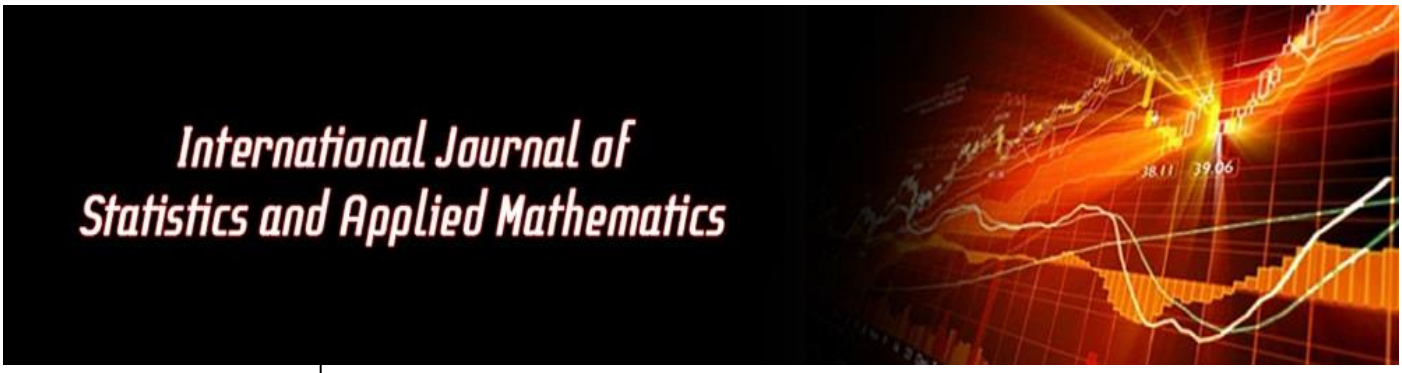
International Journal of Statistics and Applied Mathematics 2025; 10(2): 95-99




**ISSN:** 2456-1452
Maths 2025; 10(2): 95-99

https://www.mathsjournal.com
Received: 08-11-2024
Accepted: 13-12-2024

**M Naresh**
Academic Consultant, Department of Statistics, Sri Venkateswara University, Tirupati, Andhra Pradesh, India

**P Sriyshnavi**
Assistant Professor, Department of CSE, Sri Padmavati Mahila Visvavidyalayam, Tirupati, Andhra Prades

**Aashis Kumar**
Research Scholar, Department of Statistics, University of Allahabad, Prayagraj, Uttar Pradesh, India

**Corresponding Author:**
**M Naresh**
Academic Consultant, Department of Statistics, Sri Venkateswara University, Tirupati, Andhra Pradesh, India

# Prediction and analysis of gold prices

**M Naresh, P Sriyshnavi and Aashis Kumar**

**DOI:** https://dx.doi.org/10.22271/maths.2025.v10.i2b.1982

**Abstract**
It aims to forecast the price of gold using time series analysis techniques. The main objective of this study is to predict the future prices of gold so that investors can make informed decisions about buying and selling gold. Inventory forecasting plays a crucial role in the financial success of a business, and forecasting the prices of gold can be particularly challenging due to its volatility and dependence on a variety of economic and geopolitical factors. We have utilized various time series models such as Exponential smoothing and Auto Regressive Integrated Moving Average (ARIMA) to analyse historical gold price data and generate forecasts. We have also conducted an exploratory data analysis (EDA) to identify any trends or patterns in the data and to assess the stationarity of the time series. Overall, the study demonstrates the usefulness of time series analysis techniques in predicting the price of gold and can be a valuable resource for investors and financial analysts.

**Keywords:** ARIMA, SARIMA, prophet, exponential smoothing model

## Introduction
Gold has been highly valued by humans for thousands of years, dating back to ancient civilizations such as the Egyptians and the Greeks. It has been used as currency, jewelry, and even as a symbol of power and wealth.

Gold has been discovered on all continents, and it has been mined throughout history. The largest gold producer in the world is currently China, followed by Australia and Russia. In the United States, gold is mined primarily in Nevada, Alaska, and Colorado.

The use of gold as a currency has played a significant role in human history. Gold coins have been used as currency since the 6th century BC in Lydia, a kingdom of Asia Minor. The Roman Empire used gold coins as currency, and gold continued to be used as currency through the Middle Ages and the Renaissance. In 1717, the United Kingdom adopted the gold standard, which made it the first country to use gold as the basis for its currency. The gold standard was adopted by other countries as well, and it remained in place until the 20th century.

In the early 20th century, many countries moved away from the gold standard, and the use of gold as a currency declined. However, gold continued to be highly valued as an investment and a store of wealth. During times of economic uncertainty, gold has often been seen as a safe haven asset, with investors buying gold to protect themselves from inflation, currency devaluation, and other economic risks.

Gold prices are influenced by a variety of factors, including supply and demand, global economic conditions, interest rates, and inflation. As an investment, gold has been known to be volatile, with prices often fluctuating sharply in response to economic and political events.

Time series analysis is a statistical technique used to analyze data that changes over time. It can be used to forecast future values based on historical data. In the case of gold price prediction, time series analysis can be used to identify patterns and trends in gold prices over time, and to forecast future prices based on these patterns.

The objective of this study is to use time series analysis to predict gold prices over a period of time. The dataset used in this study includes the daily prices of gold from January 2014 to July 2022. The dataset contains information on the opening, closing, high, and low prices, as well as the volume of gold traded on each day.

The study consists various time series models, including Moving average, Exponential Smoothing and ARIMA to forecast gold prices. The models will be used to identify trends and patterns in the data, and to develop forecasts for future gold prices.

**Objectives**

1. To forecast the future prices of gold: The primary objective of gold price prediction is to forecast the future prices of gold using historical data and relevant predictors. This can help investors and businesses make informed decisions about buying and selling gold.
2. To identify patterns and trends in the gold market: By analysing historical data, a gold price prediction study can identify patterns and trends in the gold market. This can help investors and businesses understand the factors that influence the price of gold and make more informed decisions.
3. To develop a predictive model for gold prices: A gold price prediction project can involve the development of a predictive model for gold prices. This can be done using time-series analysis or other statistical techniques.
4. To evaluate the accuracy of gold price predictions: Another objective of a gold price prediction study is to evaluate the accuracy of the predictions made by the model. This can be done by comparing the predicted prices with the actual prices over a given time period.

**Data Descriptions**

This dataset contains information about daily gold prices, specifically the daily opening, closing, high and low prices. In addition, changes in the amount and percentage of gold traded were recorded.

**This Gold Price Data set link Gold Price from 2014 to 2022**

- Date: Recording date.
- Price: It is close price which can be considered as final price.
- Open: Price at the time of market opening at that day.
- High: Highest price during whole day.
- Low: Lowest price during whole day.
- Volume: Traded volume.
- Chg: % Change from previous price.

**Methodology**

Data Collection: The first step in the methodology is to collect the historical data of the gold price. The data can be collected from various sources such as Yahoo Finance, Google Finance, and the World Gold Council.

1. **Data Pre-processing:** The collected data needs to be pre-processed before it can be used for analysis. The data pre-processing involves cleaning the data, imputing missing values, transforming the data, and scaling the data. The data should also be checked for any outliers and removed if necessary.
2. **Data Visualization:** The pre-processed data needs to be visualized to get a better understanding of the trends and patterns in the data. Various time series plots can be used to visualize the data, such as line plots, scatter plots, and box plots.
3. **Time Series Analysis:** After visualizing the data, the next step is to perform time series analysis. The analysis involves decomposing the time series into its trend, seasonal, and residual components. The time series model can be selected based on the characteristics of the decomposed components.
4. **Model Selection:** There are various time series models that can be used to predict the gold price, such as ARIMA, SARIMA, and Prophet. The model can be selected based on the accuracy of the model, and the complexity of the model.
5. **Model Evaluation:** The selected model needs to be evaluated based on its performance. The performance can be evaluated using various statistical measures such as mean absolute error, root mean square error, and mean absolute percentage error.
6. **Forecasting:** After evaluating the model, the next step is to use the model to forecast the future gold prices. The forecasted values can be visualized using time series plots.
7. **Model Refinement:** If the performance of the model is not satisfactory, the model can be refined by adjusting the model parameters or by using a different model. The model refinement process should be repeated until the model provides satisfactory performance.

**Triple Exponential Smoothing:** Triple Exponential Smoothing is a forecasting method that can be used for time series data that exhibit trend and seasonality. The model uses three smoothing equations to forecast future values in the time series:

1. **Level Equation:** The level equation calculates the smoothed level at each time period by taking a weighted average of the current observation and the previous smoothed level, adjusted for the previous smoothed slope and seasonal index.
2. **Slope Equation:** The slope equation calculates the smoothed slope at each time period by taking a weighted average of the current slope and the previous smoothed slope.
3. **Seasonal Equation:** The seasonal equation calculates the seasonal index at each time period by taking a weighted average of the current observation and the current smoothed level, adjusted for the previous seasonal index.

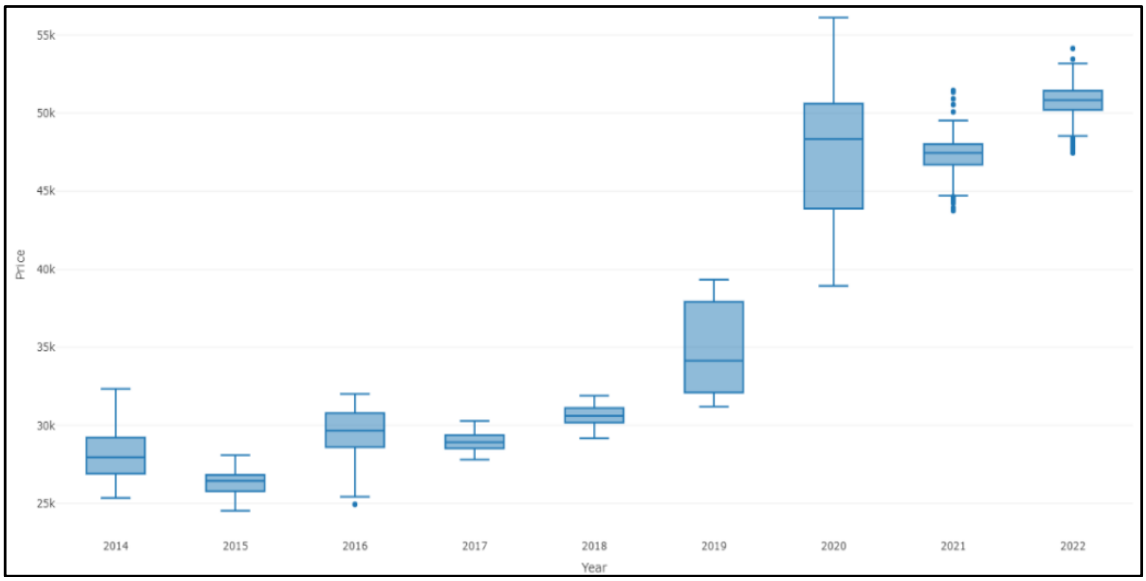
Finally, the forecast equation for Triple Exponential Smoothing combines the smoothed level, smoothed slope, and seasonal index to predict future values in the time series.

**ARIMA:** ARIMA is a forecasting method that can capture both the trend and seasonality of a time series. The model is based on three key parameters:

1. **Autoregressive (AR) part:** The AR part of the model represents the linear dependence between an observation and some number of lagged observations. This means that the current observation is a function of its previous values.
2. **Integrated (I) part:** The I part of the model represents the degree of differencing required to make the time series stationary. This means that the model tries to remove any trends or seasonal patterns that exist in the data.
3. **Moving Average (MA) part:** The MA part of the model represents the dependence between the current observation and a number of lagged white noise error terms. This means that the current observation is a function of previous errors in the model.

The ARIMA model is typically represented as ARIMA (p,d,q), where p is the autoregressive order, d is the degree of differencing, and q is the moving average order. The model uses these parameters to calculate the future values in the time series.

**Empirical Investigation**



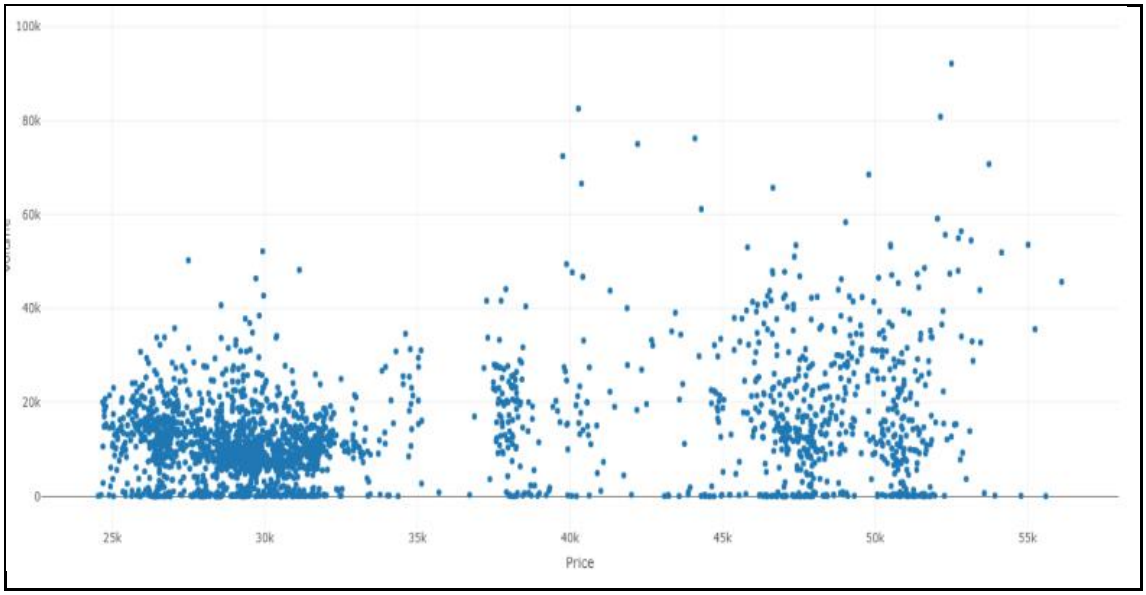

**Graph 1:** Gold price by Year

The given boxplot of gold prices suggests that the year 2020 had the highest variation in gold prices, while the years 2017 and 2018 had the least variation. This can be seen by comparing the heights of the boxes for each year, which represent the range of gold prices, and by looking at the length of the whiskers, which represent the variability in the data.

The high variation in gold prices in 2020 may be attributed to the impact of the COVID-19 pandemic, which disrupted global markets and caused significant economic uncertainty. The fluctuating demand and supply of gold, combined with changes in the global economy, likely contributed to the higher range of gold prices in 2020.

On the other hand, the lower variation in gold prices in 2017 and 2018 may be attributed to more stable market conditions during those years. During these years, there may have been less economic and political instability, leading to a more consistent demand and supply of gold.

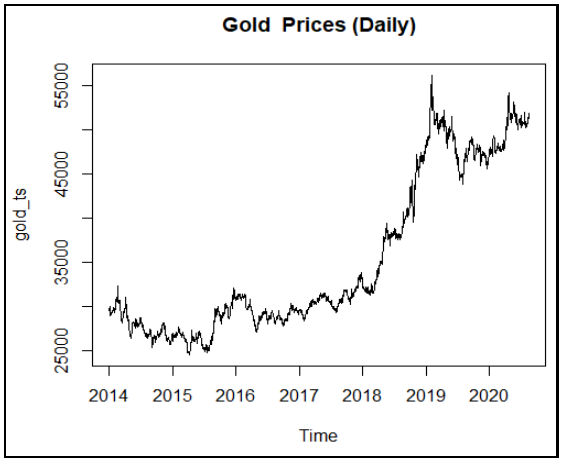
Overall, the boxplot of gold prices year wise provides valuable information about the variability of gold prices over time, and can help identify any significant trends or patterns in the data.




**Graph 2:** Gold price vs. Volume

The scatter plot of gold price from 2014 to 2022 indicates that there are two main clusters of data points with high density. The first cluster is located between 25,000 and 35,000, while the second cluster is located between 45,000 and 50,000. These clusters suggest that gold prices have been relatively stable within these ranges over the past eight years.

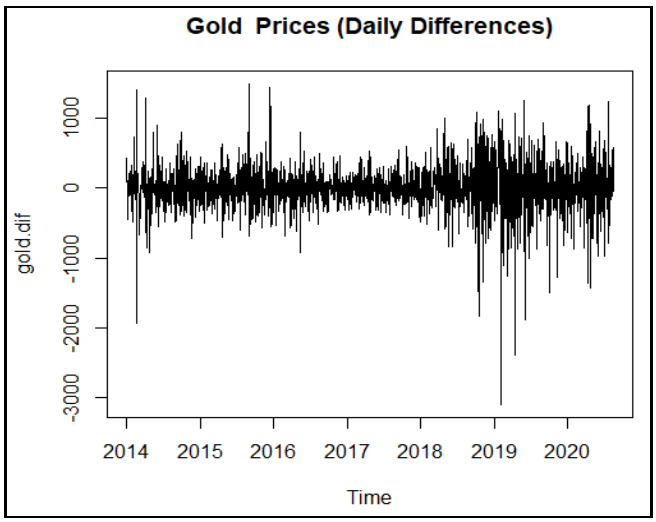
The density of data points in these clusters indicates that there have been many instances of gold prices falling within these ranges over the past eight years. This may be due to a variety of factors, including changes in supply and demand, shifts in global economic conditions, and fluctuations in investor sentiment.




**Graph 3:** Gold Prices (Dally)

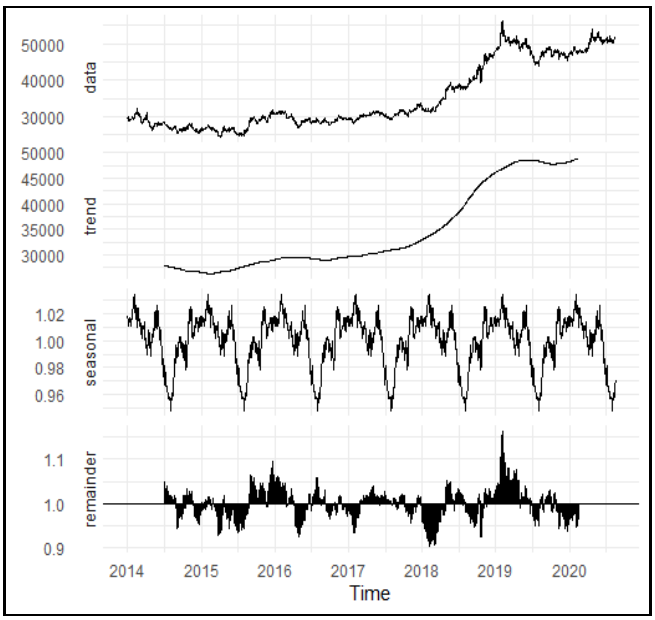
The given graph provides a visual representation of the time series data for gold prices and suggests that there is a long-term increasing trend in the data, as well as some shorter-term seasonal fluctuations. By applying a multiplicative time series model to this data, we can estimate the values of the trend and seasonal components and use these estimates to make predictions about future values of the time series.


**Graph 4:** Gold Price Candlestick Charl




**Graph 5:** Gold Price (Dally Differences)

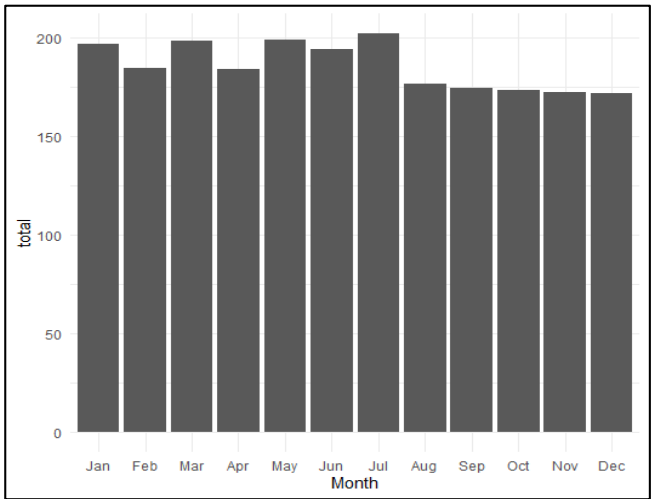
The time series does not seem to be stationary due to the presence of a trend. We will evaluate if first differences are sufficient to turn the time series stationary.



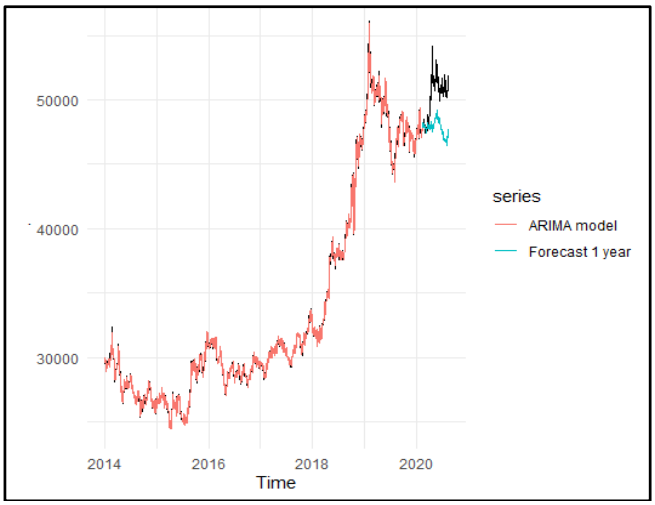

**Graph 6:** Decomposition of multiplicative time series

The increasing trend observed in the data suggests that gold prices have been on an upward trajectory over time. This may be due to a variety of factors, including changes in global economic conditions, shifts in investor sentiment, and fluctuations in supply and demand. Understanding this trend is important for predicting future price trends and developing a forecasting model that takes into account the overall upward trend in gold prices.

In addition to the trend, the time series chart also shows a seasonal pattern, indicating that there are recurring patterns in gold prices that occur at regular intervals. This may be due to factors such as changes in demand during certain times of year or shifts in the global economy during specific periods. Identifying and understanding these seasonal patterns is important for predicting future price trends and developing a forecasting model that takes into account these regular fluctuations in gold prices.




**Graph 7:** The price is high in January but going down in February and going up again in March.




**Graph 8:** Time

From the graph we can see the red line (ARIMA model) with train data is quite similar with our original data. Unfortunately, the blue line (Forecast) still having deviation from original data. We will check our error with Mean Absolute Percentage Error (MAPE)

## Conclusions

The error for ARIMA model is 0.413%, and the error for forecasting is 2.19%.The error is slightly lower than Triple Exponential Smoothing model, we will choose this ARIMA model as our final model.

To sum up our analysis, we have to return to our main objective. We already make model that can predict gold price in 12 months ahead, with forecasting error at 2.197% and still needs improvement in normality of residuals check. As an alternative, we already make another model with Triple Exponential Smoothing Method with forecasting error at 3.237%.

## References

1. Chakraborty S, Chatterjee D, Gupta P. Forecasting gold prices in India using machine learning algorithms. IIM Kozhikode Society & Management Review. 2020;9(1):64-80.
2. Dash SK, Behera SK. Forecasting gold prices in India: An application of univariate time series models. Global Business Review. 2018;19(1):1-13.
3. Estimation of parameters of a multivariate time series VAR model. International Journal of Statistics and Applied Mathematics (IJSAM). 2018;3(1):442-450.
4. Hemasekhar P, et al. Forecasting sales of Honda motor vehicles through SAS software. International Journal of Current Advanced Research. 2018;7(4):12127-12130.
5. Mandal SK, Sahu SK. Forecasting gold prices in India using Box-Jenkins ARIMA models. Journal of Advances in Management Research. 2019;16(3):342-352.
6. Mishra SK, Kumar B. A comparative study of gold price forecasting using ARIMA and neural network models in India. Investment Management and Financial Innovations. 2020;17(2):313-323.
7. Mohapatra S, Pradhan R. A comparative analysis of gold price prediction using time series and machine learning techniques: Empirical evidence from India. Journal of Economic Structures. 2021;10(1):1-21.
8. Mukherjee S, Mishra AK. Modelling and forecasting gold prices in India using ARIMA and GARCH models. IIM Kozhikode Society & Management Review. 2018;7(1):27-35.
9. Panigrahi N, Dash SK. Forecasting gold prices in India using univariate time series models. South Asian Journal of Management. 2020;27(1):138-153.
10. Prediction and analysis of Daucus Carota yields in five districts at Karnataka state. International Journal of Recent Scientific Research (IJRSR). 2018;9(2):24534-24537.
11. Sarkar S, Bhowmik S. Modelling and forecasting of gold prices in India. International Journal of Business and Emerging Markets. 2020;12(2):157-173.
12. Singh JP, Pradhan R. Forecasting gold prices in India: Application of ARIMA and artificial neural network models. Global Business Review. 2015;16(6):1157-1173.
13. Tiwari AK, Pradhan R. Modelling and forecasting of gold price in India using ARIMA and exponential smoothing models. Resources Policy. 2017;53:88-102.
14. Gupta P, Molapata HK, Madhu Sudan G, Kalaparthi NK, Kumar D, Karri S. Statistical model comparison for the National Stock Exchange (NSE) stock price prediction. Communications on Applied Nonlinear Analysis. 2025;32(3).
15. Sumbul SK, Molapata HK, Madhu Sudan G, Kalaparthi NK, Karri S. Model comparison for the prediction of stock prices in the National Stock Exchange by data science techniques. Journal of Computational Analysis and Applications. 2024;33(7).