Journal of Research Administration
Society of Research Administrators International

VOLATILITY ANALYSIS USING MACHINE LEARNING FOR IMPROVED INVESTMENT DECISIONS

Srinivas Gumparthi[1] , Venkata Vara Prasad[2], Bhargavi Rentachintala[3] ,

Vignesh G R[4]& Vaseekaran S[5]

[1]Professor, SSN School of Management, Kalavakkam, Sri Sivasubramaniya Nadar College of Engineering, Chennai 603110, Tamil Nadu, India.

[2,4,5]Professor, Sri Sivasubramaniya Nadar College of Engineering, Kalavakkam, Chengalpattu District, Chennai 603110, Tamil Nadu, India.

[3]Professor, SCOPE, VIT University, Vandalur-Kelambakkam Road, Chennai.

Abstract

Understanding and managing volatility in the stock market is crucial for investors and traders seeking to optimize risk management and decision-making. This study examines the volatility of 10 companies within the NIFTY50 index over a two-year period (2021–2023). Leveraging the GARCH model, the research predicts volatility for each company and ranks them accordingly. Results reveal that Infosys and Bajaj Finance exhibit the highest levels of volatility, with conditional variances of 3.077 and 2.896, respectively. Additionally, the study conducts a detailed analysis of monthly volatility rankings for the years 2021 and 2022, providing valuable insights for market participants navigating fluctuating market conditions.

Introduction

Volatility serves as a pivotal metric in financial markets, encapsulating the extent of price variation over time and providing insights into potential fluctuations in asset prices. It stands as a fundamental gauge for investors, traders, and risk managers alike, shaping investment decisions and risk mitigation strategies.

This paper's volatility analysis carries significant implications for stakeholders. By pinpointing the most volatile entities within the Nifty 50 index, investors gain the ability to recalibrate their portfolios, accounting for heightened risk exposure. Moreover, traders leverage volatility forecasts to discern opportune moments for trading activities, crafting strategies that capitalize on market dynamics and fluctuations.

Literature survey

The literature review presents a comprehensive exploration of research endeavors aimed at understanding volatility characteristics and modeling techniques across diverse financial markets. Each study contributes unique insights into the dynamics of market volatility and offers valuable implications for investors, traders, and risk managers.

Dana AL-Najjar [1] examined the volatility characteristics on Jordan's capital market that include clustering volatility, leptokurtosis, and leverage effect. This objective can be accomplished by selecting symmetric and asymmetric models from GARCH family models. The findings suggest that the symmetric ARCH /GARCH models can capture characteristics of ASE and provide more evidence for both volatility clustering and leptokurtic, whereas EGARCH output reveals no support for the existence of leverage effect in the stock returns at Amman Stock Exchange.

Dong et al. [2] uses diagnostic checking in which it includes observing residual plot and its ACF PACF diagram. If ACF PACF of the model residuals show no significant lags, the selected model is appropriate. They plot ACF/ PACF of the original data. If they observe no lags/ no dying down, they 'll take difference and plot ACF/ PACF of the differenced data. Their findings demonstrate that if ACF and PACF of differences of that stock doesn't have any lag then there is no trend.

Emenike et al. [3] investigated the volatility of stock market returns in Nigeria using GARCH (1,1) and the GJR-GARCH (1,1) models. They used Volatility clustering, leptokurtosis, and leverage effects to examine the NSE returns series. The NSE return series shows indications of volatility clustering, according to the GARCH (1,1) model results.

Chaiwat Kosapattarapim et al. [[4] used three emerging stock markets in Southeast Asia to examine the volatility forecasting potential of GARCH (p, q) models with six distinct types of error distributions. Their findings demonstrate that compared to a GARCH (p, q) model with a normal error distribution, a GARCH (p, q) model with non-normal error distribution typically performs better for out-of-sample forecasting. Arfa Maqsood1 et al. [5] used the General Autoregressive Conditional Heteroscedastic (GARCH) type models for the estimation of volatility of the daily returns of the Kenyan stock market. The results show that the volatility process is highly persistent, thus giving evidence of the existence of risk premium for the NSE index return series.

Mamun Miah and Azizur Rahman [6] studied the performance of simple GARCH model. We apply the Generalized Autoregressive Conditional Heteroscedasticity (GARCH) model of different lag order to model volatility of stock returns of four Bangladeshi Companies on Dhaka Stock Exchange (DSE). Akaike Information Criteria (AIC) and Bayesian Information Criteria (BIC) are used to select the best GARCH (p, q) model. Result shows that, GARCH (1,1) is the best than other GARCH (p, q) models in modeling volatility for the daily return series of DSE.

Ludwig Schmidt [7] investigated the volatility forecasting performance of symmetric and asymmetric GARCH models on Nordic indices during COVID-19. The models examined in this paper are GARCH, EGARCH, GJR and APARCH. The results of this paper are that the symmetric GARCH (1,1) on average has the worst forecasting performance during a crisis. However, the difference between the predictability of the models is in practice small. The superior forecasting models are the GJR (1,1) and EGARCH (1,1) when forecasting a crisis on Nordic indices according to the evaluation measures.

Ahmed Shamiri and Zaidi Isa [8] have used high frequency in their study to enable a useful comparison of volatility forecast models. With six error distributions (normal, skew normal, student-t, skew student-t, generalised error distribution, and normal inverse Gaussian), they compared the performance of symmetric GARCH, asymmetric EGARCH, and non-leaner asymmetric NAGARCH models.

Erginbay Ugurlu1 et al. [9] examined the use of GARCH-type models for modeling volatility of stock markets returns for four European emerging countries and Turkey. We use daily data from Bulgaria (SOFIX), Czech Republic (PX), Poland (WIG), Hungary (BUX) and Turkey (XU100) which are considered as emerging markets in finance. They found that GARCH, GJR-GARCH and EGARCH effects are apparent for returns of PX and BUX, WIG and XU whereas for SOFIX there is no significant GARCH effect.

Yuling Wang et al. [10] used the data of the Shanghai Composite Index and Shenzhen Component Index returns were selected to conduct an empirical analysis based on the generalised autoregressive conditional heteroskedasticity (GARCH)- type model. Mean absolute error (MAE) and root- mean-squared error (RMSE). The results denote that the ARMA (4,4)-GARCH (1,1) model under Student's t-distribution outperforms other models when forecasting the Shanghai Compos ite Index return series. For the return series of the Shenzhen Component Index, ARMA (1,1)-TGARCH (1,1) display the best fore- casting performance among all models.

Overall, the literature underscores the importance of volatility modeling in financial markets and provides valuable insights into modeling techniques and empirical findings across different contexts.

Architectural Design for Proposed system

The dataset obtained from Yahoo Finance is first cleaned to check for null values. It is then processed to obtain only the closing price of each company.

Then the daily returns are calculated from the preprocessed data and sent to different order of GARCH type models based on PACF analysis to predict the volatility. Finally, the best GARCH model is chosen using AIC and BIC measures.

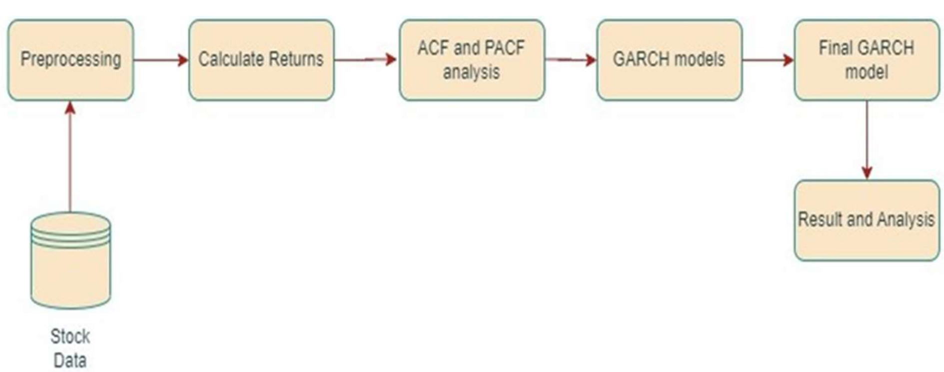

Fig. 1 Model Architecture

Modules Split-up

Calculate return.

In finance, return is a profit on an investment. It comprises any change in value of the investment which the investor receives from that investment, such as interest payments. Returns Rt is calculated by.

$$R_t = Log(p_t/p_{t-1}) \tag{1}$$

where pt is present day price and pt−1 is previous day's price of the stock.

ACF and PACF Analysis

Autocorrelation analysis is an important step in the Exploratory Data Analysis of time series forecasting. Autocorrelation analysis helps detect patterns and check for randomness. The lags in the ACF and PACF graphs are used to decide the orders (p, q) of GARCH model. Where p is AR component and Q is MA component.

GARCH Model

Generalized Autoregressive Conditional Heteroskedasticity (GARCH) is a statistical modeling technique used to help predict the volatility of returns on financial assets. GARCH is appropriate for time series data where the variance of the error term is serially autocorrelated following an autoregressive moving average process. GARCH is useful to assess risk and expected returns for assets that exhibit clustered periods of volatility in returns. In general, the GARCH (p, q) model is presented in the following formula:

$$\sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i \cdot \varepsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \cdot \sigma_{t-j}^2$$

Where i=0,1,2, 3... p, conditional volatility ω, αj, βi are non-negative constants with αj + βi > 1 it should be near to unity for an accurate model, ϵt−1 is residuals, and it is lagged conditional volatility, and the last part of the formula is the main difference in applying both ARCH and GARCH. Hence αj and ε2 t−1 are ARCH components and βj and σ2 t−1 are GARCH components.

Final Model

From the different order of GARCH models the best model is chosen based on AIC and BIC scores. The Akaike Information Criterion (AIC) is a method for scoring and selecting a model. The AIC statistic is defined for logistic regression as follows.

$$AIC = -2/N * LL + 2 * k/N \tag{3}$$

Where N is the number of examples in the training dataset, LL is the log-likelihood of the model on the training dataset, and k is the number of parameters in the model.

The Bayesian Information Criterion, or BIC for short, is a method for scoring and selecting a model. The BIC statistic is calculated for logistic regression as follows:

$$BIC = -2 * LL + \log(N) * k \quad (4)$$

Where log () has the base-e called the natural logarithm, LL is the log-likelihood of the model, N is the number of examples in the training dataset, and k is the number of parameters in the model.

Result and Analysis

The results of the volatility analysis of each company compared provide a comprehensive assessment of the performance and risk associated with investing in 20 companies.

Results

Ten companies from NIFTY50 are ranked based on their predicted volatility for the years 2021 and 2022. Table 1 and 2 shows the predicted volatility ranking for 2021 and 2022.

Infosys is a leading global provider of digital services and consulting, while Bajaj Finance is a non-banking financial company (NBFC) in India that provides consumer and business loans. Both companies operate in highly competitive and rapidly changing industries that are subject to a range of economic, technological, and regulatory risks. For example, Infosys may be affected by changes in demand for digital services,

Rank	Company	Conditional Variance
1	INFY	3.077977
2	BAJAJFINSV	2.896437
3	KOTAKBANK	2.712492
4	RELIANCE	2.614571
5	LT	2.568922
6	HDFC	2.346259
7	ITC	2.114418
8	TCS	1.577903
9	HINDUNILVR	1.489335
10	HDFCBANK	0.958922

Table 1 Volatility ranking for 2021.

Competition from other service providers, and regulatory changes that affect the out- sourcing industry. Similarly, Bajaj Finance may be affected by changes in consumer behavior, interest rates, and regulatory changes that affect the NBFC industry. The high volatility of these companies may reflect the uncertainty and risk associated with their industries, as well as the potential for high returns if they are able to navigate these challenges successfully.

Rank	Company	Conditional Variance
1	INFY	4.002195
2	BAJAJFINSV	3.699044
3	RELIANCE	2.254154
4	LT	2.181108
5	ITC	2.034492
6	TCS	1.634887
7	HDFC	1.576713
8	HINDUNILVR	1.431025
9	KOTAKBANK	1.394172
10	HDFCBANK	1.047793

Table 2 Volatility ranking for 2022.

HDFC Bank is a leading private sector bank in India, while Hindustan Unilever is a leading consumer goods company that produces a range of products including home care, personal care, and food and beverages. Both companies operate in relatively stable and mature industries that are less subject to sudden changes and disruption

HDFC Bank may be affected by changes in interest rates, credit quality, and regulatory changes, while Hindustan Unilever may be affected by changes in consumer preferences and competition. The lower volatility of these companies may reflect the relative stability and predictability of their industries, as well as their established market positions and strong financial performance.

The chart below (Figure 2) shows the predicted volatility of Infosys stock price against its true returns.

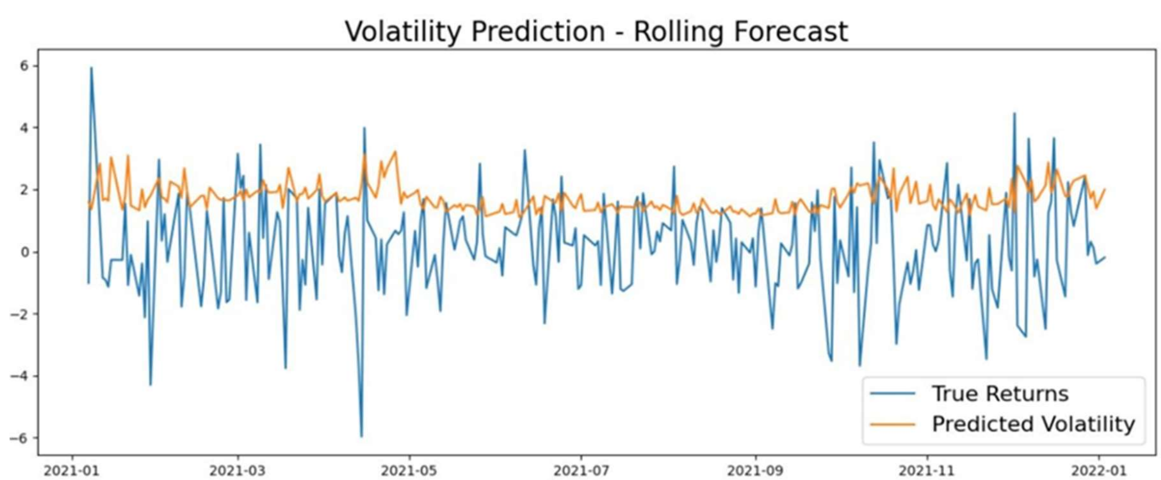

Fig. 2 Volatility chart for Infosys during 2021

The chart below (Figure 5) shows the predicted volatility of Bajaj Finance stock price against its true returns.

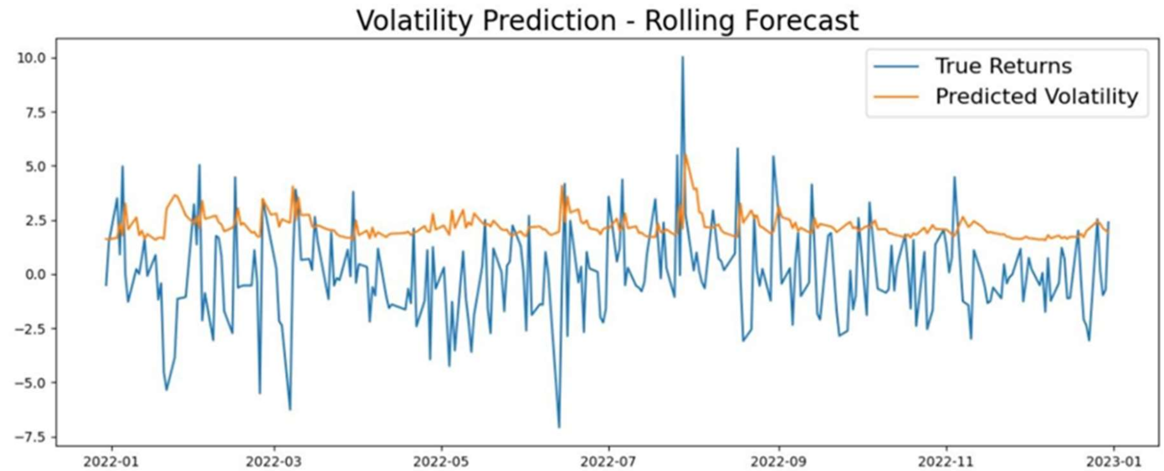

Fig. 3 Volatility chart for Bajaj Finance during 2022

The chart below (Figure 4) shows the predicted volatility of HDFC Bank stock price against its true returns.

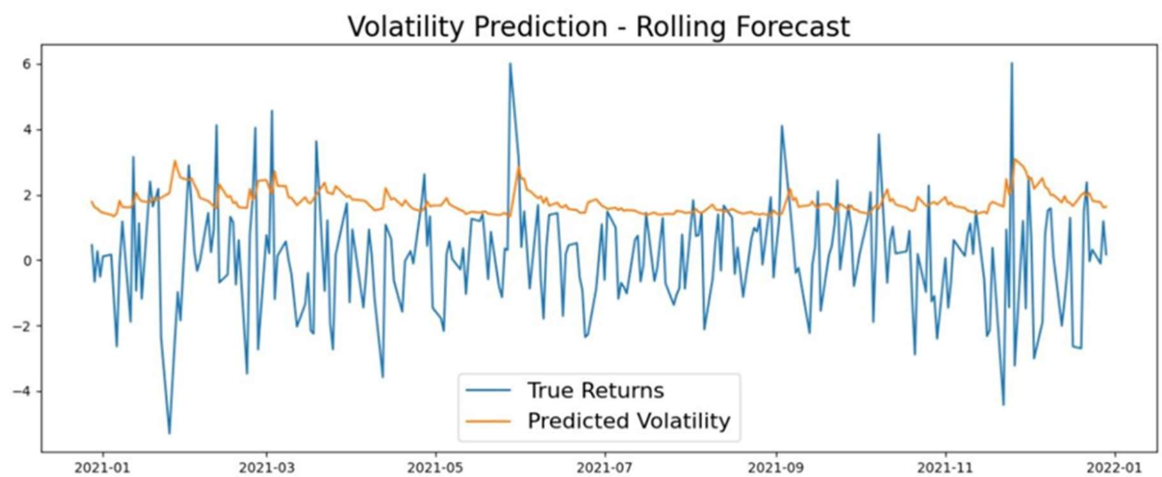

Fig. 4 Volatility chart for HDFC Bank during 2022

The chart below (Figure 5) shows the predicted volatility of Hindustan Unilever stock price against its true returns.

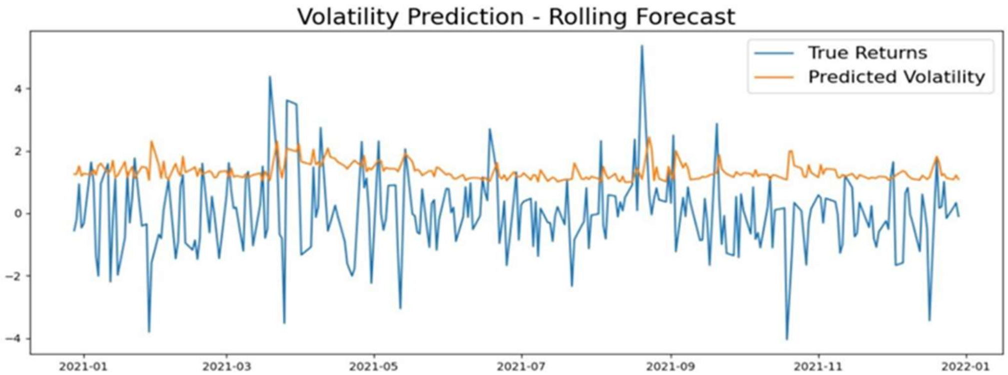

Fig. 5 Volatility chart for Hindustan Unilever during 2022

Rank 2-7	Months					
	January	February	March	April	May	June
1	**RELIANCE**	**HDFC**	**BAJAJ FINSV**	**RELIANCE**	**INFY**	**BAJAJ FINSV**
2	**HDFC**	**INFY**	**KOTAKBANK**	**KOTAKBANK**	**BAJAJ FINSV**	**KOTAKBANK**
3	**HDFC BANK**	**BAJAJ FINSV**	**HDFC**	**LT**	**TCS**	**RELIANCE**

Table 3 Monthly volatility ranking for first six months of 2021.

Rank 2-7	Months					
	July	August	September	October	November	December
1	**LT**	**INFY**	**HDFCBANK**	**RELIANCE**	**HDFC**	**INFY**
2	**RELIANCE**	**HDFC**	**BAJAJFINSV**	**HDFC**	**HINDUNILVR**	**RELIANCE**
3	**BAJAJFINSV**	**BAJAJFINSV**	**HDFC**	**INFY**	**RELIANCE**	**BAJAJFINSV**

Table 4 Monthly volatility ranking for last six months of 2021.

The four tables (Table 3 - Table 6) present data on monthly volatility ranking for the years 2021 and 2022. It can be observed that RELIANCE has a high ranking in January, April and October 2021, and March 2022. BAJAJFINSV is ranking first in March, June 2021 and May, August, September, and November 2022. TCS and LT have grabbed the first spot only once during the two years- in January 2022 and July 2021 respectively. HDFC also has the top rank in the months February and November 2021, February, October, and December 2022. INFY occupies first rank only in 2021.

Rank 2-7	Months					
	January	February	March	April	May	June
1	TCS	HDFC	RELIANCE	KOTAKBANK	BAJAJFINSV	KOTAKBANK
2	RELIANCE	INFY	INFY	ITC	INFY	BAJAJFINSV
3	HDFCBANK	BAJAJFINSV	KOTAKBANK	LT	RELIANCE	RELIANCE

Table 5 Monthly volatility ranking for first six months of 2022.

Rank 2-7	Months					
	July	August	September	October	November	December
1	KOTAKBANK	BAJAJFINSV	BAJAJFINSV	HDFC	BAJAJFINSV	HDFC
2	BAJAJFINSV	HDFCBANK	RELIANCE	BAJAJFINSV	INFY	KOTAKBANK
3	HDFC	RELIANCE	LT	LT	HDFC	HDFCBANK

Table 6 Monthly volatility ranking for last six months of 2022.

during the months of May and August. HDFCBANK also tops during September 2021. KOTAKBANK, not having been in the first place in 2021 takes place in 2022 during the months of April, June, and July.

Conclusion

In this research, we have analysed the volatility of ten companies from NIFTY50 and identified that Infosys and Bajaj finance as most volatile with 3.077 and 2.896 conditional variance respectively. Both companies operate in highly competitive and rapidly changing industries that are subject to a range of economic, technological, and regulatory risks.

Limitations

The analysis only considers volatility as a factor in ranking the companies and does not consider other important factors such as financial performance, management quality, competitive position, and market trends. These factors could have a significant impact on investment decisions and should be considered in a more comprehensive analysis. The analysis only includes ten companies, which may not be representative of the broader market or economy. A larger sample size could provide more robust insights into the volatility of different industries and sectors.

References

Dana, A. N. (2016). Modelling and estimation of volatility using ARCH/GARCH models in Jordan's stock market. Asian Journal of Finance Accounting, 8(1), 152-167.

Dong, Y., Li, S., Gong, X. (2017, April). Time series analysis: An application of ARIMA model in stock price forecasting. In 2017 International Conference on Innovations in Economic Management and Social Science (IEMSS 2017) (pp. 703-710). Atlantis Press.

Emenike, K. O. (2010). Modelling Stock Returns Volatility in Nigeria Using GARCH Models. African Journal of Management and Administration, 3(1), 8-15.

[[4] Kosapattarapim, Chaiwat; Lin, Yan-Xia, McCrae, Michael. (2011). Evaluating the volatility forecasting performance of best fitting GARCH models in emerging Asian stock markets. Centre for Statistical and Survey Methodology, University of Wollongong, Working Paper 13-11.

Maqsood, A., Safdar, S., Shafi, R., Lelit, N. J. (2017). Modeling stock market volatility using GARCH models: A case study of Nairobi Securities Exchange (NSE). Open Journal of Statistics, 7(2), 369-381.

Miah, M., Rahman, A. (2016). Modelling volatility of daily stock returns: Is GARCH (1, 1) enough. American Scientific Research Journal for Engineering, Technology, and Sciences (ASRJETS), 18(1), 29-39.

Schmidt, L. (2021). Volatility Forecasting Performance of GARCH Models: A Study on Nordic Indices During COVID-19.

Shamiri, A. and Isa, Z. (2009) Modeling and Forecasting Volatility of the Malaysian Stock Markets. Journal of Mathematics and Statistics, 5, 234-240.

Ugurlu, E., Thalassinos, E., Muratoglu, Y. (2014). Modeling volatility in the stock markets using GARCH models: European emerging economies and Turkey.

Wang, Y., Xiang, Y., Lei, X., Zhou, Y. (2022). Volatility analysis based on GARCH-type models: Evidence from the Chinese stock market. Economic Research-Ekonomska Istraˇzivanja, 35(1), 2530-2554.