



Application of ARIMA-GARCH Model to Estimating Expected Shortfall in BMRI Stocks

Alberto Simanjuntak^{1*}

¹*Bachelor Program in Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Padjadjaran, Indonesia*

**Corresponding author email: albertosimanjuntak97@gmail.com*

Abstract

Stocks are one of the best-known forms of investment and are still used today. In stock investment, it is necessary to know the movement and risk of loss that may be obtained from the stock investment so that investors can consider the possibility of profit. One way of calculating risk is to use the Expected Shortfall (ES). Because the stock movement is in the form of a time series, a model can be formed to predict the movement of the stock which can then be used for ES calculations using time series analysis. The purpose of the study was to determine the expected shortfall value of BMRI shares using time series analysis. The data used for this research is the daily closing price of shares for three years. In the time series analysis stage, the models used in predicting stock movements are Autoregressive Integrated Moving Average (ARIMA) for the mean model and Generalized Autoregressive Conditional Heteroscedasticity (GARCH) for the volatility model. The average value and variance obtained from the model are then used in calculating the ES on BMRI stock. Based on the results of the study, it was obtained that BMRI's stock had an ES of 0.008116. This means if an investment is made for BMRI shares of IDR 1,000,000.00 for 37 days (5% of 751 days) for an investment period with a 95% confidence level, the expected loss to be borne by the investor is IDR 8,116.00.

Keywords: time series analysis, ARIMA, GARCH, Expected Shortfall

1. Introduction

Investment, in general, is the investment of assets or funds by a party for a certain period to gain profits or benefits in the future. One of the most well-known forms of investment and is still being carried out today is stock investment, or rather share ownership rights.

Shares are proof of ownership of the share capital of a limited liability company that entitles them to dividends and others according to the size of the paid-up capital; or rights that people (shareholders) have over the company due to the surrender of a share of capital so that it is considered a share in ownership and control (Regionalitas, 2016).

Every investment has its advantages and risks, stocks are no exception. To make a profit, it is necessary to consider the benefits obtained from the return price of the stock with the existing risks. Stocks themselves have price movements that tend to be difficult to ascertain, making stocks have their risks compared to other forms of stock.

The risk of stock investment makes investors need to know the seasonal variance and price movements. There are several ways to estimate risk in investment, such as volatility and Expected Shortfall. This study will use the Expected Shortfall method in analyzing stock risk with a time series model.

These models are widely used for time series analysis of various types of data. Following are some previous studies regarding the use of the ARIMA-GARCH model. The first study was entitled "Forecasting Inflation in Kenya Using ARIMA-GARCH Models" by Uwilingiyimana et al. (2015). This research was motivated by the inflation rate in Kenya which at that time became out of control. This study aims to develop a model that can explain Kenya's inflation rate from 2000 to 2014 using time series analysis. Data analysis used the least-squares method and Autoregressive Conditional Heteroscedastic (ARCH). After being analyzed separately through the ARIMA model and the GARCH model, it was found that the ARIMA (1,1,12) model formed forecasts based on stationarity tests and data patterns that were more accurate than the GARCH(1,2) model. It can be concluded that the ARIMA(1,1,12)-GARCH(1,2) model produces the most accurate estimation when compared to other models.

In addition, there is also a study entitled "Estimated Value-at-Risk (VaR) on a stock portfolio with Copula" by Iriani et al. (2013). This research is motivated by the risk of investment in the form of shares which tend to be high. One method of determining investment risk is Value-at-Risk (VaR). The purpose of this study was to determine the

VaR return of several stocks from 2005 to 2010 using the Copula method. The study used the ARMA-GARCH model to obtain the residual GARCH (1,1) which was then used for copula modeling and VaR estimation. The copulas used are several Archimedean copulas, including Clayton, Frank, and Gumbel. The study showed that Clayton copula modeling as the best copula model was able to capture heavy tail better based on the resulting VaR.

In this study, the model used is the ARIMA-GARCH model to estimate the Expected Shortfall of BMRI shares. The purpose of this study is to apply the ARIMA-GARCH model in the estimation of Expected Shortfall BMRI stock data. The purpose of this study was to determine the characteristics of the analyzed stock data, to estimate the ARIMA-GARCH model from the historical data of several stocks, and to determine the amount of Expected Shortfall stock data based on the ARIMA-GARCH model. This study uses the help of Microsoft Excel and Eviews 7 applications to chart the size of the data and to estimate the stock model.

2. Literature Review

2.1. Stock

Stock is one form of investment that is considered the most promising because it has the opportunity to provide relatively large profits. However, these benefits also need to be compared with the risks involved. To analyze the stock data itself, the return value is usually used which is formulated as follows (Dowd, 2002)

$$r_t = \left(\frac{P_t - P_{t-1}}{P_{t-1}} \right), \quad (1)$$

where r_t is the value of the data return at time t , P_t is the value of data at time t , and P_{t-1} is the value of data at time $t - 1$ (1 previous time).

For stock return value analysis, usually to make the data stationary, the data is derived or differentiated with the analyzed data calculated using (Dowd, 2002)

$$r_t = \ln \left(\frac{P_t}{P_{t-1}} \right), \quad (2)$$

2.2. Normality test

The normality test has a purpose to determine whether the tested data is normally distributed. The data for the normality test in this study is the residual data return model where the normality test uses the Jarque-Berra test. The test uses the following hypothesis.

H_0 : The tested data is normally distributed.

H_1 : The tested data is not normally distributed.

The test statistic used has the following equation (Jarque and Bera, 1980)

$$JB = n \left(\frac{\zeta^2}{6} + \frac{(k-3)^2}{24} \right), \quad (3)$$

where n is the sample size, ζ is skewness, and k is kurtosis. The test criteria is that H_0 is rejected if $JB \geq \chi^2$.

2.3. Mean Model

In this section, we will discuss ARMA and ARIMA models for research. The ARMA time series model is used to briefly describe weak stationary stochastic processes, namely autoregression and moving average. This model was popularized by Box and Jenkins (1994). As the name implies, this model combines the AR model and the MA model to forecast time series data over a certain period. This model is denoted by ARMA(p, q) where p is ordered AR and q is ordered MA. The equation of the ARMA(p, q) model is as follows (Adhikari and Agrawal, 2013)

$$W_t = \beta_0 + \varepsilon_t + \sum_{i=1}^p \beta_i W_{t-i} + \sum_{i=1}^q \theta_i \varepsilon_{t-i}, \quad (4)$$

where W_t is the data at time t , β_0 is a constant, β_i is the AR model parameter coefficient which depends on the lag limit, θ_i is the MA model parameter coefficient which depends on the lag limit, and ε_t is the error data at time t .

To select good p and q , PACF is used to determine p and ACF is to determine q . In addition, AIC can also be used.

The ARIMA model is a generalization of the ARMA model. This model is used as a tool to explain the analyzed time series and predict the value of the data in the future (forecasting). This model is denoted by ARIMA(p, d, q) where p is the order for the AR process, d is the degree of data integration so that the data is stationary (the number of times the data value is different from the previous data), and q is the order for the MA process with p , d , and q are non-negative integers respectively (Uwilingiyimana et al., 2015). Process $\{W_t\}$ is ARIMA(p, d, q) if $\Delta^d W_t = (1 - L)^d W_t$ is ARMA(p, q). In general, the model is written as follows

$$\beta(L)(1 - L)^d W_t = \theta \varepsilon_t; \{\varepsilon_t\} \sim WN(0, \sigma^2), \quad (5)$$

with ε_t following the white noise (WN).

Suppose, L is the lag operator where $L^k W_t = W_{t-k}$ where the autoregressive operator and moving average are defined as follows

$$\begin{aligned} \beta(L) &= 1 - \beta_1(L) - \beta_1(L^2) - \dots - \beta_p(L^p), \\ \theta(L) &= 1 - \theta_1(L) - \theta_1(L^2) - \dots - \theta_q(L^q). \end{aligned} \quad (6)$$

The functions β and θ are autoregressive polynomials and moving averages with orders p and q in the variable L , $\theta(L) \neq 0$ if $|\theta| < 1$, $\{W_t\}$ is stationary if and only if $d = 0$, which makes the model ARMA(p, q) (Uwilingiyimana et al., 2015).

For estimating this model, the steps that need to be done are estimating the shape of the model using a correlogram, selecting the best shape, then testing the verification and validation of the model, as well as a diagnostic test (Uwilingiyimana et al., 2015).

2.4. Volatility Model

This section discusses the ARCH and GARCH models for research. The ARCH model is a statistical model that describes the variance of the analyzed time series residuals. This model is used when the error variance in the model follows the autoregressive (AR) form.

To model the time series using the ARCH(p) process, ε_t is used which denotes the residual return from the model mean, as follows

$$\varepsilon_t = \sigma_t Z_t, \{Z_t\} \sim iid N(0, 1), \quad (7)$$

where σ_t is the time-dependent standard deviation and z_t is a white noise random variable. The σ_t^2 series is modeled as follows (Uwilingiyimana et al., 2015)

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \varepsilon_{t-2}^2 + \dots + \alpha_p \varepsilon_{t-p}^2, \quad (8)$$

where $\alpha_0 > 0$, $\alpha_i \geq 0$, $i = 1, 2, \dots, p$, and Z_t where $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_t$ are independent for each t . The GARCH model is a generalization of the ARCH model developed by Bollerslev in 1986, where if ARCH is used when the residual model is in the form of AR, GARCH is used when the residual model is in the form of ARMA. The GARCH(p, q) model has a form like the ARMA model as follows

$$\varepsilon_t = \sigma_t \varepsilon_t, \varepsilon_t \sim N(0, 1),$$

where

$$\sigma_t^2 = \alpha_0 + \alpha_1 r_{t-1}^2 + \dots + \alpha_p \varepsilon_{t-p}^2 + \beta_1 \sigma_{t-1}^2 + \dots + \beta_q \sigma_{t-q}^2, \quad (9)$$

where $\alpha_0 > 0$, $\alpha_i \geq 0$, $i = 1, 2, \dots, p$, $\beta_j \geq 0$, $j = 1, 2, \dots, q$.

If $\{r_t\}$ is the return of mean, ε_t is Gaussian white noise with mean 0 and unit variance, $\{W_t\} = \{r_1, r_2, \dots, r_{t-1}\}$, then $\{r_t\}$ is GARCH(1,1) if (Uwilingiyimana et al., 2015)

$$\begin{aligned} \varepsilon_t &= \sigma_t \varepsilon_t, \varepsilon_t \sim N(0, 1), \\ \sigma_t^2 &= \alpha_0 + \alpha_1 r_{t-1}^2 + \beta_1 \sigma_{t-1}^2. \end{aligned}$$

To estimate this model from the previous ARIMA model, it must first be seen whether the residual model is heteroscedastic (containing ARCH elements), and if there is the next step, it is more or less the same as the estimation of the model means. If the diagnostic test has been carried out and the model is obtained, it must be re-checked for the heteroscedasticity of the estimated model. If the model estimate contains ARCH elements, then the model can already be used. The variance and average of these models will then be used in the next stage, namely the estimation of Expected Shortfall.

2.5. Expected Shortfall

Expected Shortfall is a method of measuring the maximum possible loss that exceeds VaR. This method was proposed by Artzner et al. (1999) when they pointed out some of the weaknesses of Value-at-Risk, including ignoring losses greater than the Value-at-Risk level and not being able to meet the axiom of coherence because Value-at-Risk is not sub-additive. Yamai and Yoshida (2002) define ES , where X is a random variable of gain or loss and $VaR(X)$ with a confidence level of $100(1 - \alpha)\%$, as follows

$$\begin{aligned} ES_{\alpha}^t(x) &= -E[X|X \leq VaR_{\alpha}(X)] \\ &= -\frac{1}{\alpha} \int_{-\infty}^{VaR_{\alpha}} xf(x)dx \\ &= -\mu_t + \sigma_t \frac{\phi(z_{1-\alpha})}{\alpha} \end{aligned} \quad (10)$$

where ϕ is the standard normal density function.

There are cases where the data distribution is not normal due to excess skewness and kurtosis resulting in deviations. So, to estimate ES, Cornish-Fisher expansion will be used so that the following formula is obtained (Situngkir, 2006)

$$\begin{aligned} F_{CF}^{-1}(\alpha) &= \phi^{-1}(\alpha) + \frac{\zeta}{6} ([\phi^{-1}(\alpha)]^2 - 1) + \frac{k-3}{24} ([\phi^{-1}(\alpha)]^3 - 3\phi^{-1}(\alpha)) \\ &\quad - \frac{\zeta^2}{36} (2[\phi^{-1}(\alpha)]^3 - 5\phi^{-1}(\alpha)), \end{aligned} \quad (11)$$

$$ES_{\alpha}^2(x) = -\hat{\mu}_t + \frac{\hat{\sigma}_t}{\alpha\sqrt{2\pi}} e^{\frac{(F_{CF}^{-1}(\alpha))^2}{2}}, \quad (12)$$

where $\hat{\mu}_t$ is the estimated mean of the data at time t ; $\hat{\sigma}_t$ is the variance of the data at time t ; $F_{CF}^{-1}(\alpha)$ is the α -quantile of the z_t distribution; $\phi^{-1}(\alpha)$ is the α -quantile of the normal distribution; ζ and k are skewness and kurtosis of $\hat{z}_t$, where $\hat{z}_t = \frac{x_t - \hat{\mu}_t}{\hat{\sigma}_t}$.

3. Results and Discussion

This section discusses the data used for analysis and the results of the analysis in the form of mean models, volatility models, and estimates of Value-at-Risk and Expected Shortfall for each selected stock.

3.1. Data

The data used for this research is historical data of daily closing prices of Bank Mandiri Persero (BMRI) shares from September 1, 2017 to August 31, 2019. Data obtained at www.financeyahoo.id.

3.2. Data Properties

The data used have descriptive statistics which are presented in Table 1.

Table 1: Descriptive Statistics of BMRI Stock

Stock	Samples (N)	Average	Median	Minimum	Maximum	Standard Deviation	Skewness	Kurtosis
BMRI	751	6014.397	5812.5	3762.5	9050	1173.498	0.283579	2.108689

Based on the table, it can be seen that the stock data has a skewness and kurtosis that deviates from the skewness, and the data kurtosis is normally distributed (0 and 3) so that the stock closing data is not normally distributed.

Then using the help of Eviews 7, it was found that the original data was not stationary. Therefore, the data must be transformed using equation (2) until the data is stationary. After one transformation, the data from the transformation of each share is stationary. So, it can be concluded that the order of d for the $ARIMA(p,d,q)$ model of BMRI stocks is 1.

3.3. Estimated Model Mean

Once it is known that the data is stationary, the shape of the model can be estimated. For the form of the mean model, the model used is the ARIMA model. After carrying out the estimation stage as described in section 2.3, it is found that the estimation of the mean stock model is in the form of ARIMA(3,1,3). The form of this model does not have a normally distributed residual. The estimation of the stock model is as follows.

The mean model for BMRI stock return data is

$$\hat{Z}_t = -0.648861Z_{t-1} + 1.289558Z_{t-2} - 0.640497Z_{t-3} - 0.698482\varepsilon_{t-1} + 0.777545\varepsilon_{t-2} + 0.479510\varepsilon_{t-3}.$$

3.4. Volatility Model Estimation

Before estimating the volatility model, it is first determined whether the mean model from the previous subsection has an ARCH element. After doing the ARCH-LM Test on each model using the help of Eviews 7, it was found that each model has an ARCH element. Because the model has an ARCH element, it is possible to estimate the volatility model. After the estimation stage, the stock volatility model is obtained in the form of GARCH(2,2). This model is significant after the validation test and the residual is white noise and is not normally distributed after the diagnostic test. Then after being checked, the models have ARCH elements so that they can be used for modeling.

The model for BMRI stock is

$$\hat{\sigma}_t^2 = 0.235788a_{t-1}^2 - 0.178785a_{t-2}^2 + 0.632555\sigma_{t-1}^2 + 0.283944\sigma_{t-2}^2.$$

The variance and mean obtained from the estimation of these models are then used for estimation and Expected Shortfall in the next subsection.

3.5. Estimated Expected Shortfall

The models of each stock option do not have a normal distribution because the residuals in the model are also not normally distributed. Therefore, to estimate the Expected Shortfall of BMRI shares, equations (11) and (12). The estimation results of BMRI's stock ES are shown in Table 2.

Table 2: Expected Shortfall of BMRI Shares

Stock	Average	Variance	Standard Deviation	Skewness z	Kurtosis z	$\phi^{-1}(5\%)$	$F_{CF}^{-1}(5\%)$	$ES_{5\%}$
BMRI	0.000832	0.000015	0.003858	0.050116	5.902366	-1.64485	-1.57197	0.008116

Based on Table 2, it can be seen that BMRI shares have an ES value of 0.008116. This means, if an investment is made for BMRI shares of IDR 1,000,000.00 for 37 days (5% of 751 days) for an investment period with a 95% confidence level, the expected loss to be borne by the investor is IDR 8,116.00.

4. Conclusions and suggestions

The first conclusion is that the original BMRI stock closing data has skewness and kurtosis that deviate from the skewness and kurtosis of the data are normally distributed (0 and 3) so that the stock closing data is not normally distributed. Then, the mean model for BMRI shares is ARIMA(3,1,3) model, while the volatility model obtained for BMRI is GARCH(2,2).

Finally, the risk level of BMRI shares has an ES of 0.008116. This means if an investment is made for BMRI shares of IDR 1,000,000.00 for 37 days (5% of 751 days) for an investment period with a 95% confidence level, the expected loss to be borne by the investor is IDR 8,116.00.

For further research, other models can be used, such as the Seasonal ARIMA (SARIMA) or ARFIMA for the mean model and the Exponential GARCH (EGARCH) or Fractionally Integrated GARCH (FIGARCH) for the volatility model.

References

- Adhikari, R., & Agrawak, K. K. (2013). *An Introductory Study on Time Series Modeling and Forecasting*. Lambert Academy Publishing.
- Artzner, P., Delbaen, Eber, J. M., & Heath, D. (1999). Coherent Measures of Risk. *Mathematical Finance*, 9, 203-228.
- Box, G. E. P., Jenkins, G. M., & Reinsel, G. C. (1994). *Time Series Analysis: Forecasting and Control*, 3rd ed. Englewood Cliffs, NJ: Prentice Hall.
- Dowd, K. (2002). *An Introduction to Market Risk Measurement*. John Wiley & Sons Inc.

- Iriani, N. P., Akbar, M. S., & Haryono. (2013). Estimasi Value at Risk (VaR) pada Portofolio Saham dengan Copula. *Jurnal Sains Dan Seni Pomits*, 2(2), 195-200.
- Jarque, C. M., & Bera, A. K. (1980). Efficient Test for Normality, Homoscedasticity, and Serial Independence of Regressions Residuals. *Economics Letter*, 6, 255-259.
- Religiositas. (2016). *Pada KBBI Daring*. Taken on July 4, 2021, from <https://kbbi.kemdikbud.go.id/entri/religiositas>
- Situngkir, H. (2006). *Value at Risk yang Memperhatikan Sifat Statistika Distribusi Return*. Bandung Fe Institute. Diakses di: <http://mpira.ub.uni-muenchen.de/895/>
- Uwilingiyimana, C., Haerimana, J. D., & Munga'tu, J. (2015). Forecasting Inflation in Kenya Using ARIMA-GARCH Models. *International Journal of Management and Commerce Innovations*, 3(2), 15-27.
- Yamai, Y., & Yoshiba, T. (2002). On The Validity of Value-at-Risk: Comparative Analysis with Expected Shortfall. *Monetary and Economic Studies*, 20, 57-86.