Use of Vector Autoregressive Model to Analyze the Stock Market Behavior in Indonesia

Agus Suharsono*, Destri Susilaningrum

Department of Statistics, Faculty of Mathematics and Natural Science, ITS Surabaya

ABSTRACT

The purpose of this study was to determine the relationship between the stock index stock market index of the world, namely the DJI index, the FTSE 100 index, the Hang Seng, Nikkei 225, Hang Seng index. The method of analysis used in this study were Multivariate Time Series, in particular the Vector Autoregression (VAR). Besides research, whether there is a case of outliers in the data. To detect outliers performed multivariate analysis using Mahalanobis Distance. This study uses monthly data from the years 2001 - 2011 for each study variable. The results of this study using T² shown Outlier no data, so that the VAR modeling is done and it can be concluded that there is a relationship between the overall stock price index, JKSE, the DJIA index, the FTSE 100 index, the Hang Seng, Nikkei 225, and the STI index.

KEYWORDS: Vector Autoregression, Stock Market, Outlier.

INTRODUCTION

In a stock exchange in the capital market, which is most in demand by investors is stock. Shares are securities that shows ownership of the company, so that shareholders have a right to dividend or other distribution of profit sharing as well as by the company to its shareholders. The capital market is an indicator of economic progress and support the economy of a country. In this decade, the stock market has experienced rapid growth due to the pressure of technological change, liberalization, and globalization. These changes affect the behavior of the stock market and cause long-term balance and encourage the creation of relations between the world's capital markets. Stated capital markets are interconnected if the two separate markets have the same movement and has a correlation between the movement of the index. Capital markets in the region tend to have the same movement and contagion effects is high (Climent and Meneu 2003). During the observation period, years 2001 - 2011, a phenomenon in which the movement of the JKSE is not always the same and has a correlation with the stock market index movement world. Hal is also supported by the differences found in the results of some previous studies.

Composite Index (ICI) is a share index of the most frequently noted investors when investing. This is due to the index lists of all shares listed on the Indonesia Stock Exchange. With the movement of the joint-stock index investors can see whether market conditions moving passionate or lethargic. Many factors affect the behavior of the stock market index coupling medium diantaranya world stock market index, ie the DJIA index, the FTSE 100 index, the Hang Seng, Nikkei 225, Hang Seng index.

Multivariate Time Series, in particular the Vector Autoregression (VAR) is a statistical method that is appropriate for modeling the behavior of the stock market in the Indonesian economy. Because the method can be known pattern of simultaneous relationship between the variables stock index, the DJI index, the FTSE 100 index, the Hang Seng, Nikkei 225, and the STI index, both the influence at the time and in times past. As for knowing each fluctuation of stock price index used approach Mahalanobis distance (Mahalanobis Distance) which exist in the multivariate analysis.

Problem in this study is How does the pattern of relationship between stock index, the DJIA index, the FTSE 100 index, the Hang Seng, Nikkei 225, and STI indices simultaneously by using a VAR and How to model the relationship pattern of ICI, the DJI index, the FTSE 100 index, the Hang Seng, Nikkei 225 and Hang Seng index in case of fluctuations in the data by using VARX (VAR model with no outliers).

MATERIALS AND METHODS

VAR Model

VAR model is actually a composite of several models of autoregressive (AR), where these models form a vector between the variables affect each other. This model describes the relationship between observations on a particular variable at a time with observations on the variable itself at the times previously and also its association with observations on other variables at previous times (Wei, 1990).

There are two important assumptions that must be considered from the time series data to be shaped into the VAR model, namely: (1) stationary, (2) normality and independence of errors. Unit root test (Unit Root Test)

*Corresponding Author: Agus Suharsono, Department of Statistics, Faculty of Mathematics and Natural Science, ITS Surabaya. Email: agus_s@statistika.its.ac.id
is one way to test stationarity. Before performing simultaneous estimation of parameters in the model, it must first be determined what the maximum lag length of the model is. The general form of the VAR model (p) according to (Hamilton, 1994) is:

\[
\begin{align*}
  y_t &= c + \Phi(B)y_t + \varepsilon_t \\
  y_t &= c + (\Phi_1 B + \Phi_2 B^2 + \Phi_3 B^3 + \ldots + \Phi_p B^p) y_t + \varepsilon_t \quad \text{with assumption} \quad \varepsilon_t \sim N(0, \Sigma)
\end{align*}
\]

**Identification of the VAR Model**

VAR model identification is done by pattern or structure of the sample correlation matrix function (MACF) and partial correlation functions of the sample matrix (MPACF) after appropriate transformation to stabilize the variance, and differencing to stabilize the average, if the data is not stationary in the variance and the average (Wei, 1990).

**Best Model Selection Criteria**

There are other ways to determine the appropriate VAR order, in addition to seeing the pattern MPACF, also consider the value of the Akaike Information Criterion (AIC). A model is said the better if its p value of the minimum AIC. The calculation of p values AIC or AIC on VAR (p) is:

\[
AIC_p = \ln \left| \hat{\Sigma}_p \right| + \frac{2pm^2}{n-p}
\]

**Parameter Estimation of VAR**

In estimating the model parameters VAR(p), there are two methods that can be performed, the method of Maximum Likelihood (MLE) and the method of Least Squares (LS) (Hamilton, 1994). Method of Maximum Likelihood (MLE) is used to estimate parameters of a model that is known density function, by maximizing the likelihood function. Density function for the observations is:

\[
f(y) = (2\pi)^{-m/2} |\Sigma|^{-1/2} \exp \left[ -1/2 (y - X\beta)^T \Sigma^{-1} (y - X\beta) \right]
\]

By taking a random sample of n observations \( y_1, y_2, \ldots, y_n \) obtained likelihood function

\[
L(\beta, \Sigma) = \prod_{t=1}^{n} (2\pi)^{-m/2} |\Sigma|^{-1/2} \exp \left[ -1/2 (y - X\beta)^T \Sigma^{-1} (y - X\beta) \right]
\]

and the log likelihood function

\[
\log L(\beta, \Sigma) = -(nm/2) \log (2\pi) + (n/2) \log |\Sigma| - (1/2) \sum_{t=1}^{n} (y - X\beta)^T \Sigma^{-1} (y - X\beta)
\]

So the estimate for \( \beta \)

\[
\hat{\beta} = \left[ \sum_{t=1}^{n} x_t' x_t \right]^{-1} \left[ \sum_{t=1}^{n} x_t' y_t \right]
\]

**Outlier detection**

Outlier is a data set that is considered to have different properties compared to most other data. Mitra (2009) using multivariate control charts to detect outliers in multivariate data sets containing outliers are indicated data. One is the use of Hotelling’s \( T^2 \).

Limit the upper limit of the graph \( T^2 \).

\[
\text{UCL} = \left( \frac{mp - mp - np + p}{mn - m - p - 1} \right) f_{a,p, \left( mn - m - p + 1 \right)}
\]

**VAR models with Additive outliers**

\[
Z_t = D_t \delta + y_t \quad ; \quad t = 1, 2, 3, \ldots, n
\]

With

\[
y_t = \mu + \Phi(B)y_t + \varepsilon_t
\]

\[
D_t = \begin{cases} 
1, & t = T_b \\
0, & t \neq T_b
\end{cases}
\]

Assumed \( \varepsilon_t \sim N(0, \Sigma) \).

**RESULTS AND DISCUSSION**

**Results of descriptive analysis.**

Before conducting the analysis using the VAR method, carried out a description of the qualitative data for the
six variables JKSE, DJIA index, FTSE 100 index, Hang Seng, Nikkei 225 and Hang Seng index. Obtained the following results:

Table 1. Descriptive of Index

<table>
<thead>
<tr>
<th>STOCK</th>
<th>Mean</th>
<th>Std Dev</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Skewness</th>
<th>Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>JKSE</td>
<td>4103.5</td>
<td>154.2</td>
<td>3791.6</td>
<td>4375.2</td>
<td>0.02</td>
<td>-1.13</td>
</tr>
<tr>
<td>DJI</td>
<td>12975.0</td>
<td>307.0</td>
<td>12119.0</td>
<td>13610.0</td>
<td>-0.03</td>
<td>-0.46</td>
</tr>
<tr>
<td>FTSE100</td>
<td>5747.7</td>
<td>159.0</td>
<td>5260.2</td>
<td>5965.6</td>
<td>-0.96</td>
<td>0.37</td>
</tr>
<tr>
<td>HSI</td>
<td>20502.0</td>
<td>1073.0</td>
<td>18502.0</td>
<td>22667.0</td>
<td>0.01</td>
<td>-0.86</td>
</tr>
<tr>
<td>N225</td>
<td>9108.9</td>
<td>490.6</td>
<td>8365.9</td>
<td>10395.2</td>
<td>0.71</td>
<td>-0.47</td>
</tr>
<tr>
<td>STI</td>
<td>2976.7</td>
<td>103.7</td>
<td>2711.0</td>
<td>3191.8</td>
<td>-0.74</td>
<td>0.16</td>
</tr>
</tbody>
</table>

Table 1 shows that the share price index by the average stock price is HSI, with an average daily stock price is 20,502, while the lowest average value owned by STI which has an average value of 2976.7. The size of data dissemination can be seen through the standard deviation. Standard deviation value indicates the level of the stock price data variability in each of these variables. The values of the standard deviation indicates that the level of diversity of daily stock price indices in the six variables. The highest level of diversity produced by HSI stock price index data for 1073.

Graph of daily stock price data movement in 2012 on six variables are shown in the form of time series plots in Figure below

![Figure 1 Plot Time Series JKSE, DJI, FTSE100, HSI, N225, STI](image)

Based on Figure 1 that plots the time series is known to have a high fluctuation. The pattern of data movement in the stock price tends to equal the six variables. Based on plots of time series data has not alleged that the stationary, because the plot shows the data with large fluctuations.

**The results of VAR analysis.**

This conjecture will be analyzed further through PACF plots. ACF and PACF of its results obtained the following results:

Schematic Representation of Partial Cross Correlations

<table>
<thead>
<tr>
<th>Var/Lag</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>y1</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>y2</td>
<td></td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>y3</td>
<td></td>
<td></td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>y4</td>
<td></td>
<td></td>
<td></td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>y5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>y6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

+ is > 2*std error, - is < -2*std error, . is between

**Figure 2. Plot PACF (Partial Autocorrelation Function)**

From the table above shows that the model has a lag of 1, so that the model is a VAR model (1). Furthermore, to see the trend of the six variables can be viewed through the area the following graph:
While the results of differencing the time series plot of the time is as follows:

From the picture above shows no data in the six variables that contain outliers, since all values below the lower limit of T² Hotelling (UCL).

While the results of differencing the time series plot of the time is as follows:

From the picture above shows that all six variables JKSE_1, DJI_1, FTSE100_1, HSI_1, N225_1 and STI_1 already shows a stationary pattern.
Processing results for the VAR model with the data that has been stationary is as follows:

Figure 3. Graph Plot area index STI, HSI, FTSE 100, HSI, N 225, and STI

Graph of the area can be seen that each of the stock price index variables have the same trend. While the results of the analysis using the VAR method obtained the following results:

\[
y_{1t} = 0.92186y_{1t-1} + 0.02813y_{2t-1}
\]

\[
y_{2t} = 0.97652y_{2t-1}
\]

\[
y_{3t} = 0.03123y_{2t-1} + 0.93019y_{3t-1}
\]

\[
y_{4t} = 0.33559y_{3t-1} + 0.98068y_{4t-1} - 0.08997y_{5t-1}
\]

\[
y_{5t} = -0.17450y_{4t-1} + 0.02994y_{5t-1}
\]

\[
y_{6t} = 0.03321y_{3t-1} + 0.93830y_{6t-1}
\]

From the equation above shows that not all variables affect each other. For example, \( y_{2t} \) only influenced by \( y_{2t-1} \) in the last period. This shows that the model is not good, since many parameters that are not included in the model. This is possible because the data may exist outliers or the data are not stationary in the average. By using T² Hotelling to see outliers, obtained the following results

Figure 4. Outliers from the T² Hotelling

From the picture above shows no data in the six variables that contain outliers, since all values below the lower limit of T² Hotelling (UCL).

While the results of differencing the time series plot of the time is as follows:

Figure 5. Plot Time Series JKSE_1, DJI_1, FTSE100_1, HSI_1, N225_1, STI_1

From the picture above shows that all six variables JKSE_1, DJI_1, FTSE100_1, HSI_1, N225_1 and STI_1 already shows a stationary pattern.
Processing results for the VAR model with the data that has been stationary is as follows:
From the above results it appears that all six variables affect each other stock price index significantly. Model VARAO or VARX planned to be compared with the VAR model is generally not done because of lack of data outliers, so that the model VARX not obtained.
CONCLUSIONS

From the discussion, it can be concluded
1. It turns out that there is a significant relationship between the six variables that world stock price index STI, HSI, FTSE100, HIS, N225 and STI.
2. No data outliers for the six variables that world stock price index STI, HSI, FTSE100, HIS, N225 and STI.
3. VAR models can not be compared with the model VARX or VARAO because the data do not contain outliers so VARX or VARAO models can not be obtained.

Acknowledgment
The authors declare that they have no conflicts of interest in this research.

REFERENCES