



Estimation of the Value-at-Risk (VaR) Using the TARARCH Model by Considering the Effects of Long Memory in Stock Investments

Nurfadhлина Abdul Halim^{a*}, Agus Supriatna^b, Adhy Prasetyo^b

^a*Faculty of Science and Technology, Universiti Sains Islam Malaysia,
Bandar Baru Nilai, 71800, Nilai, Negeri Sembilan, Malaysia*

^b*Department of Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Padjadjaran,
Jl. Raya Bandung Sumedang KM.21, Hegarmanah, Kec. Jatinangor, Kabupaten Sumedang, Jawa Barat 45363, Indonesia*

*corresponding author: nurfadhлина@usim.edu.my

Abstract

Value at Risk (VaR) is one of the standard methods that can be used in measuring risk in stock investments. VaR is defined as the maximum possible loss for a particular position or portfolio in the known confidence level of a specific time horizon. The main topic discussed in this thesis is to estimate VaR using the TARARCH (Threshold Autoregressive Conditional Heteroscedasticity) model in a time series by considering the effect of long memory. The TARARCH model is applied to the daily log return data of a company's stock in Indonesia to estimate the amount of quantile that will be used in calculating VaR. Based on the analysis, it was found that with a significance level of 95% and assuming an investment of 200,000,000 IDR, the VaR using the TARARCH model approach was 5,110,200 IDR per day.

Keywords: Long memory, VaR, TARARCH models.

1. Introduction

In the capital market, almost all investments contain an element of uncertainty or risk. Investors do not know with certainty the results that will be obtained from the investment they do. Because investors face risky investment opportunities, a measurement tool is needed to test these market risks, so that investors can safely know how far they can invest (Batuparan, 2001; Jogiyanto. 2007).

Based on the description above, the purpose of this paper is to calculate the amount of Value at Risk (VaR) of a stock, which is to calculate the maximum loss at a certain position with the level of confidence that has been known in a specific horizon time. The approach used is the TARARCH time series model to estimate the size of the quantile that will be used in the calculation of Value at Risk (VaR) (Sukono et al., 2019).

2. Formulation of the Problem

Suppose there is a closing stock data, then the log return value will be sought as follows

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

where P_t is the share price in period t and P_{t-1} is the stock price in period $t - 1$.

Based on the log return data will be tested whether the data is stationary and if it is stationary then the data will be processed to get the mean equation model and the variance equation model. The mean equation model will follow the time series model, while the variance equation model will follow the TARCh (Threshold Autoregressive Heteroscedasticity) model. After getting the mean equation model and the variance equation model, the quantile size will be calculated with a known level of confidence, for example by 95% (Ruppert, 2004).

The problem that will be discussed in this paper is how to calculate the value of VaR by modeling the mean and variance equations that produce quantile values to be calculated together with a number of investments given.

3. Literature Review

3.1. Stock Returns

Return is the income that will be received if we invest money in a financial asset (stocks, bonds) or real assets (property, land). Return calculation is done as in equation (1).

3.2. Time Series Model

There are several types of time series models that can be used in modeling the mean and variance for VaR calculations as follows.

Autoregressive (AR). For r_t log return series, the AR(p) model is

$$r_t = \phi_1 r_{t-1} + \phi_2 r_{t-2} + \Lambda + \phi_p r_{t-p} + a_t. \tag{2}$$

Moving Average (MA). For r_t log return series, the MA(q) model is

$$r_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \Lambda - \theta_q a_{t-q}. \tag{3}$$

Autoregressive Moving Average (ARMA). For r_t log return series, the ARMA(p, q) model is

$$r_t = \phi_1 r_{t-1} + \phi_2 r_{t-2} + \Lambda + \phi_p r_{t-p} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \Lambda - \theta_q a_{t-q}. \tag{4}$$

Autoregressive Integrated Moving Average (ARIMA). In general, the equation for the ARIMA($p, 1, q$) model is

$$W_t = \phi_1 W_{t-1} + \phi_2 W_{t-2} + \Lambda + \phi_p W_{t-p} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \Lambda - \theta_q a_{t-q}. \tag{5}$$

Autoregressive Conditional Heteroscedasticity (ARCH). Specifically, the ARCH(p) model is assumed as

$$a_t = \sigma_t \varepsilon_t, \quad \sigma_t^2 = \alpha_0 + \alpha_1 a_t^2 + \Lambda + \alpha_p a_{t-p}^2, \tag{6}$$

where $\{\varepsilon_t\}$ is a series of independent and identically distributed (iid) random variables with mean 0 and variance 1, $\alpha_0 > 0$, $\alpha_i \geq 0$, for $i > 0$ (Beronilla et al., 2007; Cryer, 1986; Mood et al., 1963).

Threshold Autoregressive Conditional Heteroscedasticity (TARCh). For r_t return log series, the TARCh(p, q) model for $p > 0$ and $q > 0$ are integers, defined as (Ruppert, 2004; Klienbaum et al., 1988):

$$r_t = \mu + a_t, \quad a_t = \sigma_t \varepsilon_t, \quad \sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + \beta_1 \sigma_{t-1}^2 + \Lambda + \gamma a_{t-1}^2 I_{t-1}, \tag{7}$$

where $\{ \varepsilon_t \}$ is the sequence of iid random variables with mean 0 and variance 1.

3.3. Quantiles

Quantile can be interpreted as a value that divides a set of data into several equal parts (Ruppert, 2004).

If the cumulative distribution function of X is continuous and does not go down, then the function of X has an inverse function F^{-1} . For each q between 0 and 1, $F^{-1}(q)$ is called q -quantile or the 100th percentile. The probability of a continuous X cumulative distribution function under q -quantile is exactly q .

Quantile that will be used in VaR calculation can be calculated as follows

$$\text{quantile} = \hat{r}_t - \eta \sqrt{\sigma_t^2} \quad (10)$$

where η is the value of the confidence coefficient chosen in normal distribution, \hat{r}_t is the result of forecasting the mean model, and $\sqrt{\sigma_t^2}$ is the volatility value resulting from forecasting variance modeling (Dowd, 2002; Herrhyanto, 2003).

3.4. Value at Risk (VaR)

Value at Risk which is usually abbreviated as VaR is generally defined as the maximum possible loss for a particular position or portfolio in the known confidence level of a specific time horizon (Redhead, 1997).

VaR calculations using the TARCh model can be calculated as follows

$$\text{VaR} = \text{amount of investment} \times \text{quantile} \quad (11)$$

4. Data Analysis

4.1. Data

The data used in this paper is taken from <http://finance.yahoo.com/q/hp?s=MPPA.JK> for Matahari Putra Prima Tbk daily shares. Observation data is the closing stock price for the last 1080 days (13 November 2000 to 31 December 2004) which lasts for five days a week except holidays. The characteristics of the analyzed data are log return (Continuously Compounded Return) of stock trading which is calculated from the closing price of the Matahari Putra Prima Tbk (MPPA) stock trading. The graph of the company data can be seen in Figure 1.

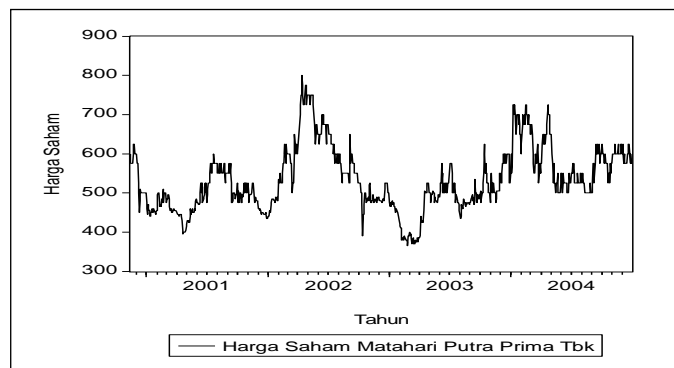


Figure 1: Plot of closing stock price of Matahari Putra Prima Tbk

4.2. Stock Return Calculation

Using equation (1) a daily log return will be calculated from Matahari Putra Prima Tbk's closing stock price data in 1080 days.

For example, the price of Matahari Putra Prima Tbk shares on November 13, 2000, and November 14, 2000, is 600 IDR and 575 IDR, then in a time horizon of 1 day (24) hours, the 1st daily log return obtained is

$$r_{M1} = \log\left(\frac{575}{600}\right) = -0.018483$$

And so on to calculate the 2nd daily log return to the 1079th daily log return at the company Matahari Putra Prima Tbk. The log return chart of the Matahari Putra Prima Tbk stock can be seen in Figure 2.

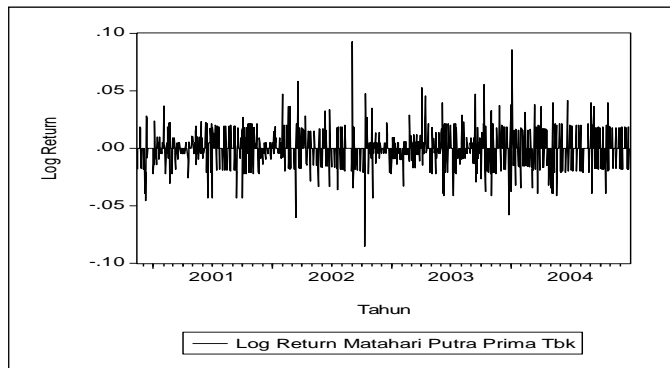


Figure 2: Plot log return data for closing stock prices of Matahari Putra Prima Tbk

4.3. TARCH Model

In modeling variance with the TARCH model, we must do the mean modeling, where the mean modeling and variance modeling is carried out simultaneously between the two. Analysis of the TARCH model is assisted with Eviews 4.1 software.

4.3.1. Modeling Equation Mean

Figure 2 shows that Matahari Putra Prima Tbk's log return data has been stationary in the mean, but this data will be performed differencing to see the long-term effect on the data. Long memory testing will be carried out using the Geweke and Porter-Hudak method. The calculation to find the value of d is by using software R, so that obtained $d = -0.182$. To identify an appropriate model can be seen from the ACF and PACF plots can be seen in Figure 3 and 4.

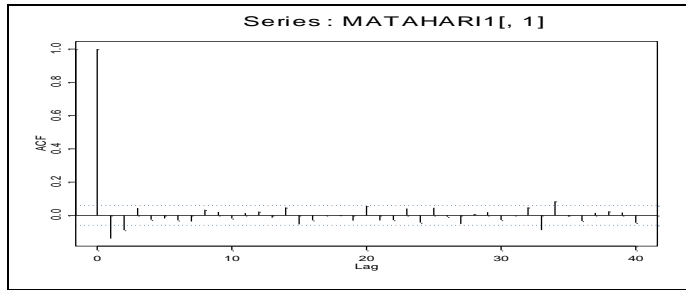


Figure 3: Plot ACF data log return Matahari Putra Prima Tbk

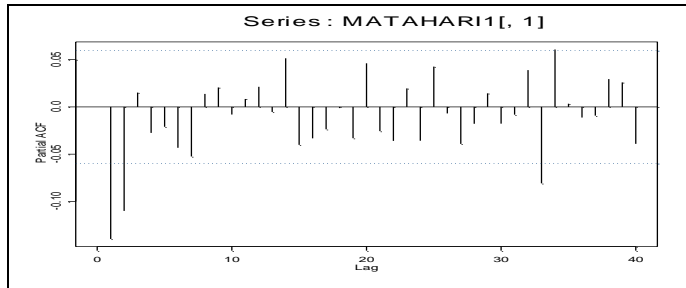


Figure 4: Plot PACF data log return Matahari Putra Prima Tbk

From the ACF and PACF plots, there are indications to model the equation of the Matahari Putra Prima Tbk log return mean data with the AR(2) or MA(2) models. Among the possibilities of the two models, the best model is AR(2), with the estimated parameters in Table 1.

Table 1: Estimated parameters of the AR(2) model

Dependent Variable: MATAHARI				
Method: Least Squares				
Date: 04/08/08 Time: 21:21				
Sample(adjusted): 15/11/2000 30/12/2004				
Included observations: 1077 after adjusting endpoints				
Convergence achieved after 3 iterations				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(1)	-0.155132	0.030315	-5.117296	0.0000
AR(2)	-0.109808	0.030315	-3.622188	0.0003
R-squared	0.031361	Mean dependent var		1.86E-09
Adjusted R-squared	0.030460	S.D. dependent var		0.015125
S.E. of regression	0.014893	Akaike info criterion		-5.574048
Sum squared resid	0.238424	Schwarz criterion		-5.564796
Log likelihood	3003.625	Durbin-Watson stat		1.996717
Inverted AR Roots	-.08+.32i	-.08 -.32i		

From Table 1, it can be seen that the parameter AR(2) has been significantly different from zero, this can be seen from the probability value that is smaller than the tolerance level of $\alpha = 5\%$ so that the model equation is obtained as follows

$$r_t = -0.155132r_{t-1} - 0.109808r_{t-2} + a_t.$$

After estimating the parameters, a standardized residual correlogram will be seen in Table 2 to see whether or not there is a serial residual correlation effect for modeling the mean AR(1) log return data.

Table 2: Residual model corelogram AR(2)

Date: 04/10/08 Time: 07:54 Sample: 15/11/2000 30/12/2004 Included observations: 1077 Q-statistic probabilities adjusted for 2 ARMA term(s)						
Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	0.002	0.002	0.0029	
		2	-0.001	-0.001	0.0041	
		3	0.008	0.008	0.0821	0.774
		4	-0.037	-0.037	1.5307	0.465
		5	-0.029	-0.028	2.4128	0.491
		6	-0.040	-0.040	4.1079	0.392
		7	-0.037	-0.037	5.6234	0.345
		8	0.027	0.027	6.4436	0.375
		9	0.020	0.019	6.8988	0.439
		10	-0.008	-0.011	6.9653	0.540
		11	0.016	0.011	7.2382	0.612
		12	0.028	0.026	8.0781	0.621
		13	-0.007	-0.007	8.1341	0.701
		14	0.037	0.038	9.6240	0.649
		15	-0.057	-0.054	13.172	0.435
		16	-0.036	-0.034	14.618	0.405
		17	-0.019	-0.020	15.024	0.450
		18	-0.004	0.002	15.043	0.522
		19	-0.031	-0.032	16.103	0.517
		20	0.048	0.044	18.601	0.417

From Table 2, it appears that there is no serial correlation in the data, this is shown by the Q-Stat probability that is greater than the tolerance level $\alpha = 5\%$ and there is no significant ACF and PACF plot until lag-20. Therefore, the AR(2) model is suitable for modeling the mean return data log of Matahari Putra Prima Tbk.

4.3.2. Variance Modeling Stage

To find out whether after modeling the mean ARCH effect is still present, an ARCH LM (Lagrange Multiplier) test will be performed which is listed in Table 3.

Table 3: ARCH LM AR(2) model tests

ARCH Test:			
F-statistic	9.184268	Probability	0.002499
Obs*R-squared	9.123353	Probability	0.002524

Table 3 shows that there is an ARCH element in Matahari Putra Prima Tbk's log return data, this can be seen from the value of the statistical F probability that is smaller than the tolerance level $\alpha = 5\%$.

Once it is known that the log return data contains ARCH elements, the data can be identified by looking at the ACF and PACF plots of squared residuals $a_t = r_t - \hat{r}_t$, which are found in Table 4.

Table 4: Correlogram residual squared AR(2) model

Date: 04/09/08 Time: 20:45 Sample: 15/11/2000 30/12/2004 Included observations: 1077 Q-statistic probabilities adjusted for 2 ARMA term(s)						
Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	0.092	0.092	9.1548	
		2	0.043	0.035	11.150	
		3	0.039	0.032	12.755	0.000
		4	0.018	0.010	13.088	0.001
		5	0.040	0.035	14.803	0.002
		6	0.090	0.089	14.803	0.006
		7	0.058	0.056	18.402	0.002
		8	0.028	0.025	18.412	0.006
		9	-0.008	-0.008	19.214	0.014
		10	0.036	0.033	20.631	0.014
		11	-0.019	-0.014	21.007	0.021
		12	-0.004	0.004	21.026	0.033
		13	-0.007	-0.005	21.084	0.049
		14	0.013	0.017	21.257	0.066
		15	0.017	0.021	21.585	0.088
		16	0.060	0.068	26.514	0.043
		17	0.032	0.023	26.630	0.046
		18	0.010	0.019	26.737	0.046
		19	-0.003	-0.010	26.746	0.084
		20	-0.003	-0.010	26.746	0.084

From Table 4, it appears that the ACF and PACF plots are significant in lag 1, so there are indications for modeling the variance equation using the TAR(1,1) model. In modeling the variance with the TAR(1,1) model and with the mean AR(2) model, the following equation is obtained

$$r_t = \phi_1 r_{t-1} + \phi_2 r_{t-2} + a_t,$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + \beta_1 \sigma_{t-1}^2 + \gamma a_{t-1}^2 I_{t-1}.$$

The estimation results of the AR (2) -TAR(1,1) model are in Table 5.

Table 5: Estimated parameters of the AR(2)-TAR(1,1) model

Dependent Variable: MATAHARI				
Method: ML - ARCH (Marquardt)				
Date: 12/30/08 Time: 09:26				
Sample(adjusted): 3 1079				
Included observations: 1077 after adjusting endpoints				
Convergence achieved after 18 iterations				
Variance backcast: ON				
	Coefficient	Std. Error	z-Statistic	Prob.
AR(1)	-0.155108	0.038218	-4.058512	0.0000
AR(2)	-0.125010	0.034677	-3.605019	0.0003
Variance Equation				
C	7.35E-05	1.02E-05	7.200332	0.0000
ARCH(1)	0.152192	0.032905	4.625142	0.0000
(RESID<0)*ARCH(1)	0.042013	0.037520	1.119734	0.2628
GARCH(1)	0.509494	0.057469	8.865474	0.0000
R-squared	0.031130	Mean dependent var	-1.49E-11	
Adjusted R-squared	0.026607	S.D. dependent var	0.015125	
S.E. of regression	0.014922	Akaike info criterion	-5.612909	
Sum squared resid	0.238483	Schwarz criterion	-5.585155	
Log likelihood	3028.552	Durbin-Watson stat	1.996370	
Inverted AR Roots	-.08+.34i	-.08 -.34i		

From Table 5, it can be seen that the AR(2)-TAR(1,1) parameter has been significantly different from zero, this can be seen from the probability value that is smaller than the tolerance level $\alpha = 5\%$. So the AR(2)-TAR(1,1) model can be written as

$$r_t = -0.155108r_{t-1} - 0.125010r_{t-2} + a_t,$$

$$\sigma_t^2 = (7.35 \times 10^{-5}) + 0.152192a_{t-1}^2 + 0.042013I_{t-1} + 0.509494\sigma_{t-1}^2.$$

After estimating the parameters, it will be seen whether the AR(2)-TAR(1,1) model still has the ARCH effect.

Table 6: LM ARCH Test MA(2)-TAR(1,1) model

ARCH Test:			
F-statistic	0.217414	Probability	0.641112
Obs*R-squared	0.217775	Probability	0.640741

From Table 6, it can be seen that there is no ARCH element in the log return data, this can be seen from the statistical F probability value greater than the tolerance level $\alpha = 5\%$.

To see whether there are still serial correlations in the model, look at the ACF and PACF squared residual plots standardized in Table 7.

Table 7: Standardized Residual Squared Correlogram

Date: 04/09/08 Time: 20:42 Sample: 15/11/2000 30/12/2004 Included observations: 1077 Q-statistic probabilities adjusted for 2 ARMA term(s)						
Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	0.092	0.092	9.1548	
		2	0.043	0.035	11.150	
		3	0.039	0.032	12.755	0.000
		4	0.018	0.010	13.088	0.001
		5	0.040	0.035	14.803	0.002
		6	0.000	-0.009	14.803	0.005
		7	0.058	0.056	18.402	0.002
		8	-0.004	-0.016	18.417	0.005
		9	-0.026	-0.029	19.142	0.008
		10	-0.008	-0.008	19.214	0.014
		11	-0.036	-0.033	20.631	0.014
		12	-0.019	-0.014	21.007	0.021
		13	-0.004	0.004	21.028	0.033
		14	-0.007	-0.005	21.084	0.049
		15	0.013	0.017	21.257	0.068
		16	0.017	0.021	21.585	0.088
		17	0.060	0.058	25.514	0.043
		18	0.032	0.023	26.630	0.046
		19	-0.010	-0.019	26.737	0.062
		20	-0.003	-0.010	26.746	0.084

From Table 7, it can be seen that in the ACF and PACF plots, all bar graphs do not cross the dotted line (Bartlett line). Besides that the statistical probability value of Q is greater than the level of significance $\alpha = 5\%$. This shows that there is no serial correlation in the model.

Next will be seen whether the residual model has a nominal distribution. By using the Tools ►Input Analyzer in ARENA 7 software, the results are shown in Figure 5.

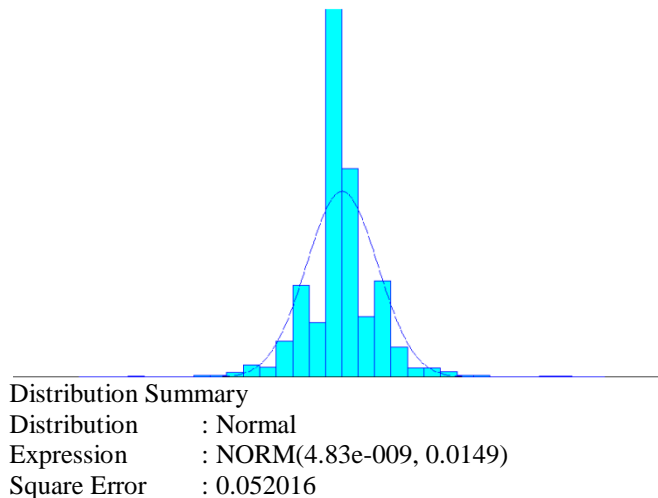


Figure 5: Residual normal distribution histogram AR(2)-TARCH(1,1) model

Figure 5 shows that the AR(2)-TARCH(1,1) residuals follow a bell curve, which means that the data is normally distributed.

5. Calculating Value at Risk (VaR)

After obtaining the results of the mean modeling stage and the variance modeling stage, the VaR amount will be calculated in the Matahari Putra Prima Tbk log return data. For example, if it is assumed to allocate funds of 200,000,000 IDR for investment in Matahari Putra Prima Tbk, the amount of VaR in Matahari Putra Prima Tbk shares will be calculated as follows.

Suppose that ε_t is the normal distribution. AR(2)-TARCH(1,1) model is

$$\hat{r}_t = -0.155108r_{t-1} - 0.125010r_{t-2},$$

$$\hat{\sigma}_t^2 = (7.35 \times 10^{-5}) + 0.152192a_{t-1}^2 + 0.042013\mathcal{I}_{t-1} + 0.509494\sigma_{t-1}^2.$$

Will be calculated \hat{r}_{1079} and $\hat{\sigma}_{1079}^2$, i.e.

$$\hat{r}_{1079} = -0.155108r_{1078} - 0.125010r_{1077} = -0.155108(-0.018483) - 0.125010(0.018483) = 0.000511$$

and σ_t variance values can be generated directly from Eviews 4.1 software. So the 1079th variance value is $\sigma_{1079}^2 = 0.000251$, so the volatility value is $\sqrt{0.000251} = 0.015843$

To calculate the quantile magnitude, look for the known \hat{r}_{1079} and $\hat{\sigma}_{1079}^2$. The quantile size will be calculated with $\alpha = 5\%$ (95% confidence coefficient), i.e.

$$\text{quantile}(0.05) = \hat{r}_{1079} - 1.645(\hat{\sigma}_{1079}) = 0.000511 - 1.645(0.015843) = -0.025551$$

where the negative sign is written as the left tail of the conditional normal distribution.

Using equation (10), obtained VaR for Matahari Putra Prima Tbk shares, i.e.

$$\text{VaR} = 200,000,000 \text{ IDR} \times 0.025551 = 5,110,200 \text{ IDR}.$$

So it can be concluded that with a 95% confidence level, there will be a loss of 5,110,200.00 IDR per day.

6. Conclusion

From the discussions that have been carried out, the following conclusions are obtained.

- 1) A fairly good time series model in modeling the Matahari Putra Prima Tbk log return data after differencing in the last 1080 days (13 November 2000 to 31 December 2004) is the AR(2) model.
- 2) The variance model that is good enough to model the Matahari Putra Prima Tbk log return data in the last 1080 days (13 November 2000 to 31 December 2004) is the TARCH(1.1) model.
- 3) After estimating the modeling of mean and variance equations, the results show that the appropriate model for the Matahari Putra Prima Tbk log return data is the AR(2)-TARCH(1,1) model. Using a 95% confidence level and assuming an investment of 200,000,000.00 IDR, then the value of the VaR Matahari Putra Prima Tbk is 5,110,200 IDR. So the loss obtained for Matahari Putra Prima Tbk shares is 5,110,200.00 IDR per day.

References

- Batuparan, S. D. (2001). Kerangka Kerja Risk Management. (www.bexi.co.id/images/_res/perbankan-Kerangka-Kerja-Risk-Management.pdf, accessed on 2 Maret 2008).
- Beronilla, L., Nikkin, & Mapa, S. D. (2007). Range-Based GARCH: A New Method of Estimating Value-at-Risk. (www.nscb.gov.ph/ncs/10thNCS/abstracts/Contributed/36_Forecasting/abstract_mapa.pdf, accessed on 2 Maret 2008).
- Cryer, J. D. (1986). Time Series Analysis. Boston: PWS-KENT Publishing Company.

- Dowd, K. 2002. *An Introduction to Market Risk Measurement*. Chichester, New York: John Wiley & Sons.
- Herrhyanto, N. (2003). *Statistika Matematis Lanjutan*. Bandung: Pustaka Setia.
- Jogiyanto. 2007. *Teori Portofolio dan Analisis Investasi*. Yogyakarta: BPFE.
- Klienbaum, D. G., Kuper, L. L. and Muller, K. E. 1988. *Applied Regression and Other Multivariabel Methods*, Second Edition. Boston: PWS-KENT Publishing Company.
- Mood, A. M., Graybill, F. A. and Boes, D. C. 1963. *Introduction to the Theory of Statistics*, Third Edition. New York: McGraw Hill Book Company.
- Ruppert, D. 2004. *Statistics and Finance an Introduction*. New York: Springer.
- Redhead, K. 1997. *Financial Derivatives: An Introduction to Future, Forwards, Options and Swaps*. Prentice Hall Europe.
- Sukono, Subartini, B., Napitupulu, H., Hidayat, Y., Putra, A.S., and Bon, A.T. 2019. Value-at-Risk and Modified Value-at-Risk under Asset Liability by Using Time Series Approach. *Proceedings of the International Conference on Industrial Engineering and Operations Management*, Bangkok, Thailand, March 5-7, 2019, 2106-2117.