Global and Stochastic Analysis **DOI: 10.64837/GSA.13.3.3**
Vol. 13 No. 3 (2026)

Received: 25th March 2026 *Revised: 22nd May 2026* *Accepted: 11th June 2026*

FORECASTING BITCOIN VOLATILITY USING A WEIGHTED SEMIPARAMETRIC EGARCH MODEL INCORPORATING CRUDE OIL MARKET VOLATILITY

ASHWINI KUMARI* AND NAVEENA

Abstract. Bitcoin price movements are highly volatile and nonlinear over time. Traditional GARCH models are widely used for volatility modeling, but they may not adequately represent nonlinear effects of exogenous variables. In this study, we develop a weighted semiparametric EGARCH model for forecasting Bitcoin volatility by combining the parametric conditional variance from an EGARCH model with a nonlinear crude oil market volatility component estimated using the Nadaraya–Watson kernel estimator. The empirical analysis uses daily Bitcoin and crude oil price data from 2017 to 2026. The forecasting performance is compared with GARCH(1,1), EGARCH(1,1), and semiparametric EGARCH models incorporating crude oil market volatility. The root mean square error, mean absolute error, quasi-likelihood loss, and Diebold–Mariano test are computed to assess the out-of-sample forecast accuracy for various training-testing split designs. Empirical results indicate that the proposed weighted semiparametric EGARCH model achieves higher forecasting accuracy and robustness for forecasting Bitcoin volatility.

1. Introduction

The financial markets are influenced by various endogenous and exogenous factors, so volatility modelling is an important research area in financial econometrics. It is particularly crucial for cryptocurrencies, which are unpredictable, volatile, and uncertain in trading actions. Bitcoin is the most widely studied digital asset and is characterized by heavy tails, volatility clustering, and nonlinear dynamics, which makes Bitcoin volatility forecasting a challenging task [10].

Volatility forecasting is an important factor in portfolio allocation, derivative pricing, risk assessment, and investment decision- making [15]. The cryptocurrency markets are highly nonlinear and dynamic, which makes it difficult to apply traditional volatility forecasting models. It highlights the importance of a flexible and accurate volatility forecasting model that captures the effect of external and internal market factors.

The widely used model for modelling time-varying volatility is the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model, proposed by Bollerslev [2]. Over the years, several variants of the GARCH model have been developed. Of these, the Exponential GARCH (EGARCH) model developed by

2000 *Mathematics Subject Classification.* 62M10;62G08;62P20.
Key words and phrases. Bitcoin Volatility, semiparametric modeling, kernel regression, crude oil volatility, volatility forecasting, Diebold–Mariano test.

Nelson [12] effectively captures the volatility clustering and the asymmetric market reaction to positive and negative shocks. But the parametric models do not adequately capture the volatility of the cryptocurrency market, which is influenced by several exogenous factors.

Crude oil is an important commodity in the global economy. Crude oil prices are fluctuating due to changes in macroeconomic factors, geopolitical events and market uncertainties [7]. The price of crude oil may give a clue to investors' sentiment and global economic uncertainty. This will affect the cryptocurrency markets via cross-market spillover effects. Hence, the volatility of the crude oil market is considered as an important exogenous factor in the Bitcoin volatility forecasting. The integration of the oil market information into the volatility models could enhance the predictive powers of these models and facilitate the understanding of interactions between the commodity and cryptocurrency markets.

Most of the existing research in the field of Bitcoin volatility forecasting has mainly focused on classical GARCH, EGARCH, TGARCH, and related parametric extensions. These models adequately capture many stylized facts of returns. The classical GARCH models are based on past returns and conditional volatility. Therefore, they might not be able to capture the nonlinear effects of several exogenous factors

Previous research also examined the connection between cryptocurrencies and commodity markets such as oil and gold. These findings highlight the significant return and volatility spillovers, hedging properties, and dynamic interdependence among these asset classes [8, 11]. However, limited attention has been given to incorporating external market information into volatility models. The combination of EGARCH models with nonparametric approaches to capture the nonlinear effects of the crude oil market in cryptocurrency volatility remains underexplored.

To address these limitations, we develop a weighted semiparametric EGARCH model for forecasting Bitcoin volatility. First, the EGARCH framework is adopted to model Bitcoin volatility, capture volatility persistence, and volatility asymmetry. Next, the Nadarya-Watson kernel estimator is used to estimate the nonlinear effect of the crude oil market volatility. Finally, a weighted semiparametric volatility forecasting model is constructed by combining the parametric EGARCH volatility component and the kernel-estimated oil volatility component using an optimal weighting approach. The proposed framework combines the structural interpretability of the EGARCH model with the flexibility of nonparametric estimation.

The main objectives of this study are:

(1) To model Bitcoin volatility using the EGARCH framework.
(2) To estimate the nonlinear contribution of oil market volatility using the Nadaraya–Watson kernel estimator.
(3) To develop a weighted semiparametric volatility forecasting model by assigning optimal weights to the EGARCH volatility component and the kernel-estimated oil volatility component.
(4) To compare the out-of-sample forecasting performance of the benchmark models with the proposed weighted semiparametric EGARCH model.

The contributions of this study to the existing literature are as follows:

(1) The study develops a semiparametric volatility forecasting approach that combines the EGARCH model with a Nadaraya–Watson kernel estimator.
(2) Oil market volatility is incorporated as an exogenous factor in Bitcoin volatility forecasting, enabling the assessment of volatility transmission across commodity and cryptocurrency markets.
(3) Finally, we build a weighted semiparametric forecasting model by combining the parametric EGARCH volatility component and the kernel-estimated oil volatility component using an optimal weighting scheme. This proposed model increases the flexibility and accuracy of forecasting volatility.

This paper is organized as follows. Section 2 provides a review of related literature. The data and methodology are described in Section 3. Section 4 highlights the results and discussion. Finally, Section 5 concludes the paper and suggests directions for future research.

2. Literature review

2.1. Bitcoin Volatility Modeling. The modeling and forecasting of Bitcoin volatility has become a significant research area in cryptocurrency markets. To capture volatility that varies across time, Engle [5] introduced the ARCH model, and Bollerslev [2] extended this to the GARCH model which allows the past conditional variances to influence the current volatility. The limitation of the Standard GARCH model is that its inability to capture the asymmetric responses to positive and negative shocks. To overcome the limitations of the standard GARCH model, Nelson [12] proposed the EGARCH model. The EGARCH model is a flexible model to capture the leverage effect and asymmetric volatility dynamics which are commonly observed in financial markets.

Many researchers are motivated to investigate the best volatility models for Bitcoin returns due to the rapid expansion of cryptocurrency markets. Katsiampa [10] demonstrated that the AR-CGARCH model can effectively capture short-term and long-term components of Bitcoin volatility. However, the applicability of purely parametric volatility models is challenged by the nonlinear and unpredictable nature of cryptocurrency markets. This phenomenon has led to recent studies on using machine learning methods for cryptocurrency volatility forecasting. Wang et al. [19] demonstrated that the machine learning models, such as Random Forest and LSTM networks, significantly outperform traditional GARCH models in predicting cryptocurrency volatility. The findings of their study highlight the importance of internal market determinants in volatility prediction and show that external determinants provide additional predictive information. Similarly, Huang et al. [9] reported that LSTM and CNN–LSTM models exhibited superior performance in predicting Bitcoin volatility in comparison to classical GARCH-based models, demonstrating the power of machine learning techniques for predicting complex nonlinear market dynamics.

Machine learning models tend to be more accurate in prediction, but they generally suffer from limited interpretability and require a large dataset to train the model. On the other hand, GARCH-type models provide an economically meaningful interpretation of volatility persistence and asymmetric effects. This has encouraged the emergence of hybrid and semiparametric volatility model frameworks that combine the interpretability of traditional volatility models with the flexibility of nonlinear methods [4, 14, 18, 20].

2.2. Oil Market and Volatility Transmission. Only a limited number of studies have examined the relationship between the volatility transmission between the cryptocurrency and crude oil markets. Hamilton [7] found that oil price shocks have significant macroeconomic impacts and often precede economic recessions. Subsequently, Sadorsky [16] demonstrated that the oil price volatility substantially affects stock market returns and shows asymmetric effects. The results indicate that crude oil market volatility can be used as an indicator of uncertainty in macroeconomic and financial markets.

There is also recent evidence indicating that crude oil markets impact cryptocurrency markets through volatility spillovers and financial uncertainty channels. Bouazizi et al. [3] studied the dynamic relationships between cryptocurrency and commodity markets through multivariate GARCH models and found that Bitcoin, crude oil, and commodity markets have significant spillovers in returns and volatility. According to their results, there are significant cross-market linkages, and they highlight the potential of oil market information to be used in cryptocurrency risk assessment and volatility modelling. In the same way, Ali et al. [1] analyzed the impact of return and volatility spillovers between cryptocurrencies, crude oil, and stock markets. Their findings show that these markets exhibit high volatility transmission effects, as well as asymmetric effects, but no return spillovers were found. The results indicate that crude oil market volatility contains useful information about financial market uncertainty and could be used as a significant external factor in forecasting cryptocurrency volatility. But most of the existing studies consider only multivariate relations and connectedness analysis rather than using oil market information directly in univariate Bitcoin volatility forecasting models [1, 3]. Hence, the use of crude oil market volatility as an external variable to Bitcoin volatility models is still relatively under-explored.

2.3. Advanced and Hybrid Volatility Models. A semi-parametric volatility model combines parametric and nonparametric components, thereby achieving improved flexibility to capture nonlinear financial dynamics, extending the traditional GARCH model [6, 18]. These models remove restrictive functional assumptions and allow explanatory variables to affect volatility through unspecified nonlinear relationships. However, Chen and Gerlach [4] introduced a semiparametric GARCH model with Bayesian model averaging and MCMC methods to show increased flexibility in modelling the non-linear volatility dynamics. They find that the complex volatility behaviour, which is not adequately represented by conventional parametric specifications, can be captured with the use of semiparametric approaches.

Currently, hybrid volatility forecasting models use GARCH-type models combined with machine learning models. Zahid et al. [20] proposed hybrid GARCH models by combining machine learning techniques for forecasting Bitcoin volatility, and found that their models outperformed traditional models. Likewise, Pérez-Hernández et al. [14] designed a hybrid volatility forecasting model by combining an artificial neural network with a GARCH-type model and achieved more accurate volatility forecasts than conventional volatility models.

The results of these studies suggest that combining traditional volatility models along with flexible nonlinear models may provide more accurate forecasts and a more comprehensive description of the complex dynamics of financial markets. But many hybrid and machine learning models are black-box systems with little economic interpretation. However, the combination of interpretability of the GARCH-type structure and the flexibility of the nonparametric estimation technique makes semiparametric volatility models more effective. Nevertheless, empirical evidence on the use of EGARCH models with external market factors remains limited, especially for forecasting Bitcoin volatility.

2.4. Research Gap. The volatility of Bitcoin has been explored in previous research with GARCH-type models, oil market volatility spillovers, and hybrid forecasting methods separately. However, limited research has integrated EGARCH-based asymmetric volatility dynamics, kernel-estimated exogenous effects, and weighted hybrid structures within a unified semiparametric framework. Moreover, the nonlinear effect of crude oil market volatility as an exogenous variable in Bitcoin volatility models has received limited attention in the existing literature.

The existing literature either uses traditional GARCH-type models, which may fail to capture nonlinear relationships, or machine-learning methods that often lack economic interpretability. Thus, there is a need for a modelling framework that can include asymmetric volatility behaviour, nonlinear external effects, and better forecasting performance.

We developed a weighted semiparametric EGARCH Model. It integrates a nonparametrically estimated crude oil market volatility component with the EGARCH based volatility dynamics. An optimal weighting scheme is used to improve Bitcoin volatility forecasting.

3. Data and Methodology

3.1. Data Description. The data used in our study for volatility modelling and forecasting are the daily closing prices of Bitcoin and crude oil. The BTC-USD series represents Bitcoin prices, and the West Texas Intermediate (WTI) crude oil futures (CL=F) series represents crude oil prices. The daily closing prices for both series are downloaded from Yahoo Finance and analyzed using the R programming environment for the period January 1, 2017, to May 31, 2026.

Observations with missing values are removed, and the Bitcoin and crude oil price series are merged with common trading dates, to ensure consistency and appropriate alignment. The logarithmic returns are used to stabilize variance and reduce non-stationarity in the price series. Daily logarithmic returns are computed as

$$r_t = \log\left(\frac{P_t}{P_{t-1}}\right) \tag{3.1}$$

where P_t denotes the closing price at time t. The return series are used for volatility modelling and forecasting analysis

For empirical analysis, the dataset is primarily divided into 80% training observations and 20% testing observations for model estimation and forecast evaluation. In addition, a 75:25 train–test split is considered as a robustness analysis to assess the stability of the forecasting results under alternative sample partitions.

3.2. Benchmark Volatility Models.

3.2.1. *GARCH Model.* The GARCH(1,1) model proposed by Bollerslev [2] assumes that current volatility depends on both past shocks and past volatility. It is widely used to model the time-varying volatility in financial time series. The mean equation is specified as

$$r_t = \mu + \sigma_t \varepsilon_t \tag{3.2}$$

The conditional variance equation is given by

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \tag{3.3}$$

where r_t denotes the logarithmic return at time t. ε_t is the error term, and σ_t^2 represents the conditional variance.The parameter α_0 is a constant term, α_1 measures the effect of past shocks, and $beta_1$ captures volatility persistence.

3.2.2. *Exponential GARCH Model.* The Exponential GARCH (EGARCH) model introduced by Nelson (1991) extends the GARCH model to capture asymmetric volatility responses to positive and negative market shocks. Unlike the standard GARCH model, the EGARCH specification models the logarithm of conditional variance. The EGARCH(1,1) model is specified as

$$\ln(\sigma_t^2) = \alpha_0 + \alpha_1 \left| \frac{\varepsilon_{t-1}}{\sigma_{t-1}} \right| + \gamma_1 \frac{\varepsilon_{t-1}}{\sigma_{t-1}} + \beta_1 \ln(\sigma_{t-1}^2) \tag{3.4}$$

where $\gamma_1$captures the asymmetric effect of shocks on volatility.A negative value of γ_1 indicates that negative shocks have a greater impact on future volatility than positive shocks of the same magnitude. Conditional variance in the EGARCH model depends on past standardized shocks, the asymmetry effect of shocks, and the persistence of volatility.

3.3. Semiparametric EGARCH Model with Oil Market Volatility. In this study, the EGARCH model is used to capture volatility clustering, persistence, and asymmetric responses in Bitcoin return volatility. The estimated conditional variance obtained from the EGARCH model represents the parametric component of volatility.Since fluctuations in the crude oil market volatility may influence Bitcoin volatility through information transmission mechanisms and changes in investor sentiment, crude oil volatility is incorporated as an exogenous factor. To capture possible nonlinear relationships between oil market volatility and Bitcoin volatility, a semiparametric approach is adopted.

The proposed semiparametric volatility model is given by

$$\hat{\sigma}_{SP,t}^2 = \hat{\sigma}_{EGARCH,t}^2 + \hat{g}(x_t) \tag{3.5}$$

where $\hat{\sigma}_{EGARCH,t}^2$ denotes the conditional variance from the fitted EGARCH model, and $\hat{g}(x_t)$ represents the estimated nonlinear effect of oil market volatility obtained using the Nadaraya–Watson kernel estimator. The function $\hat{g}(x_t)$ captures the nonlinear oil-related component of Bitcoin volatility that is not explained by the EGARCH model.

3.4. Nadaraya–Watson Kernel Estimation. The Nadaraya–Watson kernel estimator is employed to estimate the nonlinear effect of crude oil market volatility on the component of Bitcoin volatility that is not explained by the EGARCH model. This nonparametric approach allows the relationship between oil market volatility and Bitcoin volatility to be modeled flexibly without imposing a specific functional form.

The nonlinear oil volatility component is estimated as

$$\hat{g}(x_t) = \frac{\sum_{i=1}^{n} K_h(x_t - x_i) S_i}{\sum_{i=1}^{n} K_h(x_t - x_i)} \tag{3.6}$$

where $K_h(\cdot)$ is the kernel function with bandwidth h, x_i represents crude oil market volatility, and S_i denotes the component of realized Bitcoin volatility that remains after accounting for the EGARCH conditional variance, defined as

$$S_i = RV_i - \hat{\sigma}^2_{EGARCH,i} \tag{3.7}$$

where RV_i denotes the realized variance proxy of Bitcoin measured by squared returns. The estimator assigns higher weights to observations that are closer to x_t, resulting in a locally weighted estimate of the nonlinear effect of oil market volatility on Bitcoin volatility.

3.5. Proposed weighted Semi-parametric EGARCH Model. The semiparametric volatility forecast is first obtained by combining the conditional variance from the EGARCH model with the nonlinear oil volatility component estimated through the Nadaraya–Watson kernel estimator. To further improve forecasting performance, the EGARCH volatility component and the nonlinear oil volatility component are combined using optimal weights determined by minimizing the mean squared error over the training sample. Using the semiparametric forecast defined in equation (3.5), the proposed weighted semiparametric EGARCH model is specified as

$$\hat{\sigma}^2_t = w_1 \hat{\sigma}^2_{EGARCH,t} + w_2 \hat{g}(x_t) \tag{3.8}$$

subject to

$$w_1 + w_2 = 1 \tag{3.9}$$

where w_1 and w_2 denote the weights assigned to the EGARCH volatility component and the kernel-based oil volatility component, respectively. The optimal weights are selected by minimizing the mean squared error between realized volatility and the combined forecast over the training period.

It is important to note that, in the present study, the nonlinear oil volatility component is estimated separately using the Nadaraya–Watson kernel estimator and then combined with the EGARCH conditional variance for forecasting. Therefore, the proposed framework is implemented as a hybrid semiparametric forecasting model rather than as a fully integrated semiparametric EGARCH specification in which the nonparametric component enters directly into the EGARCH variance equation.

3.6. Forecast Evaluation Measures. The forecasting performance of competing models is evaluated using three commonly used statistical error measures, such as RMSE, MAE, and QLIKE, which are widely used in volatility forecasting studies [17].

3.6.1. *Root Mean Square Error (RMSE).* RMSE measures the square root of the average squared difference between the realized variance proxy and the forecasted volatility. It is given by

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{t=1}^{n} e_t^2}, \qquad e_t = \sigma_t^2 - \hat{\sigma}_t^2 \tag{3.10}$$

where σ_t^2 and $\hat{\sigma}_t^2$ denote the realized variance proxy and forecasted volatility, respectively.

3.6.2. *Mean Absolute Error (MAE).* MAE measures the average absolute difference between realized variance proxy and forecasted volatility. It is defined as:

$$\text{MAE} = \frac{1}{n}\sum_{t=1}^{n} |e_t|, \qquad e_t = \sigma_t^2 - \hat{\sigma}_t^2 \tag{3.11}$$

where σ_t^2 and $\hat{\sigma}_t^2$ denote the realized variance proxy and forecasted volatility, respectively

3.7. Quasi-Likelihood (QLIKE). The QLIKE loss function is used to evaluate the accuracy of volatility forecasts because it is robust to outliers and measurement errors in realized volatility and is widely used in volatility forecasting studies [13, 17]. The QLIKE loss is defined as

$$\text{QLIKE} = \frac{1}{n}\sum_{t=1}^{n}\left(\log(\hat{\sigma}_t^2) + \frac{\sigma_t^2}{\hat{\sigma}_t^2}\right) \tag{3.12}$$

where σ_t^2 represents the realized variance proxy computed from squared returns and $\hat{\sigma}_t^2$ denotes the predicted conditional variance. Lower values of QLIKE indicate better forecasting performance.

3.8. Diebold-Mariano Test. The Diebold–Mariano (DM) test is employed to compare the predictive accuracy of competing volatility forecasting models using the QLIKE loss function. The null hypothesis states that the competing models have equal predictive accuracy. Rejection of the null hypothesis indicates a statistically significant difference in forecasting performance.

In this study, the DM test is conducted on the loss differentials derived from the QLIKE loss values of the competing models. A p-value less than 0.05 indicates that the difference in forecasting performance between the competing models is statistically significant.

4. Results and Discussion

4.1. Preliminary Data Analysis. Bitcoin returns are treated as the primary series and crude oil returns are included as an exogenous factor in modelling volatility. The data set is split into training subset and testing subset. Approximately 80% of the observations are used as in-sample data (training data) for model estimation. The remaining 20% are used in out-of-sample (testing) data for forecast evaluation and comparison of the competing models.

The descriptive statistics of Bitcoin and crude oil returns for the training and testing samples are reported in Table 1. The Bitcoin returns have a higher standard deviation than the crude oil returns in both samples, suggesting that the Bitcoin market is more volatile and has higher market risk. The values of skewness and kurtosis suggest the population is non-normal. The training sample shows extremely high kurtosis, particularly for crude oil returns. This indicates the presence of extreme returns and heavy tails. Similarly, the leptokurtic nature of cryptocurrency returns is supported by the observation that Bitcoin returns are leptokurtic in both the training and testing samples.

Based on a comparison of the training and testing samples, return volatility decreased during the testing period. However, the high kurtosis observed in both samples indicates that extreme price fluctuations remain a significant characteristic of both markets. These findings are consistent with the stylized facts of financial return series, including volatility clustering, asymmetry, and leptokurtosis. Such characteristics strongly support the use of GARCH-type models for capturing time-varying conditional volatility.

Table 1. Descriptive statistics of Bitcoin and crude oil returns

Index	**Train Data**		**Test Data**	
	btc_ret	**oil_ret**	**btc_ret**	**oil_ret**
Min.	-0.464730	-0.282206	-0.152334	-0.179298
1st Qu.	-0.016582	-0.011643	-0.016213	-0.014139
Median	0.001905	0.002257	0.000463	0.000731
Mean	0.002141	0.000557	0.000501	0.000129
Standard deviation	0.045574	0.029535	0.029507	0.027701
Skewness	-0.635494	0.029709	0.104385	-0.872048
Kurtosis	12.004070	30.728640	7.180081	9.678207
3rd Qu.	0.022485	0.013607	0.015551	0.014135
Max.	0.225119	0.319635	0.147392	0.115187

Figure 1 presents the daily closing price series of Bitcoin and WTI crude oil over the study period. Both series show substantial fluctuations, highlighting the dynamic nature of cryptocurrency and commodity markets. These price movements suggest the presence of time-varying volatility and motivate the analysis of return volatility.

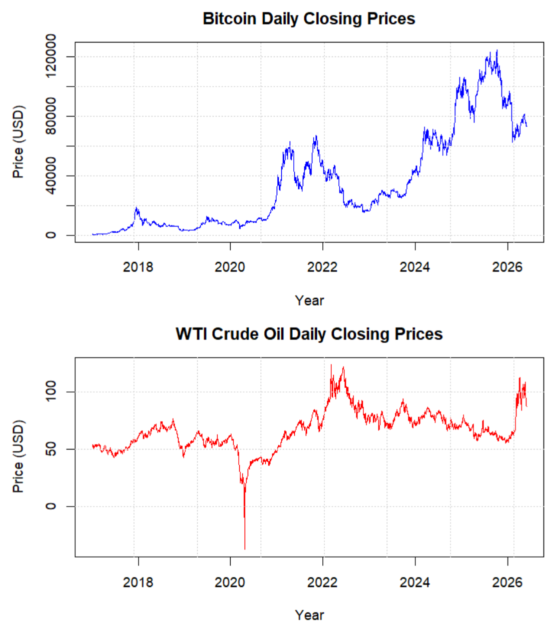

FIGURE 1. Daily closing prices of Bitcoin and WTI crude oil

Figure 2 presents the time series plots of Bitcoin and crude oil returns for the training sample. The plots show periods of high and low volatility and support the presence of volatility clustering, which is consistent with the stylized properties of financial returns and supports the use of GARCH-type models.

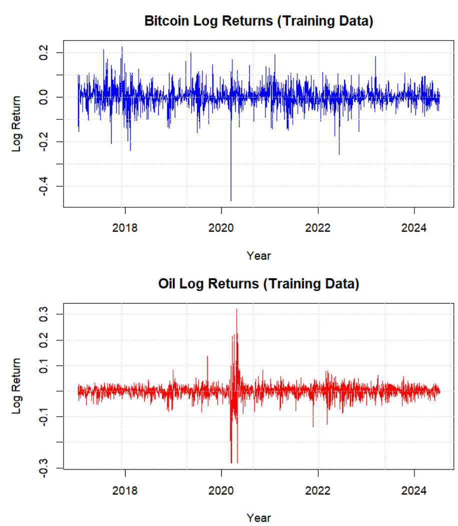

FIGURE 2. Log returns of Bitcoin and oil prices (training sample)

Also, Figure 3 shows that Bitcoin and crude oil returns have weak autocorrelation, as most spikes lie within the confidence bands. This suggests that past returns do not strongly influence current returns, although volatility may still vary over time.

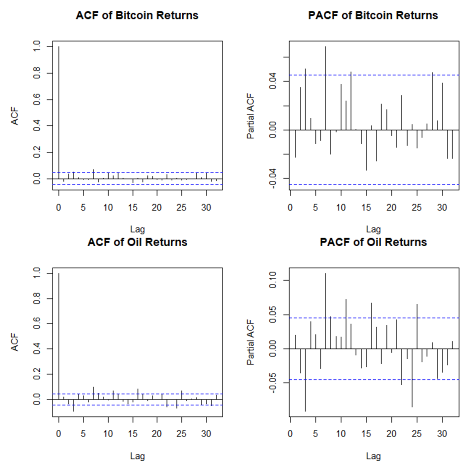

FIGURE 3. ACF and PACF plots of Bitcoin and crude oil returns

To examine time-varying volatility, the squared returns of Bitcoin and crude oil are plotted in Figure 4. The clustering of small and large fluctuations in the squared return series supports the presence of volatility clustering and the use of GARCH-type models for volatility forecasting.

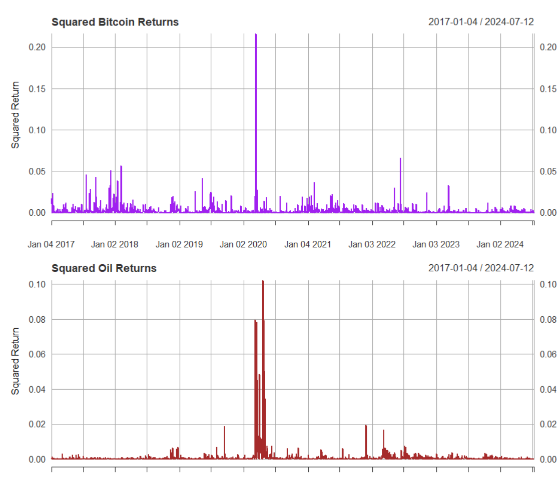

FIGURE 4. Squared returns of Bitcoin and crude oil

4.2. Stationarity Analysis. Before model estimation, the stationarity of the Bitcoin and crude oil return series is examined using the Augmented Dickey–Fuller (ADF) test.The results are reported in Table 2.

TABLE 2. Augmented Dickey–Fuller test results

Series	**ADF Statistic**	**Lag order**	**p-value**	**Conclusion**
Bitcoin Returns	-10.514	12	< 0.01	Stationary
Oil Returns	-9.8888	12	< 0.01	Stationary

The null hypothesis of a unit root is rejected at the 1% level of significance for both Bitcoin and crude oil returns. The large negative ADF statistics provide strong evidence against non-stationarity. This indicates that both return series are stationary.

Stationarity is an important prerequisite for GARCH-type modeling. Thus, the stationary nature of the data supports the validity of the subsequent volatility modelling and parameter estimation.

4.3. Volatility Modeling and Estimation. The estimated parameters of the GARCH(1,1) and EGARCH(1,1) models using the in-sample data are reported in Table 3. The parameter estimates indicate a high degree of volatility persistence in both models, as reflected in the large values of the β_1 coefficients.

TABLE 3. Parameter estimates for GARCH(1,1) and EGARCH(1,1)

Parameter	**GARCH(1,1)**	**EGARCH(1,1)**
μ	0.001363	0.001243
ω	0.000043	-0.112900
α_1	0.097530	0.023709
β_1	0.901470	0.981297
γ	–	0.226037
Shape	3.096307	2.852607

The GARCH(1,1) model shows that current volatility is influenced by both past shocks and past conditional variance. In particular, the relatively large value of β_1 indicates that volatility shocks persist over time. Similarly, the EGARCH model also exhibits strong persistence in volatility, as reflected in its high β_1 coefficient. For the GARCH(1,1) model, the sum of α_1 and β_1 is close to unity, indicating a highly persistent volatility process.

The EGARCH model includes an asymmetry parameter ($\gamma = 0.226037$). This indicates that positive and negative shocks have asymmetric effects on Bitcoin volatility. The shape parameters suggest heavy-tailed behavior in the return distribution for both models, consistent with the non-Gaussian nature of financial return series. Both models are able to capture volatility clustering, and the EGARCH model also provides additional flexibility, allowing asymmetric responses of volatility to market shocks

4.4. Out-of-Sample Forecasting Performance Evaluation. A test set consisting of 20% of the observations is used to evaluate the forecasting performance of the competing models. The forecast accuracy of the models is evaluated using Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and QLIKE loss. The out-of-sample forecasting results for the train–test split are reported in Table 4.

TABLE 4. Out-of-Sample Forecasting Performance of Competing Models (80:20 Split)

Model	Data Split	RMSE	MAE	QLIKE
GARCH	80:20	0.010108	0.008984	-4.632922
EGARCH	80:20	0.002600	0.001961	-5.692465
EGARCH+Oil Vol	80:20	0.002297	0.001378	-5.828875
Weighted EGARCH+Oil Vol	**80:20**	**0.002247**	**0.001283**	**-5.904704**

The results show that the proposed weighted semiparametric EGARCH model with crude oil market volatility achieves the lowest values of RMSE and MAE among all competing models. This indicates that the proposed model outperforms the benchmark GARCH, EGARCH, and unweighted semiparametric EGARCH models. It also records the lowest QLIKE loss, further confirming its superior performance from a likelihood-based forecasting perspective.

The above results indicate that the inclusion of asymmetric volatility dynamics based on the EGARCH model and nonlinear information from crude oil volatility improves forecasting performance. The weighted hybrid model provides more accurate and stable volatility forecasts compared to other competing models.

4.5. Robustness Analysis: Alternative Sample Split (75:25 Split). To examine the stability of the proposed model, a robustness analysis is conducted using an alternative 75:25 train–test split. This analysis is designed to evaluate whether the forecasting performance of the competing models remains consistent across different sample partitions. The out-of-sample forecasting results for the 75:25 split are presented in Table 5

TABLE 5. Out-of-Sample Forecasting Performance (75:25 Split)

Model	Data Split	RMSE	MAE	QLIKE
GARCH	75:25	0.012236	0.010976	-4.464354
EGARCH	75:25	0.002669	0.002197	-5.606998
EGARCH+Oil Vol	75:25	0.002242	0.001483	-5.839700
Weighted EGARCH+Oil Vol	**75:25**	**0.002187**	**0.001357**	**-5.887358**

The results reported in Table 5 are broadly consistent with those obtained using the 80:20 split. The proposed weighted EGARCH model with crude oil market volatility achieves the lowest RMSE and MAE values among all competing models and indicates superior forecasting accuracy under the alternative sample partition. It also records the lowest QLIKE loss, further confirming its robustness in out-of-sample volatility forecasting.

The EGARCH-based specifications demonstrate the importance of modelling volatility asymmetry in financial time series by achieving higher forecasting accuracy than the standard GARCH model. The improvement in forecasting performance observed after incorporating crude oil market volatility further suggests that exogenous market information contributes to a better understanding of Bitcoin volatility dynamics. The results of the robustness analysis further support the conclusion that the proposed weighted semiparametric EGARCH model provides stable and reliable out-of-sample forecasting performance.

4.6. Diebold–Mariano Test for Forecast Accuracy. The Diebold–Mariano (DM) test based on the QLIKE loss function is used to assess whether differences in forecasting performance between competing models are statistically significant. Table 6 presents the result of DM test for the 80:20 and 75:25 sample splits.

TABLE 6. Diebold–Mariano Test Results Based on QLIKE Loss Function

Split	**Comparison**	**DM Statistic**	**p-value**
80:20	GARCH vs EGARCH	-41.781	< 0.001
80:20	EGARCH vs EGARCH + Oil	-4.2239	< 0.001
80:20	EGARCH vs Weighted EGARCH + Oil	-8.4492	< 0.001
80:20	EGARCH + Oil vs Weighted EGARCH + Oil	0.48883	0.6252
75:25	GARCH vs EGARCH	-46.142	< 0.001
75:25	EGARCH vs EGARCH + Oil	-25.542	< 0.001
75:25	EGARCH vs Weighted EGARCH + Oil	-22.455	< 0.001
75:25	EGARCH + Oil vs Weighted EGARCH + Oil	-13.799	< 0.001

The results in Table 6 indicate that the EGARCH model significantly outperforms the traditional GARCH model in both 80:20 and 75:25 sample splits. This confirms the importance of accounting for asymmetric volatility dynamics in modelling Bitcoin return volatility. The inclusion of crude oil market volatility in the model improves forecasting performance, as evidenced by the statistically significant DM test results for the comparison between the EGARCH and EGARCH with crude oil market volatility models in both sample splits.

The proposed weighted semiparametric EGARCH model also performs significantly better than the benchmark EGARCH model in both the 80:20 and 75:25 splits. In the 75:25 split, the weighted model shows a statistically significant improvement over the unweighted EGARCH with crude oil volatility model. This indicates the weighting scheme improves forecast accuracy under the alternative sample partition. However, in the 80:20 split, the DM test for the comparison between the EGARCH with crude oil volatility model and the weighted EGARCH with crude oil volatility model is not statistically significant ($p = 0.6252$). This suggests that, although the weighted model achieves slightly better forecast evaluation measures, its improvement over the unweighted EGARCH with crude oil model is not statistically significant under the 80:20 split.

Overall, the DM test results support the conclusion that incorporating asymmetric volatility dynamics, crude oil market volatility, and a weighting mechanism improves the predictive performance of the proposed volatility forecasting framework.

5. Conclusion

The study investigates the volatility dynamics of Bitcoin returns using a combination of parametric and semiparametric volatility forecasting approaches. Specifically, GARCH (1,1), EGARCH(1,1), a semiparametric EGARCH model incorporating crude oil market volatility, and a proposed weighted semiparametric EGARCH model are examined. Daily Bitcoin and crude oil data are used for the empirical analysis,and two out-of-sample splits, 80:20 and 75:25, are used to evaluate the robustness of the forecasting performance.

The study first extends the classical EGARCH framework by incorporating a nonlinear crude oil market volatility component estimated non parametrically through the Nadaraya–Watson kernel estimator. This yields a semiparametric EGARCH specification in which the EGARCH conditional variance captures the parametric volatility dynamics of Bitcoin returns, while the kernel-based component captures the nonlinear contribution of crude oil market volatility. To further improve forecasting performance, a weighted semiparametric EGARCH model is then proposed, where the EGARCH conditional variance and the kernel-estimated oil volatility component are combined using optimal weights determined from the training sample.

Results show that EGARCH based models consistently outperform the standard GARCH model, highlighting the importance of capturing asymmetric volatility effects in cryptocurrency markets. The inclusion of crude oil market volatility as an external factor in volatility modelling further enhances accuracy of forecasts, indicating spillover effects between the energy and cryptocurrency markets.

Among the competing models, the proposed weighted semiparametric EGARCH model with crude oil volatility has the lowest RMSE values of 0.002247 and 0.002187 for 80:20 and 75:25 data splits, respectively. It also records lowest MAE and QLIKE values in both sample partitions. This indicates substantial improvement in forecasting performance. The Diebold–Mariano test further supports the statistical superiority of the proposed model, with most pairwise comparisons showing significant differences in predictive performance ($p < 0.001$). These findings show that the hybrid model consistently outperforms the competing models across different data splits.

Therefore, this study shows that combining a parametric EGARCH framework with non-parametric kernel-based crude oil volatility components and a weighting scheme yields better forecasting accuracy. It also provides more stable out-of–sample forecasts. The inclusion of both external factors and the internal volatility dynamics of cryptocurrency markets makes the proposed model more flexible and robust for volatility modeling in financial markets.

The proposed model will be useful for future research of other assets like commodities, equities, and exchange rates. It can also incorporate external factors, such as macroeconomic indicators or market sentiment, to boost the accuracy of volatility forecasting. Furthermore, an additional enhancement of the current work could be to incorporate the nonparametric oil volatility component directly into the EGARCH variance equation to estimate a fully integrated semiparametric EGARCH model, instead of combining the two components at the forecasting stage. This methodology could offer a detailed analysis of effects of asymmetric volatility dynamics and nonlinear exogenous market effects. The integration of machine learning techniques and multivariate volatility models might also provide valuable guidance for capturing more complex nonlinear relationships across financial markets.

References

1. Ali, H. H., Kumaraswamy, S., Al Balooshi, S., Abdulla, Y.: Return and volatility spillover between cryptocurrencies, oil price and stock market in GCC countries, *Cogent Economics & Finance* **13** (2025) 2453584.
2. Bollerslev, T.: Generalized autoregressive conditional heteroskedasticity, *J. Econometrics* **31** (1986) 307–327.
3. Bouazizi, T., Galariotis, E., Guesmi, K., Makrychoriti, P.: Investigating the nature of interaction between crypto-currency and commodity markets, *Int. Rev. Financ. Anal.* **88** (2023) 102690.
4. Chen, W. Y., Gerlach, R.: Semiparametric GARCH via Bayesian model averaging, *J. Bus. Econ. Stat.* **39** (2019) 437–452.
5. Engle, R. F.: Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation, *Econometrica* **50** (1982) 987–1007.
6. Hafner, C. M., Rombouts, J. V. K.: Semiparametric multivariate volatility models, *Econometric Theory* **23** (2007) 251–280.
7. Hamilton, J. D.: Oil and the macroeconomy since World War II, *J. Polit. Econ.* **91** (1983) 228–248.
8. Hanif, W., Ko, H.-U., Pham, L., Kang, S. H.: Dynamic connectedness and network in the high moments of cryptocurrency, stock, and commodity markets, *Financial Innovation* **9** (2023) 84.
9. Huang, Z.-C., Sangiorgi, I., Urquhart, A.: Forecasting Bitcoin volatility using machine learning techniques, *J. Int. Financ. Mark. Inst. Money* **97** (2024) 102064.
10. Katsiampa, P.: Volatility estimation for Bitcoin: A comparison of GARCH models, *Economics Letters* **158** (2017) 3–6.
11. Mensi, W., Gubareva, M., Ko, H.-U., Vo, X. V., Kang, S. H.: Tail spillover effects between cryptocurrencies and uncertainty in the gold, oil, and stock markets, *Financial Innovation* **9** (2023) 92.
12. Nelson, D. B.: Conditional heteroskedasticity in asset returns: A new approach, *Econometrica* **59** (1991) 347–370.
13. Patton, A. J.: Volatility forecast comparison using imperfect volatility proxies, *Journal of Econometrics* **160** (2011) 246–256.
14. Pérez-Hernández, F., Arévalo-de-Pablos, A., Camacho-Miñano, M. M.: A hybrid model integrating artificial neural network with multiple GARCH-type models and EWMA for performing the optimal volatility forecasting of market risk factors, *Expert Syst. Appl.* **243** (2024) 122896.
15. Poon, S.-H., Granger, C. W. J.: Forecasting volatility in financial markets: A review, *J. Econ. Lit.* **41** (2003) 478–539.
16. Sadorsky, P.: Oil price shocks and stock market activity, *Energy Economics* **21** (1999) 449–469.
17. Sahiner, M., McMillan, D. G., Kambouroudis, D.: Do artificial neural networks provide improved volatility forecasts: Evidence from Asian markets, *Journal of Economics and Finance* **47** (2023) 723–762.
18. Wang, L., Feng, C., Song, Q., Yang, L.: Efficient semiparametric GARCH modeling of financial volatility, *Statistica Sinica* **22** (2012) 249–270.
19. Wang, Y., Andreeva, G., Martin-Barragan, B.: Machine learning approaches to forecasting cryptocurrency volatility: Considering internal and external determinants, *Int. Rev. Financ. Anal.* **90** (2023) 102914.

20. Zahid, M., Iqbal, F., Koutmos, D.: Forecasting Bitcoin volatility using hybrid GARCH models with machine learning, *Risks* **10** (2022) 237.

Ashwini Kumari : Nitte (Deemed to be University), NMAM Institute of Technology (NMAMIT), Department of Mathematics, Nitte, Karkala, 574110, Karnataka , India.
Email address: dr.ashwini@nitte.edu.in * Corresponding author

Naveena: Department of Statistics, M.G.M. College, Udupi 576102, Karnataka, India
Email address: naveenholla99@gmail.com