Studies on VaR Estimation of China's Agricultural Product Futures Market Based on GARCH Models

Li Peizhe[1,a] Wang Yue[1,*]

[1]*College of Economics，Sichuan Agricultural University, Chengdu, 611130, China*
[a] *Email: sxczlpz@163.com*
[*]*Corresponding author. Email: wangyue@sicau.edu.cn*

ABSTRACT
Taking China Agricultural Products Futures Index (CAFI) as an example, this paper makes descriptive statistics and analysis on the logarithm yield of CAFI, discusses the calculation of VaR under the assumption that the residual is subject to the normal distribution, t-distribution, and generalized error distribution (GED), and tests the accuracy of VaR calculated by each model by Kupiec backtest. The results show that the yield of the CAFI index has the characteristics of "peak and fat tail", volatility aggregation, and asymmetric effect. GARCH (1,1) model under the assumption of normal distribution and GED distribution can truly reflect and measure the risk of the agricultural futures market.
Keywords: Agricultural futures; GARCH models; VaR; Risk measurement

1. INTRODUCTION

China's agricultural product futures market is in a process of rapid growth, but the development of the futures market and the spot market is not perfect, and there are phenomena such as unreasonable trading structures, which have led to the violent fluctuations of China's agricultural product futures market. Therefore, the development of agricultural product futures and their risks to the smooth operation of the futures market and macroeconomic risks have important theoretical and practical significance.

1.1. Related Work

In recent years, many scholars have adopted quantitative methods to study the risk measurement of financial assets. Chen Shoudong and Yu Shidian [1] use the GARCH model to calculate the VaR value of the Shanghai Composite Index and Shenzhen Composite Index under the assumption of the normal distribution, t-distribution, and GED distribution. The results show that the risk of return can be better reflected by t-distribution，and GED distribution than by normal distribution. Huang Chongzhen et al. [2] use the GARCH models to calculate the VaR value of the China CSI 300 ETF feeder fund under the normal distribution, t distribution，and GED distribution, and concludes that the GARCH-M (1, 1) model is the most suitable due to the risk of the fund measure. Yu Xiaojian et al. [3] use the mixing data of Hushen 300 to estimate the parameters of the m-realized GARCH model. The results show that the m-realized GARCH model has higher volatility prediction and VaR measurement accuracy than the traditional GARCH

model. And the m-realized GARCH model with t-distribution is the most suitable. At present, most of the research objects for calculating VaR using the GARCH models are financial assets such as stocks, funds, stock index futures, and exchange rate and interest rate risks. There is little research on the risk measurement of agricultural product futures. Wang Peng et al. [4] use two risk measures (VaR and ES) to estimate four agricultural product futures indexes (hard wheat, cotton, sugar, and soybean oil), and the results shows that China's agricultural product futures market volatility has a "leverage effect". The use of t-distribution can improve the accuracy of risk measurement index estimation, and points out that when measuring the risk of the agricultural futures market, using GARCH-SST is a more reasonable estimation model.

1.2. Our Contribution

Domestic and foreign scholars have relatively complete research on financial asset risk measurement. However, they have relatively few researches on agricultural futures risk measurement, and few scholars use the China Agricultural Products Futures Index (CAFI) index to conduct related research. Therefore, based on the previous research results, this article is innovative to use the CAFI index as a risk measurement object.

1.3. Paper Structure

The rest of the paper is organized as follows. The second section introduces the preparatory work used in this article, including the GARCH models and the calculation method of VaR. The third part is an empirical analysis. It performs

descriptive statistics on the CAFI index, uses the GARCH models to estimate, calculates the VaR of the logarithmic rate of return, and conducts Kupiec backtest. The fourth part is to summarize the full text.

2. BACKGROUND

2.1. GARCH models

Since the volatility of financial asset returns often has obvious volatility agglomeration effects, also called heteroscedasticity, Engle [5] first proposes the autoregressive conditional heteroscedasticity (ARCH) model in 1982, and gradually developes the GARCH model, TGARCH model, EGARCH, PGARCH model, etc.
The GARCH (p, q) model includes the following:

$$y_t = x_t \omega + \varepsilon_t \qquad (1)$$

Conditional variance equation:

$$\sigma_t^2 = \mu + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + + \sum_{j=1}^{P} \beta_j \sigma_{t-j}^2 \qquad (2)$$

There are three terms in the conditional variance equation: constant term (μ), ARCH term (ε_{t-i}^2), GARCH term (σ_{t-j}^2), σ_t^2 represents the conditional variance of ε_t . The conditions that each parameter should meet are: $\mu > 0$, $\alpha_i > 0$, $\beta_j > 0$, $\sum_{i=1}^{q} \alpha_i + \sum_{j=1}^{p} \beta_j < 1$. Take the GARCH (1,1) model as an example, expressed as

$$\sigma_t^2 = \mu + \alpha \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2 \qquad (3)$$

where $\alpha + \beta < 1$.
GARCH model can capture the heteroscedasticity of the return on assets based on better reflecting the time-varying characteristics of the return on financial assets. However, it has obvious defects. Conditional variance σ_t^2 in the GARCH model is an asymmetric function of the ARCH term σ_{t-1}^2. It only depends on the size of ε_{t-1}^2 and has nothing to do with the sign. This is inconsistent with the actual situation in the financial market. In order to reflect this asymmetry of financial asset returns, Zakoian [6] proposes the TGARCH model in 1990. The expression of the model is:

$$\sigma_t^2 = \mu + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 \\ + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2 + \gamma N_{t-1} + \varepsilon_{t-i}^2 \qquad (4)$$

If $\gamma \neq 0$, it shows that the influence of rising information and falling information on volatility is asymmetrical, and this effect is also called leverage effect. If $\gamma > 0$, the volatility generated by falling information is greater than the impact of rising news; if $\gamma < 0$, the volatility generated by rising news is greater than the impact of falling news.
Subsequently, in 1991, Nelson [7] proposes the EGARCH model, the expression of the conditional variance is:

$$ln(\sigma_t^2) = \mu + \sum_{i=1}^{q} \alpha_i \left| \frac{\varepsilon_{t-1}}{\sigma_{t-1}} - E\left(\frac{\varepsilon_{t-1}}{\sigma_{t-1}}\right) \right| \qquad (5)$$

$$+ \sum_{i=1}^{q} \beta_j ln(\sigma_{t-j}^2) + \sum_{i=1}^{q} \gamma_i \frac{\varepsilon_{t-i}}{\sigma_{t-1}}$$

The model can not only capture the heteroscedasticity of the time series but also describe the specific leverage effect. $\frac{\varepsilon_{t-i}}{\sigma_{t-1}} > 0$ means good news, and $\frac{\varepsilon_{t-i}}{\sigma_{t-1}} < 0$ means bad news. If $\gamma_k = 0$, there is no difference between good news and bad news, and the volatility effects of good news and bad news are the same; if $\gamma_k < 0$, the overall volatility caused by good news is less than the volatility effect caused by bad news; if $\gamma_k > 0$, the conclusion is the opposite.
Since the time series of financial asset returns often have the characteristics of sharp peak and fat tail, it cannot be simply regarded as a normal distribution. To describe this time series feature more accurately, the error term usually has three distributions, namely normal distribution, t distribution, and generalized error distribution (GED distribution).

2.2. VaR calculation based on GARCH models

VaR measures the maximum loss faced by financial assets during a certain holding period. It refers to the value of the estimated return sequence determined by the parameter method under a certain profit and loss probability distribution condition and confidence level (such as 95% or 99%). When the confidence level of the GARCH models is at time *t*, VaR can be expressed as:

$$VaR = P_{t-1} Z_\alpha \sigma_t \qquad (6)$$

In formula (6), P_{t-1} is the asset value at time *t-1*, Z_α is the corresponding quantile under the confidence level α, and σ_t is the conditional standard deviation obtained by using the GARCH models.

3. EMPIRICAL STUDY of CAFI RISK MEASUREMENT

China Agricultural Products Futures Index (CAFI) is released by the China Futures Margin Monitoring Center. The index includes 11 mature and high liquidity agricultural product futures listed on Zhengzhou and Dalian commodity futures exchanges. The research object of this paper is the CAFI index from April 9, 2015, to October 31, 2018. A total of 871 samples are processed and analyzed by Eviews 7.0 software.

3.1. Descriptive analysis

Calculate the logarithmic return sequence of the CAFI index, and draw the logarithmic return sequence diagram, as shown in Figure 1.
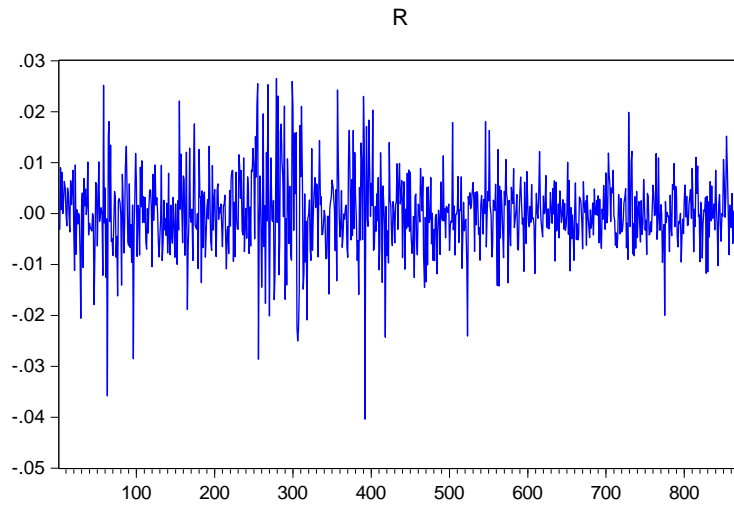
R

Figure 1 logarithmic yield of CAFI index

Draw a histogram of the logarithmic rate of return, and the result is shown in Figure 2.
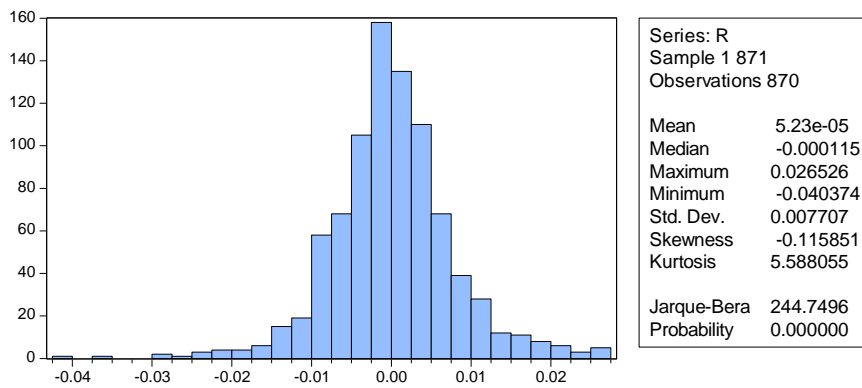

Figure 2 Histogram of Log Return of CAFI Index

The skewness of the time series of the CAFI index logarithmic return is -0.115851, and the kurtosis is 5.588055, which does not satisfy the characteristics of a skewness value of 0 and a kurtosis value of 3 under the condition of normal distribution. The time series has the characteristics of partial peak (left deviation) and "peak and fat tail".

3.2. Empirical analysis

3.2.1. Sequence stationarity test

The ADF statistic is -30.23122, which rejects the null hypothesis that there is a unit root. The CAFI index return rate has a stable time series, and the next step can be analyzed.

Table 1 ADF test results

		t statistic	P value
ADF statistic		-30.2312	0.0000
critical value	1%	-3.43765	
	5%	-2.86465	
	10%	-2.56848	

3.2.2. Sequence correlation analysis

Let the maximum lag order is 12. From the autocorrelation graph and partial autocorrelation graph of the sequence, the autocorrelation coefficient and partial autocorrelation coefficient are scattered in the middle of the two dotted lines, the absolute value is close to 0, and the *Q* statistic corresponds to The *P* values of are all significantly greater

than the significance level of 5%, indicating that the autocorrelation of the return rate series r_t is not significant.

3.2.3. ARCH effect test

Establish a simple mean value model for the return sequence r_t:

$$r_t = C_t + \varepsilon_t \qquad (7)$$

According to the estimation results, the residual sequence is obtained. The correlation diagram shows the autocorrelation

and the residual sequence has an obvious ARCH effect.

3.2.4. GARCH models analysis

This article uses three models, GARCH(1,1), TGARCH(1,1), and EGARCH(1,1). Assume that the residuals obey normal distribution, t distribution, and generalize error distribution, and build GARCH models. The estimated results are shown in the Table 2.

Table 2 Estimation results of GARCH models

	Models	μ	α	β	γ
Normal distribution	GARCH	1.32E-06**	0.067762***	0.909829***	-
	EGARCH	-0.27288***	0.114229***	0.034567*	0.981036***
	TGARCH	1.04E-06**	0.076238***	-0.064703***	0.937529***
t-distribution	GARCH	1.03E-06	0.058278***	0.924846***	-
	EGARCH	-0.18784**	0.081691**	0.069923**	0.98716***
	TGARCH	7.47E-07*	0.080095***	-0.080658**	0.948188***
GED distribution	GARCH	1.12E-06	0.060255***	0.920212***	-
	EGARCH	-0.21836*	0.092346**	0.053887**	0.984931***
	TGARCH	8.43E-07*	0.075719***	-0.072175**	0.945406***

Note: *, ** and *** indicate significance at the significance level of 0.05, 0.01 and 0.001

According to the results in Table 2, the coefficients of the three models are significant at 95% confidence level when they obey normal distribution, t distribution and GED distribution. When GARCH model is used, $\alpha + \beta < 1$ is satisfied under different distributions and the constraint of conditional variance coefficient is satisfied, which indicates that market volatility has long-term memory; for EGARCH model, the positive and negative effects of coefficient α and coefficient β are the same, which indicates that the increasing effect of upward information is greater than that of downward information; this effect is asymmetric; for TGARCH model, coefficient $\beta \neq 0$ also shows that the impact effect is asymmetric and has certain leverage effect. The ARCH LM effect was tested for the estimated residuals of each model. The results show that at 95% confidence level, there is no arch effect in the three models when the

residual sequence follows three distributions.
To sum up, the return series of CAFI index has obvious "peak and fat tail" characteristics, and GARCH effect.

3.2.5. VaR calculation based on GARCH models

Let the confidence levels α be 95% and 99%, respectively. The quantile table below is based on the degree of freedom k of the t distribution and the shape control parameters of the GED distribution obtained by the above regression. According to formula (6), the VaR values at 95% and 99% confidence levels are figured out as Table 4.

Table 3 Quantile table under three distribution hypotheses

	confidence level	GARCH	EGARCH	TGARCH
Normal distribution	95%	1.6449	1.6449	1.6449
	99%	2.3263	2.3263	2.3263
t-distribution	95%	1.9401	1.9544	1.9380
	99%	3.1335	3.1766	3.1270
GED distribution	95%	1.6509	1.6507	1.6511
	99%	2.5779	2.5838	2.5747

Table 4 The VaR values calculated by each model under the three distributions

	GARCH		EGARCH		TGARCH	
	95%	99%	95%	99%	95%	99%
Normal distribution	7.745642	10.952196	7.723185	10.920442	7.846462	12.096430
	7.631811	10.791241	7.590852	10.733327	7.722434	11.905224
	8.211890	11.611463	8.342262	11.795807	8.438256	13.008764
	8.010016	11.326016	8.101355	11.455168	8.307826	12.807689
t-distribution	9.116728	14.722286	9.135250	14.849924	9.264580	14.948577
	8.994082	14.524230	8.894294	14.458234	9.104218	14.689829
	9.592449	15.490512	9.920628	16.126604	9.994093	16.125660
	9.390978	15.165164	9.742977	15.837822	9.851128	15.894984
GED distribution	7.758632	12.120365	7.719166	12.083555	7.878424	12.286925
	7.649908	11.950520	7.547975	11.815574	7.749263	12.085490
	8.172671	12.767168	8.345576	13.064136	8.468169	13.206671
	7.992253	12.485324	8.163012	12.778350	8.347512	13.018499

3.2.6. VaR backtest check

It is necessary to establish a model to test the accuracy and validity of the VaR results. This article adopts the Kupiec failure rate test method. The LR test proposed by Kupiec can be expressed as follows:

$$LR = -2\ln\left[(1-c)^{T-N}c^{N}\right] + 2\ln\left[(1-p)^{T-N}p^{N}\right] \quad (8)$$

In Formula (8), c is the confidence of the assumed VaR and the expected failure probability, T is the actual test days, N is the failure days, and $P = N/T$ is the failure probability. Under the null hypothesis, $LR \sim \chi(1)$. In addition, the coefficient of variation of the VaR sequence obtained from each model under different distributions is also calculated. The results are shown in Table 5

Table 5 VaR backtest test results and coefficients of variation for each model

	Confidence level	Models	Failed days	Failure frequency	LR Statistics	Coefficient of Variation
Normal distribution	95%	GARCH-N	40	4.59%	0.312789	0.703476
		EGARCH-N	41	4.71%	0.160164	0.706949
		TGARCH-N	40	4.59%	0.312789	0.708117
	99%	GARCH-N	14	1.61%	2.740915	0.540455
		EGARCH-N	14	1.61%	2.740915	0.541482
		TGARCH-N	8	0.92%	0.060100	0.517715
t-distribution	95%	GARCH-T	24	2.76%	10.95723	0.614413
		EGARCH-T	26	2.99%	8.647342	0.630308
		TGARCH-T	24	2.76%	10.95723	0.635872
	99%	GARCH-T	5	0.57%	1.885601	0.447127
		EGARCH-T	6	0.69%	0.955960	0.466452
		TGARCH-T	6	0.69%	0.955960	0.471459
GED distribution	95%	GARCH-GED	39	4.48%	0.517825	0.702471
		EGARCH-GED	43	4.94%	0.007341	0.713391
		TGARCH-GED	40	4.59%	0.312789	0.711945
	99%	GARCH-GED	10	1.15%	0.184197	0.504355
		EGARCH-GED	7	0.80%	0.363526	0.514445
		TGARCH-GED	6	0.69%	0.955960	0.520250

If the sequence obeys the t-distribution, the LR statistics of the three models are greater than the critical value at 95% confidence level, and the test is unqualified, while the test passes the test at a 99% confidence level, which indicates that there may be an overestimation of risk. Assuming that the sequence obeys normal distribution or GED distribution, the VaR values obtained by the three models are less than the critical value at 95% and 99% confidence levels, so zero hypotheses are accepted. Therefore, when the sequence obeys the normal distribution or GED distribution, the effect is better than that when the sequence obeys the t-distribution. When the residual sequence obeys the normal distribution and GED distribution, the VaR coefficient of the GARCH model is smaller than that of the other two models, so the VaR value fluctuation is small. Therefore, it is better to calculate VaR according to GARCH (1,1) - N distribution model and GARCH (1,1) - GED distribution model.
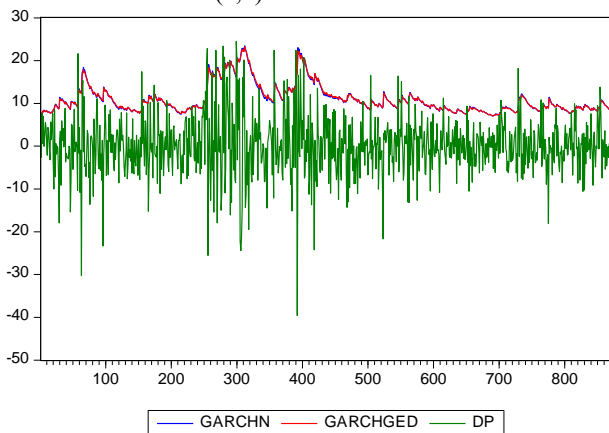

Figure 3 Schematic diagram of VaR and logarithmic returns calculated using GARCH(1,1)-N and GARCH(1,1)-GED

4. CONCLUSION

CAFI index return has the characteristics of volatility agglomeration, its histogram reflects the "peak and fat tail" characteristics, and has an obvious GARCH effect.

The return rate of the CAFI index has an asymmetric effect, which shows that the positive news has a greater impact on the subsequent volatility than the bad news, but this is not consistent with most of the situations in the financial market. When the residuals follow the t-distribution, the VaR values calculated by the three models can achieve better results at the significance level of 1%, but cannot pass the test at the significance level of 5%, which indicates that the hypothesis of t-distribution may lead to overestimation of risk. When the residuals obey the normal distribution or GED distribution, the three models can pass the test, and the GARCH (1,1) model under the assumption that the residuals obey the normal distribution or GED distribution is more suitable to measure the volatility of China's agricultural futures index CAFI.

REFERENCES

[1] S. Chen, S. Yu, Analysis of China's Stock Market Using VaR Method Based on GARCH Model, Jilin University Journal Social Sciences Edition, 2002, pp.11-17.

[2] C. Huang, Q. Cao, Risk Measurement on Open-end Fund Based on GARCH-VaR Model, Statistics and Decision, 2017, pp.152-155

[3] X. Yu X.Wang, Mix Frequency Realized GARCH Models: The Forecast of Volatility and Measure of VaR, Statistical Research, 2018, pp.104-116.

[4] P. Wang, H. Wang, Y. Wei, Research on Stylized Facts and Risk Models for Chinese Agricultural Futures Market, Journal of Industrial Engineering and Engineering Management, 2013, pp.172-182.

[5] R. F. Engle, Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of U.K. Inflation, Econometrica, 1981, pp.987-1008.

[6] J. Zakoian, Threshold heteroskedastic models, Journal of Economic Dynamics & Control, 1994, pp.931-955.

[7] D. Nelson, Conditional Heteroskedasticity in Asset Returns: A New Approach, Modelling Stock Market Volatility, 1991, pp.347-370.