Original Research Article

A Hybrid LSTM-DCC Model for Multivariate Cryptocurrency Volatility Prediction

.

ABSTRACT

|  |
| --- |
| Accurate volatility forecasting remains a central challenge in the analysis of cryptocurrency markets, where extreme price fluctuations, nonlinear dependencies and evolving cross-asset correlations complicate traditional modeling approaches. This study proposes a hybrid framework that integrates the Dynamic Conditional Correlation (DCC) GARCH model with Long Short-Term Memory (LSTM) networks to enhance forecasting accuracy. The LSTM–DCC model improves the representation of volatility clustering, structural breaks and interdependencies among digital assets by feeding the time-varying covariance matrix from the DCC process into the LSTM as an additional input at each time step. Using daily return data for Bitcoin (BTC), Ethereum (ETH) and Binance Coin (BNB) from January 2018 to March 2025, the hybrid model was developed and evaluated across multiple forecast horizons and compared with standalone LSTM and DCC-GARCH models. Forecast performance was measured using Mean Absolute Error (MAE) and Root Mean Square Error (RMSE). The LSTM–DCC model achieved the lowest average errors across all horizons, with MAE values of 0.00165 (train), 0.00243 (30-step) and 0.00371 (60-step) and RMSE values of 0.00417 (train), 0.00445 (30-step) and 0.00572 (60-step). These results outperform both standalone LSTM and DCC-GARCH models. This confirms the hybrid model’s superiority in capturing nonlinear temporal dynamics and cross-asset interactions. The findings support the adoption of integrated deep learning and econometric models for robust and reliable volatility forecasting in complex financial environments. |

*Keywords: Deep Learning, MGARCH, LSTM-DCC, Hybrid Models, Cryptocurrency, Volatility Forecasting*

1. INTRODUCTION

Rapid evolution of cryptocurrency markets over the past decade has attracted significant academic and practitioner interest, driven by their high volatility and potential for substantial financial returns. Unlike traditional financial assets, cryptocurrencies exhibit unique market dynamics characterized by extreme price fluctuations, nonlinear dependencies and pronounced cross-asset interrelations (Corbet et al., 2019). These features pose considerable challenges for effective modeling and forecasting of their return volatility, which is critical for portfolio management, risk assessment and derivative pricing.

Accurately forecasting volatility in cryptocurrency markets has become increasingly important, particularly as institutional adoption and trading volumes grow. It plays a vital role in risk management, portfolio optimization and regulatory oversight (Zhou et al., 2025). Cryptocurrencies exhibit strong interdependencies, largely driven by shared sensitivity to macroeconomic factors, market sentiment and technological developments. Understanding these dynamics is essential for managing systemic risk, enhancing portfolio diversification and anticipating contagion effects within the digital asset ecosystem. However, forecasting volatility in this market remains highly challenging due to its nonlinear characteristics, heavy-tailed return distributions and rapid price fluctuations (Sheraz et al., 2022).

Traditional econometric models, such as the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) family, have been extensively used in forecasting financial market volatility (Bollerslev, 1986). These models capture time-varying volatility effectively and their multivariate forms are capable of modeling dynamic relationships among multiple assets (Engle & Kroner, 1995). This feature is especially important in markets like cryptocurrencies, where assets are often interdependent. However, despite their mathematical rigor, GARCH models struggle to account for the nonlinear behavior and long memory commonly found in high-frequency or highly volatile markets (Cont, 2007). Their reliance on linear dynamics and fixed structures limits their responsiveness during periods of market stress.

Deep learning models, a subset of machine learning, have shown strong capabilities in identifying complex nonlinear structures and long-term dependencies in financial data (Fischer & Krauss, 2018). This makes them suitable for markets characterized by instability and structural complexity, such as cryptocurrencies (McNally et al., 2018). Nonetheless, these models face several issues. They often lack transparency, require large datasets and are sensitive to shifts in market conditions (Goodfellow et al., 2016). Additionally, they do not inherently include mechanisms to model volatility clustering, which is a central feature of traditional econometric methods (Legrand, 2019).

To address these limitations, hybrid models have been developed. They combine the mathematical discipline of econometric techniques with the adaptive learning capacity of deep neural networks (Bao et al., 2017). These models aim to integrate structured statistical assumptions with the flexibility needed to capture evolving market dynamics (Kim & Kim, 2021). In doing so, they offer a more balanced framework for forecasting volatility that can accommodate both the known properties of financial time series and the irregular, nonlinear behavior seen in modern digital asset markets. In the case of cryptocurrencies, where price dynamics are influenced by a mix of economic signals, technological shifts and speculative behavior, such integrated approaches can provide improved forecasting accuracy and greater adaptability.

2. Review of Literature

The modeling of cryptocurrency volatility has evolved from traditional econometric methods to more adaptable machine learning and hybrid approaches. Early research predominantly employed time series models such as GARCH variants to account for the time-varying nature of digital asset volatility. For instance, Bouoiyour and Selmi (2015) utilized the ARDL bounds test to investigate factors influencing Bitcoin’s value between December 2010 and June 2014. Their findings characterize Bitcoin chiefly as a speculative asset, indicating connections with investor interest, network hash rate and the Shanghai stock market. Evidence for Bitcoin’s role as a safe haven was limited, whereas some support emerged for its transactional utility and sensitivity to Chinese market fluctuations and mining activity. These outcomes emphasize speculation as the principal driver behind Bitcoin’s price dynamics and emphasize the significance of structural shifts.

Katsiampa (2017) assessed several GARCH-type models to represent Bitcoin return volatility using daily data from July 2010 to October 2016. Comparing standard GARCH(1,1), EGARCH and AR-CGARCH frameworks, the study identified the AR-CGARCH model as the best-fitting, effectively capturing both transient shocks and persistent variance patterns. Additionally, asymmetric models such as EGARCH better reflected leverage effects, revealing the differing impacts of positive and negative shocks. The results suggest that volatility models incorporating persistence and asymmetry provide a more nuanced understanding of Bitcoin’s behavior.

Subsequent studies introduced multivariate models to examine co-movements among financial assets, including cryptocurrencies. Erten et al. (2012) applied the BEKK-GARCH approach to study volatility spillovers among emerging markets during financial stress. Their results demonstrate significant time-varying volatility transmission, with spillover effects intensifying during periods of financial turmoil. The diagonal BEKK model captures the evolving dependencies and uneven volatility reactions, confirming its effectiveness for tracing contagion effects in emerging economies.

Katsiampa et al. (2019) utilized a bivariate diagonal BEKK-GARCH model to examine volatility links between Bitcoin and Ether. The findings demonstrate shifting correlations and changing volatility spillovers between the two cryptocurrencies. These spillovers grow stronger during market turbulence, illustrating their close connection. The study points to the need for models that capture joint volatility to better grasp risk transfer and diversification in digital assets. Zhang et al. (2020) applied the DCC-GARCH model to examine volatility spillovers across global financial markets during the COVID-19 pandemic. Their findings reveal significant time-varying correlations and increased volatility transmission during the crisis, demonstrating the DCC model’s effectiveness in capturing dynamic dependencies and contagion effects under financial stress.

In addition to econometric approaches, recent contributions have emphasized the utility of machine learning models for financial forecasting. Alessandretti et al. (2018) examine the predictive performance of several machine learning models, including linear regression, support vector regression and LSTM networks, for forecasting short-term cryptocurrency price movements. Using historical price data from multiple cryptocurrencies, the study reports improved forecasting accuracy using LSTM compared to traditional econometric models. LSTM models generally outperform both statistical and other machine learning methods in classification and regression tasks. The findings suggest that deep learning models can capture temporal patterns and nonlinear dynamics in cryptocurrency markets, although the authors note that performance depends on hyperparameter tuning and the quality of the training data.

Sebastião and Godinho (2021) investigate the effectiveness of deep learning under high-volatility conditions, applying an ensemble of five models to forecast and trade Bitcoin, Ethereum and Litecoin from 2015 to 2019. Their ensemble approach demonstrates strong predictive power, achieving favorable Sharpe ratios even after incorporating trading costs. Similarly, Chatterjee et al. (2022) evaluate LSTM networks for stock market volatility forecasting, showing their comparative advantage over traditional GARCH-type models in adapting to structural regime shifts and capturing volatility clustering. However, the study also notes inherent limitations in LSTM models, including substantial data demands and reduced interpretability. Umar et al. (2024) explore cross-asset dependencies using multivariate LSTM networks applied to daily exchange rates of four major currencies relative to the Nigerian Naira over a two-decade period. Their findings indicate that bidirectional LSTM models achieve superior forecasting performance compared to unidirectional (vanilla) LSTM networks.

Recent literature has increasingly explored hybrid modeling frameworks that combine the strengths of traditional econometric techniques with the flexibility of deep learning architectures. Michańków et al. (2023) developed a hybrid framework that integrates GARCH models with Gated Recurrent Unit (GRU) networks to forecast volatility and risk. Using daily returns from the S&P 500, gold and Bitcoin and applying the Garman–Klass volatility estimator, the study found that the hybrid models improved point forecast accuracy. However, enhancements in risk measures such as Value-at-Risk (VaR) and Expected Shortfall (ES) were less pronounced, suggesting the benefits of hybridization may be model-specific. Xu et al. (2024) proposed a GARCH-Informed Neural Network (GINN) model that integrates GARCH-based volatility estimates into an LSTM architecture for improved volatility forecasting. Using daily data from seven international stock indices, the study benchmarked the hybrid model against standard GARCH variants and standalone LSTM models. Empirical results showed that GINN achieved consistently lower forecast errors across multiple evaluation metrics, including mean squared error (MSE), mean absolute error (MAE) and the coefficient of determination (R²). On average, the hybrid model outperformed traditional GARCH frameworks by approximately 5% in predictive accuracy. These findings shows the potential of combining econometric structures with deep learning methods to enhance the reliability and precision of volatility forecasts in financial markets.

He et al. (2024) developed a hybrid ARIMA-LSTM model for Bitcoin price forecasting, combining ARIMA’s capacity to model linear trends with LSTM’s ability to capture nonlinear patterns. The model outperformed standalone ARIMA and LSTM approaches in prediction accuracy, indicating that integrating traditional time series methods with deep learning can improve forecasting in volatile financial markets. Li & Dai (2020) who proposed a CNN-LSTM hybrid model for short-term Bitcoin price prediction, combining transaction data with macroeconomic and sentiment indicators. Feature extraction was performed through the CNN, while the LSTM captured temporal dependencies. The hybrid model outperformed standalone CNN and LSTM networks in both value and direction prediction accuracy, demonstrating its capacity to model complex nonlinearities in financial time series.

Huang et al. (2024) assessed LSTM and a CNN-LSTM model incorporating Markov Transition Field (MTF) transformations, benchmarked against GARCH and HAR models. Their results demonstrated that the CNN-LSTM-MTF model consistently delivered superior predictive performance, particularly in short-term volatility forecasts. This supports the use of deep learning models augmented with alternative representations for modeling high-frequency market fluctuations. García-Medina and Aguayo-Moreno (2024) developed an LSTM-GARCH hybrid model in which GARCH-based volatility estimates were fed into an LSTM architecture. While both models showed strong individual performance, the hybrid approach yielded lower forecast errors and better captured structural breaks. Formal statistical tests confirmed its superiority, reinforcing the view that combining econometric and deep learning techniques can enhance predictive performance in volatile markets.

In parallel with advancements in deep learning, recent studies have introduced transformer-based architectures as a promising alternative for modeling financial time series, including cryptocurrency markets. Mahdi et al. (2025) proposed a hybrid framework integrating a Transformer encoder with GRU layers to forecast Bitcoin and Ethereum prices. Drawing on daily price data and benchmarked against BiLSTM and BiGRU models, their approach delivered superior predictive accuracy across multiple evaluation metrics, including MSE, RMSE, MAE and MAPE, with additional confirmation from formal statistical tests. Bui et al. (2025) developed a transformer-based model incorporating Time2Vec feature encoding to enhance temporal representation and asset correlation modeling in financial data. Their empirical findings suggest improved generalization and lower forecast errors, underscoring the utility of advanced temporal encoding within attention-based networks. Additionally, Brugière and Turinici (2024) examined a transformer framework applied to S&P 500 index forecasting, demonstrating the architecture’s effectiveness in capturing dynamic dependencies and long-range temporal patterns. These contributions call attention to the growing relevance of transformer-based models in financial forecasting and their potential to outperform traditional deep learning approaches in volatile and non-stationary environments such as cryptocurrency markets.

3. material and methods

3.1. Data

The empirical analysis utilizes a dataset comprising daily closing prices (denominated in USD) for three leading cryptocurrencies: Bitcoin (BTC), Ethereum (ETH) and Binance Coin (BNB), spanning the period from January 1, 2018, to January 1, 2025. The price data were sourced from CoinMarketCap..

The dataset consists of daily closing prices (in USD) for three major cryptocurrencies: Bitcoin (BTC), Ethereum (ETH) and Binance Coin (BNB), spanning the period from January 1, 2018, to January 1, 2025. The price data were sourced from CoinMarketCap. Daily logarithmic returns are computed as:

Where denotes the closing price of asset on day . The resulting multivariate return vector is used as the input for all subsequent models.

**3.2 Dynamic Conditional Correlation Generalized Autoregressive Conditional Heteroskedasticity (DCC-MGARCH) Model**

The DCC-MGARCH model proposed by Engle (2002), extends the Constant Conditional Correlation (CCC) model by allowing the conditional correlations to evolve over time. The conditional variance-covariance matrix is given by:

Here, is a diagonal matrix of conditional standard deviations and is the matrix of dynamic conditional correlations. The evolution of Dynamic Conditional Correlation matrix is specified as:

Here, denotes the unconditional covariance matrix of the standardized residuals, is the vector of standardized residuals at time , and are non-negative parameters capturing the sensitivity and persistence of the correlation dynamics.

3.3 Long Short Term Memory (LSTM) Networks

Long Short-Term Memory (LSTM) networks, introduced by introduced by Hochreiter and Schmidhuber (1997), are a type of recurrent neural network (RNN) designed to capture long-term dependencies in sequential data. Unlike traditional RNNs, LSTMs use memory cells with specialized gates; input, forget and output gates that regulate the flow of information, enabling the model to retain and update memory states over time.

1. Forget gate layer: Determines what proportion of the previous cell state to retain
2. Input Gate Layer: Controls which new information to store in the cell state:

Cell state update: Combines the retained memory and new information:

1. The output gate layer: Determines the output based on the updated cell state:

Where:

: Input at time t.

: Current and previous hidden states is the hidden state or output at time t

: Current and previous cell state

Input, forget, output and candidate vectors

: Sigmoid activation function

: Hyperbolic tangent activation function

: Weight matrices for the respective gates.

: Bias vectors for the respective gates

: Element-wise (Hadamard) product

3.3 LSTM–DCC Hybrid Model

The LSTM–DCC Hybrid model synergistically combines the strengths of the DCC-MGARCH model with the adaptability of LSTM networks. In this hybrid framework, the DCC model is first used to estimate the time-varying covariance matrix . This output is then introduced as an additional input feature to the LSTM model at each time step, enabling the network to learn from both the raw market data and the volatility dynamics captured by the DCC process. The modified LSTM gates for the hybrid model are defined as:

Here is the modeled covariance matrix at time t from the DCC model, is the candidate cell state and all other variables retain their meanings as described in Section 3.2.

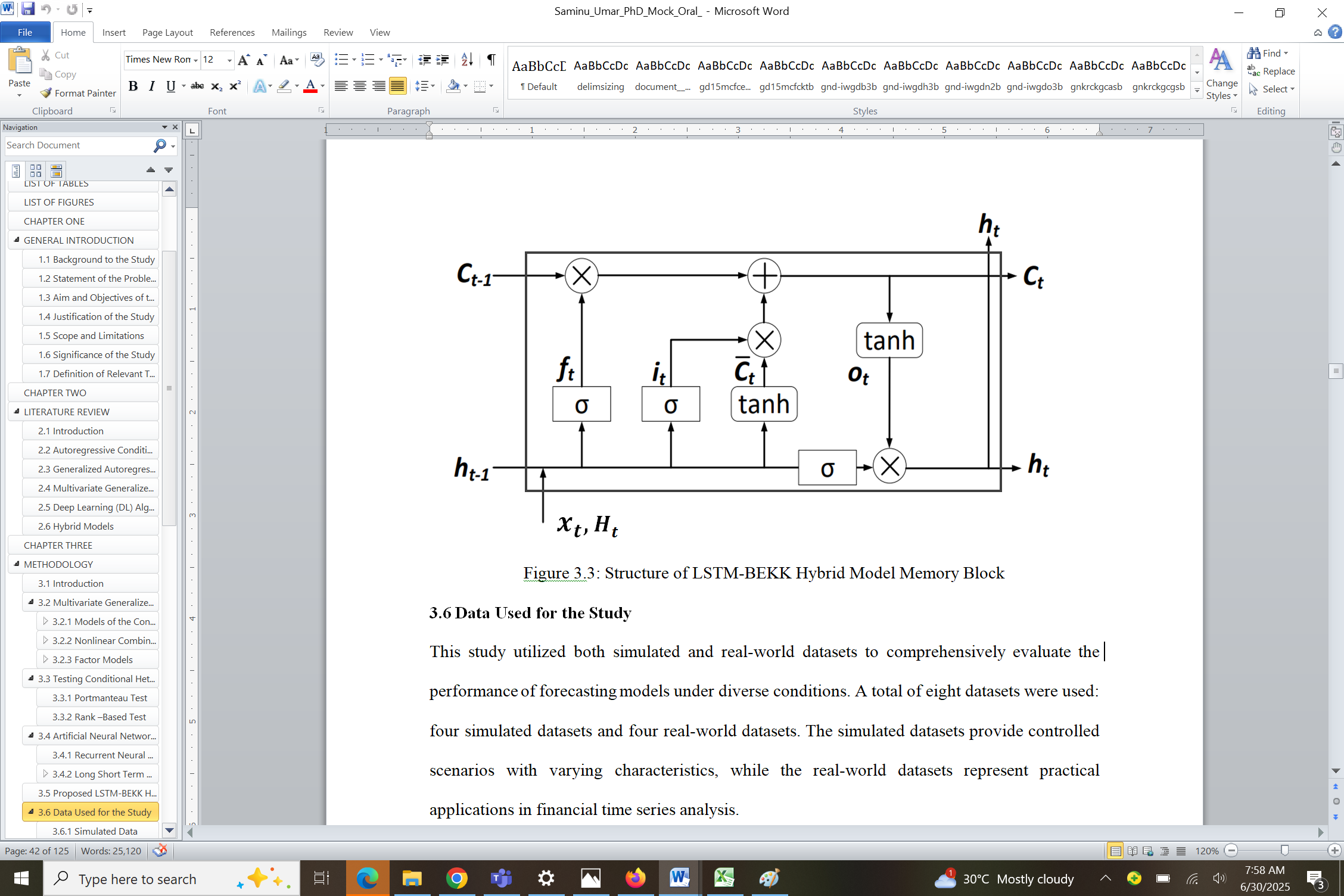


Figure 1: Structure of LSTM-DCC Hybrid Model Memory Block

4. RESULTS and DISCUSSION

This section presents the results of testing the performance of the DCC-MGARCH, LSTM and LSTM–DCC Hybrid models in forecasting cryptocurrency return volatility. It includes descriptive statistics, model performance evaluations and a comparative analysis, providing clear insights into each model’s ability to capture the complex and volatile dynamics of cryptocurrency markets.

4.1 Descriptive Statistics

Descriptive statistics provide an initial overview of the cryptocurrency data, offering insight into its general behavior and underlying characteristics. This forms the basis for further modeling and analysis.



Figure 2: Daily closing prices and returns of BTC, ETH and BNB

Figure 2 displays the daily closing prices and log returns of Bitcoin (BTC), Ethereum (ETH) and Binance Coin (BNB) from January 1 2018 to March 31 2025. The price trend and return patterns reveal periods of rapid growth, sharp declines and subsequent recoveries. Notably, BTC and ETH experienced substantial gains in 2020 and 2021, followed by corrections in 2022 and rebounds by 2024. BNB followed a similar trend, with pronounced growth in 2021, driven by the expansion of the Binance Smart Chain and continued recovery into 2025. These cryptocurrencies exhibit high volatility and correlated price movements. Their returns show large fluctuations, underscoring their speculative nature and making them suitable candidates for evaluating models designed for volatile markets.

Table 1 summarizes the descriptive statistics of the daily log returns. All three cryptocurrencies exhibit positive average returns, with BNB having the highest. ETH shows the greatest variability relative to its average return (as indicated by the coefficient of variation), whereas BNB is the least volatile. The return distributions include extreme highs and lows, with ETH and BNB experiencing particularly sharp negative drops. Skewness values indicate that BTC and ETH tend to have more frequent large negative returns, while BNB shows a slight positive bias. All three return distributions are leptokurtic, suggesting the presence of heavy tails and frequent extreme events. The Augmented Dickey-Fuller (ADF) test confirms stationarity of the return series, while the ARCH test detects volatility clustering. The Jarque-Bera (JB) test strongly rejects normality, justifying the use of models that account for non-Gaussian features.

Table 1:Descriptive Statistics of Daily Log Returns for Cryptocurrencies

| Statistic | BTC | ETH | BNB |
| --- | --- | --- | --- |
| Mean | 0.0006914895 | 0.0003410203 | 0.001619046 |
| Coefficient of Variation | 51.2047 | 133.7487 | 30.69602 |
| Minimum | -0.4647302 | -0.5507317 | -0.5430839 |
| Maximum | 0.1718206 | 0.2306952 | 0.5292179 |
| Skewness | -0.9645488 | -0.9376635 | 0.2971755 |
| Kurtosis | 13.9913 | 11.22384 | 20.49041 |
| ADF Test | -13.065 \*\*\* | -13.210 \*\*\* | -12.987 \*\*\* |
| ARCH Test | 63.593 \*\*\* | 84.460 \*\*\* | 258.950 \*\*\* |
| JB Test | 22,041 \*\*\* | 14,309 \*\*\* | 46,426 \*\*\* |

**Note** \*\*\* indicates the rejection of the null hypotheses at the 1% level

4.2 DCC-MGARCH Model Estimation

Table 2 reports the estimated parameters of the DCC-MGARCH model for BTC, ETH and BNB returns. The mean returns are positive, with statistical significance only for BTC, while AR(1) terms are negative and insignificant, suggesting limited short-term autocorrelation. Volatility parameters α₁ (ARCH) and β₁ (GARCH) are positive and highly significant across all assets, indicating strong volatility clustering and persistence. BTC and BNB exhibit higher α₁ estimates compared to ETH, whereas ETH has the highest β₁, reflecting subtle differences in volatility dynamics. The constant variance term ω is significant for BTC and BNB but not for ETH, suggesting a baseline volatility level is more pronounced in these assets. The joint DCC parameters reveal a highly persistent dynamic correlation process, consistent with existing literature on cryptocurrency market interdependencies. Overall, these results confirm the DCC-MGARCH model’s capacity to capture the complex conditional variance and correlation structures inherent in cryptocurrency returns.

Table 2: Parameter Estimates of the DCC-MGARCH Model

| Parameter | Estimate | Std. Error | t-value | p-value |
| --- | --- | --- | --- | --- |
| BTC Mean (μ) | 0.00142 | 0.00065 | 2.16 | 0.031 |
| BTC AR(1) | -0.03048 | 0.03396 | -0.90 | 0.369 |
| BTC Alpha (α₁) | 0.10014 | 0.04653 | 2.15 | 0.031 |
| BTC Beta (β₁) | 0.84985 | 0.04470 | 19.01 | <0.01 |
| BTC Omega (ω) | 0.00007 | 0.00003 | 2.69 | 0.007 |
| ETH Mean (μ) | 0.00096 | 0.00076 | 1.26 | 0.208 |
| ETH AR(1) | -0.02576 | 0.02184 | -1.18 | 0.238 |
| ETH Alpha (α₁) | 0.07416 | 0.03124 | 2.37 | 0.018 |
| ETH Beta (β₁) | 0.90885 | 0.03787 | 23.99 | <0.01 |
| ETH Omega (ω) | 0.00004 | 0.00003 | 1.45 | 0.146 |
| BNB Mean (μ) | 0.00105 | 0.00068 | 1.55 | 0.121 |
| BNB AR(1) | -0.03646 | 0.02638 | -1.38 | 0.167 |
| BNB Alpha (α₁) | 0.15345 | 0.05163 | 2.97 | <0.01 |
| BNB Beta (β₁) | 0.84478 | 0.04164 | 20.29 | <0.01 |
| BNB Omega (ω) | 0.00005 | 0.00002 | 2.18 | 0.029 |
| Joint dcca1 | 0.04202 | 0.00756 | 5.56 | <0.01 |
| Joint dccb1 | 0.94351 | 0.01285 | 73.41 | <0.01 |

4.3 LSTM Model Estimation

Table 3 details the LSTM network architecture used for modeling normalized cryptocurrency data with a sequence length of 28 days. The model begins with an LSTM layer of 512 units employing ReLU activation, followed by a Dropout layer with a rate of 0.3 to prevent overfitting. Two subsequent Dense layers with 128 units (ReLU) and 3 units (output) map the learned features to the target variables. The model contains approximately 1.12 million fully trainable parameters, balancing complexity and regularization for effective temporal feature extraction.

Table 3. LSTM Network Architecture Summary

| Layer (Type) | Output Shape | Param # |
| --- | --- | --- |
| LSTM (lstm\_1) | (None, 512) | 1,056,768 |
| Dropout (dropout\_1) | (None, 512) | 0 |
| Dense (dense\_2) | (None, 128) | 65,664 |
| Dense (dense\_3) | (None, 3) | 387 |
| Total Parameters | — | 1,122,819 |
| Trainable Params | — | 1,122,819 |
| Non-trainable Params | — | 0 |

4.4 Hybrid LSTM–DCC Model Estimation

Table 4 outlines the architecture of the LSTM-DCC hybrid model, which extends the LSTM framework to support dynamic correlation analysis. Similar to the baseline LSTM model, the architecture begins with a 512-unit LSTM layer followed by a Dropout layer and a Dense layer with 128 units, maintaining a comparable structural backbone. However, the final Dense layer in this model outputs 12 features, in contrast to the 3-output configuration of the LSTM model (Table 3). This expanded output dimensionality is designed to generate latent representations of multiple series as inputs to the DCC model. The total number of trainable parameters increases slightly to 1.14 million, reflecting the broader output scope. The model remains fully trainable, enabling integrated feature extraction prior to the DCC estimation phase.

Table 4: LSTM-DCC Hybrid Model Architecture Summary

| Layer (type) | Output Shape | Parameters |
| --- | --- | --- |
| LSTM (lstm\_2) | (None, 512) | 1,075,200 |
| Dropout (dropout\_2) | (None, 512) | 0 |
| Dense (dense\_4) | (None, 128) | 65,664 |
| Dense (dense\_5) | (None, 12) | 1,548 |
| Total Parameters |  | 1,142,412 |
| Trainable params |  | 1,142,412 |
| Non-trainable params |  | 0 |

4.5 Comparative Analysis of Model Performance

Table 5 reports the average forecast errors for BTC, ETH and BNB returns across the DCC-MGARCH, LSTM and LSTM-DCC hybrid models, evaluated via Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) metrics over the training sample and at 30- and 60-step ahead horizons. The LSTM-DCC hybrid model exhibits consistently superior predictive accuracy, achieving the lowest average MAE and RMSE values across all horizons. This performance advantage becomes more pronounced at extended forecast horizons, underscoring the hybrid model’s enhanced ability to capture intricate nonlinear temporal dependencies and dynamic cross-asset interactions. The LSTM model also significantly outperforms the conventional DCC-MGARCH benchmark, reflecting the limitations of traditional volatility-based models in accommodating the complex behaviors characteristic of cryptocurrency markets. Collectively, these findings substantiate the efficacy of deep learning-based hybrid frameworks for robust multi-step forecasting in highly volatile and interdependent asset environments.

Table 5: Forecasting Performance on Cryptocurrency Return Series

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Model | Metric | Train | 30 points ahead | 60 points ahead |
| DCC-MGARCH | MAE | 0.03539 | 0.02365 | 0.02526 |
| RMSE | 0.04142 | 0.02413 | 0.02631 |
| LSTM | MAE | 0.00429 | 0.00612 | 0.00720 |
| RMSE | 0.00710 | 0.00706 | 0.00857 |
| LSTM-DCC | MAE | 0.00165 | 0.00243 | 0.00371 |
| RMSE | 0.00417 | 0.00445 | 0.00572 |

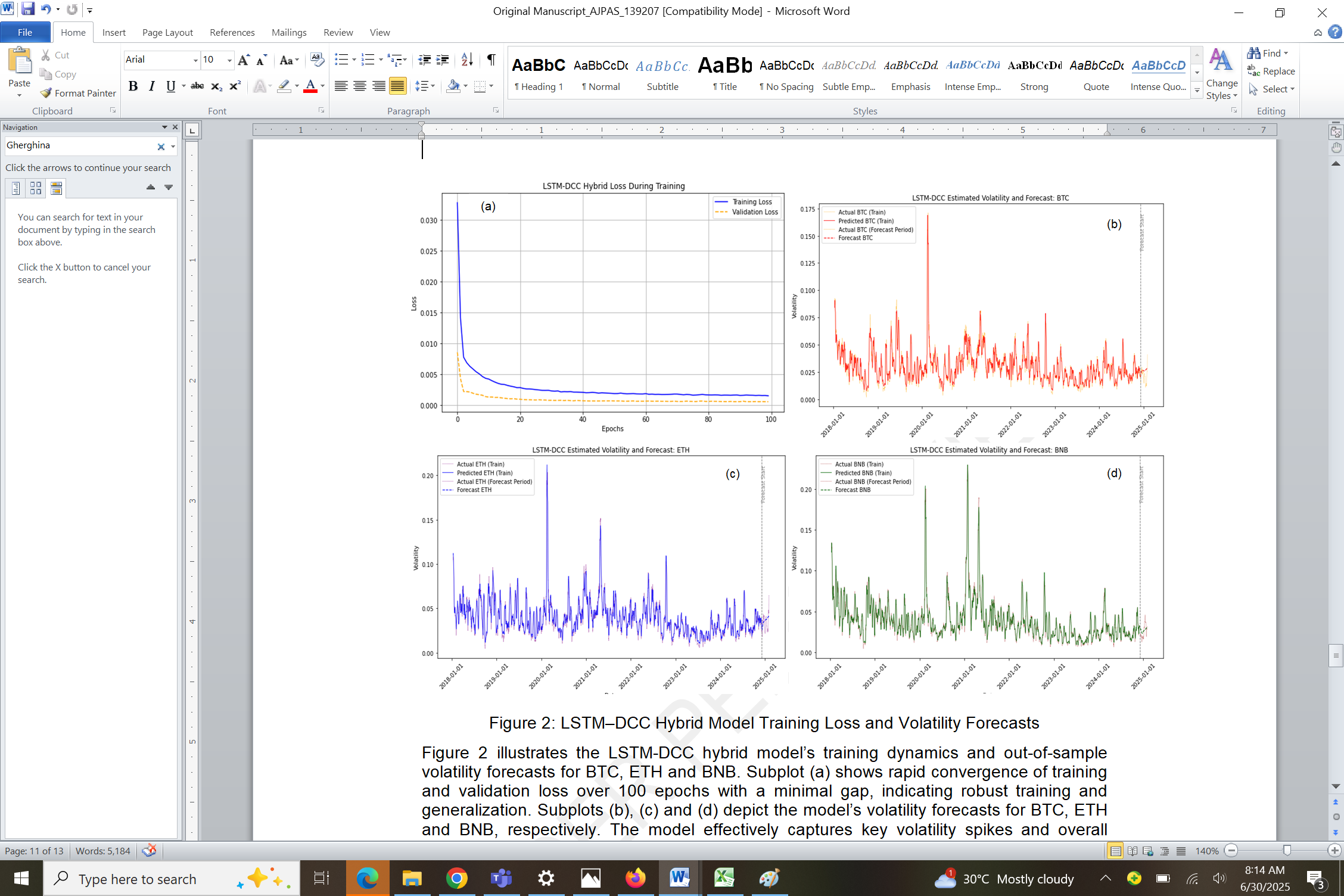


Figure 3: LSTM–DCC Hybrid Model Training Loss and Volatility Forecasts

Figure 3 illustrates the LSTM-DCC hybrid model’s training dynamics and out-of-sample volatility forecasts for BTC, ETH and BNB. Subplot (a) shows rapid convergence of training and validation loss over 100 epochs with a minimal gap, indicating robust training and generalization. Subplots (b), (c) and (d) depict the model’s volatility forecasts for BTC, ETH and BNB, respectively. The model effectively captures key volatility spikes and overall patterns across all three assets, with minor underestimations during abrupt surges. Notably, it maintains strong predictive accuracy even amid the pronounced volatility bursts characteristic of BNB. These plots complement the quantitative findings in Table 5, collectively demonstrating the hybrid model’s efficacy in modeling complex, nonlinear volatility dynamics in cryptocurrency markets.

5. CONCLUSION

This study develops and evaluates an LSTM-DCC hybrid framework for modeling and forecasting volatility dynamics in major cryptocurrencies. By integrating DCC model’s ability to characterize time-varying cross-asset correlations with the LSTM deep learning Networks capacity to capture complex nonlinear temporal patterns, the proposed approach addresses critical limitations of traditional volatility models in highly volatile and interdependent markets. Empirical evaluation across Bitcoin, Ethereum and Binance Coin demonstrates that the hybrid model consistently outperforms both traditional DCC-MGARCH and standalone LSTM models in multi-step ahead forecast accuracy, while maintaining robust generalization and training stability. These results show the significant benefits of integrating deep learning architectures with established econometric models to enhance volatility modeling in complex financial environments. The findings contribute to the growing literature on hybrid approaches and provide a rigorous methodological foundation for improved risk management and asset allocation in cryptocurrency portfolios. Future work may explore extensions incorporating alternative neural network structures and high-frequency data to further advance forecasting precision and adaptability.

**Disclaimer (Artificial intelligence)**

Authors hereby declare that NO generative AI technologies such as Large Language Models (ChatGPT, COPILOT, etc.) and text-to-image generators have been used during the writing or editing of this manuscript.

**Competing Interests**

Authors have declared that no competing interests exist.

References

Alessandretti, L., ElBahrawy, A., Aiello, L. M., & Baronchelli, A. (2018). Anticipating cryptocurrency prices using machine learning. *Complexity*, *2018*(1), 8983590.

Bao, W., Yue, J., & Rao, Y. (2017). A deep learning framework for financial time series using stacked autoencoders and long-short term memory. *PloS one*, *12*(7), e0180944.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, *31*(3), 307-327.

Bouoiyour, J., & Selmi, R. (2015). What does Bitcoin look like?. *Annals of Economics & Finance*, *16*(2).

Brugiere, P., & Turinici, G. (2025, March). Transformer for time series: An application to the s&p500. In *Future of Information and Communication Conference* (pp. 511-528). Cham: Springer Nature Switzerland.

Bui, N. K. H., Chien, N. D., Kovács, P., & Bognár, G. (2025). Transformer Encoder and Multi-features Time2Vec for Financial Prediction. *arXiv preprint arXiv:2504.13801*.

Chatterjee, A., Bhowmick, H., & Sen, J. (2022, October). Stock volatility prediction using time series and deep learning approach. In *2022 IEEE 2nd Mysore Sub Section International Conference (MysuruCon)* (pp. 1-6). IEEE.

Cont, R. (2007). Volatility clustering in financial markets: empirical facts and agent-based models. In *Long memory in economics* (pp. 289-309). Berlin, Heidelberg: Springer Berlin Heidelberg.

Corbet, S., Lucey, B., Urquhart, A., & Yarovaya, L. (2019). Cryptocurrencies as a financial asset: A systematic analysis. *International Review of Financial Analysis*, *62*, 182-199.

Engle, R. (2002). Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of business & economic statistics*, *20*(3), 339-350.

Engle, R. F., & Kroner, K. F. (1995). Multivariate simultaneous generalized ARCH. *Econometric theory*, *11*(1), 122-150.

Erten, I., Tuncel, M. B., & Okay, N. (2012). Volatility spillovers in emerging markets during the global financial crisis: diagonal BEKK approach.

Fischer, T., & Krauss, C. (2018). Deep learning with long short-term memory networks for financial market predictions. *European journal of operational research*, *270*(2), 654-669.

García-Medina, A., & Aguayo-Moreno, E. (2024). LSTM–GARCH hybrid model for the prediction of volatility in cryptocurrency portfolios. *Computational Economics*, *63*(4), 1511-1542.

Gherghina, Ş. C., & Simionescu, L. N. (2023). Exploring the asymmetric effect of COVID-19 pandemic news on the cryptocurrency market: evidence from nonlinear autoregressive distributed lag approach and frequency domain causality. *Financial Innovation*, *9*(1), 21.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). Regularization for deep learning. *Deep learning*, 216-261.

He, X., Huang, J., Ma, K., & He, H. (2024, November). Bitcoin price prediction based on ARIMA-LSTM model. In *Proceedings of the 5th International Conference on Artificial Intelligence and Computer Engineering* (pp. 655-659).

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, *9*(8), 1735-1780.

Huang, Z. C., Sangiorgi, I., & Urquhart, A. (2023). Forecasting Bitcoin volatility using machine learning techniques. *Journal of International Financial Markets, Institutions and Money*, *97*, 102064.

Katsiampa, P. (2017). Volatility estimation for Bitcoin: A comparison of GARCH models. *Economics letters*, *158*, 3-6.

Katsiampa, P. (2019). Volatility co-movement between Bitcoin and Ether. *Finance Research Letters*, *30*, 221-227.

Kim, T., & Kim, H. Y. (2019). Forecasting stock prices with a feature fusion LSTM-CNN model using different representations of the same data. *PloS one*, *14*(2), e0212320.

Legrand, N. (2019). The empirical merit of structural explanations of commodity price volatility: Review and perspectives. *Journal of Economic Surveys*, *33*(2), 639-664.

Li, Y., & Dai, W. (2020). Bitcoin price forecasting method based on CNN‐LSTM hybrid neural network model. *The journal of engineering*, *2020*(13), 344-347.

Mahdi, E., Martin-Barreiro, C., & Cabezas, X. (2025). A Novel Hybrid Approach Using an Attention-Based Transformer+ GRU Model for Predicting Cryptocurrency Prices. *Mathematics*, *13*(9), 1484.

McNally, S., Roche, J., & Caton, S. (2018, March). Predicting the price of bitcoin using machine learning. In *2018 26th euromicro international conference on parallel, distributed and network-based processing (PDP)* (pp. 339-343). IEEE.

Michańków, J., Kwiatkowski, Ł., & Morajda, J. (2023). Combining Deep Learning and GARCH Models for Financial Volatility and Risk Forecasting. *arXiv preprint arXiv:2310.01063*.

Sebastião, H., & Godinho, P. (2021). Forecasting and trading cryptocurrencies with machine learning under changing market conditions. *Financial Innovation*, *7*, 1-30.

Sheraz, M., Dedu, S., & Preda, V. (2022). Volatility dynamics of non-linear volatile time series and analysis of information flow: Evidence from cryptocurrency data. *Entropy*, *24*(10), 1410.

Umar, S., Lamido, A. M., & Aminu, M. K. (2024). Multivariate foreign exchange rate prediction with Long-Short Term Memory deep learning networks. International Journal of Novel Research in Marketing Management and Economics, 11(2).

Xu, Z., Liechty, J., Benthall, S., Skar-Gislinge, N., & McComb, C. (2024, November). GARCH-Informed Neural Networks for Volatility Prediction in Financial Markets. In *Proceedings of the 5th ACM International Conference on AI in Finance* (pp. 600-607).

Zhang, D., Hu, M., & Ji, Q. (2020). Financial markets under the global pandemic of COVID-19. *Finance research letters*, *36*, 101528.

Zhou, Y., Xie, C., Wang, G. J., Gong, J., & Zhu, Y. (2025). Forecasting cryptocurrency volatility: a novel framework based on the evolving multiscale graph neural network. *Financial Innovation*, *11*(1), 87.