

Time Series Forecasting of Gold Futures Prices Using ARIMA

Tejaswini Shekar

MSDA, Western Governor's University



Problem Statement

"Is it possible to effectively predict gold futures prices based on historical data from the past 10 years using an ARIMA model?" The price of gold has increased by 25% in the past year (CBS News, 2025) and does not appear to be slowing down. Geopolitical tensions worldwide have resulted in individual investors and nations buying gold to protect themselves from incurring losses. In such a time where gold is more popular than ever, it becomes relevant to explore methods of predicting variations in its price. By analyzing data from the past 10 years, the patterns underlying the growth of gold futures can be uncovered. It is crucial to fully understand a potential investment and its projected growth before making the commitment to invest.

Hypotheses

o H0 (Null Hypothesis):

The ARIMA model is not able to effectively predict gold prices with a low error (MAPE is not < 20%).

o H1 (Alternate Hypothesis):

The ARIMA model can predict gold prices effectively with a low error (MAPE < 20%).

Data Analysis Process:

1. Data Collection

The data set used for the analysis is publicly available at <https://www.investing.com/commodities/gold-historical-data>. It was downloaded as a CSV file by entering the data range of 01/01/2015 to 01/01/2025.

The data set has a total of 2564 records, starting from 01/02/15 till 12/31/24. There is a "Date" column and six other variables, including the closing price ("Price") variable which is relevant for the analysis.

2. Data Preparation

The programming language used was Python in a Jupyter Notebook. A range of open-source libraries were also used for data preparation, including Pandas, NumPy, Datetime, Matplotlib, RE and Sci-Kit Learn.

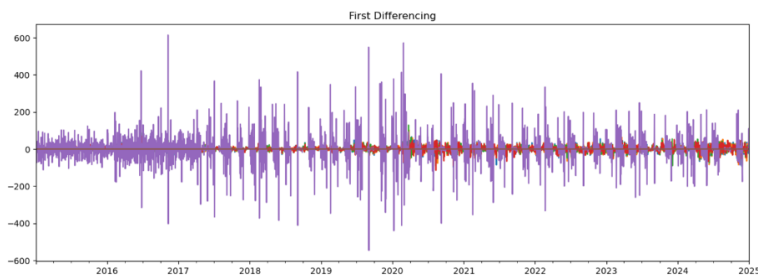
After the CSV file was read into a pandas data frame, the data set was cleaned. Punctuation and irregular characters were removed, and the "Date" column was converted to "DateTime" format. Missing dates/records were added, and their closing prices were imputed by forward filling from the previous day's closing price. The data set was then split into training and testing data sets in an 80-20 split as per industry standards.

Exploratory data analysis was done on the cleaned data. The closing prices for gold ranged from \$1,049.60 to \$2,800 per ounce, with a mean of \$1,606.88 over the past 10 years. The daily closing price of gold is visualized below:



3. Time-Series Analysis

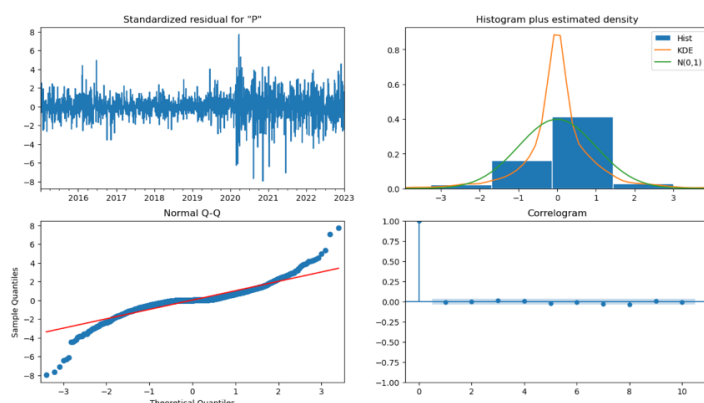
An ARIMA model was deemed appropriate for this analysis because it can use historical price data to predict future prices. The autoregressive, differencing and moving average parts of ARIMA together can capture the changing patterns and trends in stock return data (Hardikkumar, 2024). Prior to building the ARIMA model, the data was first checked for stationarity using the Augmented Dickey-Fuller test. The test showed that the data was not stationary. The test was rerun after taking the first difference of the daily closing prices ($d=1$) to ensure that the data was stationary. A visualization of the data after first differencing also confirmed that the differenced data was stationary (mean of the data is zero, and the variance is uniform).



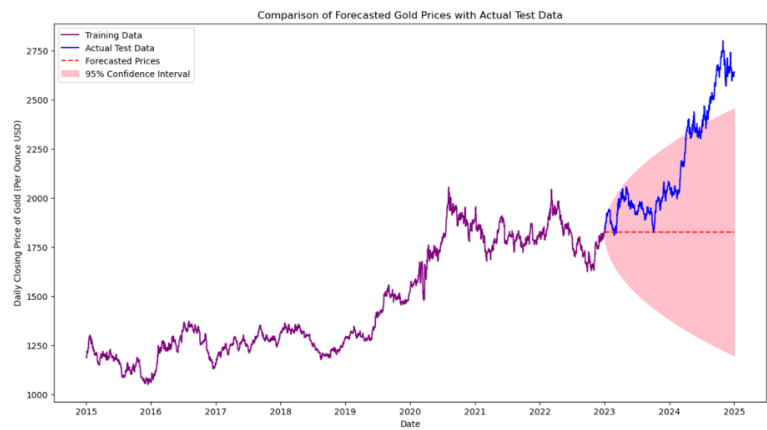
The data set was also decomposed to visualize the trends and seasonality in the data. The gold prices showed an upward trend with yearly seasonal variation. Autocorrelation functions, ACF and PACF, were also plotted for the differenced data. The ACF plot quickly dropped to zero further confirming that the differenced data was stationary. Finally, the best order of ARIMA was determined based on the combination of parameters resulting in the lowest AIC (Akaike Information Criterion). Various values of p and q were run in a for loop to find the lowest AIC score for each combination. The final ARIMA model was fit and trained on the training data set using the best combination of parameters ($p=2$, $d=1$ and $q=2$).

Outline of Findings

The results of the analysis show that the ARIMA model was able to capture the trends in the gold futures prices well as per the evaluation methods used. The AIC and BIC of the final model were 22826.480 and 22856.378, respectively. The residual errors were plotted, showing a mean of zero with uniform variance and no significant patterns in the residuals. The histogram had a normal distribution, and the ACF plot did not show any autocorrelation. This implies that the ARIMA model was able to fully capture the trends and variations in the gold closing price data.



On comparing the model forecast with the actual test data set, while the first half of the test data is within the forecasted 95% confidence interval, the latter half shows a sharp rise well above the predicted confidence interval. It can be inferred that although the ARIMA model may be useful for capturing general trends and linear relationships, its performance is inadequate when making precise, reliable predictions.

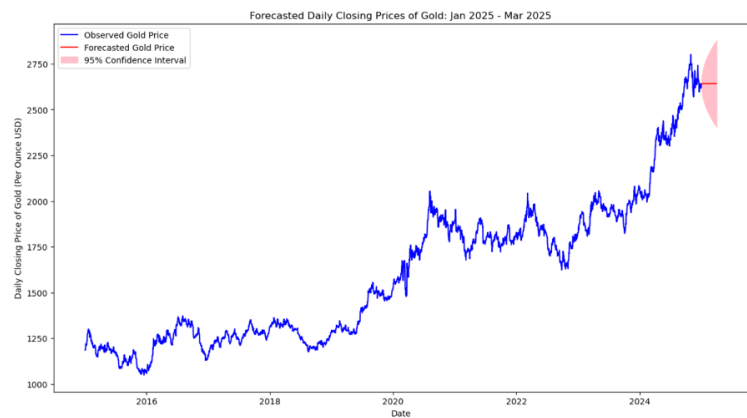


The criterion for determining the model efficacy was MAPE. It is the error expressed as a percentage of the actual values and is scale-independent. In general, a lower MAPE indicates a better model and a MAPE between 10%-20% is good (Juan et al, 2013).

The MAPE calculated for the final ARIMA model was 14.68%. Since this is <20%, we can reject the null hypothesis and accept the alternate hypothesis that it is possible to build an ARIMA model that can effectively forecast gold prices with MAPE <20%.

Model Forecasting:

The ARIMA model forecast predicted the daily closing price of gold to be about \$2,641 per ounce for the next 3 months (Jan-Mar 2025).



Limitations of the Tools/Techniques Used

- One limitation of the analysis is that ARIMA assumes linearity and stationarity. The true nature of gold prices is much more erratic, with numerous global factors contributing to its spikes and dips. These complex relationships cannot be accounted for by the ARIMA model.
- ARIMA is also highly sensitive to outliers and missing values. These factors influence the final accuracy of the model and the reliability of the results

Proposed Actions

- Explore alternative forecasting methods and compare predictive accuracies: One possibility for further analysis is the implementation of neural networks. Compared to ARIMA, these models are better suited for complex, non-linear

relationships and may produce more precise and accurate forecasts. It is important to compare the accuracies of various models using metrics like RMSE and MAPE to find the model with the best outcome.

- Utilize other variables in the data: The data set contains other relevant variables (e.g. per cent change in price) that were not considered in the analysis. These variables can be utilized to build more complex and reliable forecasting models.
- Expand the data set: The data set can be further expanded to include 15-20 years of historical data. This provides the model with more data to learn from and may result in the model capturing long-term trends and variations in pricing. It is important to remember that when making investment decisions, it is best not to rely on any single model or source of information. Additional research is always both recommended and necessary.

Expected Benefits of the Study

- Guidance for Individual Investors: One of the key benefits of the study is that the individual investor can use it as a guide to buying and selling gold futures. Rather than making financial decisions blindly or based on speculation, investors can use historical trends in the data and the model's forecasted prices. The insights into the patterns of price variation in gold futures allow for informed, data-driven decision-making.
- Foundation for Advanced Modeling Techniques: The study also paves the way for more sophisticated modelling methods, such as neural networks/LSTM (Long Short-Term Memory) models, which are more suited for complex, non-linear relationships.

The study provides a prepared and processed data set that acts as a foundation for these more complex models. The analysis also acts as a benchmark for comparing the predictive accuracies of various forecasting models.

Sources

1. Data Science Wizards. (3 November 2023). Preprocessing and Data Exploration for Time Series – Handling Missing Values. Medium. <https://medium.com/@datasciencewizards/preprocessing-and-data-exploration-for-time-series-handling-missing-values-e5c507f6c71c>.
2. Galarnyk, Michael. (3 February 2025). Train Test Split: What It Means and How to Use It. BuiltIn. <https://builtin.com/data-science/train-test-split>.
3. Hardikkumar. (4 December 2024). Stock Market Forecasting using Time Series Analysis with ARIMA Model. Analytics Vidhya. <https://www.analyticsvidhya.com/blog/2021/07/stock-market-forecasting-using-time-series-analysis-with-arma-model/>.
4. Investing.com (n.d.) Gold Futures Historical Data. <https://www.investing.com/commodities/gold-historical-data>.
5. Maxwell, Tim. (24 February 2025). What's the Gold Price Outlook for the Rest of 2025? CBS News. <https://www.cbsnews.com/news/whats-the-gold-price-outlook-for-the-rest-of-2025/>.
6. Montaña, Juan & Palmer, Alfonso & Sesé, Albert & Cajal, Berta. (2013). Using the R-MAPE index as a resistant measure of forecast accuracy. Psicothema. 25. 500-506. 10.7334/psicothema2013.23. https://www.researchgate.net/publication/257812432-Using_the_R-MAPE_index_as_a_resistant_measure_of_forecast_accuracy.
7. Patha, Abdulla. (14 September 2024). What are the advantages and disadvantages of using ARIMA models for forecasting? LinkedIn. <https://www.linkedin.com/advice/0/what-advantages-disadvantages-using-arma#:~:text=While%20ARIMA%20models%20are%20useful,missing%20data%20require%20careful%20preprocessing>.