

La GUBU, PhD (corresponding author)

la.gubu@uho.ac.id

Halu Oleo University, Kendari, Indonesia

Muhamad Rashif HILMI, PhD Candidate

rashif.hilmi@uin-suka.ac.id

Sunan Kalijaga State Islamic University, Yogyakarta, Indonesia

Beyond Mean-Variance Markowitz Portfolio Selection: A Comparison of Mean-Variance-Skewness-Kurtosis Model and Robust Mean-Variance Model

Abstract. *In this paper, two developments of the classical Markowitz Mean-Variance (MV) portfolio model are presented, namely, the Mean-Variance-Skewness-Kurtosis (MVSK) portfolio model and the robust Mean-Variance (MV) portfolio. The robust MV portfolio model uses two robust estimations: the robust Fast Minimum Covariance Determinant (FMCD) and the robust Constrained M (CM). The robust MV portfolios that use FMCD estimation yield the MV_{FMCD} portfolio model, while the robust MV portfolios that use CM estimation yield the MV_{CM} portfolio model. The MVSK model is intended to overcome the fact that most stock returns in the capital market do not follow the normal distribution, and there are skewness and excessive kurtosis, while the robust MV portfolio model is intended to overcome the presence of outliers in the data. An empirical study revealed that robust MV portfolios outperforms MVSK and classical MV Markowitz models. Besides, we also found that for $\gamma = 0.5$ to $\gamma = 20$, the portfolio using the robust MV_{FMCD} model outperforms the portfolio using the robust MV_{CM} model. However, for $\gamma > 20$, the robust MV_{CM} model outperforms the portfolio using the robust MV_{FMCD} model.*

Keywords: *robust estimation, robust portfolio, MVSK portfolio, Sharpe ratio, portfolio performance.*

JEL Classification: C61, G11.

1. Introduction

Since its initial introduction by Markowitz in 1952, the Mean-Variance (MV) model has become a widely recognised and useful tool for portfolio optimisation in modern portfolio theory. Many studies on portfolio selection have been conducted based on only the first two moments of return distributions, demonstrating how the MV model has completely changed the way people think about asset portfolios since Markowitz's seminal work was published. The MV model operates under the presumption that stock returns follow a normal distribution (Naqvi et al., 2017). As noted by several earlier researchers: Lai et al. (2006), Gotoh et al. (2018), Naqvi et al. (2017), Metaxiotis (2019), and Lu et al. (2019), they discovered that stock returns are not normally distributed and can be skewed either positively or negatively with excess kurtosis. According to Khan et al. (2020), stocks that

exhibit negative skewness indicate a higher likelihood of a negative return than a positive return. Many researchers, including Chen et al. (2020), Marques and Benasciutti (2020), and Khan et al. (2020), have taken into account the third- and fourth-order moments, skewness, and kurtosis when choosing the best portfolio. Naqvi et al. (2017) and Díaz et al. (2022) explained that skewness and kurtosis play a critical role in portfolio formation. The model of portfolio that considers aspects of skewness and kurtosis is called the Mean-Variance-Skewness-Kurtosis (MVSK) model. The objectives of this model are to minimise risk and excess kurtosis while maximising profit and skewness.

Another problem with the MV portfolio model is that the mean vector and variance-covariance matrix must be estimated from very fluctuating data, and it is not uncommon for outlier data to appear. There are many different estimation techniques available for parameter estimation, and any method used will inevitably result in estimation errors. Estimation errors play a crucial role in the creation of MV portfolio models and have a substantial impact on the outcomes of optimum portfolio formation. Best and Grauer (1991), Chopra and Ziemba (1993), Bengtsson (2004), and Ceria and Stubbs (2006) have all conducted studies pertaining to estimation errors and their relationship with optimal portfolio formation. These studies lead to the conclusion that while the MV model has strong theoretical backing and makes computation easier, it has certain drawbacks, such as the poorly diversified optimal portfolio it generates. A limited number of assets typically make up the majority of the resulting portfolio. Furthermore, the variance-covariance matrix and the mean vector, which are the input parameters of the MV model, are extremely sensitive to changes in these parameters. In order to decrease the error of the estimated vector mean and the variance-covariance matrix in the MV portfolio model, some researchers have constructed a robust portfolio. Robust estimation is one of the common methods used to create an optimal robust portfolio. Several studies have been conducted by DeMiguel and Nogales (2008), Kusch (2012), Hu (2012), Supandi (2017), Ghahtarani and Najafi (2018), Gubu et al. (2020), and Gubu et al. (2021) regarding the creation of optimal portfolios through robust estimation. The type of robust estimation employed in these studies differs from one another. All of the study's results stated that the portfolio's performance with robust estimation is better than the classical portfolio. Nevertheless, comparing the optimal portfolio performance constructed using the classic MV, MVSK, and robust MV models has not yet been covered in the literature mentioned above.

In this paper, as our new contribution, we compare the portfolio's performance constructed by MV classic, MVSK, and robust MV models. In robust MV models, we employ the Fast Minimum Covariance Determinant (FMCD) and Constrained M (CM) robust estimation techniques to determine the mean and covariance of the data.

2. Material and Method

This section will provide the material and methods used in this research, including Mean-Variance portfolio, Mean-Variance-Skewness-Kurtosis portfolio, outlier detection, robust estimation, and Sharpe ratio.

2.1 Mean-Variance portfolio

The relationship between risk and return in investment management is sturdy and linear. The return will also be high if the risk is high, and vice versa if the return is low. In the 1950s, Harry M. Markowitz pioneered the Modern Portfolio Theory. This theory posits that investments have two components: risk and return. It suggests that diversification and assembling a portfolio of different investment instruments can reduce risk. The Journal of Finance published a widely read version of the theory in 1952.

The foundation of Markowitz's portfolio theory is the mean-variance approach, in which risk is measured by variance, and expected return is measured by the mean (Markowitz, 1952). Consequently, the mean-variance model (MV) is another name for Markowitz's portfolio theory. To select and construct an optimal portfolio, this model strongly emphasises attempting to maximise expected return and minimise risk. Supandi (2017) states that the following optimisation problem can be solved to formulate the mean-variance portfolio:

$$\max_{\mathbf{w}} \mathbf{w}'\boldsymbol{\mu} - \frac{\gamma}{2} \mathbf{w}'\boldsymbol{\Sigma}\mathbf{w} \tag{1}$$

$$\mathbf{w}'\mathbf{e} = 1 \tag{2}$$

where $\boldsymbol{\Sigma}$ is the covariance matrix, \mathbf{e} is the column matrix with all elements equal to 1, $\boldsymbol{\mu}$ is the mean vector, $\gamma \geq 0$ is the relative risk avoidance measure, and \mathbf{w} represents the portfolio's weight.

To achieve a specific level of return, every investor seeks a particular degree of risk. Investors should select the appropriate γ to balance the trade-off between risks and returns, since returns offset risks. Investors seek to lower risk (loss) and increase the rate of return (profit) in two extreme scenarios. Equation (1) gives the maximum rate of return when $\gamma = 0$, regardless of the risk that must be assumed. In the meantime, investors will select the lowest risk option regardless of return level if $\gamma = \infty$.

The optimisation problems in equations (1) and (2) have been solved by (Gubu et al., 2020) using the Lagrange method and the solution is:

$$\mathbf{w} = \frac{1}{\gamma} (\boldsymbol{\Sigma}^{-1} - \boldsymbol{\Sigma}^{-1}\mathbf{e}(\mathbf{e}'\boldsymbol{\Sigma}^{-1}\mathbf{e})^{-1}\mathbf{e}'\boldsymbol{\Sigma}^{-1})\boldsymbol{\mu} + \boldsymbol{\Sigma}^{-1}\mathbf{e}(\mathbf{e}'\boldsymbol{\Sigma}^{-1}\mathbf{e})^{-1} \tag{3}$$

Equation (3) reveals that input $\boldsymbol{\mu}$ dan $\boldsymbol{\Sigma}$ determines the optimal portfolio (\mathbf{w}).

2.2 Mean-Variance-Skewness-Kurtosis Portfolio

Investors take into account the mean, variance, skewness, and kurtosis when building a portfolio using the MVSK model. It is commonly known that covariance between asset returns and asset variance also contributes to the variance of the portfolio. The returns of the portfolio's assets cannot be taken to be independent since they typically move in tandem. Co-skewness and co-kurtosis return assets are also included in skewness and kurtosis portfolios, albeit in slightly different ways. The following is the formula of the co-skewness and co-kurtosis:

$$\begin{aligned} \sigma_{ijk} &= E[(r_i - \mu_i)(r_j - \mu_j)(r_k - \mu_k)] \\ &= E[(r_i r_j r_k)] - \mu_i \sigma_{jk} - \mu_j \sigma_{ik} - \mu_k \sigma_{ij} - \mu_i \mu_j \mu_k \end{aligned}$$

and

$$\begin{aligned} \sigma_{ijkl} &= E[(r_i - \mu_i)(r_j - \mu_j)(r_k - \mu_k)(r_l - \mu_l)] \\ &= E[(r_i r_j r_k r_l)] - \mu_i \sigma_{jkl} - \mu_j \sigma_{ikl} - \mu_k \sigma_{ijl} - \mu_l \sigma_{ijk} \\ &\quad - \mu_i \mu_j \sigma_{kl} - \mu_i \mu_k \sigma_{jl} - \mu_i \mu_l \sigma_{jk} - \mu_j \mu_k \sigma_{il} - \mu_j \mu_l \sigma_{ik} - \mu_k \mu_l \sigma_{ij} - \mu_i \mu_j \mu_k \mu_l \end{aligned}$$

where r_i is return of stock i and μ_i is mean return of stock i .

Let us assume that the portfolio contains p assets. The co-skewness matrix (\mathbf{M}_3) is a $p \times p^2$ matrix with the entry σ_{ijk} . While the co-kurtosis matrix (\mathbf{M}_4) is $p \times p^3$ matrix with the entry σ_{ijkl} . More clearly can be written as follows

$$\mathbf{M}_3 = \begin{bmatrix} \sigma_{111} & \dots & \sigma_{1p1} & \dots & \sigma_{11p} & \dots & \sigma_{1pp} \\ \vdots & \ddots & \vdots & \dots & \vdots & \ddots & \vdots \\ \sigma_{p11} & \dots & \sigma_{pp1} & \dots & \sigma_{p1p} & \dots & \sigma_{ppp} \end{bmatrix}$$

and

$$\mathbf{M}_4 = \begin{bmatrix} \sigma_{1111} & \dots & \sigma_{1p11} & \dots & \sigma_{11p1} & \dots & \sigma_{1pp1} & \vdots & \dots & \dots \\ \vdots & \ddots & \vdots & \dots & \vdots & \ddots & \vdots & \vdots & \dots & \dots \\ \sigma_{p111} & \dots & \sigma_{pp11} & \dots & \sigma_{p1p1} & \dots & \sigma_{ppp1} & \vdots & \dots & \dots \\ \dots & \dots & \vdots & \sigma_{111p} & \dots & \sigma_{1p1p} & \vdots & \dots & \sigma_{11pp} & \dots & \sigma_{1ppp} \\ \dots & \dots & \vdots & \vdots & \ddots & \vdots & \vdots & \dots & \vdots & \ddots & \vdots \\ \dots & \dots & \vdots & \sigma_{p11p} & \dots & \sigma_{pp1p} & \vdots & \dots & \sigma_{p1pp} & \dots & \sigma_{pppp} \end{bmatrix}$$

Furthermore, the skewness portfolio (\mathbf{s}_{port}) and kurtosis portfolio (\mathbf{k}_{port}) are defined as the third and fourth moments around the mean respectively.

$$\mathbf{s}_{port} = E(R_p - E(R_p))^3 = \mathbf{w}^T \mathbf{M}_3 (\mathbf{w} \otimes \mathbf{w}) \tag{4}$$

and

$$\mathbf{k}_{port} = E(R_p - E(R_p))^4 = \mathbf{w}^T \mathbf{M}_4 (\mathbf{w} \otimes \mathbf{w}) \tag{5}$$

In this case, \otimes is Kronecker product and $\mathbf{w}^T = [w_1 \dots w_p]$, w_i is the weight for the stock i that will be determined.

The primary challenge when optimising a portfolio with the MVSK model is to figure out how much money should go into each stock in order to produce a portfolio with a high mean, positive skewness, lower variance, and the least amount of excess kurtosis possible when all the money is invested. According to Lai et al. (2006), mathematically, it can be expressed as:

$$\text{maximise } R_p = \mathbf{r}^T \mathbf{w} \tag{6}$$

$$\text{minimise } \rho_{port}^2 = \mathbf{w}^T \mathbf{M}_2 \mathbf{w} \tag{7}$$

$$\text{maximise } s_{port} = \mathbf{w}^T \mathbf{M}_3 (\mathbf{w} \otimes \mathbf{w}) \tag{8}$$

$$\text{minimise } k_{port} = \mathbf{w}^T \mathbf{M}_4 (\mathbf{w} \otimes \mathbf{w} \otimes \mathbf{w}) \tag{9}$$

$$\text{with constraint } \mathbf{w}^T \mathbf{1}_p = 1 \tag{10}$$

The linear combination can be formed by providing the four weighted coefficients, a_1, a_2, a_3 , and a_4 , based on the objective function (6)-(9). It can be expressed as follows:

$$\begin{aligned} \text{minimise } & -a_1 \mathbf{r}^T \mathbf{w} + a_2 \mathbf{w}^T \mathbf{M}_2 \mathbf{w} - a_3 \mathbf{w}^T \mathbf{M}_3 (\mathbf{w} \otimes \mathbf{w}) \\ & + a_4 \mathbf{w}^T \mathbf{M}_4 (\mathbf{w} \otimes \mathbf{w} \otimes \mathbf{w}) \end{aligned} \tag{11a}$$

$$\text{with constraint } \mathbf{w}^T \mathbf{1}_p = 1 \tag{11b}$$

To minimise the objective function and the given constraint, apply the Lagrange function as follows:

$$\begin{aligned} L = & -a_1 \mathbf{r}^T \mathbf{w} + a_2 \mathbf{w}^T \mathbf{M}_2 \mathbf{w} - a_3 \mathbf{w}^T \mathbf{M}_3 (\mathbf{w} \otimes \mathbf{w}) + a_4 \mathbf{w}^T \mathbf{M}_4 (\mathbf{w} \otimes \mathbf{w} \otimes \mathbf{w}) \\ & + \alpha (\mathbf{w}^T \mathbf{1}_p - 1) \end{aligned} \tag{12}$$

Let $a_1 = s$, $a_2 = \lambda$, $a_3 = u$ and $a_4 = v$, where $s, u, v \geq 0$ and $\lambda > 0$.

To optimise (12), the step is derived L to \mathbf{w} and equates to zero, and obtained:

$$\mathbf{w} = \frac{1}{2\lambda} \mathbf{M}_2^{-1} (s\mathbf{r} + 3u\mathbf{M}_3 (\mathbf{w} \otimes \mathbf{w}) - 4v\mathbf{M}_4 (\mathbf{w} \otimes \mathbf{w} \otimes \mathbf{w}) - \alpha \mathbf{1}_p) \tag{13}$$

By substituting Equation (13) into Equation (11b), we obtain

$$\begin{aligned} \alpha = & \frac{s}{\mathbf{1}_p^T \mathbf{M}_2^{-1} \mathbf{1}_p} \mathbf{1}_p^T \mathbf{M}_2^{-1} \mathbf{r} + \frac{3u}{\mathbf{1}_p^T \mathbf{M}_2^{-1} \mathbf{1}_p} \mathbf{1}_p^T \mathbf{M}_2^{-1} \mathbf{M}_3 (\mathbf{w} \otimes \mathbf{w}) - \\ & \frac{4v}{\mathbf{1}_p^T \mathbf{M}_2^{-1} \mathbf{1}_p} \mathbf{1}_p^T \mathbf{M}_2^{-1} \mathbf{M}_4 (\mathbf{w} \otimes \mathbf{w} \otimes \mathbf{w}) - \frac{2\lambda}{\mathbf{1}_p^T \mathbf{M}_2^{-1} \mathbf{1}_p} \end{aligned} \tag{14}$$

The MVSK portfolio's weight has become:

$$\mathbf{w}_{MVSK} = -\frac{1}{2\lambda} \mathbf{M}_2^{-1} (s\mathbf{r} + 3u\mathbf{M}_3 (\mathbf{w} \otimes \mathbf{w}) - 4v\mathbf{M}_4 (\mathbf{w} \otimes \mathbf{w} \otimes \mathbf{w}) - \alpha \mathbf{1}_p)$$

where α as given in Equation (14).

By choosing $s=1$, $u=1$, $v=1$, and λ as varied risk aversion coefficients, the MVSK portfolio weight can be determined for given risk aversion using the Newton-Rapson method, as has been done by Agustina et al. (2022).

2.3 Outlier detection

Observations that significantly deviate from the remaining data points in a population sample are known as outliers. The higher the analytical result of a sample, the greater the observation's distance from the centre of all observations. Outliers frequently have significant distances from the centre of all observations (Filzmoser et al., 2005). The distance of an observation from the data centroid and the form of the data must be considered in multivariate situations.

Multivariate data's size and form are quantified by the covariance matrix. One widely used distance measure that takes the covariance matrix into account is the Mahalanobis distance. The Mahalanobis distance for a multivariate sample $\mathbf{x}_1, \dots, \mathbf{x}_n$, in p -dimensions is defined as follows:

$$MD_i = ((\mathbf{x}_i - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x}_i - \boldsymbol{\mu}))^{1/2} \quad \text{for } i = 1, \dots, n \quad (15)$$

here, $\boldsymbol{\mu}$ stands for the estimated multivariate location, and $\boldsymbol{\Sigma}$ for the covariance matrix (Filzmoser et al., 2005). $\boldsymbol{\mu}$ is commonly used to represent the multivariate arithmetic mean, while $\boldsymbol{\Sigma}$ stands for the sample covariance matrix.

In multivariate normally distributed data, the values of MD_i^2 are approximately chi-square distributed with p degrees of freedom (χ_p^2). Ellipsoids with the same distance from the centroid can be defined by using a quantile of χ_p^2 as a constant for the (squared) Mahalanobis distance (Gnanadesikan, 1977).

Filzmoser et al. (2005) define large (squared) Mahalanobis distance observations as outliers in multivariate statistics. Furthermore, Filzmoser et al. emphasised that in the multivariate case, a quantile of the chi-squared distribution (like the 98% quantile, $\chi_{p;0.98}^2$), might be considered an outlier.

2.4 Robust estimation

Robust estimating for multivariate data comes in a wide variety. A member of the affine equivariant class, robust estimation is used in this work.

Given a data set $(\mathbf{R} = \mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_n)$ with $\mathbf{r}_i \in \mathbb{R}^p$. The robust estimate for the mean vector is $\hat{\boldsymbol{\mu}}(\mathbf{R}) \in \mathbb{R}^p$ while the covariance matrix estimation is $\hat{\boldsymbol{\Sigma}}(\mathbf{R}) \in \mathbb{P}_p$ (the set of all symmetric matrices positive definite of size $p \times p$). An estimation satisfies the requirements of the following definition to be classified as belonging to the affine equivariant class (Maronna et al., 2006).

Definition 2.1 (Maronna et al., 2006)

Given \mathbf{Q} is an invertible matrix of size $p \times p$, vector $\mathbf{v} \in \mathbb{R}^p$ and data sets $\mathbf{R} \in \mathbb{R}^{p \times n}$.

- i. $\hat{\boldsymbol{\mu}}$ is affine equivariant if $\hat{\boldsymbol{\mu}}(\mathbf{QR} + \mathbf{v}) = \mathbf{Q}\hat{\boldsymbol{\mu}}(\mathbf{R}) + \mathbf{v}$
- ii. $\hat{\boldsymbol{\Sigma}}$ is affine equivariant if $\hat{\boldsymbol{\Sigma}}(\mathbf{QR} + \mathbf{v}) = \mathbf{Q}\hat{\boldsymbol{\Sigma}}(\mathbf{R}) + \mathbf{v}$

Constrained M (CM) and Fast Minimum Covariance Determinant (FMCD) are two examples of robust estimation that are part of the affine equivariant class used in this study. A brief discussion of the FMCD and CM estimation will be given below. Additionally, we described how to calculate $\hat{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{\Sigma}}$ using the robust FMCD and robust CM estimation techniques.

2.4.1 Robust FMCD estimation

The goal of the minimum covariance determinant (MCD) estimation is to identify robust estimation where the covariance matrix has the smallest determinant, based on observations of total observations (n). The MCD estimation is a pair of $\hat{\boldsymbol{\mu}} \in \mathbb{R}^p$ and $\hat{\boldsymbol{\Sigma}}$ (a symmetric positive definite matrix with a dimension of $p \times p$ from a sample of h observation)

$$\text{where } \frac{(n+p+1)}{2} \leq h \leq n \tag{16}$$

with

$$\hat{\boldsymbol{\mu}} = \frac{1}{h} \sum_{i=1}^h \mathbf{r}_i \tag{17}$$

Solving the following equation yields the estimation of the covariance matrix.

$$\hat{\boldsymbol{\Sigma}} = \frac{1}{h} \sum_{i=1}^h (\mathbf{r}_i - \hat{\boldsymbol{\mu}})(\mathbf{r}_i - \hat{\boldsymbol{\mu}})' \tag{18}$$

Because MCD must look at every possible subset of h from a number of n data, this method's calculations can become very complex as the data dimensions increase. Thus, Fast MCD (FMCD), a faster MCD calculation algorithm, was discovered by Rousseeuw and Driessen (1999). The C-Step theorem explained below forms the foundation of the FMCD method.

Theorem 2. 1 (Rousseeuw and Driessen, 1999)

If H_1 is the set of size h taken from data of size n , the sample statistics are:

$$\hat{\boldsymbol{\mu}}^1 = \frac{1}{h} \sum_{i \in H_1} \mathbf{r}_i \tag{19}$$

$$\hat{\boldsymbol{\Sigma}}^1 = \frac{1}{h} \sum_{i \in H_1} (\mathbf{r}_i - \hat{\boldsymbol{\mu}}^1)(\mathbf{r}_i - \hat{\boldsymbol{\mu}}^1)' \tag{20}$$

If $|\hat{\boldsymbol{\Sigma}}^1| > 0$ than distance $d_i = (\mathbf{r}_i; \hat{\boldsymbol{\mu}}^1, \hat{\boldsymbol{\Sigma}}^1)$. Next, specify H_2 is subset consist of the observation with the smallest distance d_i , namely $\{d_1(i) | i \in H_2\} = \{(d_1)_1, \dots, (d_1)_h\}$ where $(d_1)_1 \leq (d_1)_2 \leq \dots \leq (d_1)_n$ is a sequential distance. Based on H_2 , calculate $\hat{\boldsymbol{\mu}}^2$ and $\hat{\boldsymbol{\Sigma}}^2$ using equations (19) and (20), so that

$$|\hat{\boldsymbol{\Sigma}}^2| \leq |\hat{\boldsymbol{\Sigma}}^1| \tag{21}$$

Equation (21) will be the same if $\hat{\boldsymbol{\mu}}^1 = \hat{\boldsymbol{\mu}}^2$ and $\hat{\boldsymbol{\Sigma}}^1 = \hat{\boldsymbol{\Sigma}}^2$.

The C-Step theorem is applied repeatedly until either $|\hat{\boldsymbol{\Sigma}}_{new}| = 0$ or $|\hat{\boldsymbol{\Sigma}}_{new}| = |\hat{\boldsymbol{\Sigma}}_{old}|$.

Nevertheless, there is no guarantee that the final result of the iteration process will yield a new $\hat{\Sigma}_{new}$ with a global minimum of the objective function, which is the MCD estimation. In order to obtain the smallest determinant, the MCD solution approach can be carried out by choosing a number of initial sets of H_1 , and then applying C-Step to each set. In order to solve this issue, Rousseeuw and Driessen (1999) developed the Fast Minimum Covariant Determinant (FMCD) algorithm. The following algorithm provides an explanation of how the FMCD estimation is calculated.

1. Consider the subset of the matrix \mathbf{R} represented by H_1 and made up of the observations $h = (n + p + 1)/2$.
2. Determine both the covariance matrix ($\hat{\Sigma}^1$) and the mean vector ($\hat{\mu}^1$).
3. Calculate mahalanobis distance $d_1(i) = (\mathbf{r}_i - \hat{\mu}_1)' \Sigma_1^{-1} (\mathbf{r}_i - \hat{\mu}_1)$
4. Sort $d_1(i)$ from the smallest to the largest value
5. Define the new subset with H_2 , such that $\{d_1(i); i \in H_2\} \{(d_1)_{1:n}, (d_1)_{2:n}, \dots, (d_1)_{h:n}\}$ where $(d_1)_{1:n} \leq (d_1)_{2:n} \leq \dots \leq (d_1)_{h:n}$
6. Compute the covariance matrix ($\hat{\Sigma}^2$), the mean vector ($\hat{\mu}_2$), and $d_2(i)$.
7. Steps 1 through 6 should be repeated until $|\hat{\Sigma}^2| \leq |\hat{\Sigma}^1|$.

2.4.2 Robust constrained M estimation

Robust Constrained M (CM) estimation is an extension of M Estimation. According to Kent and Tyler (1996), the advantage of M estimation is that it has good robustness characteristics both locally and globally. M estimation has the advantage of local robustness properties, such as good efficiency and a bounded influence function, but the weakness of this estimation is that it has a small breakdown point. To overcome this problem, Kent and Tyler (1996) proposed another estimation, namely CM Estimation. CM estimation has all the advantages of robustness properties both locally and globally.

Definition 2.2 (Kent dan Tyler,1996)

Given $\{\mathbf{r}_i, i = 1, \dots, n\}$ is data set in \mathbb{R}^p and P_p is set of symmetric matrices positive definite with size $p \times p$. CM estimation for measure of location $\hat{\mu} \in \mathbb{R}^p$ and dispersion $\hat{\Sigma}(R) \in P_p$ is a pair of $\hat{\mu}$ and $\hat{\Sigma}(R)$ that minimised the objective function

$$L(\boldsymbol{\mu}, \boldsymbol{\Sigma}; \mathbf{r}) = \frac{1}{n} \sum_{i=1}^n \rho(\mathbf{d}_i) + \frac{1}{2} \log |\boldsymbol{\Sigma}| \tag{22}$$

with constraint

$$\frac{1}{n} \sum_{i=1}^n \rho(\mathbf{d}_i) \leq \epsilon \rho(\infty) \tag{23}$$

where ϵ is the breakdown point. The breakdown point is the tolerance for the proportion of incorrect/erroneous observations (extreme/outlier observations). An estimation is more robust if it has a higher breakdown point.

Furthermore, Kent and Tyler (1996) explain that if ρ is differentiable, then estimates of location size and scale using the CM estimation method can be obtained by solving the following equations.

$$\boldsymbol{\mu} = \frac{\sum_{i=1}^n \psi(\mathbf{d}_i) \mathbf{r}_i}{\sum_{i=1}^n \psi(\mathbf{d}_i) r_i} \tag{24}$$

$$\boldsymbol{\Sigma} = p \frac{\sum_{i=1}^n \psi(\mathbf{d}_i) (\mathbf{r}_i - \boldsymbol{\mu})(\mathbf{r}_i - \boldsymbol{\mu})'}{\sum_{i=1}^n W(\mathbf{d}_i)} \tag{25}$$

and

$$\frac{1}{n} \sum_{i=1}^n W(\mathbf{d}_i) = p \tag{26}$$

or

$$\frac{1}{n} \sum_{i=1}^n \rho(\mathbf{d}_i) = \epsilon \rho(\infty) \tag{27}$$

where $\psi(\mathbf{d}) = 2\rho'(\mathbf{d})$ and $W(\mathbf{d}) = \mathbf{d}\psi(\mathbf{d})$

Equations (24), (25), and (26) apply if the constraints in equation (23) are in inequality form. Meanwhile, equations (24), (25), and (27) apply if the constraints in equation (23) become equations. Equations (24), (25), and (26) appear as critical points for equation (22).

2.5 Sharpe ratio

Several metrics, such as the Sharpe ratio or the Sharpe index, can be used to assess the performance of stocks or portfolios. The Sharpe ratio measures excess return (or risk premium) per unit of risk in an asset (Sharpe, 1994). Moreover, Sharpe (1994) claims that the Sharpe ratio illustrates how well asset returns cover investors' risk-taking expenses. By dividing the difference between stock returns (R) and the risk return-free rate (R_f) by the standard deviation of stock returns (σ), the Sharpe ratio (SR) is computed. This can be expressed as follows:

$$SR = \frac{R - R_f}{\sigma} \tag{28}$$

If the Sharpe ratio measures portfolio performance, the stock return and stock risk in Equation (28) are replaced with portfolio return and portfolio risk. The higher the Sharpe ratio of a stock/portfolio, the better the stock/portfolio.

3. Empirical Study

3.1 Data description

The data used in this research is the daily data of stock prices listed on the Indonesia Stock Exchange, which are included in the LQ45 index for February-July 2023. The daily stock prices for each stock are obtained online at <https://finance.yahoo.com>. Four of the 45 stocks with the best performance were selected to form an optimum portfolio. A more detailed description of the data used in this research is given in Tables 1-2 and Figure 1.

Table 1. Mean, standard deviation, skewness, kurtosis, and Sharpe ratio of stocks return

Stock	Mean	Standard Dev.	Skewness	Kurtosis	SR
BBRI	0.0018	0.0123	0.0802	3.8104	0.1303
ACES	0.0042	0.0331	1.1357	5.5378	0.1208
BRIS	0.0022	0.0284	1.2967	10.2879	0.0733
ASII	0.0013	0.0158	0.2860	7.2291	0.0731

Source: Own calculations based on the daily stocks returns of the best four stocks of the LQ45 index February-July 2023 period using R 3.6.1.

Table 2. Normality of stocks return data

Stock	p-value	Conclusion
BBRI	0.031745	Not normal distribution
ACES	0.000083	Not normal distribution
BRIS	0.000003	Not normal distribution
ASII	0.000008	Not normal distribution

Source: Own calculations based on the daily stocks returns of the best four stocks of the LQ45 index February-July 2023 period using R 3.6.1.

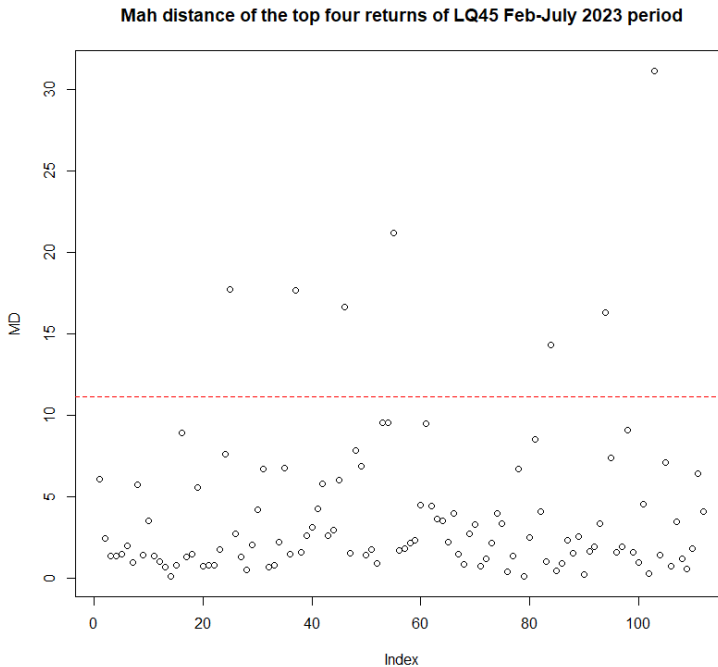


Figure 1. Outliers of stocks return

Source: Graphic illustration of outliers based on returns of the best four stocks of the LQ45 index February-July 2023 period using R 3.6.1.

3.2 Portfolios Weighting

The present study employed the MVSK and robust MV portfolio models to construct an optimal portfolio. In the robust MV model, we use two robust estimations, FMCD estimation, and CM estimation, to compute the optimal portfolio. The robust MV portfolio that uses FMCD estimation yield the MV_{FMCD} portfolio model, while the robust MV portfolio that uses CM estimation yield the MV_{CM} portfolio model.

3.2.1 Weight of MV classic and MVSK portfolio

For weighting the MVSK portfolio, we use **rootSolve function** in R packages for various risk aversion γ . We also determine the portfolio constructed using the MV classic portfolio model ($MV_{Classic}$) as a comparison. The MV classic portfolio model is a special case of the MVSK portfolio model, namely by taking $a_3 = 0$ and $a_4 = 0$ in Equation 11a, section 2.2. The stocks utilised are the best four stocks of LQ45 in the February-July 2023 period, as shown in Table 1. Tables 3-4 show the resulting portfolio weights.

Table 3. Weight of MV classic portfolio

γ	BBRI	ACES	BRIS	ASII
0.5	0.40102	4.78461	1.15354	-5.33917
1	0.47230	2.42497	0.61125	-2.50852
2	0.50794	1.24515	0.34010	-1.09319
5	0.52933	0.53726	0.17741	-0.24400
10	0.53646	0.30129	0.12318	0.03907
15	0.53883	0.22264	0.10511	0.13342
20	0.54002	0.18331	0.09607	0.18060
25	0.54073	0.15972	0.09065	0.20891
30	0.54121	0.14398	0.08703	0.22778
35	0.54155	0.13275	0.08445	0.24126
40	0.54180	0.12432	0.08251	0.25136
45	0.54200	0.11777	0.08101	0.25923
50	0.54216	0.11252	0.07980	0.26552

Source: Own calculations based on vector returns and covariance matrix (classic) of the best four stocks of the LQ45 index February-July 2023 period using R 3.6.1.

Table 4. Weight of MVSK portfolio

γ	BBRI	ACES	BRIS	ASII
0.5	0.01953	2.68255	0.56625	-2.26833
1	0.44235	1.29067	0.33519	-1.06821
2	0.51723	0.66164	0.20357	-0.38244
5	0.53593	0.30170	0.12303	0.03934
10	0.54000	0.18334	0.09603	0.18063
15	0.54123	0.14398	0.08701	0.22778

γ	BBRI	ACES	BRIS	ASII
20	0.54183	0.12431	0.08250	0.25136
25	0.54219	0.11251	0.07979	0.26551
30	0.54242	0.10465	0.07798	0.27494
35	0.54259	0.09903	0.07669	0.28168
40	0.54272	0.09482	0.07573	0.28674
45	0.54281	0.09154	0.07497	0.29067
50	0.54289	0.08892	0.07437	0.29382

Source: Own calculations based on vector returns, covariance, co-skewness, and co-kurtosis of the best four stocks of the LQ45 index February-July 2023 period using R 3.6.1.

3.2.2 Weight of robust MV_{FMCD} and MV_{CM} portfolios

We utilise the **CovMcd** and **CovMest** functions in R packages to weight the robust MV_{FMCD} and MV_{CM} portfolios for various risk aversion γ . The stocks utilised are the best four stocks of the LQ45 index in the February-July 2023 period, as shown in Table 1. Tables 5-6 show the resulting portfolio weights.

Table 5. Weight of robust MV_{FMCD} portfolio

γ	BBRI	ACES	BRIS	ASII
0.5	37.70183	-6.34245	2.19159	-32.55097
1	19.06345	-3.11799	1.13786	-16.08333
2	9.74427	-1.50576	0.61100	-7.84951
5	4.15275	-0.53842	0.29488	-2.90921
10	2.28892	-0.21597	0.18950	-1.26245
15	1.66764	-0.10849	0.15438	-0.71353
20	1.35700	-0.05475	0.13682	-0.43906
25	1.17061	-0.02251	0.12628	-0.27439
30	1.04636	-0.00101	0.11926	-0.16460
35	0.95760	0.01434	0.11424	-0.08619
40	0.89104	0.02586	0.11047	-0.02737
45	0.83926	0.03482	0.10755	0.01837
50	0.79785	0.04198	0.10521	0.05496

Source: Own calculations based on vector returns and covariance (estimated by robust FMCD estimation) of the best four stocks of the LQ45 index February-July 2023 period using R 3.6.1.

Table 6. Weight of robust MV_{CM} portfolio

γ	BBRI	ACES	BRIS	ASII
0.5	33.24738	-5.40152	5.30094	-32.14681
1	16.83357	-2.64755	2.68538	-15.87140
2	8.62667	-1.27056	1.37760	-7.73370
5	3.70253	-0.44437	0.59293	-2.85108
10	2.06115	-0.16898	0.33137	-1.22354
15	1.51402	-0.07718	0.24419	-0.68103

γ	BBRI	ACES	BRIS	ASII
20	1.24045	-0.03128	0.20059	-0.40977
25	1.07632	-0.00374	0.17444	-0.24702
30	0.96689	0.01462	0.15700	-0.13851
35	0.88873	0.02774	0.14455	-0.06101
40	0.83011	0.03757	0.13520	-0.00289
45	0.78452	0.04522	0.12794	0.04232
50	0.74804	0.05134	0.12213	0.07849

Source: Own calculations based on vector returns and covariance (estimated by robust CM estimation) of the best four stocks of the LQ45 index February-July 2023 period using R 3.6.1.

3.3 Portfolios performance comparison

The portfolio's returns, risks, and Sharpe ratios can then be determined using the portfolio weights, mean vectors, and covariance matrices. The portfolio's returns, risks, and Sharpe ratios are shown in Table 7.

Table 7. Return, risk, and Sharpe ratio of portfolios

γ	$MV_{Classic}$			MV_{SK}			MV_{FMCD}			MV_{CM}		
	Return	Risk	SR	Return	Risk	SR	Return	Risk	SR	Return	Risk	SR
0.5	0.0162	0.1700	0.0944	0.0095	0.0917	0.1019	0.1776	0.5951	0.2981	0.1676	0.5779	0.2898
1	0.0090	0.0853	0.1037	0.0055	0.0444	0.1202	0.0890	0.2976	0.2986	0.0841	0.2890	0.2905
2	0.0054	0.0433	0.1212	0.0036	0.0230	0.1501	0.0448	0.1490	0.2996	0.0424	0.1447	0.2920
5	0.0032	0.0190	0.1621	0.0025	0.0121	0.1957	0.0182	0.0600	0.3011	0.0174	0.0583	0.2950
10	0.0025	0.0121	0.1957	0.0022	0.0096	0.2090	0.0094	0.0307	0.3000	0.0090	0.0299	0.2961
15	0.0023	0.0103	0.2064	0.0020	0.0091	0.2081	0.0064	0.0213	0.2946	0.0062	0.0207	0.2928
20	0.0022	0.0096	0.2090	0.0020	0.0089	0.2059	0.0050	0.0167	0.2862	0.0048	0.0163	0.2862
25	0.0021	0.0093	0.2090	0.0020	0.0088	0.2040	0.0041	0.0142	0.2758	0.0040	0.0138	0.2774
30	0.0020	0.0091	0.2081	0.0019	0.0087	0.2024	0.0035	0.0125	0.2644	0.0034	0.0123	0.2674
35	0.0020	0.0089	0.2070	0.0019	0.0087	0.2011	0.0031	0.0115	0.2527	0.0030	0.0112	0.2569
40	0.0020	0.0089	0.2059	0.0019	0.0087	0.2001	0.0027	0.0107	0.2412	0.0027	0.0105	0.2465
45	0.0020	0.0088	0.2049	0.0019	0.0087	0.1993	0.0025	0.0101	0.2303	0.0025	0.0100	0.2365
50	0.0020	0.0088	0.2040	0.0019	0.0087	0.1986	0.0023	0.0097	0.2199	0.0023	0.0096	0.2270

Source: Own calculations based on vector returns, covariance, co-skewness, and co-kurtosis of the best four stocks of the LQ45 index February-July 2023 period using R 3.6.1.

4. Discussions

Table 1 shows the four stocks used in this research, namely BBRI, ACES, BRIS, and ASII stocks, which are the four stocks with the best Sharpe ratio out of the 45 LQ45 stocks for February-July 2023. The Sharpe ratios for these stocks are 0.1303, 0.1208, 0.033, and 0.0731, respectively. Table 1 also shows that the four stocks have skewness and kurtosis, with BBRI stock having the smallest skewness, namely 0.0802, and the stock with the highest kurtosis is BRIS stock, 10.2879. Meanwhile, Table 2 shows the normality of the stocks used in this research. From Table 2 it can be seen that all the stocks used are not normally distributed. According to the literature review presented in Section 1, it is not enough to use the

first and second moments (mean and covariance) in forming a portfolio; it must also involve the third and fourth moments (skewness and kurtosis).

Figure 1 shows a distribution plot of data used in a multivariate manner. From Figure 1, it can be seen that the multivariate data used in this research contains outliers. Based on the literature review discussed in Section 1, an estimate resistant to outliers, namely a robust estimate, must be used to obtain a good estimate of the mean and covariance matrix. This study used two robust estimations for the mean and covariance matrix, namely, the FMCD robust and the CM robust estimations.

Table 3 shows the portfolio weighting using the classic MV portfolio model. Table 3 shows that the stock with the smallest Sharpe ratio (ASII) has a weight starting from negative (moving towards positive as the value γ increases). The same thing also happens to stocks with the highest Sharpe ratio (BBRI). Meanwhile, the weight of the other two stocks, namely ACES and BRIS, decreases as the value γ increases. The same thing also happens in portfolio weighting using the MVSK model, as seen in Table 4.

Table 5 shows the portfolio weighting using the robust MV_{FMCD} portfolio model. Table 5 shows that the stock with the smallest Sharpe ratio has the smallest (negative) weight at $\gamma = 0.5$, and moves towards positive as the value of γ increases, reaching a positive weight at $\gamma = 45$. The same thing also happened to the other three stocks. The weighting using the robust MV_{CM} portfolio model is in line with the weighting using the robust MV_{FMCD} portfolio model, the results of which can be seen in Table 6.

The return is not the only factor to consider when evaluating the performance of a portfolio, the risks also need to be taken into account. The Sharpe ratio is one of several metrics that can be used to assess portfolio performance. Table 7 shows the returns, risks, and Sharpe ratios of portfolios formed using the MV classic, MVSK, and robust MV portfolio model. From Table 7, it can be seen that the portfolio's performance using the robust portfolio model outperforms the portfolio using the MV classic portfolio model and the MVSK portfolio model. Furthermore, from Table 7, it can also be seen that for $\gamma = 0.5$ to $\gamma = 20$, the performance of the portfolio using the robust MV_{FMCD} model outperforms the performance of the portfolio using the robust MV_{CM} model. However, for $\gamma > 20$, the portfolio performance using the robust MV_{CM} model outperforms the portfolio using the robust MV_{FMCD} model.

5. Conclusions and Future Research

This paper presents two developments of the MV Markowitz portfolio model, namely the MVSK portfolio model and the robust MV portfolio model. In the robust MV portfolio model, there are two robust estimations used, namely the robust FMCD estimation and the robust CM estimation, which then produces two robust portfolio models, namely the robust MV_{FMCD} model and the robust MV_{CM} portfolio model. The portfolios formed using these models are then compared in performance using the Sharpe ratio. The research results show that the performance portfolio formed using the robust MV model outperforms those formed using the MV classic and MVSK models. Furthermore, it is also found that for $\gamma = 0.5$ to

$\gamma = 20$, the portfolio using the robust MV_{FMCD} model outperforms the portfolio using the robust MV_{CM} model. However, for $\gamma > 20$, the portfolio using the robust MV_{CM} model outperforms the portfolio using the robust MV_{FMCD} model.

For future research, there is the possibility of constructing portfolio models that involve skewness, kurtosis, and the presence of outliers in the data.

Acknowledgements: *For allowing us to use their facilities to conduct this research, we are grateful to the Head of the Mathematics Laboratory at Halu Oleo University's Faculty of Mathematics and Natural Sciences.*

References

- [1] Agustina, D., Sari, D.P., Winanda, R.S., Hilmi, M.R., Fakhriyana, D. (2022), *Comparison of Portfolio Mean-Variance Method with the Mean-Variance-Skewness-Kurtosis Method in Indonesia Stocks*. *Eksakta*, 22(2), 88-97.
- [2] Bengtsson, C. (2004), *The Impact of Estimation Error on Portfolio Selection for Investor with Constant Relative Risk Aversion*. Working Paper, Department of Economics Lund University.
- [3] Best, M.J., Grauer, R.R. (1991), *On the Sensitivity of Mean-Variance Efficient Portfolios to Changes in Asset Means: Some Analytical and Computational Results*. *Review of Financial Studies*, 4(2), 315-342.
- [4] Ceria, S., Stubbs, R.A. (2006), *Incorporating Estimation Errors into Portfolio Selection: Robust Portfolio Construction*. *Journal of Asset Management*, 7(2), 109-127.
- [5] Chen, B., Zhong, J., Chen, Y. (2020), *A Hybrid Approach for Portfolio Selection with Higher Order Moments: Empirical Evidence from Shanghai Stock Exchange*, *Expert Systems with Applications*, 145, 1-26.
- [6] Chopra, V.K., Ziemba, W.T. (1993), *The Effects of Errors in Means, Variances, and Covariances on Optimal Portfolio Choice*. *Journal of Portfolio Management*, 19(2), 6-11.
- [7] De Miguel, V., Nogales, F.J. (2008), *Portfolio Selection with Robust Estimation*. *Journal of Operation Research*, 57, 560-577.
- [8] Díaz, A., Esparcia, C., López, R. (2022), *The Diversifying Role of Socially Responsible Investments During the COVID-19 Crisis: A Risk Management and Portfolio Performance Analysis*. *Economic Analysis and Policy*, 75, 39-60.
- [9] Filzmoser, P., Garrett, R.G., Reimann, C. (2005), *Multivariate Outlier Detection in Exploration Geochemistry*. *Computers & Geosciences*, 31, 579-587.
- [10] Ghahtarani, A., Najafi, A.A. (2018), *Robust Optimization in Portfolio Selection by m-MAD Model Approach*. *Economic Computation and Economic Cybernetics Studies and Research*, 52(1), 279-291.
- [11] Gnanadesikan, R. (1977), *Methods for The Statistical Data Analysis of Multivariate Observations*. Wiley, New York.
- [12] Gotoh, J., Kim, M.J., Lim, A.E.B. (2018), *Robust Empirical Optimization is Almost the Same as Mean-Variance Optimization*. *Operations Research Letters*, 46(4), 448-452.

- [13] Gubu, L., Rosadi, D., Abdurrahman (2020), *Robust Mean-Variance Portfolio Selection with Ward and Complete Linkage Clustering Algorithm*. *Economic Computation and Economic Cybernetics Studies and Research*, 54(3), 111-127.
- [14] Gubu, L., Rosadi, D., Abdurrahman (2021), *A New Approach for Robust Mean-Variance Portfolio Selection Using Trimmed k-Means Clustering*. *Industrial Engineering & Management Systems*, 20(4), 782-794.
- [15] Hu, J. (2012), *An Empirical Comparison of Different Approaches in Portfolio Selection*. *Project Report, Uppsala Universitet*.
- [16] Kent, J.T., Tyler, D.E. (1996), *Constrained M-Estimation for Multivariate Location and Scatter*. *Journal Annals of Statistic*, 24(3), 1346-1370.
- [17] Khan, K.I., Waqar, S.M., Naqvi, A., Ghafoor, M.M. (2020), *Sustainable Portfolio Optimization with Higher-Order Moments of Risk*. *Sustainability*, 12(5), 1-14.
- [18] Kusch, P. (2012), *Portfolio Optimization with Robust Mean-Variance and Mean-Conditional Value at Risk Strategies*. *Project Report, Department of Economics Umea University*.
- [19] Lai, K.K., Yu, L., Wang, S. (2006), *Mean-Variance-Skewness-Kurtosis-based Portfolio Optimization*. *First International Multi-Symposiums on Computer and Computational Sciences (IMSCCS'06)*, 2, 292-297.
- [20] Lopuhaa, H.P. (1989), *On the Relation Between S-estimators and M-Estimators of Multivariate Location and Covariance*. *The Annals of Statistics*, 17, 1662-1683.
- [21] Lu, X., Liu, Q., Xue, F. (2019), *Unique Closed-Form Solutions of Portfolio Selection Subject to Mean-Skewness-Normalization Constraints*. *Operations Research Perspectives*, 6, 1-15.
- [22] Markowitz, H.M. (1952), *Portfolio Selection*. *Journal of Finance*, 7, 77-91.
- [23] Maronna, R.A., Martin, R.D., Yohai, V.J. (2006), *Robust Statistics: Theory and Methods*. *John Wiley and Sons*.
- [24] Marques, J.M.E., Benasciutti, D. (2020), *More on Variance of Fatigue Damage in Non-Gaussian Random Loadings-Effect of Skewness and Kurtosis*. *Procedia Structural Integrity*, 25(1), 101-111.
- [25] Metaxiotis, K. (2019), *A Mean-Variance-Skewness Portfolio Optimization Model*. *International Journal of Computer and Information Engineering*, 13(2), 85-88.
- [26] Naqvi, B., Mirza, N., Naqvi, W.A., Rizvi, S.K.A. (2017), *Portfolio Optimisation with Higher Moments of Risk at the Pakistan Stock Exchange*, *Economic Research-Ekonomska Istrazivanja*, 30(1), 1594-1610.
- [27] Rousseeuw, P.J., Van Driessen, K. (1999), *A Fast Algorithm for The Minimum Covariance Determinant Estimator*. *Technometric*, 41, 212 - 223.
- [28] Sharpe, W.F. (1994), *The Sharpe Ratio*. *The Journal of Portfolio Management*, 21, 49-58.
- [29] Supandi, E.D. (2017), *Developing of Mean-Variance Portfolio Modeling Using Robust Estimation and Robust Optimization Method*. *Ph.D. Dissertation, Mathematics Department Gadjah Mada University, Indonesia*.