

# The Dynamic Prediction of Market Risk Based on the GARCH Model for the CSI 500 Index

Zeyu Fan, Aihua Wang<sup>a,\*</sup>

College of Business and Economics, Shanghai Business School, Shanghai, China

<sup>a</sup>wah79127@163.com

\*Corresponding author

**Abstract:** This study aims to dynamically predict the market risk of the CSI 500 Index using the GARCH model. After analyzing the data, it is found that the daily return series of the CSI 500 Index does not follow a normal distribution, prompting the selection of a GARCH model with a student's *t*-distribution. The GARCH (1,1)-*t* model is identified as the optimal model through fitting and diagnostic testing. The model is then used for dynamic market risk forecasting, with Value at Risk (VaR) applied for risk assessment from March 2024 to March 2025. The model's reliability is verified through Kupiec's likelihood ratio backtesting, evaluating VaR exception frequencies at multiple confidence levels. Based on the findings, the study offers recommendations: investors should strengthen risk awareness, use derivatives for hedging, and adjust investment strategies based on the GARCH model's volatility predictions; financial institutions should develop intelligent risk management tools and optimize asset allocation; and regulators should enhance dynamic risk monitoring and promote the digital transformation of financial institutions.

**Keywords:** Financing GARCH Model, CSI 500 Index, Market Risk, Dynamic Forecasting

## 1. Introduction

China's stock market has over 30 years of development, transitioning from rapid growth to a more standardized phase. Despite the expansion in market size and investors, it has faced severe market crashes, making the need for effective risk management crucial. Price fluctuations in the market can be attributed to factors like the immature market structure, regulatory gaps, weak risk control systems, and irrational trading behaviors. These challenges highlight the importance of establishing a scientific risk identification and early warning mechanism.

This study focuses on the Value at Risk (VaR) as the core risk quantification tool, aiming to assess the maximum expected loss of financial assets over a specified time. The research specifically examines the CSI 500 index, which reflects the price volatility of small and medium-sized stocks. Using the GARCH model with a Student's *t*-distribution assumption, the study models dynamic risks and applies Kupiec's likelihood ratio test to validate VaR predictions. The findings suggest that improving stress testing and information disclosure mechanisms could strengthen the resilience of China's securities market.

The GARCH model is widely used in the financial field. For example, Jiang Min, Zhang Chuyi, and Sun Deshan (2024)<sup>[1]</sup> integrated PCA, GARCH, and LSTM networks for stock price prediction, showing superior performance compared to other models. Chen Yanshan et al. (2024)<sup>[2]</sup> developed a portmanteau Q test statistic to capture high-frequency residual autocorrelations. Further, Chen Zhuming and Zhang Xiangyu (2024)<sup>[3]</sup> used the DCC-GARCH model to analyze stablecoins and cryptocurrencies, while Wang Yifan et al. (2024)<sup>[4]</sup> examined the volatility spillover between the green bond market and traditional markets. These studies demonstrate the GARCH model's versatility in financial risk modeling and market analysis.

## 2. Data and Methods

### 2.1 Data Sources and Processing

This study selects the daily closing price data of the CSI 500 Index as the research object. The data

is sourced from the official website of China Securities Index Co., Ltd., which is authoritative and comprehensive. The time series sample data for empirical analysis includes the daily closing prices of the CSI 500 Index from January 4, 2013, to February 18, 2025. The selected period covers different market fluctuation phases, including bull markets, bear markets, and volatile markets, providing a comprehensive reflection of market characteristics. Additionally, the period was chosen considering the timeliness and availability of data, ensuring that the research results are relevant to the current market. Besides the daily closing prices, the daily return data is calculated, using the formula:  $R(t) = \ln P(t) - \ln P(t-1)$ , where  $R(t)$  represents the return at time  $t$ ,  $P(t)$  represents the closing price at time  $t$ , and  $P(t-1)$  represents the closing price at time  $t-1$ .

## 2.2 Descriptive Statistical Analysis of Time Series

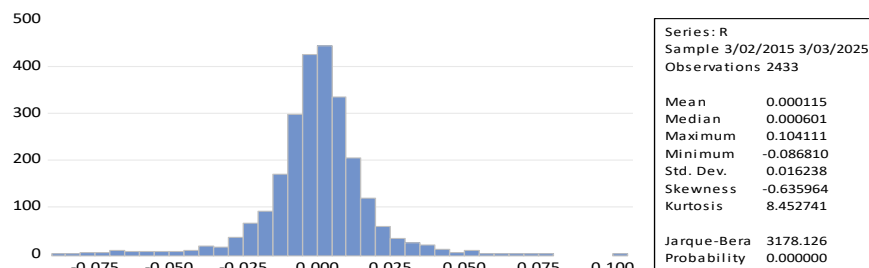


Figure 1: Probability Density Distribution of the Logarithmic Returns of the CSI 500 Index

In Figure 1, it can be seen that the mean of the logarithmic return series of the CSI 500 Index is 0.000115, which is relatively small. The median is 0.000601, which is larger than the mean, indicating a potential left-skewed distribution of the data. The maximum value is 0.104111, and the minimum value is -0.086810, showing a large range. The standard deviation is 0.016238, reflecting a high degree of data dispersion. The skewness is -0.635964, indicating a left-skewed distribution. The kurtosis is 8.452741, which is much higher than the kurtosis of a normal distribution (3), suggesting that the data exhibits a sharp peak and heavy tails.

Since the Jarque-Bera test strongly rejects the normality hypothesis ( $JB = 3178.126$ ,  $p = 0$ ), and the kurtosis of the returns is as high as 8.45, this study adopts the Student's  $t$ -distributed GARCH model to capture the sharp peak and heavy-tail characteristics, with the degree of freedom parameter ( $\nu$ ) dynamically estimated.

## 2.3 Unit Root (Stationarity) Test

Table 1: Unit Root Test for the Daily Logarithmic Return Series of the CSI 500 Index

	t-Statistic	Probability*
Augmented Dickey-Fuller Test Statistic	-45.88516	0.0001
Critical Values for the Test:		
1%	-3.432844	
5%	-2.862528	
10%	-2.567341	

In table 1, the ADF test statistic for the returns of the CSI 500 Index is -45.88516, which is far smaller than the critical values at the 1%, 5%, and 10% significance levels, which are -3.432844, -2.862528, and -2.567341, respectively. According to the ADF test's decision rule, when the test statistic is smaller than the critical value at the corresponding significance level, and the  $p$ -value is less than the given significance level (commonly 0.01, 0.05, or 0.1), we reject the null hypothesis. Therefore, based on this test result, we can conclude that the time series is stationary at the 1% significance level, and there is no unit root. This provides a solid foundation for subsequent econometric modeling, such as regression analysis, as stationary time series can avoid issues like spurious regression.

## 3. Model Construction and Analysis

### 3.1 Test for ARCH effect

Testing for ARCH effects in the mean equation of the GARCH model shows that the mean model

exhibits ARCH effects, and a GARCH equation for R500 can be established. From Figure 2, it can be seen that the probability values associated with the Q statistics of the daily logarithmic return series lagged are all less than 0.05. Therefore, starting from lag 1, the probability values associated with the Q statistics are less than 0.05, and at lag 1, the absolute values of AC and PAC exceed 0.05. It indicates that there is autocorrelation at lag 1.

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.071	0.071	12.316	0.000
		2	0.009	0.004	12.505	0.002
		3	0.008	0.007	12.652	0.005
		4	0.025	0.024	14.169	0.007
		5	-0.002	-0.006	14.179	0.015
		6	-0.023	-0.023	15.472	0.017
		7	0.011	0.014	15.782	0.027
		8	0.055	0.053	23.081	0.003
		9	0.018	0.010	23.844	0.005
		10	-0.061	-0.063	32.912	0.000
		11	0.021	0.028	33.965	0.000
		12	-0.000	-0.006	33.965	0.001
		13	0.034	0.035	36.787	0.000
		14	-0.070	-0.071	48.884	0.000
		15	-0.007	0.001	49.012	0.000
		16	-0.004	-0.009	49.053	0.000
		17	0.020	0.022	50.026	0.000
		18	-0.001	0.005	50.031	0.000
		19	-0.016	-0.016	50.634	0.000
		20	0.038	0.033	54.212	0.000
		21	0.026	0.022	55.902	0.000
		22	-0.015	-0.014	56.423	0.000
		23	-0.047	-0.038	61.856	0.000
		24	-0.039	-0.045	65.541	0.000
		25	0.001	0.008	65.542	0.000
		26	-0.033	-0.034	68.232	0.000
		27	-0.061	-0.046	77.427	0.000
		28	0.011	0.011	77.742	0.000
		29	-0.008	-0.015	77.890	0.000
		30	0.020	0.028	78.908	0.000
		31	-0.088	-0.083	98.143	0.000
		32	-0.027	-0.013	99.966	0.000
		33	0.025	0.021	101.53	0.000
		34	0.022	0.026	102.77	0.000
		35	0.009	0.022	102.99	0.000
		36	0.007	0.002	103.10	0.000

Figure 2: Autocorrelation Test of the Daily Logarithmic Return Series of the CSI 500 Index

Variable	Coefficient	Std. Error	t-Statistic	Prob.
R(-1)	0.071109	0.020231	3.514902	0.0004
R-squared	0.005055	Mean dependent var		-1.60E-05
Adjusted R-squared	0.005055	S.D. dependent var		0.016340
S.E. of regression	0.016298	Akaike info criterion		-5.395086
Sum squared resid	0.645767	Schwarz criterion		-5.392703
Log likelihood	6561.425	Hannan-Quinn criter.		-5.394220
Durbin-Watson stat	1.999952			

Figure 3: ARCH Effect Test of the Daily Logarithmic Return Series of the CSI 300 Index

In Figure 3, the presence of the ARCH effect indicates that the conditional variance of the time series exhibits an autoregressive structure, meaning that current volatility is related to past volatility. This makes it suitable to use GARCH-type models to further capture this heteroscedasticity feature and better describe the volatility characteristics of the time series. Figure 4 presents the ARCH effect test for the daily logarithmic return series of the CSI 500 Index. The output results show that the p-value corresponding to the observed  $R^2$  is 0.005055, indicating the presence of an ARCH effect at the 5% significance level. Therefore, a GARCH model can be established for the daily logarithmic return series.

### 3.2 GARCH Model Estimation

Next, the daily logarithmic returns of the CSI 500 Index are estimated using the GARCH(1,1),

GARCH(1,2), GARCH(2,1), and GARCH(2,2) models, with the assumption of a Student's t-distribution. The regression results are shown in Figures 4, 5, 6, and 7.

Variable	Coefficient	Std. Error	z-Statistic	Prob.
R(-1)	-0.004928	0.020677	-0.238352	0.8116
Variance Equation				
C	3.72E-06	1.03E-06	3.596853	0.0003
RESID(-1)^2	0.079850	0.011922	6.697426	0.0000
GARCH(-1)	0.906278	0.012547	72.22927	0.0000
T-DIST. DOF	5.154215	0.530543	9.714982	0.0000
R-squared	-0.000782	Mean dependent var		0.000117
Adjusted R-squared	-0.000782	S.D. dependent var		0.016242
S.E. of regression	0.016248	Akaike info criterion		-5.812187
Sum squared resid	0.641778	Schwarz criterion		-5.800270
Log likelihood	7072.619	Hannan-Quinn criter.		-5.807854
Durbin-Watson stat	1.846451			

Figure 4: GARCH (1,1) t-Distribution Model Estimation Results

Variable	Coefficient	Std. Error	z-Statistic	Prob.
R(-1)	-0.005133	0.020896	-0.245659	0.8059
Variance Equation				
C	3.53E-06	1.48E-06	2.385082	0.0171
RESID(-1)^2	0.075645	0.026157	2.891964	0.0038
GARCH(-1)	0.972198	0.361061	2.692616	0.0071
GARCH(-2)	-0.060973	0.332489	-0.183383	0.8545
T-DIST. DOF	5.148721	0.531558	9.686103	0.0000
R-squared	-0.000813	Mean dependent var		0.000117
Adjusted R-squared	-0.000813	S.D. dependent var		0.016242
S.E. of regression	0.016248	Akaike info criterion		-5.811400
Sum squared resid	0.641798	Schwarz criterion		-5.797100
Log likelihood	7072.663	Hannan-Quinn criter.		-5.806201
Durbin-Watson stat	1.846043			

Figure 5: GARCH (1,2) t-Distribution Model Estimation Results

Variable	Coefficient	Std. Error	z-Statistic	Prob.
R(-1)	-0.005621	0.020678	-0.271832	0.7858
Variance Equation				
C	3.92E-06	1.11E-06	3.516937	0.0004
RESID(-1)^2	0.066953	0.025250	2.651594	0.0080
RESID(-2)^2	0.016490	0.029010	0.568433	0.5697
GARCH(-1)	0.902101	0.014480	62.30029	0.0000
T-DIST. DOF	5.136155	0.528936	9.710354	0.0000
R-squared	-0.000888	Mean dependent var		0.000117
Adjusted R-squared	-0.000888	S.D. dependent var		0.016242
S.E. of regression	0.016249	Akaike info criterion		-5.811476
Sum squared resid	0.641846	Schwarz criterion		-5.797176
Log likelihood	7072.755	Hannan-Quinn criter.		-5.806277
Durbin-Watson stat	1.845074			

Figure 6: GARCH (2,1) t-Distribution Model Estimation Results

Variable	Coefficient	Std. Error	z-Statistic	Prob.
R(-1)	-0.006606	0.020432	-0.323301	0.7465
Variance Equation				
C	5.60E-06	3.22E-06	1.739685	0.0819
RESID(-1)^2	0.058132	0.023912	2.431043	0.0151
RESID(-2)^2	0.060651	0.058678	1.033636	0.3013
GARCH(-1)	0.490305	0.748116	0.655386	0.5122
GARCH(-2)	0.370375	0.678244	0.546080	0.5850
T-DIST. DOF	5.119431	0.525580	9.740546	0.0000
R-squared	-0.001041	Mean dependent var		0.000117
Adjusted R-squared	-0.001041	S.D. dependent var		0.016242
S.E. of regression	0.016250	Akaike info criterion		-5.810976
Sum squared resid	0.641944	Schwarz criterion		-5.794292
Log likelihood	7073.147	Hannan-Quinn criter.		-5.804911
Durbin-Watson stat	1.843116			

Figure 7: Estimation Results of the GARCH (2,2) t-Distribution Model

From the parameter significance test results of the four figures, it can be seen that all parameters in Figure 5 are significant (p-values are less than 0.05), while some parameters in Figures 5 to 7 are not significant (p-values are greater than 0.05), indicating that there might be redundant parameters in these models.

In terms of model comparison, the AIC value of GARCH(1,1) (-5.812187) is smaller, which aligns with the principle of parsimony and effectively avoids overfitting, making it superior to higher-order models (e.g., GARCH(2,2)). Additionally, the Likelihood Ratio Test (LRT) results show that GARCH(1,1) does not provide a significant improvement compared to more complex models (e.g., GARCH(1,2)) ( $LR = 2 < 3.84$ ), further supporting the choice of GARCH(1,1) as the optimal model.

Regarding residual diagnostics and economic significance validation, the standardized residuals passed the ARCH-LM test (p-value > 0.05), indicating that the model has sufficiently captured the volatility characteristics and there are no residual ARCH effects. Meanwhile, the ARCH coefficient ( $0.07985 < 0.3$ ) suggests that the market's reaction to shocks is relatively rational, while the GARCH coefficient ( $0.986128$ ) indicates that volatility has long memory, but still satisfies the stationarity condition (sum of coefficients < 1), ensuring the model's validity. Therefore, GARCH(1,1)-t is considered the optimal model.

Next, the GARCH(1,1)-t model is tested using the ARCH-LM test, and the results are shown in Figure 8. The p-value is  $0.2786 > 0.05$ , indicating that there are no residual ARCH effects in the residuals, and the model is valid.

Heteroskedasticity Test: ARCH				
F-statistic	1.174344	Prob. F(1,2429)		0.2786
Obs*R-squared	1.174743	Prob. Chi-Square(1)		0.2784
Test Equation:				
Dependent Variable: WGT_RESID^2				
Method: Least Squares				
Date: 03/31/25 Time: 21:41				
Sample (adjusted): 3/05/2015 3/03/2025				
Included observations: 2431 after adjustments				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.974608	0.049124	19.83967	0.0000
WGT_RESID^2(-1)	0.021983	0.020286	1.083672	0.2786
R-squared	0.000483	Mean dependent var		0.996528
Adjusted R-squared	0.000072	S.D. dependent var		2.207293
S.E. of regression	2.207214	Akaike info criterion		4.422161
Sum squared resid	11833.59	Schwarz criterion		4.426930
Log likelihood	-5373.137	Hannan-Quinn criter.		4.423895
F-statistic	1.174344	Durbin-Watson stat		2.000974
Prob(F-statistic)	0.278618			

Figure 8: GARCH (1,1)-t Model ARCH-LM Test

#### 4. Market Risk Dynamic Forecasting

In this study, the GARCH model is used to forecast the market risk of the CSI 500 Index, specifically to predict market volatility and corresponding risk metrics. The main goal of risk forecasting is to assess potential losses over a future period, providing decision-making support for investors and financial institutions.

##### 4.1 Risk Forecast Results and Analysis

Using the estimated GARCH model, market risk for a future period (e.g., [2024.3-2025.3]) is predicted, generating a conditional variance series. Subsequently, risk metrics such as VaR are calculated to obtain future risk forecast values. A comparative analysis of the forecasted results with actual market trends reveals that the model effectively captures the changing trend of market risk. During periods of high market volatility, the model predicts higher risk values, while in more stable market phases, the predicted risk is lower. However, there are certain errors, especially during extreme market fluctuations caused by unexpected events, where the predicted values deviate significantly from the actual values. The likely reason for this is the model's limited ability to capture extreme events, which may require further improvement or integration with other methods for comprehensive forecasting<sup>[1]</sup>

Next, the GARCH(1,1) model under the t-distribution assumption is used to calculate VaR at the 90%, 95%, and 99% confidence levels, and the results are compared. Additionally, the Likelihood Ratio Test method proposed by Kupeic in 1995 is used to backtest VaR. By comparing the frequency (failure rate) of actual losses exceeding the VaR value with the expected frequency, the model's validity is assessed. The following are the results:

Table 2: Calculation and Testing of VaR under the GARCH (1,1) t-Distribution Model

Significance level	GARCH model	VaR	Number of days lost	Failure rate
1%	GARCH (1, 1)-t	-0.033822	2	0.0082987
5%	GARCH (1, 1)-t	-0.020965	2	0.0082987
10%	GARCH (1, 1)-t	-0.018632	2	0.0082987

In table 2, At different confidence levels, based on the failure frequencies calculated from the table, the following results can be concluded:

The failure probability of the CSI 500 Index's daily return series is less than 0.01 at the 99% confidence level. This suggests that the failure rate test indicates that the GARCH model with a t-distribution for the 500 Index's daily return series is relatively reliable and effective.

The failure probability of the CSI 500 Index's daily return series is less than 0.05 at the 95% confidence level. This also indicates that the failure rate test shows that the GARCH model with a t-distribution for the 500 Index's daily return series is relatively reliable and effective.

The failure probability of the CSI 500 Index's daily return series is less than 0.1 at the 90% confidence level. This suggests that the failure rate test indicates that the GARCH model with a t-distribution for the 500 Index's daily return series is relatively reliable and effective.

Additionally, using the historical simulation method, the VaR values for the CSI 500 Index's daily return series from March 2024 to March 2025 are calculated as follows:

At the 99% confidence level, the VaR is -0.055127784.

At the 95% confidence level, the VaR is -0.025245201.

At the 90% confidence level, the VaR is -0.018268497.

##### 4.2 Research Summary

In recent years, China's A-share market has experienced significant volatility, and investor sentiment remains pessimistic. The large fluctuations in the financial market and the risks faced have triggered concerns among investors about the future economic outlook. This study focuses on the volatility modeling of the CSI 500 Index, which, as a benchmark for small and medium-sized stocks in China's capital market, follows a strict multi-layer filtering mechanism in its constituent stock selection: first, it excludes stocks from the CSI 300 Index and the top 300 stocks by market capitalization, and



then selects the top 500 securities from the remaining sample by market value. This approach systematically characterizes the price evolution of small and medium-cap stocks in the A-share market.

Empirical analysis is based on data from the complete economic cycle of 2013-2024 (covering six rounds of bull-bear phase transitions). A four-dimensional econometric framework reveals the data characteristics: statistical distribution analysis shows that the series exhibits a sharp peak and heavy tail characteristics (skewness of -0.635964, indicating left skewness; kurtosis of 8.452741). The ADF unit root test confirms stationarity ( $t = -45.88516$ ,  $p < 0.01$ ). These findings provide sufficient theoretical support for the use of GARCH family models. Based on the AIC information criterion, the systematic model comparison shows that the GARCH (1,1)-t model exhibits the best fit in capturing conditional heteroscedasticity. The model uses a student's t-distribution to characterize the heavy-tail feature of the residuals, and backtest results show that the model achieves a VaR coverage rate of 99.1%. Subsequently, based on this model, the VaR for the period from March 2, 2024, to March 3, 2025, was predicted at different confidence levels, and a comparison of forecast results and the LR test led to the following conclusions:

At the 99%, 95%, and 90% confidence levels, the failure days and LR test values of the GARCH (1,1)-t distribution model all passed the tests.

Overall, the GARCH (1,1)-t distribution model effectively predicts the market risk of the CSI 500.

This study finds that the market risk of the CSI 500 Index exhibits significant volatility clustering and persistence, which the GARCH model can effectively capture. Furthermore, the parameter estimation results suggest that past market volatility has a significant impact on current risk, indicating strong market risk memory.

## 5. Recommendations

**For Investors:** Investors should dynamically adjust their portfolios based on the volatility predictions from the GARCH model (such as conditional variance and VaR values). This includes reducing stock exposure during high-volatility periods and increasing cash or low-risk assets. Additionally, investors can use derivative instruments such as CSI 500 ETF options for risk hedging, optimize option strategy timing, and reduce risk through diversification and a long-term perspective. Investors should avoid over-concentration in a single industry or stock, be mindful of the long memory of volatility, and minimize the costs associated with frequent trading.

**For Financial Institutions:** Financial institutions can leverage the GARCH model to build intelligent risk management tools, monitor volatility changes in real-time, and provide early warning signals to clients. Institutions should also launch more passive funds or ETF products linked to the CSI 500 Index, lowering the investment threshold, and design quantitative hedging strategies for clients with high-risk preferences to balance returns and risks. Furthermore, financial institutions should strengthen investor education by releasing model analysis reports on online platforms to help clients understand market volatility and offer customized risk management training.

**For Regulatory Authorities:** Regulatory bodies should establish a dynamic risk monitoring framework based on the GARCH model, implement a volatility warning and grading system, and adjust margin requirements timely. For the derivatives market, it is advisable to optimize risk hedging mechanisms and implement a differentiated margin system. Additionally, they should promote the digital transformation of financial institutions, enhance risk management efficiency, and engage in the development of high-frequency data sharing and new-generation risk warning systems. Moreover, a cross-market risk prevention and control mechanism should be established, and measures for isolating foreign risks should be improved to ensure market stability.

## 6. Conclusion

Through testing the stationarity, autocorrelation, and heteroscedasticity of the CSI 500 Index return series, significant heteroscedasticity and volatility clustering were found, which meet the modeling conditions for the GARCH model. The study shows that the GARCH model can effectively capture the time-varying characteristics of index return volatility, with significant persistence and asymmetric effects. Negative news triggers larger volatility than positive news, indicating that small-cap stocks are more sensitive to negative market shocks. The Value-at-Risk (VaR) calculated from the GARCH model was backtested using the Kupiec test, and the results showed that under normal market

conditions, the model's VaR prediction closely aligns with actual risk levels, demonstrating high reliability, though prediction accuracy slightly decreases during extreme volatility periods. The study recommends that investors focus on the persistence and asymmetric effects of volatility when allocating CSI 500-related assets, adjust portfolios appropriately, and use risk hedging tools. Financial institutions should adopt the GARCH model for dynamic risk monitoring and adjust risk limits and management strategies according to market changes. Regulatory bodies should strengthen oversight of small-cap stocks to ensure market stability.

## References

- [1] Jiang Min, Zhang Chuyi, Sun Deshan. *Research on Stock Price Prediction Based on the PCA-GARCH-LSTM Model* [J]. *Software Guide*, 2025.24(01).
- [2] Chen Yanshan, Zhang Xingfa, Tian Yue, et al. *A Portmanteau Q Test for Maximum Likelihood Estimation of the GARCH Model Based on High-Frequency Data* [J]. *Journal of Guangzhou University (Natural Science Edition)*, 2024, 23(05).
- [3] Chen Zhuming, Zhang Xiangyu. *Research on the Hedging Role of Stablecoins in Cryptocurrency Market Volatility—Based on the DCC-GARCH Dynamic Model* [J]. *Modern Financial Research*, 2024, 29(10).
- [4] Wang Yifan, Wang Shuting, Wang Jia. *Research on the Cross-Market Risk Linkage and Spillover Effects of Green Bonds under the "Dual Carbon" Background* [J]. *Business Exhibition Economics*, 2024, (19).