Portfolio Optimization Based on Markowitz Investment Theory and Monte Carlo Simulation

Senfan Ding[1,*]

[1]School of Computer Science and Information Engineering, Hefei University of Technology, Xuancheng, 242000, China

Abstract. In the current global economic recovery, the market still has a certain degree of volatility, in the case of volatility or bad market how to go to the portfolio and optimize it is a very critical task. In this paper, based on Markowitz's investment theory, Monte Carlo algorithm is applied to achieve portfolio optimization, which is verified by the dataset of five A-share data selected as a representative dataset in the period of global epidemic. For the model and the algorithm, good results are achieved for the objective market, and the market conditions are successfully represented; for the portfolio optimization, the high-risk and high-return portfolio with the highest Sharpe ratio is finally found, and compared with the portfolio with minimized risk, and the portfolio with the highest Sharpe ratio is concluded as the optimal portfolio for this experiment, which is the optimal portfolio for this experiment, and should not be pursued for low risk but will be the best portfolio for this experiment in the turbulent and downturn market. In a volatile and depressed market, it should not be pursued for low risk but will result in a loss of return, and it should be maximized with a certain degree of volatility.

1 Introduction

In recent years, the global economy has been recovering amid volatility, but due to the recurrence of epidemics and supply chain bottlenecks, the growth rate will gradually fall back to the norm, and inflationary pressures will intensify, increasing the risk of global "stagflation". Monetary policy tightening in developed economies, cross-border capital flow back to the United States, the dollar index and U.S. bond yields rose, offshore dollar liquidity tightened, financial market volatility risk [1].

China's economy has shown a "V"-shaped trend, with an annual growth rate of around 3.2%, under the impact of a variety of domestic and international factors that exceeded expectations [2]. The soundness of China's financial system has been fully assessed. The soundness of China's financial system was comprehensively assessed, and the financial system persisted in serving the real economy, preventing and controlling financial risks, and deepening financial reforms, creating a favourable monetary and financial environment for promoting high-quality economic development. The global financial system is still facing a high degree of uncertainty, and the banking sector is still feeling tremendous pressure in its operations. Credit risk and liquidity risk appeared particularly prominent, while sluggish demand for credit led to higher lending standards, which in turn slowed the expansion of banks' asset size. Financial market volatility showed a divergent trend, and banks encountered some obstacles in capital replenishment. In addition, the recovery of non-interest-bearing business has been unsatisfactory, resulting in a decline in the overall profitability of banks. This series of factors together constitute the complex challenges facing the banking industry today [3].

Against this market backdrop, it is particularly important for individual investors to make good portfolio decisions and optimize them accordingly. In more volatile markets, portfolio optimization is of utmost importance as it is important to adopt the right portfolio to maximize the benefits. Portfolio optimization has always been a challenging proposition in finance and management. Portfolio optimization helps in portfolio selection in volatile market conditions [4]. Numerous intelligent algorithms based on artificial intelligence and machine learning have been used for optimization, such as the Bayesian network model proposed by Shenoy is well suited for scenarios that combine qualitative and quantitative information [5]. Among the classical approaches, there are many fruitful models, like the multi-stage portfolio optimization using var as a risk measure proposed by Xu, et al. has advantages in short-term investment, but for long-term investment, there are more uncertainties [6]. Cong and Oosterlee also proposed a Monte Carlo-based multi-period mean-variance portfolio optimization model, which obtains more satisfactory allocations under the constraints of control variables for dynamic portfolio management problems [7]. Shadabfar and Cheng also used a probabilistic approach for optimal portfolio selection using hybrid Monte Carlo simulation and Markowitz model and finally achieved good results [8]. Monte Carlo algorithm has been a widely used mathematical

* Corresponding author: 2021218369@mail.hfut.edu.cn

method for stochastic simulation of complex systems in the financial field, which can simulate the consideration of changes in different influencing factors, provide a reliable reference for asset allocation, and finally achieve portfolio optimization. Some traditional techniques, the efficient frontier and the Monte Carlo simulation are classical methods used in portfolio optimization and are the basis of the methodology of this paper [9, 10].

In this paper, it is through the Monte Carlo model, based on Markowitz's portfolio theory by calculating the Sharpe ratio and other indicators to finally find the efficient frontier to achieve the optimization of the portfolio, and at the same time, the corresponding visualization is carried out to prove that the modeling algorithm has usability and reliability on the given data set.

2 Methodology

2.1 Data selection

The dataset studied in this paper comes from tushare, imported through the code writing interface, and studied the data in the five stocks '002955.SZ' - Honghe Technology, '300033.SZ' - Flush, '601127.SH' - Sailix, '002594.SZ' - BYD, '603000 .SH' - People's Daily, five stocks for the period from January 1, 2019 to December 31, 2022, for this particular period, the characteristics of the A-share data also simultaneously make the results more relevant.

2.2 Modeling

2.2.1 Markowitz portfolio theory

The "mean and variance portfolio model" proposed by Markowitz in 1952 is to analyze the mean and variance of the returns of individual stocks in a portfolio of assets under the conditions of prohibition of securities financing and absence of risk-free borrowing, to find out the effective boundary of the portfolio, i.e., the portfolio with the smallest variance under a certain level of return, and to select a portfolio from it. investment portfolio.

The methodology for realizing this model includes the following steps.

Definition of Return and Risk: The expected return of a portfolio as a measure of return and the variance as a measure of risk.

Risk-Return Tradeoff: Finding the portfolio that maximizes expected return for a given level of risk or minimizes risk for a given level of expected return by solving a quadratic programming problem.

Analysis of Correlation Characteristics: A measure that analyzes the extent to which a portfolio's expected and probable returns deviate from their expectations, using variance as the primary tool of analysis.

Identification of Efficient Portfolios: Identify theoretically feasible efficient portfolios by properly analyzing a security's expected return, the variance of the return, and the interrelationship of the security's return with other securities.

$$\begin{aligned}
&\tilde{r} = (\widetilde{r_1}, \ldots, \widetilde{r_n}) \\
&E(\tilde{r}) = \left(E(\widetilde{r_1}), \ldots, E(\widetilde{r_n})\right) \\
&\sigma^2(\tilde{r}) = \left(\sigma^2(\widetilde{r_1}), \ldots, \sigma^2(\widetilde{r_n})\right) \\
&V = \left(\mathrm{cov}\left(\widetilde{r_i}, \widetilde{r_j}\right)\right) \\
&W = (W_1, \ldots, W_n), W^{'}1 = 1W^{'} = 1 \\
&\widetilde{r_p} = W^{'}\tilde{r} \\
&E\left(\widetilde{r_p}\right) = W^{'}E(\tilde{r}) \\
&\sigma^2\left(\widetilde{r_p}\right) = W^{'}VW
\end{aligned} \tag{1}$$

Portfolio returns:

$$\mathrm{r_p} = (\mathrm{W_1} - \mathrm{W_0})/\mathrm{W_0} \Rightarrow \mathrm{W_0}\left(1 + \mathrm{r_p}\right) = \mathrm{W_1} \tag{2}$$

The utility function of the investor:

$$\mathrm{u(W)} = \mathrm{a} + \mathrm{bW} + \mathrm{CW^2} \tag{3}$$

$\tilde{r}$ represents the gain vector, and $E(\tilde{r})$ represents the expectation vector, and $\sigma^2(\tilde{r})$ represents the variance vector, and V represents the covariance matrix. W represents the weight vector. $\widetilde{r_p}$ represents the portfolio return. $E\left(\widetilde{r_p}\right)$ represents the portfolio expectation, and $\sigma^2\left(\widetilde{r_p}\right)$ represents the portfolio variance, and the above are the formulas used in Markowitz's theory to compute it.

2.2.2 Monte Carlo algorithm

The basic idea: when the problem is to solve the probability of the occurrence of a random event, or the expected value of a random variable, through some kind of "experiment" method, the frequency of the occurrence of such events to estimate the probability of this random event, or to get some of the numerical characteristics of this random variable, and will be the solution to the problem.

First of all, it is necessary to set the necessary parameters of the model, including the standard deviation of the simulated returns σ_{δ}, the total amount of funds M, the benchmark interest rate μ and so on.

Generate a sample of random numbers by Monte Carlo algorithm for a given stock market:

Determine the n dimensional white noise vector based on the model parameters:

$$\epsilon = (\epsilon_1, \cdots, \epsilon_n)^T \tag{4}$$

Calculations $\sigma_{\epsilon}^{-2}\exp\left(-r\tau + 0.5\tau^{'}V_{\epsilon}(\tau)/\lambda\right)$ where r is the prime rate and τ is the time scale. $V_{\epsilon}(\tau)$ is the covariance matrix between the two future time scales and λ is the parameter of the Poisson harmonic curve.

For each time scale t, compute the Monte Carlo distribution corresponding to the expected return.

$$r_{ij}(t) = \frac{1}{\sqrt{N}}\left(\mu_j + \sum_{l-1}^{N-1} \frac{1}{\sqrt{N-l}}\psi\left(l - \frac{(t-1)N}{N}\right)\left(\sigma_\epsilon r_j(t+(l-1)\Delta t) + \sqrt{N-l}\sigma_i r_j(t)\right)\right) + \epsilon_j(t) \quad (5)$$

Calculate the rate of return:

$$Rate\ of\ return = \frac{current\ price - purchase\ price}{purchase\ price} \quad (6)$$

Calculate volatility:

$$\sigma_p = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(P_i - \bar{P}\right)^2} \times \sqrt{252} \quad (7)$$

Assume that $P_1, P_2 \dots P_n$ is a continuous stock price and that 252 days is the number of trading days in a year.

Calculate the Sharpe ratio:

$$R_p = \left(\frac{current\ price / puchase\ price}{n}\right)^{\frac{1}{year}} - 1 \quad (8)$$

$$Sharp\ ratio = \frac{R_p - R_f}{\sigma_p} \quad (9)$$

Where R_p is the annualized rate of return, R_f is the risk-free rate, andσ_p is the annual volatility.

Sharpe ratio, volatility, and yield are three of the most important metrics essential to the application of the Monte Carlo algorithm based on Markowitz's investment theory, and the final efficient frontier will be derived from a combination of them.

3 Empirical results

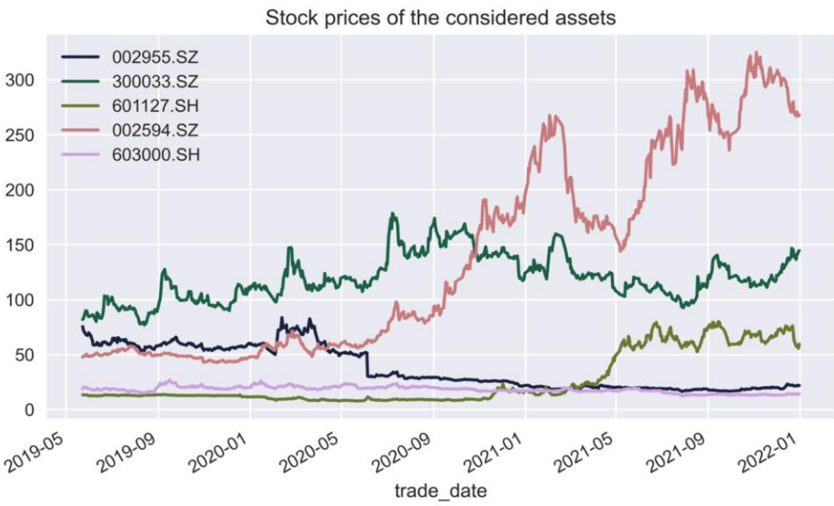


Fig. 1. Visualization of data for several stocks (Photo/Picture credit: Original).

For the overall data visualization between May 2019 and January 2022 as shown in Figure 1, it can be seen that until May 2020 the stock prices of the five stocks were significantly lower due to the objective impact and influence of the global epidemic on the financial markets at that time. After that, 002594.SZ (BYD) with a large surge, 300033.SZ (Flush) is a slight increase to maintain a certain degree of stability; in January 2021 after 601127.SH (Sailix) also had a more significant increase, but after May 2020 002955.SZ (Hong and Technology) and 603000.SH (People's Daily) but a slight downward trend.) are on a slight downward trend. Overall, these five stocks are unstable and highly variable over the time implied by the dataset, a characteristic that leads to some error or uncertainty in finding efficient frontiers later on, and will require the Monte Carlo algorithm to simulate much more often.

Table 1. Portfolio Performance (a).

Start date	End date	Total months	Annual return	Cumulative returns	Annual volatility	Sharp ratio	Calmar ratio	Stability
2021-12-30	2019-05-23	30	-8.4%	-20.0%	32.4%	-0.11	-0.21	0.45

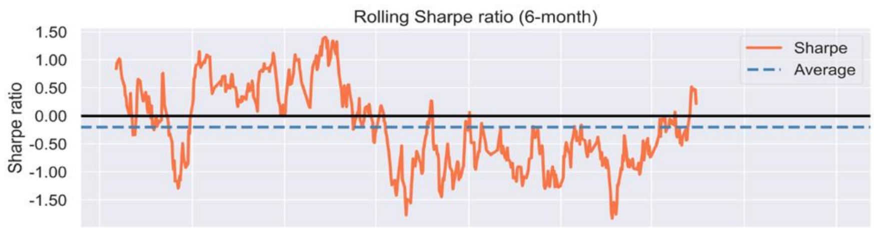
Table 2. Portfolio Performance (b).

Max drawdown	Omega ratio	Sortino ratio	Skew	Kurtosis	Tail ratio	Daily value at risk
-40.7%	0.98	-0.17	0.70	2.53	1.00	-4.1%

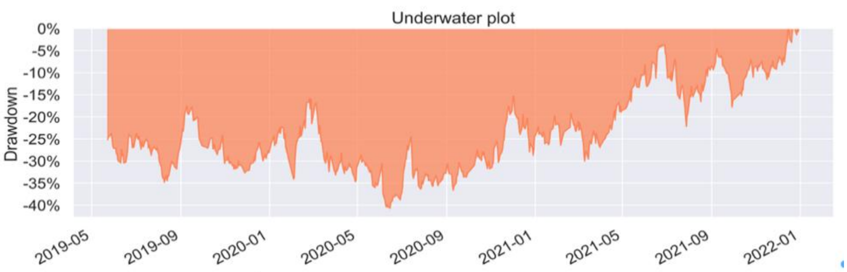
Derived from a function in python that creates a simple portfolio performance are shown in Tables 1 and 2. Firstly, the Cumulative and Annual Returns are both negative and the Cumulative Return of -20.0% is a testament to the volatility of the financial markets and the high level of disruption to all aspects of trading during this period. The annual volatility of 32.4% indicates a high level of risk; the Sharpe Ratio is negative, indicating poor balance and excessive risk; the Kalmar Ratio is negative, indicating poor return performance based on the maximum retracement; the maximum retracement of -40.7% is a good indication of a large relative peak-to-trough loss over a period of time that experienced a large movement; and the Omega Ratio is close to 1, indicating a positive return over a period of time that experienced a large movement. suggests that the portfolio performs better during the time period of positive returns and is a more appropriate portfolio; the Sortino Ratio is negative, suggesting that it performs poorly during downside volatility; the skewness of 0.7 indicates some right skewness, suggesting that the chances of a positive return are slightly higher than a negative one; the kurtosis is 2.53, which suggests that the portfolio has a high degree of spikiness relative to a normal distribution, with a potentially heavier tail; and a negative value of the daily value-at-risk is indicates that the maximum degree of loss is not low.

Overall, there was a high exposure to risk during the backtesting period and an overall underperformance of returns (see Figure 2). The investment strategy needs to be carefully evaluated and adjustments need to be made to improve portfolio performance and risk management.

(a) Cumulative return



(b) Rolling Sharpe ratio



(c) Underwater plot

Fig. 2. Performance of investment portfolio (Photo/Picture credit: Original).

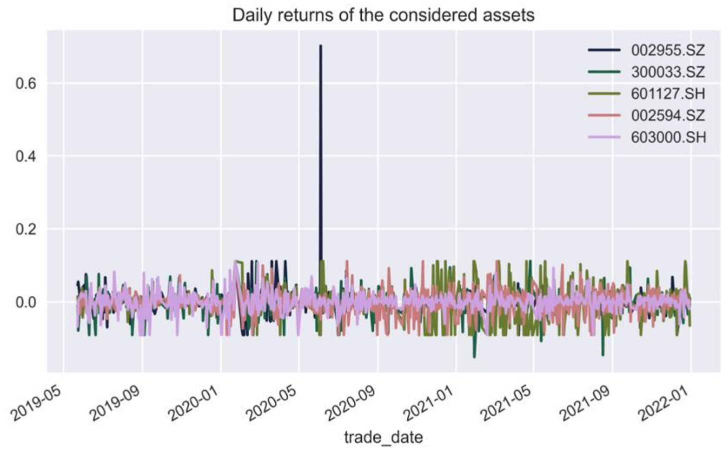


Fig. 3. Daily returns of the considered assets (Photo/Picture credit: Original).

As can be seen from Figure 3 for this period, several stocks have returned above and below 0 and have never had higher returns, again illustrating the slump and turmoil in the market during this period. Therefore, portfolio and portfolio optimization will become even more important.

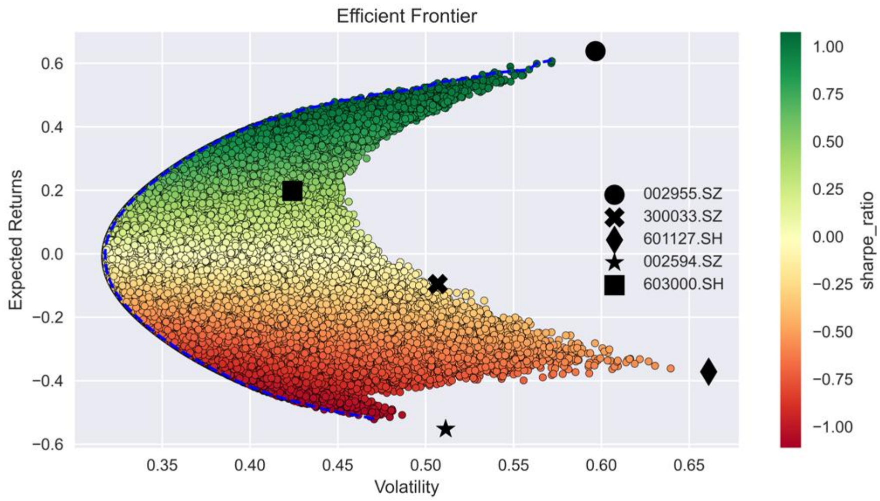


Fig. 4. Efficient frontier map corresponding to five stocks (Photo/Picture credit: Original).

Each point in Figure 4 represents a portfolio, the x-axis represents volatility, and the y-axis represents the income i.e. return. Also portfolios with higher Sharpe ratios are usually more efficient. Meanwhile, the black dots corresponding to the five stocks are their performance as separate assets in terms of return and risk. It can be seen that for the stock 002955.SZ, it has the best return despite the higher volatility. As can be seen from Figure 1, its stock price is relatively stable and has declined somewhat, which suggests that it is and is more likely to contribute to a high-quality portfolio; compared to 002955.SZ, 603000.SH and 300033.SZ ensured a certain level of return with lower volatility, while the two 002594.SZ and 601127.SH performed SZ and 601127.SH are poor, in the case of volatility is not low return is still not optimistic.

From here it is already basically clear how the five stocks corresponded to the situation during this period, in general the performance is relatively bad, the higher volatility represents the instability of the market, the low returns indicate the slump during this period, the pursuit of certain returns can only be carried out in a risky situation.

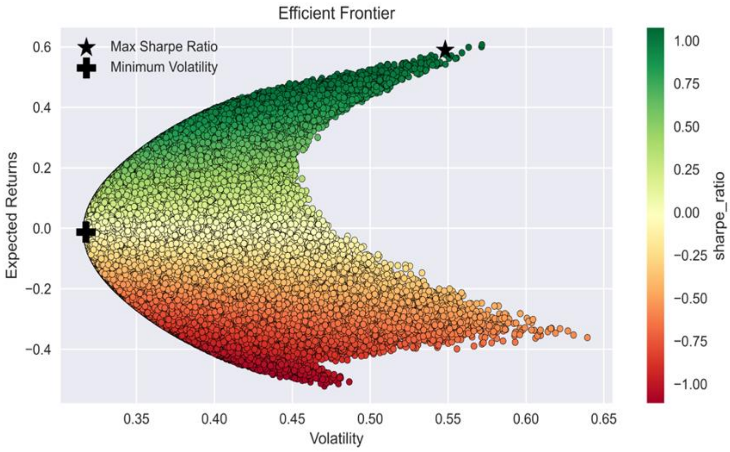


Fig. 5. Two key portfolios (Photo/Picture credit: Original).

Figure 5 is the most important one in the whole portfolio optimization and the result of the portfolio optimization. In the experimental calculations carried out in this paper, the proportion of each stock in the portfolio corresponding to the maximum Sharpe ratio and the proportion of each stock in the minimum risk are calculated as follows:

Maximum Sharpe Ratio portfolio: returns: 58.97% volatility: 54.82% sharpe ratio: 107.57

Weights: 002955.SZ: 89.75% 300033.SZ: 0.57% 601127.SH: 0.21% 002594.SZ: 0.16% 603000.SH: 9.31

Analyzing this result, for this particular period, the return alone at 58.97% is already among the better performers in the market, and a volatility higher than 50% indicates a higher risk and requires careful consideration. In terms of its weight allocation, the vast majority of the percentage is given to 002955.SZ, while the others account for very little. It can be seen that Honghe Technology has a better performance in Figure 4, with the highest return but also higher risk, which makes the portfolio when the Sharpe Ratio is the highest when the portfolio is exposed to greater risk and high sensitivity for this stock, which will affect the performance of the entire portfolio once it receives the impact may lead to the worst case scenario of high risk and low return. 603000 has the same weighting as 002955.SZ in Figure 4, except 002955.SZ has the same weighting as 002955.SZ, which has the same weighting as 002955.SZ. Figure 4 has the best overall performance except for 002955, which guarantees low volatility and relatively high returns, but if extreme returns are sought, it should be the second target, and is therefore the second in the portfolio. For the other underperforming stocks, it is correct and factual to consider a lower percentage. Overall, the returns were significant but accompanied by a high degree of concentration and risk.

For a portfolio that minimizes risk, the calculations are as follows:

Minimum Volatility portfolio : returns: -1.25% volatility: 31.75% sharpe_ratio: -3.95

Weights: 002955.SZ: 17.73% 300033.SZ: 15.60% 601127.SH: 12.50% 002594.SZ: 22.93% 603000.SH: 31.24%.

In this portfolio, the order of percentage is 603000>002594>002955>300033>601127, which is in line with the return-risk profile of each stock in Figure 4. For 603000, its risk is the lowest and its return is guaranteed, but for 002594, it is actually biased, its performance is the worst, but in order to ensure a lower risk still occupies a large proportion, which directly leads to the final overall return is negative and the indicators are not good. From these results, it is easy to see that the portfolio actually underperformed as the returns were negative even when risk was minimized and the volatility was not kept low and the Sharpe ratio was negative. Although a more even distribution of stocks would have avoided concentration, it was not a good portfolio with the goal of high return and low risk. During this period of time, even if one were to aim for low risk, one would still incur losses. Overall, it was not a good investment choice to pursue low risk at the cost of poorer returns.

In summary, the impact of the global epidemic and certain international factors, these five stocks have been affected by the more serious, and they as a few of the typical A-share stocks also shows that the market downturn at the time, in the optimization of the portfolio can only get high-risk and high-return portfolio, and did not achieve the ideal low-risk and high-return optimal portfolio.

4 Conclusion

Based on the above experimental data, it can be concluded that the Monte Carlo algorithm based on Markowitz's investment theory successfully demonstrates the financial market downturn and turbulence during this special period and optimizes the portfolios for the five stocks, and finally obtains the optimal high-risk-high-return portfolio with the highest Sharpe ratio. For this special period of time, this paper does not introduce the risk factor model, which may lead to a certain degree of randomness or bias in the results, the future will consider adding the risk model to the portfolio optimization, so that in the case of multiple condition constraints can still be found in a variety of market environments, that is, to optimize the portfolio; in addition to the selection of the data set is also very important, this experiment is specially Selected this special period of data for portfolio optimization, obviously, there is a certain degree of difficulty, and finally arrived at the minimization of risk portfolio has a certain degree of irrationality, in the face of such a market situation still want to pursue low returns and high yield is obviously unrealistic, and even in the case of minimizing the risk of the yield is still negative, it can be imagined that this period of time is not optimistic about the market situation, should be taken more appropriate investment measures. More appropriate investment measures and allocations should be adopted.

Algorithmic models, also due to the specificity of the market, resulting in even the Monte Carlo algorithm in the number of simulations as high as 10 to the seventh power still failed to find a better portfolio, so consider the applicability of the model for such a special market is not particularly ideal. Currently many intelligent artificial intelligence and machine learning model algorithms are more advanced and reasonable, the future will carry out research on the application of deep learning or machine learning algorithms model for portfolio optimization, especially in the processing of time series data above the more advantageous, the results will be more accurate.

References

1. Bank of China. 2023 Fourth Quarter Economic and Financial Outlook Report. https://www.boc.cn/aboutboc/bi1/202310/t20231008_23843447.html.
2. Bank of China. Economic and Financial Outlook Report for the Third Quarter of 2023. https://www.boc.cn/aboutboc/bi1/202211/t20221130_22168786.html.
3. Bank of China. Economic and Financial Outlook Report for the Third Quarter of 2023. https://www.boc.cn/aboutboc/bi1/202307/t20230703_23314068.html.
4. A. Gunjan, & S. Bhattacharyya. AIR, 56, 3847-3886 (2023).
5. C. Shenoy, & P. P. Shenoy. Bayesian network models of portfolio risk and return. MIT Press (2000).
6. C. Xu, J. Wang, & N. Shiba, IJIC, 3(3), 709-724 (2007).
7. F. Cong, & C. W. Oosterlee. JEDC, 64, 23-38 (2016).
8. M. Shadabfar, & L. Cheng. AEJ, 59(5), 3381-3393 (2020).
9. M. J. Best, & J. Hlouskova. MMOR, 52(2), 195-212 (2000).
10. M. Pedersen. SSRN 2438121 (2014).