International Journal of Mathematics and Computer Research

ISSN: 2320-7167

Volume 14 Issue 05 May 2026, Page no. – 6379-6385

Index Copernicus ICV: 57.55, Impact Factor: 8.751

DOI: 10.47191/ijmcr/v14i5.04

MULTI-OBJECTIVE STOCK PORTFOLIO OPTIMIZATION USING THE PARTITIONING AROUND MEDOIDS (PAM) APPROACH ON THE IDXESGL INDEX

Herni Herawati [1], Abdul Hoyyi[2*], Yuciana Wilandari[3]

[1,2,3] Department of Statistics, Universitas Diponegoro, Indonesia

ARTICLE INFO

Published Online:
07 May 2026

Corresponding Author:
Abdul Hoyyi

ABSTRACT

Investment refers to allocating funds into financial assets, including stocks, to increase future value. To reduce potential losses, investors apply diversification through portfolio formation. This study optimizes stock portfolio construction within the IDXESGL index, which comprises 30 stocks, by combining Partitioning Around Medoids clustering and multi-objective optimization. Clustering employed 2024 financial ratios Return on Assets, Earnings per Share, and Debt to Equity Ratio while considering outliers. Portfolio optimization used daily closing prices from June 20 to November 30, 2025. The multi-objective approach balanced expected return and risk based on investor preferences, with risk measured using Value at Risk under the Historical Simulation method. PAM produced three optimal clusters with a higher silhouette coefficient. MAPA and BBNI represented Clusters 1 and 2, while Cluster 3 was excluded due to negative expected return. For risk seekers, the portfolio was dominated by MAPA (0.939) and BBNI (0.061), generating an expected return of 0.001155026 and a 1-day VaR of -4.9182%. Risk neutral investors obtained more balanced weights, yielding an average expected return of 0.000617052 and a 1-day VaR between -3.0614% and -2.2197%. For risk averters, BBNI (0.759) and MAPA (0.241) produced an expected return of 0.000534355 and a 1-day VaR of -2.218%.

KEYWORDS: Investment; Portfolio; Partitioning Around Medoids (PAM); Multi-objective Optimization; Value at Risk.

I. INTRODUCTION

Along with the development of the financial system, investment has become a fundamental component. Investment is the allocation of funds to assets with the objective of increasing their value in the future to support long-term financial growth and stability [1]. Stocks represent financial assets that serve as investment instruments, as they represent ownership of a portion of a company's assets, and their movements on the Indonesia Stock Exchange are reflected in the ESG Leaders Index (IDXESGL). The ESG Leaders Index is an index launched on November 20, 2020, designed to track the stock price movements of companies that receive favorable assessments in environmental, social, and governance (ESG) aspects [2].

Investors aim to mitigate various risks associated with their investment activities, both in the short and the long term [3]. [4] states that investment risk may be reduced through diversification by allocating funds across several negatively correlated securities, thereby forming an investment portfolio. A portfolio is a collection of stocks constructed to achieve an optimal balance between return and risk [5]. An optimal portfolio is a portfolio constructed to achieve the most ideal balance between return and risk.

The selection of an optimal portfolio requires not only an understanding of risk and potential returns but also a systematic selection process to ensure more accurate investment decisions. The study conducted by [6] explores alternative methods for stock selection and portfolio diversification that do not rely solely on correlation but also consider price movement patterns and companies' fundamental factors. Clustering is one approach that may assist in stock selection for portfolio construction by grouping stocks based on their financial ratios [7].

In this study, the Partitioning Around Medoids (PAM) clustering method is used to group companies listed in the IDXESGL Index into several clusters. The PAM method determines cluster centers based on medoids, which are data points that best represent the members of each cluster. The use of medoids results in the clustering outcomes being less affected by the presence of outliers [8]. This characteristic makes the PAM method more robust to the presence of outliers or extreme data compared to K-means, which tends to be sensitive to such extreme values.

Portfolio construction always involves two main aspects, namely the level of return and the associated risk. The close relationship between return and risk requires investors to carefully evaluate the trade off between potential gains and the level of risk they are willing to bear. Multi-objective optimization is the process of simultaneously optimizing multiple objectives, even though these objectives are often conflicting [9]. The level of risk may be analyzed using various approaches, one of which is the Value at Risk (VaR) method.

The results of the study by [10] entitled A Multi-objective Approach to Portfolio Optimization, indicate that multi-objective optimization offers an alternative comparable to the mean–variance method and can be applied to investors with varying levels of risk tolerance, as measured by the risk aversion index. The findings of [8] show that, by applying time series clustering using the PAM method with the Dynamic Time Warping (DTW) distance measure, 45 stocks listed on the Indonesia Stock Exchange and included in the LQ45 Index were successfully grouped into six clusters. This study develops a novel approach to portfolio construction by integrating the Partitioning Around Medoids (PAM) clustering method with multi-objective optimization.

II. THEORETICAL FRAMEWORK

According to [11], fundamental analysis is conducted by examining key financial indicators to assess a company's intrinsic value. Investment ratios commonly utilized in fundamental stock analysis include Return on Assets (ROA), Earnings per Share (EPS), and the Debt to Equity Ratio (DER). ROA reflects the extent to which a company's assets generate profitability.

$$\text{ROA} = \frac{\text{Net Income}}{\text{Total Assets}} \quad (1)$$

EPS represents the amount of earnings allocated to each outstanding share.

$$\text{EPS} = \frac{\text{Net Income After Tax}}{\text{Number of Outstanding Shares}} \quad (2)$$

DER represents the proportion of a company's total liabilities relative to its equity.

$$\text{DER} = \frac{\text{Total Liabilities}}{\text{Total Equity}} \quad (3)$$

Data may be standardized using several methods. The Z-Score is a commonly used standardization method because it yields relatively stable values even in the presence of outliers. The formula for calculating the Z-Score is presented in Equation (4).

$$Z_{ij} = \frac{x_{ij} - \bar{x}_j}{\sigma_j} \quad (4)$$

where Z_{ij} is the standardized value of the i-th observation for the j-th financial ratio variable, x_{ij} is the i-th observation of the j-th financial ratio variable with $i = 1,2, \ldots N$, $\bar{x}_j$ is the mean of the j-th financial ratio variable with $j = 1,2, \ldots p$ and σ_j is the standard deviation of the j-th financial ratio variable.

Clustering refers to a method used to group data or objects into clusters based on the similarity of their characteristics. According to [12], cluster analysis requires that the sample be representative and the absence of multicollinearity among the variables used. A representative sample may be assessed using the Kaiser-Meyer-Olkin (KMO) measure, which is used to evaluate whether the data are suitable for analysis. [13] states that KMO may be expressed as shown in Equation (5).

$$KMO = \frac{\sum_{j=1}^{p}\sum_{l=1,l\neq j}^{p} r_{x_j x_l}^2}{\sum_{j=1}^{p}\sum_{l=1,l\neq j}^{p} r_{x_j x_l}^2 + \sum_{j=1}^{p}\sum_{l=1,l\neq j}^{p} \rho_{x_j x_l, x_u}^2} \quad (5)$$

where $r_{x_j x_l}$ is the correlation between variables x_j and x_l. $\rho_{x_j x_l}$ is the partial correlation between variables x_j and x_l after controlling for x_u. The sample is considered representative of the population if the resulting KMO value falls within the range of 0.5 to 1.

Multicollinearity is an issue in multivariate analysis because it complicates the identification of the effects of highly correlated variables, which can lead to hidden weighting among variables. According to [14], one method to detect multicollinearity is to calculate the Variance Inflation Factor (VIF). The VIF value is calculated using Equation (6).

$$VIF_j = \frac{1}{(1 - R_j^2)} \quad (6)$$

A VIF value greater than 10 indicates the presence of multicollinearity among the variables.

Outliers are observations that deviate from the general pattern in the data, and the most commonly used measure for detecting outliers is the Mahalanobis distance.

$$d_{MD}^2(i) = (\boldsymbol{x_i} - \bar{\boldsymbol{x}})^T \boldsymbol{\Sigma}^{-1} (\boldsymbol{x_i} - \bar{\boldsymbol{x}}) \quad (7)$$

where $d_{MD}^2(i)$ is the Mahalanobis squared distance of the i-th stock observation, $\boldsymbol{x_i}$ is the $1 \times p$ data vector of the i-th observation, $\bar{\boldsymbol{x}}$ is the $p \times 1$ mean vector of each variable, and $\boldsymbol{\Sigma}$ is the $p \times p$ covariance matrix. The i-th observation is identified as an outlier if $d_{MD}^2(i) > \chi_{p,1-\alpha}^2$ where $\chi_{p,1-\alpha}^2$ represents the outlier threshold at a probability of $1 - \alpha$.

The distance measures are used to assess similarity by evaluating how close the values of observations are across cluster variables [12]. However, technically, distance represents dissimilarity, as larger distances indicate lower similarity. The euclidean distance is often chosen as the most commonly used distance measure. Euclidean distance is commonly used calculated by summing the squared differences of each value and then taking the square root [15]. The formula for calculating Euclidean distance is presented in Equation (8).

$$d(x_i, m_c) = \sqrt{\sum_{j=1}^{p} (x_{ij} - m_{cj})^2} \quad (8)$$

where x_{ij} is the value of the i-th stock for the j-th variable ($i = 1,2,3, \ldots, N$) and m_{cj} is the value of the c-th cluster center for the j-th variable ($c = 1,2,3, \ldots, k$).

The data grouping process in the Partitioning Around Medoids (PAM) method is carried out by selecting medoids that best represent the cluster centers, thereby minimizing the distance or dissimilarity among objects within the clusters. An advantage of Partitioning Around Medoids lies in its flexibility, allowing clustering to be performed using various types of distance matrices, making the algorithm more robust to the presence of outliers [16]. The algorithm used in the Partitioning Around Medoids method consists of the following steps [17]:

1. Determine the number of k clusters to be formed;
2. Randomly select k objects from the dataset as the initial medoids;

3. Calculate the distance between non-medoid objects and the medoid objects using the Euclidean distance formula;
4. Assign each object to the nearest medoid, forming temporary cluster members;
5. Compute the total distance D' obtained;
6. Reselect k objects from the dataset as new medoids, choosing those that minimize the total distance to members within each cluster;
7. Reassign objects to the nearest new medoids to form a new cluster arrangement;
8. Calculate the change in total distance (D);
9. If $D < 0$, swap the medoids again to form a new set of k medoids;
10. Repeat steps 2 to 9 until $D \geq 0$ the iteration stops when no further medoid swaps are required.

One indicator used to assess cluster quality is the silhouette coefficient, which combines information on how closely cluster members are grouped together and how far they are from other clusters [18] :

$$SC = \frac{1}{N}\sum_{i=1}^{N} s(i) \tag{9}$$

where

$$s(i) = \frac{b(i)-a(i)}{\max\{a(i),b(i)\}};$$

$$a(i) = \frac{1}{N_c-1}\sum_{h=1,h\neq i}^{N_c} d(x_i,x_h);$$

$$b(i) = \min_{l\neq k}\left[\frac{1}{N_l}\sum_{h=1,h\neq i}^{N_l} d(x_i,x_h)\right]$$

$a(i)$ is the average distance between the i-th stock and all objects within the same cluster, and $b(i)$ is the minimum of the average distances between the i-th stock and stocks in other clusters.

Return refers to the gain obtained from an investment. The return is one of the primary factors attracting investors to invest, as it reflects the actual price changes [19]. The return of a stock can be calculated using Equation (10).

$$R_{it} = \ln\left(\frac{s_{it}}{s_{i,t-1}}\right) \tag{10}$$

Expected return refers to the anticipated level of profit to be obtained, calculated as the weighted average of various possible investment outcomes. The expected return of a stock can be calculated using Equation (11).

$$E(R_i) = \frac{\sum_{t=1}^{n} R_{it}}{n}; i = 1,2,3,\dots N \tag{11}$$

where $E(R_i)$ is the mean return of the i-th stock, and R_{it} is the return of the i-th stock in the t-th period.

Risk is defined as the deviation between the expected return and the actual return obtained. In stock investment, risk increases as the variance grows, whereas a smaller variance indicates a lower investment risk [20]. Investment risk can be calculated using the following variance formula.

$$\sigma_i^2 = \frac{1}{n-1}\sum_{t=1}^{n}(R_{it} - E(R_i))^2 \tag{12}$$

The investors' risk preferences may be categorized into three types. Risk seekers are willing to take high risks and have an optimistic view of future prospects. Risk neutrals exhibit a more balanced attitude toward risk. Risk averters avoid risk and prefer to invest in assets with relatively stable returns [1].

Visual inspection of the multivariate normality assumption may be conducted by creating a Chi-Square plot. This involves comparing the ordered Mahalanobis distances $d_{MD}^2(t)$ of each observation with the quantile values of the Chi-Square distribution. According to [21], multivariate normality can be formally tested using the Kolmogorov-Smirnov method. If the Chi-Square plot forms a straight-line pattern and the Mahalanobis distances are shown to follow the Chi-Square distribution based on the Kolmogorov-Smirnov test, it can be concluded that the stock return data are multivariate normally distributed.

Portfolio optimization aims to determine the asset allocation that maximizes return while minimizing risk [22]. The portfolio return is calculated by summing the products of each asset's weight and its return, as shown in Equation (13) [19].

$$R_p = \sum_{i=1}^{m} w_i R_{it} \tag{13}$$

The return of a portfolio can be computed using a matrix formulation.

$$R_p = w_1R_{1t} + w_2R_{2t} + \cdots + w_mR_{mt}$$

$$\begin{bmatrix} w_1 & w_2 & \cdots & w_m \end{bmatrix}\begin{bmatrix} R_{1t} \\ R_{2t} \\ \vdots \\ R_{mt} \end{bmatrix} = \boldsymbol{w^T}\mathbf{R} \tag{14}$$

The return of each asset is assumed to have a certain expected value, expressed as the mean return of the asset :

$$E(R_i) = \mu_i$$

Based on the expected return of each asset, the portfolio's expected return can be expressed as formulated in Equation (15).

$$E(R_p) = \sum_{i=1}^{m} w_i\,\mu_i = w_1\mu_1 + w_2\mu_2 + \cdots + w_m\mu_m =$$

$$\begin{bmatrix} w_1 & w_2 & \cdots & w_m \end{bmatrix}\begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_m \end{bmatrix} = \boldsymbol{w^T\mu} \tag{15}$$

The risk of a portfolio is defined as the variance of the portfolio return. The equation for calculating the portfolio variance can be expressed as shown in Equation (16).

$$\sigma_p^2 = \sum_{i=1}^{m}\sum_{h=1}^{m} w_i w_h\,\sigma_{ih} \tag{16}$$

where σ_p^2 is the portfolio return variance, w represents the portfolio weights, and σ_{ih} is the covariance between the i-th and h-th stocks.

Using a matrix formulation, the portfolio risk can be computed as presented in Equation (17).

$$\sigma_p^2 = \begin{bmatrix} w_1 & w_2 & \cdots & w_m \end{bmatrix}\begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1m} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{m1} & \sigma_{n2} & \cdots & \sigma_{mm} \end{bmatrix}\begin{bmatrix} w_1 \\ w_1 \\ \vdots \\ w_m \end{bmatrix}$$

$$= \boldsymbol{w^T\Sigma w} \tag{17}$$

where $\boldsymbol{\Sigma}$ is the variance-covariance matrix.

The multi-objective approach combines several objectives $f_1(x), f_2(x), f_3(x), \dots f_n(x)$ into a single objective function by assigning weight coefficients to each objective. Multi-objective optimization aims to maximize the expected return while simultaneously minimizing the portfolio risk.

Maximizing return $R_p = \sum_{i=1}^{m} w_i R_i = \boldsymbol{w}^T \boldsymbol{\mu}$

Minimizing risk $\sigma_p^2 = \sum_{i=1}^{m} \sum_{j=1}^{m} w_i w_h \sigma_{ih} = \boldsymbol{w}^T \boldsymbol{\Sigma} \boldsymbol{w}$ (18)

The effort to maximize portfolio return while simultaneously minimizing its risk can be expressed by minimizing the negative portfolio return and the portfolio risk. The multi-objective optimization function is formulated as follows [10].

$$\begin{cases} \text{minimize} : -\boldsymbol{\mu}^T \boldsymbol{w} + \boldsymbol{w}^T \boldsymbol{\Sigma} \boldsymbol{w} \\ \text{subject to} : \sum_{i=1}^{m} w_i = \mathbf{1}_m^T \boldsymbol{w} = 1 \end{cases} \quad (19)$$

Multi-objective optimization is solved using a scalarization technique, which transforms a multi-objective problem into a single-objective problem, making it easier to identify Pareto optimal points. Positive weight coefficients a_1 and a_2 where $a_1, a_2 > 0$.

$$\begin{cases} \text{minimize} : -a_1 \boldsymbol{\mu}^T \boldsymbol{w} + a_2 \boldsymbol{w}^T \boldsymbol{\Sigma} \boldsymbol{w} \\ \text{subject to} : \sum_{i=1}^{m} w_i = \mathbf{1}_m^T \boldsymbol{w} = 1 \end{cases} \quad (20)$$

The coefficients are defined as $a_1 = 1$ and $a_2 = K > 0$.

$$\begin{aligned} &\text{minimize} : -\boldsymbol{\mu}^T \boldsymbol{w} + K \boldsymbol{w}^T \boldsymbol{\Sigma} \boldsymbol{w} \\ &\text{subject to} : \sum_{i=1}^{m} w_i = \mathbf{1}_m^T \boldsymbol{w} = 1 \end{aligned} \quad (21)$$

The weight coefficient K represents the level of risk an investor is willing to bear in order to achieve the desired expected return [23]. Multi-objective optimization can be solved using the Lagrange function. In this approach, the Lagrange multiplier λ acts as a parameter that combines multiple objective functions with the constraints that must be satisfied.

$$L(w, \lambda) = -\boldsymbol{\mu}^T \boldsymbol{w} + K \boldsymbol{w}^T \boldsymbol{\Sigma} \boldsymbol{w} + \lambda(\mathbf{1}_m^T \boldsymbol{w} - 1) \quad (22)$$

The optimal value of $\boldsymbol{w}$ is obtained by differentiating Equation (22) with respect to $\boldsymbol{w}$ and setting the derivative equal to zero. From this procedure, the equation for the optimal portfolio weights in multi-objective optimization as presented below.

$$\boldsymbol{w} = \frac{1}{2K} \boldsymbol{\Sigma}^{-1} \left(\boldsymbol{\mu} - \left(\frac{\mathbf{1}_m^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu} - 2K}{\mathbf{1}_m^T \boldsymbol{\Sigma}^{-1} \mathbf{1}_m} \right) \mathbf{1}_m \right) \quad (23)$$

The proof that the Lagrange function in the multi-objective portfolio optimization problem yields an optimal solution is carried out by analyzing the second derivative of the Lagrange function.

$$\frac{\partial^2 L}{\partial w^2} [-\boldsymbol{\mu} + 2K \boldsymbol{\Sigma} \boldsymbol{w} + \lambda \mathbf{1}_m] > \mathbf{0}$$

$$2K\boldsymbol{\Sigma} > \mathbf{0} \quad (24)$$

The positive parameter $K > 0$ and the positive semi definite covariance matrix $\boldsymbol{\Sigma}$ result in a positive second derivative of the Lagrange function. This indicates that the Lagrange function is convex, implying that the obtained solution represents a global minimum and is optimal.
The Value at Risk (VaR) is one of the most commonly used indicators for measuring risk, representing the maximum estimated loss over a certain period [23]. VaR can be calculated using the historical simulation method based on historical return data. The mathematical formulation of VaR under the historical simulation method is provided in Equation (25) [19].

$$\text{VaR} = V_0 \times \text{percentile}\,\alpha \times \sqrt{T} \quad (25)$$

where V_0 is the investment amount, the $\alpha\%$ percentile of returns represents the $\alpha\%$ quantile derived from historical data, and T is the holding period.

III. RESEARCH METHOD

The data used in this study are secondary and obtained from the 2024 annual financial statements of 30 stocks included in the IDX ESG Leaders Index (IDXESGL), sourced from the Indonesia Stock Exchange https://www.idx.co.id/ as well as daily closing price data obtained from https://finance.yahoo.com. The variables analyzed include ROA, EPS, and DER for 2024, along with stock returns calculated from daily closing prices over the period from 20 June to 30 November 2025, resulting 112 observations.

This study was conducted using RStudio version 2025.09.1+401. The data analysis was carried out through the following stages.

1. The data were prepared for Return on Assets (ROA), Earnings per Share (EPS), and Debt to Equity Ratio (DER) from companies listed on the IDX ESG Leaders Index (IDXESGL) as the variables to be analyzed.
2. The variables were standardized using the Z-Score method according to Equation (4).
3. The assumptions were examined for cluster analysis, namely the representativeness of the data using Equation (5) and the absence of multicollinearity using Equation (6) for the variables ROA, DER, and EPS.
4. Outliers were identified in the data according to Equation (7).
5. Clustering was performed using the Partitioning Around Medoids (PAM) algorithm as defined in Equation (9).
6. The number of clusters was validated using the Silhouette Coefficient.
7. The return was calculated of each stock based on Equation (10).
8. The expected return was computed according to Equation (11) and select stocks with positive expected returns from each cluster as portfolio candidates.
9. The securities were determined of each cluster to be included in the portfolio.
10. The variance-covariance matrix was constructed based on combinations of stock returns.
11. The multivariate normality test was conducted on the combined stock return data.
12. If multivariate normality is not satisfied, select alternative stocks from each cluster and repeat the test until a combination of returns meeting multivariate normality is obtained.

13. The weights were calculated of each stock using multi-objective optimization with various values of the parameter K according to Equation (23).
14. The portfolio return was computed formed through the multi-objective optimization method using Equation (15).
15. Value at Risk (VaR) was estimated for each portfolio obtained from the multi-objective optimization using Equation (25).
16. The results were compared across all generated portfolios

IV. RESULT AND DISCUSSION

Cluster analysis began with testing the suitability of the data to ensure that the sample used was representative and free from multicollinearity among the variables. The adequacy of the data for clustering analysis was assessed through a sample representativeness test to verify that the data sufficiently represents the population. The Kaiser-Meyer-Olkin (KMO) test was conducted in RStudio using the *psych* package, resulting in a KMO value of 0.56, which exceeds the minimum threshold of 0.5, indicating that the data are adequate and representative for cluster analysis. A non-multicollinearity test was performed to ensure that there are no strong relationships among variables that could compromise the accuracy of the clustering results. Multicollinearity was detected using the Variance Inflation Factor (VIF) in RStudio, and the VIF values for each variable are presented in Table 1

Table 1. VIF Values for Detecting Non-Multicollinearity

Variable	VIF
ROA	1,0280
EPS	1,0303
DER	1,0336

The Partitioning Around Medoids (PAM) method is known to be relatively robust to the presence of outliers. However, outlier identification was still conducted as a confirmation step to ensure the suitability of the method for cluster analysis. Outliers were detected quantitatively using the mahalanobis distance with a significance level of 5% and the number of variables $p = 3$, resulting in a critical value of $\chi^2_{3,(1-0,05)} = 7{,}814728$. The analysis revealed several extreme observations in the ROA, EPS, and DER variables, with four data points GOTO, PGEO, SIDO, and UNVR exceeding the critical mahalanobis distance. Nevertheless, the presence of these outliers did not hinder the application of PAM, and the clustering process continued without excluding the outlier data.

The stock clustering process was carried out using the Partitioning Around Medoids (PAM) algorithm with Euclidean distance measured in RStudio for various numbers of clusters, namely $k = 2,3,4,5,6,7,8$. The determination of the optimal number of clusters was performed through validation using the silhouette coefficient. Higher values, approaching 1, indicate better clustering quality. The calculated silhouette coefficient values for each number of clusters are presented in Table 2.

Table 2. Silhouette Coefficient Values

Number of Clusters	Silhouette Coefficient
2	0,5131
3	0,5177
4	0,4260
5	0,4379
6	0,4533
7	0,2849
8	0,2914

Based on Table 2, three clusters yielded the highest silhouette coefficient value of 0.5177 compared to other cluster numbers. Therefore, this study adopted three clusters, which subsequently served as the basis for portfolio formation, with one stock selected as a representative from each cluster.

Table 3. Medoids for 3 Clusters

Cluster	ROA	EPS	DER	Number of Members
1	0,0738	114,3452	0,7818	23
2	0,0569	396,2100	5,5960	6
3	0,0535	2494,0100	0,4921	1

Table 3 presents the characteristics of the variables resulting from the clustering process based on the cluster centroids and the number of stocks in each cluster. The clustering results indicated that Cluster 1 (23 stocks) has ROA of 0.0739, EPS of 114.35, and DER of 0.78, indicating relatively stable companies with moderate financial risk. Cluster 2 (6 stocks) has ROA of 0.0570, EPS of 396.21, and DER of 5.60, indicating high profit potential but with higher financial risk. Meanwhile, Cluster 3 (1 stock) has ROA of 0.0535, EPS of 2,494.01, and DER of 0.49, indicating strong financial performance with low financial risk.

After determining the optimal number of clusters as three, the next step was portfolio formation by selecting one stock from each cluster with a positive expected return. The calculation results showed that not all stocks met this criterion. In Cluster 1, several stocks exhibited positive expected returns, whereas in Cluster 2, only BBNI, BNGA, and UNVR showed positive expected performance. The stock in Cluster 3, PGEO, had a negative expected return and was therefore excluded from the portfolio. Based on these considerations, the portfolio was formed focusing on clusters containing assets with positive expected returns, with MAPA chosen to represent Cluster 1 with an expected return of 0.0012092676, and BBNI representing Cluster 2 with an expected return of 0.0003200547.

A multivariate normality test was conducted to ensure that the returns of the stocks forming the portfolio followed a multivariate normal distribution. Based on the test using R software at a 5% significance level, a p-value of 0.05833 was obtained, indicating that the returns of MAPA and BBNI stocks are multivariate normally distributed. In the multi-objective approach, the weighting coefficient K is used to reflect how investors perceive the trade-off between risk and expected return, where a higher K value indicates a more risk-averse attitude, while a lower K value indicates a risk-seeking tendency. In this study, K was set at various levels, namely $K = 0{,}5$; $K = 0{,}7$; $K = 1$; $K = 5$; $K = 10$; $K = 50$; $K = 100$; $K = 500$; $K = 1000$; dan $K = 2000$, and the optimal weights of each stock in the portfolio were determined using the Lagrange method by minimizing the formulated objective function. The calculated portfolio weights are presented in Table 4.

Table 4. Multi-Objective Portfolio Stock Weights

Investor Type	K	Stock Weights	
		MAPA	BBNI
Risk Seeker	0,5	0,939	0,061
	0,7	0,739	0,261
Risk Neutral	1	0,590	0,410
	5	0,311	0,689
	10	0,276	0,724
	50	0,248	0,752
	100	0,245	0,755
Risk Averter	500	0,242	0,758
	1000	0,241	0,759
	2000	0,241	0,759

The expected return of a portfolio represents the anticipated level of profit for investors based on the combination of asset weights and their risk preferences. For illustration, with an investment of IDR 100,000,000, the expected return for each portfolio at various weighting coefficient K levels is presented in Table 5.

Table 5. Expected Return Portofolio and Profit Level

Investor Type	K	Expected Return Portofolio	Profit Level
Risk Seeker	0,5	0,001155026	Rp115.502,600
	0,7	0,000977183	Rp97.718,300
Risk Neutral	1	0,000844690	Rp84.469,030
	5	0,000596600	Rp59.659,990
	10	0,000565478	Rp56.547,750
	50	0,000540580	Rp54.057,950
	100	0,000537912	Rp53.791,190
Risk Averter	500	0,000535244	Rp53.524,500
	1000	0,000534355	Rp53.435,500
	2000	0,000534355	Rp53.435,500

The Value at Risk (VaR) was calculated using the historical simulation method for various values of the risk preference coefficient K at a 5% confidence level. The VaR represents the maximum potential loss that could occur under worst-case market conditions based on historical data. This analysis was conducted for holding periods of 1, 5, and 20 days, assuming an initial investment of IDR 1. The results are presented in Table 6.

Table 6. VaR for Various Holding Periods

Investor Type	K	Holding Period		
		1	5	20
Risk Seeker	0,5	-4,9182%	-10,9974%	-21,9948%
	0,7	-3,9664%	-8,8692%	-17,7385%
Risk Neutral	1	-3,0614%	-6,8455%	-13,6909%
	5	-2,2496%	-5,0302%	-10,0604%
	10	-2,2501%	-5,03114%	-10,0628%
	50	-2,2227%	-4,9700%	-9,9400%
	100	-2,2197%	-4,9634%	-9,9269%
Risk Averter	500	-2,2182%	-4,9601%	-9,9202%
	1000	-2,2179%	-4,9593%	-9,9187%
	2000	-2,2179%	-4,9593%	-9,9187%

For risk-seeking investors, the VaR was −4.9182% for a 1-day holding period and increased to −21.9948% for a 20-day holding period. This indicates that investors with a high-risk preference face relatively larger potential losses as the investment horizon extends. In contrast, for risk-neutral investors, the VaR was relatively lower and more stable. For a 1-day holding period, VaR ranged from −3.0614% to −2.2197%, reflecting a more controlled level of risk compared to risk-seeking investors. Meanwhile, for risk-averse investors, the VaR exhibited a converging pattern, approximately between −2.2182% and −2.2179% for a 1-day holding period.

V. CONCLUSION

This study demonstrates that the integration of Partitioning Around Medoids (PAM) clustering, multi-objective portfolio optimization, and Value at Risk (VaR) measurement using the historical simulation method is capable of producing an optimal stock portfolio within the IDXESGL index while considering the balance between return and risk according to investor preferences. The PAM method effectively forms three stock clusters based on fundamental characteristics, whereas multi-objective optimization generates portfolio compositions adaptable to risk-seeking, risk-neutral, and risk-averse investors. Additionally, the VaR analysis provides a quantitative assessment of potential maximum losses, thereby making this approach a relevant strategy for constructing sustainable ESG-based stock portfolios.

REFERENCES

[1] Santoso, A., Syahputri, A., Pustpita, G., Nurhikmat, M., Dewi, S., Arisandy, M., Nugraha, A., Anggraeni, I. S. K., Azizi, E., Yulaikah, Novyarni, N., Nurlia, Zahara, V. M., & Sasmiyati, R. Y. (2023). *Manajemen Investasi dan Portofolio*. Eureka Media Aksara.

[2] Kumajas, L. I., Saerang, I. S., Maramis, J. B., & Gamaliel, H. (2023). *Manajemen Investasi: Aplikasi pada Keuangan Berkelanjutan*. Bintang Semesta Media.

[3] Fahmi, I., & Hadi, Y. L. (2011). *Teori portofolio dan analisis investasi*. Alfabeta.

[4] Lubis, T. A. (2016). Manajemen investasi dan perilaku keuangan: Pendekatan teoritis dan empiris. Salim Media Indonesia.

[5] Dewi, N. K. A., & Candradewi, M. R. (2020). Pembentukan Portofolio Optimal Pada Saham Indeks IDX80 Dengan Menggunakan Model Markowitz. E-Journal Manajemen, 9(4), 1614–1633.

[6] Bolos, M.-I., Rusu, S., Leordeanu, M., Sabau-Popa, C. D., Perticas, D. C., & Crisan, M. I. (2025). K-Means Clustering for Portfolio Optimization. Symmetry, 17(847), 1–29.

[7] Basri, H. R., & Wibowo, W. (2025). Optimasi portofolio dan analisis risiko saham IDX30 menggunakan K-Means, Markowitz dan VaR. Jurnal Ilmiah Akuntansi dan Keuangan, 4(2), 123–142.

[8] Gubu, L., Rosadi, D., & Abdurakhman. (2021). Pembentukan Portofolio Saham Menggunakan Klastering Time Series K-Medoid dengan Dynamic Time Warping. Jurnal Aplikasi Statistika & Komputasi Statistik, 13, 35–46.

[9] Mishra, S. K., Panda, G., & Majhi, B. (2016). Prediction based mean-variance model using multiobjective evolutionary algorithms. Swarm and Evolutionary Computation, 28, 117–130.

[10] Duan, Y. C. (2007). A Multi-Objective Approach to Portfolio Optimization. Rose-Hulman Undergraduate Mathematics Journal, 8(1).

[11] Jones, C. P., Utama, S., Frensidy, B., Ekaputra, I. A., & Budiman, R. U. (2009). Investments Analysis and Management. Salemba Empat.

[12] Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). Multivariate Data Analysis (7th ed.). Pearson.

[13] Widarjono, A. (2010). Analisis Statistika Multivariat Terapan. UPP STIM YKPN.

[14] Gujarati, D. N. (2009). Dasar-Dasar Ekonometrika Jilid 2. Erlangga.

[15] Supranto, J. (2004). Analisis Multivariat Arti & Interpretasi. PT Rineka Cipta.

[16] Botyarov, M., & Miller, E. E. (2022). Partitioning Around Medoids. Results in Engineering, 15, 100544.

[17] Han, J., & Kamber, M. (2006). Data Mining: Concepts and Techniques. Elsevier.

[18] Vercellis, C. (2009). Business Intelligence. John Wiley & Sons.

[19] Maruddani, D. A. I. (2019). Value at Risk untuk Pengukuran Risiko Investasi Saham. Wade Group.

[20] Aliwu, R. R., Rahmi, E., Nuha, A. R., Yahya, L., Wungguli, D., & Arsal, A. (2025). Optimal Portfolio Formation Using Multi-Objective Optimization. Jambura Journal of Mathematics, 7(1), 49–56.

[21] Daniel, W. W. (1989). Statistika Nonparametrik Terapan. PT Gramedia.

[22] Lobato, L., Godinho, P., & João, M. (2017). Mean-semivariance portfolio optimization. Expert Systems with Applications, 79, 33–43.

[23] Primajati, G., Amrullah, A. Z., & Ahmad, A. (2019). Analisis Portofolio Investasi dengan Metode Multi Objektif. Jurnal Varian, 3(1), 6–12.

[24] Jorion, P. (2007). Value at Risk: The New Benchmark for Managing Financial Risk. McGraw-Hill.