

Portfolio Allocation Using Monte Carlo Simulation and ARIMA Model Targeting Chinese Companies Trading on the US Stock Exchange

Yuxuan Chen^{1,*†}, Tingsong Li^{2,*†}, Leyao Lin^{3,*†}
¹*s192092f@yokohama-cu.ac.jp, ²*tingsong.li@case.edu, ³*leyaol@sfu.ca

¹Faculty of Science, Yokohama City University, Yokohama, Kanagawa, Japan, ²Weatherhead School of Management, Case Western Reserve University, Cleveland, OH, USA, ³The Department of Statistic and Actuarial science, Simon Fraser University, Burnaby, BC, Canada

*Statement: All the authors are the co-first authors, and they are the co-responding authors.

†These authors contributed equally.

Abstract— Against the background of the COVID-19 pandemic, stock markets around the world have been greatly impacted. The Chinese stocks face the risk render by pandemics and face the crackdown from the Chinese government. This study aims to evaluate the portfolio allocation with the U.S.-listed Chinese stocks. In this paper, first, we calculate the Sharpe ratio of all Chinese companies that trade on U.S. major stock markets and choose six stocks with the largest Sharpe ratio to do asset allocation. The stocks that we choose are SPI, AACG, BTB, MOXC, DQ and RENN. Then we plot the efficient frontier by the Monte Carlo simulation method to find the maximum Sharpe Ratio portfolio and minimum volatility portfolio, and view the performance of each portfolio. We get a similar composition of stocks in these two portfolios, which have a high Sharpe ratio and high volatility. In conclusion, we can say that allocations with pure Chinese stocks may not be suitable for risk-averse people. Finally, we use the ARIMA model to predict the future 21 days returns of each stock. We calculate the AIC to evaluate the model and plot the prediction of stocks' returns and compare it with the real value which the stocks show. As a result, we find that the ARIMA model only has limited accuracy in predicting the future.

Keywords-component; Portfolio allocation; Chinese Corporation Analysis; US stock market; Monte Carlo simulation; ARIMA model

1 INTRODUCTION

In such an era as the COVID-19 pandemic has been a global menace to human beings, the global economy has been greatly affected, primarily for retailing and tourist industries, and the negative impact may further be amplified with the uncertain trend of the pandemic [1]. As a result of economic globalization, all of the stock markets around the world have been impacted by the COVID-19 [2]. Especially for Chinese stocks, not only are they influenced by the pandemic but are also facing severe political pressure in China [3]. In this case, it is a challenge for us to form a portfolio with US-listed Chinese stocks, given the aforementioned risks rendered by the pandemic as well as the crackdown from the Chinese government [4].

Since 1994, there is an increasing number of Chinese stocks listed in US major stock markets. As per the data from USCC, as of May 2021, there have been 248 China-related companies

listed in the US, with a total market capitalization of 2.1 trillion dollars [5]. The main reason is that those companies are attracted by a higher degree of market activity. Also, they are seeking more opportunities to familiarize themselves with global investors as well as avoiding the relatively strict regulation in China [6]. Indeed, investors have been attracted by the integration of the US and Chinese stock markets as of decades ago [7]. However, we did not find plenty of published research focusing on this particular group of stocks.

Regarding the high risks to Chinese stocks listed in US stock markets, we would like to maximize the return of the portfolio by maximizing the return of the unit risk. One of the most common ways to measure the performance of assets under risk is the Sharpe ratio, which calculates the return of a unit risk [8]. For portfolio optimization methodology, we are inspired by the Monte Carlo simulation, which is extensively used in portfolio management. Similar to the method of exhaustion, Monte Carlo simulation assigns different weights to each asset and derives as many conditions as possible, which has been seen as an ideal way to solve portfolio management problems [9]. We also utilize the efficient frontier to acquire efficient combinations of weights with the variety of returns and risks for a particular portfolio, which has proved to be useful since the last century [10]. For trend forecasting, we referred to the autoregressive integrated moving average (ARIMA) model. ARIMA is a widely used time series model. It is based on the autoregressive moving average model with integration to make the time series forecast more stationary and more robust in a short-term forecast horizon. Many scholars used ARIMA in the field of finance and economy, such as forecasting the future value of a currency and future demand for supply chain management [11, 12]. However, all those methods are flawed in that there is still no research showing the application of those methods on US-listed Chinese stocks.

In this paper, we will select 6 assets with the Sharpe ratio as our portfolio. We will then figure out the specific weight of each asset in the portfolio by Monte Carlo simulation and select one combination of weights that is suitable for us. We would finally do the trend forecast via the ARIMA model to predict our portfolio.

2 METHOD

2.1 Data preparation

We get a list of Chinese companies that trade on the U.S stock market from Wind [13]. After that, we select the company which is listed on market for at least 4 years. Then we got those companies' historical data from the website Yahoo finance [14]. Each company's daily stock price data is selected from January 1, 2017, to June 30, 2021. We annualized daily return and daily return's standard deviation then calculated the annual sharp ratio of all those companies to pick the sock. The annual sharp ratio is defined as:

$$SharpeRatio = \frac{E(R_a) - R_f}{\sqrt{Var(R_a)}} \quad (1)$$

where R_a is the annualized asset return, R_f is the risk-free return which is assumed as constant. According to the date after calculation, we selected the six companies with the highest Sharpe ratio. The six companies are SPI Energy Co., Ltd., ATA Creativity Global, Bit Brother Limited,

Moxian, Inc., Daqo New Energy Corp., and Renren Inc. For more information about those companies, see Table 1. The calculation and analysis through Python.

TABLE 1. THE STOCK INFORMATION

Stock Symbol	Company	Sector	Industry	Market Cap. (Until August 12, 2021, \$)
SPI	SPI Energy Co., Ltd.	Technology	Solar	128.14M
AACG	ATA Creativity Global	Consumer Defensive	Education&Training Services	88.76M
BTB	Bit Brother Limited	Consumer Cyclical	Restaurants	41.046M
MOXC	Moxian, Inc.	Communication Services	Internet Content & Information	148.543M
DQ	Daqo New Energy Corp.	Technology	Semiconductor Equipment & Material	4,162B
RENN	Renren Inc.	Consumer Cyclical	Auto & Truck Dealerships	245.718M

2.2 Portfolio optimization through Monte Carlo

Then we plot the efficient frontier. We used the Monte Carlo method to find efficient frontier and portfolio allocation. The first step of this method is generating lots of random numbers as the weight of each stock. Then we constructed many different asset allocations. For each asset allocation, we calculated the annual expected returns, volatility, and sharp ratio. The annual expected return can be calculated by equation (2),

$$ExpectedReturn = E(R_p) = \sum w_i r_i \quad (2)$$

where w_i is the weight of the $asset_i$. The volatility of each allocation is the covariance matrix of each stock return multiply by the weight and can be calculated by equation (3).

$$Volatility = \sigma^2(r_p) = \sum \sum w_i w_j cov(r_i, r_j) \quad (3)$$

2.3 Trend forecast

2.3.1 Pre-process

We used the compound return for trend forecast and time series analysis. The compound return is defined as:

$$CompoundReturn = \ln\left(\frac{V_{t+1}}{V_t}\right) \quad (4)$$

where V_t is the price at time t. Then, we combined the histogram and scattered plot to check whether the compound return of each stock is invariant or not. For the histogram method, each

stock data will be split in half, and then compare the distribution of 2 plots. According to the plot of each stock, the compound return of all stocks is invariant.

2.3.2 ARIMA model for forecasting

We use the autoregressive integrated moving average (ARIMA) model to predict the future trend of each stock. The ARIMA (p, d, q) model's concise form is shown as:

$$\phi(B)(1 - B)^d x_t = \theta(B)w_t \quad (5)$$

where B is the backshift operator; x_t is the time series data; w_t is "white noise" random shock variable that follows an independent and identically distribution $(0, \sigma_w^2)$; $\phi(B) = 1 - \phi_1 B^1 - \phi_2 B^2 - \dots - \phi_p B^p$ and $\theta(B) = 1 + \theta_1 B^1 + \theta_2 B^2 + \dots + \theta_q B^q$ represent the autoregressive operator and moving average operator with order p and q ; $(1 - B)^d$ is the difference process with order d [15]. The order of autoregressive process (p) and moving average process (q) could be obtained by partial autocorrelation (PACF) graphs and autocorrelation (ACF) graphs. For a more accurate result, we fitted the various model with different order combinations. The range of order p and q is from 0 to 5, and the range of d is 0 to 3. We used the Akaike information criterion (AIC) to compare each model. The AIC can be calculated by the formula (6),

$$AIC = 2k - 2\ln(\hat{L}) \quad (6)$$

where k is the number of estimated parameters in the model and \hat{L} is the maximum value of the likelihood function for the model. The model with the lowest AIC fitting the trend best. We also do the residual analysis to check whether the residual is normally distributed. By observing the time series plot, which includes plots of residual data, ACF, PAF, QQ and probability, it is proved that the residual is independent.

3 RESULTS AND DISCUSSION

3.1 Efficient frontier and portfolio allocation

In figure 1, the blue line shows the efficient frontier of the 5 stocks that we chose. The x-axis shows volatility, and the y-axis shows the expected returns on the left side and the Sharpe ratio on the right side. The marks represent the volatility and Sharpe ratio of each stock when the stock's weight is 100% in the portfolio. We can see that most dots have high volatility and are red, which means most combinations have a low Sharpe ratio.

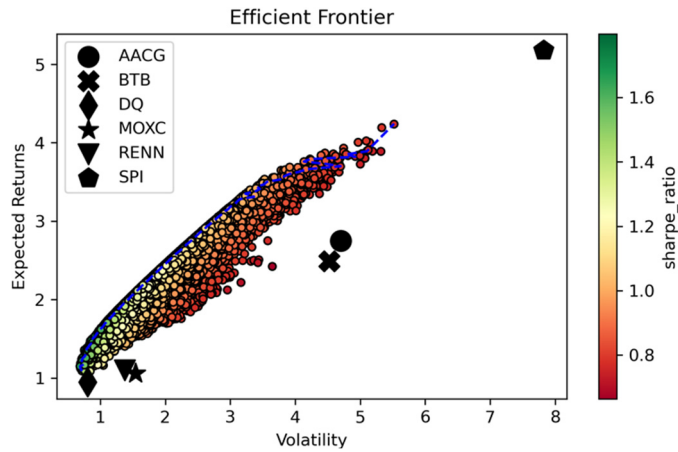


Figure 1. Efficient Frontier 1.

In figure 2, the max Sharpe ratio portfolio and minimum volatility portfolio are shown by the marks on the efficient frontier. The details of each portfolio are shown in the figures below.

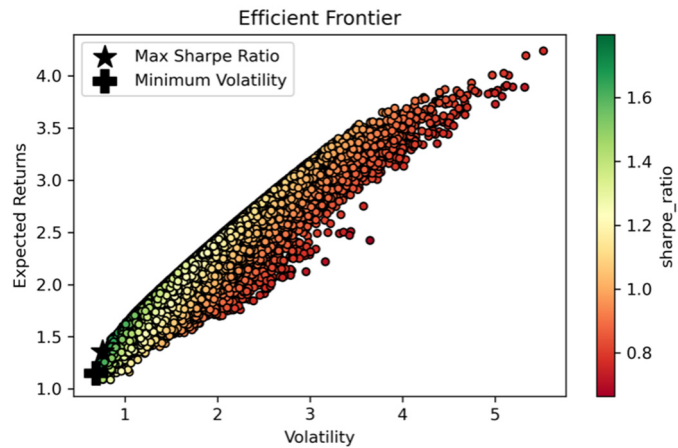
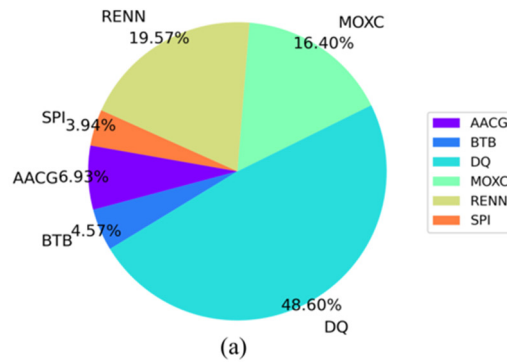


Figure 2. Efficient Frontier 2.

Figure 3, (a) represents the percentages of each stock in the maximum Sharpe ratio portfolio and (b) represents the percentages of each stock in the minimum volatility Sharpe ratio. The composition of these two portfolios, which are oriented in different strategies, is similar. The result may be due to the method which we used to choose stocks by the highest Sharpe ratio.

Maximum Sharpe Ratio portfolio



Minimum Volatility portfolio

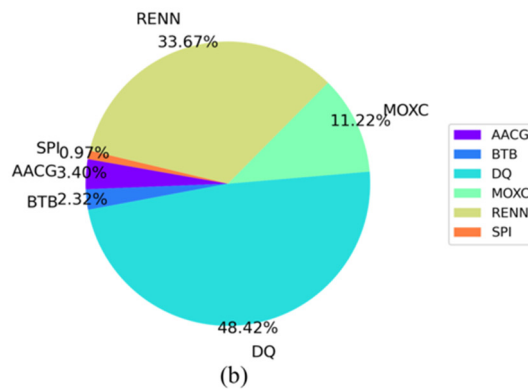


Figure 3. Weights of each stock in portfolios

Table 2 shows the performance of the two portfolios. Because of the similar allocation of stocks, they show a similar result correspondingly. The volatility of both portfolios is typically high even they have a high Sharpe ratio at the same time. This result shows that high risk always comes with high returns.

TABLE 2. PERFORMANCE OF PORTFOLIO

	Returns	Volatility	Sharpe ratio
Maximum Sharpe Ratio portfolio (SR)	135.79%	75.55%	179.73%
Minimum Volatility portfolio (MV)	114.99%	68.80%	167.15%

3.2 Results of ARIMA-based trend prediction

Table 3 shows the AIC of the model that we use to predict the stocks, which can show us how well the model fit. We got a minus relatively low value. Therefore, we can say that the model we used fit well.

TABLE 3. AIC OF STOCK'S MODEL

	SPI	AACG	BTB	MOXC	DQ	RENN
ARIMA	-1069.2712	-1966.7606	-1867.201252	-2204.1646	-3455.9884	-2656.1765

3.3 Price Prediction based on Arima

In figure4, the x-axis shows the date, and the y-axis shows the six stocks' compound return (log return). Graph (a) shows the predicted compound return which we calculated by the ARIMA model and graph (b) shows the real compound return which is calculated by the yahoo finance data. According to the result, the patterns are not as similar as the training data set. The model that we used does not have so much accuracy in prediction and still has the potential to be improved to obtain better precision. The reason may be that the sample size is too small. In future work, we need to increase the sample size or try other models.

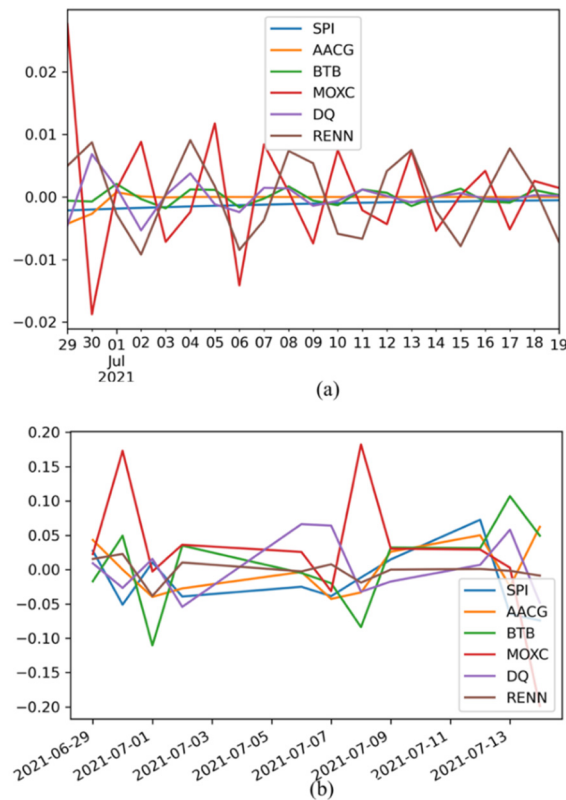


Figure 4. Prediction of the stocks

4 CONCLUSION

In this paper, six US-listed Chinese stocks were selected to form different asset portfolios. The selection is based on the annual sharp ratio. The optimal portfolio allocation is selected by efficiency frontier, calculated based on the Monte Carlo method. According to the efficiency frontier plot, no allocation with single stock is the optimal allocation. The final portfolio allocations are asset allocation with the maximum sharpe ratio (SR) and minimum volatility (MV) allocation. Both have a high return and high risk, which shows that these Chinese companies may not be good assets. The allocation with pure Chinese companies' stock is not good for the risk-averse people and may need to add some other stocks to balance the risk. The ARIMA model has some accuracy in estimating the future, but not much for the forecasting model. This can be shown by comparing trend plots of predicted data and real data. The patterns of forecasting return and real return are relatively similar since both fluctuate around zero. However, the real data is more volatile than the forecast data. The model predicts MOXC's stock most accurately for single stock forecasting and SPI's stock least accurately. The SPI's return shows as a straight line in the predicted data, but it shows more volatility in the real data. Thus, the ARIMA model could only be used as a reference for return prediction. To further reinforce this research, future research could examine the portfolio allocation with different stock selection indicators (which is the shape ratio in this paper). Increasing the size of data and using an enhanced model with more accuracy could also be considered in future work.

REFERENCES

- [1] Leduc, Sylvain, and Zheng Liu. "The Uncertainty Channel of the Coronavirus." FRBSF Economic Letter, March 2020, pp.1-5.
- [2] Ashraf, Badar Nadeem. "Stock Markets' Reaction to Covid-19: Cases or Fatalities." April 2020.
- [3] Wu, Xianbo, and Xiaofeng Hui. "The Impact of COVID-19 on the Dependence of Chinese Stock Market." *Discrete Dynamics in Nature & Society*, May 2021, pp.1–11.
- [4] Cheung, Rachel. "China's Crackdown on Tech Platforms Is Just Getting Started." *World Politics Review* (Selective Content), Aug. 2021, pp.1.
- [5] *U.S.-China Economic and Security Commission*. "Chinese Companies listed on U.S. Major Stock Exchanges." <https://www.uscc.gov/research/chinese-companies-listed-major-us-stock-exchanges>
- [6] Ross, John. "Risks of U.S. Listings by China's Companies." *China Today*, vol. 64, no. 1, January 2015, pp. 48–49.
- [7] Brooks, Robert, et al. "Testing the Integration of the US and Chinese Stock Markets in a Fama-French Framework." *Journal of Economic Integration*, vol. 24, no. 3, 2009, pp. 435–454.
- [8] Sharpe, William F. "Mutual Fund Performance." *The Journal of Business*, vol. 39, no. 1, 1966, pp. 119–138.
- [9] Detemple, Jérôme B., et al. "A Monte Carlo Method for Optimal Portfolios." *The Journal of Finance*, vol. 58, no. 1, 2003, pp. 401–446.
- [10] Dybvig, Philip H., and Stephen A. Ross. "Portfolio Efficient Sets." *Econometrica*, vol. 50, no. 6, 1982, pp. 1525–1546.
- [11] Amaira, Bouzid. "Future Value of the Tunisian Dinar against the Euro: Arima Approach." *Journal of Academic Research in Economics*, vol. 13, no. 1, Mar. 2021, pp. 149–162.

- [12] Hsiao, J. M., and C. J. Shieh. "Evaluating the Value of Information Sharing in a Supply Chain Using an ARIMA Model." *International Journal of Advanced Manufacturing Technology*, vol. 27, no. 5/6, Jan. 2006, pp. 604–609.
- [13] Wind website: <https://www.wind.com.cn/>
- [14] Yahoo Finance website: <https://finance.yahoo.com/>
- [15] Shumway, Robert H., and David S. Stoffer. *Time Series Analysis and Its Applications*. Springer International Publishing AG, 2017.