable for daily gold futures price data that have strong pattern changes in the testing period. ARIMA remains feasible in terms of residuals that are not autocorrelated, but its prediction results are less adaptive to sharp increases in gold prices.

4. CONCLUSION

This study compared the ARIMA and Least Square Spline methods in forecasting daily gold futures prices from 4 January 2021 to 30 December 2025. The best ARIMA model used was ARIMA(1,1,0), while the best Least Square Spline model was a first-order spline with 15 knot points. Based on the evaluation results on the testing data, ARIMA(1,1,0) produced an RMSE of 909.5807, an MAE of 761.1346, and a MAPE of 20.59894%.

Meanwhile, Least Square Spline produced an RMSE of 355.8918, an MAE of 267.4223, and a MAPE of 7.11823%. Because Least Square Spline had lower RMSE, MAE, and MAPE values, this method was proven to be more accurate than ARIMA in this study.

The advantage of Least Square Spline is due to its ability to form a more flexible curve through knot points, allowing the model to follow local pattern changes in gold futures price data. In contrast, ARIMA(1,1,0) produced flatter predictions, making it less adaptive to the sharp increase in gold prices during the testing period.

Future research may consider models that are able to handle high volatility, such as ARIMA-GARCH, LSTM, SVM, or hybrid models. In addition, external variables such as exchange rates, inflation, interest rates, world oil prices, and global market indices can be added so that the model can capture economic factors that influence gold prices more broadly. Testing with several data split scenarios can also be conducted to assess the stability of model performance.

References

[1] S. Setyowibowo, M. As'ad, Sujito, and En. Farida, "Forecasting of Daily Gold Price using ARIMA-GARCH Hybrid Model," *J. Ekon. Pembang.*, vol. 19, no. 2, pp. 257–270, 2021, doi: 10.29259/jep.v19i2.13903.

[2] Y. R. Madhika, Kusrini, and T. Hidayat, "Gold Price Prediction Using the ARIMA and LSTM Models," *Sink. J. dan Penelit. Tek. Inform.*, vol. 7, no. 3, pp. 1255–1264, 2023, doi: 10.33395/sinkron.v8i3.12461.

[3] F. Yuan, C. Lee, and C. Chiu, "Using Market Sentiment Analysis and Genetic Algorithm-Based Least Squares Support Vector Regression to Predict Gold Prices," *Int. J. Comput. Intell. Syst.*, vol. 13, no. 1, pp. 234–246, 2020, doi: 10.2991/ijcis.d.200214.002.

[4] H. R. Sari, S. Wahyuningsih, and M. Siringoringo, "Forecasting Indonesian gold prices using the ARIMA(0,1,1)-GARCH(1,0) model," *J. EKSPONENSIAL*, vol. 15, no. 2017, pp. 1–10, 2024, doi: 10.30872/eksponensial.v15i1.1265.

[5] T. Bunnag, "The Importance of Gold's Effect on Investment and Predicting the World Gold Price Using the ARIMA and ARIMA-GARCH Model," *Ekon. J. Econ.*, vol. 2, no. 1, pp. 38–52, 2024, doi: 10.60084/eje.v2i1.155.

[6] D. Makala and Z. Li, "Prediction of gold price with ARIMA and SVM," *J. Phys. Conf. Ser.*, vol. 1767, no. 2021, pp. 1–8, 2021, doi: 10.1088/1742-6596/1767/1/012022.

[7] G. S. Pillai and M. I. Mary, "Gold Price Forecast Based on the Least Square Support Vector Machine," *Int. J. Intell. Syst. Appl. Eng.*, vol. 12, no. 3, pp. 684–697, 2024.

[8] M. Yurtsever, "Gold Price Forecasting Using LSTM , Bi-LSTM and GRU," *Eur. J. Sci. Technol.*, vol. 31, no. 31, pp. 341–347, 2021, doi: 10.31590/ejosat.959405.

[9] Z. Amry and B. H. Siregar, “ARIMA Model Selection for Composite Stock Price Index in Indonesia Stock Exchange,” *Int. J. Account. Financ. Stud.*, vol. 2, no. 1, pp. 31–38, 2019, doi: 10.22158/ijafs.v2n1p31.

[10] A. Sofian, E. Hidayat, and G. Primajati, “Estimating and Forecasting Composite Index in Pandemic Era Using ARIMA-GARCH Model,” *J. Varian*, vol. 7, no. 2, pp. 115–124, 2024, doi: 10.30812/varian.v7i2.2103.

[11] A. T. Fitriyah, N. Chamidah, and T. Saifudin, “Comparison Of Least Square Spline and ARIMA Models For Predicting Indonesia Composite Index,” *BAREKENG J. Math. Its Appl.*, vol. 19, no. 3, pp. 2169–2178, 2025, doi: 10.30598/barekengvol19iss3pp2169-2178.

[12] A. Araveeporn, “The Estimating Parameter and Number of Knots for Nonparametric Regression Methods in Modelling Time Series Data,” *Modelling*, vol. 5, pp. 1413–1434, 2024, doi: 10.3390/modelling5040073.

[13] C. Andreas, F. Y. H, E. J. S, and N. Chamidah, “Analyzing The Effect of BI 7-Days Repo Rate on The Jakarta Composite Index Using Nonparametric Regression Approaches Based on Least Square Spline Estimator Berdasarkan Estimator Least Square Spline,” *J. Mat. Stat. Komputasi*, vol. 17, no. 3, pp. 447–461, 2021, doi: 10.20956/j.v17i3.13101.

[14] A. T. Fitriyah, N. Chamidah, T. Saifudin, B. Lestari, and D. Aydin, “Least Squares Spline Estimation Method in Semiparametric Time Series Regression for Predicting Indonesia Composite Index,” *Stat. Optim. Inf. Comput.*, vol. 14, pp. 3776–3803, 2025, doi: 10.19139/soic-2310-5070-2704.