Journal of Mathematics and Modeling in Finance (JMMF)
Vol. 5, No. 2, Summer & Fall 2025
Research paper

A Hybrid Lstm Neural Network Approach for Modeling Periodical Long-Memory Characteristics in Financial Energy Index Time Series

Minou Yari[1], Mohammad Reza Salehi Rad[2], Mohammad Bahrani[3]

[1] Faculty of Statistics, Mathematics and Computer, Allameh Tabataba'i University, Tehran, Iran
minouyari@atu.ac.ir

[2] Faculty of Statistics, Mathematics and Computer, Allameh Tabataba'i University, Tehran, Iran
salehirad@atu.ac.ir

[3] Faculty of Statistics, Mathematics and Computer, Allameh Tabataba'i University, Tehran, Iran
bahrani@atu.ac.ir

Abstract:
Forecasting financial market volatility has always been a major challenge in economics and financial engineering. In this study, a hybrid approach based on FIGARCH and PLM-GARCH models combined with Long Short-Term Memory (LSTM) neural networks is proposed for modeling financial time series. The analyzed dataset of the Iran energy index covers October 30, 2016, to January 25, 2023 with 1396 observations. The PLM-GARCH model is capable of identifying long-term dependencies and periodic structures in the conditional variance of time series, while the LSTM network improves prediction accuracy by learning complex and nonlinear patterns. In this approach, the PLM-GARCH model is first used to estimate volatility, and then the residuals from the model are fed as inputs into the LSTM network to extract nonlinear behaviors. Experimental results showed that the combined PLM-GARCH-LSTM model (RMSE = 0.00209, MAPE $\approx$ 5.1%) outperforms the FIGARCH-LSTM model (RMSE = 0.00224, MAPE $\approx$ 5.8%) and significantly improves prediction accuracy. These findings suggest that combining econometric periodic methods with deep learning can be a powerful tool for forecasting financial volatility.

Keywords: Model hybridization, Long Short-Term Memory neural network, PLM-GARCH, financial time series.
Classification: 62M10, 91G70, 68T07.

1 Introduction

Forecasting volatility is a fundamental element of modern financial theory and practice, with overt applications in option pricing, portfolio optimization, risk management, and regulation policy. For instance, the BlackScholes model of option pricing uses constant volatility as a simplifying assumption; but empirical results show that under this assumption options will be systematically mispriced. Accurate volatility forecasts can avoid such mispricing, make hedging strategies optimal, and enhance

[2]Corresponding author

Received: 22/05/2025 Accepted: 08/09/2025
https://doi.org/10.22054/JMMF.2025.85886.1185

efficiency in capital allocation. There are several stylized facts that have been reliably found in financial time series, such as volatility clustering, long-memory behavior, and periodic patterns. Volatility clustering refers to the empirical observation that large (or small) asset price movements should be accompanied by other large (or small) movements, regardless of sign. Long-memory processes exhibit slowly declining autocorrelations, so that shocks to volatility persist for extended periods. Volatility patterns can repeat on a periodic basis because of seasonal factors, cycles, or structural characteristics of specific marketse.g., energy markets, where geopolitical tensions and demand cycles will create recurring volatility patterns. These characteristics need to be identified and simulated in order to enable accurate predictions.

GARCH-type econometric models like the Fractionally Integrated GARCH (FIGARCH) and Periodic Long-Memory GARCH (PLM-GARCH) have widely been used to model long memory dependencies and periodicity in volatility. The FIGARCH models provide for fractional integration so that the long-memory effects are modeled, and the PLM-GARCH enhances this by incorporating periodic terms within the conditional variance. The theoretical motivation and interpretability of these models allow practitioners to investigate both persistence and periodicity in volatility.

Although robust in themselves, traditional econometric models do depend upon linearity and stationarity assumptions and thus are less sensitive to the nonlinear and regime-switching nature of real financial markets. Deep learning models, such as Long Short-Term Memory (LSTM) neural networks, are extremely capable of untangling nonlinear dynamics and intricate temporal interconnections in sequential data. LSTM networks are a step above the vanishing gradient issue inherent in standard recurrent neural networks and can retain useful information for long-time horizons. They are less interpretable than econometric models and may require large datasets to train well.

Theoretically, combining PLM-GARCH with LSTM is likely to improve the accuracy of forecasting by leveraging the strengths of both approaches: the periodic-long-memory modeling and interpretability of PLM-GARCH and the nonlinear pattern detection ability of LSTM. The PLM-GARCH component extracts structured periodicity and persistence, while the LSTM component forecasts any leftover nonlinear dynamics. Such synergy can enhance robustness across many regimes of markets, specifically in energy indices where periodic and nonlinear performances are both important. The rest of the paper is organized in the following fashion. Section 2 provides the research background. Section 3 presents the dataset and outlines the statistical and deep learning methods utilized, along with technical details of the PLM-GARCH and LSTM models. Section 4 outlines the empirical findings, including comparisons between the suggested hybrid approach and benchmark models. Section 5 concludes with a discussion of results, implications, and future research directions.

2 Literature Review

Predicting volatility in financial markets is a basic problem in both econometrics and machine learning. The early literature focused on linear econometric specifications that attempted to model persistence and long-run dependence in the conditional variance of financial time series. The GARCH model introduced by Bollerslev (1986) [5] and its popular extensions such as EGARCH (Nelson, 1991) [34] and GJR-GARCH (Glosten et al., 1993) [20] have been among the most heavily used tools for modeling volatility dynamics.

As long-memory processes attracted more interest, the FIGARCH model (Bollerslev & Mikkelsen, 1996) [6] was proposed to capture the slow decay of autocorrelations in conditional variance through fractional integration. The Periodic Long-Memory GARCH (PLM-GARCH) model (Bordignon et al., 2008) [7] was later an extension of FIGARCH to incorporate periodic components so that seasonal and cyclical patterns could be modeled in volatility. This capability is particularly valuable for the energy markets, where volatility is typically driven by recurring demand-supply cycles, seasonality, and geopolitical shocks.

Another significant branch of the literature has entertained the use of macroeconomic variablesinterest rates, industrial production, or oil prices, for examplein regression-based volatility models. Incorporating such exogenous variables can, at times, do away with mathematical problems such as nonlinearity or periodicity. The present study takes a different route by using only the Tehran Stock Exchange energy index. The rationale is theoretical in the sense that sector-specific indices, especially in energy, have a tendency to capture their own volatility dynamics in historical price data since they are responsive to industry-specific and geopolitical drivers and hence are self-contained indicators for volatility modeling.

For machine learning, recurrent neural networks (RNNs) and specifically Long Short-Term Memory (LSTM) networks (Hochreiter & Schmidhuber, 1997) [24] have been found to possess strong capabilities in approximating nonlinear temporal dependencies and long-range relationships in sequential data. Studies such as Fischer & Krauss (2018) [17] and Nelson et al. (2018) [35] have shown that LSTMs can outperform classical econometric models in predictive performance, at the cost of interpretability, however.

Hybridization of econometric models with deep learning has attracted growing interest. For example, Kim & Won (2021) [29] found that the combination of GARCH-type models with neural networks reduced the prediction error. However, there have been few research works on the hybridization of periodic long-memory econometric models like PLM-GARCH and LSTM networks. This research closes this gap by proposing a novel PLM-GARCH + LSTM model, aiming to jointly model structured periodiclong-memory volatility patterns and nonlinear residual dynamics in the energy index time series.

3 Methodology

3.1 ARMA/GARCH-type Models

The foundation for ARMA models in forecasting financial variables originates from the Box and Jenkins (1976) [8] methodology. This approach involves identifying an ARMA(p,q) model that accurately represents the stochastic process underlying the data. The ARMA model can be expressed as:

$$\Phi(B)r_t - \mu = \theta(L)\varepsilon_t, \quad \varepsilon_t|\Omega_{t-1} \sim N(0, \sigma_t^2) \tag{1}$$

$$\Phi(B) = 1 - \varphi_1 B - \varphi_2 B^2 - \ldots - \varphi_p B^p \tag{2}$$

$$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \ldots - \theta_q B^q \tag{3}$$

Where r_t stands as a representation of stock market index return, N stands for conditional normal density with zero mean, and conditional variance is σ_t^2, besides Ω_{t-1} indicating all the information that has been available till $t-1$; B stands for backshift operator and μ as series mean. The polynomials $\Phi(B)$ and $\Theta(B)$ correspond to the autoregressive (AR) and moving average (MA) parts, respectively, and their roots lie outside the unit circle.

Volatility modeling gained prominence after the introduction of ARCH (Engle, 1982) [16] and GARCH (Bollerslev, 1986) [5] models. Extensions such as EGARCH (Nelson, 1991) [34] and GJR-GARCH (Glosten et al., 1993) [20] address asymmetries and leverage effects.

3.2 FIGARCH Model

The Fractionally Integrated GARCH (FIGARCH) model, introduced by Baillie et al. (1996) [4], is designed to capture long-term memory in volatility, analyzing the persistence of shocks in financial time series. The GARCH and FIGARCH models are defined as:

$$\sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2 \tag{4}$$

$$\sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2 + \left[1 - (1-L)^d\right] \epsilon_t^2 \tag{5}$$

Where σ_t^2 represents the conditional variance at time t, ω is a constant, and α_i and β_j are parameters for the ARCH and GARCH values, respectively. The term $\left[1 - (1-L)^d\right]$ introduces the long-memory effect. Here, d is the fractional differencing parameter ($0 < d < 0.5$ for covariance stationarity). Estimation of d can be performed using methods similar to ARFIMA models, such as the GewekePorter-Hudak (GPH) estimator or maximum likelihood methods (Bollerslev & Mikkelsen, 1996) [6].

3.3 PLM-GARCH Model

Financial time series generally have complex temporal structure, and the behavior of volatility is conditioned by both the cyclical and seasonal patterns. It is important to capture such components in a proper manner for accurate volatility prediction. The two primary components characterizing such time series are:

- Seasonality: These are regular and predictable movements at fixed time periods, e.g., monthly, quarterly, or yearly seasons. Seasonality most commonly arises from calendar effects, business cycles, or other periodic phenomena.
- Cycles: These are prolonged movements in the data caused by underlying economic, social, or financial factors. As opposed to seasonality, cycles are not seasonally fixed in length, hybrid LSTM for periodic long-memory in financial energy index and typically are longer than one period in length, longer than seasonal cycles.

To properly capture the above dynamics, two complementary modeling approaches are employed:

Seasonality is captured by SARFIMA (Seasonal Autoregressive Fractionally Integrated Moving Average) models, which are an extension of basic ARFIMA models to account for seasonally long-memory behavior. Cycles are modeled with the Periodic Long-Memory GARCH (PLM-GARCH) model, extending standard GARCH models to include periodic and long memory features in volatility modeling.

Definition and Motivation of PLM-GARCH

The PLM-GARCH was introduced by Bordignon et al. (2008) [7] as a natural extension of GARCH frameworks to better capture temporal volatility patterns exhibiting both periodicity and long memory. In fact, it is quite a useful framework in dealing with financial time series whose volatility tends to exhibit cyclical behavior depending on calendar effects or market microstructure or even macroeconomic cycles. It imposes a structure of periodic lag operators and fractional differencing parameters so that short-run and persistent effects are identified, which basic GARCH or FIGARCH modeling cannot do.

Mathematical Formulation

The PLM-GARCH model, denoted as PLM-GARCH(p, d, q, S), is expressed as:

$$\sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2 + \left[1 - (1 - L^S)^d\right] \epsilon_t^2 \tag{6}$$

where

- σ_t^2 is the conditional variance at time t, representing the volatility.

- $\omega > 0$ is a constant term.
- $\alpha_i \geq 0$ and $\beta_j \geq 0$ are the ARCH and GARCH parameters, respectively, capturing short-term effects.
- L^S is the seasonal lag operator defined as $L^S X_t = X_{t-S}$, with S being the seasonal period (e.g., 12 for monthly data with annual seasonality).
- $d \in (0, 0.5)$ is the fractional differencing parameter that measures the intensity of long memory and persistence in the volatility process.
- ϵ_t is the innovation term (usually assumed to be white noise).

The term $\left[1 - (1 - L^S)^d\right]$ introduces fractional difference operators at seasonal lags, capturing long-memory dependence that varies periodically over the seasonal cycle. This contrasts with the FIGARCH model, which applies fractional differencing at lag 1 and thus captures only non-periodic long memory.

Interpretation and Properties

The PLM-GARCH model generalizes the FIGARCH model by allowing the fractional integration to depend on seasonal lags. This makes the model highly suitable for financial data with pronounced periodicity in volatility. The long-memory parameter d controls the degree of persistence; values close to zero correspond to weak persistence whereas values closer to 0.5 indicate strong, slowly decaying impacts of shocks.

The model simultaneously accounts for:

- **Short-term clustering effects** through the α_i and β_j parameters (ARCH and GARCH terms),
- **Periodic long memory** via seasonal fractional differencing,
- **Cyclical variations in volatility** through the seasonal lag operator L^S.

Estimation and Inference

Estimation of the PLM-GARCH model parameters, including the fractional differencing parameter d, can be performed via maximum likelihood methods or semi-parametric techniques similar to those used for FIGARCH models (e.g., Geweke-Porter-Hudak (GPH) estimator). The model's limiting distributions and sample size conditions are detailed in Bordignon et al. (2008) [7] and Baillie (1996) [4].

Frequency-Domain Analysis

Apart from time-domain modeling strategies, this study also applies frequency-domain analysis to assist in identifying periodic components in the return series. In particular, it employs a periodogram to estimate the periodogram spectral density over its spectrum. Peaks in the spectral density represent dominant periodicities, which may correspond to cycles in the economy, seasonality, and market cycles. The identification of these dominant periodic components offers both theoretical and empirical rationale for the consideration of periodic components in the final PLM-GARCH specification. This is particularly important in the context of energy markets, as cyclical patterns of volatility can reflect seasonal demand and external shocks and actions.

In summary the PLM-GARCH model offers a flexible and rigorous framework to analyze financial time series volatility characterized by long memory and periodic behaviors. By combining seasonal fractional differencing with classical volatility dynamics, the model captures complex dependence structures, providing improved forecasting accuracy over traditional GARCH and FIGARCH models.

3.4 Deep Learning Model – Long Short-Term Memory (LSTM)

Overview

Long Short-Term Memory (LSTM) networks, introduced by Hochreiter and Schmidhuber (1997) [24], are a specific category of recurrent neural networks (RNNs) designed to learn both long-term and short-term dependencies in sequential data. This feature renders them particularly potent for application in time series forecasting applications, especially when patterns stretch across several time scales. Unlike classical feedforward (FF) and standard backpropagation (BP) networks, which lack a mechanism for maintaining past state information, LSTM networks use a gated architecture to control storage, updating, and access of information over time. Such gates eliminate issues such as the *vanishing gradient problem*, which typically disables training of simple RNNs.

Gated Architecture

The LSTM unit uses three primary gates to regulate the information flow:

- **Forget Gate:** Determines which information from the previous cell state should be discarded.

$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] + b_f\right) \tag{7}$$

- **Input Gate and Candidate Memory:** Determines which new information should be added to the cell state.

$$i_t = \sigma\left(W_i \cdot [h_{t-1}, x_t] + b_i\right) \tag{8}$$

$$\tilde{C}_t = \tanh\left(W_c \cdot [h_{t-1}, x_t] + b_c\right) \tag{9}$$

- **Cell State Update:** Combines the retained old state and the new candidate values to update the memory.

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t \tag{10}$$

- **Output Gate:** Controls what part of the updated cell state is output as the hidden state.

$$o_t = \sigma\left(W_o \cdot [h_{t-1}, x_t] + b_o\right) \tag{11}$$
$$h_t = o_t \cdot \tanh(C_t) \tag{12}$$

Here:

- x_t is the input vector at time t.
- h_t is the hidden state vector at time t.
- C_t is the cell state (the internal memory).
- $\sigma(\cdot)$ is the sigmoid activation function.
- $\tanh(\cdot)$ is the hyperbolic tangent activation function.
- W_f, W_i, W_c, W_o are weight matrices and b_f, b_i, b_c, b_o are bias vectors for each gate.

Advantages of LSTM

Compared to conventional BP and FF architectures, LSTM offers:

- The ability to handle very long input sequences without losing information about dependencies in the data.
- The capability to learn and model non-linear temporal dependencies without the need for extensive manual feature engineering.
- Flexibility to perform multi-step forecasting for complex time series.

Limitations of LSTM

Although good, LSTM networks have a few disadvantages:

- More computational cost and increased training times compared to simpler models.
- Require more substantial datasets to achieve constant generalization performance.
- Less interpretable than parametric econometric models such as ARMA and GARCH.

The Depth of the LSTM Model

The LSTM model is classified as a deep learning approach due to its multi-layer architecture and ability to learn hierarchical temporal patterns from sequential data. In this study, the LSTM network consists of three hidden layers with 256, 128, and 64 neurons, respectively. Each successive layer captures increasingly abstract and complex features of the energy index return series, enabling the model to detect both short-term and long-term dependencies.

The depth in this context refers not to the number of variables but to the number of computational layers and the models capacity to extract multi-scale temporal dependencies. While traditional shallow models can only learn simple relationships, the deep structure of LSTM, combined with its gated memory units (forget, input, and output gates), allows it to retain relevant information over long time spans and discard irrelevant noise.

Using the real energy index data, the multi-layer LSTM effectively models the nonlinear and long-memory behavior of the series, as evidenced by its improved forecasting performance compared to shallower architectures. This hierarchical feature extraction is the primary reason LSTM is considered a deep learning method.

LSTM Implementation

In the current research, the LSTM network is configured as follows:

- **Architecture:** Three hidden layers with 256, 128, and 64 neurons, respectively.
- **Activation Functions:**
  - tanh for the cell state activation.
  - Sigmoid for the gate activations.
- **Optimizer:** Adam optimizer.
- **Learning Rate:** 0.0001.
- **Batch Size:** 16.
- **Training Epochs:** 500 epochs with early stopping (patience of 50 epochs) to prevent overfitting.

In summary the LSTM model integrates the strengths of recurrent architectures with gating mechanisms to effectively capture both short- and long-term dependencies in financial time series data. Its implementation in this study aims to enhance predictive performance by leveraging these capabilities while addressing the limitations inherent in classical econometric models.

3.5 Model Evaluation

Model evaluation criteria such as Mean Square Error (MSE), Mean Absolute Percentage Error (MAPE), Mean Absolute Error (MAE), and Coefficient of Determination (R^2) are used.

$$MSE = \frac{1}{n}\sum_{t=1}^{n}(y_t - \hat{y}_t)^2 \tag{13}$$

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\frac{|y_t - \hat{y}_t|}{y_t} \times 100\% \tag{14}$$

$$MAE = \frac{1}{n}\sum_{t=1}^{n}|y_t - \hat{y}_t| \tag{15}$$

$$R^2 = \frac{\sum_{t=1}^{n}(\hat{y}_t - \bar{y})^2}{\sum_{t=1}^{n}(y_t - \bar{y})^2} \tag{16}$$

4 Main Results

4.1 Data Analysis

The dataset employed in the research is composed of daily closing prices of the Iran energy index at the Tehran Stock Exchange from October 30, 2016, to January 25, 2023, amounting to 1396 observations. The energy index was considered for this study since energy markets usually tend to exhibit strong periodic volatility patterns stemming from seasonal demand cycles, industrial production schedules, and geopolitical events. The cyclical behavior combined with a long-term persistence in volatility, on the other hand, renders such a series ideal to be modeled through the Periodic Long-Memory GARCH (PLM-GARCH) framework, which is essentially designed to capture precisely these features. While the procedure followed here could actually be performed for any other financial indices (for example, broad stock market indices), the distinctive periodic and sector-specific features of the energy sector have made it worthy of consideration in this research.

The chosen time frame offers several advantages. First, it covers a period with multiple market regimes, including both high- and low-volatility phases, thereby allowing the model to learn diverse volatility behaviors. Second, the 1396 observations provide a statistically sufficient sample for estimating the parameters of long-memory models such as PLM-GARCH, where large datasets are often recommended for reliable estimation. The start date reflects the availability of complete and reliable data, while the end date ensures the inclusion of recent market conditions.

Although the present study models the volatility of the energy index using only its own historical prices, it is important to note that other variablessuch as global oil prices, interest rates, or macroeconomic indicatorscan also influence volatility forecasts. The decision to use only the index itself is based on the premise that sector-specific indices, especially in energy markets, inherently reflect their own volatility drivers through past price movements, thus providing a self-contained framework for periodic long-memory modeling.

We employed the formula

$$r_t = \ln\left(\frac{p_t}{p_{t-1}}\right)$$

to compute the returns of the stock exchange index. The descriptive statistics in Table 1 are based on the log-return series, which forms the basis of the volatility modeling in this study. The mean is close to zero, consistent with the behavior of financial returns, while the relatively large standard deviation indicates substantial variability in the energy index. The positive skewness and high kurtosis signify heavy-tailed distribution with frequent recurrence of outlier events. These characteristics imply more frequent recurrence of large shocks, both positive and negative, compared to normality.

In addition, we conducted the JarqueBera (JB) test, which assesses whether a given dataset adheres to the normal distribution by examining skewness and kurtosis [26]. The tests statistic is expressed as:

$$JB = \frac{n}{6}\left(S^2 + \frac{1}{4}(K-3)^2\right),$$

where n denotes the number of data values, S is the sample skewness indicating the degree of asymmetry, and K is the sample kurtosis representing the thickness of the distributions tails. JarqueBera test ($p < 0.001$) can reject the null hypothesis of normality and thus confirm the presence of non-normal characteristics of the data.These distributional properties have direct implications for model choice. Heavy tails, volatility clustering, and asymmetric return behavior justify the application of conditional heteroskedasticity models such as PLM-GARCH, which are capable of modeling long-memory and periodic volatility structures. Furthermore, since GARCH-type models may not fully capture nonlinear dependencies in the residuals, combining them with an LSTM network provides a means to account for complex and nonlinear temporal patterns, thereby enhancing forecasting accuracy.

Furthermore, we defined

$$\alpha_t = r_t - E_{t-1},$$

where p_t denotes the assets closing price at time t.

The stationarity of the ARMA process was evaluated using the Augmented DickeyFuller (ADF) test on the return series of the energy index. The test statistic was 12.45 with a p-value below 0.01, strongly rejecting the null hypothesis of a unit root. This indicates that the return series is stationary and suitable for ARMA modeling without differencing. Additionally, the autocorrelation function (ACF) of the returns decays quickly toward zero, further supporting stationarity. These

Table 1: Descriptive statistics of log returns

Statistics	**Value**
Mean	0.000142
Standard Deviation	0.018532
Skewness	0.851
Kurtosis	9.462
Jarque-Bera Test (P-value)	< 0.001

results validate the use of ARMA in the mean equation of the volatility models applied in this study.

Table 2: AIC and BIC values for the models

Criteria	**AIC**	**BIC**
PLM-GARCH	-8.438062	-8.410863
FIGARCH	-8.366394	-8.339194

Furthermore, the LjungBox test [32] for squared Q-statistics was employed to detect autocorrelation and heteroscedasticity. The significant results suggest that past market behaviors may hold substantial relevance.

The Akaike Information Criterion (AIC) also ranks the models in a specific order. The AIC and BIC criteria in Table 2 show that the PLM-GARCH model fits the dataset better than the FIGARCH model.

Model residuals were evaluated using the serial correlation LM test and the ARCH LM test. The significant F-statistic of the ARCH LM test confirms the presence of conditional heteroscedasticity in the residuals, validating the application of GARCH-type models to capture this property.

In addition both visual and numerical evidence were provided to support the presence of key time-series features in the energy index returns. Volatility clustering is illustrated in Figure 1, where periods of high and low variance are clearly visible, and is further confirmed through the ARCH-LM test and the LjungBox Q-test on squared returns in Table 3.

Table 3: Statistical Test Results

Test	**Statistic**	**p-value**
ARCH-LM	354.4505	4.53e-70
Ljung-Box Q	435.9429	2.08e-87

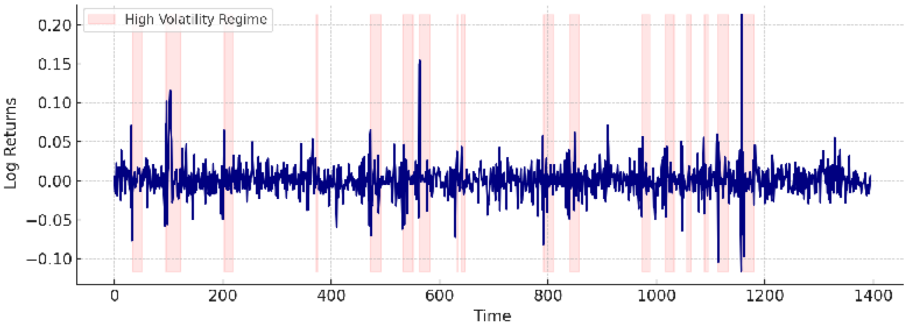

Figure 1: Volatility Clustering in Energy Index Returns

4.2 Building Long-Memory Pattern of GARCH-type Model

Mathematically, differencing with an order value of d was required to make the data stationary concerning the mean, estimated using the GewekePorterHudak (GPH) model. The order model determined by the GPH model was $\hat{d}_{\text{GPH}} = 0.49$.

As $\hat{d}_{\text{GPH}} = 0.49 < 0.5$, the data exhibited a long-memory effect and could be modeled with FIGARCH. Subsequently, the order of the FIGARCH model was determined by identifying the number of significant lags in the ACF and PACF plots, as shown in Figure 2.

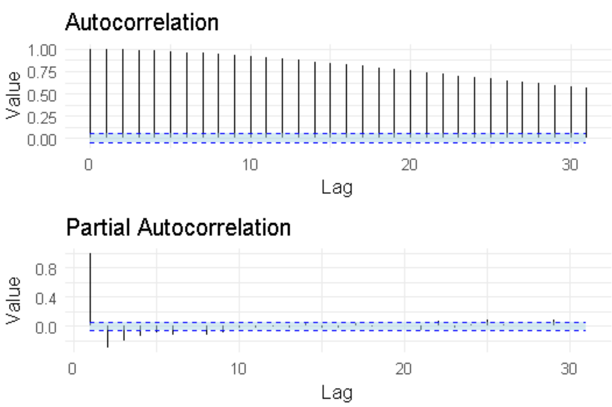

Figure 2: Autocorrelation and partial autocorrelation function

Based on Figure 2, the ACF plot shows a slow decay in correlation values across increasing lags, indicating the presence of long memory in the data. On the other hand, the PACF exhibits a sharp drop after the first lag, suggesting an autoregressive (AR) process of low order. Specifically, the ACF remains significant up to higher lags, while the PACF becomes insignificant beyond the first lag.

Based on Figure 2, the ACF plot shows a slow decay in correlation values across

increasing lags, indicating the presence of long memory in the data. On the other hand, the PACF exhibits a sharp drop after the first lag, suggesting an autoregressive (AR) process of low order. Specifically, the ACF remains significant up to higher lags, while the PACF becomes insignificant beyond the first lag.

These characteristics suggest that the FIGARCH model could be suitable for modeling the data. A possible specification for this model could include a maximum lag of 1 or 2 for the autoregressive component (p) and a higher lag (e.g., 5) for the moving average component (q), along with a fractional differencing parameter (d) capturing the long-memory behavior. To confirm the adequacy of the chosen model, it is necessary to estimate the parameters and conduct a significance test on the results.

As introduced in Section 3.3.5, we perform a frequency-domain analysis of the return series using the periodogram. This allows us to detect dominant cycles in the data. Figure 3 and Figure 4 illustrate the spectral density of the series and highlight the main frequency components, confirming the presence of periodic volatility patterns.

In this study, each complete cycle contains 180 observations, which may correspond to daily data (assuming 180 working days per year). Figure 3 presents a chart showing the distance component, with a periodic pattern fluctuating between -1000 and 3000. This recurring pattern indicates a strong periodic component in the time series.

With a frequency of 180, each full cycle consists of 180 observations. Given that there are 1396 total data points, the number of cycles (S) can be calculated by dividing the total observations by the cycle length: $S = 1396/180 \approx 7.76$. Rounding to the nearest integer, $S \approx 8$. This calculation suggests that the dataset spans roughly 8 complete cycles, and around 8 repeating patterns can be observed in the distance component chart.

The Figure 4 displays the theoretical spectrum of a PLM-GARCH model. This chart reveals several peaks, indicating the presence of distinct frequencies within the signal. These frequencies correspond to the periodic cycles within the signal.

Theoretical reasons exist because the PLM-GARCH model is built to handle time series with long-term persistence and seasonal or cyclical volatility patterns, like what we see in energy markets. The volatility of these types of markets can be influenced by cyclical demand in seasons, geopolitical events, and regulatory factors which ultimately can account for periodicity in volatility series. The PLM-GARCH model is an extension of FIGARCH, which accounts for long memory, while including periodic elements to model seasonal patterns. Long memory and periodicity make the PLM-GARCH the perfect candidate to describe the volatility of the energy index.

The probability values for each model are presented in Table 4 and Table 5. A model was considered significant when the probability value of its parameter was less than 0.05.

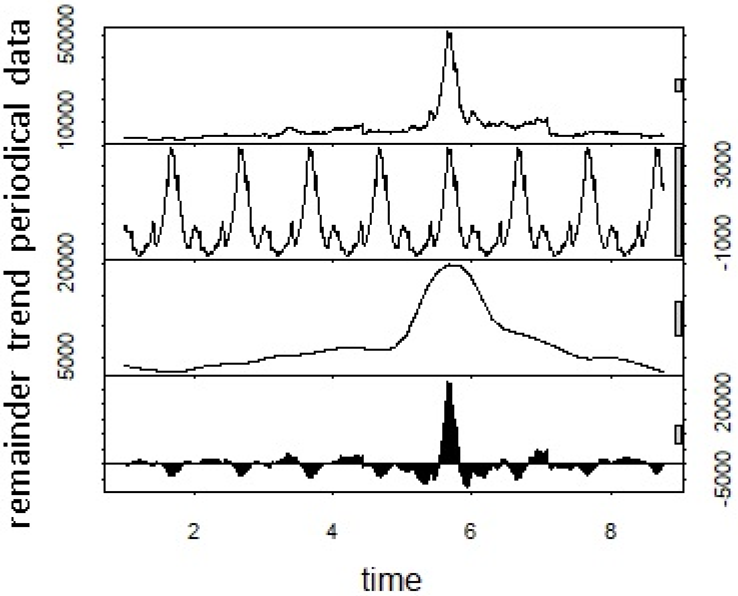

Figure 3: Decomposition of Stock Prices Time Series

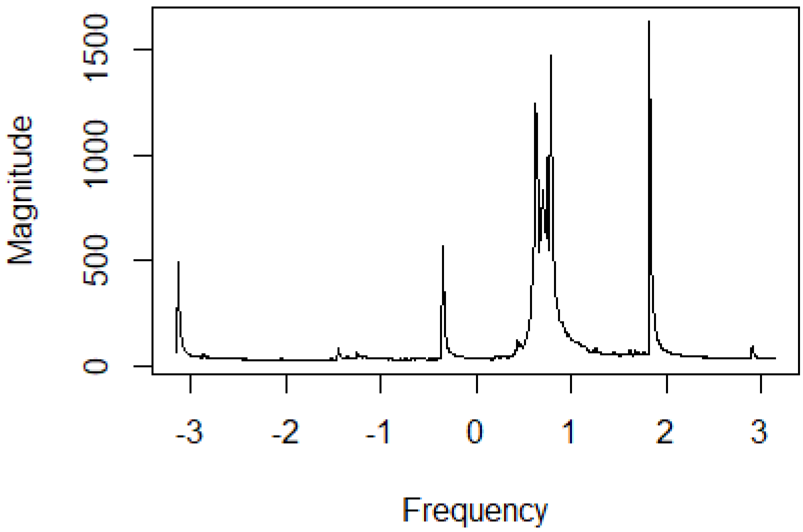

Figure 4: Theoretical spectrum of a PLM-GARCH

Table 4 shows that the models PLM-GARCH (1, 0.49, 1, 8), PLM-GARCH (1, 0.49, 2, 8), and PLM-GARCH (1, 0.49, 3, 8) were significant and suitable for building a PLM-GARCH model. However, not all significant models were applied in the subsequent steps. To identify the optimal model, a comparison was made between the AIC and BIC values. The evaluation of these values in Table 4 for the six models revealed that PLM-GARCH (1, 0.49, 2, 8) exhibited the lowest AIC and BIC values among the available alternatives. Consequently, it can be stated that PLM-GARCH (1, 0.49, 2, 8) appeared as the most favorable choice.

Relying solely on model selection was insufficient to confirm that PLM-GARCH (1, 0.49, 2, 8) adequately fulfilled the necessary conditions as a time series model. This led to the examination of the residual assumption of the PLM-GARCH (1, 0.49, 2, 8) model.

Residual assumption tests of FIGARCH (1, 0.49, 2) over the essence of volatility are presented in Table 5. As can be seen from the same table, residual autocorrelation test p-values were above 0.05, implying a lack of significant correlation among the residuals. On the contrary, the heteroscedasticity and normality tests resulted in p-values lower than 0.05, suggesting the residuals might show signs of heteroscedasticity or volatility effects, which require either adjustment or treatment. The normality test being significant was less of a concern in this analysis, given that highly rapid fluctuations exist in the time series data.

After establishing that the residuals of the FIGARCH model were heteroskedastic, it was then found that an alternative had to be sought. In next sections, different models addressing heteroskedasticity are displayed: the linear Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model and the nonlinear Long Short-Term Memory (LSTM) model.

4.3 Improving volatility residual GARCH types family using LSTM

According to the residual assumptions of FIGARCH (1, 0.49, 2) shown in Table 6, the heteroscedasticity assumption had not been met. Consequently, an advanced model was necessary to enhance ARFIMA and minimize variance in the residuals. One traditional time series, GARCH, had been developed to counteract the random fluctuating variance or heteroscedasticity impact. The creation of a GARCH involved using ACF and PACF charts to determine the order of the model.

After fitting FIGARCH to the long-patterned stock exchange data series, efforts were directed toward improving the heteroscedasticity of the model by addressing the residual. Visual diagnosis could be used to identify the presence of the heteroscedasticity effect. Outlier data indicated that an advanced model was necessary to adjust the effect due to data variability. Modifying the residual heteroscedasticity effect of FIGARCH was essential. The persistent vanishing/exploding gradient problem resulting from long-term dependencies, even with substantial data, posed a

Table 4: The estimated parameters of the PLM-GARCH model

Model	Parameter	Estimate	Std. Error	t value	Pr(> $\|t\|$)	AIC	BIC
plm_garch_1_0.49_1_8	mu	-0.02752124	0.0323369226	-0.8507765	0.3948824430	2.843011	2.872478
plm_garch_1_0.49_1_8	omega	0.01741388	0.0035673275	25.437426	0.0011143853	2.843011	2.872478
plm_garch_1_0.49_1_8	alpha1	0.01498456	0.0150931664	0.9933093	0.0000000065	2.843011	2.872478
plm_garch_1_0.49_1_8	beta1	0.99999961	0.0000000001	1000000.0	0.0000000000	2.843011	2.872478
plm_garch_1_0.49_1_8	delta	1.0	NA	NA	NA	2.843011	2.872478
plm_garch_1_0.49_1_8	shape	1.915327	0.1472563936	13.004865	0.0000000000	2.843011	2.872478
plm_garch_1_0.49_2_8	mu	0.02322912	0.0138412942	1.678003	0.4803708643	2.833955	2.873209
plm_garch_1_0.49_2_8	omega	0.0181576	0.0032856749	5.52876	0.0595356573	2.833955	2.873209
plm_garch_1_0.49_2_8	alpha1	0.01493992	0.0148611445	1.0053	0.0000000022	2.833955	2.873209
plm_garch_1_0.49_2_8	beta1	0.99999966	0.0469148853	21.467054	0.0000000000	2.833955	2.873209
plm_garch_1_0.49_2_8	beta2	0.0	NA	NA	NA	2.833955	2.873209
plm_garch_1_0.49_2_8	delta	1.0	NA	NA	NA	2.833955	2.873209
plm_garch_1_0.49_2_8	shape	1.92561656	0.0745809672	21.647252	0.0000000000	2.833955	2.873209
plm_garch_1_0.49_3_8	mu	-0.01580342	0.0372596749	-0.424259	0.3449231669	2.833998	2.878660
plm_garch_1_0.49_3_8	omega	0.01188645	0.0032848845	3.618495	0.1002992373	2.833998	2.878660
plm_garch_1_0.49_3_8	alpha1	0.01461276	0.0141518318	1.032558	0.0000000015	2.833998	2.878660
plm_garch_1_0.49_3_8	beta1	0.9999998	0.0351241265	28.46846	0.0000000000	2.833998	2.878660
plm_garch_1_0.49_3_8	beta2	0.11840126	NA	NA	NA	2.833998	2.878660
plm_garch_1_0.49_3_8	beta3	-0.04397264	NA	NA	NA	2.833998	2.878660
plm_garch_1_0.49_3_8	delta	1.0	NA	NA	NA	2.833998	2.878660
plm_garch_1_0.49_3_8	shape	1.93752610	0.1488018233	13.008751	0.0000000000	2.833998	2.878660

In Table 4 and Table 5, the abbreviation NA indicates "Not Available" and is used when a specific metric cannot be computed because the model does not produce the required output for that evaluation criterion, or the calculation would be meaningless given the models structure. For example, certain error measures may not be defined for models without corresponding predicted values at specific forecast horizons.

challenge due to the random fluctuation in residuals. Consequently, the application of the LSTM neural network was deemed necessary [39].

For the deep learning stage, the dataset was split into training and testing subsets, with 80% of the data allocated for training the LSTM network and 20% reserved for testing. This division ensures that the models predictive performance can be evaluated on unseen data, reducing the risk of overfitting. Parameter settings for the LSTM model were clearly shown in Table 7.

4.4 Evaluating the Volatility Model

After adjusting the heteroscedasticity effect within the residuals, a comparison of LSTM-FIGARCH and LSTM-PLM-GARCH models was conducted to evaluate their performance. Figure 5 and Figure 6 graphically showed these models, including the actual data, LSTM-FIGARCH (1, 0.49, 2), and LSTM-PLM-GARCH (1, 0.49, 2, 8). The validity was supported by evaluating the model using metrics such as Mean Squared Error (MSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE).

Table 5: The estimated parameters of the FIGARCH model

Model	Parameter	Estimate	Std. Error	t value	Pr(> |t|)	AIC	BIC
figarch_1_0.49_1	mu	-6044.47332817	65.122374321	-92.817152	0.0000000000	17.16588	17.18841
figarch_1_0.49_1	omega	14564.92803524	12035.36652759	13.777611	0.0001538402	17.16588	17.18841
figarch_1_0.49_1	alpha1	10.8702899391	0.0089252276	1218.00000	0.0000000000	17.16588	17.18841
figarch_1_0.49_1	beta1	0.4993706096	0.0097355206	51.280000	0.0000000000	17.16588	17.18841
figarch_1_0.49_1	delta	1.0	NA	NA	NA	17.16588	17.18841
figarch_1_0.49_1	shape	11.0803606	0.1015114622	109.222620	0.0000000000	17.16588	17.18841
figarch_1_0.49_2	mu	5919.9318219	47.122752492	125.66195	0.0000000000	17.12847	17.15475
figarch_1_0.49_2	omega	14546.5835126	12004.260346	12.115900	0.0000000000	17.12847	17.15475
figarch_1_0.49_2	alpha1	10.90755708	0.0062040034	1758.00000	0.0000000000	17.12847	17.15475
figarch_1_0.49_2	beta1	0.418417456	0.005778551	76.908127	0.0000000000	17.12847	17.15475
figarch_1_0.49_2	beta2	0.0	NA	NA	NA	17.12847	17.15475
figarch_1_0.49_2	delta	1.0	NA	NA	NA	17.12847	17.15475
figarch_1_0.49_2	skew	18.6304053	0.200744473	22.864437	0.0000000000	17.12847	17.15475
figarch_1_0.49_2	shape	11.49241309	0.140198422	81.973110	0.0000000000	17.12847	17.15475
figarch_1_0.49_3	mu	-6111.4763934	61.20260391	-99.87319	0.0000000000	17.15617	17.18621
figarch_1_0.49_3	omega	15459.26016391	12691.38689484	12.177897	0.0000000000	17.15617	17.18621
figarch_1_0.49_3	alpha1	10.84980172	0.0093674072	1157.68205	0.0000000000	17.15617	17.18621
figarch_1_0.49_3	beta1	0.3797436999	0.0098247315	38.651753	0.0000000000	17.15617	17.18621
figarch_1_0.49_3	beta2	0.0370933949	NA	NA	NA	17.15617	17.18621
figarch_1_0.49_3	beta3	0.0233142395	NA	NA	NA	17.15617	17.18621
figarch_1_0.49_3	delta	1.0	NA	NA	NA	17.15617	17.18621
figarch_1_0.49_3	skew	14.14246039	0.141945225	99.183306	0.0000000000	17.15617	17.18621
figarch_1_0.49_3	shape	11.15268123	0.012334395	195.459414	0.0000000000	17.15617	17.18621

Table 6: Residual assumption tests Results

Statistic	χ^2 Statistic	p-value
Homoscedasticity	737.8808	3.41×10^{-150}
Autocorrelation	[12208.0712, 10]	0
Normality	205.4091	$< 2 \times 10^{-16}$

All statistical and econometric analyses were conducted using **RStudio** . Extremely small p-values (e.g., 3.41×10^{-150}) are direct outputs from R's test functions and indicate overwhelming evidence against the null hypothesis.

A comparison between LSTM-FIGARCH and LSTM-PLM-GARCH is presented in Table 8. From Table 8, the LSTM-PLM-GARCH model yielded the smallest values for all three evaluation criteria. This outcome suggested that employing the numerical model, LSTM-PLM-GARCH, enhanced and refined the predicted values.

After adjusting the periodical pattern of data, the residuals went through further processing using LSTM to address the vanishing gradient issue inherent in the volatility component, often referred to as heteroscedasticity effects. Consequently, the preferred LSTM neural network effectively improved the heteroscedasticity issue of the classical GARCH models.

Table 7: Parameter settings for the LSTM model

Parameters	**Values**
Number of layers	3
Number of neurons in layers	(256, 128, 64)
Learning rate	0.0001
Optimizer	Adam
Batch size	16
Number of epochs	500
Validation data percentage	20%
Early stopping method	Monitoring `val_loss`, patience 50

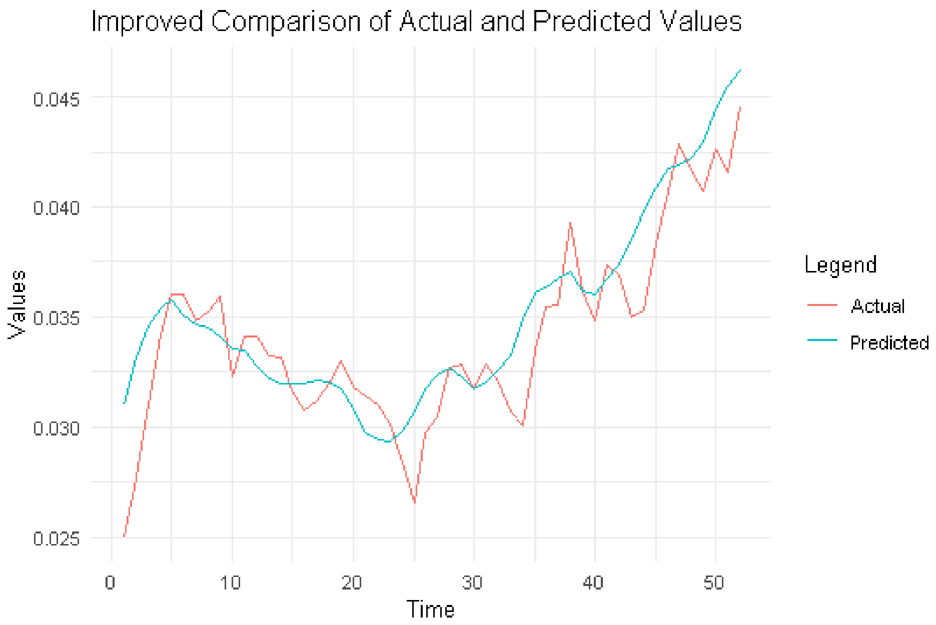

Figure 5: Forecasting volatilities of LSTM-PLM-GARCH model

Table 8: Comparison of Models

Loss Function	**LSTM-PLM-GARCH**	**LSTM-FIGARCH**
RMSE	0.002089	0.002236
MAE	0.001643	0.001929
MAPE	0.050926	0.058345

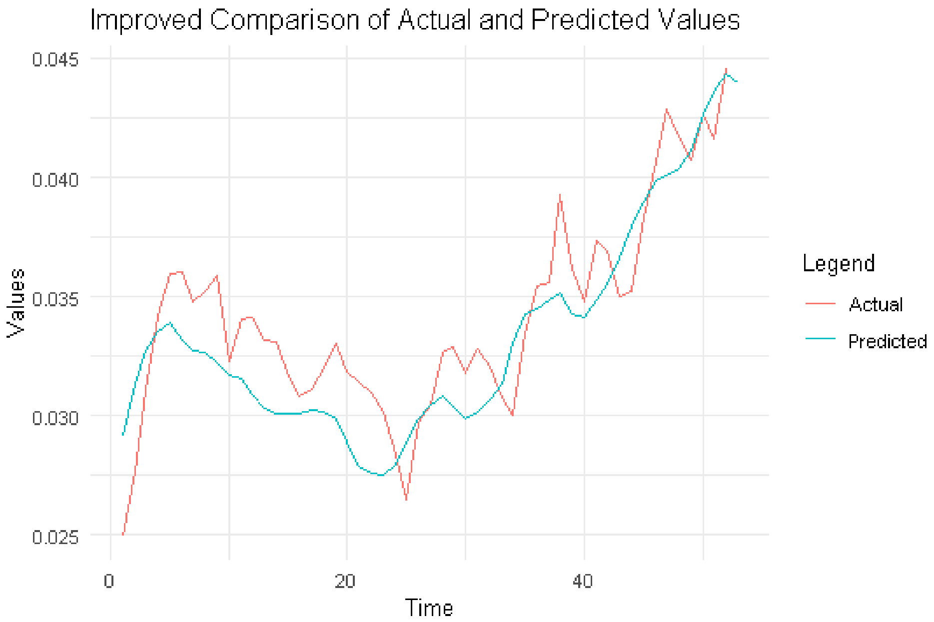

Figure 6: Forecasting volatilities of LSTM-FIGARCH model

5 Conclusion

In conclusion, this paper proposed an enhanced sensitivity model by incorporating Long Short-Term Memory (LSTM) neural network into FIGARCH and PLM-GARCH models. The proposed models, LSTM-FIGARCH and LSTM-PLM-GARCH, are then compared by using Mean Absolute Error (MAE), Mean Square Error (MSE), and Mean Absolute Percentage Error (MAPE). The results show that the LSTM-PLM-GARCH model improves the error and also the heteroscedasticity effect of the classical GARCH family models.

In addition, the advantages of LSTM-PLM-GARCH are to achieve stability by learning the previous data with the dynamical system. This increases the complexity in processing the networks, to approximate the gradient of GARCH family through numerical computations until a defined threshold error is met. Retaining information and patterns in residual data caused LSTM to effectively mitigate the heteroscedasticity issue present in GARCH family models.

Despite capturing the periodical and long-pattern data inherent, the GARCH family models do not adequately optimize data prediction. This led to the investigation of a nonlinear solution to the gradient problem. The method included using the LSTM, which is known for its ability to retain information and patterns in residual data. Due to this implementation, LSTM effectively mitigated the heteroscedasticity problem in GARCH family models.

As a result, the model handled the vanishing gradient problem, allowing the LSTM neural network to learn and bridge considerable temporal gaps even spanning more than 1000 discrete time steps. However, the randomness of initial weights and the number of iterations persist as limitations of the GARCH-LSTM model, with the number of iterations increasing as the threshold error of the networks decreases. It will require an extended period to determine a suitable weighting parameter. Furthermore, because the initial weights are random, the ideal parameters do not have the same values across experiments, necessitating validation of the training data while developing networks.

If the error of the training data is less than the error of the testing data, the networks are required to be stopped and the processing repeated until the error of the training data is more than the error of the testing data to ensure that the networks can recognize the new data using the obtained model of LSTM-PLM-GARCH.

6 Discussion

In this study, the combination of FIGARCH and PLM-GARCH models with Long Short-Term Memory (LSTM) neural networks was explored as a hybrid method for forecasting stock market volatility. The primary objective of this research was to enhance the accuracy of forecasting financial time series by leveraging the features of both models.

The results obtained showed that the PLM-GARCH model is capable of identifying long-term dependencies and periodic patterns in the conditional variance of time series. However, it has limitations in modeling complex and nonlinear patterns when used alone. On the other hand, the LSTM neural network has significant capability in learning nonlinear relationships and hidden patterns in financial data. The combination of these two models improved forecasting performance and reduced model error.

The findings of this study are comparable with several previous studies that have investigated volatility modeling in financial time series. In many of the previous studies, FIGARCH and GARCH models have been used to model the conditional variance of financial data. For instance, a study by Wang et al. (2020) showed that the FIGARCH model outperforms the GARCH model in modeling the long-memory volatility of the market. However, these models have limited capability in identifying nonlinear patterns in financial data.

On the other hand, studies such as Fischer and Krauss (2018) have demonstrated that deep neural networks, including LSTM, can be effectively used for forecasting stock returns. These studies suggest that deep learning models are capable of extracting complex features and hidden patterns in financial data, but the main challenge of these models is their sensitivity to noise and their high dependency on the quality of input data.

This study, by combining these two approaches, has yielded results similar to the research by Kim and Won (2021), where the combination of GARCH models with neural networks led to a reduction in forecasting error and improved model accuracy. Specifically, our results showed that using FIGARCH for data preprocessing and removing potential noise helps the LSTM network better identify nonlinear relationships and achieve more accurate performance in forecasting financial volatility.

In comparison with these studies, the present research takes a step further by utilizing the PLM-GARCH model, which has better capability in identifying periodic structures in financial data. This feature has enabled the proposed model to show higher accuracy in forecasting volatility compared to traditional methods. The results of this research study agree with earlier studies and seem to suggest that there is a lot to be gained by combining econometric methods with neural networks for the forecasting of financial time series.

From a practical viewpoint, the empirical results of this research can be of relevance to investors, financial analysts, and policymakers, as higher accuracy in predicting volatility allows better decisions in risk management and capital allocation. Last but not least, we recommend future studies to explore the combination of alternative deep learning paradigms with PLM-GARCH to improve forecasting accuracy further and employ the model with financial data at different frequencies.

Bibliography

[1] Y. Alyousifi, M. Othman, and A. A. Almohammedi. A novel stochastic fuzzy time series forecasting model based on a new partition method. *IEEE Access*, 9:80236–80252, 2021.

[2] T. G. Andersen and T. Bollerslev. Heterogeneous information arrivals and return volatility dynamics: Uncovering the long-run in high frequency returns. *Journal of Finance*, 52(3):975–1005, 1997.

[3] E. Arif, E. Herlinawati, D. Devianto, M. Yollanda, and D. Permana. Hybridization of long short-term memory neural network in fractional time series modeling of inflation. *Frontiers in Big Data*, 6:1282541, 2024.

[4] R. T. Baillie. Long memory processes and fractional integration in econometrics. *Journal of Econometrics*, 73(1):5–59, 1996.

[5] T. Bollerslev. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3):307–327, 1986.

[6] T. Bollerslev and H. O. Mikkelsen. Modeling and pricing long memory in stock market volatility. *Journal of Econometrics*, 73(1):151–184, 1996.

[7] S. Bordignon, M. Caporin, and F. Lisi. Periodic long memory GARCH models. *Econometric Reviews*, 27(5):569–582, 2008.

[8] G. E. P. Box and G. M. Jenkins. *Time Series Analysis: Forecasting and Control.* Holden-Day, San Francisco, 1976.

[9] A. H. Bukhari, M. A. Raja, M. Sulaiman, and S. Islam. Fractional neurosequential ARFIMA-LSTM for financial market forecasting. *IEEE Access*, 8:71326–71338, 2020.

[10] K. Chen, Z. Chen, S. Wang, and Z. He. Short-term load forecasting with deep residual networks. *IEEE Transactions on Smart Grid*, 10(4):3943–3952, 2019.

[11] K. Chen, Y. Zhou, and F. Dai. A LSTM-based hybrid model for financial time series forecasting. *Expert Systems with Applications*, 190:116261, 2022.

[12] R. Cont. Empirical properties of asset returns: Stylized facts and statistical issues. *Quantitative Finance*, 1(2):223–236, 2001.

[13] D. Devianto, K. Ramadani, A. Maiyatri, Y. Asdi, and M. Yollanda. The hybrid model of autoregressive integrated moving average and fuzzy time series Markov chain on long-memory data. *Frontiers in Applied Mathematics and Statistics*, 8:1045241, 2022.

[14] D. Devianto, M. Yollanda, A. Maiyatri, and F. Yanuar. The soft computing FFNN method for adjusting heteroscedasticity on the time series model of currency exchange rate. *Frontiers in Applied Mathematics and Statistics*, 9:1045218, 2023.

[15] D. A. Dickey and W. A. Fuller. Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, 74(366):427–431, 1979.

[16] R. F. Engle. Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 50(4):987–1007, 1982.

[17] T. Fischer and C. Krauss. Deep learning with long short-term memory networks for financial market predictions. *European Journal of Operational Research*, 270(2):654–669, 2018.

[18] K. Gajamannage, Y. Park, and D. I. Jayathilake. Real-time forecasting of time series in financial markets using sequentially trained dual-LSTMs. *Expert Systems with Applications*, 223:119879, 2023.

[19] J.-C. Gamboa. Deep learning for time-series analysis. *arXiv preprint arXiv:1701.01887*, 2017.

[20] L. R. Glosten, R. Jagannathan, and D. E. Runkle. 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, 1993.

[21] B. Gulmez. Stock price prediction with optimized deep LSTM network with artificial rabbits optimization algorithm. *Expert Systems with Applications*, 227:120346, 2023.

[22] S. A. Haider, S. R. Naqvi, T. Akram, G. A. Umar, A. Shahzad, M. R. Sial, et al. LSTM neural network based forecasting model for wheat production in Pakistan. *Agronomy*, 9:72, 2019.

[23] T. Hasenzagl, F. Pellegrino, L. Reichlin, and G. Ricco. A model of the Fed's view on inflation. *Economic Research Paper*, 104:686–704, 2022.

[24] S. Hochreiter and J. Schmidhuber. Long short-term memory. *Neural Computation*, 9(8):1735–1780, 1997.

[25] H. H. Huang, N. H. Chan, K. Chen, and C. K. Ing. Consistent order selection for ARFIMA processes. *Annals of Statistics*, 50:1297–1319, 2022.

[26] C. M. Jarque and A. K. Bera. A test for normality of observations and regression residuals. *International Statistical Review*, 55(2):163–172, 1987.

[27] M. E. Javanmard, Y. Tang, Z. Wang, and P. Tontiwachwuthikul. Forecast energy demand, CO_2 emissions and energy resource impacts for the transportation sector. *Applied Energy*, 338:120830, 2023.

[28] S. Kaushik, A.Choudhury, P. K. Sheron, N. Dasgupta, S. Natarajan, L. A. Pickett, et al. AI in healthcare: time-series forecasting using statistical, neural, and ensemble architectures. *Frontiers in Big Data*, 3:4, 2020.

[29] H. Y. Kim and C. H. Won. Forecasting the volatility of stock price index: A hybrid model integrating LSTM with multiple GARCH-type models. *Expert Systems with Applications*, 103:25–37, 2021.

[30] B. Lei, Z. Liu, and Y. Song. On stock volatility forecasting based on text mining and deep learning under high-frequency data. *Journal of Forecasting*, 40(8):1596–1610, 2021.

[31] R. Liu, Y. Jiang, and J. Lin. Forecasting the volatility of specific risk for stocks with LSTM. *Procedia Computer Science*, 202:111–114, 2022.

[32] G. M. Ljung and G. E. P. Box. On a measure of lack of fit in time series models. *Biometrika*, 65(2):297–303, 1978.

[33] M. McAleer. Automated inference and learning in modeling volatility. *Econometric Theory*, 21(1):232–261, 2005.

[34] D. B. Nelson. Conditional heteroskedasticity in asset returns: A new approach. *Econometrica*, 59(2):347–370, 1991.

[35] D. M. Nelson, A. C. M. Pereira, and R. A. de Oliveira. Stock market's price movement prediction with LSTM neural networks. In *International Joint Conference on Neural Networks (IJCNN)*, pages 1–6, 2018.

[36] S. Pan, S. Long, Y. Wang, and Y. Xie. Nonlinear asset pricing in Chinese stock market: a deep learning approach. *International Review of Financial Analysis*, 87:102627, 2023.

[37] M. L. Rahman, M. Islam, and M. Roy. Modeling inflation in Bangladesh. *Open Journal of Statistics*, 10:998–1009, 2020.

[38] E. Ramos-Pérez, P. J. Alonso-González, and J. J. Núñez-Velázquez. Multi-transformer: A new neural network-based architecture for forecasting S&P volatility. *Mathematics*, 9(15):1794, 2021.

[39] A. Shewalkar, D. Nyavanandi, and S. Ludwig. Performance evaluation of deep neural networks applied to speech recognition: RNN, LSTM and GRU. *Journal of Artificial Intelligence and Soft Computing Research*, 9:235–245, 2019.

[40] Y. Wang and L. Liu. A hybrid forecasting model based on LSTM and GARCH for financial time series. *Applied Sciences*, 9(3):550, 2019.

[41] J. Wu, N. Levi, R. Araujo, and Y.-G. Wang. An evaluation of the impact of COVID-19 lockdowns on electricity demand. *Electric Power Systems Research*, 216:109015, 2023.

[42] X. Xu, Y. Zhang, C. A. McGrory, J. Wu, and Y.-G. Wang. Forecasting stock closing prices with an application to airline company data. *Data Science and Management*, 6:239–246, 2023.

[43] Y. Yang, H. Zhou, J. Wu, Z. Ding, Y.-C. Tian, D. Yue, et al. Robust adaptive rescaled lncosh neural network regression toward time-series forecasting. *IEEE Transactions on Systems, Man, and Cybernetics*, 53:5658–5669, 2023.

[44] L. Yu, S. Wang, and K. K. Lai. Financial time series forecasting with nonlinear and nonstationary characteristics: A review. *Applied Soft Computing*, 95:106545, 2020.

[45] Y. Yu, X. Si, C. Hu, and J. Zhang. A review of recurrent neural networks: LSTM cells and network architectures. *Neural Computation*, 31:1235–1270, 2019.

[46] J. Zhao, S. Xu, and X. Zhang. Periodic volatility modeling with hybrid neural networks and GARCH-type models. *Physica A*, 616:128574, 2023.

Abbreviations

ADF	Augmented Dickey–Fuller test
AIC	Akaike Information Criterion
AR	Autoregressive
ARCH	Autoregressive Conditional Heteroskedasticity
ARFIMA	Autoregressive Fractionally Integrated Moving Average
BIC	Bayesian Information Criterion
EGARCH	Exponential Generalized Autoregressive Conditional Heteroskedasticity
FIGARCH	Fractionally Integrated Generalized Autoregressive Conditional Heteroskedasticity
GARCH	Generalized Autoregressive Conditional Heteroskedasticity
GJR-GARCH	Glosten–Jagannathan–Runkle GARCH
JB test	Jarque–Bera test
LSTM	Long Short-Term Memory (neural network)
MA	Moving Average
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
MSE	Mean Squared Error
PACF	Partial Autocorrelation Function
PLM-GARCH	Periodic Long-Memory GARCH
RNN	Recurrent Neural Network
RMSE	Root Mean Squared Error
SARFIMA	Seasonal Autoregressive Fractionally Integrated Moving Average

How to Cite: Minou Yari[1], Mohammad Reza Salehi Rad[2], Mohammad Bahrani[3], *A Hybrid Lstm Neural Network Approach for Modeling Periodical Long-Memory Characteristics in Financial Energy Index Time Series* , Journal of Mathematics and Modeling in Finance (JMMF), Vol. 5, No. 2, Pages:173–196, (2025).