BAREKENG: Journal of Mathematics and Its Applications
December 2024 Volume 18 Issue 4 Page 2563–2576
P-ISSN: 1978-7227 E-ISSN: 2615-3017
https://doi.org/10.30598/barekengvol18iss4pp2563-2576

COMPARATIVE ANALYSIS OF TIME SERIES FORECASTING MODELS USING ARIMA AND NEURAL NETWORK AUTOREGRESSION METHODS

Melina[1*], Sukono[2], Herlina Napitupulu[3], Norizan Mohamed[4], Yulison Herry Chrisnanto[5], Asep ID Hadiana[6], Valentina Adimurti Kusumaningtyas[7], Ulya Nabilla[8]

[1,5,6]Department of Informatics, Faculty of Science and Informatics, Universitas Jenderal Achmad Yani
Jln. Terusan Jend. Sudirman, Cimahi, 40531, Indonesia

[2,3]Department of Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Padjadjaran
Jln. Raya Bandung Sumedang KM.21, Jatinangor, 45363, Indonesia

[4]Department of Mathematics, Faculty of Ocean Engineering Technology and Informatics,
Universiti Malaysia Terengganu
Kuala Nerus Terengganu, 21030, Malaysia

[7]Department of Chemistry, Faculty of Science and Informatics, Universitas Jenderal Achmad Yani
Jln. Terusan Jend. Sudirman, Cimahi, 40531, Indonesia

[8]Department of Mathematics, Faculty of Engineering, Universitas Samudra
Jln. Prof. Dr. Syarief Thayeb, Langsa, 24416, Indonesia

*Corresponding author's e-mail: ** ***melina@lecture.unjani.ac.id***

ABSTRACT

Article History:

Received: 11th, May 2024
Revised: 3rd, July 2024
Accepted: 8th,August 2024
Published: 14th, October 2024

Keywords:

ARIMA;
Forecasting;
NNAR;
Non-Linear;
Time series.

Gold price fluctuations have a significant impact because gold is a haven asset. When financial markets are volatile, investors tend to turn to safer instruments such as gold, so gold price forecasting becomes important in economic uncertainty. The novelty of this research is the comparative analysis of time series forecasting models using ARIMA and the NNAR methods to predict gold price movements specifically applied to gold price data with non-stationary and non-linear characteristics. The aim is to identify the strengths and limitations of ARIMA and NNAR on such data. ARIMA can only be applied to time series data that are already stationary or have been converted to stationary form through differentiation. However, ARIMA may struggle to capture complex non-linear patterns in non-stationary data. Instead, NNAR can handle non-stationary data more effectively by modeling the complex non-linear relationships between input and output variables. In the NNAR model, the lag values of the time series are used as input variables for the neural network. The dataset used is the closing price of gold with 1449 periods from January 2, 2018, to October 5, 2023. The augmented Dickey-Fuller test dataset obtained a p-value = 0.6746, meaning the data is not stationary. The ARIMA(1, 1, 1) model was selected as the gold price forecasting model and outperformed other candidate ARIMA models based on parameter identification and model diagnosis tests. Model performance is evaluated based on the RMSE and MAE values. In this study, the ARIMA(1, 1, 1) model obtained RMSE = 16.20431 and MAE = 11.13958. The NNAR(1, 10) model produces RMSE = 16.10002 and MAE = 11.09360. Based on the RMSE and MAE values, the NNAR(1, 10) model produces better accuracy than the ARIMA(1, 1, 1) model.

How to cite this article:

Melina, Sukono, H. Napitupulu, N. Mohamed, Y. H. Chrisnanto, A. ID. Hadiana, V. A. Kusumaningtyas and U. Nabilla., "COMPARATIVE ANALYSIS OF TIME SERIES FORECASTING MODELS USING ARIMA AND NEURAL NETWORK AUTOREGRESSION METHODS," *BAREKENG: J. Math. & App.*, vol. 18, iss. 4, pp. 2563-2576, December, 2024.

Journal homepage: *https://ojs3.unpatti.ac.id/index.php/barekeng/*
Journal e-mail: *barekeng.math@yahoo.com; barekeng.journal@mail.unpatti.ac.id*

Research Article · Open Access

1. INTRODUCTION

Gold is one of the precious metals that function as jewelry and is an investment favored by the public because gold is easy to liquidate, can be used for simple transactions, and is not easily damaged [1]. As an investment that is not eroded by inflation, prices tend to rise, and the process of buying and selling gold is flexible, causing people to be more interested in investing in gold than stocks. In a globalized economy, gold is widely used as a haven [2]. Gold price fluctuations always increase investment risk, while the causes of such fluctuations are complex, and gold price trends are influenced by many factors [3]. As a financial asset, gold protects against extreme movements in value compared to other types of assets. Although the price of gold is relatively volatile, it is trending upwards, so investors should be wise in investing in gold and be able to forecast future gold price opportunities. In statistics, several models can predict data in the future. Time series analysis is one of the most popular forecasting models because it uses past data that will be used as a reference for forecasting the future [4]. Time series data is displayed based on time, with the characteristic of time series data being within a certain period. Traditional approaches to time series forecasting, such as Box-Jenkins or auto-regressive integrated moving averages (ARIMA), are linear process models with limitations for non-stationary data [5]. Time series forecasting is the forecasting of events that will occur in the future based on previous data [6].

ARIMA and neural network autoregressive (NNAR) can resolve periodic series in time series data. ARIMA can only be applied to time series data that are already stationary or converted into stationary form through differentiation. However, ARIMA may struggle to capture complex non-linear patterns in non-stationary data, where NNAR models become valuable. Unlike ARIMA, NNAR can handle non-stationary data more effectively by modeling the complex non-linear relationship between input and output variables. Using lagged values of time series data as inputs, NNAR can complement the limitations of ARIMA, providing a more robust approach to forecasting in the presence of non-linearity and non-stationarity [7].

Previous studies have conducted forecasting using the ARIMA and NNAR models. One of them is a study that forecasts the death rate due to COVID-19 infection in Brazil. The research data was obtained from the total data of confirmed cases and the total data of deaths due to COVID-19. Data on COVID-19 cases in Brazil from February 15, 2020, to April 30, 2020. The NNAR and ARIMA models were applied to data for 76 days. The accuracy of forecasting is checked through MSE. The result is that the $NNAR(1, 1)$ model outperforms the $ARIMA(0, 2, 1)$ model with an error rate of 6.85% for the NNAR model and 7.11% for the ARIMA model [8].

Research designing a forecasting model for rice production in Andhra Pradesh. The dataset was taken over the period 1982 to 2022 (for 40 years). The research results show that the ARIMA model (0,1,1) produces the best accuracy with RMSE = 53.1440 and MAPE = 4.8300. On the other hand, the $NNAR(1, 1)$ model obtained a value of RMSE = 64.8142 and MAPE = 6.0901. Forecasting rice production for 2023 to 2030, the $NNAR(1, 1)$ model produces a higher rice production forecast value, whereas the $ARIMA(0, 1, 1)$ model shows lower production [9].

Forecasting gold price fluctuations in time series is difficult because the data is neither stationary nor linear. The novelty of this research is that it provides a comparative analysis of time series forecasting models using ARIMA and NNAR methods specifically designed to predict gold price movements characterized by non-stationary, volatility, and non-linear data. The models are measured for performance using a comprehensive set of evaluation metrics across the most recent dataset covering 1449 periods from January 2, 2018, to October 5, 2023. During this period, in early March 2020, the COVID-19 pandemic shook the financial markets [10]. By examining how each model handles the inherent non-stationarity and volatility of gold prices and testing the residuals of the ARIMA outputs, this study aims to identify the strengths and limitations of ARIMA and NNAR when using non-stationary, volatile, and non-linear characterized data. Both models are measured for performance using a series of evaluation metrics. The evaluation metrics are based on obtaining root mean square error (RMSE) and mean absolute error (MAE) values [11]. The ARIMA model is very effective for short-term forecasting with the requirement of stationarity of time series data. The NNAR model is a non-linear model that uses lag values in time series data for input variables to the NN. In the NNAR model, setting parameters by changing the network architecture greatly affects the model's accuracy. Generally, previous studies did not conduct parameter-setting experiments to get the best NNAR model. This research offers valuable insights for financial analysts, researchers, banks, governments, and investors seeking more reliable forecasting tools in the face of market uncertainty.

2. RESEARCH METHODS

2.1 Dataset

The spot price is the most widely used gold price and is the most real-time gold price as it is constantly updated. The cost of gold in the spot market uses the United States Dollar, and the unit is a troy ounce (oz). Historical gold price data is obtained from the website http://www.finance.yahoo.com, with 1449 periods from January 2, 2018, to October 5, 2023.

2.2 ARIMA

The ARIMA model is a traditional statistical model used in forecasting, which considers correlations in time series and considers various assumptions, such as stationarity and white noise, that must be met [12]. The ARIMA method is also known as the Box-Jenkins method [6]. The Box-Jenkins method fits a mixed ARIMA model to a given data set. ARIMA, which stands for 'Auto-Regressive Integrated Moving Average,' is a model class that describes a particular time series based on past values. An ARIMA model is characterized by 3 terms namely p is the order of the autoregressive (AR) terms, q is the order of the moving average (MA) terms, and d is the number of differences required to make the time series stationary stationer [13]. The general p th-order AR(p) process is [12].

$$Z_t = \emptyset_1 Z_{t-1} + \emptyset_2 Z_{t-2} + \cdots + \emptyset_p Z_{t-p} + a_t, \tag{1}$$

or

$$(1 - \emptyset_1 B - \emptyset_2 B^2 - \cdots - \emptyset_p B^p) Z_t = a_t \tag{2}$$

The general q th-order $MA(p)$ process is

$$Z_t = (1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q) a_t \tag{3}$$

The ARMA model is a combination of AR and MA models. The general form of the ARMA(p, q) model is

$$\left(1 - \emptyset_1 B - \cdots - \emptyset_p B^p\right) Z_t = \theta_0 + (1 - \theta_1 B - \cdots - \theta_q B^q) a_t \tag{4}$$

The general non-stationary of the ARIMA(p, d, q) model with $d \neq 0$ is [14], [15]

$$\emptyset_p(B)(1 - B)^d Z_t = \theta_0 + \theta_q(B) a_t, \tag{5}$$

or

$$\left(1 - \emptyset_1 B - \cdots \emptyset_p B^p\right) Z_t (1 - B)^d = \theta_0 + (1 - \theta_1 B - \cdots - \theta_q B^q) a_t, \tag{6}$$

where, Z_t is time series at time t, $\emptyset_t$ is autoregressive coefficient, θ_t is moving average coefficient, B is a backshift operator, a_t is residual at time t, p is autoregressive component, q is moving average component, d is differencing component.

2.3 NNAR

A neural network (NN) is an information-processing technique or approach inspired by how the nervous system of human brain cells works in processing information [16]. The key element of this technique is the structure of the information processing system, which is unique and diverse for each application. The term artificial here is used because this NN is implemented using a computer program that can complete several calculation processes during the learning process [17].

NNs with multiple layers have one or more layers located between the input and output layers called hidden layers (HL) [18]. These multi-layer networks can solve more challenging and rigorous problems with non-linear data. Figure 1 shows the concept of a multi-layer NN with n inputs and HL [19]. The coefficients attached to these input variables are called weights. The forecast is obtained from a linear combination of the inputs. The weights are selected in the NN framework using a learning algorithm until a minimum cost function value such as mean absolute error (MAE) or mean squared error (MSE) is obtained.

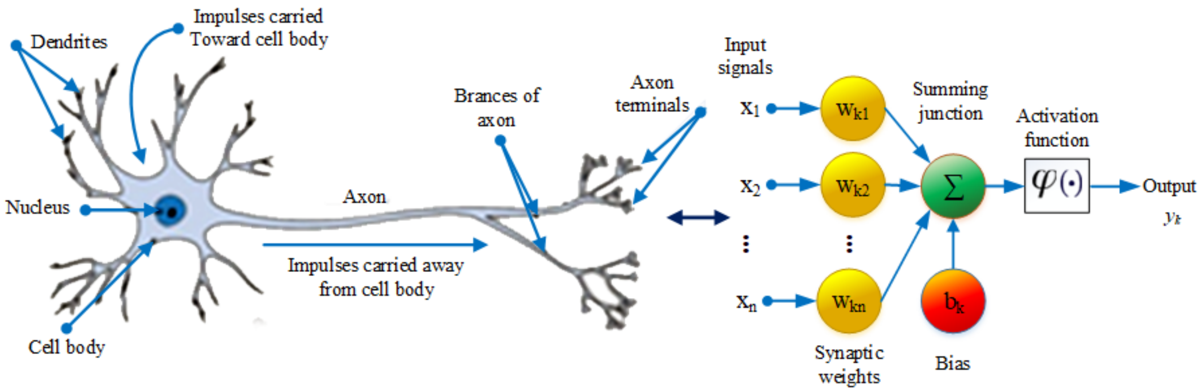

Figure 1. Neural network concept

NNAR is an NN method that uses lag values from time series data as input variables to the NN. Rob J Hyndman and George Athanasopoulos introduced the NNAR method in 2018 [20]. The model is denoted by NNAR(p, k), where p denotes the p-lag as input and k as nodes in the HL. The NNAR(p, k) model indicates that there are p-lags of the last observation $(y_{t-1}, y_{t-2}, \dots, y_{t-p})$ used as input variables to forecast the output Z_t, and k nodes in the HL. In general, NNAR(p, P, k) and k neurons or nodes in the HL. Models that omit the HL with the setting $k = 0$, NNAR$(p, P, 0)$ are analogous to the seasonal ARIMA$(p, 0, 0)$ $(P, 0, 0)$ model, and NNAR$(p, 0)$ is the same as the AR(p) model but with a non-linear function [21].

2.4 Methodology

In this research, the computational process uses the RStudio program. The methodology carried out in this study can be seen in Figure 2 as follows.

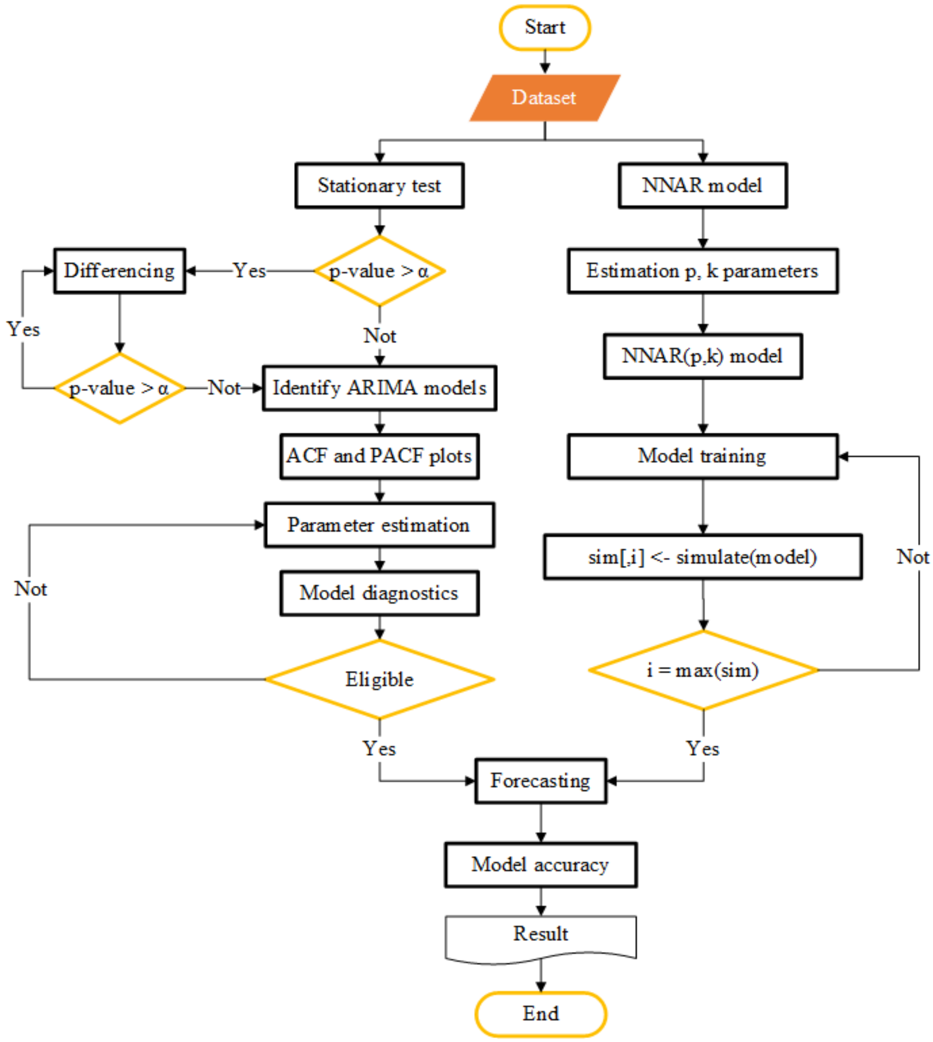

Figure 2. Methodology

3. RESULTS AND DISCUSSION

The first step taken is data pre-processing. Data pre-processing is done because the data obtained is incomplete, contains noise, or is missing so that the dataset used is complete and of high quality. Furthermore, the ARIMA and NNAR models use the dataset for input variables. Characteristic testing is done by utilizing the terasvirta test function available in RStudio. The hypothesis is:

H_0: Linear data.
H_1: Non-linear data.

The terasvirta neural network test results on the dataset obtained p-value = 0.003218. Based on the test results, where p-value < 0.05, H_0 is rejected, and H_1 is accepted. The conclusion is that the data is non-linear, so it is aligned with the research objectives. Furthermore, the dataset is used for input variables in the ARIMA and NNAR models.

3.1 ARIMA Model

Generally, time series data is non-stationary. In the ARIMA model, time series data must be stationary; not stationary data is changed to stationary by differencing. A plot of the time series dataset is done to visually determine stationary data and see if the data distribution has a trend or seasonal pattern. The following plot of historical daily gold price data is shown in Figure 3.

Figure 3. Historical gold price

Stationarity tests of time series are identified visually using plots of the autocorrelation function (ACF), partial autocorrelation function (PACF), and the Augmented Dickey-Fuller (ADF) test. When the dataset is not stationary, appropriate transformation and differentiation are then performed to produce a stationary time series. The number of ACF and PACF lags in the stationary time series is then used to determine the order of the ARIMA model. Parameter estimation is then performed on several ARIMA models to be tested. Diagnostics are carried out to check the parameters' significance and whether the forecasting residuals have averages that are not significantly different from zero and are uncorrelated. If these assumptions are met, the model can be used for forecasting. Otherwise, it can return to the identification stage for a better model. The following hypothesis criteria are carried out.

H_0: Stationary data.
H_1: Data is not stationary.

The Augmented Dickey-Fuller test results on the dataset obtained Dickey-Fuller value = -1.7725, Lag order = 11, p-value = 0.6746. Based on the test results, p-value = 0.6746, where p-value > 0.05, then H_0 is rejected and H_1 is accepted. The conclusion is that the data is not stationary and must be differenced to make the data stationary.

The initial step of ARIMA time series modeling is identifying the dataset on the variables used. This identification step aims to determine whether the data used has met the assumption of stationary in the mean. The ARIMA model uses stationary time series data to produce accurate short-term forecasting. The results of the stationary test on the dataset have been carried out; it is concluded that the dataset is not stationary, so the differencing stage must be carried out so that all variables used meet the assumption of stationary in the mean. The dataset plot after differencing is presented in **Figure 4** as follows.

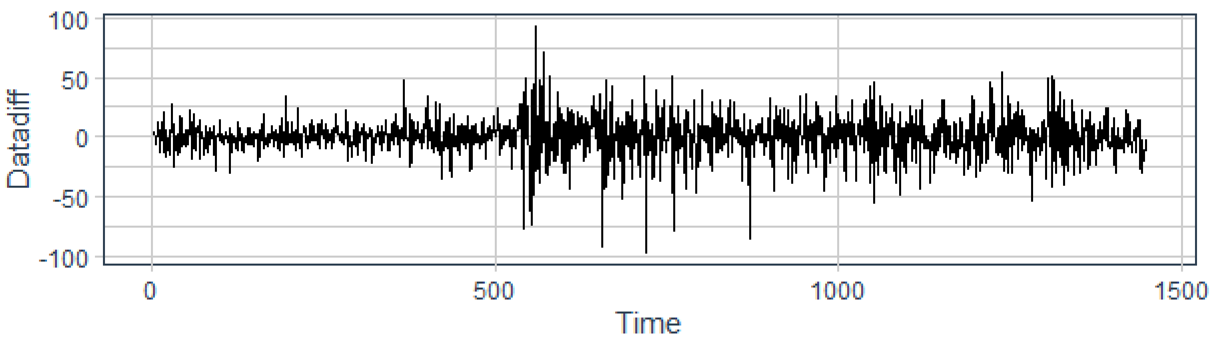

Figure 4. Differencing 1

Figure 4 shows that after differing the dataset once, the variables have relatively small fluctuations, and the trend pattern disappears. Visually, it can be indicated that the data is stationary in the mean. However, some data still spikes very low and very high. The ARIMA model identification stage has been carried out, so the data has met the assumption of stationarity in the mean. The ADF test on the differencing dataset will further ensure the data is stationary. The hypothesis criteria are.

H_0: The data is stationary.
H_1: The data is not stationary.

Based on the ADF test results on the differencing dataset, the Dickey-Fuller value = -10.92, Lag order = 11, p-value = 0.01. Based on the ADF test results, p-value = 0.01, where p-value < 0.05, then H_1 is rejected. The conclusion is that the data is stationary at differencing = 1, as well as getting the parameter estimation value d = 1. Determination of the ARIMA order can be done with ACF and PACF plots on the differencing dataset. **Figure 5** shows the different ACF and PACF plots of the dataset.

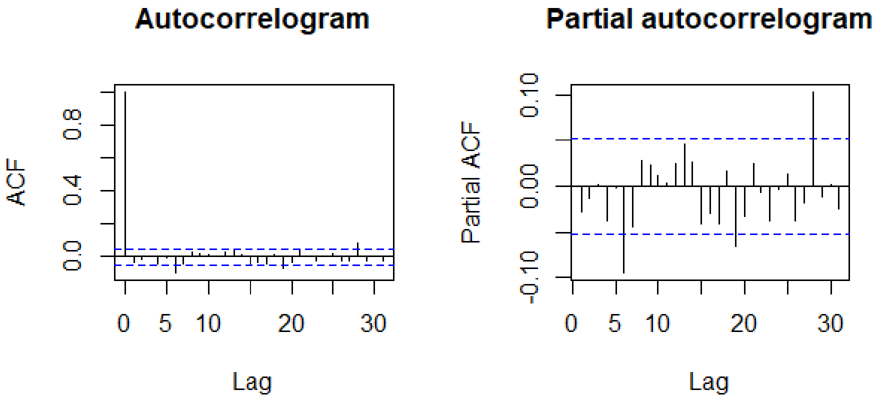

Figure 5. ACF and PACF plot of differencing 1

After identifying the parameters, several conjecture models or candidate ARIMA model parameters are determined. Here, in **Table 1**, the estimated parameters of the ARIMA model are shown.

Table 1. Parameters Estimates

No	Parameters			Model candidate
	p	***d***	***q***	
1	1	1	1	ARIMA(1,1,1)
2	2	1	2	ARIMA(2,1,2)
3	3	1	2	ARIMA(3,1,2)
4	2	1	3	ARIMA(2,1,3)

The estimated parameters must then be tested to determine their significance in the model, which is to see whether the parameters of the estimated results are significant in the model or insignificant. The estimated parameters of the ARIMA model are then estimated using the z-test of coefficients method. The parameter estimation results on the candidate model can be seen in **Table 2** as follows.

Table 2. Z test of Coefficients Parameter

Models	*z test of coefficients*:					Results
	Parameters	**Estimated**	**Std. Error**	***z* value**	$Pr(>\|z\|)$	
ARIMA(1,1,1)	AR1	0.830230	0.135380	6.1328	8.64×10^{-10}	Significant
	MA1	-0.858740	0.124630	-6.8900	5.58×10^{-12}	Significant
ARIMA(2,1,2)	AR1	0.027764	0.237268	0.1170	0.9068466	Not
	AR2	0.644950	0.205705	3.1353	0.0017167	Significant
	MA1	-0.047449	0.229233	-0.2070	0.8360167	Not
	MA2	-0.680039	0.200373	-3.3939	0.0006892	Significant
ARIMA(3,1,2)	AR1	0.066627	0.225347	0.2957	0.7674869	Not
	AR2	0.653999	0.185394	3.5276	0.0004193	Significant
	AR3	0.017726	0.030085	0.5892	0.5557243	Not
	MA1	-0.095618	0.223887	-0.4271	0.6693199	Not
	MA2	-0.687324	0.179890	-3.8208	0.0001330	Significant
ARIMA(2,1,3)	AR1	0.082392	0.232201	0.3548	0.7227148	Not
	AR2	0.666670	0.184653	3.6104	0.0003057	Significant
	MA1	-0.110325	0.232679	-0.4742	0.6353921	Not
	MA2	-0.698990	0.178264	-3.9211	8.81×10^{-05}	Significant
	MA3	0.016834	0.029856	0.5638	0.5728623	Not

Based on **Table 2**, the results of the z-test of coefficients on the candidate models show that only one model has significant parameter coefficients, namely the $\text{ARIMA}(1,1,1)$ model. This means that this model was chosen for forecasting. Furthermore, the candidate ARIMA model's akaike information criterion (AIC) value is presented.

Table 3. AIC Value

No	Models	Log likelihood	AIC
1	$\text{ARIMA}(1,1,1)$	-6088.21	12182.41
2	$\text{ARIMA}(2,1,2)$	-6087.95	12185.90
3	$\text{ARIMA}(3,1,2)$	-6087.77	12187.54
4	$\text{ARIMA}(2,1,3)$	-6087.79	12187.59

The best model is determined based on the smallest AIC value. The AIC acquisition results in **Table 3** show that the AIC value of the $\text{ARIMA}(1,1,1)$ model is smaller than the AIC of other model candidates. This indicates that $\text{ARIMA}(1,1,1)$ is the best model compared to other model candidates, this is also reinforced by the results of the Z test of coefficient, where this model has all significant coefficients.

Based on the results of the Z test of coefficients and the smallest AIC value, the $\text{ARIMA}(1,1,1)$ model is the most effective forecasting model. Next, model diagnostics will be carried out on $\text{ARIMA}(1,1,1)$. Model diagnostics are intended to determine whether the residuals of the model built have met the modeling assumptions. To ensure the suitability of the $\text{ARIMA}(1,1,1)$ model, several assumption tests are performed on the residuals.

1. The autocorrelation test is used to detect autocorrelation in residuals using the Ljung-Box test. The hypothesis criteria are:
   H_0: There is no autocorrelation in the residuals.
   H_1: There is autocorrelation in the residuals.
2. Test the mean of residuals to see whether the mean of residuals is equal to zero using a one-sample t-test. Hypothesis criteria are:
   H_0: $\mu = 0$.
   H_1: $\mu \neq 0$.
3. Normality test to test the normality of residuals using the Kolmogorov-Smirnov test. The hypothesis criteria are:
   H_0: Residuals are normally distributed.
   H_1: Residuals are not normally distributed
4. The variance homogeneity test is used to test the homogeneity of variance using the Breusch-Pagan test. Hypothesis criteria are:
   H_0: Residual variance is constant (homoscedasticity).
   H_1: Residual variance is not constant (heteroscedasticity).

The following **Table 4** presents the results of the assumption test on the residuals of the $\text{ARIMA}(1,1,1)$.

Table 4. Residual test

Test	$p-value$	Hypothesis
Ljung-Box	0.9537	H_0 is accepted
One Sample	0.3177	H_0 is accepted
Kolmogorov-Smirnov	2.2×10^{-16}	H_0 is rejected
Studentized Breusch-Pagan	2.6×10^{-10}	H_0 is rejected

Table 4 shows the results of the assumption test on the residuals of the $\text{ARIMA}(1,1,1)$ model. The autocorrelation test yields a p-value of 0.9537, since the p-value is greater than 0.05, H_0 is accepted. This proves that there is no autocorrelation in the residuals. The residual mean test p-value is 0.3177, meaning that the mean of the residuals is equal to zero. The normality test results in a p-value = 2.2×10^{-16}, this means the residuals are not normally distributed. Therefore, H_0 is rejected. The homogeneity of variance test resulted in a p-value = 2.6×10^{-10}. The p-value not exceeding 0.05 indicates heteroscedasticity in the residuals.

Only the autocorrelation and residual mean tests are fulfilled based on the residual test. The normality test and homogeneity of variance test are not fulfilled. This indicates that the residuals of the $\text{ARIMA}(1,1,1)$ model have heteroscedasticity, or the residual variance is not constant. The inability to fulfill the normality assumption is a signal that the ARIMA model may fail to read the data pattern properly. This is an indication that the ARIMA model may not fully fit the dataset. Since the purpose of this study is to analyze short-term forecasting comparisons and the ARIMA$(1,1,1)$ model is the only significant parameter coefficient, this model is used for forecasting.

3.2 NNAR Model

In time series data, the lag value of the time series can be used as an input variable in the NN. The NNAR model is denoted by NNAR(p,k), where p denotes the lag p used as input. k denotes the number of neurons in one HL. In this study, the nnetar function available in the RStudio library is used to identify the values of p and k. NNAR is a non-linear autoregressive model, and it should be too difficult to solve analytically due to the complexity of the model. Therefore, the model uses simulation to simulate future sample paths iteratively. Here is a simulation of five possible future sample paths for the gold price dataset. Each sample path covers the following 30 periods after the observed data, as shown in **Figure 6**.

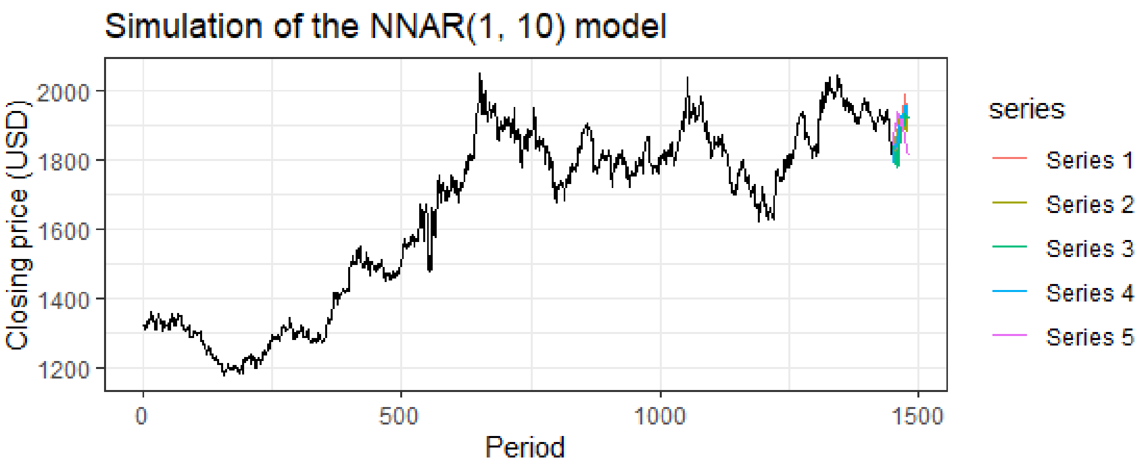

Figure 6. Simulation of the NNAR$(1,10)$ model

In the NNAR model, repeated simulations of the sample path are performed, so that the model can build knowledge about the distribution of all future values based on the fitted NN. Therefore, it is recommended that the NN model be run multiple times, and 20 is the minimum requirement. Experiments were conducted by changing the number of neurons (k) to get the best NNAR model. **Table 5** shows the experiments and the accuracy of the NNAR model.

Table 5. Experiments NNAR models

Model	NNAR$(1,1)$	NNAR$(1,2)$	NNAR$(1,10)$	NNAR$(1,20)$
RMSE	16.17115	16.12521	16.10002	16.10444
MAE	11.16473	11.11188	11.09360	11.09739

Table 5 shows the experiments by increasing the number of k of the NNAR model. It can be seen that increasing the number of k affects the accuracy of the model, where the more k the better the accuracy. However, too many k will cause bias due to high computational load, resulting in poor accuracy. This can be seen from the NNAR$(1,10)$ model, which obtained the best accuracy and decreased when the k increased as seen in the NNAR$(1,20)$ model.

3.3 Forecasting

Based on **Table 3**, the selected candidate ARIMA model is the ARIMA$(1,1,1)$ model. This model will be used in forecasting gold prices for the next four periods. The forecasting results of the ARIMA$(1,1,1)$ model will be compared with the forecasting results of the NNAR$(1,10)$ model which is the best NNAR model based on **Table 5.** The following **Table 6** shows the forecasting results of the ARIMA$(1,1,1)$ and NNAR$(1,10)$ models for the next four periods.

Table 6. Forecasting Results in Four Periods Ahead

Period	ARIMA(1, 1, 1)	NNAR(1, 10)
1450	1818.140	1816.811
1451	1819.418	1817.021
1452	1820.480	1817.230
1453	1821.361	1817.438

The following **Figure 7** and **Figure 8** respectively, show plots of actual and forecast data with the ARIMA(1, 1, 1) and NNAR(1, 10) models as follows.

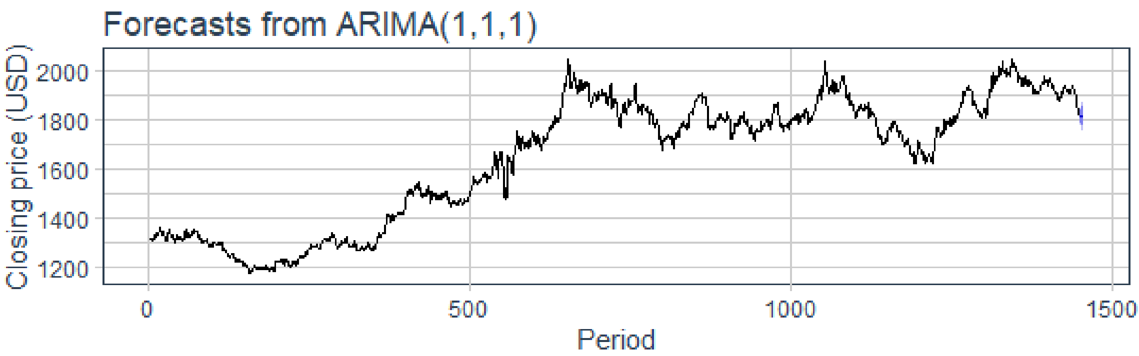

Figure 7. Data and forecasts of ARIMA(1, 1, 1) model

Figure 8 shows a graph of actual data and the results of forecasting for the following four periods using the ARIMA(1, 1, 1) model. It can be seen that the forecasting results of this model increase slowly. The ARIMA(1, 1, 1) model in period 1450 produced a forecast of 1818.140. This value is greater than the output of the NNAR(1, 10) model in the same period. Overall, the results of ARIMA(1, 1, 1) forecasting in each period are greater than those of the NNAR(1, 10). **Figure 9** displays the dataset and forecasting results using the NNAR(1, 10) model as follows.

Figure 8. Data and forecasts of NNAR(1, 10) model

Based on **Figure 7** and **Figure 8**, the plot of forecasting results from both models tends to rise slowly. This means that gold prices are expected to increase from 1450 to 1453. The following **Table 7** shows the accuracy obtained from each model. Model accuracy based on RMSE and MAE values.

Table 7. Model accuracy

Model	RMSE	MAE
ARIMA(1, 1, 1)	16.20431	11.13958
NNAR(1, 10)	16.10002	11.09360

Table 7 presents the accuracy of the two models used, namely the ARIMA(1, 1, 1) model and the NNAR(1, 10) model. The ARIMA model works by combining three components: autoregression (AR), differencing (I), and moving average (MA). The parameters (1, 1, 1) indicate that the model uses one lag in the AR and MA components and one level of differencing to make the data stationary. The ARIMA(1, 1, 1) model obtained forecasting accuracy RMSE = 16.20431 and MAE = 11.13958. The NNAR model is a variant of NN adapted for time series analysis. The NNAR model utilizes the ability of NNs to capture non-linear patterns in the data, which often cannot be captured by traditional linear models such as ARIMA. The

parameters $(1,10)$ indicate that the model utilizes one input lag and ten neurons in the HL. The NNAR$(1,10)$ model produces forecasting accuracy with a value of RMSE = 16.10002 and MAE = 11.09360. **Table 5** shows the experimental results of the NNAR model by changing the parameter settings. The best model obtained is the NNAR$(1,10)$ model and the NNAR $(1,1)$ model is the model with the worst accuracy when compared to other NNAR models. It can be seen that all NNAR models produce better accuracy when compared to the ARIMA$(1,1,1)$ model. The results show that the NNAR model is slightly more accurate than the ARIMA model regarding RMSE and MAE, although the difference is insignificant. **Figure 9** shows the last ten periods of actual data and their forecast values, as well as the results of forecasting the gold closing price for the next four periods using the ARIMA$(1,1,1)$ and NNAR$(1,10)$ models.

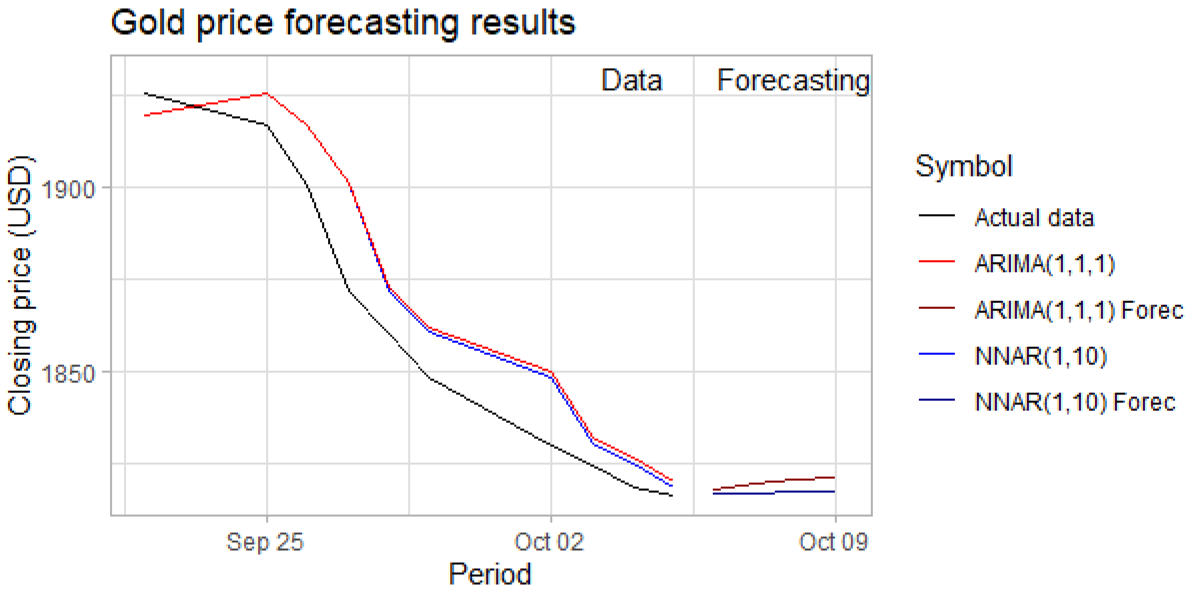

Figure 9. Gold price forecasting results for the next four periods

Figure 9 shows that the forecasting results of the NNAR$(1,10)$ model on the data are closer to the actual data than the forecasting results of the ARIMA$(1,1,1)$ model. Then these two models will perform gold price forecasting for the following four periods, starting from October 6, 2023 to October 9, 2023. The ARIMA$(1,1,1)$ model forecasting are symbolized by a dark red line, and the results of the NNAR(1,10) model forecasting are symbolized by dark blue. The ARIMA$(1,1,1)$ model produces forecasts that increase from period 1450 to period 1453, and the NNAR$(1,10)$ model produces forecasting values that increase slowly. This demonstrates the potential of NNs to improve the accuracy of time series forecasting, especially in the case of data that may have non-linear components, making it suitable for gold price forecasting tasks, especially in volatile markets due to economic uncertainty. This research demonstrates the potential use of NN models in time series analysis.

4. CONCLUSIONS

In this study, gold price forecasting was carried out using the ARIMA$(1,1,1)$ and NNAR$(1,10)$ models. Based on the accuracy value obtained, the NNAR model outperforms the ARIMA$(1,1,1)$ model. The NNAR$(1,10)$ model produces RMSE = 16.10002 and MAE = 11.09360. This is smaller than the accuracy value of the ARIMA$(1,1,1)$ model with RMSE = 16.20431 and MAE = 11.13958. Although the parameter coefficients of the ARIMA$(1,1,1)$ model are all significant, the results of the residual normality test and the variance homogeneity test are not fulfilled. This indicates patterns in the data that have not been captured by the ARIMA$(1,1,1)$ model. If the variance is not constant or heteroscedasticity, it indicates that the ARIMA$(1,1,1)$ model does not fit the data or other factors have not been captured by the model. Heteroscedasticity can cause the forecasting accuracy of the model to be less accurate because it is susceptible to outliers. Based on the analysis results, the ARIMA method may not be suitable for datasets with complex or non-linear characteristics.

The NNAR model is suitable for time series data with non-linear characteristics that cannot be captured well by linear models such as ARIMA. Not bound by the assumptions of linearity or normality, NNAR models are often more robust in dealing with complex and time-varying data. The main focus of NNAR models is on good predictive ability based on accuracy metrics.

From the results of this study, both ARIMA and NNAR models have advantages in time series forecasting. ARIMA is better at handling linear patterns and stationary data. The NNAR model is better at capturing non-linear patterns. NNAR is very flexible and can capture non-linear patterns in the data, so this model is less dependent on classical assumptions such as residual normality. The NNAR model focuses more on the ability to effectively forecast time series data, especially in cases where the data has non-linear patterns. Therefore, model selection should be tailored to the characteristics of the data and forecasting objectives.

ACKNOWLEDGMENT

The authors are grateful to Lembaga Penelitian dan Pengabdian kepada Masyarakat (LPPM) Universitas Jenderal Achmad Yani (UNJANI) for providing an internal research grant.

REFERENCES

[1] M. As'ad, S. Sujito, and S. Setyowibowo, "Neural Network Autoregressive For Predicting Daily Gold Price," *J. Inf.*, vol. 5, no. 2, p. 69, 2020, doi: https://doi.org/10.25139/inform.v0i1.2715.

[2] P. Hajek and J. Novotny, "Fuzzy Rule-Based Prediction of Gold Prices using News Affect," *Expert Syst. Appl.*, vol. 193, p. 116487, 2022, doi: https://doi.org/10.1016/j.eswa.2021.116487.

[3] I. Livieris, E. Pintelas, and P. Pintelas, "A CNN-LSTM model for gold price time series forecasting," *Neural Comput. Appl.*, vol. 32, Dec. 2020, doi: https://doi.org/10.1007/s00521-020-04867-x.

[4] I. K. Hasan and Ismail Djakaria, "Perbandingan Model Hybrid ARIMA-NN dan Hybrid ARIMA-GARCH untuk Peramalan Data Nilai Tukar Petani di Provinsi Gorontalo," *J. Stat. dan Apl.*, vol. 5, no. 2, pp. 155–165, 2021, doi: https://doi.org/10.21009/jsa.05204.

[5] G. Zhang, B. Eddy Patuwo, and M. Y. Hu, "Forecasting with artificial neural networks: The state of the art," *Int. J. Forecast.*, vol. 14, no. 1, pp. 35–62, 1998, doi: https://doi.org/10.1016/S0169-2070(97)00044-7.

[6] G. E. P. Box and G. M. Jenkins, *Time Series Analysis: Forecasting and Control*. Holden-Day, 1976.

[7] R. Thoplan, "Simple v/s Sophisticated Methods of Forecasting for Mauritius Monthly Tourist Arrival Data," *Int. J. Stat. Appl.*, vol. 4, no. 5, pp. 217–223, 2014, doi: https://doi.org/10.5923/j.statistics.20140405.01.

[8] A. Ahmar and E. Boj, "Application of Neural Network Time Series (NNAR) and ARIMA to Forecast Infection Fatality Rate (IFR) of COVID-19 in Brazil," *JOIV Int. J. Informatics Vis.*, vol. 5, p. 8, Mar. 2021, doi: https://doi.org/10.30630/joiv.5.1.372.

[9] G. Vijayalakshmi, K. Pushpanjali, and A. Mohan Babu, "A comparison of ARIMA & NNAR models for production of rice in the state of Andhra Pradesh," *Int J Stat Appl Math*, vol. 8, no. 3, pp. 251–257, 2023, doi: https://doi.org/10.22271/maths.2023.v8.i3c.1041.

[10] H. Y. Liu, A. Manzoor, C. Wang, L. Zhang, and Z. Manzoor, "The COVID-19 outbreak and affected countries stock markets response," *International Journal of Environmental Research and Public Health*, vol. 17, no. 8. 2020. doi: https://doi.org/10.3390/ijerph17082800.

[11] M. Melina, Sukono, H. Napitupulu, A. Sambas, A. Murniati, and V. A. Kusumaningtyas, "Artificial neural network-based machine learning approach to stock market prediction model on the Indonesia Stock Exchange during the COVID-19," *Eng. Lett.*, vol. 30, no. 3, pp. 988–1000, 2022.

[12] S. C. Hillmer and W. W. S. Wei, "Time Series Analysis: Univariate and Multivariate Methods," *J. Am. Stat. Assoc.*, vol. 86, no. 413, p. 245, 1991, doi: https://doi.org/10.2307/2289741.

[13] F. M. Khan and R. Gupta, "ARIMA and NAR based prediction model for time series analysis of COVID-19 cases in India," *J. Saf. Sci. Resil.*, vol. 1, no. 1, pp. 12–18, 2020, doi: https://doi.org/10.1016/j.jnlssr.2020.06.007.

[14] M. L. Ayala and D. L. L. Polestico, "Modeling COVID-19 cases using NB-INGARCH and ARIMA models: A case study in Iligan City, Philippines," *Procedia Comput. Sci.*, vol. 234, pp. 262–269, 2024, doi: https://doi.org/10.1016/j.procs.2024.03.012.

[15] T. Umairah, N. Imro'ah, and N. M. Huda, "Arima model verification with outlier Factors using control chart," *BAREKENG J. Math. Its Appl.*, vol. 18, no. 1, pp. 0579–0588, 2024, doi: https://doi.org/10.30598/ barekengvol18iss1pp0579 - 0588.

[16] D. F. J. de C. C. Sandoval U, D. F. J. de C. Bogotá U, and S. J, "Computational Models of Financial Price Prediction: A Survey of Neural Networks, Kernel Machines and Evolutionary Computation Approaches," *Ingeniería*, vol. 16, no. 2, pp. 125–133, Jul. 2011.

[17] S. Haykin, *Neural Networks and Learning Machines*, Third Edit. New York: Pearson Education, Inc, 2009.

[18] M. Melina, Sukono, H. Napitupulu, and N. Mohamed, "Modeling of Machine Learning-Based Extreme Value Theory in Stock Investment Risk Prediction: A Systematic Literature Review," *Big Data*, pp. 1–20, Jan. 2024, doi: https://doi.org/10.1089/big.2023.0004.

[19] Melina, Sukono, H. Napitupulu, and N. Mohamed, "A conceptual model of investment-risk prediction in the stock market using extreme value theory with machine learning: A semisystematic literature review," *Risks*, vol. 11, no. 3, pp. 1–24, 2023, doi: https://doi.org/10.3390/risks11030060.

[20] R. J. Hyndman and G. Athanasopoulos, *Forecasting: principles and practice*, 3rd ed. Melbourne, Australia: OTexts, 2021. [Online]. Available: https://otexts.com/fpp3/

[21] C. Twumasi and J. Twumasi, "Machine learning algorithms for forecasting and backcasting blood demand data with missing values and outliers: A study of Tema General Hospital of Ghana," *Int. J. Forecast.*, vol. 38, no. 3, pp. 1258–1277, 2022, doi: https://doi.org/10.1016/j.ijforecast.2021.10.008.