Forecasting Realized Volatility in Cryptocurrency Markets Using GARCH Models

Andrey Sadchikov, Vladimir Yurchenko, and George Erykalkin

Moscow Institute of Physics and Technology, Moscow, Russia

Abstract. This paper investigates the efficacy of Generalized Autoregressive Conditional Heteroskedasticity (GARCH) models in forecasting realized volatility of Bitcoin returns. By employing both rolling-window and expanding-window approaches, we assess the predictive performance of various GARCH-type models against naive benchmarks. Our findings indicate that GARCH-based models significantly outperform naive models. Additionally, we examine the distributional properties of returns and realized volatility, establishing normality, which justifies the choice of normal distribution in GARCH modeling. This research contributes to the understanding of volatility dynamics in cryptocurrency markets, emphasizing the importance of advanced modeling techniques for effective risk management and trading strategies.

1 Introduction

Volatility forecasting is a crucial aspect of financial econometrics, particularly in the context of rapidly evolving markets such as cryptocurrencies. The unpredictability and extreme fluctuations in asset prices pose significant challenges for investors and risk managers. The Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model, first introduced by (2), provides a robust framework for modeling time-varying volatility.

This study aims to evaluate the effectiveness of GARCH-type models (GARCH, EGARCH, APARCH, etc.) in predicting realized volatility of Bitcoin returns compared to naive models. We employ a comprehensive methodology, including distributional tests and rolling-window predictions, to provide insights into the performance of these models.

2 Historical Overview and Related Work

- **AR and MA Models (1970s)**
  Before the introduction of the GARCH family of models, autoregressive (AR) and moving average (MA) models, as well as their combination—ARMA models—were widely employed. However, these models failed to account for time-varying variance (volatility), which posed challenges when analyzing financial data. Another issue with ARMA-based models was that they failed in caputing so-callued volatility clustering phenomena, where periods of high and low volatility alternate, which is common amongst financial data.

- **ARCH Model (1982)**
  The ARCH (Autoregressive Conditional Heteroskedasticity) model was introduced in 1982 (5). This model is based on the premise that the variance of the random errors in a time series depends on past values of these errors.
- **GARCH Model (1986)** (2)
  The GARCH model extends this idea by incorporating the dependence of variance not only on past errors but also on past values of the variance itself.
- **EGARCH (Exponential GARCH, 1991)** (9) accounts for asymmetric effects, such as the observation that negative news tends to have a stronger impact on volatility.
- **GJR-GARCH (1993)** also addresses asymmetry but employs a different mechanism.
- **TARCH (Threshold GARCH, 1994)** (6) incorporates threshold effects in volatility modeling.
- **Multivariate GARCH models** (8) are used to model covariances between different assets.

These models have established GARCH as one of the fundamental tools for analyzing financial time series, particularly in the context of risk and volatility forecasting in stock market settings (4). Later on, with the emergence and development of cryptocurrencies, GARCH has also become applicable for assessing the volatility of crypto assets such as Bitcoin. This is exemplified in recent research that explores GARCH-based models for predicting Bitcoin realized volatility (3).

Subsequently, researchers began combining various machine learning techniques, like Long-Short Term Memory (7) or Transformer Networks (10) to enhance results of classical models, which is detailed in the article (1).

3 Methodology

3.1 Data Collection

We obtained hourly closing price data for BTC-USD and ETH-USD from January 1, 2023, to October 1, 2024, using the yfinance API.

3.2 GARCH Modeling

Distribution Choice We conducted a series of distributional tests (see Table 1) to determine the return distributions. To exclude heteroskedastic factors, we evaluated the distribution of $\zeta_t := \frac{R_t}{\sigma_t}$, utilizing the Shapiro-Wilk and Kolmogorov-Smirnov tests. Our analysis aimed to validate the assumption of normality for GARCH modeling. The plots in Figures 1 and 2 represent ζ_t calculated from ETH-USD data.

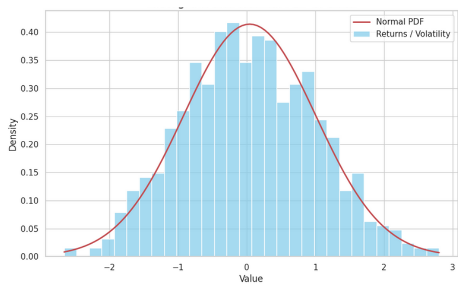

Fig. 1: Normal PDF vs. ζ_t

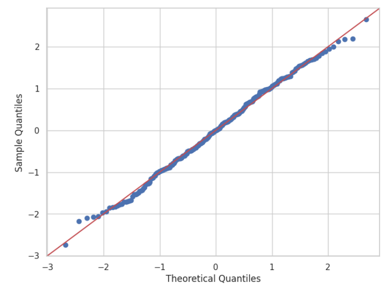

Fig. 2: QQ-Plot of ζ_t

	Shapiro-Wilk Test		**Kolmogorov-Smirnov Test**	
	Statistic	**p-value**	**Statistic**	**p-value**
BTC-USD	0.992	0.153	0.049	0.5214
ETH-USD	0.995	0.5781	0.036	0.863

Table 1: Statistical Tests Results for BTC-USD and ETH-USD

Building GARCH Predictions The GARCH model directly predicts the current volatility, represented as follows:

$$\hat{\sigma}_t^2 = \omega + \sum_{i=1}^{q} \alpha_i u_{t-1}^2 + \sum_{j=1}^{p} \beta_j \tilde{\sigma}_{t-j}^2 \tag{1}$$

Here, $\tilde{\sigma}_t^2$ denotes the conditional volatility at time t.

In order to prevent data leaks, the final predction of model was obtained in the expanding window style: model was fitted on data from the moment $t = 0$ to the moment $t - 1$ and then made estimation $\hat{\sigma}_t$. The result of such process can be witnessed at the Figure 3.

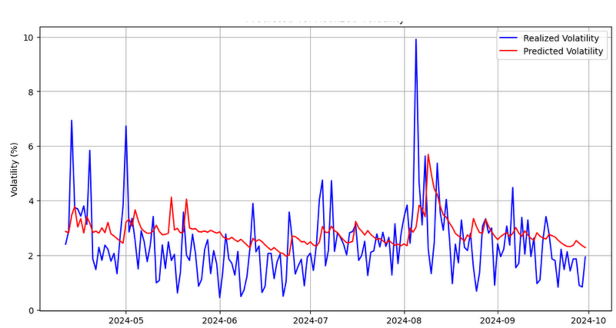

Fig. 3: Results of GARCH(1,1) Predictions

3.3 Evaluation

In order to estimate model performance, we computed Mean Squared Error (MSE) and Mean Absolute Error (MAE) of realized volatility values σ_t (i.e., ground truth) and GARCH predctions $\hat{\sigma}_t$.

$$\mathrm{MSE}(\{\sigma_t\}_{t=1}^T, \{\hat{\sigma}_t\}_{t=1}^T) := \frac{1}{T}\sum_{t=1}^{T}(\sigma_t - \hat{\sigma}_t)^2 \tag{2}$$

$$\mathrm{MAE}(\{\sigma_t\}_{t=1}^T, \{\hat{\sigma}_t\}_{t=1}^T) := \frac{1}{T}\sum_{t=1}^{T}|\sigma_t - \hat{\sigma}_t| \tag{3}$$

3.4 Realized Volatility Computation

To compute the realized volatility on day t, we calculated the daily standard deviation of hourly returns r_t^i, adjusting for the number of trading hours per day:

$$\text{Realized Volatility on day } t = \sigma_t = \mathrm{Std}(r_t^1, \ldots, r_t^{24}) \cdot \sqrt{24} \tag{4}$$

3.5 Model Comparison

We used the *volatility persistence model* as our benchmark naive model, which posits that future volatility can be predicted from its most recent value. Mathematically, this is expressed as:

$$\sigma_{t+1} = \sigma_t \tag{5}$$

In this equation, σ_{t+1} represents the expected future volatility, while σ_t is the current volatility. This model serves as a simple baseline against which we can measure the performance of more complex forecasting methods.

In order to effectively capture the difference between the perfomance of naive and advanced models, we introduce Mean Scaled Squared Error (MSSE) and Mean Absolute Squared Error (MASE) metrics:

$$\mathrm{MSSE}(\{\sigma_t\}_{t=1}^T, \{\hat{\sigma}_t\}_{t=1}^T) := \frac{\sum_{t=2}^{T}(\sigma_t - \hat{\sigma}_t)^2}{\sum_{t=2}^{T}(\sigma_t - \sigma_{t-1})^2} \tag{6}$$

$$\mathrm{MASE}(\{\sigma_t\}_{t=1}^T, \{\hat{\sigma}_t\}_{t=1}^T) := \frac{\sum_{t=2}^{T}|\sigma_t - \hat{\sigma}_t|}{\sum_{t=2}^{T}|\sigma_t - \sigma_{t-1}|} \tag{7}$$

Here $\hat{\sigma}_t$ is considered as the prediction of advanced model.

4 Results

In order to maintain readability we present only few examples of models evaluation metrics in the Tables 2 and 3. As can be evident from them, GARCH models significantly outperformed the naive volatility persistence model in our analysis. Among the various GARCH specifications tested, the **GJR-GARCH(1, 2)** model exhibited nearly the best overall performance. However, it is important to note that the differences in performance among the various GARCH models were pretty huge, typically differing up to 30%. This indicates that while GARCH models generally provide better forecasts than the naive model, we still should dedicate some time when choosing the best model for our particular data.

Table 2: Performance Metrics for Model Evaluation on BTC-USD data

Model	MSE	MAE	MSSE^{-1}	MASE^{-1}
Naive	2.0030	1.0740	1.0000	1.0000
ARCH(1)	1.7230	1.0075	1.1627	1.0661
GARCH(1,2)	1.6628	0.9595	1.2048	1.1195
EGARCH(1,1,0)	1.6292	0.9514	1.2296	1.1289
GJR-GARCH(1,2)	1.5327	0.9486	1.3071	1.1323
GJR-GARCH(3,3)	1.6977	0.9430	1.1800	1.1391
APARCH(1,2)	1.6086	0.9665	1.2454	1.1114

Table 3: Performance Metrics for Model Evaluation on ETH-USD data

MSE	MAE	MSSE^{-1}	MASE^{-1}
3.0700	1.2190	1.0000	1.0000
2.8084	1.2082	1.0931	1.0085
2.7105	1.1478	1.1326	1.0616
2.2285	1.0770	1.3775	1.1314
2.2823	1.0835	1.3451	1.1246
2.2734	1.0780	1.3504	1.1303
2.458080	1.1081	1.2489	1.0996

5 Conclusion

As it was discussed in Section 4, any GARCH-based model can be a reliable choice for volatility forecasting, especially in complex markets like Bitcoin. However, it is still important to choose optimal type of model and parameters for practical use.

Another important result was established in Subsection 3.2, stating that normalized returns are likely to normally distributed

$$\zeta_t = \frac{R_t}{\sigma_t} \overset{d}{\sim} N(0,1)$$

where R_t represents returns on day t.

Future research could apply these models to other asset classes or integrate macroeconomic factors and machine learning techniques, but overall, GARCH models remain essential for risk management and trading strategies in volatile markets.

Bibliography

[1] Amirshahi, B., Lahmiri, S.: Hybrid deep learning and garch-family models for forecasting volatility of cryptocurrencies. Machine Learning with Applications **12**, 100465 (2023). https://doi.org/https://doi.org/10.1016/j.mlwa.2023.100465, https://www.sciencedirect.com/science/article/pii/S266682702300018X

[2] Bollerslev, T.: Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics **31**(3), 307–327 (1986). https://doi.org/https://doi.org/10.1016/0304-4076(86)90063-1, https://www.sciencedirect.com/science/article/pii/0304407686900631

[3] Chu, J., Chan, S., Nadarajah, S., Osterrieder, J.: Garch modelling of cryptocurrencies. Journal of Risk and Financial Management **10**(4) (2017). https://doi.org/10.3390/jrfm10040017, https://www.mdpi.com/1911-8074/10/4/17

[4] Engle, R.: Garch 101: The use of arch/garch models in applied econometrics. Journal of Economic Perspectives **15**(4), 157–168 (December 2001). https://doi.org/10.1257/jep.15.4.157, https://www.aeaweb.org/articles?id=10.1257/jep.15.4.157

[5] Engle, R.F.: Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. Econometrica **50**(4), 987–1007 (July 1982), https://ideas.repec.org/a/ecm/emetrp/v50y1982i4p987-1007.html

[6] Guo, H., ZmeÅ¡kal, Z.: Tarch model-based dynamic hedging strategy of adr portfolios. International Journal of Economics and Accounting **11**(2), 199–211 (2022), https://EconPapers.repec.org/RePEc:ids:ijecac:v:11:y:2022:i:2:p:199-211

[7] Hochreiter, S., Schmidhuber, J.: Long short-term memory. Neural Computation **9**(8), 1735–1780 (1997)

[8] Silvennoinen, A., Teräsvirta, T.: Multivariate GARCH Models, pp. 201–229. Springer Berlin Heidelberg, Berlin, Heidelberg (2009). https://doi.org/10.1007/978-3-540-71297-8_9, https://doi.org/10.1007/978-3-540-71297-8_9

[9] St. Pierre, E.F.: Estimating egarch-m models: Science or art? The Quarterly Review of Economics and Finance **38**(2), 167–180 (1998). https://doi.org/https://doi.org/10.1016/S1062-9769(99)80110-0, https://www.sciencedirect.com/science/article/pii/S1062976999801100

[10] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, L., Polosukhin, I.: Attention is all you need (2023), https://arxiv.org/abs/1706.03762

6 Appendix

All the supporting code can be found at GitHub.