# MODELING FINANCIAL MARKET VOLATILITY

GARCH(1,1) Model

Subham Khinchi
skhinch@ncsu.edu

# Abstract

Volatility plays a central role in financial economics, risk management, and derivative pricing. This project investigates the modeling of financial market volatility using the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model, specifically focusing on the GARCH(1,1) specification. The S&P 500 index is selected as the underlying asset to demonstrate real-world applicability. Through a step-by-step empirical analysis including data cleaning, log return computation, exploratory data analysis, and model estimation, it presents a clear picture of how volatility clustering can be effectively modeled. Model diagnostics are conducted to assess fit and potential extensions are proposed.

## Table of Contents

# 1. Introduction

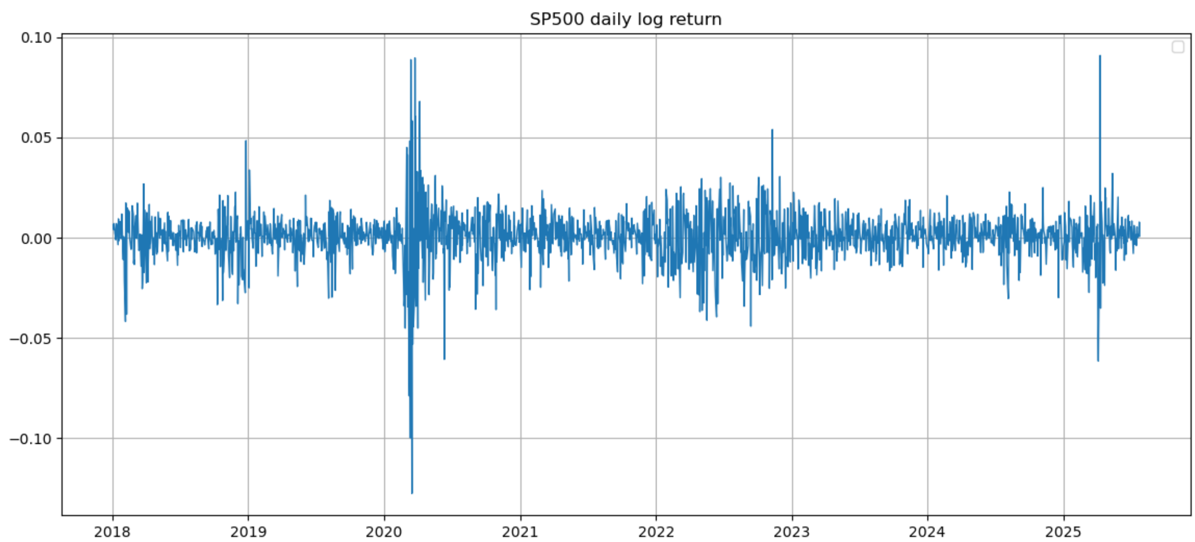
Volatility is a measurement of how varied the returns of a given security or market index are over time. It is widely used in finance to estimate the riskiness of assets. Unlike average returns, volatility reflects uncertainty and is a crucial input in portfolio optimization, derivatives pricing (such as Black-Scholes), and Value-at-Risk models. One characteristic of financial time series is that volatility tends to cluster, periods of high-volatility likely to be followed by high volatility and the same is true for low-volatility periods. This phenomenon violates the assumption of homoskedasticity, thus this project is using models such as GARCH.
The GARCH model addresses the shortcomings of traditional time series models by allowing for time-varying conditional variance. It is particularly useful in modeling heteroskedasticity observed in financial returns.

# 2. Data and Preprocessing

The daily adjusted closing prices of the S&P 500 index were obtained from Yahoo Finance, covering several recent years of market activity. To analyse return dynamics, prices were first transformed into continuously compounded daily log returns, calculated as $r_t = \log(P_t/P_{t-1})$. This transformation stabilizes variance and makes the data more suitable for time series modeling.

Exploratory Data Analysis (EDA)

An exploratory analysis was then conducted, beginning with a time series plot of the log returns to visualize broad patterns and fluctuations over time.



## 2.1. Time Series Stationarity Test

The Augmented Dickey Fuller (ADF) test was performed to assess the time series stationarity, the p-value <0.05 suggesting the log return series is stationary.

Checking for stationarity using Augmented Dickey Fuller test

Start Chat to Generate Code (⌘I)

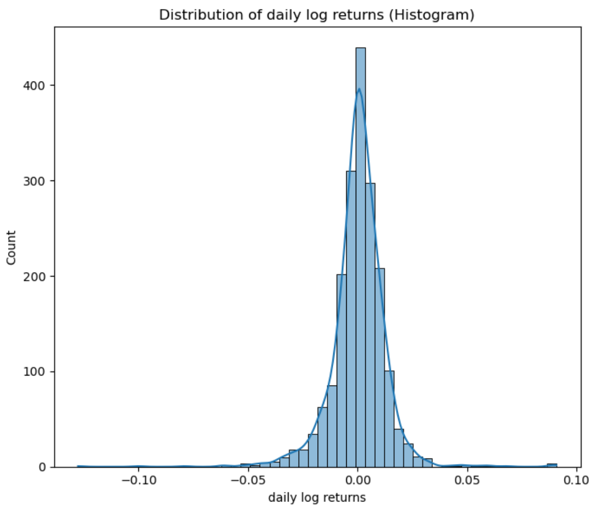
```python
from statsmodels.tsa.stattools import adfuller

adf_result = adfuller(returns)

print(f"ADF Statistic: {adf_result[0]}")
print(f"p-value: {adf_result[1]}")

if adf_result[1] < 0.05:
    print("The return series is stationary.")
else:
    print("The return series is not stationary.")
```

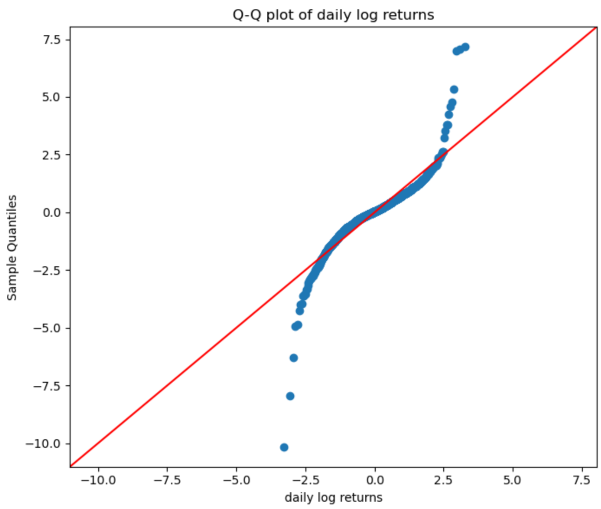
[23] ✓ 0.1s Python

```
ADF Statistic: -13.62548916333653
p-value: 1.7634550202137705e-25
The return series is stationary.
```

## 2.2. Normality of log returns

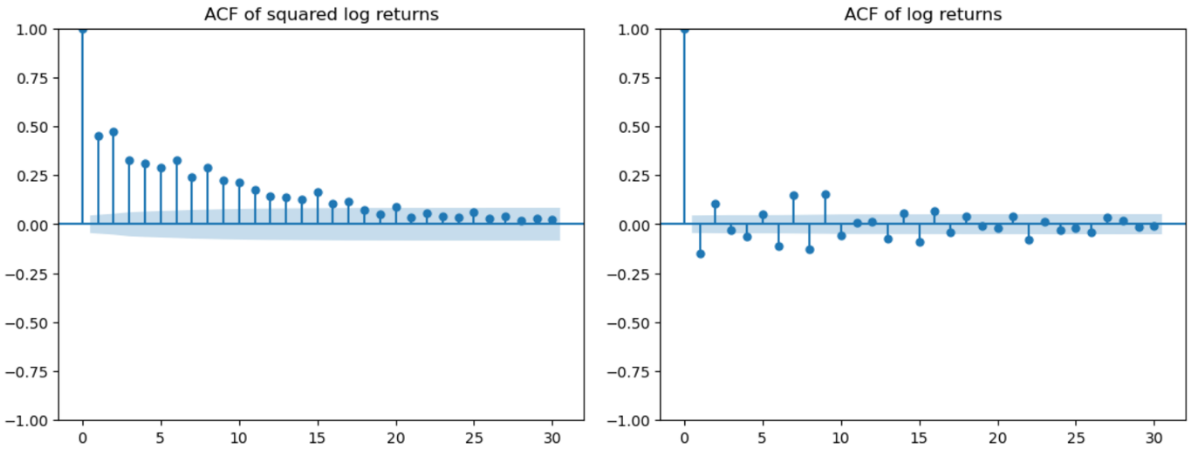
The distributional characteristics of returns were assessed using a histogram and a quantile-quantile (Q-Q) plot, revealing notable deviations from normality. The histogram exhibited heavy tails and Q-Q plot showed significant departures from the reference line, particularly in the tails, indicating the presence of excess kurtosis. Such characteristics not only violate the assumptions of normality commonly made in classical time series models but also highlights the need for models that can accommodate these features. This provides empirical motivation for modeling the volatility using GARCH-type models.



## 2.3. Autocorrelation functions(ACF) of returns

To assess the time series properties of the return series, we examined the autocorrelation functions (ACFs) of both raw log returns and their squared values. The ACF of log returns exhibits no significant autocorrelation beyond lag 0, suggesting that the returns themselves are weakly dependent and resemble a white noise process. However, the left panel reveals strong and persistent autocorrelation in the squared log returns, a clear indication of volatility clustering. This pattern justifies the application of GARCH-type models to capture the time-varying volatility structure present in the data.




# 3. GARCH(1,1) Model Specification

The Generalized Autoregressive Conditional Heteroskedasticity GARCH(1,1) model is widely used to model time-varying volatility in financial return series. It improves upon the ARCH model by incorporating a lag of the conditional variance itself, allowing for greater persistence and smoother volatility dynamics. The specification captures the volatility clustering observed in financial markets, where large shocks tend to be followed by large shocks (of either sign), and small shocks tend to be followed by small shocks.

Let $r_t$ denote the return at time t, and let the return process be decomposed as:

$$\mathrm{r}_t = \mu_t + \epsilon_t$$

Where -

$\mu_t$ is the conditional mean of returns.

$\epsilon_t$ is the residual terms, assumed to follow a conditionally normal distribution with zero mean and variance $\sigma_t^2$

Mean Equation AR(1) process

In the GARCH model, we modeled the mean return process using a simple autoregressive process of order 1, denoted AR(1) to handle the serial correlation in returns, so residuals are less autocorrelated.

$$\mu_t = \phi_0 + \phi_1 r_{t-1}$$

Variance Equation GARCH(1,1) process

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

Where-

$\omega > 0$ : long-run average volatility (constant term)

$\alpha \geq 0$ : reaction to previous squared shocks (**ARCH effect**)

$\beta \geq 0$ : persistence from previous conditional variance (**GARCH effect**)

This formulation provides a simple yet powerful framework for modeling financial time series data with time-varying volatility and has served as the foundation for more advanced models like GARCH, which extends the dependency to both past squared residuals and past variances.

## 4. Model Estimation and Results

To estimate the GARCH(1,1) model, the daily log returns of the S&P500 index were first scaled to improve numerical stability. The model was implemented using the *arch* package in Python which fits an GARCH(1,1) process by maximum likelihood estimation.

Following is the model estimation equation for conditional variance

$$\sigma_t^2 = 0.0473 + 0.1788 \varepsilon_{t-1}^2 + 0.7929 \sigma_{t-1}^2$$

- All three GARCH parameters are highly significant ($p < 0.001$).
- The persistence of volatility, measured as

$$\alpha+\beta = 0.1788 + 0.7927 = 0.9715$$

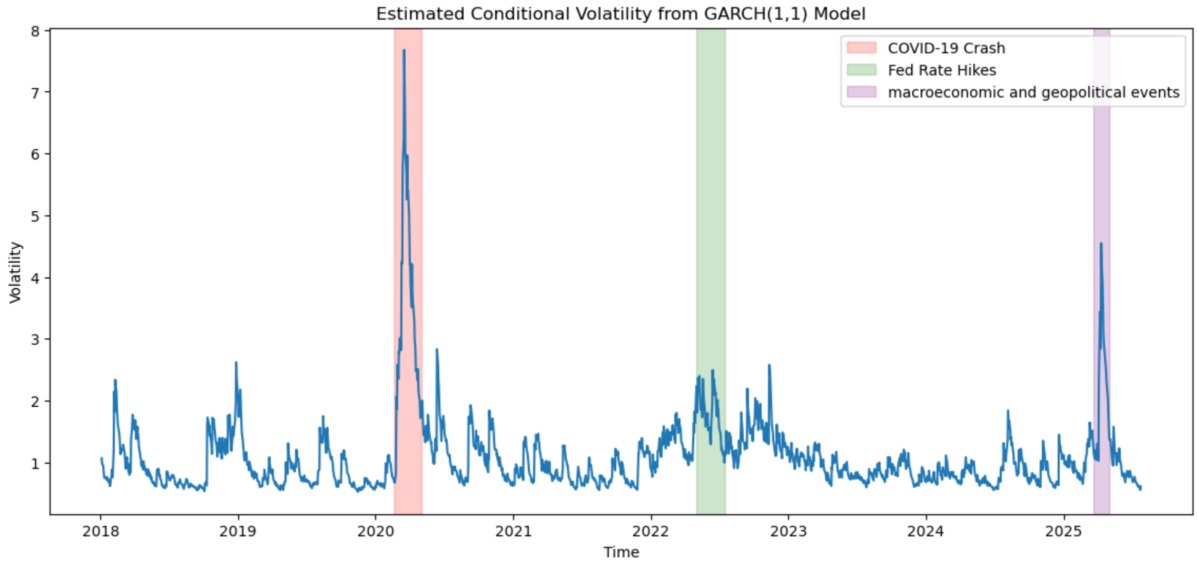
  is very close to 1, indicating strong long-term memory in volatility.

- The relatively large $\beta 1$ value implies that volatility shocks have a long-lasting effect.

```
                       AR - GARCH Model Results
==============================================================================
Dep. Variable:                  SP500   R-squared:                       0.010
Mean Model:                        AR   Adj. R-squared:                  0.009
Vol Model:                      GARCH   Log-Likelihood:               -2662.34
Distribution:                  Normal   AIC:                           5334.67
Method:            Maximum Likelihood   BIC:                           5362.41
                                        No. Observations:                 1898
Date:                Tue, Jul 29 2025   Df Residuals:                     1896
Time:                        03:39:53   Df Model:                            2
                                 Mean Model
=============================================================================
                 coef    std err          t      P>|t|       95.0% Conf. Int.
-----------------------------------------------------------------------------
Const          0.0978  1.988e-02      4.916  8.827e-07    [5.878e-02,  0.137]
SP500[1]      -0.0437  2.636e-02     -1.658  9.741e-02 [-9.534e-02,7.971e-03]
                              Volatility Model
============================================================================
                 coef    std err          t      P>|t|      95.0% Conf. Int.
----------------------------------------------------------------------------
omega          0.0473  1.398e-02      3.386  7.091e-04 [1.994e-02,7.474e-02]
alpha[1]       0.1788  2.959e-02      6.043  1.513e-09     [  0.121,  0.237]
beta[1]        0.7927  2.930e-02     27.057 3.122e-161     [  0.735,  0.850]
============================================================================

Covariance estimator: robust
```

The estimated conditional volatility over time for the S&P 500 daily returns shows Volatility Clustering around early 2020, corresponding to the COVID-19 market crash, mid 2022 as inflation soared and the Federal Reserve raised interest rate, again in 2025, likely due to macroeconomic and geopolitical events.

This chart shows that volatility is time-varying and predictable to some extent using past data. The GARCH(1,1) model is successfully capturing this dynamic. Financial markets often show periods of tranquility followed by bursts of turbulence, this is reflected in the chart through spikes and clusters in volatility.

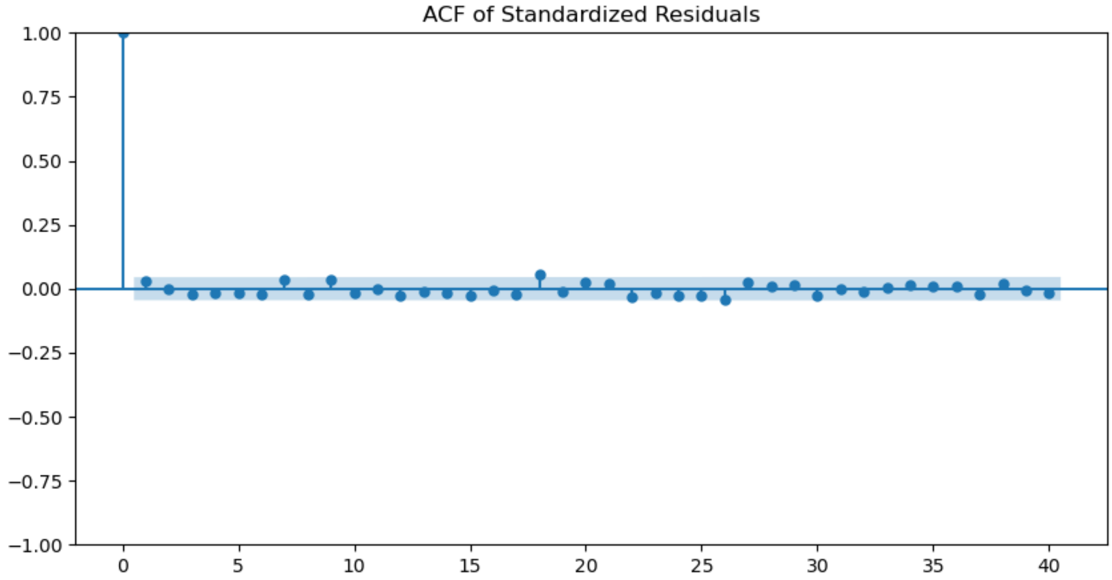



# 5. Model Diagnostics

To assess the adequacy of the GARCH(1,1) model, residual diagnostics were conducted on the standardized residuals and their squared values. These diagnostics provide insight into whether the model has successfully captured the conditional heteroskedasticity present in the returns series.

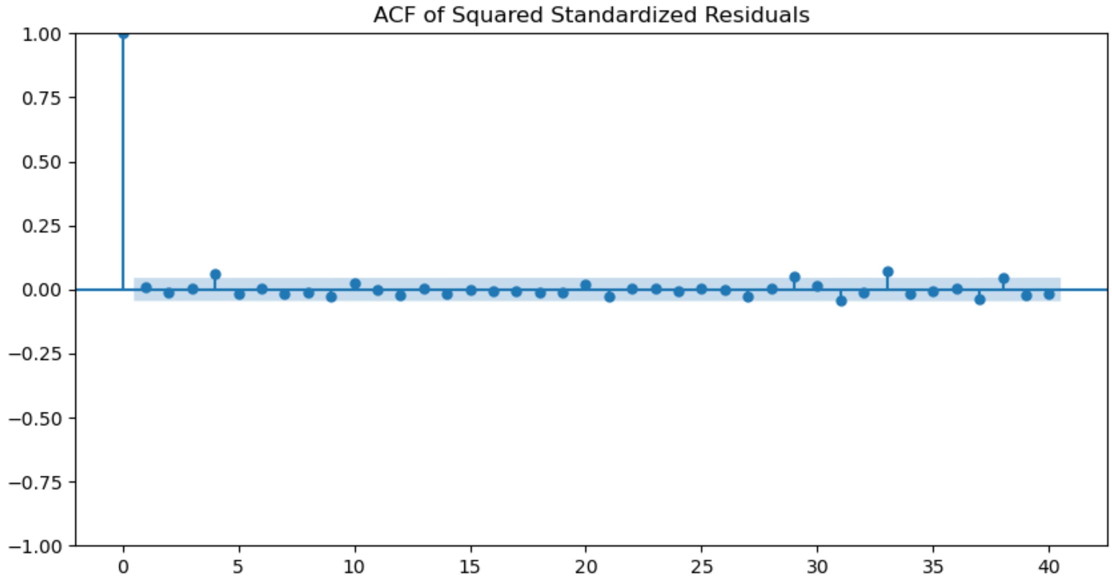
## 5.1. ACF of Standardized Residuals

The ACF plot of the standardized residuals shows no significant autocorrelation at any lag, with all autocorrelation coefficients lying well within the 95% confidence bounds. This suggests that the mean equation, in this case modeled as an AR(1) process, has effectively captured the serial dependence in the returns. The residuals appear to be approximately white noise, satisfying one of the key assumptions of the GARCH framework.




## 5.2. ACF of Squared Standardized Residuals

The ACF plot of the squared standardized residuals is a critical diagnostic to assess whether the GARCH model has successfully captured the conditional heteroskedasticity (volatility clustering). The absence of significant autocorrelation in this plot indicates that the model has effectively accounted for the volatility dynamics in the data. In our case, all autocorrelations lie within the confidence bands, supporting the adequacy of the GARCH(1,1) specification.




## 5.3. Ljung Box Statistical Test

To empirically validate the lack of autocorrelation in the standardized residuals and squared standardized residuals, the Ljung box test was performed and the p-value for standardized residual (0.441596 > 0.05) and p-value for squared standardized residual

(0.272592 > 0.05) suggested there is no autocorrelation in the standardized residuals and their squares.

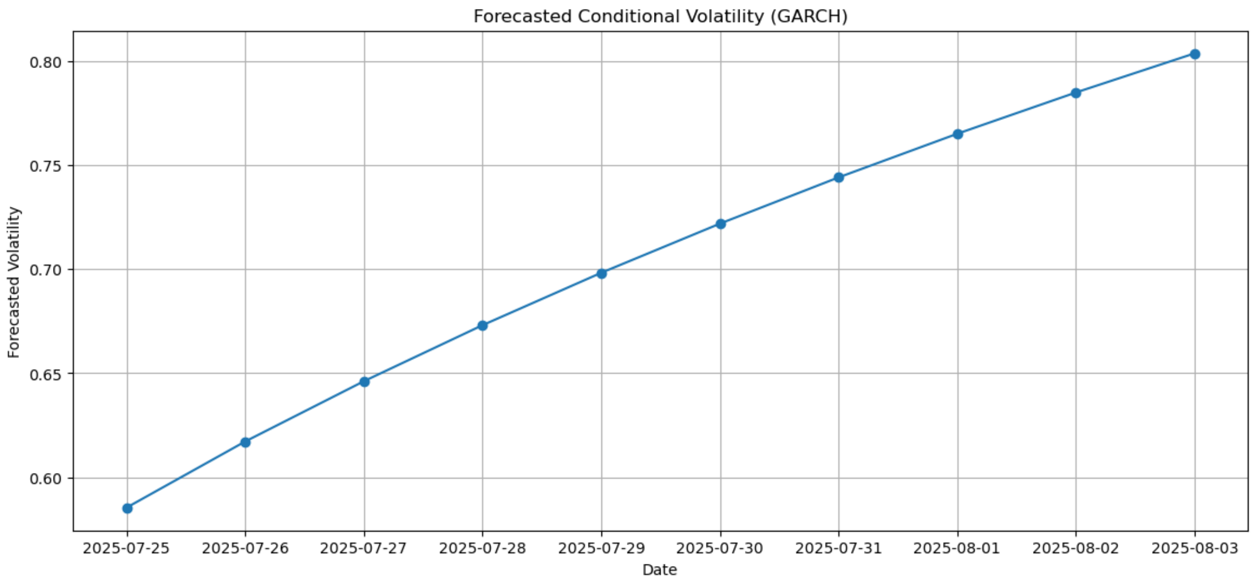
```python
from statsmodels.stats.diagnostic import acorr_ljungbox


std_res_lb_test = acorr_ljungbox(standard_residual, lags=[10], return_df=True)
std_res_sq_lb_test = acorr_ljungbox(standard_residual**2, lags=[10], return_df=True)


print(f"p-value of standard residuals : {std_res_lb_test['lb_pvalue']}")
print(f"p-value of squared standard residuals : {std_res_sq_lb_test['lb_pvalue']}")
```

```
✓ 0.0s                                                                    Python
p-value of standard residuals : 10    0.441596
Name: lb_pvalue, dtype: float64
p-value of squared standard residuals : 10    0.272592
Name: lb_pvalue, dtype: float64
```

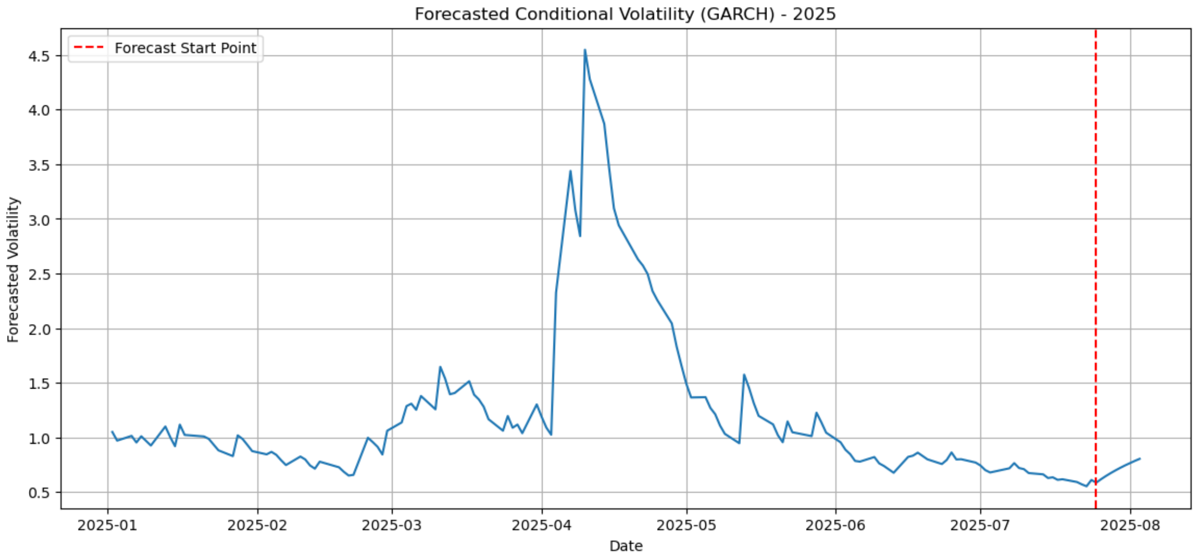
# 6. Volatility Forecast

Forecasted 10 days forward conditional volatility using the fitted GARCH(1,1) model.

The volatility forecast was generated using the *forecast()* method from the *arch* library in Python. which outputs the predicted conditional variance for future time steps. The square root of the conditional variance was taken to convert it into forecasted volatility.




The forecasted volatilities were plotted alongside the historical conditional volatilities (for year 2025) estimated by the model. This combined chart offers a visual representation of how future market uncertainty is expected to evolve based on the most recent patterns in volatility.




# 7. Conclusion and Future Work

The GARCH(1,1) model effectively captured the conditional heteroskedasticity in S&P 500 daily returns. The estimated parameters indicated strong volatility persistence, and diagnostics confirmed a good fit. Future work could involve:

- Extending the model to GARCH(p,q) or EGARCH/TGARCH to capture asymmetries.
- Considering alternative error distributions (e.g., t-distribution) to better account for fat tails.
- Applying the model to other asset classes or to a multivariate setting.