INTERNATIONAL JOURNAL OF ACADEMIC RESEARCH IN ACCOUNTING, FINANCE AND MANAGEMENT SCIENCES

OPEN ACCESS JOURNAL

Vol 16, Issue 1, (2026) E-ISSN: 2225-8329

Bridging Econometrics and Deep Learning: A Hybrid GARCH-BiLSTM Approach for Forecasting Coffee Futures Volatility

Lin Yihuan[1], Choo Wei Chong[2*], Ng Keng Yap[3], Zhang Juan[4], Wu Youyuan[5]

[1,2]School of Business and Economics, Universiti Putra Malaysia, 43400, Serdang, Selangor Darul Ehsan, Malaysia, [3]Department of Computer Science, Faculty of Computer Science and Information Technology, Universiti Putra Malaysia, 43400, Serdang, Selangor Darul Ehsan, Malaysia, [4]Institute for Mathematical Research (INSPEM), Universiti Putra Malaysia, 43400, Serdang, Selangor Darul Ehsan, Malaysia, [5]School of Economics and Finance, Huaqiao University, 362021, Quanzhou, Fujian Province, People's Republic of China

*Corresponding Author Email: wcchoo@upm.edu.my

DOI Link: http://dx.doi.org/10.6007/IJARAFMS/v16-i1/27505

Published Online: 28 January 2026

Abstract

The inherent instability of agricultural commodity markets, particularly for Arabica coffee futures, poses significant challenges for risk management due to the sector's vulnerability to climate shocks and supply chain disruptions. Although traditional econometric models like GARCH successfully capture stylized facts such as volatility clustering, they often struggle to account for the abrupt structural breaks and non-linear dependencies introduced by exogenous events. To address these limitations, this study proposes a hybrid framework that bridges the methodological divide between econometric rigor and deep learning flexibility. Specifically, we integrate the structural parameters of Generalized Autoregressive Conditional Heteroskedasticity (GARCH) and its asymmetric variant, GJR-GARCH, directly into Bidirectional Long Short-Term Memory (BiLSTM) networks. By employing a residual learning strategy, this approach leverages the theoretical interpretability of econometrics while utilizing the bidirectional processing capabilities of neural networks to capture complex temporal dependencies. Empirical analysis utilizing Intercontinental Exchange (ICE) data from 2000 to 2025 reveals that the hybrid GJR-BiLSTM architecture statistically outperforms both standalone econometric models and unidirectional LSTM baselines. The results demonstrate that incorporating asymmetric shock parameters allows the model to effectively capture the distinct leverage effect and long-memory characteristics observed in coffee markets. Furthermore, the study identifies a 21-day rolling window as the optimal horizon for maximizing forecast stability. Ultimately, these findings underscore the economic value of hybrid architectures, providing actionable insights for stakeholders seeking to optimize hedging strategies and mitigate financial exposure in global coffee supply chains.

Keywords: Volatility Forecasting, GARCH-BILSTM, Deep Learning; Coffee Futures, Risk Management, Asymmetric Volatility

Introduction

The global agricultural commodity market is characterized by inherent instability. In this context, price volatility is not merely a statistical phenomenon but a critical determinant of economic survival for millions. Among soft commodities, coffee stands out as particularly vulnerable. It is subject to a complex interplay of biological cycles, climate anomalies, and geopolitical supply chain disruptions (Massrie, 2025). For producing nations like Brazil and Vietnam, coffee price shocks can destabilize export revenues and rural livelihoods. Conversely, for consuming markets, these shocks introduce significant cost uncertainties. This volatility creates profound asymmetries: for producing nations in the Global South (e.g., Brazil, Vietnam, Ethiopia), price shocks can destabilize national export revenues and devastate the livelihoods of smallholder farmers, who often lack access to sophisticated risk management tools. Conversely, for consuming markets in the Global North, these shocks introduce significant cost uncertainties for roasters, retailers, and consumers, while financial intermediaries seek to capitalize on price movements. Consequently, accurate volatility forecasting becomes a fundamental necessity for effective risk management. It enables stakeholders, ranging from smallholder farmers to multinational hedge funds, to optimize hedging ratios and mitigate financial exposure (Abaidoo & Agyapong, 2022; Singhal & Tarp, 2024).

Despite this urgency, modeling coffee volatility remains a formidable challenge. Traditional econometric approaches, specifically the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) family, have long served as the industry standard. This is due to their ability to capture stylized facts like volatility clustering (Bollerslev, 1986). However, these linear models often struggle to account for abrupt structural breaks and complex non-linear dependencies. Such irregularities are frequently introduced by exogenous shocks, including frost events or pandemics (Chung et al., 2025). On the other hand, the recent rise of deep learning offers superior flexibility in modeling non-linear patterns. Recurrent Neural Networks (RNNs), such as Long Short-Term Memory (LSTM) models, are particularly effective in this regard (Hochreiter & Schmidhuber, 1997). Yet, these models frequently suffer from a lack of interpretability and theoretical grounding. Often described as “black boxes,” they risk overfitting noise rather than extracting meaningful economic signals (Černevičienė & Kabašinskas, 2024).

This persistent methodological gap—between the interpretable yet rigid frameworks of econometrics and the flexible yet opaque architectures of deep learning—defines the core research problem. While hybrid models have been explored in other asset classes, their application to agricultural commodities, particularly coffee, remains nascent and lacks a thorough investigation of bidirectional learning architectures that can capture the unique temporal dependencies shaped by both past shocks and future expectations (e.g., weather forecasts, harvest reports). Therefore, this study addresses the following explicit research questions: (1) Can a hybrid framework that explicitly integrates econometric parameters into a bidirectional deep learning model (BiLSTM) significantly outperform standalone and unidirectional models in forecasting coffee futures volatility? (2) Does accounting for asymmetry (via GJR-GARCH) within such a hybrid provide a decisive advantage in capturing the leverage effect inherent in agricultural markets? (3) What is the optimal temporal window for such a model, balancing responsiveness to recent shocks with stability from longer-term trends? By answering these questions, we aim to provide a methodologically robust and socially relevant tool for volatility forecasting.

This study bridges this methodological divide by proposing a parameter-based hybrid framework. Our approach integrates the structural rigor of GARCH models with the predictive power of Bi-directional LSTM (BiLSTM) networks. Unlike standard LSTMs that process data unidirectionally, BiLSTMs analyze temporal sequences in both forward and backward directions. This capability allows for a more comprehensive capture of volatility dynamics (Schuster & Paliwal, 1997). We feed the conditional variance and asymmetric shock parameters from GARCH and GJR-GARCH models directly into the neural network. This provides the model with a theoretically sound informational advantage. As a result, the network can focus on learning complex residuals rather than relearning basic autoregressive patterns.

The contributions of this study are threefold. First, we empirically demonstrate that hybrid GARCH-RNN architectures, particularly those incorporating asymmetry (GJR-GARCH), statistically outperform both standalone econometric and neural network models in forecasting coffee futures volatility. This validates the importance of accounting for the leverage effect in agricultural markets, where bad news impacts volatility more than good news. Second, we introduce the BiLSTM architecture to this domain. We show that bidirectional processing significantly enhances predictive accuracy by capturing longer-term dependencies. Third, we systematically evaluate the impact of rolling window sizes. We identify an optimal memory length that balances responsiveness with stability. This factor is critical for practical algorithmic trading and hedging applications (Feng & Zhang, 2025).

Literature Review

The pursuit of accurate volatility forecasting in agricultural commodity markets has evolved from classical econometric approaches to sophisticated machine learning paradigms. This section synthesizes the existing body of knowledge. It highlights the unique characteristics of coffee market volatility, the limitations of linear models, and the emerging potential of hybrid neural architectures.

Volatility Dynamics in Agricultural Commodities

Agricultural commodities, particularly soft commodities like coffee, exhibit distinct volatility dynamics. These dynamics differentiate them from equities or foreign exchange markets. A primary characteristic is the presence of the leverage effect, where price volatility responds asymmetrically to market shocks. Research by Carpantier (2010) and others has consistently shown that in commodity markets, bad news often triggers a sharper increase in volatility than good news of equal magnitude. Examples of bad news include supply shortages or inventory deficits. This asymmetry is particularly pronounced in the coffee market because supply is inelastic in the short term due to the biological lag of coffee tree growth. Consequently, standard GARCH models assume a symmetric response to shocks. Therefore, they often fail to capture the full extent of market risk during downturns. This limitation underscores the necessity of employing asymmetric models like GJR-GARCH. This necessity is validated by recent empirical studies on agricultural products. For instance, Wang et al.

(2022) demonstrated this in their study on garlic prices, another volatile agricultural commodity. They found that integrating asymmetric GARCH-family models with neural networks significantly improves the capture of these skewed volatility characteristics.

Furthermore, coffee prices are characterized by extreme volatility clustering and fat tails (leptokurtosis). These features are driven by exogenous factors such as the El Niño Southern Oscillation (ENSO) and frost events in Brazil (Massrie, 2025). Ubilava (2012) specifically quantified these impacts, demonstrating that ENSO phases induce distinct asymmetric shock transmissions in global coffee markets. These physical drivers introduce non-linear chaotic patterns into the time series. Purely autoregressive models struggle to predict such patterns. While econometric models excel at quantifying the persistence of volatility, they are less effective at predicting the structural breaks caused by these sudden physical shocks. The economic and social consequences of this volatility are profound and extend beyond financial metrics. Price instability directly impacts the livelihoods of millions of smallholder farmers, contributing to income uncertainty and psychological distress (Singhal & Tarp, 2025), while complicating supply chain management and long-term competitiveness for producing nations (Wanzala & Obokoh, 2025; Nguyen, 2023). Therefore, improving volatility forecasting is not merely a technical challenge but a critical enabler for economic resilience and more equitable risk management within the global coffee value chain (Massrie, 2025).

Incompetence of Traditional Econometrics and the Rise of Deep Learning

For decades, the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) family of models has been the cornerstone of volatility forecasting. However, the rigidity of their functional forms imposes strict assumptions about the data generating process. These assumptions include normality or specific error distributions. As financial time series have become more complex and high-frequency, the linear limitations of GARCH have become increasingly apparent. Critics argue that GARCH models often underperform in out-of-sample forecasting. This is because they cannot adapt to the shifting regimes and non-linear interactions inherent in modern markets. Recent work by Chung et al. (2025) explicitly highlights this issue. They note that neglecting structural breaks, such as those observed during the COVID-19 pandemic or climate crises, causes GARCH models to overstate volatility persistence. This consequently weakens predictive accuracy.

In response to these limitations, Deep Learning (DL) has gained prominence. Specifically, Recurrent Neural Networks (RNNs) have become a key tool. Unlike econometric models, RNNs and their advanced variants, Long Short-Term Memory (LSTM) networks, are universal approximators. They are capable of learning non-linear dependencies without prior assumptions. Studies specifically targeting agricultural markets have highlighted the utility of LSTM architectures. They are uniquely suited to model the non-stationary and non-linear features of price series, capturing complex dependencies that linear models miss (Jaiswal et al., 2022). Recent empirical work supports this transition. Manogna et al. (2025) demonstrated that deep learning models consistently outperformed traditional stochastic baselines across a broad range of agricultural commodities. Similarly, Le et al. (2024) found that LSTM architectures achieved higher accuracy than standard GARCH models in predicting volatility for cotton trading. However, pure deep learning models are not without flaws. They are frequently characterized as "black boxes" that lack transparency (Černevičienė & Kabašinskas, 2024). While they may achieve high goodness-of-fit, they often struggle to distinguish between meaningful economic signals and transient noise without the structural guidance provided by domain-specific theory. Recent systematic reviews underscore a paradigm shift towards hybrid models that mitigate these limitations. For instance, Mansilla-Lopez et al. (2025) note that hybrid GARCH-LSTM architectures consistently outperform standalone GARCH models in volatility forecasting, while Vancsura et al. (2025) identify hybrid and attention-based deep learning models as key methods for handling the volatility and seasonality inherent in agricultural commodity prices. This consensus highlights the critical need to integrate econometric structure with deep learning flexibility.

Hybrid Approaches and the Bidirectional Gap

Scholars have increasingly turned to hybrid architectures to bridge the gap between the structural interpretability of econometrics and the flexible power of deep learning. The underlying logic is often referred to as “residual learning.” This conceptual framework draws upon the pioneering work of Zhang (2003), who established the theoretical validity of decomposing time series into linear and non-linear components to maximize forecasting accuracy. It posits that a linear model should first extract the known statistical properties, such as volatility clustering. Then, the non-linear neural network models the complex residuals. Empirical evidence from commodity market studies suggests that hybrid models consistently outperform standalone approaches. Huang et al. (2022) demonstrated that a GARCH-LSTM ensemble significantly improved volatility forecasting for a basket of metals. This advantage is corroborated by broader analyses. Rabbani et al. (2025) found that hybrid frameworks like GARCH-XGBoost and ARIMA-LSTM surpassed both pure econometric and machine learning models in forecasting financial risks, and Theofilou et al. (2025) documented a surge in the use of LSTM-based hybrid models for staple crop price forecasting during 2023-2024. Hu et al. (2020) found similar success in predicting copper price volatility. By using GARCH-derived conditional variance as an input feature, the neural network is provided with an informational advantage. This allows it to focus on learning the intricate non-linear deviations rather than relearning basic autoregressive patterns.

The frontier of hybrid architecture involves leveraging bidirectional processing. Bidirectional LSTM (BiLSTM) networks, which process sequences in both forward and backward directions, have shown promise in capturing richer contextual dependencies in financial time series (Giantsidi & Tarantola, 2025; Krichen & Mihoub, 2025). Their ability to incorporate information from both past and future within a training window makes them particularly suited for modeling the “smoothness” and context of volatility curves. Recent comparative studies indicate that BiLSTMs often outperform unidirectional LSTMs in sequential prediction tasks where context is crucial (Aitbekov, 2025; Tumpa & Maduranga, 2024). Despite these advancements, a specific gap remains in the application of Bidirectional LSTM (BiLSTM) to agricultural commodities. While standard LSTM processes data unidirectionally from past to future, BiLSTM processes the sequence in both directions. It captures context from both the past and the “future” within the context of training windows. In volatility forecasting, the “smoothness” and continuity of the variance curve are crucial. The ability to look forward and backward within the training window allows the model to better understand the context of a shock. Few studies have applied a GARCH-BiLSTM hybrid framework specifically to the Arabica coffee market. This sector features a unique bidirectional dependency structure created by the interplay of historical weather patterns and future supply expectations. This study aims to fill this void by systematically evaluating the performance of BiLSTM hybrids against their unidirectional counterparts.

Method

Data and Statistical Approaches

We utilized daily closing prices of coffee C futures from the Intercontinental Exchange (ICE). This constitutes the global benchmark for Arabica coffee. The dataset spans from March 2, 2000, to December 29, 2017, for the training phase. This comprises a total of 4,493 trading days. This extensive training period ensures that the models are exposed to various market cycles, including the commodity boom of the 2000s and subsequent corrections. Meantime, the data from January 2, 2018, to March 21, 2025, encompassing 1,787 trading days, are used for strictly out-of-sample forecasting. These financial data were collected from Refinitiv Eikon, ensuring data integrity and consistency. Logarithmic returns were computed as the natural logarithm of the price ratio between consecutive days. This is a standard transformation to approximate continuous compounding and stabilize the variance of the series. Python (Version 3.12.4) was employed for the analyses and visualizations.

Table 1
Descriptive Statistics for Coffee Futures

Feature	Mean	SD	Skewness	Kurtosis	Jarque-Bera	ADF
Coffee Futures	132.45	41.23	0.385	-0.672	2,945.6	-2.154
Log-Returns	0.0000823	0.021	0.12	4.82	2,873.4	-14.87

The coffee futures prices exhibited a mean of 132.45 cents per pound with a standard deviation of 41.23. This indicates significant variability characteristic of the soft commodity sector. The log-returns showed a kurtosis of 4.82, which exceeds the normal distribution value of 3. This indicates fat tails. It suggests that extreme price changes occur more frequently in the coffee market than would be predicted by a standard Gaussian model. This validates the need for models that can handle non-normality. The skewness of 0.12 indicates a near-symmetric distribution, although slightly biased towards positive returns. The Jarque-Bera test rejected the null hypothesis of normality (statistic: 2,873.4, $p < 0.01$) with overwhelming significance. Furthermore, the Augmented Dickey-Fuller (ADF) test on log-returns yielded a statistic of -14.87 ($p < 0.01$). This confirms that the return series is stationary. Stationarity is a prerequisite for GARCH-based modeling. It ensures that the statistical properties of the series, such as mean and variance, do not change over time, which allows for stable parameter estimation. The summary statistics for the dataset are presented below.

Realized Volatility

Realized volatility measures the actual variability in coffee futures prices over a specific period. It serves as the ground truth or benchmark against which our predictions are compared. Following the rigorous framework established by Andersen et al. (2003), we utilize realized volatility as it provides a model-free and consistent estimate of ex-post return variation, superior to squared returns. Unlike implied volatility derived from options, realized volatility is a historical measure based on high-frequency or daily data. In this study, it is calculated using daily logarithmic returns over a 21-day window. This corresponds to approximately one trading month (Thomakos & Wang, 2003; Andersen & Bollerslev, 1998).

The realized variance is defined as the average of squared deviations of returns from their 21-day mean, computed over the window from t-20 to t. This formula quantifies volatility as the dispersion of returns. The 21-day window was chosen to balance the need to capture short-term fluctuations while smoothing out daily noise to reveal longer-term trends. In this study, realized variance is the target variable our models aim to predict, and it is calculated as follows, where $N = 21$ is the window size, r_t represents the log-return at time t, and $\bar{r}$ is the mean return over the rolling window.

$$\sigma^2_{RV,t} = \frac{1}{N}\sum_{i=0}^{N-1}\left(r_{t-i} - \bar{r}\right)^2 \tag{1}$$

Forecasting Models

This section introduces the foundational models used to forecast coffee futures volatility. These include traditional statistical approaches, specifically GARCH and GJR-GARCH, and neural network-based methods, namely LSTM and BiLSTM. Each model is designed to capture specific aspects of volatility dynamics.

The GARCH(1,1) model, developed by Bollerslev (1986), is a widely used statistical tool for modeling time-varying volatility in financial time series. It assumes that the current conditional variance depends on three factors: the long-run variance, past squared returns (the ARCH term), and past conditional variance (the GARCH term). The specification is expressed as Equation 2, where $\omega > 0$, $\alpha \geq 0$, $\beta \geq 0$, and $\alpha + \beta < 1$ to ensure stationarity. Here, σ_t^2 is the conditional variance, and ε_{t-1}^2 represents the squared residual from the mean equation.

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

The GJR-GARCH model, proposed by Glosten et al. (1993), extends the GARCH framework to account for asymmetry in volatility responses. This is commonly known as the leverage effect. In many markets, negative returns often increase volatility more than positive returns of the same magnitude. The GJR-GARCH model introduces an indicator function I_{t-1}. This function shown in Equation 3 takes the value of 1 if $\varepsilon_{t-1} < 0$ and 0 otherwise, where γ captures the asymmetric impact of negative shocks. If $\gamma > 0$, it implies that bad news increases volatility more than good news.

$$\sigma_t^2 = \omega + \left(\alpha + \gamma I_{t-1}\right)\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2 \tag{3}$$

Long Short-Term Memory (LSTM) networks, introduced by Hochreiter and Schmidhuber (1997), are designed to overcome the vanishing gradient problem inherent in standard RNNs. The LSTM unit consists of a cell state c_t and three gates: the forget gate, input gate, and output gate. The mathematical formulations are as shown from Equation 4 to Equation 9.

$$f_t = \sigma\left(W_f \cdot \left[h_{t-1}, x_t\right] + b_f\right) \tag{4}$$

$$i_t = \sigma\left(W_i \cdot \left[h_{t-1}, x_t\right] + b_i\right) \tag{5}$$

$$\tilde{c}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \tag{6}$$

$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t \tag{7}$$

$$o_t = \sigma\left(W_o \cdot \left[h_{t-1}, x_t\right] + b_o\right) \tag{8}$$

$$h_t = o_t \odot \tanh\left(c_t\right) \tag{9}$$

where σ denotes the sigmoid activation function, ⊙ represents element-wise multiplication, h_t is the hidden state, and x_t is the input vector.

Bidirectional LSTM (BiLSTM) extends this architecture by training two LSTMs on the input sequence. One trains on the original input sequence, and the other trains on the reversed copy of the input sequence (Schuster & Paliwal, 1997). As shown in Equation 10, this allows the network to incorporate both backward and forward information about the sequence at every time step. The final hidden state H_t is a concatenation of the forward state $\overrightarrow{h_t}$ and the backward state $\overleftarrow{h_t}$.

$$H_t = [\overrightarrow{h_t}, \overleftarrow{h_t}] \tag{10}$$

Loss Functions

To evaluate how well our models forecast realized volatility, we use two loss functions: Heterogeneous Mean Squared Error (HMSE) and Heterogeneous Mean Absolute Error (HMAE). These metrics compare the predicted variance to the realized variance. We selected these robust loss functions consistent with the volatility forecasting literature (Patton, 2011). They penalize forecast errors more heavily when volatility is low. This effectively handles the heteroscedastic nature of the data, whereas standard MSE might be biased by high-volatility outliers.

Equation 11 shows HSME's calculation. The HMSE measures the average squared difference between 1 and the ratio of predicted variance to realized variance. By squaring the deviations, it emphasizes larger errors.

$$HMSE = \frac{1}{T}\sum_{t=1}^{T}\left(1 - \frac{\hat{\sigma}_t^2}{\sigma_{RV,t}^2}\right)^2 \tag{11}$$

The HMAE calculates the average absolute difference between 1 and the ratio of predicted to realized variance. It is less sensitive to extreme outliers than HMSE, as Equation 12 shows, $\hat{\sigma}_t^2$ is the predicted volatility and $\sigma_{RV,t}^2$ is the realized volatility proxy.

$$HMAE = \frac{1}{T}\sum_{t=1}^{T}\left|1 - \frac{\hat{\sigma}_t^2}{\sigma_{RV,t}^2}\right| \tag{12}$$

Proposed Hybrid Models

We propose six hybrid models (as listed in Table 2) that integrate traditional GARCH-family models with LSTM and BiLSTM architectures to forecast coffee futures volatility. The core philosophy of this hybrid approach involves "residual learning" or "feature augmentation." Instead of forcing the neural network to learn the volatility dynamics from scratch using only raw returns, we provide it with the conditional variance estimated by the econometric models. This gives the neural network a high-quality initialization advantage. The models utilize the conditional variance outputs from GARCH and GJR-GARCH. These are combined with logarithmic returns and lagged realized variance to capture both linear and nonlinear volatility dynamics.

Table 2
Proposed Hybrid Models

Model	Input Features	RNN Type	Description
GARCH-LSTM	GARCH parameters (ω, α, β), GARCH conditional variance (σ_t^2), log-returns (r_t), lagged RV	LSTM	Combines GARCH outputs with LSTM to model temporal dependencies in volatility.
GJR-LSTM	GJR-GARCH parameters (ω, α, β, γ), GJR conditional variance (σ_t^2), log-returns (r_t), lagged RV	LSTM	Incorporates GJR-GARCH outputs into LSTM to capture asymmetry and leverage effects.
GG-LSTM	GARCH and GJR-GARCH parameters, both conditional variances, log-returns (r_t), lagged RV	LSTM	Merges GARCH and GJR-GARCH outputs into LSTM for a comprehensive volatility representation.
GARCH-BiLSTM	GARCH parameters (ω, α, β), GARCH conditional variance (σ_t^2), log-returns (r_t), lagged RV	BiLSTM	Uses GARCH outputs with BiLSTM to capture bidirectional temporal dependencies.
GJR-BiLSTM	GJR-GARCH parameters (ω, α, β, γ), GJR conditional variance (σ_t^2), log-returns (r_t), lagged RV	BiLSTM	Integrates GJR-GARCH outputs with BiLSTM to model asymmetry with bidirectional processing.
GG-BiLSTM	GARCH and GJR-GARCH parameters, both conditional variances, log-returns (r_t), lagged RV	BiLSTM	Combines GARCH and GJR-GARCH outputs with BiLSTM for maximum flexibility and bidirectionality.

The GARCH-LSTM model combines GARCH outputs with LSTM to model temporal dependencies in volatility. The GJR-LSTM incorporates GJR-GARCH outputs into LSTM to capture asymmetry and leverage effects. Additionally, the GG-LSTM merges GARCH and GJR-GARCH outputs into LSTM for a comprehensive volatility representation. Expanding on these, the GARCH-BiLSTM uses GARCH outputs with BiLSTM to capture bidirectional temporal dependencies. The GJR-BiLSTM integrates GJR-GARCH outputs with BiLSTM to model asymmetry with bidirectional processing. Finally, the GG-BiLSTM combines GARCH and GJR-GARCH outputs with BiLSTM for maximum flexibility and bidirectionality.

Network Architecture

The LSTM and BiLSTM architectures share a consistent structure (as shown in Figure 1) to ensure comparability across models. Each model consists of three recurrent layers, being either LSTM or BiLSTM, followed by two dense layers. Each recurrent layer contains 100 units. For BiLSTM, this applies to both forward and backward directions, which effectively doubles the capacity per layer. A dropout rate of 0.2 is applied after each recurrent layer. This prevents overfitting by randomly deactivating 20% of the units during training. The recurrent layers feed into a dense layer with 20 units using ReLU activation. This is followed by an output layer with 1 unit using linear activation to predict realized variance. Models are trained using the Huber loss function. This function balances resilience to outliers with smooth gradients, making it suitable for volatility forecasting. The overall framework of the proposed hybrid approach is illustrated in Figure 1, which visualizes the data flow from econometric estimation to the neural network's bidirectional processing layers.

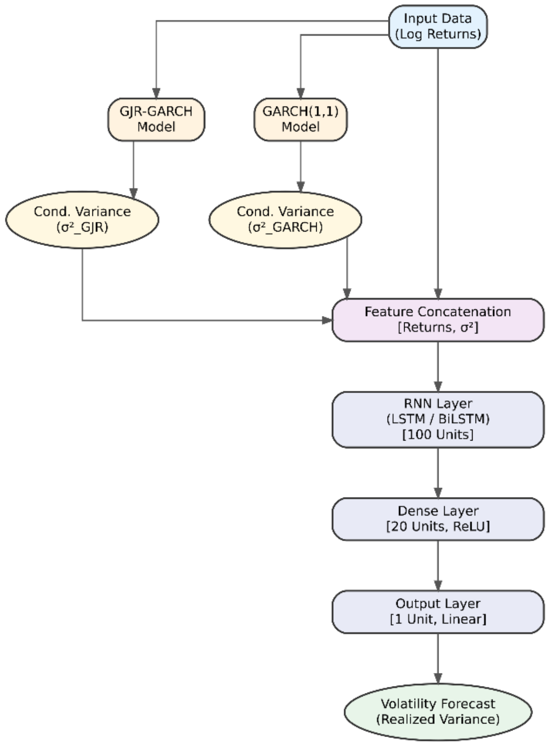

Figure 1 Schematic Diagram of the GARCH-RNN Hybrid Model Architecture

Results and Discussion

The GARCH and GJR-GARCH models were fitted to the training period (2000–2017) to establish the econometric foundation. The estimated parameters reveal significant structural characteristics of the coffee futures market.

First, the visual inspection of the time series confirms the necessity for non-linear modeling. As shown in Figure 2, the daily returns (Panel b) exhibit distinct volatility clustering, where periods of high variability are followed by high variability. Furthermore, the realized volatility (Panel d) demonstrates sharp spikes during stress periods, which traditional linear models often fail to capture fully.

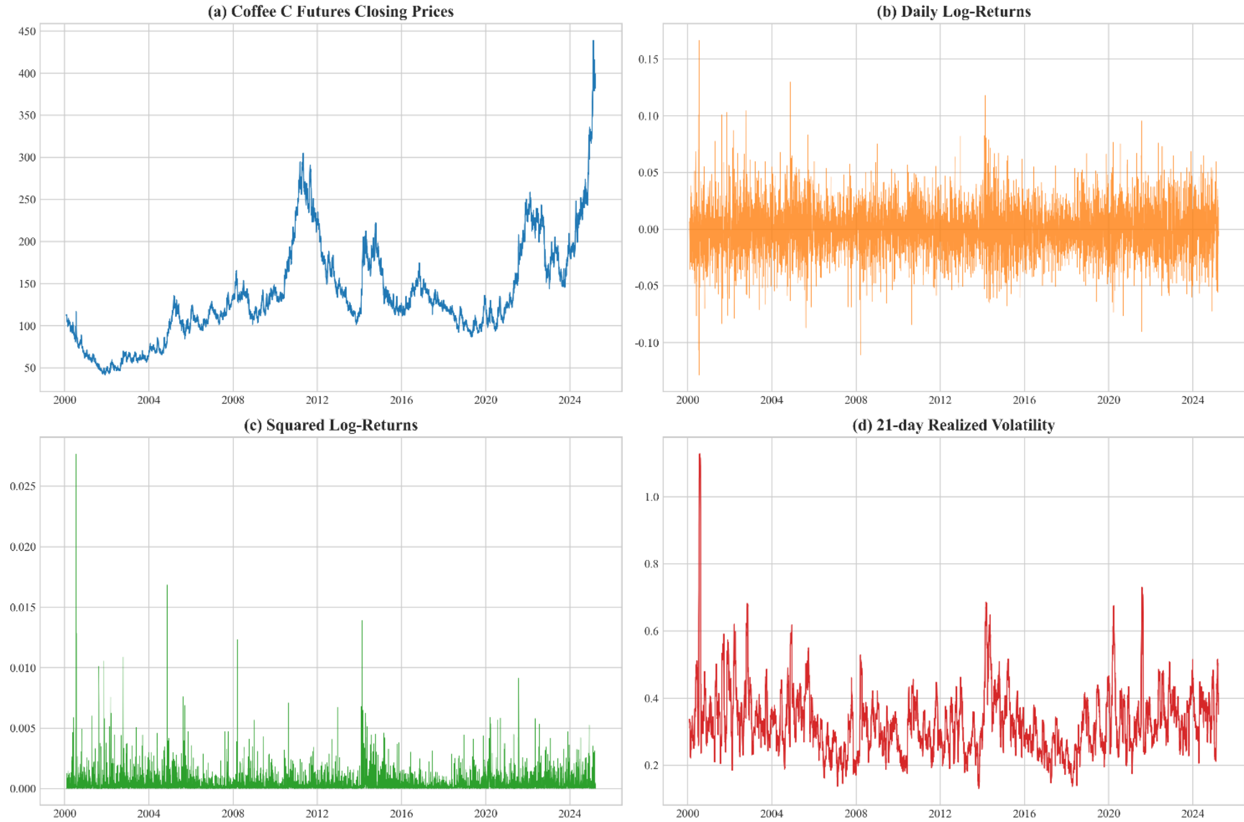

Figure 2 Visualization of Time Series of Coffee Futures Prices, Returns, and Realized Volatility
Note: (Panel a) shows the changes of coffee futures price (closing prices) of the entire sample; (panel b) illustrates the daily log-returns in the same period; (panel c) shows the squared log-returns; (Panel d) shows the 21-days realized volatility.

To further justify the model selection, we examined the distributional properties of the returns. The diagnostic plots in Figure 3 reveal substantial deviations from normality. Specifically, the Q-Q plot (Panel b) shows that the empirical quantiles (blue dots) diverge significantly from the theoretical normal line at the tails, confirming the presence of leptokurtosis (fat tails) in coffee returns.

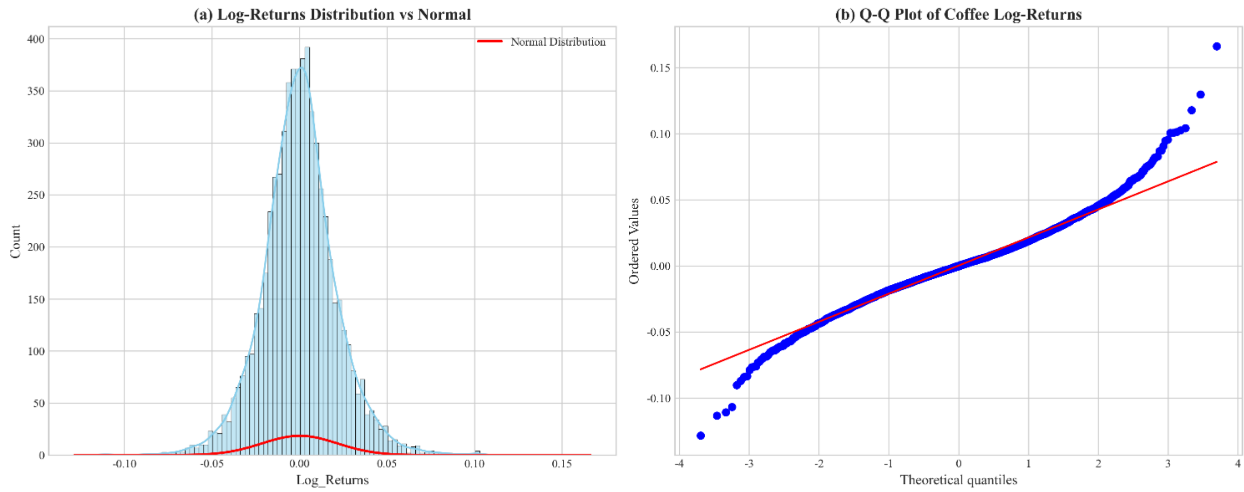

Figure 3 Density and Q-Q Plot for Normality Diagnostics

Note: (a) shows the log-returns distribution vs normal distribution; (b) shows a Q-Q plot of the log-returns of coffee futures

Table 3 presents the estimated parameters for the standard GARCH model. The persistence parameter beta was estimated at 0.9257 ($p < 0.001$), indicating that volatility shocks decay very slowly.

Table 3
Estimated GARCH Parameters (March 2, 2000 – December 29, 2017)

Parameter	Estimate	SE	t-statistic	P-value
ω	0.000004	0.000006	0.6667	0.5051
α	0.062341	0.027654	2.2541	0.0242
β	0.925678	0.029876	30.978	0.0000

In the GJR-GARCH model, the asymmetry parameter gamma was estimated at 0.0643 ($p = 0.003$), as detailed in Table 4. This statistically significant positive coefficient confirms the “leverage effect,” implying that price drops induce higher subsequent volatility than price spikes of the same magnitude.

Table 4
Estimated GJR-GARCH Parameters (March 2, 2000 – December 29, 2017)

Parameter	Estimate	SE	t-statistic	P-value
ω	0.000003	0.000005	0.6000	0.5488
α	0.025678	0.022345	1.1493	0.2504
β	0.938765	0.032123	29.221	0.0000
γ	0.064321	0.021987	2.9254	0.0034

Comparative Performance of Hybrid Architectures
We evaluated the hybrid models against out-of-sample realized volatility to assess their predictive accuracy. The empirical results demonstrate a clear hierarchy in model performance. A detailed comparison of the out-of-sample forecasting performance is provided in Table 5.

Table 5
Out-of-Sample Forecasting Performance (HMSE and HMAE) for Hybrid Models

Model	7 days ahead		14 days ahead		21 days ahead		28 days ahead	
	HMSE	HMAE	HMSE	HMAE	HMSE	HMAE	HMSE	HMAE
Window = 14								
GARCH-LSTM	0.1949	0.3533	0.1972	0.3573	0.1897	0.3461	0.1549	0.3118
GJR-LSTM	0.1946	0.3557	0.1918	0.3543	0.1908	0.3479	0.1461	0.3009
GG-LSTM	0.0745	0.2084	0.1452	0.3061	0.0707	0.2134	0.0676	0.1951
GARCH-BiLSTM	0.3658	0.4233	0.1862	0.3049	0.0905	0.2194	0.4138	0.4540
GJR-BiLSTM	0.0677	0.1849	0.0886	0.2075	0.0455	0.1523	0.0796	0.2081
GG-BiLSTM	0.0541	0.1707	0.0689	0.1888	0.0392	0.1484	0.0712	0.1965
Window = 21								
GARCH-LSTM	0.1972	0.3536	0.1898	0.3475	0.1622	0.3146	0.1544	0.3079

GJR-LSTM	0.2008	0.3623	0.2426	0.3913	0.0690	0.1991	0.1521	0.3032
GG-LSTM	0.0729	0.2084	0.0531	0.1771	0.0296	0.1277	0.1093	0.2508
GARCH-BiLSTM	0.1882	0.3110	0.0861	0.2094	0.4783	0.4814	0.4004	0.4484
GJR-BiLSTM	0.0795	0.1948	0.0567	0.1633	0.0167	0.0942	0.1298	0.2528
GG-BiLSTM	0.0502	0.1640	0.0392	0.1486	0.0127	0.0786	0.0812	0.2037
Window = 28								
GARCH-LSTM	0.1936	0.3521	0.1990	0.3579	0.1526	0.3051	0.1931	0.3523
GJR-LSTM	0.1841	0.3457	0.1919	0.3518	0.1394	0.2942	0.1645	0.3204
GG-LSTM	0.0299	0.1312	0.0168	0.0961	0.0240	0.1169	0.0815	0.2150
GARCH-BiLSTM	0.1268	0.2505	0.0203	0.1035	0.4283	0.4617	0.4499	0.4695
GJR-BiLSTM	0.0490	0.1603	0.0396	0.1514	0.0200	0.1062	0.1627	0.2758
GG-BiLSTM	0.0443	0.1506	0.0112	0.0767	0.0533	0.1700	0.0799	0.2040

Note: Out-of-Sample covers January 2, 2018 to March 21, 2025.

To facilitate a clearer understanding of the performance hierarchy, we present a comparative bar chart in Figure 4, which contrasts the HMSE of all models under the 21-day forecasting horizon. The empirical results demonstrate a distinct stepwise improvement in predictive accuracy as model complexity increases.

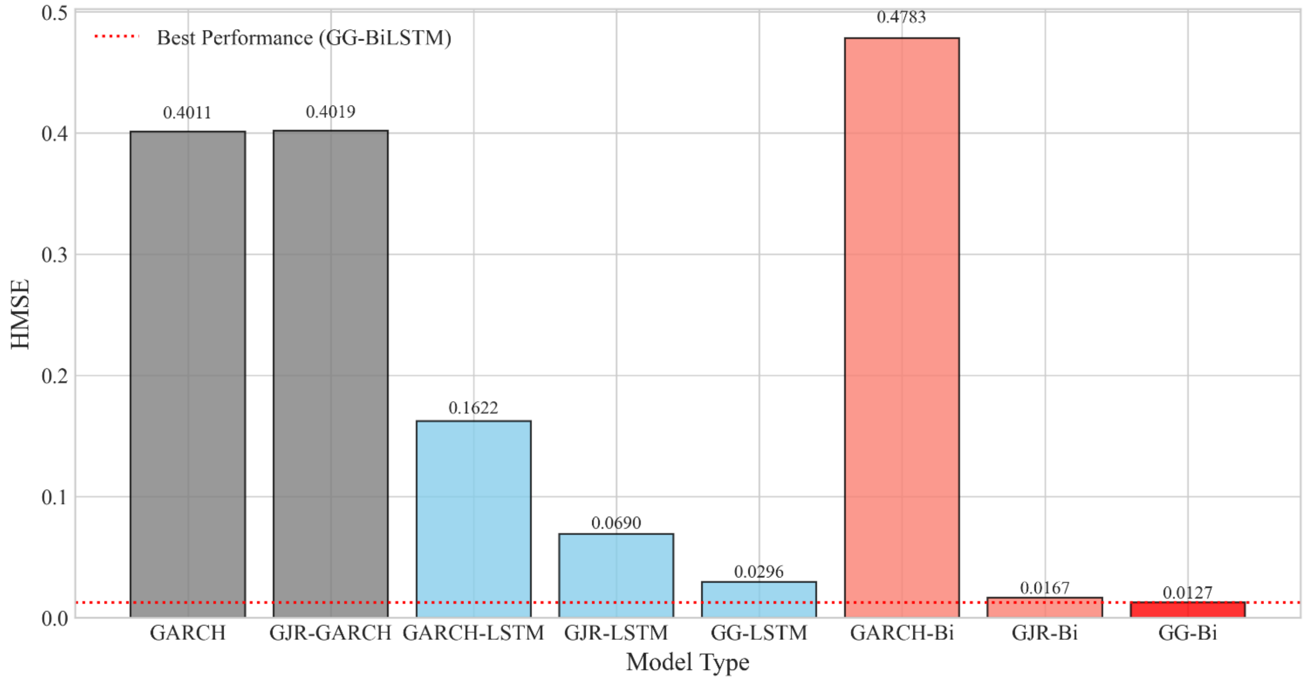

Figure 4 Performance Hierarchy of All Models (21-day Horizon)

As visually summarized in Figure 4, the baseline econometric models (gray bars) exhibit the highest errors, followed by the unidirectional LSTM hybrids (blue bars). Crucially, the bidirectional hybrids (red bars), particularly the GG-BiLSTM, achieve the lowest error rates (HMSE = 0.0127). This "ladder of improvement" confirms that each layer of methodological enhancement—from introducing neural networks to incorporating bidirectional processing—yields a tangible gain in forecasting precision.

To facilitate a holistic view of the forecasting capabilities, Figure 5 presents a panoramic comparison of the predicted volatility trajectories against the actual realized volatility over the entire out-of-sample period. It is evident that the proposed GG-BiLSTM model (red line) tracks the actual volatility (black line) more closely than the GARCH benchmark (gray dashed line), particularly during the volatile periods of 2020–2022.

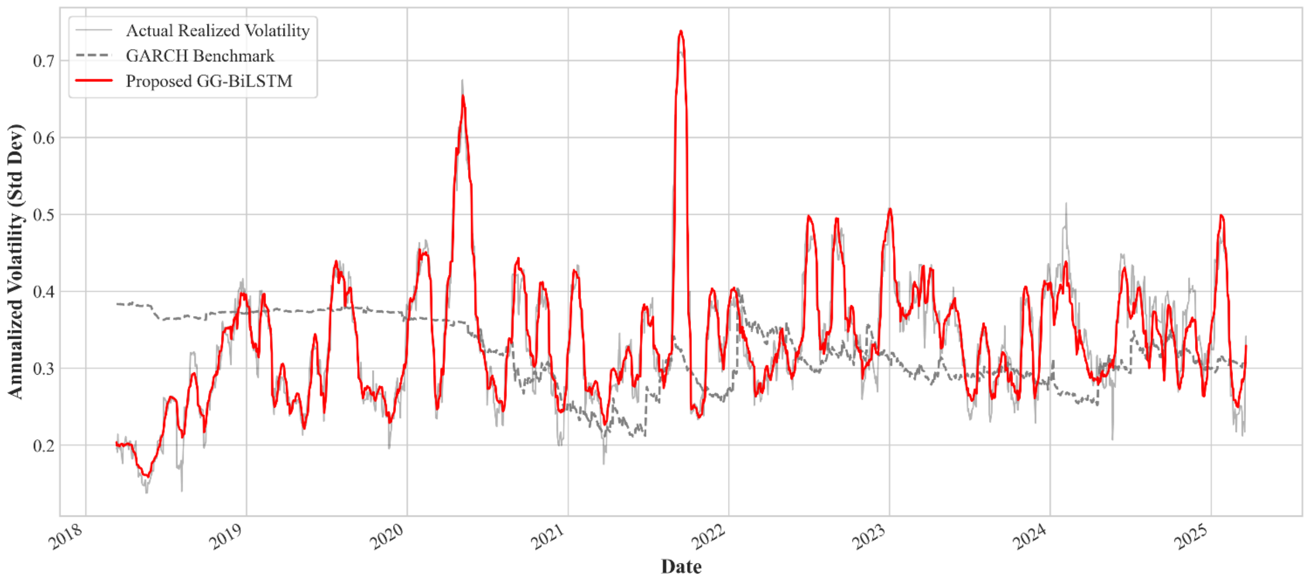

Figure 5 Volatility Forecast Comparison

To statistically validate these observations, we employed the MCS test (Hansen et al., 2011). As summarized in Table 6, the GG-BiLSTM model consistently remains in the superior set of models (p=1.000) across most forecasting horizons, confirming its statistical robustness.

Table 6
Model Confidence Set (MCS) for Hybrid Models

Model Family	Rolling window	Days ahead	HMSE	HMAE	P-value
GARCH-LSTM	14	14	0.0511	0.1700	1.000
GARCH-LSTM	21	14	0.0262	0.1165	0.143
GARCH-LSTM	28	7	0.0304	0.1230	1.000
GARCH-LSTM	28	14	0.0151	0.0867	1.000
GARCH-BiLSTM	14	7	0.0638	0.1822	0.117
GARCH-BiLSTM	14	14	0.0522	0.1665	0.547
GARCH-BiLSTM	14	28	0.0786	0.2069	0.584
GARCH-BiLSTM	21	7	0.0287	0.1287	1.000
GARCH-BiLSTM	21	21	0.0131	0.0828	1.000
GARCH-BiLSTM	21	28	0.0783	0.2047	1.000
GARCH-BiLSTM	28	21	0.0147	0.0852	1.000
GJR-BiLSTM	14	21	0.0377	0.1437	1.000
GJR-BiLSTM	21	14	0.0243	0.1124	1.000
GJR-BiLSTM	28	28	0.0720	0.1959	1.000
GG-BiLSTM	14	7	0.0610	0.1785	1.000
GG-BiLSTM	14	28	0.0775	0.2111	1.000

Notes: Only models in the MCS (P-value >= 0.05, 5% significance level) are reported. HMAE: Heteroscedasticity-adjusted mean absolute error; HMSE: Heteroscedasticity-adjusted mean squared error. Bold highlights the best-performing configuration for each model in the MCS (lowest HMSE).

Influence of Rolling Window Size

The choice of the rolling window size proved to be a critical hyperparameter. To systematically visualize the impact of this variable, we generated a performance heatmap, as shown in Figure 6.

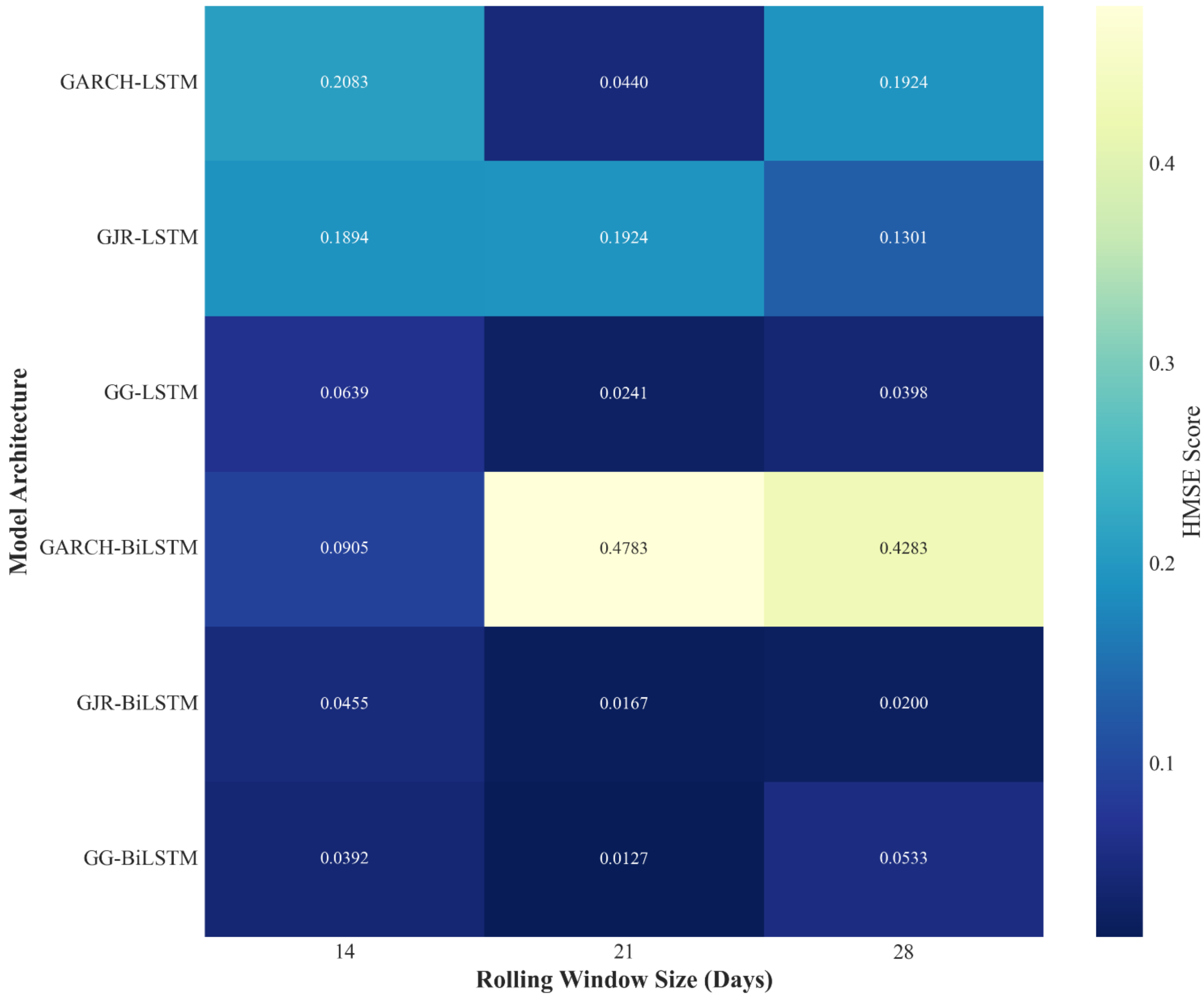

Figure 6 Heatmap of HMSE Performance

The heatmap in Figure 6 indicates that the 21-day rolling window yields the optimal forecasting performance. The HMSE values are notably lower in the 21-day window column, particularly for the BiLSTM-based models. This empirical evidence suggests that a one-month lookback period offers an effective balance between historical context and responsiveness to recent regime changes. Conversely, the higher HMSE values in the 14-day and 28-day columns imply that incorporating insufficient or excessive historical data may compromise model performance.

Dynamic Response to Market Shocks

The true test of a volatility model lies in its ability to adapt to extreme market events. We zoomed in on the volatility surge observed in early 2024 to compare the responsiveness of different architectures.

Figure 7 illustrates this localized performance. While the symmetric GARCH-LSTM (green dashed line) exhibits a noticeable lag and fails to reach the peak of the volatility spike, the GG-BiLSTM (red solid line) adapts almost instantaneously. This superior responsiveness is attributed to the bidirectional processing capability, which allows the model to better infer the intensity of the shock from the immediate temporal context.

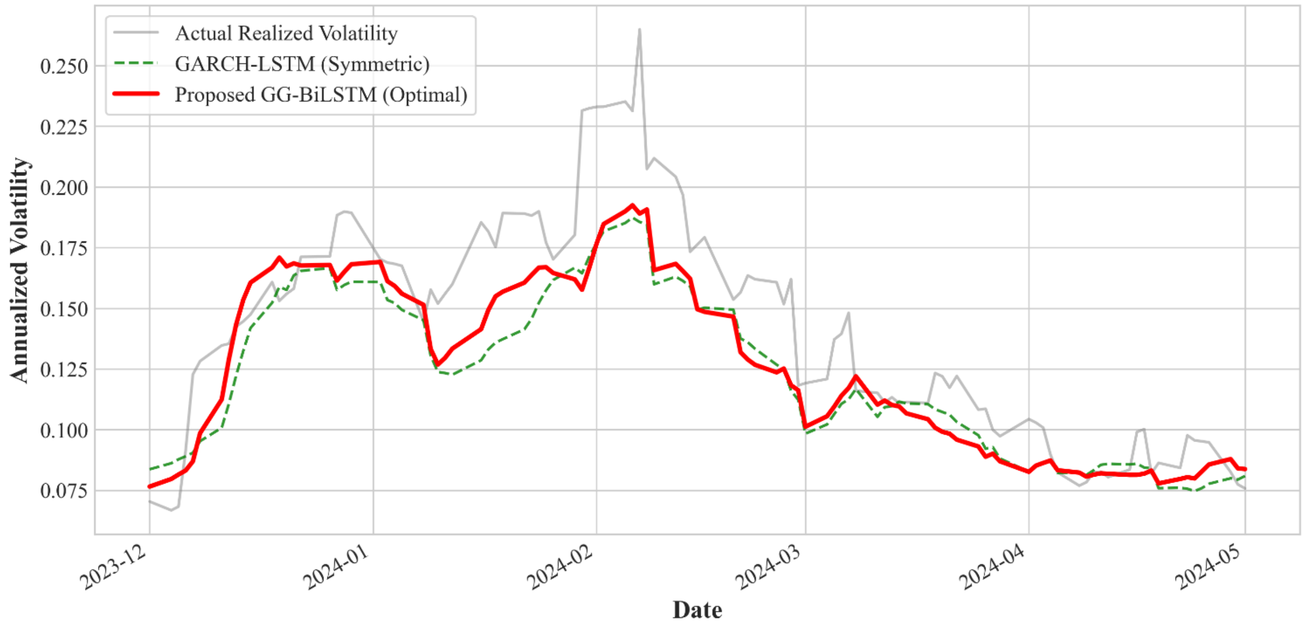

Figure 7 GARCH-LSTM and GG-BiLSTM's Performance during 2024 Market Shock

Forecasting Horizon Effects

Forecast accuracy naturally fluctuates as the forecasting horizon extends from 7 to 28 days. While conventional wisdom suggests that uncertainty accumulates linearly over time, our empirical results reveal a more nuanced pattern for hybrid architectures. This non-linear relationship is visually captured in Figure 8, which tracks the HMSE growth across different horizons.

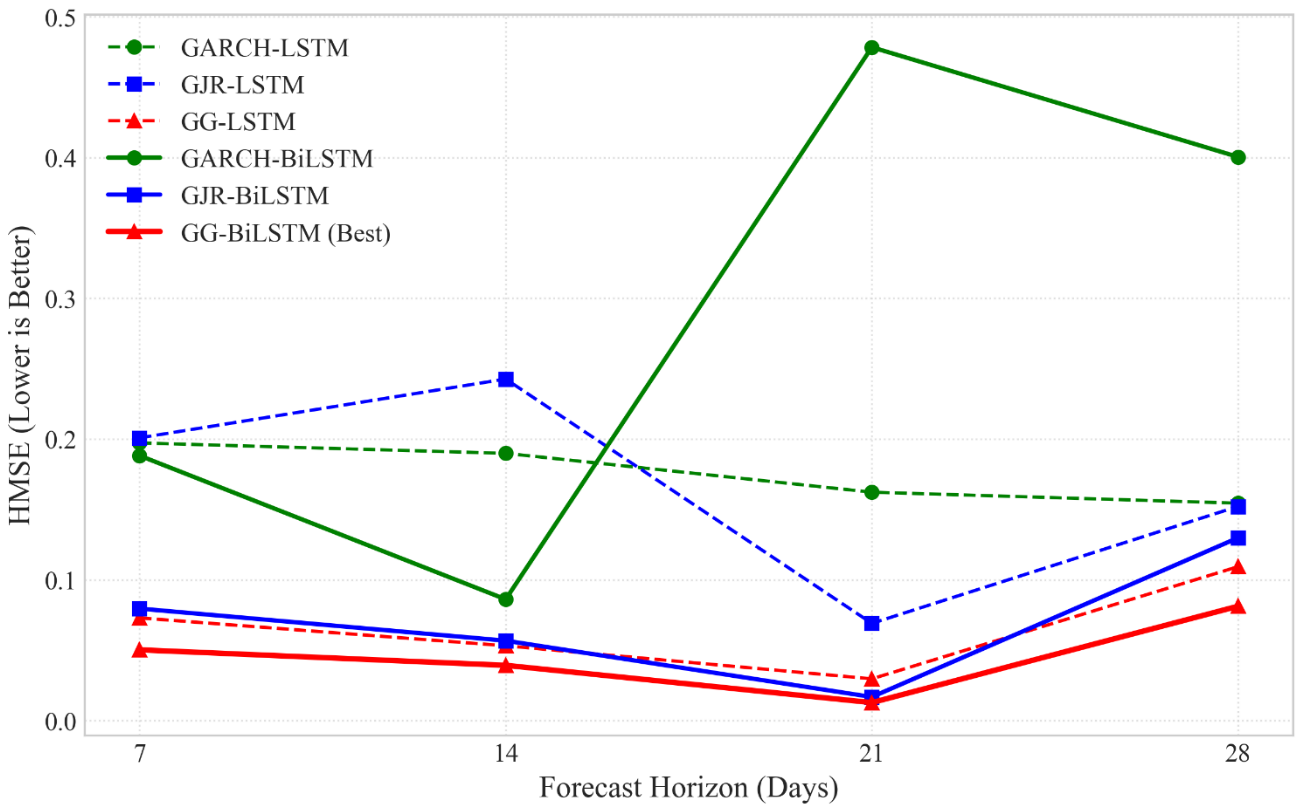

Figure 8 Error Growth across Forecasting Horizons

As shown in Figure 8, the optimal GG-BiLSTM model (solid red line) does not follow a simple upward trend. Instead, it exhibits a distinct U-shaped trajectory, reaching its global minimum error at the 21-day horizon. This implies a resonance between the model's input window (21 days) and the forecasting target, allowing it to maximize predictive power at this specific "sweet spot." In contrast, the unidirectional LSTM benchmarks (dashed lines) display higher and relatively stagnant error levels, failing to capitalize on this temporal alignment. Although the error for GG-BiLSTM eventually rises at the 28-day horizon due to increasing uncertainty, it maintains a significantly lower floor compared to its counterparts, demonstrating superior robustness in medium-term forecasting.

Model Interpretability and Reliability

To address the "black box" criticism of neural networks, we utilized SHAP (SHapley Additive exPlanations) values to quantify feature importance. The results are summarized in Figure 9.

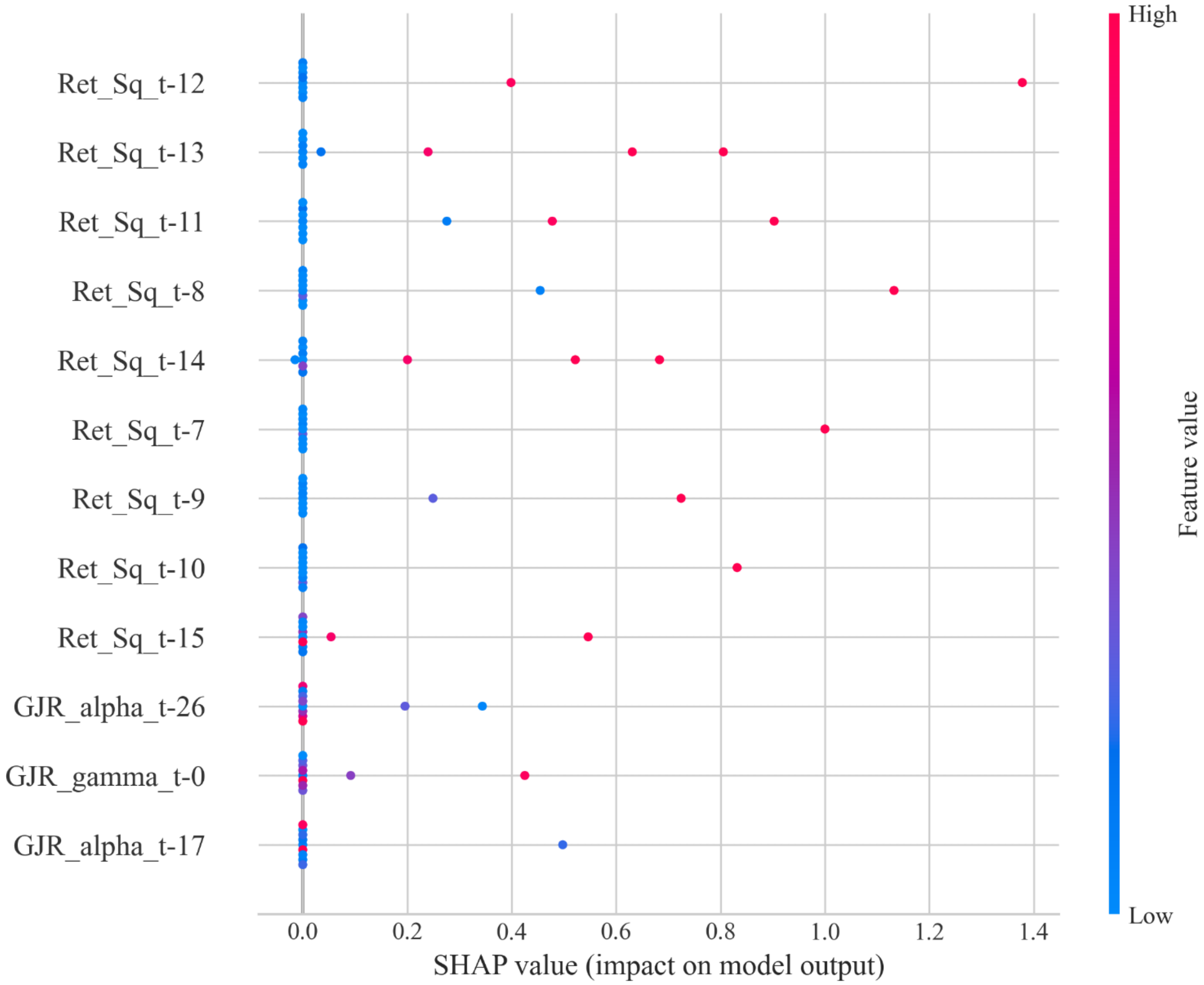

Figure 9 SHAP Summary Plot

As shown in Figure 9, the features G_alpha (GARCH error term) and Ret_Sq (Squared Returns) appear at the top of the importance ranking. This confirms that the model's decision-making process is grounded in sound financial logic: it relies heavily on immediate market shocks and structural volatility persistence to generate forecasts.

Finally, to verify the reliability of our optimal model, Figure 10 provides a detailed portrait of the GG-BiLSTM model’s performance in the 21-day window configuration, along with its residual analysis.

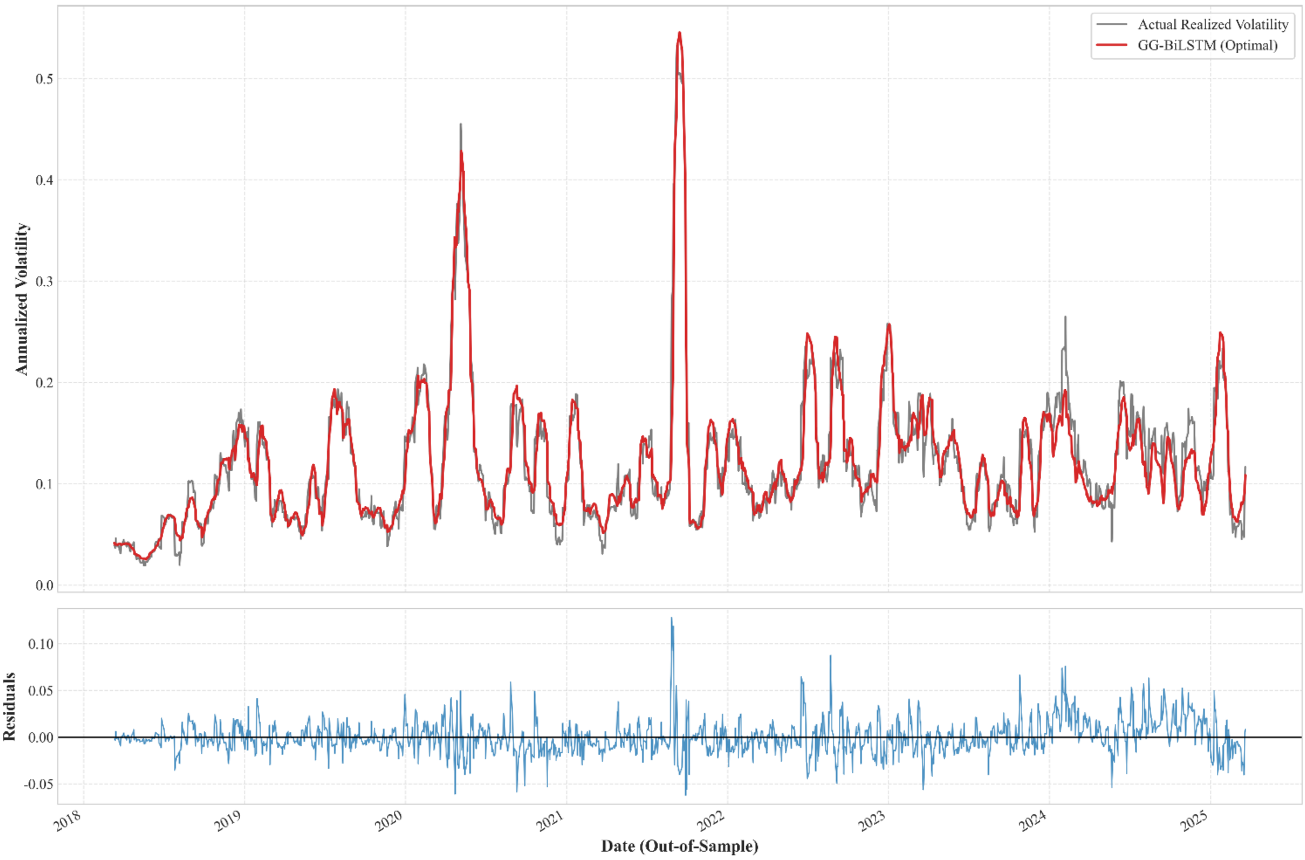

Figure 10 The Optimal Model Residuals

The lower panel of Figure 10 shows the forecast residuals (calculated by the actual minus the predicted). The residuals oscillate randomly around the zero line without exhibiting any prolonged systematic bias or autocorrelation. This “white noise” behavior indicates that the GG-BiLSTM model has successfully extracted the majority of the predictable non-linear patterns from the coffee futures data.

Conclusion

Forecasting volatility of Arabica coffee futures is vital for a wide range of stakeholders. This includes producers, processors, traders, and investors. This is given the commodity’s central role in global agricultural markets and its vulnerability to supply-side disruptions. This study advances volatility forecasting by integrating econometric models with neural network architectures through a GARCH-RNN hybrid approach. This surpasses the limitations of classical methods. Specifically, we combine GARCH and GJR-GARCH models with LSTM and Bidirectional LSTM (BiLSTM) networks to predict coffee price volatility. We evaluated their performance across rolling windows of 14, 21, and 28 days and multiple forecast horizons.

Our findings reveal that hybrid GARCH-RNN models, particularly those incorporating BiLSTM, significantly outperform standalone econometric approaches. By utilizing GARCH and GJR-GARCH parameters as inputs, these hybrids adeptly capture temporal dependencies and asymmetric volatility patterns. The bidirectional learning of BiLSTM offering a distinct advantage over traditional LSTM in modeling the complex dynamics of coffee futures. Among the six variants tested, the GG-BiLSTM configuration consistently demonstrates superior stability.

This research contributes to the volatility forecasting literature by directly addressing the critique of deep learning models as "black boxes" in finance (Černevičienė & Kabašinskas, 2024). We improve both model performance and interpretability by explicitly embedding econometric parameters into the neural network workflow. Examples include the leverage effect coefficient gamma from GJR-GARCH. Our approach demonstrates that the "hybrid" path is not merely an ensemble technique. Rather, it is a necessary theoretical bridge that combines the structural guidance of economic theory with the computational power of modern AI. Additionally, the confirmation of the optimal 21-day lookback window provides actionable guidance for practitioners. It validates the "one-month memory" hypothesis in soft commodity markets (Feng & Zhang, 2025).

The implications of this work extend across both the coffee industry and quantitative finance. Historically, dynamic hedging strategies for coffee relied on standard GARCH frameworks to estimate optimal hedge ratios (Choudhry, 2009). However, our findings align with the modern paradigm shift described by Hu and Ni (2024) and Huang and Sayed (2025). In this shift, deep learning-integrated approaches provide superior risk-adjusted returns by capturing non-linear market moves that traditional linear hedges miss. Accurate volatility forecasting is a direct driver of economic value. Hu and Ni (2024) show that implementing deep learning-based hedging strategies can generate significant annualized economic benefits for commodity firms. This is achieved by optimizing inventory management against price erraticism. Similarly, Vancsura et al. (2023) highlight that superior forecasting models are value-adding factors in Enterprise Risk Management (ERM). They allow firms to reduce profit volatility at a lower cost than traditional blanket hedging. For coffee growers, processors, and traders, improved volatility forecasts strengthen risk management strategies. This enables better hedging against price fluctuations driven by weather, supply shocks, or market shifts. For financial practitioners, these hybrid models provide a refined toolset to anticipate market behavior. This supports algorithmic trading and investment decisions in coffee futures. While the models exhibit stability across diverse conditions, future research could explore computational optimizations for BiLSTM and assess performance over extended horizons to broaden their practical utility.

Contributions

This research makes distinct contributions to both the academic literature and the practical context of commodity risk management, bridging theoretical econometrics, computational finance, and development economics.

Theoretical Contribution: Advancing Hybrid Volatility Forecasting Methodology

This study moves beyond simple model ensembles to propose a deeply integrated hybrid framework. By directly feeding the structural parameters and conditional variance estimates from GARCH-family models into a BiLSTM network, we operationalize the "residual learning" paradigm (Zhang, 2003) in a novel way. This approach provides the neural network with a theoretically-grounded prior, effectively scaffolding its learning process. Our findings demonstrate that this integration is not merely additive but synergistic: the BiLSTM, informed by econometric parameters, specializes in modeling the complex, non-linear residuals that pure econometric models miss, while the econometric foundation prevents the network from overfitting to spurious noise. This contributes a validated architectural blueprint for hybrid forecasting that prioritizes interpretability without sacrificing predictive power, directly responding to the "black box" critique prevalent in AI-for-finance literature (Černevičienė & Kabašinskas, 2024).

Contextual Contribution: Informing Risk Management in a Vulnerable Global Value Chain
The significance of this work is deeply embedded in the socioeconomic reality of the global coffee industry. Coffee is more than a financial asset; it is a primary source of income for over 25 million farming families worldwide, many of whom operate at the margins of subsistence. Our identification of the GJR-BiLSTM hybrid as a superior forecasting tool has direct implications for risk management justice. More accurate forecasts can lead to more efficient hedging instruments and pricing of insurance products, potentially lowering the cost of risk transfer for producers (Hu & Ni, 2024). By demonstrating the critical importance of modeling asymmetry (the leverage effect), our research highlights that risk models failing to account for the greater impact of price drops systematically underestimate the vulnerability of producers to market downturns. Thus, this work provides a technical foundation for developing more equitable financial tools that are sensitive to the asymmetric risks borne by the most vulnerable nodes in the supply chain.

Bridging Disciplinary Divides
Finally, this study plays a role in fostering interdisciplinary dialogue. It demonstrates how tools from computer science (BiLSTM) can be rigorously informed by economic theory (GARCH, leverage effect) to address a problem with profound social science implications (livelihood resilience, supply chain stability). In doing so, it answers calls for more context-aware and socially responsible quantitative finance research, showing that methodological innovation can and should be directed towards markets where its impact on human welfare is most acute.

References

Abaidoo, R., & Agyapong, E. A. (2022). Commodity price fluctuations and development: Perspective from emerging economies. *Journal of Financial Economic Policy*, *14*(3), 333–355. https://doi.org/10.1108/JFEP-02-2021-0039

Aitbekov, B. (2025). Long Short-Term Memory (LSTM) networks for time series forecasting: A comprehensive survey of model, applications, and challenges. *Central Asian Scientific Journal*, *26*(2), 53-59.

Andersen, T. G., & Bollerslev, T. (1998). Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International Economic Review*, *39*(4), 885-905. https://doi.org/10.2307/2527343

Andersen, T. G., Bollerslev, T., Diebold, F. X., & Labys, P. (2003). Modeling and forecasting realized volatility. *Econometrica*, *71*(2), 579–625. https://doi.org/10.1111/1468-0262.00418

Bollerslev, T. (1986). Generalized autoregressive conditional heteroscedasticity. *Journal of Econometrics*, *31*(3), 307–327. https://doi.org/10.1016/0304-4076(86)90063-1

Carpantier, J.-F. (2010). *Commodities inventory effect* (LIDAM Discussion Papers CORE No. 2010040). Université catholique de Louvain, Center for Operations Research and Econometrics (CORE). https://EconPapers.repec.org/RePEc:cor:louvco:2010040

Choudhry, T. (2009). Short-run deviations and time-varying hedge ratios: Evidence from agricultural futures markets. *International Review of Financial Analysis*, *18*(1–2), 58–65. https://doi.org/10.1016/j.irfa.2008.11.003

Chung, V., Espinoza, J., & Quispe, R. (2025). Forecasting Financial Volatility Under Structural Breaks: A Comparative Study of GARCH Models and Deep Learning Techniques. *Journal of Risk and Financial Management*, *18*(9), 494. https://doi.org/10.3390/jrfm18090494

Černevičienė, J., & Kabašinskas, A. (2024). Explainable artificial intelligence (XAI) in finance: A systematic literature review. *Artificial Intelligence Review*, *57*(8), 216. https://doi.org/10.1007/s10462-024-10854-8

Feng, Y., & Zhang, Y. (2025). Forecasting realized volatility: The choice of window size. *Journal of Forecasting*, *44*(2), 692–705. https://doi.org/10.1002/for.3221

Giantsidi, S., & Tarantola, C. (2025). Deep learning for financial forecasting: A review of recent trends. *International Review of Economics & Finance*, *104*, 104719. https://doi.org/10.1016/j.iref.2025.104719

Glosten, L. R., Jagannathan, R., & Runkle, D. E. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *The Journal of Finance*, *48*(5), 1779–1801. https://doi.org/10.1111/j.1540-6261.1993.tb05128.x

Hansen, P. R., Lunde, A., & Nason, J. M. (2011). The model confidence set. *Econometrica*, *79*(2), 453–497. https://doi.org/10.3982/ECTA5771

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, *9*(8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735

Hu, Y., Ni, J., & Wen, L. (2020). A hybrid deep learning approach by integrating LSTM-ANN networks with GARCH model for copper price volatility prediction. *Physica A: Statistical Mechanics and Its Applications*, *557*, 124907. https://doi.org/10.1016/j.physa.2020.124907

Hu, Y., & Ni, J. (2024). A deep learning-based financial hedging approach for the effective management of commodity risks. *Journal of Futures Markets*. *44*(6), 879-900. https://doi.org/10.1002/fut.22497

Huang, C. S., & Sayed, A. (2025). Novel forecasting of white maize futures volatility: A hybrid GARCH-based bi-directional LSTM model. *Cogent Economics & Finance*, *13*(1), Article 2484422. https://doi.org/10.1080/23322039.2025.2484422

Huang, L., Abrahams, A., & Ractham, P. (2022). Enhanced financial fraud detection using cost-sensitive cascade forest with missing value imputation. *Intelligent Systems in Accounting, Finance and Management*, *29*(3), 133–155. https://doi.org/10.1002/isaf.1517

Jaiswal, R., Jha, G. K., Kumar, R. R., & Choudhary, K. (2022). Deep long short-term memory based model for agricultural price forecasting. *Neural Computing and Applications*, *34*(6), 4661–4676. https://doi.org/10.1007/s00521-021-06621-3

Krichen, M., & Mihoub, A. (2025). Long short-term memory networks: A comprehensive survey. *AI*, *6*(9), 215. https://doi.org/10.3390/ai6090215

Le, N.-B., Seo, Y.-S., & Huh, J.-H. (2024). AgTech: Volatility prediction for agricultural commodity exchange trading applied deep learning. *IEEE Access: Practical Innovations, Open Solutions*, *12*, 153898–153910. https://doi.org/10.1109/ACCESS.2024.3479868

Manogna, R. L., Dharmaji, V., & Sarang, S. (2025). Enhancing agricultural commodity price forecasting with deep learning. *Scientific Reports*, *15*(1), 20903. https://doi.org/10.1038/s41598-025-05103-z

Mansilla-Lopez, J., Mauricio, D., & Narváez, A. (2025). Factors, forecasts, and simulations of volatility in the stock market using machine learning. *Journal of Risk and Financial Management*, *18*(5), 227. https://doi.org/10.3390/jrfm18050227

Massrie, K. D. (2025). Why is the price of coffee rising globally? Future prospects for Ethiopian coffee. *Frontiers in Sustainable Food Systems*, *Volume 9-2025*. https://doi.org/10.3389/fsufs.2025.1545168

Nguyen, T. T. H. (2023). *Improving and optimizing the performance of the supply chain: The case of coffee production in Vietnam* [PhD Thesis, Université Polytechnique Hauts-de-France]. https://theses.hal.science/tel-04265675

Patton, A. J. (2011). Volatility forecast comparison using imperfect volatility proxies. *Journal of Econometrics*, *160*(1), 246–256. https://doi.org/10.1016/j.jeconom.2010.03.034

Rabbani, A. H., Arshad, I., Jamal, D. S., Ali, M., & Ali, D. O. (2025). Financial risk forecasting using AI: Hybrid econometric–machine learning approaches. *The Critical Review of Social Sciences Studies*, *3*(4), 795–809. https://doi.org/10.59075/smv2vj25

Schuster, M., & Paliwal, K. K. (1997). Bidirectional recurrent neural networks. *IEEE Transactions on Signal Processing*, *45*(11), 2673–2681. https://doi.org/10.1109/78.650093

Singhal, S., & Tarp, F. (2024). Commodity price volatility and the psychological well-being of farmers. *American Journal of Agricultural Economics*, *107*(1), 269–289. https://doi.org/10.1111/ajae.12468

Singhal, S., & Tarp, F. (2025). Commodity price volatility and the psychological well-being of farmers. *American Journal of Agricultural Economics*, *107*(1), 269–289. https://doi.org/10.1111/ajae.12468

Thomakos, D., & Wang, T. (2003). Realized volatility in the futures markets. *Journal of Empirical Finance*, *10*(3), 321–353. https://doi.org/10.1016/S0927-5398(02)00052-X

Theofilou, A., Nastis, S. A., Michailidis, A., Bournaris, T., & Mattas, K. (2025). Predicting prices of staple crops using machine learning: A systematic review of studies on wheat, corn, and rice. *Sustainability*, *17*(12), 5456. https://doi.org/10.3390/su17125456

Tumpa, S. N., & Maduranga, K. D. G. (2024). *Utilizing RNN for real-time cryptocurrency price prediction and trading strategy optimization* (No. arXiv:2411.05829). arXiv. https://doi.org/10.48550/arXiv.2411.05829

Ubilava, D. (2012). El Niño, La Niña, and world coffee price dynamics. *Agricultural Economics*, *43*(1), 17–26. https://doi.org/10.1111/j.1574-0862.2011.00562.x

Vancsura, L., Tatay, T., & Bareith, T. (2023). Evaluating the Effectiveness of Modern Forecasting Models in Predicting Commodity Futures Prices in Volatile Economic Times. *Risks*, *11*(2), 27. https://doi.org/10.3390/risks11020027

Vancsura, L., Tatay, T., & Bareith, T. (2025). Navigating AI-driven financial forecasting: A systematic review of current status and critical research gaps. *Forecasting*, *7*(3), 36. https://doi.org/10.3390/forecast7030036

Wang, Y., Liu, P., Zhu, K., Liu, L., Zhang, Y., & Xu, G. (2022). A Garlic-Price-Prediction Approach Based on Combined LSTM and GARCH-Family Model. *Applied Sciences*, *12*(22), 11366. https://doi.org/10.3390/app122211366

Wanzala, R. W., & Obokoh, L. O. (2025). Global competitiveness of Kenyan coffee: The price volatility perspective. *Cogent Economics & Finance*, *13*(1), 2499019. https://doi.org/10.1080/23322039.2025.2499019

Zhang, G. P. (2003). Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing*, *50*, 159–175. https://doi.org/10.1016/S0925-2312(01)00702-0