Current Science Research Bulletin
Volume 02 (09), September 2025, pp. 276-283
CrossRef DOI: 10.55677/csrb/05-V02I09Y2025
Available on: https://csrbjournal.org/
e-ISSN: 3050-7103
p-ISSN: 3050-709X

Formation of Mean-Semivariance Portfolio with Stock Selection Using Fuzzy C-Means Clustering on the LQ45 Stock Index

Viny Debora Sandri[1], Suparti[2*], Rita Rahmawati[3]
[1,2,3]Departement of Statistics, Faculty of Science and Mathematics, Universitas Diponegoro

ABSTRACT: Forming an optimal portfolio with a focus on minimizing risk can use the Mean-Semivariance method, which is accompanied by stock selection through clustering analysis with Fuzzy C-Means. The Fuzzy C-Means aims to group stocks based on certain characteristics. The data used are financial ratio data, namely Earning per Share (EPS), Price to Earning Ratio (PER), and Return on Equity (ROE) as criteria in the process of selecting stock securities with cluster analysis, and daily return data used for the formation of optimal stocks within the LQ45 Index. The results showed that the best number of clusters was 3 and representative stock securities of each cluster are selected based on the highest expected return value, namely UNTR and BRIS. The optimal portfolio obtained by the Mean-Semivariance method produces a weight value of 63.91% in UNTR stock securities and 36.09% in BRIS stock securities. Portfolio performance with the Sharpe index is 0.088966, which shows good portfolio performance results, and Value at Risk (VaR) calculated by Historical Simulation shows at a 95% confidence level for holding periods of 1, 7, and 30 days, respectively, it is 0.021543, 0.056998, and 0.117996.

KEYWORDS: Portofolio Optimal, Fuzzy C-Means, Silhouette Coefficient, Mean-Semivariance, Indeks Sharpe, VaR

1. INTRODUCTION

Stock investments offer high profit potential but come with significant risks that many investors overlook due to a limited understanding of risk measurement. One effective strategy for minimizing risk is to create an optimal portfolio using the Mean-Semivariance method, an extension of Harry Markowitz's Mean-Variance model. Unlike variance, semivariance focuses only on negative return fluctuations, making it more suitable for risk-averse investors. To handle a large number of stocks, cluster analysis spesifically Fuzzy C-Means (FCM) can be employed. FCM allows for overlapping cluster membership and more accurate center placement, enhancing the portfolio selection process.

Previous research conducted by Hasibuan et al. (2021) found that FCM outperformed Single Linkage clustering with LQ45 data, while Gubu et al. (2023) reported a higher Sharpe ratio for FCM (0.17615) compared to K-Means (0.13240). Pakungwati et al. (2024) also showed that the Mean-Semivariance approach led to optimal mutual fund portfolio performance. This study aims to form an optimal portfolio using the Mean-Semivariance method combined with FCM clustering on consistently listed LQ45 stocks. Clustering result will be validated using the Sillhouette Coefficient, with stock selection based on the highest expected return per cluster. The portfolio's risk is measured using Value at Risk (VaR) via Historical Simulation, and performance is evaluated using the Sharpe Ratio.

2. LITERATURE REVIEW

The Number Of Stocks In The Capital Market Is Quite Large, Which Can Affect The Performance Of The Resulting Portfolio. The More Stocks That Are Formed Into A Portfolio, The More Difficult It Will Be To Determine The Investment Proportion Of Each Stock. Cluster Analysis Is An Effective Alternative In Solving This Problem. This Statistical Analysis Is Able To Group Objects Into Different Groups Based On Their Characteristics.

Data Values With Different Scales, If Not Standardized, Can Cause The Analysis Process To Be Inaccurate Because The Algorithm Is Influenced By Variables With A Larger Range (Amorim Et Al., 2022). One Way To Standardize Data Is To Use The Max Absolute Scaler Method. This Method Is Useful For Maintaining The Relationship Between The Original Data And Retaining The Influence Of Outliers.

$$x_{ij_{scaled}} = \frac{x_{ij}}{|\max(x_j)|} \quad (1)$$

Corresponding Author: Suparti

Where x_{ij} is the data value of the i-th object in the j-th variable, i: 1,2, … , n with n as the number of objects, and j: 1,2, … , m with m indicating the number of variables..

There are two cluster assumptions, namely:

1. Representative samples are analyzed using the Kaiser Meyer Olkin (KMO) value. KMO is used to measure sample validity by comparing the correlation coefficient value to the partial correlation, which can be calculated mathematically using Equation (2).

$$KMO = \frac{\sum_{j=1}^{m}\sum_{l=1,l\neq j}^{m} r_{x_j x_l}^2}{\sum_{j=1}^{m}\sum_{l=1,l\neq j}^{m} r_{x_j x_l}^2 + \sum_{j=1}^{m}\sum_{l=1}^{m} \rho_{x_j x_l, x_h}^2} \quad (2)$$

Where $r_{x_j x_l}$ is the correlation between variables x_j, x_l and $\rho_{x_j x_l, x_h}$ is the partial correlation between variables x_j and x_l while controlling for other variables (x_h)

The Value of KMO > 0,5 indicates that the sample is sufficiently representative, so cluster analysis can proceed.

2. Non-multicollinearity indicates that the variables used should not be perfectly correlated with each other as this can result in inaccurate cluster analysis. The assumption of non-multicollinearity can be fulfilled when the value of $VIF_j < 10$. The formula for calculating the VIF value is as follows (Gujarati & Porter, 2010).

$$VIF_j = \frac{1}{(1-R_j^2)} \quad (3)$$

This study uses Euclidean distance as a measure of distance. According to Johnson and Wichern (2007), Euclidean distance is defined as the distance between two objects, for example objects i and r, which are located in m-dimensional space. Systematically, this distance is written as Equation (4).

$$d_{(x_i, x_r)} = \sqrt{\sum_{j=1}^{m}(x_{ij} - x_{rj})^2} \quad (4)$$

Where $d_{(x_i, x_r)}$ is the Euclidean distance between object sequence i and sequence r, x_{ij} is the data of object sequence i on variable j, and x_{rj} is the data of object sequence r on variable j.

This study uses Fuzzy C-Means (FCM) cluster analysis. Gubu et al. (2023) state that the basic concept of the clustering process using FCM begins with identifying cluster centers by marking the average position for each cluster. In the initial stage, the cluster center positions may not be accurate, and each data point has a certain membership in each cluster. Therefore, the cluster centers and the membership of each point are updated repeatedly until the cluster centers move to the appropriate positions. The purpose of the iteration process is to produce a minimal objective function that represents the distance between the point and the cluster center, taking into account the weight based on the membership level.

The following are the steps in Fuzzy C-Means analysis (Bezdek, 1981):

1. Inputting data in the form of an n x m matrix.
2. Determining the initial parameters, namely the number of clusters (c), fuzzy weights ($w > 1$), iteration number (t = 1), maximum iterations (MaxIter), minimum error value (ε), and initial objective function (Po = 0).
3. Generate random numbers (μ_{ik}) with *i* = 1,2, … , n and k = 1,2, … , c being the number of clusters.

$$\sum_{k=1}^{c} \mu_{ik} = 1 \quad (5)$$

4. Determine the center of cluster k using the following equation:

$$V_{kj} = \frac{\sum_{i=1}^{n}(\mu_{ik})^w x_{ij}}{\sum_{i=1}^{n}(\mu_{ik})^w}; 1 \leq i \leq n; 1 \leq j \leq m; 1 \leq k \leq c \quad (6)$$

5. Measure the objective function for iteration t.

$$P_t = \sum_{i=1}^{n}\sum_{k=1}^{c}\left(\left[\sum_{j=1}^{m}(x_{ij} - V_{kj})^2\right](\mu_{ik})^w\right) \quad (7)$$

6. Calculate the change in the membership matrix.

$$\mu_{ik} = \frac{\left[\sum_{j=1}^{m}(x_{ij}-V_{kj})^2\right]^{\frac{-1}{w-1}}}{\sum_{k=1}^{c}\left[\sum_{j=1}^{m}(x_{ij}-V_{kj})^2\right]^{\frac{-1}{w-1}}}; 1 \leq i \leq n; 1 \leq j \leq m; 1 \leq k \leq c \quad (8)$$

7. Evaluate the termination condition

- If ($|P_t - P_{t-1}| < \varepsilon$) or t ≥ *MaxIter* then the iteration stops.
- If ($|P_t - P_{t-1}| > \varepsilon$) or t < *MaxIter* then t = t + 1, repeat calculating the cluster center.

Corresponding Author: Suparti

The results of the clusters that have been formed will be validated using the Silhouette Coefficient, which is formulated as follows (Struyf et al., 1997).

$$SC = \frac{1}{n}\sum_{i=1}^{n} s(i) \quad (9)$$

$$\text{Where } s(i) = \frac{b(i)-a(i)}{\max(a(i),b(i))},\ a(i) = \frac{1}{|\boldsymbol{A}|-1}\sum_{j\epsilon \boldsymbol{A}, j\neq i} d_{(x_i,x_r)},\ b(i) = \min_{\boldsymbol{C}\neq \boldsymbol{A}} d(\boldsymbol{x}_i, \boldsymbol{C})$$

$a(i)$: distance from object i to all objects in the same cluster

$b(i)$: minimum distance value from object i to other objects in clusters other than A

The Silhouette Coefficient value ranges from -1 to 1, where values closer to 1 indicate better clustering results.

After obtaining the cluster results, the return value of each stock will be calculated to form an optimal portfolio. According to Hartono (2017), realized return refers to the profits that have been achieved and obtained from past data. One way to calculate realized return is by continuously compounding return. The advantage of using this return calculation is the ease of multi-period return (Pakungwati et al., 2024).

$$R_{i,t} = ln\left|\frac{p_{it}}{p_{i(t-1)}}\right| \quad (10)$$

where:

$R_{i,t}$: stock return value in period t

p_{it} : closing price of the stock in period t

$p_{i(t-1)}$: closing price of the stock in period (t-1)

Hartono (2017) states that expected return can be used in investment decision making. Expected return is an estimate of the profit that investors expect to earn.

$$\mu_i = E(R_i) = \frac{1}{T}\sum_{t=1}^{T} R_{i,t} \quad (11)$$

where μ_i : expected return of stock i and T : time period

Furthermore, it is also necessary to calculate the expected return of the portfolio by calculating the weighted mean value of the expected return of individual stocks (single assets) in a portfolio.

$$\mu_p = E(R_p) = E(\sum_{i=1}^{N} w_i R_i) = \sum_{i=1}^{N} w_i E(R_i) = \sum_{i=1}^{N} w_i \mu_i \quad (12)$$

where:

μ_p : expected return of the portfolio

w_i : weight of stock i

N : total number of stocks

According to Hartono (2017), an optimal portfolio has a good combination of expected return and risk, where it can provide a certain expected return while minimizing risk. Mean-Semivariance was a method for forming a more optimal portfolio in minimizing risk. Semivariance measures the risk associated with below-average values (downside risk) of expected returns by taking into account the negative fluctuations of an asset or stock. This shows that semivariance focuses on value declines (losses), thereby producing a better portfolio than using variance. Markowitz (1959) proposed a formula for calculating semivariance and semicovariance of a portfolio with benchmark b, but the proposed equation remains asymmetric and endogenous because portfolio weights influence periods when the portfolio underperforms the benchmark and the calculation of semivariance-semicovariance. Estrada (2008) developed a more effective formula for calculating semivariance and semicovariance as follows:

$$\varphi_p^2 = \frac{1}{T}\sum_{t=1}^{T}\left[Min(R_{i,t} - b_t, 0)^2\right] \quad (13)$$

$$\varphi_{ij} = \frac{1}{T}\sum_{t=1}^{T}\left[Min\left(R_{i,t} - b_t, 0\right) \times Min\left(R_{j,t} - b_t, 0\right)\right] \quad (14)$$

Where φ_p^2 is the portfolio semivariance and φ_{ij} is the semicovariance of stock i and stock j.

The risk value of the mean-semivariance portfolio is formulated as follows:

$$\varphi_p^2 = \sum_{i=1}^{N}\sum_{j=1}^{N} w_i w_j \varphi_{ij} \quad (15)$$

Ernitrisnasari (2015) states that the calculation of the mean-semivariance portfolio weights can use the weights $\boldsymbol{w} = [w_1, \ldots, w_N]^T$ with the aim of minimizing the semivariance, so that the optimization function with the constraint $0 \leq w_i \leq 1, i = 1, \ldots, N$ is changed to the following form:

minimize $\boldsymbol{w^T}\sum_{sv}\boldsymbol{w}$

Corresponding Author: Suparti

with the constraint $\boldsymbol{w}^T\mathbf{1}_N = 1$

where:

$\boldsymbol{w}$: weight vector

$\boldsymbol{w}^T$: transpose of w

$\sum_{SV}$: semivariance-semicovariance matrix

$\mathbf{1}_N$: column vector with elements 1

Next, the optimal weighting will be sought, which can be solved by minimizing the Lagrange function. The Lagrange function equation is written as follows:

$$L(\boldsymbol{w}, \lambda) = \boldsymbol{w}^T \textstyle\sum_{sv} \boldsymbol{w} + \lambda(1 - \boldsymbol{w}^T \mathbf{1}_N) \quad (16)$$

where L is the Lagrange function and λ is the Lagrange multiplier.

The Lagrange function is differentiated with respect to w to obtain the optimal weight using the final formula (Ernitrisnasari, 2015):

$$\boldsymbol{w} = \frac{\sum_{sv}{}^{-1}\mathbf{1}_N}{\mathbf{1}_N^T \sum_{sv}{}^{-1}\mathbf{1}_N} \quad (17)$$

where $\sum_{sv}{}^{-1}$: inverse of the semivariance-semicovariance matrix and $\mathbf{1}_N^T$: transpose of $\mathbf{1}_N$

The VaR measurement used is Historical Simulation, which utilizes past data from portfolio returns to calculate the VaR value. The formula used to calculate VaR using the Historical Simulation method is as follows (Maruddani & Trimono, 2020).

$$\text{VaR}_{\alpha,hp}^{HS}(R_p) = V_0 \times \delta \times \sqrt{hp} \quad (18)$$

where:

R_p : *portfolio return*

α : confidence level $(1-\alpha)$

hp : holding period or investment period

V_0 : initial capital

δ : the $(\alpha \times 100)$th percentile or α-th quantile of the sorted portfolio return data

The Sharpe index is a method used to analyze portfolio performance. Portfolio performance is considered better if the Sharpe Index value is higher, which is systematically written as follows (Adnyana, 2020).

$$S_r = \frac{\mathrm{E}(R_p) - \overline{R_f}}{\varphi_p} \quad (19)$$

where $\overline{R_f}$ is the *risk free rate* and φ_p is the portfolio's semideviation.

3. DATA AND METHODOLOGY

Data for the clustering process uses secondary data from the financial ratios of each stock, such as Earnings per Share (EPS), Price to Earnings Ratio (PER), and Return on Equity (ROE) in 2024, obtained from the website https://www.idx.co.id/ for active stocks included in the LQ45 index based on a major evaluation conducted by the Indonesia Stock Exchange. Data for portfolio formation uses secondary data from stock closing prices and the Composite Stock Price Index (IHSG) for active stocks in the LQ45 index during the period from January 2 to December 30, 2024, accessed via https://finance. yahoo.com/, as well as Bank Indonesia interest rates for the 2024 period, which are sourced from the website https://www.bi.go.id/id/statistik/indikator/BI-Rate.aspx.

1. Prepare EPS, PER, and ROE data based on stocks on the LQ45 index, LQ45 index closing price data, IHSG data, and BI rate data.
2. Eliminate stocks with negative EPS, PER, and ROE values.
3. Standardize EPS, PER, and ROE data using Max Absolute Scaler.
4. Test the assumptions of representative sample and non-multicollinearity cluster analysis.
5. Perform Fuzzy C-Means cluster analysis.
6. Validate the number of clusters (c = 2,3,4,5) formed using the Silhouette Coefficient.
7. Calculate the stock return value in the selected cluster based on the highest Silhouette Coefficient value.
8. Calculate the expected return value for each stock.
9. Removing stocks with expected return values ≤ 0.
10. Determining the best stock from each cluster based on the highest expected return value.
11. Forming the optimal portfolio using the Mean-Semivariance method.
    a. Calculating the semivariance value of each selected stock.
    b. Calculating the semicovariance between selected stocks.

Corresponding Author: Suparti

c. Calculating the weight of each selected stock using Mean-Semivariance.
d. Calculating the portfolio's return and risk.

12. Calculating the portfolio's VaR using the Historical Simulation method.
13. Measuring portfolio performance using the Sharpe Index.

4. RESULTS AND DISCUSSION

Descriptive statistics on the cluster variables EPS, PER, and ROE as attributes used in the clustering process with the Fuzzy C-Means method can be seen in Table 1.

Table 1. Descriptive Statistics of EPS, PER, and ROE

Variable	N	Minimum	Maximum	*Mean*	Standard Deviation
EPS	37	2.71	5590.87	548.713	1180.679374
PER	37	3.07	339.61	34.213	73.778746
ROE	37	0.41	105.40	15.262	16.923393

Based on Table 1, it can be seen that the EPS variable has a mean value of 548.713, which is a positive value. The results show that the average of the 37 stocks listed on the LQ45 index generated a profit of 548.713 rupiah per share. For the PER variable, the mean value is positive at 34.213, indicating that, on average, the 37 stocks have a stock value 34.213 times higher than the profit per share. The ROE variable has a positive mean value of 15.262, indicating that the average of the 37 stocks listed on the LQ45 index generates a profit of 15.262% of total equity.

In this study, using EPS, PER, and ROE as clustering variables, it is known that the data scales are different, so data standardization is required. The researcher used the Max Absolute Scaler method for data standardization. Data standardization was obtained with the help of Google Colaboratory software so that cluster analysis could be continued.

The assumptions for cluster analysis based on Google Colaboratory software processing are:

a. Representative Sample

Based on the obtained output, the results show that the overall KMO value is 0.540233. This value indicates that the assumption of a representative sample is met.

b. Non-Multicollinearity

The output results of the non-multicollinearity test show that the VIF values for EPS are 1.031425, PER are 1.086626, and ROE are 1.062081. This indicates that all data variables have VIF values below 10, indicating that there is no correlation between the variables EPS, PER, and ROE, and the assumption of non-multicollinearity is met.

After the cluster assumption is met, data clustering is performed using the Fuzzy C-Means method with a known number of clusters (c = 2,3,4,5), and the cluster results are validated using the Silhouette Coefficient as follows.

Table 2. Silhouette Coefficient Calculation Results

Number of Clusters (c)	Silhouette Coefficient
2	0.752578
3	**0.812212**
4	0.231534
5	0.253940

Based on the results in Table 2, a cluster analysis was obtained which showed that the highest silhouette coefficient value for a number of clusters compared to other numbers of clusters was found in the number of clusters (c) = 3 with the highest silhouette coefficient value of 0.812212.

Based on the best number of clusters, the members of each cluster can be seen as follows:

Table 3. Best Fuzzy C-Means Clustering Results for 3 Clusters

Cluster k	Stock Members
1	ARTO, BRPT
2	ITMG, UNTR

Corresponding Author: Suparti

3	ACES, ADRO, AKRA, AMRT, ANTM, ASII, BBCA, BBNI, BBRI, BBTN, BMRI, BRIS, CPIN, ESSA, EXCL, ICBP, INCO, INDF, INKP, INTP, KLBF, MAPI, MBMA, MEDC, MTEL, PGAS, PGEO, PTBA, SIDO, SMGR, TLKM, TOWR, UNVR

The final cluster center with c = 3 is:

$$\mathbf{V} = \begin{bmatrix} 0.001027 & 0.939982 & 0.008120 \\ 0.937337 & 0.016143 & 0.203216 \\ 0.052851 & 0.048363 & 0.152355 \end{bmatrix}$$

These values are still in standardized form. Below are the cluster center results in their original values:

$$\mathbf{V} = \begin{bmatrix} 5.25 & 325.855 & 0.84 \\ 5261.945 & 5.1 & 21.495 \\ 296 & 18.302 & 15.758 \end{bmatrix}$$

Based on the final cluster center results, the following information is obtained:

For LQ45 stocks in cluster 1, the average EPS is 5.25 rupiah, PER is 328.855, and ROE is 0.84%, The results indicate that LQ45 stocks have the ability to utilize capital to generate net profit and have the lowest earnings per share value but the highest stock value compared to stocks in other clusters.

For LQ45 stocks in Cluster 2, the average EPS is 5,261.945 rupiah, the P/E ratio is 5.1, and the ROE is 21.495%. The results indicate that LQ45 stocks have the ability to utilize capital to generate net profits and have the highest earnings per share value but the lowest stock value compared to stocks in other clusters.

For LQ45 stocks in cluster 3, the average EPS is 296 rupiah, a P/E ratio of 18.302, and a ROE of 15.758%. The results indicate that LQ45 stocks have the ability to utilize capital to generate net profit and have a high earnings per share value, as well as a moderate stock value, making them relatively high compared to stocks in other clusters.

Based on the analysis, the optimal number of clusters is c = 3, indicating that there are 3 stocks forming the portfolio. Stock representatives are selected based on high expected returns within each cluster. The highest expected returns for each cluster are presented in Table 4.

Table 4. Highest Expected Return Values for Each Cluster

Cluster k	Stock	Expected Return
1	ARTO	-0.000793
2	UNTR	0.000557
3	BRIS	0.001908

The results in Table 4 show that in cluster 1, ARTO shares have the highest expected return value of -0.000793, which means that these shares cannot be used as portfolio components because the negative expected return does not meet the requirements of the Mean-Semivariance method. The shares selected to form a Mean-Semivariance portfolio are UNTR and BRIS.

Table 5. Descriptive Statistics of Selected Stocks

No.	Stock	Expected Return	Skewness	Kurtosis	Standard Deviation
1	UNTR	0.000557	-0.279011	2.296183	0.016757
2	BRIS	0.001908	0.420716	1.544763	0.026778

Table 5 presents the descriptive statistics for UNTR and BRIS. The results show that the expected returns for both stocks are positive, with the highest value obtained by BRIS at 0.001908. It can also be seen that the skewness value of UNTR stock is negative, indicating that UNTR stock has a negative curve or is skewed to the left. Meanwhile, BRIS stock has a positive skewness value, meaning it has a positive curve or is skewed to the right. The kurtosis values of both stocks are less than three, meaning that UNTR and BRIS stocks have flat curves (platykurtic). The standard deviation of BRIS stock is higher than that of UNTR. This means that BRIS stock has a higher potential risk than UNTR.

Next, the semivariance value is calculated based on the stock return value and the IHSG as a benchmark using the formula $Min(R_{i,t} - b, 0)^2$ with the following results:

Table 6. Semivariance Value

Stock	Semivariance Value

Corresponding Author: Suparti

UNTR	0.000133
BRIS	0.000208

Next, calculate the semicovariance value using the formula $Min\left(R_{i,t} - b, 0\right) \times Min\left(R_{j,t} - b, 0\right)$. The semicovariance value obtained is 0.000035.

The calculation of the semivariance and semicovariance values results in a semivariance-semicovariance matrix that is used for the optimal weight calculation.

$$\sum_{sv} = \begin{bmatrix} 0.000133 & 0.000035 \\ 0.000035 & 0.000208 \end{bmatrix}$$

Calculate the inverse of the matrix.

$$\sum_{sv}{}^{-1} = \begin{bmatrix} 7882.0185 & -1316.2352 \\ -1316.2352 & 5023.7826 \end{bmatrix}$$

Next, determine the portfolio weights according to Equation (17).

$$\mathbf{w} = \frac{\sum_{sv}{}^{-1}\mathbf{1_2}}{\mathbf{1}_2^T\sum_{sv}{}^{-1}\mathbf{1_2}}$$

$$\mathbf{w} = \frac{\begin{bmatrix} 7882.0185 & -1316.2352 \\ -1316.2352 & 5023.7826 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix}}{\begin{bmatrix} 1 & 1 \end{bmatrix}\begin{bmatrix} 7882.0185 & -1316.2352 \\ -1316.2352 & 5023.7826 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix}}$$

$$\mathbf{w} = \frac{\begin{bmatrix} 6565.78323 \\ 3707.54739 \end{bmatrix}}{10273.3306}$$

$$\mathbf{w} = \begin{bmatrix} 0.63911 \\ 0.36089 \end{bmatrix}$$

Based on the portfolio weight calculations, it can be concluded that the optimal stock investment can be achieved by allocating 63.911% of the funds to UNTR shares and 36.089% to BRIS shares.

The portfolio return value is calculated by multiplying each stock return by the portfolio weight obtained. The calculation was performed using Google Colaboratory software. Based on the output, the expected return value obtained is 0.001045 and the portfolio return semivariance obtained is 0.000097, so it can be concluded that when investors want to invest their funds in the formed portfolio, they will obtain future profits of 0.001045 and will bear a risk of 0.000097.

The VaR calculation results using the Historical Simulation method show that at a 95% confidence level with holding periods of 1, 7, and 30 days, respectively, the values obtained are 0.021543, 0.056998, and 0.117996. If the initial investment capital is Rp10,000,000, then the maximum potential loss that may occur will not exceed Rp215,431 over a 1-day period, Rp569,977 over a 7-day period, and Rp1,179,964 over a 30-day period.

The Sharpe ratio can be calculated using the formula in Equation (19).

$$\overline{R_f}_{daily} = \frac{\frac{1}{12}(6\% + 6\% + 6\% + 6{,}25\% + \cdots + 6\%)}{365} = 0.0167\% = 0.000167$$

After the risk-free rate is obtained, the Sharpe ratio can be calculated as follows.

$E\left(R_p\right)$ = 0.001045

$$\varphi_p = \sqrt{\varphi_p{}^2} = \sqrt{0.000097} = 0.009866$$

$$\text{Sharpe} = \frac{0.001045 - 0.000167}{0.009866} = 0.009866$$

The calculation yields a Sharpe ratio of 0.088966, indicating that the mean-semivariance portfolio has good portfolio performance.

5. CONCLUSIONS

Corresponding Author: Suparti

Cluster analysis using the Fuzzy C-Means method on LQ45 index stocks obtained the best cluster number of 3 with the highest silhouette coefficient value of 0.812212. The number of members obtained was 2 stock members in cluster 1, 2 stock members in cluster 2, and 33 stock members in cluster 3.

The selection of stocks to form the Mean-Semivariance portfolio was done by taking the representative stocks from each cluster with the highest expected return values, namely UNTR from cluster 2 and BRIS from cluster 3. The investment weights obtained based on the selected stocks were 63.91% for UNTR and 36.09% for BRIS. The portfolio performance measurement using the Sharpe Index yielded a value of 0.088966, indicating good portfolio performance.

Based on VaR measurement using the Historical Simulation method, the VaR values at a 95% confidence level with holding periods of 1, 7, and 30 are 0.021543, 0.056998, and 0.117996, respectively. If an investment fund of Rp10,000,000.00 is allocated, the maximum potential loss that may occur at a 95% confidence level is Rp215,431.00 with a holding period of 1 day, Rp569,977.00 with a holding period of 7 days, and Rp1,179,964.00 with a holding period of 30 days.

REFERENCES

1. Adnyana, I. M. (2020). *Manajemen Investasi dan Portofolio*. Jakarta Selatan: Lembaga Penerbitan Universitas Nasional.
2. Amorim, L. B., Cavalcanti, G. D., & Cruz, R. M. (2022). *The Choice of Scalling technique matters for classification performance*. https://arxiv.org/pdf/2212.12343.
3. Bezdek, J. C. (1981). *Pattern Recognition with Fuzzy Objective Function Algorithms*. Plenum Press.
4. Entrisnasari, F. V. (2015). Analisis Portofolio Optimum Saham Syariah Menggunakan Mean Semivarian. *Jurnal Fourier*, *4*(1), 31-42.
5. Estrada, J. (2008). Mean-Semivariance Optimization: A Heuristic Approach. *Journal of Applied Finance*. 57-72.
6. Gubu, L., Cahyono. E., Arman., Budiman, H., & Djafar, K. M. (2023). Optimasi Portofolio Mean-Variance dengan Analisis Klaster Fuzzy C-Means. *Jurnal Gaussian*, *11*(4), 593-604.
7. Gujarati, D. N., & Porter, D. C. (2010). *Dasar-Dasar Ekonometrika Edisi 5 Buku 1*. Jakarta: Salemba Empat.
8. Hartono, J. (2017). *Teori Portofolio dan Analisis Investasi*. Edisi sebelas, BPFE Yogyakarta, Jakarta.
9. Hasibuan, R. A., Rizki, S. W., & Perdana, H. (2021). Perbandingan Metode *Fuzzy Clustering* Means dan *Single Linkage* pada Pengelompokan Saham LQ45. *Jurnal Untan*, *10*(3), 361-368.
10. Johnson, R. A., & Wichern, D. W. (2007). *Applied Multivariate Statistical Analysis Six Edition*. New Jersey: Pearson Prentice Hall.
11. Markowitz, H. (1959). Portofolio Selection: *Efficient Diversification of Investment.* Yale University Press.
12. Maruddani, D. A. I., & Trimono. (2020). *Microsoft Excel untuk Pengukuran Value at Risk: Aplikasi pada Risiko Investasi Saham*. Semarang: Undip Press.
13. Pakungwati, R. A., Wilandari, Y., & Maruddani, D. A. I. (2024). Metode Mean-Semivariance dalam Pembentukan Portofolio Reksa Dana Saham Terbaik Barometer Bareksa. *Jurnal Gaussian*, *13*(1), 250-259.
14. Struyf, A., Hubert, M., & Rousseeuw, P. (1997). Clustering in an Object-Oriented Environment. *Journal of Statistical Software*, *1*(4), 1-30.
15. Tandelilin, E. (2017). *Pasar Modal Manajemen Portofolio dan Investasi*. Yogyakarta: Penerbit PT Kanisius.

Corresponding Author: Suparti