Application of Vector Error Correction Model (VECM) and Impulse Response Function for Daily Stock Prices

S. Winarno¹, M. Usman¹, Warsono.¹, D. Kurniasari¹, Widiarti¹

¹1) Departement of Mathematics, Faculty of Science and Mathematics, Lampung University, Indonesia

email: salsanursabilanw@gmail.com

Abstract. Vector Error Correction Model is a cointegrated VAR model. This idea of Vector Error Correction Model (VECM), which consists of a VAR model of the order p - 1 on the differences of the variables, and an error-correction term derived from the known (estimated) cointegrating relationship. Intuitively, and using the stock market example, a VECM model establishes a short-term relationship between the stock prices, while correcting with the deviation from the long-term comovement of prices. An Impulse Response Function traces the incremental effect of a 1 unit (or one standard deviation) shock in one of the variables on the future values of the other endogenous variables. Impulse Response Functions trace the incremental effect of the marketing action reflected in the shock. The data used in this analysis are 4 (four) daily plantation stock prices in Indonesia with time period of January to July in three years which are 2018, 2019, and 2020. The objective of this study is to determine the relationship among 4 (four) stocks prices with VECM and to know the behaviour of each stocks prices with Impulse Response.

Keyword: Impulse Response Function, VAR, VECM, Granger Causality

1. Introduction
A time series data is a series of data listed in time order. The multivariate time series data is a time series data that has more than one time-dependent variable. In the stock market, there are various multivariate time series data. In this study, plantation stock prices analyzed with Vector Error Correction Model. According to Medvegyef (2015) Vector Error Correction Model is a cointegrated VAR model. The standard VAR (Vector Autoregression) models can only be estimated when the variables are stationary [6]. However not all data is stationary, with that reason VECM model is made. In this study the plantation stock prices which are going to be analyzed are daily stock prices of PT. Provident Agro Tbk, PT. PP London Sumatera Indonesia Tbk, PT. Sampoerna Agro Tbk, and PT. Sawit Sumbermas Sarana. Due to COVID-19 there is some shock in stock market, in this study the stock prices are going to be analyzed in three times period, with time period of January to July in three years which are 2018, 2019, and 2020 to see if the stock prices is affected by the shock. The data used in this paper cited from Yahoo Finance (2020) [9].
2. Statistical Model

2.1. Cointegration

According to Engle and Granger (1987), a cointegration test is used to establish if there is a correlation between several time series in the long term [3]. According to Lütkepohl (2005), The Johansen test can be seen as a multivariate generalization of the augmented Dickey-Fuller test [5]. The generalization is the examination of linear combinations of variables for unit roots. If there are three variable search with unit roots, there are at most two cointegrating vectors.

More generally, if there are n variables which all have unit roots, there are at most n - 1 cointegrating vectors. The Johansen test provides estimates of all cointegrating vectors. The Johansen tests are based on eigenvalues of transformations of the data and represent linear combinations of the data that have maximum correlation (canonical correlations). To repeat, the eigenvalues used in Johansen’s test are not eigenvalues of the matrix Π directly, although the eigenvalues in the test also can be used to determine the rank of Π and have tractable distributions.

Suppose that eigenvalues for the Johansen test have been computed. Order the n eigenvalues by size so λ₁ ≥ λ₂ ≥ ... ≥ λₙ and recall that λ_i ≥ 0 for all i. If λ₁ = 0, then the rank of Π is zero and there are no cointegrating vectors. If λ₁ ≠ 0, then the rank of Π is greater than or equal to one and there is at least one cointegrating vector. The Johansen tests are likelihood-ratio tests. There are two tests: 1. the maximum eigenvalue test, and 2. the trace test. For both test statistics, the initial Johansen test is a test of the null hypothesis of no cointegration against the alternative of cointegration. The tests differ in terms of the alternative hypothesis

2.1.1 Maximum Eigenvalue Test

The maximum eigenvalue test examines whether the largest eigenvalue is zero relative to the alternative that the next largest eigenvalue is zero. The first test is a test whether the rank of the matrix Π is zero. The null hypothesis is that rank(Π) = 0 and the alternative hypothesis is that rank(Π) ≥ 1. For further tests, the null hypothesis is that rank(Π) = 1, 2... and the alternative hypothesis is that rank(Π) ≥ 3.

\[ \lambda_{\text{max}}(r, r + 1) = -T \ln(1 - \hat{\lambda}_i) \]  

(2.1)

2.1.2 Trace Test

The trace test is a test whether the rank of the matrix Π is r₀. The null hypothesis is that rank(Π) = r₀. The alternative hypothesis is that r₀ < rank(Π) ≤ n, where n is the maximum number of possible cointegrating vectors. For the succeeding test if this null hypothesis is rejected, the next null hypothesis is that rank(Π) = r₀ + 1 and the alternative hypothesis is that r₀ + 1 < rank(Π) ≤ n.

\[ Tr(r) = -T \sum_{i=r+1}^{n} \ln(1 - \hat{\lambda}_i) \]  

(2.2)

\( \hat{\lambda}_i \): The estimation of Eigen values

T : Number of observations,

k : Number of endogenous variables.

2.2. Vector Autoregression (VAR)

Vector Autoregression (VAR) models were introduced by the macroeconometrician Christopher Sims (1980) to model the joint dynamics and causal relations among a set of macroeconomic variables. According to R S Tsay (2010) VAR(p) model is a multivariate version of Yule-Walker equation of a univariate AR(p) model [8]. The VAR(p) model can be written in the form

\[ x_t = \phi^*(x_{t-1}) + b_t \]  

(2.3)

where:

\( x_t \): the element vector of at time

\( \phi^* \): Matrix order kp x kp which the elements are the coefficient of the vector \( x_{t-1} \)

\( b_t \): Random vector of shock.
2.3. Vector Error Correction Model (VECM)

The standard VAR model discussed earlier can only be estimated when the variables are stationary. The conventional way to remove unit root model is to first differentiate the series. However, in the case of cointegrated series, this would lead to over differencing and losing information conveyed by the long-term comovement of variable levels. For that reason, the cointegrated VAR model is build. According to Medvegyev (2015) this idea of Vector Error Correction Model (VECM), which consists of a VAR model of the order p - 1 on the differences of the variables, and an error-correction term deriving from the known (estimated) cointegrating relationship [6]. Intuitively, and using the stock market example, a VECM model establishes a short-term relationship between the stock prices, while correcting with the deviation from the long-term comovement of prices. An appropriate VECM model can be formulated as follows

\[
\Delta y_t = a\beta'y_{t-1} + \Sigma_{i=1}^{p-1} \Gamma_i \Delta y_{t-i} + \epsilon_t
\]  

(2.4)

where:

\( \Delta \) : Operator differencing, where \( \Delta y_t = y_t - y_{t-1} \)

\( y_{t-i} \) : Vector variable endogenous with the 1-st lag.

\( \epsilon_t \) : Vector residual.

\( \Gamma_i \) : Matrix with order k x k of coefficient Endogenous of the i-th variable.

\( \alpha \) : Vector adjustment, matrix with order (k x r)

\( \beta \) : Vector cointegration (long-run parameter) matrix (k x r)

2.4 The Length of The Optimal Lag

Minimum values of the criteria is used to determine the length of the lag to be chosen. Some commonly used criteria are as follows:

a. Akaike Information Criterion (AIC)

\[
AIC(p) = \ln(|\hat{\Sigma} u(p)|) + (k + pk^2) \frac{2}{T}
\]

(2.5)

b. Final Prediction Error (FPE)

\[
FPE(p) = \left[ \frac{T + kp + 1}{T - kp - 1} \right]^k |\hat{\Sigma} u(p)|
\]

(2.6)

c. Bayesian Criterion of Gideon Schwartz (SBC)

\[
SC(p) = |\hat{\Sigma} u(p)| + (k + pk^2) \frac{2\ln(T)}{T}
\]

(2.7)

d. Hannan-Quinn Criterion (HQC)

\[
HQ = \ln(|\hat{\Sigma} u(p)|) + (k + pk^2) \frac{\ln(T)}{T}
\]

(2.8)

Where \( \hat{u} \) are denotes the residuals estimation from the model VAR(p), k number of dependent variables, T is number of observations and p is the length of model (Kirchgassner and Wolters, 2007).

2.5 Model Stability

According to Lütkepohl (2005) a condition for stability for VAR(p) requires that all the eigenvalues of A (The AR Matrix of the comparison from \( Y_t \)) are smaller than one in modulus or all the roots larger than one [5]. Consider the VAR(p) model

\[
Y_t = c + \varnothing_1 Y_{t-1} + \cdots + \varnothing_p Y_{t-p} + \epsilon_t
\]

(2.9)

Substituting t=1 we obtain

\[
Y_1 = c + \varnothing_1 Y_0 + \epsilon_1
\]

\[
Y_2 = c + \varnothing_1 Y_1 + \epsilon_2
\]

\[
= c + \varnothing_1(c + \varnothing_1 Y_0 + \epsilon_1) + \epsilon_2
\]

\[
= (I_k \varnothing_1)c + \varnothing_1^2 Y_0 + \varnothing_1 \epsilon_1 + \epsilon_2
\]

\[
Y_t = \left( I_k + \varnothing_1 + \cdots + \varnothing_1^{t-1} \right)c + \varnothing_1^t Y_0 + \sum_{i=0}^{t-1} \varnothing_1^i \epsilon_{t-i}
\]

(2.10)

Therefore, if the eigenvalues are smaller than one in modulus then \( Y_t \) has the following representation

\[
Y_t = (I - \varnothing)^{-1} + \sum_{i=0}^{j} \varnothing_1^i \epsilon_{t-i}
\]

(2.11)
Note that the eigenvalues of $\emptyset$ satisfy $\det(I_k + \emptyset) \neq 0$ for $|z| \leq 1$ therefore VAR(p) is called stable if
\[
\det(I_k + \emptyset) = \det(I_k - \emptyset_{1z} - \emptyset_{p}z^p) \quad \text{for } |z| \leq 1 \quad (2.12)
\]

2.6 Impulse Response Function
According to J A Petersen and V Kumar (2012), impulse response function traces the incremental effect of a 1 unit (or one standard deviation) shock in one of the variables on the future values of the other endogenous variables [7]. Consider the VAR(p) model
\[
Y_t = A_0 + A_1X_{t-1} + e_t \quad (2.13)
\]
Where $A_0 = B^{-1} \Gamma_0$, $A_1 = B^{-1} \Gamma_1$ and $e_t = B^{-1} e_t$

Vector error can be written as:
\[
\begin{bmatrix}
e_1t \\
e_2t \\
\vdots \\
e_{3t}
\end{bmatrix} = \frac{1}{\det(A_1)} \sum_{i=0}^{\infty} [a_{11} \quad a_{12} \quad a_{13}]^i \times adj(A_1) \times \begin{bmatrix}
e_{1t-i} \\
e_{2t-i} \\
\vdots \\
e_{3t-i}
\end{bmatrix} \quad (2.14)
\]
det(A_1) is a determinar value of $A_1$ and adj(A_1) is adjoint matrix of $A_1$, therefore:
\[
\begin{bmatrix}
x_t \\
y_t \\
z_t
\end{bmatrix} = \begin{bmatrix}
\hat{x} \\
\hat{y} \\
\hat{z}
\end{bmatrix} + \frac{1}{\det(A_1)} \sum_{i=0}^{\infty} [a_{11} \quad a_{12} \quad a_{13}]^i \times adj(A_1) \times \begin{bmatrix}
e_{1t-i} \\
e_{2t-i} \\
\vdots \\
e_{3t-i}
\end{bmatrix} \quad (2.15)
\]

With $\phi$ matrix :
\[
\begin{bmatrix}
x_t \\
y_t \\
z_t
\end{bmatrix} = \begin{bmatrix}
\hat{x} \\
\hat{y} \\
\hat{z}
\end{bmatrix} + \sum_{i=1}^{\infty} \begin{bmatrix}
\phi_{11}(i) & \phi_{12}(i) & \phi_{13}(i) \\
\phi_{21}(i) & \phi_{22}(i) & \phi_{23}(i) \\
\phi_{31}(i) & \phi_{32}(i) & \phi_{33}(i)
\end{bmatrix} \times \begin{bmatrix}
e_{1t-i} \\
e_{2t-i} \\
\vdots \\
e_{3t-i}
\end{bmatrix} \quad (2.16)
\]

With elemen $\phi_{jk}(i)$:
\[
\phi_{j} = \frac{1}{\det(A_1)} \sum_{i=0}^{\infty} [a_{11} \quad a_{12} \quad a_{13}]^i \times adj(A_1) \quad (2.17)
\]

That can also be written as:
\[
Z_t = \mu + \sum_{i=0}^{\infty} \phi_{j} \varepsilon_{t-i} \quad (2.18)
\]
The coefficient $\phi_{jk}(i)$ is called Impulse Response Function (IRF). $\phi_{jk}(i)$ plot is the best way to visualize the response toward the shocks [2].

2.7 Granger Causality
Granger causality is a method that attempts to determine whether one series is likely to influence a change in the other. This is done by taking different lags of one series and using this to models the change in the second series [1].

\begin{align*}
H_0 &\implies a_{2,i} = 0 \quad \text{for each } i = 1, 2, ..., p \quad (y_{2t} \text{ not } \text{"Granger-Cause" } y_{1t}) \\
H_1 &\implies a_{2,i} \neq 0 \quad \text{for at least one } i = 1, 2, ..., p \quad (y_{2t} \text{ "Granger-Cause" } y_{1t})
\end{align*}

\[
F - \text{Test} = \frac{\text{RSS}_2 - \text{RSS}_1}{\text{RSS}_1/(T-2p-1)} \quad (2.19)
\]

3. Data Analysis

3.1 Stationary
Hypothesis:
\begin{align*}
H_0 &\implies \text{Data is nonstationary} \\
H_1 &\implies \text{Data stationary}
\end{align*}
### Table 1. ADF Test for 2018 Data

**Augmented Dickey-Fuller Unit Root Tests**

| Stock | Type   | Lags | Rho   | Pr < Rho | Tau  | Pr < Tau | F    | Pr > F |
|-------|--------|------|-------|----------|------|----------|------|--------|
| PALM  | Zero Mean | 3    | -0.1603 | 0.6453  | -0.94 | 0.3059   |      |        |
|       | Single Mean | 3    | -10.5313 | 0.1126 | -2.36  | 0.1553  | 3.15 | 0.2672 |
|       | Trend    | 3    | -52.7932 | 0.0005 | -4.38  | 0.0031  | 9.66 | 0.0010 |
| LSIP  | Zero Mean | 3    | -0.3119 | 0.6109  | -1.05  | 0.2652   |      |        |
|       | Single Mean | 3    | -2.2689 | 0.7440  | -0.90  | 0.7875  | 0.85 | 0.8551 |
|       | Trend    | 3    | -16.3044 | 0.1314 | -2.80  | 0.1992  | 4.03 | 0.3730 |
| SGRO  | Zero Mean | 3    | -0.0834 | 0.6627  | -1.60  | 0.1039   |      |        |
|       | Single Mean | 3    | -1.6539 | 0.8171  | -1.15  | 0.6963  | 1.87 | 0.5941 |
|       | Trend    | 3    | -6.2856 | 0.7152  | -1.43  | 0.8468  | 1.27 | 0.9232 |
| SSMS  | Zero Mean | 3    | -0.1735 | 0.6423  | -0.88  | 0.3348   |      |        |
|       | Single Mean | 3    | -2.8502 | 0.6720  | -1.12  | 0.7065  | 0.94 | 0.8306 |
|       | Trend    | 3    | -8.0612 | 0.5691  | -1.94  | 0.6303  | 1.89 | 0.8003 |

### Table 2. ADF Test for 2019 Data

**Augmented Dickey-Fuller Unit Root Tests**

| Stock | Type   | Lags | Rho   | Pr < Rho | Tau  | Pr < Tau | F    | Pr > F |
|-------|--------|------|-------|----------|------|----------|------|--------|
| PALM  | Zero Mean | 3    | -0.1218 | 0.6540  | -0.58  | 0.4660   |      |        |
|       | Single Mean | 3    | -9.3751 | 0.1502  | -1.77  | 0.3923  | 1.70 | 0.6379 |
|       | Trend    | 3    | -19.9705 | 0.0608 | -2.88  | 0.1736  | 4.50 | 0.2785 |
| LSIP  | Zero Mean | 3    | -0.2352 | 0.6283  | -0.81  | 0.3659   |      |        |
|       | Single Mean | 3    | -3.9810 | 0.5370  | -1.31  | 0.6249  | 1.09 | 0.7931 |
|       | Trend    | 3    | -12.3674 | 0.2811 | -2.39  | 0.3844  | 2.87 | 0.6043 |
| SGRO  | Zero Mean | 3    | -0.0556 | 0.6690  | -0.36  | 0.5516   |      |        |
|       | Single Mean | 3    | -11.5680 | 0.0867 | -2.21  | 0.2017  | 2.50 | 0.4345 |
|       | Trend    | 3    | -18.5452 | 0.0825 | -2.84  | 0.1857  | 4.07 | 0.3642 |
| SSMS  | Zero Mean | 3    | -0.2006 | 0.6361  | -1.49  | 0.1264   |      |        |
|       | Single Mean | 3    | -2.9328 | 0.6618  | -1.60  | 0.4793  | 2.24 | 0.4991 |
|       | Trend    | 3    | -9.3859 | 0.4671  | -1.82  | 0.6913  | 2.14 | 0.7504 |
Table 3. ADF Test for 2020 Data

Augmented Dickey-Fuller Unit Root Tests

| Stock  | Type         | Lags | Rho  | Pr < Rho | Tau  | Pr < Tau | F     | Pr > F |
|--------|--------------|------|------|----------|------|----------|-------|--------|
| PALM   | Zero Mean    | 3    | 0.1504 | 0.7167  | 0.40 | 0.7982   |       |        |
|        | Single Mean  | 3    | -19.2267 | 0.0116  | -2.12 | 0.2368  | 2.38  | 0.4641 |
|        | Trend        | 3    | -17.5139 | 0.1017  | -2.01 | 0.5887  | 3.01  | 0.5757 |
| LSIP   | Zero Mean    | 3    | -0.6709 | 0.5329  | -1.16 | 0.2236  |       |        |
|        | Single Mean  | 3    | -5.1382 | 0.4163  | -2.09 | 0.2507  | 2.42  | 0.4543 |
|        | Trend        | 3    | -3.7364 | 0.8992  | -1.25 | 0.8952  | 2.39  | 0.7009 |
| SGRO   | Zero Mean    | 3    | -0.2830 | 0.6173  | -1.45 | 0.1367  |       |        |
|        | Single Mean  | 3    | 0.7804  | 0.9840  | 0.30  | 0.9778  | 1.13  | 0.7824 |
|        | Trend        | 3    | -1.9714 | 0.9700  | -0.62 | 0.9758  | 1.20  | 0.9370 |
| SSMS   | Zero Mean    | 3    | -0.0094 | 0.6794  | -0.04 | 0.6676  |       |        |
|        | Single Mean  | 3    | -4.6704 | 0.4623  | -1.49 | 0.5359  | 1.11  | 0.7870 |
|        | Trend        | 3    | -4.5276 | 0.8501  | -1.42 | 0.8507  | 1.13  | 0.9499 |

Tables 1-3 show that data from 2018, 2019, and 2020 do not pass through the significance $\alpha = 0.05$, this means that the p-values are greater than 0.05. Thus, it is not sufficient evidence to reject $H_0$, so we can conclude that the data from 2018, 2019, and 2020 are nonstationary. Next, in order to make the data are stationary, we need to perform differencing on data. After the first differencing, then the stationary data can be rechecked through the tables below.

Table 4. ADF Test for 2018 Data after the First Differencing

Augmented Dickey-Fuller Unit Root Tests

| Stock  | Type         | Lags | Rho  | Pr < Rho | Tau  | Pr < Tau | F     | Pr > F |
|--------|--------------|------|------|----------|------|----------|-------|--------|
| PALM   | Zero Mean    | 3    | -1934.64 | 0.0001  | -8.07 | <.0001  |       |        |
|        | Single Mean  | 3    | -2857.50 | 0.0001  | -8.13 | <.0001  | 33.02 | 0.0010 |
|        | Trend        | 3    | -2960.23 | 0.0001  | -8.10 | <.0001  | 32.84 | 0.0010 |
| LSIP   | Zero Mean    | 3    | -89.2492 | <.0001  | -5.22 | <.0001  |       |        |
|        | Single Mean  | 3    | -95.7418 | 0.0012  | -5.32 | <.0001  | 14.13 | 0.0010 |
|        | Trend        | 3    | -97.2380 | 0.0005  | -5.31 | 0.0001  | 14.12 | 0.0010 |
| SGRO   | Zero Mean    | 3    | -465.289 | 0.0001  | -7.27 | <.0001  |       |        |
|        | Single Mean  | 3    | -690.552 | 0.0001  | -7.50 | <.0001  | 28.15 | 0.0010 |
|        | Trend        | 3    | -726.613 | 0.0001  | -7.53 | <.0001  | 28.36 | 0.0010 |
| SSMS   | Zero Mean    | 3    | -94.9362 | <.0001  | -5.35 | <.0001  |       |        |
|        | Single Mean  | 3    | -98.3650 | 0.0012  | -5.38 | <.0001  | 14.47 | 0.0010 |
|        | Trend        | 3    | -98.2752 | 0.0005  | -5.36 | 0.0001  | 14.38 | 0.0010 |
Table 5. ADF Test for 2019 Data after the First Differencing

| Stock  | Type       | Lags | Rho       | Pr < Rho   | Tau       | Pr < Tau | F        | Pr > F   |
|--------|------------|------|-----------|------------|-----------|----------|----------|----------|
| PALM   | Zero Mean  | 3    | -248.008  | 0.0001     | -6.70     | <.0001   |          |          |
|        | Single Mean| 3    | -252.229  | 0.0001     | -6.70     | <.0001   | 22.42    | 0.0010   |
|        | Trend      | 3    | -268.015  | 0.0001     | -6.74     | <.0001   | 22.70    | 0.0010   |
| LSIP   | Zero Mean  | 3    | -129.900  | 0.0001     | -8.03     | <.0001   |          |          |
|        | Single Mean| 3    | -131.033  | 0.0001     | -8.04     | <.0001   | 32.34    | 0.0010   |
|        | Trend      | 3    | -131.302  | 0.0001     | -8.01     | <.0001   | 32.14    | 0.0010   |
| SGRO   | Zero Mean  | 3    | -306.768  | 0.0001     | -12.29    | <.0001   |          |          |
|        | Single Mean| 3    | -307.189  | 0.0001     | -12.25    | <.0001   | 75.06    | 0.0010   |
|        | Trend      | 3    | -307.392  | 0.0001     | -12.21    | <.0001   | 74.57    | 0.0010   |
| SSMS   | Zero Mean  | 3    | -234.265  | 0.0001     | -6.57     | <.0001   |          |          |
|        | Single Mean| 3    | -277.215  | 0.0001     | -6.73     | <.0001   | 22.68    | 0.0010   |
|        | Trend      | 3    | -301.812  | 0.0001     | -6.82     | <.0001   | 23.28    | 0.0010   |

Table 6. ADF Test for 2020 Data after the First Differencing

| Stock  | Type       | Lags | Rho       | Pr < Rho   | Tau       | Pr < Tau | F        | Pr > F   |
|--------|------------|------|-----------|------------|-----------|----------|----------|----------|
| PALM   | Zero Mean  | 3    | -189.928  | 0.0001     | -6.06     | <.0001   |          |          |
|        | Single Mean| 3    | -191.414  | 0.0001     | -6.06     | <.0001   | 18.42    | 0.0010   |
|        | Trend      | 3    | -215.824  | 0.0001     | -6.21     | <.0001   | 19.27    | 0.0010   |
| LSIP   | Zero Mean  | 3    | -126.595  | 0.0001     | -5.70     | <.0001   |          |          |
|        | Single Mean| 3    | -131.959  | 0.0001     | -5.73     | <.0001   | 16.43    | 0.0010   |
|        | Trend      | 3    | -174.924  | 0.0001     | -6.07     | <.0001   | 18.44    | 0.0010   |
| SGRO   | Zero Mean  | 3    | -227.389  | 0.0001     | -6.47     | <.0001   |          |          |
|        | Single Mean| 3    | -275.118  | 0.0001     | -6.67     | <.0001   | 22.22    | 0.0010   |
|        | Trend      | 3    | -359.712  | 0.0001     | -6.95     | <.0001   | 24.20    | 0.0010   |
| SSMS   | Zero Mean  | 3    | -119.655  | 0.0001     | -5.58     | <.0001   |          |          |
|        | Single Mean| 3    | -119.722  | 0.0001     | -5.56     | <.0001   | 15.45    | 0.0010   |
|        | Trend      | 3    | -120.569  | 0.0001     | -5.56     | <.0001   | 15.51    | 0.0010   |

Tables 4-6 show that data from 2018, 2019, and 2020 pass through the significance α = 0.05, this means that the p-values are not greater than 0.05. Thus, it is sufficient evidence to reject Ho, so we can conclude that the data from 2018, 2019, and 2020 are stationary.
3.2 Cointegration Test

Hypothesis:
- \( H_0 \): Data is not cointegrated
- \( H_1 \): Data is cointegrated

Table 7. Johansen Cointegration

| Year | H0: Rank=r | H1: Rank>r | Eigenvalue | Trace | Pr > Trace |
|------|-------------|-------------|------------|-------|------------|
| 2018 | 0           | 0           | 0.6470     | 540.7853 | <.0001     |
|      | 1           | 1           | 0.6203     | 384.5839 | <.0001     |
|      | 2           | 2           | 0.5783     | 239.3111 | <.0001     |
|      | 3           | 3           | 0.5190     | 109.7858 | <.0001     |
| 2019 | 0           | 0           | 0.7279     | 618.2802 | <.0001     |
|      | 1           | 1           | 0.6882     | 423.0466 | <.0001     |
|      | 2           | 2           | 0.6124     | 248.2388 | <.0001     |
|      | 3           | 3           | 0.5070     | 106.0744 | <.0001     |
| 2020 | 0           | 0           | 0.6637     | 459.2331 | <.0001     |
|      | 1           | 1           | 0.5688     | 305.5922 | <.0001     |
|      | 2           | 2           | 0.5434     | 186.9920 | <.0001     |
|      | 3           | 3           | 0.4185     | 76.4410  | <.0001     |

\( H_0 \) is not rejected if the value \( \lambda_{\text{trace}} \) < critical values. Table 7 \( \lambda_{\text{trace}} \) < Critical values in 2018, 2019, and 2020 data. Thus, we can conclude that the all variables in each year have cointegration.

3.3 Model Estimation

The first step to be taken is the VECM model to determine the optimum lag by comparing every lag to the criteria used. In the table below the minimum criteria for each information criterion value are given with star sign (*). Here are information criterion value for each year to determine the optimum lag of VECM(p) model.

Table 8. Information Criterion for VECM(p) Model

| Year | Lag | AIC       | SBC       | HQC       | FPEC     | AICC     |
|------|-----|-----------|-----------|-----------|----------|----------|
| 2018 | 1   | 22.13694* | 22.458074*| 22.267407*| 4.11124E9*| 22.142785*|
|      | 2   | 22.223397 | 22.86854  | 22.485508 | 4.48413E9| 22.247768 |
|      | 3   | 22.258744 | 23.230813 | 22.653693 | 4.65018E9| 22.315977 |
|      | 4   | 22.42171  | 23.723667 | 22.950709 | 5.4844E9 | 22.528061 |
|      | 5   | 22.501565 | 24.136418 | 23.165844 | 5.96097E9| 22.675516 |
| 2019 | 1   | 24.295844*| 24.616979*| 24.426311*| 3.561E10 | 24.301689*|
|      | 2   | 24.382818 | 25.027961 | 24.644929 | 3.886E10 | 24.407189 |
|      | 3   | 24.488468 | 25.460537 | 24.883418 | 4.3234E10| 24.545702 |
|      | 4   | 24.534311 | 25.836268 | 25.06331  | 4.5354E10| 24.640662 |
|      | 5   | 24.675438 | 26.310291 | 25.339716 | 5.241E10 | 24.849389 |
| 2020 | 1   | 25.268928*| 25.603539*| 25.404902*| 9.4229E10*| 25.275554*|
|      | 2   | 25.277927 | 25.950303 | 25.55116  | 9.5122E10| 25.305633 |
|      | 3   | 25.38567  | 26.399013 | 25.797466 | 1.0607E11| 25.450928 |
|      | 4   | 25.439323 | 26.79689  | 25.991005 | 1.122E11 | 25.560967 |
|      | 5   | 25.555422 | 27.260521 | 26.248333 | 1.2654E11| 25.75506   |
The table above indicates that the optimal lag for each year VECM(p) model is at lag 1, therefore the VECM(p) model used is VECM(1) for each year.

For 2018 data, the model VECM(1) is:

\[
\begin{bmatrix}
Y_1 \\
Y_2 \\
Y_3 \\
Y_4
\end{bmatrix} =
\begin{bmatrix}
-1.27844 & 0.00467 & 0.03808 & -0.02819 \\
-0.14635 & -1.00449 & 0.01351 & -0.09519 \\
-0.06397 & -0.05995 & -1.24355 & 0.06276 \\
-0.24856 & 0.17273 & -0.11893 & -1.12800
\end{bmatrix}
\begin{bmatrix}
Y_{t-1} \\
Y_{t-2} \\
Y_{t-3} \\
Y_{t-4}
\end{bmatrix}
+ 
\begin{bmatrix}
\varepsilon_1 \\
\varepsilon_2 \\
\varepsilon_3 \\
\varepsilon_4
\end{bmatrix}
\]

For 2019 data, the model VECM(1) is:

\[
\begin{bmatrix}
Y_1 \\
Y_2 \\
Y_3 \\
Y_4
\end{bmatrix} =
\begin{bmatrix}
-1.43099 & 0.01240 & 0.00789 & -0.05430 \\
0.24740 & -1.07110 & -0.02757 & -0.15429 \\
-1.04578 & 0.23566 & -1.25534 & -0.02919 \\
0.22213 & -0.06016 & -0.06541 & -1.21535
\end{bmatrix}
\begin{bmatrix}
Y_{t-1} \\
Y_{t-2} \\
Y_{t-3} \\
Y_{t-4}
\end{bmatrix}
+ 
\begin{bmatrix}
\varepsilon_1 \\
\varepsilon_2 \\
\varepsilon_3 \\
\varepsilon_4
\end{bmatrix}
\]

For 2019 data, the model VECM(1) is:

\[
\begin{bmatrix}
Y_1 \\
Y_2 \\
Y_3 \\
Y_4
\end{bmatrix} =
\begin{bmatrix}
-1.09804 & 0.01673 & 0.02657 & 0.03201 \\
-0.42123 & -0.80048 & -0.13059 & 0.02852 \\
-0.32258 & 0.13501 & -1.27918 & 0.09241 \\
-0.45767 & 0.01594 & -0.00788 & -1.08631
\end{bmatrix}
\begin{bmatrix}
Y_{t-1} \\
Y_{t-2} \\
Y_{t-3} \\
Y_{t-4}
\end{bmatrix}
+ 
\begin{bmatrix}
\varepsilon_1 \\
\varepsilon_2 \\
\varepsilon_3 \\
\varepsilon_4
\end{bmatrix}
\]

3.4 Model Stability

### Table 9. Characteristic of Roots AR

| Year | Index | Real   | Imaginary | Modulus | Radian  | Degree    |
|------|-------|--------|-----------|---------|---------|-----------|
| 2018 | 1     | -0.07445 | 0.11823   | 0.1397  | 2.1328  | 122.1981  |
|      | 2     | -0.07445 | -0.11823  | 0.1397  | -2.1328 | -122.1981 |
|      | 3     | -0.23480 | 0.00000   | 0.2348  | 3.1416  | 180.0000  |
|      | 4     | -0.27079 | 0.00000   | 0.2708  | 3.1416  | 180.0000  |
| 2019 | 1     | -0.00644 | 0.00000   | 0.0064  | 3.1416  | 180.0000  |
|      | 2     | -0.26927 | 0.15934   | 0.3129  | 2.6073  | 149.3852  |
|      | 3     | -0.26927 | -0.15934  | 0.3129  | -2.6073 | -149.3852 |
|      | 4     | -0.42779 | 0.00000   | 0.4278  | 3.1416  | 180.0000  |
| 2020 | 1     | 0.12942  | 0.00000   | 0.1294  | 0.0000  | 0.0000    |
|      | 2     | -0.07509 | 0.15612   | 0.1732  | 2.0191  | 115.6858  |
|      | 3     | -0.07509 | -0.15612  | 0.1732  | -2.0191 | -115.6858 |
|      | 4     | -0.24327 | 0.00000   | 0.2433  | 3.1416  | 180.0000  |

Table above shows that the modulus of characteristic roots for every lag is < 1 for each year. Hence, the VECM(1) model for each year has high stability.
3.5 Impulse Response Function

Figure 1. Graph of 2018 Data IRF

Figure 1 shows every variables impulse response to other variables in data of 2018. As shown above, the stock prices give each other shocks, even themselves. The fluctuations of the shock response mostly ended in lag 4. Hence, the stock prices become stable after lag 4.
Figure 2 shows every variables impulse response to other variables in data of 2019. As shown above, the stock prices give each other shocks, even themselves. The fluctuations of the shock response mostly ended in lag 4 and lag 6. Hence, the stock prices become stable after lag 4 and lag 6.
Figure 3. Graph of 2020 Data IRF

Figure 3 shows every variables impulse response to other variables in data of 2020. As shown above, the stock prices give each other shocks, even themselves. The fluctuations of the shock response mostly ended in lag 4. Hence, the stock prices become stable after lag 4.

### 3.6 Granger Causality

| Test | Group Variable | Pr > ChiSq | Granger-cause |
|------|----------------|------------|---------------|
| 1    | Group 1 Variables : PALM  
Group 2 Variables : LSIP, SGRO, SSMS | <.0001 | Yes |
| 2    | Group 1 Variables : PALM  
Group 2 Variables : LSIP, SGRO | 0.0006 | Yes |
| 3    | Group 1 Variables : PALM  
Group 2 Variables : LSIP, SSMS | 0.0027 | Yes |
| 4    | Group 1 Variables : PALM  
Group 2 Variables : SGRO, SSMS | <.0001 | Yes |
| 5    | Group 1 Variables : PALM  
Group 2 Variables : LSIP | 0.0006 | Yes |
| 6    | Group 1 Variables : PALM  
Group 2 Variables : SGRO | 0.0007 | Yes |
| 7    | Group 1 Variables : PALM  
Group 2 Variables : SSMS | 0.0923 | No |
Table 10 shows that data of 2018 has 9 granger tests has p-value less than $\alpha = 0.05$, hence $H_0$ is rejected and group 1 variables influenced by the group 2 variables.
## Table 11. Granger Causality Test of 2019 Data

| Test | Group 1 Variables | Group 2 Variables | Pr > ChiSq | Granger-cause |
|------|------------------|------------------|------------|--------------|
| 1    | Group 1 Variables: PALM | Group 2 Variables: LSIP, SGRO, SSMS | 0.0092 | Yes |
| 2    | Group 1 Variables: PALM | Group 2 Variables: LSIP, SGRO | 0.0102 | Yes |
| 3    | Group 1 Variables: PALM | Group 2 Variables: LSIP, SSMS | 0.0035 | Yes |
| 4    | Group 1 Variables: PALM | Group 2 Variables: SGRO, SSMS | 0.0046 | Yes |
| 5    | Group 1 Variables: PALM | Group 2 Variables: LSIP | 0.0024 | Yes |
| 6    | Group 1 Variables: PALM | Group 2 Variables: SGRO | 0.9036 | No |
| 7    | Group 1 Variables: PALM | Group 2 Variables: SSMS | 0.0016 | Yes |
| 8    | Group 1 Variables: LSIP | Group 2 Variables: PALM, SGRO, SSMS | 0.0146 | Yes |
| 9    | Group 1 Variables: LSIP | Group 2 Variables: PALM, SGRO | 0.8905 | No |
| 10   | Group 1 Variables: LSIP | Group 2 Variables: PALM, SSMS | 0.0145 | Yes |
| 11   | Group 1 Variables: LSIP | Group 2 Variables: SGRO, SSMS | 0.0051 | Yes |
| 12   | Group 1 Variables: LSIP | Group 2 Variables: PALM | 0.6307 | No |
| 13   | Group 1 Variables: LSIP | Group 2 Variables: SGRO | 0.9222 | No |
| 14   | Group 1 Variables: LSIP | Group 2 Variables: SSMS | 0.0036 | Yes |
| 15   | Group 1 Variables: SGRO | Group 2 Variables: SSMS | 0.0626 | No |
| 16   | Group 1 Variables: SGRO | Group 2 Variables: PALM, LSIP | 0.7215 | No |
| 17   | Group 1 Variables: SGRO | Group 2 Variables: PALM, SSMS | 0.0687 | No |
| 18   | Group 1 Variables: SGRO | Group 2 Variables: LSIP | 0.0482 | Yes |
| 19   | Group 1 Variables: SGRO | Group 2 Variables: LSIP, SSMS | 0.9011 | No |
| 20   | Group 1 Variables: SGRO | Group 2 Variables: PALM | 0.5755 | No |
| 21   | Group 1 Variables: SGRO | Group 2 Variables: LSIP | 0.0771 | No |
| 22   | Group 1 Variables: SSMS | Group 2 Variables: SSMS | 0.5101 | No |
| 23   | Group 1 Variables: SSMS | Group 2 Variables: PALM, LSIP, SGRO | 0.3234 | No |
Table 11 shows that data of 2019 has 11 granger tests has p-value less than $\alpha = 0.05$, hence $H_0$ is rejected and group 1 variables influenced by the group 2 variables.

Table 12. Granger Causality Test of 2020 Data

| Test | Group Variable | $Pr > \text{ChiSq}$ | Granger-cause |
|------|----------------|----------------------|--------------|
| 1    | Group 1 Variables: PALM | 0.0020 | Yes |
|      | Group 2 Variables: LSIP, SGRO, SSMS | | |
| 2    | Group 1 Variables: PALM | 0.2370 | No |
|      | Group 2 Variables: LSIP, SGRO | | |
| 3    | Group 1 Variables: PALM | 0.1221 | No |
|      | Group 2 Variables: LSIP, SSMS | | |
| 4    | Group 1 Variables: PALM | 0.0675 | No |
|      | Group 2 Variables: SGRO, SSMS | | |
| 5    | Group 1 Variables: PALM | 0.3360 | No |
|      | Group 2 Variables: LSIP | | |
| 6    | Group 1 Variables: PALM | 0.1848 | No |
|      | Group 2 Variables: SGRO | | |
| 7    | Group 1 Variables: PALM | 0.3061 | No |
|      | Group 2 Variables: SSMS | | |
| 8    | Group 1 Variables: LSIP | 0.4668 | No |
|      | Group 2 Variables: PALM, SGRO, SSMS | | |
| 9    | Group 1 Variables: LSIP | 0.3209 | No |
|      | Group 2 Variables: PALM, SGRO | | |
| 10   | Group 1 Variables: LSIP | 0.7509 | No |
|      | Group 2 Variables: PALM, SSMS | | |
| 11   | Group 1 Variables: LSIP | 0.4564 | No |
|      | Group 2 Variables: SGRO, SSMS | | |
| 12   | Group 1 Variables: LSIP | 0.5126 | No |
|      | Group 2 Variables: PALM | | |
| 13   | Group 1 Variables: LSIP | 0.2093 | No |
|      | Group 2 Variables: SGRO | | |
| 14   | Group 1 Variables: LSIP | 0.5590 | No |
|      | Group 2 Variables: SSMS | | |
| 15   | Group 1 Variables: SGRO | 0.7649 | No |
|      | Group 2 Variables: PALM, LSIP, SSMS | | |
| 16   | Group 1 Variables: SGRO | 0.7259 | No |
|      | Group 2 Variables: PALM, LSIP | | |
|    | Group 1 Variables |  | Group 2 Variables |  |
|----|-------------------|---|-------------------|---|
| 17 | SGRO              | 0.6393 | PALM, SSMS | No |
| 18 | SGRO              | 0.9479 | LSIP, SSMS | No |
| 19 | SGRO              | 0.4237 | PALM | No |
| 20 | SGRO              | 0.8359 | LSIP | No |
| 21 | SGRO              | 0.7456 | SSMS | No |
| 22 | SSMS              | 0.0977 | PALM, LSIP, SGRO | No |
| 23 | SSMS              | 0.0487 | PALM, LSIP | Yes |
| 24 | SSMS              | 0.0900 | PALM, SGRO | No |
| 25 | SSMS              | 0.0831 | LSIP, SGRO | No |
| 26 | SSMS              | 0.0276 | PALM | Yes |
| 27 | SSMS              | 0.0500 | LSIP | No |
| 28 | SSMS              | 0.6338 | SGRO | No |

Table 12 shows that data of 2020 has 3 granger tests has $p$-value less than $\alpha = 0.05$, hence $H_0$ is rejected and group 1 variables influenced by the group 2 variables.

4. Conclusions

Based on the discussion and results detailed above, we can conclude that the data of daily stock prices of PT. Provident Agro Tbk, PT. PP London Sumatera Indonesia Tbk, PT. Sampoerna Agro Tbk, and PT. Sawit Sumbermas Sarana from January to July in 2018, 2019, and 2020 have cointegration relationship and can be modeled by using Vector Error Correction Model (1). In impulse response function we can conclude that each variables give each other shock response, even themselves.

Acknowledgement

The author thank yahoo finance for providing data in this study

References

[1] Chinnamgari S and Lesmeister C 2019 Advanced Machine Learning with R. (Packt Publishing, United Kingdom)
[2] Enders W 2015 Applied Econometric Time Series (John Wiley and Sons Interscience Publication, New York)
[3] Engle FR, Granger CWJ 1987 Cointegration and error correction: Representation, estimation and testing Econometrica 55(2):251–76
[4] Kirchgassner G and Wolters J 2007 Introduction to Modern Time Series Analysis (Springer, Berlin)
[5] Lütkepohl H 2005 New Introduction to Multiple Time Series Analysis (Springer-Verlag, Berlin)
[6] Medvegyev P, Edina B and Ferenc I 2015 Mastering R For Quantitative Finance (Packt Publishing, United Kingdom)
[7] Petersen J and Kumar V 2012 *Statistical Method in Customer Relationship Management* (John Wiley and Sons, United Kingdom)

[8] Tsay RS 2010 *Analysis of Financial Time Series 3rd Ed.* (John Wiley and Sons, New Jersey)

[9] Yahoo Finance (2020), Daily Stock Prices History, https://finance.yahoo.com, accessed on August 2nd 2020.