# Forecasting Bitcoin Volatility Using GARCH model

**Authors : Hugo Michel and William Arnaud**



## 1. Methodology

The non-crypto scientific papers we have identified are as follows:

- Francisco Joao Matos Costa, (2017), ***Forecasting volatility using GARCH models*** (link)
- Erginbay Ugurlu et al.,(2017), ***Modeling Volatility in the Stock Markets using GARCH Models: European Emerging Economies and Turkey*** (link)

The scientific paper on crypto that we have identified is as follows:

- Mamoona Zahid et al., (2022), ***Forecasting Bitcoin Volatility Using Hybrid GARCH Models with Machine Learning*** (link)

## 2. Data source

The data we used is the bitcoin price between 2014-11-30 and 2022-03-31. We chose this data because we wanted to study the COVID impacts on the evolution of Bitcoin prices.

To collect the data, we used the CoinMarketCap platform.
From this platform, it's easy to extract the bitcoin price in .csv format by selecting a date range.

## 3. Financial Analysis

We have applied a GARCH model to forecast the volatility of the Bitcoin price.

In finance, volatility is a statistical measure of the dispersion of asset returns over time. It is often calculated as the standard deviation or variance of price returns. Volatility describes the uncertainties surrounding the potential evolution of financial asset prices. It is an essential concept, widely used in risk management, portfolio optimization and other areas. It is also one of the most active areas of research in empirical finance and time series analysis. In general, the higher the volatility, the riskier a financial asset.

The GARCH model deals with stationary (i.e. constant mean around 0 and variance constant over time).

The GARCH model embeds 3 concepts:

- **Autoregressive :** The current value can be expressed as a function of previous values, i.e. they are correlated.
- **Conditional:** This indicates that the variance is based on past errors.
- **Heteroscedasticity:** This implies that the series displays an unusual variance (variable variance).

Firstly, the model is autoregressive in nature. It attempts to estimate volatility at time $t$ on the basis of information known at time $t-1$. Secondly, it estimates volatility as a weighted average of past information.

For a GARCH(1,1) process to be realistic, two conditions must be met:

- Firstly, all parameters $\omega$, $\alpha$, and $\beta$ must be non-negative. This ensures that the variance cannot be negative.
- Secondly, $\alpha+\beta<1$ means that the model estimates are mean reversions of the long-term variance.
- The long-term variance is equal to $\omega/(1-\alpha-\beta)$.

The empirical rules for model parameters is that the larger the parameter $\alpha$, the greater the immediate impact of shocks. Here, shocks are expressed as residuals or prediction errors. If fixed, the larger the $\beta$ is, the longer the duration of the impact, i.e. periods of high or low volatility tend to persist.

**GARCH(1,1) equation**

$$GARCH(1,1) : \sigma_t^2 = \omega + \alpha\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2$$

**GARCH(p,q) equation**

$$\text{GARCH}(p,q) : \sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i\epsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j\sigma_{t-j}^2$$

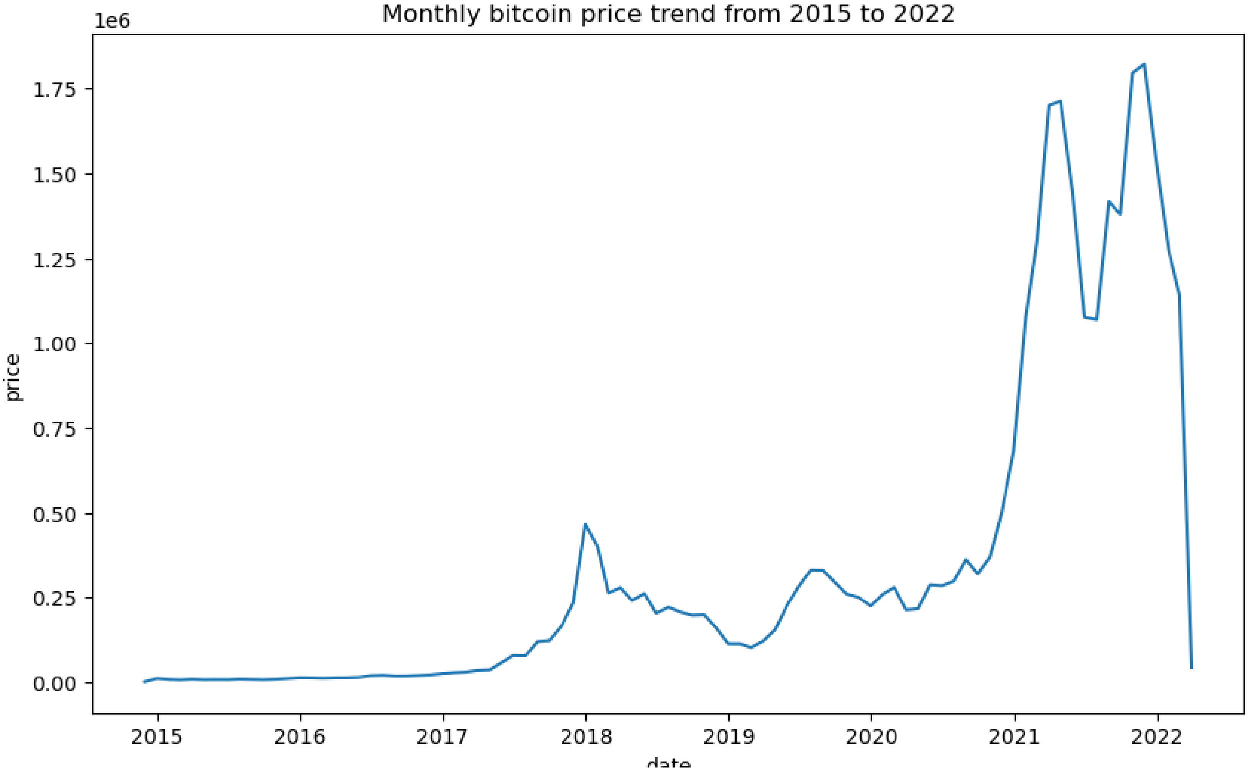
- $\sigma_t^2$ represents the conditional variance at time $t$
- $\omega$ is the constant or intercept term
- $\alpha_i$ are the parameters for the autoregressive conditional variance terms, where $i$ range from $1$ to $p$
- $\epsilon_{t-i}^2$ represents the squared returns at time $t-i$
- $\beta_j$ are the parameters for the conditional variance lag terms, where $j$ ranges from $1$ to $q$
- $\sigma_{t-j}^2$ represents the conditional variance at time $t-j$

Intuitively, the GARCH variance forecast can be interpreted as a weighted average of three different variance forecasts. One is a constant variance that corresponds to the long-term average. The second is the new information that was not available when the previous forecast was made. The third is the forecast made in the previous period. The weights of these three forecasts determine how quickly the variance changes with new information, and how quickly it returns to its long-term average.

In expectation, $X_t^2$ is the marginal variance of the log returns. The relationship between variance and expected value of $var(X) = E[X^2] - E[X]^2$. We know that $E[X_t] = 0$ because expected returns are trivially 0 due to the efficient market hypothesis. Thus, E[X□]$^2$ = 0, and we are left with the marginal variance $\sigma = E[X^2]$, which is constant and does not depend on $t$.

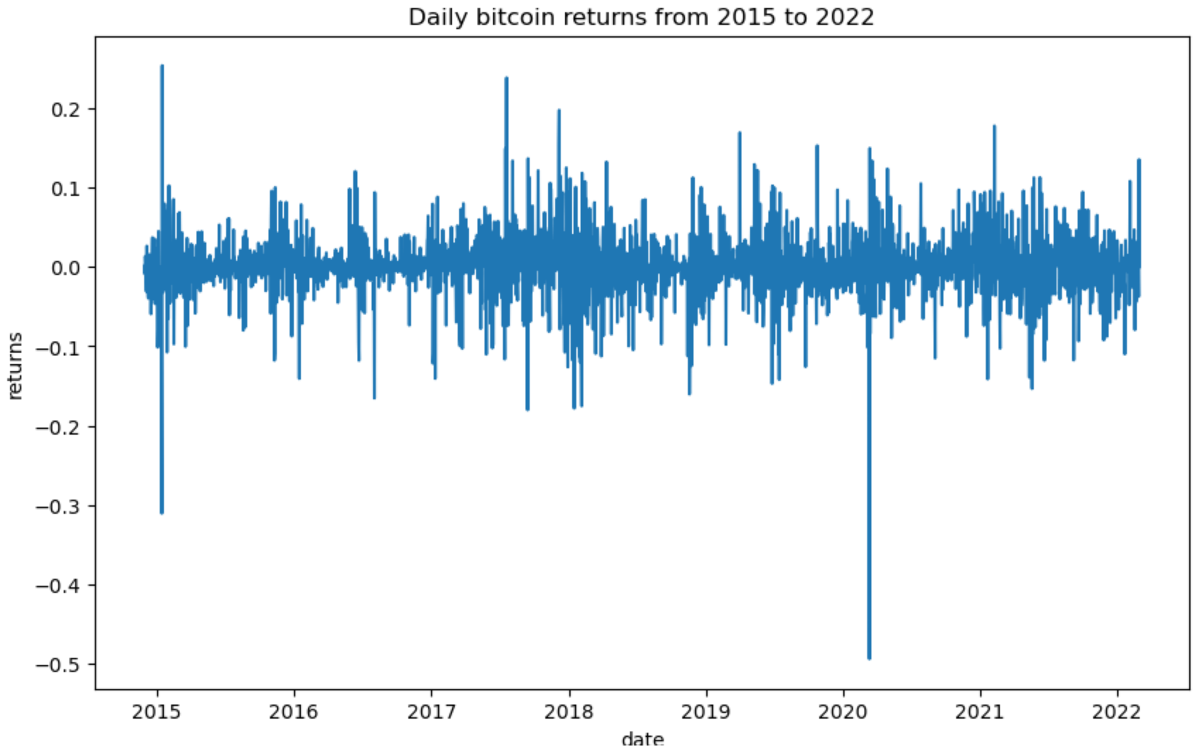
We build on the concept of constant marginal variance to incorporate heteroskedasticity by modeling the volatility at time $t$ ($\sigma_t^2$), which is the conditional variance of the time series and is directly influenced by the squared log return $X_t^2$

**<u>Monthly bitcoin price trend from 2015 to 2022</u>**




**Bitcoin is likely to fall after 2022.**

**<u>Daily bitcoin returns from 2015 to 2022</u>**




**Bitcoin's returns are a stationary series.**

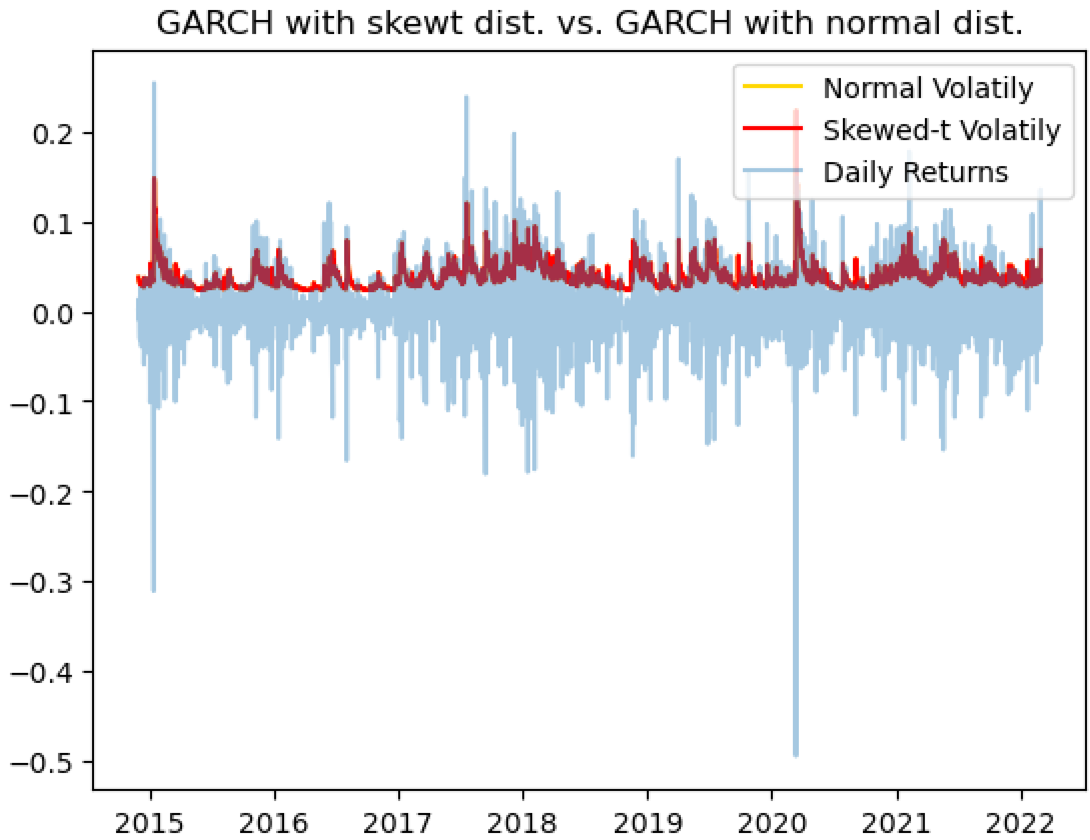
**<u>Find the best distribution parameters for GARCH</u>**

Price-return distribution curves often have wide tails. In other words, the probability of observing very large returns, whether positive or negative, is higher than in the case of a normal distribution.

Returns distribution also tends to exhibit asymmetry. Negative asymmetry means that the left-hand tail of the distribution curve is longer, so that the mass of the distribution is on the right-hand side of the figure. Conversely, positive asymmetry means that the right-hand tail is longer, and the mass of the distribution is concentrated on the left.

Let's select which distribution our model fits the best. To do that we have two possibility

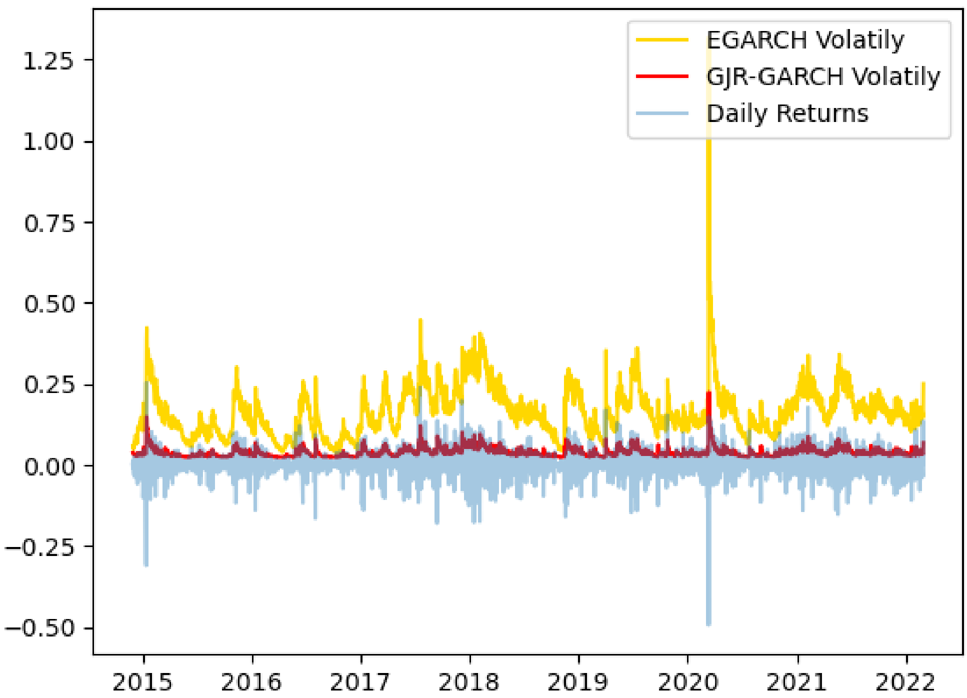
- normal distribution
- skew t-distribution (t-student asymmetric)




As we can see, with the assumption of an asymmetrical t-Student distribution, the GARCH model estimate is more in line with actual observations.

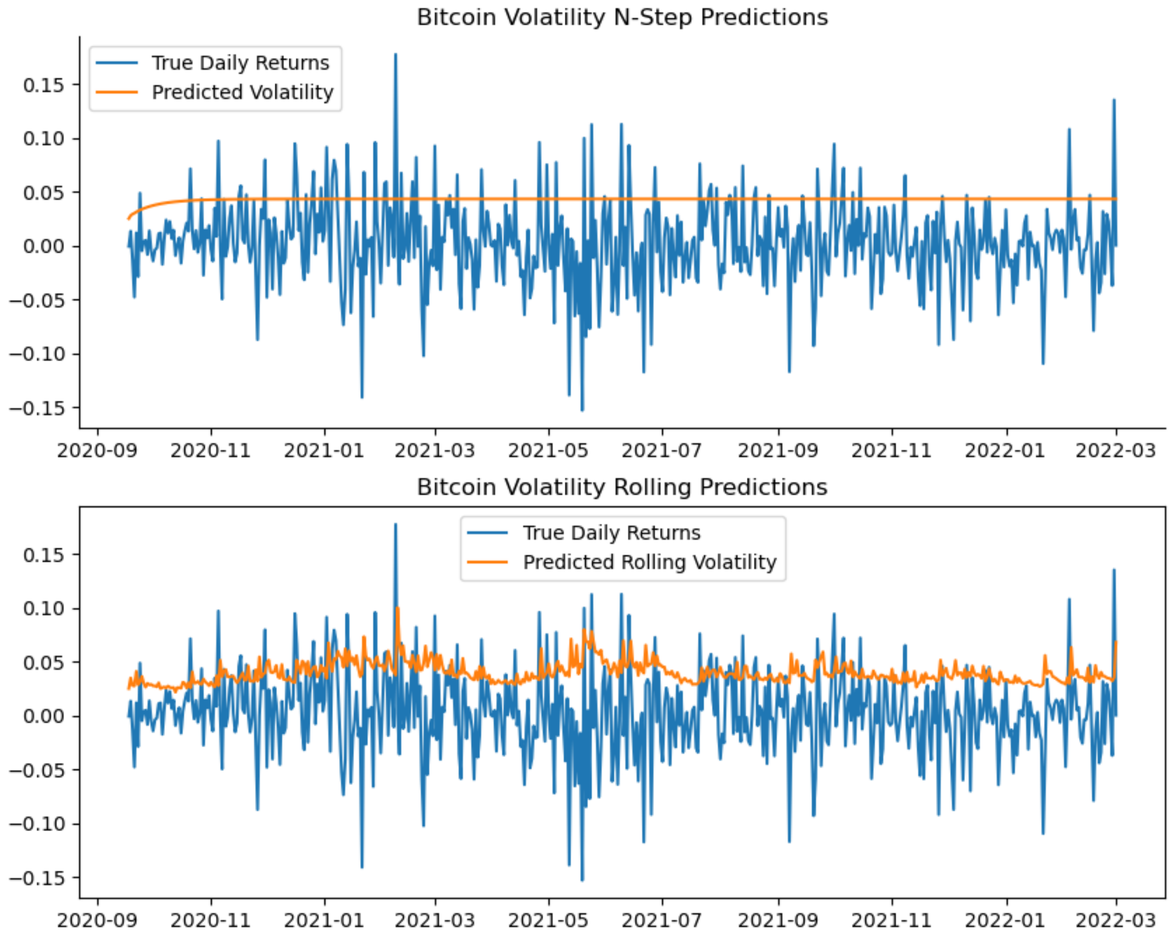
Hence, we will chose the following parameter is ***dist = skewt***

**<u>GJR-GARCH vs. EGARCH</u>**



Overall, the GJR-GARCH and EGARCH models fitted the real data well. Comparatively, the GJR-GARCH model is more conservative in estimating volatility when applied to the bitcoin dataset.

**<u>Predicted Rolling Volatility (on the Covid period)</u>**




See the full analysis on the github repo:
https://github.com/hugo-mi/bitcoin_volatility_forecast/blob/main/ToB_Bitcoin_vol_forecast_HugoMichel.ipynb

# 4. Conclusion

This analysis shows that the bitcoin market is highly volatile. This phenomenon became even more pronounced during the covid pandemic. We therefore conclude that it is difficult to predict the volatility of the bitcoin market. Knowing the fact that it is possible but sometimes complicated, forecasting the volatility of the Bitcoin is useful information for market-participants to better manage the uncertainties and react in advance to the potential evolution of bitcoin prices. Volatility is a useful concept for risk management and better asset allocation in order to optimize the portfolio returns.