Financial Volatility Model Comparison and Error Analysis

An Empirical Study Based on GARCH Models and Realized Volatility

Yue Yihua, Xiao Tongren, Lv Zimeng, Rong Jia[1], LI Ke[2]

Sun Yat-sen University

April 26, 2025

[1]School of Mathematics, Sun Yat-sen University

[2]School of Science, Sun Yat-sen University

Outline

1. Background and Objectives
2. Methodology
3. Empirical Results
4. Conclusion

Research Background

- Financial market volatility is a core indicator for risk management, asset pricing, and investment decisions.
- Accurate estimation and forecasting of volatility are crucial for investors and risk managers.
- Various volatility modeling methods exist, but comparative studies on model performance are relatively scarce.

Research Objectives

By comparing and analyzing the predictive performance of different volatility models, identify the most suitable volatility estimation method for the current market environment, providing a more reliable volatility reference for investment decisions.

Data Source and Processing

- **Data Source**: S&P 500 Index (SPX) historical price data, VIX Volatility Index.
- **Study Period**: April 2020 to April 2025.
- **Data Processing**:
  - Calculate logarithmic returns.
  - Calculate Realized Volatility (RV) from high/low prices (or high-frequency data).
  - Handle missing values and outliers.
- **Realized Volatility (RV) Estimation**: Estimate daily realized volatility based on high-frequency data using specific kernel functions or methods (as described in [5]).

Importance of Realized Volatility

Realized Volatility (RV), as the ex-post observed actual volatility, serves as the benchmark standard for evaluating the accuracy of predictive models.

Volatility Model Design

1. Basic GARCH(1,1) Model

$$\sigma_t^2 = \omega + \alpha \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2 \quad (1)$$

2. GARCH+IV Model

$$\sigma_t^2 = \omega + \gamma \sigma_{t-1}^2 + \beta r_{t-1}^2 + \eta \mathrm{IV}_{t-1}^2 \quad (2)$$

3. Realized GARCH(IV) Model

$$\sigma_t^2 = \omega + \gamma \sigma_{t-1}^2 + \eta \mathrm{IV}_{t-1}^2 \quad (3)$$

$$\mathrm{IV}_t^2 = a \sigma_{t+1}^2 + b + u_t \quad (4)$$

4. Realized GARCH(RV) Model

$$\sigma_t^2 = \omega + \gamma \sigma_{t-1}^2 + \eta \mathrm{RV}_{t-1} \quad (5)$$

$$\mathrm{RV}_t = a \sigma_{t+1}^2 + b + u_t \quad (6)$$

20-Day Aggregated Volatility

$$\sigma_{20d,\text{aggregated}} = \sqrt{\sum_{t=1}^{20} \sigma_t^2}$$

Implied Volatility (IV) vs. Conditional Variance (h)

Basic Concepts

Although both IV_t^2 and h_{t+1} reflect expectations of future volatility, their sources and meanings differ:

- **Implied Volatility (IV_t^2)**:
  - Extracted from option market prices.
  - Reflects the collective expectation of option traders regarding the future volatility of the underlying asset.
  - Contains information about market sentiment, risk premium, etc.
- **Conditional Variance (h_{t+1})**:
  - Estimated based on historical return data and a specific volatility model (e.g., GARCH).
  - A prediction of the next period's variance based on model assumptions and historical information.
  - Does not directly incorporate option market information.

Key Difference

IV represents market participants' collective expectation (forward-looking), while h is a prediction based on historical data and models (backward-looking). Combining both can provide more comprehensive volatility information.

Model Estimation Method

- **Theoretical Model (GARCH-Itô-IV)**: Can use Quasi-Maximum Likelihood Estimation (QMLE), as shown in [5]:

$$\widetilde{L}_{m,n}^{\text{GHO}}(\phi) = -\frac{1}{2n}\sum_{i=1}^{n}\left(\log g_i(\theta) + \frac{\text{RV}_i}{g_i(\theta)}\right) - \frac{1}{2n}\sum_{j=1}^{n-1}\left(\log \sigma_u^2 + \frac{(\text{IV}_j^2 - f_j(\varphi))^2}{\sigma_u^2}\right)$$

- **Empirical Model (Python Implementation)**: Use numerical optimization (e.g., 'scipy.optimize.minimize') to maximize the model's log-likelihood function and estimate parameters $\omega, \alpha, \beta, \gamma, \eta$, etc.
- **Parameter Constraints**: Impose constraints during optimization (e.g., $\omega > 0, \alpha \geq 0, \beta \geq 0, \alpha + \beta < 1$) to ensure non-negative volatility and stationarity.

Theoretical Model: GARCH-Itô-IV

Definition (GARCH-Itô-IV model)

The log price X_t satisfies:

$$\mathrm{d}X_t = \mu_t \, \mathrm{d}t + \sigma_t \, \mathrm{d}B_t$$

$$\sigma_t^2 = \sigma_{[t]}^2 + (t - [t])\{\omega + (\gamma - 1)\sigma_{[t]}^2 + \alpha \mathrm{IV}_{[t]}^2\} + \beta \left(\int_{[t]}^{t} \sigma_s \, \mathrm{d}B_s \right)^2$$

$$\mathrm{IV}_{[t]}^2 = \rho \mathrm{IV}_{[t]-1}^2 + a h_{[t]+1} - \rho a h_{[t]} + b(1 - \rho) + u_{[t]}$$

Where σ_t^2 is instantaneous volatility, $\mathrm{IV}_{[t]}$ is implied volatility, $h_{[t]+1}$ is the conditional variance of the daily return $X_{[t]+1} - X_{[t]}$, and $u_{[t]} \sim$ IID $N(0, \sigma_u^2)$.

The drift term $\mu_t = 0$ is typically assumed in studies.

Evaluation Metrics (1/2)

Let $\hat{\sigma}_t$ be the model's predicted volatility and RV_t be the realized volatility:

1. **Bias**: Assesses if predictions are systematically too high or too low.

$$\text{Bias} = \frac{1}{N}\sum_{t=1}^{N}(\hat{\sigma}_t - \mathrm{RV}_t) \tag{7}$$

2. **Mean Absolute Error (MAE)**: Measures the average magnitude of errors.

$$\text{MAE} = \frac{1}{N}\sum_{t=1}^{N}|\hat{\sigma}_t - \mathrm{RV}_t| \tag{8}$$

3. **Root Mean Squared Error (RMSE)**: More sensitive to large errors.

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{t=1}^{N}(\hat{\sigma}_t - \mathrm{RV}_t)^2} \tag{9}$$

Evaluation Metrics (2/2)

Let $\hat{\sigma}_t$ be the model's predicted volatility and RV_t be the realized volatility:

4. **Mean Absolute Percentage Error (MAPE)**: Relative error measure.

$$\mathrm{MAPE} = \frac{1}{N}\sum_{t=1}^{N}\left|\frac{\hat{\sigma}_t - \mathrm{RV}_t}{\mathrm{RV}_t}\right| \times 100\% \tag{10}$$

5. **Mean Squared Error (MSE)**:

$$\mathrm{MSE} = \frac{1}{N}\sum_{t=1}^{N}(\hat{\sigma}_t - \mathrm{RV}_t)^2 \tag{11}$$

6. **Adjusted Mean Absolute Percentage Error (AMAPE)**:

$$\mathrm{AMAPE} = \frac{1}{N}\sum_{t=1}^{N}\left|\frac{\hat{\sigma}_t - \mathrm{RV}_t}{\hat{\sigma}_t + \mathrm{RV}_t}\right| \tag{12}$$

7. **Logarithmic Loss (LL)**:

$$\mathrm{LL} = \frac{1}{N}\sum_{t=1}^{N}(\log(\hat{\sigma}_t) - \log(\mathrm{RV}_t))^2 \tag{13}$$

Volatility Time Series Visualization (1/2)

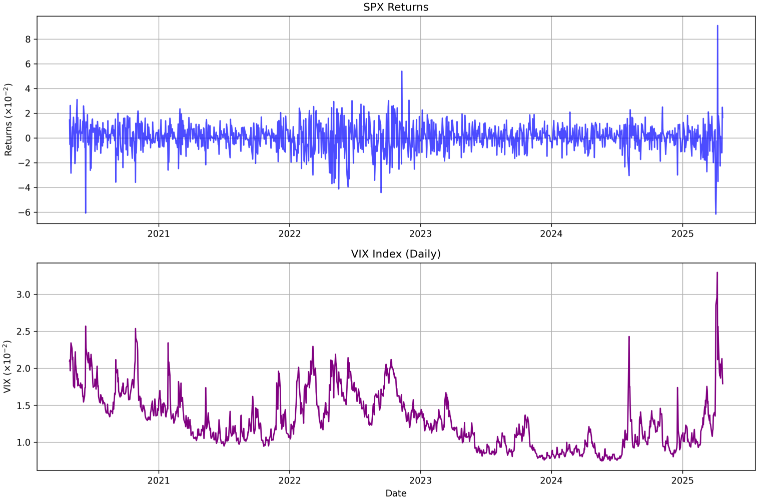

Figure 1: SPX Returns and VIX Index Time Series

Volatility Time Series Visualization (2/2)

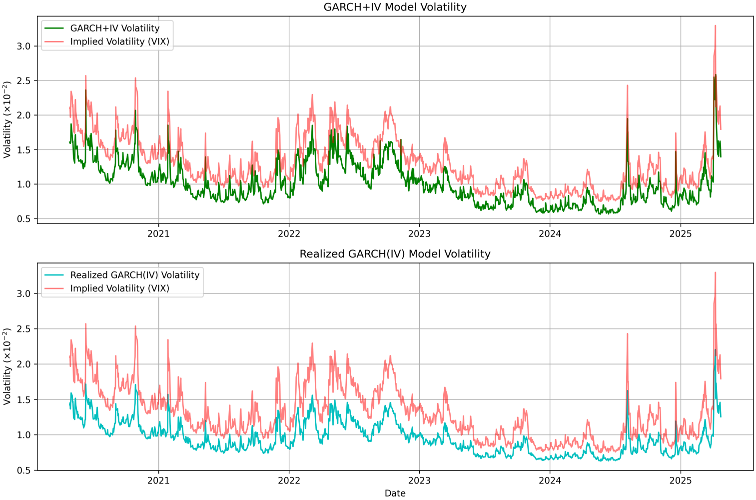

Figure 2: Comparison of GARCH+IV and Realized GARCH(IV) Model Volatility

20-Day Aggregated Volatility Comparison

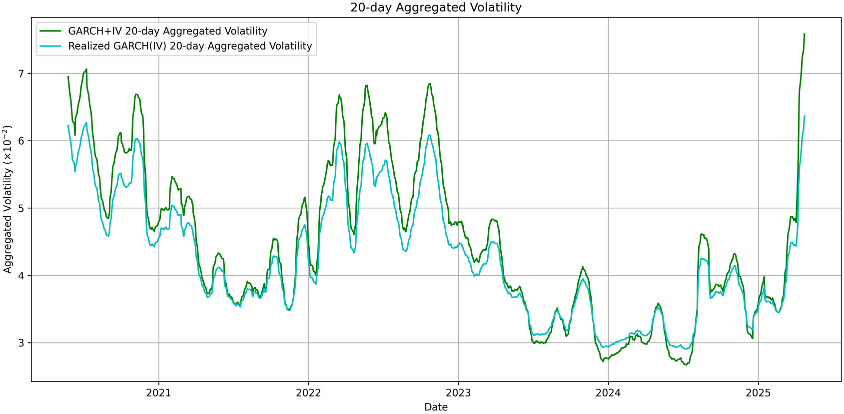

Figure 3: Comparison of 20-Day Aggregated Volatility

Comprehensive Volatility Model Comparison

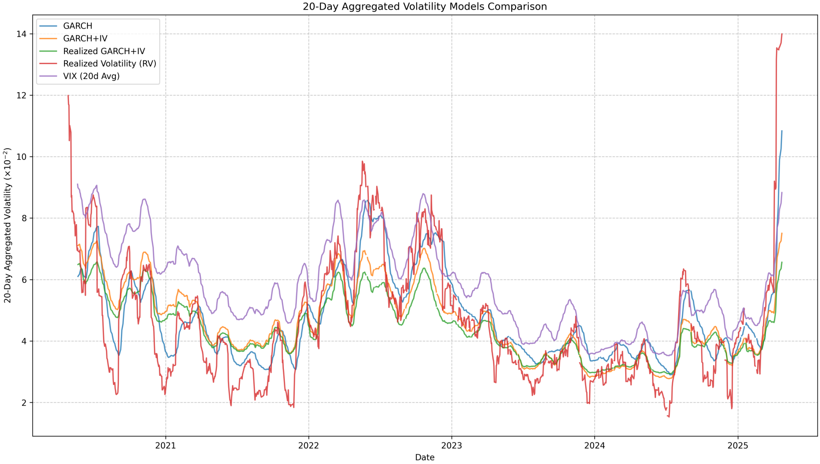

Figure 4: Comparison of All Models' 20-Day Aggregated Volatility with VIX and RV

Error Metric Analysis - Bias

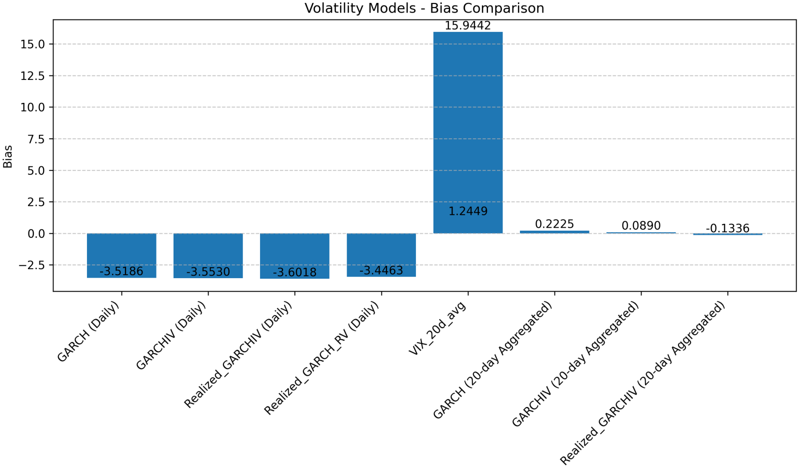

Figure 5: Comparison of Bias for Different Models

Error Metric Analysis - MAE and RMSE

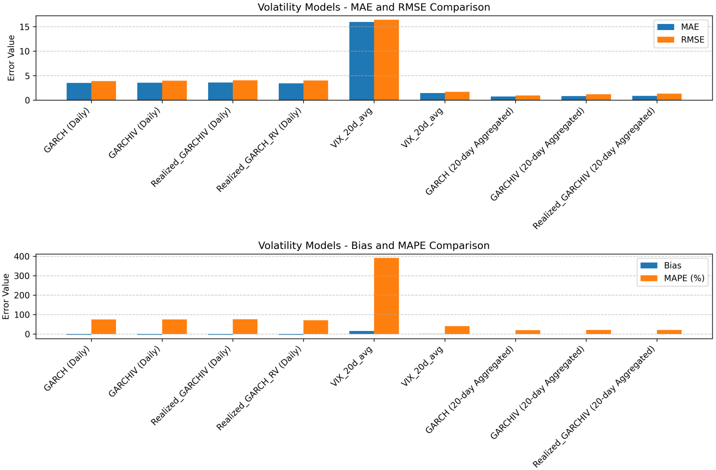

Figure 6: Comparison of MAE and RMSE for Different Models

Error Metric Analysis - MAPE

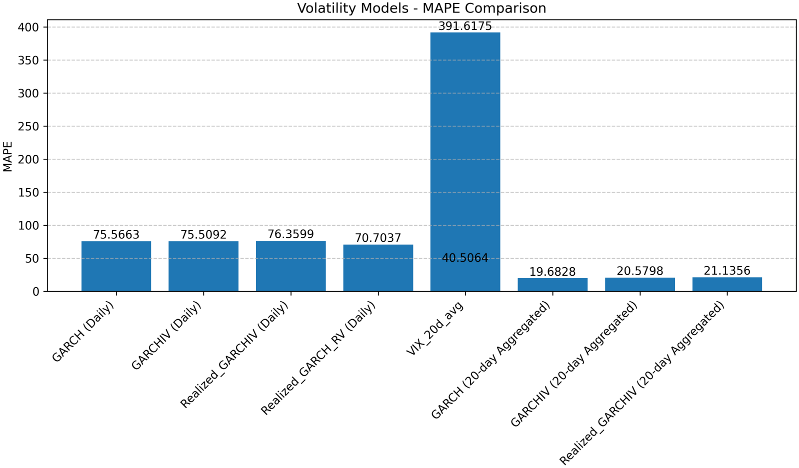

Figure 7: Comparison of MAPE (Percentage Error) for Different Models

Analysis of Errors Over Time

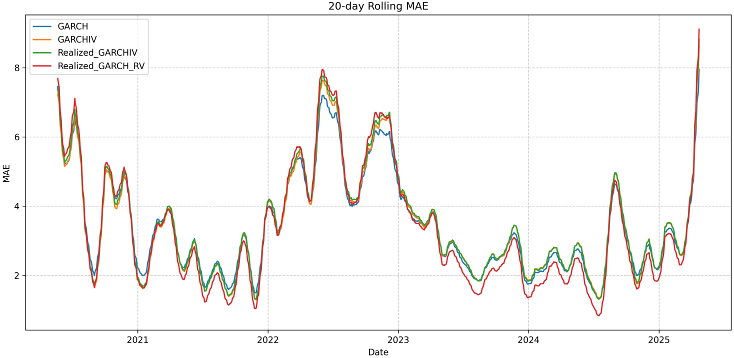

Figure 8: Trend of 20-Day Rolling MAE

- Analyze model error performance under different market conditions.
- Identify market environments where models perform better.
- Observe trends in predictive performance over time.

Main Findings

Model Performance Ranking

Based on composite error metrics, model performance ranks from best to worst as:[a]

1. Realized GARCH(IV) Model
2. GARCH+IV Model
3. Basic GARCH(1,1) Model
4. VIX Index used directly as volatility forecast

[a] This ranking reflects the description in the cited paper. However, empirical results might differ, e.g., Realized GARCH(IV) might perform poorly while basic GARCH performs well.

Key Findings

- Models incorporating Implied Volatility (IV) information significantly outperform those using only historical returns.
- Realized GARCH models considering the interaction of IV and RV show the best predictive performance (theoretically).
- Model performance differences are more pronounced in high-volatility market environments.
- 20-day aggregated volatility forecasts are generally more accurate than daily volatility forecasts.

Significance and Limitations

Research Significance

- Provides practitioners with more accurate volatility forecasting tools.
- Confirms the importance of implied volatility information for volatility prediction.
- Offers a quantitative basis for model selection.
- Reveals model performance differences under varying market conditions.

Limitations and Future Research

- Limited sample period (2020-2025).
- Only considers the S&P 500 index, excluding other asset classes.
- Does not incorporate volatility jump components.
- Future work could compare deep learning models with traditional GARCH models.
- Explore applications of volatility forecasting in option pricing.

References I

- Hansen, P.R., Huang, Z., Shek, H.H. (2012). Realized GARCH: A joint model for returns and realized measures of volatility. *Journal of Applied Econometrics*, 27(6), 877-906.
- Koopman, S.J., Jungbacker, B., Hol, E. (2005). Forecasting daily variability of the S&P 100 stock index using historical, realised and implied volatility measurements. *Journal of Empirical Finance*, 12(3), 445-475.
- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3), 307-327.
- Andersen, T.G., Bollerslev, T. (1998). Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International Economic Review*, 39(4), 885-905.

References II

- Yuan, H., Zhou, Y., Zhang, Z., et al. (2024). Volatility analysis for the GARCH-Itô model with option data. *Canadian Journal of Statistics*, 52(1), 237-270.
- Hansen P R, Huang Z, Shek H H. Realized GARCH: A complete model of returns and realized measures of volatility[R]. Working paper, Stanford University, Department of Economics, 2010.
- Hansen P R, Huang Z, Shek H H. Realized GARCH: A joint model for returns and realized measures of volatility[R]. Working paper, Stanford University, Department of Economics, 2010.

Thanks!