Correlation Analysis and Optimal Portfolio in UK and US Stock Markets- Based on Copula-VaR Function

Yuxin Long *

Southwestern University of Finance And Economics, Chengdu, Sichuan, China

*Corresponding author: 42050074@smail.swufe.edu.cn

Abstract. Studying the correlation between the various stock and selecting the risk of the optimal portfolio are the key topics of current research. In the macro field, it is a further problem with delving into the correlation in the different stock markets. The researchers found that the copula function can break the restrictions of traditional normal hypothesis and linear correlation, which has a significant advantage in studying the correlation of the stock market to choose the optimal portfolio. This study employed S&P500 to represent the American stock market, FTSE to represent the British stock market, model and select the best copula function through copula function, and applied risk management (VaR) analysis to obtain the best portfolio to select British and American stocks. The research results show that the student's t Copula function of the five copula functions best fits the two stock markets in Britain and the United States. The correlation between them reveals a descendent trend from the extreme upward stage to the extreme downward stage. Finally, it is suggested that investors choose 60% FTSE and 40% S&P500 for their investment portfolio.

Keywords: Copula; VaR model; GARCH; Stock Market.

1. Introduction

Under the influence of the background of the epidemic and the trend of financial globalization in recent years, the stock markets of various countries have been subject to different degrees of volatility, they are interconnected and interacted. Among all, the stock markets in the United States and the United Kingdom, as capitalized markets, large-scale, mature, and standardized operated, were taken as samples to further explore their correlation and linkage. Correlation research in financial markets plays an important role in investment portfolios. Shi Daoji & Li Fan [1] applied the copula function to model the dependence structure of the stock exchange on the basis of the data of the S&P index and FSTE from 2000 to 2005, conducted VaR analysis and determined the optimal portfolio to reach utility maximization. This experiment aims to verify their structured correlation law, observe the interdependence changes between the two stock markets in recent years, reestablish a new model, compare the distinction, and identify the possible causes to provide investors with updated portfolio recommendations. This experiment selects S&P500 to represent the average stock prices in the American market and FTSE to represent England., applying the Copula-GARCH model to overcome the traditional assumption of normality and the drawbacks of linear correlation [2] while analyzing the correlation. This study opts for the best-fitting copula function and exploits risk management (VaR) analysis to obtain the best portfolio.

2. Related work

Copula Function is a mathematical method that connects the univariate marginal distribution with its multivariate joint distribution as correlation analysis and multivariate statistical analysis tool. Copula was originally proposed by Sklar [3], who defined that the correlation properties of joint distribution are determined by its copula function. After the proposition of the Sklars theorem, the copula function was not put into practice until computer technology, statistical inference methods, and multivariate modeling developed in the 90s. Nelsen [4] systematically introduced its definition, construction method, Archimedean Copulas, and dependency. Embrechts, McNeil, and Straumann [5] firstly induced copula function into financial risk management, then Bouye E, Durrleman V, and Nikeghbali A [6] utilized copula and observed its outstanding advantages of it in the field of risk management, and systematically introduced some applications of it in the financial field. Since then, the Copula function has been widely used in the financial field. Roberto De Matteis [7] summarized the copula, especially the Archimedean Copulas. Claudio Romano [8] analyzed the earning rate of the Italian stock market by applying the copula function. Forbes [9] compiled and presented various related models of Copula and applied them in financial risk management. Jondeau and Rockinger [10] first proposed and established the Copula-GARCH model. They argued that the GARCH model can portray the volatility aggregation phenomenon of financial time series combining the copula model facilitates the observation of the linkage of stock indices and based on this model analyze the correlation between four major international stock markets. Mendes and Marques [11] studied a set of financial assets and their efficient allocation for management based on pair-Copula with high-dimensional correlation modeling, stating that pair-Copula based on the minimum portfolio of risk provides the best long-term portfolio. Hence, many scholars have made in-depth research on portfolios, credit risk models, risk aggregation, and so on by utilizing the copula method, and have obtained many results.

Value at Risk (VaR) model accurately defines and measures financial risks, which was proposed by group of thirty (G30) in the report on The Practice and Rules of Derivatives. Later, JP. Morgan [12] launched the Risk Metrics model for calculating VaR. Beder [13] applied an empirical analysis of the portfolio, followed by Hendrick's [14] evaluation of the risk of the foreign exchange portfolio using VaR measures. The VAR model is based on the assumption of normal distribution of financial data. However, the actual financial data has the characteristics of high peaks and fat tails, which makes the traditional risk measurement model which was constructed on the assumption of statistical normal distribution unable to accurately estimate the real risk. The copula function can solve the defects of traditional risk measurement models by describing the joint distribution characteristics between different distribution characteristic variables. Then the studies by Claudio Romano [8], Embrechts [15], Qichen Liu, and Pin Wang [16] respectively proved that the VaR calculation method based on the Copula theory can achieve better results than the traditional model. Areski Cousin [17], extended the VAR model from a bivariate model to a multivariable model by using a copula function.

3. Experiment

3.1 Data sample selection

This paper selects the S&P 500 index and FTSE (Financial Times of London) as the study's research objects. The S&P 500 index is a stock index of 500 listed companies in the U.S and has the characteristics of wide sampling, high representativeness, high accuracy, diversified risk, and good continuity, which can reflect a wider range of market changes. The FTSE 100 index listed the stocks of the largest 100 companies on the London Stock Exchange, comprising approximately 80% of its market value, reflecting the overall stock market situation in a comprehensive manner. The data sample is the daily stock closing prices from August 29, 2019, to September 9, 2022, a total of 1496 data, mainly using R language to estimate and analyze related models. [18, 19] Closing prices on weekends and public holidays are not included in the data set. The yield can be calculated according to the following formula:

$$Y_t = \ln(X_t) - \ln(X_{t-1}) \tag{1}$$

Y_t represents daily yield; X_t represents the closing price of S&P or FTSE on day t.

3.2 Data sample analysis

Firstly, the Sequence Diagram visualizes the trend of stock prices and daily returns of the S&P 500 index and the FTSE index in the recent three years, to observe the similarity of their stock price trends and the overlap degree of their daily returns, to better reveal their interactive relationship.

As shown in Figure 1, the closing price trends of the S&P and FTSE indices are relatively consistent, with a strong up-and-down effect. According to Figure 2, their logarithmic yields indicate the trend of great fluctuations followed by great ones and small swings followed by small ones, which had a strong volatility aggregation.

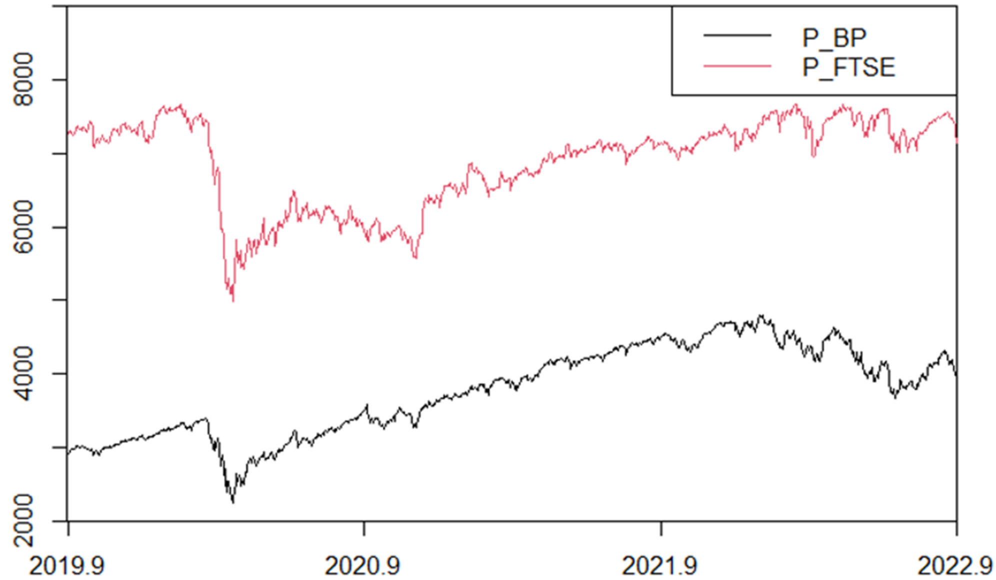

Fig. 1 US S&P, UK FTSE Closing Price Time-Series Charts

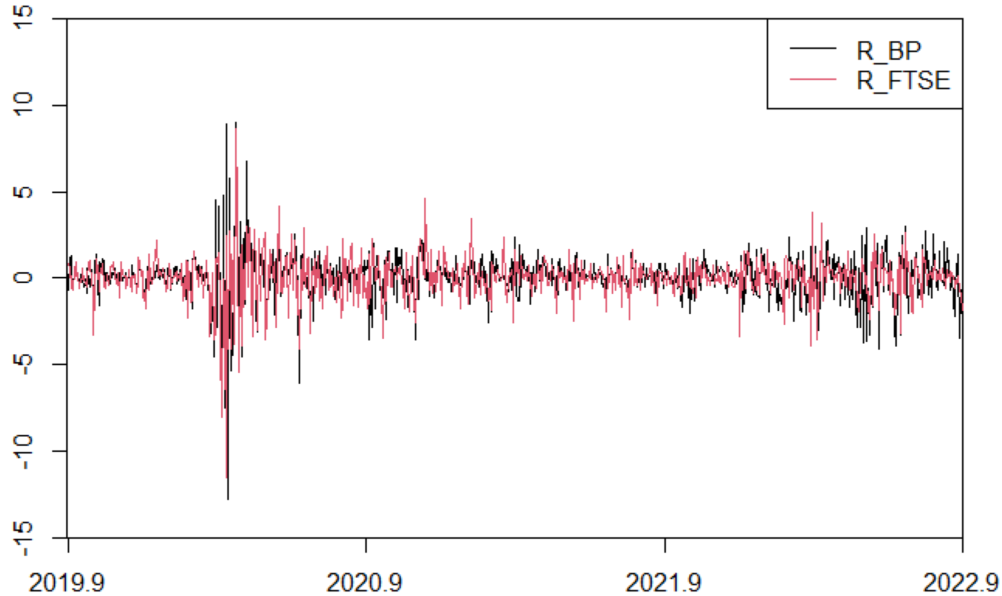

Fig. 2 US S&P, UK FTSE Logarithmic Yields Time Series Charts

Secondly, Descriptive Statistics is employed to check whether the sample data conforms to a normal distribution from their skewness and kurtosis.

Table 1. Descriptive Statistics: US S&P, FTSE Closing Prices, and Log Yield

Descriptive statistics	Mean	Median	Max	Min	Standard Deviation	Skewness	Kurtosis	Max-Min
S&P (Price)	3769.803	3836.430	4793.540	2237.400	604.282	-0.187	-1.157	2556.140
FTSE (Price)	6878.455	7088.700	7674.560	4993.890	600.472	-0.800	-0.437	2680.670
S&P (Yield)	0.041	0.132	8.968	-12.765	1.548	-0.936	13.149	21.734
FTSE (Yield)	-0.001	0.077	8.667	-11.512	1.318	-1.12	12.701	20.179

As shown in Table 1, Rows 1-2 are the descriptive statistics of the closing prices of the S&P and the FTSE. Rows 3-4 are the descriptive statistics of the log returns. It can be detected that the mean value of the US S&P returns is positive and the skewness is negative, while the mean value of the UK FTSE returns is negative and the skewness is also negative. Their two columns of returns have kurtosis greater than 12, indicating a high peaks and fat tails distribution characteristic.

Lastly, Augmented Dickey-Fuller Test (ADF) is applied to test the Stability.

The initial assumption of the ADF unitary root test is that a unit root exits and the series is nonstationary. P-value is calculated as a measurement of statistical significance which means that the shown difference between the attitudes of two groups is due to systematic factors rather than accidental factors, for the intention to prove the null hypothesis is false or rejected. The P-value corresponding to the US S&P and UK FTSE closing price smoothness tests are 0.7745 and 0.3683 separately, which are both greater than 0.1. The number of that indicates that the original hypothesis cannot be dismissed, and the 2-column closing prices are non-stationary series. The P-values corresponding to the smoothness tests of the US S&P and UK FTSE log returns are less than 0.01, stating that at the 1% level of significance, the original assumption is rejected, the 2 columns of returns are smooth series perform well statistical properties.

3.3 GARCH model

This part begins with the Autocorrelation Test, which is employed to test whether the subsequent step requires adopting the ARMA model to ameliorate.

Based on the results of the autocorrelation test shown, the autocorrelation test is employed to test whether autocorrelation exit with an original hypothesis of nothingness. This test detected that the P-value corresponding to the log return of the U.S. S&P index is 8.327e-15, which is smaller than 0.01, reflecting that the basic assumption is refuted at the 1% significance level. Series has significant autocorrelation, so in the subsequent process, GARCH model means equation should be used to improve the autocorrelation using the ARMA model. The P-value corresponding to the log return of the U.K. FTSE index is 0.6675, which cannot reject the original hypothesis. There is no autocorrelation, so the subsequent GARCH model mean-equation only needs to include the constant term.

Then, ARCH (Autoregressive Conditional Heteroskedasticity Model) effects are applied to test whether or not the 2 index indices satisfy the prerequisites for conducting GARCH modeling.

The ARCH effect test is performed on the residuals of the mean equation, and the scientific hypothesis of the ARCH effect test is that the residual series does not have ARCH effects. In this test, the P-values of the residual tests of the averaged equations of their log returns are both lower than 0.01. It indicates the residuals of these two mean equations have a significant ARCH effect, satisfying the prerequisites for conducting GARCH modeling.

Table 2. Parameter estimation results of the GARCH model for the U.S. S&P index

Value	Estimate	Standard Error	T-value	Pr (>|t|)
Mu	0.135536	0.027707	4.8918	0.000001
AR	-0.565413	0.203744	-2.7751	0.005518
MA	0.492401	0.214903	2.2913	0.021948
Omega	0.045599	0.017055	2 6737	0.007503
Alpha	0.254412	0.050892	4 9990	0.000001
Bata	0.744587	0.043571	17.089	0.000000
Shape	6.666400	1.480817	4.5018	0.000007

Lastly, the results of GARCH parameter estimation are shown in Table 2. For the U.S. S&P GARCH model, rows 1-3 are the mean equation constant term, AR term, and MA term, respectively, rows 4-6 are the variance equation constant term, ARCH term, and GARCH term, respectively, and row 7 is the t-distribution degrees of freedom parameter. For the UK FTSE GARCH model shown in Table 3, row 1 is the mean equation constant term, rows 2-4 are the variance equation constant term, ARCH term, and GARCH term, respectively, and row 5 is the t-distribution degrees of freedom parameter.

The p-values can be detected corresponding to the key parameters of both GARCH models, ARCH term and GARCH term parameters, which are less than 0.01, indicating that they are 1% leveled statistically significant. Furthermore, both ARCH term and GARCH term parameters are greater than 0 and add up to less than 1. It illustrates that the stability conditions of the GARCH model are satisfied and proves that the estimation results are robust.

Table 3. Parameter estimation results of the UK FTSE GARCH model

Value	Estimate	Standard Error	T-value	Pr (>|t|)
Mu	0.06349	0.02936	2.1624	0.030585
Omega	0.06736	0.027547	2.4453	0.014475
Alpha	0.19489	0.056124	3.4725	0.000516
Bata	0.78860	0.05093	15.4841	0.000000
Shape	4.09594	0.661826	6.1888	0.000000

4. Results

4.1 Copula correlation analysis

The results of Copula parameter estimation are shown in Table 4. Through Copula modeling data were obtained by probability integral transformation to the standardized residuals of the GARCH model. Gaussian Copula, Clayton Copula, Gumbel Copula, and Frank Copula have one parameter each, and student t-Copula has two parameters. The values of all five types of Copula parameters divided by the standard errors are greater than 3, indicating that they are statistically significant at the 1% level of significance.

Table 4. Copula parameter estimation results

Type	Parameter 1	Parameter 2
Gaussian Copula	0.45（0.03）	-
Student-t Copula	0.45（0.03）	6.28（1.81）
Clayton Copula	0.52（0.05）	-
Gumbel Copula	1.45（0.04）	-
Frank Copula	2.88（0.23）	-

Note: Standard errors are in round brackets.

The results of Copula GOF (goodness-of-fit) are evinced in Table 5; LL refers to the log-Likelihood value which is meaningful in antithesis or more models and with insignificance in only one value, AIC indicates the Akaike Information Criterion that has detected affiliated with LL, and BIC is the Bayesian information criterion; the larger the LL and the smaller the AIC and BIC, the more suitable the model is. Student t-Copula is observed to have the largest LL value and the smallest AIC value, Gumbel Copula has the smallest BIC value, and Student t-Copula is the best-fitting Copula type under the comprehensive comparison.

Table 5. Copula goodness of fit

Type	LL	AIC	BIC
Gaussian Copula	87.69	-173.39	-168.78
Student-t Copula	**95.96**	**-187.92**	-178.71
Clayton Copula	70.63	-139.26	-134.65
Gumbel Copula	93.65	-185.31	**-180.70**
Frank Copula	77.18	-152.35	-147.74

Note: The Copula type with the best fit is bolded.

Copula correlation coefficients are shown below. Since the Student t-Copula is ideal, this paper employed the student t-Copula to assess the overall rank correlation. Furthermore, Normal Copula and Frank Copula are not sensitive to tail correlation and Student t-Copula has symmetric tails, these three Copula types are not suitable for measuring tail correlation. Therefore, Clayton Copula and Gumbel Copula are used to measure the lower-tail and upper-tail correlation coefficients, respectively. The rank correlation measures the overall correlation, the lower tail correlation is a metric of the downside risk of stocks, and the upper tail correlation measures when a skyrocketing of stock occurs. The Kendall rank correlation coefficient between the S&P and the FTSE is 0.30, the upper tail correlation coefficient is 0.39, and the lower tail correlation coefficient is 0.26. It indicates that the correlation becomes strong when an extreme upmarket (bull market) occurs and becomes weak when an extreme downmarket (bear market) occurs, and the holistic Rank Correlation Coefficient is between these two.

Table 6. Copula correlation coefficient

Correlation Structure	Rank Correlation	Upper Tail Correlation	Lower Tail Correlation
US S&P-UK FTSE	0.30	0.39	0.26

4.2 VaR analysis

The student t-Copula is adopted to calculate the risk VaR since it has the best fitting effect. Monte Carlo simulated 10000 times with the simulated yield data obtained by inverse transformation of the marginal distribution, which in turn calculates the risk. In this paper, nine portfolios with different weights are constructed, and it can be discovered that the lowest risk of the portfolio is -1.6733 when the S&P index accounts for 40% plus the U.K. FTSE index accounts for 60%.

Table 7. Portfolio(95% VaR) results

US S&P Index Weighting	UK FTSE Index Weighting	95%VaR
10%	90%	-1.7593
20%	80%	-1.6998
30%	70%	-1.6878
40%	60%	-1.6733
50%	50%	-1.6824
60%	40%	-1.7146
70%	30%	-1.7637
80%	20%	-1.8444
90%	10%	-1.9336

5. Conclusion

The main finding of this study: firstly, the student-t-copula function performs the best fitting among all five copulae modeling. It has the largest LL value, the smallest AIC value, and the suboptimal BIC value. Then the model shows the stock markets are more correlated in bull markets and less correlated in bear markets in U.S. and U.K. Finally, the VaR analysis suggests that investors should choose a portfolio with 60% FTSE and 40% S&P500. The possible reason may be the fact that their overall trend in stock closing prices is similar and in an uptrend. Over the last three years, the biggest bear market was the period between the end of 2019 to March 2020, which was when the epidemic situation just outbreak. With this complication, the least coincident yield movement curves were between March 5, 2020, and March 23, 2020, with daily yields in the U.S. below 2.8% and overall negative for this bear market, while U.K. yields fluctuate widely between -12.0% and 9.1% during this period, showing more times of positive values during that period. Thus, the correlation with the stock market in the UK and the US is weak under the largest bear market. The stock markets in the UK and the US have shown an upward trend in the last three years, with the average daily increase in bull markets being smaller than the average daily decrease in bear markets, which implies that the shock from the epidemic needs a slow recovery. It reflects a strong correlation in the bull.

In general, this study has certain reference significance to research and analysis of the correlation in American and England's stock markets in this field's forthcoming future. Then it provides the investors with the optimum decision for the portfolio using the latest data. However, the data only for the last three years, and the correlation for long-term stock development is yet to be verified, and it is not accurate to use S&P 500 index and the FTSE index to represent and calculate the overall stock exchange transaction demand and supply of the two countries. After that, this experiment only uses the best-fit copula function and 95% confidence level to do the VaR analysis without verifying more cases. The shortcomings of this experiment can be improved in the future, and the method can also be followed to extend the copula multivariate function to analyze the correlation and portfolios of multiple stock markets.

References

[1] Daoji Shi & Fan Li. (2007). Stock Market VaR and Optimal Portfolio Analysis Based on Copula Journal of Tianjin University of Technology (03), 13-16.

[2] Longin, F.M., &Solnik, B. (2000). Extreme Correlation of International Equity Markets. Capital Markets: Asset Pricing & Valuation eJournal.

[3] Sklar, M.J. (1959). Fonctions de repartition a n dimensions et leursmarges.

[4] Nelsen, R.B. (1997). Dependence and Order in Families of Archimedean Copulas. Journal of Multivariate Analysis, 60, 111-122.

[5] Embrechts, P., McNeil, A.J., & Straumann, D. (1999). Correlation: Pitfalls and Alternatives.

[6] Bouyé, E., Durrleman, V., Nikeghbali, A., Riboulet, G., & Roncalli, T. (2000). Copulas for Finance - A Reading Guide and Some Applications. Risk Management.

[7] Roberto De Matteis. (2001).Fitting Copulas to Data.IMU.Zurich.

[8] Clemente, A.D., & Romano, C.A. (2021). Calibrating and Simulating Copula Functions in Financial Applications. Frontiers in Applied Mathematics and Statistics.

[9] Forbes, K.J., &Rigobón, R. (2002). No Contagion, Only Interdependence: Measuring Stock Market Comovements. Journal of Finance, 57, 2223-2261.

[10] Jondeau, E., &Rockinger, M. (2002). Conditional Dependency of Financial Series: The Copula-Garch Model. Capital Markets: Asset Pricing & Valuation eJournal.

[11] [Mendes, B.V., & Marques, D.S. (2012). Choosing an optimal investment strategy: The role of roust pair-copulas based portfolios. Emerging Markets Review, 13, 449-464.

[12] Li, D. (1999). On Default Correlation: A Copula Function Approach. Journal of Financial Abstracts eJournal.

[13] Beder, T.S. (1995). VAR: Seductive but Dangerous. Financial Analysts Journal, 51, 12-24.

[14] Hendricks, D. (1996). Evaluation of Value-at-Risk Models Using Historical Data.

[15] Embrechts, P., Höing, A., &Juri, A. (2003). Using copulae to bound the Value-at-Risk for functions of dependent risks. Finance and Stochastics, 7, 145-167.

[16] Chengqi Hou,& Pin Wang (2008). Research on integrated risk measurement based on connection function Statistical Research, 25 (11), 9.

[17] Bielecki, T.R., Cousin, A., Crépey, S., &Herbertsson, A. (2013). A Bottom-Up Dynamic Model of Portfolio Credit Risk with Stochastic Intensities and Random Recoveries. Communications in Statistics - Theory and Methods, 43, 1362 - 1389.

[18] Investing.com. 2022. SP500 VIX Futures Enhanced Roll TR (SPVIXETR). https://cn.investing.com/indices/sp500-vix-futures-enhanced-roll-tr-historical-data.

[19] Investing.com. 2022. FTSE. https://cn.investing.com/indices/uk-100-historical-data.