

Portfolio Based on AR Model Using 5 Stocks

Jiayi Gao *

Department of Economics and Management, Beijing Jiaotong University, Beijing, China

* Corresponding Author Email: 1318043825@qq.com

Abstract. An investment portfolio is a collection of stocks, bonds, and financial derivatives held by investors or financial institutions, aimed at diversifying risks and maximizing returns. This article selects stocks of Manulife, BOA, Apple, PLC, and National Pharmaceutical Group from October 21, 2022 to August 9, 2023. This article uses AR model, minimum variance model, and maximum Sharpe's ratio model to predict stock prices, and establishes investment portfolios based on the predicted values. Firstly, based on the first 75% of the data, this article predicts the last 25% of the data. Secondly, based on the prediction, the minimum variance model and the maximum Sharpe ratio model are applied and compared with SP500 within the same time frame. However, the results indicate that the SP500 index performs best in terms of yield. This study suggests that purchasing market index funds is an important investment decision for investors during the volatile US dollar interest rate hike cycle.

Keywords: Mean-variance model; AR model; SP500.

1. Introduction

The economic globalization and the diversification of the financial market have intensified the volatility and investment risk of the financial market [1]. In order to hedge this risk, the portfolio investment strategy proposed by Markowitz has become the main risk management and avoidance means [1]. In financial market, project of portfolio investment optimization facilitates weighing up risks against returns for various investment products [2]. Portfolio is also a hot topic aiming to make the best effort to distribute wealth efficiently among the variety of assets for academic researchers, in order to maximize the effect of revenue [3]. At present, the global economy is in a cycle of US dollar interest rate hikes, which highlights the importance of portfolio construction.

The recent years have witnessed a number of researches about portfolio investment. For instance, Laila's group [4] identifies what constituted an valid portfolio of ten Islamic banks financed investments. Cho and Moon [5] conduct a comparative study to find a better investment portfolio within the narrow range near the designated threshold of FDI by comparing FDI enterprises with foreign securities investment enterprises. Yang and Li [6] put forward portfolio selection model with transaction costs, and obtain condition for its solution as well as the general numerical solution. Furthermore, some researchers are also looking to find new approaches to improve portfolios in order to achieve greater returns. For example, Ait-Sahalia and Brandt [7] study the optimal portfolio as certain conditions are partially predictable at some time. Mercurio's group [8] proposes a novel approach that uses entropy as a risk to construct optimal portfolios. Campbell and Viceira [9] demonstrate the use of the latest results of analytical and numerical methods to understand and explore solutions to portfolio selection problems of financiers in long term. However, researches on portfolio application are still scarce under the background of US dollar interest rate rise. Meanwhile, current research on the combination of forecasting models and portfolio models is still insufficient.

The purpose of this article is trying to utilize AR model to generate a more informed prediction of future returns of investment products, which optimizes investment portfolio and minimums the risks. The empirical process is shown below, this paper first selects five properties of Manulife, BOA, Apple, PLC, Sinopharm Group from American stock market between October 21st, 2022, and August 9th, 2023. Second, the paper regards the previous 150 days of the data as the set trained to practice AR model to calculate the next day's rate of the returns. Based on the value of prediction, the optimal daily weights of the 50 days (the maximum Sharpe ratio pattern and the minimize variance pattern) are recorded at the same time. Third, calculate the yield of daily returns by using the optimal daily

weights, and then get the result of the yield of cumulative returns based on the yield of daily returns. Forth, through comparing the data of the maximum Sharpe ratio pattern, the minimize variance pattern and SP500, the paper got the final best result whose rate of cumulative returns is highest, and the results show that investing the market index is a preferred choice.

The structure of the remainder of this article is as follows. Data sections introduces stocks selection and data reduction in the study. Method sections employed in this study are elaborated in detail, including AR prediction model, mean-covariance model, and other technologies. Result sections displayed the results of prediction and accuracy of the model by tables and figures. Eventually, in conclusion sections, the paper concluded the whole content of the article and evaluated the article.

2. Data

The paper selects five representative stocks based on diverse aspects including insurance company, banks, pharmaceutical industry and electronics technology, which closely relate to daily life. The ticker of 5 stocks is MFC, BOA, Apple, PLC and CNPGC. The daily stock data in this essay is provided by Sina Finance (<https://finance.sina.com.cn/>) and the data of SP500 is acquired by Investing Com (<https://cn.investing.com/indices/us-spx-vix-futures-historical-data>). Closing rates of returns from October 21st,2022, to August 9th,2023, are separated into training set and test set. Eventually, data are collected in the Table 1.

Table 1. 5 selected stocks

Stock Symbol	Company
MFC	Manulife Financial Company
BOA	Bank of America
Apple	Apple Inc.
PLC	Prudential Investment, Loan, and Assurance Company
CNPGC	China National Pharmaceutical Group Corporation

In order to simplify the name of stocks to identify them readily, the paper will then use the abbreviations above to symbolize them.

Meanwhile, the paper records the rates of returns and calculates some basic information like mean, variance, maximum, minimum to utilize the descriptive statistics on the data displayed in Table 2.

Table 2. Descriptive statistics of five stocks

	'MFC'	'BOA'	'Apple'	'PLC'	'CNPGC'
Mean	0.0011	-0.0003	0.0012	0.0016	0.0026
Std	0.0135	0.017	0.0176	0.0262	0.0275
Max	0.0505	0.0497	0.0890	0.1178	0.1641
Min	-0.0383	-0.0620	-0.0480	-0.1071	-0.0984

3. Method

3.1. Monte-Carlo model

Monte Carlo simulations can measure statistics over repeated experiments to solve those complicated problems [10]. Based on probability and statistical theory, Monte Carlo simulation can directly simulate the actual complex system, which is suitable for solving the problems of high dimension, complexity and time continuity [11]. The study uses Monte Carlo simulations to simulate the weights of asset mentioned below.

3.2. Mean-Variance model

In 1952, mean-variance model was firstly raised by H.M. Markowitz as one of the models to solve risk and investment problems, which is the first application of mathematical statistics to researching

portfolio selection and makes multi-objective optimization of benefits and risks achieve the best balance effect. Thus the model is appropriate for the portfolio optimization. There are three important indexes in the mean-variance model including return, risk and Sharpe ratio.

Expected return refers to the profits on assets that may be achieved in the future under the situation of indeterminate portfolio. It is the sum of the value in each stock multiplied by that stock's weight. Let w_i be the weight of the i -th asset such that $\sum_i w_i$ and μ_i be the expected return of the i -th asset. Then the expected return of portfolio is:

$$E(R_p) = \sum_i w_i \mu_i \quad (1)$$

In Mean-Variance model, risk is typically a symbol as the variance of the portfolio returns. The variance of a portfolio is the sum of the weighted variances of every asset in the portfolio, plus twice the weighted covariance between each pair of assets. Let

σ_i be the standard deviation of the i -th asset and ρ_{ij} be the correlation between the returns of the i -th and j -th asset. Then the portfolio return variance is:

$$\sigma_p^2 = \sum_i w_i^2 \sigma_i^2 + \sum_i \sum_{j \neq i} w_i w_j \sigma_i \sigma_j \rho_{ij} \quad (2)$$

Sharpe Ratio is standard index of fund performance evaluation. When getting positive value, it means that the rate of cumulative return is higher than the risk. Otherwise, the risk is higher. Let R_f is the current risk-free rate of the market and σ_p is the standard deviation of the portfolio's return, then the Sharpe Ratio is:

$$\text{Sharpe ratio} = \frac{R_p - R_f}{\sigma_p} \quad (3)$$

3.3. AR model

AR model, autoregressive model, is the model that utilize itself as a regression variable to make the analyzation and indication. In other words, the model is the frequent formation of the time series, applying the earlier linear combination of random variables to predict the linear regression of random variables at later time. The equations for the AR model are as follows:

In autoregressive model $AR(p)$, if time series x satisfies:

$$x_t = \phi_0 + \sum_{i=1}^p \phi_i x_{t-i} + \varepsilon_t \quad (4)$$

where x_t is the value of time series x at t time, ϕ_0 is constant term, $\phi_1, \phi_2, \dots, \phi_p$ are autoregressive coefficient and ε_t is independently distributed series of random variables, which satisfies:

$$E(\varepsilon_t) = 0, \text{Var}(\varepsilon_t) > 0 \quad (5)$$

Then it said that time series x obey p -order autoregressive model $AR(p)$.

Yule-Walker Equation:

Set $\rho_i, i = 1, 2, \dots, p$ are i -order autoregressive coefficient, then:

$$\left\{ \begin{array}{l} \rho_1 = \phi_1 + \phi_2 \rho_2 + \dots + \phi_p \rho_p \\ \rho_2 = \phi_1 \rho_1 + \phi_2 + \dots + \phi_p \rho_p \\ \dots \\ \rho_p = \phi_1 \rho_1 + \phi_2 \rho_2 + \dots + \phi_p \end{array} \right. \quad (6)$$

Solve the equations, and then gain the autoregressive coefficient $\rho_i, i = 1, 2, \dots, p$.

Then use the recurrence relationship to find the autocorrelation coefficients of each order:

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} + \dots + \phi_p \rho_{k-p}, k \geq p + 1 \quad (7)$$

Here is the absolute process of AR prediction. First of all, the paper determines whether the time series are stationary by graph test method. Moreover, make sure the proper time series model is

$AR(p)$ and determine the model order. Furthermore, estimate the parameter and generate the model. Finally utilize the model to make the prediction and evaluate the result from May 26th, 2023, to August 9th, 2023.

3.4. MSE

The paper use MSE to calculate the prediction accuracy of the model. MSE, Mean-Squared-Error, is a typical measure of the difference between the predicted value and the actual value. Its value can be obtained by squaring the difference between the predicted value and the true value and averaging it. Therefore, MSE is one of the most important indexes to measure the accuracy of prediction model. It can help us judge the predictive power of the model and the size of the error. In general, the smaller the MSE, the better the predictive power of the model. There is the calculation formula:

$$mse = \frac{1}{n} * \sum (y_i - \hat{y}_i)^2 \tag{8}$$

Where y_i is the actual value, \hat{y}_i is the predicted value and n is the days number of predicted value.

4. Results

The study first investigates the fitting performance of AR model that use sliding windows to predict rate of returns. Based on these data, the paper then predicts yield of each asset and constructs the five comparison figures between indicated yield and effective yield from Figure 1 to Figure 5.

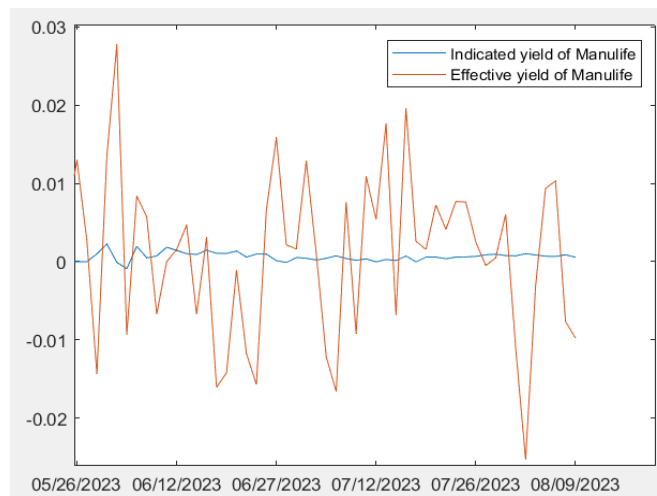


Fig. 1 Comparison between indicated yield and effective yield of Manulife.

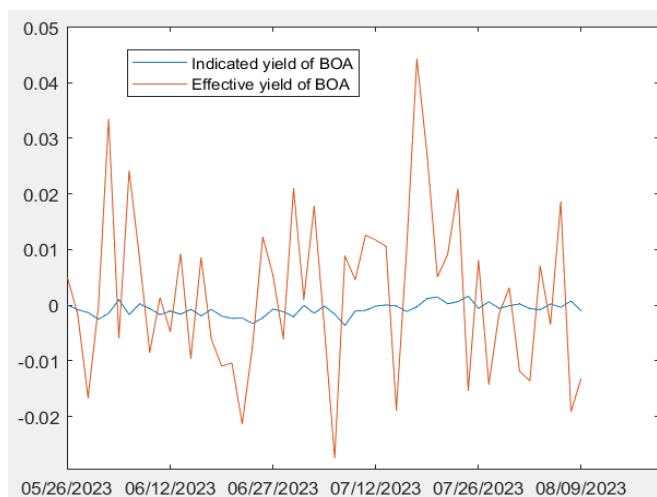


Fig. 2 Comparison between indicated yield and effective yield of BOA.



Fig. 3 Comparison between indicated yield and effective yield of Apple

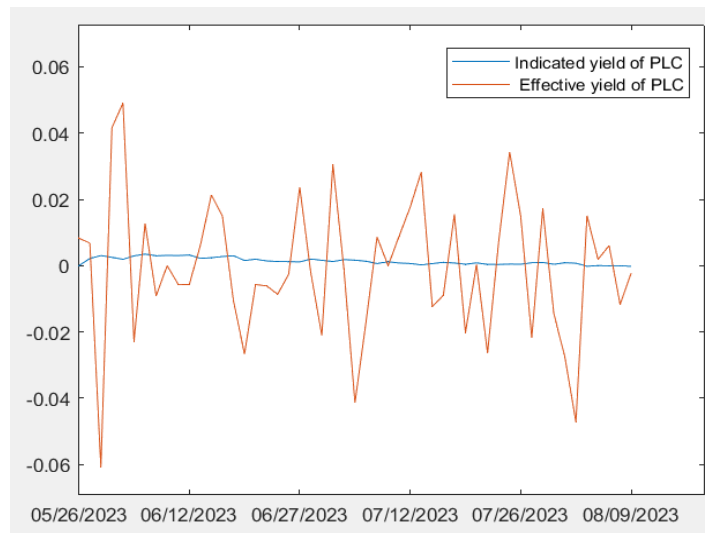


Fig. 4 Comparison between indicated yield and effective yield of PLC.

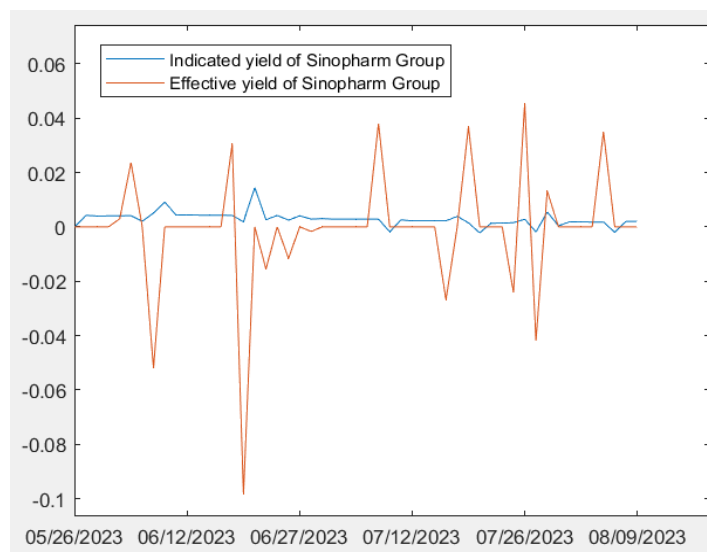


Fig. 5 Comparison between indicated yield and effective yield of CNPGC.

At the same time, in order to evaluate the error of mean square between effective value and indicated value as can be seen in Table 3, the paper calculates MSE of five assets respectively.

Table 3. The value of MSE of the final forecast and five assets

	Manulife	BOA	Apple	PLC	CNPGC
value of MSE	0.000113	0.000188	0.00014	0.000469	0.000482

From Table 3, the whole of five assets' MSE is small and performs well. Among this, the asset Manulife has the best performance as it has the smallest value of MSE, which illustrates the forecast for this asset is the most accurate. Meanwhile, the predictions of BOA and Apple perform are also great that has the relatively small value of MSE.

Based on these daily predicted returns, this study constructed time-varying asset weights. Afterwards, based on the actual return status of the assets, the daily real return of the portfolio during the prediction period was calculated. Through comparing the results minimum covariance, maximum Sharpe Ratio and SP500, the paper the comparison of the three lines from May 26th, 2023, to August 9th, 2023, shown in Figure 6:

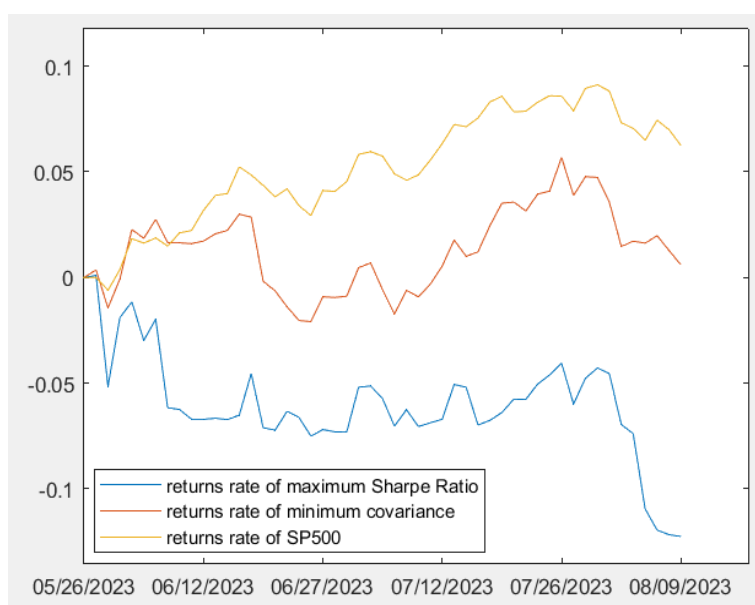


Fig. 6 Comparison of three cumulative returns

From Figure 6, it can be seen that the cumulative yield of the SP500 index achieved the highest return, indicating that in a turbulent financial context, directly purchasing market indices is a better choice.

5. Conclusion

In conclusion, this essay firstly selects five stocks including Manulife, BOA, Apple, PLC, Sinopharm Group from October 21st, 2022, to August 9th, 2023, as the researching data. Secondly, the essay predicts the last 25% of the data based on the first 75% of the data. And then based on the predicted data on each day, the paper utilizes the minimum variance model and the maximum Sharpe ratio model to make the comparison with SP500 during the same period. According to the final result, it can draw the conclusion that from the perspective of yield, the SP500 index performs best, which illustrates that buying market index funds is significant for investors to make the decisions under the volatile situation of US dollar rate hikes.

However, potential deficiencies exist. For example, in the field of forecasting, numerous models can be selected, adopting alternative models which aims to beat the market benchmark is still attracting and deserving further investigations.

References

- [1] Ren Xianling, Zhang Shiyong. Portfolio risk analysis based on kernel estimation and multiple Archimedes copula. *Management science*, 2007, 20(5): 6.
- [2] Shantian Yang, Deep reinforcement learning for portfolio management. *Knowledge-Based Systems*, 2023, 110905, ISSN 0950-7051.
- [3] Fereydooni, Ali, Barak Sasan, Asaad Sajadi Seyed Mehrzad. A novel online portfolio selection approach based on pattern matching and ESG factors. *Omega*, 2024.
- [4] Nisful Laila, Karina A. Saraswati, Himmatul Kholidah. Efficient portfolio composition of Indonesian Islamic bank financing. *Entrepreneurship and Sustainability Issues*, 2019, 34-43.
- [5] Hyo Eun Cho, Jon Jungbien Moon. Foreign direct investment vs. portfolio investment: evidence from China. *Int. J. of Multinational Corporation Strategy* (3/4), 2018, 193-217.
- [6] Zhaojun, Y., Zhizhong, L. Portfolio selection with transaction costs. *MATHEMATICAL THEORY AND APPLICATION*, 1999.
- [7] Ait-Sahalia, Y., Brandt, M. W.. Variable selection for portfolio choice. *Journal of Finance*, 2001, 56(4), 1297-1351.
- [8] Peter Joseph Mercurio, Yuehua Wu, Hong Xie. An Entropy-Based Approach to Portfolio Optimization. *Entropy* (3), 2020, 332-332.
- [9] Campbell, J. Y., Viceira, L. M.. Strategic asset allocation: portfolio choice for long-term investors. *Economic Journal*, 2010, 113(488), F408-F409.
- [10] Korn, G. A. *Advanced dynamic-system simulation: Model replication and monte carlo studies*. John Wiley & Sons, Incorporated, 2013.
- [11] Zhu Yu, Zhao Xiangmo, Xu Zhigang, Wang Runmin. Automatic generation algorithm of virtual test scene of expressway lane change for unmanned vehicles based on Monte Carlo simulation. *China Journal of Highway and Highway* (03), 2022, 89-100.