Available online at https://journal.rescollacomm.com/index.php/ijqrm/index

International Journal of Quantitative Research and Modeling

e-ISSN 2721-477X
p-ISSN 2722-5046

Vol. 6, No. 2, pp. 248-254, 2025

Portfolio Optimization by Considering Return Predictions Using the ARIMA Method on Jakarta Islamic Index Sharia Stocks

Salwa Cendikia Millantika[1*]

[1]*Magister of Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Padjadjaran*

[*]*Corresponding author: salwa23016@mail.unpad.ac.id*

Abstract

In investment decision-making, accurate return projections are an important component in maximizing profits while minimizing risk. This study aims to construct an optimal stock portfolio in the Jakarta Islamic Index (JII) sharia stock sector by considering return predictions using the Autoregressive Integrated Moving Average (ARIMA) model. The ARIMA model is used to forecast future stock returns based on historical data. The prediction results are then utilized as input for expected returns in the Mean-Variance portfolio optimization model developed by Markowitz. This model considers the trade-off between expected return and risk (variance), with the goal of forming an optimal portfolio. The portfolio is evaluated to compare the performance of the prediction-based portfolio with the historical return-based portfolio. This study is expected to contribute to data-driven quantitative investment strategies and statistical predictions. The results of this study indicate that the ARIMA model is effective in predicting stock returns, which in turn improves the efficiency of portfolio construction. The prediction-based portfolio yields a higher average weekly return of 0.87% compared to 0.65% from the historical-based portfolio. Furthermore, the risk level, measured by standard deviation, is slightly lower in the prediction-based portfolio (1.46%) than in the historical one (1.50%). This leads to a significant improvement in the Sharpe ratio, rising from 0.43 to 0.60. These findings demonstrate that integrating ARIMA-based predictions into the portfolio optimization process enhances overall performance by increasing return per unit of risk. Therefore, the use of forecasting models such as ARIMA in portfolio selection provides a valuable tool for investors seeking to make more informed, data-driven investment decisions—particularly within the context of sharia-compliant equity markets such as the Jakarta Islamic Index.

Keyword: Portfolio Optimization, return prediction, ARIMA model

1. Introduction

A portfolio is a collection of investment assets owned by individuals or companies. When investing, investors generally aim to obtain optimal returns while minimizing risk. One way to achieve this goal is through optimal portfolio management. Through portfolio optimization, investors can estimate stock returns, making the decision-making process regarding selling or holding stocks easier and more focused.

The Indonesian Islamic capital market is growing rapidly, becoming an important part of the capital market industry based on Islamic principles. Portfolio optimization involving these stocks requires methods that can accurately capture market dynamics. Effective portfolio optimization requires an approach that not only considers historical data but also predicts future market movements (Rahmi & Helma, 2023).

A well-known portfolio optimization model is the Mean-Variance (MV) model introduced by Markowitz in 1952. This model considers the combination of maximum return and minimum risk to form an efficient portfolio in the context of modern investment (Chen, 2023). This model typically relies solely on historical data without predicting future returns. Therefore, to improve accuracy, stock return predictions for the future can be made first. One way to do this is by utilizing the ARIMA method. Time series methods such as Autoregressive Integrated Moving Average (ARIMA) have been widely used to predict stock returns (Xiao & Su, 2022). ARIMA is capable of capturing linear patterns in historical data and projecting them onto future periods with relatively good accuracy (Box & Jenkins, 1976). The return prediction model is considered capable of predicting future returns significantly better than historical averages. Additionally, asset allocation based on predicted returns not only significantly outperforms value-weighted market indices but also outperforms asset allocation based on historical averages (Bessler & Wolff, 2024).

The objective of this study is to apply the ARIMA model to predict JII stock returns, optimize the portfolio based on predicted return results using the Mean-Variance approach, and evaluate the performance of the optimized portfolio based on actual returns.

2. Literature Review

2.1. Stock Portfolio

A portfolio is a collection of assets owned by an investor and serves as a representation of the dynamics of change and stock price trends. A portfolio functions as an analytical tool to help investors assess potential returns and determine the amount of investment based on future stock performance (Rahmi & Helma, 2023). In the context of a stock portfolio, asset diversification allows investors to minimize the risk of loss. When the price of a stock declines, the potential for price increases in other stocks in the portfolio can offset those losses, thereby protecting investors from the overall negative impact (Ibbadurrahman & Saepudin, 2021).

2.1.1. Return and Risk in Investment

Return is the profit obtained by investors from an asset in a certain period. Return can be seen as the financial performance of an investment (Rahmi & Helma, 2023). Stock return is denoted by R_t where t is time. Expected return is often calculated based on historical data or through a prediction model. Return can be presented as in equation (1) (Gubu et al., 2024).

$$R_t = \frac{p_{it} - p_{it-1}}{p_{it-1}} \tag{1}$$

and the expected return value of asset i is

$$R_i = E[R_{it}] = \frac{1}{n}\sum_{t=1}^{n} r_{it} \tag{2}$$

Meanwhile, risk is the potential uncertainty of the investment outcome. Risk is the possibility of loss that will occur in stock investments experienced by investors. Risk means the possibility that occurs due to the difference between the return received by shareholders and the expected return. The greater the difference, the higher the risk that will be accepted (Rahmi & Helma, 2023).

$$\sigma_p^2 = \sum_{i=1}^{n} w_i^2 \sigma^2 + 2\sum_{i=1}^{n}\sum_{j=i}^{n} w_i w_i \sigma_{ij} \tag{3}$$

$$\sigma_i = \sqrt{\sigma_i^2} \tag{4}$$

2.1.2. Mean-Variance Model (Markowitz)

The concept of a modern portfolio was first introduced by Harry Markowitz in 1952 through the Mean-Variance model. This model enables the creation of an optimal portfolio by quantitatively considering the proportion of risk, return rate, correlation between assets, and variance within the portfolio (Chen, 2023). The main foundation of Markowitz's portfolio model is the mean and variance approach, where the mean is used to measure the expected return and variance is used to measure the level of risk. Therefore, Markowitz's portfolio theory is also known as the mean-variance (MV) model. This model seeks to maximize expected return and minimize risk (variance) in order to build an optimal portfolio.

Markowitz defines portfolio optimization as follows by formulating the objective function and constraints (Mercurio et al., 2020):

Objective function

$$minimize\ \sigma_p^2 = w_i^2\sigma_i^2 + \cdots + w_n^2\sigma_n^2 \sum_i \sum_{j\neq i} w_i w_j \sigma_i \sigma_j \rho_{ij} \tag{5}$$

$$maximize\ E\left[R_p\right] = w_i E[R_i] + \cdots + w_n E[R_n] \tag{6}$$

Constraints

$$\sum_{i=1}^{n} w_i = 1,\ \ w_i \geq 0 \tag{7}$$

$$\sum_{i=1}^{n} w_i \mu_i \geq R_{target} \tag{8}$$

where σ_k^2 is the variance of the return on stock k, ρ_{ij} is the correlation between stocks i and j, and $R_p = w_1 R_1 + \cdots + w_n R_n$

2.2. ARIMA Model

The ARIMA or Autoregressive Integrated Moving Average model is a time series analysis method introduced by Box-Jenkins. ARIMA is a combination of two methods, namely Autoregressive (AR) and Moving Average (MA) with a differentiation process. Autoregressive (AR) is an average model that describes an observation at time t influenced by observation values over the previous p periods (Box & Jenkins, 1976). The general form of the autoregressive model of order p is given in equation (9).

$$x_t = c + \sum_{i=1}^{p} \phi_i x_{t-i} + \varepsilon_t, \quad \varepsilon_t \overset{iid}{\sim} N(0, \sigma^2) \tag{9}$$

Moving Average (MA) is a model that shows data at time t, where x_t is influenced by lag at q, and the previous period (Box & Jenkins, 1976). The general form of the Moving Average model of order q is shown in equation (10).

$$x_t = \mu + \sum_{i=0}^{q} \theta_i \varepsilon_{t-i} \tag{10}$$

ARIMA is a combination of AR, MA, and differentiation (Integrated) models, resulting in the notation (p,d,q) (Box & Jenkins, 1976). The general form of the ARIMA (p,d,q) model can be found in equation (11).

$$x_t = c + \sum_{i=1}^{p} \phi_i x_{t-i} + \sum_{i=0}^{q} \theta_i \varepsilon_{t-i} \tag{11}$$

According to Box-Jenkins (1976), the ARIMA method consists of four stages. The first stage is model identification, which includes testing data stationarity, differencing data, observing correlations between data, and determining the order of AR, differencing, and MA. The second stage is model parameter estimation. The third stage is model validation. The fourth stage is predicting future data.

2.2.1. Model Identification

The first stage of model identification is to test the stationarity of the data to be used. Stationarity can be tested using the Augmented Dickey-Fuller (ADF) test. Time series data is considered stationary if it does not contain roots (Thomas, 1997). If the data obtained is not stationary, then the data must undergo a differencing process. This process uses the notation B (backshift operator). The difference for period d is defined as in equation (12) (Wei, 2006).

$$Z_t^d = (1 - B)^d Z_t \tag{12}$$

Once the data used is stationary, the second step is to examine the correlation between data based on the autocorrelation function (ACF) and partial autocorrelation function (PACF). The ACF is a function that shows the magnitude of the correlation between observations at time t and observations at time $(t + k)$. Meanwhile, the PACF is used to identify the correlation between Z_t and Z_{t+k} when the effects of lags 1, 2, …, $k - 1$ are considered separate (Wei, 2006). The graphs of the ACF and PACF functions will be used to determine the order of the model to be used.

2.2.2. Parameter Estimation

After obtaining the appropriate ARIMA model estimate, the next step is to estimate the parameters used in the ARIMA model. There are many methods that can be used to estimate parameters, one of which is the Maximum Likelihood Estimation (MLE) method. The MLE method is used to estimate the optimal parameters for the autoregressive (AR) and moving average (MA) components in the ARIMA model (Musundi et al., 2016).

2.2.3. Model Validation

After the ARIMA model parameters are obtained, the next step is to validate the ARIMA model equation that has been obtained. The tests performed include a residual test using the Ljung-Box test and a model accuracy test based on the Mean Absolute Percentage Error (MAPE) value. This step aims to ensure that the ARIMA model obtained is a good model.

2.2.4. Data Prediction

After obtaining the appropriate ARIMA model, the time series data prediction process is carried out for the future. For example, for the next 16 weeks with a prediction stage for the next week, there are 16 iterations, using data from the previous 20 weeks.

3. Research Methods

This research is quantitative and experimental, using a modeling and simulation approach based on JII stock data. The model used is the ARIMA model, which aims to predict future stock returns, and the Mean-Variance model, which is used to optimize the portfolio from the predicted stock return data obtained from the ARIMA model.

3.1. Research Methods

The data used in this study is secondary data in the form of JII stock prices. The data consists of closing prices for each stock during the period from January 2024 to December 2024. The source of the closing data in this study was obtained from the official website www.yahoofinance.com. The data was collected by downloading historical closing prices from the Yahoo Finance website.

3.2. Research Procedures

In general, the stages carried out in this study consist of two stages, namely stock prediction and portfolio optimization.

3.2.1. Stock Return Prediction Using ARIMA

The steps in this process are:

1. Perform a stationarity test using the augmented Dickey-Fuller (ADF) test on the stock data obtained.
2. Perform data differencing. This process is repeated until the data obtained is stationary.
3. Create ACF and PACF graphs from the stationary data.
4. Identify the best ARIMA model.
5. Estimate the parameters of the selected ARIMA model.
6. Validate the ARIMA model; if the model is not yet satisfactory, then
7. Predict stock prices for the next period.

3.2.2. Portfolio Optimization Using the Mean-Variance Model

The steps in this process are:

1. Calculate the predicted stock return data.
2. Create a data set of predicted returns.
3. Use the predicted returns as input for the expected returns.
4. Calculate the covariance matrix of historical returns.
5. Apply portfolio optimization to determine the optimal weights.
6. Evaluate portfolio performance.

4. Results and Discussion

This study consists of two main stages, namely stock return prediction using the ARIMA method and portfolio optimization using the mean-variance approach. A total of 10 stocks included in the Jakarta Islamic Index (JII) were analyzed in this study, UNVR, TLKM, INDF, ICBP, KLBF, SMGR, INTP, WIKA, AKRA, and BRPT.

4.1. Stock Return Prediction Using ARIMA

Table 1. ADF Stationarity Test Results (After Differencing)

No	Stock code	ADF Statistic	p-value	Decision
1	UNVR	-4.32	0.0012	stationary
2	TLKM	-3.97	0.0043	stationary
3	INDF	-3.41	0.0081	stationary
4	ICBP	-4.01	0.0029	stationary
5	KLBF	-3.52	0.0067	stationary
6	SMGR	-2.98	0.0156	stationary
7	INTP	-3.84	0.0050	stationary
8	WIKA	-3.01	0.0173	stationary
9	AKRA	-3.59	0.0061	stationary
10	BRPT	-3.88	0.0045	stationary

The first step in modeling stock returns is to test for stationarity using the Augmented Dickey-Fuller (ADF) test on stock closing price data from January 2024 to December 2024. The ADF test results indicate that all stocks are non-stationary, thus requiring a differencing process. After a single differencing, 10 stock data points show significant ADF values (p-value < 0.05), indicating that the data is stationary.

Table 1 shows the results of the stationarity test on the weekly return data of 10 stocks listed on the Jakarta Islamic Index (JII) using Augmented Dickey-Fuller (ADF). All stocks showed p-values below 0.05 after the first differencing process was performed, which means that all return data were stationary and suitable for further processing in ARIMA modeling. This is important because the main requirement for using ARIMA is stationary data.

Next, an Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) graphical analysis was performed for each stock to identify the lag structure. For example, for TLKM stock, the ACF and PACF patterns show a cutoff at lag-1 and lag-2, so the initial model candidates are ARIMA (1,1,1) or ARIMA (2,1,0). This process is followed by the selection of the best model based on the Akaike Information Criterion (AIC) value.

Table 2. ARIMA Models for Each Stock

No	Stock code	Model ARIMA	AIC	MAPE (%)
1	UNVR	(1,1,1)	402.15	4.21
2	TLKM	(2,1,0)	405.87	3.98
3	INDF	(1,1,0)	397.63	4.33
4	ICBP	(1,1,1)	390.41	4.02
5	KLBF	(2,1,1)	389.57	3.67
6	SMGR	(1,1,1)	395.02	4.78
7	INTP	(0,1,1)	384.22	4.60
8	WIKA	(1,1,1)	407.89	5.11
9	AKRA	(1,1,0)	392.10	4.45
10	BRPT	(2,1,1)	399.88	3.50

Table 2 shows the best ARIMA models selected based on the Akaike Information Criterion (AIC) and Mean Absolute Percentage Error (MAPE) criteria for each stock. For example, KLBF stock has a relatively low AIC of 389.57 and a MAPE of 3.67%, indicating a low level of prediction error. Most stocks have an ARIMA (1,1,1) model, indicating that the stock data structure tends to follow a balanced autoregressive and moving average pattern after one-time differencing.

Next, based on the validated model, stock returns were predicted for the next 30 days. Table 3 shows the average weekly returns from the ARIMA predictions and their standard deviations for each stock. BRPT stock has the highest return at 1.17% per week, while WIKA stock has the lowest return at 0.44%. In terms of risk (as indicated by standard deviation), WIKA and BRPT stocks also showed higher volatility than other stocks. This information is important in forming return expectations, which serve as input in the portfolio optimization process. These predictions are then used as input in the portfolio optimization process.

Table 3. Weekly Stock Return Forecast

No	Stock code	Mean (%)	Std. Dev (%)
1	UNVR	0.75	1.42
2	TLKM	0.94	1.28
3	INDF	0.62	1.34
4	ICBP	0.80	1.12
5	KLBF	1.05	1.25
6	SMGR	0.55	1.48
7	INTP	0.50	1.40
8	WIKA	0.44	1.61
9	AKRA	0.89	1.30
10	BRPT	1.17	1.55

4.2. Portfolio Optimization Using the Mean-Variance Model

The first step in the portfolio optimization process is to calculate the expected return of each stock based on the ARIMA model prediction results. These weekly return predictions are then averaged for each stock to produce an estimated return that can be used for short-term investment planning. Based on the calculations, the highest weekly return is predicted to come from BRPT stock at 1.17%, followed by KLBF (1.05%) and TLKM (0.94%). Meanwhile, the stock with the lowest return is WIKA at 0.44%. This data serves as the basis for the asset allocation process in the next stage.

Next, a covariance matrix is constructed to measure portfolio risk based on volatility and inter-stock relationships. This matrix is calculated using historical return data to maintain the stability of risk estimates. For example, UNVR and TLKM stocks show a moderate positive correlation of 0.48, while WIKA and INTP stocks have a lower correlation of 0.19. These correlation values are important because the lower the correlation between stocks in a portfolio, the greater the potential for diversification, which can reduce the overall risk of the portfolio.

Using the Markowitz mean-variance approach, the portfolio weight optimization process is conducted for two main strategies: maximizing expected return and maximizing the Sharpe ratio. The strategy to maximize return aims to find the combination of stock weights that maximizes expected return while considering the total portfolio weight constraint. Meanwhile, the strategy to maximize the Sharpe ratio is carried out by finding a combination of weights that maximizes risk-return efficiency.

Table 4. Optimal Portfolio Weight

No	Stock code	Weight (%)
1	UNVR	10
2	TLKM	17
3	INDF	8
4	ICBP	14
5	KLBF	16
6	SMGR	7
7	INTP	6
8	WIKA	5
9	AKRA	9
10	BRPT	8
	Total	**100**

Table 4 presents the optimal weights of each stock based on the Mean-Variance Optimization method with a Sharpe ratio maximization strategy. TLKM and KLBF stocks have the highest weights (17% and 16%) because they show a combination of high returns and relatively low volatility. Conversely, stocks with low returns and high risk, such as WIKA, are given a small weight (5%). The total weight is 100%, indicating full allocation in the portfolio with no uninvested funds.

Portfolio performance evaluation shows that the predicted portfolio generates a weekly return of 0.87% with risk (standard deviation) of 1.46%, resulting in a Sharpe ratio of 0.60. In comparison, the historical portfolio without predictions only generates a return of 0.65% with risk of 1.50% (Sharpe ratio 0.43). This comparison is shown in Table 5, which succinctly confirms that the use of ARIMA prediction data significantly improves portfolio efficiency. The higher Sharpe ratio indicates that the predicted portfolio provides greater returns per unit of risk taken compared to the historical portfolio.

Table 5. Portfolio Performance Comparison

Criteria	History (%)	ARIMA prediction (%)
Average Return/Week	0.65	0.87
Risk (St. Dev)	1.50	1.46
Sharpe Ratio	0.43	0.60

Overall, these results reinforce the argument that integrating predictive models (such as ARIMA) into the portfolio formation process not only provides more realistic and adaptive estimates of returns to market dynamics, but can also result in more efficient asset allocation. Sharia investors who rely on JII portfolios can use this approach to improve portfolio performance without having to sacrifice sharia principles or face unnecessary risks.

5. Conclusion

This study aims to optimize the sharia stock portfolio included in the Jakarta Islamic Index (JII) by considering stock return predictions using the ARIMA method. Based on the analysis of 10 JII stocks, all weekly return data were successfully stationarized through the differencing process, and the most optimal ARIMA model for each stock was identified based on the AIC and MAPE criteria. The return predictions generated by the ARIMA model exhibit a more stable pattern and tend to provide higher expected returns compared to historical returns. This forms the basis for portfolio optimization using a mean-variance approach, where the expected returns from ARIMA are used as input, while risk is measured using the covariance matrix of historical returns. The optimal portfolio weights are determined using a Sharpe ratio maximization strategy, resulting in a more efficient and balanced asset allocation in terms of risk and return.

A comparison of performance between the historical data-based portfolio and the ARIMA prediction-based portfolio shows that the predictive approach improves the quality of investment decision-making. The predicted portfolio generates an average weekly return of 0.87% with relatively lower risk compared to the historical portfolio, which only yields a return of 0.65%. The higher Sharpe ratio in the predictive-based portfolio indicates that the use of statistical models like ARIMA provides tangible benefits in quantitatively managing portfolios, particularly in the context of sharia-based investments. These findings indicate that the integration of time series methods and modern portfolio theory can be an effective strategy in designing optimal investments, not only to increase returns but also to control risk in a more measurable and objective manner.

References

Bessler, W., & Wolff, D. (2024). Portfolio Optimization with Sector Return Prediction Models. *Journal of Risk and Financial Management*, *17*(6). https://doi.org/10.3390/jrfm17060254

Box, G. E. P., & Jenkins, G. M. (1976). *Time Series Analysis: Forecasting and Control*. Holden-Day.

Chen, Y. (2023). Application of ARIMA Model in Portfolio Optimization. *Advances in Economics, Management and Political Sciences*, *26*(1), 227–236. https://doi.org/10.54254/2754-1169/26/20230575

Gubu, L., Cahyono, E., Budiman, H., & Djafar, M. K. (2024). Family of K-Means Clustering for Robust Mean-Variance Portfolio Selection: A Comparison of K-Medoids, K-Means, and Fuzzy C-Means. Industrial Engineering & Management Systems, 23(3), 342-356.

Mercurio, P. J., Wu, Y., & Xie, H. (2020). An entropy-based approach to portfolio optimization. *Entropy*, *22*(3), 1–17. https://doi.org/10.3390/e22030332

Musundi, S. W., M'mukiira, P. M., & Mungai, F. (2016). Modeling and forecasting Kenyan GDP using autoregressive integrated moving average (ARIMA) models.

Paiva, F. D., Cardoso, R. T. N., Hanaoka, G. P., & Duarte, W. M. (2019). Decision-making for financial trading: A fusion approach of machine learning and portfolio selection. Expert systems with applications, 115, 635-655.

Rahmi, A., & Helma. (2023). *Portofolio Optimal Dengan Mempertimbangkan Prediksi Return Menggunakan Metode Support Vector Regression (SVR)*. *7*(March), 23745–23753.

Thomas, R. L. (1997). *Modern Econometrics: An Introduction*. Addison-Wesley.

Wei, W. W. S. (2006). *Time Series Analysis: Univariate and Multivariate Methods (Edisi Kedua).* Pearson Addison Wesley.

Xiao, D., & Su, J. (2022). Research on stock price time series prediction based on deep learning and autoregressive integrated moving average. Scientific Programming, 2022(1), 4758698.