Application of GARCH model in the price inflation of foodstuff in West Java

J Nahar*, E Hertini, and A K Supriatna

Departemen of Mathematics, Universitas Padjadjaran, Jl. Raya Bandung Sumedang km 21 Jatinangor, Sumedang 45363, Indonesia

*julita.nahar@unpad.ac.id

Abstract. The risk and uncertainty faced by consumers and producers due to fluctuations in food prices in recent years have continued to increase. One of them is the fluctuation in the price of foodstuffs which are basic human needs. The process of increasing prices in general and continuously related to the market mechanism which can be caused by various factors is called inflation. Inflation can be measured by calculating the change in the percentage rate of change in a price index. In this study, the index used is the Consumer Price Index (CPI), which is an index that measures the average price of certain goods purchased by consumers. The food price index tends to have high volatility, this can be seen from the sharp increase and decrease in the price of food commodities in Indonesia. Data that has high volatility will be very risky if you use the basic time series method which has stationary data and costly variance, for that we need a method to deal with data that has non-constant and changing variance, or is called heteroscedasticity. Therefore in this study using the GARCH model application to analyze data on foodstuff price inflation in West Java. The results showed that the best GARCH model for foodstuff price inflation data is the ARCH (1) model, which means that the variance of foodstuff price inflation in period t is determined by a constant (1.565) and the remaining square in the previous period with a proportion of 74.41. %.

1. Introduction
The risks and uncertainties faced by both consumers and producers due to fluctuations in food prices in recent years have continued to increase. One of them is the fluctuation in the price of foodstuffs which are basic human needs. The process of increasing prices in general and continuously related to the market mechanism which can be caused by various factors is called inflation. Inflation is a process of increasing prices in general and continuously related to market mechanisms which can be caused by various factors, including increased public consumption, excess liquidity in the market that triggers consumption or even speculation, including the result of the improper distribution of goods [1,2]. Inflation can be measured by calculating the change in the percentage rate of change in a price index. In this study, the index used is the Consumer Price Index (CPI), which is an index that measures the average price of certain goods purchased by consumers. The food price index tends to have high volatility, this can be seen from the sharp increase and decrease in the price of foodstuff commodities in Indonesia. Data that has high volatility will be very risky if you use the basic time series method which has stationary data and costly variance, for that we need a method to deal with data that has a variance that is not constant and changing, or is
called heteroscedasticity. The appropriate econometric model to estimate such behavior is called the Autoregressive Conditional Heteroscedasticity (ARCH) model. This model was first developed by Engle [3]. The ARCH model was then refined by Bollerslev [4] by including not only the error term in the past but also variants of the error term in the past. This Bollerslev model is then called the Generalized Autoregressive Conditional Heteroscedasticity (GARCH) model [5]. Therefore, in this study, the application of the GARCH model is used to analyze the inflation price data for foodstuffs in West Java.

2. Method

2.1 The Box-Jenkins ARIMA Model

ARIMA (Auto Regressive Integrated Moving Average) model is a Box-Jenkins variant of the ARMA model where the model requires the analyzed data to move along a constant (stationary) average [makridakis1998,7]. If the forecasting data is not stationary, then the data stationary process is carried out using the differencing process. Differencing is an operation in which new time series data are built by taking sequential differences from successive prices, \( Z_t - Z_{t-1} \) along a non-stationary time series data pattern to obtain stationary time series data. In general, the ARIMA model \((p, d, q)\) consists of three parts, namely: the differencing process component so that the non-stationary time series becomes stationary after the differencing process; auto regressive AR component \((p)\) and moving average MA \((q)\) component. Gujarati [8] describes the Box-Jenkins methodology into four steps, namely identification, estimation, diagnostic examination, and forecasting. In general, the equations for the ARIMA model \((p, d, q)\) can be seen in the equation below:

\[
\phi_p(B)(1-B)Z_t = \theta_q(B)e_t
\]  

with the Autoregressive (AR) model is \( \phi_p(B)Z_t = e_t \), a function of \( Z_t = \phi_1Z_{t-1} + \phi_2Z_{t-2} + \ldots + \phi_pZ_{t-p} + e_t \) and the Moving Average (MA) model is \( Z_t = \theta_q(B)e_t \), a function of \( Z_t = e_t - \theta_{1}e_{t-1} - \theta_{2}e_{t-2} - \ldots - \theta_{q}e_{t-q} \), where \( B \) is backshift operator or operator delay, \( p \) is order AR, \( q \) is order from MA, and \( d \) is order differencing. In some economic and financial data, it generally has residual variances that are not constant (heteroscedasticity) [9], such as inflation which has high fluctuations which causes the residual variance to change over time. Therefore, the model that can be used in overcoming the residual variance condition that always changes every time (heteroscedasticity) is Autoregressive Conditional Heteroscedasticity (ARCH) and some of the development of this model in overcoming some of its weaknesses.

2.2 ARCH-GARCH Model

The Autoregressive Conditional Heteroscedasticity (ARCH) model was first developed by Engle [3]. In this model, it is assumed that the variance of the current period data is influenced by the residual data of the previous period. In general, the ARCH \((p)\) model can be expressed in the equation:

\[
\sigma_i^2 = \alpha_0 + \alpha_1e_{i-1}^2 + \alpha_2e_{i-2}^2 + \ldots + \alpha_pe_{i-p}^2
\]  

where:

- \( \sigma_i^2 \) : the variance of the t-period residuals
- \( p \) : order ARCH
- \( \alpha_0 \) : constant components
\( \alpha_i \) : ARCH parameter

\( \varepsilon^2_{t-p} \) : the quadrature of the period t-p residual

To calculate the volatility or risk of increasing or decreasing inflation in the ARCH \((p)\) model, it can be calculated from the root of Equation (2). This model has difficulty estimating parameters in high order so that the ARCH model is then refined by Bollerslev [4] which is then called the model Generalized Autoregressive Conditional Heteroscedasticity (GARCH). The GARCH model is used to handle the large number of parameters in the ARCH model required for modeling the volatility process. In the GARCH model, it states that the residual variant depends not only on the residual of the previous period but also the variance of the residual of the previous period. In general, the GARCH model \((p,q)\) can be expressed in the equation:

\[
\sigma^2_t = \alpha_0 + \alpha_1 \varepsilon^2_{t-1} + \ldots + \alpha_p \varepsilon^2_{t-p} + \beta_1 \sigma^2_{t-1} + \ldots + \beta_q \sigma^2_{t-q}
\]  

(3)

where:

\( p \) : orde ARCH

\( q \) : orde GARCH

\( \sigma^2_t \) : the variance of the t-period residuals

\( \alpha_0 \) : constant components

\( \alpha_i \) : parameters from ARCH

\( \beta_j \) : parameters from GARCH

\( \varepsilon^2_{t-p} \) : the quadrature of the period t-p residual

\( \sigma^2_{t-q} \) : variance of the period t-q residuals

To calculate the volatility or risk of increasing or decreasing inflation in the GARCH model \((p,q)\), it can be calculated from the root of Equation (3).

2.3 ARCH-Effect Test

Engle developed a test to find out the heteroscedasticity problem in time series data using the Langrange Multiplier test as follows.

\[
LM = N \times R^2
\]  

(4)

where:

\( N \) : the number of observations

\( R^2 \) : the coefficient of determination

The Langrange Multiplier test follows the distribution of \( \chi^2 \) with degrees of freedom of \((a, p+q)\). If the probability value is less than the degree of confidence, there is an ARCH effect in the model. If there is an ARCH effect in the model, the estimation can be done using ARCH-GARCH.

2.4 Best Model Selection

The selection of the feasibility of an ARCH-GARCH model was carried out by first doing a further diagnosis of the residuals using the ARCH-LM (to make sure that no ARCH effect remained), examining the correlogram Q-Statistic (CQS), and correlogram squared residual (CSR). The model is said to be feasible if the residuals do not have autocorrelation between the residues for all lag \(k\). Also, this test is conducted to see whether the ARCH effect is still present in the model. The selection of the best model is done if there is more than one time series model that is suitable for use, namely by
calculating the value of *Akaike Information Criteria* (AIC) or *Schwartz Information Criteria* (SIC). The smaller the value obtained, the better the model is used in forecasting. Several types of forecasting might be used. In this study, researchers used the ARCH-GARCH model as the analysis used because this model considers the existence of a period where there is very high volatility and there are other periods where the volatility is very low, such a pattern is called the heteroscedasticity problem because there are variants of the error whose magnitude depends on it. on the volatility of errors in the past so that the ARCH-GARCH model is considered relevant to the research conducted compared to the classical forecasting model which still assumes that time series data does not have heteroscedasticity problems.

3. Result and Discussion

3.1 Research Data

The data source was taken from Badan Pusat Statistik (BPS) West Java province. Inflation rate of foodstuff prices from 2016-2019.

3.2 Research Stages

The steps taken in analyzing the data in this study are as follows:

1. Stationarity test
   The stationarity test using the Augmented Dickey Fuller (ADF) test statistic is as follows.
   **Hypothesis Test:**
   - $H_0: \delta = 0$ (Data is not stationary)
   - $H_1: \delta < 0$ (Data is stationary).
   **Test Statistics:**
   \[
   t = \frac{\delta}{se(\delta)} \tag{6}
   \]
   reject $H_0$ if $|t| > t_{(a,n)}$.

2. Finding ACF and PACF Plots
   Example $z_t$ represents the data for the t-specific period, $z_{t-1}$ represents the data for the previous period, and $z_{t-2}$ represents the two data for the previous period. Then autocorrelation can be formulated as follows:
   \[
   \hat{\rho}_k = \frac{\sum_{t=1}^{T-k}(z_t-\bar{z})(z_{t+k}-\bar{z})}{\sum_{t=1}^{T}(z_t-\bar{z})^2} \tag{7}
   \]
   $\bar{\bar{z}} = \frac{\sum_{t=1}^{T}z_t}{T}, \hat{\rho}_k$: autocorrelation for time-lag 1,2,3,..., $k$, $T$: number of observation times.
   Partial autocorrelation is used in time series analysis to measure the degree of closeness between $z_t$ and $z_{t+k}$, when the effects of lag 1,2,3,... are considered separate and can also help in determining the appropriate forecasting model. Partial autocorrelation or $\varphi_{kk}$ is stated by Wei [10]
   \[
   \varphi_{kk} = \text{Corr} (z_t, z_{t+k} | z_{t+1}, z_{t+2}, ..., z_{t+k-1}) \tag{8}
   \]

3. Identify and estimate the ARIMA model
   After the data obtained becomes stationary, it is followed by identification of the AR and MA orders according to the correlogram obtained. Table 1 can be used to identify the $p$ and $q$ levels of a time series data based on the shape of the ACF and PACF patterns.
After the model is identified, the model parameters are estimated using the maximum likelihood. From several possible models, the best model is selected. The selection of the best model is done by looking at the smallest AIC and SIC values.

**Table 1.** The ACF and PACF chart patterns of the process are stationary

| Process   | ACF                                      | PACF                                      |
|-----------|------------------------------------------|------------------------------------------|
| AR(p)     | Dies down (descending rapidly exponentially / sinusoidally) | Cuts off after lag p                      |
| MA(q)     | Cuts off after lag q                      | Dies down (descending rapidly exponentially / sinusoidally) |
| ARMA(p,q) | Dies down after lag (q-p) (get off fast after lag (q-p)) | Dies down after lag (p-q) (get off fast after lag (p-q)) |

4. Diagnostic Checking
The model is said to be good if the residual sequence \( \{\varepsilon_t\} \) is white noise. A sequence is said to be white noise if \( \varepsilon_t \) is not autocorrelated and normally distributed. To test the autocorrelation of the residual \( \{\varepsilon_t\} \) series, the Q Ljung-Box test was used with the following hypothesis:

\[
H_0: \rho_1 = \rho_2 = \ldots = \rho_k = 0 \quad \text{(Residue data white noise)}
\]

\[
H_1: \text{there is at least one } \rho_k \neq 0 \quad \text{(Residue data is not white noise)}
\]

Test statistics:

\[
Q = n(n + 2) \sum_{k=1}^{K} \frac{\hat{\rho}_k^2}{n-k}
\]

\( K \): maximum lag performed
\( n \): the number of observations
\( \hat{\rho}_k \): residual ACF samples at the k-lag

reject \( H_0 \) if \( Q > \chi^2_{(\alpha;k-p-q)} \), accept in other cases.

5. Determine the ARIMA model through the AIC and SIC values

\[
AIC = n \ln(\hat{\sigma}^2_a) + 2m
\]

\[
SIC = n \ln(\hat{\sigma}^2_a) + m \ln(n)
\]

where:

\( m \): the number of parameters in the model
\( n \): the number of observations
\( \hat{\sigma}^2_a \): the maximum likelihood estimate of \( \sigma^2_a \)

The smaller the AIC and SIC values, the better the model is obtained.

6. Heteroscedasticity Effect Test
Testing the heteroscedasticity effect was carried out using the Lagrange Multiplier (LM) test statistic. The test is performed by regressing the residual value squared with a constant up to the m-lag, thus forming a regression equation \( a_t^2 = a_0 + a_1 a_t^2 + \ldots + a_m a_t^2 \).

The value of the LM test statistic is obtained by multiplying the number of observations \( (N) \) by the coefficient of determination \( (R^2) \).

Where the test hypothesis is:
H0: α1 = ⋯ = αₘ = 0 (There is no heteroscedasticity effect)
H1: there is at least one αₘ ≠ 0 (There is a heteroscedasticity effect)
The test statistic used uses Equation (2.8) where H₀ is rejected if \( LM > \chi^2_{(α,k−p−q)} \)

7. Determine the ARCH-GARCH model estimate

From the results of the heteroscedasticity effect test, the ARCH-GARCH model was formed by looking at the ACF and PACF plots of the squared residuals. From several possible models, parameter estimation and selection of the best model will be carried out using AIC and SIC.

The best ARCH-GARCH model as formulated by Equation (1) and Equation (2) is carried out by Diagnostic Checking and testing whether there is a heteroscedasticity effect on the residual of the model. If the Diagnostic Checking has been fulfilled, then the model will be used to calculate the inflation value in period t.

3.3 Data Identification

The stationarity test in the average data was carried out by using the Augmented Dickey Fuller (ADF) test. ADF value for inflation data can be seen in Table 2. Tests are carried out using a significance level α=5%.

| Inflation       | ADF       | Critical Value (5%) |
|-----------------|-----------|---------------------|
| Food material   | -6.1742   | -2.89               |

In Table 2, you can get the value of | ADF | on foods material inflation> | Critical Value (5%) |, so H₀ is rejected. This means that the foodstuff inflation data is stationary in the average, so that the foodstuff inflation data can be used in time series analysis.

3.4 ARIMA Model Formation

The first step in building the model is identifying or estimating the model. Identification of the ARIMA model can be done by looking at the significant ACF and PACF plots.

Figure 1. Plot ACF (1a) and PACF (1b)

In the picture above, the ACF and PACF plots are significant at the 1st lag, so that the model estimates are ARIMA (3,0,0), ARIMA (0,0,1), and ARIMA (3,0,1). From several possible models, the
best model can be determined by looking at the smallest AIC value of each model formed. The smallest AIC value can be seen in Table 3.

**Table 3. AIC value in each model**

| Model           | AIC  |
|-----------------|------|
| ARIMA (3,0,0)   | 142.62 |
| ARIMA (0,0,1)   | 145.22 |
| ARIMA (3,0,1)   | 134.29 |

From Table 3, it is obtained that the best ARIMA model with the smallest AIC value is owned by the ARIMA model (3,0,1), with the equation model as follows:

\[ Z_t = 0.426 + 0.075\epsilon_{t-1} + \epsilon_t - 0.2391Z_{t-3}. \]  
(12)

To see the suitability of the model that has been obtained, it can be done by testing the residual of the model whether it is white noise / random or not. The calculation results can be seen in table 4 below.

**Table 4. Diagnostic checking**

| Model           | \( \lambda^2 \) | \( p \)-value |
|-----------------|-----------------|---------------|
| ARIMA (3,0,1)   | 0.0046          | 0.945         |

Table 4 shows that the model has a \( P\)-value \( > \alpha \) which indicates that \( H_0 \) is accepted, meaning that with a significant level of 5%, it can be concluded that the model has met sufficient requirements or the residual has met the requirements for white noise. The best ARIMA model is then performed a heteroscedasticity test on the residual squared. If there is heteroscedasticity in the quadratic residual data, it will be modeled with the ARCH-GARCH model.

### 3.5 ARCH-GARCH Model Building

Before modeling with ARCH-GARCH is done, first test the effect of heteroscedasticity on the squared residual data obtained from the best ARIMA model. ARCH Effect testing on inflation data can be seen in table 5.

**Table 5. ARCH-Effect testing on data**

| Lag to- | \( \lambda^2 \) | \( p \)-value |
|---------|-----------------|---------------|
| to-6    | 14,477          | 0.025         |
| to-7    | 14,54           | 0.043         |

From table 5 above, it is found that the inflation of foodstuff price data has a smaller \( p\)-value against \( \alpha \) by 5%, namely at the lag-6 and lag-7, so that \( H_0 \) is rejected, meaning that it indicates that there is a heteroscedasticity effect on the quadratic residual data. So, the next step is to build the ARCH-GARCH model. To see the possible models formed, it can be seen from the ACF and PACF plots that will be used to estimate the ARCH-GARCH model. From the calculation results, it is found that the best ARCH-GARCH model for foodstuff price inflation data is the ARCH model (1). The ARCH (1) model can be written in the form of a mathematical equation as follows:
This means that the variance of foodstuff data inflation in the \( t \)-period is determined by a constant \((1.565)\) and the remaining squared in the previous period with a proportion of 74.41\%. To see whether the model is appropriate for use, a diagnostic test will be carried out, namely re-testing whether there is still a heteroscedasticity effect in the ARCH model (1). From the ARCH-Effect test using the Ljung-Box test on the squared residual is not significant to \( \alpha \) by 5\% so that \( H_0 \) is accepted, meaning that food inflation data has no heteroscedasticity effect on the ARCH model (1), so the ARCH (1) model is formed has met the assumptions and can be used. After the model is formed, it can be used to forecast food material inflation.

4. Conclusion

From the results of the analysis carried out, the ARIMA model that is formed is the ARIMA model (3,0,1), with the equation model as follows:

\[
\sigma^2_{t(1)} = 1.565 + 0.7441\varepsilon^2_{t-1}.
\]  

This means that the variance of foodstuff data inflation in the \( t \)-period is determined by a constant \((1.565)\) and the remaining squared in the previous period with a proportion of 74.41\%.

Acknowledgments

Authors would like to thanks to Rector of Universitas Padjadjaran for full support to the research conducted through Academics Leadership Grant 2020 with Contract no. 147/UN6.3.1/LT/2020

References

[1] Mishkin S 2004 The economics of money, banking, and financial markets 7th ed (Columbia: The Addyson- Wesley)
[2] Diha D P 2011 Metode Persamaan Simultan Dalam Memodelkan Tingkat Inflasi di Jawa Barat (Bandung: Universitas Padjadjaran)
[3] Engle R 2001 GARCH 101: The Use of ARCH/GARCH Models in Applied Econometric Journal of Econometric Prespective 15 157-168
[4] Bollerslev T 1986 Generalized Autoregressive Conditional Heteroskedasticity Journal of Econometrics 31 307-27
[5] Widarjono A 2002 Aplikasi Model ARCH Kasus Inflasi di Indonesia Jurnal Ekonomi Pembangunan 7 71-82.
[6] Makridakis W S 1998 Forecasting: Method and Applications (New York:Wiley)
[7] Tsay R S 2005 Analysis of Financial Time Series (New York: John Willey and Sons, Inc)
[8] Gujarati D N 2003 Basic Econometrics Fourth Edition (USA: McGraw-Hill)
[9] Lo M S 2003 Generalized Autoregressive Conditional Heteroscedastic Time Series Models (Doctoral dissertation, Theses, Department of Statistics and Actuarial Science, Simon Fraser University)
[10] Wei W W 2006 Time series Analysis, Univariate and Multivariate Methods (Canada: Addison Wesley Publishing Company)