International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

Volatility of Fish Prices in Manado City Using Multivariate GARCH Model

Nelson Nainggolan[1], Hanny Andrea Huibert Komalig[2], Tohap Manurung[3], Dingse Pandiangan[4]

[1,2,3]Department of Mathematics, Faculty of Mathematics and Natural Sciences, Sam Ratulangi University, Manado, Indonesia
[4]Department of Biology, Faculty of Mathematics and Natural Sciences, Sam Ratulangi University, Manado, Indonesia

Abstract
Time series data modeling is generally performed using the assumption of homoscedasticity or residual variance that is constant over time. However, the homoscedasticity assumption cannot answer the problem of volatility in economic and business time series data, because generally data on economics and business have residual variances that always change over time. Therefore, a model developed with the assumption of not constant variance is known as heteroskedasticity model. Multivariate GARCH model is a development of the univariate GARCH model. The multivariate GARCH model can be viewed as a conditional heteroskedasticity model in a multivariate time series. This paper discusses the parameterization of covariance matrices such as Vech model representation, BEEK model and Constant Correlation model. For parameter estimation the maximum likelihood method is used. Furthermore, multivariate GARCH model application is applied for multivariate model. In this research, the data used are data on the prices of fresh fish in Manado. The fish price data is the price of Tuna, Skipjack (Cakalang), Cobe (Tongkol), Kite (Malalugis) and Selar fish (Tude). The ARCH effect test results show that the data meet the assumption of non-constant variance (heteroscedasticity). GARCH mutivariate modeling in the form of a VEC diagonal model gives the results of the price volatility of fresh fish in Manado.

Keywords: Multivariate GARCH Model, Fish Prices, Volatility

1. Introduction

Food ingredients are materials that can be used as food ingredients that are safe, have a good taste and are healthy for humans. Vegetable foods include rice, corn, peanuts, vegetables, fruit, cooking oil and white sugar. Meanwhile, animal foods include beef and buffalo, chicken, eggs, milk and fish. Food is material that makes humans grow and develop well.

Fresh fish is fish that still has the same characteristics as live fish, both in appearance, taste and texture. In other words, fresh fish is fish that has just been caught and has not undergone any preservation or further processing. Fish that has not undergone physical or chemical changes or that still has the same properties when caught.

Manado City, the capital of North Sulawesi province, is famous for being rich in marine fish. Fresh fish can often be found in this city and is served in Manado restaurants. Fresh fish that are often found include snapper, grouper, fight, tuna, skipjack and others.

In an effort to increase the contribution of fisheries and marine businesses in the economy and implement fisheries revitalization, it is deemed necessary to provide information. The success of managing fisheries and marine resources, including tourism, is determined by the availability of accurate and always up-to-date information, because accurate information will help formulate policies or plan resource management so that the output produced is more effective and on target. One form of providing information is in mathematical modeling, namely time series modeling.

The Time Series model can analyze stationary and non-stationary data, as well as seasonal or non-seasonal data. Time Series-based models include the Autoregressive Moving Average (ARMA) model, a combination of the Autoregressive Model (AR) and the Moving Average (MA) Model. This model includes a stationary time series model [1]. Autoregressive Integrated Moving Average (ARIMA) modeling is an extension of the ARMA model for non-stationary data. However, the ARMA and ARIMA models only use one variable (univariate), so if you want simultaneous modeling with several variables, it cannot be done with this model but using a multivariate time series model. The multivariate model can explain the relationship between observations on certain variables at a time with observations on the variable itself at previous times, and also its relationship with observations on other variables at previous times.

The conventional time series model assumes that the variance error is constant over time. The assumption of constant variance is an ideal assumption rarely encountered in real situations, especially related to the financial field [2]. Therefore, a model developed with the assumption of unstable variance is known as heteroskedasticity model [3]. The heteroskedastic model is not only in univariate form but also in multivariate form [4]. The Generalized of space time autoregressive (GSTAR) model with ARCH error has also been developed [5].

Multivariate GARCH model is a combination of several univariate time series simultaneously, meaning that one univariate time series does not only depend on its own variable but also depends on other univariate time series variables [6][7]. The VAR model generalizes to a single variable (univariate autoregressive model by allowing for multivariate time series. VAR models are often used in economics and the natural sciences. The time series analysis of the VAR model was also developed by Hamilton [8], Enders [9]. Application of the Autoregressive Vector Model on International Tourist Arrivals is also provided by Lee [10]. Economic globalization and rapidly developing communication technology make prices from one market to another and commodities move to influence each other [11]. Therefore, multivariate time series modeling is very necessary.

Price volatility analysis is very necessary for economic actors because the results of the analysis can be used to make decisions regarding business risk issues. As stated previously, agricultural commodity prices have very high volatility. The impact that arises from data with high volatility is that the error variable has a variance that is not constant. Therefore, another method for modeling the behavior of data with high volatility can use the Autoregressive Conditional Heteroscedasticity (ARCH) and Generalized Autoregressive Conditional Heteroscedasticity (GARCH) models [3].

2. ARIMA Model

Autoregressive (AR) model is a model which states that the value of the current observation depends on the value of the observations at previous times. The value of Z_t at time t is regressed to the previous values of itself plus white noise. Autoregressive process of order p written with AR(p) is of the form [2]:

$$Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \ldots . + \phi_p Z_{t-p} + \varepsilon_t \quad (1)$$

with assumption

$$\varepsilon_t \overset{iid}{\sim} N(0, \sigma^2) .$$

When using the Backshift B operator, the AR(p) model can be written as

$$(1 - \phi_1 B - \phi_2 B^2 - \ldots . + \phi_p B^p) Z_t = \varepsilon_t$$

The second stationary time series model that will be discussed is the Moving Average (MA). The Moving Average process is a linear process which is expressed in the form

$$Z_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \cdots - \theta_q \varepsilon_{t-q}$$
$$Z_t = (1 - \theta_1 B - \ldots\ldots\ldots - \theta_q B^q)\varepsilon_t$$

where

ε_t : white noise with zero mean and variance σ2 ,

$\theta_1, \cdots, \theta_q$: are parameters

The Autoregressive Moving Average (ARMA) process is a mix of AR and MA models. The ARMA(p,q) model can be written as

$$Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \ldots . + \phi_p Z_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_q \varepsilon_{t-q}$$
$$Z_t - \phi_1 Z_{t-1} - \phi_2 Z_{t-2} - \ldots . - \phi_p Z_{t-p} = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_q \varepsilon_{t-q}$$
$$\phi_p(B) Z_t = \theta_q(B)\varepsilon_t \quad (2)$$

where

$$\varepsilon_t \overset{iid}{\sim} N(0, \sigma^2)$$
$$\phi_p(B) = 1 - \phi_1 B - \phi_2 B^2 - \ldots . + \phi_p B^p$$
$$\theta_q(B) = 1 - \theta_1 B - \ldots\ldots\ldots - \theta_q B^q$$

The analyzed data does not always show stationary, because non-stationary data are often found in real life. A time series whose mean is not constant is not stationary. One way of stationary data is through differencing. The Autoregressive Integrated Moving Average (ARIMA) process is a time series model for non-stationary data.

Stationary processes can be obtained from the results of the differentiation of non-stationary data. The random variable $\{Z_t\}$ is said to be an Autoregressive Integrated Moving Average model, written ARIMA (p,d,q), if it is distinguished d times and is a stationary ARMA process. Usually encountered in practice, the degree of differentiation: d = 1 or d = 2 is sufficient.

In general, the model equation for ARIMA (p,1, q) is:

$$Z_t - Z_{t-1} = \phi_1 (Z_{t-1} - Z_{t-2}) + \phi_2 (Z_{t-2} - Z_{t-3}) + \ldots . + \phi_p (Z_{t-p} - Z_{t-p-1})$$

$$+\varepsilon_t - \theta_1\varepsilon_{t-1} - \cdots - \theta_q\varepsilon_{t-q} \qquad (3)$$

In particular, the ARIMA (1,1,1) model satisfies the equation:

$$Z_t - Z_{t-1} = \phi_1(Z_{t-1} - Z_{t-2}) + \varepsilon_t - \theta_1\varepsilon_{t-1}$$

3. Univariate ARCH and GARCH Model

The Autoregressive Conditional Heteroscedastic (ARCH) model was introduced by Engle (1982) [3] which is a time series model that can accommodate heteroscedastic properties. Then Bollerslev (1986) developed the ARCH model into a Generalized Autoregressive Conditional Heteroscedastic (GARCH) model [4].

In conventional time series models such as the autoregressive moving average (ARMA) model it is assumed that the error variance (ε_t) is constant, that is Var (ε_t) = σ^2. Suppose the conditional variance of ε_t is not constant, then the variance of Y_t conditional on Y_{t-1} is not constant, , $\boldsymbol{Var\ (\varepsilon_t) = \sigma_t^2}$. One strategy is to model conditional variance as AR(q) process through the preceding error square, that is

$$\sigma_t^2 = \alpha_0 + \alpha_1\varepsilon_{t-1}^2 + \cdots + \alpha_{t-q}\varepsilon_{t-q}^2 + \eta_t \qquad (4)$$

with η_t is a white-noise process. For this reason, (4) is called an autoregressive conditional heteroscedastic (ARCH) model [3]. Engle then proposes a scheme in which heteroscedasticity depends on previous Y_t values, namely

$$Y_t = \eta_t\sqrt{h_t} \text{ and } h_t = \alpha_0 + \alpha_1 Y_{t-1}^2 \qquad (5)$$

with η_t iid.N(0.1). The values of $\alpha_0 > 0$ and $\alpha_1 > 0$. Then (5) is called the ARCH (1) model. Note that the variance of Y_t conditional on Y_{t-1} is

$$Var(Y_t|Y_{t-1}) = Var\left(\eta_t\sqrt{h_t}\,|Y_{t-1}\right) = h_t Var(\eta_t) = h_t$$

If it is related to the application of the model, let Y_t be inflation then a process ARCH (1) states that high inflation in the past period will result in great variance at the present time.

Furthermore, if the ARCH(q) process is included lag of $h_t = \sigma_t^2$ then obtained model GARCH (p,q) [2], namely

$$h_t = \alpha_0 + \beta_1 h_1 + \cdots + \beta_p h_{t-p} + \alpha_1 Y_{t-1}^2 + \cdots + \alpha_q Y_{t-9}^2$$

where p denotes lag on σ_t^2 and q states lag on Y_t^2. Specifically for p = 1 and q = 1 obtained the GARCH model (1,1) ie

$$h_t = \alpha_0 + \alpha_1 Y_{t-1}^2 + \beta_1 h_{t-1}.$$

4. Multivariate ARCH and GARCH Model

Multivariate time series or time series vector can be interpreted as a combination of several univariate time series, where one univariate time series does not only depend on its own variables but also depends on other univariate time series variables. In multivariate time series, each univariate time series does not stand alone but the variables between univariate time series influence each other.

This multivariate time series has a pattern similar to a system of linear equations with two or more variables. Therefore, in studying multivariate time series, it is necessary to have sufficient understanding of the concepts of matrices and vectors. As usual in writing, bold letters are used, namely bold capital letters to represent matrices and bold lowercase letters to represent vectors.

Economic globalization and rapidly developing communication technology make prices from one market to another and commodities move to influence each other. Therefore, multivariate time series modeling becomes very necessary.

Let $\mathbf{Z}_t$ be a time series multivariate process with dimension m, that is

$$\boldsymbol{Z}_t = \begin{pmatrix} Z_{1,t} \\ Z_{2,t} \\ \vdots \\ Z_{m,t} \end{pmatrix}, \qquad t = 0, \pm 1, \pm 2, \ldots.$$

$\mathbf{Z}_t$ is said to be weakly stationary, if it satisfies:

i. Mean: $E(\mathbf{Z}_{i,t}) = \mu_i$ constant, for $i = 1, 2, 3, \ldots, m$.
ii. Cross-covariance between $\mathbf{Z}_{i,t}$ and $\mathbf{Z}_{j,t}$ is a function of only the time difference..

In other words, the vector $\mathbf{Z}_t$ said to be weakly stationary if the first and second moments of $\mathbf{Z}_t$ do not depend on time t or are constant over time.

The expansion of the ARCH/GARCH univariate model into the m-variate model requires the conditions that random variables $\boldsymbol{\varepsilon}_t$ have m-dimension, zero mean and conditional variant-covariance matrices of $\boldsymbol{\varepsilon}_t$ depend on the elements of the information set of historical data.

Let $\{\boldsymbol{\eta}_t\}$ be a random variable vector i.i.d. sized $(m \times 1)$ with the following characteristics. Suppose $\{\boldsymbol{\varepsilon}_t\}$ is a randomly sized (mx1) random vector,

$$\varepsilon_t = \eta_t \sqrt{H_t}$$

where:

$$E_{t-1}(\varepsilon_t) = 0$$
$$E_{t-1}(\varepsilon_t \varepsilon_t') = H_t$$

and H_t is the positive definite matrix of size $(m\times m)$ and measured to the set of information F_{t-1}, ie σ-field generated by the past information: $\{\varepsilon_{t-1}, \varepsilon_{t-2}, \ldots\}$. Parameterization of H_t as an ARCH (GARCH) multivariate, as a function of the information set of F_{t-1}, then the elements of H_t are dependent on the lag-q of the ε_t and cross-products ε_t squares. Therefore, the elements of the covariance matrix follow an ARMA process vector in the square and cross-products of disturbances (error). Parameterization of H_t as a multivariate ARCH (GARCH) is given in three forms, ie Vech model, BEKK model and Consonant Correlation model.

5. Vech Model

Vech is the vector-half operator (ie half-vector) which is piling the elements of the lower triangle of the m×m matrix into the vector sized (m(m+1)/2) ×1. Vech representation is often called full parameterization. Since the covariance matrix H_t is a symmetric matrix, then vech (H_t) contains elements in H_t singly. Thus, the expansion of GARCH multivariate model (p,q) in vech representation can be written as

$$vech\,(H_t) = W + \sum_{i=1}^{q} A_i\, vech(\varepsilon_{t-i}\varepsilon_{t-i}) + \sum_{j=1}^{p} B_j\, vech\big(H_{t-j}\big)$$

where: W is a vector of sizes (m(m+1)/2) ×1, whereas Ai and Bj are matrices (m(m+1)/2) × (m(m+1)/2). The number of parameters in the general formulation of vech representation is as much as {m(m+1)/2+(p+q)(m(m+1)/2)2}. For example, let m = 2, and p = q = 1, then the number of parameters is 21 pieces. The form of the vech (Ht) model for this example is as follows:

$$\begin{bmatrix} h_{11,t} \\ h_{21,t} \\ h_{22,t} \end{bmatrix} = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} + \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} \varepsilon_{1,t-1}^2 \\ \varepsilon_{1,t-1}\varepsilon_{2,t-1} \\ \varepsilon_{2,t-1}^2 \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix} \begin{bmatrix} h_{11,t-1} \\ h_{21,t-1} \\ h_{22,t-1} \end{bmatrix}$$

In this case, the element to (i, j) in H_t depends on the (i, j) element corresponding in $\varepsilon_t\varepsilon_t'$ and H_{t-1}. In this case assume that the H_t matrix is positive definite.

To reduce the large number of parameters then simplified the vech representation. One way is to select the matrices A and B in diagonal form. This is called the *vech diagonal* model. This simplification reduces the number of parameters where many parameters become (m(m+1)/2)(1+p+q). Suppose that for m=2, and p=q=1, then the diagonal vech model can be written as:

$$\begin{bmatrix} h_{11,t} \\ h_{21,t} \\ h_{22,t} \end{bmatrix} = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} + \begin{bmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{bmatrix} \begin{bmatrix} \varepsilon_{1,t-1}^2 \\ \varepsilon_{1,t-1}\varepsilon_{2,t-1} \\ \varepsilon_{2,t-1}^2 \end{bmatrix} + \begin{bmatrix} b_{11} & 0 & 0 \\ 0 & b_{22} & 0 \\ 0 & 0 & b_{33} \end{bmatrix} \begin{bmatrix} h_{11,t-1} \\ h_{21,t-1} \\ h_{22,t-1} \end{bmatrix} \quad (6)$$

By multiplying the above matrix is obtained

$$h_{11,t} = w_1 + a_{11}\varepsilon_{1,t-1}^2 + b_{11}h_{11,t-1}$$
$$h_{21,t} = w_2 + a_{22}\varepsilon_{1,t-1}\varepsilon_{2,t-1} + b_{22}h_{21,t-1}$$
$$h_{22,t} = w_3 + a_{33}\varepsilon_{2,t-1}^2 + b_{33}h_{22,t-1}.$$

In this case the number of parameters is reduced from 21 simplified to only 9.

6. Results and Discussion

The data in this research is secondary data taken at the Tumumpa Fisheries Harbor Hall Office. The data on fish prices in the center are the prices of Tuna (Panitang), skipjack (Cakalang/Maesang), Cobe (Tongkol/Deho), kite (Layang/Malalugis) and Selar (Tude). The data taken is daily data (excluding

Sundays) for 1 January 2019 to 31 March 2020, namely 391 data. The data graph for the prices of Tuna, Skipjack, Cobe, Kite and Selar fish can be seen in Figure 1.

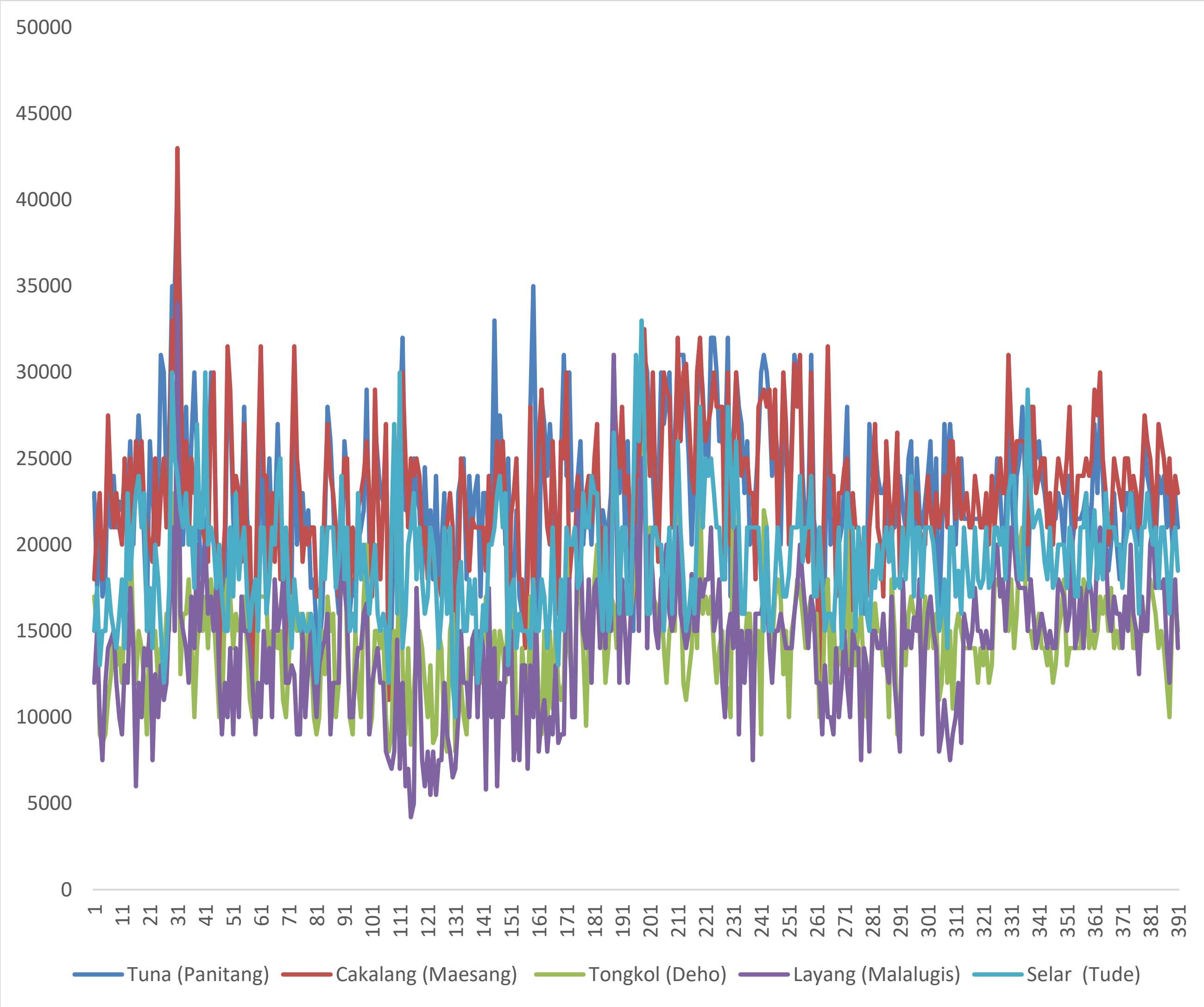

Figure 1. Data graph of daily prices for fresh fish in Manado City.

Furthermore, the fresh fish price data can also be presented in the following bar diagram (histogram) (Figure 2).

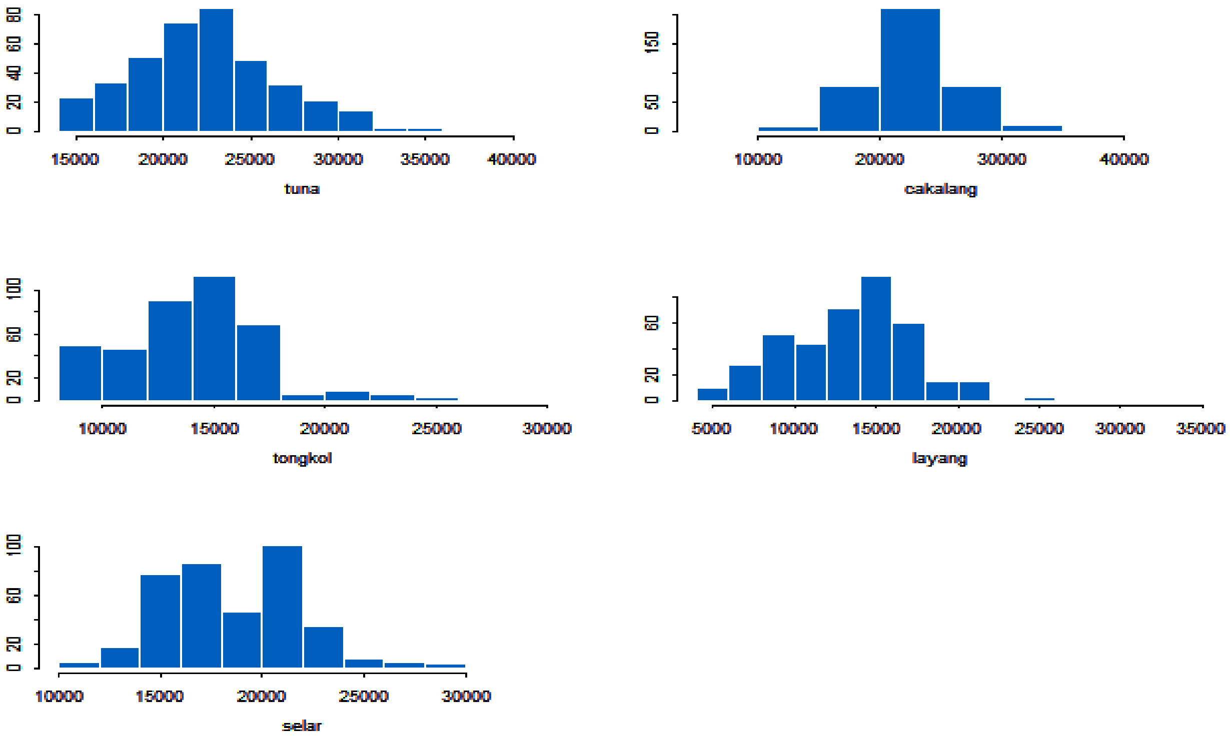

Figure 2. Histogram of Daily Price Data for Fresh Fish in Manado City.

From the ACF and PACF patterns show that the ACF pattern shows that the data is stationary, except for the price of flying fish, where the pattern is slowly decreasing, which shows that the data is not stationary. Furthermore, the PACF cut off pattern is at lag-1, except for the kite fish price data at lag-4, this shows that the mean data model is likely to follow a 1st order autoregression model.

Multivariate GARCH modeling requires data stationarity requirements. To ensure data stationarity, an ADF test is carried out to see whether it has a unit root or not. If there is a unit root then the data is not stationary. The ADF test results can be seen in Table 1 below.

Table 1. ADF Test Results for Fish Price Data in Manado City

No	Fish	p-value	Stationarity
1	Tuna	0,04367	Stationary
2	Skipjack	0.03561	Stationary
3	Cobe	0.008787	Stationary
4	Kite	0.08777	Not stationary
5	Selar	3.249e-05	Stationary

From Table 1 it shows that the p-value for flying fish is 0.08777 > 0.05. Therefore, for the $\alpha = 0.05$ level, the price data for flying fish is not stationary, while other fish have a p-value of less than 0.05, thus fulfilling the assumption of stationarity. This is also in accordance with the ACF pattern of flying fish. Therefore, for further analysis the data used is stationary data, namely Tuna fish, Skipjack tuna, Tuna fish and Selar fish, while Flying fish are not included in the analysis. For flying fish, it is necessary to first carry out stationary, where the analysis is different from data that is already stationary.

Next, a heteroscedasticity test is carried out to see whether the data meets the heteroscedasticity assumption (variance changes every time) or not. The results of the ARCH effect test can be seen in Table 2 below.

Table 2. ARCH Effect Test Results for Fish Price Data in Manado City

No	Fish	p-value
1	Tuna	0,000
2	Skipjack	0.000
3	Cobe	0.000
4	Selar	0.000

From Table 2 it can be seen that the p-value = 0.000 < 0.05 for all types of fish. This means that the fish price data meets the heteroscedasticity assumption so it can be modeled using a multivariate GARCH model. Furthermore, the results of the estimation of the multivariate GARCH model DVEC parameters were obtained by computer program.

From the results of these parameter estimates, a multivariate GARCH model with diagonal vector form (DVEC) is obtained so that volatility (standard deviation of the model) can be calculated. The volatility graph of the DVEC multivariate GARCH model can be seen in Figure 3.

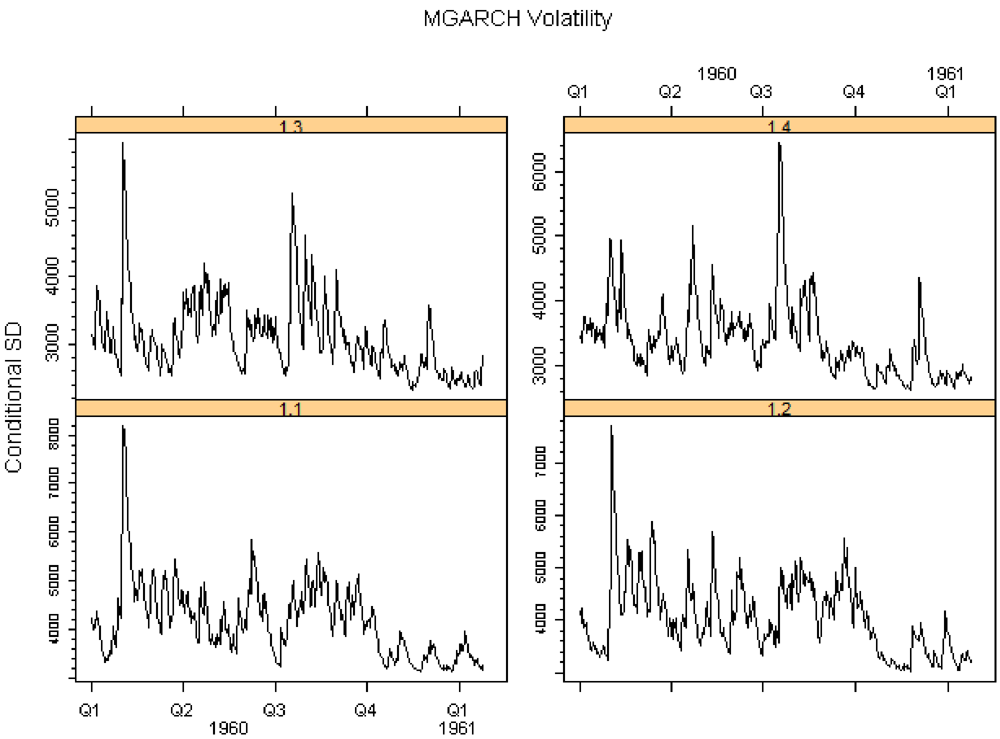

Figure 3. Volatility Graph of the DVEC Multivariate GARCH Model for Fish Price Data in Manado City: 1.1: Volatility of Tuna, 1.2: Volatility of Skipjack, 1.3: Volatility of Cobe, 1.4: Volatility of Selar.

Next, to see the suitability of the model, you can see the QQ-plot graph of standardized residuals. The qq-plot graph of standardized residuals can be seen in Figure 4.

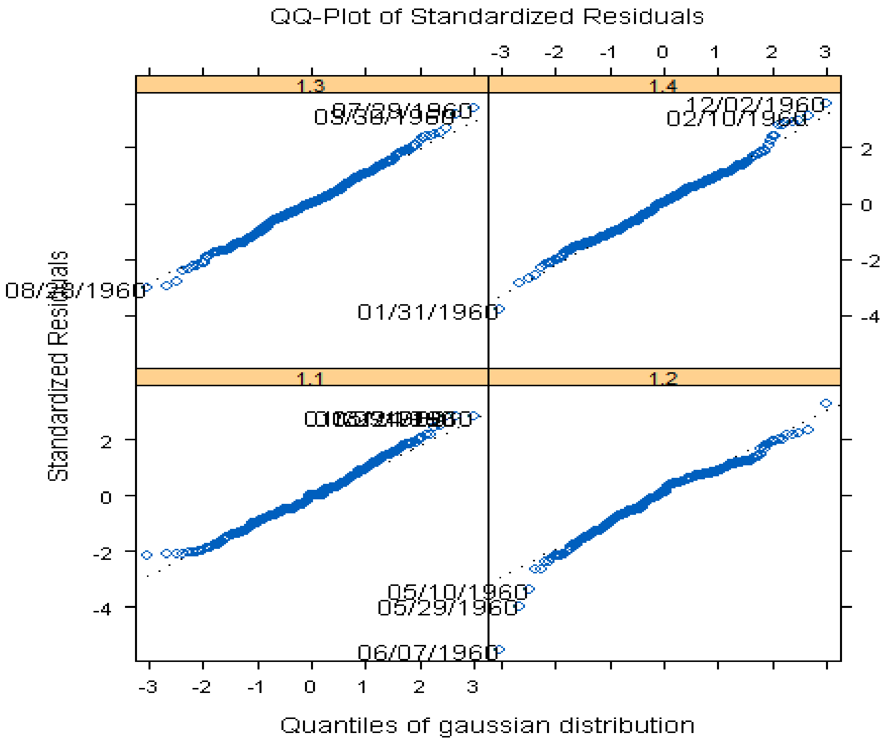

Figure 4. QQ-Plot Graph of Standardized Residuals: 1.1. QQ-Plot of Tuna, 1.2. QQ-Plot of Skipjack, 1.3. QQ-Plot of Cobe, 1.4. QQ-Plot of Selar

7. Conclusion

The volatility of the price of fresh fish in Manado City is influenced by the volatility of the price of fresh fish in the previous period, as well as the volatility of the price of one type of fish is influenced by the volatility of the price of other types of fish.

8. Acknowledgement

The authors would like to thank the Ministry of Research Technology and Higher Education of Indonesia for funding this research through Sam Ratulangi University.

9. References

1. W.W.S. Wei, “Time Series Analysis: Univariate and Multivariate Methods 2nd Edition”, Addison Wesley Publishing Company Inc., New York, 2006.
2. W.H. Greene, “Econometric Analysis”, 4th Ed., Printice Hall, USA, 2002
3. R.F. Engle, “Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation”, Econometrica, 1982, 50, 987 – 1008.
4. L. Bauwens, L. Sebastian, J.V.K. Rombouts, “Multivariate GARCH Models: a Survey”, Journal of Applied Econometrics, 2006, 21: 79 - 109.

5. N. Nainggolan, B.N. Ruchjana., S. Darwis and R. E. Siregar, “GSTAR Models With ARCH Errors and The Simulations”, Proceeding of The Third International Conference on Mathematics and Natural Sciences (ICMNS), ITB. Bandung, 2010, p.1075-1084.
6. E. Zivot, J. Wang, “Modelling Financial Time Series with S-PLUS”, Springer Science+Business Media, LLC. USA, 2002.
7. R. S. Tsay, “Analysis of Financial Time Series”, John Wiley & Sons Inc, Canada, 2002.
8. J. D. Hamilton, “Time Series Analysis”, Princeton University Press, New Jersey, 1994.
9. W. Enders, “Applied Econometric Time Series”, Third ed, John Wiley & Sons. pp. 272–355. New York, 2010.
10. J. W. Lee, M. Kwag, “Vector Autoregressive Models with Multivariate Time Series: Spillovers of International Tourist Arrivals on a Local Economy. Quantitative Tourism Research in Asia. Perspectives on Asian Tourism”, Springer, 2019, pp 273-295.
11. S. Borovkova, H.P. Lopuhaa and B.N. Ruchjana, “Consistency and Asymptotic Normality of Least Squares Estimator in Generalized STAR Models”, Statistica Neerlandica, 2008, 62, nr.4, pp. 482-508.