Vol. 10 (2025): Proceedings of the 2nd International Conference on Business, Management and Sustainability (ICBMS 2025)
E-ISSN: 3105-0522 P-ISSN：3105-0514
DOI: https://doi.org/10.70267/icbms.2502.383390

ARIMA-GARCH and ARIMA-EGARCH-t Models in Fitting and Forecasting Volatility of the Shanghai Composite Index

Qinming Yao*

School of Statistics and Data Science, Capital University of Economics and Business, Beijing, 100070, China

**Corresponding author: Qinming Yao*

Abstract

Stock market volatility lies at the heart of risk management. Existing research indicates that A-share returns exhibit fat-tail characteristics, yet further refinement is required in balancing the model's depiction of fat tails and asymmetry with predictive robustness. This paper utilises daily data from the Shanghai Composite Index spanning January 2010 to June 2025 as its sample. It compares the Autoregressive Integrated Moving Average-Generalised Autoregressive Conditional Heteroskedasticity (1,1)-Normal (ARIMA-GARCH (1,1)-Normal) and Autoregressive Integrated Moving Average-Exponential Generalised Autoregressive Conditional Heteroskedasticity (1,1)-t (ARIMA-EGARCH (1,1)-t) models in terms of volatility fitting and forecasting performance. Methodologically, the mean correlation was first filtered using ARIMA (2,0,2). Maximum likelihood estimation was employed for modelling, with performance evaluated through rolling forecasts. The research encompasses sequence stationarity testing and model construction, evaluating in-sample fit through metrics such as the Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), and log-likelihood values. Out-of-sample forecasting performance is measured using Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and Quantile Likelihood (QLIKE) values. The results indicate that the EGARCH-t model exhibits superior in-sample fit with a higher log-likelihood value (10,606.9), successfully capturing fat tails (ν=4.7236) and the leverage effect (γ=-0.0216). However, the GARCH model demonstrated greater robustness in out-of-sample forecasting, with lower error metrics such as MSE (0.000000) and RMSE (0.000523). Research indicates that model selection requires balancing goodness-of-fit with generalizability, while EGARCH-t is well-suited for capturing historical volatility mechanisms, GARCH holds greater predictive utility.

Keywords

Shanghai composite index, GARCH model, EGARCH model, fitting, forecasting

1. Introduction

Stock market volatility serves as a critical input for risk management, asset pricing, and regulatory early warning systems. As a barometer of China's capital markets, the Shanghai Composite Index reached 5,166.35 points on 12 June 2015 before plunging 28.4% over the subsequent fortnight, underscoring the pressing need for robust volatility measurement and forecasting methodologies (Xu, 2001).

Extensive literature indicates that the distribution of Chinese stock returns exhibits significant deviation from normality, presenting a narrow peak and heavy tails. Against this backdrop, GARCH family models have become mainstream due to their ability to capture volatility clustering. Furthermore, in the Chinese market, employing heavy-tailed errors (such as the Studentised t-distribution) typically outperforms the normal assumption (Wang et al., 2022, Zhou, 2004).

To further reflect the asymmetry whereby negative shocks tend to amplify volatility, the EGARCH model is widely adopted. Cross-market evidence also indicates that in scenarios with pronounced asymmetry, EGARCH (or its variants) can enhance forecasting performance (Lin et al., 2020, Lama et al., 2015). For instance, in the Cotlook A Index price series for international cotton, EGARCH achieved an RMSE of 14.41, lower than the 15.38 recorded by GARCH (Lama et al., 2015).

However, comparative studies on the Shanghai Composite Index also reveal that under the studentised t-distribution, the out-of-sample forecasting performance of the ARMA(4,4)-GARCH(1,1) model surpasses that of the EGARCH(1,1) model, suggesting a trade-off between fat tails and model complexity (Wang et al., 2022). Moreover, the Chinese market exhibits both leverage effects and coexisting short-term and long-term drivers, necessitating models capable of handling asymmetry while maintaining robust forecasting capabilities (Han, 2021). ARIMA models may be employed for linear filtering of mean equations, yet they struggle to address conditional heteroskedasticity independently; consequently, their integration with GARCH-type models holds greater practical value (Shi et al., 2014).

Based on this, this paper employs daily data from the Shanghai Composite Index spanning January 2010 to June 2025 as its sample. It first employs ARIMA(2,0,2) to filter out mean-related effects, then systematically compares the performance of ARIMA-GARCH(1,1)-Normal and ARIMA-EGARCH(1,1)-t in both in-sample fitting and out-of-sample rolling forecasting. This endeavour seeks to provide empirical evidence for volatility characterisation and applications in the Chinese market, balancing heavy-tailed asymmetry with robust forecasting.

2. Method

2.1 Data Sources and Indicator Selection

All data utilised in this paper is sourced from the Guotai An (CSMAR) database, covering the period from January 2010 to June 2025. The research subject comprises daily data of the Shanghai Composite Index (CSMAR, n.d.), primarily including indicators such as trading date, opening price, highest price, lowest price, and closing price. To facilitate volatility modelling analysis, this paper first calculates logarithmic returns based on closing prices, employing these as the primary analytical variable for subsequent time series modelling. The entire sample is divided into a training set (January 2010 to December 2023) and a test set (January 2024 to June 2025), utilised respectively for model fitting and predictive performance evaluation.

The core analytical variable in this paper is the logarithmic return r_t, namely:

$$r_t = \ln(P_t) - \ln(P_{t-1}) \tag{1}$$

Here, P_t denotes the closing price of the Shanghai Composite Index on day t. Compared to the original price sequence, the logarithmic return exhibits advantages such as stationarity and normality approximation, rendering it suitable for time series modelling. Furthermore, by applying ARIMA filtering to the logarithmic returns, the resulting residual sequence is extracted for subsequent volatility modelling using GARCH and EGARCH models. The conditional variance output from these models serves as the estimated value for the forecast volatility.

2.2 Methodology Overview

To satisfy the prerequisite assumption of sequence stationarity for the ARIMA model, this paper first visualises the logarithmic returns of the Shanghai Composite Index and conducts a unit root test. As shown in Figure 1, the yield fluctuates around zero without exhibiting a discernible trend. Further Augmented Dickey-Fuller (ADF) tests, presented in Table 1, significantly reject the unit root hypothesis at the zero-order difference level. Consequently, d=0 is adopted for subsequent mean modelling.

Figure 1: Time Series Chart of Logarithmic Yields (Photo/Picture credit: Original).

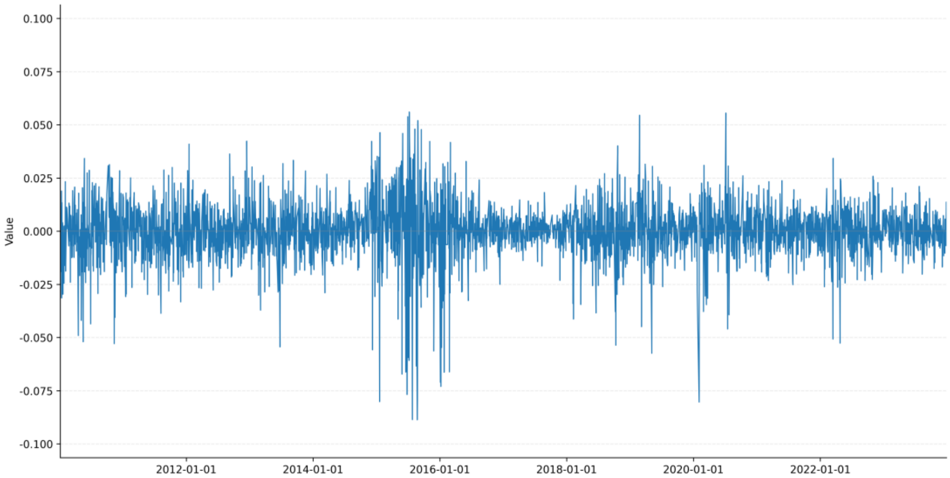


Table 1: *ADF Test Table*

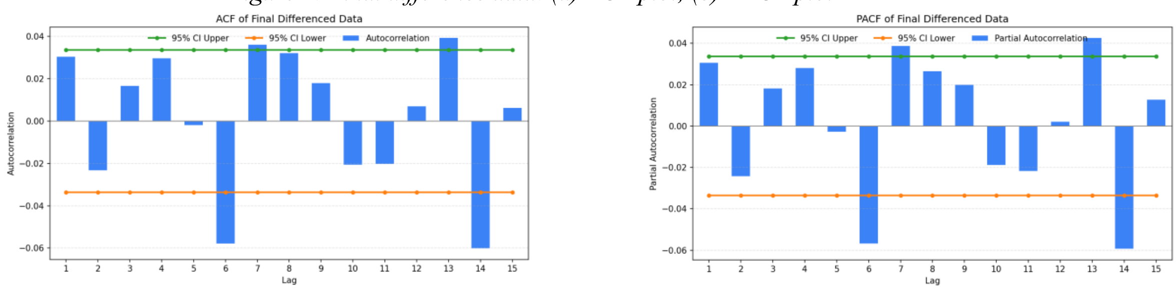
Variable	Differential order	t	P	AIC	threshold		
					1%	5%	10%
Logarithmic yield	0	-10.817	0.000***	-19859.613	-3.432	-2.862	-2.567
	1	-18.029	0.000***	-19739.504	-3.432	-2.862	-2.567
	2	-24.402	0.000***	-19469.779	-3.432	-2.862	-2.567

To determine the autoregressive order (p) and moving average order (q) of the ARIMA model, an autocorrelation analysis was conducted on the stationary log return series. By plotting the Autocorrelation Function Test (ACF) and Partial Autocorrelation Function (PACF) plots of the sequence (see Figure 2, i.e., the ACF and PACF plots of the final differenced data), the optimal ARIMA model was ultimately determined to be ARIMA (2,0,2) through automatic parameter optimisation. The ARIMA model integrates three components: autoregression (AR), differencing (I), and moving average (MA). Differencing stabilises the sequence while enabling analysis of stationary time series by fitting historical observations to residuals. Its fundamental form is:

$$r_t = \mu + \phi_1 r_{t-1} + \phi_2 r_{t-2} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} \tag{2}$$

Here, r_t denotes the logarithmic return at time t, μ represents the constant term, ϕ_1 and ϕ_2 denote the autoregressive coefficients,θ_1 and θ_2 denote the moving average coefficients, and ε_t denotes the white noise residual term.

Figure 2: Final difference data: (a) ACF plot; (b) PACF plot



(a) ACF of final differenced data

(b) PACF of final differenced data

Subsequently, an ARCH-LM test was conducted on the ARIMA residuals, revealing significant conditional heteroskedasticity. This necessitated the introduction of a GARCH-type model to account for volatility. During the training phase, this study constructed both an “ARIMA-GARCH(1,1)-Normal” model and an “ARIMA-EGARCH(1,1)-t” model. The conditional variance equation for the GARCH(1,1) model is:

$$\sigma_t^2 = \omega + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2 \tag{3}$$

Here, σ_t^2 denotes the conditional variance at time t, ω>0 represents the constant term, α≥0 is the ARCH coefficient reflecting the impact of past shocks on current volatility, and β≥0 is the GARCH coefficient indicating volatility persistence. This model is suitable for modeling symmetric financial volatility structures with insignificant fat tails.

To further characterize the asymmetry and heavy-tailed properties in financial markets, the EGARCH(1,1) model is introduced, whose conditional variance equation is:

$$\ln(\sigma_t^2) = \omega + \beta \ln(\sigma_{t-1}^2) + \alpha \left(\left| \frac{\varepsilon_{t-1}}{\sigma_{t-1}} \right| - E[|Z|] \right) + \gamma \left(\frac{\varepsilon_{t-1}}{\sigma_{t-1}} \right) \tag{4}$$

Here, $\ln(\sigma_t^2)$ denotes the logarithmic conditional variance, $\frac{\varepsilon_{t-1}}{\sigma_{t-1}}$ represents the standardized residual, $E[|Z|]$ is the expected value under the standard normal or t-distribution,α measures the intensity of the absolute impact of shocks, and γ indicates the degree of asymmetric volatility response. If $\gamma < 0$, it indicates that negative shocks exert a stronger influence on volatility.

Model estimation employs the maximum likelihood estimation (MLE) method. During the forecasting phase, the ARIMA model is first fitted using training set samples. A one-day rolling forecast filter is applied to the test set, extracting the dynamic residual sequence as input. Subsequently, the pre-trained GARCH and EGARCH models are invoked to perform conditional volatility rolling forecasts. The forecast results are saved for model accuracy evaluation and comparative analysis during the test period.

3. Results

According to the ARIMA (2,0,2) model test table (Table 2), the Q-statistic test for this model shows that the P-value for Q12 is 0.147 (>0.05), indicating that the residuals approximate white noise and satisfy the ARIMA model's requirement for purely random residuals. However, the model's coefficient of determination R^2 is only 0.01, indicating that the model poorly fits the logarithmic returns of the Shanghai Composite Index and struggles to effectively capture the dynamic characteristics of the time series.

***Table 2:** ARIMA Model (2,0,2) Test Statistics*

Item	Symbol	Value
	Df Residuals	3395
Sample size	N	3400
Q-statistic	Q6(P)	5.893(0.015**)
	Q12(P)	9.499(0.147)
	Q18(P)	18.494(0.102)
	Q24(P)	28.766(0.051*)
	Q30(P)	46.047(0.004***)
Information Guidelines	AIC	-20022.893
	BIC	-19986.104
Goodness of fit	R^2	0.010

*Note: ***, **, and * denote significance levels of 1%, 5%, and 10%, respectively.*

The ARCH-LM test was performed on the residuals of the ARIMA (2,0,2) model (with a lag order of 12, see Table 3, i.e., the ARCH-LM test table for residual variables). The results show an LM statistic of 406.3673 with a corresponding p-value of 0.0000 (less than 0.05), and an F-statistic of 38.3316 with a p-value also of 0.0000. This indicates significant volatility clustering in the residual sequence. This indicates that a simple ARIMA model cannot capture the heteroskedasticity of returns, necessitating the integration of GARCH-type models to enhance volatility forecasting accuracy.

Based on the residual sequence, this paper constructed the ARIMA-GARCH (1,1)-Normal and ARIMA-EGARCH (1,1)-t models, with parameter results shown in Table 3. The GARCH model achieved a log-likelihood value of 10446.7, corresponding to AIC=20887.3 and BIC=20868.9. In its variance equation, the constant term $\omega \approx 3.26\times10^{-6}$, the ARCH term α=0.10, and the GARCH term β=0.88 all passed significance tests, indicating significant volatility clustering and strong persistence.

In comparison, the EGARCH model demonstrated superior goodness-of-fit. The log-likelihood value increased to 10,606.9, while the AIC and BIC values decreased to -21,203.7 and -21,173.1, respectively. In its exponential equation, $\omega = -0.0841$ (significant at the 5% level), $\alpha = 0.1307$, and $\beta = 0.9896$ (both significant at the 1% level), highlighting the strong memory effect of volatility shocks. The asymmetric term $\gamma = -0.0216$ (borderline significant at the 10% level) with a negative sign confirms the leverage effect, indicating negative shocks amplify volatility more strongly. The degree-of-freedom parameter $\nu = 4.7236$ (significant at the 1% level) further validates the fat-tail characteristics of the SSE Index returns, with the Student's t-distribution better fitting its residual distribution.

Table 3: Training Set Model Parameters

Indicators	GARCH(1,1)-Normal	EGARCH(1,1)-t
Log-Likelihood (LogL)	10446.700	10606.900
Akaike Information Criterion (AIC)	-20887.300	-21203.700
Bayesian Information Criterion (BIC)	-20868.900	-21173.100
Distribution Assumptions	Normal distribution	Student's t-distribution
ω (Constant Term)	3.2556×10^{-6}（p<0.001）	-0.0841（p=0.0128）
α_1 (ARCH Term)	0.1000（p<0.001）	0.1307（p<0.001）
β_1 (GARCH Term)	0.8800（p<0.001）	0.9896（p<0.001）
γ (Skewness Term)	——	-0.0216（p=0.0681）
ν (Degrees of Freedom Parameter)	——	4.7236（p<0.001）
Sample Size	3400	3400
Covariance Estimation Method	Robust	Robust

On the test set (January 2024 to June 2025), daily rolling forecasts were generated using the GARCH (1,1)-Normal and EGARCH (1,1)-t models fitted from the training set, yielding their respective forecast metrics (Table 4). The results indicate that the GARCH model delivers more stable forecasts with smaller prediction errors, demonstrating superior generalization capabilities. In terms of evaluation metrics, the MSE, RMSE, and MAE of the GARCH model were all lower than the corresponding values of the EGARCH model (Table 4). This indicates that the GARCH model exhibits significantly smaller prediction errors.

Table 4: Test Set Model Metrics Table

Model Type	MSE	RMSE	MAE	QLIKE
GARCH model	0.000000	0.000523	0.000171	-7.973346
EGARCH model	0.000004	0.001947	0.001791	-6.309844

4. Discussion

Specifically, the GARCH model exhibits smoother conditional variance changes when forecasting volatility and demonstrates a high degree of alignment with actual volatility, indicating that the model effectively adapts to the volatility patterns observed in the test set. This phenomenon may stem from the discrepancy between overfitting and the model's generalization capability.

In contrast, while the EGARCH model demonstrates superior fit on the training set, effectively capturing volatility's asymmetry and fat-tail characteristics, its performance on the test set is notably weaker. This reversal—strong in-sample but weak out-of-sample—has also been observed in comparisons of other volatility models with fat-tail assumptions. For instance, in a comparison of heavy-tailed distributions in high-frequency Realized-GARCH models, authors Wang Tianyi and Huang Zhuo noted that skewed-t exhibits overfitting, leading to poor out-of-sample forecasting performance. This suggests that complex distributions/structures do not always translate to stronger generalization capabilities (Wang and Huang, 2012). This may stem from the EGARCH model in the training set better capturing the complex volatility characteristics (e.g., asymmetry) of the data while simultaneously learning noise, thereby reducing its adaptability to new data.

Additionally, GARCH models are typically simpler than EGARCH models, as they only consider historical volatility values (GARCH terms) and current shocks (ARCH terms). Consequently, they are less prone to overfitting and exhibit greater robustness when applied to out-of-sample data. Using the Shanghai Composite Index and its representative constituent stocks as examples, empirical research based on ARIMA+GARCH reveals that the GARCH variance term exhibits significant and persistent volatility memory ($\alpha+\beta$ close to 1), consistent with the "volatility clustering" phenomenon observed in the Chinese market. Furthermore, the relative error in short-term forecasts can be maintained at a low level, demonstrating good usability and robustness (Nong, 2020).

At the industry index level, ARIMA-GARCH forecasts also effectively capture directional shifts: while point value fits are not perfect, predictions align with actual trends in direction, offering practical reference value. This further demonstrates that the simpler GARCH structure exhibits stronger generalization capabilities across numerous application scenarios.

In summary, when the test set lacks significant asymmetric shocks or extreme volatility, more complex EGARCH models may overemphasize extreme samples from the training period, thereby weakening their extrapolation capabilities. In contrast, GARCH models, due to their structural simplicity, achieve superior QLIKE in empirical tests, demonstrating the advantage of robust forecasting.

Finally, this study only compared two typical models. Future research could further incorporate additional heavy-tailed distributions (such as GED) or more complex asymmetric models (such as APARCH) for expanded analysis (Chen and Yang, 2003). Additionally, dynamic changes in market structure within the sample period may impact model performance. Subsequent studies could integrate rolling window or time-varying parameter models to enhance the dynamic capture of market volatility patterns.

5. Conclusion

This study examines daily data from the Shanghai Composite Index spanning January 2010 to June 2025. By constructing ARIMA-GARCH (1,1)-Normal and ARIMA-EGARCH (1,1)-t models to compare their performance in volatility fitting and forecasting. This investigation explores the effectiveness of heavy-tailed asymmetric models in capturing and predicting the volatility characteristics of the SSEI. The main conclusions are as follows:

In terms of model fitting capability, the ARIMA-EGARCH (1,1)-t model demonstrated superior performance on the training set (2010–2023). This model effectively captures the fat-tail characteristics of returns by incorporating a Student’s t-distribution (degree of freedom parameter $\nu=4.7236$, significantly non-zero). The asymmetric term $\gamma=-0.0216$ validates the market leverage effect, indicating that negative shocks exert a stronger influence on volatility than positive shocks, thereby better aligning with the actual volatility patterns observed in the Chinese stock market. Its log-likelihood, AIC, and BIC values all outperform those of the ARIMA-GARCH (1,1)-Normal model, demonstrating superiority in characterizing the complex volatility structure (asymmetry and fat tails).

Out-of-sample forecasting performance shows a reversal, with the ARIMA-GARCH (1,1)-Normal model demonstrating greater robustness in rolling forecasts for the test set (first half of 2024-2025). Its MSE, RMSE, and MAE are significantly lower than those of the EGARCH-t model, while its QLIKE value is superior, indicating stronger generalization capabilities. This reveals a trade-off between model fit quality and prediction robustness. EGARCH-t tends to overfit by excessively capturing specific training set noise, whereas the structurally simpler GARCH model demonstrates superior generalization during the test period with stable volatility by “simplifying” data characteristics.

The study implies that model selection should be tailored to specific objectives. For capturing historical volatility mechanisms (such as leverage effects and fat-tail characteristics), the ARIMA-EGARCH (1,1)-t model is preferable. To enhance out-of-sample forecasting accuracy and robustness, the ARIMA-GARCH (1,1)-Normal model offers greater practical value. This highlights that financial forecasting requires balancing model complexity and generalization capability, avoiding sacrificing predictive utility for the sake of fitting quality.

In summary, this study empirically verifies the effectiveness of combining the Student’s t-distribution with the EGARCH model in capturing the historical volatility characteristics of the Shanghai Composite Index. It also demonstrates the superiority of simple models in terms of volatility prediction generalization capabilities, providing valuable insights for financial risk management, asset pricing, and policy formulation.

References

Chen, X. and Yang, H., (2003). VaR-APARCH model and quantitative analysis of securities investment risk. *Chinese Journal of Management Science*, no. 1, pp. 22-27.

CSMAR, (n.d.). *Domestic index daily quotes file (Shanghai Composite Index data. 2025-08-12)* [Online]. Available: www.csmar.com [Accessed 24 December 2025].

Han, Y., (2021). Volatility analysis of the Shanghai Composite Index based on the asymmetric GARCH-MIDAS model. *Statistics and Decision Making,* vol. 37, no. 24, pp. 157-160.

Lama, A., Jha, G. K., Paul, R. K. and Gurung, B., (2015). Modelling and forecasting of price volatility: an application of GARCH and EGARCH models §. *Agricultural Economics Research Review,* vol. 28, no. 1, pp. 73-82.

Lin, Y., Xiao, Y. and Li, F., (2020). Forecasting crude oil price volatility via a HM-EGARCH model. *Energy Economics,* vol. 87, p. 104693.

Nong, X., (2020). *Forecasting the Shanghai Composite Index and portfolio investment research based on ARIMA and GARCH models.* Master's Thesis, Hunan University.

Shi, H., You, Z. and Chen, Z., (2014). Analysis and forecasting of the Shanghai Composite Index using ARIMA models based on wavelet analysis. *Practice and Understanding of Mathematics,* vol. 44, no. 23, pp. 66-72.

Wang, T. Y. and Huang, Z., (2012). High-frequency volatility modeling: Realized GARCH model based on heavy-tailed distributions. *Journal of Quantitative Economics and Technical Economics,* vol. 29, no. 5, pp. 149-161.

Wang, Y., Xiang, Y., Lei, X. and Zhou, Y., (2022). Volatility analysis based on GARCH-type models: Evidence from the Chinese stock market. *Economic Research-Ekonomska Istraživanja,* vol. 35, no. 1, pp. 2530-2554.

Xu, L., (2001). Empirical study on the steady-state characteristics of stock returns in the Chinese stock market. *Financial Research*, no. 6, pp. 36-43.

Zhou, Z., (2004). Comparative study on the volatility forecasting capabilities of GARCH models based on different distribution assumptions for the Shanghai Composite Index. *Value Engineering*, no. 3, pp. 70-72.

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgment

This paper is an output of the science project.

Open Access