Gold price modeling in Indonesia using ARFIMA method

D Safitri¹, Mustafid¹, D Ispriyanti¹, Sugito¹

¹Department of Statistics, Diponegoro University
Jl. Prof. Soedharto, SH, Tembalang, Semarang 50275, Indonesia
E-mail: diahsafitri.fifi@gmail.com

Abstract. Gold investment is the best choice to control finance. Gold is easy to resell if there is a financial need at the unpredictable moment. The data of gold price in Indonesia is a long-term memory data series or a time series data that has a long-term dependency. ARFIMA model is an appropriate model for such long-term memory data series since ARFIMA model has a different parameter value (d) of integer while the value of d in ARFIMA model is non-integer value due to the long-term memory or the long-term dependency. This research aims to obtain the best ARFIMA model of gold price data in Indonesia. It is obtained the value of ARFIMA model (1,d,[3]) with d=1,05716 as the best model.

1. Introduction
Gold industry does not experience any fluctuations even though it is in the sluggish economy because people like to buy gold for investment and social status. One of the benefits of gold investment is, when people have a financial need or when they want to change the model of jewelry, it is easy to sell gold. Gold price forecast is important since there are lots of gold transactions (purchase and sale) in the community.

Forecasting is to estimate the value of a future variable with reference to the present and past values. Time series method is a forecasting method that always experiences development. At the time series, past data sets are used as a reference for forecasting.

The Autoregressive Integrated Moving Average (ARIMA) model is a model for non-stationary time series data with Autocorrelation Function (ACF) plot that decreases exponentially. In addition, there is a non-stationary time series data whose ACF plot decreases non-exponentially but slowly, such data is called long-memory data. The Autoregressive Fractionally Integrated Moving Average (ARFIMA) model is the right model for time series data with long-term memory. The ARIMA model has a differencing parameter (d) with an integer value, whereas d in the ARFIMA model is not an integer due to time series with long-term memory or long-term dependency.

The gold price in Indonesia can be forecasted for some time in the future based on the previous sets of gold prices. The results of such forecasting are useful for gold users, either those using gold for jewelry or those using it for investment. Besides, the results can also serve as an input to any relevant agencies in issuing any gold-related policies. Based on this background of study, the researchers were interested in conducting a study entitled "Gold Pricing Model in Indonesia using ARFIMA Model."

Gold price data in Indonesia is a time series data with long-term memory or time series data which has a long-term dependency. ARFIMA model is the right model for time series data with long-term memory because the differencing parameter (d) in this model is fractional (not integer) due to the time series data with long-term memory or long-term dependency. The problem in this study is what is the best ARFIMA model from the gold price data in Indonesia.
This study aimed to create a time series model for the gold price in Indonesia, more specifically to create ARFIMA model based on gold price data in Indonesia. ARFIMA model is used because the gold price in Indonesia is a time series data with long-term memory.

The study to obtain a gold price model in Indonesia is important because gold investment is the best option to manage finances. It is easy to sell gold when there are any financial needs at any time, but many people do not understand the fluctuations in gold prices.

The results of this study are expected to serve as an input to develop a time series model for gold prices in Indonesia, especially the ARFIMA model. The targeted findings and outcomes of this study are time series model, particularly the best ARFIMA time series model for gold prices in Indonesia

2. Literature Review

Research on gold price forecasting in Indonesia has been conducted several times. [1] analyzed ARIMA modeling for gold price forecasting; [2] analyzed gold price forecasting using Feedforward Neural Network with backpropagation algorithm; [3] analyzed world’s gold price prediction to encourage gold stock investment decisions using data mining techniques; [4] conducted a research on gold price forecasting using fuzzy time series Markov chain model.

Forecasting is an important element in decision making; whether a decision is effective or not depends on several factors that are not observable when the decision is made. Time series analysis is one of the forecasting methods. A characteristic of time series analysis is: a series of observations on a variable are seen as the realization of a random variable with a joint distribution [5].

In [6] stationary process \( \{Z_t\} \) has a mean \( E\{Z_t\} = \mu \), variance \( Var(Z_t) = E(Z_t - \mu)^2 = \sigma^2 \), covariance between \( Z_t \) and \( Z_{t+k} \):

\[
\gamma_k = Cov(Z_t, Z_{t+k}) = E[(Z_t - \mu)(Z_{t+k} - \mu)]
\]

and correlation between \( Z_t \) and \( Z_{t+k} \):

\[
\rho_k = \frac{Cov(Z_t, Z_{t+k})}{\sqrt{Var(Z_t)Var(Z_{t+k})}} = \frac{\gamma_k}{\gamma_0}
\]

In a time series analysis, \( \gamma_k \) is called as autocovariance function, and \( \rho_k \) is called an autocorrelation function (ACF).

For stationary process, autocovariance function \( \gamma_k \) and autocorrelation function \( \rho_k \) have characteristics:

1. \( \gamma_0 = Var(Z_t) ; \rho_0 = 1 \)
2. \( |\gamma_k| \leq \gamma_0 ; |\rho_k| \leq 1 \)
3. \( \gamma_k = \gamma_{-k} ; \rho_k = \rho_{-k} \), for all k values.
4. \( \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j \gamma_{|t-i|} \geq 0 \) and \( \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j \rho_{|t-i|} \geq 0 \) for some time \( t_1, t_2, \ldots, t_n \) and for real numbers \( \alpha_1, \alpha_2, \ldots, \alpha_n \)

The steps undertaken in selecting the model are the postulation of general class of model, model identification, model parameter estimation, model parameter testing, model diagnostic testing, and selection of the best model [5].

According to [7] long-term memory pattern can be observed from the autocorrelation value on the ACF or PACF plot that decreases slowly for an increasing lag. [8] stated that Hurst statistics are used to determine whether a data has a long-term memory pattern or not.

According to [6] Autoregressive Fractionally Integrated Moving Average (ARFIMA) model has three parameters, namely \( p \) is autoregressive parameter, \( d \) is an integrated value which is a real number or not an integer, and \( q \) is a moving average parameter. ARFIMA model \( (p,d,q) \):
\[ \phi_p(B)(1 - B)^d Z_t = \theta_q(B)a_t \]

where:
- \( d \) = differencing parameter
- \( \phi_p(B) = (1 - \phi_1 B - \ldots - \phi_p B^p) \) is operator AR (p)
- \( \theta_q(B) = (1 - \theta_1 B - \ldots - \theta_q B^q) \) is operator MA (q)
- \((1 - B)^d\) is differencing operator

Time series model with long-memory is in which deviations from the long run mean decay more slowly than an exponential decay [9]

In [10] mentioned that Hurst Exponent could be calculated using rescaled range analysis (R/S analysis). For time series data \( Z_1, Z_2, \ldots, Z_n \), R/S analysis model can be used using the following steps:

1. Calculating the value of \( \bar{Z} \):
   \[ \bar{Z} = \frac{1}{n} \sum_{t=1}^{n} Z_t \]

2. Calculating the mean adjusted
   \[ Y_t = Z_t - \bar{Z}, \quad t = 1, 2, \ldots, n \]

3. Calculating the cumulative deviation
   \[ G_t = \sum_{j=1}^{t} Y_j, \quad t=1, 2, \ldots, n \]

4. Calculating the value of Range
   \[ R_t = \max(G_t, G_2, \ldots, G_t) - \min(G_1, G_2, \ldots, G_t), \quad t = 1, 2, \ldots, n \]

5. Calculating the standard deviation
   \[ S_t = \sqrt{\frac{1}{t} \sum_{j=1}^{t} (Z_j - \bar{Z})^2}, \quad t = 1, 2, \ldots, n \]
   where \( \bar{u} \) is the mean of \( X_1 \) to \( X_t \)

6. Calculating the rescaled range (R/S)
   \[ (R/S)_t = \frac{R_t}{S_t}, \quad t = 1, 2, \ldots, n \]

7. Determining the Hurst (H) exponent using R/S statistics from time series data.
   \[ (R/S)_t = c \cdot t^H \quad \text{with} \; t=1,2,\ldots,n \]
   where
   - \( c = \text{constant} \)
   - \( H = \text{Hurst Exponent} \)

The method used in estimating differentiating parameters (d) is the Geweke Porter-Hudak (GPH) method. In [8] GPH steps are as follows

1. Determine the value of \( m \) dan \( \omega \)
   \[ m = n^{0.5} \]
   \[ \omega_j = \frac{j}{n}, \quad j = 1, 2, \ldots, m \]
   where:
   - \( n \) = the amount of observation
   - \( \pi = 3.14 \)

2. Determine the value of \( \gamma_0 \)
3. Determine the value of $I_z(\omega_j)$

$$I_z(\omega_j) = \frac{1}{2\pi} (\gamma_0 + 2 \sum_{t=1}^{n-1} \gamma_t \cos(t \omega_j)), \quad j = 1, 2, \ldots, m$$

4. Determine $X_j$ as an independent variable

$$X_j = \ln \left( \frac{1}{4\sin^2(\omega_j/2)} \right), \quad j = 1, 2, \ldots, m$$

5. Determine $Y_j$ as a dependent variable

$$Y_j = \ln(I_z(\omega_j)), \quad j = 1, 2, \ldots, m$$

6. Form a regression equation between $X_j$ and $Y_j$, the value of the coefficient in parameter $X_j$ is the value of the differencing parameter estimate

3. Methodology of research

3.1. Data sources and research variables

This study uses secondary data, i.e. gold price data in Indonesia from PT. Antam (persero) Tbk from 2015 to 2017 in [11].

3.2. Data analysis method

The data in this study are processed using software Eviews 9, Minitab and R. The steps are undertaken in analyzing the data are as follows

a. Conducting stationarity tests in variance
b. Conducting stationarity tests in mean
c. Identifying whether the data had long-term memory pattern using Hurst (H) statistics
d. Estimating the differencing parameter (d) using Geweke & Porter-Hudak (GPH) method
e. Creating ACF plot from differencing data
f. Creating PACF plot from differencing data
g. Estimating parameter $\phi$ and $\theta$
h. Testing the parameter significance of ARFIMA model
i. Examining if the model was sufficient
j. Selecting the best model based on the lowest AIC value
k. Calculating the value of Mean Absolute Percentage Error (MAPE) for the established evaluation model.

4. Results and Discussions

4.1. Stationarity Test in Variance (Box-Cox)

Figure 1 shows that the rounded value = -1.00 ($\lambda=-1.00$), so it can be concluded that the gold price data in Indonesia was not stationary in variance, thus Box-Cox transformation is conducted using formula $1/Z_t$. 
Figure 1. Box-Cox plot of gold

Figure 2 shows that, after the transformation, the rounded value = 1.00 (\(\lambda=1\)) so the data is stationary in variance.

Figure 2. Box-Cox plot of data transformation

4.2. Unit Root Test (Pre-Differencing Data)
Augmented Dickey-Fuller (ADF) test is conducted using software Eviews 9 to find out if the data is stationary in mean.

Hypothesis:

- \(H_0: |\phi| = 1\) (data is not stationary)
- \(H_1: |\phi| < 1\) (data is stationary)

Significance Level:
\[ \alpha = 5\% \]

Test Statistics:
\[ DF = \phi - 1 \]
\[ SE(\phi) = -0.667256 \]
\[ P-value = 0.8523 \]

Decision:
\( H_0 \) is not rejected at a significance level of 5% because of \( P\)-value > \( \alpha \)

Conclusion:
Gold price data in Indonesia is not stationary in mean.

4.3. Identification of Long Memory Pattern
To determine whether the data has long-memory pattern can be done by looking at the ACF plot and calculating the Hurst statistics value. The ACF plot can be seen in Figure 3 (using software Minitab 14), it is observable that there was a long memory because the ACF plot decreased slowly and hyperbolically.

\[ \text{Figure 3. Long Memory Identification} \]

The Hurst value is at an interval \( 0.5 < H < 1 \); it can be concluded that there is a long memory. The calculation of Hurst statistics value is as follows:

1. Calculating the mean:
\[ \bar{Z} = \frac{1}{n} \sum_{i=1}^{n} Z_i = \frac{98.51487}{539} = 0.18277 \]

2. Calculating the mean adjusted
\[ Y_1 = Z_1 - \bar{Z} \]
\[ Y_1 = Z_1 - 0.18657 = -0.18277 - 0.00380 \]
\[ Y_2 = Z_2 - \bar{Z} = 0.18727 - 0.18277 = 0.00450 \]
\[ Y_3 = Z_3 - \bar{Z} = 0.16867 - 0.18277 = -0.01410 \]
\[ \ldots \]
\[ Y_{539} = Z_{539} - \bar{Z} = 0.16867 - 0.18277 = -0.01410 \]
3. Determining cumulative deviation

\[ G_t = \sum_{i=1}^{t} Y_i \]

\[ G_1 = \sum_{i=1}^{1} Y_i = Y_1 = 0.00380 \]

\[ G_2 = \sum_{i=1}^{2} Y_i = Y_1 + Y_2 = 0.00380 + (0.00450) = 0.00830 \]

\[ \vdots \]

\[ G_{539} = \sum_{i=1}^{539} Y_i = Y_1 + Y_2 + \ldots + Y_{539} = 0.00380 + 0.00450 + \ldots + (-0.01411) = 0.0000 \]

4. Calculating the Range (R) from cumulative deviation

\[ R_t = \max(G_1, G_2, \ldots, G_t) - \min(G_1, G_2, \ldots, G_t) \]

\[ R_1 = \max(G_1) - \min(G_1) = 0.00380 - 0.00380 = 0 \]

\[ R_2 = \max(G_1, G_2) - \min(G_1, G_2) = 0.00830 - 0.00380 = 0.00450 \]

\[ \vdots \]

\[ R_{539} = \max(G_1, G_2, \ldots, G_{539}) - \min(G_1, G_2, \ldots, G_{539}) = 1.50733 - 0.0000 = 1.50733 \]

5. Calculating the standard deviation

\[ S_t = \sqrt{\frac{1}{t} \sum_{i=1}^{t} (Z_i - \overline{Z})^2} \]

\[ S_1 = \sqrt{\frac{1}{1} \sum_{i=1}^{1} (Z_i - \overline{Z})^2} = 0.00380 \]

\[ S_2 = \sqrt{\frac{1}{2} \sum_{i=1}^{2} (Z_i - \overline{Z})^2} = 0.00416 \]

\[ \vdots \]

\[ S_{539} = \sqrt{\frac{1}{539} \sum_{i=1}^{539} (Z_i - \overline{Z})^2} = 0.00750 \]

6. Calculating the rescaled range (R/S)

\[ (R/S)_t = R_t / S_t \]

\[ (R/S)_1 = R_1 / S_1 = 0.0000 / 0.00380 = 0.0000 \]

\[ (R/S)_2 = R_2 / S_2 = 0.00450 / 0.00416 = 1.08050 \]

\[ \vdots \]

\[ (R/S)_{539} = R_{539} / S_{539} = 1.50733 / 0.00750 = 200.9212 \]

7. Determining the Hurst (H) exponent value through R/S statistics from the time series data.
\[
\frac{R}{S} = c \cdot t^H \\
\ln(R/S) = c + H \ln t
\]

The value of \(H\) was obtained using Ordinary Least Square (OLS) method:
\[
\hat{H} = \frac{\sum_{t=1}^{n} (X_t - \overline{X})(Y_t - \overline{Y})}{\sum_{t=1}^{n} (X_t - \overline{X})^2} = 0.77922
\]

There was long memory in the data because \(0.5 < 0.77922 \times 1\).

### 4.4. Estimating the Differencing Parameter (\(d\))

Differencing parameter \(d\) value is used for data differentiation intended to meet the stationarity in the mean. The estimation of \(d\) value is done using software R and GPH (Geweke and Porter-Hudak) method. The value of \(d\) is 1.05716. Then differentiation is conducted using the value of \(d\) that is obtained.

### 4.5. Unit Root Test (Post-Differencing Data)

Augmented Dickey-Fuller test is done to differencing data as follows:

**Hypothesis:**
- \(H_0: \phi = 1\) (data is not stationary)
- \(H_1: \phi < 1\) (data is stationary)

**Significance Level:**
\(\alpha = 5\%\)

**Test Statistics:**
\[
DF = \frac{\hat{\phi} - 1}{SE(\phi)} = -17.12295
\]

**p-value = 0.0000**

**Decision:**
\(H_0\) is not rejected at a significance level of 5\% because of p-value \(< \alpha\)

**Conclusion:**
Gold price data in Indonesia is stationary in the mean.

After the assumption of stationarity in the mean and variance had been fulfilled, then the ARFIMA model is identified by looking at the ACF and PACF plots from the transformation and differencing data, the possible ARFIMA models are ARFIMA (1,d,0), ARFIMA ([3],d,0), ARFIMA (0,d,1), ARFIMA (0,d,[3]), ARFIMA (1,d,1), ARFIMA (1,d,[3]), ARFIMA (1,d,[1,3]), ARFIMA ([3],d,1), ARFIMA ([3],d,[3]), ARFIMA ([3],d,[1,3]), ARFIMA ([1,3],d,1), ARFIMA ([1,3],d,[3]) and ARFIMA ([1,3],d,[1,3]).

### 4.6. Parameter Estimation and Significance Test of ARFIMA Model

The equation of ARFIMA model which is used is as follows:
\[
\phi_p (B)(1 - B)^d Z_t = \theta_q (B)a_t
\]

Parameter estimation is done by Exact Maximum Likelihood (EML) method using EViews 9. The results are shown in table 1.

| Model          | Parameter | Parameter Estimation |
|----------------|-----------|----------------------|
| ARFIMA (1,d,0) | d         | -0.138347            |
Table 2. P-value and conclusion

| Model                  | Parameter | P-value | Conclusion       |
|------------------------|-----------|---------|------------------|
| ARFIMA (1,d,0)         | d         | 0,999999| H0 rejected      |
|                        | $\phi_1$  | 0,999977|                  |
|                        | $\phi_3$  | 0,389042|                  |
| ARFIMA (0,d,1)         | $\theta_1$| 1,000000|                  |
|                        | d         | 0,498572|                  |
| ARFIMA (0,d,[3])       | $\theta_3$| 1,000000|                  |
|                        | d         | -0,082996|                 |
| ARFIMA (1,d,1)         | $\phi_1$  | 0,999997|                  |
|                        | $\theta_1$| -0,094386|                |
|                        | d         | -0,117320|                 |
| ARFIMA (1,d,[3])       | $\phi_1$  | 0,999999|                  |
|                        | $\phi_3$  | -0,086074|                |
|                        | d         | -0,012442|                 |
| ARFIMA (1,d,[3])       | $\theta_1$| -0,159231|                |
|                        | $\theta_3$| -0,110125|                |
|                        | d         | 0,290816 |                  |
| ARFIMA ([3],d,1)       | $\phi_3$  | 0,995364|                  |
|                        | $\theta_1$| 0,231465 |                  |
|                        | d         | 0,499828 |                  |
| ARFIMA ([3],d,[3])     | $\phi_3$  | 0,953258|                  |
|                        | $\theta_3$| -0,770119|                |
|                        | d         | 0,499914 |                  |
| ARFIMA ([3],d,[1,3])   | $\phi_3$  | 0,848057|                  |
|                        | $\theta_1$| 0,212895 |                  |
|                        | $\theta_3$| -0,678079|                |
|                        | d         | -0,071250|                 |
| ARFIMA ([1,3],d,1)     | $\phi_1$  | 1,008706|                  |
|                        | $\phi_3$  | -0,008709|                |
|                        | $\theta_1$| -0,116598|                |
|                        | d         | -0,031851|                 |
|                        | $\phi_1$  | 0,908543|                  |
| ARFIMA ([1,3],d,[3])   | $\phi_3$  | 0,091450 |                  |
|                        | $\theta_3$| -0,127247|                |
|                        | d         | 0,048745 |                  |
| ARFIMA                  | $\phi_1$  | 0,941086|                  |
| ([1,3],d,[1,3])        | $\phi_1$  | 0,058889|                  |
|                        | $\theta_1$| -0,157884|                |
|                        | $\theta_3$| -0,137381|                |
| Model                  | Parameter | Value | $H_0$ Status |
|------------------------|-----------|-------|--------------|
| ARFIMA ([3],d,0)       | $d$       | 0.0000| $H_0$ rejected |
|                        | $\phi_3$ | 0.0000| $H_0$ rejected |
| ARFIMA (0,d,1)         | $\theta_1$ | 0.2033| $H_0$ accepted |
| ARFIMA (0,d,[3])       | $\theta_3$ | 0.0429| $H_0$ rejected |
|                        | $d$       | 0.0585| $H_0$ accepted |
| ARFIMA (1,d,1)         | $\phi_1$ | 0.0000| $H_0$ rejected |
|                        | $\theta_1$ | 0.1039| $H_0$ accepted |
|                        | $d$       | 0.0000| $H_0$ rejected |
| ARFIMA (1,d,[3])       | $\phi_3$ | 0.0000| $H_0$ rejected |
|                        | $\theta_3$ | 0.0124| $H_0$ rejected |
|                        | $d$       | 0.8398| $H_0$ accepted |
|                        | $\phi_1$ | 0.0000| $H_0$ rejected |
| ARFIMA (1,d,[1,3])     | $\theta_1$ | 0.0386| $H_0$ rejected |
|                        | $\theta_3$ | 0.0073| $H_0$ rejected |
|                        | $d$       | 0.0000| $H_0$ rejected |
| ARFIMA ([3],d,1)       | $\phi_3$ | 0.0000| $H_0$ rejected |
|                        | $\theta_1$ | 0.0000| $H_0$ rejected |
|                        | $d$       | 0.0000| $H_0$ rejected |
| ARFIMA ([3],d,[3])     | $\phi_3$ | 0.0000| $H_0$ rejected |
|                        | $\theta_1$ | 0.0000| $H_0$ rejected |
|                        | $d$       | 0.0000| $H_0$ rejected |
| ARFIMA ([3],d,[1,3])   | $\phi_3$ | 0.0000| $H_0$ rejected |
|                        | $\theta_1$ | 0.0000| $H_0$ rejected |
|                        | $\theta_3$ | 0.0000| $H_0$ rejected |
|                        | $d$       | 0.1537| $H_0$ accepted |
|                        | $\phi_1$ | 0.0000| $H_0$ rejected |
| ARFIMA ([1,3],d,1)     | $\phi_3$ | 0.0000| $H_0$ rejected |
|                        | $\theta_1$ | 0.0651| $H_0$ accepted |
|                        | $d$       | 0.2445| $H_0$ accepted |
|                        | $\phi_1$ | 0.0000| $H_0$ rejected |
| ARFIMA ([1,3],d,[3])   | $\phi_1$ | 0.0000| $H_0$ rejected |
|                        | $\phi_3$ | 0.0000| $H_0$ rejected |
|                        | $\theta_3$ | 0.0003| $H_0$ rejected |
|                        | $d$       | 0.4932| $H_0$ accepted |
|                        | $\phi_1$ | 0.0000| $H_0$ rejected |
| ARFIMA ([1,3],d,[1,3]) | $\phi_3$ | 0.0000| $H_0$ rejected |
|                        | $\theta_1$ | 0.0647| $H_0$ accepted |
|                        | $\theta_3$ | 0.0015| $H_0$ rejected |

4.7. Diagnostic Test on Residuals of ARFIMA Model

The diagnostic test on the residuals is done by testing the white noise residual assumptions. The test for white noise residual is the Ljung-Box test using Eviews 9 software. The result is below:

Hypothesis:

$H_0$: $\rho_1 = \rho_2 = \ldots = \rho_k = 0$ (residual is white noise)
H1: There is at least one $\rho_k \neq 0$; k = 1, 2, ..., k (residual is not white noise)

Significance Level:
$\alpha = 5\%$

Test Statistics:
$$Q = n(n + 2) \sum_{i=1}^{k} (n - k)^{-1} \rho_k^2$$

Decision:
At a significance level of 5%, $H_0$ is accepted for ARFIMA model (1,d,[3]) because p-value > $\alpha$ for each lag. Meanwhile for ARFIMA (1,d,0), ARFIMA ([3],d,0), ARFIMA (0,d,[3]), ARFIMA ([3],d,1), ARFIMA ([3],d,[3]) and ARFIMA ([3],d,[1,3]) $H_0$ is rejected at a significance level of 5% because p-value < $\alpha$ for each lag.

Conclusion:
The residual of ARFIMA model (1,d,[3]) is white noise. On the other hand, the residuals of ARFIMA (1,d,0), ARFIMA ([3],d,0), ARFIMA (0,d,[3]), ARFIMA ([3],d,1), ARFIMA ([3],d,[3]) and ARFIMA ([3],d,[1,3]) are not white noise.

4.8. Selection of Best Model

The model which has a significant parameter and meets the assumption of white noise is ARFIMA model (1, d, [3]), so that the model is selected as the best model:

$$(1 - \phi_q B)(1 - B)^{k} Z_t = \theta_q (B) a_t$$

$$(1 - \phi_q B)(1 - B)^{k} Z_t^* = (1 - \theta B) a_t$$

$$(1 - \phi_q B)(1 - B)^{k} Z_t^* = a_t - \theta a_{t-1}$$

$$(1 - 0.9999999 B)(1 - B)^{-0.117320} Z_t^* = a_t + 0.086074 a_{t-1}$$

with $Z_t^* = 1/Z_t$, so to return to the initial data $Z_t = 1/Z_t^*$, with $Z_t^*$ is the gold price data after transformation and $Z_t$ is the gold price data.

5. Conclusion

The best model for the gold price data in Indonesia is ARFIMA model (1,d,[3]):

$$(1 - 0.9999999 B)(1 - B)^{-0.117320} Z_t^* = a_t + 0.086074 a_{t-1}$$

This model is a model in time series with long memory because data contains visual effects of long-term memory, the ACF plot drops slowly hyperbolic and the Hurst value is 0.77922

References

[1] Santoso D A 2013 Pemodelan ARIMA untuk Peramalan Harga Emas (Surabaya: STIKOM)
[2] Kusumadewi F 2014 Peramalan Harga Emas Menggunakan Feedforward Neural Network dengan Algoritma Backpropagation (Yogyakarta: FMIPA UNY).
[3] Mahena Y, Rusli M and Winarso E 2015 Kalibiscientia Jurnal Sains dan Teknologi 2 36.
[4] Rukhansah N, Muslim M A and Arifudin R 2016 Komputaki 1 56.
[5] Soejoeti Z 1987 Analisis Runtun Waktu (Jakarta: Karunika).
[6] Wei W W S 2006 Time Series Analysis: Univariate and Multivariate Methods (New York: Pearson).
[7] Darmawan G 2009 Statistika 9.
[8] Idris S, Goejantoro R and Nasution Y N 2014 Jurnal EKSPONENSIAL 5 137.
[9] Hongnogc T 2014 Asian Journal of Finance & Accounting 6 171.
[10] Qian B and Rasheed K 2004 Proceedings of the 2nd IASTED International Conference on Financial Engineering and Applications.
[11] PT. Antam (Persero) Tbk. Harga Emas, Buyback, Silver Hari ini. http://www.pusatdata.kontan.co.id. Accessed on 28 February 2018.