Journal of Engineering and Applied Sciences 13 (10): 3418-3422, 2018
ISSN: 1816-949X

Application of Generalized Autoregressive Conditional Heteroscedasticity (GARCH) Model for Forecasting

[1]M. Yusuf S. Barusman, [2]Mustofa Usman, [2]Riyama Ambarwati and [3]Erica Virginia
[1]Faculty of Economics, Universitas Bandar Lampung (UBL), Bandar Lampung, Indonesia
[2]Department of Mathematics, Universitas Lampung, Bandar Lampung, Indonesia
[3]Department of Accounting, President University, Bakasi, Indonesia

Abstract: Financial data sometimes have not only high volatility but also heterogeneous variances. The Box Jenkins method cannot be used to overcome a model which has an effect of heteroscedasticity. One of the models can be used to overcome the effect of heteroscedasticity is GARCH Model. The aims of this study are to find the best model, to estimate the parameters of the best model and to predict the share price data of JAPFA Comfeed Indonesia over the period of June 2015 to October 2016. The best model which fits to the data is ARIMA (0, 1, 2) and GARCH (1, 1). The application of the two models for forecasting the share price data of JAPFA Comfeed Indonesia for the next 5 weeks period is very sound and all the forecast values are within 95% confidence interval.

Key words: Volatility, heteroscedasticity, ARIMA, GARCH Model, forecasting, Comfeed Indonesia

INTRODUCTION

Time series is an ordered sequence of observation. It is usually through time, particularly in terms of some equally spaced time interval (Wei, 2006). Modeling for time series data commonly uses Autoregressive (AR), Moving Average (MA) or the combination both known as Autoregressive and Moving Average (ARMA) with the assumption that the data have a constant variance or homoscedasticity. However in the case of financial data, like share prices generally tend to fluctuate rapidly from time to time, so that the variance of its error will always change from time to time (heterogeneous). One of the methods can be used to overcome the problem of heterogeneity of variance is a method of Autoregressive Conditional Heteroscedasticity (ARCH) introduced by Engle (1982). The model was generalized by Bollerslev (1986) and called Generalized Autoregressive Conditional Heteroscedasticity (GARCH) Model. This model is used to analyze share price data of JAPFA Comfeed Indonesia over the period of June 2015 to October 2016. In this study, the application of the GARCH model for forecasting data of the next periods will also be presented.

MATERIALS AND METHODS

Time series modeling: ARIMA Model is one of the models that can be used to analyze nonstationary time series data. ARIMA Model (Tsay, 2005; Brockwell and Davis, 1991, 2002; Pankratz, 1991) is presented as follows:

$$\phi_p(B)(1-B)^d x_t = \theta_0 + \theta_q(B)\varepsilon_t, \quad \{\varepsilon_t\} \sim WN(0,\sigma^2) \tag{1}$$

where, $\phi_p(B) = (1-\phi_1 B, ..., \phi_1 B^p)$, $\theta_q(B) = (1-\theta_1 B, ..., \theta_1 B^q)$ and $(1-B)^d$ is differencing non seasonal of order d. The equation can be written as ARIMA (p, d, q).

In the estimation of time series model, the procedure of the ARIMA Box Jenkins Model is used. The data have to be stationary in mean and variance. To check the stationary data, both time series plot and Augmented Dickey Fuller (ADF) test can be applied:

- H_0: $\phi = 0$ there is a unit root or the data are nonstationary)
- H_0: $\phi<0$ (there is no a unit root or the data are stationary)

$$\Delta y_t = \phi y_{t-1} + \sum_{j=1}^{p-1} \alpha_j \Delta y_{t-j} + u_t \tag{2}$$

where, $\phi = -\alpha(1)$ $\alpha_j^* = -(\alpha_{j+1}+, ..., +\alpha_p)$ with the degrees of freedom n and the level of significance α. The ADF t-statistic is:

$$ADF_t = t = \frac{\hat{\varphi}-1}{Se(\varphi)} \tag{3}$$

Corresponding Author: M. Yusuf S. Barusman, Faculty of Economics, Universitas Bandar Lampung (UBL), Bandar Lampung, Indonesia

If H_0 is rejected, the data are stationary. In time series data analysis, the main tool to identify a model of the data to be forecast is by using Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) (Montgomery *et al*., 2008). The parameters of the found model are then estimated and tested. The residuals are also tested to see whether or not they have the properties of white noise. To test the white noise, Ljung box test (Wei, 2006) is used with the following hypotheses:

- H_0: $\rho_1 = \rho_2 = \rho_3 = , ..., = 0$ (the residuals are not autocorrelated or white noise)
- H_1: $\exists\rho_k \neq 0$, $k = 1, 2, \ldots, K$ (the residuals are autocorrelated or not white noise)
- The level of significance $\alpha = 5\%$

The statistic test is:

$$Q_k = T(T+2)\sum_{k=1}^{K}\frac{\hat{\rho}k^2}{T\text{-}k} \tag{4}$$

To reject or not to reject H_0, the value of Q and that of $X^2_{a,\ df}$ are compared. To select the best model, the criteria of Akaike Info Criterion (AIC) is used:

$$AIC = -2\left(\frac{1}{T}\right)+2\left(\frac{k}{T}\right) \tag{5}$$

The ARIMA Model to be selected is based on the smallest values of AIC or SC.

ARCH and GARCH Models: The Autoregressive Conditional Heteroscedasticity (ARCH) is a function of autoregression with an assumption that the variance is changed over time and the value of the variance is affected by some previous data. The ARCH Model is used for modeling the volatility. The ARCH model with order q or ARCH(q) is defined as follows:

$$\sigma_t^2 = \lambda_0+\lambda_1\varepsilon_{t-1}^2+\lambda_2\varepsilon_{t-2}^2+, ...,+\lambda_q\varepsilon_{t-q}^2 \tag{6}$$

Sometimes in ARCH model, so many parameters are involved and it makes the model to become complex. To overcome the complexity, Bollerslev (1986) introduced GARCH Model. In GARCH Model, the change of the variance is affected by previous data and previous variance. The GARCH Model is defined as follows:

$$x_t = \delta+\sum_{i=1}^{p}\phi_i x_{t-i} + \varepsilon_t - \sum_{i=1}^{q}\theta_i\varepsilon_{t-i}$$
$$\varepsilon_t \sim N(0, \sigma^2) \tag{7}$$
$$\sigma_t^2 = \lambda_0 + \sum_{i=1}^{q}\lambda_i\varepsilon_{t-i}^2 + \sum_{j=1}^{q}\beta_j\sigma_{t-j}^2$$

where, x_t is the equation of conditional mean.

Research method: The data used in this study is the weekly closing share price data of JAPFA Comfeed Indonesia over the period of June 2015 to October 2016. The data were taken from www.finance.yahoo.com. Some steps are carried out in the process of the data analysis. The first step is to check the stationary data. The stationary in mean is checked through the plot of the data and statistical test using Augmented Dickey Fuller (ADF) test while the stationary in variance is through the plot of the data. If the data are nonstationary, differencing and transformation of the data are used. When the data have been stationary, Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) are applied to estimate the order of ARIMA. The second step is to estimate and test the parameters, to diagnose and test the residuals and to select the best model based on the criteria of the smallest values of AIC or SC. The residuals obtained from the best ARIMA model are checked by using Lagrange Multiplier (LM) test to know whether or not they have ARCH effect. If there is ARCH effect, the data are modeled by using ARCH or GARCH Model. The order of ARCH or GARCH Model is found through the plot of the squared residuals of PACF. The third step is to estimate and to test the parameters of the model and to forecast the weekly share price data.

RESULTS AND DISCUSSION

Identification: The first step of the data analysis is to check the assumption of stationarity. To check the stationarity, two approaches: the plot of the data and Augmented Dickey Fuller (ADF) test are used. Figure 1 shows the movement of the data shows that the data are nonstationary. This results from the pattern of the share prices that tends to rise and decline very sharply and the result of the mean of the data at each lag which is not constant.

Since, $p = 0.4802>0.05$, there is no enough evidence to reject H_0. Thus, the data is nonstationary. This means that the share price data of JAPFA Comfeed Indonesia has unit root implying that the data is

Fig. 1: Plot of share price data of JAPFA Comfeed Indonesia over the period of June 2015 to October 2016

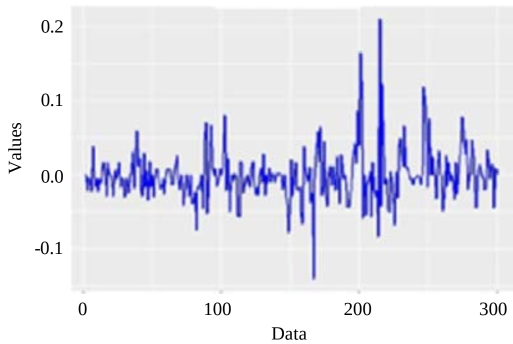

Fig. 2: Graph of share price data of JAPFA Comfeed Indonesia over the period of June 2015 to October 2016 after differencing

Table 1: Result of ADF test of share price data of JAPFA Comfeed Indonesia

Variables	Values
Augmented Dickey-Fuller test	
Parameter	
Lag order	0
Statistic	
Dickey-Fuller	-0.4271
P-value	0.4802

Table 2: Results of ADF test of share price data of JAPFA Comfeed Indonesia over the period of June 2015 to October 2016 after differencing

Variables	Values
Augmented Dickey-Fuller test	
Test results	
Parameter	
Lag order	0
Statistic	
Dickey-Fuller	-14.5304
P-value	0.0100

nonstationary. Therefore, differencing or transformation of the data to make it stationary is needed. For this, the method to be used is Difference Stationary Processes (DSP). By DSP, the transformation of the data is conducted by differencing k-th lag with the -(k-1) lag of the data. The graph after differencing is given as in Fig. 2 and Table 1.

Table 2 shows ADF test Output of share price data of JAPFA Comfeed Indonesia after differencing. Table 2 shows that, $p = 0.01 < 0.05$, so H_0 is rejected. This means that the share price data of JAPFA Comfeed Indonesia does have not unit root implying that the data is stationary.

Estimation of ARIMA Model: The data have been stationary, the next step is to identify the order and estimate the ARIMA Model. Table 3 shows the results of model estimation along with the best ARIMA model for the data by using Software R 3.2.3.

Table 3 shows that, the best model based on the smallest value of AIC is ARIMA (0, 1, 2) Model

Table 3: Results of model estimation and best model for share price data of JAPFA Comfeed Indonesia based on the smallest value of AIC

Variables	Values
ARIMA (2,1,2) with drift	-1158.878
ARIMA (0,1,0) with drift	-1151.728
ARIMA (1,1,0) with drift	-1157.552
ARIMA (0,1,1) with drift	-1156.376
ARIMA (0,1,0)	-1153.599
ARIMA (1,1,2) with drift	-1161.500
ARIMA (1,1,1) with drift	-1159.609
ARIMA (1,1,3) with drift	-1159.932
ARIMA (2,1,3) with drift	-1157.121
ARIMA (1,1,2)	-1163.429
ARIMA (0,1,2)	-1165.851
ARIMA (0,1,1)	-1158.273
ARIMA (0,1,3)	-1164.588
ARIMA (1,1,3)	-1161.845
ARIMA (0,1,2) with drift	-1163.931

Best model: ARIMA (0, 1, 2)

Table 4: Results of parameter estimation for ARIMA (0,1,2) model

Coefficient	Estimation	t-values	p-values
MA (1)	0.1349	2.3961	0.0172
MA (2)	0.1838	3.1527	0.0018

with the value of AIC = -1165.851. The estimation for the mean model of ARIMA (0, 1, 2) is given in Table 4.

The estimation of parameters shows that the coefficient for MA (1) is 0.1349 with the standard error of 0.0172 and for MA (2) is 0.1838 with the standard error of 0.0018. Thus, the mean model for ARIMA can be written as follows:

$$x_t = 0.1349\varepsilon_{t-1} + 0.1838\varepsilon_{t-2} + \varepsilon_t$$

Evaluation of ARIMA Model: At the stage of model evaluation, whether the residuals of ARIMA (0, 1, 2) have the properties of white noise and normal distribution have to be checked. Test for white noise. To view whether the residuals are white noise, they may be tested by using the Ljung-box test. The results of Ljung-box are presented.

Figure 3 shows that p-value for Ljung-box test is above 0.05 which means that the residuals are not correlated and the residuals of ARIMA (0, 1, 2) Model have the properties of white noise.

Normality test: Figure 4 shows that some residuals are not in the area of 95% confidence interval and the distribution of data does not form a straight line. Thus, the residuals for ARIMA (0, 1, 2) Model are not normal. In addition to using normal QQ plot, the normality test can also be conducted by using Jarque-Berra test as follows:

$$JB = \left[\left(\frac{T}{6}\right)S^2 + \left(\frac{T}{24}\right)(K+3)^2\right]$$

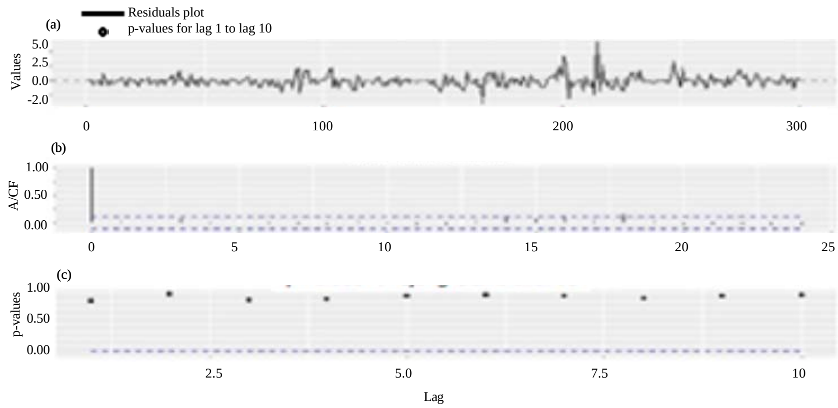

Fig. 3: Results of Ljung-Box test for the residuals of ARIMA (0, 1, 2) model: a) Standardized residuals; b) ACF of residuals and c) p-values for Ljung-box statistics

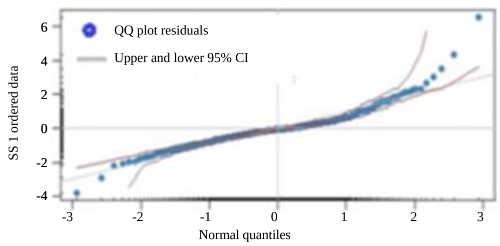

Fig. 4: Normal QQ plot with Confidence Interval (CI) of 95% for ARIMA (0, 1, 2)

Table 5: Results of Jarque-Berra test for ARIMA (0,1,2)

Variables	Values
Jarque-bera normalality test	
Test results	
Statistic χ^2	
p-value	830.0186
Asymptotic p-value	< 2.2e-16

Where:
T = Number of observations
S = Skewness
K = Kurtosis

Since, p-value is ($2.2e^{-16}$) <0.05, H_0 is rejected. Thus, the residuals are not normally distributed (Table 5).

Lagrange Multiplier (LM) test: Time series data have both a problem of autocorrelation and heteroscedasticity. The Lagrange Multiplier (LM) test can be used to detect the presence or the existence of ARCH effect or heteroscedasticity. The null and alternative hypotheses are as follows:

- $H_0 = \lambda_0 = \lambda_1 = , ..., = \lambda_p = 0$ (There is no ARCH effect)
- $H_1 = \lambda_0 \neq 0$ or $\lambda_1 \neq 0$ or, ..., or $\lambda_p \neq 0$ (There is an ARCH effect)
- Level of significance $\alpha = 0.05$
- Criteria: H_0 is rejected if $p < 0.05$

Table 6: ARCH-Lagrange Multiplier test for ARIMA (0, 1, 2)

Variables	Values
Chi-squared	21.80500
df	12.00000
p-value	0.03976

ARCH LM-test; null hypothesis: no ARCH effects; Data: diff_t

Table 7: Results of parameter estimation of GARCH (1,1) Model

Parameters	Coefficient	SE	p-values
δ	-0.004765	0.001751	0.00649
θ_1	0.057256	0.077790	0.46171
θ_2	0.075336	0.073068	0.30252
a_0	0.000297	0.000078	0.00014
λ_1	0.605359	0.176351	0.00059
β_1	0.280349	0.101281	0.00564

Table 6 shows that p-value = 0.03976<0.05, so, the test is significant. Thus, it can be concluded that the ARIMA (0, 1, 2) Model has ARCH effects. Therefore, the modeling by using ARCH/GARCH is very recommended.

GARCH Model: The results of parameters estimation can be seen in the following Table 2. From the results given in Table 7, GARCH (1, 1) Model can be written in the following equation:

$$x_t = -0.004765+0.057256\varepsilon_{t-1}+0.075336\varepsilon_{t-2}$$

And:

$$\sigma_t^2 = 0.000297+0.605359\varepsilon_{t-1}^2+0.280349\sigma_{t-1}^2$$

where, x_t is the equation of conditional mean.

Forecasting: Table 8 shows the result of forecasting share price data of JAPFA Comfeed Indonesia for the next 5 weeks period.

Table 8 show the forecast values of the share price data of JAPFA Comfeed Indonesia for the next 5 weeks period. Table 8 shows the forecast values are close to the real data. It is shown by all values of the forecast ranging

Table 8: Forecasting of share price data of JAPFA Comfeed Indonesia

				Confidence interval forecasting 95%	
Periods	Date	Real data	Forecasting	Lower bound	Upper bound
359	27-10-2016	1800	1749.99645	1670.204022	1820.63581
360	28-10-2016	1950	1749.99191	1652.475323	1836.53227
361	31-10-2016	1885	1749.98715	1640.092376	1849.54519
362	01-11-2016	1960	1749.98238	1630.632284	1860.26517
363	02-11-2016	1915	1749.97762	1622.943573	1869.06802

from the upper to lower bound of the 95% confidence interval. Thus, it can be concluded that GARCH (1, 1) Model can be used to forecast the share price data of JAPFA Comfeed Indonesia in the next week period.

CONCLUSION

Based on the results and the detail of data analysis using GARCH Model for the share price data of JAPFA Comfeed Indonesia over the period of June 2015 to October 2016, it can be concluded as follows: In the application of GARCH Model for forecasting share price data of JAPFA Comfeed Indonesia over the period of June 2015 to October 2016 with the total number of observation 358, it is found that the best model is GARCH (1, 1) Model. The conditional mean model is ARIMA (0, 1, 2) and the conditional variance model is GARCH (1, 1) as follows:

Conditional mean model; ARIMA (0, 1, 2):

$$x_t = 0.1349\varepsilon_{t-1}+0.1838\varepsilon_{t-2}+\varepsilon_t$$

Conditional variance model; GARCH (1, 1):

$$x_t = -0.004765+0.057256\varepsilon_{t-1}+0.075336\varepsilon_{t-2}$$

And:

$$\sigma_t^2 = 0.000297+0.605359\varepsilon_{t-1}^2+0.280349\varepsilon^2_{t-1}$$

The forecast values of share price data of JAPFA Comfeed Indonesia for the next 5 week period are very close to the real values: $X^*_{359} = 1749.996$, $X^*_{360} = 1749.991$, $X^*_{361} = 1749.987$, $X^*_{362} = 1749.982$ and $X^*_{363} = 1749.977$. This shows all the forecast values are within 95% confidence interval. Thus, GARCH (1 ,1) Model is more acceptably used to predict the share price data of JAPFA Comfeed Indonesia for the next periods.

ACKNOWLEDGEMENTS

The researchers would like to thank Yahoo Finance and JAPFA Comfeed Indonesia for providing the weekly closing share price data used in this study. The authors would also like to thank unanimous reviewers for their suggestions and improvement to make this study better.

REFERENCES

Bollerslev, T., 1986. Generalized autoregressive conditional heteroscedasticity. J. Econometrics, 31: 307-327.

Brockwell, P.J. and R.A. Davis, 1991. Time Series: Theory and Methods. 2nd Edn., Springer-Verlag, New York, ISBN: 0387974296.

Brockwell, P.J. and R.A. Davis, 2002. Introduction to Time Series and Forecasting. 8th Edn., Springer, Berlin, ISBN: 978-0-387-95351-9, Pages: 469.

Engle, R.F., 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. Econometrica, 50: 987-1007.

Montgomery, D.C., C.L. Jennings and M. Kulachi, 2008. Introduction Time Series Analysis and Forecasting. John Wiley & Sons, Hoboken, New Jersey, ISBN:9780471653974, Pages: 472.

Pankratz, A., 1991. Forecasting with Dynamic Regression Models. John Wiley & Sons, Hoboken, New Jersey, ISBN:9780471615286, Pages: 400.

Tsay, R.S., 2005. Analysis of Financial Time Series. 2nd Edn., Wiley-Interscience, New York.

Wei, W.W.S., 2006. Time Series Analysis: Univariate and Multivariate Methods. 2nd Edn., Pearson Addison Wesley, USA., ISBN-13: 9780321322166, Pages: 614.