Asian Journal of
Science, Technology, Engineering, and Art

e-ISSN: 3025-4507
p-ISSN: 3025-5287
Index: Harvard, Boston, Sydney University, Dimensions, Lens, Scilit, Semantic, Google, etc

https://doi.org/10.58578/AJSTEA.v3i4.5926

Modeling Volatility Using Bayesian GARCH with Student-*t* and Generalized Error Distributions: A Case Study of Bitcoin

Adashu Jacob Daniel & Anule Aondolum Josaphat
Federal University Wukari, Taraba State, Nigeria
adashu@fuwukari.edu.ng; anuleaondolum@gmail.com

Article Info:

Submitted:	Revised:	Accepted:	Published:
Apr 21, 2025	May 18, 2025	May 30, 2025	Jun 4, 2025

Abstract

This study investigates the optimal model for capturing and forecasting volatility in the cryptocurrency market, with a specific focus on Bitcoin (BTC). Various Generalized Autoregressive Conditional Heteroscedasticity (GARCH) models are evaluated to determine the most effective approach for modeling the stylized facts commonly observed in financial time series data. While the Maximum Likelihood Estimation (MLE) method is widely employed for estimating GARCH model parameters, this study introduces a Bayesian framework, utilizing the Metropolis-Hastings algorithm to estimate parameters of the symmetric GARCH(1,1) model. Under this approach, model parameters are treated as random variables with known prior distributions. The analysis is based on 2,000 daily BTC observations from January 2018 to June 2023, obtained from Yahoo Finance. Model selection criteria, including the Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), and Hannan–Quinn Criterion (HQC), identified the EGARCH(1,1) model under the Student-t and Generalized Error Distributions as the most suitable for capturing BTC volatility. Results further indicate the presence of volatility asymmetry and persistence, characteristic of cryptocurrency markets. In terms of predictive performance, the Bayesian GARCH(1,1) model under the

Generalized Error Distribution and the EGARCH(1,1) model under the Student-t distribution exhibited the lowest values for RMSE, MAE, MAPE, and ME, confirming their suitability for future volatility forecasting in the cryptocurrency space.

Keywords: ARCH; GARCH; EGARCH; IGARCH; Bayesian GARCH; Cryptocurrency Volatility; Bitcoin (BTC)

INTRODUCTION

Modeling volatility has become an important issue in financial markets and has drawn the attention and interest of so many researchers and investors over the past three decades. This is because the concept of volatility is applicable to portfolio optimization, risk management and asset pricing according to Arum and Uche (2019). The ability to estimate and forecast volatility an asset leads to a better understanding of current and future risk of holding that asset.

The Economic Times defined volatility as the rate at which the price of security increases or decreases for a given set of returns. According to Kuhe and Agaibe (2018) it measures the uncertainty and risk that play major role in modern financial analysis. Volatility is an important variable in the valuation of derivatives; it allows the construction of stock market indices which serves as tool for measuring the levels of uncertainty among investors in financial markets as studied by Ouael and Abdelati (2018).

One of the main problems often encountered by scholars and researchers in financial time series when modeling financial data is the problem of nonstationarity. Nonstationarity occurs when the rules that generate the time series change over time and in most cases without any prior indication. This makes the modeling process a bit difficult for the traditionally used autoregressive moving average (ARMA) models which assumes stationarity. Also, financial data such as stock prices, exchange rate, interest rate, commodity prices and others, are associated with high volatility. The implication of this is that variance of the error is not constant but heteroscedastic. In order to tackle these problems, several approaches to measure residual volatility were attempted by expert researchers in finance and economics, but the breakthrough came with the advent of the Autoregressive Conditional Heteroscedasticity (ARCH) by Robert Engle in 1982 and later its extensions extensions by Bollerslev (1986).

The most commonly used method of estimation of the parameters of GARCH-type models is the frequentist approach which is based on maximum likelihood estimation (MLE). However, according to Hamilton, (1994) this approach is limited in the sense that it only provides point estimates of parameters and the uncertainty in these estimates is typically based on asymptotic theory where the MLE is assumed to be normally distributed. Also, the frequentist approach can only produce point estimation of forecasts of volatility.

To address these limitations, a Bayesian approach to statistical inference for the GARCH class of models was developed. The Bayesian approach is fast-growing and is believed by some scholars to offer more advantages over the frequentist approach. In Bayesian inference, the parameter θ is treated as a random variable and all inferences about θ were made based on a posterior distribution. In the case of GARCH-type models, due to the recursive nature of the conditional variance, the joint posterior and the full conditional densities are of unknown forms, that is, analytical evaluation of posterior density is difficult. Hence, sampling directly from a posterior density of interest cannot be implemented.

Markov chain Monte Carlo (MCMC) sampling procedures, as commonly used in Bayesian inference, offers an effective way for solving this type of problem. Metropolis (1953) introduced the idea of the MCMC sampling and was later generalized by Hastings (1970). The sampling strategy relies on the construction of a Markov chain with realizations $\psi^{(0)}$, … , $\psi^{(j)}$ … in the parameter space. Under appropriate regularity conditions, asymptotic results guarantee that as j tends to infinity, $\psi^{(j)}$, tends in distribution to a random variable whose density is the posterior. After removing the burn-in of the draws, the values obtained from the chain can then be used to make inference on the joint posterior.

Literature Review

According to Chuan (2005) the ARCH models was the first successful attempt in econometrics to capture volatility clustering in time series data. Until the 1980s, researchers and financial experts used models where volatility was assumed to be constant over time. Engle (1982) used conditional variance to characterize volatility and postulate a dynamic model for it. A Generalized Autoregressive Conditional Heteroscedasticity (GARCH) model was later introduced by Bollerslev in 1986 as an extension of the ARCH models. It makes use of the values of the past squared returns and past variances to model the current variance of a financial data at time t. He developed the GARCH model primarily to overcome the problem of large number of parameters needed in the ARCH process to model volatility.

Bayesian Estimation of GARCH-Type Models

Bayesian inference is a statistical approach where all uncertainties are expressed in terms of probabilities. In the Bayesian approach the parameters are assumed to be random variables which are characterized by a prior density function. The prior distribution captures the beliefs of the researcher about the experiment before any data are collected. Moreover, the prior density can be more or less informative according to the researcher's conviction. because of the recursive nature of the conditional variance, the joint posterior and the full conditional densities are of unrecognizable form, which makes analytical evaluation of posterior density difficult. Hence, sampling directly from a posterior density of interest cannot be implemented.

The concept of Metropolis-Hastings (MH) algorithm was introduced by Metropolis (1953) and generalized by Hastings (1970). This approach constructs a Markov chain by generating draws from a candidate density; the candidate draw is then accepted (or rejected) based on an acceptance probability. If the candidate is accepted, the chain moves to the new value, otherwise the chain stays in the current state. After a *burn-in* period, which is required to make the influence of initial values negligible, draws from the Markov chain are considered as (correlated) draws from the joint posterior distribution of interest. Two variants of the MH approach that are commonly used are;

- the independence chain MH and
- the random-walk MH.

In the former case, candidate draws are generated from an unconditional candidate distribution whereas in the latter draws are generated from a distribution conditional on (and around) the current value of the chain. In both variants the candidate distribution must be tuned to achieve a reasonable acceptance rate and to explore sufficiently the domain of the posterior distribution. This tuning process requires preliminary runs and some knowledge of MCMC techniques from the user. In addition, for interpreted languages such as R (R Development Core Team, 2009) or MATLAB (The MathWorks Inc., 2009), the MH algorithm can be significantly slower than the importance sampling strategy. This is so because the probability of acceptance of the new draw depends on the current state of the Markov chain; hence, the algorithm cannot be vectorized. Finally, the

MH algorithm creates a sequence of correlated draws from the posterior distribution. Therefore, robust techniques must be used to assess the precision of the estimators and more draws are required to achieve the same degree of (numerical) precision as the IS approach.

An interesting strategy has been proposed to automate the MH algorithm and improve its efficiency. This approach is proposed by Nakatsuma (1998 & 2000) and consists of an independence chain MH algorithm where the proposal distributions are constructed from auxiliary ARMA processes on the squared observations. In addition to be fully automatic and more efficient than naive MH approaches, the methodology can be extended to regime-switching GARCH models; see Ardia (2008b, chapter 7) and Ardia (2009a). Note however that the construction of the proposal distributions highly depends on the form of the scedastic function and is not applicable to all GARCH-type models. Moreover, the algorithm requires filtering steps which increases computational burden in a situation where models are highly parametrized. This approach is used by Nakatsuma (1998, 2000), Kaufmann and Scheicher (2006) and other researchers. The algorithm is implemented in the R package bayesGARCH (Ardia, 2008a, 2009b) for the GARCH(1,1) model with Student-*t* disturbances.

METHODOLOGY

Suppose a_t denotes the value of a stock at time t. Then, the return y_t of the stock at time t is

$$y_t = \frac{a_t - a_{t-1}}{a_{t-1}}$$

Autoregressive Conditional Heteroscedasticity, ARCH (P) Model

The basic model as suggested by Robert Engle assumes that

$$\left.\begin{aligned} y_t &= \sigma_t \mathcal{E}_t, \qquad \mathcal{E}_t \sim N(0,1) \\ \sigma_t^2 &= \omega + \alpha_1 y_{t-1}^2 + \alpha_2 y_{t-2}^2 + \alpha_3 y_{t-3}^2 + \dots + \alpha_p y_{t-p}^2 \end{aligned}\right\} \qquad (3.1)$$

Where the parameters $\omega > 0$, $\alpha_i \geq 0$ where $i = 1,2,3,\dots,p$ to ensure that the conditional variance of y_t is positive. The model is capable of capturing volatility clustering in a financial data.

Generalised Autoregressive Conditional Heteroscedasticity (GARCH (P,Q))

The GARCH model is the extension of the ARCH model, it allows σ_t^2 to have an additional autoregressive structure within itself. The model is given by

$$\left.\begin{aligned} y_t &= \sigma_t \varepsilon_t \qquad\qquad e_t \sim N(0,1) \\ \sigma_t^2 &= \omega + \alpha_1 y_{t-1}^2 + \alpha_2 y_{t-2}^2 + \cdots + \alpha_p y_{t-p}^2 + \beta_1 \sigma_{t-1}^2 + \beta_2 \sigma_{t-2}^2 + \beta_q \sigma_{t-q}^2 \end{aligned}\right\} \quad (3.2)$$

$$\sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i y_{t-i}^2 + \sum_{i=1}^{q} \beta_i \, \sigma_{t-i}^2$$

With the conditions $\omega > 0, \alpha_i > 0,\ \beta_j > 0\ \ i = 1,2,\dots,p\ \ j = 1,2,\dots,q$ to guarantee positive conditional variance.

GARCH with Student t-distributed Innovations

When the assumption of normality is violated, it leads to series of problem such as inconsistent parameters estimates and difficulty in providing valid conditional forecasting intervals for y_{T+l} given I_T. Because of this, it is useful to consider a distribution that is leptokurtic according to Herwartz, (2004), hence we considered the student-t distribution.

If ε_t is t-distributed with degrees v of freedom, has a zero mean and a variance of σ_t^2, then its probability density function is given by

$$f(\varepsilon_t \backslash \theta, v) = \frac{v^{\frac{v}{2}} \Gamma\left(\frac{v+1}{2}\right)}{\sqrt{\pi}\,\Gamma\left(\frac{v}{2}\right)\sqrt{\frac{(v-2)\sigma_t^2}{v}}} \left(v + \frac{v y_t^2}{(v-2)\sigma_t^2}\right)^{-\frac{v+1}{2}}$$

where $\Gamma(.)$ is the gamma function given by

$$\Gamma(h) = \int_0^\infty x^{h-1} e^{-x} dx \qquad h > 0$$

The contribution of an observation to the log-likelihood function is given by

$$l(\theta, v) = \ln f(\varepsilon_t \backslash I_{t-1})$$

$$= \ln\left(v^{\frac{v}{2}} \Gamma\left(\frac{v+1}{2}\right)\right) - \ln\left(\sqrt{\pi}\,\Gamma\left(\frac{v}{2}\right)\sqrt{\frac{(v-2)\sigma_t^2}{v}}\right) - \frac{v+1}{2} \ln\left(v + \frac{v y_t^2}{(v-2)\sigma_t^2}\right)$$

Bayes' Theorem

Suppose $y = [y_1, y_2, \dots, y_n]$ is a random sample from the probability density function $f(x;\theta)$, where the parameters θ is a random variable with known probability function $g(\theta)$ and the likelihood as $L(y,\theta)$. Then the Bayesian Rules for obtaining the posterior density is given as

$$L(y,\theta) = \prod f(y_i \backslash \theta)$$

Using the Bayes rule to get the posterior density is as follows;

$$p(\theta \backslash y) \propto L(y \backslash \theta) g(\theta) \qquad (3.20)$$

The expression simply states that updated knowledge combines prior knowledge with the data at hand. The prior would have greater influence on the posterior distribution for small samples than in case of large samples.

The method that is widely used in the classical estimation of the parameters of GARCH-type models is the maximum likelihood estimation. Bayesian approach to these types of models is also fast-growing and is believed to offer advantages over the classical approach. In Bayesian inference, we treat the parameter θ as a random variable and all inferences about θ are made based on a posterior distribution. In the case of GARCH-type models, because of the recursive nature of the conditional variance, the joint posterior and the full conditional densities are of unrecognizable form, which makes analytical evaluation of posterior density difficult. Hence, sampling directly from a posterior density of interest cannot be implemented.

Common

choices for ε_t are Normal and Student-t disturbances. The Student-t specification is particularly useful, since it can provide the excess kurtosis in the conditional distribution that is often found in financial

time series processes (unlike models with Normal

innovations).

The Bayesian approach offers an attractive alternative which enables small sample results, robust estimation, model discrimination, model combination, and probabilistic statements on (possibly nonlinear) functions of the model parameters.

GARCH Model and the Priors

The GARCH (*1*, *1*) model with student-*t* innovation for the process y_t can be written as

$$y_t = \varepsilon_t \left(\left(\frac{v-2}{v}\right) \eta_t \sigma_t^2\right)^{(1/2)} \qquad \text{for } t = 1, \dots, T,$$

where ε_t ~ *iid* N(0, 1), η_t ~ *iid Inverse Gamma* , $\left(\frac{v}{2}, \frac{v}{2}\right)$, and

$$\sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i y_{t-i}^2 + \sum_{i=1}^{q} \beta_i \sigma_{t-i}^2$$

With the conditions $\omega > 0, \alpha_i > 0, (i = 1,2, \dots, p)\ \beta_j > 0\ (j = 1,2, \dots, q)$

According to Ardia and Hoogerheide (2009), the restriction on the degrees of freedom parameter *v* ensures the conditional variance to be finite and the restrictions on the GARCH parameters ω, α_1 and β_1 guarantee its positivity. It is also essential to empathize that only positivity constraints are implemented in the MH algorithm; no stationarity conditions are imposed in the simulation procedure. To obtain the likelihood function we define the following vectors

$y = (y_1, y_2, \dots, y_T)'$, $\alpha = (\omega, \alpha_1, \alpha_2, \dots, \alpha_p)'$, $\beta = (\beta_1, \beta_2, \dots, \beta_q)'$ $\eta = (\eta_1, \eta_2, \dots, \eta_T)'$,

$\Theta = (\alpha, \beta, \nu)$ and a $T \times T$ diagonal matrix $\Sigma = \Sigma(\theta) = diag(\{\eta_t \left(\frac{v-2}{v}\right) \sigma_t^2\}_{t=1}^{T})$, hence we have

$$L(\Theta, \eta \backslash y) \propto (det\Sigma)^{-\frac{1}{2}} \exp\left[-\frac{1}{2} y' \Sigma^{-1} y\right]$$

The Bayesian approach considers (Θ, η) as a random variable which is characterized by a prior density denoted by $p(\Theta, \eta)$.

We propose the following proper priors on the parameters α and β of the preceding models;

$$g(\alpha) \propto N_2(\alpha \backslash \mu_\alpha \Sigma_\alpha) \mathbb{I}_{[\alpha > 0]}$$

$$g(\beta) \propto N_1(\beta \backslash \mu_\beta \Sigma_\beta) \mathbb{I}_{[\beta > 0]}$$

Where $\mu_\bullet$ and $\Sigma_\bullet$ are hyper parameters and $_{[\bullet]}$ is the indicator function which equals unity if the constraint holds, and zero otherwise. N_d is the d-dimensional Normal distribution (d > 1) We also assume that there exist prior independence between α and β.

The prior distribution of vector η_t conditional on v is found by noting that the components η_t are independent and identically distributed from the inverted gamma density, which gives.

$$P(\eta \backslash v) = \left(\frac{v}{2}\right)^{\frac{Tv}{2}} \left[\Gamma\left(\frac{v}{2}\right)\right]^{-T} \prod_{t=1}^{T} \eta_T^{-\frac{v}{2}-1} exp\left\{-\frac{1}{2}\left(\frac{v}{\eta_t}\right)\right\}$$

We agree with Deschamps (2006) in the choice of the prior distribution on the degrees of freedom parameter. The distribution is obtained as translated exponential with parameters $\lambda > 0$ and $\delta \geq 2$.

$$P(\nu) = \lambda \exp\{-\lambda(\nu - \delta)\} I\{\nu > \delta\}\}$$

The joint prior distribution is then formed by assuming prior independence between the parameters,

$$P(\Theta, \eta) = g(\alpha) g(\beta) P(\eta \backslash v) P(\nu)$$

Because of the recursive nature of the conditional variance, the joint posterior and the full conditional densities are of unrecognizable form, which makes analytical evaluation of posterior density difficult. Hence, sampling directly from a posterior density of interest cannot be implemented. We resort to a more reliable Markov Chain Monte Carlo (MCMC) simulation technique to obtain the posterior. Markov chain Monte Carlo (MCMC) sampling procedures, as commonly used in Bayesian inference, offers an effective way for solving this type of problem. One of these known methods is the Metropolis-Hastings (MH) algorithm that was introduced by Metropolis (1953) and generalized by Hastings(1970).

The MH algorithm consists of the following steps.

- Initialize the iteration counter to $j = 1$ and set an initial value $\theta^{[0]}$;
- Move the chain to a new value $\theta^{[*]}$ generated from a proposal (candidate) density $q(\bullet \mid \theta^{[j-1]})$;
- Evaluate the acceptance probability of the move from $0^{[j-1]}$ to $0^{[*]}$ given by:

$$min\left\{\frac{p(\theta^* \backslash y) q(\theta^{j-1} \backslash \theta^*)}{p(\theta^{j-1}) q(\theta^* \backslash \theta^{j-1})}\right\}$$

If the move is accepted, set $\theta^{[j]} = \theta^*$, if not, set $\theta^{[j]} = \theta^{[j-1]}$ so that the chain does not move.

- Change counter from *j* to *j*+1 and go back to step 2 until convergence is reached.

RESULTS

We applied our Bayesian estimation methods to daily closing price of Bitcoin(BTC) peered against United State Dollar(USD) log-returns. The sample period is from January 1, 2018, to May 8, 2023, for a total of 2000 observations. The descriptive statistics of the returns is shown in the Table 1 below.

Table 1: Descriptive statistics for the returns of Bitcoin

N	Max	Min	Mean	Sd
1999	0.1874647	-0.3716954	0.001121109	0.003753884

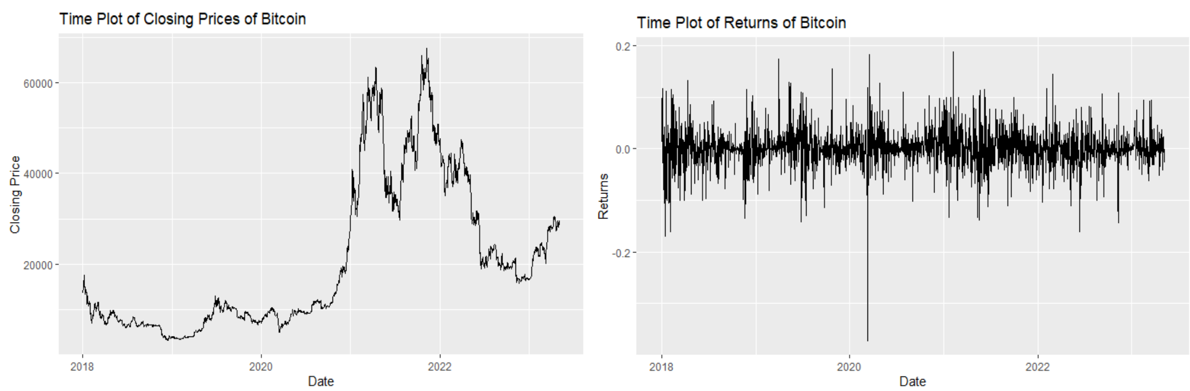

Figure 1: Plots of the daily closing price (left) and log-returns (right) for Bitcoin

In order to justify the use of the application of the GARCH model with student t error term, we first have to prove that the data is non-normally distribution. Hence, we carried out Ljung-Box test for ARCH effect, ADF test for stationarity and Jarque-Bera test of normality. The results presented in Table 2 and Table 3.

Table 2: Box-Ljung test for ARCH effect and ADF test for stationarity.

X-squared	df	p-value	*ADF*	*Lag order*	*p-value*
28984	33	2.2e-16	*-11.842*	*12*	*0.01*

The Ljung-Box test result reveals a clear evidence of arch effect in the return series of Bitcoin, with the p-values less than $\alpha = 0.05$, we reject the null hypothesis. Also, at 5% level of significance, the p-values of ADF test is lower, suggesting that the return series for Bitcoin, is stationary. There we proceed to implement the GARCH model.

Table 3: Jarque Bera Test for Normality for the returns

	X-squared	Df	p-value
Bitcoin	4642.3	2	2.2e-16

The test statistic (X-squared) values is 4642.3 is large, suggesting that there is a significant departure from the normal distribution. In this case, the df is fixed at 2, because the test is based on kurtosis and skewness. Hence, we reject the null hypothesis that the return series is normally distributed.

Proceeding with the parameters estimation of the GARCH(1,1) model we found the results presented in Table 4.

Table 4: GARCH(1,1) parameters

	Estimate	Std. Error
mu	**0.000755**	**0.000532**
omega	**0.000017**	**0.000009**
alpha1	**0.081443**	**0.012800**
beta1	**0.917557**	**0.013933**

After the parameters estimation, we test the normal of the model using Q-Q plot of the residuals. The residual shape in the plot shows that the normality assumption of the residual seems to be violated. This result gave us additional evidence to why we applied the student-t error term in the analysis.

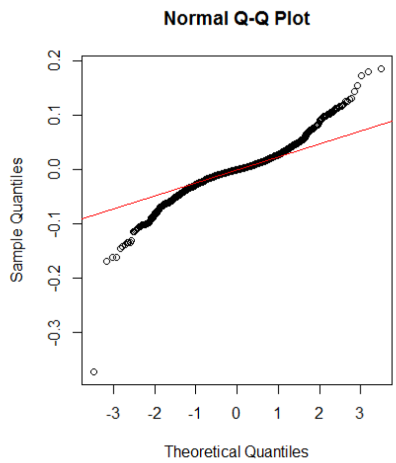

Figure 2: Q-Q Plots of residuals from GARCH (1,1)

Table 5 Posterior mean estimate for the parameters of GARCH (1,1) model

omega	alpha	beta	nu
Min. :2.643e-06	Min. :0.02298	Min. :0.9475	Min. :5.228
1st Qu.:2.643e-06	1st Qu.:0.02439	1st Qu.:0.9475	1st Qu.:5.230
Median :2.643e-06	Median :0.04607	Median :0.9475	Median :5.230
Mean :7.323e-06	Mean :0.03529	Mean :0.9506	Mean :5.232
3rd Qu.:8.872e-06	3rd Qu.:0.04607	3rd Qu.:0.9532	3rd Qu.:5.234
Max. :8.893e-05	Max. :0.04607	Max. :0.9651	Max. :5.234

The table represents the estimated parameters for the Bayesian GARCH model using Metropolis Hastings Algorithm. The results that the parameters are statistically different from zero.

Figure 3, below shows the running mean plot for MH chains. The reseals that for the parameter

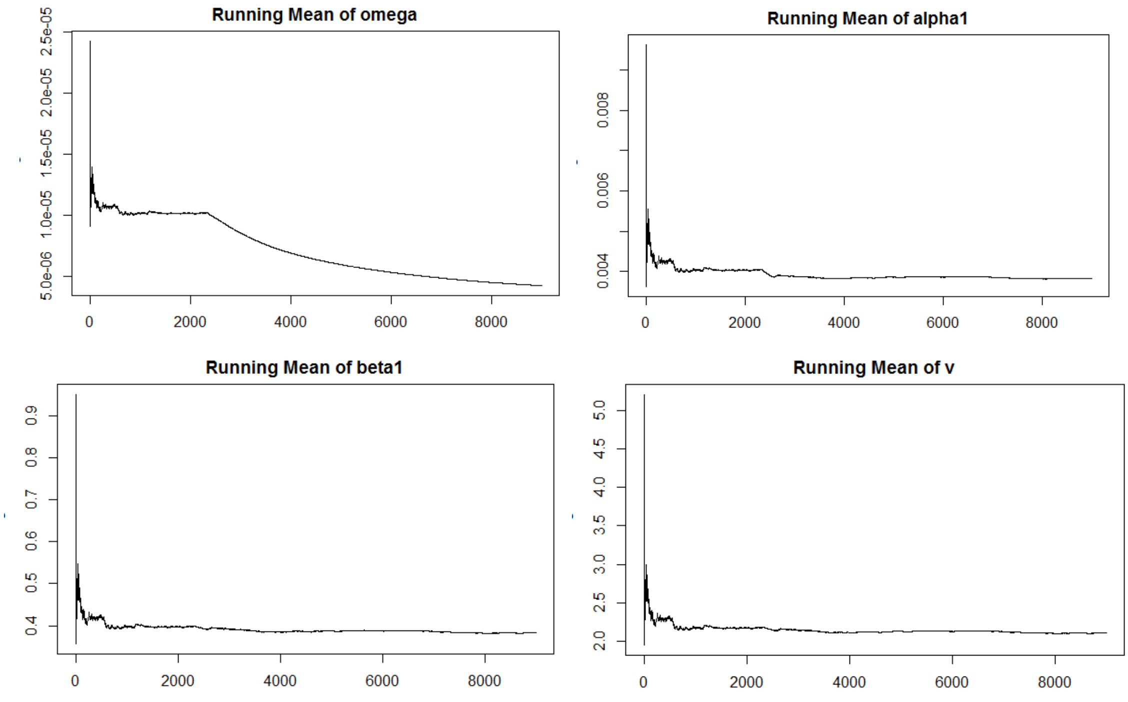

Figure 3. Running mean of BTC for MH-chain of over 10,000 iterations, showing convergence to the stationary point under student error distribution.

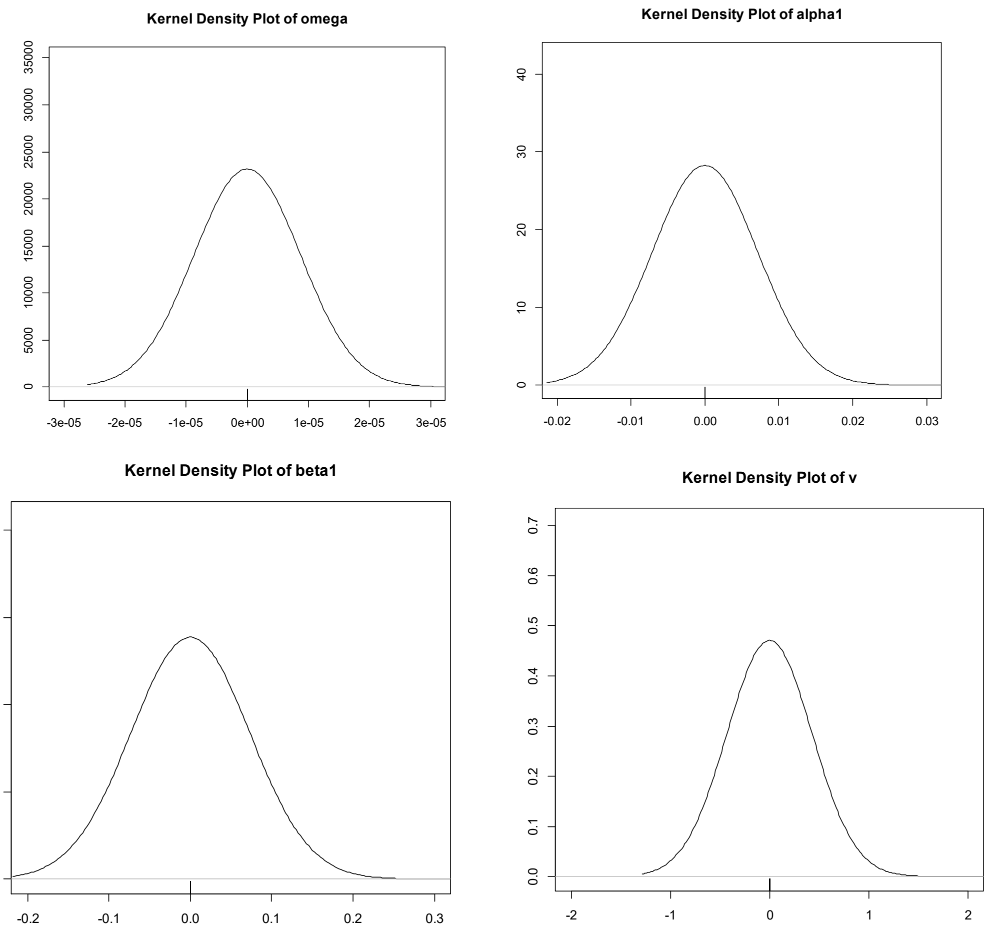

Figure 4. Kernel density function constructed for the parameter of GARCH (1, 1) model with student-t innovation, using the chains of MH for BTC.

Table 6 Results of the forecasting performance Classical GARCH (1,1) and Bayesian GARCH(1,1)

Models	GARCH		Bayesian GARCH	
	Normal	Student-t	Normal	Student-t
RMSE	0.037533	0.037531	0.031585	0.030647
MAE	0.000506	0.025224	0.030484	0.029271
MAPE	1.062524	107.9278	0.812075	0.779745
ME	0.000506	0.000376	-0.030342	-0.029171

CONCLUSION

Because of the nature of returns being subject to time-varying variance it can be cannot be argued that GARCH models are more consistent in estimating the conditional variance than order traditional methods of estimation. Estimation of GARCH models from

Bayesian viewpoint is relatively recent and can be considered very promising due to its advantages when compared to classical technique. The Bayesian GARCH enables small sample results, robust estimation, it combines model forecasts, thus accounting for model risk in the predictions, which is crucial from a risk management perspective. We compared it forecasting performance against the classical GARCH in predicting the prices of Bitcoin in terms of MAE, MAPE, RSME and ME, the result shows that the Bayesian GARCH had a higher predicting accuracy than the classical GARCH.

REFERENCES

Ardia, D. (2008). *Financial Risk Management with Bayesian Estimation of GARCH Models: Theory and Applications*. Springer-Verlag.

Ardia, D., & Hoogerheide, L. F. (2010). Bayesian Estimation of the GARCH(1,1) Model with Student-t Innovations. *The R Journal*, *2*(2), 41–47.

Attah, A. A., & Kuhe, D. A. (2018). Modelling Volatility Dynamics of Foreign Exchange Rates in Nigeria: Application of Univariate GARCH Models. *Journal of Scientific and Engineering Research*, *5*(5)

Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics*, *31*(3), 307–327.

Chuan, C. H., & Chew, E. (2005, July 6–8). Polyphonic Audio Key Finding Using the Spiral Array CEG Algorithm. In *Proceedings of the IEEE International Conference on Multimedia and Expo*.

Deschamps, P. J. (2006). A Flexible Prior Distribution for Markov Switching Autoregressions with Student-t Errors. *Journal of Econometrics*, *133*(1), 153–190.

El Jebari, O., & Hakmaoui, A. (n.d.). GARCH Family Models vs EWMA: Which Is the Best Model to Forecast Volatility of the Moroccan Stock Exchange Market?

Engle, R. F., & Bollerslev, T. (1986). Modelling the Persistence of Conditional Variances. *Econometric Reviews*, *5*, 1–50.

Geweke, J. F. (1993). Bayesian Treatment of the Independent Student-t Linear Model. *Journal of Applied Econometrics*, *8*(S1), S19–S40.

Hamilton, J. D. (1994). *Time Series Analysis*. Princeton University Press.

Hastings, W. K. (1970). Monte Carlo Sampling Methods Using Markov Chains and Their Applications. *Biometrika*, *57*(1), 97–109.

Kingsley, A., & Peter, U. (2019). Volatility Modelling Using ARCH and GARCH Models (A Case Study of Nigerian Exchange Rate). *International Journal of Mathematics Trends and Technology (IJMTT)*, *65*(4)

Kaufmann, S., & Scheicher, M. (2006). A Switching ARCH Model for the German DAX Index. *Studies in Nonlinear Dynamics and Econometrics*, *10*(4), Article 3. http://www.bepress.com/snde/vol10/iss4/art3/

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. (1953). Equation of State Calculations by Fast Computing Machines. *Journal of Chemical Physics*, *21*, 1087–1092.

Nakatsuma, T. (1998). A Markov-Chain Sampling Algorithm for GARCH Models. *Studies in Nonlinear Dynamics and Econometrics*, *3*(2), 107–117.

Nakatsuma, T. (2000). Bayesian Analysis of ARMA-GARCH Models: A Markov Chain Sampling Approach. *Journal of Econometrics*, *95*(1), 57–69. https://doi.org/10.1016/S0304-4076(99)00029-9