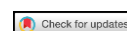




ROBUST MEAN-VARIANCE PORTFOLIO SELECTION USING CLUSTER ANALYSIS: A COMPARISON BETWEEN KAMILA AND WEIGHTED K-MEAN CLUSTERING



La Gubu¹
 Dedi Rosadi^{2*}
 Abdurakhman³

¹Department of Mathematics, Gadjah Mada University, Yogyakarta, Indonesia; Department of Mathematics, Haluoleo University, Kendari, Indonesia.

¹Email: lagubu2014@gmail.com Tel: +62-81341636559

^{2,3}Department of Mathematics, Gadjah Mada University, Yogyakarta, Indonesia.

²Email: dedirosadi@gadjahmada.edu Tel: +62-81328471898

³Email: rachmanstai@ugm.ac.id Tel: +62-81578978557



(+ Corresponding author)

ABSTRACT

Article History

Received: 30 June 2020
Revised: 4 August 2020
Accepted: 1 October 2020
Published: 19 October 2020

Keywords

KAMILA clustering
Weighted k-means clustering
Robust estimation
FMCD estimation
S estimation
Outliers
Portfolio optimization.

JEL Classification:

G11.

This study presents robust portfolio selection using cluster analysis of mixed-type data. For this empirical study, the daily price data of LQ45 index stocks listed on the Indonesia Stock Exchange were employed. First, six stocks clusters are formed by using the KAMILA algorithm on a combination of continuous and categorical variables. For comparison purposes, weighted k-means cluster analysis was also undertaken. Second, stocks that were representative of each cluster, those with the highest Sharpe ratios, were selected to create a portfolio. The optimum portfolio was determined through classic (non-robust) and the robust estimation methods of fast minimum covariance determinant (FMCD) and S estimation. Using a robust procedure enables the best-performing portfolio to be created efficiently when selecting assets from a large number of stocks, especially as the results are largely unaffected in the presence of outliers. This study found that the performance of the portfolio developed with the KAMILA clustering algorithm and robust FMCD estimation outperformed those created by other methods.

Contribution/Originality: This study presents robust portfolio selection using cluster analysis of mixed-type data, with the possible presence of outliers. The results reveal the portfolio developed with the KAMILA clustering algorithm and robust FMCD estimation outperformed portfolios created by other methods.

1. INTRODUCTION

Portfolio management is one issue that attracts the interest and attention of financial researchers, mainly in relation to developing portfolios of the best securities to maximize investors' profits. The fundamental theory of portfolio optimization can be traced back to the Markowitz (1952), who advised the selection and allocation of investments based on mean-variance analysis.

More recently, practitioners seek efficient time and cost management in creating the optimum portfolio, which can be achieved by first applying such techniques as cluster analysis to selecting securities. This technique has been undertaken by several researchers (e.g., Guan and Jiang, 2007; Tola, Lillo, Gallegati, and Mantegna, 2008; Chen and Huang, 2009; Nanda, Mahanty, and Tiwari, 2010; Long, Wisitpongphan, Meesad, and Unger, 2014), using different methods of cluster analysis and for selecting securities, although all adopting the Markowitz mean-

variance (MV) portfolio model. All the studies reported that cluster analysis to be very efficient in creating an optimum portfolio when there is a large number of securities are available for selection.

However, the MV portfolio model depends on mean vectors and variance–covariance matrices being estimated from highly volatile data, while there were a range of parameter estimation techniques available, which inevitably contain estimation errors. As an important factor in forming the MV portfolio model, estimation errors will therefore affect portfolio optimization significantly. Several studies have been conducted on the relationship between estimation errors and portfolio optimization (Best and Grauer, 1991; Broadie, 1993; Chopra and Ziemba, 1993; Bengtsson, 2004; Ceria and Stubbs, 2006), concluding that although the MV portfolio model had a strong theoretical basis and was relatively easy to compute, it contained weaknesses. For instance, the optimum portfolios created were not well diversified but concentrated within a specific sector. In addition, the model is highly sensitive to changes in mean vectors and variance–covariance matrices.

Therefore, some researchers have investigated robust portfolios, which reduce the effect of outliers on the estimated vector means and the variance–covariance matrices. Several studies have adopted the standard robust estimation approach to developing an optimum robust portfolio is (Victoria-Feser, 2000; Lauprete, 2001; Vaz De Melo and Camara, 2003; Zhou, 2006; Welsch and Zhou, 2007; DeMiguel and Nogales, 2008; Hu, 2012; Kusch, 2012; Supandi, 2017). These studies have used different robust estimates to optimize the portfolio, but all reported that robust portfolios perform better than the classical one in the presence of outliers. However, there has been no research into combining cluster analysis and robust estimation in creating an optimum portfolio. In addition, cluster analysis uses only continuous data, stock return data, excluding categorical data such as sector, listing board, and market capitalization of stocks.

This study bridges this gap in the literature, therefore, by applying the KAMILA algorithm to cluster analysis of mixed data: continuous and categorical, the results of which are combined with robust estimation to create the optimum portfolio. Fast minimum covariance determinant (FMCD) and S estimation were used to calculate the mean and covariance of the mixed data.

The rest of this paper is organized as follows: Section 2 we explains concepts necessary to the discussion; Section 3 presents the empirical results; Section 4 discusses these results; and Section 5 details the conclusions reached.

2. MATERIALS AND METHOD

Following the literature review, it became evident that the problem of selecting the best securities can be resolved more efficiently by clustering stocks and then selecting stock clusters to optimize the portfolio.

In the current study, the KAMILA algorithm was first applied to a combination of continuous—stock prices and trading volumes—and categorical—sector, listing board, and market capitalization of stocks—data to produce several clusters of stocks. Thereafter, the returns and risk are calculated for each stocks cluster from the historical data, followed by the Sharpe ratio to determine the performance of every stock within each cluster. Those stocks with the highest Sharpe ratio are then selected to represent each cluster in creating the optimum portfolio. Finally, FMCD and S estimation were used to determine the weighting of each stock in the portfolio by providing a robust estimation of the mean and variance of those stocks.

A comparison was undertaken between the performances of the portfolios produced by both the KAMILA algorithm and weighted k-means clustering, which both employed the same robust estimation method, to determine the advantages of the former.

The subsequent subsections summarized other concepts related to this study.

2.1. Mean-Variance Portfolio

Markowitz's portfolio theory is based on the mean–variance approach, in which the mean measures the level of the expected return and variance measures the level of risk (Markowitz, 1952); thus, Markowitz's mean–variance (MV) was produced, emphasizing the maximization of expected returns and minimization of risk in developing the optimum portfolio. According to Supandi (2017), the mean–variance portfolio can be created by resolving the following optimization problems:

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

$$\mathbf{w}'\mathbf{e} = 1 \quad (2)$$

where \mathbf{w} indicates the weighting of the portfolio, $\boldsymbol{\mu}$ the mean vector, $\boldsymbol{\Sigma}$ the covariance matrix, \mathbf{e} the column matrix (all being 1), and $\gamma \geq 0$ the risk-aversion parameters, or the relative size of risk avoidance. Investors select a MV portfolio based on utility function criteria, which reflect their risk–averse attitudes and wish for the maximum expected utility $E(U)$. Thus, the MV portfolio can be expressed as:

$$\max E(U(\mathbf{w})) \quad (3)$$

$$\mathbf{w}'\mathbf{e} = 1 \quad (4)$$

Given an initial capital of W_0 , in a portfolio with a weight vector of \mathbf{w} , the capital at the end of the period becomes $W_0(1 + R_p)$, where R_p denotes portfolio returns (Supandi, 2017). The utility function of $W_0(1 + R_p)$ is $U(W_0(1 + R_p))$, which according to Supandi (2017), can be expanded using Taylor's second-order approximation:

$$\begin{aligned} U(W_0(1 + R_p)) &= U(W_0) + U'(W_0)W_0R_p + \frac{1}{2}U''(W_0)(W_0R_p)^2 + \frac{1}{6}U'''(W_0)(W_0R_p)^3 + \dots \\ &= U(W_0) + W_0U'(W_0)R_p + \frac{1}{2}W_0^2U''(W_0)R_p^2 + O(R_p^3) \end{aligned} \quad (5)$$

Taking the expected value of Equation 5, the following approximation can be calculated:

$$\begin{aligned} E\left(U(W_0(1 + R_p))\right) &= U(W_0) + W_0U'(W_0)\mu_p + \frac{1}{2}W_0^2U''(W_0)\sigma_p^2 \\ &= U(W_0) + W_0U'(W_0)\left(\mu_p + \frac{1}{2}W_0\frac{U''(W_0)}{U'(W_0)}\sigma_p^2\right) \end{aligned} \quad (6)$$

Based on Equation 6, maximizing the expected utility function is equivalent to maximizing:

$$\mu_p - \frac{1}{2}\gamma\sigma_p^2 \quad (7)$$

where $\gamma = -W_0 \frac{U''(W_0)}{U'(W_0)}$ expresses the relative size of risk aversion (Supandi, 2017).

Since $\mu_p = \mathbf{w}'\boldsymbol{\mu}$ and $\sigma_p^2 = \mathbf{w}'\boldsymbol{\Sigma}\mathbf{w}$, Equation 7 can be rewritten as:

$$\mathbf{w}'\boldsymbol{\mu} - \frac{\gamma}{2}\mathbf{w}'\boldsymbol{\Sigma}\mathbf{w} \quad (8)$$

Therefore, Equation 1 is equivalent to Equation 8, with the constraint of $\mathbf{w}'\mathbf{e} = 1$.

Nevertheless, every investor accepts a certain amount of risk to obtain a certain level of return. As profits compensate for risks, investors should balance the profits and risks by choosing the right γ . There are two extreme situations. First, where investors wish to increase returns (profits) and reduce risk (losses), $\gamma = 0$; Equation 1 provides the maximum rate of return without regard to the risk. Second, where investors opt for the minimum amount of risk regardless of the level of return, $\gamma = \infty$.

The optimization problems shown in Equations 1 and 2 can be resolved through the Lagrange method (Winston & Goldberg, 2004). First, the Lagrange function is formed:

$$L = \mathbf{w}'\boldsymbol{\mu} - \frac{1}{2}\gamma\mathbf{w}'\boldsymbol{\Sigma}\mathbf{w} + \lambda(\mathbf{w}'\mathbf{e} - 1) \quad (9)$$

Based on the Kuhn–Tucker theorem (Winston & Goldberg, 2004), the necessary conditions for the optimal of Equation 9 are:

$$\frac{\partial L}{\partial \mathbf{w}} = 0 \quad (10)$$

$$\frac{\partial L}{\partial \lambda} = 0 \quad (11)$$

From Equations 9–11, the following is derived:

$$\mathbf{w} = \frac{\boldsymbol{\Sigma}^{-1}}{\gamma}(\boldsymbol{\mu} + \lambda\mathbf{e}) \quad (12)$$

and

$$\mathbf{e}'\mathbf{w} = 1 \quad (13)$$

Substituting Equation 12 into Equation 13 gives:

$$\lambda = \gamma(\mathbf{e}'\boldsymbol{\Sigma}^{-1}\mathbf{e})^{-1} - (\mathbf{e}'\boldsymbol{\Sigma}^{-1}\mathbf{e})^{-1}\mathbf{e}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} \quad (14)$$

Then, substituting Equation 14 into Equation 12 produces:

$$\begin{aligned}
\mathbf{w} &= \frac{\Sigma^{-1}}{\gamma} \{ \boldsymbol{\mu} + \mathbf{e}[\gamma(\mathbf{e}'\Sigma^{-1}\mathbf{e})^{-1} - (\mathbf{e}'\Sigma^{-1}\mathbf{e})^{-1}\mathbf{e}'\Sigma^{-1}\boldsymbol{\mu}] \} \\
&= \frac{1}{\gamma} (\Sigma^{-1} - \Sigma^{-1}\mathbf{e}(\mathbf{e}'\Sigma^{-1}\mathbf{e})^{-1}\mathbf{e}'\Sigma^{-1})\boldsymbol{\mu} + \Sigma^{-1}\mathbf{e}(\mathbf{e}'\Sigma^{-1}\mathbf{e})^{-1}
\end{aligned} \tag{15}$$

Consequently, Equation 15 shows that the optimum portfolio \mathbf{w} depends on inputs $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$, and also assumes multivariate normality.

2.2. Cluster Analysis

Cluster analysis is a statistical method that groups together data objects sharing similar characteristics, aiming for within-group homogeneity, or at least as small a variation between the data objects as possible. A brief overview of the two types of cluster analysis, or segmentation, used in this study is now provided.

2.2.1. KAMILA Algorithm

KAMILA is actually the acronym for k-means for mixed large data sets, meaning that the KAMILA algorithm is a development of the k-means clustering method for use with a mixture of continuous and categorical data. There are several other clustering methods capable of dealing with mixed-type data, but there are drawbacks with each: strong parametric assumptions, such as the Gaussian multinomial mixture models (Foss and Markatou, 2018); cannot minimize or maximize the contribution of individual variables, such as Modha–Spangler weighting (Foss, Markatou, Ray, and Heching, 2016); or based on arbitrary weighting to determine the contribution of continuous and categorical variables, such as dummy coding and Gower's distance (Foss and Markatou, 2018).

The KAMILA algorithm, however, combines the two best clustering methods for large mixed-type data: k-means clustering method and Gaussian multinomial mixture models. KAMILA and k-means are similar in not needing parametric assumptions for continuous data; however, KAMILA holds the advantage because it can deal with different levels of overlap for individual variables, unlike K-means that relies on arbitrary weighting to calculating the Euclidean distance at different overlap levels.

2.2.1.1. Model

Supposing a data set of N independent and identically distributed (i.i.d.) observations of a $(P + Q)$ dimensional vector of random variables $(\mathbf{V}^T, \mathbf{W}^T)^T$, which may be dependent, in a finite mixture distribution with G components, where \mathbf{V} is a dimensional vector of the P continuous random variable and \mathbf{W} a Q vector of the categorical random variable, and where the q th element of \mathbf{W} has L_q categorical levels, with $q = 1, 2, \dots, Q$.

Given cluster membership in the g th cluster, Foss and Markatou (2018) state that, under the local independence assumption, the joint density of $(\mathbf{V}^T, \mathbf{W}^T)^T$ is:

$$f_{\mathbf{V}, \mathbf{W}, g}(\mathbf{v}, \mathbf{w}; \mu_g, \Sigma_g, \theta_{gq}) = f_{\mathbf{V}, g}(\mathbf{v}; \mu_g; \Sigma_g) \prod_{q=1}^Q m(w_q; \theta_{gq})$$

while overall density not dependent on cluster membership is:

$$f_{\mathbf{V}, \mathbf{W}}(\mathbf{v}, \mathbf{w}) = \sum_{g=1}^G \pi_g f_{\mathbf{V}, \mathbf{W}, g}(\mathbf{v}, \mathbf{w}; \mu_g, \Sigma_g, \theta_{gq}) \quad (16)$$

where π_g denotes the previous probability of observing the g th cluster.

2.2.1.2. Radial Kernel Density Estimation

Foss et al. (2016) states that if $\mathbf{X} \in \mathbb{R}^p$ is in a spherically symmetric distribution with center μ , then:

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{f_R(r) \Gamma(\frac{p}{2} + 1)}{p r^{p-1} \pi^{\frac{p}{2}}} \quad (17)$$

where $r = \sqrt{(\mathbf{x} - \mu)^T (\mathbf{x} - \mu)}$, $R = \sqrt{(\mathbf{X} - \mu)^T (\mathbf{X} - \mu)}$, and f_R is the probability density of R .

Foss and Markatou (2018) proceeded to construct \hat{f}_R using (univariate) kernel density estimation, which then substitutes \hat{f}_R in Equation 17. Univariate kernel density estimation avoids the drawbacks of being too expensive to calculate and tending to overfit data points (Foss & Markatou, 2018) found in multivariate kernel density estimation (Scott, 1992).

2.2.1.3. Algorithm Description

KAMILA introduces an iterative process to estimate the unknown parameters in Equation 16 (Foss et al., 2016). Foss and Markatou (2018) explained that the iterative process consisted of two stages: partition and estimation. Partition assigns each observation to a cluster, while estimation re-estimates the parameters of interest using these new clusters.

At the t th iteration of the algorithm, $\hat{\mu}_g^{(t)}$ denotes the estimator for the centroid (center) of population g , and

$\hat{\theta}_{gq}^{(t)}$ denotes the estimator for the multinomial distribution parameters corresponding to the q th discrete random variable drawn from population g . The parameter estimates are initialized using either uniformly random samples from the observed data points or centroids selected through another mechanism (e.g., a preliminary clustering round).

Given $\hat{\mu}_g^{(t)}$, g , and $\hat{\theta}_{gq}^{(t)}$ at the t th iteration, Foss and Markatou (2018) defined the Euclidean distance from

observation i to each $\hat{\mu}_g^{(t)}$ as $d_{ig}^{(t)} = \sqrt{\sum_{p=1}^P [(v_{ip} - \hat{\mu}_{gp}^{(t)})]^2}$. The minimum distance for the i th observation is

then calculated as $r_i^{(t)} = \min_g (d_{ig}^{(t)})$. Finally, the kernel density estimation of that minimum distances is

expressed as:

$$\hat{f}_R^{(t)}(r) = \frac{1}{Nh^{(t)}} \sum_{i=1}^N k\left(\frac{r-r_i^{(t)}}{h^{(t)}}\right) \tag{18}$$

where $k(\cdot)$ is the kernel function and $h^{(t)}$ the corresponding bandwidth at iteration t . Foss and Markatou (2018)

used the Gaussian kernel with a bandwidth of $h = 0.9An^{-1/5}$, where $A = \min(\hat{\sigma}, \hat{q}/1.34)$, in which $\hat{\sigma}$ is the

sample standard deviation and \hat{q} the sample interquartile range. The $\hat{f}_R^{(t)}(r)$ function here is used to construct

$\hat{f}_V^{(t)}(r)$ as described in Section 2.2.1.2.

Foss and Markatou (2018) assumed that the Q categorical variables are independent within a given population

g (i.e., local independence), and calculated the probability of observing the i th vector of categorical variables within

a given population as $c_{ig}^{(t)} = \prod_{q=1}^Q m(w_{iq}; \hat{\theta}_{gq}^{(t)})$, where $m(\cdot; \cdot)$ is the multinomial probability mass function.

The i th object was thus assigned to population g , which maximized the function as follows:

$$H_i^{(t)} = \log\left(\hat{f}_V^{(t)}\left(d_{ig}^{(t)}\right)\right) + \log\left(c_{ig}^{(t)}\right) \tag{19}$$

In each iteration t , the latest partition of the N observations was used to calculate $\hat{\mu}_{gp}^{(t+1)}$ and $\hat{\theta}_{gq}^{(t+1)}$ for g, p , and

q . If $\Omega_g^{(t)}$ denotes the index set of observations assigned to population g at iteration t , then the parameters are

estimated:

$$\hat{\mu}_g^{(t+1)} = \frac{1}{|\Omega_g^{(t)}|} \sum_{i \in \Omega_g^{(t)}} v_i$$

$$\hat{\theta}_{gql}^{(t+1)} = \frac{1}{|\Omega_g^{(t)}|} \sum_{i \in \Omega_g^{(t)}} I\{w_{iq} = l\}$$

where $I\{\cdot\}$ denotes the indicator function and $|A|$ the cardinality of set A .

The KAMILA R package applies the simple rule of stopping once group membership remains unchanged between iterations.

2.2.2. Weighted K-Means Clustering

Weighted k-means clustering develops k-means clustering by adding a weight parameter (Kerdprasop, Kerdprasop, & Sattayatham, 2005) and assigning a different weight to each data point. Similar to k-means, weighted k-means clustering partitions N samples into k clusters, in which centroid \mathbf{c}_j minimizes the number of squares in the cluster, the within-cluster sum of squares (WCSS), which is defined as:

$$WCSS = \sum_{i=0}^N \min_{\mu_j \in \mathcal{C}} \left(w(x) \|\mathbf{x}_j - \mathbf{c}_j\|^2 \right) \quad (20)$$

where $\|\mathbf{x}_j - \mathbf{c}_j\|^2$ is the Euclidean distance between the sample and centroid—the center, or sample mean—of

the cluster \mathcal{C} and $w(x)$ the weight of the sample X (Kerdprasop et al., 2005).

The weighted k-means and k-means algorithms are the same, as follows (Kerdprasop et al., 2005):

1. Determine the desired number of k clusters.
2. Initialize k cluster centroids.
3. Allocate data objects to the nearest cluster based on the centroid produced by Equation 21:

$$\mathbf{c}_j = \frac{1}{\sum_{x \in \mathcal{C}} w(x)} \sum_{x \in \mathcal{C}} x w(x) \quad (21)$$

4. Reallocate each data point after every iteration until the centroid no longer changes (i.e., Equation 20 is satisfied), at which point cluster analysis is complete.

2.3. Sharpe Ratio

Once the clusters are formed, the performance of every stock within each cluster is assessed using the Sharpe ratio, also known as the Sharpe index, which measures excess return (or risk premium) per unit risk in an asset (Sharpe, 1994): it characterizes how well asset returns compensate investors for the risks they have taken. The

Sharpe ratio (SR) is calculated by comparing the difference between stock returns (R) and risk return free rate (R_f)

with a standard deviation of stock return (σ), which is expressed as follows:

$$SR = \frac{R - R_f}{\sigma} \quad (22)$$

In general, the higher the value of a stock's Sharpe ratio, the better the stock's performance.

2.4. Outlier Detection

Outliers are data points that are a significant distance from the majority of other observations and the cluster centroid, or deviate from the general pattern of data in some way. The higher the value of an outlier, the greater its distance from the centroid—outliers typically lie at large distances (Filzmoser, Garrett, & Reimann, 2005). However, not only the distance between an observation and the centroid but also the overall shape of the data should be considered with multivariate data.

The shape and size of multivariate data are quantified by the covariance matrix, which the Mahalanobis distance measure accommodates. For a p -dimensional multivariate sample $\mathbf{x}_1, \dots, \mathbf{x}_n$, the Mahalanobis distance is defined as:

$$MD_i = ((\mathbf{x}_i - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x}_i - \boldsymbol{\mu}))^{1/2} \quad \text{for } i = 1, \dots, n \quad (23)$$

where $\boldsymbol{\mu}$ is the estimated multivariate location and $\boldsymbol{\Sigma}$ the estimated covariance matrix (Filzmoser et al., 2005); usually, $\boldsymbol{\mu}$ is the multivariate arithmetic mean, and $\boldsymbol{\Sigma}$ the sample covariance matrix. For normally distributed multivariate data, the MD_i^2 values indicate the approximate chi-square distribution with p degrees of freedom (χ_p^2). By setting the (squared) Mahalanobis distance as equal to a certain constant (i.e., a certain quantile of χ_p^2), it is possible to define ellipsoids with the same Mahalanobis distance from the centroid (Gnanadesikan, 1977).

Multivariate outliers are those observations with a large (squared) Mahalanobis distance; moreover, with multivariate data, a quantile of the chi-squared distribution (e.g., the 98% quantile $\chi_{p;0.98}^2$) can also be considered an outlier (Filzmoser et al., 2005).

2.5. Portfolio Selection Using Robust Estimation

In this study, the robust FMCD and robust S estimation methods were used to determine the weighting of the stocks selected for the optimum portfolio. Both methods will now be briefly described.

2.5.1. Robust FMCD Estimation

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

observations, where $\frac{(n+p+1)}{2} \leq h \leq n$ with:

$$\hat{\boldsymbol{\mu}} = \frac{1}{h} \sum_{i=1}^h r_i \quad (24)$$

Thus, the covariance matrix can be estimated by solving the following equation:

$$\hat{\Sigma} = \frac{1}{h} \sum_{i=1}^h (r_i - \hat{\mu})(r_i - \hat{\mu})' \quad (25)$$

However, MCD estimation becomes complicated as data dimensions increase, due to all possible subsets of h observations from a total of n observations being examined. Therefore, Rousseeuw and Driessen (1999) developed a faster algorithm, FMCD, which is based on the C-step theorem. In this theorem, if H_1 is the set of size h from the data of size n , the sample statistics would be:

$$\hat{\mu}^1 = \frac{1}{h} \sum_{i \in H_1} r_i \quad (26)$$

$$\hat{\Sigma}^1 = \frac{1}{h} \sum_{i \in H_1} (r_i - \hat{\mu}^1)(r_i - \hat{\mu}^1)' \quad (27)$$

Here, if $|\hat{\Sigma}^1| > 0$, then distance $d_i = (r_i; \hat{\mu}^1, \hat{\Sigma}^1)$. Then, specify H_2 as a subset containing the observation with the smallest distance d_i : $\{d_1(i) | i \in H_2\} = \{(d_1)_1, \dots, (d_1)_h\}$, where $(d_1)_1 \leq (d_1)_2 \leq \dots \leq (d_1)_n$ is a sequential distance. Based on H_2 , Equations 25 and 26 produces:

$$|\hat{\Sigma}^2| \leq |\hat{\Sigma}^1| \quad (28)$$

Equation 28 will be the same as though $\hat{\mu}^1 = \hat{\mu}^2$ and $\hat{\Sigma}^1 = \hat{\Sigma}^2$.

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

2.5.2. Robust S-Estimation

Robust S estimation was first introduced by Rousseeuw and Yohai (1984), and later developed by Davies (1987) and Lopuhaa (1989). Davies' (1987) definition stated that given $\{r_i, i = 1, \dots, n\}$ is the data set in \mathbb{R}^p and \mathbf{P}_p is a set of symmetric matrices of size $p \times p$, the S estimation for location $\hat{\mu} \in \mathbb{R}^p$ and dispersion $\hat{\Sigma}(R) \in \mathbf{P}_p$ is a pairing of $\hat{\mu}$ and $\hat{\Sigma}(R)$ that minimizes $|\Sigma|$ with the following condition:

$$\frac{1}{n} \sum_{i=1}^n \rho[(r_i - \mu)' \Sigma^{-1} (r_i - \mu)]^{1/2} = b_0 \quad (29)$$

where ρ is the loss function and b_0 the constant. The constant must be determined precisely because its value affects the estimation result; therefore, if the distribution of the data is unknown, then the constant is set as $b_0 = E\{\rho\|r\|\}$.

S estimation is calculated iteratively by the following equations:

$$\frac{1}{n} \sum_{i=1}^n u(d_i)(r_i - \mu) = 0 \quad (30)$$

$$\frac{1}{n} \sum_{i=1}^n p u(d_i)(r_i - \mu)(r_i - \mu)' - v(d_i)\Sigma = 0 \quad (31)$$

where $d_i = (r_i - \mu)' \Sigma^{-1} (r_i - \mu)$, $\psi(d_i) = \frac{\partial \rho}{\partial d}$, $u(d_i) = \psi(d_i)/d_i$, while

$$v(d_i) = \psi(d_i)d_i - \rho(d_i) + b_0.$$

According to Hardin (2000) the algorithm for S estimation is:

1. Determine the initial estimation of the mean vector $\hat{\mu}_0$ and covariance matrix $\hat{\Sigma}_0$.
2. Calculate $d_i = (r_i - \hat{\mu}_0)' \hat{\Sigma}_0^{-1} (r_i - \hat{\mu}_0)$.
3. Determine k_0 so that $\frac{\sum \rho(d_i/k_0)}{n} = b_0$.
4. Calculate $\tilde{d}_i = \frac{d_i}{k_0}$.
5. Determine $\hat{\mu} = \frac{\sum \psi(\tilde{d}_i) r_i}{\sum \psi(\tilde{d}_i)}$ and $\hat{\Sigma} = \frac{p \sum \psi(\tilde{d}_i) (r_i - \mu)(r_i - \mu)'}{\sum \psi(\tilde{d}_i)}$.
6. Repeat steps 2–3 until $\hat{\mu}$ and $\hat{\Sigma}$ converge.

3. RESULTS AND DISCUSSION

3.1. Stocks Clustering

The current study used the daily prices of all stocks in the LQ45 index for the Indonesia Stock Exchange (see e.g., <https://finance.yahoo.com>). For the weighted k-means clustering, the weightings of the continuous and categorical variables was varied: first, combining (0.8, 0.5, 0.2) for the continuous variables, and then, (0.2, 0.5, 0.8) for the categorical variables. The kamila and wkmeans functions in the R package (R Core Team, 2020) revealed that the LQ-45 index stocks could be grouped into six clusters, which are presented in Tables 1–4.

Table-1. Stocks clusters using KAMILA algorithm.

Cluster	Stocks									
1	AALI	AKRA	BMTR	LPPF	LSIP	MNCN	SSMS			
2	ADRO	ANTM	BSDE	BUMI	INCO	MYRX	PTBA	PWON	WSKT	
3	BBCA	GGRM	UNTR	UNVR						
4	ASII	EXCL	HMSP	ICBP	JSMR	PGAS	TLKM			
5	ADHI	BBNI	BBRI	BBTN	BJBR	BMRI	LPKR	PPRO	PTPP	SCMA
	SMRA	SRIL	WIKA							
6	BRPT	INDF	INTP	KLBF	SMGR					

Table-2. Stocks clusters using weighted k-means 0.8 : 0.2.

Cluster	Stocks									
1	LPKR	AKRA	AALI	MNCN	PPRO	LSIP	ADHI	PTPP	SSMS	WIKA
	SMRA	LPPF	BMTR	SRIL	BJBR					
2	BUMI	TLKM	PWON	MYRX						
3	BBTN	INDF	KLBF	BSDE	SMGR	JSMR	INTP	EXCL	SCMA	PGAS
	WSKT	BRPT								
4	UNVR	GGRM	UNTR							
5	ASII	BMRI	BBNI	ICBP	BBRI	HMSP	BBCA			
6	ADRO	ANTM	PTBA	INCO						

Table-3. Stocks clusters using weighted k-means 0.5 : 0.5.

Cluster	Stocks									
1	LPKR	AKRA	AALI	MNCN	PPRO	LSIP	ADHI	PTPP	SSMS	WIKA
	SMRA	LPPF	BMTR	SRIL	BJBR					
2	BUMI	PWON	MYRX							
3	INDF	KLBF								
4	BSDE	SMGR	JSMR	INTP	EXCL	SCMA	PGAS	BBTN	WSKT	BRPT
	ADRO									
5	ASII	TLKM	UNVR	BMRI	BBNI	ICBP	BBRI	GGRM	HMSP	UNTR
	BBCA									
6	ANTM	PTBA	INCO							

Table-4. Stocks clusters using weighted k-means 0.2 : 0.8.

Cluster	Stocks									
1	LPKR	AKRA	AALI	MNCN	PPRO	LSIP	ADHI	PTPP	SSMS	WIKA
	SMRA	LPPF	BMTR	SRIL	BJBR					
2	PWON	MYRX								
3	INDF	KLBF	BSDE	JSMR	EXCL	SCMA	PGAS	BBTN	WSKT	ADRO
4	SMGR	INTP	BRPT							
5	ASII	TLKM	UNVR	BMRI	BBNI	ICBP	BBRI	GGRM	HMSP	UNTR
	BBCA									
6	BUMI	ANTM	PTBA	INCO						

3.2. Representative Stocks from Clusters

Following the formation of the six clusters, the Sharpe ratio was calculated for every stock in each cluster, using the latest Bank Indonesia rate of 5.25% per year as the risk return free rate. Based on these calculations, stocks that were representative of each cluster were identified to create the optimum portfolio, as presented in Tables 5–8.

Table-5. Representative stocks from clusters using KAMILA algorithm.

Cluster	Representation	Return	Risk	Sharpe Ratio
1	BMTR	-0.000164	0.029596	-0.010416
2	INCO	0.002580	0.026729	0.091156
3	BBCA	0.000878	0.012855	0.057126
4	HMSP	0.000664	0.021462	0.024240
5	BBRI	0.000311	0.018511	0.009023
6	BRPT	0.000425	0.022963	0.012221

Table-6. Representative stocks from clusters using weighted k-means 0.8 : 0.2.

Cluster	Representation	Return	Risk	Sharpe Ratio
1	BJBR	0.000225	0.022459	0.003619
2	MYRX	0.000428	0.032915	0.008622
3	BRPT	0.000425	0.022963	0.012221
4	UNTR	0.000805	0.022303	0.029621
5	BBCA	0.000878	0.012855	0.057126
6	INCO	0.002580	0.026729	0.091156

Table-7. Representative stocks from clusters using weighted k-means 0.5 : 0.5.

Cluster	Representation	Return	Risk	Sharpe Ratio
1	BJBR	0.000225	0.022459	0.003619
2	MYRX	0.000428	0.032915	0.008622
3	KLBF	-0.000857	0.018216	-0.054946
4	ADRO	0.000490	0.028072	0.012363
5	BBCA	0.000878	0.012855	0.057126
6	INCO	0.002580	0.026729	0.091156

Table-8. Representative stocks from clusters using weighted k-means 0.2 : 0.8.

Cluster	Representation	Return	Risk	Sharpe Ratio
1	BJBR	0.000225	0.022459	0.003619
2	MYRX	0.000428	0.032915	0.008622
3	ADRO	0.000490	0.028072	0.012363
4	BRPT	0.000425	0.022963	0.012221
5	BBCA	0.000878	0.012855	0.057126
6	INCO	0.002580	0.026729	0.091156

3.3. Detection of Outliers in Representative Stocks from Clusters

The Mahalanobis distance, introduced in Section 2.4, was used to determine the outliers in the representative stocks. The mahalanobis and qchisq functions in the R package were used, with a 97.5% threshold for the distance, on 260 daily return data of the LQ45 index stocks for the period August 2017–July 2018. Table 9 presents the number of outliers for the representative stocks identified by the four clustering methods, while Figure 1 depicts the outliers from KAMILA clustering.

Table-9. Number of outliers in representative stocks from clusters.

Clustering Method	Number of Outliers	Percentage
KAMILA	16	6.15
Weighted k-means 0.8 : 0.2	15	5.77
Weighted k-means 0.5 : 0.5	18	6.92
Weighted k-means 0.2 : 0.8	18	6.92

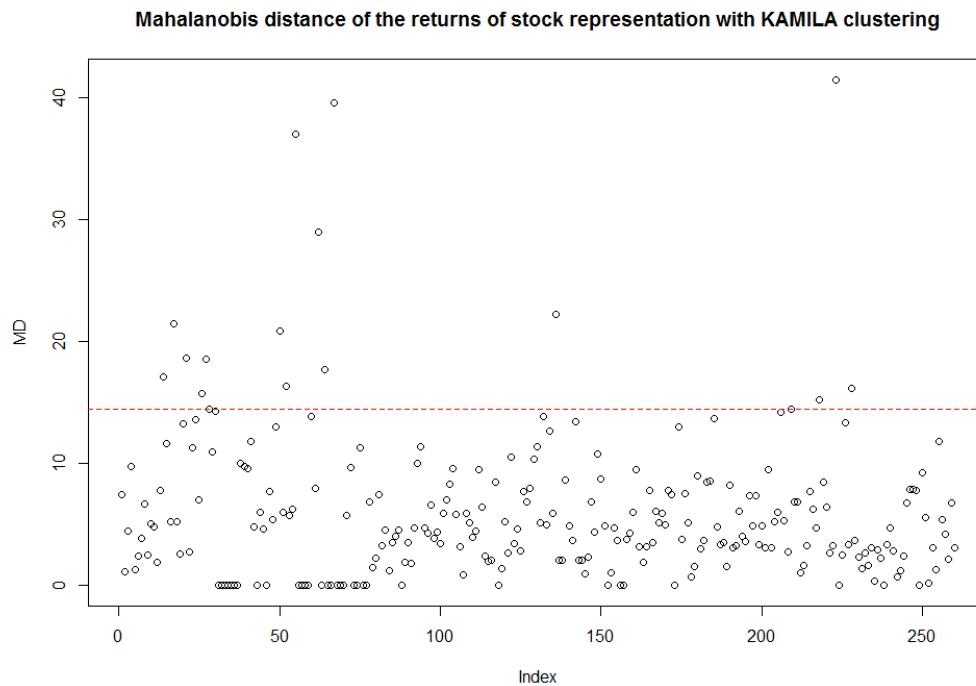


Figure-1. Outliers of Stocks Representation with KAMILA Clustering.

3.4. Portfolio Weightings and Comparison of Portfolio Performance

The optimum portfolio in this study was determined by the MV portfolio model based on both robust FMCD (MV_{FMCD}) and S (MV_S) estimations. The first stage is to set the portfolio weightings in each estimation model for various risk aversion values γ , using the **CovMcd** and **CovSest** functions in the R package. For comparison purposes, the same procedure was undertaken for the optimum portfolio determined by the classic MV portfolio model ($MV_{Classic}$). The optimum portfolio comprised the stocks that were representative of each cluster, as presented in Tables 1–4, and the resulting portfolio weightings are presented in Tables 10–13.

Table-10. Portfolio weight of MV classic and robust portfolio using KAMILA algorithm.

Model	γ	BMTR	INCO	BBCA	HMSP	BBRI	BRPT
$MV_{Classic}$	0.5	-2.169794	5.180605	3.730208	-0.081726	-3.980619	-1.678674
	1	-1.048586	2.647133	2.134911	0.019339	-1.966693	-0.786105
	2	-0.487982	1.380397	1.337263	0.069873	-0.959730	-0.339821
	5	-0.151619	0.620355	0.858674	0.100193	-0.355553	-0.072050
	10	-0.039498	0.367008	0.699144	0.110299	-0.154160	0.017207
MV_{FMCD}	0.5	-12.892920	8.580200	-2.769171	4.651188	3.405259	0.025444
	1	-6.441513	4.295077	-1.020066	2.364864	1.756319	0.045319
	2	-3.215808	2.152516	-0.145514	1.221703	0.931849	0.055256
	5	-1.280387	0.866979	0.379217	0.535805	0.437167	0.061218
	10	-0.635246	0.438467	0.554128	0.307173	0.272273	0.063206
MV_S	0.5	-6.369058	8.047326	-0.269468	3.788822	-2.865004	-1.332618
	1	-3.168092	4.048475	0.152108	1.934055	-1.349535	-0.617010
	2	-1.567609	2.049049	0.362896	1.006671	-0.591801	-0.259207
	5	-0.607319	0.849394	0.489368	0.450241	-0.137160	-0.044524
	10	-0.287222	0.449509	0.531526	0.264764	0.014387	0.027036

Table-11. Portfolio weight of MV classic and robust portfolio using weighted k-means 0.8 : 0.2.

Model	γ	BJBR	MYRX	BRPT	UNTR	BBCA	INCO
	0.5	-2.486973	-0.543940	-2.232721	-1.622755	1.852398	6.033991
	1	-1.151905	-0.243585	-1.065875	-0.772379	1.184499	3.049245
<i>MV_{Classic}</i>	2	-0.484370	-0.093408	-0.482452	-0.347191	0.850550	1.556872
	5	-0.083850	-0.003301	-0.132399	-0.092079	0.650180	0.661448
	10	0.049657	0.026734	-0.015714	-0.007041	0.583390	0.362974
	0.5	-11.310331	-6.042790	-2.089360	6.315137	6.636202	7.491142
	1	-5.492350	-2.983140	-1.019847	3.182170	3.566533	3.746633
<i>MV_{FMCD}</i>	2	-2.583359	-1.453315	-0.485091	1.615686	2.031699	1.874379
	5	-0.837964	-0.535419	-0.164237	0.675796	1.110799	0.751026
	10	-0.256166	-0.229454	-0.057286	0.362499	0.803832	0.376575
	0.5	-8.918575	-10.173089	-0.645023	4.678080	9.867617	6.190990
	1	-4.308953	-5.028871	-0.271714	2.348799	5.161014	3.099725
<i>MV_S</i>	2	-2.004142	-2.456762	-0.085059	1.184159	2.807713	1.554092
	5	-0.621256	-0.913497	0.026934	0.485374	1.395733	0.626712
	10	-0.160294	-0.399075	0.064265	0.252446	0.925072	0.317585

Table-12. Portfolio weight of MV classic and robust portfolio with weighted k-means 0.5 : 0.5.

Model	γ	BJBR	MYRX	KLBF	ADRO	BBCA	INCO
	0.5	-0.727967	-0.002456	-11.693996	-2.334734	8.864260	6.894892
	1	-0.273182	0.028654	-5.782720	-1.164226	4.697497	3.493977
<i>MV_{Classic}</i>	2	-0.045789	0.044209	-2.827081	-0.578973	2.614116	1.793519
	5	0.090646	0.053542	-1.053698	-0.227820	1.364087	0.773244
	10	0.136125	0.056653	-0.462571	-0.110770	0.947411	0.433153
	0.5	-14.226208	-10.046229	9.691180	0.234813	8.346215	7.000230
	1	-6.981442	-4.968158	4.846326	0.115790	4.479496	3.507989
<i>MV_{FMCD}</i>	2	-3.359060	-2.429122	2.423899	0.056278	2.546136	1.761868
	5	-1.185630	-0.905701	0.970443	0.020571	1.386120	0.714196
	10	-0.461154	-0.397894	0.485958	0.008669	0.999448	0.364972
	0.5	-12.823344	-10.459752	2.031370	3.270296	12.501207	6.480223
	1	-6.265823	-5.164758	1.048243	1.635188	6.488352	3.258797
<i>MV_S</i>	2	-2.987062	-2.517261	0.556679	0.817635	3.481925	1.648084
	5	-1.019806	-0.928763	0.261741	0.327103	1.678069	0.681656
	10	-0.364054	-0.399263	0.163428	0.163592	1.076783	0.359514

Table-13. Portfolio weight of MV classic and robust portfolio with weighted k-means 0.2 : 0.8.

Model	γ	BJBR	MYRX	ADRO	BRPT	BBCA	INCO
	0.5	-2.502745	-0.449469	-2.233956	-2.397207	2.392938	6.190438
	1	-1.155284	-0.194802	-1.115651	-1.142252	1.470317	3.137671
<i>MV_{Classic}</i>	2	-0.481554	-0.067468	-0.556498	-0.514774	1.009006	1.611288
	5	-0.077316	0.008933	-0.221007	-0.138288	0.732220	0.695458
	10	0.057430	0.034399	-0.109176	-0.012792	0.639957	0.390181
	0.5	-10.543469	-11.993346	2.789438	3.930087	11.139218	5.678072
	1	-5.124550	-5.947702	1.423867	1.993446	5.805781	2.849158
<i>MV_{FMCD}</i>	2	-2.415091	-2.924881	0.741082	1.025126	3.139062	1.434702
	5	-0.789415	-1.111188	0.331410	0.444134	1.539031	0.586028
	10	-0.247523	-0.506624	0.194853	0.250469	1.005688	0.303136
	0.5	-10.905638	-10.208993	4.415846	1.437402	9.354582	6.906801
	1	-5.301508	-5.046395	2.208353	0.765807	4.907432	3.466311
<i>MV_S</i>	2	-2.499443	-2.465096	1.104607	0.430009	2.683856	1.746066
	5	-0.818204	-0.916316	0.442359	0.228531	1.349711	0.713919
	10	-0.257791	-0.400056	0.221610	0.161371	0.904996	0.369870

Based on the portfolio weightings, as well as the mean vectors and covariance matrices, the Sharpe ratio was calculated for the three portfolio models, as presented in Table 14.

Table-14. The Sharpe ratio of MV classic portfolio and robust portfolios with KAMILA algorithm and weighted k-means.

γ	KAMILA			Weighted k-means								
	Classic	FMCD	S	Weight = (0.8 : 0.2)			Weight = (0.5 : 0.5)			Weight = (0.2 : 0.8)		
				Classic	FMCD	S	Classic	FMCD	S	Classic	FMCD	S
0.5	0.0883	0.1967	0.1356	0.0855	0.1574	0.1588	0.1316	0.1914	0.1873	0.0886	0.1755	0.1773
1	0.0921	0.1946	0.1349	0.0890	0.1556	0.1568	0.1333	0.1887	0.1855	0.0921	0.1734	0.1754
2	0.0986	0.1901	0.1330	0.0948	0.1519	0.1526	0.1360	0.1830	0.1817	0.0980	0.1689	0.1713
5	0.1081	0.1751	0.1244	0.1037	0.1392	0.1383	0.1394	0.1648	0.1686	0.1071	0.1542	0.1573
10	0.1062	0.1468	0.1051	0.1022	0.1151	0.1115	0.1337	0.1324	0.1434	0.1058	0.1269	0.1310

4. DISCUSSIONS

Having formed six stocks clusters from the LQ45 index through KAMILA and weighted k-means clustering algorithms, shown in Tables 1–4, the Sharpe ratio was calculated for every stock in each cluster to identify the stocks that were representative of each cluster, presented in Tables 5. Those representative stocks with the highest Sharpe ratios, those identified as the best performing in the cluster, were selected for inclusion in the portfolio. Thus, following KAMILA clustering algorithm, Table 5 shows that BMTR stocks, with a Sharpe ratio of -0.010416 , were chosen from Cluster 1, and INCO, BBCA, HMSP, BBRI, and BRPT stocks for Clusters 2, 3, 4, 5, and 6, respectively. The same process was followed for the weighted k-means clusterings, the results of which can be seen in Tables 6–8.

Table 9, and Figure 1 in relation to KAMILA clustering, indicates the number of outliers found for each clustering method: 16 (6.15%) using KAMILA clustering, and 15 (5.77%), 18 (6.92%), and 18 (6.92%) for weighted k-means 0.8 : 0.2, 0.5 : 0.5, and 0.2 : 0.8 clustering, respectively. Based on these results, it is reasonable to employ a robust estimation method to create optimum portfolios.

All the clustering methods produced stocks with negative returns (e.g., Table 10 shows negative weighting for BMTR stock), denoting short selling, in all four portfolio models, across every risk aversion value γ (see Tables 11–13). In contrast, stocks with large returns are always shown with positive weightings in all four portfolio models (e.g., INCO stocks in Table 10). Table 10 also illustrates that the greater the γ value, the smaller the weightings of the stocks.

When assessing portfolio performance, though, the risks taken by the investor should also be considered. One way to measure both returns and risks is to calculate the Sharpe ratio, which is shown in Table 14 for all four portfolio models developed through KAMILA and weighted k-means clusterings. These results reveal that the MV portfolio based on KAMILA clustering and robust FMCD estimation outperformed the other alternatives in this empirical study.

5. CONCLUSIONS

The current study demonstrated how to integrate clustering techniques for mixed-type data—continuous and categorical—into portfolio management and create an optimum portfolio. To establish the advantages of this proposed method, the performance of the optimum portfolio was compared to that of the optimum portfolio created through weighted k-means clustering, which can also be applied to mixture-type data. The final results revealed that the best portfolio performance was obtained by combining KAMILA clustering with robust FMCD estimation.

Funding: This study received no specific financial support.

Competing Interests: The authors declare that they have no competing interests.

Acknowledgement: The first author wishes to thank the Indonesia Endowment Fund for Education (Lembaga Pengelola Dana Pendidikan; LPDP) from the Indonesian Ministry of Finance for the scholarship funding of his Doctoral Program at the Faculty of Mathematics and Natural Sciences of Gadjah Mada University. The authors also acknowledge the financial support from PDD 2020 Program from Deputy of Research and Development, Ministry of Research and Technology / National Agency for Research and Innovation of Republic of Indonesia.

REFERENCES

- 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.
- 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. *The Review of Financial Studies*, 4(2), 315-342. Available at: <https://doi.org/10.1093/rfs/4.2.315>.
- Broadie, M. (1993). Computing efficient frontiers using estimated parameters. *Annals of Operations Research*, 45(1), 21-58. Available at: <https://doi.org/10.1007/bf02282040>.
- Ceria, S., & Stubbs, R. A. (2006). Incorporating estimation errors into portfolio selection: Robust portfolio construction. *Journal of Asset Management*, 7(2), 109-127. Available at: <https://doi.org/10.1057/palgrave.jam.2240207>.
- Chen, L.-H., & Huang, L. (2009). Portfolio optimization of equity mutual funds with fuzzy return rates and risks. *Expert Systems with Applications*, 36(2), 3720-3727. Available at: <https://doi.org/10.1016/j.eswa.2008.02.027>.
- 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.
- Davies, P. L. (1987). Asymptotic behaviour of S -estimates of multivariate location parameters and dispersion matrices. *The Annals of Statistics*, 15(3), 1269-1292. Available at: <https://doi.org/10.1214/aos/1176350505>.
- DeMiguel, V., & Nogales, F. J. (2008). Portfolio selection with robust estimation. *Journal of Operation Research*, 57(3), 560 - 577.
- Filzmoser, P., Garrett, R. G., & Reimann, C. (2005). Multivariate outlier detection in exploration geochemistry. *Computers & Geosciences*, 31(5), 579-587. Available at: <https://doi.org/10.1016/j.cageo.2004.11.013>.
- Foss, A. H., & Markatou, M. (2018). kamila: clustering mixed-type data in R and Hadoop. *Journal of Statistical Software*, 83(1), 1-44.
- Foss, A., Markatou, M., Ray, B., & Heching, A. (2016). A semiparametric method for clustering mixed data. *Machine Learning*, 105(3), 419-458. Available at: <https://doi.org/10.1007/s10994-016-5575-7>.
- Gnanadesikan, R. (1977). *Methods for the statistical data analysis of multivariate observations*. New York: Wiley.
- Guan, H. S., & Jiang, Q. S. (2007). *Cluster financial time series for portfolio*. Paper presented at the Int. Conf. on Wavelet Analysis and Pattern Recognition, ICWAPR, Beijing, China.
- Hardin, J. S. (2000). *Multivariate outlier detection and robust clustering with minimum covariance determinant estimation and s-estimation*. Dissertation. University of California.
- Hu, J. (2012). An empirical comparison of different approaches in portfolio selection. Project Report 2012:7. Uppsala Universitet.
- Kerdprasop, K., Kerdprasop, N., & Sattayatham, P. (2005). Weighted K-means for density-biased clustering. In: Tjoa A.M., Trujillo J. (eds) Data Warehousing and Knowledge Discovery. DaWaK 2005. Lecture Notes in Computer Science (Vol. 3589). Berlin, Heidelberg: Springer.
- Kusch, P. (2012). Commodities or not commodities. Portfolio optimization with robust mean-variance and mean-conditional value at risk strategies. Project Report. Umea University Department of Economics.
- Lauprete. (2001). *Portfolio risk minimization under departures from normality*. Dissertation. Massachusetts Institute of Technology.

- Long, N. C., Wisitpongphan, N., Meesad, P., & Unger, H. (2014). Clustering stock data for multi-objective portfolio optimization. *International Journal of Computational Intelligence and Applications*, 13(02), 1-13. Available at: <https://doi.org/10.1142/s1469026814500114>.
- Lopuhaa, H. P. (1989). On the relation between S-estimators and M-estimators of multivariate location and covariance. *The Annals of Statistics*, 17(4), 1662-1683. Available at: <https://doi.org/10.1214/aos/1176347386>.
- Markowitz, H. M. (1952). Portfolio selection. *Journal of Finance*, 7(1), 77-91.
- Nanda, S., Mahanty, B., & Tiwari, M. (2010). Clustering Indian stock market data for portfolio management. *Expert Systems with Applications*, 37(12), 8793-8798. Available at: <https://doi.org/10.1016/j.eswa.2010.06.026>.
- R Core Team. (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Retrieved from: <https://www.R-project.org/>.
- Rousseeuw, P. J., & Driessen, K. V. (1999). A fast algorithm for the minimum covariance determinant estimator. *Technometrics*, 41(3), 212-223. Available at: <https://doi.org/10.1080/00401706.1999.10485670>.
- Rousseeuw, P., & Yohai, V. (1984). Robust regression by means of S-estimators. In *Robust and Nonlinear Time Series Analysis*. Lecture Notes in Statistics (Vol. 26, pp. 256 - 272). New York: Springer.
- Scott, D. W. (1992). *Multivariate density estimation: Theory, practice, and visualization*. New York: John Wiley & Sons Inc.
- Sharpe, W. F. (1994). The sharpe ratio. *Journal of Portfolio Management*, 21(1), 49-58. Available at: <https://doi.org/10.3905/jpm.1994.409501>.
- Supandi, E. D. (2017). *Developing of mean-variance portfolio modeling using robust estimatin and robust optimization method*. Ph.D. Dissertation. Mathematics Department Gadjah Mada University, Indonesia.
- Tola, V., Lillo, F., Gallegati, M., & Mantegna, R. N. (2008). Cluster analysis for portfolio optimization. *Journal of Economic Dynamics and Control*, 32(1), 235-258.
- Vaz De Melo, B., & Camara, R. P. (2003). Robust multivariate modelling in finance. Working Paper Series 355, Federal University at Rio de Janeiro, Rio de Janeiro, Brazil.
- Victoria-Feser, M. P. (2000). Robust portfolio selection. Research Report. Universite of Geneve.
- Welsch, R. E., & Zhou, X. (2007). Application of robust statistics to asset allocation models. *REVSTAT-Statistical Journal*, 5(1), 97-114.
- Winston, W. L., & Goldberg, J. B. (2004). *Operations research: Applications and algorithms* (4th ed.). Canada: Thomson Learning Inc.
- Zhou, X. (2006). *Application of robust statistics to asset allocation models*. Thesis. Massachusetts Institute of Technology.

